Methods in
Molecular Biology 3105

Springer Protocols

Sean P. Curran *Editor*

Aging

Methods and Protocols

Second Edition

Humana Press

Volume 3105

Methods in Molecular Biology

Series Editor
John M. Walker
School of Life and Medical Sciences, University of Hertfordshire, Hatfield, Hertfordshire, UK

For further volumes: http://www.springer.com/series/7651

For over 35 years, biological scientists have come to rely on the research protocols and methodologies in the critically acclaimed Methods in Molecular Biology series. The series was the first to introduce the step-by-step protocols approach that has become the standard in all biomedical protocol publishing. Each protocol is provided in readily-reproducible step-by-step fashion, opening with an introductory overview, a list of the materials and reagents needed to complete the experiment, and followed by a detailed procedure that is supported with a helpful notes section offering tips and tricks of the trade as well as troubleshooting advice. These hallmark features were introduced by series editor Dr. John Walker and constitute the key ingredient in each and every volume of the Methods in Molecular Biology series. Tested and trusted, comprehensive and reliable, all protocols from the series are indexed in PubMed.

Manuscript Instructions https://drive.google.com/file/d/1K4Az5vpAFVzYadJ3jKx2-QSo7Dqa_SNA/view?usp=sharing

Editor
Sean P. Curran

Aging
Methods and Protocols Second Edition

Second Edition 2027



Editor
Sean P. Curran
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

ISSN 1064-3745 e-ISSN 1940-6029
Methods in Molecular Biology
ISBN 978-1-0716-5702-7 e-ISBN 978-1-0716-5703-4
https://doi.org/10.1007/978-1-0716-5703-4

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2020, 2027

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Unless otherwise expressly agreed by the Publisher in writing, the Publisher does not authorize use or reproduction of all or any part of this work in any manner for the purpose of training, fine-tuning, evaluating, or developing artificial intelligence models, technologies, or systems. In accordance with Article 4(3) of the EU Directive 2019/790 on copyright and related rights in the Digital Single Market, the Publisher expressly reserves all rights with respect to this work under the text and data mining exception of the directive.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Humana imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.
The registered company address is: 1 New York Plaza, New York, NY 10004, U.S.A.

Preface

The biology of aging has emerged as one of the most dynamic and interdisciplinary areas of modern biomedical research. Advances in genetics, molecular biology, physiology, computational sciences, and translational medicine have transformed our understanding of the mechanisms that govern lifespan and health span. Central to these discoveries has been the development of robust experimental approaches across diverse model systems. This second edition volume brings together established and emerging methodologies that enable investigators to study aging at molecular, cellular, organismal, and population levels. Written in the highly successful *Methods in Molecular Biology* series format, chapters include introductions to their respective topics, lists of the necessary materials and reagents, step-by-step, readily reproducible laboratory protocols, and key tips on troubleshooting and avoiding known pitfalls.

Authoritative and cutting-edge, *Aging: Methods and Protocols, Second Edition* aims to provide practical guidance and a valuable resource for researchers and experts seeking to advance the science of aging and promote healthy longevity.

Sean P. Curran
Los Angeles, CA, USA

Contents

Preface
Contributors

Chapter 1Calculating Epigenetic Clocks from DNA Methylation Data to Index Accelerated Epigenetic Aging
Harshanna Badhesha, Thalida Em Arpawong, Eric T. Klopack and Eileen M. Crimmins

Chapter 2A Standardized Workflow for Single-Cell/Nuclei RNA-Seq Analysis in Aging Research
Clayton Baker, Eyael Tewelde, Ella Schwab and Bérénice A. Benayoun

Chapter 3Isolation and Functional Screening of Plant-Derived Natural Products for Assessing Healthy Aging and Neuroprotective Activities in *Caenorhabditis elegans*
Chatrawee D. Shepard

Chapter 4Translational Profiling of Adult *C. elegans* Glia Using Translating Ribosome Affinity Purification (TRAP)
Karmen Mah and Ashley E. Frakes

Chapter 5An Integrated Pipeline for Quantifying Survival, Cell Morphology, and Nuclear Integrity During Chronological Aging of Fission Yeast
Kristen Hairapatian, Kayla Mallari, Thomas Riffle, Annika Schrader, Emily M. Sánchez, Júlia Guil Rodríguez, Shanjika Nagendrakumar, Martin Vo and Wilber Escorcia

Chapter 6Practical Considerations for Automated Lifespan Measures in *C. elegans*
Shivay Parihar, Athena Alcala and Gilberto Garcia

Chapter 7Assessing Age-Related Changes in *Caenorhabditis elegans* Behavior Using the Forced Food Choice Assay
Nicole L. Stuhr and Sean P. Curran

Chapter 8Analyzing Spatial Learning Behavior and Cognitive Function in Aging Mice Using the Barnes Maze
Amy Christensen

Chapter 9Quantification of *C. elegans* Movement and Muscle Function Using WormLab-Derived Maximum Dynamic Amplitude
Brandy A. Weathers, Tripti Nair and Sean P. Curran

Chapter 10Voluntary Wheel Running as a Noninvasive and Low-Stress Exercise Paradigm in Aging Mice
Jocelyne Leon, Karel Aceituno and Constanza J. Cortes

Chapter 11Characterizing Gut Diversity During Aging
Prerana Vaddi, Denise B. Lynch and Cristal M. Hill

Chapter 12Massively Parallel, Ultra-accurate RNA Sequencing with CircSeq
Claire S. Chung, Jean-François Goût and Marc Vermulst

Chapter 13RNAscope in the Brain of the Naturally Short-Lived African Turquoise Killifish
Rapheal G. Williams, George M. Cardenas, Hengyu Zhou and Bérénice A. Benayoun

Chapter 14Practical Guide to RNA Sequencing Analysis in Aging Research
Gee-Yoon Lee and Seung-Jae V. Lee

Chapter 15Mitochondrial-Wide Association Studies: Integrating Population-Based Data and Mitochondrial Genetics
Yanxi Li, Changyeob Lee and Thalida Em Arpawong

Chapter 16Imaging of Organelle and Actin Cytoskeletal Morphology During Aging
Aeowynn J. Coakley, Eric K. F. Donahue, Juri Kim, Maxim Averbukh, Gilberto Garcia, Ryo Higuchi-Sanabria and Kristopher Burkewitz

Chapter 17Measurements of Nuclear Pore Function in Living *Caenorhabditis elegans*
Yifei Zhou and Alexander A. Soukas

Chapter 18Measurement of Skeletal Muscle Mitochondrial DNA Synthesis Using Stable Isotope Labeling
Robert V. Musci, Morgan Johnson, Noelle Morrow, Clover D. Crenshaw, Morgan K. Daniel, Frederick F. Peelor, Benjamin F. Miller, Johanna Pham, Andrea L. Hevener and Jonathan Wanagat

Chapter 19Mnemonic Similarity Task Variants for Probing Age-Related Memory Impairments Across Species and Modalities
Lorena A. Ferguson, Chih-Yuan Jessie Chien, Stephanie L. Leal, Craig Stark, Sara N. Burke and Teal Eich

Chapter 20Guide to Extending Findings from Selected Samples Using Inverse Odds of Selection Weights
Eleanor Hayes-Larson

Contributors

Karel Aceituno
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Athena Alcala
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Thalida Em Arpawong
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Maxim Averbukh
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Harshanna Badhesha
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Clayton Baker
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA
Molecular and Cellular Biosciences Department, USC Dornsife College of Letters, Arts and Sciences, Los Angeles, CA, USA

Bérénice A. Benayoun
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA
Pharmacology and Pharmaceutical Sciences Department, Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California, Los Angeles, CA, USA

Cancer Biology Department, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA
USC Norris Comprehensive Cancer Center, Epigenetics and Gene Regulation, Los Angeles, CA, USA
USC Stem Cell Initiative, Los Angeles, CA, USA
Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California, Los Angeles, CA, USA

Sara N. Burke
Department of Neuroscience, University of Arizona, Tucson, AZ, USA
BIO5 Institute, University of Arizona, Tucson, AZ, USA

Kristopher Burkewitz
Department of Cell and Developmental Biology, Vanderbilt University, School of Medicine, Nashville, TN, USA

George M. Cardenas
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Chih-Yuan Jessie Chien
Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Amy Christensen
Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Claire S. Chung
University of Southern California, Leonard Davis School of Gerontology, Los Angeles, CA, USA

Aeowynn J. Coakley
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Constanza J. Cortes
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Clover D. Crenshaw
Loyola Marymount University, Los Angeles, CA, USA

Eileen M. Crimmins
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Sean P. Curran
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Morgan K. Daniel
Loyola Marymount University, Los Angeles, CA, USA

Eric K. F. Donahue
Department of Cell and Developmental Biology, Vanderbilt University, School of Medicine, Nashville, TN, USA

Teal Eich
Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Wilber Escorcia
Department of Biology, California State University, Los Angeles, CA, USA

Lorena A. Ferguson
Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Ashley E. Frakes
Genetics and Biochemistry Branch, NIDDK, National Institutes of Health, Bethesda, MD, USA

National Institute of Neurological Disorders and Stroke (NINDS), National Institutes of Health, Bethesda, MD, USA

Gilberto Garcia
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Jean-François Goût
Mississippi State University, Department of Biological Sciences, Starkville, MS, USA

Kristen Hairapatian
Department of Biology, California State University, Los Angeles, CA, USA

Eleanor Hayes-Larson
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Andrea L. Hevener
Division of Endocrinology, Department of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

Ryo Higuchi-Sanabria
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Cristal M. Hill
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Morgan Johnson
Division of Endocrinology, Department of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

Juri Kim

Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Eric T. Klopack
School of Public Health-Bloomington, Indiana University, Bloomingto, USA

Stephanie L. Leal
Integrative Biology and Physiology, University of California, Los Angeles, CA, USA

Changyeob Lee
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Gee-Yoon Lee
Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon, South Korea

Seung-Jae V. Lee
Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon, South Korea

Jocelyne Leon
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Yanxi Li
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Denise B. Lynch
One Codex, Wilmington, DE, USA

Karmen Mah

Genetics and Biochemistry Branch, NIDDK, National Institutes of Health, Bethesda, MD, USA

Kayla Mallari
Department of Biology, California State University, Los Angeles, CA, USA

Benjamin F. Miller
Aging and Metabolism Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA
Oklahoma City Veterans Affairs Medical Center, Oklahoma City, OK, USA

Noelle Morrow
Division of Endocrinology, Department of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

Robert V. Musci
Loyola Marymount University, Los Angeles, CA, USA

Shanjika Nagendrakumar
Department of Biology, California State University, Los Angeles, CA, USA

Tripti Nair
University of Southern California, Leonard Davis School of Gerontology, Los Angeles, CA, USA

Shivay Parihar
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Frederick F. Peelor
Aging and Metabolism Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA

Johanna Pham
Division of Geriatrics, Department of Medicine, University of California —Los Angeles, Los Angeles, CA, USA

Thomas Riffle
Department of Biology, California State University, Los Angeles, CA, USA

Júlia Guil Rodríguez
Department of Biology, California State University, Los Angeles, CA, USA

Emily M. Sánchez
Department of Biology, California State University, Los Angeles, CA, USA

Annika Schrader
Department of Biology, California State University, Los Angeles, CA, USA

Ella Schwab
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA
Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California, Los Angeles, CA, USA

Chatrawee D. Shepard
The Appalachian Natural Products Research Program, Department of Biomedical Sciences, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV, USA
Center of Excellence on Natural Products for Neuroprotection and Anti-ageing (Neur-Age NatChula), Department of Clinical Chemistry, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand

Alexander A. Soukas
Center for Genomic Medicine and Diabetes Unit, Endocrine Division, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
Broad Institute of Harvard and MIT, Cambridge, MA, USA

Craig Stark
Department of Neurobiology and Behavior, University of California, Irvine, CA, USA

Nicole L. Stuhr
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Eyael Tewelde
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA
Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California, Los Angeles, CA, USA

Prerana Vaddi
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Marc Vermulst
University of Southern California, Leonard Davis School of Gerontology, Los Angeles, CA, USA

Martin Vo
Department of Pediatrics, University of Hawai'i, Honolulu, HI, USA

Jonathan Wanagat
Division of Geriatrics, Department of Medicine, University of California —Los Angeles, Los Angeles, CA, USA
Geriatric Research, Education and Clinical Center (GRECC), VA Greater Los Angeles Healthcare System, Los Angeles, CA, USA

Brandy A. Weathers
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA
University of Southern California, Dornsife College of Letters, Arts and Sciences, Los Angeles, CA, USA

Rapheal G. Williams

Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Hengyu Zhou
Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

Yifei Zhou
Department of Pharmaceutical Analysis, School of Pharmacy, China Pharmaceutical University, Nanjing, China

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_1

1. Calculating Epigenetic Clocks from DNA Methylation Data to Index Accelerated Epigenetic Aging

Harshanna Badhesha[1], Thalida Em Arpawong[1] ✉, Eric T. Klopack[2] and Eileen M. Crimmins[1]

(1) Leonard Davis School of Gerontology, University of Southern California, 3715 McClintock Ave., Los Angeles, CA 90089, USA

(2) School of Public Health-Bloomington, Indiana University, 1025 E. Seventh Street, Bloomingto, IN 47405, USA

✉ **Thalida Em Arpawong**
Email: arpawong@usc.edu

Abstract
Reliable biomarkers to assess biological age are important in geroscience research. Reliable estimators of biological aging have been quantified from epigenetic data, specifically DNA methylation, a modification of DNA that does not change the sequence but can impact gene regulation. DNA methylation patterns quantified into summary markers of epigenetic aging that use years of age as their metric are called epigenetic clocks because they represent the ticking of biological time. As more DNA methylation data become available through large population-based studies, biobanks, and clinical studies, there is a growing need to document methodological protocols for calculating epigenetic clocks for all researchers, including those who do not typically work in the epigenetic space. Some of the most widely used clocks in geroscience research include Horvath, Hannum, PhenoAge, and GrimAge. This chapter provides a step-by-step guide for researchers new to DNA methylation analysis and introduces the principles of epigenetic clock calculation methods in the R programming language. This protocol provides guidance for calculating the clocks in addition to elucidating the necessary steps for quality control analysis of human DNA methylation data.

Keywords DNA methylation – Epigenetic clocks – Epigenetic age – R

1 Introduction

Chronological age, defined as years since birth, is one of the strongest predictors of mortality, disease burden, and declines in function [1, 2]. On the other hand, biological age reflects the functional health of an organism and can serve as a stronger indicator of an individual's risk of age-related health outcomes [3]. Both molecular (e.g., DNA methylation, telomere length, transcriptomics) and clinical biomarkers (e.g., inflammatory markers, blood pressure, glycated hemoglobin) have emerged as strong quantitative proxies for biological age [1]. This chapter focuses on DNA methylation (DNAm), or chemical modifications to DNA that alter gene expression without impacting the DNA sequence [4]. Measures based on DNAm have been shown to be reliable proxies of accelerated aging [1–4].

DNAm changes occur with the addition of a methyl group to DNA. It can influence gene transcription and cellular function [2]. Altered DNAm patterns have been associated with cancer [5], metabolic disease [6], mortality [7], and other health outcomes. Importantly, aggregated DNAm levels can serve as quantitative markers for estimating epigenetic age. Epigenetic clocks generally focus on methylation at locations where cytosine is followed by guanine (so-called CpG sites). Epigenetic clocks are developed using predictive models based on machine learning to identify CpG sites across the genome whose methylation levels are differentially associated with aging outcomes [2–4]. The human genome contains approximately 29 million CpG sites from which only a small subset is selected for clock estimation [8]. These clocks provide an estimate of an individual's epigenetic aging rate, or degree of accelerated/decelerated epigenetic aging when calculated as a deviation from one's chronological age [3]. DNAm data, from which these clocks are typically developed, come from microarray platforms such as the Illumina 450k or EPICv1 and EPICv2 arrays. These arrays measure methylation levels at hundreds of thousands of CpG sites, with the current technology measuring nearly one million sites. The choice of array can impact the clock results because each array measures a different set of CpG sites [9].

Since the introduction of the first epigenetic clock, the Horvath pan-tissue clock in 2013, well over 50 clocks have been developed [2], reflecting rapid methodological expansion in the field. The Horvath clock was trained to predict chronological age. It was developed using 7844 samples through a combination of 82 Illumina DNA methylation array datasets across multiple tissue types, identifying 353 CpG sites most relevant for predicting age. The Horvath clock provided the first evidence that DNAm patterns could serve as a robust marker for indexing age [10]. The Hannum blood-based clock, also released in 2013, was trained to predict chronological age using DNAm patterns measured in whole blood. The model identified 71 CpG sites and demonstrated strong correlations between predicted epigenetic age and chronological age, further supporting the idea that DNAm patterns can be a marker for indexing age [11]. Since then, additional generations of clocks have been developed, trained on different outcomes, including mortality and other biological markers, using more advanced algorithms to capture distinct dimensions of epigenetic aging related to health outcomes. While there are multiple epigenetic clocks, this chapter focuses on several of

the most widely used clocks in geroscience research (Horvath1, Hannum, PhenoAge, GrimAge1, and GrimAge2) and their calculation with the MethylCipher software, available on GitHub [12]. We also describe the calculation of two more recent clocks, DunedinPace and PhysAge, as examples of how to use the developer's repositories.

The ability for more researchers to calculate these clocks enables more widespread use of these useful estimators of epigenetic aging at multiple levels, from use in population-based work to group-level comparisons in research settings. Having a step-by-step guide to calculate these clocks moves the field forward, so researchers can effectively integrate these clocks into their studies.

---

2 DNA Methylation Data

The production of a DNAm dataset that can be used for epigenetic clock calculation requires the completion of several quality control steps, including sex verification of samples, sample integrity checks, array or sequencing quality assessment, normalization, and correction for batch effects [13]. The U.S. Health and Retirement Study (HRS) documentation provides excellent detail on performing quality control on DNAm data (hrs.isr.umich.edu/sites/default/files/genetic/HRS_DNAm_OCT2023.pdf). The HRS is a population-based study of individuals over age 50, sponsored by the National Institute on Aging (NIA U01AG009740) and conducted by the University of Michigan [13]. Ultimately, after quality control steps are taken, a beta matrix is produced, containing beta values for each CpG site for each study sample. Beta values reflect the proportion of methylated signal relative to the total methylated and unmethylated signal detected at each of the CpG sites (ranging from 0 to 1). This guide provides a detailed workflow for calculating epigenetic clocks from a beta matrix, referencing the critical quality control procedures that can affect the beta matrix values and thus results from clock estimation.

2.1 Finding and Accessing DNA Methylation Data

There are several DNAm datasets well-suited for calculating epigenetic clocks. The desired input for clock calculation is a beta matrix. Examples include the DNAm dataset provided by the Broad Institute (https://www.broadinstitute.org/epigenomics/dna-methylation-data) and HRS Venous Blood Study 2016 DNA methylation data (https://niagads.scrollhelp.site/support/application-instructions). Other datasets, such as Series GSE77716 and GSE42861, are available through Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/).

While some publicly available data require setting up an account and completing specific training, other datasets require authorization, which can take up to several months. Access to these datasets is typically granted through a proposal submission process and requires adherence to data use agreements. Detailed instructions for gaining access are available on the respective data repositories.

2.2 Training and Certificates

Accessing human genomic data requires prior authorization to ensure that participant privacy and data security protocols are strictly followed. Researchers must complete required training and obtain certifications to demonstrate compliance with confidentiality requirements, Institutional Review Board (IRB) approvals, and data use agreements. These steps are mandatory for anyone seeking access to human data and familiarize researchers with best practices for protecting individual identities. Additional information on NIH guidelines can be found at https://grants.nih.gov/policy-and-compliance/policy-topics/sharing-policies/accessing-data/certification-agreement and https://grants.nih.gov/policy-and-compliance/policy-topics/human-subjects/training-and-resources). NIAGADS-specific trainings and requirements are available at https://niagads.scrollhelp.site/support/data-application).

2.3 Data Use Agreements

In addition to required training and certifications, researchers must follow strict protocols to protect health information, including guidelines for data storage, access, and sharing. Information about NIH guidelines can be found at https://grants.nih.gov/policy-and-compliance/policy-topics/sharing-policies). Establishing secure computational environments and controlled access is essential to prevent unauthorized data access and ensure that individual participants cannot be identified.

---

3 Calculating Epigenetic Clocks in R

Researchers interested in calculating epigenetic clocks should have familiarity with the R programming language, basic data management, and analysis skills. Some data wrangling is required, and a basic understanding of data frames is necessary.

Because beta matrices are typically very large, containing hundreds of thousands of CpG sites and potentially thousands of individuals, the files generally cannot be opened and manipulated on a local computer. Often a high-performance computing environment is needed to use these beta matrices. To manage and analyze these data, this protocol described here uses R. Many of the relevant packages for calculating epigenetic clocks have been developed in the R language. R is widely used both for quality control of DNAm data and for calculating epigenetic clocks. Its open-source functionality allows access to multiple packages that facilitate data analysis, including MethylCipher [12], which will be used in this protocol. The latest documentation and download options for R and RStudio are available at www.r-project.org. Researchers planning to run epigenetic clocks should become familiar with R and practice uploading and exploring a beta matrix. Multiple online resources for learning R are available (e.g., RStudio Education, Codecademy, Coursera).

Online repositories host code and functions that allow for quick calculation of epigenetic clocks. These are typically functions that apply the clock algorithm directly to the input beta matrix without the need to code the entire algorithm manually.

MethylCipher is available for download via GitHub. The package includes simple, one-step functions that require a beta matrix to calculate epigenetic clocks, making it a quick and accessible option for researchers new to clock calculations. MethylCipher is actively maintained and supports multiple epigenetic clocks. Information on download and documentation can be found at https://github.com/HigginsChenLab/methylCIPHER.

3.1 Understanding the Beta Matrix

A fully quality-controlled (QC) beta matrix will have rows corresponding to samples and columns corresponding to CpG sites. Depending on the dataset, the beta matrix may need to be transposed to match this format. Two QC steps are particularly important for calculating epigenetic clocks: Detection *P*-values and bead counts. Detection *P*-values compare the total methylated and unmethylated signal for each CpG site to the background signal level; CpGs that fail the chosen threshold are removed, which can lead to missing values [14]. Bead counts represent the number of individual beads on the array used to measure a specific CpG site. Higher bead counts indicate more reliable measurements, while CpGs with bead counts below a specified threshold are either marked as missing (NA) or removed. QC documentation accompanying the beta matrix should describe the thresholds used and how CpGs failing these criteria were handled (see Sect. 3.2).

In this protocol, we use a beta matrix in which the sub-threshold beads were marked as missing (NA), thereby creating "holes" of missing values in the beta matrix [13, 15]. The thresholds are determined through the quality control analysis process and may differ by study. It is important for the researcher to become familiar with the QC process implemented in producing the beta matrix to be used for clock calculation (*see* **Notes**). The example provided below utilizes the HRS 2016 beta matrix and MethylCipher.

Code to transpose beta matrix:

```
## Load in beta matrix (called beta)
load("…/beta.RData")
## Check number of samples and CpGs
dim(beta)
## [1] 836660 4018
## This indicates 836660 CpGs(rows) and 4018 samples(columns)
## Transform to get samples corresponding to rows and CpGs
corresponding to columns
tbetas_median_imp<-t(beta)
dim(tbetas_median_imp)
## [1] 4018 836660
## This indicates 4018 samples(rows) and 836660 CpGs(columns)
```

3.2 Handling Missing CpGs from Some Samples

As noted above, missing beta values may be present in the beta matrix due to bead counts detected during quality control checks. Two common approaches for addressing missing CpG values are mean imputation and median imputation. Mean imputation replaces missing beta values with the average beta value of the respective CpG site across all samples. This approach is most appropriate for beta matrices with approximately symmetric distributions; however, it can lead to regression toward the mean and may introduce bias when calculating epigenetic clocks. Median imputation replaces missing beta values with the median beta value of the respective CpG site across all samples. This method is better suited for skewed beta value distributions and can reduce the influence of outliers, potentially resulting in less bias in epigenetic clock calculations [16]. Median imputation is used here to align with preprocessing approaches employed with second-generation epigenetic clocks within the HRS framework [13].

Median imputation code with the input file (transposed beta matrix called *tbetas_median_imp*) formatted as samples in rows and CpGs in columns:

```
## Calculate column medians
all_column_median <- apply(tbetas_median_imp, 2, median,
na.rm=TRUE)
## Impute median value where cell is NA
for(i in colnames(tbetas_median_imp)) tbetas_median_imp[,i]
[is.na(tbetas_median_imp[,i])] <- all_column_median[i]
## Save Median Imputed file. This will be the file used for
calculating clocks.
save(tbetas_median_imp, file ="…/tbetas_median_imp.RData")
```

Median imputation can be run for the entire beta matrix at once. Running median imputation may take several hours. For a beta matrix with ~4000 samples and ~900,000 CpG sites, and approximately 30 GB of RAM on a computing cluster, it will take approximately 6 h of processing time.

3.3 Handling Missing CpGs from the Entire Dataset

Required CpGs may be missing entirely from the beta matrix. Include all CpGs required for each clock to ensure accurate results. This step can be done before or after checking for missing CpGs for some samples (Sec. 3.2). Table 1 shows the number of CpG sites required to calculate each epigenetic clock. Missing CpGs can be dealt with using imputation methods that are recommended by the developers of the clocks or by replacing missing values with a gold-standard beta value applied across all samples. These values are available at https://dnamage.clockfoundation.org/) in the file *datMiniAnnotation4_fixed.csv* (https://www.dropbox.com/scl/fi/d3k9s3t4xsoxt8dmldnu5/datMiniAnnotation4_fixed.csv?rlkey=64gmyhj3wlbpcqyvpra4xg7sw&e=1&dl=0).

Please cite https://pmc.ncbi.nlm.nih.gov/articles/PMC7778979/) when using these resources.

Table 1 The number of CpG sites required for epigenetic clocks: Horvath1, Hannum, PhenoAge, GrimAge1/GrimAge2, DunedinPace, and PhysAge

Datasets may require manually adding missing CpGs prior to implementing the software commands. Constant values for missing CpGs can be added from those provided in *datMiniAnnotation4_fixed.csv*. This file contains duplicates and NA values. Rows with NA values should be removed, which will also remove the duplicates. Some CpGs may not be in this file, and NA will take the place of the beta value.

For GrimAge calculations (GrimAge1 or GrimAge2), a phenotype file containing "Age" and "Female" (coded as female = 1, male = 0) is also required. Variable names are case-sensitive, and the methylation (datMeth) and phenotype (datPheno) files must be sorted by sample and contain the same number of samples.

MethylCipher has a function to check for 28 clocks if all the required CpGs are in the beta matrix. The results of the first 6 clocks are shown below using the median imputed beta matrix:

```
## Install package
## devtools::install_github("HigginsChenLab/methylCIPHER")
library(methylCIPHER)
getClockProbes(tbetas_median_imp)
## Clock Total.Probes Present.Probes Percent.Present
## 1 Alcohol 450 447 99%
## 2 BMI 1109 1106 100%
## 3 Bocklandt 1 1 100%
## 4 Bohlin 251 231 92%
## 5 DNAmClockCortical 347 347 100%
## 6 DNAmTL 140 139 99%
```

To calculate epigenetic clocks not included in MethylCipher, missing values will need to be checked manually.

The example below shows how to manually check the 71 CpG sites in the Hannum clock. The same steps can be done for each clock. Replace 'hannum_cpgs' with your CpG list.

```
## Load in CpG list for each epigenetic clock of interest
hannum_cpgs <- read.csv("hannum_cpgs.csv")
library(dplyr)
## Convert to a data frame
tbetas_median_imp <- as.data.frame(tbetas_median_imp)
```

```
missinghannum <- setdiff(hannum_cpgs$cpg,
colnames(tbetas_median_imp))
## missinghannum
## [1] "cg07927379" "cg09651136" "cg14361627" "cg18473521"
"cg21139312" "cg24079702"
## For missing CpGs, the gold standard beta value will be
added in the next step.
We will need to add values for the missing CpGs in the beta
matrix using the "gold standard beta value" before calculating
the clock. This value will be the same for each sample:
## Load in datMiniAnnotation4_fixed file
datMiniAnnotation4 <- read.csv("datMiniAnnotation4_fixed.csv")
## Add in missing CpG constant values
lookup <- setNames(

   datMiniAnnotation4_fixed$ProbesMedian,
   datMiniAnnotation4_fixed$Name)

vals <- lookup[missinghannum]
valid_cols <-!is.na(missinghannum)
## This will add in the constant values to the median imputed
data matrix
tbetas_median_imp[missinghannum[valid_cols]] <-

   lapply(vals[valid_cols], function(v) rep(v,
nrow(tbetas_median_imp))
```

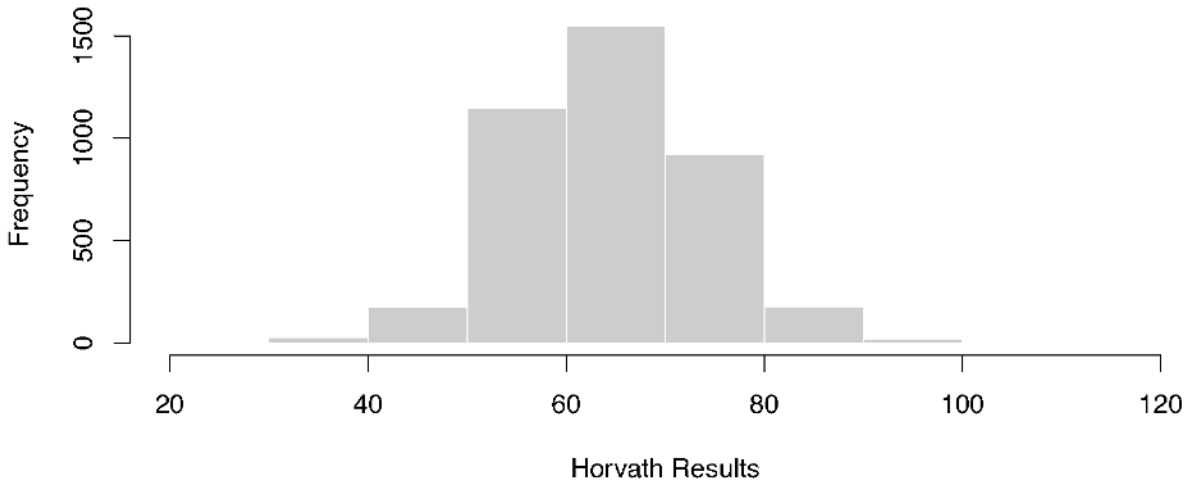
3.4 Running MethylCipher

Sections 3.2 and 3.3 addressed missing CpGs for some samples by using median imputation and missing CpGs for all samples by adding a constant value using the gold standard annotation file. The beta matrix can now be used to calculate epigenetic clocks. As mentioned above, MethylCipher currently has 28 clocks available; the full list can be found at https://github.com/MorganLevineLab/methylCIPHER.

MethylCipher has a simple function formatted "calc(CLOCKNAME)" for all 28 clocks. Depending on the size of the beta matrix, this step is generally quick.

A. **Horvath1**

The function to calculate the Horvath clock is calcHorvath1 and requires the input file to contain beta values with samples corresponding to rows and CpGs corresponding to columns. The median imputed file in the example is set up this way, with clock summary statistics shown in Fig. 1.



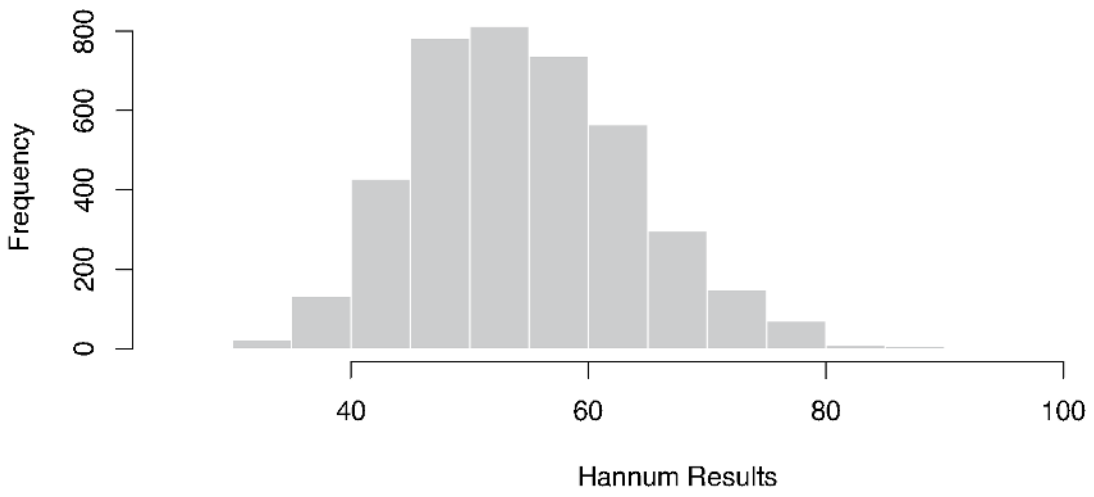
Min	Median	Mean	Max
22.07	64.01	64.49	113.30

Fig. 1 Expected summary statistics and Horvath1 clock age distribution for the 2016 HRS DNAm data

```
## Install packages if necessary
## devtools::install_github("MorganLevineLab/methylCIPHER")
library(methylCIPHER)
## Use MethylCipher to calculate Horvath1. 'Horvath1' will
contain the calculated age.
horvath1 <-
as.data.frame(methylCIPHER::calcHorvath1(tbetas_median_imp))
```

B. **Hannum Clock**

The function to calculate the Hannum clock [17] is calcHannum and requires the input file to contain beta values with samples corresponding to rows and CpGs corresponding to columns. The median imputed file in the example is set up this way, with clock summary statistics shown in Fig. 2.



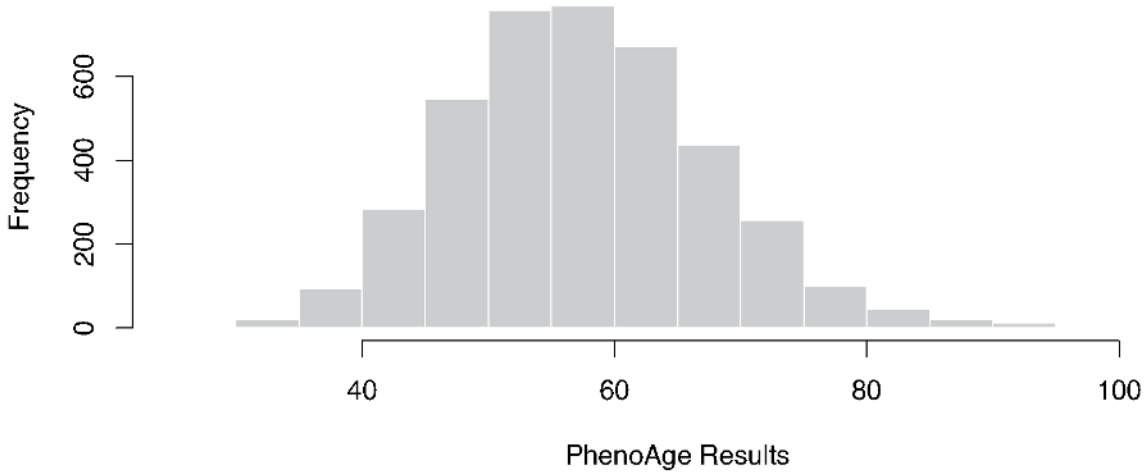
Min	Median	Mean	Max
25.06	53.91	54.59	107.79

Fig. 2 Expected summary statistics and Hannum clock age distribution for the 2016 HRS DNAm data

```
## Install packages if necessary
## devtools::install_github("MorganLevineLab/methylCIPHER")
library(methylCIPHER)
## Use MethylCipher to calculate Hannum. 'Hannum' will
contain the calculated age.
hannum <-
as.data.frame(methylCIPHER::calcHannum(tbetas_median_imp))
```

C. **PhenoAge**

The function to calculate the PhenoAge clock [18] is calcPhenoAge and requires the input file to contain beta values with samples corresponding to rows and CpGs corresponding to columns. The median imputed file in the example is set up this way, with the clock summary statistics shown in Fig. 3.



Min	Median	Mean	Max
26.69	56.84	57.47	101.68

Fig. 3 Expected summary statistics and PhenoAge clock age distribution for the 2016 HRS DNAm data

```
## Install packages if necessary
## devtools::install_github("MorganLevineLab/methylCIPHER")
library(methylCIPHER)
## Use MethylCipher to calculate PhenoAge. 'PhenoAge' will
contain the calculated age.
PhenoAge <-
as.data.frame(methylCIPHER::calcPhenoAge(tbetas_median_imp))
```

D.

GrimAge1

The GrimAge clock function can be found in a different version of MethylCipher (https://github.com/HigginsChenLab/methylCIPHER/tree/main). This version will need to be installed to calculate GrimAge. Unlike other epigenetic clocks, GrimAge1 or GrimAge2 requires a phenotype file in addition to the beta matrix. The median imputed beta matrix is used (although researchers may make different decisions on the type of imputation performed) in addition to an input file containing "Age" and "Female" (female = 1, male = 0) [19]. The beta matrix must be a matrix type, and the phenotype file must be a data frame.

```
## Install packages if necessary
## devtools::install_github("HigginsChenLab/methylCIPHER")
library(methylCIPHER)
## Use MethylCipher to calculate GrimAge1. 'GrimAge1' will
contain the calculated age along with 8 other surrogates.
tbetas_median_imp <- as.matrix(tbetas_median_imp)
```

```
GrimAge1 <-
as.data.frame(methylCIPHER::calcGrimAge1(tbetas_median_imp,
phenotype_df))
## Use MethylCipher to calculate GrimAge2. 'GrimAge2' will
contain the calculated age along with 8 other surrogates.
GrimAge2 <-
as.data.frame(methylCIPHER::calcGrimAge2(tbetas_median_imp,
phenotype_df))
```

3.5 Summary of Coding Steps

An overview of the coding steps required is presented below, along with line-by-line annotations explaining the purpose and function of each segment of code. Hannum is used as an example; however, Steps 6 and 7 must be repeated for each clock.

Step 1. Load the beta matrix

Load the beta matrix (called beta):

```
load("…/beta.RData")
```

Step 2. Transpose the matrix

Transpose the matrix so that samples correspond to rows and CpG sites correspond to columns:

```
tbetas_median_imp <- t(beta)
```

Step 3. Perform median imputation

Compute the median for each CpG (column), excluding missing values:

```
all_column_median <- apply(tbetas_median_imp, 2, median, na.rm
= TRUE)
Replace missing values in each column with the corresponding
column median:
for (i in colnames(tbetas_median_imp)) {

   tbetas_median_imp[, i][is.na(tbetas_median_imp[, i])] <-
all_column_median[i]

}
```

Note: This step may take several hours depending on the dataset size.

Step 4. Install and load the required package

```
devtools::install_github("HigginsChenLab/methylCIPHER")
```

```
library(methylCIPHER)
```

Step 5. Identify required CpG sites
Determine which CpG sites required for clock calculation are missing:

```
getClockProbes(tbetas_median_imp)
```

Step 6. Add missing CpG values (Hannum clock example)
Load required data and libraries:

```
library(dplyr)
hannum_cpgs <- read.csv("hannum_cpgs.csv")
tbetas_median_imp <- as.data.frame(tbetas_median_imp)
Identify CpG sites missing from the dataset:
missinghannum <- setdiff(hannum_cpgs$cpg,
colnames(tbetas_median_imp))
Load reference values and create a lookup table:
datMiniAnnotation4_fixed <-
read.csv("datMiniAnnotation4_fixed.csv")
lookup <- setNames(

   datMiniAnnotation4_fixed$ProbesMedian,
   datMiniAnnotation4_fixed$Name)

Retrieve values for missing CpGs:
vals <- lookup[missinghannum]
valid_cols <-!is.na(missinghannum)
Add missing CpGs to the dataset using reference values:
tbetas_median_imp[missinghannum[valid_cols]] <-

   lapply(vals[valid_cols], function(v) rep(v,
nrow(tbetas_median_imp)))
```

Step 7. Calculate the epigenetic clock (Hannum)

```
hannum <-
as.data.frame(methylCIPHER::calcHannum(tbetas_median_imp))
```

---

4 Resources to Calculate Other Epigenetic Clocks

MethylCipher does not include all epigenetic clocks that researchers may be interested in using. Other clocks, for example, include DunedinPACE (https://github.com/danbelsky/DunedinPACE) [20], which is an epigenetic proxy for one's annual rate of

aging, and PhysAge (https://github.com/em-arpy/DNAm_PhysAge) [21], which calculates epigenetic surrogates of indicators of physiological health (e.g., lung function, inflammation, kidney function, blood sugar) and a composite epigenetic estimator of biological age. Both GitHub repositories provide clear documentation and outline the required beta matrix formats for running each clock in R. The beta matrix should have all the necessary CpGs for these clocks as well, with each developer providing recommendations on how to handle the missing CpG values. These packages include functions that produce epigenetic age estimates efficiently within minutes.

A. **DunedinPACE**

DunedinPace requires a beta matrix structured with rows as CpGs and columns as samples. DunedinPace is based on 173 CpGs but uses an additional 19827 CpGs for normalization [20].

To get the 173 required CpGs, use:

```
requiredprobes <- getRequiredProbes(backgroundList = FALSE)
```

To get the 173 required CpGs and 19827 CpGs required for normalization, use:

```
requiredprobes <- getRequiredProbes(backgroundList = TRUE)
```

Code to calculate the DunedinPACE clock:

```
## Install package
## devtools::install_github("danbelsky/DunedinPACE",
build_vignettes = TRUE)
#dunedinpace
library(DunedinPACE)
## Transpose dataframe so that rownames correspond to CpGs
and columns correspond to samples.
tbetas <- t(tbetas_median_imp)
## Calculate DunedinPace clock
dunedinpaceresults <- PACEProjector(tbetas_median_imp)
```

B. **PhysAge**

To calculate the PhysAge clock, code can be copied from: https://github.com/em-arpy/DNAm_PhysAge/blob/main/Calculate_DNAmPhysAge_surrogates.R. The beta matrix should be formatted so that rows are samples and columns are CpG sites. The code guides researchers through the calculations and can be run straight through; however, note the two locations in the code that require changes in order to point to the desired beta matrix. Additionally, if there are CpGs missing from all samples, the developer recommends researchers use the CpG means calculated from the clock's

training sample to replace missing values prior to clock calculation. These means are provided on GitHub, and sample code is provided below.

```
# Add the HRS training set means (HRS_cpg_training_means.txt)
for any CpGs missing from the entire dataset

    cpgs <- tbetas_median_imp

# Get list of all CpGs needed from
'HRS_cpg_training_means.txt'
want0 <- read.table("…/HRS_cpg_training_means.txt", header=T)
want <- want0$cpg
# Get list of those you have
have<-colnames(tbetas_median_imp)
# Get list you don't have in your data
need<-setdiff(want, have)
# Subset the means file
need.betas <- want0[want0$cpg %in% need,]
# Add the constant (means) for the missing CpGs to your beta
file
rownames(need.betas)<-need.betas$cpg
need.betas$cpg<-NULL
tneed.betas<-t(need.betas)
# Impute constant
full.betas <- cbind(cpgs, tneed.betas[rep(1, nrow(cpgs)),])
dim(full.betas)
```

After running the code, the output will include eight surrogates in addition to the DNAmPhysAge clock, calculated in a metric of years [21].

---

5 Notes

QC documentation for DNAm analysis should include, at a minimum, information specified by the NIA Biomarker Network, a project sponsored by the National Institute on Aging. The suggested minimum reporting documentation can be downloaded here: https://biomarker-network.isr.umich.edu/wp-content/uploads/2025/04/Minimum-Reporting-Requirements-for-Epigenetic-Data.docx. This documentation should accompany a beta matrix and provide information on all QC and filtering steps performed, including detection *P*-value and beta count thresholds, the proportion of CpG sites or samples affected, and how the values failing these criteria were handled.

Multiple epigenetic clocks have been developed for predicting health-related outcomes and indicators, including cancer and mitotic rate clocks, pediatric age predictors, and trait-associated clocks. Many of these can be calculated using the same

steps outlined here using MethylCipher. Accurate estimation of epigenetic age relies on standardized preprocessing steps, rigorous quality control, and following recommended guidelines for imputing missing CpG values. With the combination of standardized QC and documentation and clock-specific guidance, researchers can improve reproducibility and strengthen the interpretation of epigenetic age across multiple research contexts.

Funding

Contributions to this work were partially supported by funding from the National Institutes of Health, with support from a National Institute on Aging grant for the Center for Biodemography and Population Health (P30-AG017265) and from a National Institute on Aging grant for the Nathan Shock Center for the Excellence in the Biology of Aging (P30-AG068345).

Competing Interests

All authors declare that they have no conflict of interest.

References

1. Jylhävä J, Pedersen N, Hägg S (2017) Biological age predictors. EBioMedicine 21:29–36. https://doi.org/10.1016/j.ebiom.2017.03.046
[PubMed][PubMedCentral]

2. Chervova O, Panteleeva K, Chernysheva E et al (2024) Breaking new ground on human health and well-being with epigenetic clocks: a systematic review and meta-analysis of epigenetic age acceleration associations. Ageing Res Rev 102:102552. https://doi.org/10.1016/j.arr.2024.102552
[PubMed]

3. Teschendorff AE, Horvath S (2025) Epigenetic ageing clocks: statistical methods and emerging computational challenges. Nat Rev Genet 26:350–368. https://doi.org/10.1038/s41576-024-00807-w
[PubMed]

4. Duan R, Fu Q, Sun Y et al (2022) Epigenetic clock: a promising biomarker and practical tool in aging. Ageing Res Rev 81:101743. https://doi.org/10.1016/j.arr.2022.101743
[https://doi.org/10.1038/s41576-024-00807-w][PubMed]

5. Lakshminarasimhan R, Liang G (2016) The role of DNA methylation in cancer. Adv Exp Med Biol 945:151–172. https://doi.org/10.1007/978-3-319-43624-1_7
[PubMed][PubMedCentral]

6. Illum LRH, Bak ST, Lund S et al (2018) DNA methylation in epigenetic inheritance of metabolic diseases through the male germ line. J Mol Endocrinol 60(2):R39–R56. https://doi.org/10.1530/JME-17-0189
[PubMed]

7. Zhang Y, Wilson R, Heiss J et al (2017) DNA methylation signatures in peripheral blood strongly predict all-cause mortality. Nat Commun 8:14617. https://doi.org/10.1038/ncomms14617
[PubMed][PubMedCentral]

8. Youk J, An Y, Park S et al (2020) The genome-wide landscape of C:G > T: a polymorphism at the CpG contexts in the human population. BMC Genomics 21:270. https://doi.org/10.1186/s12864-020-6674-1
[PubMed][PubMedCentral]

9.

Weinhold L, Wahl S, Pechlivanis S et al (2016) A statistical model for the analysis of beta values in DNA methylation studies. BMC Bioinformatics 17(1):480. https://doi.org/10.1186/s12859-016-1347-4 [PubMed][PubMedCentral]

10. Horvath S (2013) DNA methylation age of human tissues and cell types. Genome Biol 14(10):R115. doi: 10.1186/gb-2013-14-10-r115. Erratum in: Genome Biol. 2015 May 13;16:96. doi: 10.1186/s13059-015-0649-6 [https://doi.org/10.1186/gb-2013-14-10-r115][PubMed][PubMedCentral]

11. Hannum G, Guinney J, Zhao L et al (2013) Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell 49(2):359–367. https://doi.org/10.1016/j.molcel.2012.10.016 [PubMed]

12. Thrush K, Higgins-Chen A, Liu Z et al (2022) R methylCIPHER: a methylation clock investigational package for hypothesis-driven evaluation & research. BioRxiv. https://doi.org/10.1101/2022.07.13.499978

13. Faul JD, Crimmins E, Munro S et al (2023) HRS DNA methylation data—VBS 2016. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/sites/default/files/genetic/HRS_DNAm_OCT2023.pdf. Accessed 22 January 2026

14. Heiss JA, Just AC (2019) Improved filtering of DNA methylation microarray data by detection p-values and its impact on downstream analyses. Clin Epigenetics 11(1):15. https://doi.org/10.1186/s13148-019-0615-3 [PubMed][PubMedCentral]

15. Sahoo K, Sundararajan V (2024) Methods in DNA methylation array dataset analysis: a review. Comput Struct Biotechnol J 23:2304–2325. https://doi.org/10.1016/j.csbj.2024.05.015 [PubMed][PubMedCentral]

16. Alam S, Ayub M, Arora S et al (2023) An investigation of the imputation techniques for missing values in ordinal data enhancing clustering and classification analysis validity. Decis Anal J 9. https://doi.org/10.1016/j.dajour.2023.100341

17. Hannum G, Guinney J, Zhao L et al (2013) Genome-wide methylation profiles reveal quantitative views of human aging rates. Molecular Cell 49(2):359–367. https://doi.org/10.1016/j.molcel.2012.10.016 [PubMed]

18. Levine ME, Lu AT et al (2018) An epigenetic biomarker of aging for lifespan and healthspan. Aging 10(4):573–591. https://doi.org/10.18632/aging.101414 [PubMed][PubMedCentral]

19. McCrory C, Fiorito G, Hernandez B et al (2021) GrimAge outperforms other epigenetic clocks in the prediction of age-related clinical phenotypes and all-cause mortality. J Gerontol A Biol Sci Med Sci 76(5):741–749. https://doi.org/10.1093/gerona/glaa286 [PubMed][PubMedCentral]

20. Belsky DW, Caspi A, Corcoran DL et al (2022) DunedinPACE, a DNA methylation biomarker of the pace of aging. Elife 11:e73420. https://doi.org/10.7554/eLife.73420 [PubMed][PubMedCentral]

21. Arpawong TE, Hernandez B, Potter C et al (2025) Physiological health age (PhysAge): a novel multi-system molecular timepiece predicts health and mortality in older adults. Geroscience https://doi.org/10.1007/s11357-025-01832-1 [PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_2

2. A Standardized Workflow for Single-Cell/Nuclei RNA-Seq Analysis in Aging Research

Clayton Baker[1, 2], Eyael Tewelde[1, 3], Ella Schwab[1, 3] and Bérénice A. Benayoun[1, 3, 4, 5] ✉

(1) Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA
(2) Molecular and Cellular Biosciences Department, USC Dornsife College of Letters, Arts and Sciences, Los Angeles, CA 90089, USA
(3) Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California, Los Angeles, CA 90089, USA
(4) USC Norris Comprehensive Cancer Center, Epigenetics and Gene Regulation, Los Angeles, CA 90033, USA
(5) USC Stem Cell Initiative, Los Angeles, CA 90033, USA

✉ **Bérénice A. Benayoun**
Email: berenice.benayoun@usc.edu

Abstract
Aging is the main risk factor for many chronic diseases, including cancer, neurodegeneration, and immune dysfunction, and is accompanied by extensive changes in gene expression across tissues and cell types. Traditional bulk RNA-sequencing is unable to distinguish whether observed transcriptional changes reflect shifts in underlying cell-type composition of the tissue of interest, or true gene expression changes within specific cell populations. Single-cell and single-nucleus RNA-sequencing (scRNAseq and snRNAseq) are able to overcome this limitation by enabling transcriptomic analysis at cellular resolution. In this chapter, we outline a robust and reproducible computational pipeline for analysis of single-cell transcriptomic datasets in aging research, using a publicly available mouse aging hypothalamus snRNAseq dataset to demonstrate analytical workflows and expected results. We describe an end-to-end workflow starting from raw FASTQ files that includes data processing, cell-type annotation, integration, cell state and population-level comparative analyses, pseudobulking, and differential gene expression analysis. The goal of this chapter is to provide a practical and repeatable computational framework for assessing age-associated transcriptional changes at the cellular level, suitable for both novice and experienced computational biologists.

Keywords Aging – Computational biology – Single-cell – Single-nucleus – 10× Genomics

1 Introduction

Aging is a complex biological process, which is generally described as the gradual decline in physiological function over time, and the leading risk factor for many chronic diseases in mammals, such as immune dysfunction, neurodegenerative disorders, and cancer [1]. It is accompanied by widespread changes in tissue function, cellular composition, and gene expression that occur across multiple organ systems [2]. Understanding and uncovering these age-related changes at a molecular/mechanistic level is critical for developing interventions to promote healthy aging and enable the use of personalized medicine [3].

Single-cell and single-nucleus RNA-sequencing (scRNAseq and snRNAseq) are technologies that measure gene expression at the cellular level, enabling the study of cellular heterogeneity and cell-type-specific transcriptional programs within tissues [4]. Unlike bulk RNA-sequencing, which measures the average gene expression across all cells in a tissue, these methods allow researchers to distinguish whether observed transcriptional differences come from changes in cell-type proportions or from real gene expression

changes within specific cell populations [4]. Additionally, these techniques enable the characterization of rare or previously unrecognized cell populations that may be key for understanding tissue-specific aging or disease-driving processes [4].

Applying scRNAseq and snRNAseq to aging studies enables the identification of cell-type-specific transcriptional programs altered with age [5]. By comparing samples across different ages (young and old, for example), researchers can map how aging impacts gene expression within individual cell populations and across tissues to look for common or unique signatures shared across cell types and tissues [5]. These approaches provide a foundation for uncovering molecular mechanisms of aging and generate insights that are not accessible using conventional bulk sequencing strategies [5].

In this chapter, we used a publicly available hypothalamic snRNAseq dataset of aging in female mice [6] as a reference to demonstrate a complete computational workflow, from raw FASTQ files (accessible via GEO accession GSE188646) to differential gene expression analysis (*see* **Note 15**). The major steps in this workflow include (i) generating a gene-cell count matrix with Cell Ranger, (ii) removing ambient RNA signal using SoupX, (iii) transcriptomic preprocessing and quality control with Seurat, (iv) doublet identification and removal using DoubletFinder and scds, (v) cell type annotation using SingleR, (vi) integration across samples using RPCA, CCA, and Harmony, (vii) cell proportion shift analyses using scProportionTest, (viii) cell type prioritization using Augur, (ix) single cell gene signature scoring with UCell, (x) pseudobulking with muscat, (xi) batch correction and technical noise filtering using SVA, and (xii) cell-type specific differential gene expression analysis with DESeq2 (Fig. 1).

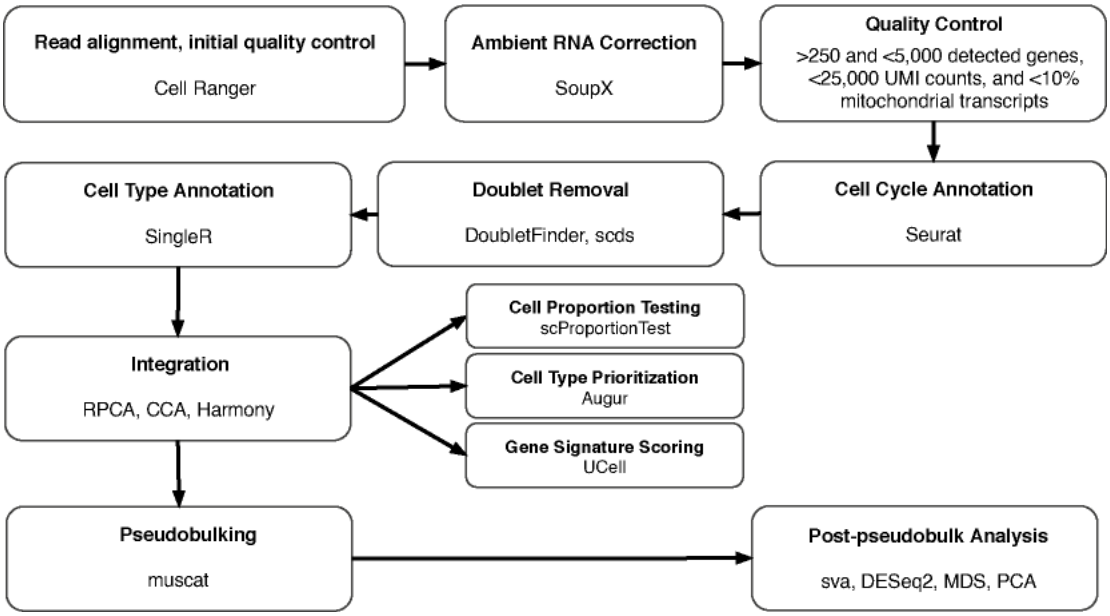


Fig. 1 Schematic overview showing single-cell workflow including preprocessing, single-cell level analysis, and pseudobulking. The pipeline is made of four primary stages: (i) Preprocessing, including read alignment, quality control, cell cycle annotation, doublet removal, and cell type annotation; (ii) Integration of sequencing chemistries; (iii) Single-cell level analysis, featuring cell proportion testing, cell type prioritization, and gene signature scoring; and (iv) post-pseudobulk analysis

Together, this workflow enables drawing robust, reproducible conclusions based on any single-cell/nuclei RNA-sequencing dataset from aging animals.

2 Materials

2.1 Example Dataset

GSE/SRA accession: GSE188646/SRP345673 (*see* **Note 16**)

2.2 Standalone Software

Cell Ranger (v10.0.0) [7] (*see* **Note 14**)
R (v.4.4.2)
R Studio (v.2024.12.1+563)

2.3 R Packages

Augur (v.1.0.3) [8]
celda (v.1.22.1) [9]
celldex (v.1.16.0) [10]
clusterProfiler (v.4.16.0) [11]
clustree (v.0.5.1) [12]
DESeq2 (v.1.48.2) [13]
DoubletFinder (v.2.0.6) [14]
Dplyr (v.1.1.4) [15]
enrichplot (v.1.28.4) [11]
ggplot2 (v.3.5.2) [16]
GSEABase (v.1.70.1) [17]
Harmony (V.1.2.4) [18]
limma (v.3.64.3) [19]
muscat (v.1.20.0) [20]
org.Mm.eg.db (v.3.21.0) [21]
Patchwork (v.1.3.0) [22]
S4Vectors (v.0.44.0) [23]
scales (v.1.4.0) [24]
Scater (v.1.34.1) [25]
scds (v.1.22.0) [26]
scProportionTest (v.0.0.0.9000) [27]
sctransform (v.0.4.2) [28]
Seurat (v.5.4.0) [29]
SeuratObject (v.5.3.0) [23]
SeuratWrappers (v.0.4.0) [23]
SingleCellExperiment (v.1.28.1) [30]
singleCellTK (v.2.16.1) [31]
SingleR (v.2.8.0) [10]
SoupX (v.1.6.2) [32]
sva (v.3.56.0) [33]
UCell (v2.14.0) [34]

2.4 Scripts

cellranger_v10_webb.sh
Hypo_webb_singlecell_workflow_V4.R

---

3 Methods

3.1 File Structure

snRNAseq Read Format

sc/snRNAseq files typically come as FASTQ file. (*see* **Note 17**) Included in the FASTQ files are sequencing quality scores, cDNA sequences, cell barcode information, and unique molecular identifiers (UMIs) (Fig. 2). Typically, the naming convention and corresponding file structure are as follows (see Table 1).

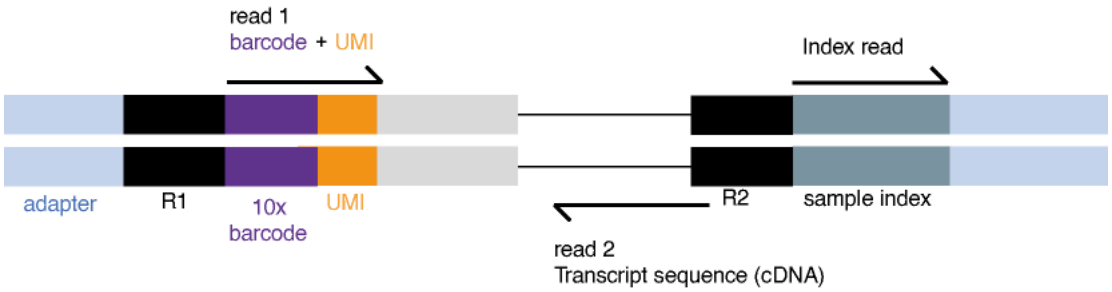


Fig. 2 Architecture of the 10× genomics single-cell sequencing library. Schematic representation of the final library construct, highlighting the arrangement of the adapters, 10× Barcode (cell identification), and Unique Molecular Identifier (UMI) (transcript quantification) on Read 1. The insert contains the cDNA transcript sequence targeted by Read 2, while the Sample Index allows for multiplexed sequencing

Table 1 File structure and contents of snRNAseq FASTQ files

UMIs are a unique tag for every cDNA molecule. This helps to identify PCR duplicates later on. Indices are used to demultiplex samples when a library preparation protocol involves multiplexing.

In order for Cell Ranger to map the FASTQ files, they should conform to the naming conventions of bcl-convert, bcl2fastq, and mkfastq:

[Sample Name]_S1_L00[Lane Number] _[Read Type]_001.fastq.gz

OR

[Sample Name]_S1_[Read Type]_001.fastq.gz

Read Type can be any of the following (*see* **Note 1**):

I1: Sample index read

I2: Sample index read

R1: Read 1

R2: Read 2

4 Read Mapping with Cell Ranger

4.1 Choosing a Reference Genome Version

Before mapping reads, it's important to select an appropriate reference genome that matches your experimental organism and the version to be used in downstream analyses. You can choose to map to a standard reference genome or build a custom reference genome.

Unless there is compelling reason (e.g., needing to quantify a transgene, poor annotation quality of an existing reference, etc.), we recommend using the most recent genome reference for your species of interest (e.g., mm39 for mouse), as they will best capture experimental data in the context of up-to-date knowledge.

Download the Appropriate Pre-Built Reference Genome

The 10× Genomics website maintains genome reference packages [35] can then be extracted to make them usable for read mapping and counting. Although the website requires registration, it is free to access for academic users.

```
# Download reference
wget "https://cf.10xgenomics.com/supp/cell-exp/refdata-gex-GRCm39-2024-
A.tar.gz"
# Extract the reference
tar -xvzf refdata-gex-GRCm39-2024-A.tar.gz
```

4.2 Read Mapping and Counting

Raw sc/snRNAseq data can be processed using Cell Ranger (v10.0.0; 10× Genomics), a software suite designed for the analysis of droplet-based sc/snRNAseq data generated by the Chromium platform. Cell Ranger performs demultiplexing, read alignment, barcode processing, and gene expression quantification to generate cell-by-gene expression matrices (*see* **Note 2**).

The paired-end FASTQ files are provided as input to the `cellranger count` pipeline. Reads will be aligned to the mouse reference genome (here, GRCm39) using an internal spliced aligner optimized for single-cell data. To capture nascent and unspliced transcripts, it is recommended to include intronic reads with the flag—`include-introns=true` during alignment, which is recommended for RNA-poor samples (i.e., nuclei, neutrophils) or cell types with expected high levels of pre-mRNA. Although not strictly necessary, it is good practice to also provide the pipeline with the expected cell number using the flag—`expect-cells` (here, 5000 per library).

For each library, the pipeline produces gene expression matrices both unfiltered (i.e., all droplet barcodes) and filtered (i.e., only droplets that are compatible with cell capture), some summary metrics, and alignment files. This step is run in a Linux shell, usually on a high-performance computing cluster, as it is costly in memory usage. Note that, to our knowledge, the software will not run on macOSX. The recommended resources to run the software efficiently are RAM 64GB memory and 1.5TB storage. Here, we show our usage (32 G memory and 500GB storage), which we found was sufficient to run the cellranger pipeline for the example dataset (and most datasets we have generated in-house in the lab).

```
cellranger count \

--id=SAMPLE_count \

--transcriptome=REF_DIR \

-fastqs=FASTQ_DIR \

--sample=SAMPLE \

--localcores=8 \

---localmem=32 \

--expect-cells=5000 \-

--include-introns=true \

--create-bam=true \

--disable-ui
```

After running Cell Ranger, open the web summary HTML file located in the outs/ directory to assess overall data quality and mapping success. Important metrics to review include the “Estimated Number of Cells,” which are given as a big green numbers and should be reasonably close to the number targeted in the experiment, and the “Mean Reads per Cell,” which for 10× Genomics data is typically recommended to be in the range of 25,000–50,000 reads per cell, depending on the assay. The “Barcode Rank Plot” is another key quality-control visualization. A high-quality dataset will show a clear knee and cliff, indicating a sharp drop-off between true cell-associated barcodes and background barcodes. This clear separation reflects good cell integrity and reliable cell calling.

---

5 Ambient RNA Contamination Correction

From this point on, subsequent analytical steps are performed using the R statistical software, on any machine with sufficient memory to handle the dataset. To note, in our experience, a dataset with up to 100–

200 K cells can be readily analyzed on a macOSX workstation with 64 GB of RAM. All the steps hereafter are run in R using the Rstudio IDE for analytical ease.

The first key preprocessing step to analysis of sc/snRNAseq datasets from aging libraries will be to reduce the impact of contamination from ambient RNA inherent to droplet-based sc/snRNAseq methods [36]. To address this experimental caveat, computational methods can be used to identify ambient RNA molecules, estimate their relative amount, and generate corrected expression matrices. Both sample-level (SoupX) [32] and droplet-level (DecontX) [37] methods exist for this purpose, both with unique advantages and drawbacks. In this example, we demonstrate the use of sample-level ambient RNA correction, using the SoupX algorithm [32].

For each sample, load raw count matrices directly from Cell Ranger outs/directories into R, which include both filtered and unfiltered droplet matrices quantified from FASTQ files. Using these matrices, SoupX estimates the overall fraction of ambient RNA in the library and generates sample-wise corrected estimated gene expression counts (Fig. 3; *see* **Note 3**). Corrected counts must then be rounded to yield integers, a step required to maintain compatibility with downstream single-cell analysis tools that expect integer count values. This step helps ensure that downstream analyses reflect true cell-specific gene expression rather than background RNA contamination.

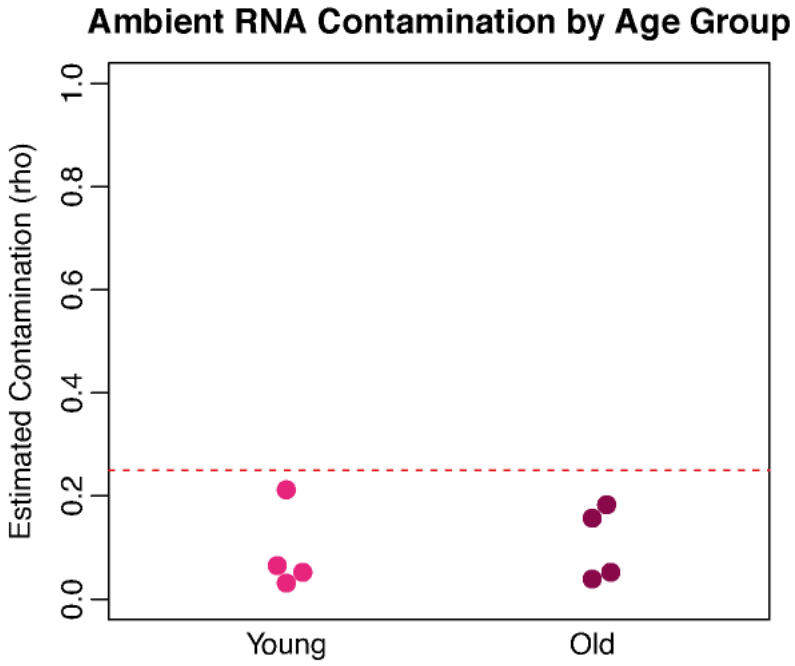


Fig. 3 A beeswarm plot of *rho* values showing estimated ambient RNA contamination for each animal/library grouped by age (each dot represents an individual library). The horizontal dashed line indicates the commonly used threshold of 0.25; all samples fell below this cutoff and were thus used for downstream analysis. Dark pink indicates old females (OF) and light pink indicates young females (YF)

```
read 10x libraries cell ranger gene barcode matrices for SoupX
############################################################
cellRanger.YF.data.1  <- load10X("/path_to_counts/counts/Young1_count/outs")
# estimate RNA soup from bottom 10% droplets
############################################################
cellRanger.YF.data.1 <- autoEstCont(cellRanger.YF.data.1)
# adjust counts based on RNA soup (and round)
############################################################
out.YF.data.1 <- adjustCounts(cellRanger.YF.data.1, roundToInt=TRUE)
```

6 Creating a Seurat Object with all Experimental Samples

Next, we create Seurat objects for each library from SoupX-corrected counts using `CreateSeuratObject` retaining genes detected in at least one cell (`min.cells = 1`) to allow for downstream quality control. Subsequently, library-level Seurat objects should be merged with unique sample identifiers appended to cell barcodes using the `add.cell.ids` parameter (*see* **Note 19**).

```
# get seurat objects
seurat.YF.data.1 <- CreateSeuratObject( out.YF.data.1, min.cells = 1)
# Merge cleaned seurat objects
hypo.webb <- merge(seurat.YF.data.1,

y = c(seurat.YF.data.2, …),

add.cell.ids = c("YF1",

"YF2", …),

project = "hypo_webb_Aging")
```

7 Populate Relevant Metadata in the Seurat Object

For subsequent analytical steps, it is critical to assign and store library- and cell-level metadata. For this dataset, each cell was labeled according to its originating library (i.e., one animal per library), the 10× Genomics chemistry used for library preparation (i.e., v2 vs v3), and its biological condition (i.e., young vs old). Incorporating cell-level metadata is essential for downstream quality control, visualization, and analysis. These annotations enable library-wise quality control, inform batch-correction and data integration (*see* **Note 4**).

```
############################################################
# Create a per-cell library identifier
# This assigns each cell to its original sequencing library
# based on patterns in the cell barcode names.
############################################################
Library <- rep("NA", length(colnames(hypo.webb@assays$RNA)))
Library[ grep("YF1", colnames(hypo.webb@assays$RNA))] <- "YF1"
# Convert to a data frame and ensure rownames match cell barcodes
Library <- data.frame(Library)
rownames(Library) <- colnames(hypo.webb@assays$RNA)
# Add library information to the Seurat object as metadata
hypo.webb <- AddMetaData(object = hypo.webb,

metadata = as.vector(Library),

col.name = "Library")
```

8 Single Cell Quality Control

8.1 Initial Quality Control

Initial QC is then performed to remove low-quality cells/nuclei and help account for technical artifacts. Several standard metrics are useful to examine in single-cell datasets to filter out low-quality cells (i.e., dead/dying cells, compromised cells/nuclei, etc): (i) the number of detected genes, (ii) the number of RNA molecules (UMIs), and (iii) the percentage of UMI derived from the mitochondrial genome. The first two

metrics are automatically calculated by Seurat upon object creation, while the third one must be manually calculated for each dataset based on the gene annotation.

For the 10× Genomics provided reference annotations (and generally Ensembl-derived standard mouse and human genome references), this calculation is made easy by the existence of the "mt" prefix in mitochondrial gene names. To note additional steps may be required when using custom, non-standard annotations for this step. Here, we calculate the percentage of mitochondrial transcripts per cell using a regular expression to identify mitochondrial genes (^mt-) and the `PercentageFeatureSet` Seurat function.

```
### QC on mitochondrial reads
hypo.webb[["percent.mito"]] <- PercentageFeatureSet(hypo.webb, pattern =
"^mt-")
```

For QC filtering, cells are then filtered to remove cells likely to represent artifacts (e.g., with abnormally high or low gene and UMI counts, abnormally high mitochondrial gene-derived UMIs, etc.). Specific filtering thresholds are going to vary based on each dataset due to both technical features (e.g., chemistry, sequencing saturation, etc) and biological features (e.g., metabolically active cells with higher mitochondrial content, etc). Although there are rough guidelines that can be followed (Table 2), these thresholds are best evaluated and derived from the specific dataset analyzed (Fig. 4A–C).

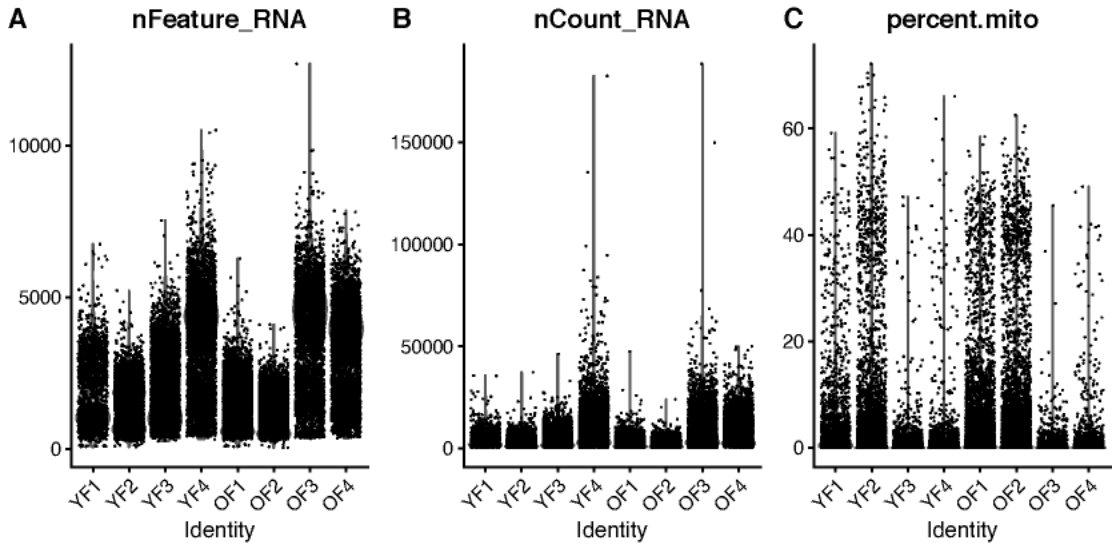


Fig. 4 Violin plots showing the distribution of (**A**) number of features (nFeature_RNA), (**B**) UMI counts (nCount_RNA), and (**C**) percent mitochondrial RNA (percent.mito) across all cells. These plots provide a visual assessment of cell quality and variability, allowing dataset-specific QC thresholds to be determined

Table 2 Quality control metrics and threshold considerations for sc/snRNAseq preprocessing

```
hypo.webb <- SetIdent(hypo.webb, value = "Library")
### QC plots
VlnPlot(object = hypo.webb, features = c("nFeature_RNA", "nCount_RNA",
"percent.mito"), ncol = 3)
plot1 <- FeatureScatter(hypo.webb, feature1 = "nCount_RNA", feature2 =
"percent.mito")
plot2 <- FeatureScatter(hypo.webb, feature1 = "nCount_RNA", feature2 =
"nFeature_RNA")
plot1 + plot2
```

In this case, based on QC metric distributions, we retained cells with 250–5000 genes, less than 25,000 total counts, and <10% mitochondrial transcripts. These thresholds should help remove low-quality, dead/dying cells and egregious likely doublets, while preserving most of the biologically meaningful cellular diversity.

```
# get clean cells (remove dead or dying cells and obvious doublets based on
counts)
hypo.webb <- subset(hypo.webb, subset = nFeature_RNA > 250 & nFeature_RNA <
5000 & percent.mito < 10 & nCount_RNA < 25000)
```

8.2 Cell Cycle Annotation

To account for variation in gene expression driven by cell cycle state, it can be useful to infer the specific phase of the cell cycle for each cell in the dataset using the Seurat cell cycle scoring algorithm. Prior to cell cycle scoring, RNA expression data are normalized using Seurat's `NormalizeData` function to correct for cell-to-cell differences in library size and to render expression values comparable across cells. Using predefined S phase and G2/M phase gene sets originally described by Tirosh et al. [38], each cell was assigned S and G2/M scores, and a corresponding cell cycle phase (G1, S, or G2/M). These curated marker genes are provided as external input files on the Seurat website and are used to compute per-cell S and G2/M scores based on relative expression levels.

After computation, scores and phase assignments are stored as metadata for use in downstream normalization and regression. Including these metrics ensures that cell cycle-associated variability does not confound analyses of other biological signals.

```
# Compute Cell-Cycle Scores and phase assignments
DefaultAssay(hypo.webb) <- "RNA"
Layers(hypo.webb[["RNA"]])
hypo.webb <- NormalizeData(hypo.webb, assay = "RNA", verbose = FALSE)
hypo.webb <- CellCycleScoring(object = hypo.webb, s.features = s.genes,
g2m.features = g2m.genes, set.ident = F)
```

8.3 Normalization, Dimensionality Reduction, and Clustering

After initial QC, we next normalize the data and perform dimensionality reduction and clustering to identify transcriptionally distinct cell populations and visualize cellular relationships. Gene expression counts were normalized and variance-stabilized using SCTransform [28], regressing out technical effects (UMI counts, detected genes, mitochondrial reads) and cell-cycle variation. This approach models sequencing depth–dependent noise using regularized negative binomial regression. Compared with standard log-normalization, SCTransform improves comparability across cells, enhances the detection of biologically meaningful variation, and provides robust input for downstream dimensionality reduction, clustering, and visualization.

```
### SCTransform normalization and regression of technical effects
hypo.webb <- SCTransform(object = hypo.webb, vars.to.regress =
c("nFeature_RNA", "nCount_RNA", "percent.mito", "Phase"))
```

A key step in analyzing single cell/nuclei data is dimensionality reduction for exploratory data analysis. As a first step, it is crucial to determine the dimensionality of your specific dataset in order to calculate proper embedding for visual representation and cell clustering. To determine dataset dimensionality, we first perform principal component analysis (PCA) with an artificially high dimension number, then identify the principal components that (i) explain the majority of the variance in the dataset (exceeding 90% cumulative variance while each contributing less than 5% individually) and (ii) the point where the percent change in variation between the consecutive PCs is less than 0.1%, as being informative. The dimensionality of the dataset is the smallest number of PCs satisfying these conditions (Fig. 5).

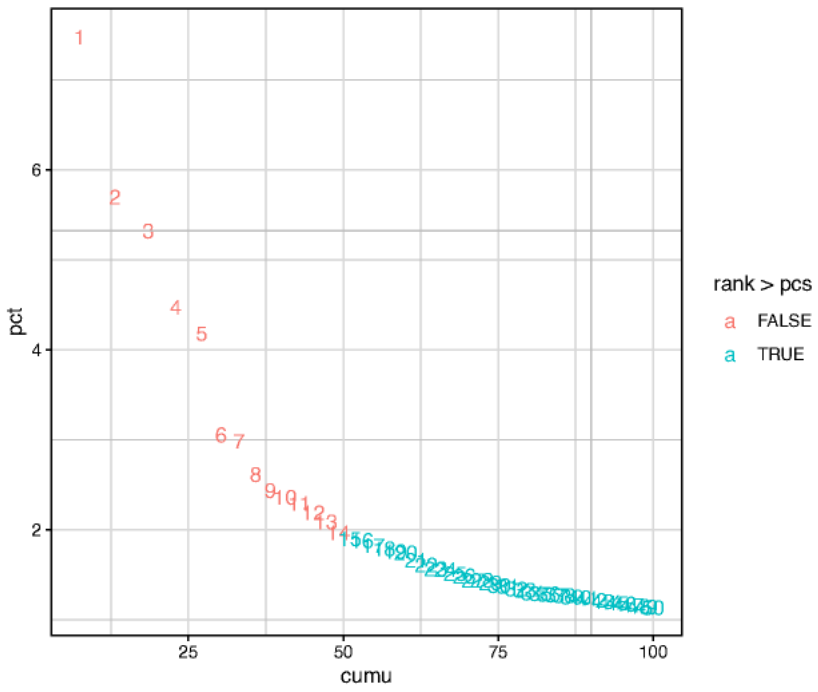


Fig. 5 Elbow plot showing the proportion of variance explained by principal components (PCs). X-axis represents the cumulative proportion of variance explained as PCs are added, while Y-axis shows the variance contribution of individual PCS. The dimensionality of the dataset resulted in 14 PCs being retained for downstream analyses

```
### Run PCA
hypo.webb <- RunPCA(hypo.webb, npcs = 50, verbose = FALSE)
### Visualise PCA (Elbow plot)
ElbowPlot(hypo.webb, ndims = 50)
### Determine number of informative PCs
pct <- hypo.webb[["pca"]]@stdev / sum(hypo.webb[["pca"]]@stdev) * 100
cumu <- cumsum(pct)
co1 <- which(cumu > 90 & pct < 5)[1]
co2 <- sort(which((pct[-length(pct)] - pct[-1]) > 0.1), decreasing = TRUE)
[1] + 1
pcs <- min(co1, co2)
### Visualize PC variance thresholds
# Create a dataframe with values
plot_df <- data.frame(pct = pct,
     cumu = cumu,
     rank = 1:length(pct))
# Elbow plot to visualize
ggplot(plot_df, aes(cumu, pct, label = rank, color = rank > pcs)) +

geom_text() +

geom_vline(xintercept = 90, color = "grey") +

geom_hline(yintercept = min(pct[pct > 5]), color = "grey") +

theme_bw()
```

Once data dimensionality is determined, here with 14 dimensions, we then construct a shared nearest-neighbor graph, enabling unsupervised clustering of cells based on transcriptional similarity. Finally, UMAP is applied for non-linear dimensionality reduction to visualize cellular relationships in low-dimensional space. Clustering and UMAP embeddings are inspected globally and color-coded by library, sequencing chemistry, and experimental conditions to assess technical effects/structure (Fig. 6; *see* **Note 5**).

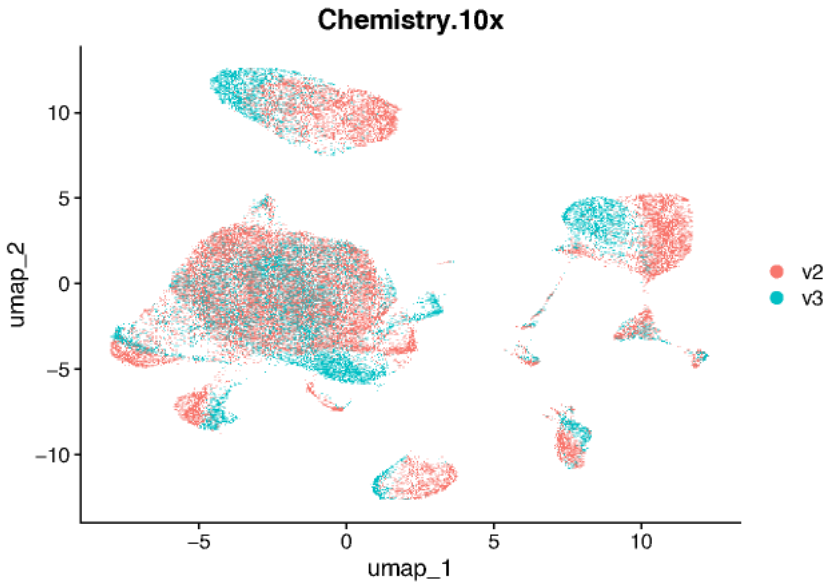


Fig. 6 UMAP visualization of single-cell transcriptomes following clustering based on a shared nearest-neighbor graph constructed from principal components (PCs). Cells are color-coded by sequencing chemistry (v2 vs v3) to inspect potential technical effects. Clusters derived from unsupervised analysis show clear separation between v2 and v3 chemistries for most cell types, demonstrating both biological and technical structure in the dataset (*see* **Note 20**).

```
### Run dimensionality reduction analysis (UMAP, neighbors, and clustering)
hypo.webb <- RunUMAP(hypo.webb, dims = 1:pcs, verbose = FALSE)
hypo.webb <- FindNeighbors(hypo.webb, dims = 1:pcs, verbose = FALSE)
hypo.webb <- FindClusters(hypo.webb, verbose = FALSE)
### UMAP visualizations
DimPlot(hypo.webb, reduction = "umap", group.by = "Chemistry.10x", raster =
TRUE)
```

8.4 Doublet Removal

Doublets arise when two cells are inadvertently captured and sequenced under the same cell barcode, causing the resulting gene expression profile to reflect a mixture of two distinct cell profiles. In fact, the proportion of doublet increases linearly with increased capture rate, and expected rates are provided by 10× Genomics in their user manual for the corresponding chemistry (Fig. 7A, B; *see* **Note 6**).

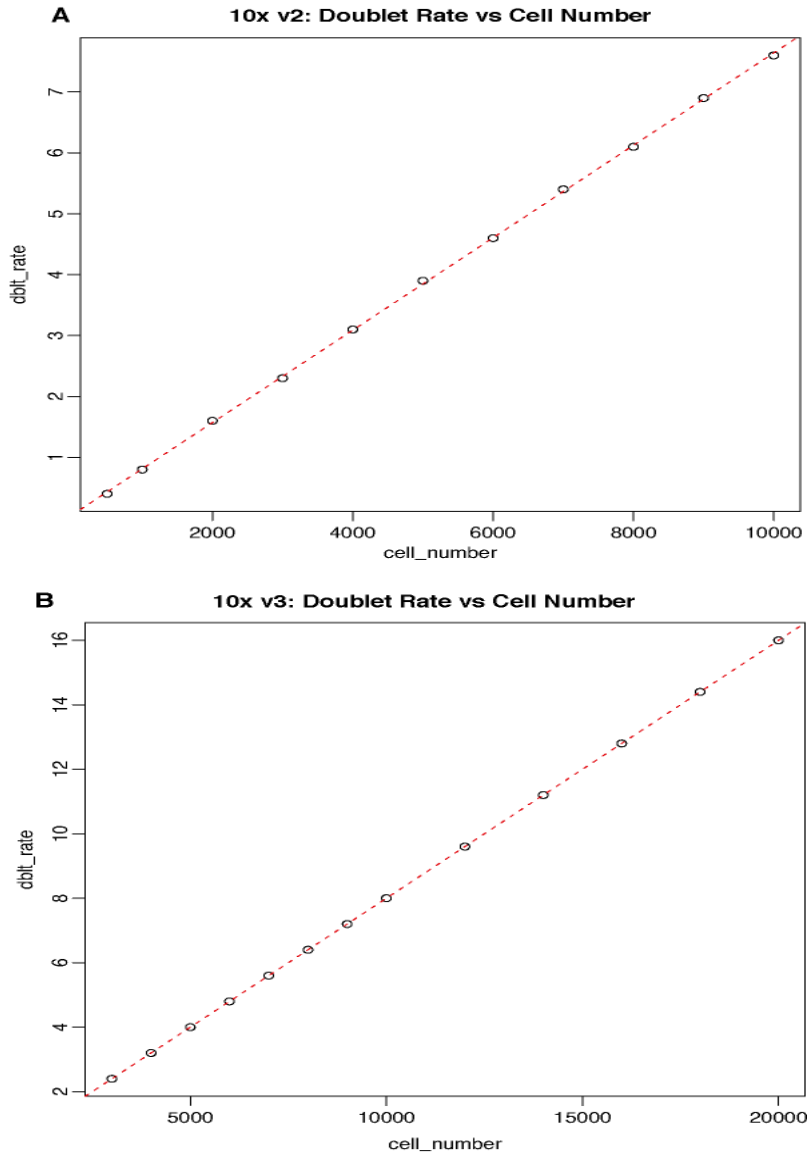


Fig. 7 Predicted doublet rates for 10× Genomics 3′ (**A**) v2 and (**B**) v3 chemistries as a function of captured cell number. Red dashed lines indicate the linear trend of doublet formation

```
# Predicted doublet rates for 10X v2 chemistry
pred.10x.dblt.v2 <- data.frame(
 "cell_number" = c(500,1000,2000,3000,4000,5000,6000,7000,8000,9000,10000),
 "dblt_rate" = c(0.4,0.8,1.6,2.3,3.1,3.9,4.6,5.4,6.1,6.9,7.6))
pred_dblt_lm.v2 <- lm(dblt_rate ~ cell_number, data = pred.10x.dblt.v2)
# Plot predicted doublet rates for 10X v2 chemistry
plot(dblt_rate ~ cell_number,
  data = pred.10x.dblt.v2,
  main = "10x v2: Doublet Rate vs Cell Number")
abline(pred_dblt_lm.v2, col = "red", lty = "dashed")
# Predicted doublet rates for 10X v3 chemistry
pred.10x.dblt.v3 <- data.frame(
 "cell_number" =
c(3000,4000,5000,6000,7000,8000,9000,10000,12000,14000,16000,18000,20000),
```

```
 "dblt_rate" = c(2.4,3.2,4.0,4.8,5.6,6.4,7.2,8.0,9.6,11.2,12.8,14.4,16.0))
pred_dblt_lm.v3 <- lm(dblt_rate ~ cell_number, data = pred.10x.dblt.v3)
#plot predicted doublet rates for 10X v3 chemistry
plot(dblt_rate ~ cell_number,
 data = pred.10x.dblt.v3,
 main = "10x v3: Doublet Rate vs Cell Number")
abline(pred_dblt_lm.v3, col = "red", lty = "dashed")
```

Doublets are technical artifacts that can confound downstream analyses and generate misleading biological conclusions, making their identification and removal an essential step during preprocessing. To remove potential doublets, we use a stringent approach, annotating cells using two complementary methods: *DoubletFinder* and *scds*. Since their underlying models are distinct, they will capture different types of doublets. For optimal stringency (preferring to discard real cells over retaining artefactual cells), only cells classified as singlets by both methods are retained for downstream analyses.

Importantly, doublet finding algorithms **must be run independently for each sample**, not on the merged object. Thus, this section starts by splitting the data by sample of origin—doublet call information will be transferred back to the merged object after predictions are completed.

```
# Split Seurat object by original 10X library to identify true doublets
seurat.list <- SplitObject(hypo.webb, split.by = "Library")
# Assign predicted doublet rates based on chemistry and cell numbers
pred.dblt.rate <- numeric(length(seurat.list))
names(pred.dblt.rate) <- names(seurat.list)
for (i in seq_along(seurat.list)) {
 ncells <- 1.15 * ncol(seurat.list[[i]])
 chem <- unique(seurat.list[[i]]@meta.data$Chemistry.10x)
 if (chem == "v2") {
 pred.dblt.rate[i] <- predict(pred_dblt_lm.v2, data.frame(cell_number =
ncells)) / 100
 } else if (chem == "v3") {
  pred.dblt.rate[i] <- predict(pred_dblt_lm.v3, data.frame(cell_number =
ncells)) / 100
 } else {
  stop(paste("Unknown 10x chemistry:", chem))
 }
}
```

Doublet Finder

DoubletFinder identifies putative doublets by modeling expected doublet formation based on cell loading, chemistry version, and library size (*see* **Note 18**). The workflow involves estimating the optimal classification parameter (neighborhood size; pK), calculating homotypic doublet proportions, and scoring each cell using a simulated doublet population.

Here, doublet identification is performed independently for each sample. Principal components used for clustering are reused for doublet identification to ensure consistency in transcriptional space. For each sample, an empirical parameter sweep is conducted to identify the optimal pK by maximizing the BC metric (the bimodality coefficient, which quantifies the degree of separation between artificial doublets and real cells).

```
#loop over samples
for (i in 1:length(seurat.list)) {
 # Parameter sweep to find optimal pK
 sweep.res.list <- paramSweep(seurat.list[[i]], PCs = 1:pcs, sct = TRUE)
```

```
 sweep.stats <- summarizeSweep(sweep.res.list, GT = FALSE)
 bcmvn   <- find.pK(sweep.stats)
 pk.hypo <- as.numeric(as.character(
  bcmvn[as.numeric(bcmvn$pK[bcmvn$BCmetric == max(bcmvn$BCmetric)]), "pK"]))
 # Estimate homotypic doublet proportion
 homotypic.prop <- modelHomotypic(seurat.list[[i]]@meta.data$seurat_clusters)
 # Calculate expected doublets
 nExp_poi <- round((pred.dblt.rate[i]) *
length(seurat.list[[i]]@meta.data$Library))
```

DoubletFinder is applied using SCT-normalized data, and cells are classified as singlets or doublets. Expected doublet numbers were calculated based on sample-specific cell counts and platform-recommended doublet rates.

```
# Run DoubletFinder
 seurat.list[[i]] <- doubletFinder(
  seurat.list[[i]], PCs = 1:pcs, pN = 0.25, pK = pk.hypo, nExp = nExp_poi,
sct = TRUE)
 # Rename classification column for clarity
 my.DF.res.col <- colnames(seurat.list[[i]]@meta.data)[
  grep("DF.classifications_0.25", colnames(seurat.list[[i]]@meta.data))]
 colnames(seurat.list[[i]]@meta.data)[
  grep("DF.classifications_0.25", colnames(seurat.list[[i]]@meta.data))] <-
"DoubletFinder"
}
```

Distribution of doublets are then visualized with UMAP plots and doublet classifications are stored in the Seurat metadata and used for downstream filtering (Fig. 8A).

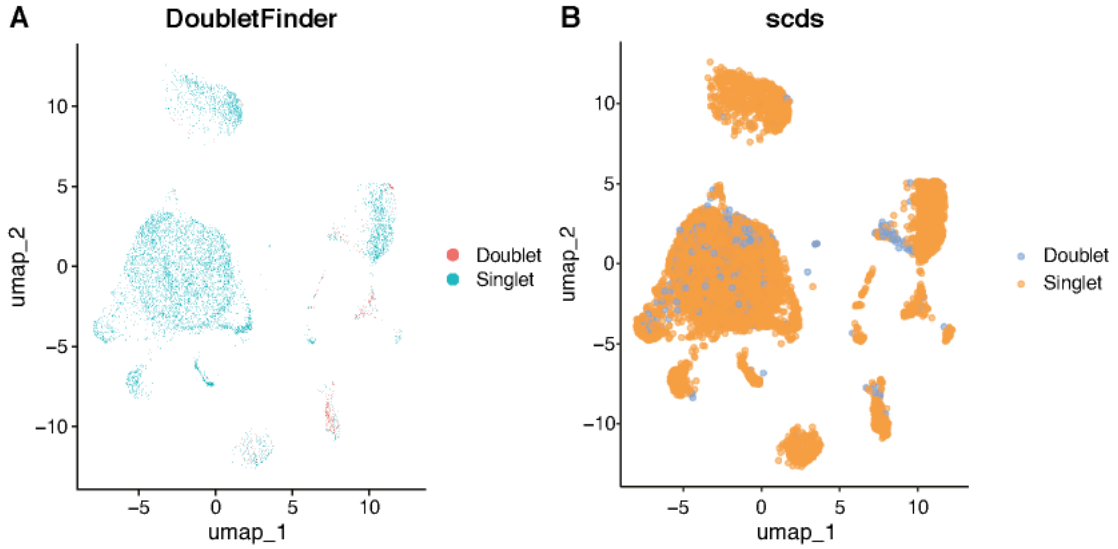


Fig. 8 Side-by-side UMAP comparison of doublet detection methods for representative library (animal-Young Female 2). **A**, Doublets and singlets identified using DoubletFinder; (**B**) doublets and singlets identified using scds

```
# UMAP plots for each sample
for (i in 1:length(seurat.list)) {
 print(DimPlot(seurat.list[[i]], reduction = "umap", group.by =
"DoubletFinder", raster = TRUE))
```

```
}
# Merge back all samples into a single Seurat object
hypo.webb.singlets.doublets.annot <- merge(
 seurat.list[[1]],
 y = c(seurat.list[[2]],…),
 project = "hypo_webb_snRNAseq")
```

Single-Cell Doublet Scoring (scds)

Single-cell doublet scoring (scds) is another doublet-detection framework based on gene-expression structure, rather than neighborhood density. The scds framework applies a hybrid approach combining cxds (co-expression of gene pairs) and bcds (binary classification using artificial doublets) to independently predict doublets.

First, to ensure that scds doublet classification is performed independently of DoubletFinder, pANN (proportion of artificial nearest neighbors) columns are recommended to be removed from the cell metadata before analysis. Next, we convert Seurat objects to `SingleCellExperiment` format.

```
# Remove 10X lane-specific pANN columns
hypo.webb.singlets.doublets.annot@meta.data <-
hypo.webb.singlets.doublets.annot@meta.data[,-grep("pANN",
colnames(hypo.webb.singlets.doublets.annot@meta.data))]
# Convert Seurat objects to SingleCellExperiment for scds
seurat.list.scds <- lapply(seurat.list, as.SingleCellExperiment)
```

Doublet scores can then be computed using the hybrid cxds–bcds method. For each sample, the expected number of doublets will be estimated based on the predicted doublet rate for the corresponding sequencing library. Cells will be ranked by hybrid doublet score, and the top-ranked cells corresponding to the expected number of doublets will be classified as doublets, with all remaining cells labeled as singlets.

```
# Run hybrid doublet scoring (cxds + bcds)
for (i in 1:length(seurat.list.scds)) {
# Annotate doublets
 seurat.list.scds[[i]] <- cxds_bcds_hybrid(seurat.list.scds[[i]])
# Assign top predicted doublets based on expected rate
 n.db <- round((pred.dblt.rate[i]) * ncol(seurat.list.scds[[i]]))
 srt.db.score <- sort(seurat.list.scds[[i]]$hybrid_score, index.return =
TRUE, decreasing = TRUE)
 seurat.list.scds[[i]]$scds <- "Singlet"
 seurat.list.scds[[i]]$scds[srt.db.score$ix[1:n.db]] <- "Doublet"
}
```

Cells labeled as singlets or doublets can be visualized with UMAP, and these annotations can then be mapped back onto the Seurat object (Fig. 8B).

```
# UMAP plots for scds classification
for (i in 1:length(seurat.list.scds)) {
 p <- plotReducedDim(seurat.list.scds[[i]], dimred = "UMAP", colour_by =
"scds", raster = TRUE)
 plot(p)

dev.off()

}
# Map scds hybrid annotations back to Seurat object
```

```
hypo.webb.singlets.doublets.annot@meta.data$scds_hybrid <- NA
for (i in 1:length(seurat.list.scds)) {
 hypo.webb.singlets.doublets.annot@meta.data[
  colnames(seurat.list.scds[[i]]), "scds_hybrid"] <-
seurat.list.scds[[i]]$scds
}
```

Integrating Doublet Predictions and Filtering Doublets

After predicting doublets using both DoubletFinder and scds, we compare doublet predictions across all cells and within each library (animal) by generating cross tabulations (tables). This allows us to assess the level of agreement between the two independent doublet-detection methods and whether disagreement or doublet rates vary by library (animal) as expected (*see* **Note 7**). Comparing doublet calls between DoubletFinder and scds acts as an extra quality control step to ensure that both methods are performing as expected.

Note that since the two methods use different assumptions for doublet identification, it is expected that a proportion of droplets will be annotated as doublets only by one of the methods (in our experience, 20%–40% of doublet calls are method specific). Depending on the desired level of stringency, once can either decide to (i) consider as a doublet any droplet called as such by either method (union), or (ii) only label as a doublet droplets called as such by both method (intersection). The decision will be based on whether it is preferred to be more conservative in cell identification (discarding potentially real cells at the cost of making sure all doublets are eliminated from analyses; our preference) or less conservative (only discard *bona fide* doublets, with the risk of retaining some real doublets in downstream analyses).

```
############################################################
# C. Merge and summarize doublet findings
# Compare doublet calls from DoubletFinder and scds (hybrid)
############################################################
# Generate a global contingency table comparing doublet
# classifications from the two methods across all cells
table(
 hypo.webb.singlets.doublets.annot@meta.data$DoubletFinder,
 hypo.webb.singlets.doublets.annot@meta.data$scds_hybrid)
############################################################
# Summarize doublet overlap within each sequencing library
# This allows assessment of library-specific doublet rates
############################################################
by(
 hypo.webb.singlets.doublets.annot@meta.data,
 hypo.webb.singlets.doublets.annot@meta.data$Library,
 function(df) {
  table(df$DoubletFinder, df$scds_hybrid)
})
```

We next create a combined DoubletCall metadata column by labeling a cell as a doublet if either DoubletFinder or scds predicts it to be a doublet, using the more conservative "union" approach (*see* **Note 8**). A useful quality control measure at this point in the workflow is to summarize the relative doublet abundance in your dataset (percentage of doublets within total cells). This should roughly match the expected proportions in Fig. 7, as these numbers are used to tell the algorithm how many doublets to call.

```
############################################################
# Define a unified doublet call using a conservative union
# A cell is labeled as a doublet if *either* method predicts
```

```
# it as a doublet
############################################################
hypo.webb.singlets.doublets.annot@meta.data$DoubletCall <-
 ifelse(
bitOr(
hypo.webb.singlets.doublets.annot@meta.data$DoubletFinder == "Doublet",
  hypo.webb.singlets.doublets.annot@meta.data$scds_hybrid == "Doublet") > 0,
 "Doublet",
 "Singlet")
# Summarize final doublet versus singlet counts
table(hypo.webb.singlets.doublets.annot@meta.data$DoubletCall)
# Save doublet-annotated object
save(
 hypo.webb.singlets.doublets.annot,
 file = paste0("10x_webb_hypo_Seurat_object_with_AnnotatedDoublets.RData"))
```

We then filter out cells labeled as doublets in the DoubletCall metadata column to retain only high-confidence singlets, which will be used for all downstream preprocessing and analyses.

```
############################################################
# Remove doublets and retain only high-confidence singlets
# This filtered object will be used for downstream analyses
############################################################
hypo.webb.annot.singlets <- subset(
 hypo.webb.singlets.doublets.annot,
 subset = DoubletCall %in% "Singlet")
# Inspect the resulting singlet-only Seurat object
hypo.webb.annot.singlets
```

8.5 Singlet Quality Control

Dimensionality Reduction, Clustering, and Cell Type Marker Visualization of Singlets

Following doublet removal, we carry out clustering and visualization on the singlet-only object. Dimensionality reduction is carried out using the previously determined number of principal components (see above). A key parameter in the `FindClusters` function is resolution, which governs clustering granularity: higher values generate more, smaller clusters, while lower values produce fewer, bigger clusters. Generally, a resolution between 0.2 and 0.5 captures broad clusters representing major cell types; 0.6–1.0 provides moderate granularity for identifying subpopulations within major populations; and values >1.0 yield high granularity, potentially revealing fine subclusters (rare cell types) at the cost of over-splitting of biologically similar cells. For this workflow, we select a resolution of 0.5, as we aim to focus on major cell types rather than finer subpopulations. Different resolutions can be evaluated using marker gene expression and UMAP visualizations to identify the clustering that is most biologically interpretable and relevant to the research question (see below).

At this stage, a key sanity check is to generate a UMAP with cells grouped by their DoubletCall annotation. Because all doublets were removed in the previous step, the resulting UMAP should display only one group, singlets. If any cells are still labeled as doublets, it is important to review the previous steps in the workflow and code for potential errors before proceeding with downstream analyses.

```
############################################################
# 1. Run UMAP, construct neighbor graph, and perform clustering
############################################################
hypo.webb.annot.singlets <- RunUMAP(hypo.webb.annot.singlets, dims = 1:pcs)
```

```
hypo.webb.annot.singlets <- FindNeighbors(hypo.webb.annot.singlets, dims =
1:pcs)
hypo.webb.annot.singlets <- FindClusters(object = hypo.webb.annot.singlets,
resolution = 0.5)
# Quick check: UMAP colored by doublet call (sanity check)
DimPlot(hypo.webb.annot.singlets, reduction = "umap", group.by =
"DoubletCall", raster = TRUE)
```

To corroborate the clustering results and ensure that distinct cell types are appropriately separated, we next examine the expression of key canonical marker genes for each expected cell type (for the brain, these include neurons, microglia, astrocytes, and oligodendrocytes). Here, we show the UMAP visualization of neurons as an example, which allows us to assess whether clusters equate to biologically meaningful cell types (*see* **Note 9**). By evaluating marker expression across clusters, we can confirm that clusters reflect the expected biology, identify potential misannotations, which will become important later in the workflow, and gain confidence in the integrity of downstream analyses (Fig. 9A–D). (Fig. 9A–D). This strategy can be applied to any sc/snRNAseq dataset by selecting marker genes relevant to the system or tissue being studied. For example, in bone marrow, major cell types include hematopoietic stem and progenitor cells, monocytes, neutrophils, T cells, and B cells.

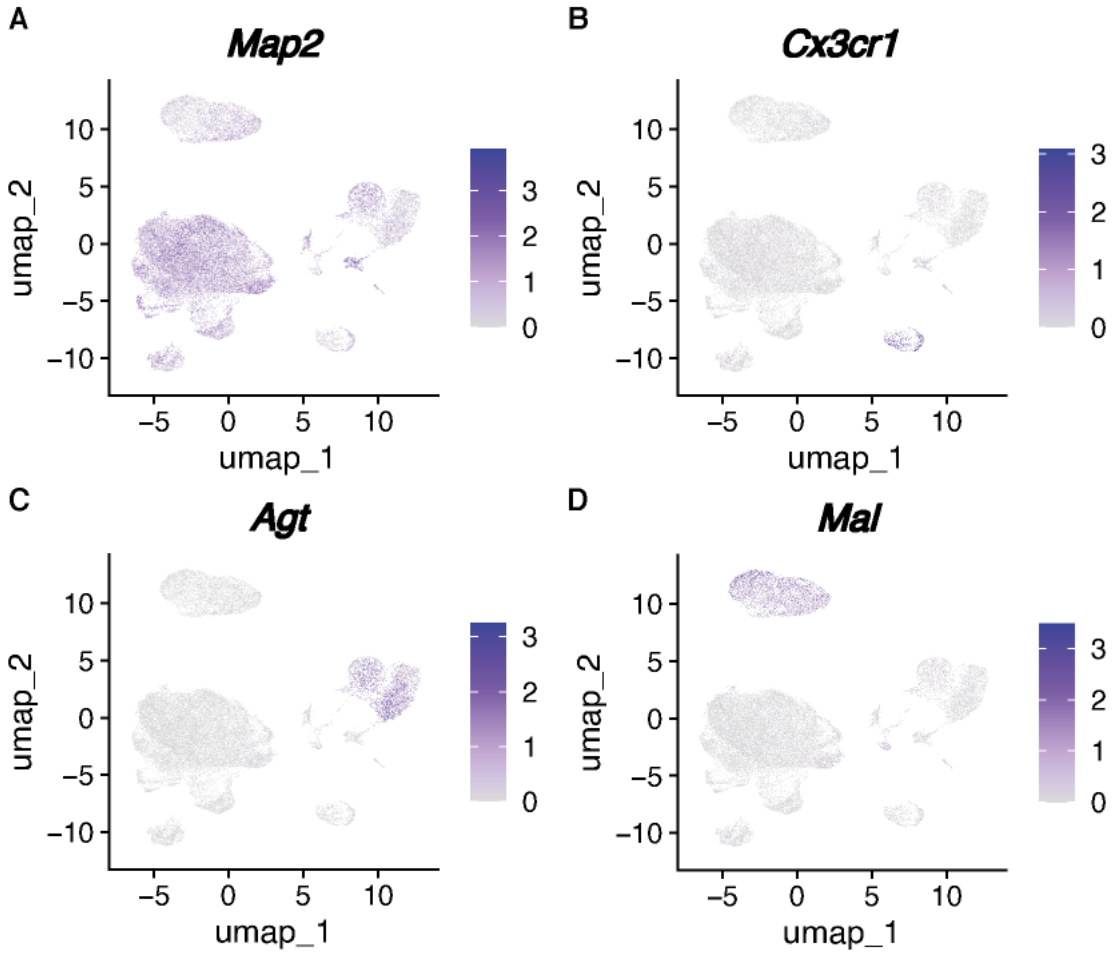


Fig. 9 UMAP plots showing expression of canonical marker genes for major brain cell types. **A,** *Map2* (neurons), (**B**) *Cx3cr1* (microglia), (**C**) *Agt* (astrocytes), and (**D**) *Mal* (oligodendrocytes). Expression patterns across clusters show that the unsupervised clustering captures biologically meaningful populations

##########

```
# 2. Visualize expression of key marker genes to validate clustering (do
cell types # separate?)
# Separate canonical markers for neurons, microglia, astrocytes, and
oligodendrocytes
############################################################
neuron.markers <- c("Map2", "Slc17a7", "Gad1", "Snap25")
microglia.markers <- c("P2ry12", "Ptprc", "Tmem119", "Cx3cr1")
astrocyte.markers <- c("Agt", "Gfap", "Aldh1l1")
oligo.markers <- c("Olig1", "Cldn11", "Mal", "Ptgds")
# Save FeaturePlots for neuron markers to a PDF
# Generate individual FeaturePlots
plots <- lapply(neuron.markers, function(gene) {
 FeaturePlot(hypo.webb.annot.singlets, features = gene, raster = TRUE)
})
# Combine all four plots into a single grid
combined_plot <- wrap_plots(plots, ncol = 2) # 2 columns, 2 rows
# Display the combined plot
combined_plot
# NOTE: Repeat the same process for microglia, astrocytes, and
oligodendrocytes
# using their respective marker sets (microglia.markers, astrocyte.markers,
# oligo.markers)
```

---

9 Cell Type Annotation and Integration

9.1 Annotating Singlet Cells with SingleR

After validating the quality of the singlet-only dataset, we next employ automated cell-type annotation. Cell-type annotation can be conducted using one of two approaches: (i) marker-based or (ii) reference-based. In addition, these strategies can also be combined to increase confidence in final annotations.

With marker-based annotation, cell identities are assigned based on the expression of predefined canonical marker genes, commonly curated from the public databases or literature (e.g., PanglaoDB [39]). Tools such as scSorter [40] and scCATCH [41] implement this marker-based strategy. Marker-based approaches are highly interpretable and work well for well-characterized tissues, in well-characterized species, with clearly defined markers, but their accuracy can be limited when cell states are continuous, when marker specificity is low or in less studied species. With reference-based annotation, each cell is compared to well-annotated reference transcriptomic profiles on pure cell populations to assign identities. Tools such as SingleR [10] and scMAGIC [42] enable automated, replicable classification of individual cells based on reference transcriptomes. The preferred approach will depend on tissue complexity, the availability of high-quality reference data, and dataset size. Marker-based methods may suffice for small, well-characterized datasets, whereas reference-based methods are advantageous for large, heterogeneous tissues or when standardized annotations across studies are desired.

In this workflow, we use SingleR because it provides automated, single-cell level annotation using a curated reference (for which multiple reliable reference datasets exist in mouse brain), performs well for complex tissues such as the brain, and allows downstream analyses within/between cell types due to robust and reproducible cell-type labels.

To use SingleR for cell-type annotation, the sc/snRNAseq dataset must first be prepared to meet its input requirements. This includes ensuring that the RNA assay contains log-normalized expression data while also converting the Seurat object into a SingleCellExperiment, a standardized data structure designed to store single-cell expression data including metadata information. SingleR uses this structure to access both expression values and cell-level information. Preparing the object in this way enables SingleR to assign accurate cell-type labels to individual cells based on a given reference dataset.

```
## -----------------------------
## 1. Prepare RNA assay for SingleR
## -----------------------------
DefaultAssay(hypo.webb.annot.singlets) <- "RNA"
hypo.webb.annot.singlets <- JoinLayers(hypo.webb.annot.singlets, assay =
"RNA")
# flatten SCT layers
## Make sure RNA assay has log-normalized data (SingleR requires 'data'
slot)
hypo.webb.annot.singlets <- NormalizeData(
 object = hypo.webb.annot.singlets,
 assay = "RNA",
 normalization.method = "LogNormalize",
 scale.factor = 1e4)
## -----------------------------
## 2. Convert Seurat to SingleCellExperiment for SingleR
## -----------------------------
sce <- SingleCellExperiment(
 assays = list(
  counts  = SeuratObject::GetAssayData(hypo.webb.annot.singlets, assay =
"RNA", layer = "counts"),
logcounts = SeuratObject::GetAssayData(hypo.webb.annot.singlets, assay =
"RNA", layer = "data")))
colnames(sce) <- colnames(hypo.webb.annot.singlets)  # keep names aligned
```

Once the sc/snRNAseq dataset is converted to a SingleCellExperiment, SingleR assigns cell-type labels to individual cells by comparing their expression profiles to a curated mouse reference dataset, MouseRNAseqData in this case [43]. We chose this reference because it contains 358 mouse RNA-seq samples annotated to 18 main cell types, including all major cell types relevant to the brain, and provides a curated selection of reference expression datasets suitable for automated reference-based single-cell annotation with SingleR (*see* **Note 10**). The resulting annotations are then transferred back to the original singlet-only Seurat object, allowing cell-type labels to be linked with each cell and summarized across libraries, ensuring that assigned labels correspond to known brain cell types rather than unrelated populations (e.g., hepatocytes).

The SingleR object is saved to allow reproducible access to the assigned cell-type labels for downstream analyses, facilitating reuse, verification, or further exploration without rerunning the annotation, which can take a significant amount of time depending on the complexity and size of the reference and query datasets.

```
## -----------------------------
## 3. Run SingleR using brain reference
## -----------------------------
ref <- celldex::MouseRNAseqData()  # reference dataset
my.singler <- SingleR(test = sce, ref = ref, labels = ref$label.main)
## Save SingleR object for later
save(my.singler, file = paste(Sys.Date(), "hypo_webb_SingleR_object.RData",
sep = "_"))
## -----------------------------
## 4. Transfer labels back to Seurat object
## -----------------------------
hypo.webb.annot.singlets$SingleR_MouseRNAseq <- my.singler$labels
## Summarize results across libraries
```

```
print(table(hypo.webb.annot.singlets$SingleR_MouseRNAseq,
hypo.webb.annot.singlets$Library))
```

One important way to visualize the cell-type annotations from SingleR (or other annotation tools) is through UMAP embeddings (Fig. 10A). Here, we expect to see clear separation of clusters corresponding to different cell types. A key sanity check at this stage is to compare the UMAP with the marker gene expression UMAPs generated previously (see above), ensuring that canonical marker expression aligns with the annotated cell-type labels. Poor agreement between marker expression and cell-type annotation may indicate that the annotation tool or reference dataset is not optimal for your dataset, and alternative references or methods should be evaluated.

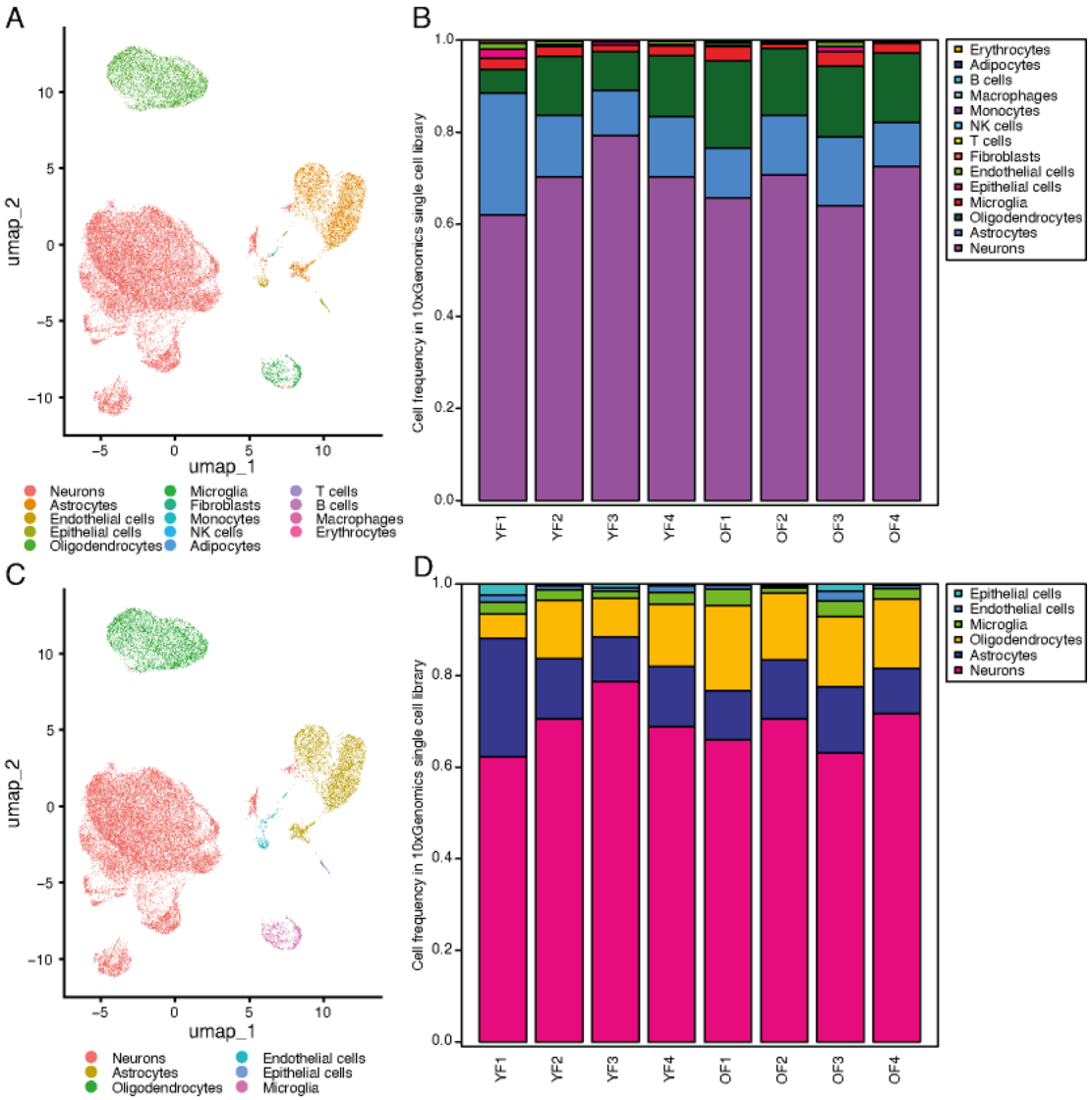


Fig. 10 Visualization and quantification of SingleR-assigned cell types before and after annotation refinement. **A,** UMAP projection of cells colored by initial SingleR-assigned cell types; (**B**) Barplot showing the relative proportions of each initial SingleR-assigned cell type within each library; (**C**) UMAP projection after refinement of annotations and removal of low-confidence or ambiguous labels; (**D**) Barplot of relative cell-type proportions per library after annotation refinement

To examine the relative cell-type proportions across libraries, we next calculate the proportion of each SingleR-assigned cell type within each library and visualize the results as a stacked barplot (Fig. 10B). Within experimental conditions, cell-type proportions are generally expected to be relatively consistent

across libraries, indicating minimal technical artifacts and supporting the credibility of the dataset. Discrepancies between conditions (e.g., young vs old in this case) may reflect true biological variation and are not an immediate cause for concern.

```
# ----------------------------
# 5. Set Seurat identities and UMAP visualization
# ----------------------------
hypo.webb.annot.singlets <- SetIdent(hypo.webb.annot.singlets, value =
"SingleR_MouseRNAseq")
DimPlot(hypo.webb.annot.singlets, reduction = "umap", raster = TRUE)
# ----------------------------
# 6. Barplot of SingleR cell-type proportions
# ----------------------------
my.freq.table <- prop.table(table(hypo.webb.annot.singlets$SingleR_MouseRNAsec
hypo.webb.annot.singlets$Library), margin = 2)
my.freq.table.av <- apply(my.freq.table,1,mean)
my.freq.table.av.sort <- sort(my.freq.table.av, decreasing = TRUE, index.retur
TRUE)
my.cols <- c("blue3","dodgerblue","cyan3",
"chartreuse3","deeppink2","darkgoldenrod1", "coral1",
    "aquamarine2","firebrick1","darkviolet","darkorchid1","deepskyblue","darkg
barplot(my.freq.table[my.freq.table.av.sort$ix,], las = 2,
   legend.text = rownames(my.freq.table)[my.freq.table.av.sort$ix],
   col = my.cols[my.freq.table.av.sort$ix],
   ylab = "Cell frequency in 10x Genomics single-cell library")
box()
```

Refining Cell-Type Annotations Using Cluster-Level Consensus

To improve interpretability and reduce cell-level annotation noise, we next create a cluster-level consensus cell-type annotation based on the SingleR predictions. For each cluster defined at a resolution of 0.5 using the `FindClusters` function described above (in this case, 21 clusters), the most frequently assigned SingleR cell type is identified and uniformly applied to all cells within that cluster. For example, if 90% of cells in a given cluster are annotated as neurons and the remaining 10% are distributed among microglia, oligodendrocytes, and astrocytes, the entire cluster is assigned a neuronal identity (100% now annotated as neurons).

This consensus-based approach dampens random cell-to-cell misclassifications while preserving biologically important cluster structure.

```
##############################################################################
# Create a cleaned SingleR annotation based on cluster-level majority voting
##############################################################################
# Initialize a new metadata column to store the cleaned SingleR labels
hypo.webb.annot.singlets@meta.data$SingleR_MouseRNAseq_CLEAN <- NA
# Loop through each cluster defined at resolution 0.5
# (clusters are numbered 0 through 20 in this object)
for (i in 0:20) {
 # Subset metadata for cells belonging to the current cluster
 cluster.cells <- hypo.webb.annot.singlets@meta.data[
  hypo.webb.annot.singlets@meta.data$SCT_snn_res.0.5 == i,]
 # Count how many cells in the cluster were assigned to each SingleR cell type
 id.summary <- aggregate(
  cluster.cells$SingleR_MouseRNAseq,
```

```
 by = list(cluster.cells$SingleR_MouseRNAseq),
 FUN = "length")
# Rename columns for clarity
colnames(id.summary) <- c("CellType", "Count")
# Calculate the proportion of each SingleR label within the cluster
id.summary$proportion <- id.summary$Count / sum(id.summary$Count)
# Assign the majority (most frequent) SingleR label
majority.label <- id.summary$CellType[which.max(id.summary$proportion)]
# Apply the majority label to all cells in the current cluster
hypo.webb.annot.singlets@meta.data$SingleR_MouseRNAseq_CLEAN[
 hypo.webb.annot.singlets@meta.data$SCT_snn_res.0.5 == i] <- majority.label
}
```

The refined annotations are then visualized on a UMAP to ensure coherent cell-type organization across libraries and clusters (Fig. 10C; *see* **Note 11**). Additionally, we again visualize the relative cell type proportions across libraries after the refinement using a stacked barplot (Fig. 10D). This refinement facilitates robust downstream analyses within individual cell types as well as comparative analyses between/across cell types.

We have discovered that saving the Seurat object after refining the cell-type annotation is a good practice. In our experience, SingleR annotation can take a substantial amount of time, ranging from several minutes to hours depending on the number of cells and the size of the reference dataset, and subsequent steps, including dataset integration, can also take up to an hour depending on the tool used. Saving the object at this point prevents loss of work and allows analyses to resume quickly in case of downstream issues or computational complications.

```
##############################################################################
# Set the cleaned SingleR labels as the active cell identity
##############################################################################
hypo.webb.annot.singlets <- SetIdent(
 hypo.webb.annot.singlets,
 value = "SingleR_MouseRNAseq_CLEAN")
##############################################################################
# Visualize cleaned SingleR annotations on a UMAP
##############################################################################
DimPlot(
 object = hypo.webb.annot.singlets,
 reduction = "umap")
##############################################################################
# Visualize cleaned SingleR annotation proportions on a barplot
##############################################################################
my.freq.table <-
prop.table(table(hypo.webb.annot.singlets$SingleR_MouseRNAseq_CLEAN,hypo.webb.
margin = 2)
my.freq.table.av <- apply(my.freq.table,1,mean)
my.freq.table.av.sort <- sort(my.freq.table.av, decreasing = TRUE, index.retur
my.cols <- c("blue3","dodgerblue","cyan3","chartreuse3","deeppink2","darkgolde
    "aquamarine2","firebrick1","darkviolet","darkorchid1","deepskyblue","darkg
barplot(my.freq.table[my.freq.table.av.sort$ix,], las = 2,
   legend.text = rownames(my.freq.table)[my.freq.table.av.sort$ix],
   col = my.cols[my.freq.table.av.sort$ix],
   ylab = "Cell frequency in 10x Genomics single-cell library")
box()
```

```
##############################################################################
# Summarize cleaned cell-type annotations across libraries
##############################################################################
print(
 table(
  hypo.webb.annot.singlets$SingleR_MouseRNAseq_CLEAN,
  hypo.webb.annot.singlets$Library))
##############################################################################
# Save updated Seurat object with cleaned SingleR annotations
##############################################################################
save(
 hypo.webb.annot.singlets,
 file = paste(Sys.Date(),
     "hypo_webb_seurat_singlets_withSingleR_CLEAN.RData",
```

9.2 Integration

Following cell-type annotation, we next execute integration on the dataset. Data integration at the single-cell level corrects for technical differences between libraries while maintaining biological variation, allowing accurate joint analysis, comparison, and visualization of cell populations across the entire dataset. Integration for this specific dataset is necessary because half of the samples (libraries) were generated with 10× v2 chemistry and the other half with 10× v3 chemistry, introducing batch and technical differences that are not biologically relevant. These differences are readily observed in UMAP visualizations, as previously shown (Fig. 6), where clusters and cell types separate clearly based on the 10× chemistry used for each library. For this workflow, we follow the *Seurat v5 vignette for Integrative Analysis* [44] and implement and compare the first three integration methods showcased there: RPCA, CCA, and Harmony.

Prior to integration, expression data are normalized and organized to ensure that technical differences between libraries do not confound downstream analyses. Library-specific layers are generated so that normalization can be applied independently to each sample, accounting for variation introduced by library preparation, sequencing chemistry, or other technical factors. Normalization also adjusts for common covariates such as total counts, mitochondrial content, feature counts, and cell cycle effects (as previously described in this workflow). This prepares a set of scaled, comparable expression measurements across all libraries, enabling accurate integration and robust joint analysis of cell populations.

```
##############################################################################
## 1. Set RNA assay and split data into sample-specific layers (Seurat v5)
##############################################################################
# Ensure the RNA assay is set as the default assay
DefaultAssay(hypo.webb.annot.singlets) <- "RNA"
# Check existing layers in the RNA assay
Layers(hypo.webb.annot.singlets[["RNA"]])
# Expected output prior to splitting: "counts" and "data"
# Split the RNA assay into library-specific layers (v5 layer-based workflow)
# This creates separate 'counts' and 'data' layers for each library/sample
# and prepares the object for integration
hypo.webb.annot.singlets[["RNA"]] <- split(

hypo.webb.annot.singlets[["RNA"]],

f = hypo.webb.annot.singlets$Library)

# Verify that layers were split correctly
Layers(hypo.webb.annot.singlets[["RNA"]])
```

```
# Expected output: data.YF1, data.YF2, …, counts.OF4
################################################################################
## 2. SCTransform normalization (performed per layer)
################################################################################
# Normalize and scale the data using SCTransform
# Technical covariates are regressed out to reduce unwanted variation
# SCTransform is applied independently to each library-specific layer
hypo.webb.annot.singlets <- SCTransform(

hypo.webb.annot.singlets,

vars.to.regress = c("nFeature_RNA", "nCount_RNA", "percent.mito", "Phase"),

verbose = FALSE)

# Warning messages here are expected when re-running SCT on an existing
object
```

Now that the expression data have been organized and normalized to account for technical (non-biologically meaningful) differences between libraries, we perform PCA on the library-specific layers to assess the structure of the dataset and facilitate visualization. Using previously described metrics, we select the optimal number of principal components, which provide a reduced-dimensional representation of the data. These PCs serve as the core for UMAP visualizations, allowing comparison of the organization of cells before and after integration, which will be a key determinant in the selection of which integration method is ideal for a given dataset (see below).

```
################################################################################
## 3. PCA on unintegrated (layered) data
################################################################################
# Perform PCA on the SCT-normalized data
# This step is done prior to integration to assess dataset dimensionality
hypo.webb.annot.singlets <- RunPCA(
 hypo.webb.annot.singlets,
 npcs = 50,
 verbose = FALSE)
# Visualize variance explained by each principal component
ElbowPlot(hypo.webb.annot.singlets, ndims = 50)
################################################################################
## Determine the optimal number of PCs using quantitative metrics
################################################################################
# Calculate percent variance explained by each PC
pct <- hypo.webb.annot.singlets[["pca"]]@stdev /
  sum(hypo.webb.annot.singlets[["pca"]]@stdev) * 100
# Calculate cumulative variance explained
cumu <- cumsum(pct)
# Identify the first PC where cumulative variance exceeds 90%
# and the individual PC explains less than 5% of total variance
co1 <- which(cumu > 90 & pct < 5)[1]
# Identify the last PC where the drop in variance between PCs is > 0.1%
co2 <- sort(
 which((pct[1:length(pct) - 1] - pct[2:length(pct)]) > 0.1),
 decreasing = TRUE)[1] + 1
# Select the more conservative cutoff
```

```
pcs <- min(co1, co2)
```

We are now ready to perform integration using each of the three methods independently (*see* **Note 12**). Each method addresses batch correction differently while aiming to conserve biological variation: RPCA provides a fast, conservative integration [23], CCA performs a more aggressive alignment [23], and Harmony offers flexible, rapid batch correction [18, 23]. There is no universally "best" method, in fact the optimal approach must be discerned for each dataset to ensure biological structure is maintained and technical variation is minimized (*see* **Note 13**). Integration is especially important for publicly available datasets, which can often include samples generated with differing protocols, at different times, or using divergent chemistries or platforms, introducing technical variation unrelated to biology [23]. Applying integration lessens these artifacts, enabling downstream single-cell analyses to focus on real biological differences as opposed to batch or technical effects [23]. In this workflow, we utilize all three methods to compare integration efficacy and choose the most appropriate approach for this dataset.

```
##############################################################################
## 4. Data integration using multiple methods
##############################################################################
# Integrate data using RPCA (fast and conservative; recommended for large
datasets)
hypo.webb.annot.singlets.rpca <- IntegrateLayers(
 object = hypo.webb.annot.singlets,
 method = RPCAIntegration,
 normalization.method = "SCT",
 orig.reduction = "pca",
 new.reduction  = "integrated.rpca",
 verbose = FALSE)
# Runtime ~2 minutes
# Integrate data using CCA (more aggressive alignment; slower)
hypo.webb.annot.singlets.cca <- IntegrateLayers(
 object = hypo.webb.annot.singlets,
 method = CCAIntegration,
 normalization.method = "SCT",
 orig.reduction = "pca",
 new.reduction  = "integrated.cca",
 verbose = FALSE)
# Runtime ~40 minutes
# Integrate data using Harmony (fast and flexible batch correction)
hypo.webb.annot.singlets.harmony <- IntegrateLayers(
 object = hypo.webb.annot.singlets,
 method = HarmonyIntegration,
 normalization.method = "SCT",
 orig.reduction = "pca",
 new.reduction  = "harmony",
 verbose = FALSE)
# Runtime ~1 minute
```

Selecting the Optimal Integration Approach

Although three integration methods, including RPCA and CCA, were applied in the workflow above for comparison, for simplicity we will focus the subsequent evaluation and visualization on Harmony, which we selected based on previous experience in our lab and because it demonstrated the best integration efficacy on this specific dataset (see explanations below). Integration quality is determined using two unique but

complementary visualization techniques, providing a foundation to juxtapose how well batch effects are reduced while preserving biologically meaningful structure.

The first visualization technique to evaluate integration performance examines how well Harmony integrates cells across the two 10× chemistry versions (v2 vs v3), which we previously identified as the primary source of batch differences between the libraries. In the unintegrated dataset, cells are clearly separated by chemistry across all clusters and cell types (Fig. 11A). Importantly, after Harmony integration, this separation is significantly reduced, exemplifying effective correction of chemistry-driven batch effects while preserving biological structure (Fig. 11B).

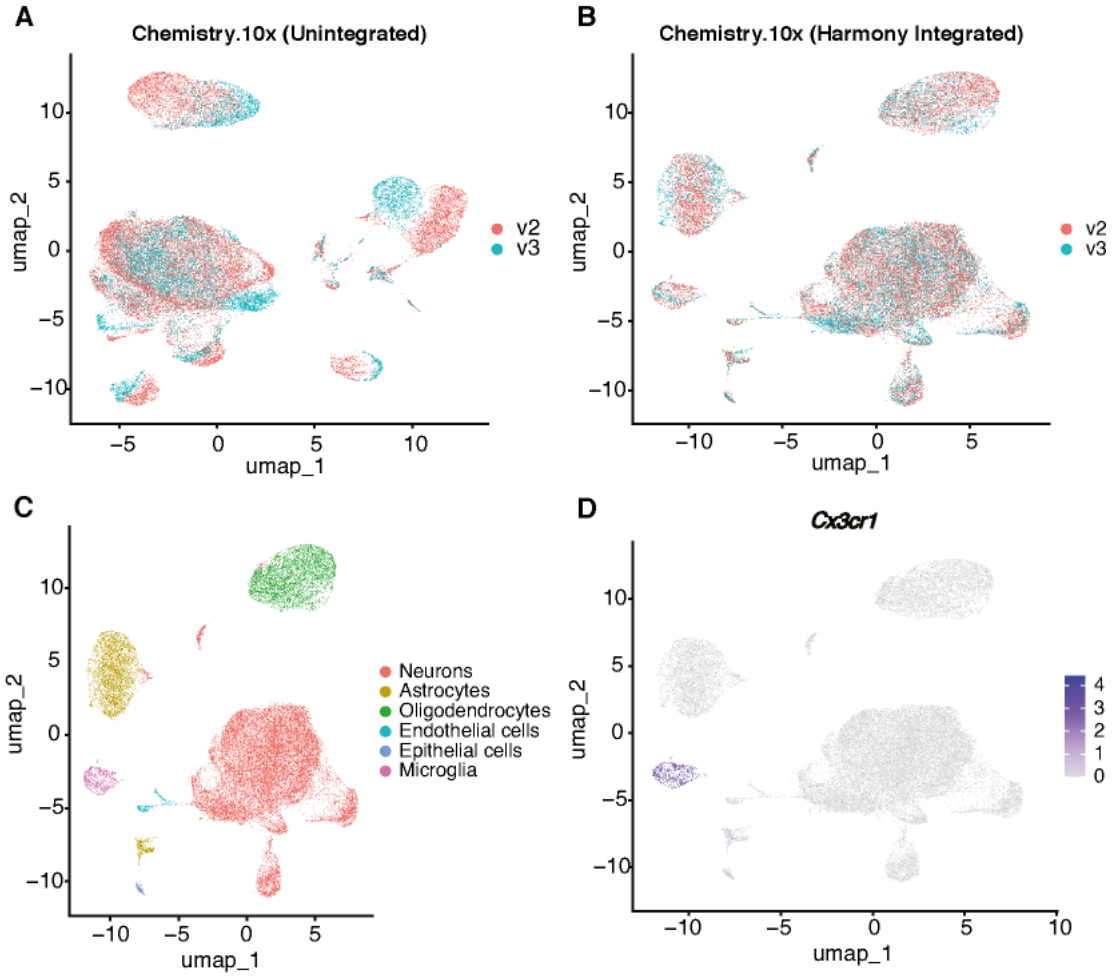


Fig. 11 Evaluation of Harmony integration performance across 10x Genomics v2 and v3 chemistries. **A,** UMAP projection of the unintegrated dataset, showing separation by v2 vs v3 indicating strong batch effects. **B,** UMAP projection after Harmony integration, showing reduced separation by chemistry while preserving biological clustering, demonstrating batch correction. **C,** UMAP visualization colored by cell-type annotation (SingleR), illustrating that biological identities are maintained after integration. **D,** UMAP showing expression of the microglia marker C×3cr1 post-integration, confirming that canonical marker expression remains specific to the corresponding cell-type cluster

```
## ----------------------------
## 5. Graph construction, clustering, and UMAP for integrated and
unintegrated data
## Purpose: Compare integration methods (only Harmony shown here) to
determine which ## best removes batch effects while preserving biological
structure.
## ----------------------------
# ----------------------------
```

```
# 5a. Unintegrated dataset (baseline)
# ----------------------------
# Construct nearest-neighbor graph, cluster, and run UMAP for the
unintegrated data
# This provides a baseline to see batch effects and library separation
before integration
hypo.webb.annot.singlets <- FindNeighbors(hypo.webb.annot.singlets, dims =
1:pcs)
hypo.webb.annot.singlets <- FindClusters(object = hypo.webb.annot.singlets)
hypo.webb.annot.singlets <- RunUMAP(hypo.webb.annot.singlets, dims = 1:pcs)
# ----------------------------
# 5b. Harmony-integrated dataset
# ----------------------------
# Harmony performs fast, flexible batch correction
hypo.webb.annot.singlets.harmony <-
FindNeighbors(hypo.webb.annot.singlets.harmony, reduction = "harmony", dims
= 1:pcs)
hypo.webb.annot.singlets.harmony <-
FindClusters(hypo.webb.annot.singlets.harmony, resolution = 0.5,
cluster.name = "harmony_clusters")
hypo.webb.annot.singlets.harmony <-
RunUMAP(hypo.webb.annot.singlets.harmony, reduction = "harmony", dims =
1:pcs)
# ----------------------------
# 6. Visualize integration quality
# ----------------------------
# Purpose: Assess how well each method mixes libraries/chemistries and
preserves cell # type structure
# Unintegrated UMAP plots
DimPlot(hypo.webb.annot.singlets, group.by = "Chemistry.10x", raster =
TRUE)  # check 10x chemistry effect
# Harmony-integrated plots
DimPlot(hypo.webb.annot.singlets.harmony, group.by = "Chemistry.10x", raster
= TRUE)
# Repeat for other integration methods as needed…
# Repeat for other metadata columns (e.g. library or cell annotation) as
needed…
```

The second visualization technique we enact to evaluate integration efficacy examines the expression of canonical marker genes, with microglia shown here as an illustrative example. Harmony is able to maintain canonical marker expression exclusively within the microglia-annotated clusters (SingleR annotation), suggesting that Harmony integration preserves biological structure while also correcting for technical differences (Fig. 11C, D).

```
# ----------------------------
# 7. Visualize key marker gene expression for Harmony integration method
# ----------------------------
# Purpose: Confirm that biological structure is preserved (cell types remain
distinct) # after integration
DimPlot(hypo.webb.annot.singlets.harmony, group.by = "
SingleR_MouseRNAseq_CLEAN ", raster = TRUE)  # check SingleR annotations
allow visualization of marker genes
```

```
# Define canonical markers for major brain cell types
neuron.markers <- c("Map2", "Slc17a7", "Gad1", "Snap25")
microglia.markers <- c("P2ry12", "Ptprc", "Tmem119", "Cx3cr1")
astrocyte.markers <- c("Agt", "Gfap", "Aldh1l1")
oligo.markers <- c("Olig1", "Cldn11", "Mal", "Ptgds")
# Example for Harmony-integrated object (repeat for CCA or RPCA as needed)
print(
 patchwork::wrap_plots(
  lapply(microglia.markers, function(g)
FeaturePlot(hypo.webb.annot.singlets.harmony, features = g, raster =
TRUE)),
  ncol = 2))
# Repeat for other cell types and integration methods as needed…
```

Harmony outperforms, or performs equally, to both RPCA and CCA integration across both visualization techniques, therefore we will continue with downstream analyses using the Harmony-integrated dataset at both the single-cell and pseudobulk levels.

```
# ----------------------------
# 8. Save integrated Harmony object for downstream analysis
# ----------------------------
save(hypo.webb.annot.singlets.harmony, file =
paste0(Sys.Date(),"_Harmony_integrated.RData"))
# Since Harmony performed best in mixing chemistries while preserving cell
type # structure, we join its layers to create a single consolidated Seurat
object for # downstream analysis
hypo.webb.annot.singlets.harmony.joined <-
JoinLayers(hypo.webb.annot.singlets.harmony)
```

---

10 Downstream Analysis

10.1 Cell Proportion Test with scProportionTest

After integration, we can see if the cell type proportions change between old and young samples in our dataset with scProportionTest [26]. Permutation testing is used to calculate a *p* value for each cell type (cluster) and then bootstrapping is used to estimate the degree of uncertainty in the proportion changes.

Formatting the Data for scProportionTest

First, we need to format the data for scProportionTest.

```
### 4. scProportionTest (global/integrated) all cell types
# create prop test object
hypo.webb.prop_test <- sc_utils(hypo.webb.annot.singlets.harmony)
```

Permutation Testing and Bootstrapping

Next, we can run the permutation test using the `permutation_test()` function.

```
# Once the object is created, the permutation testing and bootstrapping can
be run.
hypo.webb.prop_test.F <- permutation_test(hypo.webb.prop_test,
         cluster_identity = "SingleR_MouseRNAseq_CLEAN",
         sample_1 = "YF",
         sample_2 = "OF",
```

```
        sample_identity = "Condition")
```

The permutation test works by reshuffling the labels of each sample and re-estimating the difference in cell type proportion.

For each cell type, our null hypothesis is: "There is no difference in cell proportion between old female samples and young female samples."

Our alternative hypothesis is: "There is a difference in cell proportion between old female samples and young female samples."

In this example, let's say we want to understand if the proportion of microglia changes between old and young samples. So, our observed difference might be 0.60% or 60% for microglia. If you were to randomly shuffle the labels (OF, YF) across all cells and recompute the difference in proportion thousands of times, you would get the difference you would expect by chance alone. So, if the observed difference is more extreme than your null difference, the proportion of cells is significantly changing with age.

```
> hypo.webb.prop_test.F@results$permutation
Key: <clusters>
clustersOFYFobs_log2FDpvalFDR
Astrocytes0.115774840.13673650-0.240076560.0009990010.001498501
Endothelial cells0.008061150.00940182-0.221953330.0859140860.085914086
Epithelial cells0.004401210.00706714-0.683227620.0029970030.003596404
Microglia0.027704420.018172640.608348370.0009990010.001498501
Neurons0.693073890.72274104-0.06046964 0.0009990010.001498501
Oligodendrocytes0.150984480.105880870.511958330.0009990010.001498501
```

In this case, there is a significant difference in the proportion of microglia between old and young samples.

Data Visualization

We can modify the plotting function to customize color and layout and visualize our results with a point-range plot to determine which cell type proportions change with age (Fig. 12).

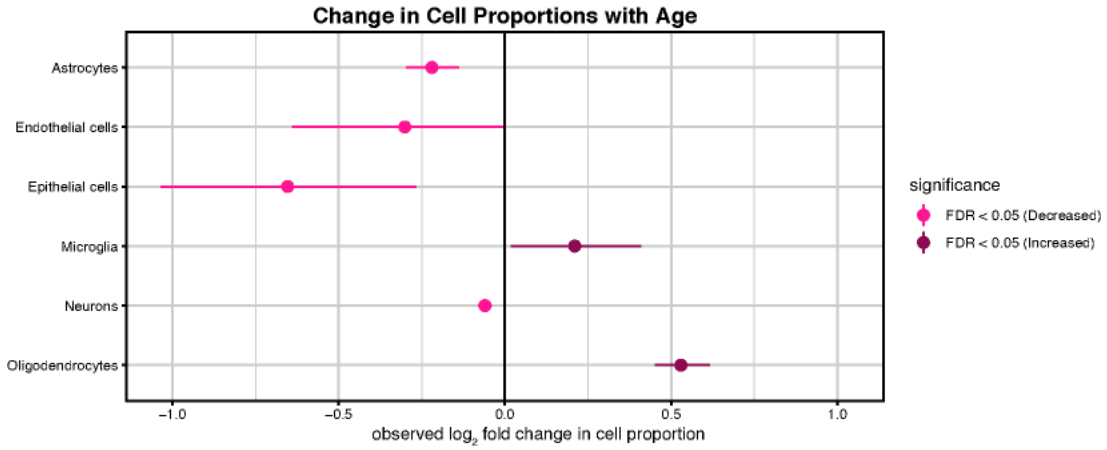


Fig. 12 Point-range plot showing the observed log2 fold change in cell type proportions with age (old vs young females) for astrocytes, endothelial cells, epithelial cells, microglia, neurons, and oligodendrocytes. Dark pink indicates a proportion significantly higher in old females, while light pink indicates a proportion significantly higher in young females (FDR < 5%)

```
## Modify function
permutation_plot_mod <- function (sc_utils_obj, FDR_threshold = 0.05,
cols_vals = c("deeppink","deeppink4"), my.title) {
 plot_data <- sc_utils_obj@results$permutation
```

```
 # fix order alphabetically
 plot_data$clusters <- factor(plot_data$clusters, levels =
rev(plot_data$clusters))
 # get significance and color scale
 plot_data$significance <- "n.s."
 plot_data$significance[plot_data$FDR < FDR_threshold &
plot_data$boot_mean_log2FD > 0] <- paste("FDR <", FDR_threshold, "
(Increased)")
 plot_data$significance[plot_data$FDR < FDR_threshold &
plot_data$boot_mean_log2FD < 0] <- paste("FDR <", FDR_threshold, "
(Decreased)")
 plot_data$significance <- factor(plot_data$significance, levels =
c(paste("FDR <", FDR_threshold, "(Decreased)"), paste("FDR <",
FDR_threshold, "(Increased)"), "n.s."))
 # fix range for symmetry
 max_range <- max(abs(plot_data[,7:8]))
 # plot
 p <- ggplot(plot_data, aes(x = clusters, y = obs_log2FD)) + theme_bw() +
  geom_pointrange(aes(ymin = boot_CI_2.5, ymax = boot_CI_97.5, color =
significance)) +
  geom_hline(yintercept = 0) + scale_color_manual(values = cols_vals) +
ylim(c(-max_range,max_range)) +
  coord_flip() + ggtitle(my.title)
 # p
 return(p)
}
# A point-range plot of the results can then be created.
perm.f <- permutation_plot_mod(hypo.webb.prop_test.F, cols_vals =
c("deeppink","deeppink4", "grey"), my.title = "Female Aging")
```

10.2 Cell Type Prioritization with Augur

Augur can be used to identify which cell types are most affected by aging. In our previous analysis, we annotated each cell in our single-cell dataset using cell type–specific markers and SingleR. We have already corrected for technical variation arising from different sequencing chemistries using Harmony, allowing us to focus on downstream analyses. Augur uses machine learning to predict, for each cell type, the experimental group of origin for individual cells. This approach enables us to determine which cell types are most impacted by aging.

Cell types are not represented in equal numbers in single-cell RNA-seq datasets. To avoid confounding cell-type abundance with model performance, Augur repeatedly subsamples a fixed number of cells from each experimental condition within each cell type. For each subsample, a classifier is trained to predict experimental condition from gene expression and evaluated using cross-validation. Model performance is summarized using the area under the ROC curve (AUC). The mean cross-validated AUC for each cell type reflects how well cells of that type can be distinguished by condition, with higher AUC values indicating stronger aging-associated transcriptomic changes (e.g., aging effects).

Calculating AUC Across Cell Types

First, we calculate the AUC for each cell type using `calculate_auc()`.

```
hypo.webb.annot.singlets.harmony
augur.celltypes <- calculate_auc(
 hypo.webb.annot.singlets.harmony@assays$SCT@data,
 hypo.webb.annot.singlets.harmony@meta.data,
```

```
 cell_type_col = "SingleR_MouseRNAseq_CLEAN",
 label_col = "Condition",
 n_threads = 1)
augur.celltypes$AUC
# # A tibble: 6 × 2
# cell_typeauc
# <chr><dbl>
# 1 Microglia0.644
# 2 Epithelial cells0.615
# 3 Oligodendrocytes0.606
# 4 Astrocytes0.558
# 5 Endothelial cells0.524
# 6 Neurons0.510
```

Here, we can see that microglia have the highest AUC value, meaning that those cells are the most impacted by aging at the transcriptomic level in our dataset.

Data Visualization

We can visualize the results either with UMAP (Fig. 13A) or a lollipop plot (Fig. 13B) to see clearly which cell types are most affected by aging.

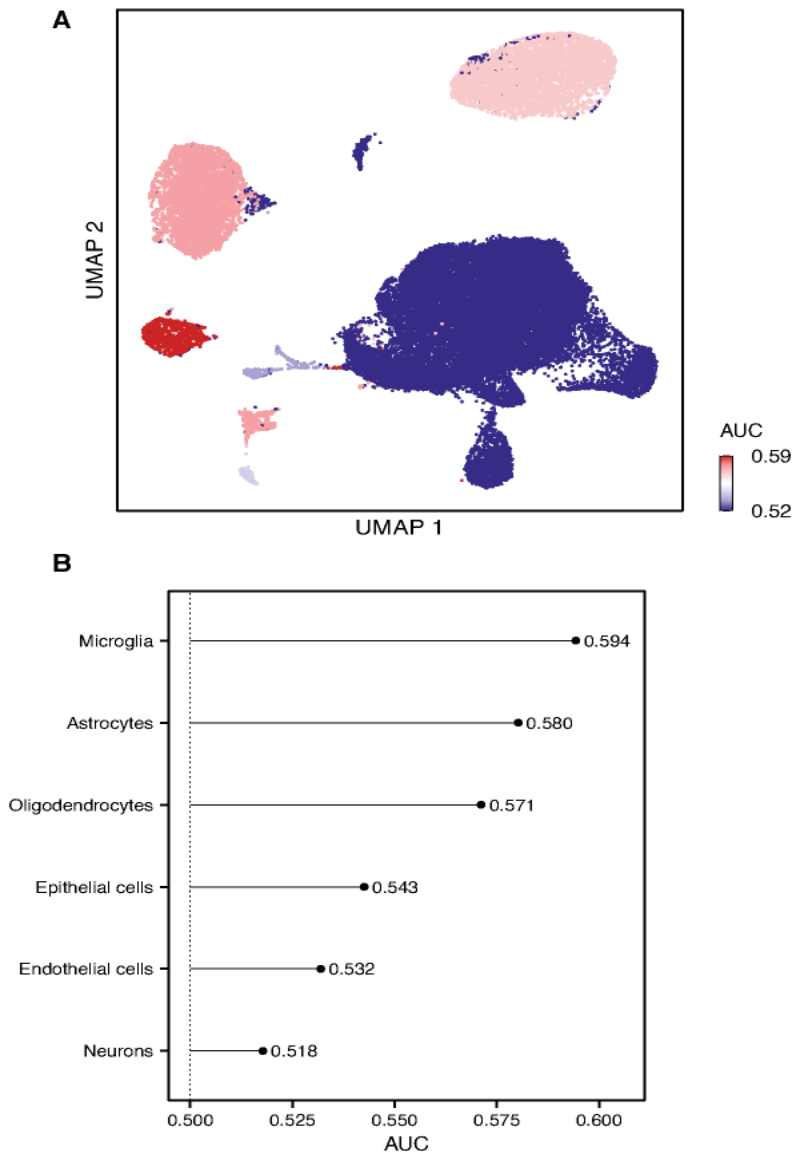


Fig. 13 Cell types most affected by aging. **A,** UMAP of cell clusters annotated with Augur AUC scores. **B,** Lollipop plot showing the age-associated effect size (AUC) for each cell type

```
plot_umap(augur.celltypes,hypo.webb.annot.singlets.harmony,cell_type_col="Sinc
plot_umap(augur.celltypes,hypo.webb.annot.singlets.harmony,cell_type_col="Sinc
(50))
plot_lollipop(augur.celltypes)
```

10.3 Calculating Gene Signature Scores with UCell

For gene signatures of interest, it is possible to obtain cell-level scores which indicate the relative activity of a specific gene program. For this purpose, we recommend a tool called UCell to estimate gene signature scores. UCell is a rank-based gene set scoring method designed specifically for single-cell RNA-seq data. Unlike approaches that rely on absolute expression values, UCell operates on within-cell gene expression ranks. Importantly, UCell computes scores independently for each cell, allowing us to estimate aging-specific gene expression signatures across diverse cell types. Finally, out of all the algorithms we tested in the lab for this purpose UCell uses the least amount of memory and runs the fastest, making it an ideal choice for larger datasets with more cells/nuclei.

Loading the Gene List

For this example, we will calculate senescence gene signature scores using the mouse-adapted SenMayo gene list from Saul et al. [45]. You can either load in the gene list via a csv or Excel file, or just manually enter the gene list of interest.

```
gene_sets <- list(
 SenMayo = c(
  "Acvr1b","Ang","Angpt1","Angptl4","Areg","Axl","Bex3","Bmp2","Bmp6",
  "C3","Ccl1","Ccl2","Ccl20","Ccl24","Ccl26","Ccl3","Ccl4","Ccl5",
  "Ccl7","Ccl8","Cd55","Cd9","Csf1","Csf2","Csf2rb","Cst10","Ctnnb1",
  "Ctsb","Cxcl1","Cxcl10","Cxcl12","Cxcl16","Cxcl2","Cxcl3","Cxcr2",
  "Dkk1","Edn1","Egf","Egfr","Ereg","Esm1","Ets2","Fas","Fgf1","Fgf2",
  "Fgf7","Gdf15","Gem","Gmfg","Hgf","Hmgb1","Icam1","Icam5","Igf1",
  "Igfbp1","Igfbp2","Igfbp3","Igfbp4","Igfbp5","Igfbp6","Igfbp7",
  "Il10","Il13","Il15","Il18","Il1a","Il1b","Il2","Il6","Il6st",
  "Il7","Inha","Iqgap2","Itga2","Itpka","Jun","Kitl","Lcp1","Mif",
  "Mmp10","Mmp12","Mmp13","Mmp14","Mmp2","Mmp3","Mmp9","Nap1l4",
  "Nrg1","Pappa","Pecam1","Pgf","Pigf","Plat","Plau","Plaur",
  "Ptbp1","Ptger2","Ptges","Rps6ka5","Scamp4","Selplg","Sema3f",
  "Serpinb3a","Serpine1","Serpine2","Spp1","Spx","Timp2","Tnf",
  "Tnfrsf11b","Tnfrsf1a","Tnfrsf1b","Tubgcp2","Vegfa","Vegfc",
  "Vgf","Wnt16","Wnt2"))
```

Calculating Ucell Score

Next, we will calculate enrichment score for a given gene list using `AddModuleScore.UCell()`.

```
hypo.webb.annot.singlets.harmony.ucell       <-
AddModuleScore_UCell(hypo.webb.annot.singlets.harmony,
                        features = gene_sets,

name = NULL)
```

Now we need to split the object by cell type so that *p* values can be calculated and enrichment can be plotted for each cell type individually. After splitting, statistical testing is performed independently for each cell type, allowing us to determine whether senescence-associated gene score differs between experimental conditions within that specific cellular context. For statistical inference, we use the Wilcoxon rank–sum test to compare UCell scores between old female (OF) and young female (YF) cells. The Wilcoxon test is a non-parametric method that compares the distributions of two groups without assuming normality or homoscedasticity. This choice is particularly appropriate for UCell scores, which are rank-based and often exhibit skewed, non-normal distributions.

```
# split seurat object by cell type
hypo.webb.annot.singlets.harmony.ucell <-
SplitObject(hypo.webb.annot.singlets.harmony.ucell, split.by =
"SingleR_MouseRNAseq_CLEAN")
# create pval matrix
hypo.webb.stats <- data.frame(
 Cell_Type = names(hypo.webb.annot.singlets.harmony.ucell),
 Mean_SenMayo_YF = NA,
 Mean_SenMayo_OF = NA,
 Delta_SenMayo = NA,
 Wilcoxon_P_SenMayo = NA,
```

```
 stringsAsFactors = FALSE)
for (i in seq_along(hypo.webb.annot.singlets.harmony.ucell)) {
 obj <- hypo.webb.annot.singlets.harmony.ucell[[i]]
 # subset to YF vs OF
 meta.YO <- obj@meta.data[obj$Condition %in% c("YF", "OF"),]
 # group means
 mean.YF <- mean(meta.YO$SenMayo[meta.YO$Condition == "YF"], na.rm = TRUE)
 mean.OF <- mean(meta.YO$SenMayo[meta.YO$Condition == "OF"], na.rm = TRUE)
 # store means
 hypo.webb.stats$Mean_SenMayo_YF[i] <- mean.YF
 hypo.webb.stats$Mean_SenMayo_OF[i] <- mean.OF
 hypo.webb.stats$Delta_SenMayo[i] <- mean.OF - mean.YF
 # Wilcoxon test
 hypo.webb.stats$Wilcoxon_P_SenMayo[i] <-
  wilcox.test(SenMayo ~ Condition, data = meta.YO)$p.value
}
```

Data Visualization

To visualize senescence gene enrichment score and its associated statistical significance across cell types, we use a bubble plot in which each point represents a single cell type. Bubble color encodes the direction and magnitude of enrichment, defined as the difference in mean SenMayo UCell score between old female and young female cells (OF–YF), while bubble size reflects statistical significance, represented as the—log10 FDR-adjusted Wilcoxon *p* value (Fig. 14).

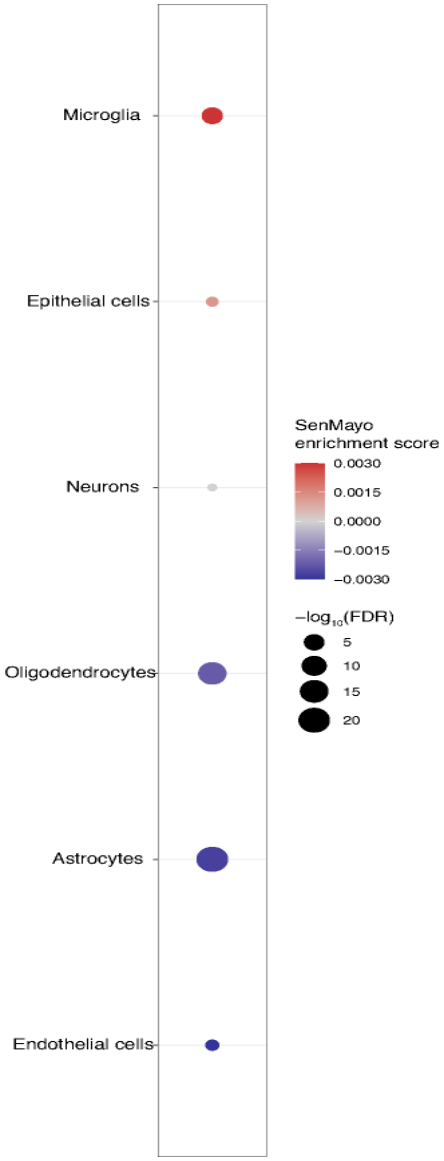


Fig. 14 Bubble plot showing SenMayo enrichment scores for each cell type, calculated using UCell. Bubble size represents the statistical significance of the enrichment (−log10 FDR), where FDR is based on Wilcoxon-adjusted *p* values

```
plot.df <- hypo.webb.stats
plot.df$log10p <- -log10(plot.df$Wilcoxon_P_SenMayo)
ggplot(plot.df,
  aes(
   x = 1,
   y = reorder(Cell_Type, Delta_SenMayo),
  color = Delta_SenMayo,
  size = log10p)) +
 geom_point() +
 scale_color_gradient2(
  low = "#333399",
  mid = "lightgrey",
  high = "#CC3333",
  midpoint = 0,
```

```
 limits = c(-0.003, 0.003),
 oob = scales::squish,
 breaks = c(-0.003, -0.0015, 0, 0.0015, 0.003),
 name = "SenMayo enrichment\n(OF - YF)") +
scale_size(
 range = c(2, 8),
 name = "-log10(FDR)") +
scale_x_continuous(breaks = NULL) +
labs(
 x = NULL,
 y = "Cell type") +
theme_bw() +
theme(
 panel.grid.major.x = element_blank(),
 panel.grid.minor.x = element_blank())
```

---

11 Pseudobulking

11.1 Pseudobulk Aggregation Across All Cell Types Using Muscat

We next perform pseudobulk analysis across all cell types together using the muscat package. Pseudobulking aggregates single-cell counts by cluster and sample (cell type and library in this case), producing a single sample-level expression profile for each cell type in each library, which reduces cell-to-cell variability and allows the application of robust bulk RNAseq statistical methods. Pseudobulking also addresses the issue of nonindependence among cells derived by the same animal. This approach is particularly useful for detecting broad, cell-type-specific effects between conditions, assessing overall cluster composition, or performing exploratory analyses such as multidimensional scaling (MDS) to visualize global differences between groups. It is important to note that the MDS generated using muscat may not be reliable for all datasets, as no batch correction has been applied at this stage (this will be covered later on in the workflow). By first aggregating all cell types together, we obtain an overview of the dataset and can identify broad trends before performing more granular analyses within individual cell types, which can reveal more subtle or cell-type-specific differences.

```
## ----------------------------
## 6. Pseudobulk analysis with muscat (all cell types together)
## ----------------------------
# Convert Seurat Harmony object to SingleCellExperiment
hypo.webb.annot.singlets.sce <-
as.SingleCellExperiment(hypo.webb.annot.singlets.harmony.joined)
# Prepare SCE for Muscat
hypo.webb.annot.singlets.sce.cl <- prepSCE(hypo.webb.annot.singlets.sce,
          kid = "SingleR_MouseRNAseq_CLEAN",
          gid = "Condition",
          sid = "Library",
          drop = TRUE)
# Store cluster and sample info
nk <- length(kids <- levels(hypo.webb.annot.singlets.sce.cl$cluster_id))
ns <- length(sids <- levels(hypo.webb.annot.singlets.sce.cl$sample_id))
names(kids) <- kids; names(sids) <- sids
# Count number of cells per cluster-sample
t(table(hypo.webb.annot.singlets.sce.cl$cluster_id,
hypo.webb.annot.singlets.sce.cl$sample_id))
# Aggregate single-cell counts into pseudobulk profiles
```

```
pb.hypo.webb <- aggregateData(hypo.webb.annot.singlets.sce.cl, assay =
"counts", fun = "sum", by = c("cluster_id", "sample_id"))
# Check available assays
assayNames(pb.hypo.webb)

   # [1] "Astrocytes"  "Endothelial cells" "Epithelial
cells" "Microglia"  "Neurons"  "Oligodendrocytes"
```

We have now pseudobulked by `cluster_id` (i.e., annotated cell type) and `sample_id` (i.e., experimental condition or age, in this case). Our pseudobulked object contains the aggregated counts for each gene across all cells within a specific sample-celltype combination. We can check the separation of our pseudobulked samples using before removing any technical variation using MDS on the complete dataset, including all cell types and samples (Fig. 15).

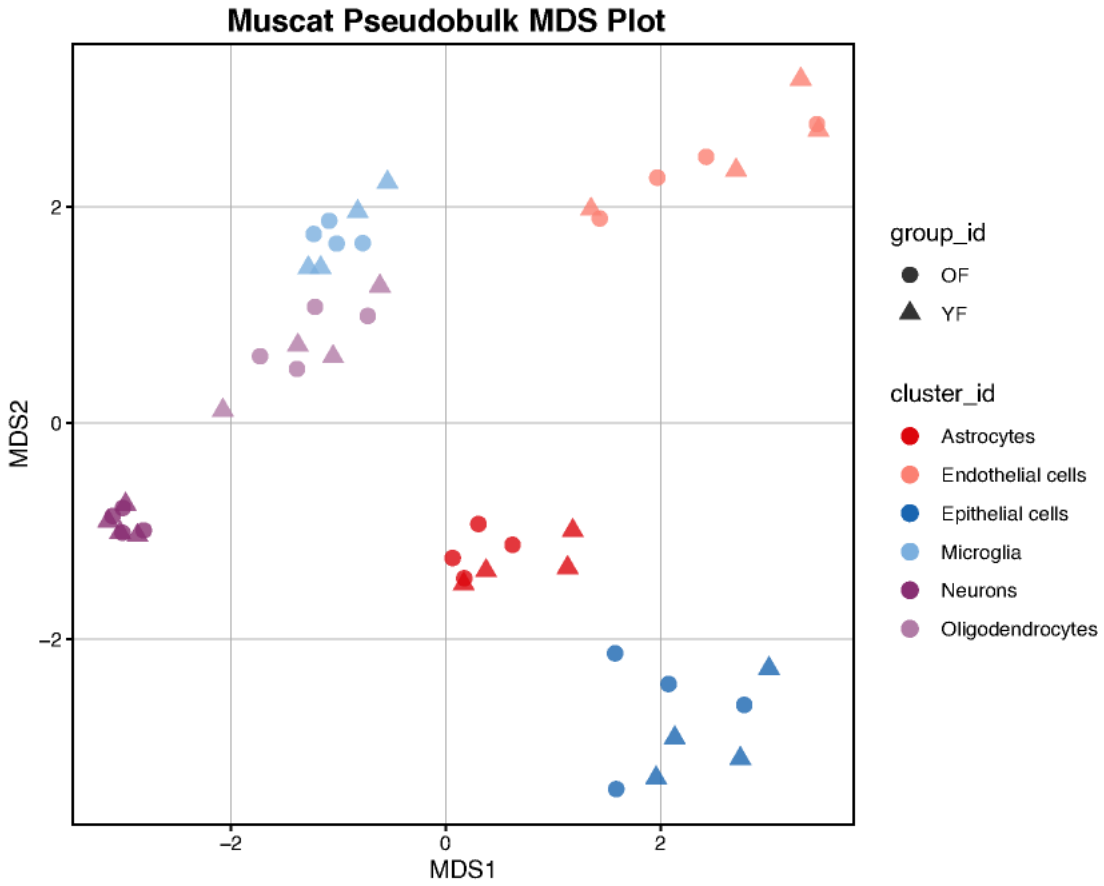


Fig. 15 Muscat pseudobulk MDS plot before correcting for technical variation. Shape indicates age group (old females [OF] vs young females [YF]), and color indicates cell type, showing variation between cell types and age groups

```
# Pseudobulk-level MDS plot
pb_mds <- pbMDS(pb.hypo.webb)
pb_mds
# Run differential state analysis
res <- pbDS(pb.hypo.webb, method = "edgeR", min_cells = 1, verbose = TRUE)
# Access results table
tbl <- res$table[[1]]
names(tbl)
# Save results to files
```

```
write.table(tbl[[1]][,1:6], file =
paste0(Sys.Date(),"_Muscat_EdgeR_neuron_genes_statistics.txt"), sep = "\t",
row.names = F, quote=F)
write.table(tbl[[2]][,1:6], file =
paste0(Sys.Date(),"_Muscat_EdgeR_astrocytes_genes_statistics.txt"), sep =
"\t", row.names = F, quote=F)
write.table(tbl[[3]][,1:6], file =
paste0(Sys.Date(),"_Muscat_EdgeR_endothelial_cells_genes_statistics.txt"),
sep = "\t", row.names = F, quote=F)
write.table(tbl[[4]][,1:6], file =
paste0(Sys.Date(),"_Muscat_EdgeR_epithelial_cells_genes_statistics.txt"),
sep = "\t", row.names = F, quote=F)
write.table(tbl[[5]][,1:6], file =
paste0(Sys.Date(),"_Muscat_EdgeR_oligodendrocytes_genes_statistics.txt"),
sep = "\t", row.names = F, quote=F)
write.table(tbl[[6]][,1:6], file =
paste0(Sys.Date(),"_Muscat_EdgeR_microglia_genes_statistics.txt"), sep =
"\t", row.names = F, quote=F)
# Save pseudobulk object and results
save(pb.hypo.webb, res, file =
paste0(Sys.Date(),"_SCE_with_Muscat_Pseudobulk_webb_hypo_object.RData"))
```

Finally, we can save our psedobulked object as a list with each distinct cell type as an element. This format will be helpful for our downstream post-pseudobulking data analysis.

```
# save the pseudobulked object as a list with bulk gene expression for each
cell type
pb.list <- lapply(assayNames(pb.hypo.webb), function(ct) {
assay(pb.hypo.webb, ct)
})
names(pb.list) <- assayNames(pb.hypo.webb)
save(pb.list, file = paste0(Sys.Date(),
"_hypo_webb_pseudobulk_by_celltype_list.rds"))
```

11.2 Single-Cell Pseudobulk Data Analysis

After pseudobulking your single-cell datasets by each cell type in the dataset, you now essentially have purified these cell types in silico. In this section, we will run through removing batch effect with SVA, getting differentially expressed genes between old and young samples with DESeq2, and visualizing separation between samples with PCA and MDS.

Getting Normalized Counts for a Given Cell Type

For simplicity, we will only show analysis pertaining to one cell type (microglia) going forward. First, we can grab the normalized counts for a given cell type by using `getVarianceStabilizedData()`.

```
# Get normalized counts
tissue.cts <- getVarianceStabilizedData(dds.deseq)
write.csv(as.data.frame(tissue.cts),file = paste0(Sy
# object corresponding to the microglia cell type - pb.list$microglia
my.matrix <- pb.list$Microglia
```

Removing Technical Noise Effects Using SVA

Single-cell and bulk RNA-seq datasets often contain unwanted technical variation due to batch effects, library preparation differences, sequencing depth, or sample processing. When these sources of variation are not fully captured by known covariates, they can obscure or confound true biological signal and, in some cases, reduce statistical power in differential expression analyses. You can determine if your dataset has identifiable technical variation by using PCA or MDS labeled by a known technical variable such as batch, sample date, or sequencing run to identify separation due to a known covariate. In addition to known covariates, it is also possible to estimate latent, unmeasured sources of technical variation using Surrogate Variable Analysis (SVA). SVA identifies surrogate variables that capture variation in the expression data not explained by the experimental design. These surrogate variables can then be incorporated into the differential expression model, allowing technical noise to be accounted for while preserving the biological contrast of interest. By modeling both known and hidden sources of variation, SVA helps reduce noisy, non-biological associations and improves the robustness of downstream differential expression results.

Next, we will remove lowly expressed genes and set up the design matrix for SVA and DESeq2. We will use the same design matrix for both (i) estimating technical noise with SVA and (ii) calculating differentially expressed genes between old and young samples. Make sure to set the young group as the reference to identify differentially expressed genes *with age* (so that the final sign matches the change with increasing age).

```
# Remove low read depth samples
colSums(my.matrix)
# YF1YF2YF3YF4OF1OF2OF3OF4
# 81208133372100499399590361247162163266650307486
my.best <- colSums(my.matrix) > 50000 # keep all samples with >50,000 counts
my.filtered.matrix <- my.matrix[, my.best]
# Build design matrix
dataDesign <- data.frame(
 row.names = colnames(my.filtered.matrix),
 age = factor(c("Y", "Y", "Y", "Y", "O", "O", "O", "O")))
# set young samples as the reference group
dataDesign$age <- relevel(dataDesign$age, ref = "Y")
# SVA processing - remove batch effect
mod1 <- model.matrix(~ age, data = dataDesign)
n.sv.be <- num.sv(my.filtered.matrix, mod1, method = "be")
if (n.sv.be > 0) {
 my.svseq <- svaseq(as.matrix(my.filtered.matrix), mod1, n.sv = n.sv.be,
constant = 0.1)
 my.clean <- removeBatchEffect(
  log2(my.filtered.matrix + 0.1),
  batch = NULL,
  covariates = cbind(my.svseq$sv),
  design = mod1)
 my.filtered.sva <- round(2^my.clean - 0.1)
} else {
 my.filtered.sva <- my.filtered.matrix
}
```

Differential Gene Expression Analysis with DESeq2

Now, we perform differential gene expression analysis with DESeq2, modeling the changes in gene expression between old and young samples. The normalized counts will be used later for MDS and PCA analysis.

```
dds <- DESeqDataSetFromMatrix(
```

```
 countData = my.filtered.sva,
 colData = dataDesign,
 design = ~ age)
dds.deseq <- DESeq(dds)
# Get normalized counts
tissue.cts <- getVarianceStabilizedData(dds.deseq)
write.csv(as.data.frame(tissue.cts),file = paste0(Sys.Date(),
"_microglia_norm_counts.csv"))
# DEGs for age (Y vs O)
res <- results(dds.deseq, contrast = c("age", "O", "Y")) # what changes in
old relative to young (reference)
microglia.degs <- as.data.frame(res)
write.csv(microglia.degs, file = paste0(Sys.Date(), "_microglia_degs.csv"))
```

PCA and MDS Analysis

An important part of the analysis is to identify how samples separate in a 2D visualization to identify the main drivers of biological variance in the experiment through dimensionality reduction. This can be done using either PCA (Fig. 16A) or MDS (Fig. 16B) depending on user preference. These techniques help us visualize transcriptional variation between samples.

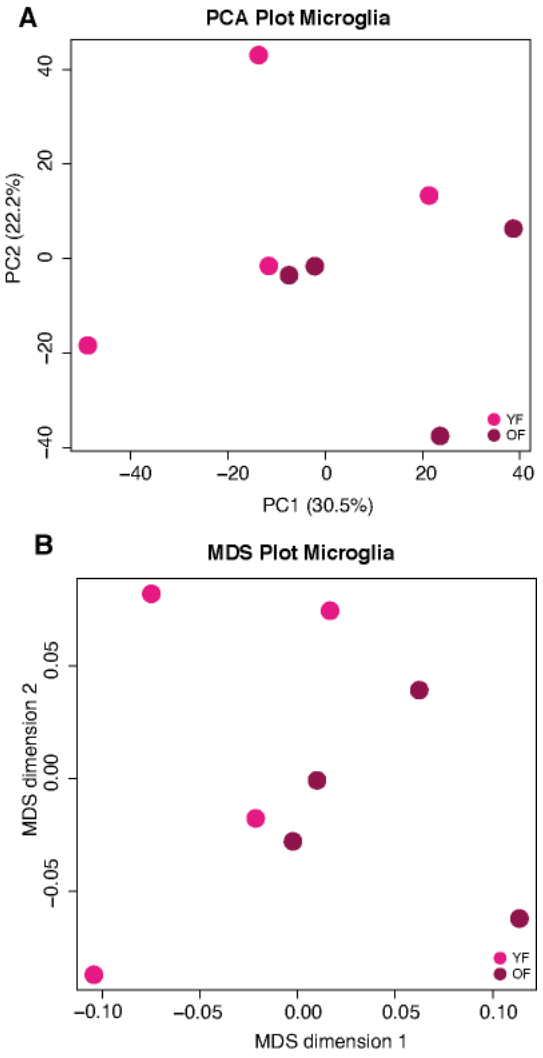


Fig. 16 Post-pseudobulk separation of samples after correcting for technical variation using SVA. **A,** PCA plot and (**B**) MDS plot showing sample separation. Dark pink indicates old females (OF) and light pink indicates young females (YF)

```
### MDS analysis
mds.result <- cmdscale(1-cor(tissue.cts,method="spearman"), k = 2, eig =
FALSE, add = FALSE, x.ret = FALSE)
x <- mds.result[, 1]
y <- mds.result[, 2]
my.colors <- ifelse(dataDesign$age == "Y", "deeppink1", "deeppink4")
plot(x, y,
  xlab = "MDS dimension 1", ylab = "MDS dimension 2",
  main = "MDS Plot",
  cex = 3, lwd = 5, pch=16,
  col = my.colors,
  xlim = c(-0.3,0.3),
  ylim = c(-0.3,0.3))
legend("bottomright",c('YF','OF'),col=c("deeppink1", "deeppink4"), bty='n',
pch = 16, pt.cex=2)
### PCA on variance-stabilized counts
pca.result <- prcomp(t(tissue.cts), scale. = TRUE)
# Extract PC1 and PC2
x <- pca.result$x[,1]
y <- pca.result$x[,2]
# Colors for plotting
my.colors <- ifelse(dataDesign$age == "Y", "deeppink1", "deeppink4")
# Plot
plot(x, y,
  xlab = paste0("PC1 (", round(100 * summary(pca.result)$importance[2,1], 1),
"%)"),
  ylab = paste0("PC2 (", round(100 * summary(pca.result)$importance[2,2], 1),
"%)"),
  main = "PCA Plot",
  cex = 3, lwd = 5, pch = 16,
  col = my.colors)
legend("bottomright", c("YF","OF"), col=c("deeppink1","deeppink4"), bty="n",
pch=16, pt.cex=2)
```

Other Bulk-Level Analyses

From this point on, any method developed to analyze bulk transcriptomic experiments can be used to analyze the data, as pseudobulked DESeq2 results will be compatible with any such workflow. This includes functional enrichment analysis methods such as Gene Set Enrichment Analysis (GSEA) and Over-representation analysis (ORA).

---

12 Notes

1. For simplicity's sake or when using publicly available data missing it, Cell Ranger can process demultiplexed FASTQs without the Index sequences. Naming conventions should otherwise be the same as outlined above.
2. The flags—`include-introns`—`create-bam`—`expect-cells` are optional while using Cell Ranger (v10.0.0) as they are redundant, the version includes this by default. In addition, the writing system of Cell Ranger might be incompatible with older servers (this can be solved by using the flag—`disable-ui`). If you have reason to believe that the automatically detected cell number may be

erroneous based on the appearance of the barcode rank plot, consider using the—`force-cells` to match the desired number.

3. In our experience, values of rho (contamination) typically range ~0.02–0.20 in our experience. Higher values are expected from samples with poor cell viability or from nuclei, as RNA leakage is unavoidable in these cases. However, any sample with >25% contamination (rho > 0.25) should be discarded to avoid substantial ambient RNA-driven artifacts, including spurious gene expression.
4. For this analysis, chemistry and biological condition were also added to the metadata (see GitHub script).
5. In our experience, biological samples rarely separate by age group at this stage, as cell type differences will tend to dwarf the effect of aging on cell transcriptomes. In fact, separation of biological groups at this stage may indicate strong batch effects, which may not be correctable. Analysis of independent sets of replicate samples is crucial at this stage to confirm appropriate sample mixing in the UMAP space.
6. This particular data set includes samples processed with both 10× v2 [46] and v3 [47] chemistries. Because of this, predicted doublet rates are calculated separately for each library type based on the expected number of captured cells. These rates are then assigned to each sample to inform downstream doublet detection. Please use the most up-to-date doublet rates based on the chemistry being used.
7. In our experience, it is crucial to use at least two doublet detection tools to ensure that true doublets are reliably identified and removed. In this workflow, we use DoubletFinder and scds, which we have regularly implemented and found to produce reproducible and robust results. These have the most different approaches to doublet finding and are thus complementary in doublet detection efforts. While many other tools exist for the detection of doublets in sc/snRNAseq datasets, such as scDblFinder [48] and DoubletDecon [49], which use alternative detection strategies, they are not explored in this workflow, but could be applied depending on specific dataset requirements, user preference or if new, better algorithms are developed in the future.
8. In our lab, we find that saving the doublet-annotated R object is a useful practice. This is mainly due to DoubletFinder taking a significant amount of time to run (for e.g., in this dataset, ~1 h). Saving the object at this stage prevents the need to rerun DoubletFinder if any issues or crashes arise later in the workflow.
9. In our experience, using ≥3 robust canonical markers per cell type (selected from the literature, prior single-cell studies, or databases such as PanglaoDB [39]) provides adequately reliable cell-type annotation for clusters in well-annotated species and can be used to assess overall data quality at this stage.
10. We have found that the choice of reference dataset strongly influences SingleR performance. Using a reference that matches the tissue type, developmental stage, and species of your dataset improves annotation accuracy and reduces mislabeling. If a reliable reference is not accessible, or if the goal is to detect rare cell types or subpopulations, marker-based annotation provides an alternative approach to ensure accurate identification.
11. We refine SingleR annotations using cluster-level majority voting to lower inconsistent labels within clusters. Using a lower clustering resolution (e.g., 0.3) may group rare cell types or subpopulations within larger populations, causing them to be lost during refinement; this is not an issue if the goal is to define major cell types. In contrast, higher resolution (e.g., 1.0) can separate these rare cell types or subpopulations into their own clusters, allowing them to be retained during cluster-level labeling. Therefore, clustering resolution should be chosen with both biological question and the cell types or

subpopulations of interest in mind and can be modified using the `FindClusters` function immediately after doublet removal.

12. In our lab, we have found that integration performs best when datasets have substantial overlap in cell types or states. In addition to strong overlap, careful experimental design and high-quality libraries remain essential to limit technical variation and guarantee that integrated data accurately reflect true biological differences.

13. In our experience, the optimal integration method often varies between datasets and can even vary within a dataset depending on the specific biological question, which shows there is no "golden standard" or one-size-fits-all approach. Due to this, if integration is needed in your dataset, it is important to compare multiple integration techniques to determine which method best corrects technical variation while preserving biologically meaningful structure [50]. In addition, integration may be run before cell annotation as well, which can help tighten biological clusters in more complex tissues.

14. Ensure that all packages used in this script correspond to the versions specified in the Materials section. Using alternative package versions may lead to errors or unexpected results, potentially affecting downstream analyses and compromising reproducibility.

15. If you encounter unexpected errors while running this script, one should first look at relevant documentation and official tutorials for the software package. Other potential solutions include searching the error or warning message in Google or using AI platforms. Stack Overflow and Biostars are valuable resources that may provide helpful answers. In some instances, the fix may be as simple as restarting your R program session.

16. Although not directly covered in this chapter, the easiest way to download raw FASTQ files from publicly available datasets is through tools with a graphical user interface such as SRA Explorer. Note, however, that this approach may not be applicable to all datasets. One can also use command-line tools such as SRA Toolkit, which are easier to use with increasing numbers of samples.

17. It is generally good practice to inspect the generated/downloaded sequencing files (.fastq.gz) to ensure they are not truncated or corrupted. Inspection typically involves verifying successful decompression, confirming expected file sizes, and checking for proper FASTQ formatting (four lines per read), all of which may indicate file integrity. Additionally, the quality of reads can also be further investigated with QC programs (i.e., FastQC).

18. Parts of this script include the use of computationally heavy tools (doublet finder, some integration tools, SingleR). We recommend increasing R's allowable global object size [e.g., `options(future.globals.maxSize = 30e9)`] to avoid memory allocation errors.

19. We recommend saving Seurat objects after each key memory-intensive step and routinely inspecting metadata and metrics throughout the workflow. This ensures that you can restart downstream analyses without recomputing everything.

20. Rasterization (by using `raster = TRUE`) transforms many individual vector points into a single pixel-based image, which greatly reduces file size and rendering time. For dense plots like UMAPs with hundreds/thousands of cells, rasterizing the points keeps figures manageable and fast while conserving vector quality for labels, annotations, and axes.

Competing Interests

The authors have no competing interests.

References

1. Guo J, Huang X, Dou L et al (2022) Aging and aging-related diseases: from molecular mechanisms to interventions and treatments. Signal Transduct Target Ther 7(1):391. https://doi.org/10.1038/s41392-022-01251-0 [PubMed][PubMedCentral]

2. Tabula Muris C (2020) A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. Nature 583(7817):590–595. https://doi.org/10.1038/s41586-020-2496-1

3. Yu M, Zhang H, Wang B et al (2021) Key signaling pathways in aging and potential interventions for healthy aging. Cells 10(3):660. https://doi.org/10.3390/cells10030660

4. Jovic D, Liang X, Zeng H et al (2022) Single-cell RNA sequencing technologies and applications: a brief overview. Clin Transl Med 12(3):e694. https://doi.org/10.1002/ctm2.694 [PubMed][PubMedCentral]

5. Fonseca Costa SS, Robinson-Rechavi M, Ripperger JA (2020) Single-cell transcriptomics allows novel insights into aging and circadian processes. Brief Funct Genomics 19(5-6):343–349. https://doi.org/10.1093/bfgp/elaa014 [PubMed][PubMedCentral]

6. Hajdarovic KH, Yu D, Hassell LA et al (2022) Single-cell analysis of the aging female mouse hypothalamus. Nat Aging 2(7):662–678. https://doi.org/10.1038/s43587-022-00246-4 [PubMed][PubMedCentral]

7. Zheng GX, Terry JM, Belgrader P et al (2017) Massively parallel digital transcriptional profiling of single cells. Nat Commun 8: 14049. https://doi.org/10.1038/ncomms14049 [PubMed][PubMedCentral]

8. Skinnider MA, Squair JW, Kathe C et al (2021) Cell type prioritization in single-cell data. Nat Biotechnol 39(1):30–34. https://doi.org/10.1038/s41587-020-0605-1 [PubMed]

9. Campbell J, Yang S, Wang Z et al (2025) celda: CEllular latent dirichlet allocation. v.1.22.1 edn. Bioconductor https://doi.org/10.18129/B9.bioc.celda

10. Aran D, Looney AP, Liu L et al (2019) Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol 20(2):163–172. https://doi.org/10.1038/s41590-018-0276-y [PubMed][PubMedCentral]

11. Yu G, Wang LG, Han Y et al (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16(5):284–287. https://doi.org/10.1089/omi.2011.0118 [PubMed][PubMedCentral]

12. Zappia L, Oshlack A (2018) Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience 7(7):giy083. https://doi.org/10.1093/gigascience/giy083

13. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550. https://doi.org/10.1186/s13059-014-0550-8 [PubMed][PubMedCentral]

14. McGinnis CS, Murrow LM, Gartner ZJ (2019) DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst 8(4):329–337. https://doi.org/10.1016/j.cels.2019.03.003 [PubMed][PubMedCentral]

15. Wickham H, François R, Henry L et al (2023) dplyr: a grammar of data manipulation, v.1.1.4 edn. R Foundation for Statistical Computing

16. Wickham H, Chang W, Henry L et al (2023) ggplot2: elegant graphics for data analysis. Posit, PBC,

17. Morgan, M, Falcon, S, Gentleman, R, 2025. GSEABase: gene set enrichment data structures and methods, v.1.70.1 edn. Bioconductor. https://doi.org/10.18129/B9.bioc.GSEABase

18. Korsunsky I, Millard N, Fan J et al (2019) Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16(12):1289–1296. https://doi.org/10.1038/s41592-019-0619-0 [PubMed][PubMedCentral]

19. Ritchie ME, Phipson B, Wu D et al (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43(7):e47. https://doi.org/10.1093/nar/gkv007 [PubMed][PubMedCentral]

20. Crowell HL, Soneson C, Germain PL et al (2020) muscat detects subpopulation-specific state transitions from multi-sample multi-condition single-cell transcriptomics data. Nat Commun 11(1):6077. https://doi.org/10.1038/s41467-020-19894-4 [PubMed][PubMedCentral]

21. Carlson M org.Mm.eg.db: genome-wide annotation for Mouse, v.3.21.0 edn. Bioconductor. https://bioconductor.org/packages/release/data/

annotation/html/org.Mm.eg.db.html. https://doi.org/10.18129/B9.bioc.org.Mm.eg.db

22. Pedersen TL (2020) patchwork: the composer of plots, v.1.3.0 edn. R Foundation for Statistical Computing

23. Hao Y, Hao S, Andersen-Nissen E et al (2021) Integrated analysis of multimodal single-cell data. Cell 184(13):3573–3587. https://doi.org/10.1016/j.cell.2021.04.048
[PubMed][PubMedCentral]

24. Wickham H, Pedersen TL, Seidel D (2025) scales: scale functions for visualization, v.1.4.0 edn. Posit, PBC

25. McCarthy DJ, Campbell KR, Lun AT et al (2017) Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 33(8):1179–1186. https://doi.org/10.1093/bioinformatics/btw777
[PubMed][PubMedCentral]

26. Bais AS, Kostka D (2020) scds: computational annotation of doublets in single-cell RNA sequencing data. Bioinformatics 36(4):1150–1158. https://doi.org/10.1093/bioinformatics/btz698
[PubMed][PubMedCentral]

27. Miller SA, Policastro RA, Sriramkumar S et al (2021) LSD1 and aberrant DNA methylation mediate persistence of enteroendocrine progenitors that support BRAF-mutant colorectal cancer. Cancer Res 81(14):3791–3805. https://doi.org/10.1158/0008-5472.CAN-20-3562
[PubMed][PubMedCentral]

28. Hafemeister C, Satija R (2019) Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 20(1):296. https://doi.org/10.1186/s13059-019-1874-1
[PubMed][PubMedCentral]

29. Hao Y, Stuart T, Kowalski MH et al (2024) Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol 42(2):293–304. https://doi.org/10.1038/s41587-023-01767-y
[PubMed]

30. Amezquita RA, Lun ATL, Becht E et al (2020) Orchestrating single-cell analysis with Bioconductor. Nat Methods 17(2):137–145. https://doi.org/10.1038/s41592-019-0654-x
[PubMed]

31. Garcia-Comendador J, Martinez-Carreras N, Fortesa J et al (2021) Combining sediment fingerprinting and hydro-sedimentary monitoring to assess suspended sediment provenance in a mid-mountainous Mediterranean catchment. J Environ Manage 299:113593. https://doi.org/10.1016/j.jenvman.2021.113593
[PubMed]

32. Young MD, Behjati S (2020) SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. Gigascience 9(12):giaa151. https://doi.org/10.1093/gigascience/giaa151

33. Leek JT, Johnson WE, Parker HS et al (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28(6):882–883. https://doi.org/10.1093/bioinformatics/bts034
[PubMed][PubMedCentral]

34. Andreatta M, Carmona SJ (2021) UCell: robust and scalable single-cell gene signature scoring. Comput Struct Biotechnol J 19:3796–3798. https://doi.org/10.1016/j.csbj.2021.06.043
[PubMed][PubMedCentral]

35. 10x Genomics I (2025) Cell Ranger reference downloads. https://www.10xgenomics.com/support/software/cell-ranger/downloads#reference-downloads

36. Janssen P, Kliesmete Z, Vieth B et al (2023) The effect of background noise and its removal on the analysis of single-cell expression data. Genome Biol 24(1):140. https://doi.org/10.1186/s13059-023-02978-x
[PubMed][PubMedCentral]

37. Yang S, Corbett SE, Koga Y et al (2020) Decontamination of ambient RNA in single-cell RNA-seq with DecontX. Genome Biol 21(1):57. https://doi.org/10.1186/s13059-020-1950-6
[PubMed][PubMedCentral]

38. Tirosh I, Izar B, Prakadan SM et al (2016) Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 352(6282):189–196. https://doi.org/10.1126/science.aad0501
[PubMed][PubMedCentral]

39. Franzen O, Gan LM, Bjorkegren JLM (2019) PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database (Oxford) 2019:baz046. https://doi.org/10.1093/database/baz046

40. Guo H, Li J (2021) scSorter: assigning cells to known cell types according to marker genes. Genome Biol 22(1):69. https://doi.org/10.1186/s13059-021-02281-7
[PubMed][PubMedCentral]

41. Shao X, Liao J, Lu X et al (2020) scCATCH: automatic annotation on cell types of clusters from single-cell RNA sequencing data. iScience 23(3):100882. https://doi.org/10.1016/j.isci.2020.100882
[PubMed][PubMedCentral]

42. Zhang Y, Zhang F, Wang Z et al (2022) scMAGIC: accurately annotating single cells using two rounds of reference-based classification. Nucleic Acids Res 50(8):e43. https://doi.org/10.1093/nar/gkab1275
[PubMed][PubMedCentral]

43. Benayoun BA, Pollina EA, Singh PP et al (2019) Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses. Genome Res 29(4):697–709. https://doi.org/10.1101/gr.240093.118
[PubMed][PubMedCentral]

44. Lab S (2023) Integrative analysis in Seurat v5 (Integration vignette). https://satijalab.org/seurat/articles/seurat5_integration

45. Saul D, Kosinsky RL, Atkinson EJ et al (2022) A new gene set identifies senescent cells and predicts senescence-associated pathways across tissues. Nat Commun 13(1):4827. https://doi.org/10.1038/s41467-022-32552-1
[PubMed][PubMedCentral]

46. 10x Genomics I (2017) Chromium Single Cell 3' Reagent Kits v2 User Guide. https://www.10xgenomics.com/support/user-guides

47. 10x Genomics I (2022) Chromium Next GEM Single Cell 3' Reagent Kits v3.1 User Guide. https://www.10xgenomics.com/support/user-guides

48. Germain PL, Lun A, Garcia Meixide C et al (2021) Doublet identification in single-cell sequencing data using scDblFinder. F1000Res 10:979. https://doi.org/10.12688/f1000research.73600.2
[PubMed][PubMedCentral]

49. DePasquale EAK, Schnell DJ, Van Camp PJ et al (2019) DoubletDecon: deconvoluting doublets from single-cell RNA-sequencing data. Cell Rep 29(6):1718–1727. https://doi.org/10.1016/j.celrep.2019.09.082
[PubMed][PubMedCentral]

50. Luecken MD, Buttner M, Chaichoompu K et al (2022) Benchmarking atlas-level data integration in single-cell genomics. Nat Methods 19(1):41–50. https://doi.org/10.1038/s41592-021-01336-8
[PubMed]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_3

3. Isolation and Functional Screening of Plant-Derived Natural Products for Assessing Healthy Aging and Neuroprotective Activities in *Caenorhabditis elegans*

Chatrawee D. Shepard[1, 2] ✉
(1) The Appalachian Natural Products Research Program, Department of Biomedical Sciences, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV, USA
(2) Center of Excellence on Natural Products for Neuroprotection and Anti-ageing (Neur-Age NatChula), Department of Clinical Chemistry, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand

✉ **Chatrawee D. Shepard**
Email: chatrawee.shepard@marshall.edu

Abstract
Plant-derived natural products are a rich source of bioactive molecules with potential benefits for healthy aging and neuroprotection. Identifying compounds with these activities requires integrated platforms capable of evaluating multiple aspects of organismal health. *Caenorhabditis elegans* provides a genetically tractable, physiologically conserved in vivo model ideally suited for functional screening of natural products. This protocol presents a comprehensive workflow for

preparing plant extracts, characterizing their antioxidant properties, and assessing their effects on stress resistance, healthspan, lifespan, and neuroprotective effects in *C. elegans*. Methods begin with plant extraction and routine maintenance of nematode and bacterial cultures to ensure consistency across experiments. Antioxidant activity is assessed through oxidative stress resistance assays, intracellular reactive oxygen species (ROS) quantification, stress-responsive reporter activation, and subcellular localization of the transcription factors DAF-16/FOXO and SKN-1/Nrf2. Healthspan assessments include lipofuscin accumulation, pharyngeal pumping, reproductive function, growth, and locomotion. Longevity effects are determined using lifespan assays. Neuroprotective potential is examined in models of β-amyloid/tau toxicity, chemotaxis impairment, polyglutamine aggregation, and α-synuclein aggregation relevant to Alzheimer's, Huntington's, and Parkinson's diseases. Collectively, this platform enables systematic functional screening of plant-derived natural products and supports the discovery of compounds that promote stress resistance, improve healthspan, and mitigate proteotoxic neurodegeneration.

Keywords *C. elegans* – Plant-derived natural products – Healthy aging – Antioxidant activity – Stress resistance – Healthspan – Lifespan – Neuroprotection – DAF-16 – SKN-1 – Protein aggregation

1 Introduction

The discovery of plant-derived natural products that promote healthy aging and protect against age-associated diseases is an active area of biomedical research [1, 2]. Plants produce structurally diverse metabolites with antioxidant [3–6], anti-inflammatory [7], and neuroprotective properties [8–13], many of which remain unexplored for their potential roles in aging [3–12, 14]. However, identifying bioactive compounds requires integrated, scalable screening platforms capable of assessing metabolic, physiological, and molecular endpoints

in vivo. Although mammalian models provide high physiological relevance, they are labor-intensive, costly, and unsuitable for early-stage, high-content screening of large natural product libraries, highlighting the need for efficient alternative model systems [15].

Caenorhabditis elegans is a powerful model organism for aging research due to its short lifespan, genetic tractability, and conserved pathways regulating stress resistance, metabolism, proteostasis, and neurodegeneration [16, 17]. Reporter strains provide real-time visualization of oxidative stress responses and transcription factor activity, including DAF-16/FOXO [18] and SKN-1/Nrf2 [19], which regulate longevity and detoxification pathways [18, 19]. Healthspan can be assessed through multiple physiological metrics such as lipofuscin accumulation, pharyngeal pumping rate, reproductive output, growth, and locomotor function, while lifespan assays provide a direct measure of longevity effects [17, 20]. In addition, *C. elegans* models of proteotoxicity enable rapid functional evaluation of neuroprotective properties relevant to human neurodegenerative disorders. These include assays for β-amyloid/tau-induced paralysis in Alzheimer's disease models [21–23], α-synuclein-associated dopaminergic dysfunction in Parkinson's disease models [24], and aggregation of expanded polyglutamine repeats associated with Huntington's disease [25].

This protocol presents an integrated workflow for isolating plant extracts and systematically assessing their antioxidant, healthspan, lifespan, and neuroprotective activities using *C. elegans*, providing a comprehensive and reproducible approach for identifying promising natural products. Together, these methods support the discovery and mechanistic characterization of plant-derived compounds that influence aging biology and may ultimately guide therapeutic development.

---

2 Materials

2.1 *C. elegans* Strains

All *C. elegans* strains and *E. coli* OP50 [26] originate from the Caenorhabditis Genetics Center (CGC), University of Minnesota, USA, the following strains were utilized: N2 Bristol wild-type (N2), TJ375 (*gpIs1[hsp-16-2::GFP*]), CF1553 (*muls84[pAD76(sod-3::GFP)*]), CF1038 (*daf-16[mu86]I*), BA17 (*fem-1[hc17]IV*), EU1 (*skn-1[zu67]*), CL2166 *([pAF15]gst-4p::GFP::NLS*), LD1(*ldls7*), CL802 (smg-1(cc546) I; rol-6(su1006) II), TJ356 (*zIs356 [daf-16p::daf-16a/b::GFP+rol-6]*), *daf-16 (mgDf47)*, CL4176 (*smg-1(cc546) I; dvIs27 [(myo-3p::A-Beta (1–42)::let-851 3'UTR) + rol-6(su1006)] X*), CL2006 (*dvIs2 [pCL12(unc-54/human Abeta peptide 1-42 minigene) + pRF4]*), CL2355 (*smg-1(cc546) dvIs50 [pCL45 (snb-1::Abeta 1–42::3' UTR(long) + mtl-2::GFP] I)*, CL2122 (*dvIs15 [(pPD30.38) unc-54(vector) + (pCL26) mtl-2::GFP]*), KAE112 *(seaIs201 [myo-3p::human tau (0N4R;V337M)::unc-54 3'UTR + vha-6p::mCherry::unc-54 3'UTR]*, AM134 (*rmIs126 [unc-54p::Q20::YFP]*), AM138 (*rmIs126 [unc-54p::Q24::YFP]*), and AM140 (*rmIs126 [unc-54p::Q35::YFP]*.

2.2 Media and Reagents

Prepare all solutions using ultrapure water and analytical-grade reagents. Unless otherwise specified, prepare and store all reagents at room temperature. Follow institutional and governmental chemical and biological waste disposal regulations when discarding any waste materials.

1. 10× M9 buffer (1 L): 30 g KH_2PO_4, 60 g Na_2HPO_4, 50 g NaCl, 10 ml 1 M $MgSO_4$, bring to 1 L with ultrapure water, autoclave or filter-sterilize.
2. NGM agar (1 L): 3 g NaCl, 17 g agar, 2.5 g peptone, 950 ml ultrapure water, autoclave and cool to 55 °C, add 1 ml 1 M $MgSO_4$, 1 ml 5 mg/ml cholesterol, 25 ml 1 M KH_2PO_4, 1 ml 1 M $CaCl_2$, 7.5 ml 2.5% streptomycin.
3. 1 M KH_2PO_4/K_2HPO_4 buffer (2 L): 60.2 g $K_2HPO_4 \cdot 3\ H_2O$, 236.2 g KH_2PO_4, bring to 2 L with ultrapure water, autoclave or filter-sterilize.

4. 5 mg/ml cholesterol: 0.5 g cholesterol dissolved in 100 ml 100% ethanol; store at 4 °C.
5. 2.5% streptomycin: 25 g streptomycin sulfate in ultrapure water to 1 L; store at 4 °C.
6. LB medium, liquid (250 ml): 5.5 g LB broth, 7.7 mg K_2HPO_4 (0.35%), ultrapure water to 250 ml, autoclave to sterilize.
7. LB broth (1 L): 22 g LB broth, 7.7 mg K_2HPO_4 (0.35%), ultrapure water to 1 L, autoclave to sterilize.
8. LB agar plates (1 L): 22 g LB broth, 20 g agar, 7.7 mg K_2HPO_4 (0.35%), ultrapure water to 1 L, autoclave to sterilize, cool to 55 °C, add 7.5 ml 2.5% streptomycin.
9. S-medium (1 L): 1 L S-basal, 3 ml 1 M $CaCl_2$, 3 ml 1 M $MgSO_4$, 10 ml 1 M potassium citrate (pH 6), 10 ml trace metals solution.
10. S-basal (1 L): 5.85 g NaCl, 1 g K_2HPO_4, 6 g KH_2PO_4, 1 ml 5 mg/ml cholesterol, ultrapure water to 1 L, autoclave to sterilize.
11. Trace metals solution (1 L): 1.86 g disodium EDTA, 0.69 g $FeSO_4 \cdot 7 H_2O$, 0.2 g $MnCl_2 \cdot 4 H_2O$, 0.29 g $ZnSO_4 \cdot 7 H_2O$, 0.025 g $CuSO_4 \cdot 5 H_2O$, ultrapure water to 1 L, autoclave to sterilize.
12. 1 M potassium citrate, pH 6 (1 L): 20 g citric acid monohydrate, 293.5 g tripotassium citrate monohydrate, ultrapure water to 1 L.
13. Bleaching solution: 37.5 ml autoclaved ultrapure water, 3 ml 50% NaOH, 10 ml bleach (freshly prepared).

3 Methods

3.1 Plant Material and Extraction

Purpose: To obtain crude plant extracts containing bioactive natural products for biological screening.

1. Collect different parts of the plant, including leaves, stems, roots, flowers, and fruits.
2. Photograph the plant and record relevant information, including the collector's name, date, time, and location of collection.
3. Submit the collected plant materials to the Department of Botany for taxonomic identification and preparation of a voucher specimen.
4. Gently wash the plant samples with deionized (DI) water two to three times until all visible dirt and debris are removed.
5. Air-dry the cleaned samples indoors for 1–2 weeks, avoiding direct sunlight to prevent degradation of bioactive compounds. Ensure samples are completely dry before extraction.
6. Weigh 40 g of dried plant material and perform sequential extraction using a Soxhlet apparatus with 400 mL of hexane, dichloromethane, and methanol (maintaining a sample-to-solvent ratio of 1:10). Reflux the sample in each solvent for 36 h, allowing complete exhaustive extraction before proceeding to the next solvent.
7. Filter each solvent extract using Whatman No. 1 filter paper to remove particulates.
8. Evaporate the filtrates at 35–45 °C using a rotary evaporator until all solvents are removed.
9. Collect crude extracts and store at −20 °C until further analysis.

3.2 *C. elegans* and *E. coli* Culture

Purpose: To maintain healthy nematode and bacterial cultures under standardized conditions.

Culture conditions: *C. elegans* are maintained in an incubator or temperature-controlled room at 20 °C (unless otherwise specified).

A. **Worm Maintenance and Culture**

1. Prepare NGM agar plates seeded with live *E. coli* OP50 as the food source.
2. Culture *C. elegans* on *E. coli* OP50-seeded plates until ready for synchronization.
3. For liquid cultures, prepare S-medium and mix with *E. coli* OP50 adjusted to OD_{600} = 1.0.
4. Transfer worms into S-medium when liquid culture experiments are required.

B. **Age Synchronization by Hypochlorite Treatment**

1. Wash gravid adults from NGM plates using 2 mL M9 buffer and transfer the suspension into a 15-mL conical tube (*see* **Notes 1 and 2**).
2. Centrifuge at 700 RCF for 30–60 s and carefully discard the supernatant.
3. Add M9 buffer to bring the total volume to 3 mL.
4. Add 3 mL of freshly prepared hypochlorite solution to the worm suspension (1:1 ratio).
5. Incubate with gentle agitation (shaker or tube rocker) for 1–5

min, until adult worm bodies are fully dissolved (*see* **Note 3**).

6. Add 6 mL M9 buffer immediately to quench the bleaching reaction.
7. Centrifuge at 700 RCF for 1 min to collect eggs.
8. Carefully discard the liquid, leaving approximately 1 mL of M9 above the egg pellet. Avoid disturbing the eggs.
9. Add 12–14 mL M9 buffer, gently resuspend the pellet, and centrifuge at 700 RCF for 30–60 s.
10. Carefully discard the liquid while keeping the egg pellet intact.
11. Wash eggs three times total as described above.
12. Resuspend the cleaned egg pellet in 6 mL M9 buffer.
13. Place the tube on a rotator overnight to allow eggs to hatch into synchronized L1 larvae under starvation conditions.
14. Centrifuge at 700 RCF for 30–60 s and carefully remove the liquid, leaving the L1 pellet intact.
15. Transfer L1 larvae onto NGM plates seeded with *E. coli* OP50 for synchronized growth.

C. **Treatment Conditions**

C1. Agar-based treatment

1. Prepare plant extracts at the desired concentrations in M9 buffer.
2. Pipette the extract solution evenly onto the *E. coli* OP50 lawn on NGM agar plates.
3. Allow plates to dry overnight at room temperature

3. Allow plates to dry overnight at room temperature.

4. For the vehicle control, prepare plates using the solvent control in M9 buffer (*see* **Note 4**).

5. Transfer synchronized worms onto extract-supplemented NGM plates (*see* **Note 5**).

6. Incubate worms under standard experimental conditions (temperature according to strain).

C2. Liquid treatment (S-Medium)

1. Prepare plant extracts at desired final concentrations in S-medium containing *E. coli* OP50 (OD600 = 1.0) (*see* **Note 6**).

2. Mix thoroughly to ensure uniform exposure.

3. Add synchronized worms to treatment S-medium for experiments conducted in liquid culture.

D. ***E. coli* OP50 Culture**

D1. Preparing an *E. coli* OP50 streak plate from frozen stock

1. Set up a sterile workspace next to a flame or inside a biosafety cabinet.

2. Label an LB agar plate (no antibiotics) on the bottom with the strain name (OP50) and date.

3. Retrieve the frozen *E. coli* OP50 stock. Using an autoclaved pipette tip or sterile inoculating loop, gently scrape a very small amount of bacteria from the surface of the frozen stock. Do not dig into the frozen material; only a tiny amount is required.

4. Streak a straight inoculation line onto the LB agar surface, then

streak across the plate to spread out cells. Avoid applying excessive pressure that could damage the agar.

5. Incubate the plate agar (side up) at 37 °C overnight (12–16 h) or until colonies or a light lawn appears.

6. Store the streak plate at 4 °C for short-term use (1–2 weeks).

D2. Preparing *E.coli* OP50 liquid culture for NGM plate seeding

1. To prepare LB broth with streptomycin for *E. coli* OP50 liquid culture, flame the neck of the LB bottle before opening.
2. Add 7.5 mL of 2.5% (w/v) streptomycin solution to 1 L LB broth.
3. Flame the bottleneck again and recap securely.
4. To inoculate LB broth with *E. coli* OP50, flame the inoculating loop until red-hot and allow it to cool.
5. Using a sterile inoculation loop, gently collect a small amount of *E. coli* OP50 from the perimeter of the streak plate, where isolated colonies are most accessible. Avoid applying excessive pressure to prevent disruption or gouging of the agar surface.
6. Transfer the bacteria into the LB broth and swirl gently to mix.
7. Flame the loop again before setting it down.
8. Incubate the culture overnight at 37 °C without shaking. The *E. coli* OP50 should grow as a stationary culture to form uniform lawns on NGM plates.
9. A successful culture will appear turbid (cloudy) the next day. Use immediately to seed NGM plates.

3.3 Assessment of Resistance to Oxidative Stress

Purpose: To determine whether plant extracts enhance the ability of *C. elegans* to tolerate oxidative stress induced by juglone, indicating antioxidant and stress-resistance properties.

1. Synchronize worms to the L1 larval stage by following Sect 3.2, Step B.
2. Add the plant extracts to S-medium supplemented with *E. coli* OP50 at the desired final concentrations (*see* **Note 7**).
3. Add 50–100 L1 larvae per group.
4. Incubate worms in S-medium with gentle shaking at 20 °C for 18–24 h.
5. Remove treatment medium and wash worms twice with S-medium.
6. Add 2 mL of 80 μM juglone solution to each group to induce oxidative stress (*see* **Note 8**).
7. Incubate the worms in the juglone-containing S-medium supplemented with *E. coli* and the plant extract treatments for 24 h at 20 °C, ensuring gentle agitation as needed to maintain culture homogeneity.
8. After 24 h juglone exposure, examine worms under a stereomicroscope.
9. Stimulate each worm by gently touching the body with a platinum wire pick or by gently tapping the plate to elicit a behavioral response.
10. Score survival using the following criteria:

10. Score survival using the following criteria:

    Alive: Responds to touch, shows movement or swimming behavior in S-medium.

    Dead: No response; body straight, rigid, and immobile.

11. Record the number of live and dead worms in each group.

12. Calculate survival percentage using the formula: Survival (%) = (Number of live worms / Total worms scored) × 100.

3.4 Quantification of Intracellular ROS

Purpose: To quantify intracellular reactive oxygen species (ROS) levels in *C. elegans* using the fluorescent probe H_2DCF-DA, allowing assessment of antioxidant properties of plant extracts.

1. Synchronize worms and prepare treatment conditions by following Sect. 3.3, Steps 1–5.
2. Pellet worms by centrifugation at 500 RCF for 2 min, and then carefully remove the supernatant.
3. Add 50 µM H_2DCF-DA (prepared fresh in S-medium) and gently resuspend the pellet.
4. Incubate in the dark at 20 °C for 1 h.
5. Remove the dye solution and wash worms twice with fresh S-medium.
6. Centrifuge at 500 RCF for 2 min and carefully remove the supernatant.
7. Add 10 mM levamisole to immobilize worms for imaging.
8. Pipette 2–5 µL of the worm–levamisole mixture onto a pre-prepared agarose pad (*see* **Note 9**).
9. Gently place a coverslip over the immobilized worms (*see* **Note**

10).

10. Randomly select worms from each treatment group and acquire fluorescence images using identical exposure settings and magnification, imaging at least 30 worms per group to ensure statistical robustness.

11. Analyze fluorescence images in Fiji by outlining the entire worm body as the region of interest (ROI) using an appropriate selection tool (e.g., freehand or polygon), and measure the mean fluorescence intensity (MFI) for each worm.

12. Determine the average MFI for each treatment group for statistical comparison.

3.5 Analysis of HSP-16.2, GST-4, and SOD-3 Reporter Expression

Purpose: To assess whether plant extracts modulate key stress-response pathways by measuring activation of the HSP-16.2 (heat-shock response), the GST-4 (SKN-1-mediated phase II detoxification pathway), and the SOD-3 (DAF-16 antioxidant defense pathway).

1. Synchronize worms and prepare treatment conditions by following Sect. 3.3, Steps 1–5 (*see* **Note 11**).

2. Prepare juglone stress solution (if applicable): Freshly prepare 20 µM juglone in S-medium immediately before use.

3. Following plant extract incubation, expose reporter strains to juglone according to the guidelines in **Note 12**.

4. After juglone exposure (or after plant extract incubation for CF1553), remove juglone solution and wash worms twice with S-medium.

5. Pellet worms by centrifugation at 500 RCF for 2 min, and then

carefully remove the supernatant.

6. Prepare worms for fluorescence imaging and quantitative analysis by following Sect. 3.4, Steps 7–12.

3.6 Analysis of DAF-16 and SKN-1 Subcellular Localization

Purpose: To assess whether plant extracts activate the DAF-16/FOXO and SKN-1/Nrf2 stress-response pathways in *C. elegans* by promoting their nuclear translocation, a hallmark of enhanced stress resistance, longevity, and metabolic regulation.

1. Synchronize worms and prepare treatments following Sect. 3.3, Steps 1–5 (*see* **Note 13**).
2. Following the 72-h treatment, remove the medium and wash worms twice with fresh S-medium.
3. Pellet worms by centrifugation at 500 RCF for 2 min, and then carefully remove the supernatant.
4. Prepare worms for fluorescence imaging and quantitative analysis by following Sect. 3.4, Steps 7–12.
5. For each worm, classify DAF-16 or SKN-1 localization into one of the following categories:

Cytosolic: GFP signal predominantly in the cytoplasm; indicates inactive transcription factor. Intermediate/Partial Nuclear: Mixed cytoplasmic and nuclear GFP; indicates moderate activation. Nuclear: GFP signal concentrated in nuclei; indicates strong pathway activation

Assessment of Healthspan Parameters

3.7 Quantification of Autofluorescent Lipofuscin

Purpose: To quantify age-associated autofluorescent lipofuscin accumulation in *C. elegans* as a biomarker of cellular aging and

proteostasis decline [20], where reduced fluorescence in treated worms indicates delayed aging or improved maintenance of cellular homeostasis.

1. Synchronize BA17 worms and prepare treatments following Sect. 3.3, Steps 1–5.
2. Incubate worms in S-medium with gentle shaking at 25 °C, replacing the S-medium (including plant extract or vehicle) every 2 days.
3. Maintain worms until day 8 and day 16 of adulthood.
4. Pellet worms by centrifugation at 500 RCF for 2 min. Carefully remove the supernatant and wash worms twice with S-medium.
5. Prepare worms for fluorescence imaging and quantitative analysis by following Sect. 3.4, Steps 7–12.
6. Image age-related autofluorescent pigment (lipofuscin) using a fluorescence microscope with the following settings:

   Excitation (λex): 360/20 nm.
   Emission (λem): 460/38 nm.

3.8 Assessment of Pharyngeal Pumping Rate

Purpose: To assess neuromuscular function and feeding behavior in aging *C. elegans* by measuring pharyngeal pumping rate, where preservation or enhancement of pumping under treatment conditions indicates improved healthspan.

1. Perform age synchronization using hypochlorite treatment following Sect. 3.2, Step B to isolate embryos.
2. Transfer synchronized L1 larvae onto NGM agar plates seeded with *E. coli* OP50.
   Incubate larvae at the appropriate strain-specific temperature

3. Incubate larvae at the appropriate strain-specific temperature until they reach the late L4 stage (~55 h post-L1 at 20 °C).
4. Pick 15–20 late L4 worms and transfer onto fresh NGM agar plates seeded with *E. coli* OP50.
5. Allow worms to lay eggs for 1–4 h, adjusting the laying period based on the egg-laying rate of the strain, until 100–200 eggs are present on each plate.
6. Remove all adult worms (L4s) from plates to maintain strict synchronization (*see* **Note 14**).
7. Incubate plates under optimal conditions for ~50 h to obtain a synchronized population of L4 larvae.
8. Transfer synchronized L4 worms (20–30 worms per plate) onto treatment plate (NGM plates seeded with *E. coli* OP50 and plant extract at desired concentration), following Sect. 3.2, Step C1.
9. Maintain worms on treatment plates throughout adulthood:

   Days 1–4 (reproductive period): Transfer daily to fresh plates to prevent mixing with progeny.

   Post-reproductive period: Transfer every other day to maintain plate quality and consistent treatment exposure.
10. Assess pharyngeal pumping on days 6, 8, 10, and 12 of adulthood.

    For each worm, count the number of pharyngeal contractions over a 60-s interval and record the pumping rate as pumps per minute (*see* **Note 15**).
11. Analyze at least 20 worms per treatment group at each time point to ensure statistical robustness.

3.9 Assessment of Brood Size

Purpose: To determine whether plant extracts influence reproductive output and whether longevity effects arise independently of reproductive toxicity.

1. Synchronize worms and establish treatment conditions by following Sect. 3.8, Steps 1–7 to obtain synchronized L4 larvae.
2. Transfer individual L4 worms (one worm per plate) onto treatment plate (NGM plates seeded with *E. coli* OP50 and plant extracts at desired concentration), following Sect. 3.2, Step C1.
3. Incubate worms at 20 °C for 24 h to allow them to reach day 1 of adulthood.
4. During the reproductive phase (day 1–day 5 of adulthood), transfer each adult worm daily to a fresh treatment plate. This prevents mixing of progeny across days and maintains consistent exposure to plant extracts.
5. After removing the adult worm, allow the eggs on each plate to hatch and develop to the L3/L4 stage (typically 40–45 h at 20 °C).
6. Count all progeny on each plate using a dissecting microscope. This number represents the daily egg production (brood size) for that time point.
7. Repeat daily progeny counting for 4–5 consecutive days to obtain a complete brood-size profile for each worm.
8. Calculate the mean total brood size for each treatment group based on the cumulative offspring number across all reproductive days.

3.10 Assessment of Body Length and Body Surface Area

Purpose: To determine whether plant extracts alter growth, metabolic status, or dietary-restriction–like responses in *C. elegans* by measuring

body length and body surface area as indicators of developmental rate and overall somatic growth.

1. Synchronize worms and establish treatment conditions by following Sect. 3.8, Steps 1–7 to obtain synchronized L4 larvae.
2. Transfer individual L4 worms (20––30 worms per plate) onto treatment plate (NGM plates seeded with *E. coli* OP50 and plant extract at desired concentration), following Sect. 3.2, Step C1.
3. Incubate worms at 20 °C for 24 h to obtain day 1 adult animals.
4. Pellet worms by centrifugation at 500 RCF for 2 min. Carefully remove the supernatant and wash worms twice with S-medium.
5. Prepare worms for imaging and quantitative analysis by following Sect. 3.4, Steps 7–12.
6. Image each worm under a bright-field microscope using a 10× objective lens to capture the full body profile.
7. Analyze images in Fiji by outlining the worm body as the region of interest (ROI) using an appropriate selection tool (e.g., freehand or polygon).
8. Record the body surface area for each worm (in $pixel^2$ or μm^2, depending on calibration settings).
9. Measure body length by tracing the entire midline of the worm (from head to tail) using the segmented line tool or freehand line tool, then record length in micrometers (μm).
10. Calculate the average body length and average body surface area for each treatment group.
11. Optional: To examine developmental progression, allow worms to grow at the designated temperature and collect samples at specific developmental time points (e.g., 60 h, 108 h, 156 h post-

specific developmental time points (e.g., 36 h, 108 h, 156 h post-L1), measuring growth parameters at each stage.

3.11 Locomotor Function Measurement Using WormLab

Purpose: To quantify locomotion (speed, amplitude, frequency) as an indicator of neuromuscular health and functional aging.

1. Synchronize worms using the standard hypochlorite treatment procedure, following Sect. 3.2, Step B.
2. Allow eggs to hatch overnight to obtain a synchronized L1 population.
3. Transfer 100–200 L1 larvae per group onto NGM plates containing the designated treatment concentrations.
4. Maintain worms on the same treatment plates throughout the developmental period.
5. Allow worms to grow until the following time points: L4, Day 2, and D4 of adulthood.
6. At each time point, gently wash worms with M9 buffer containing 0.1% Triton X-100 to remove residual bacteria and prevent tracking interference.
7. Transfer worms by dropwise pipetting onto NGM plates (without *E. coli* OP50).
8. Allow worms to roam freely for 1 h to adapt and display baseline locomotive behavior.
9. Position the plate under the MBF Bioscience WormLab microscope.
10. Record crawling and thrashing behaviors using standardized WormLab image acquisition settings.
11. Ensure each worm remains moving for at least 90% of the

recording period. Exclude worms that do not meet the 90% movement requirement from downstream analysis.

12. Import all video recordings into WormLab (version 2022), track individual worms using the automated tracking system, and extract both the average speed (μm/s) and average body length (μm) for each worm.

13. Normalize movement speed by worm size using the equation:

    Normalized Speed=Average Speed (μm/s)/Average Body Length (μm)

14. Export all quantitative data for statistical analysis.

Lifespan Analysis

3.12 Lifespan Assay in *C. elegans*

Purpose: To directly assess whether plant extracts extend lifespan under controlled conditions, establishing longevity-promoting potential.

1. Synchronize worms and establish treatment conditions by following Sect. 3.8, Steps 1–8 to obtain synchronized L4 larvae.

2. Maintain worms on treatment plates throughout adulthood: Days 1–4 (reproductive period): Transfer adults daily to fresh treatment plates to avoid progeny mixing. Post-reproductive period: Transfer every other day to maintain plate quality and treatment exposure.

3. Score lifespan daily, ideally at the same time each day.

4. Gently touch each worm with a platinum wire pick to determine viability:

    Alive: Responds to touch

    Dead: No response to touch

5. Remove dead worms from plates immediately after scoring.

6. Calculate daily percent survival for each treatment group.
7. Record and censor worms exhibiting: internally hatched progeny (bagging), extruded gonads, and crawling off the plate. These animals should be removed and excluded from all lifespan analyses.
8. Plot survival curves using Kaplan–Meier survival analysis.
9. Compare treatment and control groups using appropriate statistical tests (e.g., log-rank test).

Evaluation of Neuroprotective Effects

3.13 Paralysis Assay

Purpose: To evaluate neuroprotective effects against β-amyloid/tau–induced paralysis in *C. elegans* models of Alzheimer's by assessing delayed paralysis onset in the temperature-inducible CL4176 and constitutive CL2006 strains, which express human Aβ1–42 in body-wall muscle cells [21, 22], and in the KAE112 expressing human tau protein in body-wall muscle [23].

1. Synchronize KAE112, CL4176, CL802 (control), and CL2006 worms to the L1 larval stage by following Sect. 3.2, Step B.
2. Transfer synchronized L1 larvae onto NGM plates containing plant extracts or vehicle control, following the procedures in Sect. 3.2, Step C1.
3. Maintain CL4176 and CL802 plates at 16 °C until worms reach the L4 stage.

   Maintain KAE112, CL2006 plates at 20 °C until L4 stage.
4. At the L4 stage, transfer 30–50 worms onto fresh NGM plates containing the appropriate treatment (extract or vehicle), and then

incubate the CL4176 and CL802 worms at 16 °C for 48 h.

5. Shift CL4176 plates to 25 °C to induce temperature-dependent expression of human Aβ1–42 in muscle cells.

6. Start a countdown timer immediately following the temperature shift to standardize scoring intervals.

7. Paralysis scoring (*see* **Note 16**):

   A. CL4176 worms: At 20, 22, 24, 26, 28, and 30 h post-temperature shift, evaluate each worm: Paralyzed: No body movement or head-only movement; typically displays a localized bacterial clearing ("halo") around the head. Non-paralyzed: Exhibits any detectable body movement.

   B. KAE112 and CL2006 worms: Maintain KAE112 and CL2006 plates at 20 °C throughout the experiment. Score paralysis every two days using the same criteria. During reproductive stages: Transfer worms daily to fresh treatment plates to avoid confounding from progeny. During post-reproductive stages: Transfer worms every other day to maintain plate quality and consistent exposure.

8. Calculate percent paralysis.

Paralysis (%) = (Number of paralyzed worms/Total worms scored) × 100.

3.14 Chemotaxis Assay

Purpose: To assess sensory-neuron function and behavioral performance in *C. elegans* by measuring chemotaxis toward an attractant. Impaired chemotaxis reflects neuronal dysfunction caused by aging or Aβ toxicity, whereas improved chemotaxis following treatment indicates neuroprotection. This assay is performed using CL2122 (control) and CL2355 worms expressing human Aβ1–42 in neuronal cells [21].

1. Pour 20-mm NGM agar plates (without antibiotics) and allow them to solidify.
2. Mark two opposite quadrants on the underside of each plate as Attractant and Control (Fig. 1).

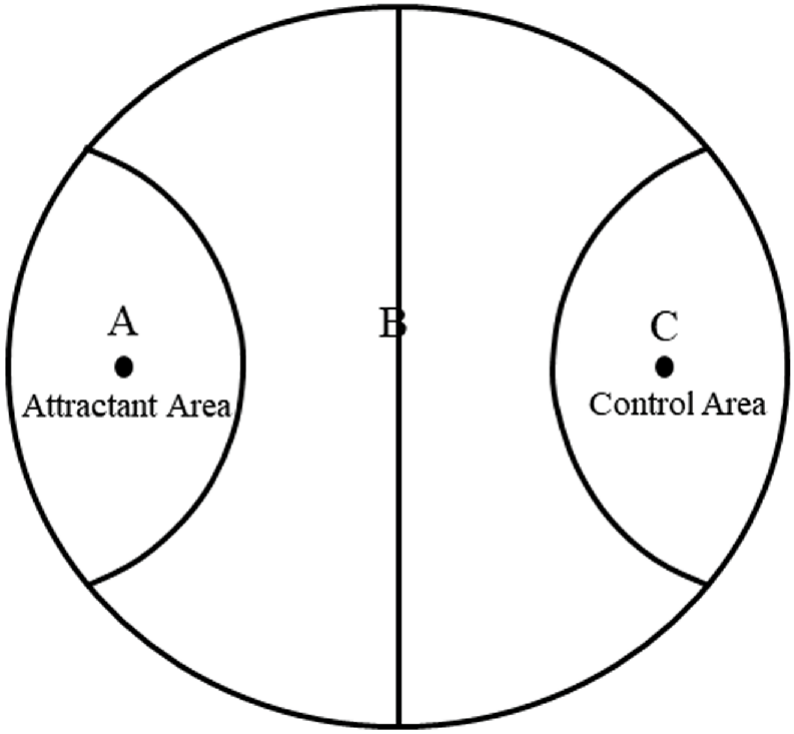


Fig. 1 Chemotaxis plate layout. (**A**) Attractant spot containing 1 µL of 0.1% diacetyl in absolute ethanol and 1 µL of 0.5 M sodium azide. (**B**) Central worm-loading (starting) position. (**C**) Control spot containing 1 µL of absolute ethanol and 1 µL of 0.5 M sodium azide

3. Prepare the following solutions: Attractant: 0.1% diacetyl in absolute ethanol; Control: absolute ethanol; Paralytic agent: 0.5 M sodium azide (NaN_3).
4. At the Attractant spot, dispense 1 µL of 0.1% diacetyl solution and 1 µL of 0.5 M sodium azide.
5. At the Control spot, dispense 1 µL of absolute ethanol and 1 µL of 0.5 M sodium azide.

6. Allow all droplets to absorb into the agar without spreading (approximately 15–30 min).
7. Synchronize CL2122 (control) and CL2355 worms to the L1 larval stage by following Section 3.2, Step B.
8. Transfer synchronized eggs onto NGM plates containing plant extracts or vehicle control, following the procedures in Sect. 3.2, Step C1.
9. Grow CL2122 (control) worms at 20 °C and CL2355 worms at 16 °C until they reach the L4 stage (approximately 36–48 h).
10. At the L4 stage, maintain treated worms for 36 h.
11. Shift worms to 25 °C for 36 h to induce neuronal expression of human Aβ1–42.
12. Wash worms off plates using M9 buffer by following Sect. 3.4, Steps 5–6.
13. Perform the Chemotaxis Assay by placing the concentrated worm suspension at the center of each chemotaxis plate (approximately 100–200 worms per plate) (*see* **Note 17**).
14. Incubate plates at 25 °C for 1 h without disturbance.
15. Count worms in each region:
    $N_{attractant}$ = Number of worms at the Attractant area.
    $N_{control}$ = Number of worms at the Control area.
    N_{total} = Total number of worms on the agar surface.
    (Exclude worms on the plastic edge; *see* **Note 18**).
16. Calculate the Chemotaxis Index (CI)

$$CI = (N_{attractant} - N_{control})/N_{total.}$$

3.15 Polyglutamine (PolyQ) Aggregation Assay

Purpose: To determine whether plant extracts mitigate Huntington's disease-related proteotoxicity in *C. elegans* by reducing fluorescent polyQ::YFP aggregates [25] in AM134 (Q20), AM138 (Q24::YFP), and AM140 (Q35::YFP) [27], strains in which expanded polyglutamine repeats form discrete, quantifiable aggregates in body-wall muscle cells.

1. Synchronize AM134 (Q20), AM138 (Q24::YFP), and AM140 (Q35::YFP) worms and prepare treatment conditions by following Sect. 3.3, Steps 1–3.
2. Maintain cultures and replace S-medium (containing plant extract and *E. coli* OP50) every second day (*see* **Note 19**).
3. On Day 3 and Day 5 of adulthood, prepare worms for imaging and quantitative analysis by following Sect. 3.4, Steps 7–12.
4. For each worm, count the number of discrete polyQ::YFP fluorescent aggregates in the body wall muscle.
5. Score approximately 30 worms per treatment group. Exclude worms that are ruptured, unhealthy, or have artifacts affecting imaging clarity.
6. Document the aggregate count for each worm and strain. Calculate and report mean ± SEM for each treatment condition.

3.16 α-Synuclein Aggregation Assay

Purpose: To determine whether plant extracts reduce α-synuclein aggregation, a key pathogenic feature of Parkinson's disease, by measuring fluorescent α-synuclein::YFP expression in the NL5901 transgenic *C. elegans* strain [9], which expresses α-synuclein::YFP in body-wall muscle cells.

1. Synchronize NL5901 worms and prepare treatment conditions by following Sect. 3.3, Steps 1–3.
2. Incubate synchronized L1 larvae at 20 °C for 96 h under treatment conditions (*see* **Note 20**).
3. Prepare worms for imaging and quantitative analysis by following Sect. 3.4, Steps 7–12.
4. Image each worm using a fluorescence microscope, focusing on the head region, where
   α-synuclein::YFP accumulation is most prominent.
5. Measure the mean fluorescence intensity (MFI) for each worm. Calculate the average MFI for each treatment group and compare across conditions for statistical analysis.

---

4 Notes

1. Cut the end of a 1000-µL pipette tip to widen the opening and prevent mechanical damage to worms during collection.
2. M9 buffer may optionally be supplemented with 0.1% Triton X-100 to improve washing efficiency by reducing surface tension.
3. Do not over-incubate in bleach. Prolonged exposure significantly reduces egg viability. Stop immediately once adult bodies dissolve.
4. When using a solvent to dilute plant extracts, ensure that its final concentration in the assay does not exceed 1% (v/v).
5. Place worms near, but not directly on, the *E. coli* OP50 lawn to avoid dilution of treatment.
6. M9 buffer may be used as an alternative to S-medium for specific steps such as worm washing or pre-treatment handling when bacterial food is not required.

7. Use 2 mL S-medium per well in a six-well plate or 2 mL per 35 × 10 mm Petri dish. Dispense worms gently using a 200-µL pipette tip under a stereomicroscope to avoid damaging larvae.
8. Prepare 80 µM juglone fresh in S-medium immediately before use, as juglone oxidizes rapidly.
9. Preparation of Agarose Pads.
   Place a drop of molten 2% agarose on a clean slide and quickly flatten it with another slide to form a thin, even pad.
10. Apply a thin layer of clear nail polish around the edges of the coverslip to prevent shifting during imaging.
11. Recommended plant extract incubation times for GFP reporter strains TJ375 (*hsp-16.2p::GFP*): 72 h CL2166 (*gst-4p::GFP*): 72 h CF1553 (*sod-3p::GFP*): 48 h.
12. Juglone exposure conditions TJ375 (*hsp-16.2p::GFP*): Expose to 20 µM juglone for 24 h at 20 °C CL2166 (*gst-4p::GFP*): Expose to 20 µM juglone for 24 h at 20 °C CF1553 (*sod-3p::GFP*): Do NOT expose to juglone
   Rationale: SOD-3 reflects baseline DAF-16-dependent antioxidant activity. Acute juglone exposure artificially inflates oxidative stress and may obscure physiologically relevant SOD-3 induction [5].
13. Incubate both TJ356 (DAF-16::GFP) and LD1 (SKN-1::GFP) strains with plant extracts for 72 h.
14. Any remaining L4s will continue laying eggs, disrupting synchronization.
15. Do NOT measure pumping on the same day worms are transferred; transfer stress transiently suppresses pumping. Ensure each worm is actively crawling on an *E. coli* OP50 lawn

Ensure each worm is actively crawling on an *E. coli* OP50 lawn before measurement. Under a dissection microscope, focus on the terminal pharyngeal bulb, as this structure provides the most reliable pump count.

16. Remove paralyzed worms immediately after scoring to prevent double-counting and to avoid confusion with desiccation or bacterial overgrowth.

    Document the number of paralyzed and total worms at each scoring time point. Do not include worms lost due to crawling off the plate or internal hatching (bagging).

    Do NOT observe paralysis on the same day worms are transferred; transfer stress transiently suppresses movement.

17. If excess M9 buffer remains and causes worms to swim rather than crawl, gently wick away the residual liquid using the edge of a Kimwipe to ensure normal chemotaxis behavior.

18. Worms on the outer plastic edge of the plate should be excluded because they are not actively participating in chemotaxis behavior.

19. If needed, add 25 μM FUDR to inhibit reproduction and maintain a synchronous aging population. Adjust culture density as necessary to prevent overcrowding.

20. Gently agitate plates or liquid culture intermittently, or place on a shaker at ~100 rpm to prevent settling. Monitor cultures daily to avoid bacterial overgrowth.

Competing Interests

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Newman DJ, Cragg GM (2020) Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. J Nat Prod 83(3):770–803. https://doi.org/10.1021/acs.jnatprod.9b01285
[PubMed]

2. Atanasov AG, Zotchev SB, Dirsch VM, Supuran CT (2021) Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov 20(3):200–216. https://doi.org/10.1038/s41573-020-00114-z
[PubMed][PubMedCentral]

3. Duangjan C, Curran SP (2022) Oolonghomobisflavans from *Camellia sinensis* increase *Caenorhabditis elegans* lifespan and healthspan. Geroscience 44(1):533–545. https://doi.org/10.1007/s11357-021-00462-7
[PubMed]

4. Zhang S, Duangjan C, Tencomnao T, Wu L, Wink M, Lin J (2022) Oolonghomobisflavans exert neuroprotective activities in cultured neuronal cells and anti-aging effects in *Caenorhabditis elegans*. Front Aging Neurosci 14:967316. https://doi.org/10.3389/fnagi.2022.967316
[PubMed][PubMedCentral]

5. Duangjan C, Rangsinth P, Gu X, Wink M, Tencomnao T (2019) Lifespan extending and oxidative stress resistance properties of a leaf extracts from *Anacardium occidentale* L. in *Caenorhabditis elegans*. Oxid Med Cell Longev 2019:9012396. https://doi.org/10.1155/2019/9012396
[PubMed][PubMedCentral]

6. Duangjan C, Rangsinth P, Gu X, Zhang S, Wink M, Tencomnao T (2019) *Glochidion zeylanicum* leaf extracts exhibit lifespan extending and oxidative stress resistance properties in *Caenorhabditis elegans* via DAF-16/FoxO and SKN-1/Nrf-2 signaling pathways. Phytomedicine 64:153061. https://doi.org/10.1016/j.phymed.2019.153061
[PubMed]

7. Rangsinth P, Sharika R, Pattarachotanant N, Duangjan C, Wongwan C, Sillapachaiyaporn C, Nilkhet S, Wongsirojkul N, Prasansuklab A, Tencomnao T, Leung GP, Chuchawankul S (2023) Potential beneficial effects and pharmacological properties of ergosterol, a common bioactive compound in edible mushrooms. Foods 12(13):2529. https://doi.org/10.3390/foods12132529

8. Duangjan C, Chang X, Seidler PM, Curran SP (2024) Oolonghomobisflavans from *Camellia sinensis* disaggregate tau fibrils across Alzheimer's disease models. bioRxiv. https://doi.org/10.1101/2024.02.26.582120

9. Zhang S, Duangjan C, Tencomnao T, Liu J, Lin J, Wink M (2020) Neuroprotective effects of oolong tea extracts against glutamate-induced toxicity in cultured neuronal cells and β-amyloid-induced toxicity in *Caenorhabditis elegans*. Food Funct 11(9):8179–8192. https://doi.org/10.1039/d0fo01072c

[PubMed]

10. Duangjan C, Rangsinth P, Zhang S, Gu X, Wink M, Tencomnao T (2021) Neuroprotective effects of *Glochidion zeylanicum* leaf extract against H(2)O(2)/glutamate-induced toxicity in cultured neuronal cells and Aβ-induced toxicity in *Caenorhabditis elegans*. Biology (Basel) 10(8):800. https://doi.org/10.3390/biology10080800

11. Duangjan C, Rangsinth P, Zhang S, Gu X, Wink M, Tencomnao T (2021) *Vitis Vinifera* leaf extract protects against glutamate-induced oxidative toxicity in HT22 hippocampal neuronal cells and increases stress resistance properties in *Caenorhabditis elegans*. Front Nutr 8:634100. https://doi.org/10.3389/fnut.2021.634100
[PubMed][PubMedCentral]

12. Duangjan C, Rangsinth P, Zhang S, Wink M, Tencomnao T (2021) *Anacardium Occidentale* L. leaf extracts protect against glutamate/H(2)O(2)-induced oxidative toxicity and induce neurite outgrowth: the involvement of SIRT1/Nrf2 signaling pathway and Teneurin 4 Transmembrane protein. Front Pharmacol 12:627738. https://doi.org/10.3389/fphar.2021.627738
[PubMed][PubMedCentral]

13. Shepard CD, Chang X, Seidler PM, Curran SP (2026) Polyphyletic screen defines distinct classes of plant-derived natural products that oppose tauopathy. Life Sci Alliance 9(2):e202503393. https://doi.org/10.26508/lsa.202503393

14. Duangjan C, Rangsinth P, Gu X, Zhang S, Wink M, Tencomnao T (2019) Data on the effects of *Glochidion zeylanicum* leaf extracts in *Caenorhabditis elegans*. Data Brief 26:104461. https://doi.org/10.1016/j.dib.2019.104461
[PubMed][PubMedCentral]

15. Justice J, Miller JD, Newman JC, Hashmi SK, Halter J, Austad SN, Barzilai N, Kirkland JL (2016) Frameworks for proof-of-concept clinical trials of interventions that target fundamental aging processes. J Gerontol A: Biol Sci Med Sci 71(11):1415–1423. https://doi.org/10.1093/gerona/glw126
[PubMed][PubMedCentral]

16. Antebi A (2007) Genetics of aging in *Caenorhabditis elegans*. PLoS Genet 3(9):1565–1571. https://doi.org/10.1371/journal.pgen.0030129
[PubMed][PubMedCentral]

17. Murphy CT, Hu PJ (2013) Insulin/insulin-like growth factor signaling in *C. elegans*. WormBook:1–43. https://doi.org/10.1895/wormbook.1.164.1

18. Tullet JM (2015) DAF-16 target identification in *C. elegans*: past, present and future. Biogerontology 16(2):221–234. https://doi.org/10.1007/s10522-014-9527-y
[PubMed]

19. An JH, Blackwell TK (2003) SKN-1 links *C. elegans* mesendodermal specification to a conserved oxidative stress response. Genes Dev 17(15):1882–1893. https://doi.org/10.1101/gad.1107803

[PubMed]

20. Son HG, Altintas O, Kim EJE, Kwon S, Lee SV (2019) Age-dependent changes and biomarkers of aging in *Caenorhabditis elegans*. Aging Cell 18(2):e12853. https://doi.org/10.1111/acel.12853
[PubMed][PubMedCentral]

21. Wu Y, Luo Y (2005) Transgenic *C. elegans* as a model in Alzheimer's research. Curr Alzheimer Res 2(1):37–45. https://doi.org/10.2174/1567205052772768
[PubMed]

22. Drake J, Link CD, Butterfield DA (2003) Oxidative stress precedes fibrillar deposition of Alzheimer's disease amyloid beta-peptide (1-42) in a transgenic *Caenorhabditis elegans* model. Neurobiol Aging 24(3):415–420. https://doi.org/10.1016/s0197-4580(02)00225-7
[PubMed]

23. Russell JC, Lei H, Chaliparambil RK, Fish S, Markiewicz SM, Lee TI, Noori A, Kaeberlein M (2021) Generation and characterization of a tractable *C. elegans* model of tauopathy. Geroscience 43(5):2621–2631. https://doi.org/10.1007/s11357-021-00436-9
[PubMed][PubMedCentral]

24. Huang Y, Wang YA, van Sluijs L, Vogels DHJ, Chen Y, Tegelbeckers VIP, Schoonderwoerd S, Riksen JAG, Kammenga JE, Harvey SC, Sterken MG (2024) eQTL mapping in transgenic alpha-synuclein carrying *Caenorhabditis elegans* recombinant inbred lines. Hum Mol Genet 33(24):2123–2132. https://doi.org/10.1093/hmg/ddae148
[PubMed][PubMedCentral]

25. Morley JF, Brignull HR, Weyers JJ, Morimoto RI (2002) The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in *Caenorhabditis elegans*. Proc Natl Acad Sci USA 99(16):10417–10422
[https://doi.org/10.1073/pnas.152161099][PubMed][PubMedCentral]

26. Brenner S (1974) The genetics of *Caenorhabditis elegans*. Genetics 77(1):71–94. https://doi.org/10.1093/genetics/77.1.71
[PubMed][PubMedCentral]

27. Silva MC, Fox S, Beam M, Thakkar H, Amaral MD, Morimoto RI (2011) A genetic screening strategy identifies novel regulators of the proteostasis network. PLoS Gen 7(12):e1002438–e1002438. https://doi.org/10.1371/journal.pgen.1002438

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_4

4. Translational Profiling of Adult *C. elegans* Glia Using Translating Ribosome Affinity Purification (TRAP)

Karmen Mah[1] and Ashley E. Frakes[1, 2] ✉
(1) Genetics and Biochemistry Branch, NIDDK, National Institutes of Health, Bethesda, MD 20814, USA
(2) National Institute of Neurological Disorders and Stroke (NINDS), National Institutes of Health, Bethesda, MD 20814, USA

✉ **Ashley E. Frakes**
Email: Ashley.frakes@nih.gov

Abstract
Glial cells are highly responsive to internal and external cues, so methods that accurately capture their molecular state are essential. However, standard transcriptomic workflows can distort native profiles, and transcript abundance does not always reflect what is ultimately translated into protein. Translational profiling helps address this gap by measuring the mRNAs actively engaged by ribosomes. Translating Ribosome Affinity Purification (TRAP) is a genetically targeted strategy that isolates translating ribosomes from specific cell types, enabling cell-type–resolved profiling of the translatome. Adapting this approach in *C. elegans* provides a useful way to profile glial translation programs and connect them to glia-driven regulation of organismal physiology.

Keywords *C. elegans* – TRAP – Translating ribosome – Immunoprecipitation – Glia – RNA – Affinity purification – mRNA translation – Translatome

1 Introduction

Glial cells are non-neuronal cells of the brain that regulate nearly every aspect of brain formation and function across species. Glia maintain homeostasis, provide trophic and metabolic support to neurons, and regulate neuronal function. Beyond these canonical roles, glia influence physiological processes that change with age, including stress-responsive programs and intercellular signaling [1, 2]. Glia exhibit dynamic molecular programs in response to physiological and environmental cues. In mammals, it has been reported that enzymatic dissociation of brain tissue during preparation of single-cell suspensions for RNA sequencing can induce ex vivo gene-expression artifacts, particularly in microglia, that can confound downstream analyses and biological interpretation [3]. Similar concerns apply to *C. elegans*, where glia represent a small population (56 cells), making cell-type-resolved gene expression measurements difficult without enrichment. Preparation of single-cell suspensions from *C. elegans* often requires chemical cuticle disruption and protease digestion and can introduce transcriptional changes, necessitating caution when interpreting stress-responsive signatures [4]. As such, approaches that preserve in situ molecular states are especially valuable for cell-type–resolved analysis. Moreover, transcription does not always equate to translation, as not all mRNAs are efficiently translated into functional proteins. Therefore, methods that directly profile translated mRNA provide an important complement to transcript-level measurements and can bypass the need for microdissection or cell sorting in many experimental contexts. Translating ribosome affinity purification (TRAP) integrates cell-type–specific transgenic labeling with affinity purification to isolate ribosomes engaged in translation and thereby enrich mRNAs undergoing active translation [5]. TRAP has been

adapted for use in *C. elegans*, modifying the GFP tag utilized in the mouse TRAP protocol to an HA or FLAG tag, and has enabled tissue-specific analyses, including alternative splicing and identification of the genes enriched in serotonergic neurosecretory motor neurons that regulate food sensing and stress-induced eating [6–8]. Therefore, we leverage glia-specific TRAP in *C. elegans* to establish a workflow for translational profiling of defined glial populations and to enable downstream discovery of glial pathways with functional relevance.

---

2 Materials

2.1 General Materials (*see* Note 1)

1. RNase-free tubes.
2. RNase-free tips.
3. RNase-free glass serological pipettes.
4. Gloves.
5. RNase AWAY Surface Decontaminant.
6. Microcentrifuge Tube Magnet Separation Rack.
7. Centrifuge.
8. Refrigerated Centrifuge.
9. AEF403:syb10734[wbmIs119[mir-228p::3xFLAG::rpl-33::SL2::wrmScarlet::unc-54 3'UTR-WBM1364]]; 8x backcross to CGC N2.

2.2 Buffers

1. M9 buffer: 3 g KH_2PO_4, 6 g Na_2HPO_4, 5 g NaCl; add water to 1 L;

sterilize and add 2.0 ml of 0.5 M $MgSO_4$ once cooled (*see* **Note 2**).

2. Bleach solution (prepare fresh): 1.8% sodium hypochlorite, 0.75 M KOH in water.
3. Phosphate-Buffered Saline: 10× pH 7.4, RNase-free.
4. Lysis Buffer A: 10 mM HEPES-KOH, pH 7.4, 150 mM KCl, 5 mM $MgCl_2$, 0.5 mM DTT, 1 mM cycloheximide in nuclease-free H_2O.
5. Low KCl Wash Buffer: 10 mM HEPES-KOH, pH 7.4, 150 mM KCl, 5 mM $MgCl_2$, 1% (v/v) IGEPAL CA-630, 0.04 U/μL RNasin Ribonuclease Inhibitor in nuclease-free H_2O.
6. Lysis Buffer B: Lysis buffer A containing 0.5% (v/v) IGEPAL CA-630, 0.4 U/μL RNasin Ribonuclease Inhibitor, 10 mM ribonucleoside vanadyl complex, 33 mM DHPC, and 1% (w/v) sodium deoxycholate, which are all added fresh before use.
7. High KCl Wash Buffer: 10 mM HEPES-KOH, pH 7.4, 350 mM KCl, 5 mM $MgCl_2$, 1% (v/v) IGEPAL CA-630, 0.04 U/μL RNAsin Ribonuclease Inhibitor in nuclease-free H_2O.

2.3 Preparation of Age-Synchronized Worms

1. NGM plates: prepare nematode growth medium agar plates using a standard NGM recipe (*see* **Note 2**).

2.4 Preparation of Beads:Antibody Complex

1. Dynabeads™ Protein G for Immunoprecipitation.
2. ANTI-FLAG® M2 antibody, Mouse monoclonal (Sigma-Aldrich; F1804-500 μg).

2.5 Preparation of Worm “Pearls”

1. Liquid nitrogen.
2. Stainless steel liquid nitrogen benchtop Dewar.
3. Cardboard insert.

2.6 Lysate Preparation

1. Prefilled 2.0 mL tubes, Zirconium Beads, 3.0 mm.
2. Bead homogenizers.

2.7 Immuno-precipitation

1. TRIzol Reagent.

2.8 RNA Isolation

1. RNeasy Micro Kit (Qiagen; 74004).
2. 2-Mercaptoethanol.
3. Molecular Biology Ethanol.
4. Isopropanol.
5. Chloroform.

2.9 RNA Quality Analysis

1. 4200 TapeStation (or similar):
    a. RNA ScreenTape (Agilent; 5067-5576).
    b. RNA ScreenTape Ladder (Agilent; 5067-5578).

c. RNA ScreenTape Sample Buffer (Agilent; 5067-5577).

2. IKA Vortexer.
3. Thermocycler.

---

3 Methods

3.1 Preparation of Age-Synchronized Worms

1. Add 5–6 mL of M9 buffer to the large worm plate. Gently rotate the plate on the benchtop for a few minutes until the worms are fully suspended.
2. Aspirate the worm suspension in M9 buffer and transfer it into a 15-mL conical (Falcon) tube. If a substantial number of worms remain on the plate, repeat Step 1 and pool the suspension in the same tube.
3. Centrifuge the tube at 1000 × *g* for 30 s.
4. Carefully aspirate the supernatant, leaving approximately 500 µL of liquid above the pellet.
5. Add approximately 3 mL of bleach solution to the tube and quickly resuspend the pellet.
6. Gently rock the tube, then briefly vortex just enough to ensure thorough mixing. Occasionally rock and allow bleaching to continue until the adult worms lyse and begin to dissolve.
7. When many eggs and some partially dissolved worm carcasses are visible, fill the tube to 15 mL with M9 buffer, invert to mix, and immediately centrifuge.
Centrifuge at 1000 × *g* for 30 s (a dense, clear/white pellet of eggs

8. Centrifuge at 2000 × *g* for 30 s (a dense, cloudy white pellet of eggs should be visible if many eggs are present).
9. Aspirate the supernatant, leaving approximately 500 µL, then fill the tube with M9 buffer and mix by inversion.
10. Centrifuge the tube at 1000 × *g* for 30 s.
11. Repeat the M9 wash (Steps 9–10) three additional times, each time filling the tube with M9 buffer, mixing, and centrifuging.
12. After the final wash, aspirate the supernatant until the desired volume of M9 remains (typically 100–300 µL). Gently mix to obtain a uniform egg suspension and count a small aliquot to obtain the number of worms per µL.
13. Plate approximately 1000 eggs per 10 cm NGM plate, spotted with 1 ml OP-50-1, which was allowed to dry at room temperature for at least one week. Prepare at least 15 plates to obtain the required 15,000 worms, with additional plates as backup in case of contamination or losses.
14. Grow worms at 20 °C until day 1 adults. This should take approximately 3 days.

3.2 Preparation of Beads:Antibody Complex

1. On the day before the immunoprecipitation, prepare a 1.5-mL microcentrifuge tube with 100 µL of Dynabeads™ Protein G and wash with 1 mL of PBS, ensuring that beads are fully resuspended. Place tube on magnetic rack for 1 min and remove supernatant.
2. To the washed beads, add 1 mL of PBS and 10 µg of FLAG antibody (10 µL) and incubate overnight with rotation end-over-end at 4 °C.
3. On the morning after, collect beads on magnet for 1 min, wash with 1 mL of Low KCl wash buffer, and resuspend in 100 µL Low KCl

wash buffer. Beads should be stored on ice until ready for use.

Prepared beads should be used immediately, or can be stored up to two weeks at 4 °C with the addition of sodium azide to a final concentration of 0.02% v/v. If using after storage with sodium azide, wash three times in Low KCl wash buffer before use.

3.3 Collection of Worms (*see* Note 3)

1. Collect 15,000 synchronized, day 1 *C. elegans* in M9 buffer. Add 9–10 mL of M9 buffer to plates containing approximately 1000 worms per plate and centrifuge at 1000 × *g* for 30 s at room temperature to collect worms. Repeat for all plates until approximately 15,000 worms are collected. M9 buffer can be moved from one plate to another to pool worms from multiple plates for quicker processing.
2. Wash worms at room temperature as follows:
   - Twice with 15 mL M9 buffer.
   - Once with 15 mL M9 buffer supplemented with cycloheximide (final concentration 1 mM) (*see* **Note 4**).
3. Wash once with 5 mL Lysis Buffer A.
4. Pellet worms at 2000 × *g* for 2 min at room temperature, then discard most of the supernatant, leaving approximately 500 µL.

3.4 Preparation of Worm "Pearls"

1. Prepare a benchtop liquid nitrogen dewar with a cardboard insert.
2. Place the empty microcentrifuge tubes into cardboard insert within the liquid nitrogen-safe vessel. Fill the dewar with liquid nitrogen so that the tubes are fully surrounded, and also fill each microcentrifuge tube with liquid nitrogen.
3. Resuspend the pellet until a uniform worm slurry is formed.

4. Shock-freeze the residual pellet (approximately 500 µL) by dropping ~20 µL aliquots into liquid nitrogen-filled microcentrifuge tubes using a P200 pipette to create frozen pearls (*see* **Note 5**).

3.5 Lysate Preparation

1. Prepare RNase-free screw-cap tubes pre-filled with zirconium beads and place them on dry ice until use.
2. Working over dry ice, transfer frozen pearls from microcentrifuge tubes into the pre-chilled bead tubes, minimizing thawing. Distribute the pearls evenly into two bead tubes per 15,000-worm sample.
3. Lyse the frozen pearls using a BioSpec Mini-Beadbeater (or equivalent) for two cycles of 1 min each, briefly returning tubes to dry ice to keep samples frozen or near-frozen (*see* **Note 6**).
4. Thaw lysed powder on ice and add Lysis Buffer B to each tube to bring the final volume of each tube to 1.75 mL, mixing gently to fully resuspend the lysate.
5. Incubate the samples on ice for 30 min.
6. Centrifuge at 12,000 × *g* for 10 min at 4 °C.
7. Carefully transfer the supernatant from each tube to a pre-chilled 15 mL conical tube on ice, avoiding disturbance of the pellet. Discard the pellet and immediately proceed with the immunoprecipitation.

3.6 Immuno-precipitation

1. Aliquot 100 µL of clarified supernatant and mix with 500 µL of

TRIzol Reagent; this constitutes the Input fraction. Samples in TRIzol reagent can be stored at –80 °C for processing at a later time.

2. Incubate the remaining supernatant with 100 µL FLAG-antibody-coupled beads for ≥3 h (no more than 4 h) at 4 °C with continuous end-over-end rotation.

3.7 Washes

1. While keeping all solutions on ice, collect beads on the chilled magnetic rack for 1 min and save the supernatant in a new tube (*see* **Note 7**).
2. Resuspend beads in 1 mL of High-KCl wash buffer and immediately collect with magnet as above.
3. Repeat wash with 1 mL of High-KCl Wash Buffer three times (for a total of four washes).

3.8 RNA Isolation from IP

1. For RNA elution, incubate washed beads in 350 µL RLT buffer (from Qiagen RNeasy Micro Kit) with 1% v/v β-mercaptoethanol for 10 min at room temperature.
2. Separate the eluate from the beads using the magnetic rack. This eluate now contains the RNA.
3. Add 350 µL 70% ethanol to the eluate and proceed with the Qiagen RNeasy Protocol for Purification of Total RNA from Animal and Human Cells, performing the optional on-column DNase digestion (*see* **Note 7**).
4. Elute RNA from the column with 15 µL of RNase-free water.

3.9 RNA Isolation from Input or Flowthrough

1. Incubate samples with TRIzol for 5 min at room temperature.
2. Add 100 µL of chloroform and securely cap the tube. Mix thoroughly by shaking or vortexing.
3. Incubate for 2–3 min.
4. Centrifuge the sample for 15 min at 12,000 × *g* at 4 °C.
5. Transfer upper (clear) aqueous phase to a new 1.5 mL microcentrifuge tube, taking care to avoid the layer underneath.
6. Add 500 µL of isopropanol to the sample, mix, and incubate at room temperature for 10 min.
7. Transfer the entire sample to the RNeasy MinElute Column (from Qiagen RNeasy Micro Kit) and centrifuge for 15 s at ≥ 8000 × *g*. Discard the flow-through (*see* **Note 8**).
8. Continue with Qiagen RNeasy Micro Kit cleanup protocol, starting with the washes with Buffer RW1, skipping initial lysis steps.

3.10 RNA Quality Analysis

1. Perform quality analysis on the input, flow-through, and IP samples by using the Agilent TapeStation platform (or similar).
2. In brief, allow the RNA sample buffer and RNA ladder to equilibrate at room temperature for 30 min.
3. Vortex reagents to mix and centrifuge before use.
4. In a strip of 8 PCR tubes, add 5 µL RNA Sample Buffer to each tube; add 1 µL RNA Ladder to the first tube and 1 µL of each RNA sample to the remaining tubes (*see* **Note 9**).
5. Spin down the strip to gather all contents at the bottom of the

tube.

6. Vortex using the IKA vortexer at 500 rpm for 1 min.
7. Centrifuge the strip once again.
8. Heat strip at 72 °C for 3 min, then incubate on ice for 2 min.
9. Centrifuge strip to gather contents to the bottom of the tube.
10. Load samples into the TapeStation instrument (*see* **Note 10**).
11. Click Start on the TapeStation Controller Software. The Agilent TapeStation Analysis Software opens after the run and displays the result.
12. Expected RNA yields are 70–100 ng/μL for the IP fraction and 500–1000 ng/μL for the input and flowthrough fractions, with RINe ≥ 7.0 across all samples (*see* **Note 11**).

---

4 Notes

1. Use RNase-free consumables and nuclease-free water throughout. Decontaminate work surfaces and wear fresh gloves to minimize RNase contamination.
2. Prepare NGM plates, M9 buffer, and culture conditions using your lab's standard *C. elegans* practices. Consistency in plate drying, bacterial lawn thickness, and temperature helps reduce variability in yield.
3. Use glass serological pipettes during worm collection to reduce worm loss from adherence to plastic pipettes.
4. Cycloheximide is included in wash buffers to stabilize ribosome–

mRNA complexes.

5. During pearl formation, keep the worm slurry well mixed so each dropped aliquot contains a similar number of worms. Store pearls at −80 °C and avoid repeated freeze–thaw cycles to limit RNA degradation.

6. Keep lysate cold at all times. Perform bead-beating in a 4 °C cold room when possible, or return tubes to dry ice between runs. Excess warming can increase RNase activity and reduce RNA quality.

7. For the flowthrough fraction, immediately after the initial bead capture, remove 100 µL of the flowthrough and mix with 500 µL TRIzol. Keep wash buffers on ice and resuspend beads completely at each wash to improve background removal.

8. If sample volume exceeds the column capacity, load the sample onto the RNeasy column in multiple rounds, centrifuging and discarding flow-through between loads. Include on-column DNase digestion when downstream assays are sensitive to DNA.

9. Input and flow-through fractions often require dilution (e.g., 1:5 to 1:10) to fall within the recommended range for the RNA ScreenTape assay. Avoid overloading the ScreenTape, which can distort RINe values.

10. Remove caps from tube strips before loading the TapeStation. Place the ladder in the correct position (typically A1) and confirm the run setup in the controller software before starting.

11. Expected yields vary by strain, tag expression level, and antibody performance. If IP yield is low, verify worm counts, lysis efficiency, bead recovery, and that the antibody recognizes the tagged protein under the lysis conditions.

Funding

This research was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) within the National Institutes of Health (NIH).

Competing Interests

The authors declare no competing interests. The contributions of the NIH author(s) are considered works of the US government. The findings and conclusions presented in this chapter are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

References

1. Frakes AE et al (2020) Four glial cells regulate ER stress resistance and longevity via neuropeptide signaling in C. elegans. Science 367(6476):436–440 [https://doi.org/10.1126/science.aaz6896][PubMed][PubMedCentral]

2. Gildea HK, Liddelow SA (2025) Mechanisms of astrocyte aging in reactivity and disease. Mol Neurodegen 20(1):21 [https://doi.org/10.1186/s13024-025-00810-7]

3. Marsh SE et al (2022) Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. Nat Neurosci 25(3):306–316 [https://doi.org/10.1038/s41593-022-01022-8][PubMed][PubMedCentral]

4. Williams RTP et al (2023) Transcriptome profiling of the Caenorhabditis elegans intestine reveals that ELT-2 negatively and positively regulates intestinal gene expression within the context of a gene regulatory network. Genetics **224**(4):iyad088

5. Heiman M et al (2014) Cell type–specific mRNA purification by translating ribosome affinity purification (TRAP). Nat Protoc 9(6):1282–1291 [https://doi.org/10.1038/nprot.2014.085][PubMed][PubMedCentral]

6. Rhoades JL et al (2019) ASICs mediate food responses in an enteric serotonergic neuron that controls foraging behaviors. Cell 176(1):85–97 [https://doi.org/10.1016/j.cell.2018.11.023][PubMed]

7. Gracida X et al (2017) An elongin-cullin-SOCS box complex regulates stress-induced serotonergic neuromodulation. Cell Rep 21(11):3089–3101 [https://doi.org/10.1016/j.celrep.2017.11.042][PubMed][PubMedCentral]
8.

Wester LE et al (2023) A single-copy knockin translating ribosome immunoprecipitation toolkit for tissue-specific profiling of actively *translated mRNAs* in *C. elegans*. Cell Rep Methods **3**(3):100433

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_5

5. An Integrated Pipeline for Quantifying Survival, Cell Morphology, and Nuclear Integrity During Chronological Aging of Fission Yeast

Kristen Hairapatian[1], Kayla Mallari[1], Thomas Riffle[1], Annika Schrader[1], Emily M. Sánchez[1], Júlia Guil Rodríguez[1], Shanjika Nagendrakumar[1], Martin Vo[2] and Wilber Escorcia[1] ✉

(1) Department of Biology, California State University, Northridge, Los Angeles, CA 91330, USA

(2) Department of Pediatrics, University of Hawai'i, John A. Burns School of Medicine, Honolulu, HI 96813, USA

✉ **Wilber Escorcia**
Email: wilber.escorcia@csun.edu

Abstract

Chronological lifespan (CLS) assays in *Schizosaccharomyces pombe* are widely used to measure post-mitotic survival using colony-forming unit (CFU) readouts. Although CFU measurements provide a robust assessment of population-level viability, they do not directly capture cellular features that accompany survival decline during aging. Here, we describe an integrated experimental pipeline that enables the coordinated collection of three complementary readouts from a single CLS time course: CFU-based survival, automated quantification of cell morphology, and fluorescence-based assessment of nuclear integrity.

During daily CLS sampling, aliquots from aging cultures are partitioned for viability plating, phase-contrast microscopy followed by automated image segmentation and statistical analysis, and ethanol fixation for DAPI staining. Automated morphology analysis yields cell-dimension distributions and normalized summary statistics aligned by aging day, while DAPI staining enables qualitative classification of nuclear morphology. This protocol allows population-level survival trajectories to be contextualized with single-cell morphological and nuclear-state measurements without requiring parallel experiments. The workflow is compatible with standard CLS assays and provides a reproducible framework for integrating viability, morphology, and nuclear integrity readouts during chronological aging in *S. pombe*.

Keywords *Schizosaccharomyces pombe* – Fission yeast – Chronological lifespan (CLS) – Colony-forming units (CFU) – Automated image segmentation – Cell morphology – CellPose – PhotoPhenosizer – DAPI staining – Nuclear integrity

---

1 Introduction

Chronological lifespan (CLS) assays in the fission yeast *Schizosaccharomyces pombe* provide a well-established experimental framework for examining cellular survival following exit from the mitotic cell cycle. In this assay, cells are cultured without nutrient replenishment, and viability is monitored over time, most commonly by determining colony-forming units (CFUs) at defined intervals (every 24 h). CLS measurements offer a direct and quantitative assessment of population-level survival during post-mitotic aging and have been widely applied to studies of stress response, metabolism, and cellular homeostasis in fission yeast.

A defining feature of *S. pombe* as a model system is the tight coupling between cellular physiology and morphology. Cell length and overall geometry are closely linked to growth state, cell-cycle progression, and environmental conditions, making morphological

measurements a biologically informative readout during aging and stress [1]. In addition, nuclear organization and integrity provide sensitive indicators of cellular stress and genome stability and can be readily visualized using fluorescence-based staining methods that are compatible with standard fission yeast workflows [2, 3].

Despite these advantages, CLS assays are often performed using CFU measurements alone, with morphological or nuclear analyses carried out in separate experiments or restricted to small, manually scored subsets of cells. Such approaches make it difficult to align population-level survival trajectories with cellular features measured at the same aging stage [4]. Integrating multiple readouts within a single CLS experiment requires careful coordination of sampling, imaging, and analysis to ensure temporal alignment and reproducibility.

Methodological work in *S. pombe* has established rigorous genetic, cytological, and imaging practices that emphasize physiological context, quantitative interpretation, and reproducibility [1–3]. Building on this foundation, recent advances in automated image segmentation and statistical analysis have enabled morphology-based measurements to be scaled while retaining biologically meaningful metrics such as cell length, width, and area. Automated pipelines allow thousands of cells to be analyzed per time point, producing full distributions rather than summary values alone [5–8].

In this chapter, we describe an integrated CLS pipeline that coordinates CFU-based viability measurements with automated cell morphology analysis and fluorescence-based assessment of nuclear integrity using daily samples drawn from the same aging cultures. The protocol is designed to be fully compatible with established *S. pombe* CLS methods while incorporating automated segmentation and standardized fluorescence staining. By consolidating these measurements within a single experimental workflow, the method enables population-level survival data to be temporally aligned with single-cell structural readouts without requiring parallel or independent assays.

2 Materials

2.1 Yeast Culture and Reagents

1. Forsburg Yeast Strain 7 (*Schizosaccharomyces pombe*, wild-type derived from 972 h^-).
2. Yeast extract plus supplements (YES medium).
3. YES agar plates (2% agar).
4. Hydrogen peroxide (H_2O_2).
5. Ethanol (100%).
6. DAPI (4′,6-diamidino-2-phenylindole).
7. Agarose (molecular biology grade).
8. Phosphate-buffered saline (PBS, 1×).
9. Sterile deionized water.

2.2 Consumables and Equipment

1. Sterile culture tubes (15 mL) and Erlenmeyer flasks (250 mL).
2. 1.5 mL microcentrifuge tubes.
3. Sterile serological pipettes and pipette tips.
4. 96-well plates.
5. Sterile spreader beads.
6. Glass microscope slides (poly-L-lysine–coated).

7. Coverslips.
8. Kimwipes or lint-free wipes.
9. Cuvettes compatible with a spectrophotometer.
10. Hemocytometer and coverslip.
11. Permanent markers.
12. Labeling tape.
13. Shaking incubator (32 °C, 220 rpm).
14. Benchtop microcentrifuge.
15. Spectrophotometer (OD_{595}).
16. Phase-contrast microscope with camera (60× or 63× objective).
17. Fluorescence microscope with DAPI filter set.

2.3 Software and Data Organization

1. CellPose (cyto2 model).
2. Photo Phenosizer (PP) pipeline.
3. R statistical computing environment.
4. Photo Phenosizer R scripts for data aggregation and analysis.
5. ImageJ or Fiji for fluorescence image inspection and nuclear

6. measurements.
   Structured data recording system (laboratory records, image storage organization, and tabular files for CFU counts, morphometric outputs, and nuclear scoring).

---

3 Methods

3.1 Overview of the Integrated CLS Workflow

This protocol is organized around a standard chronological lifespan (CLS) assay in *Schizosaccharomyces pombe* in which aging cultures are sampled daily without nutrient replenishment. At each sampling time point, aliquots are withdrawn from the same culture and partitioned for multiple downstream analyses. One aliquot is used for colony-forming unit (CFU) determination to quantify population-level survival (Fig. 1; Table 1). Parallel aliquots are processed for phase-contrast microscopy to enable automated cell morphology analysis (Fig. 2) and for ethanol fixation to preserve samples for fluorescence-based assessment of nuclear morphology (Fig. 3). All measurements are indexed to the same aging day, allowing population-level survival data to be aligned temporally with single-cell morphological and nuclear readouts (*see* **Note 1**).

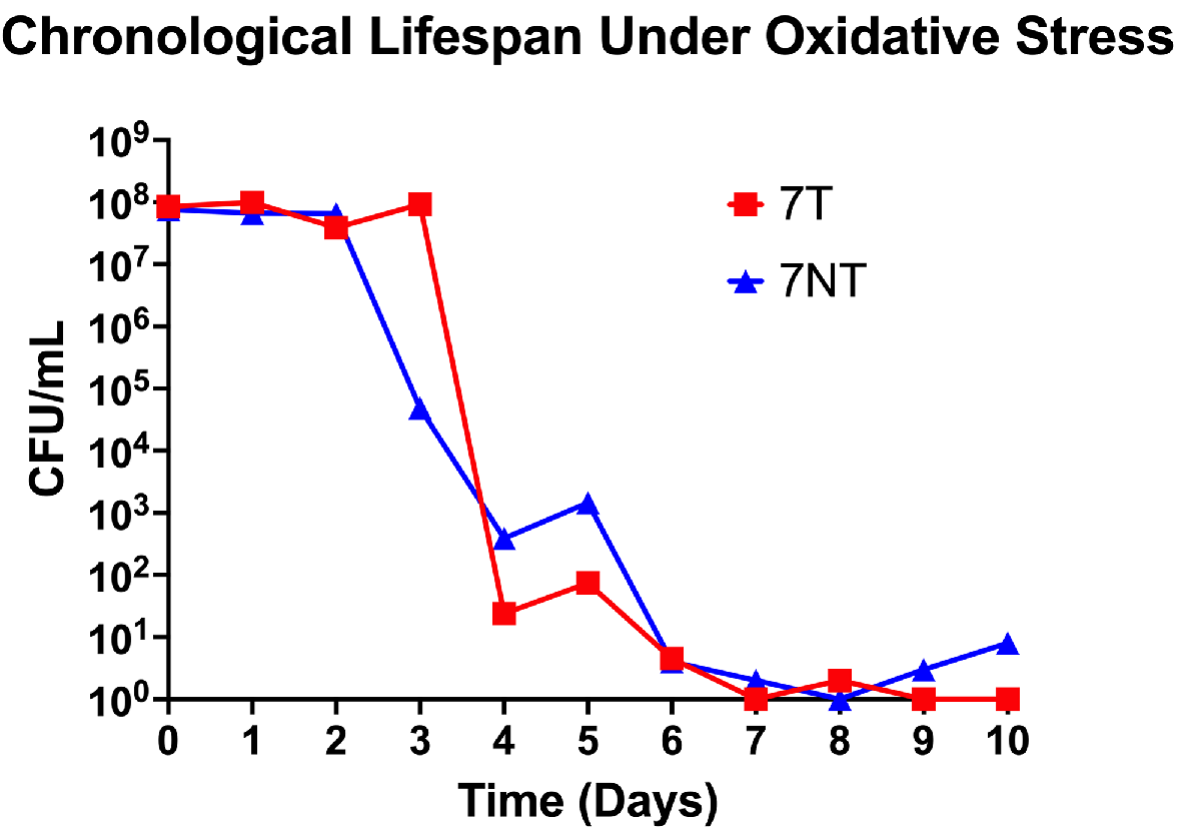


Fig. 1 Chronological lifespan (CLS) survival during oxidative stress. Wild-type *Schizosaccharomyces pombe* cultures were aged in liquid YES medium in the absence (NT; 0 mM H_2O_2) or presence (T; 1 mM H_2O_2) of oxidative stress. At the indicated days of chronological aging, aliquots from the same cultures were withdrawn and used for colony-forming unit (CFU) determination, as well as for cell morphology and nuclear integrity analyses. CFU values are expressed as CFU/mL on a logarithmic scale and plotted as a function of aging time (days). Survival trajectories illustrate the decline in population-level viability during chronological aging and the accelerated loss of viability under oxidative stress conditions. These data provide a population-level survival framework that is temporally aligned with single-cell morphological and nuclear readouts obtained from the same samples

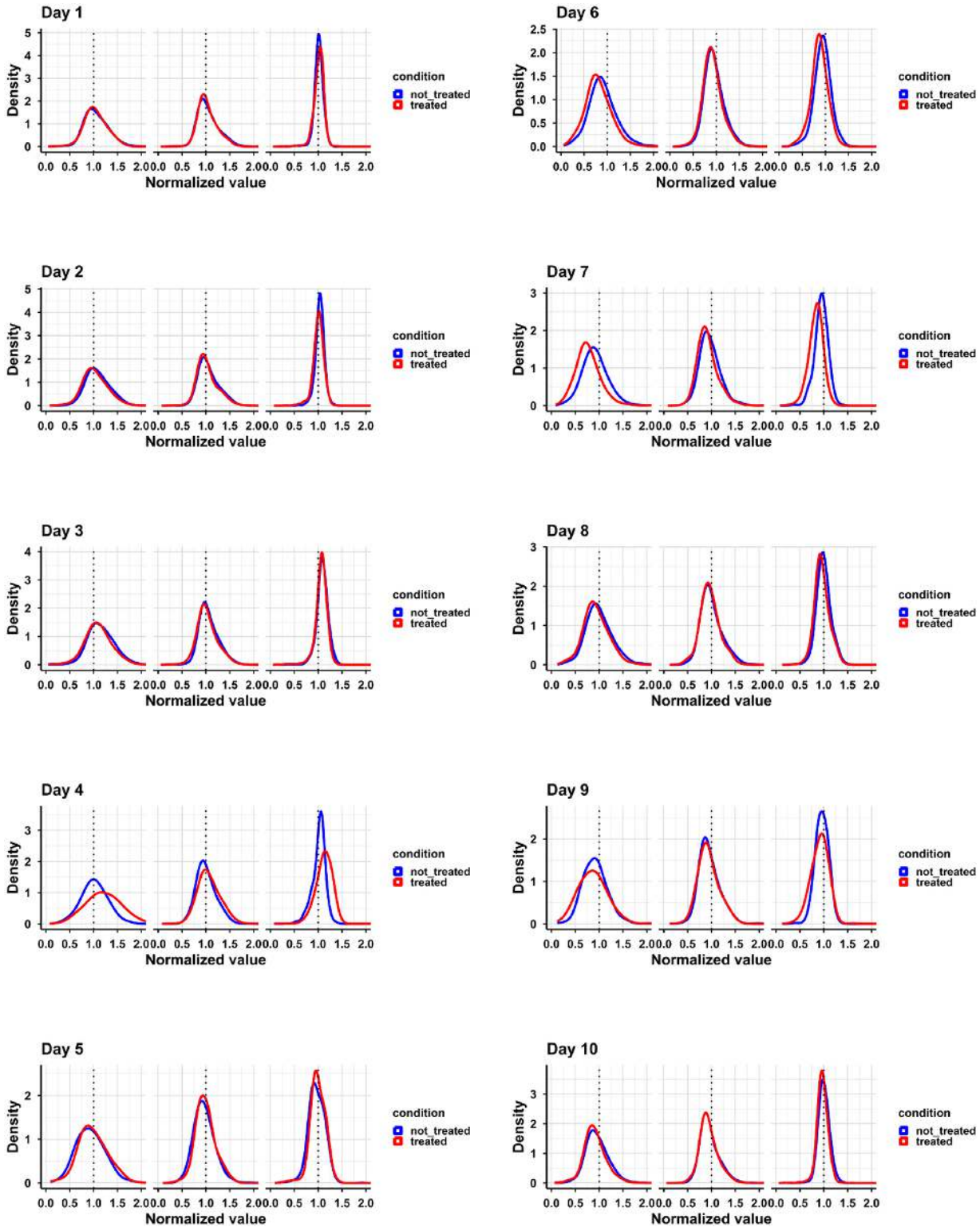


Fig. 2 Automated quantification of cell morphology during chronological aging. Phase-contrast images of aging *S. pombe* cultures were acquired daily from the same CLS samples used for CFU determination and nuclear integrity analysis. Images were segmented using the CellPose cyto2 model follwed by statistical analysis using the Photo Phenosizer pipeline. Representative

segmentation overlays illustrate automated identification of individual cells with numeric labels (left). Cell morphology metrics, including cell area, major axis length, and minor axis length, were extracted for each segmented cell and aggregated by aging day. This automated approach enables high-throughput, distribution-based analysis of morphological changes that accompany chronological aging and can be directly aligned with population-level survival and nuclear integrity measurements from the same time points

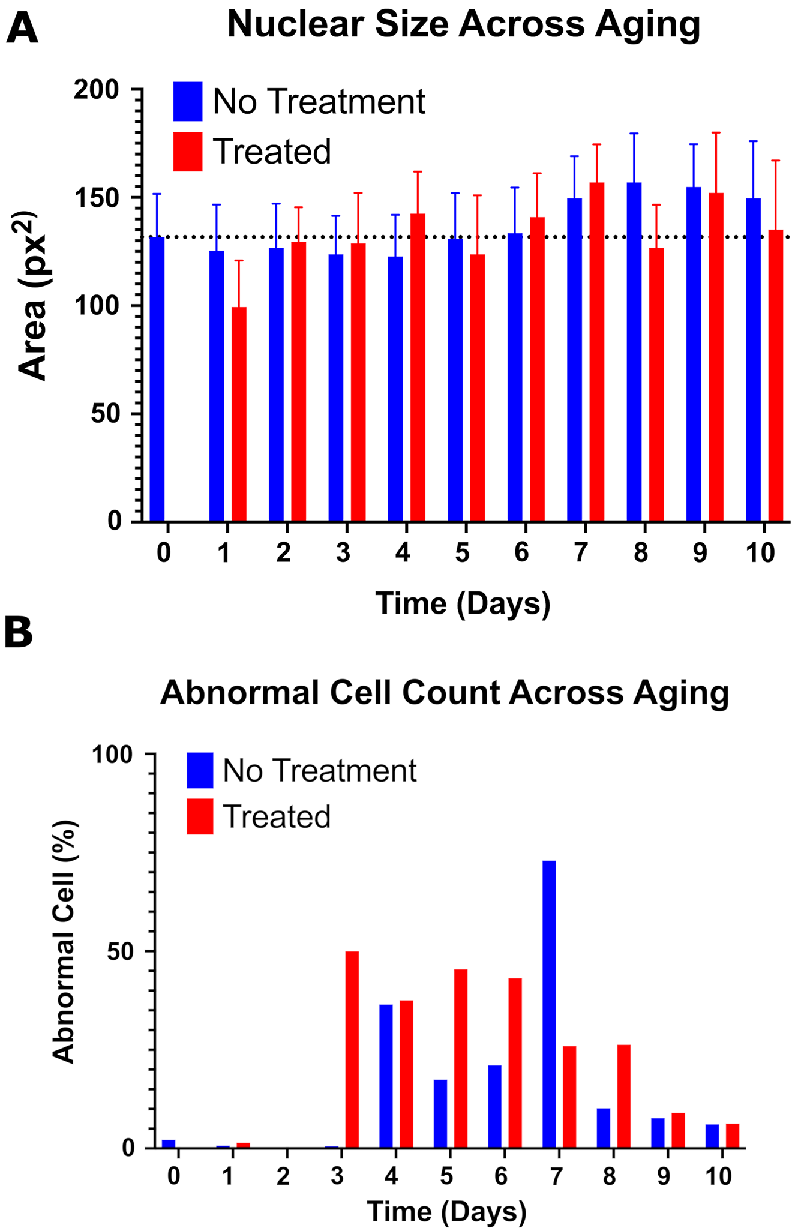


Fig. 3 Quantitative and categorical assessment of nuclear morphology during chronological aging. Cells were collected daily from aging *Schizosaccharomyces pombe* cultures during a chronological lifespan (CLS) assay and fixed in ethanol for fluorescence-based analysis of nuclear morphology. DAPI-stained nuclei were imaged by fluorescence microscopy from the same CLS samples used for CFU determination and cell morphology analysis. Nuclear integrity was evaluated using two complementary approaches. (**A**) For continuous quantification, nuclear area

was measured for individual cells using ImageJ, enabling population-level comparisons across aging days. (**B**) In parallel, nuclei were scored categorically as normal or abnormal based on qualitative features including irregular shape, fragmentation, enlargement, or diffuse DAPI staining. The proportion of cells exhibiting abnormal nuclear morphology was calculated for each time point. Together, these quantitative and categorical readouts provide complementary measures of nuclear integrity that can be temporally aligned with changes in cell morphology and population-level survival during chronological aging

Table 1 Normalized cell morphology metrics across CLS

The workflow is designed to maximize information obtained from a single CLS experiment while maintaining compatibility with established fission yeast aging protocols. No additional cultures or parallel experiments are required to obtain the integrated readouts described here (*see* **Note 15**).

3.2 CLS Culture Setup

a. Streak *S. pombe* FY7 cells from frozen stock onto YES agar plates and incubate at 32 °C until well-isolated colonies are visible (typically 2–3 days).

b. Using sterile technique, inoculate a single colony into YES liquid medium in a sterile culture tube or flask.

c. Incubate the starter culture overnight at 32 °C with shaking at approximately 220 rpm to allow cells to reach logarithmic growth (0.3–0.7 OD Abs_{595nm}).

d. On the following day, standardize the culture by counting cells using a hemocytometer and dilute in fresh YES medium to a starting density of 1000 cells/mL.

e. Prepare 100 mL liquid YES medium per strain or treatment to allow for daily sampling across the full 10-day aging period.

f. Label the culture with strain identifier, treatment condition (e.g., 1 mM H_2O_2), and start date.

g. Place cultures in a shaking incubator at 32 °C and 220 rpm. Do not replenish nutrients or dilute cultures after this point; the time of inoculation into fresh YES is designated as Day 0 of chronological aging.

3.3 Daily Sampling During Chronological Aging

a. Perform daily sampling from Day 1 through Day 10 (or as required by the experimental design) (*see* **Note 13**). At each time point, remove cultures from the incubator and mix gently to resuspend settled cells (*see* **Note 2**).

b. Withdraw a total of 5 mL of culture per condition for downstream analyses. This volume is partitioned as follows:

   1. Aliquots for CFU determination (including concentrated samples and dilution stocks).
   2. Aliquot for optical density measurement and hemocytometer-based viable cell counting.
   3. Aliquot for phase-contrast microscopy to enable image acquisition and assessment of cellular morphology.
   4. Aliquot for ethanol fixation to preserve samples for fluorescence-based staining.

c. For optical density measurements, mix 200 µL of culture with 800 µL of YES medium to generate a 1:5 dilution. Measure absorbance at OD_{595} and record the value. Use the same diluted sample to prepare further dilutions for hemocytometer counting.

d. Perform hemocytometer counts on diluted samples, counting only cells that appear intact and refractile. Record viable cell counts and calculate cells per milliliter using the appropriate dilution and hemocytometer conversion factors, applying the dilution factor to

determine the true cell density of the original culture.

e. Prepare aliquots for CFU determination and for fixation at the same time point to ensure all downstream measurements correspond to the same aging day (*see* **Note 1**).

f. Fixed samples for fluorescence-based assays are prepared by adding 420 mL culture aliquots directly to 1 mL 100% ethanol and stored at 4 °C until staining.

g. Proceed immediately with CFU plating and microscopy sample preparation or store fixed samples appropriately for later analysis.

3.4 CFU Determination: Concentration, Dilution, Plating, and Calculations

This section describes the procedure used to quantify cell viability during chronological aging by colony-forming unit (CFU) analysis. The protocol is structured to ensure reproducible sampling, accurate dilution tracking, and compatibility with parallel morphology and nuclear assays, and CFU-based survival trajectories derived from this workflow are shown in Fig. 1 (*see* **Note 1**).

Preparation of Concentrated Cell Stocks

A. From the daily CLS sample, transfer 2.0 mL of culture into two labeled 1.5 mL microcentrifuge tubes (1 mL per tube).

B. Centrifuge the sample at 6000 rpm for 1 min at room temperature to pellet cells (*see* **Note 2**).

C. Carefully remove the supernatant and resuspend the pellet in 100 µL of sterile YES medium to generate a 10× concentrated stock.

D. Mix thoroughly by gentle pipetting to ensure a homogeneous suspension.

E. Prepare a 5× concentrated stock by resuspending the second pellet with 200 µL of sterile YES.

F. Retain an aliquot of 200 µL unconcentrated culture as the 1× stock.

All three stocks (10×, 5×, and 1×) are used for CFU plating to ensure that countable colonies are obtained across the full aging time course (*see* **Note 3**). These calculations are for one technical replicate. Adjust accordingly for multiple replicates.

3.5 Serial Dilution Setup

A. Label a sterile 96-well plate to reflect dilution series, aging day, and treatment condition.

B. Add 180 µL of sterile YES medium to each well designated for dilution (five dilution wells per strain/treatment).

C. Starting with the 1× sample (well #1), perform serial 1:10 dilutions by transferring 20 µL of cell suspension into 180 µL of YES and mixing thoroughly in each dilution.

D. Continue serial dilutions as needed to span a range expected to yield countable colonies (typically 30–300 CFU per plate).

E. Record all dilution factors carefully to enable accurate back-calculation of CFU per milliliter (*see* **Note 4**).

Plating for Colony Formation

A. For each dilution, plate 100 µL of concentrated and diluted cultures onto YES agar plates.

B. Spread cells evenly across the plate surface using sterile spreader beads. Use seven beads per plate to ensure uniform volume distribution across samples.

Plate all samples in duplicate to improve accuracy and

C. Plate all samples in duplicate to improve accuracy and reproducibility.

D. Allow plates to dry briefly with lids ajar under sterile conditions before incubation.

E. Incubate plates at 32 °C until colonies are visible and well separated, typically 3–4 days.

Colony Counting and CFU Calculation

A. After incubation, select plates with well-isolated colonies suitable for counting.

B. Count all visible colonies on each plate manually or using an appropriate colony counter.

C. For each plate, calculate CFU per milliliter using the formula:

D. CFU/mL = [Number of colonies/Volume plated (mL)] × Dilution factor.

E. Average CFU values from duplicate plates for each dilution.

F. For each aging day and condition, select CFU values derived from plates within the optimal counting range.

G. Normalize CFU values as required for downstream analysis (e.g., relative to Day 1 untreated controls) while retaining raw CFU/mL values for reference.

H. Record all counts, dilution factors, and calculated CFU values in a structured spreadsheet or database to facilitate downstream plotting and integration with morphological and nuclear data.

3.6 Phase-Contrast Microscopy and Automated Cell Morphology Analysis

This section describes the preparation of microscopy samples, image acquisition, and automated analysis pipeline used to quantify cell morphology during chronological aging. Samples used for morphology analysis are drawn from the same CLS cultures and time points used for CFU determination to ensure temporal alignment across readouts, and representative segmentation outputs and derived morphometric measurements are shown in Fig. 2.

3.7 Preparation of Agarose Micropads and Microscopy Samples

A. Prepare agarose micropads by dissolving agarose in sterile water or an appropriate buffer to a final concentration of approximately 1% (w/v) [9]. Heat until fully dissolved and allow the solution to cool slightly before use.

B. Dispense 50 µL of molten agarose onto a clean microscope slide and flatten using a second slide or coverslip to create a thin, even pad. Allow the pad to solidify at room temperature (*see* **Note 9**).

C. From the daily CLS sample, withdraw a small aliquot of culture (typically 100 µL), spin for 1 min at 6000 rpm, decant the supernatant, and gently mix with 10 µL liquid YES to resuspend cells.

D. Spot the culture aliquot onto the surface of the solidified agarose pad.

E. Carefully place a coverslip over the pad, avoiding air bubbles and excessive pressure, and use a fingertip to carefully rotate the coverslip clockwise 3× to generate a cell monolayer.

F. Seal the coverslip edges with nail polish and immediately begin image acquisition to minimize dehydration or morphological distortion.

distortion.

Phase-Contrast Image Acquisition

A. Acquire phase-contrast images using a microscope equipped with a high-numerical-aperture objective suitable for *S. pombe* morphometrics (e.g., 60× or 63× oil) and a monochrome camera that produces digital images for downstream analyses.

B. Use consistent illumination, exposure settings, and optical configuration across all aging days and conditions (*see* **Note 7**).

C. Capture multiple nonoverlapping fields of view per sample to ensure representative sampling of the population.

D. Save images in a lossless format (e.g., TIFF), using a structured naming convention that encodes aging day, treatment condition, and replicates.

E. Ensure that the imaging system is calibrated prior to acquisition. Downstream analyses use calibrated pixel sizes that are specific to the microscope system employed.

3.8 Automated Image Segmentation Using CellPose

A. Organize phase-contrast image files into a hierarchical directory structure in which images are grouped by aging day and experimental conditions.

B. Load the organized image directories into the Photo Phenosizer (PP) pipeline (https://github.com/XavierCompBio/PhotoPhenosizer), which uses CellPose as the underlying segmentation engine.

C. Perform automated cell segmentation using CellPose (https://www.cellpose.org/) with the cyto2 model, applying the following parameters (*see* **Note 8**):

a. Flow threshold = 0.4

Flow_threshold = 0.4.

b. cellprob_threshold = 0.0.

c. Downsample = 4.

d. Object splitting disabled.

D. Configure the PP pipeline to process image directories recursively, enabling batch analysis across all aging days and experimental conditions.

E. For each processed image, generate segmentation masks and corresponding overlay images displaying numeric cell identifiers and total cell counts per field of view.

Morphometric Filtering and Measurement

A. Exclude objects that contact the image border to prevent partial-cell measurements (*see* **Note 6**).

B. Apply optional debris filtering if required to remove non-cellular objects or segmentation artifacts.

C. For each segmented cell, compute morphometric parameters, including:

a. Cell area.

b. Major axis length (cell length).

c. Minor axis length (cell width).

D. If absolute measurement units are required in analysis, convert all measurements from pixels to micrometers using a microscope-specific, calibrated pixel size (e.g., 0.0869 μm/px for the

microscope system used in this work).

E. Export per-image and aggregated measurements as comma-separated value (CSV) files in Photo Phenosizer format.

Statistical Aggregation and Output

A. Combine per-image CSV files using the Photo Phenosizer R scripts to generate per-day and per-condition datasets.

B. Perform statistical summaries and distribution analyses for each morphometric parameter, retaining full cell-by-cell distributions (*see* **Note 5**).

C. Generate plots and tables reporting normalized and raw measurements as required for downstream interpretation.

D. Archive all raw images, segmentation outputs, and processed data files to ensure reproducibility and traceability (*see* **Note 14**).

3.9 DAPI Staining and Nuclear Morphology Assessment

This section describes the fixation, staining, imaging, and analysis procedures used to assess nuclear morphology during chronological aging. Samples used for nuclear analysis are collected in parallel with CFU determination and cell morphology measurements to ensure temporal alignment across all readouts, and representative quantitative and categorical assessments of nuclear morphology are shown in Fig. 3.

3.10 Ethanol Fixation of CLS Samples

A. From the daily CLS sample, transfer a 420-µL aliquot of culture into a labeled 1.5 mL microcentrifuge tube.

B. Add 1 mL 100% ethanol directly to the culture aliquot to achieve fixation. Mix gently by inversion (*see* **Note 10**).

Incubate fixed samples at room temperature for at least 10 min to

C. Incubate fixed samples at room temperature for at least 10 min to allow complete fixation.

D. Store fixed samples at 4 °C. Fixed cells may be stored for extended periods prior to staining without appreciable loss of nuclear integrity.

Preparation of Fixed Cells for DAPI Staining

A. Pellet 100 µL of 70% ethanol–fixed cells by brief centrifugation (1 min) at room temperature.

B. Carefully remove the ethanol supernatant without disturbing the cell pellet.

C. Rehydrate the pellet by gently resuspending cells in 1 mL sterile water to remove residual ethanol.

D. Centrifuge briefly and remove the wash solution completely.

E. Resuspend the rehydrated cells in 20 µL sterile water to obtain a concentrated cell suspension for microscopy.

DAPI Staining

A. Prepare a DAPI mounting solution by combining 99 µL of 50% (v/v) glycerol in water with 1 µL of a DAPI working solution (100 µg/mL), yielding a final volume of 100 µL.

B. If staining intensity is excessive, dilute the mounting solution 1:10 by mixing 15 µL of the DAPI mounting solution with 135 µL of 50% glycerol in water prior to use.

C. Perform DAPI staining during the mounting step by direct application of the diluted DAPI-containing mounting solution to heat-fixed cells on the slide, eliminating the need for a separate staining incubation.

Fluorescence Microscopy of Nuclei

A. Spot 5 µL of the rehydrated cell suspension onto a poly-L-lysine–coated (or charged) microscope slide.

B. Allow cells to adhere by placing the slide on a slide warmer set to low heat until the spot is just dry around the edges, avoiding over-drying. Alternatively, allow the slide to air-dry briefly.

C. Add 5 µL of the diluted DAPI mounting solution directly onto the adhered cells and gently apply a coverslip to minimize bubble formation and prevent cell compression.

D. Seal the coverslip with nail polish to prevent sample dehydration.

E. Image samples using a fluorescence microscope equipped with a DAPI filter set and appropriate UV/blue light excitation and emission settings.

F. For each treatment condition, prepare two independent slides and acquire three non-overlapping fields of view per slide using both brightfield and DAPI fluorescence channels, maintaining consistent exposure and imaging settings across samples.

3.11 Nuclear Morphology Scoring and Quantification

A. Inspect DAPI-stained nuclei visually to assess nuclear morphology.

B. Classify nuclei into qualitative categories such as normal or abnormal, where abnormal nuclei include enlarged, fragmented, irregular, or diffuse DAPI staining patterns (*see* **Note 11**).

C. Record the number of cells in each category and calculate the proportion of abnormal nuclei for each treatment condition.

D. When quantitative analysis is required, measure nuclear area using ImageJ or Fiji by outlining individual nuclei and recording area values (*see* **Note 12**). If threshold-based segmentation is used

values (see **Note 12**). If threshold-based segmentation is used, adjust threshold parameters with reference to the original fluorescence image to ensure accurate representation of nuclear boundaries and to avoid inclusion of background signal.

E. Aggregate nuclear morphology data across fields of view and technical replicates for each condition to ensure representative sampling.

F. Archive raw fluorescence images, measurement files, and scoring records to ensure reproducibility (*see* **Note 14**).

---

4 Notes

1. Temporal alignment is essential. All CFU, morphology, and nuclear measurements should be derived from samples taken from the same culture on the same aging day. Even small timing offsets can obscure relationships between survival, cell morphology, and nuclear integrity.
2. Avoid overhandling aging cultures. As cultures age, cells become increasingly fragile. Gentle resuspension and minimal centrifugation are critical to avoid mechanically induced cell damage that can confound both CFU counts and morphological measurements.
3. Plating across multiple concentrations is strongly recommended. The use of 10×, 5×, and 1× concentrated stocks ensures that countable colonies can be obtained across the full CLS time course, particularly at later aging days when viability declines sharply.
4. Consistency in dilution tracking is critical. Errors in dilution bookkeeping are a common source of variability in CFU calculations. Maintaining a standardized dilution map and recording all transfer steps at the bench is strongly recommended.

5. Do not rely solely on summary statistics for morphology. Aging-associated changes in *S. pombe* are often reflected in distributional shifts rather than changes in the mean alone. Retaining full cell-by-cell distributions provides greater sensitivity to heterogeneity during aging.

6. Border filtering improves morphometric accuracy. Excluding cells that contact image borders prevents partial-cell measurements that would otherwise bias length, width, and area calculations.

7. Imaging conditions must be fixed across the experiment. Changes in illumination, focus, or exposure settings between aging days can introduce systematic bias into automated segmentation and downstream measurements.

8. CellPose parameters should remain constant. Segmentation parameters optimized early in the experiment should not be adjusted mid-study, as parameter drift can mimic biological effects in longitudinal analyses.

9. Agarose pad thickness matters. Pads that are too thick can impair imaging quality, while overly thin pads may dry quickly and distort cell morphology. Pads should be thin, uniform, and freshly prepared.

10. Ethanol-fixed samples are stable but not infinite. While fixed cells can be stored at 4 °C for extended periods, nuclear morphology is best assessed within a consistent time window across all samples to minimize storage-related variability.

11. Nuclear scoring benefits from blind analysis. When possible, scoring normal versus abnormal nuclei should be performed without knowledge of aging day or treatment conditions to reduce observer bias.

Quantitative and qualitative nuclear metrics are complementary

12. Quantitative and qualitative nuclear metrics are complementary. Nuclear area measurements provide objective data, while categorical scoring captures structural abnormalities that may not be reflected in size alone.
13. Expect increased heterogeneity at later aging days. Variability in CFU counts, cell dimensions, and nuclear morphology typically increases as cultures age. This is a biological feature of chronological aging, not an experimental failure.
14. Archiving raw data is essential. Raw images, segmentation outputs, and original CFU counts should be retained alongside processed datasets to allow retrospective validation and reanalysis.
15. This pipeline is modular. While described here as an integrated workflow, individual components (CFU determination, morphology analysis, or nuclear assessment) can be adapted independently for other experimental contexts.

Acknowledgments

We thank the Department of Biology at California State University, Northridge, for technical support and financial assistance provided through startup funds that made this work possible. We are grateful to members of the Escorcia Lab for their thoughtful feedback and discussions during the development of this manuscript. We also thank Celia Cruz, Betty Marion White Ludden, and Calvin Cordozar Broadus Jr. for motivation and encouragement during the execution of this work. Also, the first author would like to thank Michael Joseph Logan and Hudson James Logan for inspiration and a reminder of curiosity throughout this process.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Moreno S (1991) Molecular genetic analyses of fission yeast *Schizosaccharomyces pombe*. Methods Enzymol 194:773

2. Forsburg SL, Rhind N (2006) Basic methods for fission yeast. Yeast 23(3):173–183 [https://doi.org/10.1002/yea.1347][PubMed][PubMedCentral]

3. Sabatinos SA, Forsburg SL (2010) Molecular genetics of *Schizosaccharomyces pombe*. In: Weissman JS, Guthrie C, Fink GR (Eds). Methods in enzymology, Academic Press, Vol. 470, pp 759–795

4. Chen BR, Runge KW (2009) A new *Schizosaccharomyces pombe* chronological lifespan assay reveals that caloric restriction promotes efficient cell cycle exit and extends longevity. Exp Gerontol 44(8):493–502 [https://doi.org/10.1016/j.exger.2009.04.004][PubMed][PubMedCentral]

5. Escorcia W, Shen KF, Yuan JP, Forsburg SL (2019) Examination of mitotic and meiotic fission yeast nuclear dynamics by fluorescence live-cell microscopy. J Vis Exp (JoVE) 148:e59822

6. Stringer C, Wang T, Michaelos M, Pachitariu M (2021) Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 18(1):100–106 [https://doi.org/10.1038/s41592-020-01018-x][PubMed]

7. Vo M, Kuo-Esser L, Dominguez M, Barta H, Graber M, Rausenberger A, Escorcia W (2022) Photo Phenosizer, a rapid machine learning-based method to measure cell dimensions in fission yeast. Micropublic Biol 2022:10–17912

8. Lawson K, Skrtic S, Vo M, Escorcia W (2025) Measuring cell dimensions in fission yeast using machine learning. In *Schizosaccharomyces pombe*: methods and protocols, Springer US, New York, pp 33–46

9. Vo M, Kuo-Esser L, Dominguez M, Driggers J, Kuchinski K, Perez S, Escorcia W (2024) Quality control method and device for producing agarose micropads. Micropublic Biol 2024:10–17912

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_6

6. Practical Considerations for Automated Lifespan Measures in *C. elegans*

Shivay Parihar[1], Athena Alcala[1] and Gilberto Garcia[1] ✉
(1) Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA

✉ **Gilberto Garcia**
Email: gilberto.garcia@usc.edu

Abstract
Many strengths of *C. elegans* as a model organism make them amenable for aging studies, including a short lifespan, the ability to produce a large number of isogenic progeny, and the availability of genetic tools, including genome-wide RNA interference (RNAi) libraries and robust gene-editing strategies. These strengths are complemented by the ability to perform high-throughput lifespan measurements, including automated strategies. Given the development of several systems to automate the measurement of lifespans, deciding between these options can depend on several factors, such as pricing, customizability, throughput, troubleshooting, and more. We provide a discussion of considerations for several methods and provide detailed protocols for implementing a scanner-based lifespan machine system with minor modifications to the existing version.

Keywords Automated lifespan – *C. elegans* – Unbiased lifespan – Lifespan machine

1 Introduction

To improve our understanding of the aging process, researchers have established several features, or hallmarks, of aging that contribute to age-related pathologies [1, 2]. The progression of these features correlates with the passage of time (i.e., chronological age) but can be altered by experimental perturbations to impact the biological age of the organism. Studies in many distinct model organisms have contributed to these discoveries, including mice, fish, flies, and worms [3], each with its unique strengths and weaknesses. Among these, the nematode *C. elegans* is one of the shortest-lived models, with a maximum lifespan of approximately 3–4 weeks in standard lab conditions at 20 °C, allowing for relatively quick experimental measurements of lifespan [4]. Additionally, worms offer several advantages to experimentally test potential aging interventions, including an extensive genetic toolkit for genetic manipulation, transparent bodies for imaging, ease of handling and growth, and the ability to obtain a large number of isogenic progeny from hermaphroditic adults [5]. Their small space requirement, simple setup, and quick experimental turnaround allow for effective testing of how multiple different genetic, dietary, or pharmacological interventions affect the trajectory of aging [6–8].

Tracking statistically significant numbers of microscopic *C. elegans* throughout their lifespan brings about several challenges. In its most basic form, standard longitudinal lifespan analyses involve manual counting and tracking the living status (alive or dead) of each animal (usually around 100 per condition) continuously for 3 weeks at 20 °C [4, 9]. Their ability to self-fertilize requires the implementation of methods to select for the adults and eliminate progeny. The simplest method is the daily transfer of each hermaphrodite adult to new plates or media throughout its reproductive lifespan of about 5 days. Alternatively, pharmacological or genetic methods to prevent progeny can be implemented [10, 11]. Lifespan assays are further complicated by the functional decline in the motility of animals with age [12]. While the correlation between movement and aging status allows for a simple method to determine their living status,

it requires using a physical stimulus at later ages to induce movement, usually either through touch or light stimuli [6, 8, 13]. This increases both the time required to evaluate each animal with the conventional metal pick and the potential for lifespan-altering injury caused by excessive prodding or movement of the nematodes. Finally, there exists the potential for bias with human judgment when determining an animal's life status, as some animals may take longer to respond to stimuli than others. Therefore, methods to automate this labor-intensive assay in *C. elegans* have been developed to mitigate some of the mentioned challenges while also increasing the possible throughput of lifespan analyses [14–17].

Systems designed to track the nematode's lifespan have varying degrees of setup difficulty, automation, scalability, space requirement, and cost. Commercially available systems, such as the NemaLife system, tend to have easier setup but potentially higher costs [18]. The NemaLife system uses a small microfluidic chamber to filter out progeny from the adults with a sieve. This method allows for easy administration of pharmaceuticals for drug intervention studies while avoiding the potentially confounding use of chemicals or mutations to prevent progeny. One caveat this introduces is the factor of a fluid-submerged lifespan as opposed to the conventional agar surface-based lifespan. However, while being placed in liquid normally causes animals to constantly "thrash" and affects development and lifespan, NemaLife utilizes micropillars to allow animals to crawl and emulate aspects of growth on an agar plate. The worms are digitally monitored in the microfluidic device for downstream analysis of survival, among other possible metrics. Based on the set parameters, this method can yield more accurate worm survival data than checking their survival manually just once a day. This system greatly reduces the labor and strict time requirements of a standard lifespan, although the initial cost of the system and need of proprietary consumables should be taken into consideration. As with any commercial products, companies are incentivized to provide customer service to simplify the user experience and may continue to introduce upgrades and innovations to improve their product and lifespan metrics.

Self-assembled open-source systems for *C. elegans* lifespan automation have been developed at varying price points, space requirements, and labor demands. One of the smallest and most cost-effective of these systems is the SiViS machine [15]. The SiViS consists of a small 3D-printed container with two separate compartments. One compartment contains a single-board computer, and the other compartment houses the imaging area. NGM plates with Fluorodeoxyuridine (FUdR)-treated worms are fitted into 3D-printed pallets that hold two plates in the same orientation throughout the experiment. The plate pallets are stored at the desired growth temperatures and taken to the SiViS daily to be imaged. Multiple pallets can be assembled and sequentially imaged for larger-scale experiments. Worms are imaged using two 5MP cameras and a seven-inch display that provides adaptive backlighting to improve segmentation [19]. Users set up, capture, and process the images on a PC running the SiViS software. While this system requires daily manual swapping and imaging of the plates, the determination of the survival is automated, greatly reducing the labor requirement of the lifespan analyses. Other points to consider with this system are its use of conventional NGM plates and the convenience of needing only one SiViS machine to process multiple plates of worms sequentially.

Systems for automated large-scale experiments with plate carousels or robotic arms have also been developed, but they increase in complexity and require more space, custom parts, and robotics assembly [17, 20]. A potentially more accessible system is the *C. elegans* Lifespan Machine, an open-source scanner-based system [14]. While high-quality photo scanners do come at a premium cost, they are designed to capture images at multiple locations, thereby simplifying the automated image acquisition of multiple plates. The imaging area is limited to 16 plates, which are loaded onto a scanner and automatically scanned at regular intervals after the initial setup. A scan of all 16 plates takes about an hour, providing an increased temporal resolution for measuring survival compared to manual lifespan assays, where plates are typically checked once a day and more frequent checks increasing the labor requirement. Scans are transmitted across a local network for analysis of worm survival by the Lifespan Machine open-source software, which provides options for manual validation and survival annotation by the user. Similar to the SiViS system, the lifespan machine maintains worms on conventional NGM plates for the experiment and requires sterilization methods such as treatment with FUdR or the use of a sterile mutant. Due to the slow imaging speed and enclosed nature of the scanner, heat generation during the scanning process must be taken into

consideration. However, this issue can be resolved with affordable box fans [21]. With a limit of 16 plates per scanner, larger-scale experiments would require additional scanners. Since each scanner's location and setup may not be identical, one should include a control strain on each scanner to account for these differences, which reduces the number of experimental slots per machine.

Here, we present our methodology for a version of the scanner-based Lifespan Machine. We believe this method combines several key advantages, like reasonable cost, moderate throughput, ease of construction, and complete imaging automation after a one-time setup per experiment. It also offers users customizability for both construction and use. While an online community does exist (https://groups.google.com/u/1/g/automated_lifespan) for aid in troubleshooting issues, the activity on the group has diminished with time. Through this chapter, we aim to familiarize the reader with an automated, customizable system that can track a large number of worms for their entire lifespan after the initial experimental setup.

2 Materials

The open-sourced nature of the Lifespan Machine enables users to use different commercially available components to implement the core system. We utilized the Epson V850 scanner, which is very similar to the Epson V800 scanner previously used [21]; however, other scanners supported by the SANE-project.org drivers could potentially be used with modifications to the system code. Raspberry Pi single-board computers (RasPi) were selected as imaging and database servers due to having sufficient computational power at their price and compact size. Alternatively, other Linux systems such as those described by the developer can be utilized for those servers [14]. We chose a Synology NAS for data storage due to its ease of use, potential for expansion, and local implementation. A PC was used as an analysis server due to the processing power at the price point and the ability to run the Worm Browser program required for the system. It should be noted that various components of the Lifespan Machine transmit data over the local network. Therefore, setting up the system requires access to the network router settings to reserve IP addresses for consistent connectivity. Alternatively, hostnames can be assigned and used in place of the IP address, but we will only refer to the IP address from here on. For simplicity, a list of common items such as screwdrivers and Ethernet cables, will not be noted, but less common or dedicated components are listed in Table 1, along with their space requirements. These are based on the items required by the original developer of the Lifespan Machine [14].

Table 1 Dimensions and quantities of components of the lifespan machine

2.1 *C. elegans* Culturing Media

C. elegans are conventionally cultured on Nematode Growth Media (NGM) agar plates kept at 20 °C with lawns of various *E. coli* strains as food. The two most common strains are OP50, a vitamin B12-deficient strain generally used for standard growth, and HT115(DE3), a strain used almost exclusively for RNA interference (RNAi) [22]. RNAi plates are NGM plates with the addition of carbenicillin for RNAi target plasmid selection and Isopropyl β-D-1-thiogalactopyranoside (IPTG) for induction of the T7 polymerase to synthesize dsRNA in the HT115 bacteria. Tetracycline is not used in the RNAi plates to prevent any effects the antibiotic may have on worms but is used when culturing RNAi bacteria to select for RNAse III-deficient bacteria that have enhanced dsRNA production.

1. Lysogeny broth (LB): All standard LB recipes are acceptable for growth of *E. coli*, including homemade recipes using Bacto-tryptone, yeast extract, and NaCl.
2. Stock antibiotic solutions: Carbenicillin 100 mg/mL in water and Tetracycline 5 mg/mL in ethanol.
3. *E. coli* strains OP50 and HT115.
4. Nematode Growth Media (NGM): 1 mM $CaCl_2$, 5 μg/mL cholesterol, 19.9 mM KH_2PO_4 monobasic, 5.11 mM K_2HPO_4 dibasic, 1 mM $MgSO_4$, 0.25% (w/v) Bacto-Peptone, 51.3 mM NaCl; 2% (w/v) agar.

5. NGM RNAi plates: 1 mM $CaCl_2$, 5 μg/mL cholesterol, 19.9 mM KH_2PO_4 monobasic, 5.11 mM K_2HPO_4 dibasic, 1 mM $MgSO_4$, 2% (w/v) agar, 0.25% (w/v) Bacto-Peptone, 51.3 mM NaCl, 1 mM IPTG, and 100 μg/mL carbenicillin/ampicillin. Store at 4 °C in the dark for up to 3 months.
6. Plates: 60 mm Tritech Petri dishes, 100 mm Falcon Petri dishes.
7. M9 solution: 22 mM KH_2PO_4 monobasic, 42.3 mM Na_2HPO_4, 85.6 mM NaCl, 1 mM $MgSO_4$.
8. 5-Fluoro-20-deoxyuridine (FUdR): 10 mg/mL dissolved in M9 solution.
9. Bleaching solution: 1.8% (v/v) sodium hypochlorite, 0.375 M KOH.

2.2 Epson V850 Scanner with Modified Focus

The Epson Perfection V850 Pro scanner is a commercially available photo scanner with LED light sources and a transparency unit. The LED light sources produce less heat compared to conventional lamp illumination. This reduced cooling requirement eliminates the need to cut holes into the scanner for cooling [21]. The transparency unit (TPU) is necessary to image nematodes as dark objects on a light background for the image analysis. An 11.75-inch × 8.75-inch glass sheet purchased and cut at a local hardware store serves as a plate carrier/assembly to mount plates onto the scanner. The scanner lens focal point must be adjusted to focus on the worms that will be on NGM RNAi plates on the glass sheet. The amount of focus adjustment needed will depend on the thickness of the glass sheet as well as the height of the NGM media in the plates.

1. Disconnect the main body of the scanner from the power supply, computer, and TPU.
2. Remove the four circular screw covers (*see* **Note 1**) and remove the screws that hold the scanner lid to the chassis.
3. Ensure that the scanner locks (gray slider located at the back left of the scanner and the bottom side near the back on the TPU) are unlocked and carefully pry up the lid and place it aside, making sure the glass remains clean (*see* **Note 2**).
4. Free the scanning carriage assembly from the drive belt by pressing down on the shiny metal clip holding the belt against a bracket on the carriage, and store it in a safe location (Fig. 1). Make sure not to tilt the scanner after removing the clip, as the carriage will move freely and quickly along the rod after undoing the clip.

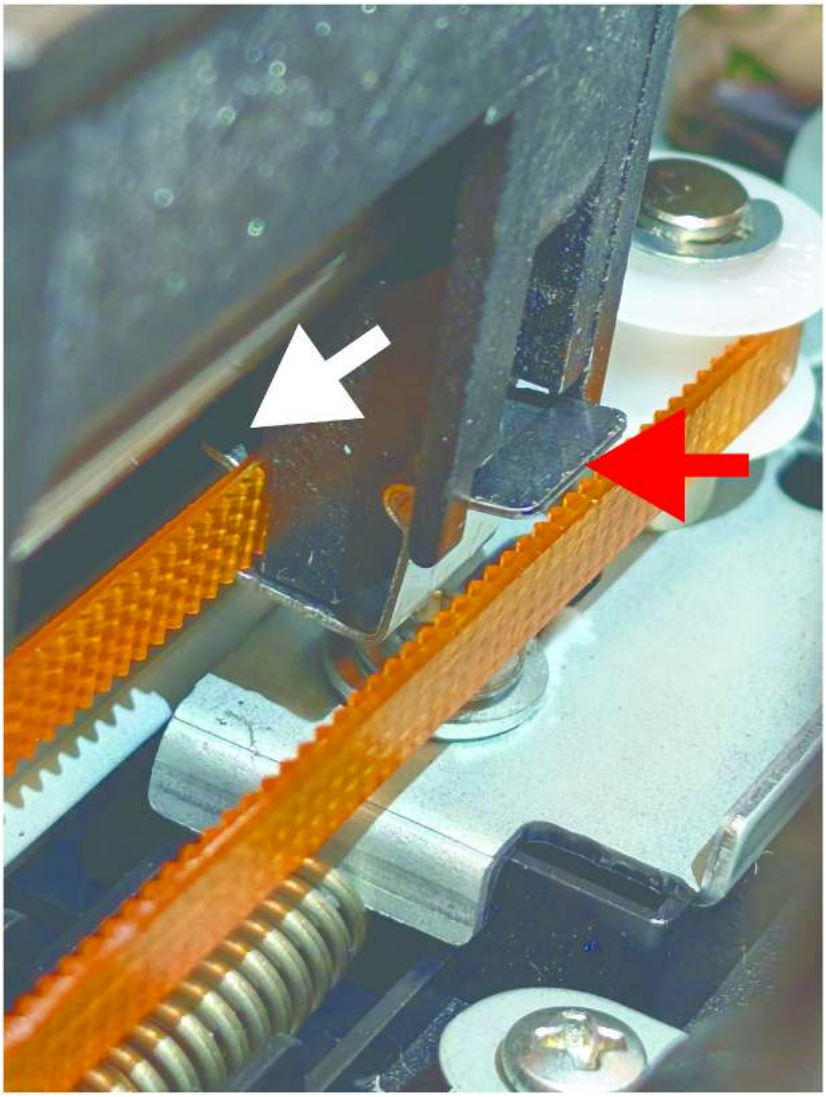


Fig. 1 Close-up image of the scanning carriage assembly clip. Image of the scanning carriage assembly metal clip securing the orange belt to the carriage bracket. The red arrow denotes the outer clip edge pushed to release the clip. The white arrow denotes the inner part of the clip correctly installed

5. Remove the screw holding the front cover in place and remove the cover.
6. Remove the nut at the front of the scanner that locks the scanner's linear rod in place. The nut will be visible from the front of the scanner once the front cover is removed.
7. Turn the backend of the linear rod about a quarter turn to release the rod from the mount.
8. Remove the rod from the mounts at both ends and flip over the scanning carriage. Hold both the rod and the carriage when flipping the parts over, while making sure not to damage the ribbon cable connecting the carriage to the chassis.
9. Remove the screws holding the bottom metal cover of the carriage and remove the cover.
10. Remove the black glue holding the lens tube in place in the square hole near the center of the carriage. The glue will look like black bubblegum. Use a small flat screwdriver to pry the glue edges and needle-

nose pliers to pull on the glue.

11. Replace the bottom cover and screws, making sure the plastic flap is tucked under the cover.
12. Flip the scanning carriage back over and place it on the guide rails while also inserting the linear rod back into its mount at both ends.
13. Turn the backend of the rod a quarter-turn to lock it back in its mount and replace the nut at the front end to secure the rod in place.
14. Secure the scanner carriage to the drive belt by reinstalling the metal clip. Make sure that the clip is holding the belt in the slot between itself and the carriage underside and not pinched between the clip and the carriage (Fig. 1). The carriage should no longer move freely (*see* **Note 3**).
15. The front cover can be replaced if desired. Then, place the scanner lid back on the chassis, but do not replace the screws.
16. Connect the scanner to the TPU, power supply, and computer, and boot up the scanner and Epson Scan utility software.
17. Lift the TPU and place the glass sheet on the scanner bed.
18. Place an inverted NGM plate with worms onto the glass sheet without its lid so the opening is sealed against the glass sheet.
19. Lift the TPU up a little before closing it over the plate so that the TPU sits a small distance above the plate. Cut packing foam or other semi-soft material into cubes that are 0.75 in^3 and affix them to the scanner lid or TPU to support the TPU evenly above the plate.
20. Set the scanning software to Document type: Film, Film Type: Positive Film, Image type 16-bit Grayscale, Resolution:9600, and press “Preview” to perform a quick scan (Fig. 2a). After the preview scan, select a small region on the plate where worms are present and press “Scan” to obtain a full-resolution scan of the small region.

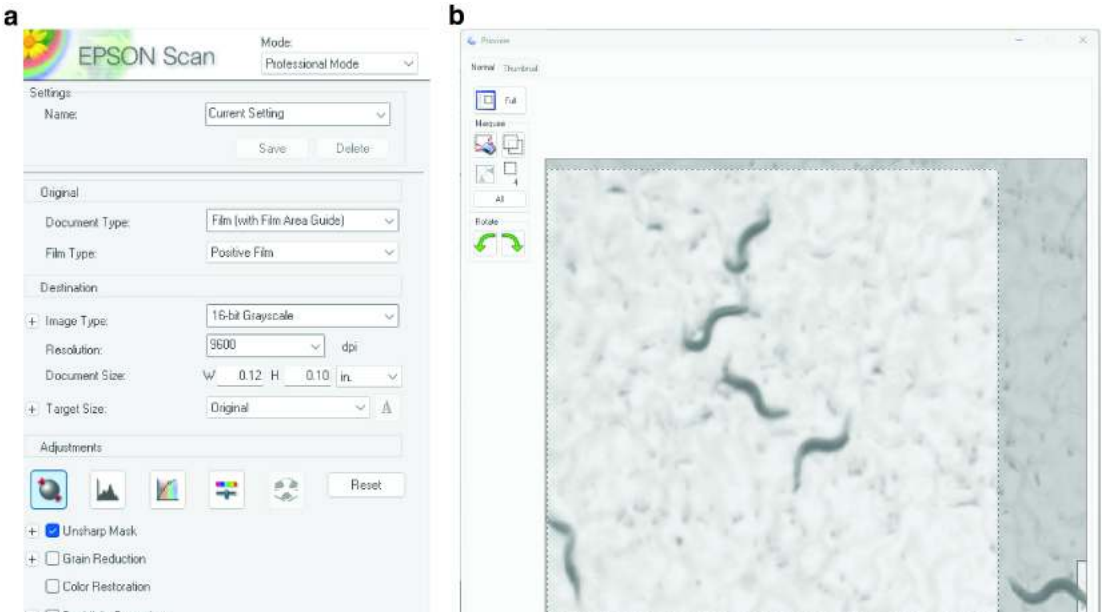




Fig. 2 Screenshots of scanning software settings and a scanned image. Image of the scanning software displaying worms in focus after adjusting the scanner lens focus. **a**, Screenshot of scanner software settings. **b**, Screenshot of worms scanned in focus

21. If the worms are in focus (Fig. 2b), the four screws can be replaced to secure the lid to the chassis. If the worms are not in focus, remove the TPU from the scanner without disconnecting it, and then remove the lid with the glass sheet and plate and place it aside.
22. Remove the thin plastic cover from the top center of the scanning carriage by using a flat screwdriver to pry it out to expose the lens tube (Fig. 3).

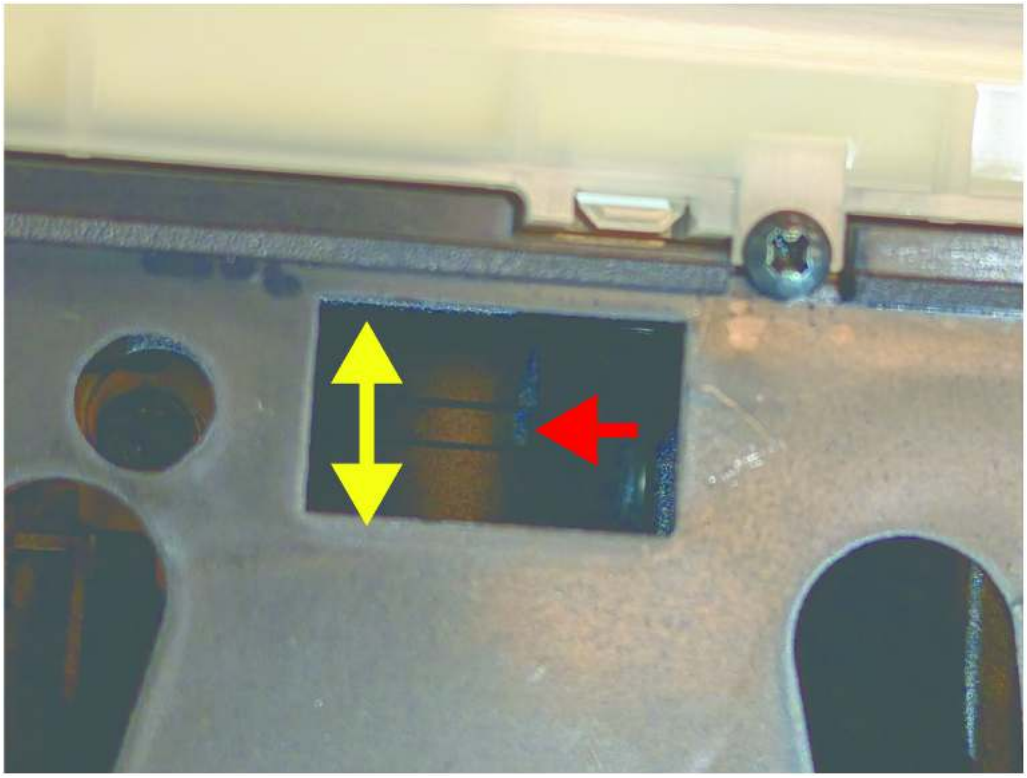


Fig. 3 Image of the scanner lens tube to be adjusted to change scanner focus. The red arrow denotes the groove used to move the lens. The yellow arrows denote the direction of movement when adjusting the lens focus

23. Mark the exposed lens tube along the edge where it disappears into the frame using a pencil.
24. Adjust the lens focus by placing a flat-tipped screwdriver on the groove on the lens tube and moving the tube parallel to its length (Fig. 3). Make very small adjustments that are barely noticeable, using the pencil marking as a reference to determine if the lens has moved toward or away from the frame.
25. Replace the lid with the glass sheet and the worm plate onto the scanner. Reinstall the TPU, making sure it is evenly supported above the plate.

26. Repeat Step 20 until the worms are adequately in focus and the screws are replaced to secure the lid.

2.3 Network Data Storage

The system requires a large network storage system connected through Ethernet cables to provide fast data transfer between the different components of the Lifespan Machine. An experiment will require about 100–125 GB of storage, so an adequate amount of storage should be provided if raw data storage will not be actively monitored and deleted (*see* **Note 4**). These instructions are for our setup using a Synology NAS and will differ if a different data storage system is used.

1. Insert the necessary hard drives into the Synology NAS, connect the device to your network, and turn the device on.
2. Follow the manufacturer's instructions to set up the system and create a shared folder to store the Lifespan Machine data.
3. Enable NFS protocol accessibility by opening the NAS control panel and selecting Shared Folder. Select the shared folder created for the LS Machine data and click Edit. In the NFS permissions tab, create new rules for each component of the LS Machine. Ensure that the correct IP addresses are input and that the Read/Write privilege is allowed. Map all users to admin with sys, krb5i, and check all the option boxes. If the IP addresses of the other components are not known, this step can be completed once the IP addresses have been established.
4. Write down the mounting path of the LS Machine data folder located at the bottom of the NFS Permissions tab (example: /volume1/NASshared).
5. Open the File Station app on the NAS admin page and select the shared folder that was created for the LS Machine data. Click "Create" and make a new folder as the main directory that will store the raw data (RawDataFolder).
6. Connect to your network router and reserve an IP address for the NAS to provide a consistent network address to the device.

2.4 Imaging and Database Servers

The imaging and database servers (DBserver) are set up on Raspberry Pi 4, 4GB single-board computers running Raspberry Pi OS on 64GB micro-SD cards. Alternatively, USB3 to SATA adapters can be used to run the computers on Solid State Drives instead of SD cards for greater stability (*see* **Note 5**). Raspberry PI OS Bullseye was installed on this system, and Buster was used at the time the original instructions for the RasPi were written. While older or new releases of the OS can be installed, packages can change with releases and may require troubleshooting if package changes for the system to function properly. The DBserver maintains a database of experimental setups, raw data file names, raw data organization, and data processing jobs, as well as running the web interface. The imaging servers acquire the raw data directly from the scanners, transfer it to the network storage, and route necessary information to the database. The minimum requirements would be one database server and one imaging server. We have run two scanners per imaging server without issues. The original RasPi instructions can be found at https://github.com/nstroustrup/lifespan/tree/flow/doc.

1. With a micro-SD adapter and two 64GB micro-SD cards (one imaging server and DBserver), use the Raspberry Pi Imager software to install Raspberry Pi OS (64-bit) Bullseye onto each SD card. Due to Bullseye being an older version of the OS, the .img file will need to be downloaded from the Raspberry Pi OS archive and installed as a custom OS with the Raspberry Pi Imager.
2. Perform the remaining steps on one computer at a time, beginning with the DBserver. Install a RasPi computer in any desired case, install the SD card, and connect it to the network with an Ethernet cord.

Connect a keyboard, a mouse, and a display to the computer. Lastly, connect the power supply.

3. Follow the prompts to create a username, unique hostname (set it to the same hostname as on the network router if one was assigned), and password. After rebooting, go to Preferences > Raspberry Pi Configuration, and allow for SSH connections. Make sure to write down the hostnames, usernames, and passwords for each server OS.
4. Connect to your network router interface and reserve an IP address for the computer to have a consistent network address.
5. Disconnect the display, mouse, and keyboard from the computer and perform the following steps over SSH. On a PC on the same network, use SSH on a terminal window to connect to the RasPi with the command:
    a. ssh username@IP address
6. Enter your password to connect to the DBserver and run the following commands to update the OS:
    a. sudo apt-get update
    b. sudo apt-get upgrade
7. Run the following commands to install the necessary packages (*see* **Note 6**):
    a. sudo apt-get install git cmake vim gcc apache2 apache2-utils php php-mysql mariadb-server mariadb-client mysql-common libmysql++-dev default-libmysqlclient-dev libjpeg-dev libpng-dev libpng++-dev libdmtx-dev
    b. sudo apt-get install zlib1g-dev libusb-1.0-0 libusb-1.0-0-dev libusb-dev cifs-utils libsane-dev screen libcurl4-gnutls-dev libxml2-dev libsnmp-dev libpoppler-glib-dev nfs-common
8. Run the following commands to configure git:
    a. git config—global user.email email@email.com #enter your email
    b. git config—global user.name “your name here” #enter your name
    c. git clone https://github.com/nstroustrup/lifespan
    d. cd lifespan
    e. git checkout flow
9. Enter the following commands to install additional dependencies:
    a. cd external_compile_libraries/libtiff-4.0.9
    b. cmake.
    c. sudo make install
    d. cd..

e. wget https://github.com/uclouvain/openjpeg/archive/refs/tags/v2.4.0.tar.gz
f. tar -zxvf v2.4.0.tar.gz
g. cd openjpeg-2.4.0
h. cmake.
i. sudo make install
j. cd..
k. wget https://gitlab.com/sane-project/backends/uploads/8bf1cae2e1803aefab9e5331550e5d5d/sane-backends-1.0.31.tar.gz
l. tar -xvzf sane-backends-1.0.31.tar.gz
m. rm -r sane-backends-1.0.27
n. cd sane-backends-1.0.31./configure BACKENDS="epson2" RELOADABLE_BACKENDS="epson2"
o. sudo make
p. sudo make install
q. cd../..

10. Install the Lifespan Machine code with the following commands:
    a. cd build
    b. cmake. -DONLY_IMAGE_ACQUISITION=1
    c. sudo make install
11. Use the following commands to create a directory and then backup and access the file system table:
    a. sudo mkdir/mnt/NASmount #any directory name can be chosen
    b. sudo cp/etc/fstab /etc/fstab.backup
    c. sudo nano/etc/fstab
12. Insert the following text as a new line at the bottom and save the file, but enter the IP address of the NAS and the correct location of the shared folder (*see* **Note 7**):
    a. NetworkStorageIPaddress/location /mnt/NASmount nfs auto,nofail,noatime,nolock,intr,tcp,actimeo=1800 0 0
13. The NAS will be mounted when the computer boots, but the following command will mount the NAS without rebooting. Once mounted, the NASmount directory should display the contents of the raw data

Without rebooting. Once mounted, the NASmount directory should display the contents of the raw data directory from the NAS:

a. sudo mount /mnt/NASmount

14. Run the following commands to copy and access the server configuration file:

a. cd..

b. sudo cp doc/ns_image_server.ini.example.pi /usr/local/etc/ns_image_server.ini

c. sudo nano /usr/local/etc/ns_image_server.ini

15. Edit and save the following lines in the file, using the correct IP addresses for the servers and locations for the directories (*see* **Note 8**):

a. host_name = serverHostname #unique display name for the current server to be displayed as on the web interface

b. long_term_storage_directory = /mnt/NASmount

c. results_storage_directory = /mnt/NASmount/Results

d. volatile_storage_directory = /home/username/volatile_storage # username of the current computer

e. central_sql_hostname = [DBserver IP address] # leave the hostname as the default “localhost” for the DBserver, otherwise, add the DBserver IP address

f. central_sql_password = password #remember the password used

g. local_buffer_sql_password = password #set the same password used for the central_sql_password

h. act_as_image_capture_server = no #set to “no” for DBserver and “yes” for the image server

i. verbose_debug_output = true

j. maximum_memory_allocation_in_mb = 2048 # “3840” for DBserver

16. Run the following command to make the volatile directory that was specified in the previous step:

a. mkdir /home/username/volatile_storage #specify the correct username

17. Access the MySQL Database configuration file with the following command:

a. sudo nano /etc/mysql/mariadb.cnf

18. Configure the Database file by adding the following lines and saving the file:

a. [client-server]!includedir /etc/mysql/conf.d/!includedir /etc/mysql/mariadb.conf.d/

b.

b. [mysqld]
c. key_buffer = 16 M
d. read_buffer = 60 K
e. sort_buffer = 1 M
f. innodb_buffer_pool_size = 64 M
g. tmp_table = 8 M
h. max_allowed_packet = 16 M
i. thread_stack = 192 K
j. thread_cache_size = 8
k. bind-address = 0.0.0.0
l. myisam-recover = BACKUP
m. max_connections = 150

19. Run the following to enable and start the database, check its status, and then exit the status screen by pressing Ctrl and C at the same time:
    a. sudo systemctl enable mariadb
    b. sudo systemctl start mariadb
    c. sudo systemctl status mariadb #Should be “active(running),” press Ctrl + C to exit
20. Create the database root user and password with this command:
    a. sudo mysqladmin -u root password #You will be prompted for a password
21. Create and configure the database with the following command. Enter “root” as the user and password if prompted:
    a. sudo ns_image_server create_and_configure_sql_db #Enter “yes” for DBserver
22. Update the database with the following command:
    a. sudo ns_image_server update_sql
23. This step and onward will only apply to the DBserver computer. Use the following command to enter the database to grant permissions:
    a. mysql -u root -p

24. Grant permissions by entering the following lines. They must be typed in, not pasted. Use the database password that was previously set:
    a. GRANT ALL on *.* TO 'image_server'@'localhost';
    b. GRANT ALL on *.* TO 'image_server'@'localhost' identified by 'password';
    c. GRANT ALL on *.* TO 'image_server'@'%';
    d. GRANT ALL on *.* TO 'image_server'@'%' identified by 'password';
25. Exit the DBserver database by typing the following:
    a. exit;
26. Update the database after the permission changes with the following command:
    a. sudo ns_image_server update_sql
27. Copy over the necessary files to run the web interface and copy the configuration file template with the following lines:
    a. sudo cp -r web_interface/* /var/www/html
    b. sudo cp /var/www/html/image_server_web/ns_image_server_website_template.ini /var/www/html/image_server_web/ns_image_server_website.ini
28. Grant access to the website with the following command:
    a. sudo chmod 755/var/www/
29. Use the following command to enter the web interface configuration file:
    a. sudo nano /var/www/html/image_server_web/ns_image_server_website.ini
30. Change the sql username to "image_server" and the password to match what was set as the database password. Save and exit the file.
31. Remove the default web interface file with the following command:
    a. sudo rm/var/www/html/index.html
32. Check that the website functions by opening a web browser on a PC on the same local network and going to the following page, but replacing "DBserverIPaddress" with your DBserver IP address: http://DBserverIPaddress/image_server_web/view_experiments.php
33. Enter the following command to create a link between the website directory and the long-term storage directory so the website can access the images:

a. sudo ln -s /mnt/NASmount /var/www/html/long_term_storage

2.5 Analysis Server and Worm Browser

The analysis server and Worm Browser program were set up on a PC desktop running Windows 11 with 32GB of RAM, an i9-10900K CPU, and a Quadro P1000 GPU. The analysis server performs the image analysis on the raw data to separate samples, identify worms, determine movement, and determine the time of death of individual worms. The Worm Browser program provides essential utilities to set up experiments and to validate worms once the experiment has been completed.

1. Map the Synology NAS shared folder (ipaddress/NASshared) as a network drive (example: Z) and check that you can access data files by going to the drive the "This PC" directory.
2. Create a folder named "server" and a folder named "Volatile_Storage" directly in the C:\ directory.
3. Download the .exe files, .dll files, and the ipp_dlls directory from https://github.com/nstroustrup/lifespan/tree/flow/binaries/windows and save them into the "server" folder created in Step 2.
4. Double-click the ns_worm_browser_x64.exe file. This will create an ns_image_server.ini file in the same directory. Open the file with a text editor and edit the following lines:
   a. host_name = [enter a unique name for the analysis server to display as on the web interface]
   b. long_term_storage_directory = Z:\RawDataFolder #use correct directories
   c. results_storage_directory = Z:\RawDataFolder\Results #use correct directories
   d. volatile_storage_directory = C:\Volatile_Storage
   e. central_sql_hostname = [enter IP address for the DBserver computer]
   f. central_sql_password = [enter the database password]
   g. local_buffer_sql_password = [enter same password as the central_sql_password]
   h. nodes_per_machine = 4 #default is 1 but can be increased if your computer has the resources
   i. verbose_debug_output = true
   j. maximum_memory_in_mb = 10240 #this will depend on your system resources
5. Double-click the ns_worm_browser_x64.exe file. A window will pop up stating that an ns_worm_browser.ini file was created. Double-click the ns_worm_browser_x64.exe file to run the Worm Browser program. A terminal window and the Worm Browser window will pop up.
6. Close the Worm Browser program and double-click the ns_image_server_x64.exe file to confirm that the analysis server will run. A terminal window will pop up and display text that will end with "Dispatcher started." The window can be closed afterward.

2.6 Custom Solutions

Various components are not absolutely necessary but aid with various issues. The glass sheet allows for assembly of the plates outside the scanner and helps keep the scanner clean since the plates are not placed directly on the scanner. We also use a second glass sheet to help the plates maintain a good seal with the

bottom sheet. Alternatively, airtight sealed plates can be used and placed directly on the scanner [21]. Instead of cutting rubber mats to hold the plates, we designed and 3D printed a plate holder for a more custom fit. This helps with plate alignment and prevents plate movement in the event the scanner is nudged in any way. While the V800/V850 scanner no longer requires modifications with cooling holes, the system still requires cooling that can be accomplished with 20-inch box fans [21]. Due to our setup inside an incubator, we needed to use smaller fans and 3D-printed holders to have the fans at the correct height, but any rigid item capable of holding the fan at the right height would suffice.

1. Custom-sized glass sheets were purchased at a local hardware store. The dimensions were 8.75 in × 11.75 in. The edges were taped for safety, except the edge meant for the back of the scanner, which the scanning head requires clear for calibration. A square located at the bottom left corner, 1.625 in wide and 1.5 in high, was drawn with a marker to designate the barcode area. A line was drawn 0.75 in from the edge for the back-end of the scanner to mark off the calibration area that must remain clear, and another line 0.75 in from the edge was meant for the front-end to designate the limit of the scanning area (Fig. 4a).

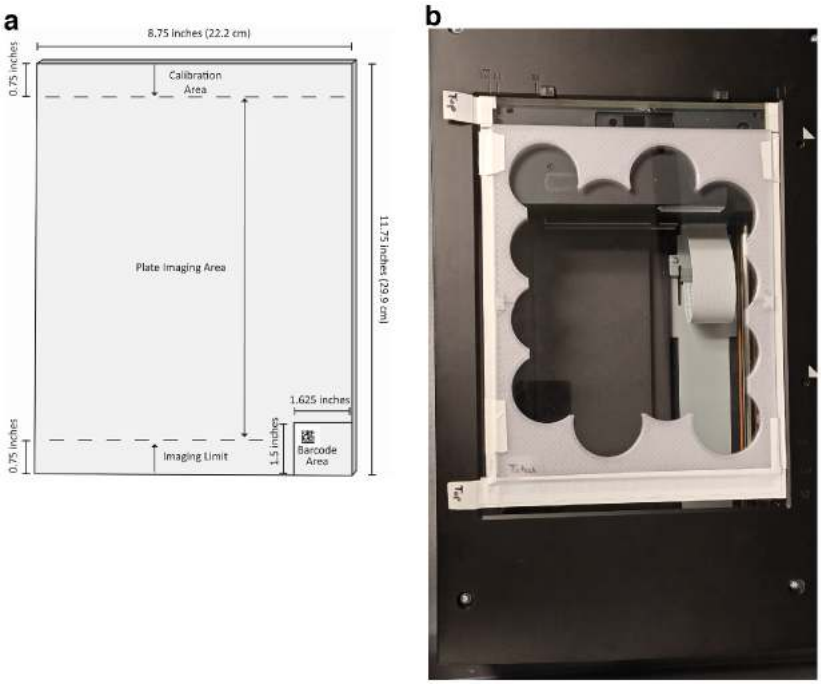


Fig. 4 Lifespan Machine glass sheet diagram as seen from the bottom and image of the glass sheet assembled with the plate holder placed on the scanner. **a**, Diagram of a glass sheet with regions and dimensions specified. **b**, Image of the glass sheet with the plate holder attached and placed on the scanner bed

2. The plate holder was designed for Tritek (T3315) plates using FreeCAD software and 3D printed with a Fused Deposition Modeling printer. The holder is attached to the glass sheet with lab tape to allow for easy removal for cleaning (Fig. 4b) (*see* **Note 9**). The STL files can be found at: https://github.com/SanabriaLab/STL-Files/tree/main/Lifespan%20Machine%20custom%20STLs
3. Small 5-inch USB-powered desk fans were used to improve cooling during the scanning process.
4. Fan holders were designed using FreeCAD software and 3D printed with a Fused Deposition Modeling printer. Fan holders were designed to allow the fan to be aimed at the plate assembly at the center of the scanning bed.

3 Methods

Once the individual Lifespan Machine components have been set up, the components must be activated in a specific order for the system to properly connect. The NAS must be active for the DBserver and imaging server to connect at boot. The DBserver must be online for the imaging server to connect to the database. Scanners must be connected to the imaging server RasPi computer prior to initializing the server for the scanners to be registered by the system through the barcodes. In the case of network failure, the imaging servers will continue imaging and store data in their volatile memory directory and transfer the data once the network has been reestablished.

3.1 Growth, Synchronization, and LS Machine Preparation of *C. elegans*

General maintenance and expansion of *C. elegans* are performed using OP50. Due to the OP50 strain's ability to grow thick lawns, we use HT115 containing the empty RNAi vector pL4440 (EV) to prevent lawn thickness from impacting the imaging of the worms. Additionally, the use of RNAi plates with antibiotics may help reduce potential contamination. It should be noted that the two bacterial diets result in differing physiological changes in worms, including but not limited to gene expression, lifespan, and metabolic profiles [23, 24]. The OP50 diet has been shown to result in defects in mitochondrial homeostasis and increased sensitivity to pathogens and stress [25]. For these reasons, we maintain worms on an OP50 diet and conduct experiments on HT115.

1. Grow OP50 cultures in LB at 22 °C–25 °C for 24–48 h to saturation.
2. Seed 200 µL of saturated OP50 culture onto 60 mm plates and 800 µL onto 100 mm plates. Allow the plates to completely dry before plating worms onto the bacterial lawn. Plates are usually left to dry overnight but can dry quicker depending on different factors, such as the dryness of NGM plates, humidity, and room temperature.
3. To maintain a steady supply of worms, move about 10–15 young worms (eggs, L1, or L2 stages) onto a new OP50 plate weekly if grown in 15 °C. Worms can be grown in 20 °C but will need to be moved more frequently (about twice a week) to prevent starvation.
4. To expand worms for the Lifespan Machine, cut out about a 1 cm × 1 cm piece of the agar where worms of interest are located and place the agar chunk face-down (worm side down) onto 100 mm OP50 plates. Chunking like this from a full 60 mm plate of animals with wild-type fecundity will generally produce a full 100 mm plate in 2–3 days at 20 °C without starvation. Strains with lower fecundity will either need bigger chunks or more chunked plates to ensure enough worms will be acquired.
5. Grow HT115 cultures in LB with tetracycline (1:500) and carbenicillin (1:1000) at 37 °C in a shaker for 16–24 h on the day that worms for the experiment were chunked (*see* **Note 10**). Ensure that the plates are dry before adding any worms to them.
6. Seed 200 µL of the HT115 culture onto at least 16 RNAi plates for the lifespan machine.
7. Check the chunked expansion plates for adult worms that are full of eggs. This denotes that the plate is ready for bleaching.
8. Collect worms by washing the plate with M9 and transferring to a 15 mL tube via a sterile serological pipette.
9. Spin the worms for 30 s at 1.1 × RCF.
10. Aspirate the M9 as much as possible without aspirating the worms.
11. Add 5 mL of bleach solution to the 15 mL tube of worms. For samples with >500 µL of worms, use 10

12. mL of bleach solution to ensure all animals are sufficiently bleached.
    At the 4-min mark, check for disintegration of worms. Worms should look like they're falling apart with free-floating eggs. If there are no worms or eggs visible, the sample might be overbleached.
13. Spin the worms for 30 s at 1.1 g and wash with M9. Repeat for a total of three washes.
14. Add M9 up to 12 mL and place in a tube rotator overnight.
15. Plate ~60 synchronized L1s onto the 60 mm HT115 plates and incubate at 20 °C once the liquid is dry.
16. After incubation for about 60 h, add 100 µL of 10 mg/mL FUDR to worms at late L4s or early adult stages (*see* **Note 11**).
17. Allow the plates to dry before placing them on the scanner.

3.2 Initializing the Lifespan Machine

Initialization of the Lifespan Machine system requires that all components be on the same local network. Once started, the system can be left on indefinitely, but imaging servers, scanners, and analysis servers can be shut down when not in use for imaging or analysis. The DBserver and NAS must remain on for annotation and downloading the data; otherwise, they can also be shut down if all experiments have been completed.

1. Ensure that the Synology NAS and DBserver computer are on.
2. Establish an SSH connection to the DBserver computer on a terminal window and ensure that the NAS is properly mounted (*see* **Note 12**).
3. If a new scanner is being added to the system, run the following command to create barcodes for the scanners to be added. The scanner names can be customized:
    a. ns_image_server_barcodes –c mybarcodes.tif scanner1 scanner2
4. To complete adding a new scanner, transfer the mybarcodes.tif file to the NAS-mounted directory and print the barcodes by accessing the NAS with a computer with printer access. Cut the barcodes out and tape them in the marked location on the glass sheets. Place the barcoded glass sheets on each scanner so that the barcode is facing down and located toward the front left side of the scanner. The glass sheet should be snug against the back right corner of the scanning bed.
5. Run the following commands to run a terminal multiplexer, start the server, and then exit the multiplexed terminal so that it will keep running once the SSH connection is terminated:
    a. screen
    b. sudo ns_image_server # text will be displayed and end with "Dispatcher started" once the server is connected and operational.
    c. press Ctrl + A
    d. press D
6. Close terminal windows and check that the system is running by going to the website interface and clicking on "Capture Devices and Image Servers." The unique hostname for the DBserver should be

7. present and in bold font.
   Ensure that the scanner is on, has the barcoded glass sheet flush against the back right corner, and is connected to the imager server RasPi. Reboot and establish an SSH connection. Ensure that the NAS is correctly mounted (*see* **Note 12**).

8. Repeat Steps 5 and 6 to start the imaging server. Once started, the imaging server will run a scan to check for barcodes. Connected scanners should be on the right side under "Image Capture Devices" after clicking on "Capture Devices and Image Servers" on the main website.

9. Repeat Steps 7 and 8 for any additional scanners.

10. On the analysis server PC, go to the "This PC" directory and ensure that the NAS is mapped as a network drive. Go to the "C:\server" directory and double-click the ns_worm_browser_x64.exe file to start the analysis server. The analysis server must be active to run any analysis jobs added to the system (*see* **Note 13**).

3.3 Implementing the Lifespan Machine

Once the system is operational, plates containing worms treated with FUdR can be loaded onto the machine to set up an experiment. Experiments should be monitored daily for possible contamination or drying issues, as well as to monitor when all worms on the plates are dead. Plates with issues should be censored from downstream analysis to prevent stalling of the experimental workflow (*see* **Note 14**). As the plates dry, cracks and agar crystals may appear on the plate and can be confused as worms by the software, which will cause stalling in the analysis (*see* **Note 15**). We recommend that an initial control lifespan of all N2 worms be performed to establish a baseline for variability in lifespans (Fig. 5a). The control lifespan will also result in large differences in lifespan if the cooling is inadequate (Fig. 5b). Due to the increased temporal resolution of the system compared to manual lifespans, differences of less than a day may be statistically significant (Table 2).

1. Clean the glass sheet and plate holder using 70% ethanol. With the glass sheet barcode facing down, use t the plate holder to the top. Ensure that the plates will be in the imaging regions of the plate.

2. Add all the plates to the glass sheet by removing the lids and inverting on to the glass plate. Write down th all the plates, using the barcode location as a reference.

3. Place a second glass sheet on top of the plates so they are sandwiched between two glass sheets. This will all the plates are fully seated and allow for easier transfer to the scanner.

4. Move the plate assembly to the scanner with the barcode located to the lower-left corner when looking at from the front. Adjust the assembly so that it is snug in the upper right corner of the scanning bed. Close a TPU so that it is an even distance from all the plates.

5. Using the computer set up with the Worm Browser, open the lifespan machine web interface, click on "Ca and Image Servers," and select the checkbox of the scanner with loaded plates. This will capture a picture scanner surface.

6. Inspect the image located in the "\partition_000\preview_captures\" subdirectory of the NAS raw data fo plates are clear and in focus, continue to the next step. If there is condensation, wait and repeat Step 5 unt condensation is gone, or remove the plate assembly to clean the condensation and replace it before repea

7. Copy the preview image into a local directory on the computer and then open the Worm Browser progran clicking the ns_worm_browser_x64.exe file.

8. Go to File > Open Image and select the preview image previously copied.

9. Define the regions to be scanned for each of the four columns of plates by clicking once on the upper-left r column and once on the lower-right region of the column. The regions do not need to be accurate, as they
10. adjusted. Four dashed-line rectangles will be superimposed on the image.
Adjust the boxes by clicking and dragging the corners of each rectangle so that the rectangles are the sma be while containing the areas where worms may be present on the plates.

11. Save the specific coordinates to the computer by going to Image Acquisition > Define Scan Areas > Save Se Areas to Disk.

12. Download the "experiment_example.xml" file from https://github.com/nstroustrup/lifespan/tree/flow/c

13. Configure the experiment by opening the file with a text editor and replacing the scanning coordinates wi in Step 11 (coordinates will begin with "<sample>" and end with "</sample>"), changing the experiment changing the scanner to the correct one ("<sample><device>" should be followed by the name of the scan the experiment duration in days after "<duration>" ("30d" by default), setting the duration between each in minutes after "<device_capture_period>" ("15 m" by default), and then saving the file with a name to id experiment.

14. On the Worm Browser program, begin the experiment by going to Image Acquisition > Submit Experimen and selecting the experiment .xml file from Step 13. Clicking "Run It" will start the experiment, and clickir Summary File" will create a.txt file showing the plate coordinates and the date/time of all the images to b

15. After a day or two, enough images will have been collected to generate a Mask Composite to specify plate regions. Open the Worm Browser program and go to File > Select Experiment > "your specific experiment. Image Acquisition > Define Sample Masks > Generate Experiment Mask Composite and save the files to th

16. Make copies of the file and store them in a subdirectory, and open the .tif mask file with ImageJ.

17. Double-click the color-picker tool and change the fill color to white (255) and the background color to bla

18. Using the oval selection tool, select the area containing the worms for one plate and fill the selection. Rep plates.

19. To remove the background, select the Magic Wand tool and hold shift while left-clicking on all the plate re the selection and press Delete.

20. Assign the column plate order by changing the fill color to a dark gray color and then selecting the Magic While holding shift, left-click on all the top plate regions. Fill the regions with the dark gray color.

21. Change the fill color to a lighter gray color and repeat Step 20 for each row of plates, change the color to a color for each row. Save the file as a.tif keeping the original filename and location.

22. Load the sample mask to the experiment by opening the Worm Browser and going to File > Select Experir specific experiment." Then go to Image Acquisition > Define Sample Masks > Analyze Plate Locations Draw Experiment Mask Composite, then select the edited .tif mask file (the .xmp file that was originally saved w must be in the same directory). The program will analyze the file and show a version of the edited mask fi Image Acquisition > Define Sample Masks > Submit Analyzed Experiment Mask Composite to Cluster.

23. Use the submitted mask to generate per-plate images. Open the web interface and click "Run Analysis" for experiment. Under Experiment-wide operations > Sample Jobs, click "New Job For all Samples." Under Sch for Individual Images, check the box for "Apply Mask" and click the button "Save Job." The analysis server applying the mask and breaking column images down to individual plates.

24. Add strain and condition information to the experiment. Open the website and click “By Position” for the
experiment. Add corresponding data for each plate (column D is closest to the scanner barcode on the gla
the bottom of the web page, click the correct scanner for the plates, then click the “Save to Selected Devic
25. save the annotations.
Add worm time zero information to the experiment. Go to the website and click “Run Analysis” for the cor
experiment. Under Experiment-wide Operations > Region Jobs, click “New Job for All Regions.” In the box
Region Information, check the box “Time at which animals had 0 Age” and add the time corresponding to
the experiment (the website interface shows the information under “Durations” with the correct format).
button “Update Selected Fields.”
26.
If this is the first analysis, add the necessary worm detection and posture models to the system. Downloa
directories from https://github.com/nstroustrup/lifespan/tree/flow/files. Place them in the NAS shared
folder/models.
27.

Configure Analysis settings. Go to the website and click “Run Analysis” for the correct experiment. Under
wide operations > Configuration, click “Configure Machine Analysis.” Under Posture Analysis Parameter S
desired model and click the button “Set Posture Analysis Models.” Under Worm Detection Parameter Sets,
desired model and click the button “Set Worm Detection Models.” Under Image Analysis Options, change a
options and change the “Maximum Number of Worms per Plate” to 25% greater than the number of actua
necessary. Leaving the default of 100 is okay for 75 or fewer animals.
28.
Begin detection of worms for each plate image. Go to the website and click “Run Analysis” for the correct
Under Experiment-wide operations > Regions Jobs, click “New Job for All Regions.” Under Schedule a Job f
Images, check “Median Filter,” “Threshold,” and “Worm Detections” (these are needed to generate lifespan
the button “Save Job” under the previously selected check boxes. Analysis can be started before all images
experiment are captured.
29.
Movement analysis should be performed once all the worms are dead. To check if the worms are dead bef
experiment set time has finished, go to the website and click on “View recently captured images.” Right-cl
image]” for the two images of the same plate column and open them in a new browser tab. Compare the t
determine if there is any worm movement. Perform the same comparison for the remaining columns. If th
movement, the experiment can be stopped by opening the website and clicking on “Cancel Pending Scans
correct experiment and then clicking on “Cancel Scheduled Captures” in the new window.
30.
Analyze movement once all the worms are dead. Go to the website and click “Run Analysis” for the correc
Under Experiment-wide operations > Regions Jobs, click “New Job for All Regions.” Under “Schedule a Job
Region,” select “Analyze Worm Movement” in the dropdown menu and click “Save Job.” To check the analy
on to the website and click on “Analysis Status” for the correct experiment. The samples will move along t
and will be complete when all samples are in the Validation column (*see* **Note 16**).
31.
Create a storyboard to validate worms. Go to the website and click “Run Analysis” for the correct experim
Experiment-wide operations > Experiment Jobs, click “[New Experiment Job].” Under Schedule a Job For a
Region, select “Generate Animal Storyboard” in the dropdown menu and click “Save Job.”
32. Validate Automated Worm Detection through Storyboard. Open the Worm Browser program and go to Fil
Current Experiment and select the desired experiment. Go to Validation > Browse Entire Experiment > Im
After Each Worm’s Death. The Worm Browser window will then display a storyboard of the worm image,
additional window will appear displaying the lifespan curves. The buttons at the bottom can be used to ch
displayed data. The images of worms on the storyboard window will be annotated with the following: Wh
Box/Outline = Excluded, Blue upper left Corner = Worm was previously left-clicked, Red bottom right corn
bottom right corner = Worm is stationary, but not long enough to be declared dead, Number in upper righ
Marked as multiple dead, and Red/Green images = Two worms (potentially the same worm) marked as be
proximity of each other. Green is the first death, and red is the second death. The following controls are us

annotate the images (click save at the bottom to save the annotations as you go):

a. Left-Clicking: will pop up the worm inspection window displaying that worm's data (Fig. 6). The left ar keyboard keys can be used to scroll through the images of that worm. The black vertical line across bo the image denotes the timepoint. Alternatively, clicking on a timepoint on either of the plots on the rig the upper annotation bar under the image will bring up the image at that timepoint. The bar colors sig following: Purple = fast-moving (alive), Yellow = changing posture (alive), Red = dead.

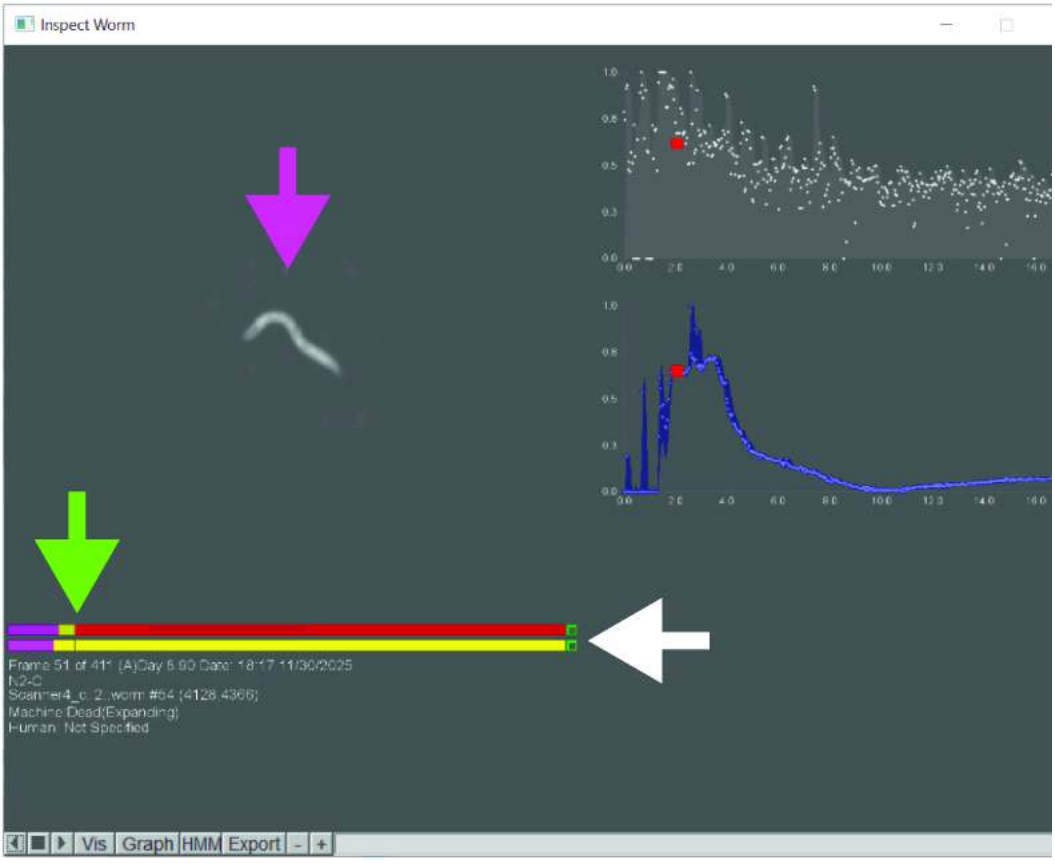


Fig. 6 Screenshot of Worm Browser worm inspection window. The magenta arrow denotes the worm image. The white arrow points to the annotation bar and lower manual annotation bar. The green arrow points to the vertical line denoting the time point in the lifespan corresponding being displayed

b. Right-Clicking: censor the worm and mark it with a white box/outline.

c.

Hold Shift and left-click: Notes multiple worms.

33. To manually annotate the death of the worm, left-click on the worm to bring up the worm inspection win

    a. Move to a timepoint when the worm is fast-moving in the upper automated annotation bar by using th instructions in Step 32. The vertical line should be in the purple area of the upper bar.

    b. Left-click on the lower bar to the right of the vertical line until the entire bar becomes purple.

    c. Move to a timepoint where the worm is no longer fast-moving but only changing its posture, and left-c lower bar to the right side of the vertical line. The right side will turn yellow.

    d. Move to a timepoint where the worm is dead and left-click on the lower bar to the right side of the ver right side will turn red.

    e. Close the timepoint window and save your annotation.

34. After validation is complete, generate survival data. Open The Worm Browser program and go to File > Se Experiment and select the desired experiment. Go to Data Files > Death Times > Generate Death Times for

    Experiment. The program will generate files in the Results directory in the data folder. The files will be in with the name of the experiment, in a subdirectory labeled "survival_data."

35. The file corresponding to the automated detection will be in the RawDataFolder/Results/experiment/survival_data/survival_simple/experiment=survival=machine_jmp_ The manually edited times of death can be found at RawDataFolder/Results/experiment/survival_data/survival_simple/experiment=survival=machine_hanc file. Any manually annotated worm death times will show "Storyboard" in the "Death Annotation Source"

36. The results can then be analyzed and input into your desired program for plotting and statistical analysis. custom Excel macro to exclude censored animals and convert the automated time of death data to a Grapl survival analysis format.

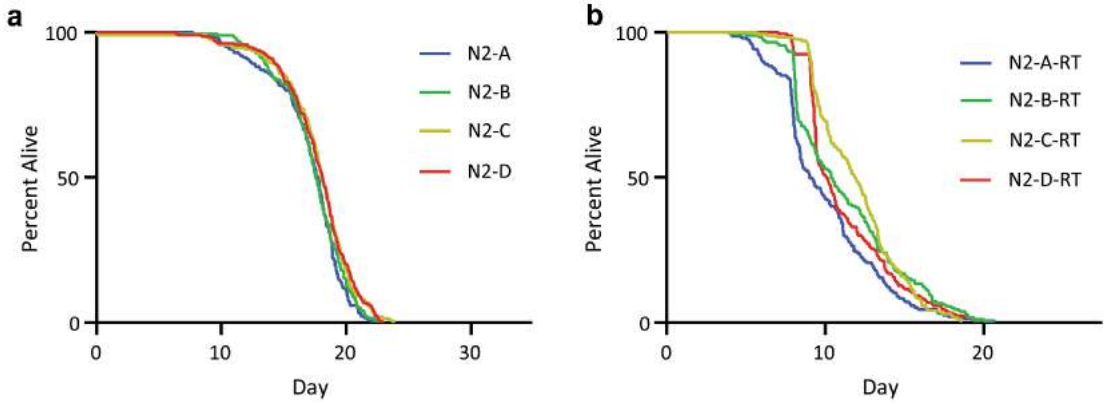


Fig. 5 Control lifespans conducted on the Lifespan Machine. **a**, Lifespan of only N2 worms grouped by plate column. **b**, Lifespan of only N2 worms grouped by plate column without proper cooling

Table 2 Control all N2 automated lifespan summary and statistics

4 Notes

1. The screw covers can be easily removed by screwing in a small, pointed screw into the cover and then using pliers to pull on the screw.
2. Wearing gloves will help prevent placing smudges on the glass lid and touching any of the sensitive internal components with bare skin.
3. Installing the clip may be difficult. One strategy to install the clip is to hold the clip with long needle-nose pliers at the end that was pressed to release it. The clip can then be held in position under its bracket, making sure the belt is between the clip and bracket. Using a finger from the other hand, press the clip up into place. Care should be taken not to damage the belt with the pliers.
4. To delete experiments, open the web browser to the experiment index page and click on the "[Edit]" link located to the far right for the experiment to be deleted. Click on the "Delete Experiment Data" button and confirm the deletion. The analysis server must be online for the deletion to occur from the system. Once the experiment is no longer visible on the web browser's experimental index page, the files can be deleted from the NAS.
5. Using an SSD is recommended for the DB server; however, crashes and corruption can still occur. Regular image backups of the database server are highly recommended to prevent loss of the database. Imaging servers can also be backed up but can easily be set up again since they do not store data long-term. Running the following command will create a system image backup of the DB server onto the NAS. The "/dev/mmcblk0" may have to be changed to whatever device is running your OS: sudo dd bs=4Mif=/dev/mmcblk0of=/mnt/NASmount/backup_directory/Date_DBserver.img.
6. All packages must be installed. If an error is displayed when installing the packages, some packages may not be installed, which can cause error messages later. In such instances, try reinstalling any packages if the error states that the package cannot be found.
7. If the NAS IP address has not been reserved yet, continue, but skip Step 14. Once the IP address has been reserved, SSH into the DBserver computer. Run the following command and perform Steps 13 and 14: sudo nano /etc/fstab.
8. If the reserved IP addresses have not been set up yet, continue without entering them. Once the IP addresses have been reserved, SSH into the DBserver computer. Run the following command and input the IP address information: sudo nano /usr/local/etc/ns_image_server.ini.
9. Tape handles can be attached to the glass sheet to aid in removal of the sheet after an experiment. Simply attach pieces of tape to the edge of the glass on both sides and allow for the overhanging parts to adhere to each other (Fig. 4b).
10. If the plates are left out too long to dry or seeded too early, the bacterial lawn may grow dense and affect the Lifespan Machine imaging quality, so growing cultures at the same time as expanding provides a good time frame to have fresh plates for the Lifespan Machine.
11. Adding FUdR too early may impact development. Adding FUdR too late will allow the development of early progeny. While the progeny will not develop into adults, an excess of progeny can potentially eat enough *E. coli* to cause the animals to starve before the experiment has completed.
12. If properly mounted, the files in the NAS raw data directory should be visible through the RasPi. The following command will list the file contents of the NASmount directory: sudo ls /mnt/NASmount.
13. We recommend changing the PC settings to prevent it from sleeping or hibernating, so any jobs added to the queue are executed as they are added.

to the queue are executed as they are added.

14. To censor plates, open the web browser to the experiment index page. For the desired experiment, click on the "[By Image]" link located under the "Annotate Plate Information" column. Look for the plate to be censored and check the box under the "Censor" column. Click the "Save" button to the right to save the annotation.

15. To censor specific plate images or all images after a specific point, open the web browser to the experiment index page and click on the "[By Image]" link located under the "Annotate Plate Information" column for the desired experiment. Look for the plate to be censored and click on the "[View images]" link under the plate name. Use the links under the "Resized Small Image" to find the desired timepoint and then click on the "[Censor_After]" or the "[Censor_This'" link to censor as desired.

16. Analysis may stall if the system detects more worms than were set in Step 3.3.27. Increasing the number of worms and restarting the analysis will allow the analysis to progress in such cases.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. López-Otín C, Blasco MA, Partridge L et al (2013) The Hallmarks of aging. Cell 153:1194–1217. https://doi.org/10.1016/j.cell.2013.05.039 [PubMed][PubMedCentral]

2. Aunan JR, Watson MM, Hagland HR, Søreide K (2016) Molecular and biological hallmarks of ageing. Br J Surg 103:e29–e46. https://doi.org/10.1002/bjs.10053

3. Cohen AA (2018) Aging across the tree of life: the importance of a comparative perspective for the use of animal models in aging. Biochim Biophys Acta BBA - Mol Basis Dis 1864:2680–2689. https://doi.org/10.1016/j.bbadis.2017.05.028

4. Sutphin GL, Kaeberlein M (2009) Measuring *Caenorhabditis elegans* life span on solid media. J Vis Exp JoVE 1152. https://doi.org/10.3791/1152

5. Johnson TE (2003) Advantages and disadvantages of *Caenorhabditis elegans* for aging research. Exp Gerontol 38:1329–1332. https://doi.org/10.1016/j.exger.2003.10.020 [PubMed]

6. Van Voorhies WA, Ward S (1999) Genetic and environmental conditions that increase longevity in *Caenorhabditis elegans* decrease metabolic rate. Proc Natl Acad Sci 96:11399–11403. https://doi.org/10.1073/pnas.96.20.11399 [PubMed][PubMedCentral]

7. Uno M, Nishida E (2016) Lifespan-regulating genes in *C. elegans*. NPJ Aging Mech Dis 2:16010. https://doi.org/10.1038/npjamd.2016.10 [PubMed][PubMedCentral]

8. Klass MR (1977) Aging in the nematode *Caenorhabditis elegans*: major biological and environmental factors influencing life span. Mech Ageing Dev 6:413–429. https://doi.org/10.1016/0047-6374(77)90043-4 [PubMed]

9. Cornwell AB, Samuelson AV (2020) Analysis of lifespan in *C. elegans*: low- and high-throughput approaches. In: Curran SP (ed) Aging: methods and protocols. Springer US, New York, NY, pp 7–27

10. Mendenhall AR, Wu D, Park S-K et al (2011) Genetic dissection of late-life fertility in *Caenorhabditis elegans*. J Gerontol A Biol Sci Med Sci 66A:842–854. https://doi.org/10.1093/gerona/glr089

11. Garcia G, Homentcovschi S, Kelet N, Higuchi-Sanabria R (2022) Imaging of actin cytoskeletal integrity during aging in *C. elegans*. Methods Mol Biol Clifton NJ 2364:101–137. https://doi.org/10.1007/978-1-0716-1661-1_5

12. Hsu A-L, Feng Z, Hsieh M-Y, Xu XZS (2009) Identification by machine vision of the rate of motor activity decline as a lifespan predictor in *C. elegans*. Neurobiol Aging 30:1498–1503. https://doi.org/10.1016/j.neurobiolaging.2007.12.007 [PubMed]

13. Mathew MD, Mathew ND, Ebert PR (2012) WormScan: a technique for high-throughput phenotypic analysis of *Caenorhabditis elegans*. PLoS One 7:e33483. https://doi.org/10.1371/journal.pone.0033483 [PubMed][PubMedCentral]

14. Stroustrup N, Ulmschneider BE, Nash ZM et al (2013) The *C. elegans* lifespan machine. Nat Methods 10:665–670. https://doi.org/10.1038/nmeth.2475

[PubMed][PubMedCentral]

15. Puchalt JC, Sánchez-Salmerón A-J, Ivorra E et al (2021) Small flexible automated system for monitoring *Caenorhabditis elegans* lifespan based on active vision and image processing techniques. Sci Rep 11:12289. https://doi.org/10.1038/s41598-021-91898-6
[PubMed][PubMedCentral]

16. Rahman M, Edwards H, Birze N et al (2020) NemaLife chip: a micropillar-based microfluidic culture device optimized for aging studies in crawling *C. elegans*. Sci Rep 10: 16190. 10.1038/s41598-020-73002-6

17. Churgin MA, Jung S-K, Yu C-C et al (2017) Longitudinal imaging of *Caenorhabditis elegans* in a microfabricated device reveals variation in behavioral decline during aging. eLife 6:e26652. https://doi.org/10.7554/eLife.26652
[PubMed][PubMedCentral]

18. Rahman M, Edwards H, Birze N et al (2020) NemaLife chip: a micropillar-based microfluidic culture device optimized for aging studies in crawling *C. elegans*. Sci Rep 10:16190. https://doi.org/10.1038/s41598-020-73002-6
[PubMed][PubMedCentral]

19. Puchalt JC, Sánchez-Salmerón A-J, Guerola PM, Martínez SG (2019) Active backlight for automating visual monitoring: an analysis of a lighting control technique for *Caenorhabditis elegans* cultured on standard Petri plates. PLoS One 14:e0215548. https://doi.org/10.1371/journal.pone.0215548
[PubMed][PubMedCentral]

20. Pitt JN, Strait NL, Vayndorf EM et al (2019) WormBot, an open-source robotics platform for survival and behavior analysis in *C. elegans*. GeroScience 41:961–973. https://doi.org/10.1007/s11357-019-00124-9
[PubMed][PubMedCentral]

21. Abbott M, Banse SA, Melentijevic I et al (2020) A simplified design for the *C. elegans* lifespan machine. J Biol Methods 7:e137. https://doi.org/10.14440/jbm.2020.332
[PubMed][PubMedCentral]

22. Timmons L, Court DL, Fire A (2001) Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. Gene 263:103–112. https://doi.org/10.1016/s0378-1119(00)00579-5
[PubMed]

23. Stuhr NL, Curran SP (2020) Bacterial diets differentially alter lifespan and healthspan trajectories in *C. elegans*. Commun Biol 3:653. https://doi.org/10.1038/s42003-020-01379-1
[PubMed][PubMedCentral]

24. Reinke SN, Hu X, Sykes BD, Lemire BD (2010) *Caenorhabditis elegans* diet significantly affects metabolic profile, mitochondrial DNA levels, lifespan and brood size. Mol Genet Metab 100:274–282. https://doi.org/10.1016/j.ymgme.2010.03.013
[PubMed]

25. Revtovich AV, Lee R, Kirienko NV (2019) Interplay between mitochondria and diet mediates pathogen and stress resistance in *Caenorhabditis elegans*. PLoS Genet 15:e1008011. https://doi.org/10.1371/journal.pgen.1008011
[PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_7

7. Assessing Age-Related Changes in *Caenorhabditis elegans* Behavior Using the Forced Food Choice Assay

Nicole L. Stuhr[1] and Sean P. Curran[1] ✉
(1) Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA

✉ **Sean P. Curran**
Email: spcurran@usc.edu

Abstract
The forced food choice assay enables the examination of *Caenorhabditis elegans* behavior over time and across life stages. Different from existing food choice assays, it requires *C. elegans* to move through one food source before leaving and encountering the second food source. The forced food choice assay provides great resolution on behavioral changes associated with aging and in the context of genetic mutation with or without dietary or pharmacologic manipulation. Here, we describe a method that removes the opportunity for "chance" decisions, allowing for the measurement of active choices to choose one food source over another.

Keywords *Caenorhabditis elegans*, – Physiology – Bacterial diet – Food choice – Forced choice – Behavior – Foraging behavior

1 Introduction

The lifespan and healthspan of *Caenorhabditis elegans* are influenced by the bacterial diet ingested [1–3]. In the wild, *C. elegans* encounters a variety of bacterial food sources in rotting fruits and vegetation [4–7]. *C. elegans* exhibits sophisticated food choice behaviors, often demonstrating preference towards beneficial and nutritious bacterial diets, including *Proteobacteria*, *Sphingomonas*, and *Bacillus*, while learning to avoid detrimental bacterial diets and pathogens that lead to accelerated aging and infection [1, 8, 9].

Although the investigation of molecular determinants of behavioral traits has been extensively studied [10–14], the analysis is often restricted to a single timepoint that may not provide adequate resolution to measure more subtle behavioral changes across time and across life stages. Traditional food choice assays typically use an open-field model where patches of distinct foods are placed on opposing sides of the environment, and animals are placed equidistant from both diets and are allowed to move towards the preferred food source [15]. These experiments are informative but can lead to chance encounters with a food source, some animals fail to encounter any of the bacterial options on the plate, and it is difficult to discern the previous encounters, if any, of animals found in between food sources [15–17].

Here, we describe an alternative assay called a Forced Food Choice Assay that requires animals to move through one food source before encountering the next. This method takes advantage of the fact that *C. elegans* prioritizes environmental exploration before switching to a foraging strategy during subsequent encounters with bacterial food sources [18–20]. The Forced Food Choice Assay allows for the analysis of age-related changes in behavior by introducing *C. elegans* initially to one bacterial diet before being able to encounter another bacterial diet option (Fig. 1). The Forced Food Choice Assay allows for the examination of the change in food preference over time and across aging populations, all on a single plate [21, 22].

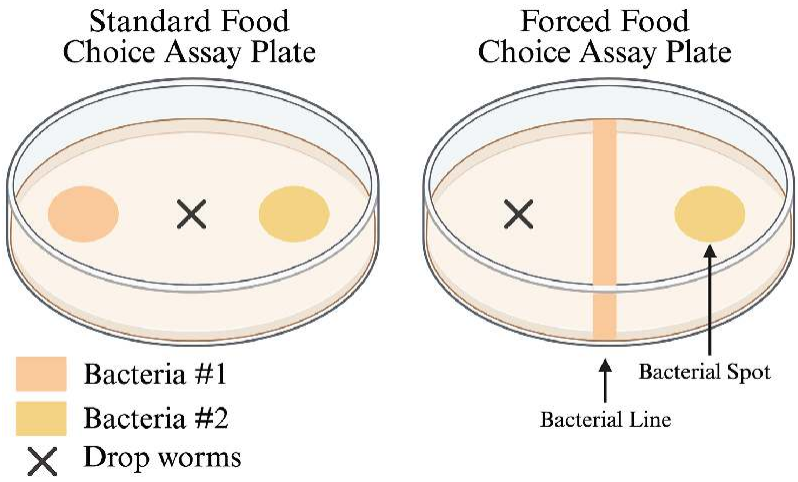


Fig. 1 Graphical schematic of key differences between the standard food choice assay plate and the forced food choice assay plate. Both assays compare *C. elegans* preference between two bacterial food sources. The key difference is that the Forced Food Choice Assay Plate has a bacterial line in the center of the plate that worms must crawl through first before having access to the other bacteria

2 Materials

Prepare all solutions using ultrapure water (prepared by purifying deionized water to attain a sensitivity of 18 MΩ-cm at 25 °C). Prepare and store all reagents at room temperature unless otherwise specified. All plate preparations should be done near a flame and in a sterile environment. Diligently follow all waste disposal regulations when disposing of waste materials.

2.1 Preparation of Forced Food Choice Plates

1. Bacterial cultures: Grow bacteria overnight in the correct growth media and, as appropriate for the bacterial strain being tested, provide antibiotic selection. Bacteria must be seeded fresh and at the same optical density (*see* **Note 1**).
2. Unseeded 6 cm nematode growth media (NGM) plates.
3. Ruler: To define the center of the plate (for the line of food source #1), measure the distance to seed the bacterial spot (for food source #2), and mark where to place worms at the start of the experiment (time = 0) (Fig. 1).

experiment (time = 0) (Fig. 1).

4. 9-inch glass Pasteur pipette: The glass Pasteur pipette will be used for "stamping" the line of bacteria onto the plates.
5. Reagent reservoir.

2.2 Forced Food Choice Assay Supplies

1. M9 buffer: Prepare 10× M9 master mix in a 4 L glass beaker with a stir bar on a hot plate. Fill to 1 L with ultrapure H_2O and add the following salts: 60 g KH_2PO_4, 120 g Na_2HPO_4, 100 g NaCl, and 40 mL 1 M $MgSO_4$. Let the salts dissolve while stirring for about 25 min before adjusting the final volume to 2 L with ultrapure H2O. For 1× stock, add 50 mL of the 10× stock into a 500 mL glass bottle designed for autoclaving and then add 450 mL of ultrapure H_2O. Autoclave on a liquid cycle for 30 min and allow to cool completely before using.
2. M9T buffer: 1× M9 buffer with 0.01% Triton-X added. For a 500 mL bottle of M9, add 50 µL of Triton-X.
3. 1.5 mL Eppendorf tube.
4. 5-Fluorodeoxyuridine (FUDR).

---

3 Methods

Carry out all procedures at room temperature unless otherwise specified.

3.1 Production of Forced Food Choice Plates

1. Two days before the assay, inoculate bacterial cultures and grow overnight in growth media, selection antibiotics (*see* **Notes 2** and **3**), and correct growth temperature(s).

2. The next day, mark lines and dots where the bacterial line, bacterial spot, and worm placement locations will be (*see* Fig. 2). For a 6 cm plate, the bacterial line is placed in the center of the plate, and the bacterial spot and worm placement point will both be 2 cm from the center of the plate.

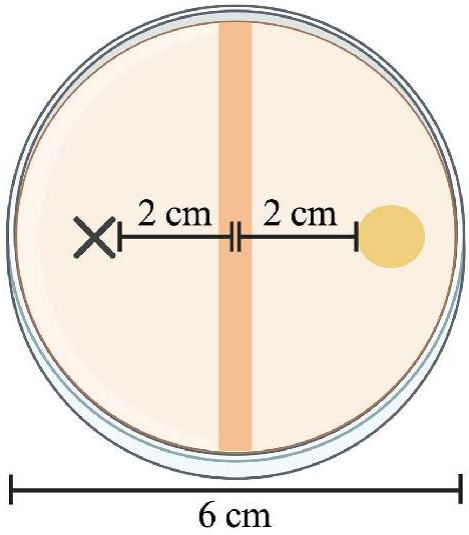


Fig. 2 How to set up the forced food choice assay plates. The bacterial line is in the center of the plate and the bacterial spot and the worm placement point are exactly 2 cm away from the bacterial line

3. Check the optical density of the bacterial cultures to make sure bacteria are being seeded at the same optical density (reminder, *see* **Note 1**).
4. Flame the nine-inch glass Pasteur pipette tip to use for "stamping" the bacterial line onto the plate. To allow for extra room for the bend of the pipette, make sure to flame at 3.75 cm away from the end of the pipette (Fig. 3). Maintain the stamp upright, near the flame, until ready to use to maintain sterility.

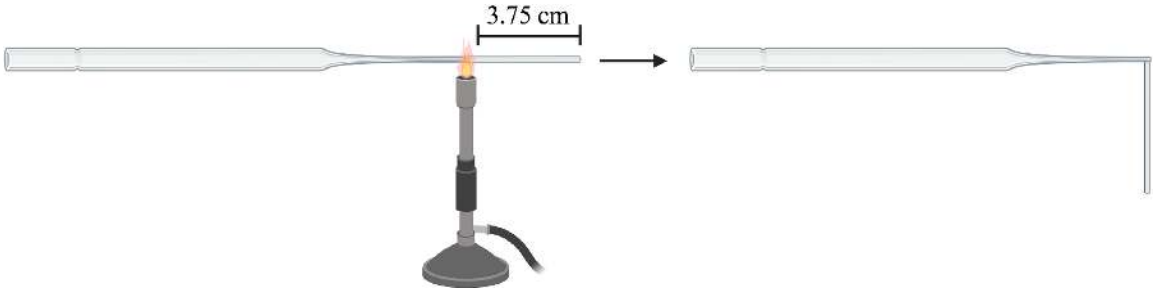


Fig. 3 The location of the nine-inch glass Pasteur pipette must be flamed (3.75 cm away from the tip of the glass pipette) to create a stamp to use for seeding the bacterial line onto the Forced Food Choice Assay Plates

5. Next to a flame, seed the bacteria line onto the forced food choice plates by stamping the bent Pasteur pipette stamping tool into the reagent reservoir with bacterial culture and removing excess culture (*see* **Note 4**) before stamping onto the nematode growth media plate where the drawn line was made (*see* **Notes 5** and **6**).
6. Seed the bacterial spot using 20 µL of bacterial culture (*see* **Note 7**) on the dot marked on the plate 2 cm from the center.
7. Dry plates in a hood completely before placing lids on. Leave plates at room temperature to grow overnight before the assay is performed the next day.

3.2 Forced Food Choice Assay

1. For each forced food choice assay, set up enough plates to have at least three technical replicates per condition. Each forced food choice assay needs 30–40 synchronized worms per plate for a total of ~100 worms per condition (*see* **Note 8**). Place worms from an overnight bleach synchronization (for L1 assays) or use M9 buffer (*see* **Note 10**) to wash worms off a growth plate used to develop a population to the correct stage or mature to a specific age (*see* **Note 9**).
2. Dropping worms onto the plate can either be done from an

overnight bleach synchronization or washing worms off the plates for testing older worms. For overnight bleach synchronization: Count worms that are present in 1 µL to get an average number of worms per microliter (*see* **Note 11**). Place worms in the overnight synchronization buffer (usually M9) onto the worm placement "X" marked on the plate. Later stages of worms require washing the worms using M9 into a 1.5 mL Eppendorf tube. Make sure to transfer the worms from the plate to the tube using an M9T-coated pipette tip so the worms do not stick to the plastic (*see* **Note 12**). For later-stage animals, it is best to cut the pipette tip to allow for a wider opening (*see* **Note 13**). Allow worms to gravity settle to the bottom of the 1.5 mL Eppendorf tube before removing excess supernatant to concentrate the worms before averaging the number of worms in 1 µL (reminder *see* **Note 11**). Place synchronized worms onto the worm placement "X" marked on the plate using an M9T-coated pipette tip.

3. Count the number of worms on each plate on the worm placement "X".

4. Keep plates open and near a flame to allow the M9 to dissipate. Once dry, close the plates and place them in the incubator until it is time to assess worm location. This timing varies depending on the assay being performed (*see* **Note 14**). For L1 to young adult assays, standard timepoints are carried out at 0, 24, 48, and 72 h. For assays starting with L4s or adults, it is best to check at least every 2 h.

5. Check assay plates at each timepoint and note the worms that are on Bacteria #1, Bacteria #2, and total not on either bacterial lawn.

3.3 Calculating the Food Choice Index

1. Once the assay is complete and all timepoints have been assessed, the food choice index can be calculated by dividing the number of

worms at each timepoint on each bacterial diet by the total number of worms on the plate (Fig. 4). Food Choice Index = [(Number of worms on Bacteria #1)−(Number of worms on Bacteria #2)]/Total Number of Worms on the Plate. If there are more worms on Bacteria #1, the food choice index will be positive. If there are more worms on Bacteria #2, the food choice index will be negative.

$$\text{Food Choice Index} = \frac{\text{(Number of Worms on Bacteria \#1)-(Number of Worms on Bacteria \#2)}}{\text{Total Number of Worms on Plate}}$$

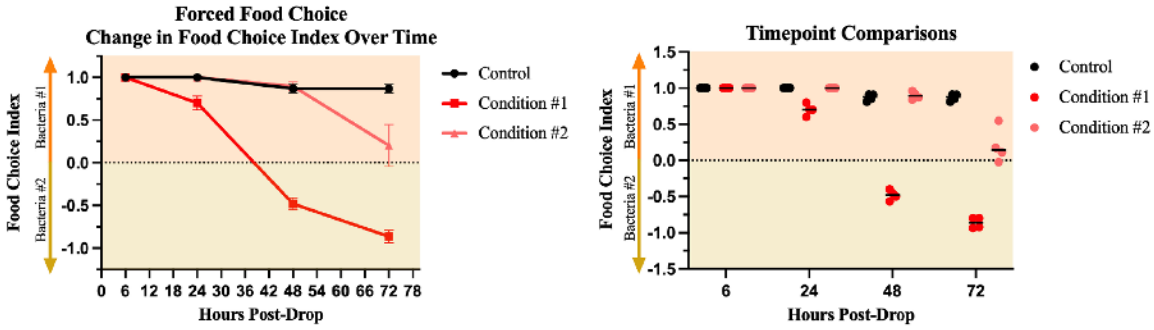


Fig. 4 Formula for calculating the Food Choice Index and representative data demonstrating how food choice indices can be graphed. The change in food choice index over time is informative on how different conditions may influence behavior, while timepoint comparisons can be made to show significant changes compared to control conditions. In this example, Control worms stayed on Bacteria #1 across the 72 h the assay was performed, while Condition #2 led to modest movement of worms from Bacteria #1 to Bacteria #2 (with more worms remaining on Bacteria #1 by the 72 h). Condition #2 shows worms that quickly encountered Bacteria #1 and moved on toward Bacteria #2 with more than 50% of the worms on Bacteria #2 at the 48 h timepoint and almost all worms on Bacteria #2 by the 72 h endpoint of the assay. (*N* = 3 technical replicates per condition, bars represent ± standard deviation)

2. There are multiple ways in which this data can be shown, but change in food choice over time is common. In addition, most people will also be interested in the data showing pairwise comparisons at each time point (Fig. 4).

---

4 Notes

1. Always check the OD of the bacteria before seeding. The bacterial culture will grow overnight after seeding onto the plate, so

normalizing the optical densities will allow bacterial cultures to be as close as possible in terms of growth conditions.

2. Selection antibiotics are very important in this step because if the bacteria used for the assay are resistant to different antibiotics, the nematode growth media plates used for the assay will not contain antibiotics to allow both seeded bacterial cultures to grow overnight.

3. The volume of bacterial culture is dependent upon the number of assay plates being prepared. In general, it is always a good idea to inoculate 10 mL because this will be more than enough culture for about 100 forced food choice assay plates.

4. It is important to make sure excess bacterial culture drips off the end of the bent Pasteur pipette before bringing it over to the plate, because excess culture will make one side of the bacterial line wider than the other.

5. Be careful when lifting the bent Pasteur pipette from the plate. Sometimes it will stick to the plate and cause the bacterial culture to splash, leaving droplets of culture on the areas surrounding the bacterial line.

6. When stamping the bacterial line, it is possible the Pasteur pipette does not cover the entire diameter of the plate, slowly drag the pipette in a line to the other edge of the plate to make sure the bacterial line spans the entire diameter.

7. The location of the bacterial spot can vary depending on the assay being set up. A general rule is to place the spot no further than 2 cm from the center of the plate. The reason for this is that the closer the spot is seeded to the edge of the plate, the more likely the spot will move due to the sides of the plates being slightly higher than the center of the plate when plates are poured.

8. Always prepare more than double the number of worms needed

for the assay due to loss of sample during aging, washing, and transferring of worms to the assay plates.

9. This assay can be performed on any stage of animals. Most assays are either from L1 to young adults (before egg laying occurs), but can be done from L4 or other adult stages if egg laying is taken into consideration. However, keep in mind that movement and muscle function decrease with age, so older animals may not perform as well as younger animals.

10. If using L4 or adult worms and wanting to look across multiple days, it is required to add FUDR to the top of the assay plates and dry before adding the worms. This makes sure that progeny will not hatch during the assay and disrupt the results. If only examining change in behavior across 8 h, this will not be necessary.

11. Concentrate worms to about ten worms per microliter to limit the volume being placed onto the assay plate, so the drying time will be shortened. Make sure that the volume used is between 5 and 10 µL and the volume is consistent across all conditions.

12. Coating the pipette tip is important to prevent loss of worms when transferring to the Eppendorf tube and when transferring to the assay plate. To coat the tip, just pipette M9T up and down a couple of times in the tip before disposing of the M9T buffer.

13. It is best to cut the pipette tip whenever working with adult worms. Cut the tip off before coating in M9T and pipetting up the worms.

14. This assay was developed to assess food choice behavior over multiple time points to look for change in food preference, which is why sodium azide is not placed on any of the bacterial cultures.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Kim DH (2013) Bacteria and the aging and longevity of *Caenorhabditis elegans*. Annu Rev Genet 47:233–246. https://doi.org/10.1146/annurev-genet-111212-133352 [PubMed]

2. Stuhr NL, Curran SP (2020) Bacterial diets differentially alter lifespan and healthspan trajectories in *C. elegans*. Commun Biol 3(1):653. https://doi.org/10.1038/s42003-020-01379-1 [PubMed][PubMedCentral]

3. Liu L, Hao X, Bai Y, Tian Y (2025) The soil mycobacterium sp. promotes health and longevity through different bacteria-derived molecules in *Caenorhabditis elegans*. Aging cell 24(3):e14416. https://doi.org/10.1111/acel.14416 [PubMed]

4. Samuel BS, Rowedder H, Braendle C, Felix MA, Ruvkun G (2016) *Caenorhabditis elegans* responses to bacteria from its natural habitats. Proc Natl Acad Sci U S A 113(27):E3941–E3949. https://doi.org/10.1073/pnas.1607183113 [PubMed][PubMedCentral]

5. Schulenburg H, Felix MA (2017) The natural biotic environment of *Caenorhabditis elegans*. Genetics 206(1):55–86. https://doi.org/10.1534/genetics.116.195511 [PubMed][PubMedCentral]

6. Avery L, Shtonda BB (2003) Food transport in the *C. elegans* pharynx. J Exp Biol 206(Pt 14):2441–2457. https://doi.org/10.1242/jeb.00433 [PubMed][PubMedCentral]

7. Worthy SE, Haynes L, Chambers M, Bethune D, Kan E, Chung K, Ota R, Taylor CJ, Glater EE (2018) Identification of attractive odorants released by preferred bacterial food found in the natural habitats of *C. elegans*. PLoS One 13(7):e0201158. https://doi.org/10.1371/journal.pone.0201158 [PubMed][PubMedCentral]

8. de Guzman ACV, Kang S, Kim EJ, Kim JH, Jang N, Cho JH, Choi SS (2022) High-glucose diet attenuates the dopaminergic neuronal function in *C. elegans*, leading to the acceleration of the aging process. ACS Omega 7(36):32339–32348. https://doi.org/10.1021/acsomega.2c03384 [PubMed][PubMedCentral]

9. Yen CA, Curran SP (2016) Gene-diet interactions and aging in *C. elegans*. Exp Gerontol 86:106–112. https://doi.org/10.1016/j.exger.2016.02.012

[PubMed][PubMedCentral]

10. Flavell SW, Raizen DM, You YJ (2020) Behavioral states. Genetics 216(2):315–332. https://doi.org/10.1534/genetics.120.303539
[PubMed][PubMedCentral]

11. Lee H, Boor SA, Hilbert ZA, Meisel JD, Park J, Wang Y, McKeown R, Fischer SEJ, Andersen EC, Kim DH (2024) Genetic variants that modify neuroendocrine gene expression and foraging behavior of *C. elegans*. Sci Adv 10(24):eadk9481. https://doi.org/10.1126/sciadv.adk9481
[PubMed][PubMedCentral]

12. Schafer WR (2005) Deciphering the neural and molecular mechanisms of *C. elegans* behavior. Curr Biol 15(17):R723–R729. https://doi.org/10.1016/j.cub.2005.08.020
[PubMed]

13. Kim DH, Flavell SW (2020) Host-microbe interactions and the behavior of *Caenorhabditis elegans*. J Neurogenet 34(3-4):500–509. https://doi.org/10.1080/01677063.2020.1802724
[PubMed][PubMedCentral]

14. Bargmann CI (1993) Genetic and cellular analysis of behavior in *C. elegans*. Annu Rev Neurosci 16:47–71. https://doi.org/10.1146/annurev.ne.16.030193.000403
[PubMed]

15. Shtonda BB, Avery L (2006) Dietary choice behavior in *Caenorhabditis elegans*. J Exp Biol 209(Pt 1):89–102. https://doi.org/10.1242/jeb.01955
[PubMed][PubMedCentral]

16. Kim AT, Li S, Kim Y, You YJ, Park Y (2024) Food preference-based screening method for identification of effectors of substance use disorders using *Caenorhabditis elegans*. Life Sci 345:122580. https://doi.org/10.1016/j.lfs.2024.122580
[PubMed]

17. Wang Y, Guo K, Wang Q, Zhong G, Zhang W, Jiang Y, Mao X, Li X, Huang Z (2024) Caenorhabditis elegans as an emerging model in food and nutrition research: importance of standardizing base diet. Crit Rev Food Sci Nutr 64(10):3167–3185. https://doi.org/10.1080/10408398.2022.2130875
[PubMed]

18. Haley JA, Chen T, Aoi M, Chalasani SH (2025) Accept-reject decision-making revealed via a quantitative and ethological study of *C. elegans* foraging. eLife 13:RP10391. https://doi.org/10.7554/eLife.103191

19. Katzen A, Chung HK, Harbaugh WT, Della Iacono C, Jackson N, Glater EE, Taylor CJ, Yu SK, Flavell SW, Glimcher PW, Andreoni J, Lockery SR (2023) The nematode worm *C. elegans* chooses between bacterial foods as if maximizing economic utility. eLife 12:69779. https://doi.org/10.7554/eLife.69779

20. Haley JA, Chalasani SH (2024) C. elegans foraging as a model for understanding the neuronal basis of decision-making. Cell Mol Life Sci 81(1):252. https://doi.org/10.1007/s00018-024-

05223-1
[PubMed][PubMedCentral]

21. Stuhr NL, Ramos CM, Turner CD, Soukas AA, Curran SP (2026) A lipid-sensitive food choice behavior influences aging outcomes from a longevity-promoting diet. FEBS J https://doi.org/10.1111/febs.70650

22. Nair T, Weathers BA, Stuhr NL, Nhan JD, Curran SP (2024) Serotonin deficiency from constitutive SKN-1 activation drives pathogen apathy. Nat Commun 15(1):8129. https://doi.org/10.1038/s41467-024-52233-5
[PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_8

8. Analyzing Spatial Learning Behavior and Cognitive Function in Aging Mice Using the Barnes Maze

Amy Christensen[1] ✉
(1) Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA

✉ **Amy Christensen**
Email: akchrist@usc.edu

Abstract
Testing cognitive functions in aging rodents presents unique challenges. Aged mice may have deficits in senses and movement that require careful consideration. Further, the translatability of rodent behavior to the human condition is difficult to assess. However, deficits in spatial memory are seen in both rodents and humans with increasing age. Thus, tests of spatial memory could be reasonably argued to be the most translationally relevant. The Barnes maze was developed to test spatial learning and memory and can be evaluated in ways that are easily adaptable to many of the issues that affect older mice. There is a wide variety of protocols used with variations of almost every step. A protocol is described here that has been adapted to the needs of aged mice.

Keywords Barnes maze – Mice – Spatial learning – Spatial memory – Aging

1 Introduction

Assessing the cognitive abilities of aging mice requires choosing tests that can be performed by older mice and that capture learning and memory deficits accurately. Tests that induce undue anxiety should be avoided, as older mice are less able to cope with extreme stress. Careful consideration should be given to the strain of mice being used. Some strains have known deficits in sight, hearing, or movement impairments with age, so tests that rely on these senses or speed may not accurately reflect the cognitive function of mice in impaired strains. Spatial learning and memory are skills that degrade with increased age in both humans and rodents, so these tests may be the most translatable [1, 2].

The Barnes maze is a test of spatial learning that was developed as an alternative to the Morris water maze. The Morris water maze requires prolonged swimming to find a hidden platform and can be very stressful, especially for older mice. The Barnes maze is a dry version that still tests the ability of mice to navigate using spatial cues. The original test was developed by Carol Barnes in 1979 and has been adapted for wild-type and transgenic mice [3–6]. Here, I will describe a protocol for the circular maze that has been successfully adapted for use in aging mice. This protocol has 1 day of habituation to the maze, followed by 4 days of training with three trials per day. The probe day follows 48 h after the final training session. Several types of useful analyses of both the training and probe days are described, which can be used to accommodate aging mice.

2 Materials

1. Barnes maze (*see* **Note 1**).
2. 15 cm diameter opaque open-ended cylinder with 20 cm height.
3. Digital video camera.

4. Tracking software like Noldus EthoVision or Stoelting ANY-maze.
5. Timer.
6. Radio tuned to static or other white noise machine (*see* **Note 2**).
7. Light (*see* **Note 3**).
8. Visual cues (*see* **Note 4**).
9. Lab notebook.
10. 70% ethanol.
11. Antimicrobial spray.

---

3 Methods

3.1 Day 1: Habituation Day

1. Bring the test mice to the room where behavior testing will be performed 30 min to 1 h prior to starting the test (*see* **Note 5**).
2. Set up the maze with no decoy or escape holes and all holes covered.
3. Wipe the test arena and cylinder with 70% ethanol (*see* **Note 6**).
4. Place the cylinder in the center of the arena.
5. Carry the mouse carefully (all feet supported) to the arena and place it in the cylinder (*see* **Note 7**).
6. After 10 s (*see* **Note 8**), remove the cylinder and allow the mouse to freely explore the arena for 3 min (*see* **Note 9**).

7. Place the mouse back in its cage.
8. Wipe the arena with 70% ethanol and repeat until all mice have experienced the arena.
9. After all mice have been tested, wipe the arena and cylinder with an antimicrobial spray and return all mice to their normal housing.

3.2 Days 2–5: Training Days

1. Bring the mice to the test room 30 min prior to the start of the test
2. Set up the maze with the escape hole, three decoy holes placed equidistant around the circle, and all other holes closed (*see* **Note 10**).
3. Wipe the test arena, cylinder, and all holes with 70% ethanol.
4. Place the cylinder in the center of the arena.
5. Turn on the bright lights and white noise.
6. Place the first mouse in the cylinder.
7. After 10 s, remove the cylinder, start video recording, and allow the mouse to explore the arena for 3 min.
8. If the mouse finds the escape hole and enters it, immediately turn off the lights and white noise and allow the mouse to remain there for 1 min.
9. If the mouse does not find the escape hole, gently guide it to the hole and help it enter. Turn off the lights and white noise and allow the mouse to remain in the escape hole for 1 min (*see* **Note 11**).

10. Carry the mouse in the escape hole back to its home cage and gently allow it to leave the hole.
11. Record the time the test ended for the mouse and the expected start time of the next trial, which will be 15 min later.
12. Clean the escape hole and arena with 70% ethanol.
13. Continue with the next mouse in the same way until all mice have been tested (*see* **Note 12**).
14. Each mouse will receive three trials per day for a total of 12 training trials over 4 days.
15. At the end of each training day, wipe the holes, arena, and cylinder with an antimicrobial spray and return all mice to their normal housing.

3.3 Day 7: Probe Day

1. 2 days after the final training trial, bring the mice to the test room 30 min prior to the start of the test.
2. Set up the maze with all the holes covered, including the escape hole (*see* **Note 13**).
3. Clean the test arena and all holes with 70% ethanol.
4. Place the cylinder in the center of the arena.
5. Turn on the bright lights and white noise.
6. Place the first mouse in the cylinder.
7. After 10 s, remove the cylinder and allow the mouse to explore the arena for 3 min.

8. Clean the arena with 70% ethanol.
9. Continue until all mice have completed the probe trial.
10. Clean the arena and cylinder with an antimicrobial spray and return all mice to their normal housing.

3.4 Analysis

The assessment of the Barnes maze is unique in that the training and probe trials can be analyzed in many ways, some of which can help account for the deficits in movement associated with aging. The most common way to analyze the training trials is averaging the time to enter the escape hole over the three trials of each day. This is the daily escape latency. Escape time should decrease over the training days if the animals have learned the task. If the mice are timid about entering the box, the training can also be analyzed as percent time in the target quadrant (*see* **Note 14**). If the mouse will not enter the escape hole but knows where it is, it will commonly spend the majority of the time near the escape hole. When the training is measured as percent time in the target quadrant, animals that know where the hole is located will show greater time in the correct quadrant and less in the opposite quadrant even if they do not enter the escape hole. Experimenters may also consider calculating the daily failure rate, which is how often the mouse has to be guided into the escape hole. This will be high on day 1 but should decrease with increasing trial numbers. Animals that have not learned (improved) by any of the measures over the 4 days of training should be excluded from further analysis.

The probe trial can also be analyzed in numerous ways and is more commonly shown with several types of analysis. Naturally, primary latency to escape the hole is generally included. Time spent in the target quadrant can also be indicative of spatial memory. In aging mice,

one concern can be movement speed. If the mice in one group are slower than the other due to size or frailty, primary latency may not be indicative of the level of learning. Instead, path length may be useful. If the mice have learned to use the spatial cues, they will walk directly to the target hole and have a shorter path length. This measure will discount any differences in walking speed. Another useful measure is primary errors or total errors. In this measure, the number of holes approached that are not the target hole is evaluated (*see* **Note 15**). This can also be a useful measure if there are differences in walking speed.

Together, analysis of the training and probe trials will help evaluate the effectiveness of treatment on learning and retention of spatial memory in aged mice.

---

4 Notes

1. The Barnes maze for mice should be a 90 cm diameter raised circular platform with 20 5-cm holes equidistant along the edges. Beneath the holes are: (1) One escape hole large enough for the mouse to be fully enclosed and not so deep that they cannot easily get into it themselves (11 cm × 6 cm × 5 cm) will be the target hole. (2) Decoy holes (3–4 is standard), which are shallower and shorter than the escape hole (5 cm × 5 cm × 3 cm), will reduce the reliance of the mouse on the visual drop that would only be correlated to the escape hole otherwise. However, they provide no escape from the noise or light. (3) Hole covers, which are flat, cover the remaining holes of the platform.
2. Noise should be >85 dB to be noxious to the mice.
3. The overhead light should be brighter than ambient lighting and equally illuminate the entire maze.
4. The mice will rely on visual cues around the arena to navigate to the correct hole. It is very important that the cues be placed where the mouse can see them and not be moved over the course of testing. The cues should be at the height of the platform and

of testing. They should be at the height of the platform and sufficiently different so that they are useful for navigation. They cannot be so far away that the mice would not see or notice them (50–200 cm from the maze edge).

5. Ideally, the behavior room would have an antechamber to decrease the amount of noise heard by the mice while not testing. If there is no antechamber, a hallway can be used if it can be made reasonably quiet and there is not much foot traffic. If neither of these are options, place the animals as far away from the testing area and noise source as possible and out of sight of the test arena.
6. Prior to the start of each day and after each tested animal, all areas of the arena should be wiped with 70% alcohol to clean the arena.
7. Ideally, the mice would be acclimated to handling. This can be done by daily handling for at least 1 week if the mice have not been previously handled.
8. Use a hand timer for this 10 s measurement. These 10 s do not need to be recorded on any of the test days and may make later analysis more difficult, as the 10 s will need to be removed from the analytic program's analysis.
9. This day does not need to be recorded, though the mice should be watched in case they fall off the arena. In the case a mouse falls, carefully replace it at the point it fell and continue the test, though the data from this trial should be excluded from analysis.
10. Different tracking software programs may require different inputs, but at this point the trial parameters should be defined within the program. Most programs will also allow setup later using the recorded videos, but check your system's requirements. Record the escape time as 180 s if the mouse is unable to find the

11. escape hole.
12. Typically, three mice can be run in one 15-min set by an experienced handler. On later test days, if the mice have learned, four mice can be run in a set.
13. Alternatively, some researchers will set up the decoy holes as in all previous trials and a decoy where the escape hole was previously.
14. Alternatively, the hole can be made slightly more palatable by adding a small piece of a nesting square. This would need to be changed after every animal. No food or other valuable reward should ever be added.
15. Many researchers include the target hole as well as the holes on either side so that all three holes are considered the "correct" hole when calculating the number of errors.

Competing Interests

The author has no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Oler JA, Markus EJ (1998) Age-related deficits on the radial maze and in fear conditioning: hippocampal processing and consolidation. Hippocampus 8:402–415 [https://doi.org/10.1002/(SICI)1098-1063(1998)8:4<402::AID-HIPO8>3.0.CO;2-I][PubMed]
2. Newman MC and Kaszniak AW (2000) Spatial memory and aging: performance on a human analog of the Morris water maze. Aging Neuropsychol Cogn 7:86–93
3. Barnes CA (1979) Memory deficits associated with senescence: a neurophysiological and behavioral study in the rat. J Comp Physiol Psychol 93:74–104
4. Bach ME, Hawkins RD, Osman M et al. (1995) Impairment of spatial but not contextual memory in CaMKII mutant mice with a selective loss of hippocampal LTP in the range of the theta

frequency. Cell 81:905–915
[https://doi.org/10.1016/0092-8674(95)90010-1][PubMed]

5. Harrison FE, Hosseini AH, McDonald MP (2009) Endogenous anxiety and stress responses in water maze and Barnes maze spatial memory tasks. Behav Brain Res 198:247–251
[https://doi.org/10.1016/j.bbr.2008.10.015][PubMed]

6. Pompl PN, Mullan MJ, Bjugstad K et al. (1999) Adaptation of the circular platform spatial memory task for mice: use in detecting cognitive impairment in the APP(SW) transgenic mouse model for Alzheimer's disease. J Neurosci Methods 87:87–95
[https://doi.org/10.1016/S0165-0270(98)00169-1][PubMed]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_9

9. Quantification of *C. elegans* Movement and Muscle Function Using WormLab-Derived Maximum Dynamic Amplitude

Brandy A. Weathers[1, 2], Tripti Nair[2] and Sean P. Curran[2] ✉

(1) University of Southern California, Dornsife College of Letters, Arts, and Sciences, Los Angeles, CA 90089, USA

(2) University of Southern California, Leonard Davis School of Gerontology, Los Angeles, CA 90089, USA

✉ **Sean P. Curran**
Email: spcurran@usc.edu

Abstract

Dynamic amplitude measures the extreme curvature point differences in a locomotor stroke as *Caenorhabditis elegans* crawls. This parameter serves as a sensitive biomechanical indicator of muscle coordination and strength and allows for detection of physiological changes caused by genetic mutations, aging, or exogenous exposure to chemical compounds. Here we describe a standardized approach to recording, processing, and analyzing maximum dynamic amplitude using WormLab software. This method facilitates the identification of phenotypes associated with the neuromuscular junction (NMJ), motor decline, and neurotransmitter signaling, providing a quantitative framework to assess locomotor health in *C. elegans*.

Keywords Dynamic amplitude – Neuromuscular junction – WormLab – Locomotor analysis – Movement

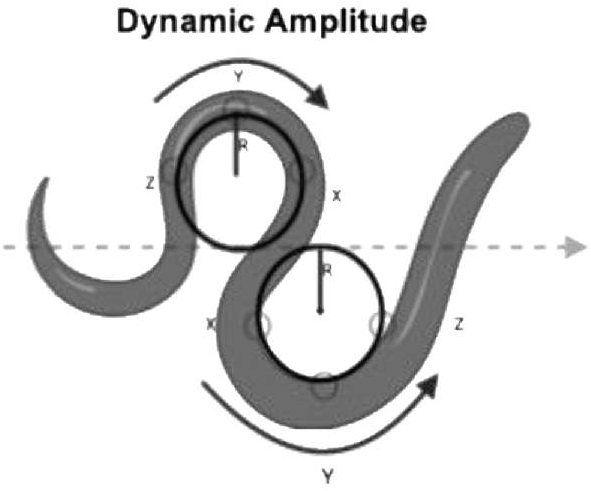
1 Introduction

Dynamic amplitude is a phenotype that has been measured in the *C. elegans* field recently for studies shaping aging and neuromuscular junctions [1]. This behavioral phenotype involves an age-related decline in the strength and coordination of muscle contractions [2]. The documented decline at the neuromuscular junction (NMJ) provides insight into how age-related changes in the nervous system contribute to the overall functional decline observed in aging [3]. Aging in the nematode has been characterized by altered structure of the muscle and reduced function. The overall decline is a result of structural disorganization and age-related mitochondrial effects leading to deterioration involving sarcomere breakdown—the basic contractile units [4]. While genetic interventions can result in changes in dynamic amplitude, the onset of this decline is stochastic, resulting in a variation even among genetically identical individuals in a uniform environment [1, 5]. Exogenous xenobiotic agents can induce measurable changes in dynamic amplitude, which can be linked to observable tissue and neuron phenotypes. This occurs through complex cellular signaling, gene expression, and metabolic pathways that are perturbed by the foreign compounds [6].

Using the program WormLab, dynamic amplitude is measured in the context of worm movement, especially during swimming or thrashing. The software analyzes video recordings of the worms to calculate the maximum curvature differences during a stroke [7]. This provides a quantitative metric for assessing the intensity of body bends.

For each frame of the video, WormLab quantifies the curvature at multiple points along the worm's body (Fig. 1), which changes as the animal moves. For each stroke, WormLab finds the two points along the worm's body that have the most extreme curvature. It then calculates

the maximum difference between these points. This value is the dynamic amplitude for that stroke.



$$\kappa Y = \frac{1}{R}$$

$$\boldsymbol{Dynamic\ Amplitude} = |\ \boldsymbol{max}(\boldsymbol{\kappa})\ | - |\ \boldsymbol{min}(\boldsymbol{\kappa})\ |$$

Fig. 1 Calculation of dynamic amplitude. Curvature (𝑘) is estimated at each midline point (excluding endpoints) by calculating the radius of the circle passing through three adjacent points, where the central point (Y) is the point of interest. The dynamic amplitude is then determined by subtracting the minimum absolute curvature from the maximum absolute curvature along the body of the animal

By quantifying dynamic amplitude, researchers can gain insights into the biomechanics and muscle function of worms in the context of aging, genetic alterations, and exogenous treatment of various compounds. It provides an objective, repeatable measurement that can be used to study subtle changes in movement.

---

2 Materials

Allow plates to calibrate to room temperature before placing animals (*see* **Note 1**). When working on animals treated exogenously with xenobiotic agents (biogenic amines, bacteria, etc.), wash animals with M9 multiple times to ensure there is no remaining residue.

2.1 Plate Preparation

1. Prepare stock solutions: 1 M $CaCl_2$, 1 M $MgSO_4$, and 0.5 M KP using doubled distilled water (ddH_2O).
2. Before placing solution in the autoclave, allocate the appropriate size glass bottles for each of the stock solutions.
    a. Add the caps but do not seal them completely, cover the caps with aluminum foil.
    b. Autoclave the bottles using the liquid program (120 °C for 20 min).
3. Prepare the Nematode Growth Medium (NGM) agar plates by adding: 2.5 g of bacto peptone, 3 g of NaCl, 17 g of agar, 0.2 g of streptomycin, and 975 mL of doubled-distilled water (ddH_2O) into a 2 L flask.
    a. Add a magnetic stir bar to help the agar remain homogenous
    b. Autoclave using the liquid program
4. After the NGM agar medium cools down to 55 °C, add 1 mL of 1 M $CaCl_2$, 1 mL of 1 M $MgSO_4$, 1 mL of 1 M 5 mg/mL of cholesterol, 1.25 mL of 10 mg/mL nystatin, 25 mL of 0.5 M KP buffer.
5. Pour the plates by adding 10 mL of NGM medium to 6 cm petri dishes using a sterile 25 mL pipette.
6. Allow plates to cool for the agar to properly solidify before use.

2.2 Sample Preparation

1. Create M9 buffer: 22 mM KH_2PO_4, 42 mM Na_2HPO_4, 86 mM NaCl,

1 mM $MgSO_4$

2. Allow animals to grow until the adult stage and recover gravid adults in 15 mL tubes by washing plates with M9 buffer.
3. Prepare the bleaching solution: 27 mL H_2O, 7.2 mL concentrated bleach (7.5% hypochlorite), 2.1 mL 0.5 M NaOH.
4. Wash gravid adults using M9 and centrifuge for 0.5 min at 1000 × *g* at room temperature. Remove the supernatant to the 3 mL mark.
5. Add 1:1 volume (3 mL) of bleach solution to the supernatant and agitate for around 1.0 min.
6. Fill tubes to the 15 mL mark with M9 and centrifuge again at 0.5 min, 1000 × *g*
7. Remove supernatant and fill to the top with M9. Resuspend the pellet in 3ml M9.
8. Repeat Steps 5–7 to double bleach and wash again.
9. Wash three to five times fully in M9, centrifuging and removing the supernatant each time.
10. Add M9 to 8 mL and spin overnight 16–20 h to allow eggs to hatch.
11. Place L1 animals onto NGM plates and allow animals to reach the desired developmental stage [e.g., juvenile L4 (40–48 h post L1 starvation (0 h)], adult stage Day (D)1 D1 (~72 h) and D7 (~168 h).

---

3 Methods

The WormLab Imaging System is a video acquisition tool solution for automated imaging and quantitative analysis of the behavior of *C.*

elegans and other nematodes.

3.1 Video Acquisition

1. Wash synchronized animals of desired developmental stage off NGM plates using M9 into a 15 mL tube.
2. Pellet worms by centrifuging for 0.5 min at 400 × *g* at room temperature. Remove the supernatant and wash one to three times to properly clear bacteria.
    a. Worms can also be washed by allowing animals to fall to the bottom of the tube before removing supernatant and washing again (gravity settling).
3. Transfer around 100 animals to an empty NGM plate (no bacterial colonies are present).
4. Allow animals to crawl freely on the plate after the droplets have fully dried, around 1 h, before image acquisition (Fig. 2).



Fig. 2 *C. elegans* on a plane NGM plate under WormLab camera. Animals are given 1 h to move freely on an NGM plate without physical limitations (e.g., liquid, clumping,

overcrowding)

5. Record using WormLab software at a frame rate of 7.5 fps for crawling (*see* **Note 2**). Set the buffer size to 128 frames (default). Maintain consistent illumination (*see* **Note 3**) and contrast settings across recordings to ensure the full animal is registered by the program.
    a. Clicking on the “display clipping” tab will show how much illumination is being registered by the program (shown as red dots).
    b. Increase the illumination until small specs of clippings (red dots) are seen. Large amounts of display clippings show that the illumination is too bright.

6. Set the time of image acquisition to 1 min across each video and ensure the strains are properly visible by adjusting the camera until the video feedback is clear.
7. After image acquisition is completed, record the ruler for at least 1 s for usage in data analysis (Fig. 3).

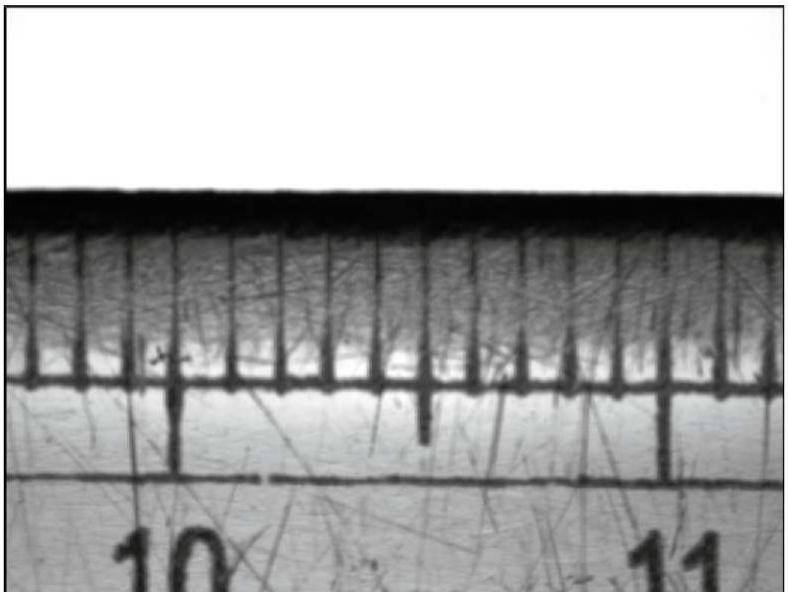




Fig. 3 Video acquisition of ruler for scale. Using WormLab software, a video of at least one second is taken. This video is loaded into WormLab, and scale is determined from clicking the "Measure" button and dragging the measurement tool from one centimeter (cm) on a ruler. Scale adjustment is manually entered in WormLab to account for the length of 1 millimeter (mm) (1 cm = 10 mm)

3.2 Data Filtering and Processing

1. Set the sequence information by using a video image with a horizontal ruler to measure the scale (μm/pixel) of each video.
    a. This will only need to be performed once if all videos are recorded at the same scale in one session.
2. Click the Measure button to set the scaling.
3. Click and drag to draw a measure line on the image. Draw 1 mm length. The software will count the number of pixels. After the mouse is released, the scale is recorded.
4. Press Save. The scale will be set for each video.
5. For each video, use Adjust Image to highlight the full body of each animal by increasing the Threshold Level.
    a. Under the Image Processing tab, hit the Hotspot illumination correction tab to balance the contrast and illumination of the video.
    b. Press Apply to see the change properly.
6. Within the Detect and Track page, under the Detection tab, click Detect worms to highlight the animals in the video. Properly detected animals will have numbers attached.

a. Any animals not detected by the program can be manually selected.

7. Under the Advanced tab (within the Detect and Track tab), change Detection frequency to 7 frames. Check the Length fitting and Width fitting to detect animals of the desired length and width.
8. Within the Detect and Track page, under the Tracking tab, filter the minimum track duration (frames) to 400 of 450 frames to be included in the analysis (*see* **Note 4**).
    a. Under the Tracking Mode, ensure that the proper mode is selected (Crawling).
9. Click on the Save Project page to properly save your Worm Project (.wpr) under the proper folder.
10. After all desired worm projects are created, select the Batch tab. Add all desired files using the Add option under File Selection. Press Start under Batch Processing Control to run all projects.

3.3 Data Analysis

1. Open worm projects under File > Open Project.
2. To view the analysis of each project, click on View > Analyze Data.
3. Under the Select an Analysis header, find Track Summary.
4. Select the Max Amplitude tab and highlight all individual animals (*see* **Note 5**).
    a. Max Amplitude quantifies the maximum curvature difference per stroke, reported as dynamic amplitude (Fig. 1).

5. Click Copy Selected Data to copy data to clipboard.
6. Paste values onto an Excel sheet or Google sheet.
7. Use the average of all recorded Max Amplitude values per genotype to remove extreme outliers from the analysis (*see* **Note 6**).
8. Export amplitude and worm length values values for each animal and compute the mean dynamic amplitude and mean worm length per genotype or treatment condition.
9. Statistical significance for the comparison of normalized dynamic amplitude (p-value ≥0.05) can be assessed using ANOVA (one-way or two-way) followed by post-hoc comparisons (e.g., Tukey's test) or using a t-test if comparing between two genotypes or conditions (*see* **Note 7**).

---

4 Notes

1. Ensure recordings are made at the same temperature to minimize behavioral variability.
2. Frame rates can be adjusted in the "settings" tab. Dynamic amplitude values can be sensitive to frame rate; maintain 7.5 fps consistently across all conditions.
3. Increasing illumination was acquired by checking the "Show stimulus controls" tab and using the "White LED" bar.
4. It is beneficial for the Track Filtering to be on and filtered to a minimum track duration of 50 less frames than the total frame amount. This ensures that the tracked animal was imaged for a majority of the video.
5. To extract each individual animal under only the Max Amplitude

tab, use ctrl + click Max Amplitude. It will highlight only the Max Amplitude row.

6. WormLab can make mistakes and record objects as animals incorrectly. To control for this, divide each individual value by the total average to obtain the normalized amplitude ratio, which helps identify and exclude outliers. Exclude data points that fall outside 1.5× the interquartile range (IQR) or ±2 standard deviations from the mean. You can use conditional formatting to highlight these values in red for easier visualization.

7. When comparing across genotypes, normalize amplitude to the mean of the WT control. When comparing conditions, normalize amplitude to the untreated conditions. The dynamic amplitude is subject to worm length (a variability parameter across developmental stages and across strains) [8, 9]. To account for this variability, the dynamic amplitude could be divided/normalized by the ratios of mean worm body length between strains or developmental stages.

Competing Interests

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Cho J et al (2025) Determination of health status during aging using bending and pumping rates at various survival rates in *Caenorhabditis elegans*. Sci Rep 15(1):9057 [https://doi.org/10.1038/s41598-025-93876-8][PubMed][PubMedCentral]

2. Chow DK et al (2006) Sarcopenia in the *Caenorhabditis elegans* pharynx correlates with muscle contraction rate over lifespan. Exp Gerontol 41(3):252–260 [https://doi.org/10.1016/j.exger.2005.12.004][PubMed][PubMedCentral]

3. Liu J et al (2013) Functional aging in the nervous system contributes to age-dependent motor activity decline in *C. elegans*. Cell Metab 18(3):392–402 [https://doi.org/10.1016/j.cmet.2013.08.007][PubMed][PubMedCentral]

4. Gaffney CJ et al (2018) Greater loss of mitochondrial function with ageing is associated with earlier onset of sarcopenia in *C. elegans*. Aging (Albany NY) 10(11):3382–3396 [https://doi.org/10.18632/aging.101654][PubMed][PubMedCentral]

5. Van Camp BT, Zapata QN, Curran SP (2025) WormRACER: robust analysis by computer-enhanced recording. GeroScience 47(3):5377–5387 [https://doi.org/10.1007/s11357-025-01631-8][PubMed][PubMedCentral]

6. Melo JA, RuvkunG (2012) Inactivation of conserved C. elegans genes engages pathogen- and xenobiotic-associated defenses. Cell 149(2):452–466

7. Roussel N et al (2014) Robust tracking and quantification of *C. elegans* body shape and locomotion through coiling, entanglement, and omega bends. Worm 3(4):e982437 [https://doi.org/10.4161/21624054.2014.982437][PubMed]

8. Nahabedian JF, et al (2012) Bending amplitude – A new quantitative assay of C. elegans locomotion: Identification of phenotypes for mutants in genes encoding muscle focal adhesion components. Methods 56(1):95–102. https://doi.org/10.1016/j.ymeth.2011.11.005 [PubMed]

9. Brar HK (2024) Functional recovery associated with dendrite regeneration in PVD neuron of *C. elegans*. eNeuro 11(5):0292–23.2024. https://doi.org/10.1523/eneuro.0292-23.2024

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_10

10. Voluntary Wheel Running as a Noninvasive and Low-Stress Exercise Paradigm in Aging Mice

Jocelyne Leon[1], Karel Aceituno[1] and Constanza J. Cortes[1] ✉
(1) Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA

✉ **Constanza J. Cortes**
Email: Cjc_149@usc.edu

Abstract

Voluntary wheel running (VWR) in rodents provides a physiologically relevant, low-stress model of chronic endurance exercise that recapitulates human training adaptations across multiple tissues, including enhanced mitochondrial biogenesis, oxidative capacity, and neuroprotection, making it ideal for aging and neurodegenerative disease research. Unlike forced treadmill protocols, in-cage VWR systems enable long-term and/or lifelong studies by minimizing handling stress, supporting nocturnal running behavior, and reducing experimenter burden, while accounting for sex-specific differences in running distance and motivation that influence exercise dosing. Here, we detail a standardized protocol for implementing VWR in standard vivarium housing for young and aged mice, with key considerations for experimental design, behavioral monitoring, and data interpretation to ensure robust, reproducible outcomes in preclinical exercise studies.

Keywords Exercise – Endurance training – Sex differences – Aging

1 Introduction

Exercise represents one of the most potent and reproducible interventions to counteract the physiological deterioration associated with aging. Across species, sustained physical activity attenuates sarcopenia, mitigates metabolic inflexibility, improves vascular function, and preserves cognitive performance. In rodent models, voluntary wheel running has emerged as a tractable and physiologically relevant paradigm to study these effects. Both mice and rats display a strong innate drive to engage in spontaneous wheel running, often covering several kilometers per night under standard housing conditions. This voluntary behavior provides a chronic, low-stress model of endurance exercise that recapitulates many of the molecular and physiological adaptations observed in trained humans, thereby facilitating mechanistic studies of exercise-regulated processes at the tissue, cellular, and systems levels.

From an experimental standpoint, endurance-type activity such as voluntary wheel running (VWR) [1] induces robust remodeling across multiple organ systems, including skeletal muscle, liver, adipose tissue, and brain [2, 3]. These adaptations involve enhanced mitochondrial biogenesis, coordinated activation of oxidative and angiogenic gene programs, modulation of nutrient-sensing and proteostasis pathways, and preservation of synaptic plasticity and neurogenesis [4]. Furthermore, epidemiological and clinical data link long-term endurance exercise, often spanning decades, to stable neuroprotective effects in the central nervous system [5–8]. This training effect appears critical for the accumulation of exercise-induced stable neuroprotective adaptations [9, 10], particularly in the context of neurological disease [11, 12]. In contrast, acute or short-term exercise interventions fail to fully recapitulate these benefits in both humans [9, 10, 13] and rodents [14] (as reviewed in [12]).

To date, most preclinical studies in aging and neurodegenerative disease mice have relied on short-term (≤4 weeks) or highly stressful forced treadmill exercise paradigms in symptomatic mice [15–18], often producing inconsistent outcomes [11, 12, 19]. VWR better models human long-term endurance training by minimizing stress [17, 20], enabling self-selected activity, and allowing assessment of motivation and motor deficits relevant to aging. VWR also more closely reflects the natural nocturnal activity patterns of mice and minimizes stress-related artifacts commonly associated with forced treadmill paradigms. In contrast, forced treadmill exercise often necessitates the use of aversive stimuli (such as gentle prodding or mild shocks) to maintain consistent running, which can elevate stress and inflammatory responses. These effects are particularly relevant in aging studies, wherein reduced natural locomotor activity and/or motivation might lead to increased aversive stimulus usage. Furthermore, housing mice with a running wheel in their home cage avoids the repeated stress of daily transport to a treadmill chamber (resulting in a less perturbed baseline and more physiologically relevant exercise paradigm) and enables long-term exercise studies by minimizing animal stress from repeated handling while also substantially reducing experimenter time and involvement compared to daily forced treadmill protocols. This feature is especially useful for longitudinal studies of aging, where timelines can span months to years, and minimizing chronic stress and experimenter burden is critical for maintaining animal welfare and data consistency over prolonged periods. As a result, VWR might be a more useful and physiologically relevant intervention if studying the intrinsic effects of endurance training in mice, especially for studies focusing on stress-related alterations in corticosterone, anxiety-like behavior, as well as circadian, inflammatory, or metabolic parameters.

Sex differences in voluntary wheel running are well documented in mice, with females typically running greater distances and spending more time on the wheel than males in voluntary running paradigms [21, 22]. These sex-specific patterns are most pronounced in young animals and can decline with age [23], underscoring the need to

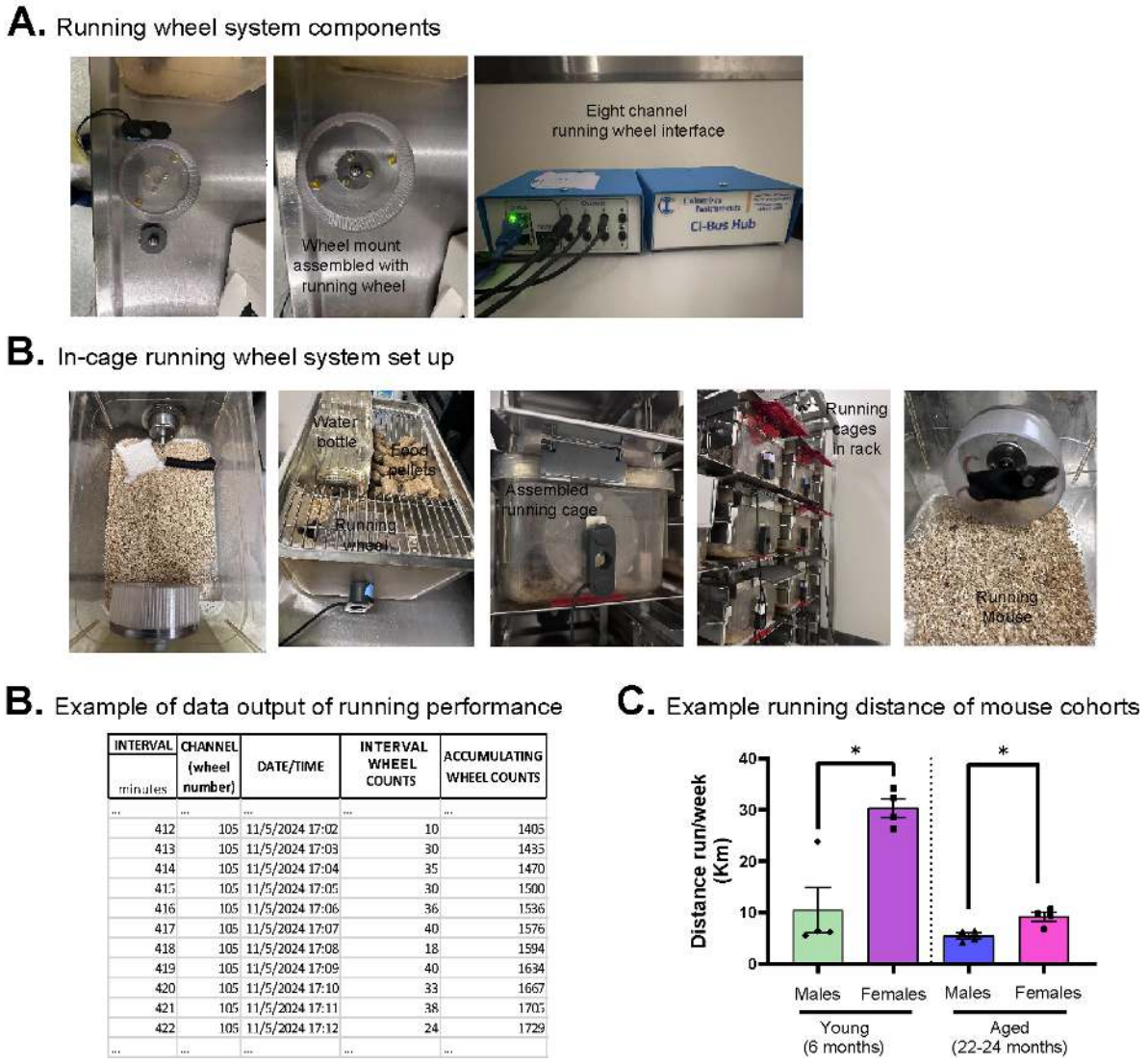
account for sex as a biological variable when designing and interpreting long-term wheel-running studies. This inherent variability can lead to heterogeneous exercise "doses" across individuals [2], creating exercise-burden differences that are not present in forced treadmill studies, where speed, duration, and frequency are tightly controlled. Consequently, careful monitoring and characterization of individual voluntary running behavior are essential for accurately interpreting physiological and behavioral outcomes in aging mice under wheel-based exercise paradigms.

Below, we outline a VWR protocol for use in an in-home cage system in standard vivarium racks, highlight important considerations for experimental design and monitoring, and provide examples of real data collected from young (6 months old) and aged (20–24 months old) wild-type mice.

---

2 Materials

- In-cage nonergometric voluntary running wheel system (Fig. 1A).
  - Wheel components described here are from the Columbus Instruments Mouse Home Cage Running Wheel system (Cat # 0294).



INTERVAL minutes	CHANNEL (wheel number)	DATE/TIME	INTERVAL WHEEL COUNTS	ACCUMULATING WHEEL COUNTS
...	...	...	...	...
412	105	11/5/2024 17:02	10	1405
413	105	11/5/2024 17:03	30	1435
414	105	11/5/2024 17:04	35	1470
415	105	11/5/2024 17:05	30	1500
416	105	11/5/2024 17:06	36	1536
417	105	11/5/2024 17:07	40	1576
418	105	11/5/2024 17:08	18	1594
419	105	11/5/2024 17:09	40	1634
420	105	11/5/2024 17:10	33	1667
421	105	11/5/2024 17:11	38	1705
422	105	11/5/2024 17:12	24	1729
...	...	...	...	...

Fig. 1 Voluntary wheel running setup and example data. (**A**) Running wheel system components highlighting the individual parts before assembly. (**B**) In-cage running wheel system setup, showing orientation of cage components, assembled running cage in the rack, and an example of active wheel utilization by the individual. (**C**) Example of data output of running performance from the Columbus monitoring software. (**D**) Example of quantified and converted data of distance run in kilometers per week for young (~6 months of age) and aged (~22–24 months of age), male and female mice (n = 4/group). Data are represented as average ± SEM. Each individual icon represents an individual mouse. *P < 0.05, t-test between age-matched sexes only

- Laptop with Microsoft Excel software.
- Animal cohorts, as needed.

2.1 Aging Animal Cohort Preparation and Considerations

1. Aging mice will run less total distance and spend less time on the wheel than young mice. Depending on the nature of the study, considerations should be taken regarding age-associated issues such as weight gain/loss, arthritis, cataracts, or tumors that may impair wheel use or confound exercise effects (*see* **Notes 1** and **2**).
2. By definition, VWR studies will contain large variability in individual animal performance. Based on our experience, we recommend utilizing at least $n = 12$ individuals/experimental cohorts to obtain biologically meaningful results. However, the use of power calculations is highly recommended to determine proper animal group sizes for the researchers' specific scientific question.
3. Perform health checks on all experimental animals before enrolling in the study for their coat appearance, gait, BAR (bright, alert, responsive) status, potential tumors, and any other physical signs that could interfere with running behavior (*see* **Note 3**). Replace such individuals from experimental cohorts.
4. Habituation of animals to the location/vivarium before initiation of the VWR protocol may be required (*see* **Note 4**).
5. Lab personnel should conduct visual health checks on experimental groups at least three times a week to monitor animal welfare, as well as potential cage issues, including water leaks and wheel blockage (*see* **Notes 5** and **6**).

---

3 Methods

3.1 In-Cage Voluntary Wheel Preparation

1. The following procedure requires the use of a standard vivarium polypropylene cage or other similarly sized enclosures.
2. Individual components include a mouse voluntary running wheel (10.16 cm in diameter), a wheel mount, a Hall effect sensor, an 8-channel running wheel interface (Catalog # 0294), and a

computer with the most recent Columbus Instruments Link monitoring software (Fig. 1A).

3. Set up the running cage by using a standard vivarium polypropylene cage (290 × 180 × 160 mm). In standard caging systems, the wheel can only be attached to the wall opposite the lixit/rack connector (see Fig. 1B).
4. Add enough sawdust material to cover the bottom (*see* **Note 7**), and approved enrichment items such as manzanita sticks and nestlets (*see* **Notes 8** and **9**). Take care to place them on the opposite side where the wheel will be (Fig. 1B).
5. Experimental cohorts should include individual runner cages (with the set-up described above) and individual sedentary cages where the running wheel is still present but locked and/or unusable (*see* **Note 10**).
6. Assemble the wheel by adding the wheel mount on the outside of the rungless running wheel (Fig. 1B).
7. Insert the assembled wheel inside the polypropylene cage and attach a Hall effect sensor on the outside of the cage (Fig. 1B).
8. Make sure the wire from the Hall effect sensor is facing down (Fig. 1B), and the wheel mount magnet has the thinnest portion also facing down (Fig. 1B).
9. Connect the assembled and mounted Hall effect sensor to the eight-channel running wheel interface (Fig. 1B).
10. Connect the eight-channel running wheel interface to the CI-Bus communication system, then connect the system to a USB port on a computer with a Windows operating system. Ensure that the CI-Bus is now listed under “Devices” on the Hardware tab of CI-Link application. In the device menu, the status of all equipment should be listed as “OK.” indicating that the device has been detected and

is ready to begin a new experiment (*see* **Notes 11** and **12**).

11. To begin a new experiment, under the Setup Experiment tab of CI-Link, create a unique experiment name and specify the desired location for the data to be stored on your device. Under "Roster Edit," include animal identifying information for each corresponding running wheel. Start the experiment recordings using the "Start" button (*see* **Note 13**).

12. To ensure the wheel is working and recording, manually spin the wheel while connected to the appropriate monitoring software (Columbus Instruments) and navigate to the "Live Data" section of the Analysis tab and visually see the counter increasing for the corresponding wheel being manually tested (*see* **Notes 14** and **15**).

13. Pre-screened mice can now be single-housed in each pre-prepared sedentary or running cage (*see* **Note 16**). If applicable, when adding the wire food rack/lid to the cage, the flat portion of the rack should face the wheel assembly to allow the wheel to freely turn inside.

14. Sedentary and running cages can be placed back on their existing rack system (Fig. 1B), ensuring connection to the wheel interface system is stable and secure.

3.2 Voluntary Running Monitoring

1. This in-cage wheel running system for mice collects the number of revolutions and the timing of the revolutions for each cage at user-determined time intervals using Columbus Instruments-provided monitoring software. An example of 1-min intervals is shown in Fig. 1C.
2. Monitoring of animal welfare and physical activity must be done daily to ensure adherence to the intervention, proper wheel

functioning, and removal of any nesting materials obstructing the free turning of the wheel.

3. We recommend downloading the running performance data every 3 days or less and moving it into long-term storage.

4. Cages should be changed every 2 weeks to ensure cleanliness and reduce animal stress. Additional cleanouts are recommended as required (i.e., cage flooding, excessive urination, etc.).

5. This in-cage voluntary wheel running system enables long-term studies, including lifelong investigations of exercise behavior and its physiological consequences in mice.

3.3 Data Output and Analysis

1. Data output records wheel revolutions at user-specified intervals.

2. Data can be viewed in real time (if needed) and exported into a CSV format at the end of the experiment (Fig. 1B) (*see* **Note 16**).

3. To export data, under the Run Experiment tab of CI-Link, select the red "Stop" button to stop the experiment. Then, in the Analysis tab, select the "Export CSV" option.

4. Each CSV export file contains an interval wheel count and a cumulative wheel count column. To calculate daily wheel running, first select a desired time of day at which to record the cumulative wheel counts of each day. We recommend right after vivarium-specific "light-outs," i.e., 6:15 p.m. every day (*see* **Note 17**). Record each day's cumulative wheel counts at this specific time in a new spreadsheet. For consecutive days recorded in the same experiment CSV export files, converting from cumulative to individual activity for each specific day can be calculated by subtracting the cumulative wheel counts of the day prior.
To convert wheel revolutions/time unit to running distance, use the

5. To convert wheel revolutions/time unit to running distance, use the following equation:

    Distance (in km) = revolutions *pi* wheel diameter (in cm)] *1*e-05

6. Example data from a small pilot cohort is shown in Fig. 1C.

3.4 Downstream Analysis and/or Tissue Collection

If possible, personnel responsible for parallel, downstream testing and any behavioral and/or biochemical analysis should be blinded to animal genotype and/or experimental condition until the running analysis has been completed for the entire cohort.

---

4 Notes

1. We recommend pairing VWR with standard aging functional testing (grip strength, rotarod, cognitive tasks), if possible, to distinguish wheel-specific effects from general physical decline.
2. Voluntary running reflects motivation as well as capacity, so reduced running in old mice may not equal reduced exercise tolerance.
3. In our hands, animal attrition due to age-associated complications is minimal if groups are properly screened before entering the study.
4. If animals are imported into the vivarium (i.e., transgenic mice, commercially purchased cohorts), allow all groups to habituate for 1 month in the local vivarium before testing.
5. All mouse handling must be performed by experienced researchers to minimize animal stress.
6. Older mice (≥16 months of age) in particular require regular health checks to ensure no development of age-associated

phenotypes that may interfere with their running performance throughout the course of the intervention.

7. Other cage materials, such as paper shreds, are not recommended, as they promote nest building around the wheel rather than running.

8. For mice that use nestlet material to clog the water lixit, water bottles can be used instead (Fig. 1B).

9. Enrichment items are strongly recommended due to single-housing nature of the experimental setup.

10. Sedentary ("control") groups must also be single-housed in cages with a wheel to control for enrichment. To prevent wheel usage but allow wheel interaction (i.e., enrichment), we recommend placing a free wheel face down at the bottom of the cage.

11. If any other status is indicated, the device may need to be updated, is not being detected, or requires other maintenance before beginning to record a new experiment. Consult the CI-Link Manual for further troubleshooting.

12. Once the software is initiated, all data collected from connected devices are recorded in .cid files, whether an experiment has been initiated or not. Buffer .wal and .shm files are also created to accompany .cid files, which can help restore data in the event of system or software crashes.

13. We recommend creating a new experiment with each data sync to prevent data loss.

14. Polypropylene thickness in home cages varies by brand/usage, and can affect the ability of the outside magnet to detect wheel movement. If the magnet sensor is not detecting wheel motion, a

thinner magnet cover from Columbus Instruments is available upon request.

15. Checking that the magnet has not moved in orientation or that the wheel is not hitting either the bottom or top of the cage is essential to ensure continuous recording and should be done daily or every other day.

16. When naming files in the Columbus Instruments monitoring software, do not use symbols such as backslash, colon, or semicolon, as they can corrupt the file.

17. Laboratory mice are nocturnal, and voluntary wheel-running behavior is almost exclusively expressed during the dark (active) phase of the light–dark cycle. Running typically begins shortly after lights-off, peaks within the first few hours of the dark period, and then gradually declines, such that the majority of daily distance and time on the wheel occur at night.

Acknowledgments

This work was supported by NIH/NIA R01 AG077536 (to C.J.C.).

Competing Interests

The authors have no competing interests to declare.

References

1. Goh J, Ladiges W (2015) Voluntary wheel running in mice. Curr Protoc Mouse Biol 5(4):283–290. https://doi.org/10.1002/9780470942390.mo140295 [PubMed][PubMedCentral]

2. Manzanares G, Brito-da-Silva G, Gandra PG (2018) Voluntary wheel running: patterns and physiological effects in mice. Braz J Med Biol Res 52(1):e7830. https://doi.org/10.1590/1414-431X20187830 [PubMed][PubMedCentral]

3.

Patterson DC, Birnbaum A, Matthews I et al (2025) Convergent skeletal muscle cytokine responses to TFEB overexpression and voluntary wheel running reflect sex-based variability to exercise adaptations in mice. Physiol Rep 13(16):e70519. https://doi.org/10.14814/phy2.70519
[PubMed][PubMedCentral]

4. Schmitt A, Herzog P, Rochner F et al (2020) Skeletal muscle effects of two different 10-week exercise regimens, voluntary wheel running, and forced treadmill running, in mice: a pilot study. Physiol Rep 8(20):e14609. https://doi.org/10.14814/phy2.14609
[PubMed][PubMedCentral]

5. Brisswalter J, Collardeau M, René A (2002) Effects of acute physical exercise characteristics on cognitive performance. Sports Med 32(9):555–566. https://doi.org/10.2165/00007256-200232090-00002
[PubMed]

6. Fenesi B, Fang H, Kovacevic A et al (2017) Physical exercise moderates the relationship of Apolipoprotein E (APOE) genotype and dementia risk: a population-based study. J Alzheimers Dis 56(1):297–303. https://doi.org/10.3233/JAD-160424
[PubMed]

7. Firth J, Stubbs B, Vancampfort D et al (2018) Effect of aerobic exercise on hippocampal volume in humans: a systematic review and meta-analysis. NeuroImage 166:230–238. https://doi.org/10.1016/j.neuroimage.2017.11.007
[PubMed]

8. Head D, Bugg JM, Goate AM et al (2012) Exercise engagement as a moderator of the effects of APOE genotype on amyloid deposition. Arch Neurol 69(5):636–643. https://doi.org/10.1001/archneurol.2011.845
[PubMed][PubMedCentral]

9. Hopkins ME, Davis FC, Vantieghem MR et al (2012) Differential effects of acute and regular physical exercise on cognition and affect. Neuroscience 215:59–68. https://doi.org/10.1016/j.neuroscience.2012.04.056
[PubMed][PubMedCentral]

10. Loprinzi PD, Roig M, Etnier JL et al (2021) Acute and chronic exercise effects on human memory: what we know and where to go from here. J Clin Med 10(21). https://doi.org/10.3390/jcm10214812

11. Miki Stein A, Munive V, Fernandez AM et al (2017) Acute exercise does not modify brain activity and memory performance in APP/PS1 mice. PLoS One 12(5):e0178247. https://doi.org/10.1371/journal.pone.0178247
[PubMed][PubMedCentral]

12. da Costa Daniele TM, de Bruin PFC, de Matos RS et al (2020) Exercise effects on brain and behavior in healthy mice, Alzheimer's disease and Parkinson's disease model-A systematic review and meta-analysis. Behav Brain Res 383:112488. https://doi.org/10.1016/j.bbr.2020.112488

[PubMed]

13. Borrega-Alonso E, Otamendi FJ (2025) Effects of acute and chronic exercise on working memory in healthy adults. An experimental investigation. Front Psychol 16:1608721. https://doi.org/10.3389/fpsyg.2025.1608721
[PubMed][PubMedCentral]

14. Shahandeh M, Roshan VD, Hosseinzadeh S et al (2013) Chronic exercise training versus acute endurance exercise in reducing neurotoxicity in rats exposed to lead acetate. Neural Regen Res 8(8):714–722. https://doi.org/10.3969/j.issn.1673-5374.2013.08.006
[PubMed][PubMedCentral]

15. Kim D, Cho J, Kang H (2019) Protective effect of exercise training against the progression of Alzheimer's disease in 3xTg-AD mice. Behav Brain Res 374:112105. https://doi.org/10.1016/j.bbr.2019.112105
[PubMed]

16. Kim D, Cho J, Lee I et al. (2017) Exercise attenuates high-fat diet-induced disease progression in 3xTg-AD mice. Med Sci Sports Exerc 49(4):676–686. https://doi.org/10.1249/mss.0000000000001166
[PubMed]

17. Svensson M, Rosvall P, Boza-Serrano A et al (2016) Forced treadmill exercise can induce stress and increase neuronal damage in a mouse model of global cerebral ischemia. Neurobiol Stress 5:8–18. https://doi.org/10.1016/j.ynstr.2016.09.002
[PubMed][PubMedCentral]

18. Zhang X, He Q, Huang T et al (2019) Treadmill exercise decreases Aβ deposition and counteracts cognitive decline in APP/PS1 Mice, possibly via hippocampal microglia modifications. Front Aging Neurosci 11:78. https://doi.org/10.3389/fnagi.2019.00078

19. Svensson M, Andersson E, Manouchehrian O et al (2020) Voluntary running does not reduce neuroinflammation or improve non-cognitive behavior in the 5xFAD mouse model of Alzheimer's disease. Sci Rep 10(1):1346. https://doi.org/10.1038/s41598-020-58309-8
[PubMed][PubMedCentral]

20. Cook MD, Martin SA, Williams C et al (2013) Forced treadmill exercise training exacerbates inflammation and causes mortality while voluntary wheel training is protective in a mouse model of colitis. Brain Behav Immun 33:46–56. https://doi.org/10.1016/j.bbi.2013.05.005
[PubMed][PubMedCentral]

21. Janowski AJ, Berardi G, Hayashi K et al (2025) The influence of sex on activity in voluntary wheel running, forced treadmill running, and open field testing in mice. Physiol Rep 13(4):e70246. https://doi.org/10.14814/phy2.70246
[PubMed][PubMedCentral]

22. Kim MJ, Carmichael PB, Bose U et al (2023) Sex differences in body composition, voluntary wheel running activity, balance performance, and auditory function in CBA/CaJ mice across the lifespan. Hear Res 428:108684. https://doi.org/10.1016/j.heares.2022.108684

[PubMed]

23. Bartling B, Al-Robaiy S, Lehnich H et al (2017) Sex-related differences in the wheel-running activity of mice decline with increasing age. Exp Gerontol 87(Pt B):139–147. https://doi.org/10.1016/j.exger.2016.04.011
[PubMed]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_11

11. Characterizing Gut Diversity During Aging

Prerana Vaddi[1], Denise B. Lynch[2] and Cristal M. Hill[1] ✉
(1) Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA
(2) One Codex, 2810N. Church St, Wilmington, DE PMB 86555, USA

✉ **Cristal M. Hill**
Email: cristalh@usc.edu

Abstract
Microbiome Protocols and Analyses are commonly reported with 16S sequencing in young animals. Our protocol aims to establish a standardized method for shotgun sequencing to improve the resolution of differentially expressed bacteria in aging mice, which otherwise have low alpha and beta diversity.

Keywords Aging – Microbiome – Shotgun sequencing – Microbiome statistics

1 Introduction

Aging poses unique challenges to gut microbiome studies. For example, irregular bowel function, age-related disease, and increased fragility contribute to impaction or constipation, making stool collection more difficult in aged animals [1,2]. This necessitates proper animal and sample handling to avoid cross-contamination and obtain accurate

reads (Fig. 1). Older animals also have lower microbial diversity and richness [3]. Leveraging methods that increase the resolution of microbial diversity and richness is optimal for microbiome studies during aging. However, the most important consideration is that aging animals have higher mortality rates [4]. The probability that sample size may drastically reduce from the beginning to the end of the experiment requires experimenters to compensate with more animals and utilize appropriate statistical methods to detect microbiome differences. As we will discuss, microbiome statistical tests are highly sensitive to sample numbers. Young adult murine models are commonly used in microbiome studies to control for age-related effects [3,5]. In these studies, 16S sequencing is primarily used for its cost efficiency and reliability. However, a new method of sequencing, called shotgun sequencing, increases the resolution of species diversity and richness [6,7]. In this protocol, we provide a revised set of shotgun sequencing methods to ensure proper study of the gut microbiome in aging animals. These methods begin with a clear study design and sample collection, followed by approaches aimed at standardized DNA or RNA extraction, library preparation, and sequencing using an appropriate platform (Fig. 2). After sequencing, the workflow typically includes data quality control, removal of contaminants, taxonomic and functional profiling, and statistical analysis to compare microbial communities across samples or conditions.

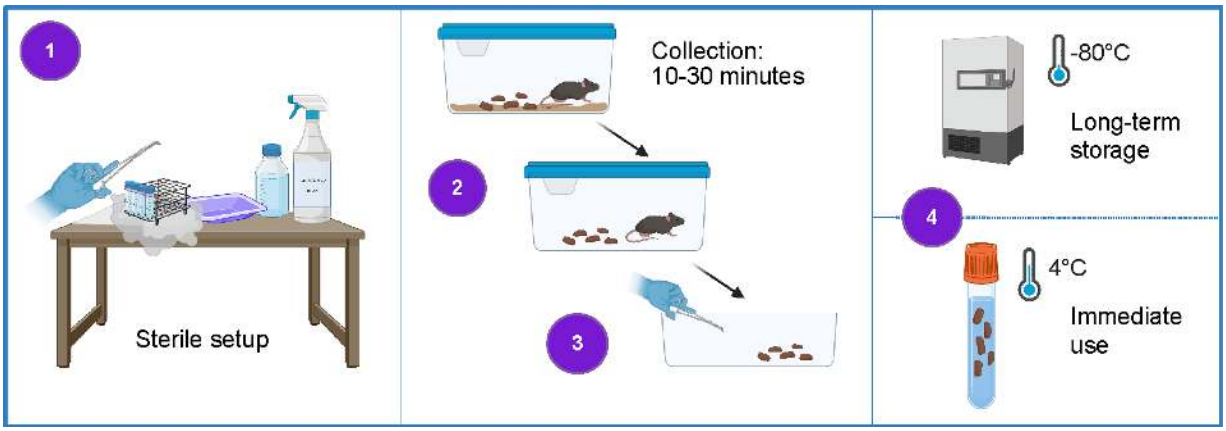


Fig. 1 Step diagram of the fecal collection technique in mice

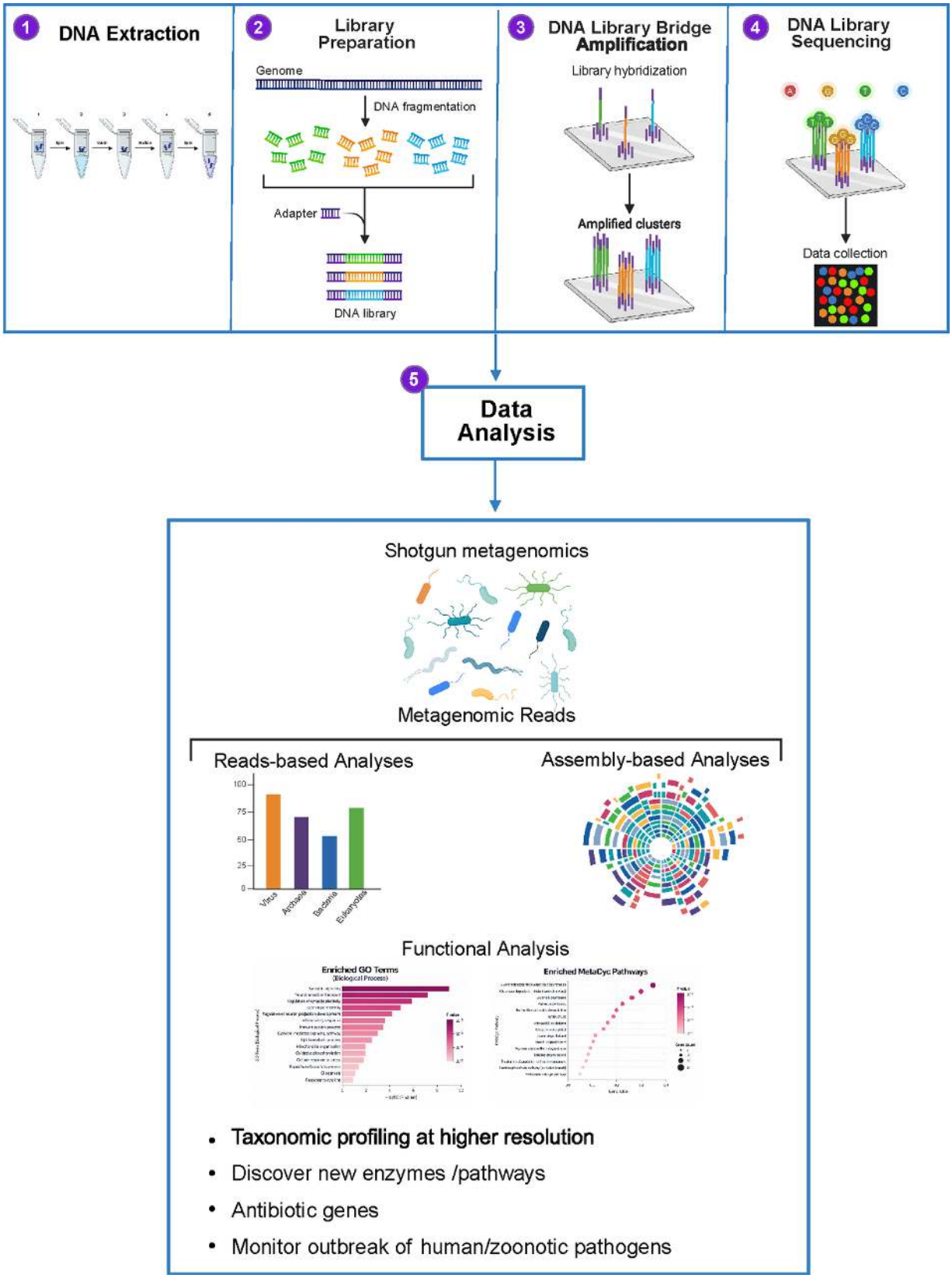


Fig. 2 Analysis workflow for shotgun metagenomic sequencing

2 Materials

2.1 Collection and Storage

- Clean, sterile microcentrifuge tubes or DNA-free cryovials, pre-labeled or readily available commercial vials with nucleic acid stabilizing buffer (Transnetyx).
- Stool Nucleic Acid Collection and Preservation System (Norgen Biotek, Catalog No. SKU 63700).
- Sterile forceps or sterile spatulas.
- Powder-free gloves.
- Clean empty collection cage or sterile holding container.
- Dry ice and/or liquid nitrogen for rapid freezing.
- 80 °C freezer.
- Ice bucket or chilled rack.
- Cage cards/sample sheet for metadata.
- Disinfectant for workspace.

2.2 DNA Extraction, Library Prep, and Shotgun Sequencing

- American Type Culture Collection (ATCC, Catalog No. MSA-2003).
- WHO International Reference Reagent for Gut Microbiome DNA Extraction, whole cell (National Institute for Biological Standards and Control, Product No. 22/210).
- DNeasy 96 PowerSoil Pro QIAcube HT (Qiagen, Catalog no. 47021).
- 10 Strain Even Mix Genomic Material (ATCC, Catalog No. MSA-1000).
- ZymoBIOMICS HMW DNA Standard (ZymoBIOMICS, Catalog No. D6322).
- KAPA HyperPlus Kit (Roche, Catalog No. 50-196-5232).
- NanoDrop Ultra spectrophotometer (Thermo Fisher Scientific, Catalog No. ND-ULTRA-GL).
- Illumina NextSeq 2000.

2.3 Microbiome Analysis

- R.

- Optional: sterile weighing boat, sterile collection paper, or Parafilm.
- Optional controls: unopened tube blank, field blank, extraction blank.

---

3 Research Methods

Plan sampling consistency. Collect all samples at the same time of day whenever possible, as fecal microbiome composition can vary with circadian cycle, feeding state, and recent cage activity [8–10]. To avoid cross-contamination, work in a clean area and wear fresh gloves. Use one sterile set of forceps per mouse, or sterilize forceps between samples using ethanol and complete drying, or use disposable sterile tools between animals or cages.

3.1 Collection and Storage

1. Pre-label all tubes with mouse ID, date, time, treatment group, and sample number (*see* **Note 1**).
2. Move the mouse to a clean, empty cage or sterile holding container with no bedding for a short collection period, usually 10–30 min. This is especially important if you have frail mice because scruffing techniques can often lead to increased stress and worsening of health outcomes (*see* **Notes 2 and 3**).
3. Wait for freshly produced fecal pellets. Collect only pellets that are newly passed and not contaminated with urine, bedding, fur, or food (*see* **Note 4**).
4. Using sterile forceps, pick up one to three fresh fecal pellets immediately upon excretion, and immediately place fecal pellets in 1 ml of nucleic acid stabilizing buffer (e.g., Stool Nucleic Acid Collection and Preservation System) at a temperature specific to the manufacturer's protocols. For long-term storage, keep in −65 to −80 freezer and make sure the buffer is compatible [11–14]. If field or facility constraints prevent this, keep samples as cold as possible on ice. Many shotgun workflows work well with a small but sufficient pellet mass, often enough to total roughly 25–100

but sufficient pellet mass, often enough to total roughly 25–100 mg if possible, though you should follow the requirements of your extraction method.

5. Transfer pellets immediately into the pre-labeled sterile tube (*see* **Note 5**).
6. Flash-freeze the sample immediately on dry ice or in liquid nitrogen.
7. Return the mouse to its home cage after collection, following your animal protocol.
8. Store samples at −80 °C as soon as possible after collection.
9. Avoid repeated freeze-thaw cycles. Aliquot only if your protocol requires it.
10. Maintain a sample inventory with the date, mouse ID, group, location in freezer, and any deviations during collection.

3.2 DNA Extraction [15]

1. Label all tubes.
2. Prepare Positive Control: Include a whole-cell mixed microbial community with known quantities. Prepare the chosen control per manufacturer's instructions, or to target approximately 5×10^8 cells in an aliquot of 200 µL (*see* **Note 6**). Proceed to the next steps.
3. Prepare Negative Control: Do not add sample and proceed with steps.
4. Take 250–300 mg of stool (roughly two pellets of stool) and 800 µL of Solution CD1 per sample. Seal the plate and vortex. Homogenize and Extract DNA per the DNeasy 96 PowerSoil Pro

5. QIAcube HT extraction kit (Qiagen, Catalog no. 47021) protocol.
6. Quantify and assess purity with a NanoDrop Ultra spectrophotometer (Thermo Fisher Scientific, Catalog No. ND-ULTRA-GL). Check for A260/A280 at 1.8–2.0 and A260/A230 at 2.0.
7. Proceed to DNA library Preparation.

3.3 DNA Library Preparation [15]

1. Prepare positive control: Include a high molecular weight DNA mixed community during library preparation. Prepare the positive control for library preparation following the manufacturer's instructions (*see* **Note 7**).
2. Prepare negative control: Include an equivalent volume of extraction elution buffer as a negative control in library preparation.
3. Using Unique Dual Indexed (UDI) adapters, follow the KAPA HyperPlus library preparation protocol.
4. Quantify the DNA of the generated libraries.
5. Pool the libraries based on their concentrations, the number of samples, and the required sequencing depths and sequencer.

3.4 DNA Library Shotgun Sequencing [15]

1. Sequence the pooled libraries using an Illumina NextSeq 2000 or NovaSeq sequencer. Target a minimum depth of two million 2 × 150 bp read pairs.
2. Demultiplex the sequencing run into FASTQ files for each sample.

3.5 Classifying Shotgun Metagenomic Results [15]

1. Classify the raw FASTQ sequence files using a *k*-mer-based classification approach, coupled with a curated database containing as many relevant microbial taxa as possible. The database or workflow should also include a filtering step for host DNA (*see* **Note 8**).

3.6 Limitations

The most significant limitation of this protocol is the cost of shotgun sequencing. 16S protocols can be leveraged to fit this protocol but should still be adapted to fit the same collection, storage, and analysis methods [16].

3.7 Analytical Approaches

Data Considerations

A common mistake is treating shotgun *k*-mer or read counts, 16S amplicon sequence variants (ASVs), or operational taxonomic units (OTUs), as equivalent to gene expression counts. Microbiome data are not continuous. They represent relative proportions within a community, and these proportions reflect a dynamic ecological system influenced by multiple interacting processes. Therefore, the analyses applied to *k*-mer or read counts, ASVs, or OTUs should accurately represent which changes are statistically valid in this highly dynamic, changing environment [17].

Although shotgun sequencing provides high resolution and can classify bacteria, viruses, and fungi while correcting for *k*-mer noise, some studies opt to preprocess data by excluding taxa with relative abundance below 1% in at least half of the samples. This threshold is optional for shotgun sequencing and can be adjusted based on study design and sequencing depth.

Shotgun sequencing also introduces challenges in data reporting. Samples often contain multiple unidentified strains. Many studies report only the highest-resolution classification (typically species

level). While species-level reporting is important due to ongoing taxonomic reclassification, focusing only on species can obscure broader community shifts when many taxa remain unidentified. Reporting both species-level and higher-order taxonomic changes (phylum, class, order, family, genus) provides a more complete view of microbiome alterations.

Report all preprocessing steps, including rationales for threshold cutoff and taxonomy selection.

Descriptive Analyses and Visualization

Initial analyses should include descriptive statistics to visualize data patterns. Before this step, batch effects should be corrected (e.g., using "removebatcheffect" from the limma R package or FAbatch from the bapred R package (R version to correct for sequencing at separate times). False discovery rate (FDR) thresholds are not required for exploratory analyses but should be applied when reporting statistical results [18,19]. Additional methods, like partial redundancy or principal variance component analysis, can be used to remove batch effects [20].

Alpha Diversity: The Shannon index, Simpson index, and observed taxa are all examples of appropriate alpha diversity measures. The interpretation of these measures requires context. Higher diversity as represented by the Shannon and Simpson indices does not necessarily indicate a healthier microbiome. These indices are useful for baseline comparisons (e.g., between genotypes or age groups). However, interventions (diet, exercise, drugs) shift community composition, enriching some taxa while reducing others due to the compositional nature of the data. In these cases, diversity indices should be interpreted cautiously. Observed taxa measurements provide a less biased measure by reflecting the number of taxa that are present rather than their proportions. Conversely, the Shannon Index identifies shifts in microbial proportions that may be biologically meaningful. However, the Shannon Index must be analyzed in conjunction with other metrics

like Observed taxa to establish relevance. Alpha diversity is typically shown using box plots or bar plots with SD or SEM [21,22].

Beta Diversity: Beta diversity assesses differences between groups. Principal coordinates analysis (PCoA) is preferred over principal component analysis (PCA) because microbiome data are not linear. PCoA supports non-Euclidean distance metrics and better captures community differences that are not linear and parametric. These analyses can be performed using packages such as QIIME2. It is also important to evaluate clustering by potential confounders (e.g., cage or litter effects) to identify bias. These data are often visualized with PCoA plots, often including ellipses or centroids to indicate clustering [21–23].

PERMANOVA: Microbiome composition varies widely, even within the same experimental group. Rather than comparing mean abundances of individual taxa, PERMANOVA provides a non-parametric method to test differences in community composition between groups and evaluates overall community structure, making it more appropriate for compositional data than traditional ANOVAs [24].

Relative Abundance: Relative abundance plots summarize community composition but do not indicate statistical significance. They should be treated as descriptive visualizations using stacked bar plots, with formal statistical tests used to identify significant differences [22].

Quantifying and Visualizing Significant Differences

ALDEx2 and ANCOM-BC: Identifying taxa that differ significantly between groups is challenging due to the compositional format of microbiome data. Methods relying on relative abundance can introduce negative correlation bias: an increase in one taxon forces a reduction in other bacteria, even if the bacteria are unaffected biologically. ANCOM-BC addresses this by evaluating log-ratio relationships between taxa, identifying those that differ relative to others in the community. ANCOM-BC extends this approach by correcting for sample and processing biases (e.g., extraction or amplification). If batch effects

were not corrected prior to analyses, one can correct for batch effects by including batches as covariates. However, ANCOM-BC requires larger sample sizes to model relationships effectively. For small studies (<10 samples per group), ALDEx2 may be more appropriate but may be less powerful for rare taxa [25]. It uses log-ratio transformations and Monte Carlo sampling to estimate differential abundance but does not model inter-species relationships directly. Method selection should be justified and clearly reported in these contexts [24,25].

However, when sample size is limited, restricting conclusions to descriptive analyses is often more appropriate and reduces overinterpretation.

Visualization of Significant Results

Common approaches include volcano plots and log2 fold-change plots to show differentially expressed bacteria, while radial phylogenetic trees and Sankey diagrams are used to depict phylogenetic differences of these bacteria. Pie charts can also be used to represent microbial proportions but should be interpreted cautiously, as they can misrepresent the microbiome by excluding unclassified or low-abundance taxa. A brief, detailed description of the pie chart's utility is important for publication [22].

Functional Analyses

There is increasing interest in linking microbial composition to functional outcomes. While experimental validation remains the gold standard, bioinformatic tools such as HUMAnN3 provide an alternative for functional profiling. HUMAnN3 [26] analyzes whole-genome shotgun sequencing data to infer pathway-level changes. After taxonomic classification (via MetaPhlAn or One Codex) [27], microbial gene content is mapped to reference databases to identify functional pathways. Outputs include gene family profiles, pathway abundance, and pathway coverage. Gene family profiles are useful for examining specific functions, pathway abundance reflects overall metabolic activity, and pathway coverage indicates pathway presence within

samples. Incorporating functional analysis improves biological interpretation of microbiome data. Other examples of functional analyses include using databases like the National Database of Antibiotic Resistant Organisms (NDARO) [28] to build a profile of antibiotic-resistant bacteria and pathogenic bacteria to assess gut health and MGnify [29], which allows one to look up KEGG pathways related to the significantly enriched bacteria.

---

4 Notes

1. **Record collection metadata carefully**. For downstream statistical analysis, metadata matters as much as the sequence data.
2. **Collect from more than one mouse in different cages whenever possible**. Pooling mice from the same cage obscures animal-to-animal variability unless the study design specifically requires it [14]. Use fresh pellets rather than old cage feces. Old pellets accumulate environmental contaminants and can distort the profile.
3. **Time of collection matters**. To control circadian rhythm effects, collect samples at a consistent time of day, preferably early in the morning just as the dark cycle ends [8,10]. Young animals are still developing their microbiomes and therefore, in animals that are less than 3 months old, being consistent with timing is critical. The timing in old animals is equally critical, as their microbiomes and circadian rhythms are more sensitive to environmental and biological changes [3]. For longitudinal studies, refer to literature citing how long an intervention takes to change the microbiome or collect microbiome samples when phenotypic changes are present. One should take a final microbiome sample as close to euthanization as possible to study the change in the microbiome from the beginning to the end of the study without causing increased stress prior to euthanization. Increased stress prior to euthanization disrupts animal behavior and physiology and can

affect downstream molecular processes and sequencing methods [18,23].

4. **Preserving DNA quality for shotgun metagenomics and preventing contamination is paramount**. Do not collect pellets contaminated with urine, as it can reduce bacterial viability. For samples that will be used for. Thus, avoidance of preservative buffers should be practiced unless downstream extraction protocols have been validated for use.
5. **Minimize time outside the freezer**. DNA quality and microbial composition are more stable when the sample is quickly frozen.
6. **Select the appropriate extraction positive control**. When selecting a positive extraction control, ensure the control is compatible with the recommended extraction protocol, as some common controls are not fully compatible. Examples include microbiome standards from American Type Culture Collection (ATCC, Catalog No. MSA-2003), or standards from the WHO International Reference Reagent for Gut Microbiome DNA Extraction whole cell (National Institute for Biological Standards and Control, Product No. 22/210). A suggested minimum of 2×10^7 cells/mL should be used, with extraction of approximately 5×10^8 cells recommended.
7. **Select the appropriate library preparation positive control:** A library preparation positive control is recommended to assess the success of library preparation. A control of high-molecular-weight DNA from mixed microbes can identify biases in library preparation. Examples of controls include 10 Strain Even Mix Genomic Material (ATCC, Catalog No. MSA-1000) and ZymoBIOMICS HMW DNA Standard (ZymoBIOMICS, Catalog No. D6322).
8. **Sequence classification pipelines and database choices**. *k-mer-*

based classification approaches are generally more computationally efficient and faster compared to alignment approaches, while also reaching more specific classifications.

a. *k*-mer-based classification approaches fragment each read into sequences of length *k* bases and use exact mapping to a *k*-mer database to identify the most specific classification for that fragment. The classifications of the various fragments (*k*-mers) in a read are then aggregated to identify the most specific and non-conflicting taxon for the entire read.
b. For each taxon identified in the sample, the number of reads that are classified as that taxon is summed.
c. Classification pipelines such as the One Codex workflow [15] include additional filtering steps of *k*-mers, reads, or taxa to reduce possible false positives due to sequencing artifacts.
d. Additional steps, including abundance estimates, can be included. For instance, One Codex estimates abundances of species in a sample based on approximate genome sizes and depths of coverage.
e. When choosing or building a database, you will need to include as many relevant genomes as possible. Including a host genome in the database allows filtering of sequences that may have come from the host. For instance, the One Codex Database includes both the mouse and human genomes, and approximately 148K microbial genomes, including 1.6K mouse gut Metagenome-Assembled Genomes (MAGs) [15] for improved sensitivity of mouse microbes.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Bosco N, Noti M (2021) The aging gut microbiome and its impact on host immunity. Genes Immun 22(5–6):289–303 [https://doi.org/10.1038/s41435-021-00126-8][PubMed][PubMedCentral]

2. Patel BA et al. (2017) The TNF-α antagonist etanercept reverses age-related decreases in colonic SERT expression and fecal output in mice. Sci Rep 7: 42754 [https://doi.org/10.1038/srep42754][PubMed][PubMedCentral]

3. Martino C et al. (2022) Microbiota succession throughout life from the cradle to the grave. Nat Rev Microbiol 20(12):707–720 [https://doi.org/10.1038/s41579-022-00768-z][PubMed][PubMedCentral]

4. Kirkwood TBL (2015) Deciphering death: a commentary on Gompertz (1825) 'On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies'. Philos Trans R Soc B Biol Sci 370(1666), 20140379. [https://doi.org/10.1098/rstb.2014.0379]

5. Ghosh TS, et al. (2020) Adjusting for age improves identification of gut microbiome alterations in multiple diseases. eLife 9, e50240 [https://doi.org/10.7554/eLife.50240][PubMed][PubMedCentral]

6. Ranjan R et al. (2016) Analysis of the microbiome: advantages of whole-genome shotgun versus 16S amplicon sequencing. Biochem Biophys Res Commun 469(4):967–977 [https://doi.org/10.1016/j.bbrc.2015.12.083][PubMed]

7. Treichel NS, et al. (2026) Benchmarking of shotgun sequencing depth reveals the potential and limitations of shallow metagenomics and strain-level analysis. Nat Microbiol 11(5):1233–1244. [https://doi.org/10.1038/s41564-026-02334-2][PubMed][PubMedCentral]

8. Thaiss CA et al. (2016) Microbiota diurnal rhythmicity programs host transcriptome oscillations. Cell 167(6):1495–1510 [https://doi.org/10.1016/j.cell.2016.11.003][PubMed]

9. Zhang M et al. (2025) Interactions between Gut Microbiota, Host Circadian Rhythms, and Metabolic diseases. Adv Nutr 16(6):100416 [https://doi.org/10.1016/j.advnut.2025.100416][PubMed][PubMedCentral]

10. Zhang Y et al. (2023) The microbiome stabilizes circadian rhythms in the gut. Proc Natl Acad Sci U S A 120(5):e2217532120 [https://doi.org/10.1073/pnas.2217532120][PubMed][PubMedCentral]

11. Fouhy F et al. (2015) The effects of freezing on fecal microbiota as determined using MiSeq sequencing and culture-based investigations. PLoS One 10(3):e0119355 [https://doi.org/10.1371/journal.pone.0119355][PubMed][PubMedCentral]

12. Wu GD et al. (2010) Sampling and pyrosequencing methods for characterizing bacterial

communities in the human gut using 16S sequence tags. BMC Microbiol 10:206 [https://doi.org/10.1186/1471-2180-10-206][PubMed][PubMedCentral]

13. Vandeputte D et al. (2017) Practical considerations for large-scale gut microbiome studies. FEMS Microbiol Rev 41(1):S154–S167 [https://doi.org/10.1093/femsre/fux027][PubMed][PubMedCentral]

14. Kim D et al. (2017) Optimizing methods and dodging pitfalls in microbiome research. Microbiome 5(1):52 [https://doi.org/10.1186/s40168-017-0267-5][PubMed][PubMedCentral]

15. Lynch D (2025) Transnetyx sequencing and analysis methods. https://docs.onecodex.com/en/articles/6891294-transnetyx-sequencing-and-analysis-methods

16. Love CJ et al. (2022) Microbiota DNA isolation, 16S rRNA amplicon sequencing, and bioinformatic analysis for bacterial microbiome profiling of rodent fecal samples. STAR Protoc 3(4):101772 [https://doi.org/10.1016/j.xpro.2022.101772][PubMed][PubMedCentral]

17. Gloor GB et al. (2017) Microbiome datasets are compositional and this is not optional. Front Microbiol 8:2224 [https://doi.org/10.3389/fmicb.2017.02224][PubMed][PubMedCentral]

18. Zhou R et al. (2023) Data pre-processing for analyzing microbiome data: a mini review. Comput Struct Biotechnol J 21:4804–4815 [https://doi.org/10.1016/j.csbj.2023.10.001][PubMed][PubMedCentral]

19. Kim M et al. (2026) Estropausal gut microbiota transplant improves measures of ovarian function in adult mice. Nat Aging 6(3):682–702 [https://doi.org/10.1038/s43587-026-01069-3][PubMed][PubMedCentral]

20. Wang Y, LêCao KA (2020) Managing batch effects in microbiome data. Brief Bioinform 21(6):1954–1970 [https://doi.org/10.1093/bib/bbz105][PubMed]

21. Kers JG, Saccenti E (2021) The power of microbiome studies: some considerations on which alpha and beta metrics to use and how to report results. Front Microbiol 12:796025 [https://doi.org/10.3389/fmicb.2021.796025][PubMed]

22. Peeters J et al. (2021) Exploring the microbiome analysis and visualization landscape. Front Bioinform 1:774631 [https://doi.org/10.3389/fbinf.2021.774631][PubMed][PubMedCentral]

23. Armstrong G et al. (2022) Applications and comparison of dimensionality reduction methods for microbiome data. Front Bioinform 2:821861 [https://doi.org/10.3389/fbinf.2022.821861][PubMed][PubMedCentral]

24. Kelly BJ et al. (2015) Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA. Bioinformatics 31(15):2461–2468 [https://doi.org/10.1093/bioinformatics/btv183][PubMed][PubMedCentral]

25. Yang L, Chen J (2022) A comprehensive evaluation of microbial differential abundance analysis methods: current status and potential solutions. Microbiome 10(1):130 [https://doi.org/10.1186/s40168-022-01320-0][PubMed][PubMedCentral]

26. Beghini F, et al. (2021) Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. eLife 10, e65088. [https://doi.org/10.7554/eLife.65088][PubMed][PubMedCentral]

27. Minot SS, Krumm N, Greenfield NB (2015) One Codex: a sensitive and accurate data platform for genomic microbial identification. bioRxiv:027607

28. Walker MK, et al. (2026) Survival trends in patients with difficult-to-treat, antibiotic-resistant, Gram-negative infections in the era of next-generation antibiotics in the USA: a retrospective cohort study. Lancet Infect Dis 26(7):659–670 [https://doi.org/10.1016/S1473-3099(26)00020-4][PubMed][PubMedCentral]

29. Mitchell AL, et al. (2020) MGnify: the microbiome analysis resource in 2020. Nucleic Acids Res 48(D1):D570–D578 [PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_12

12. Massively Parallel, Ultra-accurate RNA Sequencing with CircSeq

Claire S. Chung[1], Jean-François Goût[2] and Marc Vermulst[1] ✉
([1]) University of Southern California, Leonard Davis School of Gerontology, Los Angeles, CA 90089, USA
([2]) Mississippi State University, Department of Biological Sciences, Starkville, MS 39762, USA

✉ **Marc Vermulst**
Email: vermulst@usc.edu

Abstract
Accurate transcription of the human genome is essential for cellular health. However, it has been difficult to design tools that can accurately measure the fidelity of transcription in cellular systems. Here, we describe a protocol for CircSeq, a high-resolution next-generation sequencing method capable of detecting transcription errors in any transcript of any organism. Using rolling-circle reverse transcription, CircSeq generates multiple concatenated cDNA copies from a single circularized RNA fragment, allowing true transcription errors to be distinguished from random artifacts that arise during library preparation or sequencing. Using commercially available reagents and adaptable to diverse organisms and transcript types, CircSeq provides a robust and rigorous approach to quantifying the fidelity of whole transcriptomes.

Keywords Transcriptional mutagenesis – Molecular fidelity – RNA polymerases – Consensus sequencing – Mutation detection – Bioinformatic analysis

1 Introduction

Accurate transcription is required for the faithful expression of genetic information. Recent evidence suggests, though, that the accuracy of transcription declines with age [1, 2], which could contribute to a wide range of age-related diseases, including Alzheimer's disease and amyotrophic lateral sclerosis [2–5]. Traditional methods to measure the error rate of transcription have relied on in vitro systems [6, 7], or on single, artificial reporter constructs [8–11], the results of which may not extend to the rest of the genome. A number of Next-Generation Sequencing (NGS) methods have recently been developed, though, that allow for high-resolution, ultra-accurate sequencing of RNA species that correct for errors introduced during library construction using a consensus-sequencing approach [12–14].

Here, we describe the preparation of CircSeq libraries that can detect transcription errors in any organism, in any transcript, and at any location [3, 15]. The key to CircSeq's success is rolling circle reverse transcription, a step that allows a single RNA molecule to be reverse transcribed multiple times to create a cDNA molecule that consists of multiple concatenated cDNA copies of a template [16, 17] (Fig. 1). This strategy overcomes the inherent limitations imposed by the reverse transcriptases that are typically used for library preparation. These enzymes tend to make a mistake once every 10,000–30,000 bases [16, 17], which masks the mistakes that are made by endogenous RNA polymerases [15]. However, by reverse transcribing the same RNA molecule multiple times, these artificial errors can be discriminated from true transcription errors due to their random distribution, while true transcription errors are present at the same location in every cDNA copy. As a result, CircSeq offers a rigorous estimate of the fidelity of transcription in mRNA, rRNA, tRNA, and any other endogenous

transcripts. This modified NGS RNA-seq library preparation protocol is based on commercially available tools such as the NEBNext Ultra II DNA Library Prep Kit, and can be flexibly adapted to different sample, transcriptome, multiplexing, and sequencing needs.

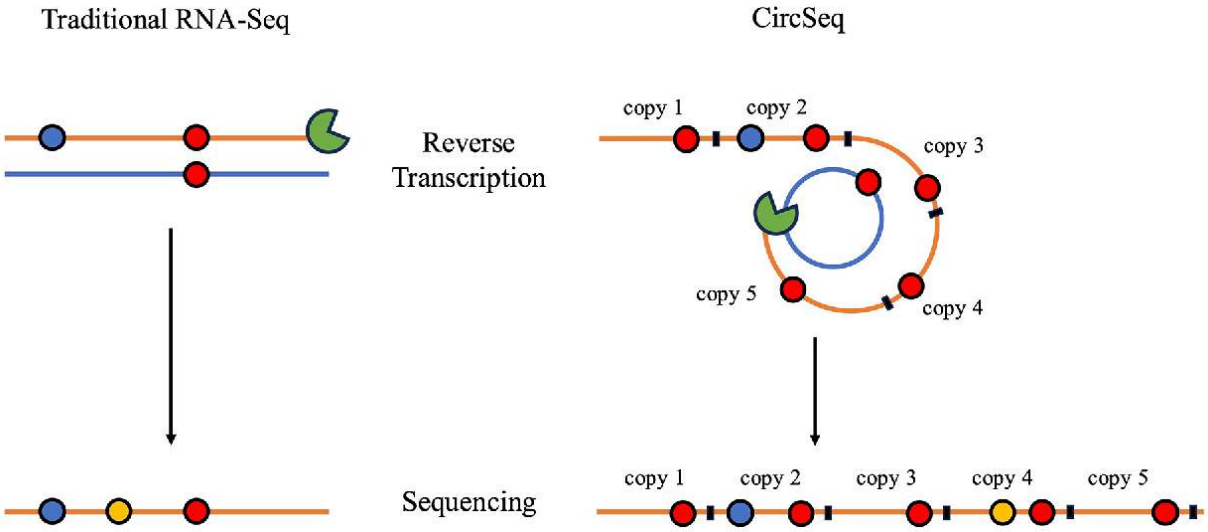


Fig. 1 Concept of the CircSeq assay. Traditional RNA-Seq assays introduce reverse transcription errors (blue circles) during the generation of cDNA libraries (orange lines; the RNA template is in blue) that are indistinguishable from true transcription errors (red circles). In addition, sequencing mistakes (yellow circles) arise during massively parallel sequencing reactions that can be misinterpreted as transcription errors. To circumvent these pitfalls, RNA molecules are ligated to themselves and circularized prior to reverse transcription. These circular molecules can then be reverse transcribed in a rolling circle fashion to generate linear cDNA molecules that consist of multiple concatenated copies of the original RNA template. If a transcription error was present in the template, this error will be present in every concatenated copy, while reverse transcription errors or sequencing mistakes will only be present in one or two copies so that true transcription errors can be distinguished from artifacts

Briefly, transcripts are harvested from cells or tissues, fragmented to generate 50–80 bp molecules, and then circularized with an RNA ligase. Afterward, these circular RNA templates are reverse transcribed with SuperScript III to create a modified RNA-Seq library with tandem repeats of cDNA. This library is then barcoded, finalized, and sequenced using 500–600 bp paired-end reads and analyzed by custom bioinformatic pipelines to distinguish true transcription errors from library construction artifacts.

2 Materials

2.1 RNA Extraction

1. RiboPure RNA Purification Kit (Ambion AM1924) or RiboPure DNA Purification Kit, yeast (AM1926).
2. TRI Reagent.
3. Chloroform.
4. ACS-grade 100% ethanol.
5. Zirconia beads.

2.2 DNase I Treatment

1. TURBO DNA-free Kit (Ambion AM1907).
2. RiboLock RNase Inhibitor.

2.3 RNA Quantity and Quality Check

1. Qubit RNA High Sensitivity Assay Kit (Invitrogen Q32852).
2. Agilent Bioanalyzer or Agilent TapeStation or equivalent.
3. NanoDrop 1000 A or equivalent.

2.4 mRNA Enrichment

1. GenElute mRNA Miniprep Kit (Sigma-Aldrich MRN70).

2.5 RNase III Fragmentation

1. RNase III.

2. Oligo Clean & Concentrator Kit (Zymo D4060).
3. ACS-grade 100% ethanol.
4. Nuclease-free water.
5. Agilent Bioanalyzer or Agilent TapeStation or equivalent.

2.6 RNA Circularization

1. T4 RNA Ligase 1 (ssRNA Ligase).
2. RiboLock RNase Inhibitor.
3. Nuclease-free water.
4. Oligo Clean & Concentrator Kit (Zymo D4060).

2.7 Rolling Circle Reverse Transcription

1. dNTP Mix (10 mM).
2. Random Hexamer (50 μM).
3. Nuclease-free water.
4. Superscript III Reverse Transcriptase.
5. Oligo Clean & Concentrator Kit (Zymo D4060).
6. ACS-grade 100% ethanol.

2.8 Second-Strand cDNA Synthesis

1. NEBNext Ultra II Non-Directional RNA Second Strand Synthesis (NEB 6111S).
2. Oligo Clean & Concentrator Kit (Zymo D4060).
3. ACS-grade 100% ethanol.
4. Nuclease-free water.

2.9 NGS Library Preparation

1. NEBNext Ultra II DNA Library Prep Kit with Sample Purification Beads (NEB E7103S); this kit is also available without sample purification beads (NEB 7645S).
2. SPRIselect DNA Size Selection Reagent (Beckman-Coulter B23317).
3. Tube magnet stand.
4. ACS-grade 80% ethanol, freshly made.
5. Tris-HCl (10 mM).
6. NEBNext Multiplex Oligos for Illumina, Index Primer Set 1, 2, 3, and 4 (NEB E7335S, E7500S, E7710S, E7730S) or NEBNext Multiplex Oligos for Illumina (96 Unique Dual Index Primer Pairs) (NEB E6440S, E7780S).
7. 0.1× TE.
8. Qubit dsDNA Quantification Assay, HS (Invitrogen Q32851).
9. Agilent Bioanalyzer or Agilent TapeStation or equivalent.

---

3 Methods

Before RNA extraction, the lab bench, pipettors, and other relevant equipment were cleaned with 70% ethanol, then with RNase decontaminant. Samples (whole organisms, tissues, or cells) were collected quickly, on ice or in a 4 °C cold room, followed immediately by RNA extraction.

3.1 RNA Extraction Using RiboPure RNA Purification Kit (Ambion AM1924) or RiboPure DNA Purification Kit, Yeast (AM1926)

1. Preheat Elution Buffer to 65 °C
2. Add 100% ethanol to Wash Solutions according to the manufacturer's protocol.
3. Collect samples quickly, on ice or in a 4 °C cold room, according to sample type. See manufacturer's protocol for tips on sample handling.
4. Homogenize samples in TRI reagent according to sample type:
    a. For tissues, use *RiboPure RNA Purification Kit (AM1924)* and follow the amended protocol below:
        i. Pre-fill a screw-cap tube with ice-cold zirconia beads until ⅓ full, and add 160 µL chloroform.
        ii. Collect tissues in microcentrifuge tubes and dounce in 10–20 volume TRI Reagent (i.e., 840 µL Tri Reagent per 50–100 mg tissue). Transfer all to screw-cap tubes with Zirconia beads.
        iii. Vortex vigorously at 4 °C for 20 min, or until tissues are visibly homogenized. Briefly spin down.
        iv. Continue to Step 5.

b. For cell culture, use *RiboPure RNA Purification Kit (AM1924)* and follow the amended protocol below:

   i. Remove cell media and wash at least once with DPBS, then remove DPBS.

   ii. Add 1 mL TRI Reagent for each 100 mm Petri dish.

   iii. Scrape the Petri dish to collect cells, and collect TRI Reagent/cell mixture in a 1.5 mL microcentrifuge tube.

   iv. Vortex briefly and incubate at room temperature for 5 min.

   v. Add 200 µL chloroform per 1 mL homogenate.

   vi. Vortex briefly and incubate at room temperature for 5 min.

   vii. Continue to Step 5.

c. For yeast and other suspension cells: Use *RiboPure RNA Purification Kit, yeast (AM1926)* and follow manufacturer's protocol, with the exception that the RNA is not exposed to temperatures higher than 65 °C (i.e substitute all higher temperature steps with 65 °C incubations). Proceed with DNase I treatment as provided by this kit. Then proceed to Sect. 3.3 for RNA quality and quantity check.

5. Centrifuge at 12,000 × *g* at 4 °C for 10 min: This step separates the mixture into a lower red organic phase containing phenol, a solid, white middle interphase, and a clear, upper aqueous phase that contains RNA.
6. Transfer 400 µL of the top aqueous phase to a new 1.5 mL microcentrifuge tube.

7. Add 200 µL of 100% ethanol; vortex immediately for 5 s.
8. Place a Filter Cartridge into a Collection Tube; transfer sample (~600 µL) into the Filter Cartridge and close the lid.
9. Centrifuge at 12,000 × *g* at room temperature for 30 s.
10. Discard the flow-through and return the Filter Cartridge to the same Collection Tube. The RNA is now bound to the Filter Cartridge.
11. Add 500 µL Wash Solution (with ethanol added) to the Filter Cartridge and close the lid.
12. Centrifuge at 12,000 × *g* at room temperature for 30 s.
13. Discard the flow-through and return the Filter Cartridge to the same Collection Tube.
14. Repeat Steps 11–13 for a second wash.
15. Centrifuge at 12,000 × *g* at room temperature for 1 min to remove residual Wash Solution.
16. Transfer the Filter Cartridge to a new Collection Tube, avoiding contact with the Filter Cartridge.
17. Add 100 µL of pre-warmed Elution Buffer (65 °C) to the Filter Cartridge and close the lid.
18. Incubate at room temperature for 2 min.
19. Centrifuge at 12,000 × *g* at room temperature for 30 s to collect RNA into the collection tube.
20. Proceed immediately to DNase I treatment.

3.2 DNase I Treatment

1. Assemble the following reaction at room temperature:

Table

Mix gently by pipetting or flicking the tube, then briefly spin down. Do not vortex.

2. Incubate reaction at 37 °C for 30 min.
3. Vortex to resuspend DNase Inactivation Reagent, then add 11 µL to sample.
4. Vortex to mix, then incubate at room temperature for 5 min.
5. Centrifuge at 16,000 × *g* at room temperature for 3 min to pellet DNase Inactivation Reagent.
6. Transfer the RNA (supernatant) to a new tube, careful not to disturb the DNase Inactivation Reagent pellet.
7. Immediately place the RNA sample on ice. Alternatively, samples can be stored at −80 °C (*see* **Note 1**).

3.3 Check RNA Quality and Quantity

1. RNA concentration can be measured via Qubit RNA High Sensitivity Assay Kit (Invitrogen Q32852), according to the manufacturer's protocol.
2. RNA quality can be assessed via Agilent 2100 Bioanalyzer or Agilent TapeStation, according to the manufacturer's protocol (Fig. 2).

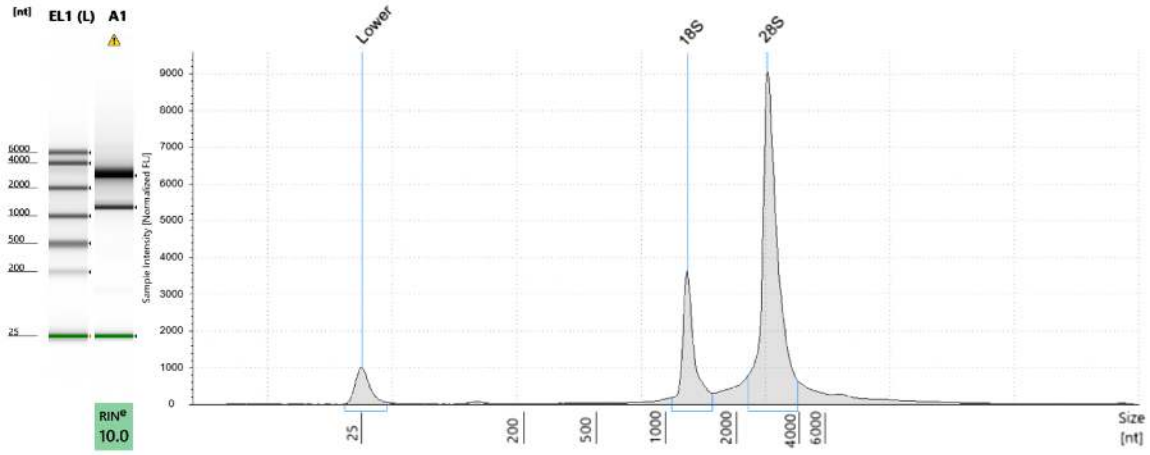


Fig. 2 TapeStation gel and trace of extracted total RNA sample, post-DNase digest. RIN (RNA Integrity Number) of total RNA can be calculated based on the 28S and 18S rRNA peaks to indicate potential RNA degradation. Ideally, CircSeq libraries should be prepared with total RNA with RIN > 9.0. RNA with RIN > 7.0 is acceptable, albeit with potential artifacts in error measurement due to high degradation of the sample. RNA with RIN < 7.0 is not recommended for CircSeq

3. RNA quality and quantity can also be checked by NanoDrop 1000 A (or equivalent spectrophotometer). However, any spectrophotometer-based measurements can only provide an estimate, as DNA contaminants will also absorb at the same wavelength of 260 nm.

3.4 mRNA Enrichment Using GenElute mRNA Miniprep Kit (Sigma-Aldrich MRN70)

1. Purify mRNA from total RNA sample (collected from Sect. 3.2) according to manufacturer's protocol, again with the exception that the RNA was not exposed to temperatures higher than 65 °C and no longer than 2.5 min.
2. Optional: An additional round of purification can be performed for further mRNA enrichment.
3. After enrichment, measure mRNA concentration using Qubit RNA High Sensitivity Assay Kit or NanoDrop 1000 A.

3.5 Fragmentation by RNase III

1. Assemble the following reaction in a PCR tube on ice:

Table

2. Mix thoroughly by pipetting or flicking the tube. Do not vortex.
3. Immediately incubate the reaction at 37 °C in a thermocycler for 27 min, targeting a fragment size of 60–80 bp. Set the thermocycler lid at 40 °C (*see* **Note 2**).
4. After fragmentation, clean up immediately with Oligo Clean & Concentrator Kit (Zymo D4060) as follows:
    a. Add 100% ethanol to Wash Buffer according to manufacturer's instructions.
    b. Transfer the fragmentation reaction to a 1.5 mL microcentrifuge tube.
    c. Add 200 µL Oligo Binding Buffer; vortex and spin down briefly.
    d. Add 800 µL 100% ethanol; vortex and spin down briefly.
    e. Assemble a spin column with a collection tube; transfer 550 µL sample mixture to the spin column.
    f. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.
    g. Transfer the remaining 550 µL sample to the spin column.
    h. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.
    i. Add 750 µL Wash Buffer (with ethanol added) to the spin

column.

j. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.

k. Repeat Steps i-j for a second wash.

l. Centrifuge at 12,000 × *g* at room temperature for 1 min to remove residual Wash Buffer.

m. Transfer spin column to a new 1.5 mL microcentrifuge tube.

n. Add 24 µL nuclease-free water to the spin column; incubate on bench for 1 min.

o. Centrifuge at 12,000 × *g* at room temperature for 1 min to elute fragmented mRNA.

5. Put samples on ice and check mRNA fragment size via Agilent 2100 Bioanalyzer or Agilent TapeStation, according to the manufacturer's protocol (Fig. 3).

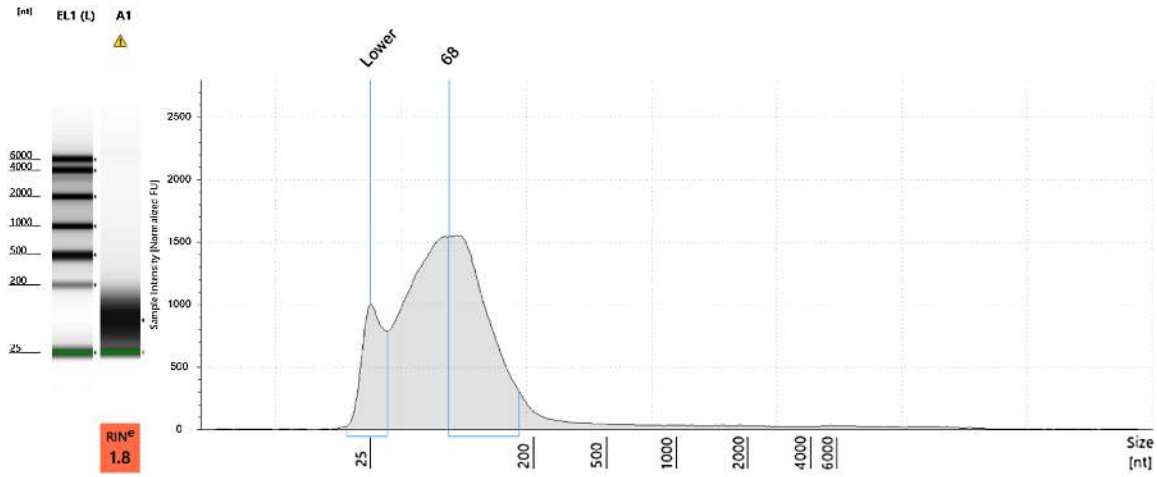


Fig. 3 TapeStation gel and trace of fragmented mRNA sample. Fragmented mRNA should have a peak size between 60 and 80 bp. Note that RIN scores do not apply to mRNA fragments

6. Proceed immediately to RNA circularization.

3.6 RNA Circularization Using T4 RNA Ligase 1

1. Transfer 20 µL sample (from Sect. 3.5) to PCR tubes.
2. Heat-denature samples at 60 °C for 1 min in a thermocycler (set thermocycler lid at 65 °C).
3. Immediately put samples on ice for >2 min.
4. Make Ligation Master Mix as follows, in the order shown:

Table

5. Add 20 µL of Ligation Master Mix to 20 µL sample (from Step 3.6.3). Mix gently by pipetting or flicking the tube. Do not vortex.
6. Incubate the ligation reaction at 25 °C in a thermocycler for 2 h, with the thermocycler lid set at 30 °C (or off). To maximize ligation efficiency, gently mix samples (pipet or flick) every 30 min. Do not vortex.
7. After ligation, clean up immediately with Oligo Clean & Concentrator Kit (Zymo D4060) as follows:
    a. Ensure ethanol has been added to Wash Buffer according to manufacturer's instructions.
    b. Transfer ligation reaction (40 µL) to a 1.5-mL microcentrifuge tube.
    c. Add 10 µL nuclease-free water to the reaction, making the total volume = 50 µL.
    d. Add 100 µL Oligo Binding Buffer; vortex and spin down briefly

Add 100 µL Oligo Binding Buffer; vortex and spin down briefly.

e. Add 400 µL 100% ethanol; vortex and spin down briefly.

f. Assemble a spin column with a collection tube; transfer all 550 µL sample mixture to the spin column.

g. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.

h. Wash with 750 µL Wash Buffer (with ethanol added) and spin the column.

i. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.

j. Centrifuge at 12,000 × *g* at room temperature for 1 min to remove residual Wash Buffer.

k. Transfer the spin column to a new 1.5 mL microcentrifuge tube.

l. Add 20 µL nuclease-free water to the spin column; incubate on bench for 1 min.

m. Centrifuge at 12,000 × *g* at room temperature for 1 min to elute circularized RNA.

8. Immediately proceed to rolling circle reverse transcription.

3.7 Rolling Circle Reverse Transcription with Superscript III

Table

Vortex to mix thoroughly

1. Add 17 µL priming reaction to a new PCR tube.

2. Transfer 9 µL of circularized RNA from Sect. 3.6 to the priming reaction. (Total volume = 26 µL) (*see* **Note 3**).

3. Heat-denature samples at 60 °C for 1 min in a thermocycler, with the thermocycler lid set at 65 °C.

4. Immediately put samples on ice for >2 min.

5. Make the Reverse Transcription Mix as follows:

Table

Mix gently by pipetting or flicking the tube. Do not vortex.

6. Add 14 µL of Reverse Transcription Mix to samples (total volume = 40 µL). Mix gently by pipetting or flicking the tube. Do not vortex.

7. Incubate the rolling-circle reverse transcription reaction in a thermocycler using the following program, with the thermocycler lid set at 45 °C (or off):

Table

8. After reverse transcription, clean up immediately with Oligo Clean & Concentrator Kit (Zymo D4060) as follows:

   a. Ensure ethanol has been added to Wash Buffer according to the manufacturer's instructions.

   b. Transfer ligation reaction (40 µL) to a 1.5-mL microcentrifuge tube.

   c.

Add 10 µL nuclease-free water to the reaction, making the total volume = 50 µL.

d. Add 100 µL Oligo Binding Buffer; vortex and spin down briefly.

e. Add 400 µL 100% ethanol; vortex and spin down briefly.

f. Assemble a spin column with a collection tube; transfer all 550 µL sample mixture to the spin column.

g. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.

h. Wash with 750 µL Wash Buffer (with ethanol added) and spin the column.

i. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.

j. Centrifuge at 12,000 × *g* at room temperature for 1 min to remove residual Wash Buffer.

k. Transfer spin column to a new 1.5 mL microcentrifuge tube.

l. Add 42 µL nuclease-free water to the spin column; incubate on bench for 1 min.

m. Centrifuge at 12,000 × *g* at room temperature for 1 min to elute cDNA.

9. Immediately place cDNA samples on ice and check distribution of single-stranded cDNA via Agilent 2100 Bioanalyzer or Agilent TapeStation, according to the manufacturer's protocol (use the protocol designed for single-stranded DNA or RNA) (Fig. 4).

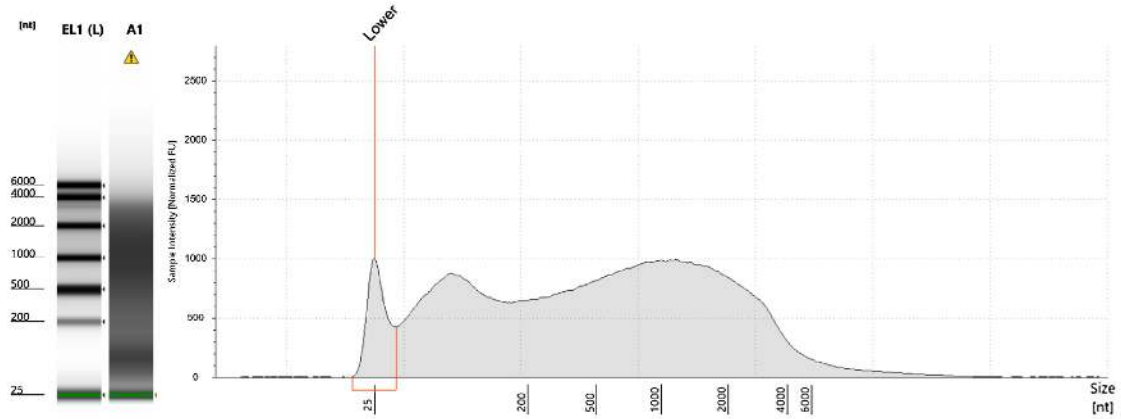


Fig. 4 TapeStation gel and trace of single-stranded cDNA from rolling circle reverse transcription. Expected cDNA size ranges from 300 bp to over 5000 bp, with a peak of around 500–1000 bp. The "bump" around 300 bp shows cDNA reverse transcribed from RNA that has not been circularized

10. Proceed immediately to second-strand cDNA synthesis (*see* **Note 1**).

3.8 Second-Strand cDNA Synthesis

1. Assemble the following Second Strand Synthesis Master Mix on ice:

Table

Mix thoroughly by pipetting or flicking the tube. Do not vortex.

2. Add 42 µL Second Strand Synthesis Master Mix to a new PCR tube.
3. Add 38 µL of single-stranded cDNA (from Sect. 3.7) to the reaction mix. (Total volume = 80 µL.)
4. Mix thoroughly by pipetting or flicking the tube. Do not vortex.
5. Incubate the second-strand synthesis reaction at 16 °C in the

thermocycler for 2.5 h, with the heated lid off.

6. After second-strand synthesis, clean up immediately with Oligo Clean & Concentrator Kit (Zymo D4060) as follows:

   a. Ensure ethanol has been added to Wash Buffer according to the manufacturer's instructions.

   b. Transfer the fragmentation reaction to a 1.5-mL microcentrifuge tube.

   c. Add 20 µL nuclease-free water to the reaction, making the total volume = 100 µL.

   d. Add 200 µL Oligo Binding Buffer; vortex and spin down briefly.

   e. Add 800 µL 100% ethanol; vortex and spin down briefly.

   f. Assemble a spin column with a collection tube; transfer 550 µL sample mixture to the spin column.

   g. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.

   h. Transfer the remaining 550 µL sample to the spin column.

   i. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.

   j. Add 750 µL Wash Buffer (with ethanol added) to the spin column.

   k. Centrifuge at 12,000 × *g* at room temperature for 30 s; discard flow-through.

   l. Repeat Steps i and j for a second wash.

   m. Centrifuge at 12,000 × *g* at room temperature for 1 min to

Centrifuge at 12,000 × *g* at room temperature for 1 min to remove residual Wash Buffer.

n. Transfer spin column to a new 1.5 mL microcentrifuge tube.

o. Add 52 µL nuclease-free water to the spin column; incubate on bench for 1 min.

p. Centrifuge at 12,000 × *g* at room temperature for 1 min to elute double-stranded cDNA.

7. Safe stopping point: Store ds-cDNA at −20 °C for <5 days and minimize freeze-thaw.

3.9 NGS Library Construction Using NEBNext Ultra II DNA Library Prep Kit (NEB E7103S or NEB E7645S)

The following protocol is based on the manufacturer's guidelines for the NEBNext Ultra II DNA Library Prep Kit. For more information, please refer to the NEB manual.

3.10 End-Repair

1. Make the End-Repair Mix as follows:

Table

Mix thoroughly by pipetting or flicking the tube. Do not vortex.

2. Add 10 µL of End-Repair Mix to a new PCR tube.
3. Add 50 µL of double-stranded cDNA (from Sect. 3.8) to the reaction mix. (Total volume = 60 µL)
4. Mix thoroughly by pipetting or flicking the tube. Do not vortex.
5. Incubate the end-repair reaction in a thermocycler using the following program. Set the thermocycler lid to 70 °C:

Table

3.11 Adapter Ligation

1. Leave purification beads/SPRI beads on the bench and let acclimate to room temperature for >30 min. Do not use beads that are not room temperature.
2. Thaw adapter on ice; dilute adapter 1:10 in 10 mM Tris-HCl and leave on ice; do not let adapter warm to room temperature.
3. To each end-repaired sample, add the following:

Table

Mix thoroughly by pipetting or flicking the tube (do not vortex).

4. Incubate adapter ligation reaction at 20 °C in thermocycler for 15 min, with the heated lid off (total volume = 93.5 µL)
5. Add 3 µL USER enzyme to the reaction mixture.
6. Mix thoroughly by pipetting or flicking the tube (do not vortex).
7. Incubate adapter ligation reaction at 37 °C in a thermocycler for 15 min, with thermocycler lid set at 40 °C (total volume = 96.5 µL).
8. After incubation, leave samples on the bench (room temperature), and proceed immediately to double-sided size selection.

3.12 Double-Sided Size Selection for cDNA of 600–800 bp Using NEBNext Sample Purification Beads or SPRIselect Beads (Beckman-Coulter B23317)

CircSeq libraries have been optimized for 600–800 bp to maximize sequence read and efficiency; to select libraries for different sizes, please refer to the NEB manual.

1. Vortex NEBNext Sample Purification Beads or SPRIselect beads vigorously for 30 s.
2. First-round selection: Add 17 µL NEBNext Sample Purification Beads or SPRIselect beads to 96.5 µL adapter-ligated sample. Mix thoroughly by pipetting up and down ten times, or by vortexing, followed by a quick spin (*see* **Note 4**).
3. Incubate at room temperature (on bench) for >5 min.
4. Place the tube on a magnetic stand for 2 min (or until the solution is clear) to separate beads from supernatant.
5. Transfer the supernatant to a new tube. Do not discard the supernatant. Discard beads that contain unwanted large fragments.
6. Second round selection: Add 10 µL NEBNext Sample Purification Beads or SPRIselect beads to the supernatant; mix thoroughly by vortexing, followed by a quick spin or pipetting.
7. Incubate at room temperature (on bench) for >5 min.
8. Place the tube on a magnetic stand for 2 min (or until the solution is clear) to separate beads from supernatant.
9. Carefully remove and discard the supernatant containing unwanted DNA fragments. Do not disturb beads.
10. Leave tubes on the magnetic stand and carefully add 200 µL of freshly made 80% ethanol without disturbing beads. Incubate at room temperature for 30 s, then carefully remove and discard the supernatant without disturbing the beads.

11. Repeat Step 10 for a second 80% ethanol wash.
12. Ensure that all residual ethanol has been removed by using a P10 pipette.
13. Leaving the tube on the magnetic stand with the lid open, air-dry the beads for up to 5 min (*see* **Note 5**).
14. Immediately remove the tube from the magnetic stand and elute DNA from beads with 18 µL Tris-HCl (10 mM). Mix thoroughly by vortexing, followed by a quick spin, or pipetting up and down ten times.
15. Incubate at room temperature (on bench) for >5 min (*see* **Note 6**).
16. Place the tube on a magnetic stand for 2 min (or until the solution is clear) to separate beads from supernatant.
17. Transfer 16 µL supernatant, which contains the size-selected DNA, to a new PCR tube; leave samples on ice and proceed to DNA quantification.

3.13 Quantify Input DNA and Determine PCR Cycles

1. Use 1 µL of size-selected DNA (from Sect. 3.9.3) to measure concentration using Qubit dsDNA Quantification Assay, HS (Invitrogen Q32851) according to manufacturer's protocol.
2. Multiply concentration by 15 µL to determine the total ng input.
3. Determine the number of PCR cycles required using the following guideline (*see* **Note 7**):

Table

3.14 PCR Amplification and Indexing of Adapter-Ligated DNA

1. Assemble the following PCR reaction individually:

   A. Single index, with individual forward and reverse primers (i.e., NEBNext Multiplex Oligos for Illumina, Index Primers Set 1, 2, 3, and 4 (NEB E7335S, E7500S, E7710S, E7730S))

   Table

   B. Dual index, with combined forward and reverse primers [i.e., NEBNext Multiplex Oligos for Illumina (96 Unique Dual Index Primer Pairs) (NEB E6440S, E7780S)]

   Table

2. Mix thoroughly by pipetting or flicking the tube (do not vortex); briefly spin down.
3. Perform PCR amplification and indexing in a thermocycler using the following program:

   Table

   Set the thermocycler lid to 105 °C.
4. Proceed immediately to Cleanup of PCR Reaction.

3.15 Cleanup of PCR Reaction Using NEBNext Sample Purification Beads or SPRIselect Beads (Beckman-Coulter B23317)

1. Vortex NEBNext Sample Purification Beads or SPRIselect beads vigorously for 30 s.
2. Add 45 µL (0.9X) NEBNext Sample Purification Beads or SPRIselect beads to 50 µL PCR reaction. Mix thoroughly by pipetting up and down ten times, or by vortexing, followed by a quick spin (*see* **Note 4**).
3. Incubate at room temperature (on bench) for >5 min.
4. Place the tube on a magnetic stand for 2 min (or until the solution is clear) to separate beads from supernatant.
5. Carefully remove and discard the supernatant containing unwanted DNA fragment; do not disturb beads.
6. Leave tubes on the magnetic stand and carefully add 200 µL of 80% ethanol without disturbing beads. Incubate at room temperature for 30 s, then carefully remove and discard the supernatant (80% ethanol) without disturbing the beads.
7. Repeat Step 6 for a second 80% ethanol wash.
8. Ensure that all residual ethanol has been removed by using a P10 pipette.
9. Leaving the tube on the magnetic stand with the lid open, air-dry the beads for up to 5 min (*see* **Note 5**).
10. Immediately remove the tube from the magnetic stand and elute DNA from beads with 18 µL 0.1X TE. Mix thoroughly by pipetting up and down ten times, or by vortexing, followed by a quick spin.
11.

11. Incubate at room temperature (on bench) for >5 min (*see* **Note 6**).
12. Place tube on magnetic stand for 2 min (or until solution is clear) to separate beads from supernatant.
13. Transfer 16 μL supernatant, which contains the final cDNA

    CircSeq libraries, to a new PCR tube; leave samples on ice and proceed to DNA quantification and quality check.
14. Samples can be stored at −20 °C.

3.16 Quantify Input DNA and Determine PCR Cycles

1. Use 1 μL of final cDNA CircSeq libraries (from Sect. 3.9.6) to measure concentration using Qubit dsDNA Quantification Assay, HS (Invitrogen Q32851) according to manufacturer's protocol (*see* **Note 8**).
2. Quality and size distribution of cDNA CircSeq libraries (from Sect. 3.9.6) can be assessed via Agilent 2100 Bioanalyzer or Agilent TapeStation, according to the manufacturer's protocol (Fig. 5) (*see* **Notes 9 and 10**).

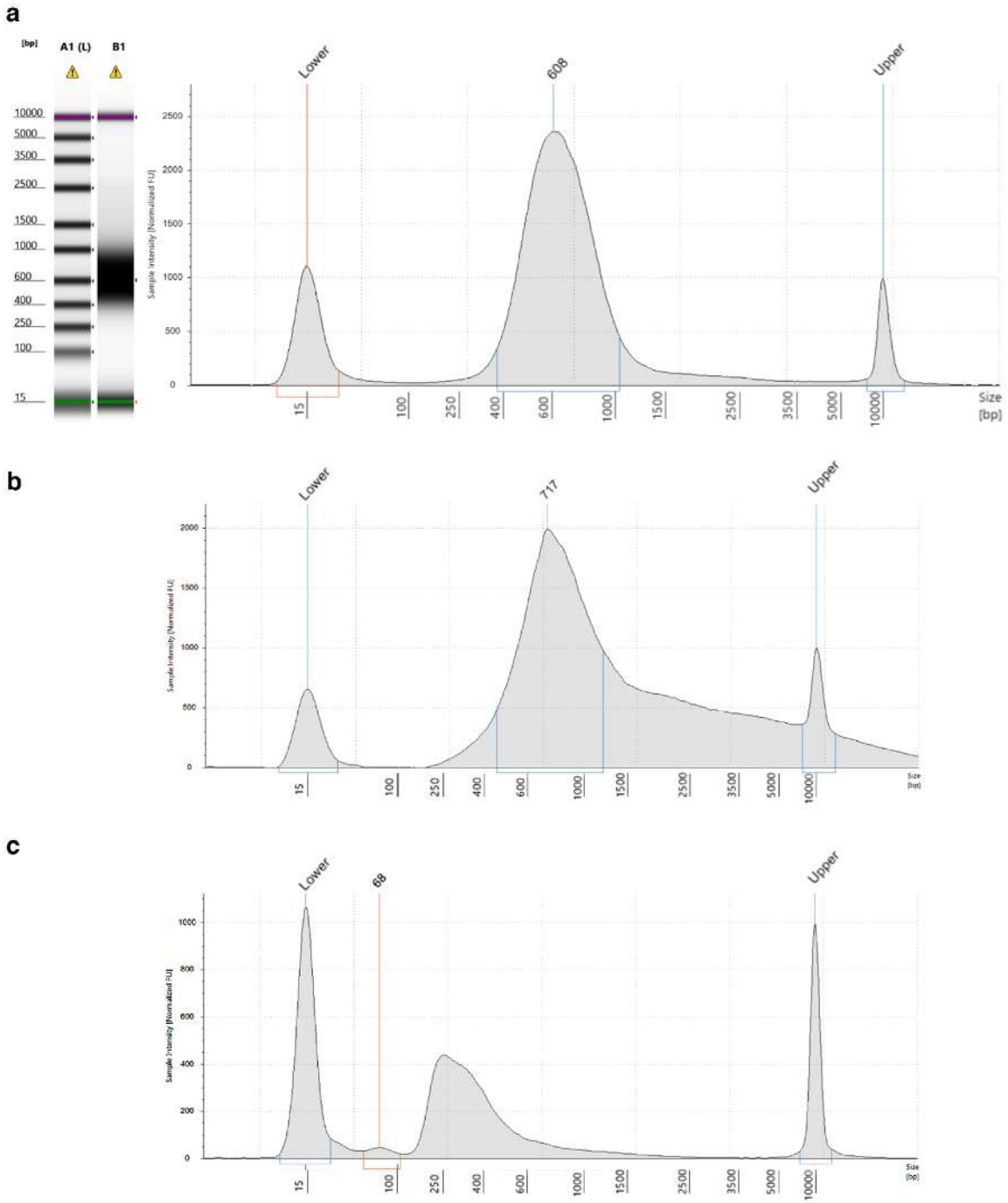


Fig. 5 TapeStation gel and trace of the final cDNA CircSeq libraries. (**a**) Final libraries should have a normal distribution, with peak size between 600 bp and 800 bp. (**b**) A library showing an overamplification "tail," with a large portion of fragments >2000 bp (*see* **Note 8**). (**c**) A library showing a small peak at <100 bp, indicating the presence of adapter dimers (*see* **Note 9**)

3. To normalize and pool libraries, calculate molarity of libraries using concentration measured by Qubit dsDNA Quantification

using concentration measured by Qubit dsDNA Quantification Assay, HS (Step 1), and average size measured from Agilent 2100 Bioanalyzer or Agilent TapeStation (Step 2), as in Fig. 6 (*see* **Note 11**).

$$\frac{(\text{concentration in } ng/\mu L)}{(660\ g/mol \times \text{average library size in } bp)} \times 10^6 = \text{concentration in } nM$$

Fig. 6 Formula for calculating library molarity from concentration and average size. dsDNA NGS libraries are typically normalized and pooled by equal molarity before sequencing. To calculate the molarity of the final CircSeq libraries, divide the cDNA concentration (in *ng/μL*, taken from Qubit measurement) by the average fragment size (in *bp*, taken from TapeStation or Bioanalyzer analysis) and 660 (in *g/mol*, the average molecular weight of one base pair). Finally, multiply by 10^6 to convert to *nM*

3.17 Sequencing on Illumina NGS Platforms

1. Normalized libraries are pooled and sequenced on Illumina's sequencing platforms with >500 cycles
   a. The Illumina NovaSeq 6000: NovaSeq SP Flow Cell v1.5 chemistry, 500-cycle kit (or 250 bp paired-end sequencing).
   b. The Illumina MiSeq: MiSeq Reagent Kit v3, 600-cycle kit (or 300 bp paired-end sequencing).

3.18 Analysis on CircSeq Pipeline

Custom pipeline can be downloaded here: https://github.com/LynchLab/TranscriptErrors

Below is a brief explanation of the bioinformatics used to identify transcription errors.

1. Identify tandem repeats within each read, with the minimum repeat size set at 30 bp, and minimum match between repeats set at 90%. A consensus sequence of the repeat is generated.

2. Using the consensus sequence, identify the RNA template by searching for the longest continuous mapping region; find the starting position of the consensus sequence that corresponds to the 5' end of the RNA template.
3. Reorganize the consensus sequence to start from the 5' end of the original RNA fragment and map it against the genome using TopHat (version 2.1.0 with Bowtie 2.1.0). All non-perfect hits then enter a refining algorithm to search for the location of the 5' end before being mapped again.
4. Each mapped nucleotide is then inspected and checked against five criteria to filter out polymorphic sites and potential RNA-editing events:
    a. It must be part of at least three repeats generated from the original RNA template.
    b. All repeats must make the same base call.
    c. The sum of all quality scores of this base must be >100.
    d. It must be >2 bp away from both ends of the consensus sequence.
    e. Each base must be covered by 200 reads or more, with <1% of these reads supporting a base call different from the reference genome.
5. Each read containing >1 mismatch is filtered through a second refining and mapping algorithm to ensure that errors in calling the position of the 5' end cannot contribute to false positives.
6. Error rate can be calculated as the number of mismatches divided by the total number of bases that passed all quality thresholds.

4 Notes

1. While this is a safe stopping point, it is recommended to avoid sample freeze-thaw and to proceed immediately to the next step. If stopping here, store samples at −80 °C for >5 days and minimize freeze–thaw.
2. RNA fragmentation time may require optimization, as fragmentation efficiency can differ by sample type and enzyme variability. Typically, fragmentation time ranges from 25 to 30 min.
3. The remaining circularized RNA can be stored at −80 °C as a backup sample. If mRNA input (from Sect. 3.4) was low (>300 ng), all circularized RNA samples can be used in lieu of water, i.e., 18 µL circularized RNA (from Sect. 3.6), 4 µL dNTP (10 mM), 4 µL random hexamer (50 ng/ µL).
4. Ensure samples are at the indicated volume prior to adding beads, as size selection relies heavily on volume. For samples less than the indicated volume, adjust by adding nuclease-free water.
5. Do not overdry beads, as it may result in lower DNA yield. Elute samples when beads are still visibly dark brown and glossy, but other visible liquid has evaporated. Beads that are light brown or cracked indicate they have been overdried.
6. Beads that were overdried would be difficult to resuspend. If beads have been overdried, incubate elution for 10 min.
7. Number of PCR cycles is determined by input amount and sample type; thus optimization may be required to generate sufficient library yield for a sequencing run without overamplification and PCR artifacts. This protocol typically yields 10–50 ng input DNA, thereby requiring seven to ten PCR cycles. Note that NEBNext adapters require a minimum of three cycles to add the complete

adapter sequence and indices for library generation.

8. Qubit and TapeStation measure all dsDNA within a sample, regardless of whether samples are barcoded and/or adapter-ligated NGS libraries. Alternatively, a qPCR test using P5 and P7 primers, such as the KAPA Library Quantification Kit (Roche), provides a comprehensive library test.

9. Libraries with a “tail” of >2000 bp indicate PCR overamplification. Size selection should be repeated to optimize for libraries around 600–800 bp, either by repeating with SPRIselect beads (see manufacturer’s protocol for size selection guidelines), or with digital size selectors like Blue Pippin, which can minimize sample loss.

10. Libraries with small peaks of <100 bp indicate the presence of adapter dimers. Adapter dimers can significantly impact sequencing efficiency and should be selected out by repeating SPRIselect bead clean-up at a 0.9× ratio of final library volume (Sect. 3.9.6).

11. If libraries have low concentration or low molarity: Libraries can be re-amplified with additional PCR cycles, using universal P5 and P7 primers (published by Illumina). To avoid overamplification, start conservatively with three cycles. Recheck concentration and molarity before re-amplification.

Competing Interests

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Chung C, Verheijen BM, Navapanich Z et al (2023) Evolutionary conservation of the fidelity of transcription. Nat Commun 14:1547 [https://doi.org/10.1038/s41467-023-36525-w][PubMed][PubMedCentral]

2. Vermulst M, Denney AS, Lang MJ et al (2015) Transcription errors induce proteotoxic stress and shorten cellular lifespan. Nat Commun 6:8065 [https://doi.org/10.1038/ncomms9065][PubMed][PubMedCentral]

3. Chung CS, Kou Y, Shemtov SJ et al (2024) Transcript errors generate amyloid-like proteins in human cells. Nat Commun 15:8676 [https://doi.org/10.1038/s41467-024-52886-2][PubMed][PubMedCentral]

4. Verheijen BM, Vermulst M (2024) linking environmental genotoxins to Neurodegenerative diseases through transcriptional mutagenesis. Int J Mol Sci 25(21): 11429

5. Verheijen BM, Chung C, Thompson B et al (2024) The cycad genotoxin methylazoxymethanol, linked to Guam ALS/PDC, induces transcriptional mutagenesis. Acta Neuropathol Commun 12:30 [https://doi.org/10.1186/s40478-024-01725-y][PubMed][PubMedCentral]

6. Erie DA, Hajiseyedjavadi O, Young MC et al (1993) Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. Science 262:867–873 [https://doi.org/10.1126/science.8235608][PubMed]

7. Izban MG, Luse DS (1992) The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'-5' direction in the presence of elongation factor SII. Genes & development 6:1342–1356 [https://doi.org/10.1101/gad.6.7.1342]

8. Petrova L, Gran C, Bjoras M, Doetsch PW (2016) Efficient and reliable production of vectors for the study of the repair, mutagenesis, and phenotypic consequences of defined DNA damage lesions in mammalian cells. PLoS One 11:e0158581 [https://doi.org/10.1371/journal.pone.0158581][PubMed][PubMedCentral]

9. Bregeon D, Doddridge ZA, You HJ et al (2003) Transcriptional mutagenesis induced by uracil and 8-oxoguanine in Escherichia coli. Mol Cell 12:959–970 [https://doi.org/10.1016/S1097-2765(03)00360-5][PubMed]

10. Viswanathan A, You HJ, Doetsch PW (1999) Phenotypic change caused by transcriptional bypass of uracil in nondividing cells. Science 284:159–162 [https://doi.org/10.1126/science.284.5411.159][PubMed]

11. Gordon AJ, Halliday JA, Blankschien MD et al (2009) Transcriptional infidelity promotes heritable phenotypic change in a bistable gene network. PLoS Biol 7:e44 [https://doi.org/10.1371/journal.pbio.1000044][PubMed][PubMedCentral]

12. Reid-Bayliss KS, Loeb LA (2017) Accurate RNA consensus sequencing for high-fidelity detection of transcriptional mutagenesis-induced epimutations. Proc Natl Acad Sci USA 114:9415–9420 [https://doi.org/10.1073/pnas.1709166114][PubMed][PubMedCentral]

13. Gout JF, Li W, Fritsch C, Li A, Haroon S, Singh L, Hua D, Fazelinia H, Smith Z, Seeholzer S, Thomas K, Lynch M, Vermulst M (2017) The landscape of transcription errors in eukaryotic

cells. Sci Adv 3:e1701484
[https://doi.org/10.1126/sciadv.1701484][PubMed][PubMedCentral]

14. Gout JF, Thomas WK, Smith Z et al (2013) Large-scale detection of in vivo transcription errors. Proc Natl Acad Sci USA 110:18584–18589
[https://doi.org/10.1073/pnas.1309843110][PubMed][PubMedCentral]

15. Chung C, Verheijen BM, Zhang X et al (2023) The fidelity of transcription in human cells. Proc Natl Acad Sci USA 120:e2210038120
[https://doi.org/10.1073/pnas.2210038120][PubMed][PubMedCentral]

16. Acevedo A, Brodsky L, Andino R (2014) Mutational and fitness landscapes of an RNA virus revealed through population sequencing. Nature 505:686–690
[https://doi.org/10.1038/nature12861][PubMed]

17. Burrill CP, Westesson O, Schulte MB et al (2013) Global RNA structure analysis of poliovirus identifies a conserved RNA structure involved in viral replication and infectivity. J Virol 87:11670–11683
[https://doi.org/10.1128/JVI.01560-13][PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_13

13. RNAscope in the Brain of the Naturally Short-Lived African Turquoise Killifish

Rapheal G. Williams[1], George M. Cardenas[1], Hengyu Zhou[1] and Bérénice A. Benayoun[1, 2, 3, 4, 5] ✉

(1) Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA
(2) Pharmacology and Pharmaceutical Sciences Department, Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California, Los Angeles, CA 90033, USA
(3) Cancer Biology Department, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA
(4) USC Norris Comprehensive Cancer Center, Epigenetics and Gene Regulation, Los Angeles, CA 90033, USA
(5) USC Stem Cell Initiative, Los Angeles, CA 90033, USA

✉ **Bérénice A. Benayoun**
Email: bbenayou@usc.edu

Abstract

Thanks to its naturally compressed lifespan, the African Turquoise killifish (*Nothobranchius furzeri*) has started to emerge as a powerful model to study aging in a vertebrate. Its genome has been recently sequenced, and using single-cell and/or single-nuclei RNA sequencing provides a transcriptomic profile to identify molecular markers of inflammation and aging of many tissues/organs, especially in the brain.

To spatially detect and validate multiple molecular targets, fluorescent in-situ hybridization (FISH), namely RNAscope (ACD), can be employed for a high signal-to-noise ratio, removing the caveats of in-house protocols (weak signal, high background, poor probe specificity)—and allowing spatial study of brain tissue in the absence of well-characterized antibodies in this species. Here, we present a protocol for the preparation and performance of RNAscope for on-slide turquoise killifish brain slices, which allows for accurate, robust, and reliable mRNA detection, visualization, and analysis in the central nervous system tissue using high-resolution and subcellular imaging techniques. Additionally, this protocol makes it possible to seamlessly integrate immunohistochemistry (IHC) with RNAscope to detect mRNA puncta within cells of interest in the killifish brain or to validate antibody specificity.

Keywords African turquoise killifish – Brain – In-situ hybridization – RNAscope – Aging – Fluorescent staining

---

1 Introduction

The African turquoise killifish (*Nothobranchius furzeri*) has emerged as a powerful vertebrate model for studying aging and its associated pathologies due to its exceptionally short lifespan, rapid maturation, and spontaneous development of immune transformations and age-associated cellular and molecular hallmarks [1–4]. Following genome sequencing of this model, its genetic architecture of lifespan appeared similar to long-lived vertebrate species, albeit with differences linked to sex-determining genes and possible influences from non-coding regions or enhancers [5]. Unlike traditional vertebrate models with long lifespans or teleost models with a high regenerative capacity in the central nervous system, the turquoise killifish enables longitudinal and interventional studies of brain aging, neuroinflammation, neurodegeneration, gliosis, immune activation, proteostasis, and cognitive decline within experimentally tractable timeframes [6–9].

Additionally, it has been shown that a candidate longevity drug, resveratrol, can partially rescue some of the changes seen in age-related changes in astrogliosis, cognitive ability, and histological markers [10–12]. Extensive neuroanatomical characterization and the development of detailed brain atlases have further supported its use in region- and cell-type–specific analyses of gene expression in the adult brain [13, 14].

Table 1 All reagents and equipment needed from ACDBio for RNAscope

Transcriptomic studies in the turquoise killifish have revealed widespread age-dependent changes in mRNA abundance across multiple tissues and organs in both an age- and sex-dependent manner [15–17]. In the brain, turquoise killifish exhibit transcriptomic changes that support protein aggregation, similar to properties seen with Alzheimer's, Parkinson's, and Huntington's diseases [3, 18]. However, bulk and single-cell RNA sequencing approaches inherently lack spatial context, limiting their ability to resolve cell-type–specific and subregional expression patterns within intact brain tissue. Spatially resolved RNA detection methods are therefore essential to bridge molecular profiling with neuroanatomical and histological analyses in this model.

RNAscope is a highly sensitive fluorescent in-situ hybridization (FISH) platform that enables single-molecule detection of mRNA transcripts while preserving tissue integrity and morphology (Table 1). Its unique dual-probe design and signal amplification strategy provide high specificity and signal-to-noise ratios, overcoming limitations of traditional ISH techniques that have historically hindered reliable RNA detection in neural tissues [19]. RNAscope and/or immunohistochemistry (IHC) has been successfully applied to insect, mammalian, and teleost brains (neuronal and non-neuronal), including zebrafish, to localize low-abundance transcripts and to study subcellular RNA localization, neuroinflammatory signaling, and cell-type–specific gene expression [20–25].

Applying RNAscope or IHC to the turquoise killifish brain, spinal cord, and other tissues (e.g., epithelium, inner ear) presents both opportunities and technical challenges [26–29]. Studies analyzing cellular and transcriptomic properties in the killifish retina have shown age-related increases in oxidative stress, inflammation, and increases in senescence-associated β-galactosidase levels (compared to longer-lived teleost models) [7, 10, 30–32]. Recent characterization of the turquoise killifish telencephalon using transcriptomics and in-situ hybridization revealed a large diversity of neurogenic niches, similar to the mammalian cortex [31]. However, aging turquoise killifish brains—as in other species—accumulate autofluorescent pigments such as lipofuscin, which can interfere with fluorescent detection and necessitate careful optimization of tissue processing and imaging conditions [11, 27]. In addition, the small size and dense cytoarchitecture of the turquoise killifish brain require precise sectioning, probe penetration, antibody availability and signal amplification strategies to ensure reproducible detection across brain regions, especially in the context of whole brain imaging with light sheet microscopy [33]. Recent methodological advances combining RNAscope with immunohistochemistry and autofluorescence-reduction strategies provide a framework for adapting this technology to aging vertebrate brain tissue, especially when transgenic turquoise killifish become a more prevalent utility [34–36].

In this chapter, we describe an optimized RNAscope protocol for detecting RNA transcripts in sections of the African turquoise killifish brain (Fig. 1). We also provide more detailed instructions for dissection, cryoprotection, brain slicing, an optional immunohistochemistry step for RNA–protein colocalization, and considerations for imaging and quantitative analysis. This method enables high-resolution, spatially defined analysis of gene expression within anatomically defined brain regions and across experimental conditions (e.g., sex, age, and strain), supporting mechanistic studies of brain aging and neurodegeneration in this emerging vertebrate model.

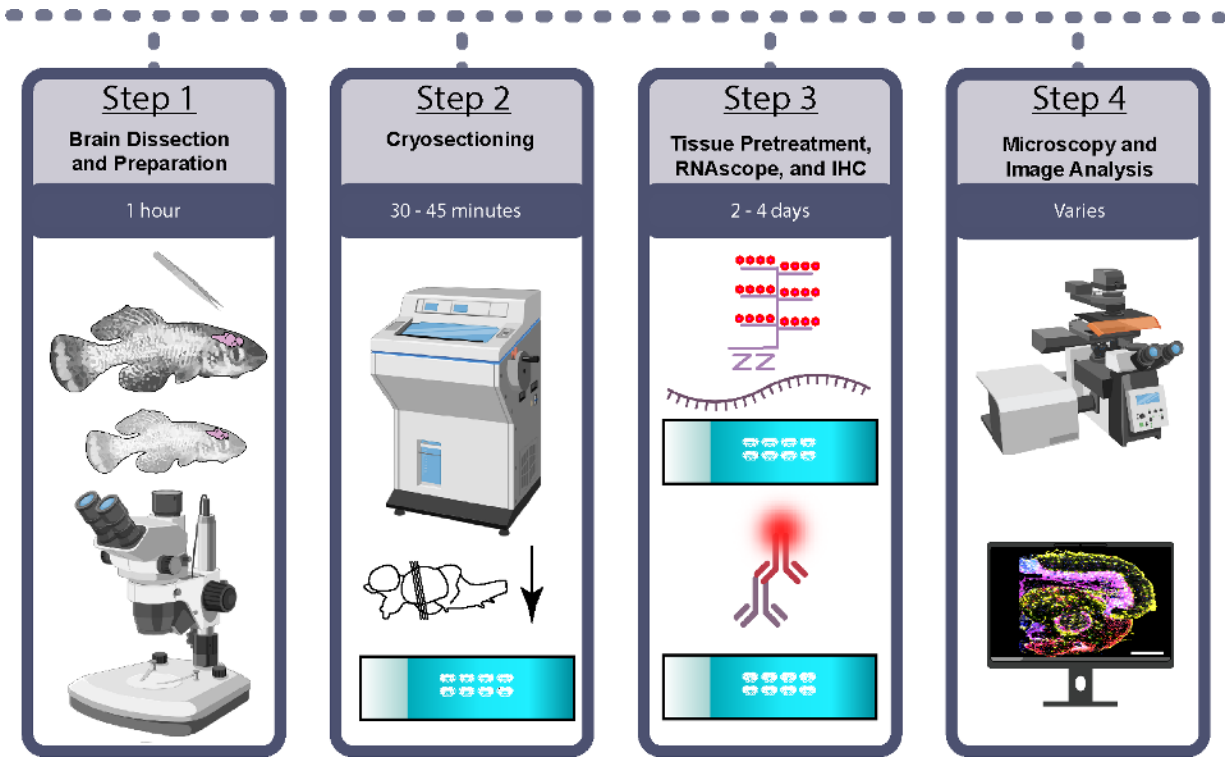


Fig. 1 Flowchart summarizing RNAscope protocol in turquoise killifish brain. These are the general steps of the whole process, from brain dissection to image analysis

2 Materials

1. Tissue preparation and collection

    a. Tricaine methanesulfonate

    b. Light-Duty Tissue Wipes

    c. Tissue-Tek® Cryomold® (REF #4566)

    d. Tissue-Tek® O.C.T. Compound (REF #4583)

    e. Dissecting Microscope: Leica EZ4 Stereo Microscope

    f. Dry Ice

    g. Fine surgical scissors

h. 2× fine-toothed curved surgical forceps

i. 4% Paraformaldehyde in PBS (10 mL per brain)

j. 10%, 20%, and 30% sucrose in PBS

2. Cryosectioning

a. Tissue-Tek® O.C.T. Compound

b. Leica Cryostat CM1860

c. Fine paintbrush

d. Precision tweezer

e. Razor blades

f. Fisherbrand™ Superfrost™ Plus Microscope Slides, 25 × 75 mm

g. Leica 100 µm Anti-Roll Plate

h. DuraEdge Microtome Blades

i. Aluminum foil

3. RNAscope Assay (days 1 and 2)

a. Probes (example probes from Fig. 5)

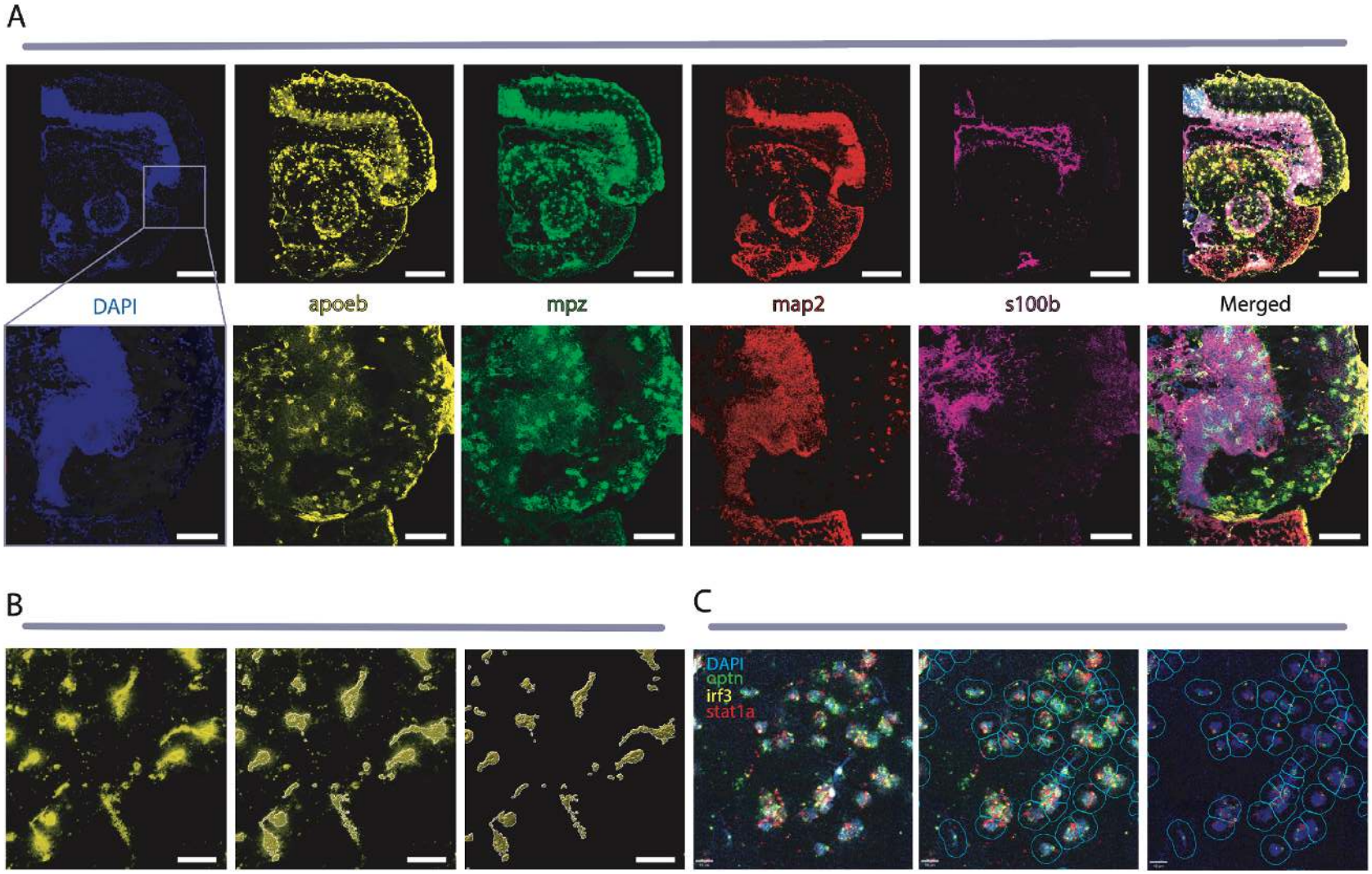

A
DAPI
apoeb
mpz
map2
s100b
Merged
B
C
DAPI
aptn
irf3
stat1a

Fig. 5 Representative image of cell type and zoomed in features. **A,** Example images of using 4-plex RNAScope to image four cell types in the killifish brain at a 1024 × 1024 resolution. The top row shows one hemisphere of a slice containing the optic tectum in a male GRZ killifish brain. DAPI is in blue, microglia (*apoeb*) are depicted in yellow, oligodendrocytes (*mpz*) are depicted in green, neurons (*map2*) are depicted in red, and astrocytes/radial glia (*s100b*) are depicted in magenta. The bottom row shows a zoomed in inset of these images. All laser intensities were set to ten and this image was taken using a bidirectional X for stacked, not serial, laser scanning. Scale bars for the top and bottom row are 300 μm and 100 μm, respectively. **B,** Using the Imaris software, it is possible to do 3D reconstruction of microglia or other cell types using the abundant *apoeb* mRNA Fig. 5 (Continued) signal. On the left is a raw image of the *apoeb* signal, in the center is an overlay of the 3D reconstruction, and on the right is an image showing just the 3D representations. Depending on your image and image processing beforehand, you can use the software's automatic thresholding to create these surfaces. This image shows manual thresholding to avoid the areas of saturated signal. Then you can filter out by volume or other properties. Scale bar is 30 μm. **C,** Using QuPath, it is possible to do 2D analysis on any or all of the Z-stacks in the image. The image shows three transcripts (*optn, irf3, stat1a*) overlayed with DAPI. The left image is the raw signal, the middle image shows the merge of both the cell detection and subcellular (spots) detection results using manual thresholding, and the right image shows the detected components alone. Scale bar is 10 μm

b. Target probe [*N. Furzeri*]—[*apoeb*]—C1: Ready to USE probe for channel 1 (3 mL)—1293241-C1

c. 50× Target probe [*N. Furzeri*]—[*mpz*]—C2: 60 μL—1235131-C2

d. 50× Target probe [*N. Furzeri*]—[*s100b*]—C3: 60 μL – 1235141-C3

e. 50× Target probe [*N. Furzeri*]—[*map2*]—C4: 60 μL— 1272201-C4 (4-plex only)

f. Target probe [*N. Furzeri*]—[*optn*]—C1: Ready to USE probe for channel 1 (3 mL)—1556591-C1

g. 50× Target probe [*N. Furzeri*]–[*irf3*]—C2: 60 μL—1556631-C2

h. 50× Target probe [*N. Furzeri*]—[*stat1a*]—C3: 60 μL— 1556581-C3

i. 3-Plex Negative Control Probe (dapB): 320871

j. 4-Plex Negative Control Probe (dapB): 321831

k. Bioinformatics service fee for probe design (if making new probes)—320269

l. RNAscope Hydrogen Peroxide: 322335 (3 mL)

m. RNAscope Protease III: 322337 (3 mL)

n. RNAscope 10× Target Retrieval solution: 322000 (70 mL × 4 bottles)

o. 4% Paraformaldehyde in PBS (200 mL for post-fixation)

p. ImmEdge Hydrophobic Barrier Pen

q. RNAscope 50× wash buffer (4 × 60 mL bottles): 310091. It is best prepared by warming in a 40 °C for 10 min before diluting with water.

r. RNAscope Multiplex Fluorescent Detection Reagents v2: 323110. Contains 3 mL bottles of each reagent for a 3-plex assay (Amp1, Amp2, Amp3, HRP-C1, HRP-C2, HRP-C3, and 3 bottles of HRP Blocker)

s. RNAscope 4-plex Ancillary kit: 323120 (contains HRP-C4 and HRP blocker)

t. Fluorophores and TSA Buffer (we use a 1:1500 fluorophore to buffer dilution)

    a. TSA Buffer Pack: 322809 (29 mL)

    b.

TSA Vivid Fluorophore 520: 323271

c. TSA Vivid Fluorophore 570: 323272

d. TSA Vivid Fluorophore 650: 323273

e. Akoya Biosciences Opal dye 620: FP1495001KT

f. Akoya Biosciences Opal dye 690: FP1497001KT

u. RNAscope Multiplex v2 DAPI: 323108

v. VWR micro cover glass

w. Prolong Gold antifade mounting medium (P36930; with DAPI, P36931)

x. HybEZ II Hybridization Oven: 321711

y. Humidity Control Tray: 310012

z. ACD EZ-Batch Slide Holder: 321716

aa. ACD EZ-Batch Wash Tray: 321717

ab. HybEZ Humidifying Paper: 310015 (15 sheets)

ac. 1 Tissue-Tek vertical slide rack

ad. 8 Tissue-Tek Staining Dishes

ae. Hamilton Beach 2-Tier Steamer

af. Digital Thermometer

ag. 100% alcohol (EtOH)

ah. Water Bath or Incubator for up to 40 °C for pre-warming probes and wash buffer

ai. Paper towels for cleanup

4. Immunohistochemistry (optional days 3 and 4)

   a. 2 L of 1× TBS with 0.1% Tween (TBST; pH 7.4–8.0).

   b. Invitrogen Normal Goat Serum.

   c. TNB Blocking Buffer (to make 50 mL and aliquot into 4 mL): Roche TSA blocking Reagent (0.25 g into 50 mL of TBST): REF #11096176001. Heat TBST to 50 °C before adding TSA blocking powder to fully dissolve.

   d. Blocking Buffer is 20% goat serum in the TNB buffer (1 mL into 4 mL).

   e. 3 mL of 1:400 Alexafluor (your choice) fluorophore to buffer dilution.

   f. ThermoFisher Pierce Immunostain Enhancer (optional primary or secondary dilution buffer for low expressing protein targets).

5. Microscopy and image analysis

   a. Confocal Images were acquired using a Leica Stellaris 5 (Leica Microsystems, Wetzlar, Germany) confocal microscope equipped with both a White Light Laser and a 405 nm laser. The stand is a DMI8 Leica stand.

   b. The objectives used are a 10× objective lens (10×/0.4 HC PL

APO CS2 506424) and a 63× oil-immersion lens (63×/1.4 HC

PL APO OIL CS2 506350). A 40×/0.8 oil-immersion lens is highly recommended if a 63× lens is not available.

c. Leica Microsystems Type F Immersion Oil.

d. Clay Adams Brand Lens Paper (Soft and Smooth).

e. QuPath-0.5.1-64× Software or any newer version[37]. Retrieved from https://qupath.github.io/

f. Imaris 11.0. (2025). Oxford Instruments. Retrieved from https://imaris.oxinst.com/

---

3 Methods

1. Before beginning the assay, you will need to order your probes. ACDBio customer service or your local sales representative will design the probes based on the sequence of your mRNA of interest (using the Refseq accession number for the transcript of interest is recommended). Additionally, the probes will have to be assigned to up to four channels (*see* **Note 1**).
2. Tissue preparation and collection

   a. Experimental fish are euthanized using 250–300 mg/L of pharmaceutical-grade Tricaine methanesulfonate (MS-222; Syndel) until opercular movement ceased. All procedures should be performed according to protocols approved by your Institutional Animal Care and Use committee (IACUC).

   b. Immediately following euthanasia, gently pat the fish dry with a paper towel to remove excess moisture. Place the fish in a dissection dish under a dissecting microscope.

c. Using fine dissection scissors, decapitate the fish just posterior to the gill region, which serves as the secondary method of euthanasia.

d. Place the decapitated head upside down (ventral side up). Using fine dissection scissors, make a ventral incision through the lower jaw/lip, cutting anterior to posterior to access the brain chamber. Carefully remove surrounding soft tissues above the skull using fine dissection tweezers, exposing the bone surrounding the brain.

e. Using two curved dissection forceps, gently crack open the skull to expose the brain. Carefully peel back bone fragments to fully expose the brain. Proceed slowly to avoid damaging the brain (*see* **Note 2**).

f. To extract the brain, using tweezers, gently push the brain out of the cranial cavity. If necessary, sever the optic nerves or remaining cranial nerve connections to fully release the brain. Carefully grasp the brain with tweezers and transfer it into 1 mL of 4% paraformaldehyde (PFA) in a labeled 1.5 mL tube (*see* **Note 3**).

g. For brain fixation, Store the extracted brains in 4% PFA at 4 °C overnight to ensure proper fixation. To cryoprotect the brain, from day 2 to day 4, incubate the samples at 4 °C in 1 mL of sucrose solution per brain in a stepwise gradient: 10% sucrose on day 2, 20% on day 3, and 30% on day 4. On day 5, brains are embedded in Optimal Cutting Temperature (OCT) compound with the brainstem oriented upward in plastic molds.

h. Fill the labeled Tissue-Tek Cryomold embedding mold with OCT compound. Place the mold directly on a few chunks of dry ice. Allow the OCT to partially solidify—firm enough to hold the tissue in place during positioning, but soft enough to allow manipulation.

i. Keep the mold on top of the dry ice chunks and transfer the fish brain into the partially solidified OCT. Under a dissecting microscope, carefully position the brain using dissection tweezers, with the brainstem facing up. Adjust the brain's orientation as needed (*see* **Note 4**). The semi-solid state of the

OCT will help maintain the positioning and reduce rotation or shifting without damaging the brain(Fig. 2A–C).

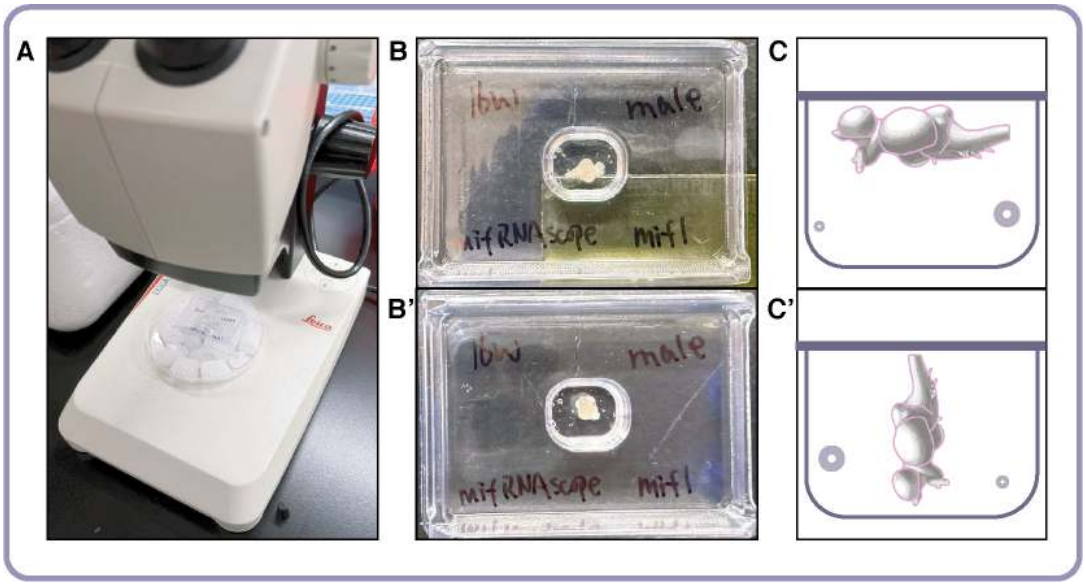


Fig. 2 Schematic of how to embed the killifish brain in the O.C.T. mounting medium. (**A**) This is a picture showing the killifish cryomold on top of a bed of dry ice to stabilize and position the brain. (**B**) An example of the initial placement of the brain inside of the mold in a horizontal orientation. **B'**) The correct placement of the brain with the most rostral part of the brain pointed downward. **(C,** Shown is the flat and incorrect way to place the brain in the mold. **C',** A cartoon of the correct placement of the brain in the mold

j. To completely freeze the sample, once proper positioning is achieved, transfer the mold to a flat surface of dry ice. Place the mold inside a freezer box or insulated container and cover with a lid to fully freeze the sample.

k. After all samples are completely frozen, transfer the molds to a −80 °C freezer for long-term storage. For best results, proceed to sectioning and staining within 6–8 months of brain fixation

to sectioning and staining within 6–8 months of brain fixation.

3. Cryosectioning

   a. Set the Cryostat to −20 °C. Place the sample-containing cryomold(s) inside the cryostat, and cool for at least 30 min.

   b. Remove a sufficiently cooled cryomold and specimen disc onto a tabletop. Extract the OCT-embedded sample onto a Kimwipe by turning over the cryomold and dropping the specimen (Fig. 3A).

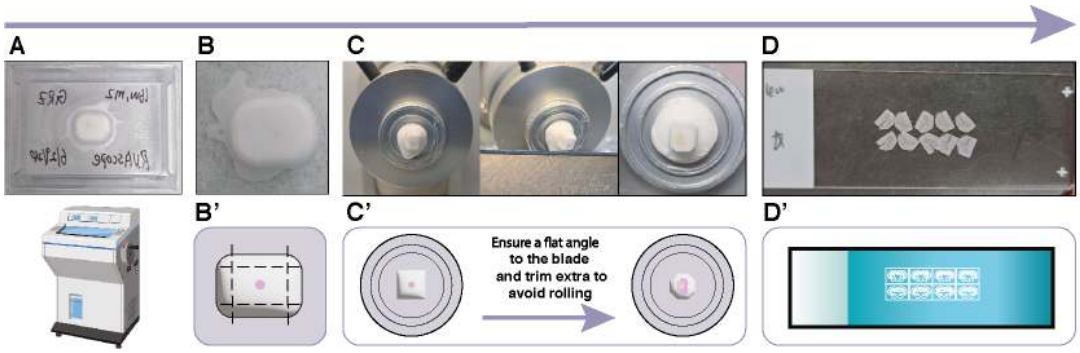


Fig. 3 Slide orientation and cryosectioning steps. **A,** A picture of the mold being flipped over and then (**B**) remove from the mold, keeping the OCT section flipped over for trimming and mounting. (**B'**) The four cuts to make with a razor blade very quickly before using a drop of OCT to secure the tissue to the slicing mount. (**C**) Example images of how the brain should be oriented on the slicing mount and then correct flat angle to have in relation to the microtome blade. Excess OCT is then trimmed off of the corners to avoid excess rolling of the slice. (**C'**) A cartoon of the process. (**D**) image and (**D'**) cartoon showing the way to orient eight to ten slices on the slide, ensuring that there is ample room to draw the hydrophobic barrier on each side

   c. Using a room-temperature razor, swiftly trim excess OCT from the specimen (Fig. 3B).

   d. On the specimen disc, place a drop of OCT to form a sphere about 3–5 mm in diameter. Using slant tweezers, quickly transfer the sample, preserving its orientation. Try to ensure that the sample sits as parallel to the tabletop as possible. During the transfer, do not flip the specimen over so that the most rostral part of the brain is pointed toward you.

e. Place the specimen disc into the cryostat to cool for at least 5–10 min.

f. Prepare the slicing environment. Place necessary tools like a fine paintbrush and tweezers inside the machine for cooling. Prepare labeled microscope slides at the bottom of the slicing stage (*see* **Note 5**).

g. Once a specimen disc is sufficiently cooled, place and fasten it to the cryostat mounting stage (*see* **Note 6**). Use a razor to trim the excess OCT from the perimeter of the sample to avoid extra rolling of the tissue (Fig. 3B').

h. Begin coarse slicing at 20 μm to establish a clean cutting plane. Switch to 15 μm once the desired brain region is reached. For our use, we focused on brain slices that contain the optic tectum—a large and recognizable area of the brain—using a previously published brain anatomical atlas [13]. If the rostral-most parts of the brain are desired (i.e., telencephalon), then it may be necessary to retain early slices and refrain from coarse slicing (Fig. 3C, C').

i. Manipulate specimen slices with your paintbrush and/or forceps to transfer and melt onto labeled microscope slides (*see* **Note 7**). Be careful to avoid agitating the brain tissue (Fig. 3D, D'; *see* **Note 8)**.

4. Tissue pretreatment (day 1)

a. This protocol is modified from the RNAScope User Manual v2 [38].

b. Pick the slides for downstream staining and imaging. For each assay, decide appropriate negative control and/or positive

control slides and label the slides with a pencil.

c. Pre-chill your fixative (4% PFA) for the post-fixation step.

d. Air dry slices for at least 1 h at room temperature or 4 °C. It is important to air dry the slides 1–2 h or the tissue may be lost by boiling off or detaching during the assay, especially during antigen retrieval (*see* **Note 9**).

e. Prepare steamer for antigen retrieval step by filling the steamer with distilled water to the "max" level of streamer.

f. Perform a PBS wash to get OCT off for 3 min, using a wash container using ~220 mL of PBS (*see* **Note 10**).

g. Bake the slices for 30 min at 60 °C in the HybEZ Oven using the "Hold-T" mode. Switch the oven back to its default "RNAScope" mode to get the oven back to 40 °C.

h. For the antigen-retrieval retrieval step, start heating the water and retrieval buffer within the steamer at the end of the baking process (set the steamer timer to 59 min; it should take ~30 min for temperature to reach 99 °C; Fig. 4A).

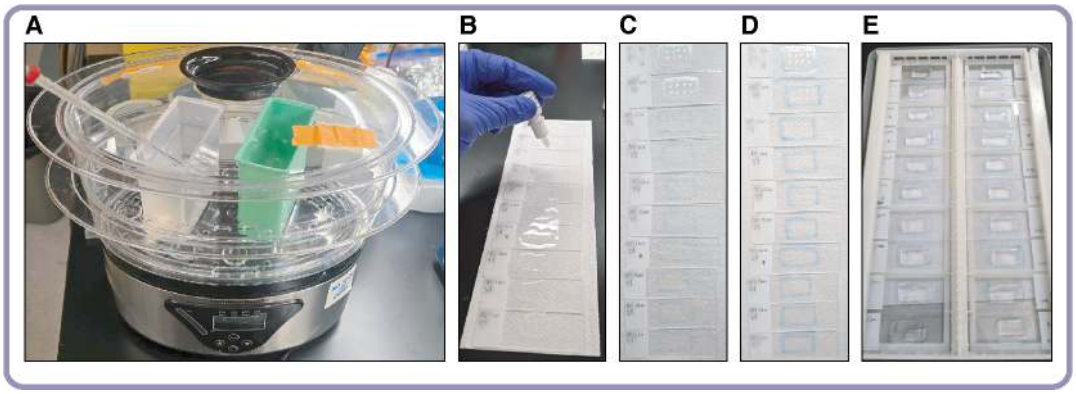


Fig. 4 Antigen retrieval steamer setup and useful tips for running the RNAscope assay. **A,** An image showing the dual-tier steamer setup. Note how the top ventilation holes are taped over except one hole which can fit a glass or digital thermometer to be put inside of the water-filled staining dish. **B,** Using the glare of the light on the slides, it is easy to visualize your tissue for applying any of the reagents, especially when

applying hydrogen peroxide without a hydrophobic barrier. **C,** An example of a hydrophobic barrier that is applied too lightly which risks losing reagent or probes when transferring into the incubation oven. **D,** The right amount of barrier to add which is about four times drawing the barrier. At the top of the image, the lights reflection assists in visualizing the tissue to avoid drawing over it. **E,** Shows the right

amount of liquid that should be added to each sample slide and how the slides are organized into the EZ Batch Slide Holder

i. Immediately following baking, place the slides into the chilled 4% PFA fixative, and incubate them for 15 min at 4 °C.

j. Perform dehydration using an ethanol gradient with four sequential baths as follows: 50% ethanol (110 mL water + 110 mL ethanol) for 5 min. 75% ethanol (50 mL water + 150 mL ethanol) for 5 min. 100% ethanol (~200 mL) for 5 min. 100% ethanol (~200 mL) for 5 min.

k. Apply approximately five to eight drops of hydrogen peroxide (provided in your RNAscope kit) to each of your slides on the bench (Fig. 4B). It is easier to take the slides from the bench if you put a paper towel underneath the slides, but make sure the slides are completely flat to avoid the solution running off.

l. Incubate the slides for 10 min. Remove extra hydrogen peroxide by flicking or tapping each slide. Wash the slides by gently moving the slide holder up and down in distilled water. Repeat washing step with fresh distilled water three to five times.

m. At this point, the steamer should have reached >99 °C to perform the antigen retrieval step.

n. Acclimate slides to the high temperature by placing them in the distilled water bath for 10 s.

o. Place slices in the pre-heated 99 °C retrieval buffer for 15 min. Keep an eye on the water level of the steamer and increase the

steamer timer if necessary to avoid the steamer turning off during the retrieval.

p. Rinse the slides for 15 s in distilled water at room temperature for the slides to cool down.

q. Transfer the slides to another staining dish filled with 100% ethanol for 3 min.

r. Dry slides on the bench for 5 min at room temperature.

s. Shake the hydrophobic pen well to ensure proper delivery of the compound and draw boundaries two to three times to make sure the barrier is thick enough to trap reagents. The border should be drawn as close to the slices as possible (using the light's reflection off the slides can help visualize the thin tissue) but avoid letting the hydrophobic wax touch any of the brain tissue (Fig. 4C, D). It is helpful to keep a paper towel on the side to blot the pen so you can avoid applying too much or too little to the slide at this step.

t. Let dry for 5 min. Although the RNAscope user manual [38] gives the option of drying the slides overnight, we do not recommend doing this, as drying the slides for an extended period can introduce variability in antigenicity, RNA integrity and overall tissue structure.

u. The slides should now be safe to load into the ACD EZ-batch tray with the white part of the slide being clamped down.

v. Add five to six drops of Protease III to each slide, ensuring that the drops completely cover each tissue slice on the section, to the point that the liquid is almost about to cross the hydrophobic barrier. Place the slides into the humidity control tray, cover with the lid, and place into the 40 °C HybEZ oven for

30 min (*see* **Note 11**). Prepare your probes during this step (Fig. 4E).

w. Gently wash the slides in 200 mL of distilled water in the EZ-Batch wash tray for 2 min. Repeat this step once with fresh distilled water for a total of two washes, then proceed to the RNAscope assay (hybridization of probes).

5. RNAscope assay (day 1 and day 2)

a. To hybridize probes, apply 150 µL of your prepared probe cocktail (*see* **Note 12**). Each kit provides 3 mL of diluted probes, so that should be enough for 20 slides to fit the spaces in the EZ-batch tray. Be sure to dispense the probe solution gently to keep it in the hydrophobic barrier and to not damage the tissue. Place the slides in the HybEZ oven at 40 °C for 2 h. During this time prepare 200–250 mL of 5× SSC solution for overnight storage.

b. Following the 2-h incubation period, turn off the HybEZ oven, and pour 200 mL of RNAscope wash buffer into the EZ-Batch wash tray.

c. Wash the slides twice with RNAscope wash buffer for 2 min each at room temperature. After each wash, pour the used buffer into a separate container, and shake the EZ-Batch slide holder to remove extra liquid.

d. Pour the 5× SSC solution into the empty wash tray. Keep the slides in the slide holder and the wash tray (covered) overnight at room temperature.

e. For Signal Amplification You will want to remove the kit from the refrigerator before starting this step, to warm reagents up to room temperature.

f. Before starting this step, turn on the HybEZ oven to the 'RNAscope' setting for 40 °C. To hybridize the first

RNAscope setting for 40 °C. To hybridize the first amplification reagent, AMP 1, dispense enough solution to cover each slide. Carefully place the slides back into the humidity control tray (making sure the paper is damp with distilled water) and incubate the slides for 30 min in the HybEZ oven. Wash the slides twice with RNAscope wash buffer for 2 min each at room temperature.

g. Remove as much liquid as possible by gently shaking the slide holder and then add enough drops of AMP 2 to each slide. Incubate the slides for 30 min in the oven at 40 °C. Wash the slides twice with RNAscope wash buffer for 2 min each at room temperature.

h. Remove as much liquid as possible by gently shaking the slide holder and then add enough drops of AMP 3 to each slide. Incubate the slides for 15 min in the oven at 40 °C. During this step, prepare fluorophores if not done already. We use a 1:1500 dilution, so for a total of 3 mL, prepare 2 µL of fluorophore diluted in 3 mL of the provided TSA buffer. Wash the slides twice with RNAscope wash buffer for 2 min each at room temperature.

i. Developing the horseradish peroxidase signal for each channel must be performed for every probe used in the multiplex design, typically three or four times. If using a four-probe design, then repeat this Step 4 times, using the appropriate fluorophore for each channel. For example, our 4-plex design for brain cell types uses TSA Vivid 520, TSA Vivid 570, Opal dye 620, and Opal dye 690. Thus, for the first channel (C1), TSA Vivid 520 is used for conjugation. Other probes follow the respective order with Opal dye 690 being used on C4.

j. Remove the excess liquid from the wash steps following the AMP 3 hybridization. Add enough drops of HRP-C1 reagent to completely cover each slide. Cover the slides with the tray lid

and place in the HybEZ oven to incubate at 40 °C temperature for 15 min. Wash the slides twice with RNAscope wash buffer for 2 min each at room temperature.

k. Gently shake the slide holder to remove excess liquid and apply 150 µL of the fluorophore reagent assigned to channel 1. Incubate the slides for 30 min at 40 °C temperature in the HybEZ oven. Wash the slides twice with RNAscope wash buffer for 2 min each at room temperature in the EZ-Batch wash tray.

l. Remove as much wash solution as possible and then cover each slide with enough drops of the HRP blocker solution. Incubate these slides for 15 min in the HybEZ oven. Wash the slides twice with RNAscope wash buffer for 2 min each at room temperature in the EZ-Batch wash tray.

m. Stop here and coverslip if only using one probe, otherwise repeat this step to perform the next fluorophore conjugation step.

6. Immunohistochemistry (optional step)

a. When performing IHC alone on killifish brain samples, it is not necessary to perform the baking, post-fixation (optional to preserve cell morphology), dehydration, hydrogen peroxide, and protease steps. Instead, after sufficiently drying the slides and performing the first PBS wash, you can proceed directly to the antigen retrieval step followed by 100% ethanol wash and hydrophobic barrier application (*see* **Note 13**).

b. Immediately following the final wash step for the final channel of your RNAscope experiment, proceed to the blocking step of the IHC protocol. Place and lock slides into the slide holder and place them on the humidity tray. Pipette ~150–200 µL blocking buffer/slide. Incubate at room temperature for 1 h.

c. Wash three times with TBST (200 mL) in the clear washing tray

c. Wash three times with TBST (200 mL) in the clear washing tray (5 min each, with gentle agitation for the first minute).

d. After washing, add ~200 μL of the diluted primary antibody in either Pierce Immunostain Enhancer (if signal is expected to be weak) or the blocking buffer to each slide. Incubate overnight in 4 °C (*see* **Note 14**).

e. Wash three times with TBST (200 mL) in the clear washing tray (5 min each, with gentle agitation for the first minute).

f. After washing, add ~150–200 μL of the diluted secondary antibody in either Pierce Immunostain Enhancer or the blocking buffer to each slide. Incubate at room temperature for 1–2 h (*see* **Note 14**).

g. Wash slices three times with TBST (200 mL) in the clear washing tray for 5 min each, with gentle agitation for the first minute.

h. Finally, use RNAscope DAPI and then coverslip with mounting medium. Alternatively, you can directly proceed to coverslipping if using mounting medium that has DAPI (Prolong Gold antifade reagent with DAPI).

i. To use the RNAscope-supplied DAPI, drop the appropriate amount (five to seven drops) onto each slide and let sit for 30 s. Flick the slide to remove extra DAPI solution and coverslip with three drops of mounting medium (Prolong Gold). For ease of dispensing the viscous mounting medium, cut the tip of the transfer pipette off. Make sure to apply adequate and even pressure to avoid bubbles around the tissue and under the coverslip.

j. Let the covered slides cure overnight in the dark at room temperature.

k. After mounting, gently wipe the extra mounting medium off with

k. After curing, gently wipe the extra mounting medium off with 70% ethanol and Kimwipes.

l. Store the slides in an opaque slide box in 4 °C until ready to image (*see* **Note 15**).

7. Microscopy and image analysis

a. To achieve subcellular resolution in your images at ~200 nm/pixel, ACD bio recommends using a high-magnification objective lens and resolution setting to visually see individual mRNA puncta. It is recommended to use a confocal microscope to lower the chances of out-of-focus light contaminating the imaging. For example, if using a 63× oil-immersion lens, the software's resolution setting for image acquisition should be set at no less than 512 × 512 to achieve the above resolution. Other lenses using 20–40× magnification should use a higher resolution setting such as 1024 × 1024 (40× oil-immersion) or 2048 × 2048 (20×).

b. In this protocol, we use a Leica DMI8 inverted microscope with the Stellaris five confocal system (*see* **Note 16**). To locate the tissue of interest for imaging, we use a 10× objective lens (10×/0.4 HC PL APO CS2 506424). Then we switch to the 63× oil-immersion lens (63×/1.4 HC PL APO OIL CS2 506350) by applying a tiny drop of immersion oil (Type F) and placing the slide face-side-down onto the slide holder. Adjust the focus and continue to the imaging. You may use either the DAPI or the FITC/TRITC setting on the DMI8 control panel. We use this step to find the highest quality tissue of the animal on the slide out of the eight slices usually on a slide.

c. After starting your software, adjust the laser intensity and assign laser wavelengths according to used fluorophores (*see* **Note 17**). If your software does not have the Opal dyes or TSA Vivid dye excitation and emission settings installed, find an installed laser setting that matches your fluorophores the best.

For example, we use FITC, Cy3, Texas Red, and Cy5.5 for each of the fluorophores mentioned above (Fig. 5A).

d. After achieving optimal settings, set your upper and lower boundaries of the desired Z-stack.

e. Image each sample using the "bidirectional X" to save time.

f. We recommend imaging each wavelength sequentially (not in parallel) to reduce crossover artifacts from wavelengths that are too close to each other.

g. To speed up imaging, a resonant scanner with line averaging at ≥2 to reduce noise can be used. This is especially important if acquiring images of the whole brain slice or one whole hemisphere (left or right).

8. Image processing and analysis

a. For image processing, such as using background reduction or applying a Gaussian smoothing filter, you can use either Image J, Imaris, or the LAS-X software that is used on your confocal computer setup.

b. To analyze the images, several softwares can be used (*see* **Note 18**).

c. The open source/free to use QuPath (Quantitative Pathology [37]) is ideal for 2D analysis, including cells and spot quantification (*see* Fig. 5C). Watershed cell segmentation using the DAPI channel is used to define cell boundaries.

d. The commercial Imaris software is recommended for 3D-reconstruction of cell surfaces, especially when combined with other staining techniques such as tissue clearing (*see* **Note 19**; Fig. 5B).

4 Notes

1. Each channel will have either a TSA Vivid or Opal dye assigned to it, so make sure the fluorophores are chosen with the imaging capability in mind, using brighter dyes to label less abundant transcripts. If combining RNAscope with IHC, make sure to choose a fluorophore that does not spectrally overlap with the fluorophores used for the mRNA probes. Negative controls usually are probes targeting bacterial genes such as *DapB*. If only doing two or three probes, prefer fluorophores that do not overlap with the naturally green autofluorescence in killifish brain tissue.
2. During the dissection, it is important to be very gentle when removing the brain as even minor scratches and tissue damage can be visualized under the microscope following the staining assay. Keep sterilized, RNase-free dissection tools to avoid contaminating or degrading samples.
3. Although the user manuals indicate that neutral buffered formalin (NBF) is an acceptable fixative for RNAscope, in our experience, NBF fixation leads to “fuzzy” signal and high background in the African turquoise killifish brain. For best signal-to-noise ratio, we recommend the use of PFA fixation.
4. When freezing the brain inside of the mold, you’ll want to place the brain in the center of the mold as much as possible so that when it is time for slicing, some brain is not cut off during trimming or removal of extra OCT.
5. The instruments should remain as cold as possible. The tweezers, especially, should be very cold and should remain at the bottom of the machine when not in use. Heat conduction from the hands is enough to occasionally warm the tweezers and cause adhesion with the slices.

6. The angle and orientation of the tissue block is important to obtain good quality slices. Having too perpendicular of an angle to the blade can allow tissue to tear. Additionally, if the tissue is tearing when slicing, it may be beneficial to rotate the tissue from a horizontal position (most dorsal part on top) to a vertical position (dorsal part of brain now either on left or right side).
7. We find it easier to flip slice over and flatten against the mounting stage before transferring to the slide. Additionally, organize the slices in a way that leaves room for your hydrophobic barrier and in small enough of an area to minimize reagent use.
8. There are two ways that you may organize tissue slices from different animals on the slides. Each slide should have eight to ten slices. You can arrange the slices from the same brain on one slide, or you can put one slice of brain from each animal on the slides. One benefit of the latter is that it reduces the possibility of having different staining integrity, but the largest drawback is that if you are not planning to slice all your animals on the same day, then you risk temperature variability and exposure to environmental confounds. Therefore, we choose the former, with one animal per slide, especially given the possibility of mounting a slice of tissue that is ripped. When having eight to ten slices on your slide from one animal, when it comes to imaging, you have eight to ten slices to choose from in case some slices are not of great quality.
9. Turquoise killifish tissues naturally accumulate autofluorescence through a number of proteins and cross-linking fixative interactions, but most prevalently the natural accumulation of lipofuscin [10, 27, 32]. Thus, the hydrogen peroxide step can potentially be combined with a photobleaching step for the tissue with high-powered LEDs using a custom light box or previously referenced LED setups [39–42]. This can be done by exposing your slides to the LEDs (~2000 lumens) during the air-drying step or through an additional PBS submerged setup.

10. Some protocols use an automatic rocker or orbital shaker for their wash steps. We discourage this since the killifish tissue is very small and prone to damage from vigorous washing steps.
11. If not using the Manual PreTreat Pro section of the RNAscope protease-free section of the manual, then shorten protease digestion times to 10 or 15 min to avoid digestion of essential cell membrane proteins.
12. It is optional to use a micropipette to apply your reagents to the slides if you are unsure about the drop requirement or amount for each slide.
13. For IHC and RNAscope together, the washes during the IHC section are paramount in getting rid of background noise of your secondary fluorophore, which could impede the RNAscope fluorescent signal. This combination is especially useful for colocalizing mRNA gene products within distal or proximal parts of cells like neurons or microglia without having transgenic lines [6, 43, 44].
14. Some primary antibodies may require 48 h of incubation. Please optimize the dilution, diluent, and incubation times for both the primary and secondary antibodies based on the specific assay.
15. For highly expressed mRNA transcripts, the time it takes for the fluorescent signal to fade and become unable to be imaged is extended to about a month. If you are unsure about the expression levels, it is recommended not to wait for more than 2 weeks to image your samples.
16. The bit depth of the images acquired can dramatically change the quality of the image. It is possible to quantify your puncta with an 8-bit image instead of 12- or 16-bit, but make sure to use a high magnification objective, as it becomes difficult to disentangle adjacent puncta.

adjacent puncta.

17. When choosing the laser intensity for each channel, ensure not to have auto-exposure turned on, as it is beneficial to see the raw image. You can use your negative control to gauge where the laser intensities should be. Additionally, ensure that the signal is not saturated and allows for a dynamic range, since crosstalk between channels can occur if one is too bright.

18. Due to possible background differences between samples and the semi-quantitative nature of possible puncta being too close to each other, it is recommended that you use the same type of thresholding standards for cell detection and cell surface creation. This applies to spot detection and segmentation algorithms, and image processing techniques in either QuPath or Imaris [45, 46]. It is best to test your settings on a few images to apply to the rest of them. Although it is best to keep thresholding conditions the same for every image, when imaging the whole brain slice that has high variability instead of smaller equidistant ROI, you may have to change thresholds for each image and have a blinded manual analysis from one or two other researchers. Additionally, QuPath does not have an automated version of their subcellular spot detection, so you must manually record each spot (0.5–1 μm) or cluster (>2 μm) detection threshold (Fig. 5C).

19. There have been advances in tissue clearing techniques for fish brains and embryos to learn more about neural circuitry or fate mapping of certain pathways, with which RNAscope could be combined [33, 47].

Acknowledgments

Figures were, in part, created using BioGDP, an open-source database for high-quality biomedical graphics [48]. The authors thank Dr. Victor Ansere for his feedback on this manuscript.

Funding

This work was funded by a SURFiN fellowship (G.C.) and Simons Foundation International grant SFI-AN-NC-AB-Research-00018042 (B.A.B.).

Conflict of Interest

The authors have no conflict of interest.

References

1. Terzibasi E, Valenzano DR, Benedetti M et al (2008) Large differences in aging phenotype between strains of the short-lived annual fish Nothobranchius furzeri. PLoS One 3(12):e3866. https://doi.org/10.1371/journal.pone.0003866 [PubMed][PubMedCentral]

2. Kim Y, Nam HG, Valenzano DR (2016) The short-lived African turquoise killifish: an emerging experimental model for ageing. Dis Model Mech 9(2):115–129. https://doi.org/10.1242/dmm.023226 [PubMed][PubMedCentral]

3. de Bakker DEM, Valenzano DR (2023) Turquoise killifish: a natural model of age-dependent brain degeneration. Ageing Res Rev 90:102019. https://doi.org/10.1016/j.arr.2023.102019 [PubMed]

4. Morabito G, Ryabova A, Valenzano DR (2024) Immune aging in annual killifish. Immun Ageing 21(1):18. https://doi.org/10.1186/s12979-024-00418-3 [PubMed][PubMedCentral]

5. Valenzano DR, Benayoun BA, Singh PP et al (2015) The African turquoise killifish genome provides insights into evolution and genetic architecture of lifespan. Cell 163(6):1539–1554. https://doi.org/10.1016/j.cell.2015.11.008 [PubMed][PubMedCentral]

6. Mitchell DM, Lovel AG, Stenkamp DL (2018) Dynamic changes in microglial and macrophage characteristics during degeneration and regeneration of the zebrafish retina. J Neuroinflammation 15(1):163. https://doi.org/10.1186/s12974-018-1185-6 [PubMed][PubMedCentral]

7. Van Houcke J, Marien V, Zandecki C et al (2021) Aging impairs the essential contributions of non-glial progenitors to neurorepair in the dorsal telencephalon of the Killifish Nothobranchius furzeri. Aging Cell 20(9):e13464. https://doi.org/10.1111/acel.13464 [PubMed][PubMedCentral]

8. Vanhunsel S, Bergmans S, Beckers A et al (2021) The killifish visual system as an in vivo model to study brain aging and rejuvenation. NPJ Aging Mech Dis 7(1):22. https://doi.org/10.1038/s41514-021-00077-4

[PubMed][PubMedCentral]

9. Bergmans S, Vanhunsel S, Moons L (2023) Visually driven behavior assays to study functionality and recovery after damage of the aged African turquoise killifish visual system. Cold Spring Harb Protoc 2023(11):108096. https://doi.org/10.1101/pdb.prot108096 [PubMed]

10. Yu X, Li G (2012) Effects of resveratrol on longevity, cognitive ability and aging-related histological markers in the annual fish Nothobranchius guentheri. Exp Gerontol 47(12):940–949. https://doi.org/10.1016/j.exger.2012.08.009 [PubMed]

11. Genade T, Lang DM (2013) Resveratrol extends lifespan and preserves glia but not neurons of the Nothobranchius guentheri optic tectum. Exp Gerontol 48(2):202–212. https://doi.org/10.1016/j.exger.2012.11.013 [PubMed]

12. Valenzano DR, Terzibasi E, Genade T et al (2006) Resveratrol prolongs lifespan and retards the onset of age-related markers in a short-lived vertebrate. Curr Biol 16(3):296–300. https://doi.org/10.1016/j.cub.2005.12.038 [PubMed]

13. D'Angelo L (2013) Brain atlas of an emerging teleostean model: Nothobranchius furzeri. Anat Rec (Hoboken) 296(4):681–691. https://doi.org/10.1002/ar.22668 [PubMed]

14. Eunjeong D, Seongsin L, Yumi K (2021) Anatomical identification of the neuroendocrine system in the Nothobranchius furzeri brain. J Vertebr Biol 70(2):21018.1-9. https://doi.org/10.25225/jvb.21018

15. Petzold A, Reichwald K, Groth M et al (2013) The transcript catalogue of the short-lived fish Nothobranchius furzeri provides insights into age-dependent changes of mRNA levels. BMC Genomics 14(185). https://doi.org/10.1186/1471-2164-14-185

16. Teefy BB, Lemus AJJ, Adler A et al (2023) Widespread sex dimorphism across single-cell transcriptomes of adult African turquoise killifish tissues. Cell Rep 42(10):113237. https://doi.org/10.1016/j.celrep.2023.113237 [PubMed][PubMedCentral]

17. Xu A, Teefy BB, Lu RJ et al (2023) Transcriptomes of aging brain, heart, muscle, and spleen from female and male African turquoise killifish. Sci Data 10(1):695. https://doi.org/10.1038/s41597-023-02609-x

18. Harel I, Chen YR, Ziv I et al (2024) Identification of protein aggregates in the aging vertebrate brain with prion-like and phase-separation properties. Cell Rep 43(6):112787. https://doi.org/10.1016/j.celrep.2023.112787 [PubMed][PubMedCentral]

19. Wang F, Flanagan J, Su N et al (2012) RNAscope: a novel in situ RNA analysis platform for

formalin-fixed, paraffin-embedded tissues. J Mol Diagn 14(1):22–29. https://doi.org/10.1016/j.jmoldx.2011.08.002
[PubMed][PubMedCentral]

20. Grabinski TM, Kneynsberg A, Manfredsson FP et al (2015) A method for combining RNAscope in situ hybridization with immunohistochemistry in thick free-floating brain sections and primary neuronal cultures. PLoS One 10(3):e0120120. https://doi.org/10.1371/journal.pone.0120120
[PubMed][PubMedCentral]

21. Morrison JA, McKinney MC, Kulesa PM (2017) Resolving in vivo gene expression during collective cell migration using an integrated RNAscope, immunohistochemistry and tissue clearing method. Mech Dev 148:100–106. https://doi.org/10.1016/j.mod.2017.06.004
[PubMed][PubMedCentral]

22. Gross-Thebing T (2020) RNAscope multiplex detection in Zebrafish. Methods Mol Biol 2148:195–202. https://doi.org/10.1007/978-1-0716-0623-0_12
[PubMed]

23. Hammond-Weinberger DR, ZeRuth GT (2020) Whole mount immunohistochemistry in zebrafish embryos and larvae. J Vis Exp 29(155):e60575. https://doi.org/10.3791/60575

24. Silva NJ, Dorman LC, Vainchtein ID et al (2021) In situ and transcriptomic identification of microglia in synapse-rich regions of the developing zebrafish brain. Nat Commun 12(1):5916. https://doi.org/10.1038/s41467-021-26206-x
[PubMed][PubMedCentral]

25. Wu M, Lambatan V, Guo P et al (2025) Protocol for multiplexed RNAscope-based imaging of mRNAs in whole-mount adult Drosophila brains. STAR Protoc 6(1):103615. https://doi.org/10.1016/j.xpro.2025.103615
[PubMed][PubMedCentral]

26. Aragona M, Porcino C, Guerrera MC et al (2021) Localization of neurotrophin specific Trk receptors in mechanosensory systems of killifish (*Nothobranchius guentheri*). Int J Mol Sci 22(19):10411. https://doi.org/10.3390/ijms221910411

27. Liu S, Wang W (2023) Immunofluorescence staining on cryosections with high autofluorescence in African killifish. Springer Protocols. In: Emerging model organisms. Neuromethods, pp 95–108. https://doi.org/10.1007/978-1-0716-2875-1_8

28. Borgonovo J, Ahumada-Galleguillos P, Onate-Ponce A et al (2021) Organization of the catecholaminergic system in the short-lived fish Nothobranchius furzeri. Front Neuroanat 15:728720. https://doi.org/10.3389/fnana.2021.728720
[PubMed][PubMedCentral]

29. Borgonovo J, Allende-Castro C, Medinas DB et al (2024) Immunohistochemical characterisation of the adult Nothobranchius furzeri intestine. Cell Tissue Res 395(1):21–38. https://doi.org/10.1007/s00441-023-03845-8
[PubMed]

30. Bergmans S, Noel NCL, Masin L et al (2024) Age-related dysregulation of the retinal transcriptome in African turquoise killifish. Aging Cell 23(8):e14192. https://doi.org/10.1111/acel.14192 [PubMed][PubMedCentral]

31. Ayana R, Zandecki C, Van Houcke J et al (2024) Single-cell sequencing unveils the impact of aging on the progenitor cell diversity in the telencephalon of the female killifish N. furzeri. Aging Cell 23(10):e14251. https://doi.org/10.1111/acel.14251 [PubMed][PubMedCentral]

32. Schöfer S, Laffer S, Kirchberger S et al (2024) Senescence-associated ß-galactosidase staining over the lifespan differs in a short- and a long-lived fish species. Eur J Histochem 68:3977. https://doi.org/10.4081/ejh.2024.397 [PubMed][PubMedCentral]

33. Lee S, Kim Y (2021) Protocol for whole-brain immunostaining of the turquoise killifish after tissue clearing. STAR Protoc 2(2):100564. https://doi.org/10.1016/j.xpro.2021.100564 [PubMed][PubMedCentral]

34. Baleriola J, Jean Y, Troy C et al (2015) Detection of axonally localized mRNAs in brain sections using high-resolution in situ hybridization. J Vis Exp 100:e52799. https://doi.org/10.3791/52799

35. Fe Lanfranco M, Loane DJ, Mocchetti I et al (2017) Combination of Fluorescent in situ Hybridization (FISH) and immunofluorescence imaging for detection of cytokine expression in microglia/macrophage cells. Bio-Protocol 7(22):e2608. https://doi.org/10.21769/BioProtoc.2608

36. Bedbrook C, Nath RD, Nagvekar R et al (2023) Rapid and precise genome engineering in a naturally short- lived vertebrate. eLife https://doi.org/10.7554/eLife.80639 [PubMed][PubMedCentral]

37. Bankhead P, Loughrey MB, Fernandez JA et al (2017) QuPath: open source software for digital pathology image analysis. Sci Rep 7(1):16878. https://doi.org/10.1038/s41598-017-17204-5 [PubMed][PubMedCentral]

38. Advanced Cell Diagnostics I (2025) UM 323100 multiplex fluorescent v2 user manual. https://acdbio.com/system/files_force/UM%20323100%20Multiplex%20Fluorescent%20v2%20User%20Manual_0.pdf?download=1

39. Sun Y, Ip P, Chakrabartty A (2017) Simple elimination of background fluorescence in formalin-fixed human brain tissue for immunofluorescence microscopy. J Vis Exp (127):e56188. https://doi.org/10.3791/56188

40. Adeniyi PA, Fopiano KA, Banine F et al (2022) Multispectral LEDs eliminate lipofuscin-associated autofluorescence for immunohistochemistry and CD44 variant detection by in situ hybridization in aging human, non-human primate, and murine brain. ASN Neuro 14:17590914221123138. https://doi.org/10.1177/17590914221123138

41.

Tsuneoka Y, Atsumi Y, Makanae A et al (2022) Fluorescence quenching by high-power LEDs for highly sensitive fluorescence in situ hybridization. Front Mol Neurosci 15:976349. https://doi.org/10.3389/fnmol.2022.976349
[PubMed][PubMedCentral]

42. Sapio MR, King DM, Maric D et al (2025) Efficient removal of naturally-occurring lipofuscin autofluorescence in human nervous tissue using high-intensity white light. J Pain 30:105359. https://doi.org/10.1016/j.jpain.2025.105359
[PubMed][PubMedCentral]

43. Narra SS, Rondeau P, Fernezelian D et al (2023) Distribution of microglia/immune cells in the brain of adult zebrafish in homeostatic and regenerative conditions: Focus on oxidative stress during brain repair. J Comp Neurol 531(2):238–255. https://doi.org/10.1002/cne.25421
[PubMed]

44. Ball JB, McNulty CJ, Green-Fulgham SM et al (2023) Combining RNAscope and immunohistochemistry to visualize inflammatory gene products in neurons and microglia. Front Mol Neurosci 16:1225847. https://doi.org/10.3389/fnmol.2023.1225847
[PubMed][PubMedCentral]

45. Avegno EM, Kasten CR, Snyder 3rdWB et al (2021) Alcohol dependence activates ventral tegmental area projections to central amygdala in male mice and rats. Addict Biol 26(4):e12990. https://doi.org/10.1111/adb.12990
[PubMed]

46. Bentley-Abbot C, Heslop R, Pirillo C et al (2024) An easy to use tool for the analysis of subcellular mRNA transcript colocalisation in smFISH data. Sci Rep 14(1):8348. https://doi.org/10.1038/s41598-024-58641-3
[PubMed][PubMedCentral]

47. Xu J, Zhu L, He S et al (2015) Temporal-spatial resolution fate mapping reveals distinct origins for embryonic and adult microglia in Zebrafish. Dev Cell 34(6):632–641. https://doi.org/10.1016/j.devcel.2015.08.018
[PubMed]

48. Jiang S, Li H, Zhang L et al (2025) Generic Diagramming Platform (GDP): a comprehensive database of high-quality biomedical graphics. Nucleic Acids Res 53(D1):D1670–D1676. https://doi.org/10.1093/nar/gkae973
[PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_14

14. Practical Guide to RNA Sequencing Analysis in Aging Research

Gee-Yoon Lee[1] ✉ and Seung-Jae V. Lee[1] ✉
(1) Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon, 34141, South Korea

✉ **Gee-Yoon Lee**
Email: gyoon@kaist.ac.kr

✉ **Seung-Jae V. Lee (Corresponding author)**
Email: seungjaevlee@kaist.ac.kr

Abstract
Aging is accompanied by widespread changes in gene expression, which underlie functional and structural declines in organisms over time. Transcriptome analysis, including RNA sequencing, has become an essential approach in modern research on biological processes, including aging, but remains challenging for many experimental scientists. Here, we provide a practical protocol for RNA sequencing analysis for aging research.

Keywords Aging – Longevity – Transcriptome – RNA-seq – Bioinformatics

1 Introduction

Aging is a progressive decline in physiological integrity that increases vulnerability to disease and death [1, 2]. Numerous studies across a wide range of model organisms have uncovered genetic and environmental factors that influence lifespan, revealing conserved aging-regulating pathways such as insulin/insulin-like growth factor-1 (IGF-1) signaling, dietary restriction, mitochondrial function, and stress responses [3–5]. Despite these advances, aging remains a complex and multifactorial process, and comprehensive approaches are required to understand the molecular mechanisms that underlie age-dependent physiological decline.

Traditional genetic and biochemical methods have generated important insights into aging and longevity. However, these techniques often focus on individual genes or pathways and provide limited information regarding global and systematic changes in the transcriptome. Recent advances in next-generation sequencing have made large-scale transcriptomic analyses widely accessible in aging research [6–8]. Transcriptome analysis offers a genome-wide perspective on gene expression and can reveal regulatory networks, signaling pathways, and cellular states associated with aging [9–13]. Because transcript abundance often reflects steady-state levels shaped by transcription and RNA turnover, transcriptome profiling has become a central method for identifying molecular signatures of aging [14–17].

RNA sequencing (RNA-seq) is one of the most widely used technologies for transcriptome analysis and is routinely applied to compare gene expression states across different ages, tissues, or genetic backgrounds [18]. Multiple transcriptomic approaches can be combined to comprehensively characterize aging-related gene expression changes. While reverse transcription-quantitative polymerase chain reaction (RT-qPCR) is widely used for measuring expression of specific candidate genes [19], and single-cell RNA-seq resolves cell-type-specific changes [20, 21], bulk RNA-seq serves as a primary tool for unbiased transcriptome-wide profiling. Bulk RNA-seq provides quantitative measurement of transcript levels and enables the identification of differentially expressed genes and pathways that

change with age. Despite its power and accessibility, many scientists in the aging research field may still find the computational aspects of RNA-seq analysis challenging. Here, we focus on bulk RNA-seq analysis and describe a practical guide for experimental biologists in aging research. This protocol is designed for researchers with limited computational experience and provides practical guidelines, recommended parameters, and analytical checkpoints to facilitate reliable and reproducible transcriptome analysis in aging research.

---

2 Materials

The following materials and computational resources are required for bulk RNA- seq analysis in aging research. Reagents should be prepared and stored following manufacturer specifications. Computational analyses can be performed on a standard Linux workstation unless otherwise indicated (*see* **Note 1**).

2.1 RNA Extraction and Library Preparation

1. Biological samples: Collected at defined aging stages (e.g., young, middle-aged, aged) or from genetic or pharmacological interventions that modulate lifespan (e.g., genetic mutants, RNAi treatments, dietary restriction, pharmacological interventions). Samples should be harvested under identical growth or culture conditions to minimize technical variation.
2. RNase-free consumables, including 1.5 ml microcentrifuge tubes, pipette tips, and RNase-free water.
3. Reagents for phenol-chloroform-based extraction or column-based purification.
4. PolyA RNA (mRNA) enrichment kit and/or ribosomal RNA (rRNA) depletion kit, depending on study design (*see* **Note 2**).
5. Stranded or non-stranded library preparation kit.

6. RNA sequencing platform for paired-end sequencing (75–150 bp).

7. Recommended sequencing depth: 30–50 million paired-end reads per sample for polyA-selected libraries. Minimum 50 million reads for complex tissues, low-expression genes, or rRNA-depleted libraries.

2.2 Software and Computational Environment

The following software and computational environments are recommended based on the Encyclopedia of DNA Elements (ENCODE) RNA-seq data standard processing pipeline. Graphical user interfaces may be preferable for researchers without command-line experience (*see* **Note 1**).

1. Workstation with at least eight CPU cores, 32 GB RAM, and 500 GB storage space.
2. FastQC v0.12.x for raw read quality assessment [22].
3. Trim Galore v0.6.x for adapter and low-quality base trimming. Cutadapt v4.x and FastQC v0.12.x are required.
4. MultiQC v1.x for summarizing quality control results across samples.
5. STAR v2.7.x for reference genome alignment (ENCODE-recommended splice aligner) [23].
6. RSEM v1.3.x (ENCODE-recommended standard transcript quantification) [24].
7. R v4.0 or higher with various packages [25], including DESeq2, edgeR, ggplot2, pheatmap, gprofiler2, and clusterProfiler installed. Other languages, such as Python, can be used depending on the user's preference and familiarity.

2.3 Reference Genomes and Annotation Files

1. Reference genome FASTA file corresponding to the species of interest, such as human GRCh38, mouse GRCm39, and *Caenorhabditis elegans* WBcel235. For example, the *C. elegans* WBcel235 genome FASTA can be downloaded from Ensembl (release 115): https://ftp.ensembl.org/pub/release-115/fasta/caenorhabditis_elegans/dna/Caenorhabditis_elegans.WBcel235.dna.toplevel.fa.gz.
2. Gene annotation file (GTF/GFF) matching the exact genome assembly version (required for ENCODE STAR-RSEM compatibility). For example, the *C. elegans* WBcel235 genome GTF file is available at Ensembl (release 115): https://ftp.ensembl.org/pub/release-115/gtf/caenorhabditis_elegans/Caenorhabditis_elegans.WBcel235.115.gtf.gz.
3. The reference genome and gene annotation files should be obtained from a consistent source (e.g., Ensembl, NCBI RefSeq, or UCSC) and match the same genome build and chromosome naming convention.

---

3 Methods/Protocols

The following RNA-seq protocol is based on the ENCODE RNA-seq data standard processing pipeline (Fig. 1; *see* **Note 1**).

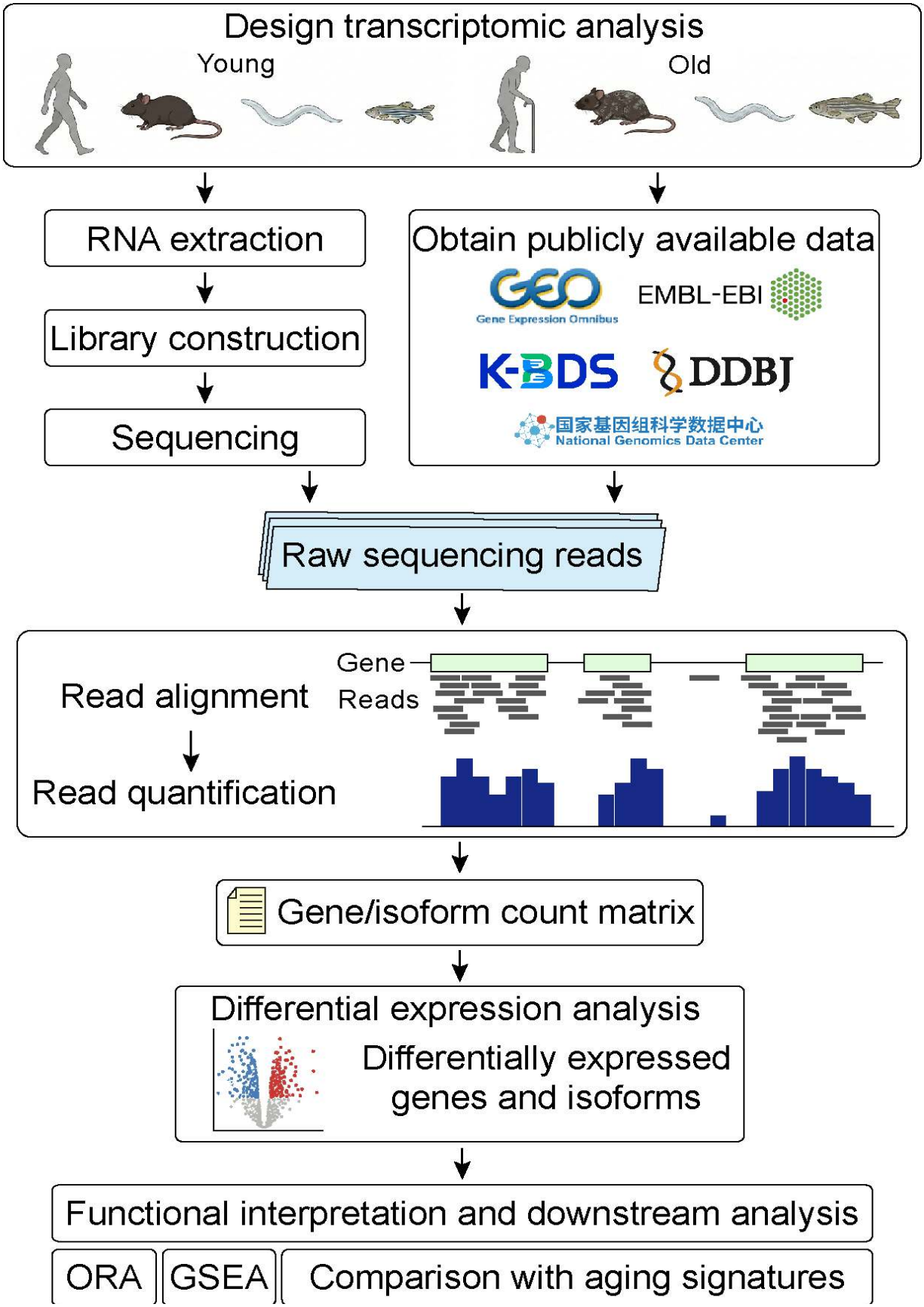

Design transcriptomic analysis
Young
Old
RNA extraction
Library construction
Sequencing
Obtain publicly available data
Gene Expression Omnibus
EMBL-EBI
K-BDS
DDBJ
国家基因组科学数据中心
National Genomics Data Center
Raw sequencing reads
Read alignment
Read quantification
Gene
Reads
Gene/isoform count matrix
Differential expression analysis
Differentially expressed genes and isoforms
Functional interpretation and downstream analysis
ORA
GSEA
Comparison with aging signatures

Fig. 1 Workflow for RNA sequencing analysis in aging samples. RNA-seq data comparing experimental conditions relevant to aging (e.g., young vs. old, long- vs. short-lived conditions, or intervention vs. control) are generated experimentally or obtained from public repositories. Reads are mapped and quantified to produce a gene- and/or isoform-level count matrix. The count matrix is used for downstream analyses, including differential expression and functional interpretation, such as over-representation analysis (ORA) for functional biological terms and pathways, gene set enrichment analysis (GSEA) against curated gene sets, and comparison with aging signatures. When genome-level alignment is not required, a pseudoalignment-based workflow (e.g., Kallisto) can be used as an alternative

3.1 Generating RNA-Seq Data or Using Public Datasets

1. Extract RNAs from various aging samples, ranging from organisms such as the nematode *C. elegans* and tissue and/or organs isolated from mice and human biopsies. After harvesting, wash samples thoroughly to remove residual contaminants such as media components, buffer residues, blood, or debris, depending on the tissue or organism (e.g., *C. elegans* washed thoroughly in M9 to remove bacterial contaminants). Immediately snap-freeze the samples in liquid nitrogen and homogenize samples for RNA extraction. Isolate total RNA using phenol-chloroform extraction or using column-based kits, followed by DNase I treatment to minimize genomic DNA contamination. RNase-free tubes, water, and filter tips should be used at all times. Prepare at least three biological replicates per condition.
2. Purify and assess RNA quality: Ensure sufficient total RNA is available for library preparation. Quantify RNA using a fluorometric assay (e.g., Qubit RNA HS/BR) rather than absorbance-based measurements. Confirm that concentration meets the sequencing facility requirements and assess RNA integrity using an electrophoretic platform (e.g., Agilent Bioanalyzer or TapeStation) before proceeding. Samples with RNA integrity number (RIN or RINe) ≥ 7 are generally suitable for polyA-selected library preparation. RIN indicates RNA quality based on the extent of degradation, typically reported on a scale from 1 (highly degraded) to 10 (intact RNA). We recommend running an agarose gel to visualize RNA integrity for examining intact rRNA bands and

detection of potential degradation or contamination that may not be fully captured by automated electrophoretic metrics.

3. Prepare RNA sequencing libraries: mRNAs can be enriched using polyA selection kits. Alternatively, rRNA can be removed using rRNA depletion kits to reduce the proportion of rRNA-derived reads and increase sequencing coverage of informative transcripts. Libraries may be prepared as stranded or non-stranded, depending on the kit used; both approaches are compatible with gene-level quantification and downstream analyses. A total of 1 to 2 μg input RNA is generally suitable, although requirements may vary depending on the library preparation kit (*see* **Note 2**).
4. Sequencing: Sequencing can be performed on widely used high-throughput next-generation sequencing platforms. Paired-end sequencing reads (75–150 bp) are recommended to improve splice-junction detection, transcript isoform quantification, and mapping accuracy in repetitive and/or ambiguous regions. A minimum depth of 30 million paired-end reads per sample is recommended. Depth of over 50 million reads may be beneficial for complex aging tissues or analyses involving transcripts with low expression and alternative splicing.
5. Archive and deposit newly generated data: Ensure that raw sequencing outputs and associated metadata are properly archived and deposited in recognized public repositories to support transparency and reproducibility (*see* **Note 3**).
6. Obtain public datasets: Download and reanalyze raw transcriptome data from public databases to extend analyses or integrate with newly generated data (*see* **Note 4**). Major repositories include Gene Expression Omnibus database [26] (GEO; https://www.ncbi.nlm.nih.gov/geo/), European Molecular Biology Laboratory's European Bioinformatics Institute [27] (EMBL-EBI; https://www.ebi.ac.uk/), Korea BioData Station [28] (K-BDS; https://kbds.re.kr/), DNA Data Bank of Japan [29] (DDBJ; https://www.ddbj.nig.ac.jp/), and

Bank of Japan [29] (DDBJ; https://www.ddbj.nig.ac.jp/), and National Genomics Data Center [30] (NGDC; https://ngdc.cncb.ac.cn/). For large-scale transcriptome datasets of aging mammalian organisms, Genotype-Tissue Expression [31] (GTEx; https://gtexportal.org/home/datasets), Tabula Sapiens [8] (https://tabula-sapiens-portal.ds.czbiohub.org/), and Tabula Muris Senis [32] (https://tabula-muris-senis.sf.czbiohub.org/) are also available.

3.2 Quality Assessment of Raw Reads

1. Inspect raw FASTQ files that were newly generated or obtained from the public databases using FastQC. Ensure that per-base Phred quality scores (Q) remain high across most read positions (e.g., Q ≥ 30), adapter contamination is minimal, and GC content is broadly consistent with the expected distribution for the species.
2. Trim adapters and low-quality bases using Trim Galore (Phred cutoff 20 and minimum read length 30 bp).
3. Re-assess the quality of trimmed reads using FastQC and summarize results with MultiQC to compare samples consistently. Flag samples with persistent quality control (QC) abnormalities for further evaluation in downstream alignment-based QC steps (*see* **Note 5**).

3.3 Alignment of Reads to the Reference Genome

This procedure follows the ENCODE-standard STAR RNA-seq alignment pipeline. Refer to the STAR alignment vignettes for detailed command lines and advanced parameters [23]. When genome-level alignment information is not required, pseudoalignment-based quantification tools such as Kallisto [33] or Salmon [34] can be used as an alternative (*see* **Note 6**).

1. Obtain reference genome FASTA and GTF for building the STAR genome index (*see* **Notes 7** and **8**).

2. Build a STAR genome index using reference FASTA and GTF.

3. Align trimmed reads using ENCODE-recommended parameters for STAR to generate genome-sorted BAM and transcriptome-aligned BAM.
4. Examine log files (e.g., Log.final.out) and confirm basic alignment quality metrics. Recommended thresholds include (1) uniquely mapped reads ≥ 70%, (2) multi-mapping ≤ 10%, and (3) mismatch rate < 1% (*see* **Note 9**).
5. Determine library strandedness from the aligned BAM files (e.g., using RSeQC or Qualimap) before running downstream quantification (*see* **Note 10**).
6. Perform post-alignment QC to identify samples with QC anomalies. Confirm that the total number of mapped reads per sample is consistent with the expected sequencing depth and comparable across biological replicates. Review additional indicators such as an unusually low mapping rate relative to the dataset, high rRNA-derived reads, abnormal exonic/intronic read distribution, strong 3′ gene body bias consistent with RNA degradation, inconsistent strandedness across samples, contamination, or extreme duplication and/or low complexity. If a sample shows clear QC anomalies, record the specific metrics and suspected cause, and consider excluding the sample from downstream analyses.
7. For mutant or transgenic samples, where the affected locus is sufficiently expressed and covered, verify expected genotypes by loading the coordinate-sorted BAM files from STAR alignment into a genome browser such as Integrative Genomics Viewer (IGV [35]). Load the reference genome and BAM files, then navigate to the genomic coordinates of the mutation site. For deletion mutants, confirm reduced or absent read coverage in the deleted region. For insertion or indel mutations, ensure that the expected variant is

present in mutant samples and absent in wild-type controls. This quality assessment helps identify potential sample mix-ups or genotyping errors before proceeding to differential expression analysis.

3.4 Transcript Quantification

This procedure follows the ENCODE-standard RSEM quantification pipeline.

1. Build RSEM reference index from genome FASTA and GTF.
2. Quantify transcript expression using RSEM from the STAR transcriptome-aligned BAM generated in Sect. 3.3. Ensure that the correct library strandedness is specified (*see* **Note 10**).
3. Examine the RSEM output files:
   (1) sample.genes.results (gene-level transcripts per million (TPM) and expected counts)
   (2) sample.isoforms.results (transcript-level TPM and expected counts)
4. Examine RSEM log files to confirm that a high proportion of reads are successfully assigned to genes or transcripts and that assignment rates are comparable across samples.
5. Compile gene-level and/or transcript-level expected counts across all samples into a count matrix for downstream statistical analyses.
6. Use expected counts as input for differential expression analysis. Estimate library-size and composition normalization factors within the statistical framework (e.g., calculate size factors in DESeq2 or perform TMM normalization in edgeR).

7. Add gene symbols or external gene names to the count matrix. Map Ensembl gene IDs to gene symbols using annotation resources (e.g., biomaRt in R). Ensure that the annotation source matches the genome/GTF version used for alignment and quantification.
8. Generate principal component analysis (PCA) and/or multidimensional scaling (MDS) plots using normalized and/or transformed values (e.g., variance-stabilized counts, log-transformed counts, or TPM) to assess sample clustering and detect potential outliers.

3.5 Differential Expression Analysis

Perform differential expression analysis using established statistical frameworks (e.g., DESeq2 [36], edgeR [37], and limma-voom [38]). The workflow below describes gene-level analysis using DESeq2. For detailed instructions and advanced options in DESeq2, refer to the DESeq2 vignette and the original publication [36].

1. Import the gene-level count matrix generated in Sect. 3.4. Filter and remove genes with low expression (e.g., genes with less than ten total read counts across all samples). Examine total read counts and library size distribution across samples before differential expression analysis to identify samples with unusually low sequencing depth or extreme library size imbalance.
2. Create a DESeq2 dataset using the count matrix and corresponding sample metadata. Design a formula specifically depending on experimental conditions and run the DESeq2 pipeline. For multifactorial experimental designs (e.g., age + batch), include all relevant variables in the design formula to avoid confounding effects and incorrect attribution of age-dependent changes.
3. Extract differential expression results and apply statistical thresholds (e.g., $|\log_2$ fold change$| \geq 1$ and adjusted p-value < 0.05).

For exploratory analyses, these thresholds may be relaxed or tightened depending on the aims and priorities of the research (*see* **Notes 11** and **12**).

4. Export a list of significantly differentially expressed genes (DEGs) for downstream functional and pathway analysis.

3.6 Functional Annotation and Pathway Analysis

1. Select enrichment analysis approach: Use over-representation analysis (ORA) when testing a predefined list of DEGs, or use gene set enrichment analysis (GSEA) [39] to detect pathway-level changes across the full ranked gene list, including genes that do not meet DEG thresholds.
2. For ORA, upload the DEG list to functional enrichment analysis tools, including DAVID [40], Enrichr [41], g:Profiler [42], and WebGestalt [43]. Identify functional terms enriched among DEGs, including Gene Ontology (GO) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and Reactome pathways/categories (*see* **Note 13**). Use organism-specific annotation resources when relevant (e.g., WormCat [44] for *C. elegans*). Use an appropriate background gene set (e.g., all expressed genes detected in the dataset) to avoid biased enrichment results. Apply multiple testing correction (e.g., Benjamini–Hochberg FDR) when interpreting enriched terms.
3. For GSEA, generate a ranked gene list from differential expression results (e.g., rank genes by signed $\log_2$ fold change or Wald statistic) and perform pathway enrichment analysis using the complete ranked gene list (*see* **Note 14**).
4. Summarize and prioritize enriched terms/pathways by effect size or enrichment statistics (e.g., adjusted *p*-value/FDR, normalized enrichment score for GSEA). Reduce term redundancy using semantic similarity clustering and compile a final shortlist for

biological interpretation in the context of aging.

3.7 Visualization of RNA-seq Data

1. Generate PCA or MDS plots to visualize sample grouping.
2. Create volcano plots displaying fold change and statistical significance of genes.
3. Construct heatmaps using gene-wise *Z*-score normalization across samples (mean-centered and scaled by standard deviation) to highlight relative expression patterns, or scale expression values relative to a reference condition (e.g., youngest age group) to visualize directional age-dependent changes.
4. Visualize enriched pathway terms using bar plots or bubble plots.
5. For selected genes of interest, visualize gene-level expression using dot/box/bar plots, preferably showing individual samples (e.g., normalized counts) to assess within-group variability (*see* **Notes 12** and **15**).

3.8 Gene Sets and Resources for Aging and Longevity Analyses

Curated gene sets associated with aging or lifespan can be used to assess whether a biological condition or intervention engages aging-related transcriptional programs through enrichment or overlap analyses and to support transcriptome-based age or stage estimation. Relevant resources are listed below.

1. Molecular Signatures Database (MSigDB) [39]: A curated collection of gene sets organized into Hallmark, curated pathway, and GO categories. Hallmark gene sets and curated KEGG or Reactome pathways can be used to evaluate pathway-level transcriptional changes associated with aging.
2. Human Ageing Genomic Resources (HAGR) [45]: An integrated

resource dedicated to the genetics and biology of aging. HAGR provides curated gene sets relevant to transcriptomic analyses, including GenAge (genes experimentally associated with lifespan modulation), GenDR (genes required for dietary restriction-mediated lifespan extension), CellAge (experimentally validated senescence-associated genes), and LongevityMap (human genetic variants associated with longevity).

3. Transcriptomic Classification via Adaptive Learning of Signature States (T-CLASS) [17]: A web-based classification tool that determines which physiological state a transcriptome resembles by comparing against reference data. T-CLASS can be particularly useful in aging research for classifying transcriptomes into longevity-promoting pathways (e.g., reduced insulin/IGF-1 signaling, dietary restriction, and reduced mitochondrial function) or types of cellular senescence to infer mechanisms underlying specific interventions.
4. SenMayo cellular senescence gene signatures [46]: The SenMayo cellular senescence gene signature is a curated gene set derived from integrative analyses of human and mouse datasets. It consists of genes consistently enriched in senescent cells and can be used to estimate senescence-associated transcriptional enrichment in bulk and single-cell RNA-seq data.
5. Mammalian Signatures of Aging and Longevity Traits (mSALT) [47]: A set of mammalian gene expression signatures linked to aging and lifespan traits across tissues and interventions.
6. Maximum lifespan-associated genes in mammals [48]: Gene sets defined based on positive (Pos-MLS) or negative (Neg-MLS) correlations between gene expression and maximum lifespan across mammalian species.
7. Aging biomarkers in *C. elegans* [3, 49, 50]: Genes that influence lifespan and longevity that can be used to evaluate age-dependent

transcriptional changes and to assess the impact of genetic or environmental interventions in *C. elegans*.

8. Real-age prediction from transcriptome staging on reference (RAPToR) [51]: A computational framework for estimating developmental or biological age from transcriptomic data by staging samples against reference time-series datasets.

3.9 Alternative Splicing Analysis

Alternative splicing generates multiple transcript isoforms from a single gene and is a key regulatory mechanism altered during aging [52]. In addition to identifying differentially expressed genes and isoforms (Sect. 3.5), analyzing changes in splicing patterns can reveal post-transcriptional regulation that contributes to aging phenotypes [14–16, 53]. Two complementary approaches exist for splicing analysis: event-based analysis identifies and quantifies specific splicing events (e.g., exon skipping, intron retention, and alternative splice sites), while transcript-level analysis examines overall shifts in isoform proportions (e.g., whether isoform A becomes more abundant than isoform B with age).

1. Perform event-based splicing analysis: Use tools such as rMATS [54], MAJIQ [55], or LeafCutter [56] with coordinate-sorted BAM files from STAR (Sect. 3.3) and the reference GTF file. Filter events by read coverage (≥ 10 junction reads), effect size ($|\Delta PSI| \geq 0.1$), and significance (FDR < 0.05). Classify events by splicing type to identify which splicing events are preferentially altered with aging (*see* **Note 16**).
2. Perform differential transcript usage analysis: Use tools such as DRIMSeq [57], IsoformSwitchAnalyzeR [58], or DEXSeq [59] with transcript-level counts from RSEM (Sect. 3.4) or Kallisto. Filter lowly expressed transcripts (e.g., TPM > 1 in ≥ 20% of samples) and genes with single isoforms. Apply statistical testing with FDR

genes with single isoforms. Apply statistical testing with FDR correction (FDR < 0.05). For tools supporting functional prediction (e.g., IsoformSwitchAnalyzeR), analyze predicted functional consequences such as the introduction of premature termination codons or upstream open reading frames (*see* **Note 17**).

3. Visualize and validate splicing changes: Generate sashimi plots using IGV with STAR BAM files to visualize junction usage for selected genes (*see* **Note 18**). Create heatmaps showing PSI values or isoform proportions across samples. Validate key splicing changes using RT-PCR with isoform-specific primers for events with potential functional consequences in aging.

---

4 Notes

1. A web-based analysis environment such as Galaxy (https://usegalaxy.org/) and THRAISE [60] can be used as an alternative to local Linux-based RNA-seq analysis. Galaxy provides graphical user interfaces for FastQC, Cutadapt, STAR alignment, featureCounts quantification, and Salmon-based transcript quantification. This option may be preferable for researchers without command-line experience, as Galaxy maintains standardized software versions, ensures consistent parameter usage, and allows reproducible execution of all major steps of the workflow. When using Galaxy, users should verify that the reference genome and annotation files selected in the tools are clearly identified and version-controlled and that the same versions are used consistently throughout the analysis. Inconsistencies between genome and annotation versions can compromise alignment accuracy and downstream quantification. Galaxy may have file size upload limits depending on server configuration. When analyzing large datasets, using FTP upload to Galaxy or performing the analysis locally may be necessary.
2. For examining noncoding RNAs, including non-polyadenylated transcripts, use total RNA library preparation rather than polyA

transcripts, use total RNA library preparation rather than polyA selection.

3. Ensure that all raw and processed data are stored and backed up properly. For publication, RNA-seq data are typically required to be deposited in a public repository such as the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) or Sequence Read Archive (SRA) (or equivalent repositories), together with sufficient processed files and metadata for reproducibility.
4. When using publicly available transcriptome data, ensure to thoroughly examine the associated metadata to confirm that the experimental design aligns with the analysis objectives. This includes verifying species, tissue type, age or condition, library preparation method, and strandedness. If the objective involves various noncoding RNAs (except long noncoding RNAs), data based on polyA selection should be avoided, as this method systematically removes non-polyadenylated transcripts during library construction.
5. QC should be performed at both the pre-alignment and post-alignment stages to detect different classes of failure. Pre-alignment QC addresses sequencing and library construction artifacts, while post-alignment QC evaluates whether reads behave as expected relative to the reference. Skipping either stage can introduce systematic bias into quantification and downstream statistical inference.
6. Pseudoalignment-based quantification (e.g., Kallisto or Salmon) can be used as an alternative to the STAR-RSEM pipeline for differential gene or isoform expression analyses when genome-level alignment information is not required. Its main advantages include substantially faster runtime, lower computational and storage requirements, and straightforward reproducibility for large-scale reanalyses of public datasets. However, this approach is generally not appropriate when read-level genome alignments

is generally not appropriate when read-level genome alignments are needed; examples include novel splice junction discovery, variant and/or RNA-editing analysis, allele-specific expression, or genome-browser visualization and detailed alignment-based quality control.

7. Reference genome and annotation files must correspond to the same genome assembly and use compatible chromosome naming conventions; mismatches reduce quantification accuracy.
8. Some species may not have complete publicly available genome annotations required for analysis. In that case, researchers can consider transferring genome annotations or coordinates from closely related species using tools such as liftOver, provided that suitable cross-species genomic alignments are available.
9. Low alignment rates may indicate contamination, poor library complexity, or incorrect strandedness.
10. Library strandedness must match aligner settings. Incorrect settings lead to inaccurate gene- or transcript-level quantification.
11. Threshold cutoffs for differentially expressed genes (DEGs) or isoforms (DEIs), such as effect size and FDR cutoffs, are study-dependent and are typically defined based on the analysis objective, experimental design, and statistical power. For exploratory analyses, these thresholds may be relaxed or tightened depending on the researcher's priorities and the desired balance between sensitivity and specificity.
12. Log_2 fold-change shrinkage in DESeq2 regularizes effect-size estimates, particularly for genes with low counts or high uncertainty, but does not provide sample-level expression values. If individual sample visualization is needed, obtain the normalized expression values calculated prior to the shrinkage step. With these normalized counts, compute condition means, calculate fold changes between groups, and use individual sample

values for custom plotting.

13. GO and KEGG analyses depend on annotation completeness; interpret low-coverage species cautiously.
14. Unlike ORA, which tests enrichment using only a DEG list, GSEA uses the ranking of all genes to test whether genes in a pathway tend to appear together near the higher- or lower-ranked part of the list, allowing pathway-level changes to be detected even when many genes do not pass the DEG cutoff individually.
15. User-friendly data visualization tools [61] may provide a more accessible alternative for users without prior experience in R.
16. Alternative splicing analysis requires high-quality RNA (RIN ≥ 7) and sufficient sequencing depth (e.g., ≥ 50 million reads per sample) for accurate isoform quantification. Ensure the same GTF file is used for STAR alignment and downstream splicing analysis to maintain annotation consistency.
17. For differential transcript usage analysis, some tools (e.g., IsoformSwitchAnalyzeR) are optimized for Salmon or Kallisto quantification and may require running these tools separately from the STAR-RSEM pipeline described in Sects. 3.3–3.4. Consult tool documentation for input requirements.
18. When visualizing splicing changes in IGV, load both the reference genome and the GTF file used for alignment to properly display gene models and annotated splice junctions alongside the observed read coverage.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (RS-2024-

00408712) to S.-J.V.L. and by the KAIST Jang Young Sil Fellow Program to G.-Y.L.

Conflict of Interest The authors declare no conflicts of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, ChatGPT-5.2 (OpenAI) and Gemini 3 (Google) were used to improve the clarity of the language. The authors reviewed and revised the content after using these tools and take full responsibility for the content of the publication.

References

1. Lee GY, Sohn J, Lee SV (2021) Combinatorial approach using *Caenorhabditis elegans* and mammalian systems for aging research. Mol Cells 44(7):425–432. https://doi.org/10.14348/molcells.2021.0080

2. Lopez-Otin C, Blasco MA, Partridge L et al (2023) Hallmarks of aging: an expanding universe. Cell 186(2):243–278. https://doi.org/10.1016/j.cell.2022.11.001 [PubMed]

3. Lee Y, An SWA, Artan M et al (2015) Genes and pathways that influence longevity in *Caenorhabditis elegans*. In: Mori N, Mook-Jung I (eds) Aging mechanisms: longevity, metabolism, and brain aging. Springer Japan, Tokyo, pp 123–169. https://doi.org/10.1007/978-4-431-55763-0_8

4. Park K, Jeon MC, Lee D et al (2024) Genetic and epigenetic alterations in aging and rejuvenation of human. Mol Cells 47(12):100137. https://doi.org/10.1016/j.mocell.2024.100137 [PubMed][PubMedCentral]

5. Kenyon CJ (2010) The genetics of ageing. Nature 464(7288):504–512. https://doi.org/10.1038/nature08980 [PubMed]

6. Rutledge J, Oh H, Wyss-Coray T (2022) Measuring biological age using omics data. Nat Rev Genet 23(12):715–727. https://doi.org/10.1038/s41576-022-00511-7

7. Ham S, Lee SV (2020) Advances in transcriptome analysis of human brain aging. Exp Mol Med 52(11):1787–1797. https://doi.org/10.1038/s12276-020-00522-6

8. Tabula Sapiens Consortium (2022) The Tabula Sapiens: a multiple-organ, single-cell transcriptomic atlas of humans. Science 376(6594):eabl4896. https://doi.org/10.1126/

science.abl4896

9. Lee C, Lee JS, Kwon Y et al (2025) Effects of heterozygous SMG1 mutations on nonsense-mediated mRNA decay in human pluripotent stem cell model. Mol Cells 48(7):100225. https://doi.org/10.1016/j.mocell.2025.100225 [PubMed][PubMedCentral]

10. Lee J, Kim JH, Chu HBK et al (2024) Comprehensive RNA-sequencing analysis of colorectal cancer in a Korean cohort. Mol Cells 47(3):100033. https://doi.org/10.1016/j.mocell.2024.100033 [PubMed][PubMedCentral]

11. Ku J, Jeong E, Gong JR et al (2025) Identification of a unique subpopulation of mucosal fibroblasts in colorectal cancer with tumor-restraining characteristics. Mol Cells 48(10):100263. https://doi.org/10.1016/j.mocell.2025.100263 [PubMed][PubMedCentral]

12. Ko WH, Kim S, Catry A et al (2025) Genome-wide statistical evidence elucidates candidate factors of life expectancy in dogs. Mol Cells 48(1):100162. https://doi.org/10.1016/j.mo1cell.2024.100162 [PubMed]

13. Tae IH, Ryu TY, Kang Y et al (2024) Negative regulation of SH2B3 by SMYD5 controls epithelial-mesenchymal transition in lung cancer. Mol Cells 47(5):100067. https://doi.org/10.1016/j.mocell.2024.100067 [PubMed][PubMedCentral]

14. Lee G-Y, Ham S, Sohn J et al (2024) Meta-analysis of the transcriptome identifies aberrant RNA processing as common feature of aging in multiple species. Mol Cells 47(5):100047. https://doi.org/10.1016/j.mocell.2024.100047

15. Ham S, Kim SS, Park S et al (2022) Systematic transcriptome analysis associated with physiological and chronological aging in *Caen-orhabditis elegans*. Genome Res 32(11–12):2003–2014. https://doi.org/10.1101/gr.276515.121

16. Ham S, Kim SS, Park S et al (2024) Combinatorial transcriptomic and genetic dissection of insulin/IGF-1 signaling-regulated longevity in *Caenorhabditis elegans*. Aging Cell 23(7):e14151. https://doi.org/10.1111/acel.14151

17. Lee SJ, Lee GY, Kim SS et al (2025) T-CLASS: an online tool for the identification and classification of aging and senescence using transcriptome data. Aging Cell 24(10):e70193. https://doi.org/10.1111/acel.70193 [PubMed][PubMedCentral]

18. Lee GY, Ham S, Lee SV (2024) Brief guide to RNA sequencing analysis for nonexperts in bioinformatics. Mol Cells 47(5):100060. https://doi.org/10.1016/j.mocell.2024.100060 [PubMed][PubMedCentral]

19. Bong D, Sohn J, Lee SV (2024) Brief guide to RT-qPCR. Mol Cells 47(12):100141. https://doi.

org/10.1016/j.mocell.2024.100141
[PubMed][PubMedCentral]

20. Lee SH, Park J, Hwang B (2024) Multiplexed multimodal single-cell technologies: from observation to perturbation analysis. Mol Cells 47(12):100147. https://doi.org/10.1016/j.mocell.2024.100147
[PubMed][PubMedCentral]

21. Kim GD, Lim C, Park J (2024) A practical handbook on single-cell RNA sequencing data quality control and downstream analysis. Mol Cells 47(9):100103. https://doi.org/10.1016/j.mocell.2024.100103
[PubMed][PubMedCentral]

22. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

23. Dobin A, Davis CA, Schlesinger F et al (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635
[PubMed]

24. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinform 12:323. https://doi.org/10.1186/1471-2105-12-323

25. Chua EW, Ooi J, Nor Muhammad NA (2024) A concise guide to essential R packages for analyses of DNA, RNA, and proteins. Mol Cells 47(11):100120. https://doi.org/10.1016/j.mocell.2024.100120
[PubMed][PubMedCentral]

26. Clough E, Barrett T, Wilhite SE et al (2024) NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. Nucleic Acids Res 52(D1):D138–D144. https://doi.org/10.1093/nar/gkad965
[PubMed][PubMedCentral]

27. Pearce M, Basutkar P, Neto RCJ et al (2025) EBI Search: providing discovery tools for biological metadata in 2025. Nucleic Acids Res 53(W1):W273–W276. https://doi.org/10.1093/nar/gkaf359
[PubMed][PubMedCentral]

28. Lee B, Hwang S, Kim PG et al (2023) Introduction of the Korea BioData Station (K-BDS) for sharing biological data. Genom Inform 21(1):e12. https://doi.org/10.5808/gi.22073

29. Fukuda A, Kodama Y, Mashima J, Fujisawa T, Ogasawara O (2021). DDBJ update: streamlining submission and access of human data. Nucleic Acids Res 49(D1):D71–D75. PMID: 33156332; PMCID: PMC7779041. https://doi.org/10.1093/nar/gkaa982

30. CNCB-NGDC Members and Partners (2025). Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2025. Nucleic Acids Res

53(D1):D30–D44. PMID: 39530327; PMCID: PMC11701749. https://doi.org/10.1093/nar/gkae978

31. Lonsdale J, Thomas J, Salvatore M et al (2013) The Genotype-Tissue Expression (GTEx) project. Nat Genet 45(6):580–585. https://doi.org/10.1038/ng.2653

32. Tabula Muris Consortium (2018) Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature 562(7727):367–372. https://doi.org/10.1038/s41586-018-0590-4

33. Bray NL, Pimentel H, Melsted P et al (2016) Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 34(5):525–527. https://doi.org/10.1038/nbt.3519 [PubMed]

34. Patro R, Duggal G, Love MI et al (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14(4):417–419. https://doi.org/10.1038/nmeth.4197 [PubMed][PubMedCentral]

35. Robinson JT, Thorvaldsdottir H, Turner D et al (2023) igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). Bioinformatics 39(1):btac830. https://doi.org/10.1093/bioinformatics/btac830

36. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550. https://doi.org/10.1186/s13059-014-0550-8 [PubMed][PubMedCentral]

37. Chen Y, Chen L, Lun ATL et al (2025) edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. Nucleic Acids Res 53(2):gkaf018. https://doi.org/10.1093/nar/gkaf018

38. Law CW, Chen Y, Shi W et al (2014) voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 15(2):R29. https://doi.org/10.1186/gb-2014-15-2-r29 [PubMed][PubMedCentral]

39. Subramanian A, Tamayo P, Mootha VK et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102(43):15545–15550. https://doi.org/10.1073/pnas.0506580102 [PubMed][PubMedCentral]

40. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T, Chang W (2022) DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res 50 (W1):W216–w221. https://doi.org/10.1093/nar/gkac194

41. Xie Z, Bailey A, Kuleshov MV et al (2021) Gene set knowledge discovery with enrichr. Curr Protoc 1(3):e90. https://doi.org/10.1002/cpz1.90 [PubMed][PubMedCentral]

42. Kolberg L, Raudvere U, Kuzmin I et al (2023) g:Profiler-interoperable web service for

functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res 51(W1):W207–W212. https://doi.org/10.1093/nar/gkad347
[PubMed][PubMedCentral]

43. Elizarraras JM, Liao Y, Shi Z et al (2024) WebGestalt 2024: faster gene set analysis and new support for metabolomics and multi-omics. Nucleic Acids Res. https://doi.org/10.1093/nar/gkae456
[PubMed][PubMedCentral]

44. Holdorf AD, Higgins DP, Hart AC et al (2020) WormCat: an online tool for annotation and visualization of *Caenorhabditis elegans* genome-scale data. Genetics 214(2):279–294. https://doi.org/10.1534/genetics.119.302919
[PubMed]

45. de Magalhães JP, Abidi Z, Dos Santos GA et al (2024) Human ageing genomic resources: updates on key databases in ageing research. Nucleic Acids Res 52(D1):D900–d908. https://doi.org/10.1093/nar/gkad927

46. Saul D, Kosinsky RL, Atkinson EJ et al (2022) A new gene set identifies senescent cells and predicts senescence-associated pathways across tissues. Nat Commun 13(1):4827. https://doi.org/10.1038/s41467-022-32552-1
[PubMed][PubMedCentral]

47. Tyshkovskiy A, Ma S, Shindyapina AV et al (2023) Distinct longevity mechanisms across and within species and their association with aging. Cell 186(13):2929–2949. https://doi.org/10.1016/j.cell.2023.05.002
[PubMed][PubMedCentral]

48. Lu JY, Simon M, Zhao Y et al (2022) Comparative transcriptomics reveals circadian and pluripotency networks as two pillars of longevity regulation. Cell Metab 34(6):836–856 https://doi.org/10.1016/j.cmet.2022.04.011

49. Son HG, Altintas O, Kim EJE et al (2019) Age-dependent changes and biomarkers of aging in *Caenorhabditis elegans*. Aging Cell 18(2):e12853. https://doi.org/10.1111/acel.12853
[PubMed][PubMedCentral]

50. Bong D, Kwon HC, Lee SV (2026) Multilayered regulation of longevity in *Caenorhabditis elegans*. Mol Cells 49(1):100308. https://doi.org/10.1016/j.mocell.2025.100308
[PubMed]

51. Bulteau R, Francesconi M (2022) Real age prediction from the transcriptome with RAPToR. Nat Methods 19(8):969–975. https://doi.org/10.1038/s41592-022-01540-0
[PubMed]

52. Jiang W, Zheng S, Chen L (2025) The implications of alternative splicing regulation for maximum lifespan. Nat Commun 16(1):10317. https://doi.org/10.1038/s41467-025-65339-1
[PubMed][PubMedCentral]

53.

Bhadra M, Howell P, Dutta S et al (2020) Alternative splicing in aging and longevity. Hum Genet 139(3):357–369. https://doi.org/10.1007/s00439-019-02094-6
[PubMed]

54. Shen S, Park JW, Lu ZX et al (2014) rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci USA 111(51):E5593–E5601. https://doi.org/10.1073/pnas.1419161111
[PubMed][PubMedCentral]

55. Vaquero-Garcia J, Aicher JK, Jewell S et al (2023) RNA splicing analysis using heterogeneous and large RNA-seq datasets. Nat Commun 14(1):1230. https://doi.org/10.1038/s41467-023-36585-y
[PubMed][PubMedCentral]

56. Li YI, Knowles DA, Humphrey J et al (2018) Annotation-free quantification of RNA splicing using LeafCutter. Nat Genet 50(1):151–158. https://doi.org/10.1038/s41588-017-0004-9
[PubMed]

57. Nowicka M, Robinson MD (2016) DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. F1000Res 5:1356. https://doi.org/10.12688/f1000research.8900.2
[PubMed][PubMedCentral]

58. Vitting-Seerup K, Sandelin A (2019) Iso-formSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. Bioinformatics 35(21):4469–4471. https://doi.org/10.1093/bioinformatics/btz247
[PubMed]

59. Anders S, Reyes A, Huber W (2012) Detecting differential usage of exons from RNA-seq data. Genome Res 22(10):2008–2017. https://doi.org/10.1101/gr.133744.111
[PubMed][PubMedCentral]

60. Heo HH, Um SJ (2025) THRAISE: an automated and reproducible web platform for RNA-seq analysis. Mol Cells 48(12):100299. https://doi.org/10.1016/j.mocell.2025.100299
[PubMed][PubMedCentral]

61. Oh BS, Kim J, Kwon M et al (2025) SimpleViz: a user-friendly, web-based tool for publication-ready data visualization in bioinformatics. Mol Cells 48(7):100222. https://doi.org/10.1016/j.mocell.2025.100222
[PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_15

15. Mitochondrial-Wide Association Studies: Integrating Population-Based Data and Mitochondrial Genetics

Yanxi Li[1], Changyeob Lee[1] and Thalida Em Arpawong[1] ✉
(1) Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA

✉ **Thalida Em Arpawong**
Email: arpawong@usc.edu

Abstract
Mitochondrial-wide association studies (MiWAS) extend genome-wide association frameworks to systematically test mitochondrial DNA variants for their associations with aging-related phenotypes. Unlike nuclear DNA, the mitochondrial genome is maternally inherited, lacks recombination, and encodes core components of oxidative phosphorylation, making it central to energy metabolism and cellular stress responses. This chapter provides a comprehensive protocol for conducting MiWAS using population-based aging cohorts, with a focus on practical implementation using publicly available datasets such as the Health and Retirement Study. We describe essential steps, including obtaining institutional approvals and data access, preparing phenotype files with appropriate covariates, extracting mitochondrial variants from genotype arrays, performing ethnic-stratified association testing, and interpreting results with multiple testing corrections. Key considerations include accounting for population stratification through mitochondrial genetic ancestry, selecting appropriate minor allele frequency thresholds, and validating findings through sensitivity analyses and replication. By making methods to assess mitochondrial genetic variation, such as MiWAS, accessible to aging researchers, this opens new opportunities to identify genetic determinants of longevity, age-related disease risk, and mitochondrial-nuclear interactions that shape healthy aging.

Keywords Mitochondrial genetics – Association studies – Human data – SNPs – Ethnicity

1 Introduction

1.1 The "Hidden" Genome of Mitochondria

Human cells contain two genetically distinct systems: the nuclear genome and the mitochondrial genome. Nuclear DNA is linear, diploid, and packaged with histones, whereas mitochondrial DNA is circular, haploid, and present in hundreds to thousands of copies per cell. The mitochondrial genome encodes 13 essential subunits of the electron transport chain (ETC), along with 22 tRNAs and 2 rRNAs required for intramitochondrial protein synthesis.

Several intrinsic features make mtDNA a hotspot for aging-related damage. Mitochondria are the primary site of oxidative phosphorylation, where electron leakage from the ETC generates reactive oxygen species (ROS). mtDNA is located in close proximity to the inner mitochondrial membrane, lacks protective histones, and has more limited DNA repair capacity than nuclear DNA. These properties contribute to higher mutation rates and accumulation of somatic mtDNA variants with age, a phenomenon long discussed within the mitochondrial theory of aging [1].

1.2 From GWAS to MiWAS

Genome-wide association studies (GWAS) systematically test common nuclear single-nucleotide polymorphisms (SNPs) across the genome for their association with phenotypes. GWAS has successfully identified thousands of loci associated with aging-related traits, yet most signals map to noncoding regions and explain only a fraction of heritability [2].

MiWAS extends this framework to test all available mitochondrial SNPs (mtSNPs) in the mitochondrial genome for their association with target phenotypes. Unlike nuclear SNPs, mtSNPs are maternally inherited, do not undergo recombination, and are present in variable proportions within cells (heteroplasmy). MiWAS therefore captures genetic variation that is not interrogated in conventional GWAS and may influence pathways central to energy metabolism, redox balance, and cellular stress responses [3].

1.3 Conceptual Distinctions and Applications of MiWAS

MiWAS follows a GWAS-like analytical approach but must be interpreted within the distinct biological framework of mitochondrial genetics. Because mtDNA encodes core structural components of the oxidative phosphorylation system, observed association signals are more likely to reflect cellular bioenergetic modulation rather than changes in gene regulatory mechanisms. In addition, heteroplasmy introduces quantitative variability in allele proportions across cells, such that the biological effect of mitochondrial variants does not follow a 0-1-2 allele effect as with nuclear genetics. This complicates assumptions about dosage effects. Finally, mitochondrial

function depends on coordinated nuclear–mitochondrial interactions, meaning that detected mtDNA associations may represent partial components of broader mitonuclear genetic architecture.

Potential applications of MiWAS include identification of mitochondrial variants associated with lifespan, age-related disease risk, and resilience to environmental stressors. In translational research, MiWAS findings can guide functional studies of mitochondrial pathways and inform stratification strategies for interventions targeting metabolism or inflammation. Several studies have applied MiWAS to investigate associations between mitochondrial genetic variation and different phenotypes of interest [3, 4].

---

2 Human Subject Research

2.1 Training, Ethical Approval, and Oversight

All investigators involved in the generation or analysis of human mitochondrial genetic data must complete institutionally required training in human subjects research and responsible conduct of research, in accordance with applicable NIH and institutional requirements (https://grants.nih.gov/policy-and-compliance/policy-topics/human-subjects/training-and-resources). This training commonly includes certification through the Collaborative Institutional Training Initiative (CITI) Program (https://about.citiprogram.org/) or an equivalent program approved by the investigator's institution. Studies involving such data require prior approval from an Institutional Review Board (IRB), including review of informed consents used to recruit participants and ensuring there is language addressing not only the key components for human subjects participation (e.g., purpose of the study, procedures, confidentiality, benefits, and risk), but also genomic data generation, storage, and sharing, and any limitations on secondary use.

2.2 Privacy, Identifiability, and Data Sharing

Because mitochondrial DNA (mtDNA) is maternally inherited and shared across extended family lineages, investigators should consider the potential for unintentional re-identification, particularly in small, isolated, or Indigenous populations. When privacy or identifiability risks are present, additional safeguards such as restricted data release, controlled-access repositories, or population-specific governance agreements should be implemented.

NIH-funded studies generating mtDNA data are subject to the NIH Data Management and Sharing (DMS) Policy (effective January 25, 2023). A Data Management and Sharing Plan (DMSP) is required and should specify the data to be generated, documentation standards, the intended repository, and whether data will be shared openly or via controlled access. Current guidance is available at:

https://grants.nih.gov/policy-and-compliance/policy-topics/sharing-policies/dms

2.3 Controlled-Access Repositories and Compliance

Human genomic data that present privacy or identifiability risks are typically shared through controlled-access mechanisms, including NIH-designated repositories (e.g., dbGaP or NIAGADS), other established data platforms, or study-specific access portals. Access is governed by Data Use Agreements (DUAs) and approval by a data access committee or authorized review group. Investigators accessing federally funded controlled-access datasets must comply with NIH security, data-use, and confidentiality requirements and remain attentive to updates in NIH genomic data sharing policy. Repository-specific guidance is available at:

https://grants.nih.gov/policy-and-compliance/policy-topics/sharing-policies/accessing-data/dbgap

---

3 Conceptual and Study-Design Preparation

3.1 Defining the Biological Question

MiWAS is a bioinformatic tool used to answer the question of whether common mitochondrial genetic variation is associated with phenotypic variation. Phenotypic variation can be defined as disease status, disease risk, or quantitative traits, including aging-related outcomes. The MiWAS tests the association between individual mtSNPs across the mitochondrial genome and target phenotypes while adjusting for relevant covariates. Although the framework is often exploratory, interpretation of significant associations should consider existing knowledge of mitochondrial biology.

3.2 Level of Mitochondrial Biology

MiWAS interrogates variation across the entire mitochondrial genome, including protein-coding genes of the electron transport chain, RNA genes, and regulatory regions such as the control region (D-loop). As a genotype-based framework, MiWAS captures DNA-level variation but does not directly assess transcriptional, proteomic, or bioenergetic phenotypes. Functional consequences therefore require independent validation.

3.3 Population Structure and Stratification

mtSNP frequencies differ across populations due to underlying ancestry structure. If overall population structure is not properly controlled, observed mtSNP–phenotype associations may reflect differences in allelic frequencies rather than biological effects. In practice, MiWAS analyses typically adjust for nuclear principal components to account for population stratification.

3.4 Nuclear–Mitochondrial Covariate Strategy

To mitigate confounding due to population stratification, MiWAS analyses typically adjust for principal components (PCs) derived from nuclear genotype data when testing associations with mtSNPs [5]. This adjustment helps ensure that observed mtSNP–phenotype associations are not driven by underlying ancestry differences.

3.5 Phenotype Definition and Time Structure

Aging phenotypes may be cross-sectional or longitudinal, continuous or categorical. Because mitochondrial dysfunction often precedes clinical disease, careful temporal alignment of phenotypes with age and exposure history is critical [6].

3.6 Environmental and Contextual Variables

Mitochondrial function is highly sensitive to environmental factors, including inflammation, diet, and toxic exposures. Incorporating contextual covariates, or confounders, can improve precision in estimating associations between mtSNPs and phenotypes of interest.

---

4 Computing Environment and Software

MiWAS are typically conducted using established GWAS toolkits such as PLINK, which support mitochondrial variant testing and regression-based association models. Users should consult official PLINK documentation to select versions compatible with their operating system and computational environment, whether on local workstations or cloud-based platforms. On macOS systems, analyses are commonly executed via the Terminal command-line interface, while high-performance computing environments enable scaling to biobank-sized cohorts.

---

5 MiWAS Protocol

5.1 Identifying and Accessing Datasets with Mitochondrial Genotype Data

Several population-based studies of aging that distribute both phenotype and genotype data provide access to approved researchers through NIH-sponsored repositories (e.g., dbGaP: https://dbgap.ncbi.nlm.nih.gov, or NIAGADS: https://www.niagads.org).

Accessing data to conduct MiWAS requires a dataset that contains both genotype and phenotype data of interest. Before applying for access to the data, researchers will want to ensure they check what technology was used to conduct genotyping or whole-genome sequencing. Many population-based studies of aging have used Illumina (https://www.illumina.com) or Affymetrix (company now called Applied Biosystems, https://www.thermofisher.com) array technology to genotype their

participants; however, not all genotyping arrays assess variants on the mitochondrial genome. Researchers will have to check the specific array names and product documentation to know whether or not mitochondrial SNPs are included, typically provided by company webpages.

One of the most widely used population-based studies of aging that includes genotyping data with mitochondrial DNA variation assessed is the U.S. Health and Retirement Study [7]. The HRS is a population-representative study of community-dwelling individuals aged 50+ and their spouses. The study began in 1992 and is ongoing, with additional data collected every 2 years. The HRS survey data (e.g., phenotypes) are available as public data and can be downloaded for use by registered users: https://hrsdata.isr.umich.edu. The first release of the HRS genotype data (collected between 2006 and 2010), which contains ~90 mitochondrial SNPs based on the Illumina Omni2.5 chip array for N~15,000 individuals, is available to researchers with an approved application to obtain HRS sensitive data (https://hrsdata.isr.umich.edu/data-products/sensitive-health) and is downloadable via dbGaP. Download instructions and information on the genotyping process are available from HRS (https://hrs.isr.umich.edu/data-products/genetic-data). Requirements and considerations for data download are provided by the dbGaP repository and outlined in Chapter 6 of the *2020 Edition Aging Methods and Protocols*. Researchers will need to gain institutional approvals and comply with Genomic Data Sharing Policies when downloading genetic data from NIH-sponsored repositories [8] (*see* **Note 1**; see also Sect. 2 for details on training requirements and data use agreements).

5.2 Trainings

Before applying for access to human genetic data, researchers must complete mandatory training in human subjects protection and responsible conduct of research. In addition, researchers working with genomic data from NIH repositories must complete specific training on the NIH Genomic Data User Code of Conduct and data security best practices. Training certifications must be current at the time of data access application (*see* **Note 2**).

5.3 Downloading and Preparing Phenotype Data

The phenotype data needed for MiWAS can be created using data manipulation software, for example, using R, SAS, STATA, or Excel. After obtaining access to survey data, researchers must identify and code their phenotype of interest. For example, in the HRS, lung disease can be defined using the wave 11 (or year 2012) survey item from the "RAND HRS Longitudinal File 2022" (https://hrsdata.isr.umich.edu/data-products/rand): "Has a doctor ever told you that you have chronic lung disease such as chronic bronchitis or emphysema?" The survey variable (R11LUNGE) uses

response options where 1 indicates having the condition, and 0 indicates not having the condition.

Recoding for PLINK format: To create a PLINK-formatted phenotype from these types of data, one would want to import a version of the data into R, then binary code the case/control dependent variable as 2 for cases, and 1 for controls. MiWAS can also be run for continuous outcomes (https://www.cog-genomics.org/plink/2.0/assoc), but here we provide an example for a case/control analysis:

```
#install the package to read SAS files
install.packages("haven")
#call the package
library("haven")
#import the SAS datafile
rand_data <- read_sas("randhrs1992_2022v1.sas7bdat")
#recode the variable
rand_data$lung_disease <- ifelse(rand_data$R11LUNG == 1, 2,
ifelse(rand_data$R11LUNGE == 0, 1, NA))
```

5.4 Selecting and Coding Covariates

Researchers should select covariates based on their relevance to the phenotype of interest. For example, age at the time of the survey (R11AGEY_M in the SAS file), biological sex reported at birth (RAGENDER), ever-smoking history as of 2012 (R11SMOKEV), and self-reported race/ethnicity (RARACEM, RAHISPAN) could be used for running within-ethnicity MiWAS.

Matching Phenotype and Genotype Data: For some studies, participant IDs must be matched to genotype IDs using a cross-reference file provided by the data repository. This ensures that phenotype and genotype data can be properly linked for each participant. For HRS, this is provided through their website for those who have approved access (https://hrsdata.isr.umich.edu/data-products/genetic-cross-reference). This matching step is crucial because the ID numbers used in phenotype files may differ from those used in genotype files.

5.5 Adding Principal Components

Principal components (PCs) are statistical measures that help control for population stratification in genetic analyses. These are often provided as precomputed quality control values alongside genotype data, or can be calculated from nuclear SNP data. For HRS, principal components PC1 through PC20 are provided by the repository and need to be merged into the phenotype file, or saved to a separate covariate file. These PCs will be used as covariates in the MiWAS to account for systematic genetic differences between subpopulations that are unrelated to the phenotype of interest, but could confound associations with mtSNPs.

5.6 Formatting Merged File for PLINK

Once all variables of interest have been identified, coded, and linked (e.g., IDs to cross-link genotypes and phenotypes, genetic ancestry PCs, phenotype, and selected covariates) into either a single phenotype data file or separate phenotype and covariate files, the file(s) must be formatted according to PLINK specifications, as shown in Fig. 1. The first two columns must contain a family ID (FID) and individual ID (IID) for the participant, followed by additional variable columns. PLINK requires specific coding conventions: binary phenotypes must be coded as 2 for cases and 1 for controls, sex must be coded as 2 for females and 1 for males, and missing data will typically be coded as −9 or −999. Researchers should consult the PLINK documentation (https://www.cog-genomics.org/plink/) for complete formatting specifications.

FID	IID	sex	Age2012	EverSmoke	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	lung_disease
10001	10001	2	62	1	0.004	0.003	0.000	0.004	-0.004	-0.004	0.004	-0.001	1
10002	10002	1	49	0	-0.002	-0.003	-0.001	0.003	0.012	0.003	-0.004	-0.006	1
10003	10003	2	81	1	0.004	0.003	0.000	0.006	-0.001	0.001	0.003	0.009	1
10004	10004	2	62	1	0.004	0.003	0.000	0.004	-0.001	-0.001	0.002	0.002	1
10005	10005	1	62	0	0.003	0.004	-0.001	0.004	0.012	0.003	-0.004	-0.006	1

Fig. 1 Representative PLINK phenotype file format for a case-control study including target covariates

The phenotype file will need to be saved as a text file. In R, this would be accomplished with the following code: `write.table(phenotype_file, "HRS_lung2012_covariates.txt", row.names = F, quote = F)`

5.7 Ethnic-Specific Analysis

Because of the need to run within-ethnicity MiWAS to address population substructure (refer to the previous section talking about mitochondrial haploidy), files also need to be created to stratify the data into subgroups. There are different ways this can be done, and researchers should understand the rationale for doing this to determine the approach that works best for their analyses based on the goals of their study and after careful consideration of the interpretation of findings. Here, we give an example of a parsimonious approach, which is to create race/ethnic-specific files based on self-identified affiliation. As shown in Fig. 2, each "keep" file will contain a list of FID and IID, one participant per line, to specify which samples to include in the within-ethnicity analysis. For example, in larger population-level studies, common stratifications include larger groups with European ancestry (or Non-Hispanic White), or Non-Hispanic Black, Hispanic/Latino, East Asian, and South Asian groups. Researchers will need to define these groups based on the participants in their study cohort, while considering study facets such as the sample size of the subgroups. Researchers are directed to PLINK documentation for proper formatting of these "keep" files (https://www.cog-genomics.org/plink/1.9/filter/).

FID	IID
10003	10003
10006	10006
10008	10008
10012	10012
10013	10013

Fig. 2 Structure of a population-stratification "keep" file for European ancestry (Non-Hispanic White) participants

5.8 Evaluating Statistical Power

Before proceeding with the analysis, check the prevalence or distribution of the phenotype within each ethnic subset to evaluate power to detect effects (*see* **Note 3**).

5.9 Accessing Mitochondrial Genotype Data

Once phenotype data have been prepared, researchers need to obtain and extract the mitochondrial SNP variants from the full genotype dataset.

5.10 Downloading Genetic Data

For HRS, genetic data are downloaded through dbGaP following the approval process described in Sect. 2.1. Detailed instructions for downloading HRS genetic data are available in the HRS tutorial at https://hrs.isr.umich.edu/data-products/genetic-data. Researchers should carefully follow the repository's download procedures to ensure data are transferred securely and in compliance with data use agreements.

5.11 Quality Control Filtering

Before extracting mitochondrial variants, researchers should refer to the Quality Control (QC) documentation provided by the data repository to identify and filter samples as recommended by the study. Most large genetic studies provide QC reports that flag samples with issues such as low genotyping call rates, sex mismatches, duplicate samples, or unexpected relatedness. These samples should typically be excluded from analysis to ensure data quality.

5.12 Extracting Mitochondrial SNPs

Once quality-controlled files have been obtained, researchers need to extract only the mitochondrial SNPs for MiWAS. In genetic data files, chromosomes are typically numbered 1–22 for autosomes and 23–24 for sex chromosomes. The mitochondrial genome is assigned a separate chromosome number, sometimes MT or chromosome 26, to provide a numeric designation. In HRS data, mitochondrial SNPs are coded as chromosome 26. Researchers should check their specific dataset documentation to determine the chromosome coding used for mitochondrial variants. The extraction is performed using the PLINK software.

The following PLINK command extracts mitochondrial SNPs from the QC'd genotype file:

```
 /plink2 --bfile HRS123_QC --chr 26 --from-bp 1 --make-bed --out HRS123_mtSNPs --recode ped --to-bp 16482
This command outputs three binary formatted PLINK files that contain only mitochondrial SNPs for subsequent MiWAS: HRS123_mtSNP.bed, HRS123_mtSNPs.bim, HRS123_mtSNPs.fam.
```

5.13 Running MiWAS Analysis

With the mitochondrial genotype data extracted and phenotype data prepared, researchers are ready to conduct the MiWAS. MiWAS for a case-control analysis can use various models to test for associations between each mitochondrial SNP and the phenotype of interest, while controlling for relevant covariates. The analysis is performed separately for each ethnic group to avoid confounding from population stratification.

The example code below provides the set of PLINK commands to run the MiWAS for the binary outcome of lung disease in a single ethnicity, here European/Non-Hispanic White ancestry. Command options will differ depending upon the desired analysis and output, and researchers should refer to PLINK documentation for more complete information on these options. The files called in by the commands may change when MiWAS is run for other ethnicities.

```
./plink2 --bfile HRS123_mtSNPs --keep HRS_lung2012_keep_nhw.txt --pheno HRS_lung2012_covariates.txt --pheno-name lung_disease --covar HRS_lung2012_covariates.txt --covar-name sex,Age2012,EverSmoke,PC1-PC8 --glm hide-covar no-x-sex --covar-variance-standardize --ci 0.95 --freq --maf 0.01 --adjust --out MiWAS_lung_nhw
```

5.14 Key Analysis Considerations

Minor Allele Frequency Filtering: The --maf option filters out rare variants that may produce unreliable association statistics due to small cell counts in contingency tables. A threshold of 0.01 (1%) or 0.05 (5%) is commonly used (*see* **Note 4**).

5.15 Multiple Testing Correction

Because MiWAS tests numerous SNPs simultaneously, there is an increased risk of false-positive findings due to multiple comparisons. The --adjust option in PLINK performs several multiple-testing corrections, including Bonferroni and false

discovery rate [9] control, to help identify which associations remain statistically significant after accounting for multiple tests.

An example output file, shown in Fig. 3, contains the association statistics for each mitochondrial SNP tested, without covariates shown (due to the hide-covar command). Each row represents the result of the logistic regression test for the respective SNP, with selected columns showing the chromosome (CHR), position (POS), SNP identifier (ID), the reference alleles for the mtSNP (REF), the alternative allele (ALT), the minor allele, which is also the effect allele tested when using PLINK2 (A1), the odds ratio (OR), confidence intervals (L95/U95), and the raw *p*-value (*P*). The odds ratio indicates the effect size of the minor allele, A1. Values above 1 suggest increased disease risk, while values below 1 suggest decreased risk. For example, MitoT1191C has an OR of 1.31, meaning the G allele is associated with 31% higher odds of lung disease compared to the A allele, adjusted for covariates. (Researchers are encouraged to consult PLINK documentation to understand all columns in the output.)

CHR	POS	ID	REF	ALT	A1	FIRTH	TEST	OBS_CT	OR	LOG.OR.SE	L95	U95	Z_STAT	P
MT	217	MitoT217C	A	G	G	N	ADD	8879	0.830	0.290	0.470	1.466	-0.641	0.5212
MT	458	MitoC458T	G	A	A	N	ADD	8852	1.027	0.155	0.757	1.392	0.170	0.8651
MT	479	MitoT479C	A	G	G	N	ADD	8879	0.907	0.215	0.595	1.383	-0.453	0.6504
MT	491	MitoT491C	A	G	G	N	ADD	8867	1.125	0.107	0.912	1.387	1.099	0.2716
MT	752	MitoG752A	G	A	A	N	ADD	8875	0.697	0.297	0.389	1.247	-1.216	0.2239
MT	1191	MitoT1191C	A	G	G	N	ADD	8874	1.309	0.136	1.003	1.707	1.985	0.0472
MT	1440	MitoG1440A	G	A	A	N	ADD	8880	0.648	0.229	0.414	1.016	-1.890	0.0587
MT	1721	MitoG1721A	G	A	A	N	ADD	8874	0.949	0.149	0.708	1.272	-0.350	0.7262

Fig. 3 Representative output from a MiWAS logistic regression analysis using the PLINK 2.0 --glm command

The example output file in Fig. 4 shows the results of basic multiple testing corrections applied to account for the number of SNPs tested, and adjustments for genomic inflation. Each row represents one SNP, sorted by the unadjusted *p*-value (UNADJ). The columns show different correction methods: Bonferroni (BONF), Holm (HOLM), Sidak (SIDAK_SS, SIDAK_SD), and false discovery rate (FDR_BH, FDR_BY). For example, MitoG16146A has an unadjusted *p*-value of 0.0063, but after FDR correction (FDR_BH = 0.2977), it no longer meets the 0.05 significance threshold. In this example, no SNPs have adjusted *p*-values below 0.05, indicating no statistically significant associations after accounting for multiple testing.

CHROM	ID	A1	UNADJ	GC	BONF	HOLM	SIDAK_SS	SIDAK_SD	FDR_BH	FDR_BY
MT	MitoG16146A	A	0.006	0.006	0.298	0.298	0.258	0.258	0.298	1
MT	MitoA10551G	G	0.035	0.035	1	1	0.815	0.808	0.441	1
MT	MitoG11378A	A	0.037	0.037	1	1	0.829	0.816	0.441	1
MT	MitoT1191C	G	0.047	0.047	1	1	0.897	0.881	0.441	1
MT	MitoG6261A	A	0.051	0.051	1	1	0.913	0.893	0.441	1
MT	MitoG1440A	A	0.059	0.059	1	1	0.942	0.921	0.441	1
MT	MitoT9699C	G	0.066	0.066	1	1	0.959	0.938	0.441	1
MT	MitoG10399A	G	0.156	0.156	1	1	1.000	0.999	0.870	1

Fig. 4 Representative multiple testing correction statistics for MiWAS results, including Bonferroni and False Discovery Rate (FDR) adjustments

5.16 Software Updates

PLINK is actively maintained and regularly updated with new features and improvements. Researchers should refer to the PLINK documentation (https://www.cog-genomics.org/plink/2.0/) for the most current command syntax and additional options that may be relevant to their specific analysis needs (*see* **Note 5**).

---

6 Post-analysis Interpretation and Validation

6.1 Signal Validation and Sensitivity Check

After identifying statistically significant associations, researchers should perform sensitivity analyses to validate the robustness of their findings. One example approach is to re-run the MiWAS analysis using different sets of covariates to ensure that the observed associations are not driven by unmeasured confounding. Associations that remain significant across multiple covariate specifications provide stronger evidence of a true genetic effect rather than confounding. However, the gold-standard approach for validation in genetic association studies is testing the SNP association with the same outcome in a completely independent cohort, with a result of a similar effect size and the same direction of allelic effect. This is essential for distinguishing true genetic signals from false positives, ensuring reproducibility, statistical significance, and biological relevance. Without such replication, any initial findings in a single cohort are considered preliminary [10].

6.2 Visualization and Reporting

Results from MiWAS can be visualized using a circular depiction of the mitochondrial genome called a solar plot, which shows the position of each SNP (as colored dots) around the circular mitochondrial genome, with the strength of association indicated by the distance from the center. This visualization allows researchers to quickly identify which regions of the mitochondrial genome harbor significant associations and whether multiple nearby SNPs show similar patterns. MiWAS results need to be imported into R, as exemplified by the code below, with some reformatting likely

necessary. Code to create a solar plot can be found online at https://github.com/stephenturner/solarplot/blob/master/solarplot.R with the example solar plot shown in Fig. 5.

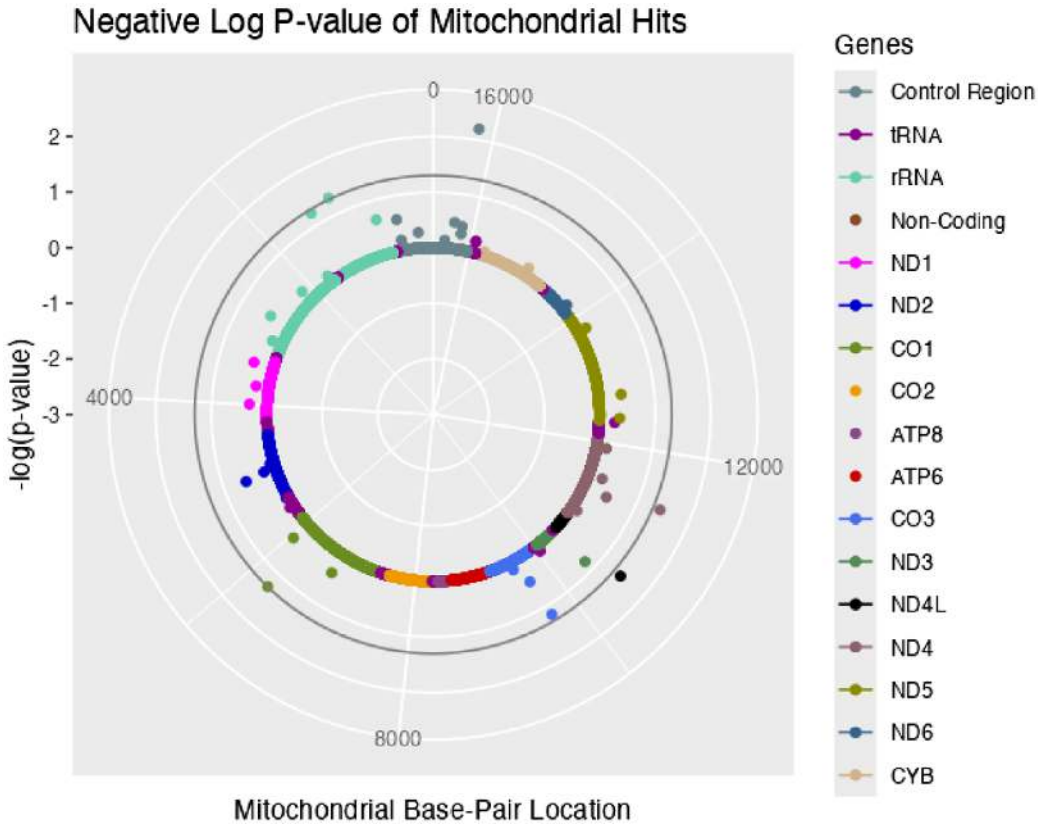


Fig. 5 Solar plot of MiWAS results for chronic lung disease. The circular layout represents the 16,569 bp mitochondrial genome, with individual SNPs plotted by genomic position

```
#first, manually remove the # of the beginning of the header
of results file or the header will not be read
results <-
read.table("MiWAS_lung_nhw.lung_disease.glm.logistic.hybrid",
header=T)
#if needed, rename position variable to fit code
results$bp <- results$POS
```

6.3 Meta-Analysis Across Ethnic Groups

Meta-analysis of results across the ethnic-specific MiWAS can be conducted to determine the generalizability of the genetic signal beyond a single ancestry group, boost statistical power by combining samples, and test for effect heterogeneity (whether the effect size differs between ethnic groups). Meta-analyses can be performed using dedicated software such as METAL (https://csg.sph.umich.edu/

abecasis/metal/), or through PLINK or R packages designed for meta-analysis of genetic association results. However, it is worth emphasizing that genetic associations found in one cohort should be followed for replication testing in an independent cohort to determine true effects.

---

7 Notes

1. The approval process typically takes several months and involves multiple steps.
2. Training certifications typically remain valid for 3 years. For detailed information on training requirements and the application process, refer to Chapter 6 Sec. 2.1 of the 2020 Edition Aging Methods and Protocols [11].
3. Power calculators such as Quanto (https://keck.usc.edu/biostatistics/software/) or R packages can help determine whether sample sizes are adequate for meaningful analysis.
4. Larger studies can reliably test variants with lower MAF, while smaller studies may need a more stringent threshold.
5. Command options and defaults may change between versions, so it is important to note which version of PLINK was used for reproducibility.

Funding

Contributions to this work were partially supported by multiple institutes and funding sources, including the National Institutes of Health through the National Institute on Aging grant (R01-AG068405), the University of Southern California Gerontology Fellowship Neurobiology Endowment (GF1014070), and the Hevolution Foundation grant (HF-AGE-23-1273964-51). The Health and Retirement Study (HRS) is sponsored by the National Institute on Aging (U01AG009740) and is conducted by the University of Michigan.

Competing Interests

All authors declare that they have no conflict of interest.

References

1. Larsson NG (2010) Somatic mitochondrial DNA mutations in mammalian aging. Nat Rev Genet 11:312–322
2. Manolio TA, Collins FS, Cox NJ et al (2009) Finding the missing heritability of complex diseases. Nature 461:747–753

[https://doi.org/10.1038/nature08494][PubMed][PubMedCentral]

3. Miller B et al (2020) A mitochondrial genome-wide association study of cataract in a Latino population. Transl Vis Sci Technol 9:25
[https://doi.org/10.1167/tvst.9.6.25][PubMed][PubMedCentral]

4. Miller B, Arpawong TE, Jiao H et al (2023) Mitochondrial DNA variation in Alzheimer's disease reveals a unique microprotein called SHMOOSE. Mol Psychiatry 28:1813–1826
[https://doi.org/10.1038/s41380-022-01769-3][PubMed]

5. Price AL, Patterson NJ, Plenge RM et al (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38:904–909
[https://doi.org/10.1038/ng1847][PubMed]

6. Wallace DC (2005) A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. Annu Rev Genet 39:359–407
[https://doi.org/10.1146/annurev.genet.39.110304.095751][PubMed][PubMedCentral]

7. Sonnega A, Faul JD, Ofstedal MB et al (2014) Cohort profile: the Health and Retirement Study (HRS). Int J Epidemiol 43:576–585
[https://doi.org/10.1093/ije/dyu067][PubMed][PubMedCentral]

8. Miller B et al (2020) Accessing and analyzing genetic data from population studies. Methods and protocols in aging research. 2nd edn. Springer, New York

9. Benjamini Y, Hochberg Y (2001) Controlling the false discovery rate in behavior genetics research. Behav Brain Res 125:279–284
[https://doi.org/10.1016/S0166-4328(01)00297-2][PubMed]

10. Gorlova OY, Xiao X, Tsavachidis S et al (2022) SNP characteristics and validation success in genome-wide association studies. Hum Genet 141:229–238
[https://doi.org/10.1007/s00439-021-02407-8][PubMed][PubMedCentral]

11. National Institutes of Health (2024) Genomic data sharing policy. National Institutes of Health, Bethesda

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_16

16. Imaging of Organelle and Actin Cytoskeletal Morphology During Aging

Aeowynn J. Coakley[1], Eric K. F. Donahue[2], Juri Kim[1], Maxim Averbukh[1], Gilberto Garcia[1], Ryo Higuchi-Sanabria[1] ✉ and Kristopher Burkewitz[2] ✉

(1) Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA

(2) Department of Cell and Developmental Biology, Vanderbilt University, School of Medicine, Nashville, TN 37232, USA

✉ **Ryo Higuchi-Sanabria (Corresponding author)**
Email: ryo.sanabria@usc.edu

✉ **Kristopher Burkewitz (Corresponding author)**
Email: kristopher.burkewitz@vanderbilt.edu

Abstract

Aging is accompanied by progressive declines in cellular and organismal function that are tightly linked to structural and functional changes in subcellular organelles. In particular, age-associated remodeling of the actin cytoskeleton, mitochondria, and endoplasmic reticulum (ER) is increasingly recognized as a hallmark of cellular aging, and numerous standardized methods exist for visualization of organelles in live cells and animals. *Caenorhabditis elegans* specifically provides a powerful model for interrogating age-dependent organelle dynamics due to its optical transparency, short lifespan, genetic tractability, and conservation of aging pathways. Here, we present a comprehensive collection of optimized protocols for imaging and quantifying actin cytoskeletal organization, mitochondrial morphology, and ER structure across tissues and throughout adulthood in *C. elegans*. We describe genetically encoded fluorescent reporters, sample preparation strategies, and microscopy approaches tailored to each organelle and tissue, including wide-field and confocal imaging. In addition, we outline quantitative and semi-quantitative analysis pipelines, ranging from blinded manual scoring to automated image analysis workflows, that enable reproducible quantitative analysis of age-associated structural changes. Where appropriate, we discuss functional assays that complement imaging-based readouts and provide insight into organelle health. Together, these methods offer a practical and scalable framework for studying organelle structure-function relationships during aging.

Keywords Aging – *Caenorhabditis elegans* – Organelle morphology – Actin cytoskeleton – Endoplasmic reticulum – Mitochondria – Fluorescence imaging

The authors "Aeowynn J. Coakley" and "Eric K. F. Donahue" contributed equally to this work.

1 Introduction

1.1 *Caenorhabditis elegans* as a Model System for Aging Biology

Caenorhabditis elegans (*C. elegans*), a nematode, is a well-established model organism for diverse research across biology. Originally founded in the 1940s [1], *C. elegans* was pioneered by Dr. Sydney Brenner in the 1960s [2] as an indispensable tool for applications in developmental biology, genetics, neuroscience, and aging. Several aspects make *C. elegans* an ideal model system in the laboratory. First, they can be easily and inexpensively cultivated in the laboratory environment due to their short life cycle and ability for a single hermaphrodite animal to produce several hundreds of isogenic progeny. In addition, the ability to grow large cultures of animals in both liquid and solid medium in a simple laboratory incubator makes them easily adaptable across diverse scientific institutions [3, 4]. For the methods described in this chapter, another ideal characteristic of *C. elegans* is its transparent and microscopic body, allowing for imaging in intact, living organisms across the lifespan [5]. Further, they are composed of multi-cellular anatomy, with defined tissues including hypodermis, muscle, digestive, excretory, reproductive, and nervous systems, which allow for tissue-specific visualization of organelles [5]. Finally, as a genetically tractable organism, genetically encoded fluorescent protein fusion tags allow for visualization of diverse organelles through aging in live animals [6].

In addition to these technical strengths, decades of research have thoroughly mapped out numerous mechanistic pathways in the worm. For example, insights from developmental biology help to shape our understanding of areas such as signaling pathways and regulatory mechanisms [7]. The cell lineage in *C. elegans* is virtually invariant, with the pattern of cell division and cell fate of each cell being similar and predictable between individual animals. Thus, developmental trajectories have been fully mapped, facilitating investigation into gene expression patterns of each cell [8]. In addition, major technological advancements in the development of a genetic toolkit allow for genetic manipulations via gene editing to create mutations or tag proteins; a fully sequenced genome allowing for large-scale, targeted multi-omics studies; and genome-wide RNA interference (RNAi) libraries to allow for mechanistic investigation into gene functions [5]. *C. elegans* also has highly conserved biology, allowing researchers to make discoveries relevant for human biology, especially in aging biology, where virtually all aging pathways are conserved from the worm to humans, including insulin/IGF-1 [9], TOR [10], MAPK [11], TGF-β [12], and Wnt [13]. These are fundamental signaling pathways that regulate metabolism, longevity, stress response, immunity, and development, and using *C. elegans*, we can better determine the mechanisms involved in each as they directly relate to human health.

Here, we offer a collection of methods that leverage the strengths of *C. elegans* as a model system for organelle imaging. Often, organelle structure and morphology can be used as a proxy for measuring the health of an animal during the aging process. As cellular processes

decline with aging [14], robust and dynamic changes occur to numerous organelles, which are directly correlated with organelle function and cellular health. For example, during aging, the actin cytoskeleton becomes disorganized [15–18], the endoplasmic reticulum (ER) deteriorates in structure and volume loss [19], and mitochondria become fragmented [20]; all of these changes can be mapped with microscopy in vivo in *C. elegans* with age.

1.2 Form and Function of Actin During Aging

The actin cytoskeleton, a dynamic structure composed of monomeric globular actin (G-actin) and filamentous actin (F-actin), the balance of which is tightly regulated by actin-binding proteins [21], is a critical component of the cell, providing the scaffolding in which all other organelles exist. Actin has well-established roles in cellular structure, transport, movement [22], and organelle dynamics, such as in mitochondria-endoplasmic reticulum contact sites [23, 24]. The actin cytoskeleton also plays emerging roles in transcription [25, 26], stress responses [27, 28], and DNA damage repair [29, 30]. Actin provides the scaffold and is critical for regulating the dynamics of several major organelles. For example, actin-mitochondria-ER contact sites regulate mitochondrial morphology through both mitochondrial fusion and fission [31–33]. The ER first positions the mitochondria for fission, marking the fission sites. Localized to the ER, an actin regulatory protein, inverted formin (Inf2), and a mitochondrial actin nucleator, Spire1c, recruit actin-myosin assemblies to mitochondria and begin constriction [31, 34–36]. Drp1, the mitochondrial division dynamin, is then recruited to these ER-mitochondria constriction sites from the cytoplasm by the ER and drives constriction around the outer mitochondria matrix via GTPase-dependent oligomerization [37]. Lastly, dynamin 2 (Dnm2), another dynamin, completes fission at the ER-mitochondria constriction site [38]. In mitochondrial fusion, the Arp2/3 complex, responsible for the branching of actin filaments [39], is required, wherein actin polymerization on the mitochondria occurs prior to recruitment of mitofusion 2 (Mfn2), a mitochondrial membrane protein, to the fusion site [31]. While the mechanism of involvement of ER in mitochondrial fusion remains to be fully elucidated, there is an enrichment of actin-ER membrane contact sites (MCS) at actin-mitochondria fusion sites, with actin-ER MCS recruited to actin-mitochondria MCS prior to or during fusion, and this is required for successful fusion [31].

A variety of imaging techniques have been developed to visualize the actin cytoskeleton across models [40]. One of the most commonly used methods is phalloidin, a small molecule isolated from the deathcap mushroom *Amanita phalloides*, which binds to F-actin with high affinity and selectivity [41]. Fluorescently-conjugated phalloidin is used to visualize actin in fixed cells and tissues [42]; similarly, standard immunofluorescence antibody staining with actin antibodies can be used as well. The actin cytoskeleton is involved in dynamic cellular processes, and thus, methods to visualize actin in living cells have been developed using actin-binding domains (ABDs) bound to fluorescent molecules. For example, SiR-act is derived from Jasplakinolide, a toxin with a high affinity for f-actin found in the marine sea sponge *Jaspis johnstoni* [43].

For genetically tractable systems, there are several genetically encoded ABDs that enable dynamic actin visualization. LifeAct, a 17-mino acid long peptide from the *S. cerevisiae* Abp140p, is the most commonly used genetic option and allows for visualization without

disruption of actin dynamics; however, if overexpressed at high enough levels, actin dynamics can be altered [44, 45]. F-Tractin is a 43-amino acid long peptide derived from the rat actin-binding inositol 1,4,5-triphophotase 3-kinase A, and while it can be used similarly to LifeAct, detrimental organismal effects have been observed, potentially due to its larger size [46]. In addition, Utrophin, a 261 amino acid peptide from the tandem calponin homology domains of utrophin, is used to visualize actin with detrimental effects being context specific to the organism or cell type [47]. An emerging technique involves the use of nanobodies, single-domain antibodies, like actin chromobody, which is a genetically encoded nanobody against actin fused to a fluorescent molecule [40]. Depending on the requisite method, there are several options for visualizing actin in a physiologically relevant context.

Actin disruption during aging is well-documented in invertebrates. In *S. cerevisiae*, the stability of F-actin declines with age, disrupting cell dynamics such as the transport of organelles, like mitochondria, from mother to daughter cell and resulting in reduced mitochondrial function and causing premature replicative senescence [16]. The actin cytoskeleton also shows a loss of structure and integrity during aging in *C. elegans* [15]. When actin and actin-binding proteins are inhibited via RNAi, their lifespan is decreased [18], and conversely, when actin function is preserved, lifespan and healthspan are improved [28]. In the *D. melanogaster* brain, F-actin increases with age along with the formation of actin-rich rod structures; actin levels correlated with metrics of healthspan, and pro-longevity interventions (dietary restriction, rapamycin) reduced actin aggregates and improved animal health [17]. Cytoskeletal abnormalities in the brain are a hallmark of neurodegenerative disorders, such as Alzheimer's disease (AD) [48].

C. elegans is an ideal model to utilize for imaging actin, and they also provide the ability to measure the functional consequences of alterations in actin. Measuring locomotory behavior involves putting the animals in a liquid medium and counting their "thrashes," the number of times their body transitions from concave to convex [49]. Considered a standard assessment of healthspan, the capacity to thrash declines with age [49], reflective of a decline in muscle function. As actin structure is an essential component of muscle function [50], thrashing measurements are an informative functional output of actin health. In addition, the actin cytoskeleton is foundational to gut integrity [51], and this intestinal barrier becomes dysregulated with age in an evolutionarily conserved manner [52, 53], reflective of the decline in actin function with aging [54]. Intestinal barrier integrity can be measured in *C. elegans* via microscopy imaging assays. In the "Smurf" assay, animals are fed their standard bacterial diet mixed with blue food coloring and the subsequent presence of blue dye in the body cavity, thus indicating a decrease in gut barrier integrity, is assessed using brightfield microscopy [55]. Similarly, *C. elegans* can be fed a diet of fluorescent bacteria, thus allowing for identification of infiltration and colonization of the gut via fluorescent microscopy [54, 56]. Beyond mapping actin alterations via imaging, these experiments provide functional readouts that demonstrate the physiological consequences of changes in actin.

1.3 Mitochondria and Aging

Mitochondria are hubs for cellular metabolism and energy production, as well as mediators for stress response, lipid oxidation, cell cycle regulation, cell death regulation, and calcium and amino acid storage, all of which often decline during aging [57–59]. During aging, mitochondria gradually lose their membrane potential, show a decline in the respiratory capacity, and accumulate mitochondrial DNA mutations [60–62]. Furthermore, mitochondria experience noticeable changes in their morphology as well, which is often associated with decreased mitochondrial capacity [63]. For example, mitochondria in human skeletal muscle from old individuals displayed remodeled morphology and a reduction in network complexity compared to younger individuals [63]. Another study reported that mitochondria in young proliferative human umbilical vein endothelial cells (HUVECs) show long, tubular structures, while they became aggregated and less uniformly tubular, and often appeared poorly organized with swollen matrix regions lacking cristae [64]. The morphology of mitochondria is also clearly altered in age-related diseases such as AD. Postmortem brain samples from AD patients show reduced size and disrupted or even lost structure of cristae, and the neurons and fibroblasts from sporadic AD patients displayed fragmented, shorter, and abnormally distributed mitochondria [65, 66]. Moreover, overexpressing amyloid-β precursor protein showed a similar change in mitochondrial morphology, suggesting the link between abnormal dynamics and AD pathology [67]. Given the numerous functions of mitochondria, impaired mitochondria with age can lead to a reduction in energy production, accumulated oxidative stress, impaired stress responses, and induction of cellular senescence [60, 68–70].

Changes in mitochondrial morphology serve as a proxy for evaluating the function of mitochondria, as morphology is often strongly correlated with functional output. For mitochondrial imaging, fluorescence imaging is the most common and high-throughput method, although more advanced methods exist, including electron microscopy (EM), atomic force microscopy (AFM), and cryo-EM for specialized applications such as resolving cristae or mitochondrial stiffness [71]. For fixed samples, immunofluorescence (IF) with mitochondrial antibodies can also be used, but IF may introduce artifacts and background noise, and can be time-consuming [71]. As an alternative, dye-based methods have also been developed that work in fixed tissue, including iridium coordination complex and PK Mito Orange FX (PKMO FX), which is specific to the mitochondrial inner membrane [71, 72]. For live imaging, there are commercially available dyes that accumulate in the negatively charged mitochondrial matrix, such as JC-1, tetramethylrhodamine ethyl ester (TMRE), and MitoTracker, which are more commonly used [73–76].

In *C. elegans*, genetically encoded mitochondrial-localized fluorophores are often used due to the advantage of their easy genetic manipulation and live imaging [77]. Previously, we developed strains expressing single-copy, matrix-localized fluorescent proteins generated by integration into safe harbors using MosSCI. The MLS of ATP-1 was fused with GFP and was expressed by cell-type-specific promoters, and we have validated that these strains can robustly visualize mitochondrial morphology in a tissue-specific manner [77]. During aging, it was shown that mitochondria of *C. elegans* display fragmentation, most markedly in the intestine, followed by the hypodermis and muscle [20]. As opposed to the linear, tubular morphology found in young animals, mitochondria in aged worms appear more truncated and spherical [20].

Quantifying changes in mitochondrial morphology can present challenges, particularly due to the variability of morphology across tissues. Therefore, automated quantification methods are attractive to increase objective and reproducible quantification across samples. For example, MitoMAPR [78] is an ImageJ/FIJI macro that uses image enhancement filters to improve the signal clarity, and skeletonizes the fluorescent mitochondrial signals in a binary way, then converts the morphology into different parameters such as object length and junction points, which represent mitochondrial length and network, respectively. This protocol can robustly quantify mitochondrial morphology in *C. elegans*, although with highly fragmented or hyperfused network, some quantification artifacts can arise [77]. Although changes in mitochondrial morphology are an excellent first-pass to determine age-associated changes, it does not provide a comprehensive profile of mitochondrial function. Therefore, imaging data can be complemented with other mitochondrial functional assays, including Seahorse assay to measure respiratory capacity, JC-1 or TMRE staining to investigate membrane potential, quantification of mtDNA copy number, and/or detecting mtDNA mutation via NGS, providing a comprehensive context of mitochondrial changes with aging [79–82]

1.4 ER Structure and Function During Aging

The ER is one of the most complex organelle networks in cells, and often the largest. The ER houses myriad essential processes, including protein synthesis and secretion, lipid and membrane synthesis, carbohydrate metabolism, xenobiotic detoxification, and calcium storage and release. More recently, the ER has also emerged as a central hub for regulating the dynamics and functions of other subcellular compartments through its propensity for forming membrane contact sites. These contact sites serve as functional microdomains with diverse roles that include acting as a signaling hub, facilitating non-vesicular transfer of lipids between membranes, and scaffolding the fission/fusion processes of other organelles [83, 84]. As described above, one of the most studied examples of inter-organelle contacts involves the formation of an actin and tubular ER scaffold around mitochondria to facilitate mitochondrial fission [35, 36, 85].

To carry out such diverse physiological functions, the ER consists of multiple structurally and functionally distinct subdomains, including the nuclear envelope, ER sheets or cisternae, and ER tubules. While the ER network is generally thought to be fully interconnected, the network and its protein and lipid composition are not homogeneous. Many proteins are enriched within particular subdomains, typically either playing a role in maintaining a distinct, subdomain-specific morphology or performing a localized biochemical activity. For example, nuclear pore complex proteins selectively localize to the nuclear envelope; ER protein translocation machinery is enriched in ER sheets; and lipid and sterol synthesis enzymes are enriched in ER tubules [86, 87]. Structurally, proteins such as p180 or CLIMP63 control the spacing of ER sheets, a morphology thought to be optimized for protein maturation and polysome docking. Notably, however, orthologs of these spacing factors have not yet been identified in *C. elegans*. On the other hand, hairpin-domain proteins such as reticulon and Dp1/Yop1p families stabilize the high curvature of ER membranes at the edges of ER sheets and throughout tubules, a geometry that supports lipid exchange and lipid droplet biogenesis [87–89]. Understanding these functionally

specialized subdomains and the proteins enriched in each has provided insights into tools to visualize them in vivo.

Many cellular processes housed partly or fully by the ER are strongly linked with aging, including unfolded protein response (UPR) signaling, protein translation and folding, calcium homeostasis, lipid metabolism, and autophagy [90, 91]. However, the morphological dynamics of the ER through age are much less explored. The best understood age-associated structural shift of the ER focuses on the nuclear envelope, which becomes progressively convoluted and distorted [92, 93]. Similar structural abnormalities are observed in human progeria models, where downstream alterations in chromatin organization and nucleocytoplasmic transport provide a mechanistic link between nuclear envelope architecture and age-associated cellular dysfunction [93, 94].

More recent work has highlighted age-associated structural dynamics in the peripheral ER subdomains as well [19, 95]. A study employing a combination of yeast, worm, and mammalian models demonstrated a widespread decline in total ER volume with age, and in certain cell types, a transition from densely packed rough ER sheets to a predominantly tubular network [19]. ER-phagy drives these structural changes, which intriguingly align with established functional shifts, including declines in rough ER-associated proteostasis and alterations in ER tubule-associated lipid processing. Related work demonstrated that ER structure in adult intestinal cells is highly plastic and extensively remodeled in response to genetically enforced activation of the UPR transcription factor XBP-1s, which extends lifespan and promotes lipophagy-dependent lipid depletion [95, 96]. Thus, both the normal aging process and interventions that extend healthspan converge to modulate ER structure. These studies highlight an exciting frontier in understanding how ER structure and function are linked to one another and how this relationship evolves across life history.

While imaging ER morphology can serve as a powerful proxy for ER health and function in *C. elegans*, for a comprehensive measurement of ER function, complementary methods should be performed. As the ER is one of the primary sites for cellular lipid synthesis [97] and the UPR^{ER} alters lipid metabolism [56, 98], measuring lipid levels can provide insight into ER function; using dehydrogenase short-chain 3, an abundant protein on the surface of intestinal lipid droplets tagged with green fluorescent protein (DHS-3::GFP), the morphology and amount of lipid droplets can be quantified in vivo [18, 56, 99]. The ER is also a site of the secretory pathway of the cell, and a common reporter for secretory capacity in *C. elegans* uses a vitellogenin protein tagged with a green fluorescent protein (VIT-2::GFP). Receptor-mediated endocytosis of VIT-2::GFP can be measured as a readout of ER secretory capacity [56, 100]. Finally, as an indirect measure of the ER's role in protein biosynthesis, measuring proteostasis serves as insight into ER function [56, 98]. Using genetic models with tissue-specific aggregation-prone polyglutamine repeats of different lengths tagged with a yellow fluorescent protein (e.g., polyQ40::YFP, polyQ44::YFP), the resulting fluorescent foci or aggregates can be quantified with a higher number of foci indicating lower proteostatic capacity of the tissue [56, 101]. Redox state can also be measured using an ER-localized roGFP measured as ratiometric quantification of luminal fluorescence of roGFP's oxidized and reduced state [102]. Ca^{2+} homeostasis can be measured using an ER-localized GCaMP measured as fluorescence intensity normalized to an ER marker [103]. Finally, measurements of organismal resistance to ER stress can be

used as an indirect readout for ER health. For example, in *C. elegans*, ER stress can be induced via treatment with tunicamycin, which inhibits N-linked glycosylation and causes a buildup of misfolded proteins in the cells [104], and with thapsigargin, with inhibits sarcoplasmic/endoplasmic reticulum calcium-ATPase (SERCA), causing an imbalance of calcium ions in the ER versus the cytoplasm [104].

Overall, while organelle morphology can be used as a quick and efficient method to profile changes to cellular health with age, comprehensive profiling of each organelle is recommended for more thorough investigations. The protocols provided here allow for relatively easy and quick interrogation that are amenable for initial discovery experiments and for large-scale screening.

---

2 Materials

2.1 Microscopes

1. Leica Stellaris 5 confocal microscope equipped with a white light laser source and spectral filters, HyD detectors, 63×/1.4 Plan ApoChromat objective, and run on LAS X software
2. Leica THUNDER Imager equipped with a 63×/1.4 Plan ApoChromat objective, standard dsRed filter (11525309), Leica DFC9000 GT camera, a Leica LED5 light source, and run on LAS X software
3. Eclipse Ti2 inverted microscope with a Yokogawa CSU-W1 spinning disc (Nikon), a Plan Apo λ objective (10×/0.45 or 100×/1.45) (Nikon), a Prime 95B sCMOS camera (Teledyne Photometrics), and run on Nikon Elements software

2.2 Reagents

1. 0.375 M KOH bleach solution: 1.8% (v/v) sodium hypochlorite
2. 100 mg/mL carbenicillin stock solution (1000×): prepared in water
3. 10 mg/mL 5-fluoro-2'-deoxyuridine (FUDR): prepared in water
4. 1 M isopropylthiogalactosidase (IPTG): prepared in water
5. Lysogeny Broth (LB): LB Media broth powder resuspended in water
6. M9 solution: 22 mM KH_2PO_4 monobasic, 42.3 mM Na_2HPO_4, 85.6 mM NaCl, 1 mM $MgSO_4$
7. Nematode Growth Media (NGM) plates: 1 mM $CaCl_2$, 5 µg/mL cholesterol, 25 mM KPO_4 pH 6.0, 1 mM $MgSO_4$, 2% (w/v) agar, 0.25% (w/v) Bacto-Peptone, 51.3 mM NaCl
8. Nematode Growth Media RNAi plates: 1 mM $CaCl_2$, 5 µg/mL cholesterol, 25 mM KPO_4 pH 6.0, 1 mM $MgSO_4$, 2% (w/v) agar, 0.25% (w/v) Bacto-Peptone, 51.3 mM NaCl, 1 mM

pH 6.0, 1 mM $MgSO_4$, 2% (w/v) agar, 0.25% (w/v) Bacto-Peptone, 51.3 mM NaCl, 1 mM IPTG, 100 μg/mL carbenicillin/ampicillin

9. 1 M sodium azide: prepared in water
10. 10 mg/mL tetracycline: stock solution (500×) in 100% Ethanol

---

3 Methods

In the laboratory, *C. elegans* are usually grown on solid agar medium in plates with the bacterium *Escherichia coli* (*E. coli*) as a food source. The two bacterial *E. coli* strains most commonly used are OP50, a B strain, for routine maintenance and HT115, a K-12 strain, for RNA interference (RNAi) experiments. HT115 carries an *rnc* null mutation, the gene that encodes for RNAse III, and an isopropylthiogalactosidase (ITPG)-inducible T7 RNA polymerase gene [105]; thus, dsRNA accumulates in the HT115 bacteria, which is then consumed by the worm. Feeding the worms dsRNA is a simple method that allows for robust knockdown of genes of interest [106] and is well-supported by the scientific community, as the Vidal [107] RNAi library contains over 19,000 RNAi constructs, covering 55% of unique protein predicted-coding genes, and the Ahringer [108] RNAi library with almost 17,000, covering 87% of annotated genes [49, 109].

While both strains are *E. coli*, these bacterial diets differ in metabolic composition, resulting in differences in physiology for animals on each diet [110], including metabolic profile, mitochondrial function, and lifespan [111]. Notably, OP50 is deficient in vitamin B12, which is linked to impaired mitochondrial hormesis and increased susceptibility to stress and pathogens [112]. Therefore, while animals can be grown on OP50 for convenience, we recommend performing all experiments using the HT115 background to avoid any of these confounding effects, regardless of RNAi requirements of the experiment. In all experiments we describe, animals are maintained and expanded on OP50, then moved to HT115 following synchronization (*see* **Note 1**).

3.1 Maintenance and Expansion

C. elegans can be maintained and expanded in liquid or on agar plates [49, 110, 113]. All experiments described here use animals grown on agar plates (*see* **Note 2**).

1. Inoculate lysogeny broth (LB) with OP50 or equivalent media and grow culture for 24 h at ambient temperature, usually between 22 °C and 25 °C.
2. Seed OP50 culture onto NGM plates: 200 μL onto 60 mm size plates and 800 μL onto 100 mm size plates. Allow bacterial lawn to dry completely at ambient temperature. Plates can be stored at 4 °C in sealed containers for approximately 2 months.
3. For standard maintenance, move 10–15 young animals (eggs, L1, and/or L2) onto 1 60 mm plate. For maintenance of strains with low fecundity, more animals should be moved. Animals are moved weekly at 15 °C and bi-weekly at 20 °C.
   For expansion, the agar of a 60 mm plate of animals can be cut into chunks and placed

4. For expansion, the agar of a 60 mm plate of animals can be cut into chunks and placed animal-side down onto the bacterial lawn on a 100 mm plate. For standard strains, a full 60 mm plate can be cut into four to six equal parts and placed onto four to six 100 mm plates; kept at 20 °C, this will generate full plates within 2–3 days.

3.2 Synchronization

Animals are synchronized via bleaching (*see* **Note 3**). One gravid adult animal will yield approximately 10–15 eggs [114]; as a full 60 mm standard NGM plate can support 100–200 adults, bleaching provides between 1000 and 2000 eggs [49].

1. Using a glass serological pipette, collect the animals into a 15 mL conical tube by first pouring a small amount of M9 solution onto the agar plate and gently agitating the plate to dislodge animals from the bacterial lawn. Glass, not plastic, pipettes are recommended as animals will stick to the plastic.
2. Pellet the animals via centrifugation for 30 s to 1 min in a centrifuge set to 1.1 RCF, then aspirate off the M9 solution.
3. Add 5 mL of bleaching solution (see Sect. 2.2) to the pelleted animals and monitor under a dissection microscope until all adult bodies have been dissolved. Shake vigorously to accelerate this process. Wild-type animals usually dissolve to release eggs within 4–6 min (*see* **Note 4**).
4. Pellet the eggs via centrifugation for 30 s to 1 min in a centrifuge set to 1.1 RCF, then aspirate off the bleach solution.
5. Add 15 mL of M9 solution to resuspend the eggs and repeat Step 5 until the eggs have undergone four washes of M9 solution to remove any residual bleach.
6. Eggs can then be plated by resuspending with 100 µL to 2 mL M9 solution, depending on the amount of worms bleached; thoroughly shake the tube to separate egg clumps and uniformly resuspend. Animals from these eggs can also be L1- arrested by resuspending the pellet in 10–12 mL M9 solution and put on a rotator for aeration for up to 24 h at ambient temperature in a 20 °C incubator.

3.3 Aging Animals

While *C. elegans* has many benefits for aging biology, one challenging aspect is that as a hermaphrodite, a single animal can produce hundreds of progeny. Thus, for aging experiments, the researcher is burdened with removing progeny for experimentation. For most imaging experiments, only a small population of animals is required, and thus, manual manipulation of adult animals away from their progeny by hand-picking adults onto a new plate is the simplest method. However, for actin imaging [15] and mitochondrial imaging [77], previous studies have shown that chemical sterilization using FUDR, which inhibits DNA replication [115], did not significantly affect organelle morphology with age, although it did slightly delay the age-associated disorganization of structures. Therefore, we provide

both manual manipulation and chemical sterilization as methods, although other methods are described elsewhere [110, 116].

Manual Manipulation of Animals from Progeny with a Standard Pick

1. Prepare plates with bacteria of choice (see Sect. 3.1).
2. Synchronize animals of choice using methods described in Sect. 3.2 and plate animals.
3. Incubate eggs at 20 °C for 3–4 days (65–96 h) until day 1 of adulthood, depending on the strain fecundity and speed of development.
4. On day 1, transfer animals onto separate plates. We recommend placing 20–25 animals on each plate.
5. For the first 3–4 days, animals should be moved away from progeny manually using a pick every day. From day 5, animals can be moved every other day until progeny are no longer visible. After progeny are no longer visible, animals can be kept on plates and only moved if there is a risk of starvation.

Chemical Sterilization of Adult Animals

1. Prepare plates spotted with bacteria of interest (see Sect. 3.1).
2. Synchronize animals of choice using methods described in Sect. 3.3 and plate animals onto spotted plates.
3. On a second set of plates, add 10 mg/mL FUDR directly to the center of the bacterial lawn and swirl the plates to ensure FUDR mix spreads evenly across the bacterial lawn. We recommend 100 µL for 60 mm plates and 500 µL for 100 mm plates. Let dry overnight.
4. Incubate eggs at 20 °C for 2 days (~56 h) to the L4 stage.
5. Wash L4 animals off plates using M9 into a 15 mL conical tube. Centrifuge at 1000 × *g* for 30 s to pellet and then aspirate the M9 solution.
6. Plate L4 animals onto the second set of plates containing FUDR. Animals can also be moved using a pick onto FUDR-containing plates. Animals can now be grown on FUDR plates until the desired age for experimentation. If there is a risk of starvation, animals can be moved onto a new FUDR-containing plate.

3.4 Preparation of Slides for Imaging

1. Prepare animals to the desired age following the protocol described in Sect. 3.3.

2. Prepare a fresh working stock of 100 mM sodium azide by dilution of a 1 M stock 1:10 into M9 solution (*see* **Note 5**).
3. Place 10–18 µL of 100 mM sodium azide onto a glass microscope slide (*see* **Note 6**).
4. Manually move 10–25 animals from NGM plates and gently place them into the sodium azide solution on the glass slide using a wire pick.
5. Gently apply a cover slip onto the sample, taking care not to introduce any air bubbles or apply pressure onto the worms, as that may cause the animals to be damaged.
6. The edges of the cover slip can be sealed with Valap wax or nail polish if desired for longer-term imaging.

3.5 Imaging of the Actin Cytoskeleton in *C. elegans*

Here, we describe optimized approaches for imaging actin structures in body wall muscle, hypodermis, and intestine of *C. elegans* using fluorescent LifeAct reporters (see **Sect. 7;** Table 1) [15, 18], with guidance on microscope selection, animal orientation, and quantification for each of these tissues. Wide-field microscopy provides sufficient resolution for actin structures in body wall muscle and hypodermis due to the superficial location and planar geometry of these tissues. Visualization of intestinal actin requires the resolution provided by confocal microscopy, given the depth of the tissue in the animal and the actin organization in this tissue. Slide preparation for actin imaging generally includes immobilization of the animals with sodium azide (see Sect. 3.4). While we provide examples of how to image actin structure in these tissues, optimization of imaging parameters specific to the microscope being used is recommended.

Table 1 Strains referred to in this protocol

Imaging of Actin Structure Using a Compound Microscope

1. Use a standard wide-field microscope for visualization of muscle and hypodermal actin structure (*myo-3p::LifeAct::mRuby* and *col-19p::LifeAct::mRuby*).
2. Optimize the imaging settings for each microscope and experimental setup. As a starting point, the following parameters are used on our Leica THUNDER Imager: Dimensions are 2048 (x)*2048 (y) pixels at a step size of 0.50 µm. *Z*-range starts from where the signal is visible but slightly out of focus and ends where it is visible and slightly out of focus. To minimize background noise in maximum-intensity projections, reduce the number of *z*-planes included in the projection where possible. A 63×/1.4NA oil-immersion objective is used and 300 ms exposure time with a 552 laser at laser power 48%.
3. Acquire images from a minimum of five animals per condition per biological replicate. It is recommended to image animals at the same location in the body each time. In our studies, we perform imaging in the anterior body region between the pharynx and the proximal gonad. For hypodermal actin imaging, orient animals in a slightly lateral

proximal gonad. For hypodermal actin imaging, orient animals in a slightly lateral position so that the seam-associated hypodermis is positioned at its maximal apparent thickness relative to the imaging plane. For muscle actin imaging, orient animals such that one body wall muscle quadrant is positioned parallel to the imaging plane, allowing visualization of sarcomere actin at maximal thickness.

4. Actin structure in the muscle is generally very stable and may require imaging on day 15 for aging experiments. We recommend imaging at: day 1, 5, 9, 13, and if necessary, days 15 and 17. Actin structure in the hypodermis goes through distinct structural changes during early adulthood and should be imaged at days 1, 4, and 8 to meaningfully capture these changes.

Imaging of Actin Structure Using a Confocal Microscope

Use a confocal microscope for visualization of intestinal actin structure (*gly-19p::LifeAct::mRuby*).

1. Optimize the imaging settings for each microscope and experimental setup. As a starting point, the following parameters for imaging were used in a Leica Stellaris 5 confocal: a white light laser was set to 85.00% Maximum Power and with a 558 nm laser line at an intensity of 4.00%. The HyD S detector was set to 563–687 nm with a gain of 2.5 in the analog setting. We unidirectionally scanned an area of 82.01 µm (x)*82.01 µm (y) with a z step size of 0.5 µm with a scanning speed of 1000 Hz, 2.25 zoom, line average of 2, and pinhole of 1 AU (95.5 µm). Collect *z*-stacks beginning slightly above the focal plane where the signal first becomes visible and ending slightly below the plane where the signal is no longer in focus. To minimize background noise in maximum-intensity projections, reduce the number of *z*-planes included in the projection where possible.
2. Acquire images from a minimum of five animals per condition per biological replicate by imaging in the anterior intestine immediately posterior to the pharynx.
3. Actin structure in the intestine is generally very stable and may require imaging on day 13 for aging experiments. We recommend imagining days 1, 5, 9, and 13.

Quantification of Muscle Actin (see Note 7)

1. Quantify muscle actin organization by manual scoring (we recommend blinding the investigator to biological conditions to avoid bias), as phenotypic variability limits reliable automated analysis.
2. For each animal, count the number of muscle cells displaying defective actin filaments and express this value as a percentage of total muscle cells.
3. Muscle cells should be scored as having defective actin if filaments display any observable differences in phenotype, such as broken, discontinuous, distorted, or wavy filaments (see Fig. 1A).

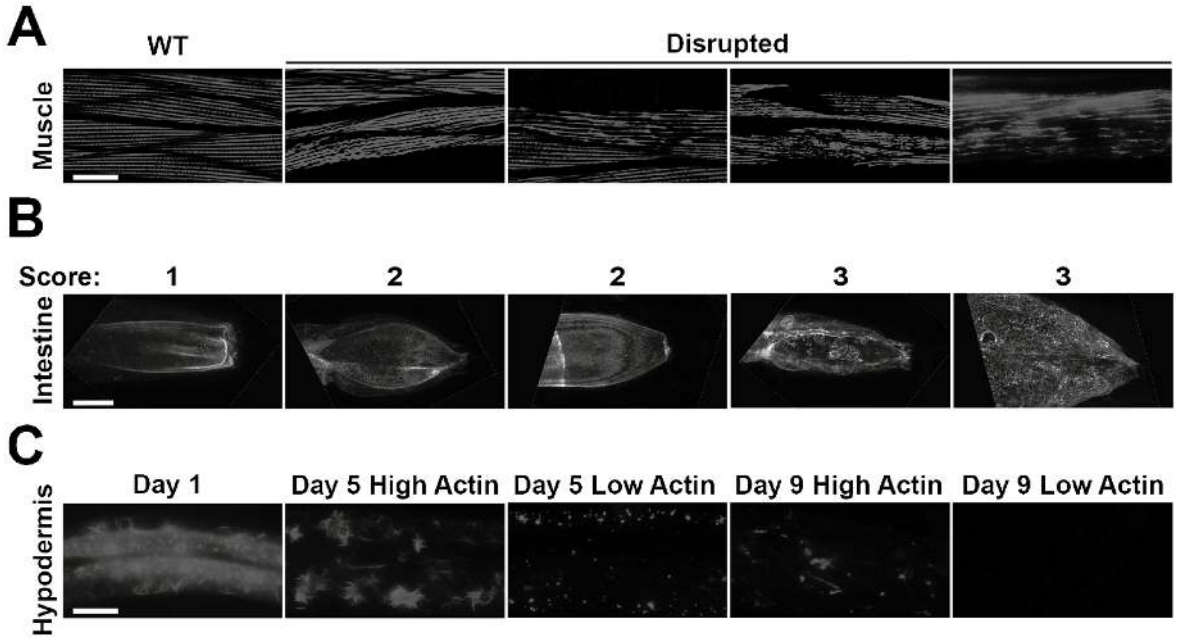


Fig. 1 Scoring actin disruption in adult *C. elegans*. **A,** Representative fluorescent images showing wild-type or disrupted actin structure in adult animals expressing LifeAct::mRuby under a muscle-specific promoter (myo-3p). **B,** Fluorescent images of intestinal actin structure representative of actin disruption scores 1–3 in adult animals expressing LifeAct::mRuby under an intestine-specific promoter (*gly-19p*). **C,** Representative fluorescent images showing wild-type or disrupted actin structure in day 1, day 5, and day 9 adult animals expressing LifeAct::mRuby under a hypodermis-specific promoter (*col-19p*). Scale bar is 10 μm

4. Calculate the percentage of muscle cells with disrupted actin per animal. Plot each biological replicate as a single data point. We recommend performing all manual scoring blind to the experimental condition to minimize observer bias.

Quantification of Intestinal Actin (see Note 8)

1. Similar to muscle, assess intestinal actin organization by manual scoring (we recommend blinding the investigator to biological conditions to avoid bias).
2. For intestinal actin, we use a three-point scale: (1) indicates normal wildtype-like morphology, (2) indicates partially disturbed structures, and (3) indicates severely disrupted structures (see Fig. 1B).
3. In young animals, wild-type intestinal actin typically appears as a thin, continuous structure lining the intestinal lumen.
4. With age or stress, actin organization becomes less distinct, with increased diffuse signal and the appearance of aggregate-like structures.

Quantification of Hypodermal Actin (see Note 9)

1. Hypodermal actin is quantified using an ImageJ Macro (provided below) by calculating

the percentage of the image area occupied by fluorescent actin signal (see Fig. 1C for examples of young animals with high levels of hypodermal actin and old animals with low levels of hypodermal actin).

2. Convert z-stacks into maximum-intensity projections prior to analysis.

3. The macro loops over all images currently open in ImageJ. Briefly, the macro converts the images to 8-bit and sets an auto-threshold ("Triangle white"). Then the macro records intensity stats over the entire image (so mean/min come from the grayscale image), while area_fraction is computed from the current threshold (fraction of pixels above threshold). Because the macro reports the fraction of non-fluorescent background area, subtract the measured area_fraction from 100 to obtain the final percentage of the image area occupied by fluorescent hypodermal actin structure.

IJM Code for Hypodermal Actin Quantification (Code Can Be Copied and Pasted Directly into an ImageJ Macro)

```
// Clear Results table
run("Clear Results");
setBatchMode(true);
// Set measurements
run("Set Measurements…","area mean min area_fraction
redirect=None decimal=3");
count = nImages;
for (i = 0; i < count; i++) {

    selectImage(i + 1);
    processImage();

}
setBatchMode(false);
updateResults();
// Ask where to save Results
outputFile = File.openDialog("Save results file");
saveAs("Results", outputFile);
function processImage() {

   prevNumResults = nResults;
   filename = getTitle();
   // Switch to channel with fluorescent images, which is channel
2 in this example
   Stack.setChannel(2);
   // Threshold + measure
   run("8-bit");
```

```
    setAutoThreshold("Triangle white");
    run("Threshold…");
    run("Measure");
    // Add filename to Results rows just created
    for (row = prevNumResults; row < nResults; row++) {
     setResult("Filename", row, filename);
    }

}
```

3.6 Imaging of the Mitochondria in *C. elegans*

Here, we provide a brief overview of optimized approaches for imaging mitochondria using genetic reporter strains (see **Sect. 7**; Tables 1 and 2***)*** and quantifying the morphology in the body wall muscle, hypodermis, and intestine of *C. elegans* (see Fig. 2). However, a much more detailed description of imaging strategies for tissue-specific mitochondria, analysis of mitochondrial morphology changes with age, and their correlation with physiological measurements can be found elsewhere [20]. Of note, immobilizing with sodium azide, which inhibits Complex IV in the electron transport chain [117], can cause mitochondrial fragmentation. Instead, we recommend using HistoBond adhesive microscope slides, which have a permanent positive charge through a chemical coating (*see* **Note 5**).

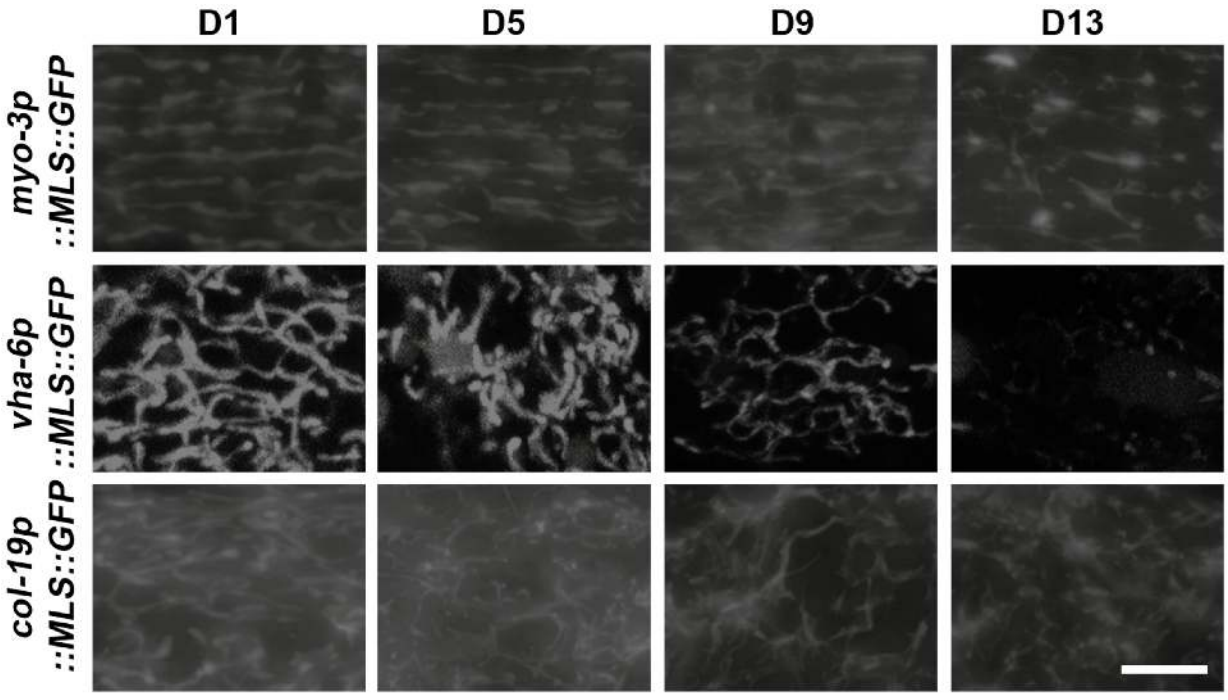


Fig. 2 Imagine mitochondrial morphology with age. Imaging of cell-type specific mitochondria in the muscle (*myo-3p*), intestine (*vha-6p*), and hypodermis (*col-19p*) using a single-copy integration of MLS::GFP using MosSCI during aging. All animals were grown on ev from the L1 stage and imaged at days 1, 5, 9, and 13 of adulthood. Arrowheads indicate examples of spherical auto-fluorescent structures in the intestine. Scale bar is 10 μm. This figure is adapted from Kim et al. Fig. 3 [77]

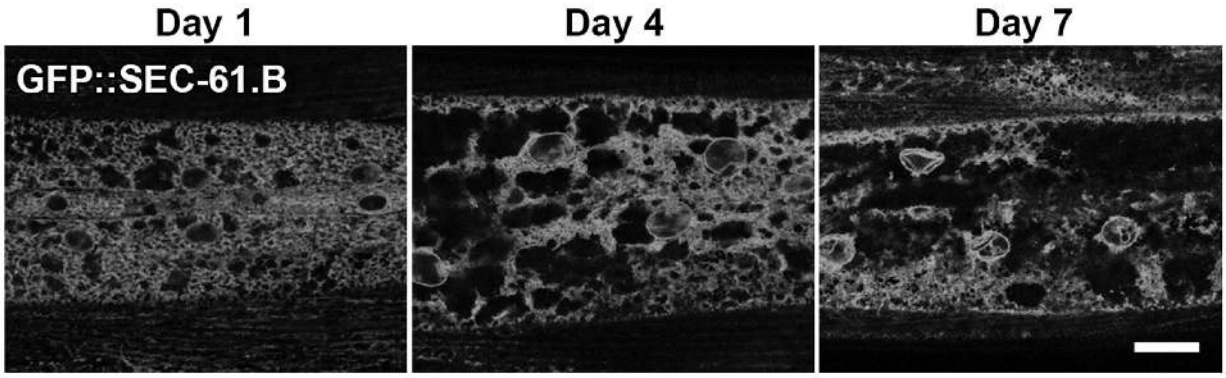


Fig. 3 Imaging ER morphology with age. Representative images of animals with GFP-tagged SEC-61.B at the endogenous gene locus and imaged at days 1, 4, and 7. Images were collected in the hypodermis. Scale bar is 10 μm

Table 2 Select genetically encoded mitochondria reporter strains in *C. elegans*

Imaging of the Hypodermal and Muscle Mitochondria Using a Compound Microscope

1. Use a standard wide-field fluorescence microscope to visualize mitochondria in the body wall muscle and hypodermal tissues. We have used animals expressing *myo-3p::MLS::GFP* and *col-19p::MLS::GFP*, respectively.
2. Optimize imaging settings for each microscope and experimental setup. In our studies, we have used a Leica Thunder microscope. As a starting point, we recommend acquiring images at a resolution of 2048 (x) × 2048 (y) pixels. When collecting *z*-stacks, use a step size of 0.50 μm for mitochondria in the hypodermis and use the "system optimized" setting for mitochondria in the muscle. Use a 63×/1.4 NA oil-immersion Plan Apochromat objective with a GFP filter set. Set the exposure time to 50 ms for both muscle and hypodermal imaging. Provide excitation using a 475 nm LED set to 50% power, setting neutral density filters to 30% for muscle and 17% for hypodermal mitochondria. Acquire *z*-stacks starting at the first visible GFP signal and ending where the signal is no longer detectable. To minimize background noise in maximum-intensity projections, reduce the number of z-planes included in the projection where possible.
3. Acquire images from a minimum of five animals per condition per biological replicate, imaging animals in the same approximate anatomical region across experiments.
4. Image mitochondrial morphology across adulthood to capture any age-associated changes. For standard aging experiments, image animals at days 1, 5, 9, and 13 of adulthood, being sure to maintain consistent anatomical regions across timepoints.

Imaging of Intestinal Mitochondria Using a Confocal Microscope

1. Use a confocal microscope to visualize intestinal mitochondria in animals, we have used expressing *vha-6p::MLS::GFP*.
Optimize imaging settings for each microscope and experimental setup. As a starting

2. Optimize imaging settings for each microscope and experimental setup. As a starting point, use a confocal system equipped with a 63×/1.4 NA oil-immersion Plan Apochromat objective, a white light laser, and HyD S detectors. In our studies, we have used a Leica Stellaris confocal microscope. Set the white light laser to 85% maximum power and excite GFP using a 485 nm laser line at an intensity of 3%. Collect emission between 490–590 nm and set the detector gain to 25 in the analog setting. Acquire images using unidirectional scanning over a 1024 (x) × 1024 (y) pixel field of view (82.01 × 82.01 µm). Acquire z-stacks with five *z*-slices at a step size of 0.495 µm, adjusting the total z-range based on tissue thickness. Set the scan speed to 1000 Hz, zoom to 2.25×, line averaging to 2, and the pinhole to 1 AU (95.5 µm). Acquire *z*-stacks beginning at the first detectable GFP signal and ending where the signal is no longer visible. Similar to the compound microscope, to minimize background noise in maximum-intensity projections, reduce the number of z-planes included in the projection, where possible.
3. Acquire images from a minimum of five animals per condition per biological replicate. Image animals at the same intestinal region across experiments to ensure consistency in mitochondrial morphology measurements.
4. For aging experiments, image intestinal mitochondria at days 1, 5, 9, and 13 of adulthood. Use consistent acquisition settings and anatomical regions across time points to allow comparison of any age-associated changes.

Quantification of Mitochondrial Morphology Using mitoMAPR

1. Install FIJI [123] and load the MitoMAPR macro by copying the ImageJ macro code provided in Source Code 1 of the original publication [78] and provided below into a new FIJI macro file and saving it for repeated use.

```
var SampleIdARRAY = newArray();
var ObjectsARRAY = newArray();
var JunctionsArray = newArray();
var JunctionsPerNetworkArray = newArray();
var NetworksARRAY = newArray();
var ObjectLengthARRAY = newArray()
var mitoAreaARRAY = newArray();
var MitoPartAreaArray = newArray();
var CoverageArray = newArray();
var CellAreaArray = newArray();
open();
imgName=getTitle();
run("8-bit");

   waitForUser ("make a selection");
   run("Clear Outside");
   selectWindow(imgName);
```

```
     rename("ROI.tif");
     electWindow("ROI.tif");
     run("Save");

// MitoMAPR-----------------------------------------------------
---------------

   macro "MitoMAPR-1.0" {

if (nImages == 0) {

      showMessage("No images are open.");
      exit();
    }

run("Duplicate…", "title=TheOne");

     selectWindow("TheOne");
     run("Grays");
     getDimensions(width, height, channels, slices, frames);
     getPixelSize(unit, pixelWidth, pixelHeight);

run("Enhance Local Contrast (CLAHE)", "blocksize=127
histogram=256 maximum=3 mask=*None*");

     run("Unsharp Mask…", "radius=2 mask=0.60");
     run("Median…", "radius=2");
     run("Duplicate…", "title=Skeletor");
     selectWindow("Skeletor");
     run("32-bit");
     run("Make Binary");
     run("8-bit");
     getStatistics(area, Mean, min, max);
     mitoArea = pow(pixelWidth, 2.0) * parseFloat(width) *
parseFloat(height) * (Mean / parseFloat(max));
     run("Restore Selection");
     getStatistics(area);
     TotalCellArea=area;
    MitoCoverage = ((mitoArea/TotalCellArea)*100);
     run("Skeletonize");
     run("Red");
     selectWindow("TheOne");
```

```
    run("Add Image…", "image=Skeletor x=0 y=0 opacity=100
zero");
    size =
toString(round(0.25*parseFloat(width)*parseFloat(pixelWidth)));
    run("Scale Bar…", "width=" + size + " height=4 font=14
color=White background=Black location=[Lower Right] bold
overlay");

{

    selectWindow("Skeletor");
    run("Analyze Skeleton (2D/3D)","prune=none show display");
    selectWindow("Tagged skeleton");
   saveAs("Tagged skeletor.tif");
   selectWindow("Skeletor-labeled-skeletons");
   saveAs("labeled skeletor.tif");
     close("Tagged skeleton");
     close("Skeletor");
     close("Skeletor-labeled-skeletons");

selectWindow("TheOne");
saveAs("TheOne.tif");
close("TheOne");
selectWindow("Results"); rows = nResults;
ObjectCounts = newArray(rows);
for (i=0; i<rows; i++) {
ObjectCounts[i] = getResult("# Branches", i);
}
JunctionCounts = newArray(rows);
for (i=0; i<rows; i++){
JunctionCounts[i] = getResult("# Junctions", i);
}
selectWindow("Results"); run("Close");
IJ.renameResults("Branch information","Results");

     selectWindow("Results"); rows = nResults;
     ObjectLengths = newArray(rows);

for (i=0; i<rows; i++) {

       ObjectLengths[i] = getResult("Branch length", i);
     }
     run("Close");
```

```
      Objects = (parseFloat(CountForMe(ObjectCounts)))*10;
      Networks =
(parseFloat(countNetworksForMe(ObjectCounts)))*10;
      JunctionPoints = (AddForMe(JunctionCounts))*10;
      JunctionsPerNetwork = round(JunctionPoints/Networks);
   MitoPartArea = mitoArea/Objects;
    ObjectLength = BeMeanForMe(ObjectLengths);

Maths = newArray( "Objects(OC)",

         "ObjectLength(OL)",
         "Networks(N)",
         "JunctionPoints(JP)",
       "JunctionsPerNetwork(JP/N)",
       "MitochondrialFootprint(MF",
       "ObjectArea(OA)",
       "MitoCoverage(MC)",
       "TotalCellArea(TCA)");

Numbers = newArray(Objects,

     ObjectLength,
         Networks,
         JunctionPoints,
         JunctionsPerNetwork,
         mitoArea,
         MitoPartArea,
         MitoCoverage,
         TotalCellArea);

Units = newArray("Counts",

      unit,
      "Counts",
      "Counts",
       "Counts",
       unit+" squared",
       unit+" squared",
       "Percent",
       unit+" squared");

Array.show("Data", Maths, Numbers, Units);
selectWindow("Data");
```

```
saveAs("Data.txt");
selectWindow("Data.csv");

    if (isOpen("Data.csv")) {
      selectWindow("Data.csv");
      run("Close");
    run("Close All");

}
}
//Do Maths for Me
function CountForMe(data) {

     entries = data.length;
     total = 0.0;

for (i=0; i<entries; i++)

       {
        total = total + 1;
       }
       return(total);

}
function BeMeanForMe(data) {

     entries = data.length;
     total = 0.0;

for (i=0; i<entries; i++) {

       total = total + data[i];
     }
     ave = total/parseFloat(entries);
     return(ave);

}
function AddForMe(data) {

     entries = data.length;
     total = 0.0;

for (i=0; i<entries; i++) {

       total = total + data[i];
```

```
    }
    sum = total;
    return(sum);

}
function countNetworksForMe(data) {

    entries = data.length;
    total = 0.0;

for (i=0; i<entries; i++) {

      if (data[i] > 1) {
        total = total + 1;
      }
      else {
      }
    }
    return(total);

}
function CountNetworkBranchesForMe(data) {

    entries = data.length;
    total = 0.0;

for (i=0; i<entries; i++) {
if (data[i] > 1) {
total = total + data[i];
}
else {
}
}

    return(total);

}
```

2. Open 3D mitochondrial image stacks in FIJI and generate maximum-intensity projections using only in-focus z-planes. Save projected images as TIFF files prior to analysis.
3. Define regions of interest (ROIs) by cropping projected images to include only the tissue area to be analyzed. Save cropped ROIs as TIFF files.

Run the MitoMAPR macro on cropped images to extract quantitative descriptors of mitochondrial morphology. When prompted, select the appropriate image directory and confirm the ROI for analysis. Export all output values from the resulting "Data" window as an Excel file for downstream analysis.

Batch Processing of Mitochondrial Morphology Using FIJI Macros

1. For large datasets, perform batch processing using FIJI macros to standardize projection, cropping, and analysis across images.
2. Convert all images to TIFF format prior to batch analysis using the macro, including the code below.

```
dir = getDirectory("Select A folder"); ids=newArray(nImages);
for (i=0;i<nImages;i++) { selectImage(i+1); title = getTitle;
print(title); Stack.setChannel(1); ids[i]=getImageID;
saveAs("tiff", dir+title.replace("/", "-"));}.
```

If images contain *z*-stacks, generate maximum-intensity projections in batch using the macro with the code below.

```
run("Z Project…","projection=[Max Intensity]").
```

3. Crop ROIs with identical dimensions across all images using cropping macros with the code described below. Define ROI dimensions once and manually position the selection for each image to capture the appropriate tissue region.

```
Make Rectangle (100, 100, 200, 200); run (Specify…, width =
100, height = 100, x = 100, y =100 scaled).
```

4. Analyze all cropped images using the batch version of the MitoMAPR macro created by copying the code below, adapted from the source code 1 of the original publication [78] under **B.). IJM code for MitoMAPR-1.0_Batch**. Save the compiled output values from the resulting "Data" window as an Excel file for statistical analysis.

```
var SampleIdARRAY = newArray();
var ObjectsARRAY = newArray();
var JunctionsArray = newArray();
var JunctionsPerNetworkArray = newArray();
var NetworksARRAY = newArray();
var ObjectLengthARRAY = newArray()
var mitoAreaARRAY = newArray();
var ObjectAreaArray = newArray();
var CoverageArray = newArray();
```

```
var CellAreaArray = newArray();
macro "MitoMAPR Batch" {

     dir = getDirectory("Choose a Directory ");
     processFiles(dir);
    SampleId = SampleIdARRAY;
    Objects = ObjectsARRAY;
    Networks = NetworksARRAY;
    JunctionsPerNetwork = JunctionsPerNetworkArray;
    JunctionPoints = JunctionsArray;
    ObjectLength = ObjectLengthARRAY;
    MitochondrialFootprint = mitoAreaARRAY;
    ObjectArea= ObjectAreaArray;
    MitoCoverage= CoverageArray;
    TotalCellArea = CellAreaArray;
     Array.show("Data",
       SampleId,
       Objects,
       Networks,
       JunctionsPerNetwork,
       JunctionPoints,
       ObjectLength,
       MitochondrialFootprint,
     ObjectArea,
     MitoCoverage,
     TotalCellArea);

}
function processFiles(dir) {

   list = getFileList(dir);
   for (i=0; i<list.length; i++) {
    if (endsWith(list[i], "/"))
     processFiles(""+dir+list[i]);
    else {
    path = dir+list[i];
    processFile(path);
    }
    }

}
function processFile(path) {
```

```
open (path);
run("Subtract Background…", "rolling=50");
run("Enhance Local Contrast (CLAHE)", "blocksize=127
histogram=256 maximum=3 mask=*None*");
run("Unsharp Mask…", "radius=2 mask=0.60");
run("Median…", "radius=2");

    run("Duplicate…", "title=TheOne");
    selectWindow("TheOne");
    run("Grays");
    getDimensions(width, height, channels, slices, frames);
    getPixelSize(unit, pixelWidth, pixelHeight);
    selectWindow("TheOne");
    run("Duplicate…", "title=Skeletor");
    selectWindow("Skeletor");
    run("32-bit");
    run("Make Binary");
    run("8-bit");
    getStatistics(area, Mean, min, max);
    run("Restore Selection");
    getStatistics(area);
    TotalCellArea=area;
   MitoCoverage = ((mitoArea/TotalCellArea)*100);
    run("Skeletonize");
    run("Red");
    selectWindow("TheOne");
    run("Add Image…", "image=Skeletor x=0 y=0 opacity=100 zero");
    size =
toString(round(0.25*parseFloat(width)*parseFloat(pixelWidth)));
    run("Scale Bar…", "width=" + size + " height=4 font=14
color=White background=Black location=[Lower Right] bold
overlay");
   {
     selectWindow("Skeletor");
     run("Analyze Skeleton (2D/3D)","prune=none show display");
     close("Tagged skeleton");
     close("Skeletor");
     close("Skeletor-labeled-skeletons");
     selectWindow("Results"); rows = nResults;
     ObjectCounts = newArray(rows);
       for (i=0; i<rows; i++) {
```

```
        ObjectCounts[i] = getResult("# Branches", i);
      }
      JunctionCounts = newArray(rows);
      for (i=0; i<rows; i++){
        JunctionCounts[i] = getResult("# Junctions", i);
      }
      selectWindow("Results"); run("Close");
      IJ.renameResults("Branch information","Results");

selectWindow("Results"); rows = nResults;
ObjectLengths = newArray(rows);
for (i=0; i<rows; i++) {
ObjectLengths[i] = getResult("Branch length",i);
}
run("Close");
Objects = (parseFloat(CountForMe(ObjectCounts)))*10;
Networks = (parseFloat(countNetworksForMe(ObjectCounts)))*10;
JunctionPoints = (AddForMe(JunctionCounts))*10;
JunctionsPerNetwork = round(JunctionPoints/Networks);
ObjectArea = mitoArea/Objects;
ObjectLength = BeMeanForMe(ObjectLengths);
SampleIdARRAY = Array.concat(SampleIdARRAY, path);
ObjectsARRAY = Array.concat(ObjectsARRAY, Objects);
NetworksARRAY = Array.concat(NetworksARRAY, Networks);
JunctionsArray = Array.concat(JunctionsArray, JunctionPoints);
ObjectLengthARRAY = Array.concat(ObjectLengthARRAY,
ObjectLength);
JunctionsPerNetworkArray = Array.concat(JunctionsPerNetworkArray,
JunctionsPerNetwork);
mitoAreaARRAY = Array.concat(mitoAreaARRAY, mitoArea);
ObjectAreaArray = Array.concat(ObjectAreaArray, ObjectArea);
CoverageArray = Array.concat(CoverageArray, MitoCoverage);
CellAreaArray = Array.concat(CellAreaArray, TotalCellArea);
titles = getList("image.titles");
for (i=0; i<titles.length; i++) {
selectWindow(titles[i]);
close();
}
}
{
SampleIdARRAY = Array.concat(SampleIdARRAY, path);
ObjectsARRAY = Array.concat(ObjectsARRAY, "");
NetworksARRAY = Array.concat(NetworksARRAY, "");
```

```
JunctionsArray = Array.concat(JunctionsArray, "");
JunctionsPerNetworkArray = Array.concat(JunctionsPerNetworkArray,
"");
ObjectLengthARRAY = Array.concat(ObjectLengthARRAY, "");
mitoAreaARRAY = Array.concat(mitoAreaARRAY, "");
ObjectAreaArray = Array.concat(ObjectAreaArray, "");
CoverageArray = Array.concat(CoverageArray, "");
CellAreaArray = Array.concat(CellAreaArray, "");
titles = getList("image.titles");
for (i=0; i<titles.length; i++) {

     selectWindow(titles[i]);
     close();
    }
   }

}
//Do Maths for Me
function CountForMe(data) {

   entries = data.length;
   total = 0.0;
   for (i=0; i<entries; i++)
    {
     total = total + 1;
    }
    return(total);

}
function BeMeanForMe(data) {

   entries = data.length;
   total = 0.0;
   for (i=0; i<entries; i++) {
    total = total + data[i];
   }
   ave = total/parseFloat(entries);
   return(ave);

}
function AddForMe(data) {

   entries = data.length;
   total = 0.0;
```

```
    for (i=0; i<entries; i++) {
     total = total + data[i];
    }
    sum = total;
    return(sum);

}
function countNetworksForMe(data) {

    entries = data.length;
    total = 0.0;
    for (i=0; i<entries; i++) {
     if (data[i] > 1) {
      total = total + 1;
     }
     else {
     }
    }
    return(total);

}
function CountNetworkBranchesForMe(data) {

    entries = data.length;
    total = 0.0;
    for (i=0; i<entries; i++) {
     if (data[i] > 1) {
      total = total + data[i];
     }
     else {
     }
    }
    return(total);

}
```

3.7 Imaging of the ER in *C. elegans*

Multiple properties of the ER impart challenges to visualizing its structures accurately. The ER fills most cells and exhibits high structural complexity, but also with very small dimensions (e.g., ~30–60 nm diameters with gaps only ~200 nm between stacked sheets) [124]. The need for high resolution across cells makes electron microscopy (EM), especially volumetric EM techniques, the gold standard for ultrastructure imaging. EM analysis of the ER has significant limitations, however, including the need to fix and stain samples through

processes that can perturb native structures, a lack of live-imaging amenability, low throughput, and challenging image segmentation and analysis in post-acquisition steps. To complement EM, advances in light microscopy and genetic techniques for labeling ER in vivo (see **Sect. 7;** Tables 1 and 3) have led to increasingly effective methods for ER visualization in live cells and animals. Here, we detail imaging the ER with age (see Fig. 3) and quantifying the ER in *C. elegans*.

Table 3 Select genetically encoded ER reporter strains in *C. elegans*

Imaging of the ER Using a Confocal Microscope

1. Mount animals for microscopy. While standard anesthetization methods are also likely appropriate for ER imaging, we employ established microbead-based mounting to minimize potential impacts of drug treatments [132]. Prepare a 10% agarose solution (1.0 g agarose in 10 mL M9 solution) and microwave it briefly (<16 s) to dissolve. Stir rapidly with a spatula and then immediately scoop a small volume onto a glass slide. Press another slide on top to create a small pad between the two. Remove the top slide and trim the edges of the pad to make a flat, square pad. Add 1.5 uL Polybead Microspheres to the pad, and pick 20–30 worms into the solution, taking care to transfer minimal bacteria. Use an eyelash pick to separate the worms and overlay a glass coverslip. Use the back of the pick to lightly press the coverslip down, allowing the liquid to distribute under and adhere to it (*see* **Note 10**).
2. Use a confocal microscope to visualize ER in animals. We have used *gfp::sec-61.b; ret-1::mKate2* (*see* **Note 11**). While acquisition settings for each microscope and experimental setup should be optimized, as a starting point, we used a Nikon Eclipse Ti2 microscope equipped with a Yokogawa CSU-W1 spinning disc, 488 and 561 nm lasers, a 10×/0.45 and a 100×/1.45 NA oil-immersion Plan Apo λ objective, and a Prime 95B sCMOS camera.
3. Ensure consistency between animals and fields of view. Because ER morphology can vary regionally within individual cells or tissues, acquire images between animals using the same location in each tissue and consistent focal planes within cells. For example, orienting the focal plane and field of view to contain the nucleus ensures capture of both perinuclear and peripheral ER structures consistently across all images. To minimize potential photobleaching artifacts, we use the DIC channel to locate our intended field of view. While all tissues are amenable to ER imaging, the thin hypodermis tends to provide the clearest views of ER morphological features. We image the anterior portion of the hypodermal syncytium (*hyp7*) between the pharyngeal bulb and anterior spermatheca (*see* **Note 12**).

Quantification of the ER Using NIS Elements

Before beginning an analysis, we recommend deconvolving images to improve ER resolution in an independent workflow. Batch pre-processing can be performed in NIS Elements

General Analysis 3 (GA3) and should be done prior to image annotation. This workflow (Fig. 4) is designed for single-plane images of the ER.

1. Open a sample image and then create a new GA3 recipe. Attach a "deconvolution" function to the color channel to be deconvolved. Double click each function to verify the correct excitation/emission wavelengths of the channel to be deconvolved, typically through either the automatic deconvolution or Richardson-Lucy algorithms. Optionally, add a denoising step, such as the "denoise.ai." function, after deconvolution.
2. After your images have been pre-processed, ensure they are blinded and use the binary editor to draw a binary region of interest around the cell/tissue of interest in each image. The binary tool is recommended over the standard region of interest (ROI) tool to allow analysis of discontinuous regions within the same cell or tissue. This enables exclusion of potentially confounding signals from other cell types (e.g., neurites passing through the hypodermis) in ubiquitous marker strains, such as endogenous GFP::SEC-61.B. For analyses focusing on peripheral ER structures, we also exclude the nuclear envelope from the ROI.
3. Once all image files have been annotated, open a sample image. Create a new GA3 recipe, which will be the basis for segmenting and measuring the ER (*see* **Note 13**). The following workflow provides a general starting point for analyzing a single reporter such as GFP::SEC-61.B, but specific settings are highly dependent on experimental conditions.
4. Add the "Binary" function and select the previously drawn cell/tissue binary ROI ("Binary Edit") to import it to the workflow (*see* **Note 14**).
5. Add the "Rolling Ball" background subtraction function to the color channel. Use the dropdown menu to set a radius larger than the biggest object in the image.
6. Add the "Threshold" function to the color channel, which will create a new binary mask ("all_ER"). Adjust the minimum intensity threshold value in the dropdown menu to optimize what is included in the mask. Optionally, use the smooth and clean operations in the menu to improve segmentation fidelity.
7. Add an "And" function. Connect "Binary Edit" to "all_ER" to create a new binary mask ("hyp_ER") (*see* **Note 15**).
8. Add the "Local Contrast" function. Set the degree and radius in the dropdown menu. Add a "Threshold" function to the contrast color channel to create a new binary ("contrast_ER"). Unlike before, optimize the *maximum* value to exclude the ER from the binary.
9. Add the "Valleys" function. Set the matrix and count in the dropdown menu. Add a "Threshold" function to the valleys color channel to create a new binary ("valley_ER").
10. Add the "Regional Minima" function. Set the matrix and count in the dropdown menu.

Add a "Threshold" function to the regional minima color channel to create a new binary ("min_ER").

11. Add the "Or" function and combine "contrast_ER," "valley_ER," and "min_ER" to create a new binary ("not_ER").
12. Add the "Subtract" function. Connect node A to "hyp_ER" and node B to "not_ER" to create a new binary ("final_ER").
13. Add a "Filter Objects" function (area < 5 pixels) to the "final_ER" binary to remove any remaining noise from the segmentation mask.
14. Add the "Filename" function to any color channel. Add the "Accumulate Records" function to the resultant table.
15. At this point, the measurements are defined. We typically quantify the fluorescence intensity, area, and perimeter in order to compare the levels of the marker, the footprint of the ER, and the perimeter:area ratio.
16. Add the "Mean Intensity" function from the field measurements section. Connect it to the background-subtracted color channel and the "final_ER" binary to measure the average pixel intensity/integrated density of the fluorescence signal in the segmented region ("ERIntensity").
17. Add two "Total Object Area" functions. Connect one to the "Binary Edit" binary to measure the area of the cell/tissue of interest ("hypArea"). Connect the other to the "final_ER" binary to measure the area of the segmented ER ("ERArea").
18. Add the "Perimeter" function from the object size measurements function. Connect it to the "finalER" binary and add the "Aggregate Rows" function to the resultant table. Use the dropdown menu to select "Sum" ("ERPerimeter") and eliminate the other rows' statistics.
19. Add the "Append Columns" function. Connect the final tables from the file name, area, and perimeter pathways. Optionally, add the "Modify Columns" function downstream to remove unneeded columns or rename other columns for clarity.
20. Add two "Calculated Column" functions. Connect the first to the table with appended columns and the second downstream. Use the first dropdown menu to create a new variable ("ERFtpt" = "ERArea"/"hypArea"). Use the second dropdown menu to create a new variable ("ERPAR" = "ERPerim"/"ERArea").
21. Add a "Results" function at the end to view the results later.
22. Add a "Table to CSV" function. Connect it to the final records table. Use the dropdown menu to set the export path and an appropriate file name. Make sure the box is unchecked to allow data from multiple files to be captured in the same CSV.

23. Save and exit the GA3 recipe.
24. Open the Batch GA3 window, select the GA3 analysis recipe, and select the annotated images or folder. Make sure the “Keep Original” box at the top of the menu is checked to prevent overwriting all images.
25. Run the analysis.

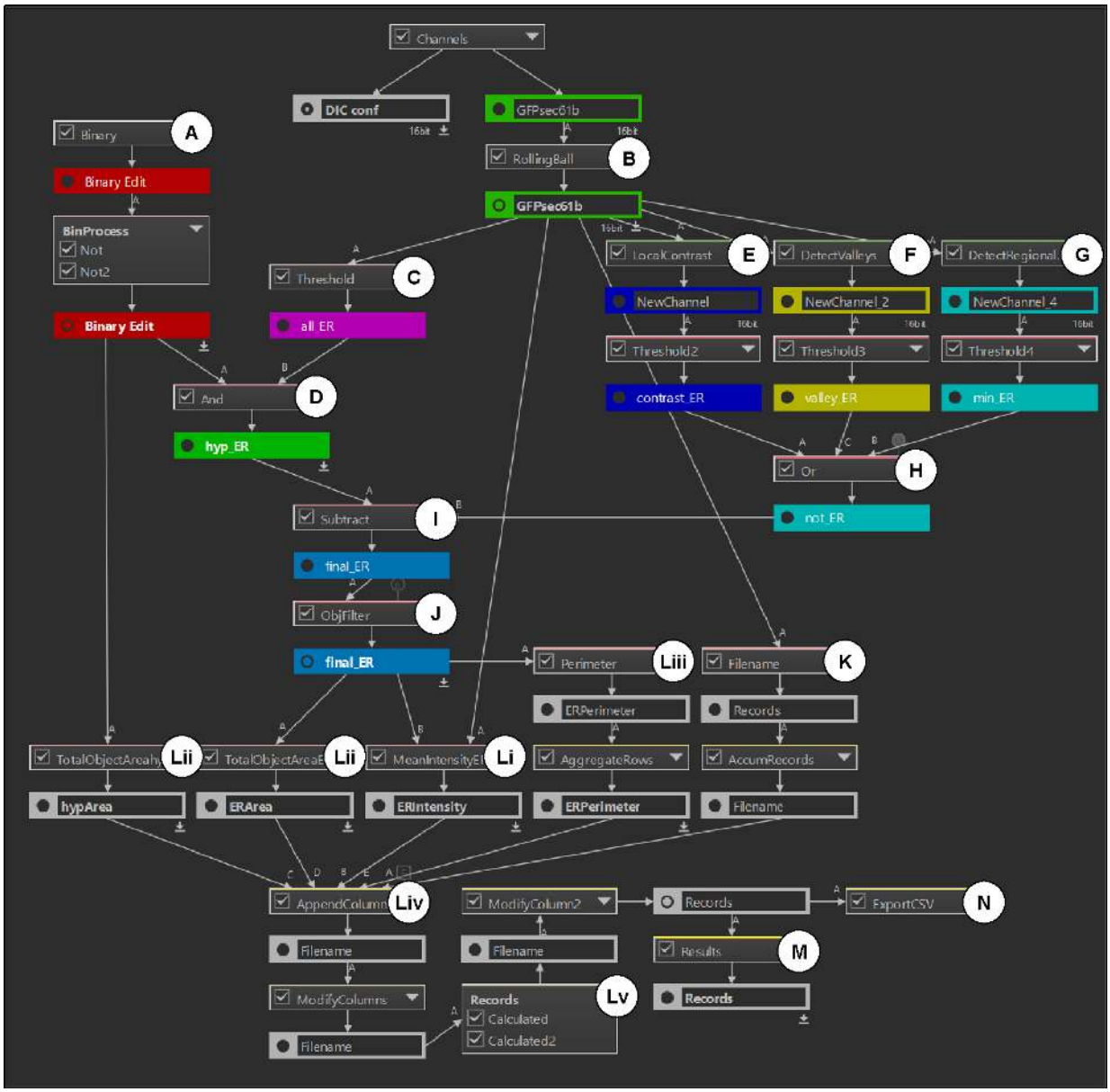


Fig. 4 ER segmentation and quantification recipe. Screenshot of sample GA3 recipe for ER segmentation and quantification. Circled letters correspond to steps in the protocol

4 Notes

1. For experiment plates, inoculate lysogeny broth (LB) or equivalent media with HT115 and add 1:1000 carbenicillin, which preserves the plasmid of the RNAi target, and 1:500 tetracycline, which selects from the T7 polymerase, both necessary for RNAi experiments. Grow culture at 37 °C in a shaking incubator, as HT115 has a slower growth rate than OP50. Seed as in Sect. 3.1.
2. We have found that different brands of agar have different stiffness and consistency, which can significantly affect lifespan [133]. Therefore, we recommend standardizing a single agar source. We recommend using BD Difco Agar (supplier number 21040).
3. Bleaching allows for robust synchronization of animals using a well-established protocol. However, overexposure to bleach can damage embryos, reduce hatch rate, and create selection pressure for bleach-resistant strains, potentially altering population genetics. Alternative methods to bleaching to avoid these caveats are egg-lay [134] or a NemaSync [135].
4. Under-bleaching may carry over contamination or leave residual bleach in remaining adult carcasses. Overbleaching may damage eggs and result in a low hatch rate (*see* **Note 3**).
5. Sodium azide can induce mitochondrial fragmentation [77]. Although short-term imaging with low concentrations of sodium azide does not induce fragmentation, we highly recommend avoiding the use of sodium azide entirely for mitochondrial imaging to avoid artifacts. For long-term imaging, it is also recommended to use an agarose pad [77].
6. The volume of sodium azide applied to each slide should be scaled according to animal size. Early adult (day 1) animals are typically immobilized with approximately 10–13 µL, whereas older animals generally require 16–18 µL for effective paralysis. Volumes may be further adjusted for smaller or larger individuals to achieve complete immobilization without excessive movement from excess liquid or tissue rupture due to insufficient volume. Slides should be sealed with Valap wax or nail polish to minimize evaporation and prevent mechanical damage to worms over time.
7. Quantification of body wall muscle actin relies on blinded manual scoring of filament organization due to the high variability in sarcomere phenotypes and the difficulty of reliably segmenting individual muscle cells computationally. This approach is inherently subjective and dependent on observer experience, which may introduce scorer-dependent bias despite blinding. In addition, this method does not capture graded changes in filament intensity, spacing, or sarcomere periodicity and may overestimate mild or early-stage defects. Variations in animal orientation, muscle quadrant alignment, and focal plane selection can further influence apparent filament continuity and should be carefully controlled during image acquisition.
8. Intestinal actin is assessed using a scoring system based on overall structural changes, but it lacks the resolution to distinguish specific types of changes (e.g., diffused or aggregated signal).

9. Quantification of hypodermal actin using area-fraction measurements provides an objective, high-throughput metric but does not directly measure filament organization or structure. This approach is sensitive to threshold selection and background fluorescence and may be influenced by imaging artifacts arising from imaging parameters. Consequently, this method is best suited for detecting changes in overall actin coverage rather than detailed alterations in filament architecture.
10. For aging experiments, we recommend imaging adults at days 1 and 7 of adulthood, though progressive changes are visible.
11. Given the high levels of ER-phagy during aging, ER markers can accumulate in lysosomes. As certain dsRed-derived fluorescent proteins (e.g., mCherry) maintain high signal in lysosomes due to a low pKa and relative protease resistance, the lysosomal fraction of these markers can become the dominant signal in cells and confound measurement of ER morphology.
12. Due to the quantity and density of ER in intestinal cells, it is challenging to resolve specific structural features with fluorescence microscopy, which may be best suited for the measurement of overall ER abundance in this tissue. Electron microscopy is a useful, complementary approach.
13. If fluorescence intensities are similar across conditions, ER image processing and segmentation workflows will work well for a batch analysis using the same parameters across all images. However, if intensity varies significantly between groups, the GA3 Wizard tool enables a semi-automated analysis where the user can manually correct thresholds on blinded images to ensure accurate ER segmentation.
14. If using an older version of Elements, imported binary ROIs may not process correctly. For a workaround, use a double negate (i.e., connecting two “Not” functions) to create an identical binary, which then proceeds as expected.
15. At this stage, we find that polygons lying between ER tubules are often included in the mask. The following steps help to eliminate empty polygons from being erroneously included in the analysis.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Nigon VM, Félix MA (2017) History of research on *C. elegans* and other free-living nematodes as model organisms. WormBook 2017:1–84. https://doi.org/10.1895/wormbook.1.181.1
2. Goldstein B (2016) Sydney Brenner on the genetics of *Caenorhabditis elegans*. Genetics 204:1–2. https://doi.org/10.1534/genetics.116.194084

[PubMed][PubMedCentral]

3. Stiernagle T (2006) Maintenance of *C. elegans*. WormBook 2006:1–11. https://doi.org/10.1895/wormbook.1.101.1

4. Riddle D, Blumenthal T, Meyer B (1997) The biological model - *C. elegans* II. In: *C. elegans* II.,2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY)

5. Corsi AK, Wightman B, Chalfie M (2015) A transparent window into biology: a primer on *Caenorhabditis elegans*. Genetics 200:387–407. https://doi.org/10.1534/genetics.115.176099

6. Heppert JK, Dickinson DJ, Pani AM et al (2016) Comparative assessment of fluorescent proteins for in vivo imaging in an animal model system. MBoC 27:3385–3394. https://doi.org/10.1091/mbc.e16-01-0063
[PubMed][PubMedCentral]

7. Zacharias AL, Murray JI (2016) Combinatorial decoding of the invariant *C. elegans* embryonic lineage in space and time. Genesis 54:182–197. https://doi.org/10.1002/dvg.22928
[PubMed][PubMedCentral]

8. Packer JS, Zhu Q, Huynh C et al (2019) A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. Science 365:eaax1971. https://doi.org/10.1126/science.aax1971
[PubMed][PubMedCentral]

9. Murphy CT, Hu PJ (2013) Insulin/insulin-like growth factor signaling in *C. elegans*. WormBook 2013:1–43. https://doi.org/10.1895/wormbook.1.164.1

10. Lapierre LR, Hansen M (2012) Lessons from *C. elegans*: signaling pathways for longevity. Trends Endocrinol Metab 23:637–644. https://doi.org/10.1016/j.tem.2012.07.007
[PubMed][PubMedCentral]

11. Andrusiak MG, Jin Y (2016) Context specificity of stress-activated mitogen-activated protein (MAP) kinase signaling: the story as told by *Caenorhabditis elegans*. J Biol Chem 291:7796–7804. https://doi.org/10.1074/jbc.R115.711101
[PubMed][PubMedCentral]

12. Gumienny TL, Savage-Dunn C (2013) TGF-β signaling in *C. elegans*. WormBook 2013:1–34. https://doi.org/10.1895/wormbook.1.22.2

13. Sawa H, Korswagen HC (2013) Wnt signaling in *C. elegans*. WormBook 2013:1–30. https://doi.org/10.1895/wormbook.1.7.2

14. López-Otín C, Blasco MA, Partridge L et al (2023) Hallmarks of aging: an expanding universe. Cell 186:243–278. https://doi.org/10.1016/j.cell.2022.11.001
[PubMed]

15. Higuchi-Sanabria R, Paul JW, Durieux J et al (2018) Spatial regulation of the actin cytoskeleton by HSF-1 during aging. MBoC 29:2522–2527. https://doi.org/10.1091/mbc.E18-06-0362
[PubMed][PubMedCentral]

16. Sing CN, Garcia EJ, Lipkin TG et al (2022) Identification of a modulator of the actin cytoskeleton, mitochondria, nutrient metabolism and lifespan in yeast. Nat Commun 13:2706. https://doi.org/10.1038/s41467-022-30045-9
[PubMed][PubMedCentral]

17. Schmid ET, Schinaman JM, Liu-Abramowicz N et al (2024) Accumulation of F-actin drives brain aging and limits healthspan in Drosophila. Nat Commun 15:9238. https://doi.org/10.1038/s41467-024-53389-w

18. Averbukh M, Nelson HT, Wang T et al (2026) Form and function of actin impacts actin health and aging. iScience 29:117256. https://doi.org/10.1016/j.isci.2026.117256

19. Donahue EKF, Hepowit NL, Ruark EM et al (2026) ER remodelling is a feature of ageing and depends on ER-phagy. Nat Cell Biol https://doi.org/10.1038/s41556-025-01860-1
[PubMed][PubMedCentral]

20. Kim J, Averbukh M, Alcala A et al (2025) Imaging and quantifying mitochondrial morphology in *C. elegans* during aging. J Vis Exp 215:e67610. https://doi.org/10.3791/67610

21.

Revenu C, Athman R, Robine S, Louvard D (2004) The co-workers of actin filaments: from cell structures to signals. Nat Rev Mol Cell Biol 5:635–646. https://doi.org/10.1038/nrm1437
[PubMed]

22. Pollard TD, Cooper JA (2009) Actin, a central player in cell shape and movement. Science 326:1208–1212. https://doi.org/10.1126/science.1175862
[PubMed][PubMedCentral]

23. Rowland AA, Voeltz GK (2012) Endoplasmic reticulum–mitochondria contacts: function of the junction. Nat Rev Mol Cell Biol 13:607–615. https://doi.org/10.1038/nrm3440
[PubMed][PubMedCentral]

24. Schiavon CR, Zhang T, Zhao B et al (2020) Actin chromobody imaging reveals sub-organellar actin dynamics. Nat Methods 17:917–921. https://doi.org/10.1038/s41592-020-0926-5
[PubMed][PubMedCentral]

25. Davidson PM, Cadot B (2021) Actin on and around the Nucleus. Trends Cell Biol 31:211–223. https://doi.org/10.1016/j.tcb.2020.11.009
[PubMed]

26. Sen B, Xie Z, Thomas MD et al (2024) Nuclear actin structure regulates chromatin accessibility. Nat Commun 15:4095. https://doi.org/10.1038/s41467-024-48580-y
[PubMed][PubMedCentral]

27. Tharp KM, Higuchi-Sanabria R, Timblin GA et al (2021) Adhesion-mediated mechanosignaling forces mitohormesis. Cell Metab 33:1322–1341. https://doi.org/10.1016/j.cmet.2021.04.017
[PubMed][PubMedCentral]

28. Garcia G, Bar-Ziv R, Averbukh M et al (2023) Large-scale genetic screens identify BET-1 as a cytoskeleton regulator promoting actin function and life span. Aging Cell 22:e13742. https://doi.org/10.1111/acel.13742
[PubMed]

29. Hurst V, Shimada K, Gasser SM (2019) Nuclear actin and actin-binding proteins in DNA repair. Trends Cell Biol 29:462–476. https://doi.org/10.1016/j.tcb.2019.02.010
[PubMed]

30. Caridi CP, Plessner M, Grosse R, Chiolo I (2019) Nuclear actin filaments in DNA repair dynamics. Nat Cell Biol 21:1068–1077. https://doi.org/10.1038/s41556-019-0379-1
[PubMed][PubMedCentral]

31. Gatti P, Schiavon C, Cicero J et al (2025) Mitochondria- and ER-associated actin are required for mitochondrial fusion. Nat Commun 16:451. https://doi.org/10.1038/s41467-024-55758-x
[PubMed][PubMedCentral]

32. Abrisch RG, Gumbin SC, Wisniewski BT et al (2020) Fission and fusion machineries converge at ER contact sites to regulate mitochondrial morphology. J Cell Biol 219:e201911122. https://doi.org/10.1083/jcb.201911122
[PubMed][PubMedCentral]

33. Wu H, Carvalho P, Voeltz GK (2018) Here, there, and everywhere: the importance of ER membrane contact sites. Science 361:eaan5835. https://doi.org/10.1126/science.aan5835
[PubMed][PubMedCentral]

34. Lewis SC, Uchiyama LF, Nunnari J (2016) ER-mitochondria contacts couple mtDNA synthesis with mitochondrial division in human cells. Science 353:aaf5549. https://doi.org/10.1126/science.aaf5549
[PubMed][PubMedCentral]

35. Manor U, Bartholomew S, Golani G et al (2015) A mitochondria-anchored isoform of the actin-nucleating spire protein regulates mitochondrial division. eLife 4:e08828. https://doi.org/10.7554/eLife.08828
[PubMed][PubMedCentral]

36. Korobova F, Ramabhadran V, Higgs HN (2013) An actin-dependent step in mitochondrial fission mediated by the ER-associated formin INF2. Science 339:464–467. https://doi.org/10.1126/science.1228360
[PubMed][PubMedCentral]

37. Labrousse AM, Zappaterra MD, Rube DA, Van Der Bliek AM (1999) *C. elegans* dynamin-related protein DRP-1 controls severing of the mitochondrial outer membrane. Mol Cell 4:815–826. https://doi.org/10.1016/S1097-2765(00)80391-3 [PubMed]

38. Lee JE, Westrate LM, Wu H et al (2016) Multiple dynamin family members collaborate to drive mitochondrial division. Nature 540:139–143. https://doi.org/10.1038/nature20555 [PubMed][PubMedCentral]

39. Machesky LM, Mullins RD, Higgs HN et al (1999) Scar, a WASp-related protein, activates nucleation of actin filaments by the Arp2/3 complex. Proc Natl Acad Sci USA 96:3739–3744. https://doi.org/10.1073/pnas.96.7.3739 [PubMed][PubMedCentral]

40. Melak M, Plessner M, Grosse R (2017) Actin visualization at a glance. J Cell Sci 130:1688–1688. https://doi.org/10.1242/jcs.204487 [PubMed]

41. Lengsfeld AM, Löw I, Wieland T et al (1974) Interaction of phalloidin with actin. Proc Natl Acad Sci U S A 71:2803–2807. https://doi.org/10.1073/pnas.71.7.2803

42. Phalloidin Conjugates for Actin Staining _ Thermo Fisher Scientific - US. In: ThermoFisher scientific. https://www.thermofisher.com/us/en/home/life-science/cell-analysis/cell-structure/cytoskeleton/phalloidin-and-phalloidin-conjugates-for-staining-actin.html

43. Lukinavičius G, Reymond L, D'Este E et al (2014) Fluorogenic probes for live-cell imaging of the cytoskeleton. Nat Methods 11:731–733. https://doi.org/10.1038/nmeth.2972 [PubMed]

44. Riedl J, Crevenna AH, Kessenbrock K et al (2008) Lifeact: a versatile marker to visualize F-actin. Nat Methods 5:605–607. https://doi.org/10.1038/nmeth.1220 [PubMed][PubMedCentral]

45. Sing CN, Yang EJ, Swayne TC et al (2022) Imaging the actin cytoskeleton in live budding yeast cells. Cytoskeleton 2364:53–80. https://doi.org/10.1007/978-1-0716-1661-1_3

46. Shatskiy D, Sivan A, Wedlich-Söldner R, Belyy A (2025) Structure of the F-tractin–F-actin complex. J Cell Biol 224:e202409192. https://doi.org/10.1083/jcb.202409192 [PubMed][PubMedCentral]

47. Burkel BM, Von Dassow G, Bement WM (2007) Versatile fluorescent probes for actin filaments based on the actin-binding domain of utrophin. Cell Motil Cytoskeleton 64:822–832. https://doi.org/10.1002/cm.20226 [PubMed][PubMedCentral]

48. Wilson DM, Cookson MR, Van Den Bosch L et al (2023) Hallmarks of neurodegenerative diseases. Cell 186:693–714. https://doi.org/10.1016/j.cell.2022.12.032 [PubMed]

49. Castro Torres T, Moaddeli D, Averbukh M et al (2022) Surveying low-cost methods to measure lifespan and healthspan in *Caenorhabditis elegans*. J Vis Exp 183:e64091. https://doi.org/10.3791/64091

50. Rayment I, Holden H, Whittaker M et al (1993) Structure of the actin-myosin complex and its implications for muscle contraction. Science 261:58–65. https://doi.org/10.1126/science.8316858 [PubMed]

51. Lechuga S, Naydenov NG, Feygin A et al.(2020) Loss of β-cytoplasmic actin in the intestinal epithelium increases gut barrier permeability in vivo and exaggerates the severity of experimental colitis. Front Cell Dev Biol 8:588836. https://doi.org/10.3389/fcell.2020.588836 [PubMed][PubMedCentral]

52. Dambroise E, Monnier L, Ruisheng L et al (2016) Two phases of aging separated by the Smurf transition as a public path to death. Sci Rep 6:23523. https://doi.org/10.1038/srep23523 [PubMed][PubMedCentral]

53. Salazar AM, Aparicio R, Clark RI et al (2023) Intestinal barrier dysfunction: an evolutionarily conserved hallmark of aging. Dis Model Mech 16:dmm049969. https://doi.org/10.1242/dmm.049969

[PubMed][PubMedCentral]

54. Egge N, Arneaud SLB, Wales P et al (2019) Age-onset phosphorylation of a minor actin variant promotes intestinal barrier dysfunction. Dev. Cell 51:587–601. https://doi.org/10.1016/j.devcel.2019.11.001 [PubMed][PubMedCentral]

55. Gelino S, Chang JT, Kumsta C et al (2016) Intestinal autophagy improves healthspan and longevity in *C. elegans* during dietary restriction. PLoS Genet 12:e1006135. https://doi.org/10.1371/journal.pgen.1006135 [PubMed][PubMedCentral]

56. Coakley AJ, Hruby AJ, Wang J et al (2025) Distinct responses to non-autonomous UPRER mediated by glutamatergic and octopaminergic neurons. Commun Biol 8:1650. https://doi.org/10.1038/s42003-025-09036-1 [PubMed][PubMedCentral]

57. McBride HM, Neuspiel M, Wasiak S (2006) Mitochondria: more than just a powerhouse. Curr Biol 16:R551–R560. https://doi.org/10.1016/j.cub.2006.06.054 [PubMed]

58. Osellame LD, Blacker TS, Duchen MR (2012) Cellular and molecular mechanisms of mitochondrial function. Best Pract Res Clin Endocrinol Metab 26:711–723. https://doi.org/10.1016/j.beem.2012.05.003

59. Shen K, Pender CL, Bar-Ziv R et al (2022) Mitochondria as cellular and organismal signaling hubs. Annu Rev Cell Dev Biol 38:179–218. https://doi.org/10.1146/annurev-cellbio-120420-015303 [PubMed]

60. Brand MD, Nicholls DG (2011) Assessing mitochondrial dysfunction in cells. Biochem J 435:297–312. https://doi.org/10.1042/BJ20110162 [PubMed][PubMedCentral]

61. Miwa S, Kashyap S, Chini E, Von Zglinicki T (2022) Mitochondrial dysfunction in cell senescence and aging. J Clin Investig 132:e158447. https://doi.org/10.1172/JCI158447 [PubMed][PubMedCentral]

62. Kujoth GC, Hiona A, Pugh TD et al (2005) Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. Science 309:481–484. https://doi.org/10.1126/science.1112125 [PubMed]

63. Scudese E, Marshall AG, Vue Z et al (2025) 3D mitochondrial structure in aging human skeletal muscle: insights Into MFN-2-mediated changes. Aging Cell 24:e70054. https://doi.org/10.1111/acel.70054 [PubMed][PubMedCentral]

64. Jendrach M, Pohl S, Vöth M et al (2005) Morpho-dynamic changes of mitochondria during ageing of human endothelial cells. Mech Ageing Dev 126:813–821. https://doi.org/10.1016/j.mad.2005.03.002 [PubMed]

65. Baloyannis SJ (2006) Mitochondrial alterations in Alzheimer's disease. JAD 9:119–126. https://doi.org/10.3233/JAD-2006-9204 [PubMed]

66. Wang X, Su B, Fujioka H, Zhu X (2008) Dynamin-Like Protein 1 reduction underlies mitochondrial morphology and distribution abnormalities in fibroblasts from sporadic alzheimer's disease patients. Am J Pathol 173:470–482. https://doi.org/10.2353/ajpath.2008.071208 [PubMed][PubMedCentral]

67. Wang X, Su B, Siedlak SL, et al (2008) Amyloid- overproduction causes abnormal mitochondrial dynamics via differential modulation of mitochondrial fission/fusion proteins. In: Proceedings of the National Academy of Sciences of the United States of America, p 105. https://doi.org/10.1073/pnas.0804871105

68. Peoples JN, Saraf A, Ghazal N et al (2019) Mitochondrial dysfunction and oxidative stress in heart disease. Exp Mol Med 51:1–13. https://doi.org/10.1038/s12276-019-0355-7 [PubMed][PubMedCentral]

69. Kim J, Dutta N, Garcia G et al (2026) Transcriptomic analysis of mitohormesis associated with lifespan extension in *Caenorhabditis elegans*. GeroScience 48:5673–5693. https://doi.org/10.1007/s11357-025-01912-2

[PubMed][PubMedCentral]

70. Hruby AJ, Higuchi-Sanabria R (2025) Mitochondrial dysfunction in cellular senescence: a bridge to neurodegenerative disease. npj Aging 11:99. https://doi.org/10.1038/s41514-025-00291-4
[PubMed][PubMedCentral]

71. Sorvina A, Bader CA, Darby JRT et al (2018) Mitochondrial imaging in live or fixed tissues using a luminescent iridium complex. Sci Rep 8:8191. https://doi.org/10.1038/s41598-018-24672-w
[PubMed][PubMedCentral]

72. Chen J, Stephan T, Gaedke F et al (2024) An aldehyde-crosslinking mitochondrial probe for STED imaging in fixed cells. Proc Natl Acad Sci U S A 121:e2317703121. https://doi.org/10.1073/pnas.2317703121
[PubMed][PubMedCentral]

73. Cottet-Rousselle C, Ronot X, Leverve X, Mayol J (2011) Cytometric assessment of mitochondria using fluorescent probes. Cytometry Pt A 79A:405–425. https://doi.org/10.1002/cyto.a.21061

74. Chen C-S, Gee KR (2000) Redox-dependent trafficking of 2,3,4,5,6-pentafluorodihydrotetramethylrosamine, a novel fluorogenic indicator of cellular oxidative activity. Free Radic Biol Med 28:1266–1278. https://doi.org/10.1016/S0891-5849(00)00265-3
[PubMed]

75. Cossarizza A, Baccaranicontri M, Kalashnikova G, Franceschi C (1993) A new method for the cytofluorometric analysis of mitochondrial membrane potential using the J-aggregate forming lipophilic cation 5,5′,6,6′-Tetrachloro-1,1′,3,3′-tetraethylbenzimidazolcarbocyanine Iodide (JC-1). Biochem Biophys Res Commun 197:40–45. https://doi.org/10.1006/bbrc.1993.2438
[PubMed]

76. Dingley S, Polyak E, Lightfoot R et al (2010) Mitochondrial respiratory chain dysfunction variably increases oxidant stress in *Caenorhabditis elegans*. Mitochondrion 10:125–136. https://doi.org/10.1016/j.mito.2009.11.003
[PubMed]

77. Kim J, Dutta N, Vega M, et al (2025) Cross-comparison of strains used for mitochondrial imaging in *Caenorhabditis elegans* during aging. Life Sci Alliance 8(8):e202403189. https://doi.org/10.26508/lsa.202403189

78. Zhang Y, Lanjuin A, Chowdhury SR et al (2019) Neuronal TORC1 modulates longevity via AMPK and cell nonautonomous regulation of mitochondrial dynamics in *C. elegans*. eLife 8:e49158. https://doi.org/10.7554/eLife.49158
[PubMed][PubMedCentral]

79. Zorova LD, Popkov VA, Plotnikov EY et al (2018) Mitochondrial membrane potential. Anal Biochem 552:50–59. https://doi.org/10.1016/j.ab.2017.07.009
[PubMed]

80. Divakaruni AS, Rogers GW, Murphy AN (2014) Measuring Mitochondrial Function in Permeabilized Cells Using the Seahorse XF Analyzer or a Clark-Type Oxygen Electrode. CP Toxicol 60(1):25.2.1–25.2.16. https://doi.org/10.1002/0471140856.tx2502s60

81. Chiang JL, Shukla P, Pagidas K et al (2020) Mitochondria in ovarian aging and reproductive longevity. Ageing Res Rev 63:101168. https://doi.org/10.1016/j.arr.2020.101168
[PubMed][PubMedCentral]

82. Sevini F, Giuliani C, Vianello D et al (2014) mtDNA mutations in human aging and longevity: controversies and new perspectives opened by high-throughput technologies. Exp Gerontol 56:234–244. https://doi.org/10.1016/j.exger.2014.03.022
[PubMed]

83. Jain A, Zoncu R (2022) Organelle transporters and inter-organelle communication as drivers of metabolic regulation and cellular homeostasis. Mol Metab 60:101481
[https://doi.org/10.1016/j.molmet.2022.101481][PubMed][PubMedCentral]

84. Voeltz GK, Sawyer EM, Hajnóczky G, Prinz WA (2024) Making the connection: how membrane contact sites have changed our view of organelle biology. Cell 187:257–270. https://doi.org/10.1016/j.cell.2023.11.040
[PubMed][PubMedCentral]

85. Friedman JR, Lackner LL, West M et al (2011) ER tubules mark sites of mitochondrial division. Science 334:358–362. https://doi.org/10.1126/science.1207385
[PubMed][PubMedCentral]

86. Wang X, Li S, Wang H et al (2017) Quantitative proteomics reveal proteins enriched in tubular endoplasmic reticulum of *Saccharomyces cerevisiae*. eLife 6:e23816. https://doi.org/10.7554/eLife.23816

87. Shibata Y, Shemesh T, Prinz WA et al (2010) Mechanisms determining the morphology of the peripheral ER. Cell 143:774–788. https://doi.org/10.1016/j.cell.2010.11.007
[PubMed][PubMedCentral]

88. Voeltz GK, Prinz WA, Shibata Y et al (2006) A class of membrane proteins shaping the tubular endoplasmic reticulum. Cell 124:573–586. https://doi.org/10.1016/j.cell.2005.11.047
[PubMed]

89. Schwarz DS, Blower MD (2016) The endoplasmic reticulum: structure, function and response to cellular signaling. Cell Mol Life Sci 73:79–94. https://doi.org/10.1007/s00018-015-2052-6
[PubMed]

90. Metcalf MG, Higuchi-Sanabria R, Garcia G et al (2020) Beyond the cell factory: homeostatic regulation of and by the UPRER. Sci Adv 6:eabb9614. https://doi.org/10.1126/sciadv.abb9614
[PubMed][PubMedCentral]

91. Donahue EKF, Ruark EM, Burkewitz K (2022) Fundamental roles for inter-organelle communication in aging. Biochem Soc Trans 50:1389–1402. https://doi.org/10.1042/BST20220519
[PubMed][PubMedCentral]

92. Haithcock E, Dayani Y, Neufeld E et al (2005) Age-related changes of nuclear architecture in *Caenorhabditis elegans*. Proc Natl Acad Sci U S A 102:16690–16695. https://doi.org/10.1073/pnas.0506955102
[PubMed][PubMedCentral]

93. Martins F, Sousa J, Pereira CD et al (2020) Nuclear envelope dysfunction and its contribution to the aging process. Aging Cell 19:e13143. https://doi.org/10.1111/acel.13143
[PubMed][PubMedCentral]

94. Scaffidi P, Misteli T (2006) Lamin A-dependent nuclear defects in human aging. Science 312:1059–1063. https://doi.org/10.1126/science.1127168
[PubMed][PubMedCentral]

95. Daniele JR, Higuchi-Sanabria R, Durieux J et al (2020) UPR ^{ER} promotes lipophagy independent of chaperones to extend life span. Sci Adv 6:eaaz1441. https://doi.org/10.1126/sciadv.aaz1441
[PubMed][PubMedCentral]

96. Taylor RC, Dillin A (2013) XBP-1 Is a cell-nonautonomous regulator of stress resistance and longevity. Cell 153:1435–1447. https://doi.org/10.1016/j.cell.2013.05.042
[PubMed][PubMedCentral]

97. Jacquemyn J, Cascalho A, Goodchild RE (2017) The ins and outs of endoplasmic reticulum-controlled lipid biosynthesis. EMBO Rep 18:1905–1921. https://doi.org/10.15252/embr.201643426
[PubMed][PubMedCentral]

98. Higuchi-Sanabria R, Durieux J, Kelet N et al (2020) Divergent nodes of non-autonomous uprer signaling through serotonergic and dopaminergic neurons. Cell Rep 33:108489. https://doi.org/10.1016/j.celrep.2020.108489
[PubMed][PubMedCentral]

99. Papsdorf K, Miklas JW, Hosseini A et al (2023) Lipid droplets and peroxisomes are co-regulated to drive lifespan extension in response to mono-unsaturated fatty acids. Nat Cell Biol 25:672–684. https://doi.org/10.1038/s41556-023-01136-6
[PubMed][PubMedCentral]

100. Grant B, Hirsh D (1999) Receptor-mediated Endocytosis in the *Caenorhabditis elegans* Oocyte. MBoC 10:4311–4326. https://doi.org/10.1091/mbc.10.12.4311
[PubMed][PubMedCentral]

101. Lange CM, Higuchi-Sanabria R, Kumsta C (2025) Autophagy in proteostasis and aging in *Caenorhabditis elegans*. Cell Stress Chaperones 30:100115. https://doi.org/10.1016/j.cstres.2025.100115 [PubMed][PubMedCentral]

102. Bazopoulou D, Knoefler D, Zheng Y et al (2019) Developmental ROS individualizes organismal stress resistance and lifespan. Nature 576:301–305. https://doi.org/10.1038/s41586-019-1814-y [PubMed][PubMedCentral]

103. Yang Y, Liu N, He Y et al (2018) Improved calcium sensor GCaMP-X overcomes the calcium channel perturbations induced by the calmodulin in GCaMP. Nat Commun 9:1504. https://doi.org/10.1038/s41467-018-03719-6 [PubMed][PubMedCentral]

104. Urban ND, Lacy SM, Van Pelt KM et al (2025) Functionally diversified *Caenorhabditis elegans* BiP orthologs control body growth, reproduction, stress resistance, aging, and autophagy. Nat Commun 16:11094. https://doi.org/10.1038/s41467-025-65998-0 [PubMed][PubMedCentral]

105. Timmons L, Court DL, Fire A (2001) Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. Gene 236:103–112. https://doi.org/doi:%2010.1016/s0378-1119(00)00579-5 [https://doi.org/10.1016/S0378-1119(00)00579-5]

106. Kamath RS, Martinez-Campos M, Zipperlen P, et al (2000) Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. Genome Biol 2(1):RESEARCH0002. https://doi.org/10.1186/gb-2000-2-1-research0002

107. Rual J-F, Ceron J, Koreth J et al (2004) Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. Genome Res 14:2162–2168. https://doi.org/10.1101/gr.2505604 [PubMed][PubMedCentral]

108. Kamath R (2003) Genome-wide RNAi screening in *Caenorhabditis elegans*. Methods 30:313–321. https://doi.org/10.1016/S1046-2023(03)00050-1 [PubMed]

109. Jagadeesan S, Hakkim A (2018) RNAi screening: automated high-throughput liquid RNAi screening in *Caenorhabditis elegans*. CP Mol Biol 124:e65. https://doi.org/10.1002/cpmb.65

110. Garcia G, Homentcovschi S, Kelet N, Higuchi-Sanabria R (2022) Imaging of actin cytoskeletal integrity during aging in *C. elegans*. In: Gavin RH (ed) Cytoskeleton. Springer US, New York, NY, pp 101–137

111. Reinke SN, Hu X, Sykes BD, Lemire BD (2010) *Caenorhabditis elegans* diet significantly affects metabolic profile, mitochondrial DNA levels, lifespan and brood size. Mol Genet Metab 100:274–282. https://doi.org/10.1016/j.ymgme.2010.03.013 [PubMed]

112. Revtovich AV, Lee R, Kirienko NV (2019) Interplay between mitochondria and diet mediates pathogen and stress resistance in *Caenorhabditis elegans*. PLoS Genet 15:e1008011. https://doi.org/10.1371/journal.pgen.1008011 [PubMed][PubMedCentral]

113. Sutphin GL, Kaeberlein M (2009) Measuring *Caenorhabditis elegans* life span on solid media. J Vis Exp 27:1152. https://doi.org/10.3791/1152

114. Gardner M, Rosell M, Myers EM (2013) Measuring the effects of bacteria on *C. elegans* behavior using an egg retention assay. J Vis Exp 80:e51203. https://doi.org/10.3791/51203

115. Mitchell DH, Stiles JW, Santelli J, Sanadi DR (1979) Synchronous growth and aging of *Caenorhabditis elegans* in the presence of fluorodeoxyuridine. J Gerontol 34:28–36. https://doi.org/10.1093/geronj/34.1.28 [PubMed]

116. Gao SM, Qi Y, Zhang Q et al (2024) Aging atlas reveals cell-type-specific effects of pro-longevity strategies. Nat Aging 4:998–1013. https://doi.org/10.1038/s43587-024-00631-1 [PubMed][PubMedCentral]

117. Zuo Y, Hu J, Xu X et al (2019) Sodium azide induces mitochondria-mediated apoptosis in PC12 cells through Pgc-1α-

associated signaling pathway. Mol Med Rep 19:2211–2219. https://doi.org/10.3892/mmr.2019.9853

118. Moehle EA, Higuchi-Sanabria R, Tsui CK et al (2021) Cross-species screening platforms identify EPS-8 as a critical link for mitochondrial stress and actin stabilization. Sci Adv 7:eabj6818. https://doi.org/10.1126/sciadv.abj6818 [PubMed][PubMedCentral]

119. Valera-Alberni M, Yao P, Romero-Sanz S et al (2024) Novel imaging tools to study mitochondrial morphology in *Caenorhabditis elegans*. Life Sci Alliance 7:e202402918. https://doi.org/10.26508/lsa.202402918 [PubMed][PubMedCentral]

120. Raiders SA, Eastwood MD, Bacher M, Priess JR (2018) Binucleate germ cells in *Caenorhabditis elegans* are removed by physiological apoptosis. PLoS Genet 14:e1007417. https://doi.org/10.1371/journal.pgen.1007417 [PubMed][PubMedCentral]

121. Liu X, Long F, Peng H, et al (2009) Analysis of cell fate from single-cell gene expression profiles in *C. elegans*. Cell 139:623–633. https://doi.org/10.1016/j.cell.2009.08.044 [PubMed][PubMedCentral]

122. Benedetti C, Haynes CM, Yang Y et al (2006) Ubiquitin-like protein 5 positively regulates chaperone gene expression in the mitochondrial unfolded protein response. Genetics 174:229–239. https://doi.org/10.1534/genetics.106.061580 [PubMed][PubMedCentral]

123. Schindelin J, Arganda-Carreras I, Frise E et al (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods 9:676–682. https://doi.org/10.1038/nmeth.2019 [PubMed][PubMedCentral]

124. Terasaki M, Shemesh T, Kasthuri N et al (2013) Stacked endoplasmic reticulum sheets are connected by helicoidal membrane motifs. Cell 154:285–296. https://doi.org/10.1016/j.cell.2013.06.031 [PubMed][PubMedCentral]

125. Thomas BJ, Wight IE, Chou WYY et al (2019) CemOrange2 fusions facilitate multifluorophore subcellular imaging in *C. elegans*. PLoS One 14:e0214257. https://doi.org/10.1371/journal.pone.0214257 [PubMed][PubMedCentral]

126. Poteryaev D, Squirrell JM, Campbell JM et al (2005) Involvement of the actin cytoskeleton and homotypic membrane fusion in ER dynamics in *Caenorhabditis elegans*. MBoC 16:2139–2153. https://doi.org/10.1091/mbc.e04-08-0726 [PubMed][PubMedCentral]

127. Joseph-Strauss D, Gorjánácz M, Santarella-Mellwig R et al (2012) Sm protein down-regulation leads to defects in nuclear pore complex disassembly and distribution in *C. elegans* embryos. Dev Biol 365:445–457. https://doi.org/10.1016/j.ydbio.2012.02.036 [PubMed][PubMedCentral]

128. Xu N, Zhang SO, Cole RA et al (2012) The FATP1–DGAT2 complex facilitates lipid droplet expansion at the ER–lipid droplet interface. J Cell Biol 198:895–911. https://doi.org/10.1083/jcb.201201139 [PubMed][PubMedCentral]

129. Li X, Itani OA, Haataja L et al (2019) Requirement for translocon-associated protein (TRAP) α in insulin biogenesis. Sci Adv 5:eaax0292. https://doi.org/10.1126/sciadv.aax0292 [PubMed][PubMedCentral]

130. Kelley CA, Triplett O, Mallick S et al (2020) FLN -1/filamin is required to anchor the actomyosin cytoskeleton and for global organization of sub-cellular organelles in a contractile tissue. Cytoskeleton 77:379–398. https://doi.org/10.1002/cm.21633 [PubMed][PubMedCentral]

131. Serot C, Scarcelli V, Pouget A et al (2025) Reticulon-dependent ER-phagy mediates adaptation to heat stress in *C. elegans*. Curr Biol 35:2365–2378. https://doi.org/10.1016/j.cub.2025.04.028 [PubMed]

132. Kim E, Sun L, Gabel CV, Fang-Yen C (2013) Long-term imaging of *Caenorhabditis elegans* using nanoparticle-mediated immobilization. PLoS One 8:e53419. https://doi.org/10.1371/journal.pone.0053419 [PubMed][PubMedCentral]

133. Oorloff M, Hruby A, Averbukh M et al (2024) Growth on stiffer substrates impacts animal health and longevity in *C. elegans*. PLoS One 19:e0302673. https://doi.org/10.1371/journal.pone.0302673 [PubMed][PubMedCentral]

134. Castro Torres T, Moaddeli D, Averbukh M, et al (2022) Surveying Low-Cost Methods to Measure Lifespan and Healthspan in *Caenorhabditis elegans*. J Vis Exp. https://doi.org/10.3791/64091

135. Rasmussen NR, Reiner DJ (2021) Nuclear translocation of the tagged endogenous MAPK MPK-1 denotes a subset of activation events in *C. elegans* development. J Cell Sci 134:jcs258456. https://doi.org/10.1242/jcs.258456 [PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_17

17. Measurements of Nuclear Pore Function in Living *Caenorhabditis elegans*

Yifei Zhou[1] and Alexander A. Soukas[2, 3] ✉
(1) Department of Pharmaceutical Analysis, School of Pharmacy, China Pharmaceutical University, Nanjing, 211198, China
(2) Center for Genomic Medicine and Diabetes Unit, Endocrine Division, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
(3) Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

✉ **Alexander A. Soukas**
Email: asoukas@mgh.harvard.edu

Abstract

The nuclear pore complex (NPC) is the sole gateway governing nucleocytoplasmic trafficking, serving as both a barrier for passive diffusion and a selective channel for active transport. Proper NPC integrity is critical for most fundamental cellular processes. However, accumulating evidence indicates that both NPC barrier function and active transport efficiency decline with aging, which contributes significantly to age-related disorders. Here, we describe in vivo methods to assess NPC function in living *Caenorhabditis elegans*, including quantitative measurements of active nuclear transport based on the nuclear accumulation of transport cargoes, as well as assays of passive diffusion derived from the diffusive dynamics of green

fluorescent protein (GFP) in living animals. These approaches enable systematic analysis of age-associated changes in NPC function and provide a practical framework for investigating the contribution of NPC dysfunction to aging and age-related disorders.

Keywords Nuclear pore complex – Aging – Nuclear transport – Nuclear localization signal – Passive diffusion – Fluorescence recovery after photobleaching (FRAP)

1 Introduction

As one of the largest protein assemblies in eukaryotic cells, the massive NPC consists of more than 1000 copies of approximately 35 distinct nucleoporins [1–3]. Embedded within the nuclear envelope, the NPC precisely governs the barrier for passive diffusion of proteins over 40 kDa while permitting the active transport of specific cargoes bearing nuclear localization signals (NLSs) or nuclear export signals (NESs). In addition, the NPC serves as the primary conduit for the export of multiple classes of RNAs [4]. Given its critical action in molecular exchange between the nucleus and the cytosol, NPC integrity is essential for a wide range of fundamental cellular processes. Increasing evidence indicates that NPC function declines during aging and in age-associated disorders [5–9]. Alterations in NPC components have been reported in multiple neurodegenerative diseases and are frequently accompanied by defects in nuclear transport [7]. Moreover, the NPC permeability barrier becomes progressively compromised with age across species, leading to dysregulated passive diffusion through the nuclear envelope [5]. This defect can be reversed by treatment with the antidiabetic metformin or the related biguanide phenformin [10]. Consistent with these observations, maintenance of NPC integrity has been shown to be required for several longevity paradigms [5, 10, 11]. Together, these findings underscore the importance of quantitatively evaluating NPC integrity across the lifespan. Such measurements

provide a valuable tool for aging research and may serve as a sensitive, cell-level biomarker of age-associated functional decline.

The dextran assay, which assesses nuclear permeability to passively transported cargoes by incubating fluorescently labeled dextrans of defined molecular weights with isolated nuclei or cells with permeabilized plasma membrane, is a widely used and well-described in vitro method for assessing the barrier function of NPC in both nematode and mammalian systems [5, 12]. However, NPC barrier properties are also influenced by the native subcellular environment. For example, mitochondrial factors have been shown to act upstream in regulating NPC barrier restriction in response to biguanide treatment [10], underscoring the need for in vivo assays to evaluate NPC function in living cells and animals. It should be noted that actively transported nuclear cargoes are also subject to passive transport through the NPC in a manner dependent upon concentration gradient. As such, changes in nucleocytoplasmic transport in aging stand to alter subcellular distribution of both actively and passively transported cargoes.

Thus, here we describe two in vivo assays to evaluate both the passive diffusion barrier and active transport functions of the NPC in living worms. Passive diffusion is assessed by fluorescence recovery after photobleaching (FRAP) of GFP in intestinal cells, whereas active nuclear transport is evaluated by quantifying the nuclear localization of NLS(SV40)::GFP [13] (Fig. 1). These assays provide several advantages over existing in vitro approaches: (1) they avoid mechanistic perturbation, which is an obligate consequence of isolation of cellular nuclei as well as over-permeabilization associated with digitonin-based treatments; (2) they preserve the native subcellular microenvironment during functional assessment; and (3) they minimize physical damage to the animals, enabling longitudinal test-rescue experiments and time-course quantification of NPC function across the lifespan of individual worms.

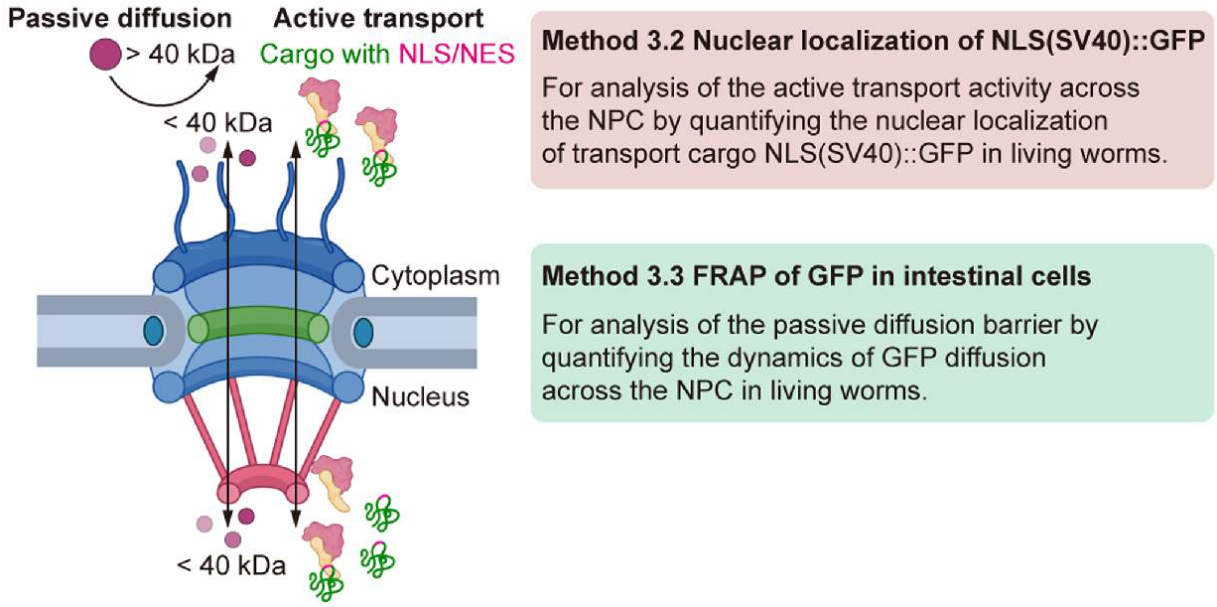


Fig. 1 Schematic overview of nucleocytoplasmic transport and experimental approaches used to assess NPC function. Small proteins traverse the NPC by passive diffusion, whereas proteins containing NLSs or NESs are transported across the NPC by active transport machinery. FRAP of GFP and analysis of NLS(SV40)::GFP nuclear localization is used to evaluate the passive diffusion barrier and active transport activity, respectively, in living worms

2 Materials

2.1 Worm Maintenance

1. Normal growth media (NGM): Add 25 g Bacto agar, 2.5 g Bacto peptone, and 3 g NaCl in 975 mL distilled H_2O. Autoclave on liquid cycle with a magnetic stir-bar in the vessel. Cool down to about 60 °C after autoclaving. Add 1 mL of 1 M $CaCl_2$, 1 mL of 1 M $MgSO_4$, 1 mL of 5 mg/mL cholesterol in ethanol, and 25 mL of 1 M KPO_4 (pH 6.0). Add streptomycin or carbenicillin based on bacterial resistance and place on heated stirring plate to maintain temperature and assure adequate mixing while pouring plates. Pour ~9 mL of medium into each 6-cm dish for worm culture using a sterile pipet or peristaltic pump outfitted with pre-sterilized tubing.
2. M9 buffer: Dissolve 3 g KH_2PO_4, 6 g Na_2HPO_4, and 5 g NaCl in distilled water to a final volume of 1 L. Autoclave, allow to cool to

room temperature, and then add 1 mL of sterile 1 M $MgSO_4$.

3. OP50: Pick a single colony of *Escherichia coli* OP50 and inoculate it into a 2-L baffled flask containing 1 L LB broth supplemented with streptomycin. Incubate at 37 °C non-shaking overnight. Seed 200 µL of the bacterial culture onto NGM plates (6-cm dish).

4. Worm strains: The *sur-5p::NLS(SV40)::GFP* transgenic strain is used to assess active nucleocytoplasmic transport, whereas the *vha-6p::GFP* strain is used to evaluate the passive diffusion barrier of the NPC (*see* **Note 1**).

2.2 Microscopy

1. Agarose (2%): Dissolve 2 g agarose powder in 100 mL distilled water and melt the agarose by heating in a microwave oven until fully dissolved.

2. Levamisole (5 mM): Prepare a 1 M levamisole stock solution in distilled water. Dilute the stock solution to 5 mM immediately before use as the working concentration (*see* **Note 2**).

3. Microscope slides and coverslips.

4. Microscopes: A Leica DM6 B microscope equipped with a THUNDER Imager, using a 5× objective and a 63× objective, is used for imaging NLS(SV40)::GFP. FRAP experiments are performed on a Zeiss LSM 800 confocal microscope equipped with an Airyscan detector and a 63×/1.4 oil immersion DIC M27 objective (*see* **Note 3**).

2.3 Data Collection and Analysis

1. Fluorescence quantification: ImageJ software, Leica LAS X, and ZEISS ZEN Microscopy Software.

2. Data analysis: GraphPad Prism software.

---

3 Methods

3.1 Preparation for Microscopic Imaging

1. For each experiment, prepare approximately 20 synchronized worms at the indicated developmental stages or ages for imaging (*see* **Note 4**).
2. Prepare molten agarose by boiling 2% (w/v) agarose in distilled water and allowing the solution to cool slightly before use.
3. Prepare agarose pads by dispensing one to two drops of molten agarose onto a glass slide and flattening it to the thickness of a single layer of laboratory tape.
4. Pipette 5 μl of 5 mM levamisole onto the agarose pad.
5. Transfer approximately 20 synchronized worms into the levamisole solution and incubate for ~5 min until the worms are paralyzed.

3.2 Nuclear localization of NLS(SV40)::GFP

1. For low-magnification imaging, arrange the immobilized worms side by side on the agarose pad. Acquire bright-field and GFP fluorescence images using a 5× objective, without placing a coverslip on the sample (Fig. 2A).

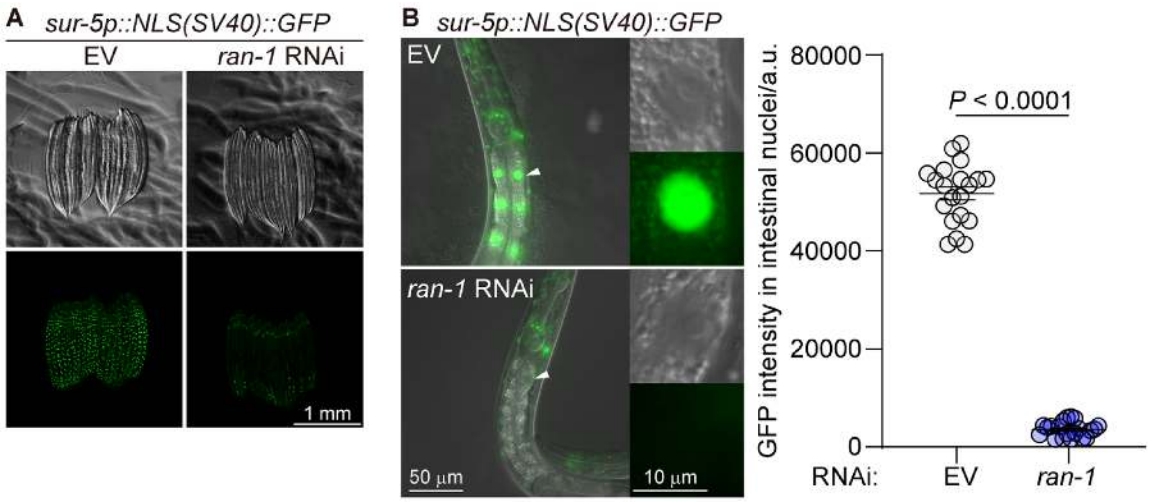


Fig. 2 Quantification of nuclear NLS(SV40)::GFP in intestinal cells. (**A**) Low-magnification and (**B**) high-magnification images of *sur-5p::NLS(SV40)::GFP* transgenic worms at the L4 stage. RNAi-mediated knockdown of *ran-1* disrupts nuclear localization of the active transport cargo NLS(SV40)::GFP in intestinal nuclei (white arrowhead). Empty vector (EV) RNAi serves as a negative control. Scale bars: 1 mm (A), 50 mm (B, left), and 10 mm (B, right). Data are presented as mean ± SD. a.u., arbitrary units. Statistical significance was determined using an unpaired *t*-test

2. For rapid assessment of nuclear localization at low magnification, evaluate GFP signals appearing as discrete puncta corresponding to nuclei within the worm body and score relative nuclear GFP intensity accordingly (Fig. 2A, *see* **Note 5**).
3. For high-magnification imaging, add an additional 5 µl of 5 mM levamisole and gently place a glass coverslip over the agarose pad, taking care to avoid air bubbles between the coverslip and the agarose surface.
4. Acquire bright-field and GFP fluorescence images using a 60× objective (Fig. 2B).
5. Quantify fluorescence intensity in the intestinal cells using ImageJ or Fiji according to standard image analysis procedures.

3.3 FRAP of GFP in Intestinal Cells

1. Pick 20 worms into the levamisole and wait about 5 min until the

worms are paralyzed.

2. Gently place a glass coverslip over the agarose pad, taking care to avoid air bubbles between the coverslip and the agarose surface.
3. Perform FRAP using a confocal microscope (e.g., Zeiss LSM 800), following Steps 4–11.
4. Focus on intestinal cell nuclei and position the sample using a 63×/1.4 oil immersion DIC M27 objective on a Zeiss LSM 800 confocal microscope equipped with an Airyscan detector.
5. In the Acquisition tab, enable **Regions, Time Series**, and **Bleaching** modes.
6. Configure the **Bleaching** panel (*see* **Note 6**) as follows:
   a. Start Bleaching after # scans: 2;
   b. Select the 488-nm laser for bleaching and set laser power to 100%;
   c. Disable all other bleaching options.
7. Configure the **Time Series** panel (*see* **Note 7**) with 30 cycles and a 5-s interval between scans.
8. Define a region of interest (ROI) within the target nucleus using the ellipse tool.
9. Select the ROI for both bleaching and fluorescence analysis.
10. Initiate image acquisition to perform the FRAP experiment.
11. Analyze FRAP data by quantifying GFP fluorescence intensity within the ROI at each time point. Normalize fluorescence

intensity to the first pre-bleach image, which is set to 100% (Fig. 3B).

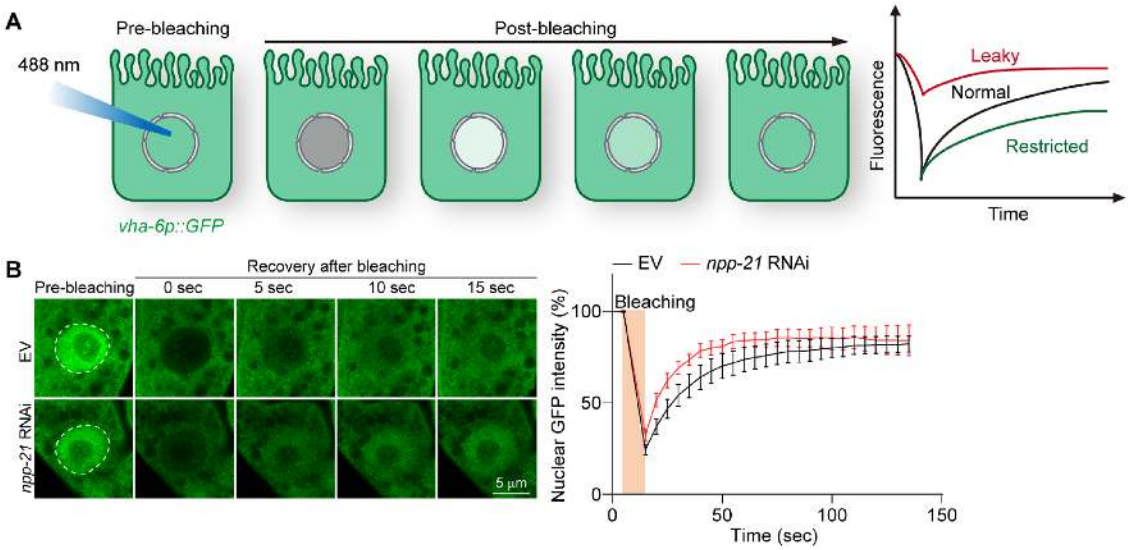


Fig. 3 FRAP analysis of GFP diffusion in intestinal cells. (**A**) Schematic illustration of the FRAP assay used to assess nuclear GFP diffusion in intestinal cells. (**B**) Quantification of nuclear GFP fluorescence in *vha-6p::GFP* transgenic worms at day 1 of adulthood (D1). RNAi-mediated knockdown of *npp-21* accelerates GFP diffusion from the cytoplasm into the nucleus. Scale bar: 5 mm. Data are presented as mean ± SEM

4 Notes

1. The worm strains used in this chapter are not limited to the two strains described here. Other transgenic strains expressing nuclear transport cargoes or fluorescent proteins can also be used to assess NPC function. Intestinal cells were selected in this chapter because of their large nuclei. However, other tissues may be examined by expressing NLS(SV40)::GFP or GFP under the control of tissue-specific promoters.
2. These assays are designed to assess NPC function in living nematodes. Therefore, sodium azide should not be used for

immobilization because of its mitochondrial toxicity and irreversible effects on cellular physiology.

3. Confocal microscopy is advantageous but not required for assessing nuclear localization of NLS(SV40)::GFP, which can be reliably evaluated using widefield fluorescence microscopy. In contrast, confocal microscopy is required for FRAP experiments to enable precise photobleaching and accurate measurement of fluorescence recovery.

4. Because short-term incubation with levamisole is minimally toxic and allows worms to remain viable, imaged animals can be recovered by gently releasing them from the agarose pad and transferring them back to nematode growth medium (NGM) plates after washing with M9 buffer. Consequently, these assays can be performed repeatedly at distinct ages in the same individuals, enabling time-course analysis of NPC function across the lifespan.

5. Nuclear localization of NLS(SV40)::GFP can be scored manually using a semi-quantitative classification system that categorizes images as “low,” “medium,” or “strong” nuclear signal. Empty vector (EV) and *ran-1* RNAi treatments may be used as representative controls for low and strong nuclear localization, respectively (Fig. 2A).

6. Bleaching parameters can be adjusted depending on the expression level and photostability of the fluorescent reporters. For example, bleaching laser power may be reduced to 50% for weak GFP reporters. For red fluorescent protein (RFP) reporters, the bleaching wavelength can be switched to 561 nm.

7. Time series acquisition parameters can be customized. The interval between scans may increase when fluorescence recovery is slow to minimize photobleaching and excessive laser exposure.

Acknowledgments

This work was supported by NIH/NIA grants R01AG058256 and R01AG069677 (to A.A.S.), the Weissman Family MGH Research Scholar Award (to A.A.S.), and a gift from Stuart and Suzanne Steele and the Obesity Research Fund at the Mass General Research Institute Center for Genomic Medicine. This research was conducted in part while Y.Z. was a Glenn Foundation for Medical Research Postdoctoral Fellow.

Conflict of Interest

Alexander A. Soukas has financial interests in Atman Health, LLC. Dr. Soukas' interest was reviewed and is managed by MGH and Mass General Brigham in accordance with their conflict of interest policies.

References

1. Strambio-De-Castillia C, Niepel M, Rout MP (2010) The nuclear pore complex: bridging nuclear transport and gene regulation. Nat Rev Mol Cell Biol 11(7):490–501. https://doi.org/10.1038/nrm2928
[PubMed]

2. Beck M, Hurt E (2017) The nuclear pore complex: understanding its function through structural insight. Nat Rev Mol Cell Biol 18(2):73–89. https://doi.org/10.1038/nrm.2016.147
[PubMed]

3. Sun J, Shi Y, Yildirim E (2019) The nuclear pore complex in cell type-specific chromatin structure and gene regulation. Trends Genet 35(8):579–588. https://doi.org/10.1016/j.tig.2019.05.006
[PubMed]

4. Kohler A, Hurt E (2007) Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol 8(10):761–773. https://doi.org/10.1038/nrm2255
[PubMed]

5. D'Angelo MA, Raices M, Panowski SH et al (2009) Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. Cell 136(2):284–295. https://doi.org/10.1016/j.cell.2008.11.037
[PubMed][PubMedCentral]

6. Bitetto G, Di Fonzo A (2020) Nucleo-cytoplasmic transport defects and protein aggregates in neurodegeneration. Transl Neurodegener 9(1):25. https://doi.org/10.1186/s40035-020-00205-2
[PubMed][PubMedCentral]

7. Grima JC, Daigle JG, Arbez N et al (2017) Mutant huntingtin disrupts the nuclear pore complex. Neuron 94(1):93–107. https://doi.org/10.1016/j.neuron.2017.03.023 [PubMed][PubMedCentral]

8. Rempel IL, Crane MM, Thaller DJ et al (2019) Age-dependent deterioration of nuclear pore assembly in mitotic cells decreases transport dynamics. Elife 8. https://doi.org/10.7554/eLife.48186

9. Martins F, Sousa J, Pereira CD et al (2020) Nuclear envelope dysfunction and its contribution to the aging process. Aging cell 19(5):e13143. https://doi.org/10.1111/acel.13143 [PubMed][PubMedCentral]

10. Wu L, Zhou B, Oshiro-Rapley N et al (2016) An ancient, unified mechanism for metformin growth inhibition in *C. elegans* and cancer. Cell 167(7):1705–1718. https://doi.org/10.1016/j.cell.2016.11.055 [PubMed][PubMedCentral]

11. Yu Y, Gao SM, Guan Y et al (2024) Organelle proteomic profiling reveals lysosomal heterogeneity in association with longevity. Elife 13. https://doi.org/10.7554/eLife.85214

12. Raices M, D'Angelo MA (2022) Analysis of nuclear pore complex permeability in mammalian cells and isolated nuclei using fluorescent dextrans. Methods Mol Biol 2502:69–80. https://doi.org/10.1007/978-1-0716-2337-4_4 [PubMed][PubMedCentral]

13. Zhou Y, Ahsan FM, Emans SW et al (2025) The nuclear pore complex connects energy sensing to transcriptional plasticity in longevity. Mol Cell 85:3605–3622. https://doi.org/10.1016/j.molcel.2025.08.035 [PubMed]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_18

18. Measurement of Skeletal Muscle Mitochondrial DNA Synthesis Using Stable Isotope Labeling

Robert V. Musci[1], Morgan Johnson[2], Noelle Morrow[2], Clover D. Crenshaw[1], Morgan K. Daniel[1], Frederick F. Peelor[3], Benjamin F. Miller[3, 4], Johanna Pham[5], Andrea L. Hevener[2] and Jonathan Wanagat[5, 6] ✉

(1) Loyola Marymount University, Los Angeles, CA 90045, USA
(2) Division of Endocrinology, Department of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA
(3) Aging and Metabolism Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA
(4) Oklahoma City Veterans Affairs Medical Center, Oklahoma City, OK 73104, USA
(5) Division of Geriatrics, Department of Medicine, University of California—Los Angeles, Los Angeles, CA 90095, USA
(6) Geriatric Research, Education and Clinical Center (GRECC), VA Greater Los Angeles Healthcare System, Los Angeles, CA 90073, USA

✉ **Jonathan Wanagat**
Email: jwanagat@mednet.ucla.edu

Abstract

Stable isotope labeling with deuterium oxide provides a robust approach for quantifying mitochondrial DNA synthesis and turnover in

vivo. This chapter describes a protocol for measuring mitochondrial DNA fractional synthesis rate and half-life in frozen samples of rodent skeletal muscle. The method combines long-term deuterium oxide labeling with isolation of mitochondria from frozen tissue, enzymatic removal of contaminating nuclear DNA using DNase, purification of mitochondrial DNA, and measurement of deuterium incorporation into deoxyribose by gas chromatography–mass spectrometry. This approach enables quantitative assessment of mitochondrial DNA turnover under physiological and experimental conditions while minimizing confounding from nuclear DNA contamination.

Keywords Mitochondrial DNA – Skeletal muscle – Stable isotope – Deuterium oxide – Fractional synthesis – Half life

1 Introduction

Mitochondrial DNA (mtDNA) turnover is the balance between the replication and loss of mtDNA. The turnover of mtDNA may contribute to mitochondrial quality control in skeletal muscle [1–3]. Measurement of mtDNA synthesis in vivo may provide insight into mitochondrial maintenance and adaptation under physiological and experimental conditions.

Stable isotope labeling with deuterium oxide offers a practical and sensitive approach for measuring DNA synthesis over extended time periods [4, 5]. Deuterium from body water is incorporated into the deoxyribose moiety during nucleotide synthesis, which is then incorporated into mtDNA. mtDNA synthesis rates can then be calculated using the body water or bone marrow deuterium enrichment to define the effective precursor pool for deoxyribose synthesis and mtDNA product enrichment. Under steady-state conditions, mtDNA synthesis approximates mtDNA loss, thus providing the turnover or half-life of mtDNA.

Accurate measurement of mtDNA synthesis requires isolation of mtDNA with minimal nuclear DNA contamination [6, 7]. Skeletal

muscle has particular challenges due to its multinucleated structure, high nuclear content, and low replication rate [8]. The protocol described here adapts an approach using enzymatic digestion of contaminating nuclear DNA using DNase following mitochondrial isolation [9], enabling enrichment of mtDNA suitable for isotope analysis using established protocols for measuring deuterium enrichment in deoxyribose.

---

2 Materials

Prepare all solutions using ultrapure water (18 MΩ·cm at 25 °C) and analytical grade reagents. Store all reagents at 4 °C unless otherwise specified.

2.1 Animals and Housing

1. Rodents of the desired strain, age, and sex.
2. Standard housing with controlled temperature and humidity.

2.2 Stable Isotope Labeling

1. Deuterium oxide–enriched saline: Prepare a solution containing 70%–99.8% deuterium oxide and supplement with sodium chloride to achieve an isotonic concentration of 0.9% (w/v) NaCl.
2. Deuterium-enriched drinking water: 4%–8% deuterium oxide in potable water.
3. 23-G hypodermic needle and syringe.

2.3 Tissue Collection and Storage

1. Isoflurane or approved anesthetic.
2. Ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA).

3. Cryogenic microfuge tubes, liquid nitrogen, dry ice (required if planning to use frozen skeletal muscle).
4. Phosphate-buffered saline (PBS), pH 7.4, sterile filtered.

2.4 Buffers and Solutions for Mitochondrial Isolation and DNase Treatment

1. Homogenization buffer: 100 mM sucrose, 0.1 mM EGTA, 50 mM Tris-HCl, pH 7.4 at 4 °C, 100 mM KCl, 1 mM KH_2PO_4, 0.2% BSA.
2. DNase I digestion buffer: 300 mM sucrose in 1× DNase I buffer, prepared in ultrapure water.
3. DNase I solution: 2000 U/mL stock.
4. RNase A solution: 20 mg/mL stock.
5. Protease digestion reagent: Proteinase, bacterial, Type XXIV, prepared fresh at 0.2 mg/mL final concentration in homogenization buffer.

2.5 mtDNA Purification Reagents

1. Mitochondrial lysis buffer: 0.5% Tween 20, 1 M Tris-HCl pH 8.5 in nuclease-free water with 200 µg/mL proteinase K for 10 h at 55 °C or overnight.
2. Sodium acetate solution: 3M, pH 5.2.
3. Molecular sieve (Type 3A beads, 8–12 mesh)
4. Ethanol: 100% (*see* **Note 1**) and 70% (v/v), prepared with ultrapure water.

2.6 Reagents for mtDNA Purity Assessment by Quantitative PCR

1. Species-specific 5′-nuclease primer and probe sets targeting mtDNA and a single-copy nuclear DNA gene (e.g., *Tfrc* or *Tert*) (*see* **Note 2**).
2. Quantitative PCR master mix (*see* **Note 3**).

2.7 Reagents for DNA Extraction of Bone Marrow for Precursor Enrichment Measure

1. QIAamp DNA Mini Kit (Qiagen cat. no. 51304). We have not validated other DNA isolation kits.
2. Molecular biology grade 100% ethanol.

2.8 Reagents for Isotope Enrichment Analysis

1. Water, molecular biology grade.
2. 0.1 N NaOH, molecular biology grade.
3. Glacial acetic acid.
4. Sodium acetate.
5. Zinc sulfate.
6. Hydrolysis buffer, 5× and 1×: Add 44 mL of ultrapure water to a sterile beaker. Dissolve 1.54 g sodium acetate and 10.75 mg zinc sulfate, then adjust the pH to 5.0 ± 0.5 with glacial acetic acid. Bring the final volume to 50 mL with ultrapure water to generate a 5× hydrolysis buffer. To prepare 1× hydrolysis buffer, dilute 1 mL of the 5× stock with 4 mL of ultrapure water.
   Potato acid phosphatase solution (1 kU): Resuspend the entire 1

7. kU in 1.0 mL of ultrapure water and keep at 4 °C until preparing the hydrolysis cocktail.
8. S1 nuclease solution: Dilute 2.5 µL of S1 nuclease solution into 2.0 mL of 1× hydrolysis buffer. Keep the diluted nuclease at 4 °C until preparing the hydrolysis cocktail.
9. Hydrolysis cocktail: Add 36.8 mL of 5× hydrolysis buffer to a 50 mL conical tube. Add 1.0 mL of the acid phosphatase solution and 1.0 mL of the diluted S1 nuclease solution, then vortex. Aliquot the resulting hydrolysis cocktail into microcentrifuge tubes (~520 µL per tube, sufficient for ten reactions) and store at −20 °C for up to 12 months.
10. O-(2,3,4,5,6-Pentafluorobenzyl) hydroxylamine hydrochloride (1 mg/mL aqueous solution) prepared fresh every week and stored at 4 °C.
11. Acetic anhydride.
12. N-methylimidazole.
13. Sodium sulfate, granular, anhydrous.
14. Methylene chloride.
15. Ethyl acetate.

2.9 Specialized Equipment

1. Glass on glass homogenizer, A “Loose” and B “Tight” (Ace 1 mL Glass Dounce Tissue Homogenizer 8343-01).
2. Refrigerated centrifuge capable of at least 10,000 × *g*.
3. Microcentrifuge.

4. Shaking platform or rotator.
5. 37 °C incubator or water bath.
6. Quantitative PCR instrument.
7. Gas chromatography–mass spectrometry system.

---

3 Methods

3.1 Stable Isotope Labeling with Deuterium Oxide

1. Calculate the volume of deuterium oxide required to achieve the desired initial body water enrichment based on animal body weight (*see* **Note 4**).
2. Prepare deuterium oxide–enriched saline by adding sodium chloride into 99.8% deuterium oxide to a concentration of 0.9% sodium chloride. Sterile filter and store at room temperature.
3. Administer warmed (30 °C) deuterium oxide bolus by intraperitoneal injection with a 23-G needle and syringe to rapidly establish body water enrichment (*see* **Note 5**).
4. Upon return to the home cage, replace normal drinking water with deuterium-enriched drinking water (4%–8% deuterium oxide).
5. Maintain animals on enriched drinking water *ad libitum* for the duration of the labeling period.
6. Continue labeling depending on expected mtDNA turnover (*see* **Notes 6 and 17**).
7. Monitor animals daily for general health, body weight, and water intake.

3.2 Tissue Harvest and Processing

1. Anesthetize the animal using an approved anesthetic protocol.
2. Collect whole blood into tubes containing EDTA (final concentration 1.8 mg/mL) as an anticoagulant. Centrifuge at 1500 × *g* for 10 min at 4 °C to separate plasma. Carefully transfer the plasma to labeled tubes and store at −80 °C until analysis. Plasma samples are used to estimate precursor enrichment.
3. From dissected muscles, remove visible connective tissue and adipose tissue.
4. Rinse tissue briefly in ice-cold saline to remove residual blood.
5. Blot tissue dry and place on a chilled surface.
6. Weigh the whole muscle.
7. Snap freeze muscle and powder under liquid nitrogen.
8. Aliquot 50–100 mg of frozen, powdered skeletal muscle for mitochondrial isolation.
9. Excise both rodent femurs and remove both ends to access bone marrow.
10. If using bone marrow to measure precursor enrichment (*see* **Note 14**), flush bone marrow from each femur with 0.5 mL of PBS and pool into a single microcentrifuge tube. Freeze bone marrow at −80 °C. The bone marrow is used to measure precursor enrichment.

3.3 Isolation of Mitochondria from Skeletal Muscle with DNase Treatment

1. Transfer 50–100 mg of frozen, powdered skeletal muscle into a

pre-chilled homogenization vessel.

2. Add 1 mL of ice-cold mitochondrial isolation buffer containing proteinase.
3. Incubate on ice for 2 min.
4. Homogenize on ice with a glass-on-glass homogenizer. Three passes with homogenizer A ("Loose") and three passes with homogenizer B ("Tight") (*see* **Note 7**).
5. Transfer homogenate to a 2-mL microfuge tube.
6. Rinse the glass homogenizer with mitochondrial isolation buffer without proteinase and transfer the rinse to the microfuge tube to bring the total volume to 2 mL.
7. Centrifuge at 700 × *g* for 10 min at 4 °C.
8. Transfer supernatant to a clean 2 mL Eppendorf tube and discard the pellet.
9. Again, centrifuge at 700 × *g* for 10 min at 4 °C.
10. Transfer supernatant to a clean 2 mL Eppendorf tube and discard the pellet.
11. Centrifuge at 10,000 × *g* for 10 min at 4 °C.
12. Discard supernatant.
13. Resuspend mitochondrial pellet in 1 mL DNase digestion buffer (mitochondrial isolation buffer supplemented with 10% (v/v) 10× DNase I buffer) with 15 µL of DNase I and 10 µL of RNase A (*see* **Note 8**).

14. Incubate at 37 °C for 120 min.
15. Pellet mitochondria at 10,000 × *g* for 10 min.
16. Discard supernatant.
17. Resuspend the mitochondrial pellet in 1 mL of DNase I digestion buffer without DNase I or RNase A.
18. Pellet mitochondria at 10,000 × *g* for 10 min.
19. Discard supernatant.
20. Resuspend mitochondrial pellet in 1 mL of DNase I digestion buffer without DNase I or RNase A.
21. Pellet mitochondria at 10,000 × *g* for 10 min.
22. Discard supernatant.
23. Resulting mitochondrial pellet can be used for DNA isolation or frozen for later use.

3.4 Purification of mtDNA

1. In a microfuge tube, lyse mitochondria in 50 µL of mitochondrial lysis buffer with proteinase K.
2. Inactivate proteinase K at 95 °C for 10 min and then spin at maximum speed (~16,000 × *g*) for 5 min.
3. Precipitate DNA with 0.5 volume (25 µL) of 7.5 M ammonium acetate (final concentration of 3.75 M ammonium acetate) and 2.0 volumes (100 µL) of 100% ethanol.
4. Incubate for 15 min at room temperature.

5. Centrifuge at maximum speed (about 16,000 × *g*) for 15 min at 4 °C.
6. Carefully remove supernatant (mtDNA pellet may not be visible; *see* **Note 9**).
7. Add 200 μL of 70% ethanol to wash the pellet.
8. Centrifuge at maximum speed for 3 min.
9. Remove 70% ethanol carefully (*see* **Note 9**).
10. Air-dry the pellet for 10 min.
11. Resuspend the mtDNA in 100 μL of nuclease-free water heated to 50 °C, vortex thoroughly, spin down, and incubate overnight at 55 °C.
12. Store isolated mtDNA at −80 °C for later analysis (Step 3.7).

3.5 Assessment of mtDNA Purity

1. Perform quantitative PCR using mitochondrial and nuclear DNA targets to assess nuclear DNA contamination. Expected results show Ct values for mtDNA ranging from 15 to 21, with nuclear DNA undetectable after 40 amplification cycles (*see* **Note 10**).

3.6 Precursor Enrichment Analysis: Bone Marrow DNA Extraction

1. Separate ~500 mg of bone marrow suspension and place in a 1.5-mL microcentrifuge tube and allow to thaw on ice.

2. From the QIAamp DNA Mini kit, add 180 µl ATL buffer and 20 µL Proteinase K to bone marrow.
3. Incubate the samples on a heat block at 56 °C and shake at 400 rpm overnight.
4. Quick-centrifuge samples up to 1000 × *g* (*see* **Note 11**).
5. Vortex AL buffer and add 200 µL AL buffer to samples and pulse vortex for 30 s (2 s on, 2 s off for 30 s).
6. Incubate at 70 °C for 15 min.
7. Quick-centrifuge samples up to 1000 × *g* (*see* **Note 11**).
8. Add 200 µL ethanol to samples and pulse vortex for 15 s.
9. Quick-centrifuge samples up to 1000 × (*see* **Note 11**).
10. Prepare the Qiagen Mini Kit and label the top of the centrifuge spin column since collection tubes will be discarded.
11. Add the sample to the centrifuge spin column (*see* **Note 12)**.
12. Centrifuge sample at 6000 × *g* for 1 min.
13. Place each spin column into a new collection tube, discarding the collection tube and eluant from the previous spin.
14. Add 500 µL AW1 buffer to samples.
15. Centrifuge sample at 6000 × *g* for 1 min at room temperature.
16. Place each spin column into a new collection tube, discarding the collection tube and eluant from the previous spin.
17. Add 500 µL AW2 buffer to samples.

18. Centrifuge sample at 6000 × *g* for 1 min at room temperature.
19. Place each spin column into a new collection tube, discarding the collection tube and eluant from the previous spin.
20. Repeat Steps 17–19.
21. Centrifuge the samples at 16,000 × *g* for 5 min at room temperature.
22. Place each spin column into a PCR clean labeled microcentrifuge tube, discarding the collection tube and eluant from the previous spin.
23. Add 200 µL molecular biology water to samples (*see* **Note 13**).
24. Incubate samples for 15 min at room temperature.
25. Centrifuge sample at 6000 × *g* for 1 min at room temperature to elute DNA.
26. Proceed to DNA derivatization (Sect. 3.7). Samples can be stored in −20 °C until DNA derivatization.

See **Note 14** for perspective regarding the use of deuterium enrichment of blood plasma instead of bone marrow.

3.7 DNA Derivatization for GC–MS Analysis

1. Incubate each extracted DNA sample with 50 µL of hydrolysis cocktail prepared in 2.7.8 at 37 °C overnight. Samples can be stored at −20 °C until derivatization.
2. Prepare a set of 2-deoxy-D-ribose standards in the range of

sample abundance on GC/MS to determine naturally occurring isotope to be subtracted during calculations (Step 3.7.3). Transfer homogenized mtDNA and prepared standards into 13 × 100 glass tubes.

3. Add 80 µL glacial acetic acid and 100 µL 1 mg/mL pentafluorobenzyl hydroxylamine to all samples and standards.
4. Cap and incubate at 100 °C for 30 min then let cool to room temperature.
5. Add 1 mL of acetic anhydride to all.
6. Carefully add 100 µL n-methylimidazole to each sample and standards. The reaction is very exothermic and may splash. Perform this step in a fume hood (*see* **Note 15)**.
7. Cap glass tubes, allow reaction to proceed, wait until samples and standards have cooled to room temperature, about 15–20 min.
8. Once the reaction has completed, add 2 mL of molecular biology grade water and 800 µL of methylene chloride to all. Cap and vortex.
9. If needed, centrifuge samples briefly for clear separation of phases.
10. Transfer the bottom (organic) phase of the sample using a disposable Pasteur pipette to a new set of glass tubes with sodium sulfate added to cover the bottom of each tube. Avoid pulling the top (aqueous) phase of the sample (*see* **Note 16**).
11. Add an additional 800 µL of methylene chloride to all. Repeat the extraction Steps 9 and 10, adding to the same sodium sulfate tube as before.
12. Once the samples have been extracted, transfer the sample from the glass tubes with sodium sulfate at the bottom to GC vials,

ensuring not to transfer sodium sulfate crystals.

13. Dry the GC vials in a SpeedVac at room temperature until completely dry.
14. Add 200 µL ethyl acetate to each GC vial, lightly cap, and vortex. Transfer ethyl acetate to GC vial insert and return the insert to the GC vial. Cap and crimp the vials.
15. For GC/MS helium is used as the carrier gas and methane as the reactant gas.
16. 1 µL of sample is injected into an inlet set to 280 °C
17. An Agilent DB-17MS (30 m, 0.25 mm, 0.25 µm) column (or similar) is used with the oven initially set to 140 °C then ramped at 40 °C/min to a final temperature of 280 °C.
18. The MS is operated in NCI SIM monitoring m/z 435 and 436 for the M0 and M1 isotopomers, respectively.
19. Two peaks corresponding to cis/trans isomers will be observed. The larger, second peak is used for quantification.

3.8 Calculation of mtDNA Fractional Synthesis Rate and Half-Life

1. Analyze mtDNA and bone marrow derivatives using GC–MS on a DB-17MS (30 m, 0.25 mm, 0.25 µm) column (or similar) with negative chemical ionization, using helium as a carrier and methane as the reagent gas.
2. Quantify the fractional molar isotope abundances at m/z 435 (M0, the mass isotopomer) and 436 (M1) of the pentafluorobenzyl triacetyl derivative of purine deoxyribose.
3. Enrichment (EM1) is calculated by the excess abundance of m+1

relative to the unenriched pentafluorobenzyl triacetyl purine dR derivative standards:

$$\mathrm{EM\,1} = \frac{(\text{abundance m/z} \quad 436)_{\text{sample}}}{(\text{abundance m/z} \quad 435, \quad 436)_{\text{sample}}} - \frac{(\text{abundance m/z} \quad 436)_{\text{standard}}}{(\text{abundance m/z} \quad 435, \quad 436)_{\text{standard}}}$$

4. Fraction new is calculated as the ratio of mtDNA enrichment to precursor enrichment:

$$\text{Fraction new} = \frac{\mathrm{EM\,1}_{(\text{product})}}{\mathrm{EM\,1}_{(\text{precursor})}}$$

5. To determine the fraction of new mtDNA, measure precursor enrichment by evaluating deuterium enrichment of plasma or bone marrow DNA, which involves extracting DNA from bone marrow and following Sect. 3.6. For measurement of plasma enrichment, refer to PMID:32213116 section “Typical analysis protocols Water analysis".
6. Once precursor enrichment is measured, divide EM1 of the sample by the EM1 of the precursor to calculate the fraction of mtDNA that is new (i.e., “fraction new”).
7. Calculate fractional synthesis rate by dividing fraction new by the labeling duration, then multiplying by 100 to express the result as a percentage.
8. While not recommended (*see* **Note 17**), to estimate half-life using a single time point, calculate half-life by first calculating the rate parameter *k* as follows:

$$k = -\ln\left(\frac{1-\text{fraction new}}{\text{label duration}}\right)$$

We assume mtDNA turnover is a first-order reaction where turnover is in a steady state. Thus, to convert *k* to half-life, perform the following calculation:

$$t_{\frac{1}{2}} = \frac{0.693}{k}$$

A single time point assumes mtDNA turnover is linear, uniform, and constant. In addition, there is an assumption that mtDNA will

eventually reach 100% new. We and others have demonstrated that not all mtDNA turns over (*see* **Note 17**) [6, 7]. Thus, we strongly recommend using a time-course approach to evaluate half-life.

9. Using statistical software (e.g., GraphPad Prism), create an *xy* data table where *x* is label duration and *y* is fraction new. Fit a non-linear, one-phase association curve with the assumption that Y = 0 and the plateau must be equal to or less than 1.0 (which reflects 100% enrichment). The resultant curve will provide a rate constant, *k*, and corresponding half-life. In addition, there will be an estimated plateau, which reflects the estimated portion of mtDNA that turns over.

---

4 Notes

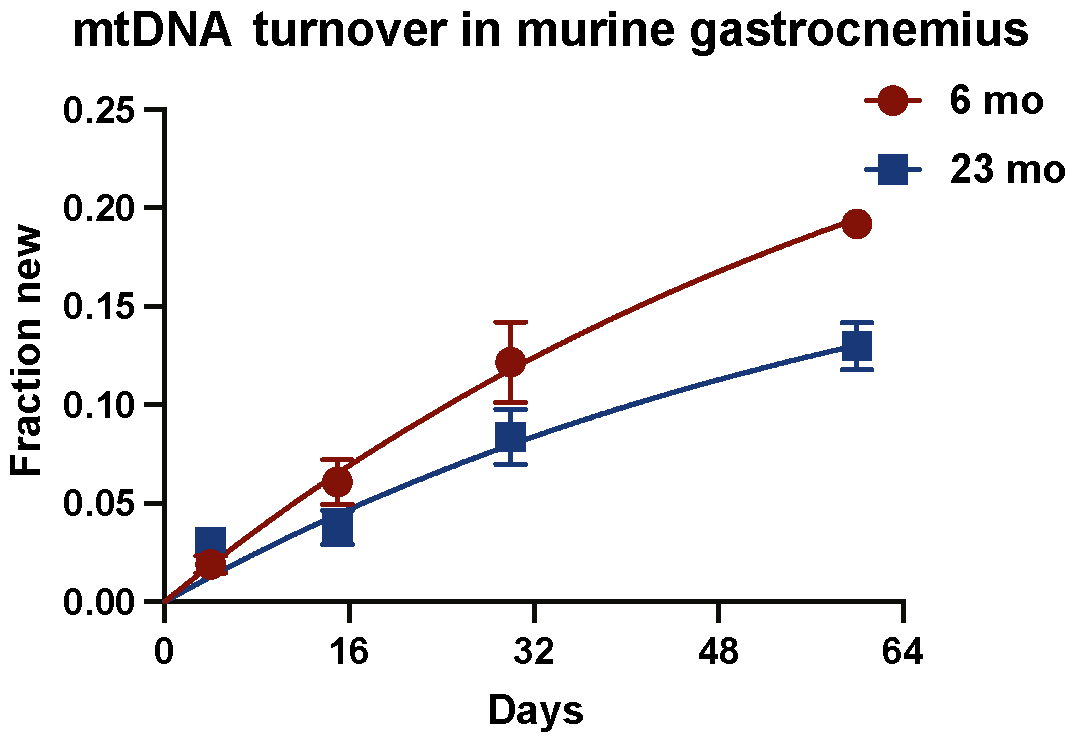
1. 100% ethanol is hygroscopic and readily absorbs water; therefore, molecular sieves are added at 10% (w/w) to maintain anhydrous conditions.
2. Using a single-copy nuclear gene target is essential for accurate quantification of mtDNA relative to nuclear DNA.
3. Consider using a qPCR master mix that includes both uracil and uracil–DNA glycosylase (UNG) to prevent carryover contamination from prior amplifications.
4. Target enrichment is typically 3%–5%.
5. Use a priming dose to achieve steady state.
6. Labeling duration affects sensitivity.
7. The extent of homogenization and number of passes with the homogenizer have been optimized for mouse and rat skeletal

muscle. For other muscle or tissue types, these parameters should be determined empirically based on the level of nuclear DNA contamination. Increasing the number of passes may increase mtDNA yield while also increasing nuclear DNA contamination.

8. DNase digestion is essential for removing nuclear DNA located outside intact mitochondria, while mtDNA remains protected from digestion within the mitochondrial membranes.
9. Position microcentrifuge tubes with the hinge facing outward in the rotor to indicate the location of the DNA pellet. When removing the supernatant, pipette from the side opposite the hinge to avoid disturbing the pellet.
10. Contamination of skeletal muscle nuclear DNA will lead to underestimation of deuterium enrichment of mtDNA [7].
11. Ensure there is no liquid buildup in the cap on the microcentrifuge tube and that no pellet has formed. If a pellet forms, resuspend pellet by vortexing and then centrifuge the samples again.
12. Avoid wetting the rim when adding buffer or sample to the spin column.
13. Do not use the elution buffer provided with the kit as it contains EDTA and can interfere with subsequent steps.
14. Accurate precursor enrichment is critical to accurate evaluation of product enrichment. While the relatively rapid turnover of bone marrow can accurately reflect precursor enrichment, the equilibration of deuterium enrichment in bone marrow DNA takes a few days. Thus, it is recommended that precursor enrichment be estimated from deuterium concentration in blood plasma when the label duration is less than a week [10].
15. N-methylimidazole can deteriorate if not properly stored; top off

with nitrogen gas to displace air and store at 4 °C.

16. Wet the tip of the Pasteur pipette with methylene chloride prior to using it to transfer the bottom (organic) layer; otherwise, the organic layer will elute out of the pipette prior to transfer.
17. A single time point assumes mtDNA turnover is linear, uniform, and constant. In addition, there is an assumption that all mtDNA turnover. Our data (Fig. 1) and others have demonstrated this to not be the case [6, 7]. Thus, we strongly recommend using a time-course approach from four to at least 60 days.



	6 mo	23 mo
Plateau	0.3361	0.2175
k	0.0144	0.01525
Half-life (days)	48.12	45.47

Fig. 1 Measuring deuterium enrichment of gastrocnemius mtDNA over time. These data

indicate the fraction of new mtDNA (i.e., mtDNA enriched with deuterium) from a time-course labeling experiment, in which cohorts of mice in the same cohort/treatment group are administered deuterium oxide over varying durations. In this example, 6-month-old and 23-month-old male C57Bl/6 mice were intraperitoneally injected with sterilized 0.9% sodium chloride mixed in 99.8% deuterium oxide equivalent to 5% of their estimated body water volume (70% of body weight) and then provided drinking water enriched with 5% deuterium oxide for 4, 15, 30, and 60 days. The initiation of labeling was adjusted to ensure mice were the same age upon euthanasia. These data demonstrate that mtDNA does not turnover at a constant rate, suggesting that there are subpopulations of mtDNA that turnover at different rates. One-phase, non-linear regressions ($r2 > 0.80$) indicate the rate constant (k) does not change with age in the gastrocnemius. Notably, the time-course design using a stable isotope reveals a slower half-life (~45 days) than previously assessed using pulse-chase experimental designs (~10 days). This is likely due to the nature of pulse-chase experimental designs, which bias towards faster turning over molecules. In addition, the non-linear regression estimated portion of mtDNA that actively turns over (plateau) declines with age, from ~33.61% in 6-month-old mice to ~21.75% in 23-month-old mice.

Funding

Dr. Musci was supported in this work by the Oklahoma Nathan Shock Pilot Grant, National Institutes of Health P30AG050911, Southwest ACSM Early Investigator Award, the Continuing Faculty, SOAR, and SURP Grants of the Frank R. Seaver College at Loyola Marymount University. The National Institutes of Health P30 AG052363 and the Department of Veterans Affairs I01 BX005592 supported Dr. Miller in this work. Dr. Wanagat is supported by funding from the National Institutes of Health R01AG055518, R01AG069924, R01AR087145, and K02AG059847 and from the Department of Veterans Affairs I01RX004521 and I01CX002838. This manuscript is the result of work supported with resources and the use of facilities at the Veterans Administration Greater Los Angeles Healthcare System.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Youle RJ, van der Bliek AM (2012) Mitochondrial fission, fusion, and stress. Science 337(6098):1062–1065. https://doi.org/10.1126/science.1219855 [PubMed][PubMedCentral]

2. Falkenberg M (2018) Mitochondrial DNA replication in mammalian cells: overview of the pathway. Essays Biochem 62(3):287–296. https://doi.org/10.1042/EBC20170100 [PubMed][PubMedCentral]

3. Lei Y, Gan M, Qiu Y et al (2024) The role of mitochondrial dynamics and mitophagy in skeletal muscle atrophy: from molecular mechanisms to therapeutic insights. Cell Mol Biol Lett 29(1):59. https://doi.org/10.1186/s11658-024-00572-y [PubMed][PubMedCentral]

4. Hellerstein MK, Neese RA (1992) Mass isotopomer distribution analysis: a technique for measuring biosynthesis and turnover of polymers. Am J Physiol 263(5 Pt 1):E988–E1001. https://doi.org/10.1152/ajpendo.1992.263.5.E988 [PubMed]

5. Busch R, Neese RA, Awada M et al (2007) Measurement of cell proliferation by heavy water labeling. Nat Protoc 2(12):3045–3057. https://doi.org/10.1038/nprot.2007.420 [PubMed]

6. Collins ML, Eng S, Hoh R et al (2003) Measurement of mitochondrial DNA synthesis in vivo using a stable isotope-mass spectrometric technique. J Appl Physiol (1985) 94(6):2203–2211. https://doi.org/10.1152/japplphysiol.00691.2002 [PubMed]

7. Musci RV, Fuqua JD, Peelor FF III et al (2025) Age-induced changes in skeletal muscle mitochondrial DNA synthesis, quantity, and quality in genetically unique rats. GeroScience 47(1):851–862. https://doi.org/10.1007/s11357-024-01344-4 [PubMed]

8. Borowik AK, Davidyan A, Peelor FF,3rd et al (2023) Skeletal muscle nuclei in mice are not post-mitotic. Function (Oxf) 4(1):zqac059. https://doi.org/10.1093/function/zqac059 [PubMed]

9. Zimmerman EG, Akins DR, Planz JV et al (1988) A rapid procedure for isolating mitochondrial DNA. Gene Anal Tech 5(5):102–104. https://doi.org/10.1016/0735-0651(88)90004-0 [PubMed]

10. Miller BF, Reid JJ, Price JC et al (2020) CORP: the use of deuterated water for the measurement of protein synthesis. J Appl Physiol (1985) 128(5):1163–1176. https://doi.org/10.1152/japplphysiol.00855.2019 [PubMed]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_19

19. Mnemonic Similarity Task Variants for Probing Age-Related Memory Impairments Across Species and Modalities

Lorena A. Ferguson[1], Chih-Yuan Jessie Chien[1], Stephanie L. Leal[2], Craig Stark[3], Sara N. Burke[4, 5] and Teal Eich[1] ✉

(1) Davis School of Gerontology, University of Southern California, Los Angeles, CA 90089, USA
(2) Integrative Biology and Physiology, University of California, Los Angeles, CA 90095, USA
(3) Department of Neurobiology and Behavior, University of California, Irvine, CA 92697, USA
(4) Department of Neuroscience, University of Arizona, Tucson, AZ 85721, USA
(5) BIO5 Institute, University of Arizona, Tucson, AZ 85719, USA

✉ **Teal Eich**
Email: teich@usc.edu

Abstract

The Mnemonic Similarity Task (MST) is a well-validated behavioral memory test for use in both humans and animal models that queries the ability to distinguish between highly similar but nonidentical stimuli ("lures") and repeated stimuli ("targets"), a process that is thought to rely upon a computational process known as pattern

separation, which is supported by the hippocampus. Both increasing age as well as pathological and neurological hallmarks of Alzheimer's disease are associated with deficits in this task. Here, we describe the methods for implementing four versions of the MST: a behavioral version for use in humans, a version that can be implemented using functional magnetic resonance imaging, an emotional version of the task, and finally a version for use in rodents.

Keywords Memory – Mnemonic similarity task – Mnemonic discrimination – Behavior – Neuroimaging – Pattern separation – Alzheimer's disease – Aging – Humans – Hippocampus – Emotion – Animal models

---

1 Introduction

Aging is accompanied by heterogeneous changes to episodic memory—the ability to recall past events along with contextual details, including spatial, temporal, emotional, and perceptual aspects [1]. For example, older adults may show perfectly normal forced-choice recognition memory for words (where one word was previously presented and one is novel), but be impaired when asked to identify which person read them a particular word [2]. Likewise, when shown pictures of objects, older adults often rely more heavily on gist rather than true episodic retrieval when making memory judgments, being prone to "false alarm" and endorsing items from the same category as a studied item [3–5]. While some of these changes reflect "normal" aging, others may be early indicators of neurodegenerative disease, including Alzheimer's disease (AD), a leading cause of morbidity and mortality that is expected to affect over 150 million people by 2050 [6]. Disentangling these trajectories is critical for identifying individuals at risk, characterizing underlying neuropathological processes, and developing effective interventions. The Mnemonic Similarity Task (MST) is a well-validated episodic memory test for use in humans [7], and additionally has been used to probe episodic-like memory using animal models [8–

12]. Unlike many standard memory tasks that measure accuracy of veridical recall or recognition of previously presented information, the MST queries the ability to distinguish similar but non-identical stimuli ("lures") from repeated stimuli ("targets"). This lure discrimination ability is attributed to the neurocomputational process known as pattern separation [7, 13], which allows for interference across overlapping memories to be reduced by making similar inputs more distinct, thereby preserving contextual detail [13, 14]. Pattern completion, in contrast, recovers stored representations from partial cues, allowing a degraded input to be "filled in" based on prior experience. Although pattern separation and pattern completion interact, they rely on dissociable subfields within the hippocampus, with separation linked to the dentate gyrus (DG) and CA3 circuitry, and completion to CA3 recurrent collaterals with output to CA1 [13, 15, 16].

There is a robust literature linking increasing age with deficits in pattern separation as-indexed by the MST [5, 17–19]. Older adults show increased false alarms to similar "lure" stimuli, dysfunctional DG/CA3 hyperexcitability [20–24], and a shift toward completion-like responding, with general target recognition (memory for identical stimuli) remaining relatively intact [5, 7]. The MST reliably differentiates cognitively healthy older adults from individuals with mild cognitive impairment (MCI) [5, 7, 25–27]—a syndrome often considered a prodromal stage of AD [28, 29]—and it is sensitive to amyloid-β (Aβ) and phosphorylated tau (p-tau), with higher levels of these key biological markers of AD associated with impaired lure discrimination, even in those without clinical diagnoses [21, 30–33]. Further, the inclusion of the MST in computational models increases the predictive probability of future AD risk and pathological burden: Vanderlip and colleagues [27] reported that the combination of behavioral performance on the MST and serum-based p-tau217 quantification improved diagnostic classification and prediction of cognitive decline relative to predictions based on p-tau217 values alone [27], and Kim et al. found that performance on the MST could predict not only cognitive decline, but also Aβ burden, outperforming

traditional neuropsychological tests [26]. The MST thus offers a scalable, easily implementable, mechanistically grounded behavioral measure suitable for both human and animal model research that links biological staging to cognitive consequences, supporting its integration into preclinical screening pipelines, trial enrichment strategies, and early risk modeling for AD. Numerous versions of the task have been used. Stimuli can include object [7], temporal [34, 35], spatial [36–38], and emotional domains [30, 39], and response options can vary from old/new to old/similar/new [5]. Some tasks utilize distinct study phases [7, 30, 36–39] while others use a continuous design [5, 34, 35, 40]. However, regardless of task version, findings broadly converge on evidence of age-related deficits in mnemonic discrimination [5].

Here, we present the methods for four different variations of the MST: a behavioral version, a version that can be used to assess brain activity using functional neuroimaging while participants complete the task, an emotional version, and finally, a version that can be used in animal models.

1.1 Behavioral Studies in Human Aging and Alzheimer's Disease

Lure discrimination declines significantly with age, whereas target recognition (memory for identical stimuli) remains relatively intact [5, 7]. However, not all types of lure stimuli are impacted in the same way. While discrimination of both object and scene stimuli is impaired with age, there are greater deficits for object stimuli [37, 41–43]. Additionally, higher levels of Aβ and tau are associated with lure discrimination impairments, even in cognitively healthy older adults [21, 30–33, 44]. Some studies found greater deficits in object discrimination [21, 32], while others found effects in spatial discrimination [44]. This discrepancy may be due to different pathologies. Tau may impair object discrimination, whereas Aβ may be more detrimental to spatial discrimination [31]. Thus, impaired performance on the behavioral MST may offer insights into early age-

and disease-related cognitive decline, as well as the neurobiological processes underlying this decline.

1.2 Functional Neuroimaging Studies in Human Aging and Alzheimer's Disease

The hippocampus is a key region involved in episodic memory [45, 46]. Two hippocampal computations support episodic memory: pattern completion and pattern separation. Pattern completion is the process by which a partial cue is completed by reactivating a previous representation [13]. In contrast, pattern separation refers to the process of orthogonalizing similar but non-identical inputs, allowing them to be stored as distinct representations [13].

Pattern separation relies primarily on the dentate gyrus (DG) subfield, although it can recruit the CA3 as well if the stimuli have a significant amount of overlap [13, 16]. The DG receives input from the cortex via the perforant path from the entorhinal cortex (EC) [13, 16, 47]. Small granule cells in the DG then project to the CA3 via the mossy fiber pathway, along which detonator cells can depolarize CA3 pyramidal cells [13, 16]. This reduces interference between stimuli by forcing distinct representations [16, 48]. This EC-DG-CA3 pathway is thought to be involved in encoding [13]. Interestingly, the DG is one of only two brain regions that undergoes adult neurogenesis. These newborn DG granule cells may be essential for pattern separation due to their high levels of synaptic plasticity and sparse firing rates [49–52]. While hippocampal pattern separation cannot be directly measured in vivo in humans, mnemonic discrimination tasks can serve as behavioral correlates [7, 13].

Pattern completion relies on the CA3 subfield of the hippocampus [13]. In addition to input from the DG, the CA3 also receives input from the EC directly through the perforant path, facilitating recall [13, 47]. CA3 pyramidal cells project onto themselves through recurrent collaterals, forming an auto-associative network [13]. This allows for associations to be formed between different stimuli [47].

CA3 pyramidal cells also project to the CA1 through Schaffer collaterals. The CA1 in turn projects to the subiculum, and both CA1 and the subiculum have projections back to the EC [13]. This pathway enables output from the hippocampus to reach the rest of the cortex [15]. The whole hippocampal circuit is heavily modulated by neurotransmitters, including norepinephrine, acetylcholine, and dopamine [53–55].

Neuroimaging studies have found that mnemonic discrimination tasks, including the MST, are sensitive to the neurobiological dynamics supporting pattern separation. Alterations in DG/CA3 activity, volumetric changes in the hippocampus, and AD pathological burden have all been associated with mnemonic discrimination performance [5, 19, 21, 33, 44]. Thus, mnemonic discrimination tasks are able to track not only memory alterations, but also underlying neurobiological function. This positions them as powerful tools for researchers and clinicians alike.

With age, there are significant neurobiological changes that are associated with lure discrimination impairments. These deficits are driven by a shift from pattern separation to pattern completion due to hippocampal network dysfunction. There is CA3 hyperexcitability with age, potentially due to reduced inhibition from GABAergic interneurons [56–60] and a reduction in perforant path input from the EC [13, 61, 62]. This strengthens the recurrent collateral network at the expense of the DG [19, 60]. CA3 hyperactivity has been associated with impairments in lure discrimination performance [20, 63, 64].

Importantly, there is not always a clear-cut distinction between healthy aging and MCI/AD. Aβ and tau begin to aggregate years before clinical diagnosis, potentially impacting cognition and neurobiological function long before disease onset [65–67]. Aβ and tau also have different progression patterns and target different brain regions, thereby impacting different aspects of cognition at different points in the pathological timeline. Tau pathology begins to aggregate in the locus coeruleus (LC) during early adulthood and midlife, before clinical diagnosis [65, 68, 69], then spreads to the EC early in disease

progression [70–73]. Neurons in layer II of the lateral EC (LEC), which project to the DG and CA3, are the initial sites of tau pathology within the cortex [74]. From there, tau progressively spreads to other limbic areas and then throughout the cortex [74]. The LEC selectively processes object information, whereas the medial EC (MEC) processes spatial information [75, 76]. Therefore, early tau accumulation in the LEC may selectively disrupt memory for objects. This is supported by findings that tau pathology is associated with impaired object discrimination [21, 31].

Aβ begins to accumulate in the frontal lobes, before spreading to the rest of the cortex, including the hippocampus [77]. Aβ may promote the development of tau [78], including in the EC [79–81], and interact with tau to influence cognition [82–85]. This Aβ and tau interaction in the EC may have impacts on the LEC. A recent study found that dysfunctional hyperconnectivity between the LEC and DG/CA3 was associated with both worse performance on an object discrimination task as well as higher Aβ levels [79]. Spread of Aβ into the MEC later in pathological progression may, in turn, impair spatial discrimination [31, 44].

Pattern separation impairments in MCI and AD may also be related to alterations in DG neurogenesis. While neurogenesis continues throughout healthy aging [86–88], it declines significantly in AD [87]. Given the importance of newborn granule cells in supporting pattern separation [49–52], the dearth of neurogenesis may contribute to mnemonic discrimination impairments in disease states [13, 89, 90]. Some research in animal models has supported this idea [91, 92]. However, studies in humans remain virtually non-existent. Thus, there is a gap in the literature that remains to be explored.

1.3 Emotional MST in Human Aging and Alzheimer's Disease

The emotional variant of the MST has revealed a gist versus detail tradeoff [93–95] in which young adults tend to remember emotional target stimuli (i.e., the emotional gist), but demonstrate greater forgetting of emotional lure stimuli (i.e., the emotional details) relative

to neutral lures. Older adults show overall worse memory performance than do young adults, but exhibit preserved memory for both emotional target and lure stimuli [96]. Thus, there may be maladaptive retention of emotional details in aging, as it may be more beneficial to retain the emotional gist at the expense of emotional details (e.g., in line with young adult performance) [96]. The emotional variant of the task is especially useful in contexts where individuals may exhibit emotional memory biases, such as in depression. Depressed individuals show evidence of a negativity bias in memory, where negative stimuli are remembered better than positive or neutral stimuli [97, 98]. Mnemonic discrimination tasks have found that depression is associated with impairments in the ability to discriminate neutral lure stimuli [99–102]. Discrimination of negative lures, however, is often enhanced with greater depressive symptoms [39, 101, 103, 104]. Importantly, depression is both a risk factor for AD and an early indicator of AD pathology [105–108]. Emotional mnemonic discrimination tasks can therefore be used to detect alterations in memory and mood that could indicate AD risk. The emotional version of the MST has also provided novel insight into how varying types of interventions may alter emotional memory biases, such as emotion regulation [103] and antidepressant use [104]. These may provide promising interventions that could alleviate depressive symptoms in those at risk of developing or in the early stages of AD.

Emotionally arousing events trigger the release of stress hormones, including epinephrine and cortisol [109–111]. The locus coeruleus (LC) is a small brainstem region that is the primary source of norepinephrine (NE) to the brain and plays an important role in arousal [111]. The DG has a high number of NE receptors, and there is some evidence that the LC-NE system plays an important role in hippocampal pattern separation [112–114]. With age, however, there is dysfunction in the LC-NE system [113, 115, 116]. This dysfunction may in turn be associated with declines in lure discrimination ability [113].

In young adults, inducing cortisol release can improve memory for negative lure stimuli at the expense of neutral stimuli [117, 118].

Similarly, higher NE release during the study phase of an emotional MST has been associated with better lure discrimination, but not target recognition [114]. However, with age, there is a rise in NE and cortisol, which are associated with memory impairments, hippocampal dysfunction, and greater AD pathological burden [119–121]. In older adults, higher levels of NE metabolism have been associated with greater levels of p-tau, inflammation, neuropsychiatric symptoms, and overall memory impairments [122]. Thus, NE and cortisol may have differing effects on hippocampal pattern separation in younger and older adults.

Functional neuroimaging studies utilizing an emotional version of the MST in young adults have found greater activation of the DG/CA3 during accurate lure discrimination of negative relative to neutral stimuli, and greater BLA activation for negative relative to neutral stimuli regardless of memory performance [123]. In contrast, older adults with subclinical memory impairment show dysfunctional hippocampal DG/CA3 hyperactivity during false alarms and reduced amygdala–hippocampal connectivity during accurate negative discrimination relative to older adults without memory impairment [124]. This suggests there may be valence-specific dysfunction in the hippocampal circuit in aging. However, more work is needed to better understand the role of emotion in hippocampal pattern separation.

1.4 Animal Models

Given that we are not able to directly assess pattern separation in humans, much of our understanding of the neurobiological underpinnings of pattern separation comes from animal models [13, 18, 47, 125]. Cross-species evidence—i.e., from mice, rats, non-human primates, and humans—supports the robust link between increasing age and pattern separation decline [13, 17, 18, 126]. Animals with age-related neurobiological changes, including impairments in DG plasticity [127–129], declining neurogenesis [50, 51, 130], and CA3 hyperactivity [131–134], have pattern separation deficits. Thus, animal models mirror findings in humans, and vice versa.

There is a significant body of literature examining spatial pattern separation in rodents [17, 125, 135–138], although there are differences in the ways pattern separation is assessed. Spatial discrimination can be assessed through behavioral outcomes such as exploration time when an object is moved to a different location [137, 138], number of errors made when exploring a radial arm maze [130, 139], or errors made when discriminating between two spatial locations in which difficulty is parametrically manipulated by varying the distance between the target and the foil [8, 140]. Pattern separation functions in animal models have also been probed behaviorally using objects as stimuli [10, 126, 131, 141]. Aged animals are found to engage in less exploration of novel objects [141] and struggle to discriminate between similar objects [10, 126]. Similar to observations with human participants [37], age-related discrimination deficits are more robust when objects rather than spatial locations are used as stimuli [10, 140]. These studies provide complementary evidence of the deleterious effects of aging on pattern separation abilities found in humans.

---

2 Materials

2.1 Behavioral MST

1. Desktop or laptop computer
2. Keyboard or mouse for response collection
3. Desk and chair
4. Quiet testing room

2.2 fMRI MST

1. 3 T or 7 T MRI scanner
2. Desktop or laptop running the fMRI version of the task. Depending on the MRI system and acquisition parameters, researchers may

need to modify the PsychoPy task script to ensure proper synchronization with fMRI data acquisition.

3. An appropriate cord (e.g., HDMI) to connect the task laptop to the response trigger box
4. MRI-compatible fiber optic response system
5. MRI-compatible audio system
6. 32-channel head coil
7. MRI projection system
8. MRI-safe corrective lenses (if needed)
9. Masking tape

2.3 Emotional MST

1. Desktop or laptop computer
2. Keyboard or mouse for response collection
3. Desk and chair
4. Quiet testing room

2.4 Animal Models

1. L-shaped track (approximate dimensions: width: 8–10 cm, stem length: 60–90 cm, start arm length: 40–60 cm, choice platform: ~30 × 30 cm, wall height: 2–5 cm).
2. Food wells: ~2 cm diameter, shallow recessed cups with one in the start arm and two in the choice platform. Wells should be small

enough to be covered by the object stimuli, and large enough to hold the food reward (e.g., Froot Loop cereal piece).

3. 3 desk lamps for illumination: 1 with a red 60 W bulb near the start area, an additional 60 W red light bulb illuminating the choice platform, and a third lamp with white 60 W bulb aimed at the ceiling for nonanxiogenic illumination to facilitate behavioral analysis.
4. Webcam or video recorder positioned ~50 cm above the choice platform on a boom stand.
5. Computer + behavioral recording software: several open-source options will work (e.g., Collector or BORIS).
6. White noise machine is used during all training/testing to minimize noise distractions.
7. Two distinct three-dimensional objects used for procedural training (plastic figurines roughly 10 cm × 6 cm × 4 cm; three copies of each swapped across trials in a session to reduce scent cueing).
8. LEGO® object sets mounted on a LEGO® base plate cut to 8 pips × 8 pips (~6.5 cm × 6.5 cm square): distinct training pair (constructed to match volume/size but distinct in frontal visual features), eight identical copies of target LEGO® object (used across all trials), three similarity-graded lure objects (two copies of each to swap during testing), one distinct control object (can be a figurine or LEGO® object with no overlapping features), and one identical target copy control.
9. 70% ethanol solution in a spray bottle for cleaning objects between sessions/tests.

---

3 Methods

3.1 Behavioral MST

This short behavioral task can be used to assess mnemonic discrimination performance in participants across a wide age range [142]. During the task, participants view 128 color images of everyday objects on a computer screen and must decide for each whether it is "new" (a foil), "old" (target), or "similar" to an item they saw before (lure). Each image is displayed for 2 s (s), followed by a 0.5 s interstimulus interval. Of the 64 foils presented, 20 are repeated as targets and the remaining 44 as lure items. The lag between the first and second presentation of an object ranges from 4 to 100 intervening trials.

Study stimuli consist of color JPEG images of everyday objects. Six sets of normed stimuli [63] are available for use from the following GitHub repository: https://github.com/celstark/MST. In each set, the similarity between targets and lures is categorized across five levels, with L1 representing very high similarity and L5 representing very low similarity. In order to set up the task, first install PsychoPy locally on the testing computer. Then download the task script ("MST_Continuous_PsychoPy.py") from GitHub (https://github.com/celstark/MST) and select a stimulus set (e.g., Set1, Set2). The same set must be used across all participants.

1. **Participant Eligibility**
    a. Normal or corrected-to-normal vision.
    b. No color blindness.
2. **Practice Task**
    a. Participants should complete the task in a quiet, distraction-free environment while seated comfortably at a viewing distance that allows clear visualization of the stimuli.
    b. Participants receive standardized instructions via verbal explanation alongside a PowerPoint presentation with visual

explanation alongside a PowerPoint presentation with visual examples. Specifically, participants are told that they will be shown a series of images of everyday objects on the computer screen and must indicate if each is “old,” “similar,” or “new” by pressing the V, B, or N key on the keyboard using their index, middle, and ring finger, respectively. Participants are instructed to respond “New” if it is the first time they have seen the image (foils); “Old” if they have seen the exact image during the previous task (targets); or “Similar” if the image closely resembles one presented earlier (lures). When explaining what constitutes a “similar” item, researchers may clarify that an image sharing perceptual features with a previously presented image should be considered “similar.” For example, the same slotted spoon shown from different perspectives would be considered “similar”, not “old.”

c. Participants are then instructed to place their fingers on the V, B, and N keys, and reminded that they should make their responses as quickly and as accurately as possible.

d. Participants are guided through five example trials with feedback and researcher clarification provided if needed, followed by four additional practice trials without researcher feedback.

e. Following the practice, researchers should answer any remaining questions, instruct the participant to start the task when they are ready, and then leave the testing room to minimize distraction.

3. **Behavioral Outcome Measures**

   a. Mnemonic discrimination performance is quantified using the Lure Discrimination Index (LDI), calculated as: *p(“Similar”|Lure) – p(“Similar”|Foil)*. Higher LDI values indicate

superior ability to discriminate similar stimuli, a process linked to hippocampal pattern separation.

b. Additional outcome measures include corrected recognition (REC), defined as: *p("Old"|Target) – p("Old"|Foil)*. Signal detection measures (d′) are also calculated to assess discrimination between (1) targets and lures and (2) targets and foils. Target recognition accuracy [*p("Old"|Target)*] and false alarm rates to lures and novel items are additionally computed.

4. **Data Exclusion Criteria**

   a. Participants should be excluded from analysis if their corrected recognition (REC) score is below 0.5, indicating insufficient engagement with or understanding of the task.

3.2 fMRI MST

The fMRI-adapted version of the MST task should be implemented using a 3 T or above magnet. Participants are administered 320 total trials divided equally among four blocks. Within each block, 40 images are foils, 15 are targets, and 25 are lures. Each image is displayed for 3 s, followed by a 0.5 s ISI. Participants' responses are recorded within this 3.5 s window using an MRI-compatible response box.

1. **Scanner Setup**

   a. Ensure that the computer screen is properly projected onto the screen inside the MRI bore.

   b. Enter the participant's ID into the scanner system.

   c. Launch the MST task on the task computer and enter the participant's ID. Select Block 1.

2. **Practice Outside the Scanner**

   a. Participants will be familiarized with the task before entering the scanner (Video instructions are available on GitHub: https://github.com/celstark/MST). Participants are told that they will be shown pictures of everyday objects on the computer screen and will be instructed to respond "old," "similar," or "new" to each image by pressing a button. Participants are instructed to respond "old" if they believe they have seen the exact image (targets), "similar" if the image resembles one presented previously (lures), and "new" if it is the first time they have seen the image (foils).

   b. Participants are then instructed to place their fingers on the keyboard to indicate their responses: index finger on V, middle finger on B, and ring finger on N. Researchers are encouraged to label these keys with 1, 2, and 3 to mimic the button box used during the actual task inside the scanner.

   c. Participants are guided through five example trials with the correct answers explained, followed by four additional practice images that require participant responses.

   d. Because participants practice using a desktop or laptop keyboard, they are reminded that during the task inside the scanner, they will use a button box resting on their chest instead of a keyboard. However, the same fingers will be used: index finger for "old," middle finger for "similar," and ring finger for "new" responses.

3. **Participant Setup in the Scanner**

   a. Conduct standard IRB-approved metal safety procedures (i.e., metal screening, no key cards or electronics, restroom use,

etc.).

b. Provide corrective MRI-safe lenses if necessary. To test lens efficacy, ask the participant to stand approximately 2 ft away from a computer screen. Adjust lens strength as needed until the image is clear.

c. Appropriate staff should escort participants into the MRI chamber.

d. Ask participants to sit on the MRI bed and remove their shoes.

e. Provide the participant with earplugs and headphones (if using).

f. Ask participants to lie supine on the bed. Padding should be placed under the participant's knees and head. A blanket can be offered if available.

g. Ask the participant to place the response box in a comfortable position on their upper body (typically chest or stomach area). Guide their fingers to the correct buttons on the response box and tape the cord so that the response box stays in place. Tape the emergency squeeze ball to the participant's lower torso.

h. Place the headcoil (a 32-channel or more is recommended) and the mirror over the participant's head.

i. Send the participant into the scanner bore and ensure that the projector is in their field of view.

j. Exit the chamber and close the MRI room door.

4. **Task Procedure Inside the Scanner**

a. Check in with the participant and remind them which buttons and fingers correspond to each response.

b. To ensure that participants have their fingers placed on the

b. To ensure that participants have their fingers placed on the correct buttons of the response box, ask participants to make

each of the three response types one at a time, confirming that the appropriate finger is on the appropriate response button. If the responses are incorrect, researchers should instruct participants to adjust their finger placement so that the recorded responses accurately reflect their intended choices.

c. Confirm that the participant is ready to start the task.

d. Start the task.

e. Check in with the participant after each block and remind the participant that the images presented in each block are independent of those shown in the previous block.

f. Once the scan is over, slowly remove the participant from the bed.

3.3 Emotional MST

This behavioral task can be used to assess mnemonic similarity performance for negative, neutral, and positive stimuli. There are several variants of the task, but the methods described below are for the original emotional MST [39]. During the task, participants viewed a total of 149 scene images during encoding and 290 scene images during retrieval, all presented on a computer screen with a black background. Each image is displayed for 2.5 s, followed by a 0.5 s interstimulus interval. During encoding, participants perform an incidental task by rating each scene for emotional valence as negative (button 1), neutral (button 2), or positive (button 3). Below each image, the response options are shown as a reminder in white text. Images are roughly balanced across negative, neutral, and positive stimuli based on average ratings for each image.

During retrieval, images are also distributed across stimulus types: targets (repeated images), lures (similar but not identical images), and foils (brand new images). For each stimulus, participants are instructed to indicate whether the image was "exact same," or "new/different." Participants are instructed to respond "exact same" if the image was exactly the same as a previously presented image and to respond "new/different" if the image was new or different in any way. Examples are provided of each type of trial to ensure understanding. Responses are scored as correct when target items were labeled "old," lure items were labeled "new/different," and foils were labeled "new/different."

Study stimuli consist of color JPEG images of everyday kinds of scenes (e.g., cooking dinner, walking by the beach, dogs playing). Images and task materials are available at https://github.com/lealmemorylab/Emotional-MDT-OG-. This task can also be conducted in the MRI scanner with very little changes needed (e.g., changing how the task starts to trigger with scanner, usually a 5). Screen size may need to be adjusted depending on the computer used. To install the task, download PsychoPy locally on the testing computer. Then copy the folder onto the local desktop and run the three.py files (study.py, test1.py, test2.py).

1. **Participant Eligibility**
    a. Normal or corrected-to-normal vision.
    b. No color blindness.
2. **Participant Preparation**
    a. Participants should complete the task in a quiet, distraction-free environment while seated comfortably at a viewing distance that allows clear visualization of the stimuli.
    b. Vision correction must be worn as needed.

3.

Main Task Procedure: Encoding Phase

a. Participants receive standardized instructions via verbal explanation alongside a PowerPoint presentation with visual examples. Specifically, participants are told that they will be shown a series of images on the computer screen for 2.5 s and will be instructed to determine whether they think the image is negative, neutral, or positive (1 = negative, 2 = neutral, 3 = positive). They must respond while the image is still on the screen.

b. Participants are then instructed to place their fingers on the keyboard to indicate their responses to be able to respond quickly. Other versions of the emotional MST have used a 1-9 scale response of emotional valence or other responses such as "indoor" and "outdoor"—the responses during encoding are flexible.

c. Researchers should answer any remaining questions and then leave the testing room to minimize distraction.

d. Participants are alone in the testing room to complete each section of the emotional MST without interruption.

4. **Main Task Procedure: Retrieval Phase**

a. After encoding, participants receive standardized instructions via verbal explanation alongside a PowerPoint presentation with visual examples. Specifically, participants are told that they will be shown a series of images on the computer screen for 2.5 s and are instructed to respond "Exact Same," or "New/Different" to each image while the image is still on the screen. Participants are instructed to respond "Exact Same" if they believe they have seen the exact scene during the task

(targets), and “New/Different” if it is new or different in any way—this could mean a similar but not identical image to one during encoding or a completely new image (lures and foils).

b. Participants are instructed to place their fingers on the keyboard to indicate their responses: 1 = exact same and 2 = new/different.

c. Participants are guided through an example to ensure understanding of how to treat the similar lure images.

d. Researchers should answer any remaining questions and then leave the testing room to minimize distraction.

e. Participants are alone in the testing room to complete each section of the MST without interruption.

f. The test portion is broken up into two parts to minimize fatigue (test 1 and test 2).

5. **Behavioral Outcome Measures**

a. Performance can be quantified using a Lure Discrimination Index (LDI), calculated as: *LDI = p(“New/different”|Lure) – p(“New/different”|Target)*. Higher LDI values indicate superior ability to discriminate similar stimuli, a process linked to hippocampal pattern separation.

b. Additional outcome measures include target recognition (d'), defined as: *d' = z [p(“Exact same”|Target)] – z[p(“Exact same”|Foil)]*.

3.4 Animal Models

Of the tasks that have been developed for testing mnemonic similarity in rodents, the one that most closely parallels the human MST is the

LEGO-block-based mnemonic similarity task. Notably, this variant of the task has been used to causally link hippocampus hyperactivity [143] and perforant pathway fiber damage [144] to behavioral decline. The following step-by-step experimental protocol will therefore focus on the LEGO block-based rodent MST methodology.

1. **Animal strain consideration and preparation**
    a. Young (5–9 months) and Aged (21–25 month) rats. Note, Fischer 344 × Brown Norway hybrid rats were used in published studies due to their superior visual acuity relative to other strains [145], and the access to aged animals from the NIA colony. However, Long Evans or other pigmented strains could also be used. Inbred Fischer 344 rats cannot discriminate between LEGO® objects and should not be used.
    b. Rats are individually housed for food restriction and should be maintained on a reversed 12-h light-dark cycle.
    c. Training and testing occur during the dark phase of the diurnal cycle.
    d. Rats are food-restricted to 85% of their body weight, corresponding to an optimal body condition of 3 [146]. For young animals, this is frequently the baseline weight during ad libitum feeding. For aged rats, in which a lifetime of ad libitum feeding has produced excessive fat reserves and over conditioning, a higher percentage of initial weight loss is usually required before motivation to retrieve the reward is achieved.
2. **Habituation and Shaping**
    a. Purpose: to familiarize the rats with the track and shape to traverse between the start arm and choice platform, ensuring similar levels of motivation between young and old age groups

similar levels of motivation between young and old age groups.

b. Testing sessions should occur 5–6 days per week at approximately the same time each day for consistency.

c. Habituate each rat to the track for 1–3 days for 10 min/day, encouraging exploration by spreading pieces of Froot Loop cereal or other appetitive rewards around the track.

d. After habituation, begin shaping sessions to train rats to shuttle between the start arm and the choice platform.

e. Bait one of the two food wells on the choice platform with a single piece of Froot Loop cereal. Randomize the side of the baited food well (left vs. right) across trials.

f. The rat is required to exit the start area, traverse the full length of the track arm, and retrieve the food reward from the baited well on the choice platform.

g. Deliver a second food reward in the start area upon return.

h. Continue shaping until the rat successfully alternates between the start area and the choice platform 32 times within a 20-min session. If a rat does not reach shaping criteria within 7 days, they require additional food restriction or a new reward should be tried.

3. **Procedural Training: Standard object discrimination**

a. Purpose: Train rats to push objects to obtain a reward and ensure that the task rule (one rewarded target and one unrewarded lure) is acquired prior to LEGO block mnemonic similarity task testing.

b. Designate one object in the pair as the rewarded stimulus (S+), and the alternate object as the non-rewarded stimulus (S−).

c. At the start of each trial, place the rat in the start area. Allow

c. At the start of each trial, place the rat in the start area. Allow the rat to exit the start area and traverse the arm to the choice platform.
d. Present the two objects simultaneously, each covering a food well on the choice platform.
e. Pseudo-randomize the side of the S+ object/baited food well (left vs. right) across trials.
f. Ensure equal numbers of left- and right-rewarded trials within each session.
g. Permit the rat to make a choice by displacing one object.
h. If the rat displaces the S+ object, the rat consumes the food reward from the choice platform. Upon returning to the start area, provide a second food reward (e.g., a piece of Froot Loop cereal; Kellogg's).
i. If the rat displaces the S− object, immediately remove both objects and the unretrieved food reward from the choice platform.
j. Allow the rat to return to the start area. Do not provide a food reward in the start area.
k. Conduct 32 trials per session for each rat and continue training daily until the rat achieves ≥26 correct responses out of 32 trials (≥81.3% correct) on two consecutive days.
l. Use two or three identical copies of each object, rotated across trials within sessions, and thoroughly clean all object copies with 70% ethanol between test sessions to ensure that scent cues are not impacting performance.

4. **Main Task Procedure: Target Training Phase**

a. Purpose: to pre-train rats to identify the Target LEGO object (S+) that would be rewarded throughout MST testing.
b. Designate a single LEGO object as the target (S+), which remains rewarded for the entirety of the experiment.
c. Use a relatively distinct LEGO object as the lure (S−) during pre-training.
d. LEGO pips can be used to calculate feature overlap. A pip was defined as the smallest visible surface unit corresponding to a single LEGO stud.
e. Construct both objects with similar overall dimensions for volume, shape, and texture, and design the objects to have 38% visible feature overlap on the front-facing surface (21 of 56 LEGO "pips"). By counting the number of front-facing LEGO pips visible to the rat as it approached the choice platform.
f. For each object, all visible pips on the front-facing surface are counted, and overlap was calculated as the number of pips shared between the two objects at corresponding spatial locations divided by the total number of visible pips. This measure provided a standardized, perceptually relevant estimate of object similarity based solely on visual surface features. Note that color is not figured into the calculation, as rats do not have color vision.
g. Present the target (S+) and distinct lure (S−) simultaneously on the choice platform.
h. Continue target training until each rat achieves ≥81.3% correct responses (≥26 correct out of 32 trials) over a full single-day session.

5. **Main Task Procedure: The LEGO Mnemonic Similarity Task**

a. After reaching the criterion in target training, give rats 2 days off before the first test session.

b. Conduct testing every 3 days and continue this schedule until four total test sessions are completed. This spacing is to prevent over-familiarization with lure objects.

c. Each test session consists of 50 discrimination trials, all presented in pseudorandom order, including:

    i. Ten trials with a distinct control object (nonLEGO figurine) for the positive control condition to ensure that both age groups can perform the procedural aspects of the task similarly.

    ii. Ten trials each with three similar lure objects constructed from LEGO® blocks with ~50%, ~70%, and ~90% visual feature overlap (30 trials total).

    iii. Ten trials with an identical copy of the target object to serve as a negative control (rats should be at chance for this condition, or they are using odor cues to solve the task).

d. Rotate eight identical copies of the target object across trials within each session.

e. Rotate two identical copies of each lure object across trials.

f. Clean all object copies thoroughly with 70% ethanol between test sessions.

6. **Behavioral Outcome Measures**

a. Assess MST ability as percent correct responses for each trial type

type.

b. Record all sessions and review videos using open-source behavioral analysis software (e.g., Collector/Minion; Burke/Maurer Laboratories or BORIS).

c. Score behavior trial-by-trial, including: reaction time and response selection behavior/checking behavior.

d. Score response selection as a binary variable (checking vs. no checking).

e. Define checking as hesitation or pausing in front of one object, followed by withdrawal and investigation or displacement of the alternate object.

f. Define reaction time as the interval between the video frame when the rat's nose crosses the choice platform threshold and the video frame when the rat initiates object displacement.

g. Finally, calculate the response bias for each object pair (that is, overlap condition). This measures the tendency to select an object on a particular side of the choice platform regardless of the object characteristics. Response bias is the absolute value of left side choices minus right-side choices, divided by the total number of choices for that overlap condition. A value of 0 indicates no bias and a value of 1 is the maximal bias, indicating that the same side was selected for every trial.

---

4 Conclusion

Mnemonic discrimination tasks, including the MST, are powerful tools for assessing not only memory, but also underlying neurobiology and pathology [5, 7, 13, 19]. This task is sensitive to perturbations caused by healthy aging, dementia, hippocampal lesions, and mental health conditions [5, 13, 19], and has been validated across a range of species, including mice, rats, non-human primates, and humans [13, 17, 18].

The MST and its variants, therefore, have utility as a means of conducting everything from basic research to clinical trials. In the current paper, we have outlined how to conduct several types of mnemonic discrimination tasks. Our aim is to foster research utilizing these tasks, as they have huge promise for advancing the fields of memory and aging.

Funding sources

The Alzheimer's Association grants AARG-22-928813, 24HPE-1287002, and A2022040S, the National Institute of Health, National Institute on Aging grants R01AG079512 and R01AG049722, and the Brain & Behavior Research Foundation Young Investigator Grant 30897.

Competing Interests

The authors declare no competing interests.

References

1. Tulving E (1972) Episodic and semantic memory. In: Organization of memory. Academic Press, Oxford, England, pp xiii, 423–xiii, 423

2. Ferguson SA, Hashtroudi S, Johnson MK (1992) Age differences in using source-relevant cues. Psychol Aging 7:443–452
[PubMed]

3. Koutstaal W, Schacter DL (1997) Gist-based false recognition of pictures in older and younger adults. J Mem Lang 37:555–583

4. Schacter DL, Koutstaal W, Johnson MK, Gross MS, Angell KE (1997) False recollection induced by photographs: a comparison of older and younger adults. Psychol Aging 12:203–215
[PubMed]

5. Stark SM, Kirwan CB, Stark CEL (2019) Mnemonic similarity task: a tool for assessing hippocampal integrity. Trends Cogn Sci 23:938–951
[PubMed][PubMedCentral]

6. Nichols E et al (2022) Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 7:e105–e125

7. Stark SM, Yassa MA, Lacy JW, Stark CEL (2013) A task to assess behavioral pattern separation (BPS) in humans: data from healthy aging and mild cognitive impairment. Neuropsychologia

51:2442–2449
[PubMed][PubMedCentral]

8. Gilbert PE, Kesner RP, DeCoteau WE (1998) Memory for spatial location: role of the hippocampus in mediating spatial pattern separation. J Neurosci Off J Soc Neurosci 18:804–810

9. Gilbert PE, Kesner RP, Lee I (2001) Dissociating hippocampal subregions: a double dissociation between dentate gyrus and CA1. Hippocampus 11:626–636
[PubMed]

10. Johnson SA et al (2017) Rodent age-related impairments in discriminating perceptually similar objects parallel those observed in humans. Hippocampus 27:759–776
[PubMed][PubMedCentral]

11. McAllister KAL, Saksida LM, Bussey TJ (2013) Dissociation between memory retention across a delay and pattern separation following medial prefrontal cortex lesions in the touchscreen TUNL task. Neurobiol Learn Mem 101:120–126
[PubMed][PubMedCentral]

12. Talpos JC, McTighe SM, Dias R, Saksida LM, Bussey TJ (2010) Trial-unique, delayed nonmatching-to-location (TUNL): a novel, highly hippocampus-dependent automated touchscreen test of location memory and pattern separation. Neurobiol Learn Mem 94:341–352
[PubMed][PubMedCentral]

13. Yassa MA, Stark CEL (2011) Pattern separation in the hippocampus. Trends Neurosci 34:515–525
[PubMed][PubMedCentral]

14. McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol Rev 102:419–457
[PubMed]

15. Knierim JJ (2015) The hippocampus. Curr Biol CB 25:R1116–R1121
[PubMed]

16. Treves A, Rolls ET (1994) Computational analysis of the role of the hippocampus in memory. Hippocampus 4:374–391
[PubMed]

17. Holden HM, Gilbert PE (2012) Less efficient pattern separation may contribute to age-related spatial memory deficits. Front Aging Neurosci 4:9

18. Leal SL, Yassa MA (2015) Neurocognitive aging and the hippocampus across species. Trends Neurosci 38:800–812
[PubMed][PubMedCentral]

19.

Leal SL, Yassa MA (2018) Integrating new findings and examining clinical applications of pattern separation. Nat Neurosci 21:163–173
[https://doi.org/10.1038/s41593-017-0065-1][PubMed][PubMedCentral]

20. Bakker A et al (2012) Reduction of hippocampal hyperactivity improves cognition in amnestic mild cognitive impairment. Neuron 74:467–474
[PubMed][PubMedCentral]

21. Berron D et al (2019) Higher CSF tau levels are related to hippocampal hyperactivity and object mnemonic discrimination in older adults. J Neurosci 39:8788–8798
[PubMed][PubMedCentral]

22. Corona-Long CA et al (2020) Comparison of male and female patients with amnestic mild cognitive impairment: hippocampal hyperactivity and pattern separation memory performance. Alzheimers Dement Diagn Assess Dis Monit 12:e12043

23. Putcha D et al (2011) Hippocampal hyperactivation associated with cortical thinning in alzheimer's disease signature regions in non-demented elderly adults. J Neurosci 31:17680–17688
[PubMed][PubMedCentral]

24. Yassa MA et al (2011) Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. Hippocampus 21:968–979
[PubMed]

25. Belliart-Guérin G, Planche V (2023) Mnemonic discrimination performance in a memory clinic: a pilot study. J Alzheimer's Dis 94:1527–1534

26. Kim S et al (2023) Examining the diagnostic value of the mnemonic discrimination task for classification of cognitive status and amyloid-beta burden. Neuropsychologia 191:108727
[PubMed][PubMedCentral]

27. Vanderlip CR, Lee MD, Stark CEL (2024) Cognitive modeling of the Mnemonic Similarity Task as a digital biomarker for Alzheimer's disease. Alzheimers Dement 20:6935–6947
[PubMed][PubMedCentral]

28. Jack CR et al (2018) NIA-AA Research framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement J Alzheimers Assoc 14:535–562

29. Jack CR et al (2024) Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's association workgroup. Alzheimers Dement 20:5143–5169
[PubMed][PubMedCentral]

30. Leal SL, Ferguson LA, Harrison TM, Jagust WJ (2019) Development of a mnemonic discrimination task using naturalistic stimuli with applications to aging and preclinical Alzheimer's disease. Learn Mem 26:219–228
[PubMed][PubMedCentral]

31. Maass A et al (2019) Alzheimer's pathology targets distinct memory networks in the ageing

brain. Brain 142:2492–2509
[PubMed][PubMedCentral]

32. Marks SM, Lockhart SN, Baker SL, Jagust WJ (2017) Tau and β-amyloid are associated with medial temporal lobe structure, function, and memory encoding in normal aging. J Neurosci 37:3192–3201
[PubMed][PubMedCentral]

33. Vanderlip CR et al (2025) Amyloid-beta deposition in basal frontotemporal cortex is associated with selective disruption of temporal mnemonic discrimination. J Neurosci 45(10):e1605242025

34. Azab M, Stark SM, Stark CEL (2014) Contributions of human hippocampal subfields to spatial and temporal pattern separation. Hippocampus 24:293–302
[PubMed]

35. Paleja M, Girard TA, Herdman KA, Christensen BK (2014) Two distinct neural networks functionally connected to the human hippocampus during pattern separation tasks. Brain Cogn 92:101–111

36. Reagh ZM et al (2014) Spatial discrimination deficits as a function of mnemonic interference in aged adults with and without memory impairment. Hippocampus 24:303–314
[PubMed]

37. Reagh ZM et al (2016) Greater loss of object than spatial mnemonic discrimination in aged adults. Hippocampus 26:417–422
[PubMed][PubMedCentral]

38. Stark SM, Yassa MA, Stark CEL (2010) Individual differences in spatial pattern separation performance associated with healthy aging in humans. Learn Mem 17:284–288
[PubMed][PubMedCentral]

39. Leal SL, Tighe SK, Yassa MA (2014) Asymmetric effects of emotion on mnemonic interference. Neurobiol Learn Mem 111:41–48
[PubMed][PubMedCentral]

40. Stark SM, Stevenson R, Wu C, Rutledge S, Stark CEL (2015) Stability of age-related deficits in the mnemonic similarity task across task variations. Behav Neurosci 129:257–268
[PubMed][PubMedCentral]

41. Berron D et al (2018) Age-related functional changes in domain-specific medial temporal lobe pathways. Neurobiol Aging 65:86–97
[PubMed]

42. Güsten J, Ziegler G, Düzel E, Berron D (2021) Age impairs mnemonic discrimination of objects more than scenes: a web-based, large-scale approach across the lifespan. Cortex 137:138–148
[PubMed]

43.

Stark SM, Stark CEL (2017) Age-related deficits in the mnemonic similarity task for objects and scenes. Behav Brain Res 333:109–117
[PubMed][PubMedCentral]

44. Webb CE, Foster CM, Horn MM, Kennedy KM, Rodrigue KM (2020) Beta-amyloid burden predicts poorer mnemonic discrimination in cognitively normal older adults. Neuroimage 221:117199
[PubMed][PubMedCentral]

45. Dickerson BC, Eichenbaum H (2010) The episodic memory system: neurocircuitry and disorders. Neuropsychopharmacology 35:86–104
[PubMed][PubMedCentral]

46. Eichenbaum H (2004) Hippocampus: cognitive processes and neural representations that underlie declarative memory. Neuron 44:109–120
[PubMed]

47. Rolls ET (2013) The mechanisms for pattern completion and pattern separation in the hippocampus. Front Syst Neurosci 7:74

48. Rolls ET (2007) An attractor network in the hippocampus: theory and neurophysiology. Learn Mem 14:714–731
[PubMed]

49. Lodge M, Bischofberger J (2019) Synaptic properties of newly generated granule cells support sparse coding in the adult hippocampus. Behav Brain Res 372:112036
[PubMed]

50. Nakashiba T et al (2012) Young dentate granule cells mediate pattern separation, whereas old granule cells facilitate pattern completion. Cell 149:188–201
[PubMed][PubMedCentral]

51. Sahay Aet al (2011) Increasing adult hippocampal neurogenesis is sufficient to improve pattern separation. Nat 472:466–470

52. Sahay A, Wilson DA, Hen R (2011) Pattern separation: a common function for new Neurons in Hippocampus and Olfactory Bulb. Neuron 70:582–588
[PubMed][PubMedCentral]

53. Cahill L, McGaugh JL (1996) Modulation of memory storage. Curr Opin Neurobiol 6:237–242
[PubMed]

54. Hasselmo ME (2006) The role of acetylcholine in learning and memory. Curr Opin Neurobiol 16:710–715
[PubMed][PubMedCentral]

55. Smith CC, Greene RW (2012) CNS dopamine transmission mediated by noradrenergic innervation. J Neurosci 32:6072–6080
[PubMed][PubMedCentral]

56. Carello-Collar G et al (2023) The GABAergic system in Alzheimer's disease: a systematic review with meta-analysis. Mol Psychiatry 28:5025–5036
[https://doi.org/10.1038/s41380-023-02140-w][PubMed]

57. Jiménez-Balado J, Eich TS (2021) GABAergic dysfunction, neural network hyperactivity and memory impairments in human aging and Alzheimer's disease. Semin Cell Dev Biol 116:146–159
[PubMed][PubMedCentral]

58. Li Y et al (2016) Implications of GABAergic neurotransmission in Alzheimer's disease. Front Aging Neurosci 8:31

59. Solas M, Puerta E, Ramirez J (2015) M. Treatment options in Alzheimer´s disease: the GABA story. Curr Pharm Des 21:4960–4971
[PubMed]

60. Wilson IA, Gallagher M, Eichenbaum H, Tanila H (2006) Neurocognitive aging: prior memories hinder new hippocampal encoding. Trends Neurosci 29:662–670
[PubMed][PubMedCentral]

61. Barnes CA, Rao G, Houston FP (2000) LTP induction threshold change in old rats at the perforant path–granule cell synapse. Neurobiol Aging 21:613–620
[PubMed]

62. Yassa MA, Muftuler LT, Stark CEL (2010) Ultrahigh-resolution microstructural diffusion tensor imaging reveals perforant path degradation in aged humans in vivo. Proc Natl Acad Sci U S A 107:12687–12691
[PubMed][PubMedCentral]

63. Yassa MA et al (2010) High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnestic Mild Cognitive Impairment. Neuroimage 51:1242–1252
[PubMed][PubMedCentral]

64. Riphagen JM et al (2020) Associations between pattern separation and hippocampal subfield structure and function vary along the lifespan: A 7 T imaging study. Sci Rep 10:7572
[https://doi.org/10.1038/s41598-020-64595-z][PubMed][PubMedCentral]

65. Braak H, Thal DR, Ghebremedhin E, Del Tredici K (2011) Stages of the pathologic process in alzheimer disease: age categories from 1 to 100 years. J Neuropathol Exp Neurol 70:960–969
[PubMed]

66. Jagust W (2018) Imaging the evolution and pathophysiology of Alzheimer disease. Nat Rev Neurosci 19:687–700
[https://doi.org/10.1038/s41583-018-0067-3][PubMed][PubMedCentral]

67. Jia J et al (2024) Biomarker changes during 20 years preceding Alzheimer's disease. N Engl J Med 390:712–722
[PubMed]

68. Andrés-Benito P et al (2017) Locus coeruleus at asymptomatic early and middle Braak stages of neurofibrillary tangle pathology. Neuropathol Appl Neurobiol 43:373–392 [PubMed]

69. Ehrenberg AJ et al (2017) Quantifying the accretion of hyperphosphorylated tau in the locus coeruleus and dorsal raphe nucleus: the pathological building blocks of early Alzheimer's disease. Neuropathol Appl Neurobiol 43:393–408 [PubMed][PubMedCentral]

70. Braak H, Del Tredici K (2011) The pathological process underlying Alzheimer's disease in individuals under thirty. Acta Neuropathol (Berl) 121:171–181 [PubMed]

71. Guo JL, Lee VMY (2011) Seeding of normal tau by pathological tau conformers drives pathogenesis of Alzheimer-like tangles. J Biol Chem 286:15317–15331 [PubMed][PubMedCentral]

72. Jacobs HIL et al (2021) In vivo and neuropathology data support locus coeruleus integrity as indicator of Alzheimer's disease pathology and cognitive decline. Sci Transl Med 13:2511

73. Sun W et al (2020) A probabilistic atlas of locus coeruleus pathways to transentorhinal cortex for connectome imaging in Alzheimer's disease. Neuroimage 223:117301 [PubMed][PubMedCentral]

74. Braak H, Braak E (1991) Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol (Berl) 82:239–259 [PubMed]

75. Knierim JJ, Neunuebel JP, Deshmukh SS (2014) Functional correlates of the lateral and medial entorhinal cortex: objects, path integration and local–global reference frames. Philos Trans R Soc B Biol Sci 369:20130369

76. Reagh ZM, Yassa MA (2014) Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. Proc Natl Acad Sci U S A 111:E4264–E4273 [PubMed][PubMedCentral]

77. Thal DR, Rüb U, Orantes M, Braak H (2002) Phases of Aβ-deposition in the human brain and its relevance for the development of AD. Neurology 58:1791–1800 [PubMed]

78. Brier MR et al (2016) Tau and Ab imaging, CSF measures, and cognition in Alzheimer's disease. Sci Transl Med 8(338):338ra66

79. Adams JN, Harrison TM, Maass A, Baker SL, Jagust WJ (2022) Distinct factors drive the spatiotemporal progression of tau pathology in older adults. J Neurosci 42:1352–1361 [PubMed]

80. Sanchez JS et al (2021) The cortical origin and initial spread of medial temporal tauopathy in Alzheimer's disease assessed with positron emission tomography. Sci Transl Med

13:eabc0655
[PubMed][PubMedCentral]

81. Schöll M et al (2016) PET imaging of Tau deposition in the aging human brain. Neuron 89:971–982
[PubMed][PubMedCentral]

82. Hanseeuw BJ et al (2019) Association of Amyloid and Tau with cognition in preclinical Alzheimer Disease: a longitudinal study. JAMA Neurol 76:915–924
[PubMed][PubMedCentral]

83. Leal SL, Lockhart SN, Maass A, Bell RK, Jagust WJ (2018) Subthreshold amyloid predicts Tau deposition in aging. J Neurosci 38:4482–4489
[PubMed][PubMedCentral]

84. Nelson PT et al (2012) Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature. J Neuropathol Exp Neurol 71:362–381
[PubMed][PubMedCentral]

85. Sperling RA et al (2019) The impact of amyloid-beta and tau on prospective cognitive decline in older individuals. Ann Neurol 85:181–193
[PubMed][PubMedCentral]

86. Boldrini M et al (2018) Human hippocampal neurogenesis persists throughout aging. Cell Stem Cell 22:589–599.e5
[PubMed][PubMedCentral]

87. Moreno-Jiménez EP et al (2019) Adult hippocampal neurogenesis is abundant in neurologically healthy subjects and drops sharply in patients with Alzheimer's disease. Nat Med 25:554–560
[https://doi.org/10.1038/s41591-019-0375-9][PubMed]

88. Spalding KL et al (2013) Dynamics of hippocampal neurogenesis in adult humans. Cell 153:1219–1227
[PubMed][PubMedCentral]

89. Choi SH, Tanzi RE (2023) Adult neurogenesis in Alzheimer's disease. Hippocampus 33:307–321
[PubMed]

90. Teixeira CM et al (2018) Untold new beginnings: adult hippocampal neurogenesis and Alzheimer's disease. J Alzheimer's Dis 64:S497–S505

91. Ding Y et al (2023) Electroacupuncture promotes neurogenesis in the dentate gyrus and improves pattern separation in an early Alzheimer's disease mouse model. Biol Res 56:65
[https://doi.org/10.1186/s40659-023-00472-z][PubMed][PubMedCentral]

92. Hollands C et al (2017) Depletion of adult neurogenesis exacerbates cognitive deficits in Alzheimer's disease by compromising hippocampal inhibition. Mol Neurodegener 12:64

[PubMed][PubMedCentral]

93. Adolphs R, Denburg NL, Tranel D (2001) The amygdala's role in long-term declarative memory for gist and detail. Behav Neurosci 115:983–992
[PubMed]

94. Denburg NL, Buchanan TW, Tranel D, Adolphs R (2003) Evidence for preserved emotional memory in normal older persons. Emotion 3:239–253
[PubMed]

95. Kensinger EA, Garoff-Eaton RJ, Schacter DL (2007) Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. J Mem Lang 56:575–591

96. Leal SL, Yassa MA (2014) Effects of aging on mnemonic discrimination of emotional information. Behav Neurosci 128:539–547
[PubMed][PubMedCentral]

97. Hamilton JP, Gotlib IH (2008) Neural substrates of increased memory sensitivity for negative stimuli in major depression. Biol Psychiatr 63:1155–1162

98. LeMoult J, Gotlib IH (2019) Depression: a cognitive perspective. Clin Psychol Rev 69:51–66
[PubMed]

99. Déry N et al (2013) Adult hippocampal neurogenesis reduces memory interference in humans: opposing effects of aerobic exercise and depression. Front Neurosci 0:66

100. Dillon DG, Pizzagalli DA (2018) Mechanisms of memory disruption in depression. Trends Neurosci 41:137–149
[PubMed][PubMedCentral]

101. Leal SL, Noche JA, Murray EA, Yassa MA (2017) Disruption of amygdala–entorhinal–hippocampal network in late-life depression. Hippocampus 27:464–476
[PubMed][PubMedCentral]

102. Shelton DJ, Kirwan CB (2013) A possible negative influence of depression on the ability to overcome memory interference. Behav Brain Res 256:20–26
[PubMed]

103. Hayes BK et al (2023) Emotion regulation during encoding reduces negative and enhances neutral mnemonic discrimination in individuals with depressive symptoms. Neurobiol Learn Mem 205:107824
[PubMed]

104. Phillips TO et al (2023) Perceived antidepressant efficacy associated with reduced negative and enhanced neutral mnemonic discrimination. Front Hum Neurosci 17:1225836
[PubMed][PubMedCentral]

105. Chi S, Yu J-T, Tan M-S, Tan L (2014) Depression in Alzheimer's disease: epidemiology, mechanisms, and management. J Alzheimer's Dis 42:739–755

106. Livingston G et al (2024) Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet 404:572–628 [PubMed]

107. Ly M et al (2021) Late-life depression and increased risk of dementia: a longitudinal cohort study. Transl Psychiatry 2021 111 11:1–10

108. Panza F et al (2010) Late-Life depression, mild cognitive impairment, and dementia: possible continuum? Am J Geriatr Psychiatr 18:98–116

109. Ferguson LA, Leal SL (2022) Interactions of emotion and memory in the aging brain: neural and psychological correlates. Curr Behav Neurosci Rep 2022:1–11. https://doi.org/10.1007/S40473-021-00245-6

110. McGaugh JL (2002) Memory consolidation and the amygdala: a systems perspective. Trends Neurosci 25:456–461 [PubMed]

111. Sara SJ (2009) The locus coeruleus and noradrenergic modulation of cognition. Nat Rev Neurosci 10:211–223 [https://doi.org/10.1038/nrn2573][PubMed]

112. Clewett D, Huang R, Davachi L (2025) Locus coeruleus activation "resets" hippocampal event representations and separates adjacent memories. Neuron 113:2521–2535.e8 [PubMed]

113. Mather M, Harley CW (2016) The Locus Coeruleus: essential for maintaining cognitive function and the aging brain. Trends Cogn Sci 20:214–226 [PubMed][PubMedCentral]

114. Segal SK, Stark SM, Kattan D, Stark CE, Yassa MA (2012) Norepinephrine-mediated emotional arousal facilitates subsequent pattern separation. Neurobiol Learn Mem 97:465–469 [PubMed][PubMedCentral]

115. Betts MJ et al (2019) Locus coeruleus imaging as a biomarker for noradrenergic dysfunction in neurodegenerative diseases. Brain 142:2558–2571 [PubMed][PubMedCentral]

116. Clewett DV et al (2016) Neuromelanin marks the spot: identifying a locus coeruleus biomarker of cognitive reserve in healthy aging. Neurobiol Aging 37:117–126 [PubMed]

117. Cunningham TJ, Leal SL, Yassa MA, Payne JD (2018) Post-encoding stress enhances mnemonic discrimination of negative stimuli. Learn Mem 25:611–619 [PubMed][PubMedCentral]

118. Szollosi Á, Kéri S, Racsmány M (2022) The key to superior memory encoding under stress: the relationship between cortisol response and mnemonic discrimination. Learn Mem 29:7–15

[PubMed]

119. Lupien SJ et al (1998) Cortisol levels during human aging predict hippocampal atrophy and memory deficits. Nat Neurosci 1998 11 1:69–73

120. Mather M (2021) Noradrenaline in the aging brain: promoting cognitive reserve or accelerating Alzheimer's disease? Semin Cell Dev Biol 116:108
[PubMed][PubMedCentral]

121. Ouanes S, Popp J (2019) High cortisol and the risk of dementia and alzheimer's disease: a review of the literature. Front Aging Neurosci 11:43
[PubMed][PubMedCentral]

122. Jacobs HIL, Riphagen JM, Ramakers IHGB, Verhey FRJ (2021) Alzheimer's disease pathology: pathways between central norepinephrine activity, memory, and neuropsychiatric symptoms. Mol Psychiatr 26:897–906

123. Leal SL, Tighe SK, Jones CK, Yassa MA (2014) Pattern separation of emotional information in hippocampal dentate and CA3. Hippocampus 24:1146–1155
[PubMed][PubMedCentral]

124. Leal SL, Noche JA, Murray EA, Yassa MA (2017) Age-related individual variability in memory performance is associated with amygdala-hippocampal circuit function and emotional pattern separation. Neurobiol Aging 49:9
[PubMed]

125. Hunsaker MR, Kesner RP (2013) The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. Neurosci Biobehav Rev 37:36–58
[PubMed]

126. Burke SN et al (2011) Age-associated deficits in pattern separation functions of the perirhinal cortex: a cross-species consensus. Behav Neurosci 125:836–847
[PubMed][PubMedCentral]

127. Bekinschtein P et al (2013) BDNF in the dentate gyrus is required for consolidation of "Pattern-Separated" memories. Cell Rep 5:759–768
[PubMed][PubMedCentral]

128. Kannangara TS et al (2015) GluN2A−/− Mice Lack bidirectional synaptic plasticity in the dentate gyrus and perform poorly on spatial pattern separation tasks. Cereb Cortex 25:2102–2113
[PubMed]

129. McHugh TJ et al (2007) Dentate gyrus NMDA Receptors mediate rapid pattern separation in the hippocampal network. Science 317:94–99
[PubMed]

130. Clelland CD et al (2009) A functional role for adult hippocampal neurogenesis in spatial

pattern separation. Science 325:210–213
[PubMed][PubMedCentral]

131. Maurer AP et al (2017) Age-related changes in lateral entorhinal and CA3 neuron allocation predict poor performance on object discrimination. Front Syst Neurosci 11:49

132. Robitsek J, Ratner MH, Stewart T, Eichenbaum H, Farb DH (2015) Combined administration of levetiracetam and valproic acid attenuates age-related hyperactivity of CA3 place cells, reduces place field area, and increases spatial information content in aged rat hippocampus. Hippocampus 25:1541–1555
[PubMed][PubMedCentral]

133. Thomé A, Gray DT, Erickson CA, Lipa P, Barnes CA (2016) Memory impairment in aged primates is associated with region-specific network dysfunction. Mol Psychiatr 21:1257–1262

134. Wilson IA, Ikonen S, Gallagher M, Eichenbaum H, Tanila H (2005) Age-associated alterations of hippocampal place cells are subregion specific. J Neurosci Off J Soc Neurosci 25:6877–6886

135. Oomen CA et al (2013) The touchscreen operant platform for testing working memory and pattern separation in rats and mice. Nat Protoc 8:2006–2021
[https://doi.org/10.1038/nprot.2013.124][PubMed][PubMedCentral]

136. Reichelt AC et al (2021) The spontaneous location recognition task for assessing spatial pattern separation and memory across a delay in rats and mice. Nat Protoc 16:5616–5633
[https://doi.org/10.1038/s41596-021-00627-w][PubMed]

137. van Goethem NP, van Hagen BTJ, Prickaerts J (2018) Assessing spatial pattern separation in rodents using the object pattern separation task. Nat Protoc 13:1763–1792
[https://doi.org/10.1038/s41596-018-0013-x][PubMed]

138. van Hagen BTJ, van Goethem NP, Lagatta DC, Prickaerts J (2015) The object pattern separation (OPS) task: a behavioral paradigm derived from the object recognition task. Behav Brain Res 285:44–52
[PubMed]

139. Morris AM, Churchwell JC, Kesner RP, Gilbert PE (2012) Selective lesions of the dentate gyrus produce disruptions in place learning for adjacent spatial locations. Neurobiol Learn Mem 97:326–331
[PubMed][PubMedCentral]

140. Johnson SA et al (2016) Discrimination performance in aging is vulnerable to interference and dissociable from spatial memory. Learn Mem 23:339–348
[PubMed][PubMedCentral]

141. Burke SN, Wallace JL, Nematollahi S, Uprety AR, Barnes CA (2010) Pattern separation deficits may contribute to age-associated recognition impairments. Behav Neurosci 124:559–573
[PubMed][PubMedCentral]

142. Stark CEL, Noche JA, Ebersberger JR, Mayer L, Stark SM (2023) Optimizing the mnemonic similarity task for efficient, widespread use. Front Behav Neurosci 17:1080366

143. Johnson SA et al (2018) Experience-dependent effects of muscimol-induced hippocampal excitation on mnemonic discrimination. Front Syst Neurosci 12:72 [PubMed]

144. Burke SN, Turner SM, Desrosiers CL, Johnson SA, Maurer AP (2018) Perforant path fiber loss results in mnemonic discrimination task deficits in young rats. Front Syst Neurosci 12:61 [PubMed][PubMedCentral]

145. Prusky GT, Harker KT, Douglas RM, Whishaw IQ (2002) Variation in visual acuity within pigmented, and between pigmented and albino rat strains. Behav Brain Res 136:339–348 [PubMed]

146. Hickman DL, Swan M (2010) Use of a body condition score technique to assess health status in a rat model of polycystic kidney disease. J Am Assoc Lab Anim Sci. JAALAS 49:155–159 [PubMed][PubMedCentral]

© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2027
S. P. Curran (ed.), *Aging*, Methods in Molecular Biology 3105
https://doi.org/10.1007/978-1-0716-5703-4_20

20. Guide to Extending Findings from Selected Samples Using Inverse Odds of Selection Weights

Eleanor Hayes-Larson[1] ✉
(1) Leonard Davis School of Gerontology, University of Southern California, 3715 McClintock Ave., Los Angeles, CA 90089, USA

✉ **Eleanor Hayes-Larson**
Email: hayeslar@usc.edu

Abstract
When study samples, such as randomized controlled trial participants or individuals in biomarker studies, come from highly selected groups, results may not generalize to broader populations. Statistical methods called transportability methods enable extension of findings from such samples to specified target populations under certain assumptions. Here, we describe the implementation of one transportability method, inverse odds of selection weighting, to extend findings from selected trial samples. We discuss assumptions and data requirements, including harmonization and concatenation of trial and target datasets, and describe estimation, evaluation, and application of inverse odds of selection weights. Throughout, we highlight practical considerations and commonly encountered challenges.

Keywords External validity – Generalizability – Transportability – Inverse odds of selection weights – Selection bias

1 Introduction

In biomedical research, investigators typically collect data from a *study sample*—individuals who participate in a given study—but aim to draw conclusions about a broader *target population*, defined as the population to which the inferences are intended to apply [1]. When the study sample differs systematically from the target population, study findings may lack external validity, meaning they may not generalize beyond the study participants [1, 2].

This challenge arises across many areas of biomedical and aging research but is particularly salient in randomized controlled trials (RCTs) and in biomarker-based aging research, where individuals who are willing and able to enroll in trials and contribute biological samples may differ systematically from the populations ultimately targeted for intervention. For example, participants in many Alzheimer's disease and dementia studies are predominantly white, highly educated, and more likely to have a family history of dementia than the general population of older adults at risk for Alzheimer's disease [3–6].

Statistical approaches can account for differences between a study sample and a given target population, thereby improving the external validity of results. These methods, called generalizability and transportability methods, have been an area of rapid development in the epidemiologic and other literatures [2, 7–11]. Here, we use the term "transportability" to refer broadly to methods that extend inferences from a study sample to a target population. In some settings, transportability is distinguished from generalizability, where the latter refers to extending findings to the population from which the study sample was drawn. There are slight technical differences in the methods for generalizing versus transporting, but the goals and broader conceptual framework are shared [12].

In this chapter, we provide a practical, step-by-step guide to implementing a commonly used transportability method: inverse odds of selection weighting (overview in Fig. 1) [13]. We describe assumptions, data preparation, including harmonization and

concatenation of trial and target datasets, as well as estimation, evaluation, and final application of inverse odds of selection weights (IOSW). Throughout, we highlight practical considerations and common challenges encountered in applied settings.

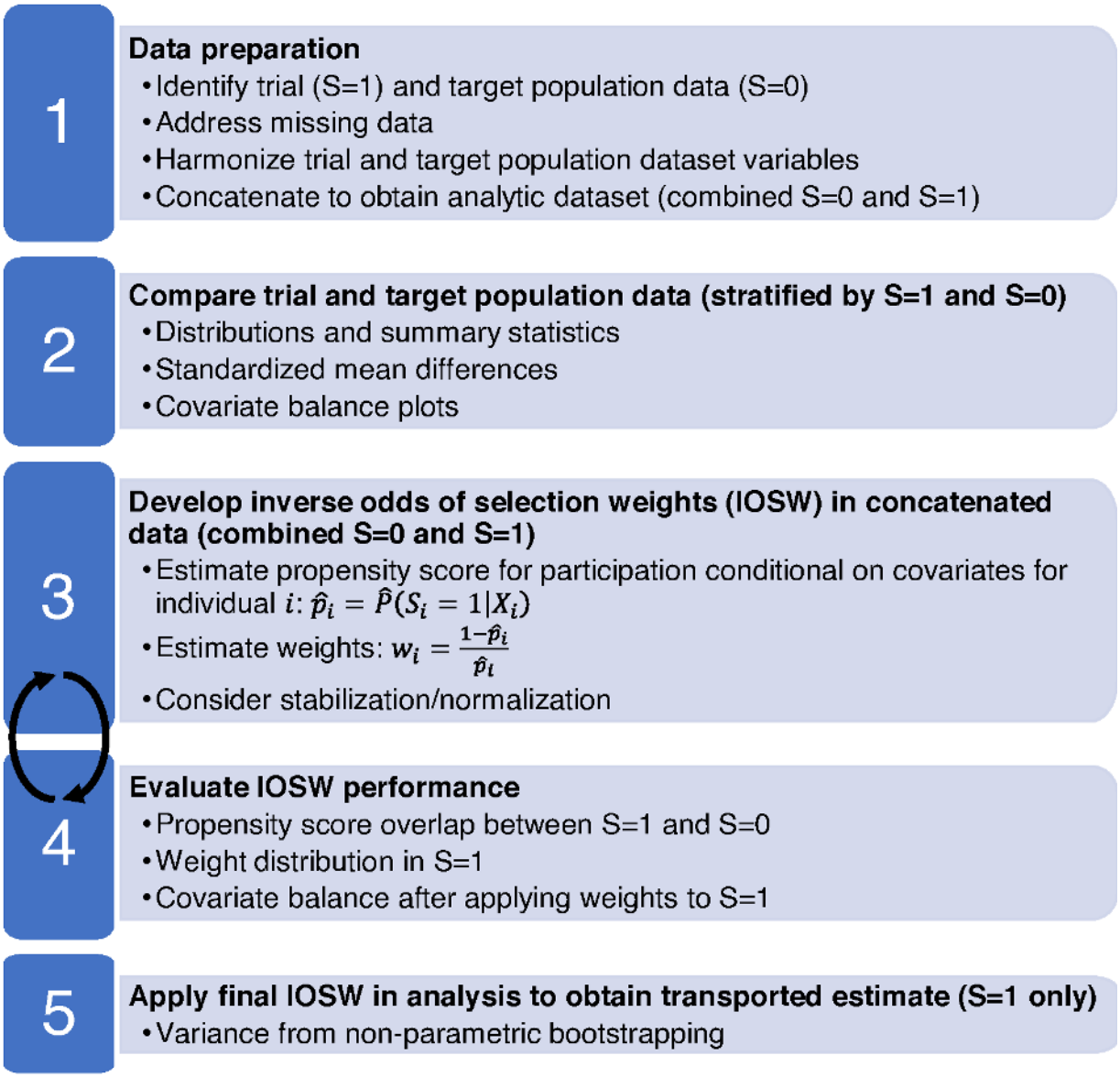


Fig. 1 Overview of data preparation and analysis steps in inverse odds of selection weighting procedure

2 Estimand and Assumptions

In this chapter, we define the estimand of interest as the average treatment effect (ATE) in the target population, that is, the difference in mean potential outcomes under treatment and control in the target population of individuals not in the trial:

$$\mathrm{ATE}_{\text{target}} = E\left[Y^1 - Y^0 \middle| S = 0\right]$$

where Y^a denotes the potential outcome under treatment assignment $a \in \{0,1\}$, and the expectation is taken over the distribution of covariates in the target population of trial non-participants ($S = 0$, with $S = 1$ indicating trial participants) [2].

To identify this estimand, several assumptions are required [2]. Although some violations can be detected in the data, the assumptions are not empirically testable and must be justified using subject-matter knowledge.

1. **Internal validity of the randomized trial**. Identification of the ATE in the trial is a prerequisite for transporting the ATE to the target population. We do not provide formal definitions here, as internal validity is expected in well-conducted randomized controlled trials [8].
2. **Conditional exchangeability across samples**. Formally, this is given by:

$$E\left[Y^1 - Y^0 | X = x, \quad S = 1\right] = E\left[Y^1 - Y^0 | X = x, \quad S = 0\right]$$

Informally, this assumption implies that after conditioning on baseline covariates X, the conditional ATE is the same in the trial and target population. Practically, this requires that all effect modifiers of the treatment that differ between the trial and target populations are measured and accounted for in the weights modeling selection (i.e., trial participation). Failure to accurately measure such effect modifiers, or measuring them with error, may result in biased transported estimates. Because effect modification is scale-dependent (additive vs. multiplicative) [14], this assumption must be evaluated with respect to the scale of the estimand (e.g., here, additive). Practically, directed acyclic graphs

(DAGs) or selection diagrams are often helpful for clarifying which variables are required [9, 11].

3. **Positivity of trial participation**. Formally, for all covariate values x with positive joint density in the target population $S = 0$,

$$P(S = 1, |, X = x) > 0$$

This means that every subgroup jointly defined by baseline covariates that is included in the target population must also be represented in the trial sample. If certain covariate patterns that occur in $S = 0$ do not appear in the $S = 1$ trial participants, the treatment effect for those subgroups is not identifiable and cannot be transported.

Violations of this assumption may be structural or nonstructural. Structural violations occur when certain subgroups in the target population are fundamentally ineligible for or excluded from the trial (e.g., due to age restrictions or comorbidities), making transport to those groups impossible without additional assumptions. Non-structural violations arise when overlap between the trial participants and target population is limited due to chance or design features, resulting in sparse representation of some subgroups in the trial.

4. **Consistency of potential outcomes**. For each individual, the observed outcome under the treatment actually received is equal to that individual's potential outcome under that treatment level. Formally, this requires that for individuals receiving treatment $A = a$:

$$Y = Y^a$$

For transportability analyses, this further requires that observed outcomes in the trial correspond to the outcomes that would occur under the same conditions in the target population. In other words, the intervention delivered in the trial must be the same intervention in the target population. This may seem tautological but could be violated if

there are trial engagement effects (trial participation influences the outcome, as in Hawthorne effects), or if treatment protocol, adherence patterns, or implementation context differ meaningfully between settings [2].

For other estimands (e.g., potential outcome mean in the target population, given by $E[Y^a|S=0]$, or an ATE in observational data), similar sets of assumptions are required [2, 9]. In addition, identification and estimation generally rely on assumptions of no measurement error, appropriate handling of missing data, and correct specification of statistical models [8].

---

3 Computational Considerations

All procedures described in this chapter assume access to a statistical software capable of data management and statistical modeling. These methods can generally be implemented in multiple programs (e.g., R, Stata, SAS, Python) with equivalent results, and specific software decision is secondary to maintaining a transparent and reproducible analytic process. Investigators should document data preparation and modeling decisions, software and package versions, and archive code and outputs to ensure that results can be reproduced.

---

4 Data Preparation

Transportability analyses require two distinct data sources: a trial sample, from which the original estimates are derived, and a target population dataset that represents the conceptual target population for which inferences are desired. In this section, we describe practical steps for defining, preparing, and aligning these data sources to facilitate subsequent analyses (Fig. 1**, Step 1**).

4.1 Trial Data

Identify the randomized trial from which the treatment effect will be transported. Apply any additional inclusion or exclusion criteria

necessary to define the analytic sample for the research question. It is not necessary to restrict the trial sample solely to mirror the target population (defined in Sect. 4.2), as individuals whose covariate patterns do not occur in the target population will receive weights approaching zero in estimation. However, any restrictions that define the causal estimand (e.g., age range or clinical eligibility relevant to the research question) should be applied. At this stage, retain all variables that may be useful for harmonization, construction of analytic covariates, or missing data imputation; restriction to the final analytic variable set should occur after harmonization across datasets.

4.2 Target Population Data

The target population is primarily a conceptual entity: the group for which inferences are desired. Target populations are specific to a given research question and are typically defined by person (individual characteristics like age, sex, race/ethnicity, and socioeconomic and clinical factors), place (geographical characteristics like catchment area, state, urban, or rural), and time (e.g., calendar years of interest). Even though the target population is defined conceptually, it is often necessary to consider practical factors such as analytic feasibility, particularly to avoid structural violations of the positivity assumption [2]. For example, if seeking to extend estimates from a trial sample that did not include any individuals with a history of cancer for a research question where cancer history is a relevant effect modifier, it may be necessary to redefine the target population to also exclude people with a history of cancer.

After the conceptual target population is identified, an empirical dataset must be selected to represent it, considering several factors. First, the dataset should be theoretically aligned with and representative of the conceptual target population. When using survey data for the target population dataset, representativeness may be achieved through the use of survey sampling weights, and all analyses should account for the survey design and incorporate sampling weights

so that the covariate distribution in the data reflects the intended target population.

The second factor to consider is whether the candidate dataset contains the variables required for the particular analysis. This typically includes baseline covariates that may modify the treatment effect and are therefore needed to achieve exchangeability across samples, as well as any variables needed to restrict the target population dataset to avoid structural positivity violations as described above. In settings involving selection on mediators (e.g., differential adherence), the target population dataset may also require measures of the mediators [9, 15]. Additionally, all required variables must be measured in a way that can be harmonized with the trial data (see Sect. 4.4). However, the target population dataset does not need to contain the treatment variable or outcome variable.

Finally, the target population dataset should have a sufficient sample size to enable modeling selection (trial participation) with good precision. Because estimation of IOSW relies on modeling trial participation as a function of covariates, sparse representation of key subgroups in the target population dataset may lead to unstable estimates.

Practically, there are often trade-offs between possible target population datasets, including in representativeness, variable availability, and sample sizes [3]. In some cases, a dataset with stronger conceptual alignment but fewer covariates may be preferable to one with broader variable availability but poorer representativeness. The variables that most strongly influence the estimand (i.e., strongest effect modifiers) are typically most important to have in the target population dataset [16].

Once a target population dataset has been selected, apply any restrictions that define the causal estimand and retain all variables that may be useful for harmonization, construction of analytic covariates, or missing data imputation, including any survey sampling weights. As with the trial data, restriction to the final analytic variable set should occur after harmonization across datasets.

4.3 Address Missing Data

Missing data within the trial dataset and the target population dataset should be handled separately, as the joint distributions of variables may vary across settings [17]. Standard missing data approaches can be used, although multiple imputation with chained equations is typically preferred [18]. Imputing on the original (pre-harmonized) variables within each dataset (versus variables after harmonization in Sect. 4.4) is standard, but there may be rare cases in which imputing a harmonized version of a variable is preferred (e.g., due to small cell sizes). If using multiple imputation, imputed datasets from the trial and target population samples can be randomly paired for the concatenation and analysis steps. Variance estimation accounting for multiple imputation should be implemented (Sect. 5.4) [19].

4.4 Data Harmonization

Harmonization refers to the process of ensuring that variables representing the same underlying constructs are defined, measured, and coded comparably across datasets [20]. In this case, harmonization across the trial and target population datasets is essential for valid transportability analyses because these methods use shared covariates to model the probability of trial participation. Inaccurate harmonization can induce measurement error that leads to residual imbalance and violations of the exchangeability assumption. Harmonization decisions should be guided by substantive knowledge, literature review, and potentially, consultation with data owners.

As described in Sect. 4.2, availability of key measures in both the trial and target population datasets can be a contributing factor in the choice of target population dataset. In this vein, generate a list of all constructs needed for transportability that are captured in both datasets. Importantly, at this stage, variables do not need to be measured identically. For example, the construct of “educational attainment” might be measured with a “years of education” variable in the trial but with a “highest degree” variable in the target population dataset.

Practically, the harmonizability of constructs across datasets can be ambiguous. For example, consider whether self-reported diagnosis in one dataset versus International Classification of Disease (ICD) diagnosis codes in another can be considered comparable constructs. For some conditions, or in populations with high health literacy, correspondence between self-report and ICD coding is high, and the two measures may be considered harmonizable; in other settings, such differences in ascertainment may mean constructs cannot be easily harmonized (e.g., [21, 22]). Similarly, self-reports and objectively measured variables (e.g., functional limitations) may have poor correspondence in some cases [23, 24]. Any differences can introduce differential measurement error in harmonized variables across datasets. Differential measurement error in effect modifiers can lead to unseen violations of exchangeability that may bias transported estimates; the direction and magnitude of this bias can be difficult to predict. Notably, the role of measurement error in transportability analyses is an area of ongoing research.

Once all harmonizable constructs have been enumerated, begin the harmonization to align the measurement of constructs between the trial and the target population datasets. Where possible, consider pre-statistical harmonization approaches, including categorizing continuous variables, recoding categorical variables to common categories, standardizing scales, or converting units of measurement [20]. When these approaches are not possible, statistical approaches, including regression modeling and use of item response theory, are alternatives [25, 26]. In some cases, external data can be used to develop appropriate "crosswalks" [27].

In the harmonization process, the practical considerations require careful evaluation of how specific decisions (e.g., category cutpoints or crosswalk rules) affect subgroup sizes, covariate distributions, and preservation of meaningful heterogeneity. For example, when categorizing variables, ensuring that resulting categories have sufficiently large sample sizes in both datasets to avoid positivity

violations can lead to collapsing across categories with substantively meaningful heterogeneity.

Analysts should assess whether alternative reasonable coding strategies materially affect covariate distributions or the resulting transported estimates in sensitivity analyses. Importantly, all harmonization decisions—including variable transformations, recoding rules, assumptions, and any constructs deemed not harmonizable—should be documented in a reproducible manner (e.g., annotated code, data dictionaries). Transparent documentation is particularly important when harmonization decisions require subjective judgment.

Once the harmonization is complete, ensure that the harmonized variables have identical names, coding schemes, and data types in both datasets to enable concatenation of the final analytic dataset.

4.5 Final Analytic Dataset

After harmonization, retain the final analytic variables in both the trial and target population datasets, as well as the treatment indicator and outcome variable in the trial dataset and any survey sampling weights in the target population data. Create a new variable called S that is defined as $S = 1$ for individuals in the trial dataset and $S = 0$ for individuals in the target population dataset.

The final analytic dataset is created by concatenating the harmonized trial and target population datasets (Fig. 2). The resulting structure should include rows for all individuals in both datasets, and columns for all harmonized variables, including the new S variable. In addition, there should be columns for the treatment and outcome variables, which will be missing in the target population dataset. Because the need for transportability arises due to the absence of a trial conducted in the target population, these missing values are expected. If survey sampling weights are used to preserve representativeness of the target population dataset, this variable will have a missing value for trial participants after the initial concatenation; these individuals should be assigned a sampling weight of 1 in the concatenated dataset so that the weights can be applied in the participation model (Sect. 5.2).

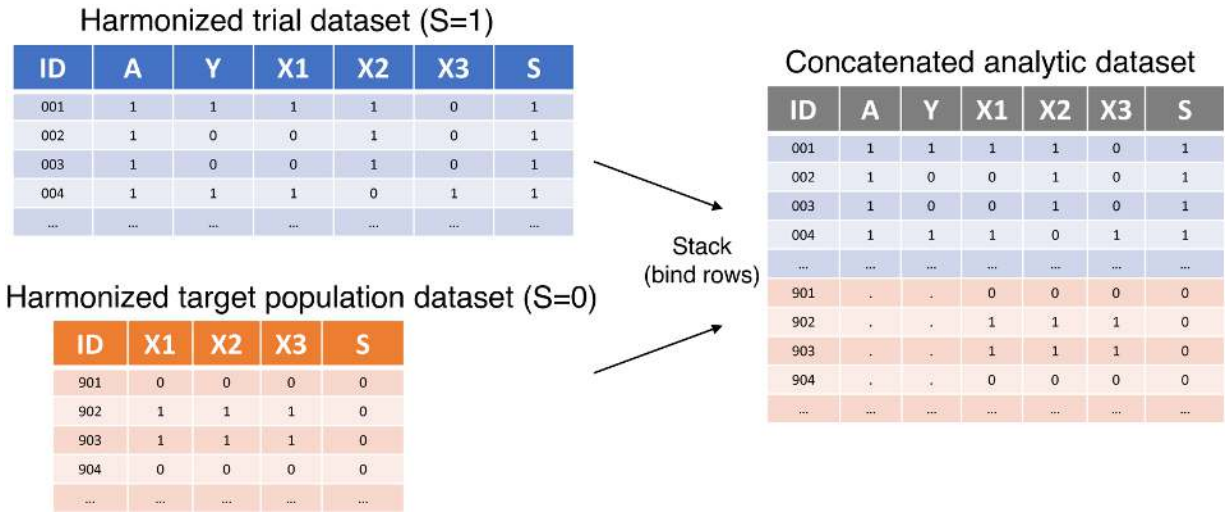


Fig. 2 Illustration of creation and structure of final analytic dataset via concatenation of harmonized datasets in the trial sample ($S = 1$) and target population data ($S = 2$). For simplicity, all baseline covariates are shown as binary variables, but in practice these harmonized variables may be of any type

5 Statistical Analysis

5.1 Describe and Compare Trial and Target Population Samples Prior to Transporting

Before constructing the IOSW, describe and compare the trial sample and the target population dataset (Fig. 1, **Step 2**). This essential step characterizes differences between the study and target population samples, which in turn informs the plausibility of assumptions (i.e., positivity and exchangeability) and choice of variables to be included in the participation model.

Conduct all comparisons on the concatenated dataset created in Sect. 4.5. If the target population dataset includes survey sampling weights, all summaries for $S = 0$ described in this section should incorporate those weights so that covariate distributions reflect the intended target population.

Start with a descriptive table with columns for both the trial and target population summarizing the distributions of baseline covariates in the two datasets [28]. An illustrative example for a limited variable set from simulated data is given in Table 1. Continuous variables may

be described using means and standard deviations (or medians and interquartile ranges, if skewed), and categorical variables using counts and proportions. Statistical testing of differences is not recommended, as small differences that do not threaten exchangeability may be significant in large samples, and meaningful differences not significant in small samples [28]. In addition to summary statistics, histograms, density plots, boxplots, or plots of empirical cumulative distribution functions may be used to visually compare distributions of continuous covariates.

Table 1 Illustrative example of a simple descriptive table comparing the trial sample to the target population data

Next, evaluate differences in covariate balance quantitatively. Standardized mean differences (SMDs) are most commonly used because they are scale-free and therefore facilitate comparison across variables [29]. For continuous variables, the SMD is typically calculated as the difference in means between the trial and the target population divided by a pooled standard deviation or the standard deviation in the target population dataset. Binary variables can be treated as if continuous, or compared with a simple difference in proportions. Visualization of differences, such as Love plots showing SMDs (e.g., Fig. 3, using the same simulated data as Table 1), is often very useful [30].

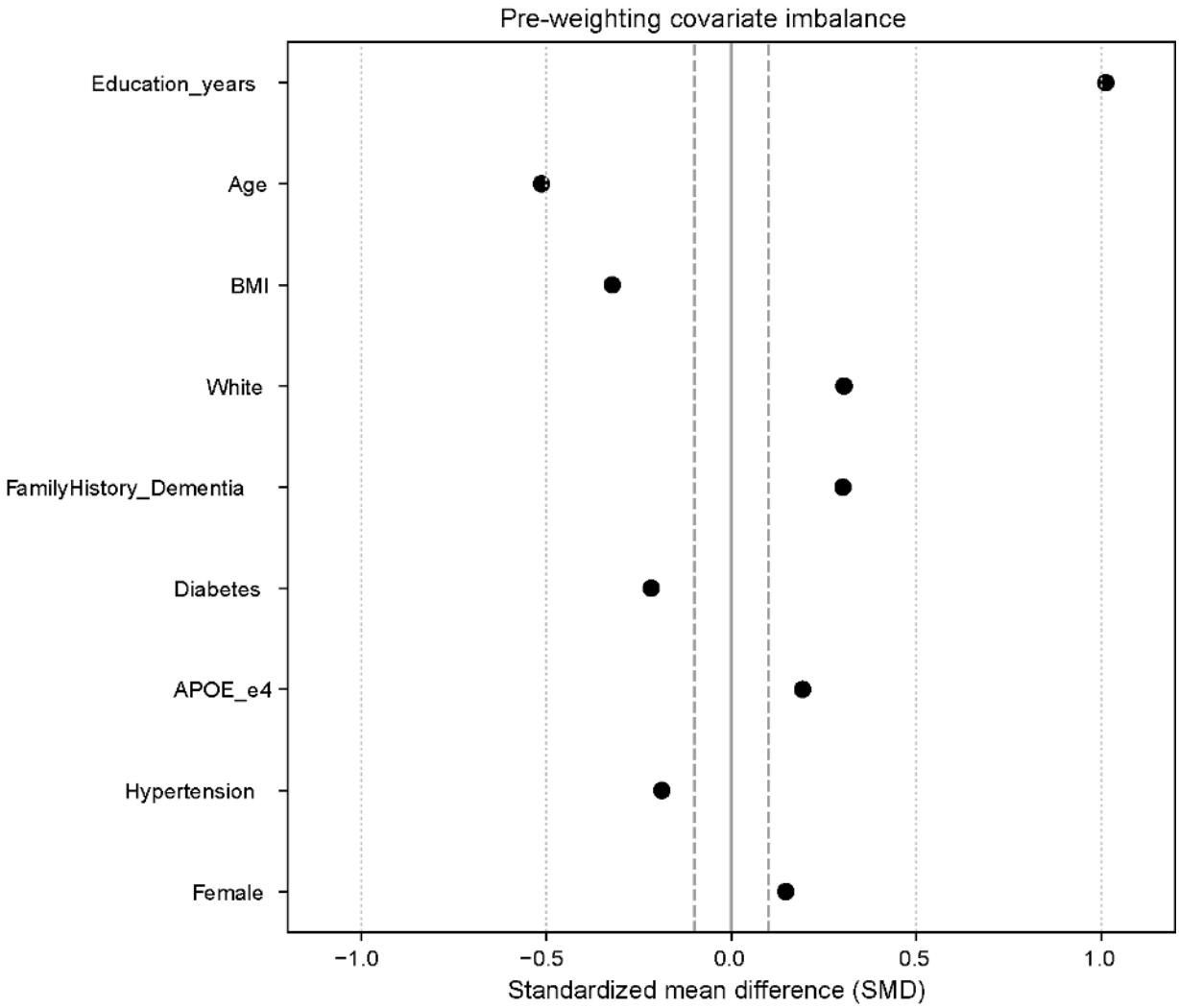


Fig. 3 Example of Love plot showing standardized mean differences comparing trial sample to target population data prior to inverse odds of selection weighting

Interpretation of these comparisons can inform the plausibility of the assumptions outlined in Sect. 2. Larger SMDs indicate greater divergence between the trial and target populations. Although there is no definitive threshold, SMD less than (0.1) are often considered indicative of minimal imbalance between the trial and target population [31], though such differences do not guarantee exchangeability. Conversely, large differences in covariate distributions may signal potential violations of the positivity assumption, especially if particular levels of a covariate represented in the target population dataset do not appear in the trial sample. When lack of overlap reflects

structural eligibility restrictions in the trial, redefining the target population or revisiting harmonization decisions may be necessary (Sect. 4).

Importantly, although SMDs and other summary metrics are useful for characterizing marginal differences between the trial and target populations, they do not guarantee that the assumptions for transportability can be satisfied. These comparisons assess balance one variable at a time, whereas the exchangeability assumption requires exchangeability over the joint distribution of all relevant covariates. Apparent balance in marginal distributions may mask imbalance in combinations of covariates (e.g., age by comorbidity interactions). In addition, measurement error in harmonized covariates can obscure true differences, creating the appearance of balance when underlying constructs differ across datasets, or apparent imbalance even when the samples are comparable. Finally, these diagnostics are limited to observed variables; differences in unmeasured variables cannot be detected and may threaten the validity of transported estimates. As such, descriptive and quantitative comparisons should be interpreted as informative regarding the plausibility of transportability assumptions, but not definitive.

In practice, substantial differences between the trial and target populations are expected in many applications and do not preclude transportability, provided there remains sufficient overlap in covariate space and the positivity assumption is not structurally violated. In these cases, the next step is to develop inverse odds of sampling weights to reweight the trial sample so that its covariate distribution more closely resembles that of the target population.

5.2 Developing Inverse Odds of Selection Weights (IOSW)

Conceptually, inverse odds of selection weighting reweights the trial sample so that it more closely mirrors the target population [13]. Intuitively, individuals who are underrepresented in the trial relative to the target population receive larger weights, while those who are overrepresented receive smaller weights. If the participation model is

correctly specified and the required assumptions hold, applying these weights yields a weighted trial sample that resembles the target population.

Estimating IOSW involves three steps: (1) fitting a model for trial participation in the concatenated dataset, (2) using that model to estimate the probability of trial participation for each individual, and (3) calculating inverse odds weights from the estimated probabilities. Each step is described below.

Fit Model for Trial Participation

First, fit a model for trial participation in the concatenated dataset (Fig. 1, **Step 3**). The simplest modeling approach is to use logistic regression:

$$\text{logit}\{P(S_i = 1, |, X_i)\} = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik}$$

where S_i is an indicator of trial participation for individual *i*, X_i is a vector of covariates measured in both the trial and target datasets, β_0 is the intercept, and β_1 to β_k are coefficients for covariates in the vector X_i (i.e., X_{i1} to X_{ik}). Note that alternatives to logistic regression, including machine learning methods, may improve model specification but can complicate inference (e.g., standard non-parametric bootstrapping for variance estimation may be invalid). Regardless of the model chosen, if the target population data includes survey sampling weights, the participation model should be fit using those weights to ensure that the target sample properly represents the underlying population.

The choice of variables to include in X_i in the participation model can be complex. To achieve exchangeability, weights should include all effect modifiers of treatment that differ between the trial and target population, which can be determined (for measured variables) from the Love plots and SMDs discussed in Sect. 5.1. In practice, it is most important to account for variables that strongly influence the estimand (here, the ATE) [31, 32]. To assess this, use domain knowledge and consider a Variable Importance for Treatment Transport plot, which shows effect sizes for trial participation on the X-axis and effect size for the outcome on the Y-axis [30]. From this plot, variables independently

associated with both participation and the outcome of interest should be retained for inclusion in weight development.

Practically, the best approach is often to start with a relatively simple model including a core set of variables, model trial participation, and evaluate weights as outlined below, and iteratively refine the model to inform the choice of a final optimal specification.

Estimated Predicted Probabilities of Trial Participation

From the fitted model, predict the probability of trial participation conditional on covariates for all individuals in the concatenated dataset (trial and target population). This estimated probability (often referred to as the propensity score for trial participation) is given by:

$$\widehat{p}_i = \widehat{P}(S_i = 1, |, X_i)$$

For example, when using logistic regression,

$$\widehat{p}_i = \frac{\exp\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{i1} + \widehat{\beta}_2 X_{i2} + \ldots + \widehat{\beta}_k X_{ik}\right)}{1 + \exp\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{i1} + \widehat{\beta}_2 X_{i2} + \ldots + \widehat{\beta}_k X_{ik}\right)}$$

Calculate Inverse Odds of Selection Weights (IOSW)

Finally, compute the IOSW for trial participants. The weight for individual *i* is given by the inverse odds of the probability of trial participation:

$$w_i = \frac{1 - \widehat{p}_i}{\widehat{p}_i}$$

Only trial participants receive these weights; because treatment and outcome data are not observed in the target dataset, target population participants do not contribute directly to the final weighted analysis, and therefore receive a weight of 0.

An optional step is to rescale the IOSW so that their mean equals one. This ensures that the weighted sample size corresponds approximately to the original trial sample size. Two approaches can be used, and importantly, both approaches rescale weights by constants and do not change the relative weighting across individuals.

Weights can be stabilized using the marginal odds of trial participation, where the stabilized weight for trial participant *i* is given by [13]:

$$w_i^{\text{stab}} = \frac{P(S=0)}{P(S=1)} \times w_i$$

Essentially, stabilization multiplies the inverse odds of participation by a constant derived from the marginal probability of trial participation (relative trial and target population sizes).

Weights can also be normalized using the average estimated weight, which yields a mean weight of exactly one:

$$w_i^{\text{norm}} = \frac{w_i}{w}$$

where w is the average IOSW w_i among trial participants. Normalization is a post-estimation rescaling step, rather than relying on the known marginal probability of trial participation. Normalization is particularly useful in situations where IOSW are combined with other weights (e.g., inverse probability of treatment weights used to adjust for confounding), as it helps ensure that the average final combined weight is equal to one.

5.3 Evaluate Weight Performance

After estimating the IOSW, evaluate the performance of the participation model and resulting IOSW (Fig. 1**, Step 4**). A number of diagnostics are relevant and help assess overlap, stability, and balance.

Propensity Score Overlap

Evaluate overlap in propensity score for trial participation to further diagnose positivity violations. Plot overlaid density plots of $\widehat{p}_i$ stratified by $S = 1$ and $S = 0$ (e.g., Fig. 4). Areas with substantial density in the target population ($S = 0$) but no density in the trial ($S = 1$) indicate potential positivity violations. Propensity scores near 0 can also indicate a potential positivity violation, although we note that when the target population is much larger than the trial, the probability of trial participation is always small. When there is poor overlap or other concerns about positivity violations, reconsider the target population

definition (Sect. 4.2), consider collapsing sparse categories in covariates (Sect. 4.4), or refine or simplify the participation model to improve specification.

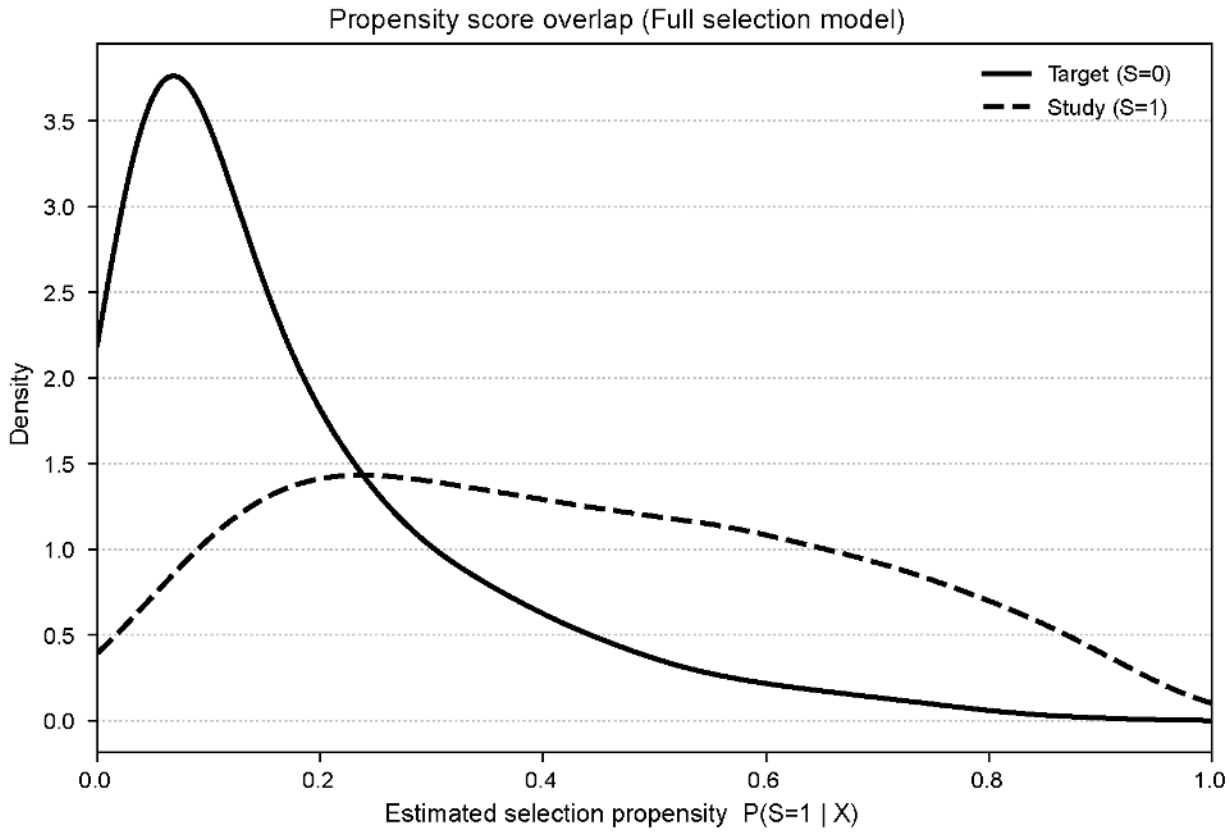


Fig. 4 Example of density plot showing overlap in propensity scores for trial participation comparing trial sample to target population data. The solid line represents the target population ($S = 0$) and the dashed line represents the trial sample. There is good overlap in density across the full range of propensity score values, although some propensities are close to zero, which could indicate positivity or model specification issues

Weight Distribution

Among trial participants only, examine the distribution of the IOSW using a histogram, density plot, and summary statistics such as mean, standard deviation, minimum and maximum, and percentiles (e.g., 1st, 10th, 25th, 50th, 75th, 90th, 99th). Excessive weight variability may reflect limited overlap (and potentially a positivity violation) or misspecification of the participation model. In particular, highly right-skewed distributions indicate instability and potential near-violation of

the positivity assumption, as large weights are generated by very small propensity scores.

Individuals with very large weights contribute disproportionately to the final estimates and may increase estimator variance. There is no universally accepted threshold for defining very large weights; for example, they can be defined relative to the median weight, using pre-specified thresholds (e.g., >10, >20), or using percentiles.

The preferred remedy for extreme weights is to adjust the participation model. Refine model specification by reconsidering functional form (e.g., nonlinearities) or inclusion of theoretically justified interactions. Remove any sparsely supported interaction terms, as these may increase instability. An alternative remedy is to trim (i.e., truncate) weights, by capping weights at set thresholds, e.g., 99th percentile or absolute weight of 10 [33]. Trimming weights is an ad-hoc approach that reduces variance but can introduce bias by altering the implicit target population [2]. Whenever trimming is applied, sensitivity analyses using alternative thresholds should be reported. In general, improving model specification is preferred to reliance on trimming alone [33].

Post-weighting Covariate Balance

A final diagnostic check is to re-evaluate covariate balance after applying the weights to the trial sample to check whether similarity between the trial and target population data on measured variables is improved. To reiterate, balance is not required for all covariates, only those that are true effect modifiers of the treatment. Comparisons like SMDs and Love plots outlined in Sect. 5.1 are useful, with the same caveats as discussed previously (marginal vs. joint balance, focus on means vs. distributions). Improved balance on measured variables not included in the participation model can improve confidence but do not provide definitive evidence regarding exchangeability on unmeasured variables. Sensitivity analyses can be used to evaluate robustness to unmeasured effect modifiers [34].

Important variables that remain imbalanced may indicate the need for refinement of the participation model. For example, Fig. 5 adds to Fig. 3, and shows covariate balance following application of IOSW derived from two different participation models. Using a model including age, sex, and education, imbalance remained on other variables even after applying IOSW; IOSW derived from a model with all covariates achieves good balance on all variables. This illustrates that initial models may reduce but not eliminate imbalance, and iterative refinement is often required (Fig. 1, **Steps 3 and 4**).

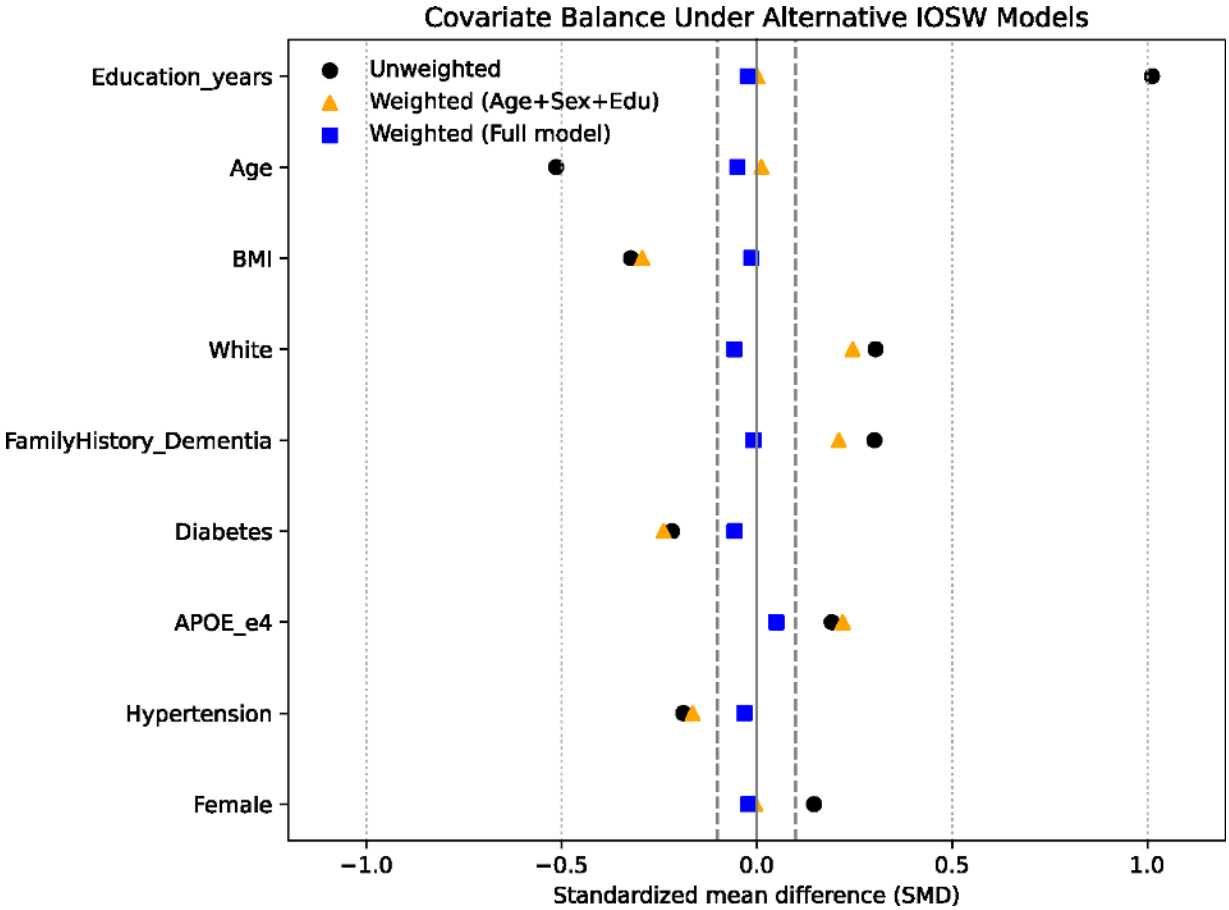


Fig. 5 Example of Love plot showing standardized mean differences comparing trial sample to target population data across multiple inverse odds of selection weighting (IOSW) models. Black circles indicate comparisons in the unweighted trial sample, as in Fig. 3. Orange triangles indicate comparisons applying IOSW based on participation models including age, sex, and education only. Blue squares indicate comparisons applying IOSW based on participation models including all covariates shown

When iterating participation models, key specification adjustments include considering alternate functional forms (e.g., splines, polynomials), adding or removing interactions, and adding or removing variables in the model. In many applications, there is not an obvious best model. For example, it may not be possible to achieve balance on all relevant variables without introducing extreme or unstable weights (e.g., when overlap is limited or sample sizes are small). In these situations, prioritize weights that yield balance on the strongest or most plausible effect modifiers and transparently report any remaining imbalance and its potential implications. Sensitivity analyses with alternative weight specifications can provide support for model choice and robustness of inferences.

5.4 Final Estimation

Once the final weights have been estimated, conduct the analysis to estimate the target population ATE in the weighted trial sample (Fig. 1, **Step 5**). Assuming correct specification of the participation model, and that other identification assumptions are met, the resulting estimate represents the estimand (ATE in the target population as defined in Sect. 2). As described previously, only trial participants contribute to this final analysis; target population participants, who receive final weights of 0, are used solely for estimation of the participation model.

Estimate variance using non-parametric bootstrapping, resampling from both the trial sample and target population dataset. Repeat the entire estimation procedure, including fitting the participation model based on the final chosen specification, weight computation, and outcome estimation, within each bootstrap. Ensure that variance estimates account for both bootstrapping and any multiple imputation procedures [19]. Additionally, when the target population comes from a complex survey sample, bootstrapping needs to account for the survey design [35]. However, when machine learning approaches are used to estimate the participation model, standard bootstrap procedures may not provide valid inference [36].

6 Summary and Conclusions

In this chapter, We summarized the implementation of inverse odds of selection weighting to transport findings from a randomized controlled trial to a target population. Transportability to strengthen external validity of study results is important in aging research, where study participants frequently differ systematically from the populations in which inferences are desired. We emphasized key steps: defining the target population, selecting and preparing an analytic dataset, and developing, evaluating, and implementing IOSW. Throughout, we highlighted tools to support transparent evaluation of assumptions and analytic feasibility. Because transportability relies on strong, untestable assumptions, analyses are often iterative, and sensitivity analyses are valuable at multiple stages.

Although this chapter focused on weighting approaches, alternative transportability methods, including outcome modeling and doubly robust estimation, are also available [2]. While the underlying identification assumptions are shared, because these alternative methods rely on different models, model misspecification may have different implications for bias than in the IOSW approach. As a result, presenting results from multiple methods—alongside clear diagnostics and sensitivity analyses—can allow for triangulation and strengthen the credibility of transported inferences.

Acknowledgments

The author thanks Suhani Amin for her assistance with manuscript preparation, including careful review of the full draft. ChatGPT 5.2 (OpenAI) was used to generate code to produce the simulated example data, table, and corresponding figures, and to edit portions of this manuscript for writing style and methodological clarity. The author verified all output and takes sole responsibility for all content.

Conflict of Interest

The author has no conflicts of interest to declare that are relevant to the content of this chapter.

References

1. Lesko CR, Ackerman B, Webster-Clark M et al (2020) Target validity: bringing treatment of external validity in line with internal validity. Curr Epidemiol Rep 7:117–124 [https://doi.org/10.1007/s40471-020-00239-0][PubMed][PubMedCentral]

2. Dahabreh IJ, Robertson SE, Steingrimsson JA et al (2020) Extending inferences from a randomized trial to a new target population. Stat Med 39:1999–2014 [https://doi.org/10.1002/sim.8426][PubMed]

3. Hayes-Larson E, Mobley TM, Mungas D et al (2022) Accounting for lack of representation in dementia research: generalizing KHANDLE study findings on the prevalence of cognitive impairment to the California older population. Alzheimers Dement 18:2209–2217 [https://doi.org/10.1002/alz.12522][PubMed][PubMedCentral]

4. Petersen RC, Aisen PS, Beckett LA et al (2010) Alzheimer's disease neuroimaging initiative (ADNI). Neurology 74:201–209 [https://doi.org/10.1212/WNL.0b013e3181cb3e25][PubMed]

5. Budd Haeberlein S, Aisen PS, Barkhof F et al (2022) Two randomized phase 3 studies of aducanumab in early Alzheimer's disease. J Prev Alzheimers Dis 9:197–210 [https://doi.org/10.14283/jpad.2022.30][PubMed]

6. van Dyck CH, Swanson CJ, Aisen P, et al (2023) Lecanemab in early Alzheimer's disease. N Engl J Med 388:9–21

7. Lesko CR, Buchanan AL, Westreich D et al (2017) Generalizing study results: a potential outcomes perspective. Epidemiology 28:553–561 [https://doi.org/10.1097/EDE.0000000000000664][PubMed][PubMedCentral]

8. Degtiar I, Rose S (2023) A review of generalizability and transportability. Annu Rev Stat Appl 10:501–524

9. Hayes-Larson E, Zhou Y, Rojas-Saunero LP et al (2024) Methods for extending inferences from observational studies: considering causal structures, identification assumptions, and estimators. Epidemiology 35:753–763 [https://doi.org/10.1097/EDE.0000000000001780][PubMed][PubMedCentral]

10. Ung L, VanderWeele TJ, Dahabreh IJ (2025) Generalizing and transporting causal inferences from randomized trials in the presence of trial engagement effects. Epidemiology 36:500–510 [https://doi.org/10.1097/EDE.0000000000001863][PubMed]

11. Pearl J, Bareinboim E (2014) External validity: from do-calculus to transportability across populations. Statist Sci 29:579–595

12. Dahabreh IJ, Robertson SE, Tchetgen EJ et al (2019) Generalizing causal inferences from individuals in randomized trials to all trial-eligible individuals. Biometrics 75:685–694 [https://doi.org/10.1111/biom.13009][PubMed][PubMedCentral]

13. Westreich D, Edwards JK, Lesko CR et al (2017) Transportability of trial results using inverse odds of sampling weights. Am J Epidemiol 186:1010–1014 [https://doi.org/10.1093/aje/kwx164][PubMed][PubMedCentral]

14. VanderWeele T (2015) Explanation in causal inference: methods for mediation and interaction. Oxford University Press, Oxford, New York

15. Rudolph KE, Laan MJ (2017) Robust estimation of encouragement-design intervention effects transported across sites. J R Stat Soc Series B Stat Methodol 79:1509–1525 [https://doi.org/10.1111/rssb.12213][PubMed]

16. Hayes-Larson E, Zhou Y, Wu Y et al (2025) Estimating dementia incidence in insured older Asian Americans and Pacific Islanders in California: an application of inverse odds of selection weights. Am J Epidemiol 194:1304–1313 [https://doi.org/10.1093/aje/kwae182][PubMed][PubMedCentral]

17. Mayer I, Josse J, Traumabase Group (2023) Generalizing treatment effects with incomplete covariates: identifying assumptions and multiple imputation algorithms. Biom J 65:e2100294 [https://doi.org/10.1002/bimj.202100294][PubMed]

18. van Buuren, S, 2018. Flexible Imputation of Missing Data, 2nd edn. Chapman and Hall/CRC, New York [https://doi.org/10.1201/9780429492259]

19. Schomaker M, Heumann C (2018) Bootstrap inference when using multiple imputation. Stat Med 37:2252–2266 [https://doi.org/10.1002/sim.7654][PubMed][PubMedCentral]

20. Tomescu-Dubrow I, Wolf C, Słomczyński KM (eds) et al (2023) Survey data harmonization in the social sciences. Wiley, Hoboken, New Jersey

21. Okura Y, Urban LH, Mahoney DW et al (2004) Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. J Clin Epidemiol 57:1096–1103 [https://doi.org/10.1016/j.jclinepi.2004.04.005][PubMed]

22. St.Clair P, Gaudette É, Zhao H et al (2017) Using self-reports or claims to assess disease prevalence: it's complicated. Med Care 55:782–788 [https://doi.org/10.1097/MLR.0000000000000753][PubMed][PubMedCentral]

23. Buta B, Zheng S, Langdon J et al (2022) Agreement between standard and self-reported assessments of physical frailty syndrome and its components in a registry of community-dwelling older adults. BMC Geriatr 22:705 [https://doi.org/10.1186/s12877-022-03376-x][PubMed][PubMedCentral]

24.

Wilfong JM, Badley EM, Power JD et al (2020) Discordance between self-reported and performance-based function among knee osteoarthritis surgical patients: variations by sex and obesity. PLoS One 15:e0236865 [https://doi.org/10.1371/journal.pone.0236865][PubMed][PubMedCentral]

25. Fortier I, Raina P, Van den Heuvel ER et al (2017) Maelstrom research guidelines for rigorous retrospective data harmonization. Int J Epidemiol 46:103–105 [PubMed][PubMedCentral]

26. Chan KS, Gross AL, Pezzin LE et al (2015) Harmonizing measures of cognitive performance across international surveys of aging using item response theory. J Aging Health 27:1392–1414 [https://doi.org/10.1177/0898264315583054][PubMed][PubMedCentral]

27. Wu Y, Hayes-Larson E, Zhou Y et al (2025) Statistical harmonization of versions of measures across studies using external data: self-rated health and self-rated memory. Ann Epidemiol 102:86–90 [https://doi.org/10.1016/j.annepidem.2025.01.002][PubMed][PubMedCentral]

28. Hayes-Larson E, Kezios KL, Mooney SJ et al (2019) Who is in this study, anyway? Guidelines for a useful Table 1. J Clin Epidemiol 114:125–132 [https://doi.org/10.1016/j.jclinepi.2019.06.011][PubMed][PubMedCentral]

29. Stuart EA (2010) Matching methods for causal inference: a review and a look forward. Stat Sci 25:1–21 [https://doi.org/10.1214/09-STS313][PubMed][PubMedCentral]

30. Lund JL, Webster-Clark MA, Westreich D et al (2024) Visualizing external validity: graphical displays to inform the extension of treatment effects from trials to clinical practice. Epidemiology 35:241–251 [https://doi.org/10.1097/EDE.0000000000001694][PubMed]

31. Austin PC (2009) Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med 28:3083–3107 [https://doi.org/10.1002/sim.3697][PubMed][PubMedCentral]

32. Ho DE, Imai K, King G et al (2007) Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Polit Anal 15:199–236

33. Lee BK, Lessler J, Stuart EA (2011) Weight trimming and propensity score weighting. PLoS One 6:e18174 [https://doi.org/10.1371/journal.pone.0018174][PubMed][PubMedCentral]

34. Nguyen TQ, Ackerman B, Schmid I et al (2018) Sensitivity analyses for effect modifiers not observed in the target population when generalizing treatment effects from a randomized controlled trial: assumptions, models, effect scales, data scenarios, and implementation details. PLoS One 13:e0208795 [https://doi.org/10.1371/journal.pone.0208795][PubMed][PubMedCentral]

35.

Ackerman B, Lesko CR, Siddique J et al (2021) Generalizing randomized trial findings to a target population using complex survey population data. Stat Med 40:1101–1120 [https://doi.org/10.1002/sim.8822][PubMed]

36. Chernozhukov V, Chetverikov D, Demirer M et al (2018) Double/debiased machine learning for treatment and structural parameters. J Econom 21:C1–C68

Component	Amount
Size-selected DNA (from Sect. 3.9.3)	15 µL
NEBNext Ultra II Q5 Master Mix*check for precipitation	25 µL
Universal PCR Primer P5 (10 µM)	5 µL
Indexing Primer with P7 (10 µM)	5 µL

Cycle step	Temperature	incubation time	Cycles
Initial denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	As determined in Sect. 3.9.4
Annealing/ Extension	65 °C	1 min 15 s	
Final extension	65 °C	5 min	1
Hold	4 °C	∞	

Component	Amount
10× reaction buffer	4 µL
ATP (10 mM)	4 µL
RiboLock RNase inhibitor	1 µL
Nuclease-free water	1 µL
50% PEG 8000	8 µL
Vortex to mix	
T4 RNA Ligase 1 (10 U/µL)	2 µL
Mix gently by pipetting or flicking the tube. Do not Vortex.	

Component	Amount
NEBNext Ultra II End Prep Reaction Buffer *Check buffer for precipitation	7 μL
NEBNext Ultra II End Prep Enzyme Mix	3 μL

Component	Amount
NEBNext 10× second strand synthesis reaction buffer	8 µL
NEBNext second strand synthesis enzyme Mix	4 µL
Nuclease-free water	30 µL

Table 1 Dimensions and quantities of components of the lifespan machine

Component	Dimensions (H × W × D) (inches)	Quantity
Raspberry Pi w/case	1 × 3.5 × 2.125	2+
NAS	6.5 × 11 × 9.125	1
Scanner	6.825 × 19.5 × 12.125	1+
Desktop PC	Dependent on user desktop format	1
PC peripherals (display/keyboard/mouse)	Dependent on user peripherals	1

Temperature	Incubation time
20 °C	30 min
65 °C	30 min

Table 1 File structure and contents of snRNAseq FASTQ files

File name	Read type	Contents
***_I1_*.fastq.gz**	Index Read 1	Sample index for demultiplexing
***_R1_*.fastq.gz**	Read 1	Cell barcode + UMI (10× Genomics)
***_R2_*.fastq.gz**	Read 2	Transcript sequence (cDNA)

Table 1 Normalized cell morphology metrics across CLS

sample_day	condition	*cell number*	area_mean	area_sd	area_median	length_mean	length_sd	length_median	width_mean	width_sd
Day 1	not_treated	1649	1.059	0.244	1.026	1.047	0.213	0.999	1.015	0.091
Day 1	treated	2433	1.048	0.237	1.026	1.034	0.198	0.997	1.017	0.103
Day 2	not_treated	2535	1.094	0.254	1.066	1.060	0.218	1.018	1.035	0.095
Day 2	treated	2225	1.035	0.256	1.011	1.031	0.214	0.987	1.006	0.113
Day 3	not_treated	3131	1.157	0.272	1.134	1.064	0.202	1.029	1.089	0.122
Day 3	treated	4713	1.098	0.291	1.087	1.030	0.223	1.000	1.063	0.130
Day 4	not_treated	1317	1.019	0.273	1.015	1.004	0.200	0.976	1.013	0.138
Day 4	treated	1476	1.209	0.361	1.203	1.067	0.231	1.038	1.125	0.166
Day 5	not_treated	1359	0.938	0.298	0.921	0.964	0.212	0.948	0.970	0.166
Day 5	treated	1294	0.990	0.299	0.964	0.991	0.206	0.977	0.994	0.158
Day 6	not_treated	4477	0.899	0.298	0.873	0.955	0.217	0.928	0.941	0.180
Day 6	treated	4416	0.802	0.278	0.783	0.922	0.210	0.905	0.869	0.181
Day 7	not_treated	1361	0.918	0.278	0.895	0.969	0.227	0.942	0.950	0.138
Day 7	treated	1445	0.767	0.255	0.743	0.919	0.216	0.888	0.837	0.149
Day 8	not_treated	1188	1.000	0.271	0.967	1.000	0.217	0.960	1.000	0.147
Day 8	treated	1421	0.932	0.259	0.909	0.969	0.206	0.951	0.961	0.153
Day 9	not_treated	1168	0.913	0.249	0.902	0.957	0.212	0.920	0.961	0.135
Day 9	treated	437	0.868	0.289	0.866	0.960	0.354	0.914	0.915	0.174
Day 10	not_treated	2474	0.958	0.236	0.931	0.962	0.200	0.926	1.005	0.116
Day 10	treated	2833	0.914	0.232	0.892	0.947	0.199	0.916	0.972	0.113

Normalized cell morphology metrics derived from automated segmentation of phase-contrast images across the CLS time course. For each aging day and condition, mean, standard deviation, and median values for cell area, major axis length, and minor axis length are reported after normalization to Day 1 untreated controls. These normalized values enable cross-day and cross-condition comparisons while retaining consistency with the underlying distribution-based analyses. All measurements are derived from the same CLS cultures used for CFU determination and nuclear morphology assessment, enabling direct comparison across survival, morphology, and nuclear integrity datasets

Component	Amount
Enriched mRNA sample (from Sect. 3.4)	1.1 µg in varying volume (up to 85 µL)
10× reaction buffer	10 µL
RNase III enzyme (1 U/µL)	5 µL
Nuclease-free water	Up to 100 µL

Table 1 Strains referred to in this protocol

Strain name	Genotype	Source
RHS41	*uthSi7[myo-3p::LifeAct::mRuby::unc-54 3'UTR, Cbr-unc-119(+)] IV*.	Caenorhabditis Genetics Center (CGC)
RHS42	*uthSi10[col-19p::LifeAct::mRuby::unc54 3'UTR + Cbr-unc-119(+)] IV*.	Caenorhabditis Genetics Center (CGC)
RHS43	*uthSi13[gly-19p::LifeAct::mRuby::unc-54 3'UTR + Cbr-unc-119(+)] IV*.	Caenorhabditis Genetics Center (CGC)
RHS191	*uthSi17[myo-3p::MLS::GFP::unc-54 39UTR, cb-unc-119(+)] I; unc-119(ed3) III;*	Caenorhabditis Genetics Center (CGC)
AGD2883	*unc-119(ed3) III; uthSi90[myo-3p::MLS::mRuby::unc-54 39UTR cb-unc-119(+)] IV;*	Caenorhabditis Genetics Center (CGC)
RHS192	*uthSi83[col-19p:MLS::GFP(65 C)::unc-54 39UTR, cb-unc119(+)] I; unc-119(ed3) III;*	Caenorhabditis Genetics Center (CGC)
RHS193	*uthSi84[vha-6p::MLS::GFP(65 C)::unc-54 39UTR, cb-unc119(+)] I; unc-119(ed3) III;*	Caenorhabditis Genetics Center (CGC)
BUZ14	*sec-61.b(wbm3[GFP::sec61.b]) IV; ret-1(bug2[ret-1::mKate2])*	Burkewitz Lab, Vanderbilt University

Component	Amount
RNA sample (from Sect. 3.1)	100 µL
10x DNase I buffer	11 µL
DNase I (2 U/µL)	4 µL
RiboLock RNase inhibitor	2 µL

Make the priming reaction as follows: Component	**Amount**
dNTP (10 mM)	4 μL
Random Hexamer (50 μM)	4 μL
Nuclease-free water	9 μL

Component	Amount
5× First-strand buffer	8 µL
DTT (100 mM)	2 µL
SuperScript III Reverse Transcriptase (200 U/µL)	4 µL

Input DNA	Number of PCR cycles required for library prep, yielding 5–50 nM
1 µg	3–4
500 ng	4–5
100 ng	6–7
50 ng	7–8
10 ng	9–10
5 ng	10–11
1 ng	12–13
0.5 ng	14–15

Temperature	Incubation time
20 °C	10 min
42 °C	20 min

Table 2 Select genetically encoded mitochondria reporter strains in *C. elegans*

Strain	Genotype	Expression patterns	Reference
RHS191	*uthSi17[myo-3p::MLS::GFP::unc-54 3′UTR, cb-unc-119(+)] I; unc-119(ed3) III;*	Expressing MLS::GFP in the muscle mitochondrial matrix	Moehle et al. [118]
RHS192	*uthSi83[col-19p:MLS::GFP(65 C)::unc-54 3′UTR, cb-unc119(+)] I; unc-119(ed3) III;*	Expressing MLS::GFP in the hypodermis mitochondrial matrix	Moehle et al. [118]
RHS193	*uthSi84[vha-6p::MLS::GFP(65 C)::unc-54 3′UTR, cb-unc119(+)] I; unc-119(ed3) III;*	Expressing MLS::GFP in intestine mitochondrial matrix	Moehle et al. [118]
RHS41	*uthSi7[myo-3p::LifeAct::mRuby::unc-54 3′UTR, Cbr-unc-119(+)] IV.*	LifeAct::mRuby labels F-actin in body wall muscle cells	Higuchi-Sanabria et al. [15]
RHS42	*uthSi10[col-19p::LifeAct::mRuby::unc54 3′UTR + Cbr-unc-119(+)] IV.*	LifeAct::mRuby labels F-actin in hypodermis (late L4 and adulthood)	Higuchi-Sanabria et al. [15]
RHS43	*uthSi13[gly-19p::LifeAct::mRuby::unc-54 3′UTR + Cbr-unc-119(+)] IV.*	LifeAct::mRuby labels F-actin in intestinal cells	Higuchi-Sanabria et al., [15]
SJ4103	*zcIs14 [myo-3::GFP(mit)].*	Expressing GFP at high levels in mitochondria of body wall muscle	Valera-Alberni et al., [119]
WBM671	*wbmEx289[myo-3p::tomm-20 (aa1-49)::GFP::unc54 3′UTR]*	Expressing GFP localized to mitochondria outer membrane of body wall muscle	Valera-Alberni et al., [119]
WBM1231	*wbmIs97[eft-3p::tomm-20::GFP::unc-54 3′UTR, *wbmIs65]*	Expressing GFP in mitochondria in all tissues	Valera-Alberni et al., [119]
WBM1232	*wbmIs98[eft-3p::tomm-20::mCherry::unc-54 3′UTR, *wbmIs65]*	Expressing mCherry in mitochondria in all tissues	Valera-Alberni et al., [119]
WBM1444	*tomm-70(wbm81[tomm-70::GFP]) III.*	Expressing GFP in mitochondrial outer membrane	Valera-Alberni et al., [119]

Strain	Genotype	Expression patterns	Reference
WBM1688	*scpl-4(wbm108[TIMM-50::scarlet]) V*	Expressing wrmScarlet in mitochondrial inner membrane	Valera-Alberni et al., [119]
WBM1689	*tomm-70(wbm81[TOMM-70::GFP) III; scpl-4(wbm108[TIMM50::scarlet]) V*	Expressing GFP in mitochondrial outer membrane andexpressing wrmScarlet in mitochondrial inner membrane	Valera-Alberni et al., [119]
JJ2586	*cox-4(zu476[cox-4::eGFP::3xFLAG]) I*	Expressing GFP in mitochondria in all tissues	Raiders SA, et al., [120]
SD1550	*ccIs4251 [myo-3p::GFP::LacZ::NLS + myo-3p::mitochondrial GFP + dpy-20(+)] I.*	Expressing GFP in mitochondria in body wall muscle	Liu, X, et al., [121]
SJ4143	*zcIs17 (ges-1::GFP (mit)).*	Expressing GFP in mitochondria in intestinal cells	Benedetti, et al., [122]

Component	Amount
1:10 Diluted Adapter	2.5 µL
NEBNext Ultra II Ligation Master Mix	30 µL
NEBNext Ultra II Ligation Enhancer	1 µL

Table 2 Quality control metrics and threshold considerations for sc/snRNAseq preprocessing

Metric	Description	Threshold range	Range adjustment	
			Higher (more permissive)	**Lower (less permissive)**
nFeature_RNA	Genes detected per cell	~250–5000	To include rare/low-complexity cells	To exclude empty droplets/noisy cells
nCount_RNA	UMI counts per cell	~800–25,000	To keep potentially low-UMI biological cells (e.g., quiescent cells)	To remove low-complex or empty droplets and high-count doublets
percent.mito	Percent mitochondrial reads	~5%–10%	To retain cell types with intrinsically elevated mitochondrial RNA reflects normal biology rather than low-quality cells (e.g., heart cells)	To exclude stressed/damaged/dying cells, where elevated mitochondrial RNA is indicative of compromised cell integrity.

Table 1 The number of CpG sites required for epigenetic clocks: Horvath1, Hannum, PhenoAge, GrimAge1/GrimAge2, DunedinPace, and PhysAge

Epigenetic clock	# of CpG sites	List of CpG sites
Horvath1	353	https://link.springer.com/article/10.1186/gb-2013-14-10-r115#MOESM3
Hannum	71	https://www.sciencedirect.com/science/article/pii/S1097276512008933#mmc2
PhenoAge	513	https://pmc.ncbi.nlm.nih.gov/articles/PMC5940111/#SD2
GrimAge1/ GrimAge2	1030	https://pmc.ncbi.nlm.nih.gov/articles/PMC6366976/
DunedinPace	173	https://github.com/danbelsky/DunedinPACE (use getRequiredProbes() function)
PhysAge	1711	https://github.com/em-arpy/DNAm_PhysAge/blob/main/CpG%20lists%20and%20weights/HRS_cpg_training_means.txt

Table 3 Select genetically encoded ER reporter strains in *C. elegans*

Strain	Genotype	Expression patterns	Reference
Extrachromosomal expression array			
VK2664	*vkEx2664 [nhx-2p::CemOrange2::tram-1 + myo-2p::GFP]*	pan-ER (membrane), intestine-specific promoter	Thomas et al., [125]
VK2674	*vkEx2674 [nhx-2p::CemOrange2::pisy-1 + myo-2p::GFP]*	pan-ER (membrane), intestine-specific promoter	Thomas et al., [125]
Integrated array			
WH0327	*unc-119(ed3); ojIs23 [SP12::GFP + unc-119(+)]*	Pan-ER (membrane), germline-specific promoter	Poteryaev et al., [126]
OCF15	*unc-119(ed3); ocfIs2[pie-1 prom::mCherry::sp12::pie-1 3'UTR + unc119 (+)]*	Pan-ER (membrane), germline-specific promoter studies in embryos	Joseph-Strauss et al., [127]
VS25	*hjIs14 [vha-6p::GFP::SPCS-1) + unc-119(+)]*	Pan-ER (membrane), intestine-specific	Xu et al., [128]
AGD2192	*uthSi60[vha-6p::ER-signal-sequence::mRuby::HDEL::unc-54 UTR, cb-unc-119(+)] I; unc-119(ed3) III*	Pan-ER (lumen), intestine-specific	Daniele et al., [95]
Endogenous fusion protein			
BQ5	*trap-1(dp674[trap-1::mCherry]) I*	ER sheet-enriched, including nuclear envelope (membrane), ubiquitous. Accumulates in lysosomes with age	Li et al., [129]
WBM875	*sec-61.b(wbm3[GFP::sec61.b]) IV*	ER sheet-enriched, stronger labeling of nuclear envelope (membrane), ubiquitous	Kelley et al., [130]
BUZ12	*ret-1(bug1[ret-1::GFP]) V*	ER tubule-enriched, no nuclear envelope (membrane), ubiquitous	Donahue et al., [19]
BUZ13	*ret-1(bug2[ret-1::mKate2]) V*	ER tubule-enriched, some ER sheet labeling, no nuclear envelope (membrane), ubiquitous	Donahue et al., [19]
RD412	*ret-1(pp125)[GFP::ret-1c/d] V*	ER tubule-enriched, some ER sheet labeling, no nuclear envelope (membrane), isoform-specific tag expressed preferentially in the nervous system	Serot et al., [131]

Strain	Genotype	Expression patterns	Reference
PHX4377	*hsp-3(hsp-3::wrmScarlet) X*	Pan-ER (lumen). Accumulates in lysosomes with age	Urban et al., [104]
PHX4415	*hsp-4 (hsp-4::wrmScarlet) II*	Pan-ER (lumen), ubiquitous but at lower basal levels than HSP-3.Accumulates in lysosomes with age	Urban et al., [104]

Table 2 Control all N2 automated lifespan summary and statistics

Strain group (Fig. 5a)	**Median survival (days)**	**Number of worms**	**Logrank (Mantel–Cox) test *P*-value with Group N2-A**
N2-A	17.7297	185	N/A
N2-B	17.5318	195	0.3618
N2-C	18.0005	204	0.0028
N2-D	18.0943	210	0.0034

Table 1 Illustrative example of a simple descriptive table comparing the trial sample to the target population data

Variable [mean, (SD), or *n* (%)]	**Trial sample (*S* = 1)**	**Target population data (*S* = 0)**
	***N* = 1500**	***N* = 5000**
Age, years	68.9 (6.0)	72.0 (6.0)
Education, years	16.0 (2.0)	14.0 (2.0)
BMI, kg/m^2	26.5 (3.3)	27.6 (3.5)
Female	942 (62.8%)	2780 (55.6%)
White	1241 (82.7%)	3497 (69.9%)
APOE ε4 carrier	489 (32.6%)	1199 (24.0%)
Hypertension	532 (35.5%)	2232 (44.6%)
Diabetes	165 (11.0%)	933 (18.7%)
Family history of dementia	525 (35.0%)	1079 (21.6%)

Component	Amount
Size-selected DNA (from Sect. 3.9.3)	15 µL
NEBNext Ultra II Q5 Master Mix *check for precipitation	25 µL
Index Primer Mix (10 –M)	10 µL

Table 1 All reagents and equipment needed from ACDBio for RNAscope

Reagent/Item name	Catalog number	Volume
Target probes		
RNAscope 50× Target Probe—apoeb (C1)	1293241-C1	3 mL (Ready-to-use)
RNAscope 50× Target Probe—mpz (C2)	1235131-C2	60 µL
RNAscope 50× Target Probe—s100b (C3)	1235141-C3	60 µL
RNAscope 50X Target Probe—map2 (C4)	1272201-C4	60 µL
RNAscope Target Probe—optn (C1)	1556591-C1	3 mL (Ready-to-use)
RNAscope 50× Target Probe—irf3 (C2)	1556631-C2	60 µL
RNAscope 50× Target Probe—stat1a (C3)	1556581-C3	60 µL
3-Plex Negative Control Probe (dapB)	320871	3 mL
4-Plex Negative Control Probe (dapB)	321831	3 mL
Bioinformatics Service Fee (Probe Design)	320269	Service
Pretreatment reagents		
RNAscope Hydrogen Peroxide	322335	3 mL
RNAscope Protease III	322337	3 mL
RNAscope 10× Target Retrieval Solution	322000	70 mL × 4 bottles
RNAscope Wash Buffer (50×)	310091	4 × 60 mL bottles
4% paraformaldehyde in PBS		200 mL
ImmEdge Hydrophobic Barrier Pen (Vector Labs)	H-4000	
Detection and amplification		
RNAscope Multiplex Fluorescent Detection Reagents v2	323110	3 mL bottles (Amp1–3, HRP-C1–3, HRP Blocker ×3)
RNAscope 4-Plex Ancillary Kit	323120	Contains HRP-C4 + HRP Blocker
RNAscope Multiplex v2 DAPI	323108	3 mL
Fluorophores		

Reagent/Item name	Catalog number	Volume
TSA Buffer Pack	322809	29 mL
TSA Vivid Fluorophore 520	323271	100 µL
TSA Vivid Fluorophore 570	323272	100 µL
TSA Vivid Fluorophore 650	323273	100 µL
Akoya Biosciences Opal Dye 620 (order via Akoya)	FP1495001KT	100 µL
Akoya Biosciences Opal Dye 690 (order via Akoya)	FP1497001KT	100 µL
Hybridization equipment		
HybEZ II Hybridization Oven	321711	
Humidity Control Tray	310012	
HybEZ Humidifying Paper	310015	15 sheets
ACD EZ-Batch Slide Holder	321716	
ACD EZ-Batch Wash Tray	321717	