Theoretical Biology

Philip K. Maini
Hiraku Nishimori
Toshio Sekimura
Tomohiko Yamaguchi *Editors*

Patterns and Forms in Nature

From Organisms to Collective Behaviors in Biological Systems

OPEN ACCESS

Springer

Theoretical Biology

Series Editor

Yoh Iwasa, Kyushu University
Fukuoka, Japan

The "Theoretical Biology" series publishes volumes on all aspects of life sciences research for which a mathematical or computational approach can offer the appropriate methods to deepen our knowledge and insight.

Topics covered include: cell and molecular biology, genetics, developmental biology, evolutionary biology, behavior sciences, species diversity, population ecology, chronobiology, bioinformatics, immunology, neuroscience, agricultural science, and medicine.

The main focus of the series is on the biological phenomena whereas mathematics or informatics contribute the adequate tools. Target audience is researchers and graduate students in biology and other related areas who are interested in using mathematical techniques or computer simulations to understand biological processes and mathematicians who want to learn what are the questions biologists like to know using diverse mathematical tools.

Philip K. Maini • Hiraku Nishimori
Toshio Sekimura • Tomohiko Yamaguchi
Editors

Patterns and Forms in Nature

From Organisms to Collective Behaviors in Biological Systems

Springer

Editors
Philip K. Maini
Mathematical Institute
Wolfson Centre for Mathematical Biology
Oxford, UK

Toshio Sekimura
Chubu University
Kasugai, Aichi, Japan

Hiraku Nishimori
Meiji Institute for Advanced Study
of Mathematical Sciences
Meiji University
Tokyo, Japan

Tomohiko Yamaguchi
Meiji Institute for Advanced Study
of Mathematical Sciences
Meiji University
Tokyo, Japan



ISSN 2522-0438 ISSN 2522-0446 (electronic)
Theoretical Biology
ISBN 978-981-92-0330-7 ISBN 978-981-92-0331-4 (eBook)
https://doi.org/10.1007/978-981-92-0331-4

This work was supported by Meiji University and Yamada Science Foundation.

© The Editor(s) (if applicable) and The Author(s) 2026. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this book or parts of it.
The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
This work is subject to copyright. All commercial rights are reserved by the author(s), 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. Regarding these commercial rights a non-exclusive license has been granted to the publisher.
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 Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

Preface

This book, titled *Patterns and Forms in Nature: From Organisms to Collective Behaviors in Biological Systems*, addresses the issues of patterns and forms in biological phenomena from various perspectives—genetics, biology, chemistry, physics, and mathematical sciences. Each chapter is written by scientists who describe the philosophy of their research and future perspectives as well as respective findings. Readers might wish to attend and listen to their talks if all the senior authors were to gather in one place. In fact, they were the invited speakers of the 6th Yamada Symposium held in Tokyo in 2023. It goes without saying that the discussions at the venue were lively and insightful.

This book is based on 10 of the 12 lectures delivered at the symposium and is divided into 3 parts. Abstracts of all the lectures are included in the Appendix. It is our hope that this book will serve as a guide for exploring the diverse patterns and forms observed in living organisms.

The timing of this publication is not unrelated to the fact that human activity is increasingly beginning to surpass nature, so let us briefly explain the background.

The world around us is filled with patterns and forms of various sizes. These extend to artificial objects that once did not exist in nature. While being surrounded by man-made objects has made our lives more convenient, it has also brought about environmental changes that leave significant marks reaching deep into the Earth's strata, such as global warming, microplastic pollution, the accumulation of radioactive substances, and loss of biodiversity.

The loss of biodiversity clearly has negative effects not only on the economy and environmental services but also on our mental activities. Since ancient times, humans have been deeply fascinated by the patterns and forms found in diverse living beings. From this allure, countless forms of art have emerged, and the intellectual curiosity to understand how these patterns and shapes arise has given birth to numerous fields of study. Biodiversity has provided strong motivation for these intellectual endeavors. Humans are characterized by their intellectual activities; therefore, biodiversity must not be lost.

Now, let us briefly introduce each chapter, beginning with Part I: Diversity of Biological Patterns and Forms in Nature:

– Dr. Ball approaches the issues of patterns and forms in nature from a physical perspective of symmetry, writing in a lively and engaging style. Additionally, the significance of the work of Alan Turing, known for his contributions to the mathematical study of biological pattern formation, is considered within a broad historical scientific context, ranging from Aristotle to the present day.
– Dr. Maini and his colleagues review how the synthesis of experimental and theoretical studies has advanced our understanding of collective neural crest migration, and offer an outlook on how these approaches may be further combined to shed light on the mechanisms that underlie collective cell migration. The neural crest serves as an important paradigmatic system for understanding collective cell migration across a range of application areas.

Then there are 3 chapters on butterfly patterns. Two of these chapters address the issue of female-limited mimicry (Batesian mimicry): one deals with genetics, and the other with population dynamics. The third focuses on tracing backward the evolution of wing patterns through numerical simulations.

– Dr. Fujiwara describes a detailed comparative study of the genomic regions involved in Batesian mimicry. Summarizing recent insights into the genetic mechanisms behind the formation of mimetic wing patterns, he discusses the evolutionary processes that produced this remarkable mimicry.
– Dr. Sekimura addresses the mimicry of the butterfly *P. polytes*, which inhabits the Sakishima Islands located at the southwestern end of Japan, and explains the advantages of mimicry using a mathematical model. Butterfly mimicry is a survival strategy in which one species reduces its predation rate by imitating the patterns of another unpalatable species. In this case, only one of the two female forms exhibits mimicry, known as Batesian mimicry. Sekimura notes, "Batesian mimicry is one of the most fascinating biological phenomena in the natural world."
– Dr. Nijhout and Dr. Gawne present research on exploring the "primitive patterns" in the evolution of butterfly wing patterns. This is an intriguing study that vividly demonstrates the extensive patterning potential of reaction-diffusion equations.

Part II: Morphology and Evolution, is composed of 3 chapters:

– Dr. Kuroda discusses the factors that determine the right and left sides of organisms, based on detailed research at the genetic level. Surprisingly, it was found that the shell coiling of the freshwater snail *Lymnaea stagnalis* is determined by a single gene. However, genes are not everything. Even more remarkably, it has been demonstrated that by mechanically reversing the direction of spiral cleavage at the third cell division when chirality first appears in the fertilized egg, a mirror-image organism can be produced! Genes are important, but are not whole story.
– Dr. Gerber presents the functional morphological space approach from a mathematical and analytical perspective. He provides an overview of the geometric structures of various morphological spaces found in the literature and discusses

mathematical structures that make it possible to evolutionarily consider sets of morphological variations.
– Dr. Ubukata applies a functional morphological space approach to data on the shapes and functions of ancient mollusks with shells. The focus is on shell construction efficiency, hydrodynamic properties in cephalopods, postural stability in gastropods, and the collision of swollen structures in bivalves, etc., discussing the major evolutionary trajectories of mollusk groups.

Finally, we have Part III Modeling Biomorphs in Nature, which contains 2 chapters:

– Dr. Shima examines plant shapes and functions from a mechanical viewpoint. Focusing on the unique forms of bamboo, butterbur, rice, and melon meshes, he discusses mathematically the relation between the forms and the function (i.e., the mechanical strength).
– Dr. Steinbock and his colleagues introduce the chemical world, which also touches upon the problem of the origins of life. From a chemists' perspective, they discuss phenomena such as the growth of inorganic semipermeable membranes (chemical gardens), slowly rotating chemical spirals (the BZ reactions), corrosion patterns and precipitation patterns on metal surfaces, and inorganic structures that have attracted attention for their life-like forms. As Dr. Ball also points out, the "forms" often considered characteristic of life are not necessarily unique to living organisms.

Readers may start reading from any chapter. We hope this book will deepen our understanding of patterns, forms, and a broad range of life phenomena.

Oxford, UK — Philip K. Maini
Tokyo, Japan — Hiraku Nishimori
Kasugai, Aichi, Japan — Toshio Sekimura
Tokyo, Japan — Tomohiko Yamaguchi

Contents

Part I Diversity of Biological Patterns and Forms in Nature

1 Symmetry Breaking from Physics to Biology 3
Philip Ball

2 Integrating Mathematical Modeling and Experiments in the Study of Collective Neural Crest Cell Migration 21
Samuel W. S. Johnson, Ruth E. Baker, and Philip K. Maini

3 Mechanisms and Evolution of Color Pattern Formation in Female-Limited Batesian Mimicry of *Papilio* Butterflies 43
Haruhiko Fujiwara

4 A Model for the Population Dynamics of the Mimetic Butterfly *Papilio polytes* in the Sakishima Islands, Japan: Mathematical Analysis and Computer Simulations 63
Toshio Sekimura

5 The Development and Evolution of Butterfly Wing Patterns: Finding the Primitive Pattern 77
H. Frederik Nijhout and Richard Gawne

Part II Morphology and Evolution

6 Snail Coiling: CRISPR Editing of a Single Gene Turns Righties into Lefties 93
Reiko Kuroda

7 Morphospace Approaches to Evolution: Geometric and Developmental Considerations 115
Sylvain Gerber

8 Functional Morphospace Analysis of Molluscan Shells 135
Takao Ubukata

Part III Modeling Biomorphs in Nature

9 Mechanical Optimality Hidden in the Structure of Plants 155
Hiroyuki Shima

10 Spiral Patterns and Morphology in Chemical Systems 175
Oliver Steinbock, Pamela Knoll, and Amrutha S.V.

Appendix (*Courtesy of the Yamada Science Foundation*) 191

Part I
Diversity of Biological Patterns and Forms in Nature

Chapter 1
Symmetry Breaking from Physics to Biology



Philip Ball

Abstract In this chapter I discuss the concept of symmetry breaking across various fields, from physics to biology, and its role in the emergence of patterns and forms in nature.

Keywords Origin of symmetry breaking · Emergence of patterns and forms · Turing patterns · Ising model

1.1 Introduction

In 1952 the British mathematician Alan Turing, best known for his work on the theory of computation, published the only paper he ever wrote on the life sciences (Turing 1952). It outlined a startlingly novel theory of morphogenesis.

The question that Turing set out to tackle is one of the oldest and most fundamental in all of biology: Where does form come from? How do you go from a spherical embryo to an organism like a human with an extremely complex form?
It is a question that preoccupied natural philosophers from Aristotle to William Harvey to Rudolf Virchow and beyond (Needham 1959). In general, early ideas assumed that biological form was either already manifest in some way within the embryo—the preformationist idea, famously illustrated by Nicolaas Hartsoeker's drawing of the homunculus in a sperm in the seventeenth century—or it emerged via the guidance of a kind of life force or soul, as Aristotle asserted: the position called epigenesis, which was essentially teleological. We might now regard both theories as ways to circumvent rather than to explain what is now seen as a kind of symmetry breaking.

In his approach to the origins of embryogenesis, Turing was going against the grain. His work came at a time when embryology, arguably a key focus of biology at the end of the nineteenth century, had fallen out of fashion. Just 10 years earlier, Julian Huxley published his book that introduced the term the *Modern Synthesis* to

P. Ball (✉)
Freelance Writer, London, UK

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_1

describe the reconciliation between Darwinian evolution and Mendelian genetics: the advent of Neodarwinism (Huxley 1942). One year after Turing's paper appeared, of course, Francis Crick and James Watson took the emerging picture of genetics to the next level by revealing the chemical structure of DNA (Watson and Crick 1953). Questions about form in biology were becoming questions about genes: the phenotype was considered to be programmed into the developing organism in the genotype. This could be regarded as a kind of *informational homunculus*: form already exists in the embryo but in an encoded format, which Erwin Schrödinger famously called the "code-script" of life (Schrödinger 1944).

It is probably partly, although not entirely, because it did not fit into this paradigm that Turing's paper had little impact for several decades. In fact, Conrad Waddington, whose book on the genetics of biological patterning Turing cited in his very short reference list (Waddington 1940), complained that development had been left out of the Modern Synthesis entirely. In the 1940s, Waddington introduced his famous epigenetic landscape (Allen 2015) (Fig. 1.1), the classic representation of symmetry breaking in biology, specifically to try to reintroduce developmental biology—as embryology was then becoming called—into the picture. That attempt was not terribly successful.

Waddington was part of the Theoretical Biology Club at the University of Cambridge, an informal collection of innovative thinkers that also included J. Desmond Bernal, Joseph Needham and J. B. S. Haldane. All were sympathetic to the idea that to understand morphogenesis, biology might benefit from a reductive "physicist's" approach, rather than the purely statistical methods of classical genetics.

Their perspective was influenced by the work of the Scottish zoologist D'Arcy Wentworth Thompson, whose seminal book *On Growth and Form* in 1917 is another of Turing's references (Thompson 1919). Thompson called for a more mathematical approach to biology that took account of the physical forces and constraints on an organism's development. He was skeptical about how the resurgent Darwinism of his time seemed to answer every question about biological form by calling it adaptive, with little thought given to the actual physical processes by which it came

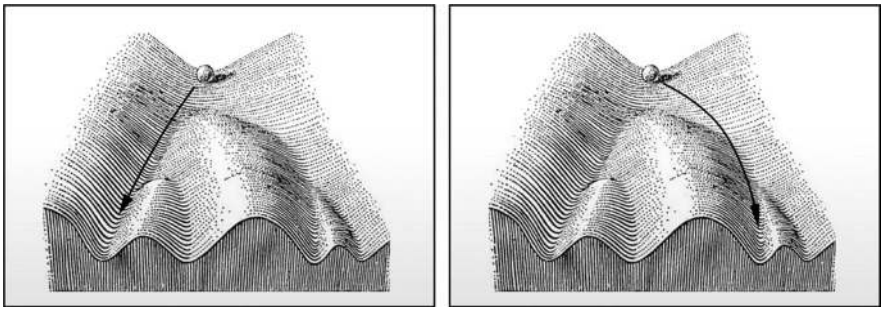


Fig. 1.1 The Waddington landscape, the paradigmatic representation of symmetry breaking in biology

into being. It was one thing to explain the shape of a limb or organism on the basis of how it enhanced fitness, but quite another to explain how it came to be shaped in the developing fetus.

Turing's treatment was highly mathematical and abstract. Mentioning that at one point "the cells are idealized into geometrical points," he then talks about Newton's laws of motion, about elasticities, and imaginary numbers, which doubtless left some biologists puzzled, or worse.

Yet his paper follows the physical scientist's instinct to express the problem at hand in the simplest possible terms, and then to explore what the minimal requirements are of a model that captures the essential phenomena. Turing points out that the question of how an embryo develops and acquires shape and form—bilateral symmetry, say, and the budding of limbs—is indeed one of *symmetry breaking*. How, he asks, can the embryonic blastula escape the spherical symmetry with which it starts, if all it has to work with is diffusion of its chemical ingredients? That, Turing wrote, certainly cannot result in an organism like a horse—which, he points out somewhat wryly, "is not spherically symmetrical."

1.2 The Origins of Symmetry Breaking

The notion of symmetry breaking was by that time already widely appreciated by anyone who was thinking about how form and differentiation emerge from uniformity. One sometimes sees suggestions that symmetry breaking as a concept was introduced into physics by particle physicists around the same time, in the 1950s. But this is not the case. On the contrary, high-energy physics at that time began to import ideas that had long existed in other branches of physics, and which Turing evidently had some inkling of.

Perhaps the first intimation of the role of symmetry in physical processes was expressed in 1894 by Pierre Curie in a paper titled "On symmetry in physical phenomena" (Curie 1894). Here he asked the question: What is the relationship between the physical and the symmetry properties of a system? Curie stated the principle of symmetry breaking by saying that some phenomena occur when a symmetry is *removed*: he said that "asymmetry is what creates the phenomenon". Where historians now disagree is whether Curie anticipated *spontaneous* symmetry breaking or, on the contrary, denied it by asserting that "When certain effects show a certain asymmetry, this asymmetry must be found in the causes which gave rise to it."

Curie's focus here was on the properties of crystals. In 1880 he and his brother Jacques discovered how symmetry breaking induced by pressure can cause piezoelectricity in some crystals. His doctorate in the early 1890s was on the subject of magnetism, and he found that ferromagnets such as iron will lose their magnetism when heated. The temperature at which the magnetism vanishes is now called the Curie temperature (Fig. 1.2a).

Spontaneous magnetization by cooling through the Curie temperature is a symmetry-breaking phenomenon: it is the point at which the random orientations of

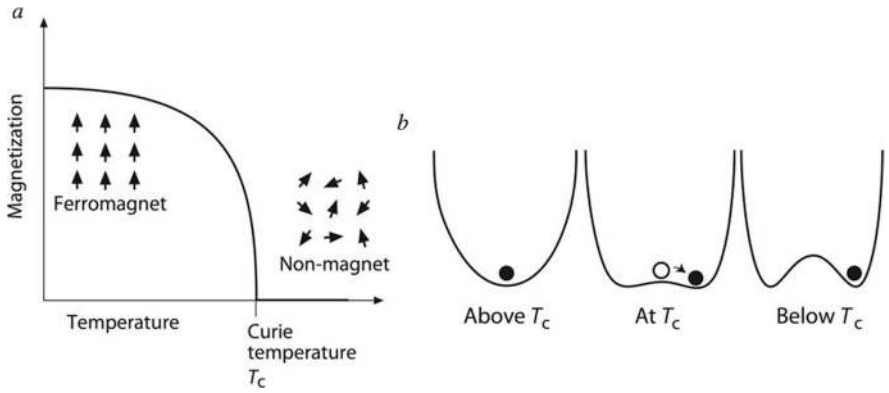


Fig. 1.2 The symmetry-breaking magnetization transition of a ferromagnet corresponds to the temperature at which the tendency of neighboring spins on magnetic atoms to align overcomes the disordering effect of thermal motion (**a**). It corresponds to the switch from a single state of minimal energy to a double well: a bifurcation (**b**)

atomic spins due to thermal motion switch into a state of mutual alignment. This process picks out one direction as "special" in a previously isotropic medium, and it is an example of a phase transition—specifically, a critical phase transition, which is driven by fluctuations and involves a divergence of the correlation length between the states of the components. The loss of symmetry to which Curie referred can be thought of as a bifurcation at the critical point at which a single stable state splits into two, so that the system is forced to make a choice of one or the other (Fig. 1.2b).

In 1920 the German physicist Wilhelm Lenz proposed a simple model for this phase transition of a ferromagnet. He imagined a regular array of spins—like the magnetic atoms in a crystal lattice—that can point just in one direction or the opposite (Fig. 1.3a). The magnetic interaction of neighboring spins will encourage them to align, so that they both point in the same direction. Lenz wondered under what conditions this tendency to align will dominate over the randomizing influence of heat and produce overall alignment throughout the lattice. In 1924 his student Ernst Ising figured out the answer for the simplest possible case: a one-dimensional lattice consisting just of a row of spins, with one neighbor on each side (Ising 1924). Somewhat to his disappointment, Ising found that the transition temperature in this case is zero: *any* amount of thermal disruption is enough to destroy the aligned magnetic state. For the two-dimensional model, Lars Onsager found an analytical solution in the 1940s that did show a finite-temperature transition (Onsager 1944). We know from simulations that a transition exists in 3D too, but there remains no analytical solution that describes it.

It is worth pointing out that Onsager's solution of the 2D Ising model provided the first rigorous, non-phenomenological indication that symmetry-breaking phase transitions are a genuine feature of many-body dynamics. That makes it a rather significant result—according to historian of science Stephen Brush, it was "one of the most important events in theoretical physics of the 20th century" (Brush 1983).

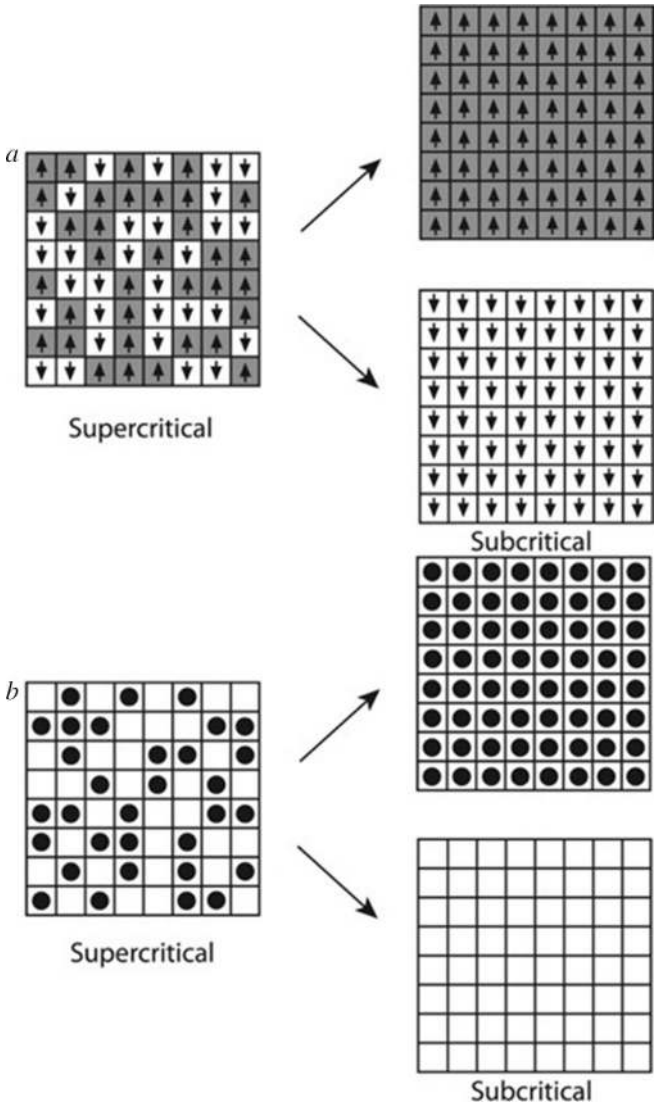


Fig. 1.3 (**a**) The Ising model of ferromagnets (here in two dimensions). (**b**) The same lattice model can represent the phase transitions of liquids and gases, where each lattice point is occupied in the liquid state and is empty in the gas state. Above the critical point, there is a patchy mixture of occupied and unoccupied grid elements

It was quickly appreciated that the Ising model could serve as a very crude model of the gas and the liquid states of matter. Here the presence of a particle at a point in the lattice denotes the liquid, while its absence denotes a gas (Fig. 1.3b). Cooling a magnet through its Curie temperature then looks equivalent to cooling a fluid through its critical point, which Thomas Andrews investigated in the 1860s

(Andrews 1869) and Johannes Diderik van der Waals explained phenomenologically in 1873 (Van de Waals 1988).

This connection was arguably the first glimpse of *universality* in physics: the idea that there is a deep parallel between systems that look superficially as though they share nothing in common. It stems from the fact that many-body behavior often does not discriminate between the fine differences in the natures of those bodies: all that matters are the general features of how they interact, such as how many neighbors feel the influence of each body. It is remarkable how deeply such universality runs. Despite the apparent crudeness with which the Ising model represents liquids and gases as filled or empty lattice elements, simulations for the 3D model show that the critical exponents that describe divergences of thermodynamic quantities at the critical point are essentially identical to those determined experimentally for real fluids (Stanley 1971).

These correspondences between the phase transitions of magnets, those of liquids and gases, and those in a rather simple lattice model, point to the deeper truth that many real-world phenomena are *collective* and have their own distinctive but often universal rules of self-organization.

1.3 Symmetry, Conservation, and Unification

The fact that the basic physical laws obey certain symmetry properties is *the* most important feature they display. For example, these laws remain unchanged from one part of space to another: they have translational symmetry. The law of gravitation is the same if measured in London, Tokyo, or on the Moon. In physics this invariance of a law is called a *gauge symmetry.* In one of the most profound insights in twentieth-century physics, the German mathematician Emmy Noether showed in 1915 that there is a deep connection between the existence of a symmetry in a physical system and the existence of a conservation law (Noether 1918). In short, Noether's theorem showed that if some quantity is conserved, there is a corresponding symmetry in the way the system behaves. For example, the fact that momentum is conserved is a consequence of the translational symmetry of space, while the conservation of energy turns out to be a result of the fact that mechanical laws have time symmetry. Many physicists consider Noether's theorem to be as deep and beautiful a discovery as Einstein's general relativity. Einstein himself said that "In the judgment of the most competent living mathematicians, Fräulein Noether was the most significant creative mathematical genius thus far produced since the higher education of women began."

It used to be thought that the physical laws are indifferent to the symmetry operation of mirror inversion. But in 1956 it was discovered that this is not always so: processes involving the weak interaction can violate this so-called principle of parity. (Incidentally the two Chinese American physicists who proposed this, Tsung Dao Lee and Chen Ning Yang, also studied phase transitions of the Ising model and coined the term "lattice gas".)

The experimental confirmation of parity violation in the weak nuclear interaction (the force responsible for beta decay) by Chien-Shiung Wu in 1956 (Wu et al. 1957) opened the path for the unification of the electromagnetic and weak interactions. The idea is that a spontaneous symmetry-breaking transition in the early universe led to the distinction between these two forces, rather as a supercritical fluid separates into two distinct states, the liquid and gas, as it passes through the critical point.

This idea was developed in the late 1950s and early 1960s by Sheldon Glashow and Abdus Salam (Glashow 1959; Salam and Ward 1964), who could only get symmetry breaking by introducing it by hand. They found that it predicted three massive new bosons, later identified with the W and Z bosons that mediate the weak force, along with the massless photon. In the later 1960s, Steven Weinberg showed how this symmetry breaking could happen spontaneously (Weinberg 1967). The theory was verified by the discovery of the W and Z bosons in CERN's Proton-Antiproton Collider in 1983 (Di Lella and Rubbia 2015).

This Nobel-winning work on electroweak symmetry breaking was intimately tied up with the theory of the Higgs field and the Higgs boson. The fundamental insight here came in the late 1950s from theoretical physicist Yoichiro Nambu, who saw an analogy between the theory of superconductivity and quantum field theory (Nambu and Jona-Lasinio 1961). Vitaly Ginzburg and Lev Landau devised a phenomenological model of superconductivity in the 1950s that presented it as another symmetry-breaking critical transition, which Bardeen, Copper, and Schrieffer (BCS) then developed into a fully-fledged theory, the BCS theory (Bardeen et al. 1957), involving electron–phonon interactions in 1957. Nambu realized that the BCS theory seemed to violate the gauge symmetry of quantum electrodynamics, but he solved the riddle by showing that here the symmetry-breaking transition gives rise to particular symmetry properties in the behavior of the electrons that become *hidden*. That is to say, the symmetry breaking is not obvious in the way the electrons behave. Nambu showed that the same kind of symmetry breaking of the quantum vacuum predicted that the particles of atomic nuclei—protons and neutrons—would gain mass.

The idea won Nambu the 2008 Nobel prize for "the discovery of the mechanism of spontaneous broken symmetry in subatomic physics." But his theory raised new problems—and it was by resolving them in the early 1960s that Peter Higgs and others figured out that space must be permeated by the Higgs field, which exerts a drag on the particles and so affords them mass (Close 2022). Higgs was the first to push the idea that this field should have an associated massive particle, the possible detection of which supplied a way to test the idea. The rest, of course, is history: the Higgs boson was discovered by the Large Hadron Collider at CERN in 2012 (ATLAS Collaboration 2012). Meanwhile, Weinberg used Nambu's ideas, and subsequently those of Higgs, to develop the 1967 model that ultimately led him to his theory of electroweak unification.

The key point here is the unity brought to many seemingly different areas of physics by the concept of symmetry breaking. This was not so much a matter of importing ideas from one field to another. Rather, many of those working on these issues, such as Landau, Nambu, Weinberg, Philip Anderson, Sidney Coleman, and

Gerard 't Hooft, considered that in the end they were working on the *same* problem, applied to different physical systems.

1.4 Turing Patterns in Biology

So let us now return to the living world. Symmetry breaking offers a kind of order for free: a distinction between a *here* and a *there* in what was initially a homogeneous system. Evolution has evidently made much use of such a cheap and useful source of organization. Cells of all kinds, for example, contain temporary, membraneless compartments caused by a spontaneous, symmetry-breaking phase transition between liquid phases of different composition (Shin and Brangwynne 2017). These liquid–liquid phase transitions are implicated in cell processes ranging from the regulation of transcription to the splicing of mRNA and the storage of cell components in response to stress.

While the mechanisms for this phase separation are still being debated, at root it seems to be nothing more than a kind of oil-in-water immiscibility, produced by rather generic differences in the affinities of the biomolecular constituents. In other words, to the extent that the process is programmed at all, it involves a fuzzy, collective molecular logic, for example involving nonspecific associations between intrinsically disordered proteins and/or non-coding RNA molecules. There is nothing much here that van der Waals or Curie would not have recognized in terms of first-order phase transitions.

The key point is not just that spontaneous symmetry breaking supplies biology with a way to create structures on length scales much greater than the molecular. It is that it does so in a way that is *robust to noise*. That is essential because living cells and organisms and their environments are very noisy; not least, there can be a high degree of cell-to-cell variation in the expression rates of genes even for nominally identical cells (Raj et al. 2006), as well as spatial heterogeneity of the components. Spontaneous symmetry breaking tends to be insensitive to those fine details. Liquid–liquid phase separation, for example, does not depend on precisely which molecules are present or on highly selective interactions between them.

And just as the macroscale behavior is not highly sensitive to the microscale details, so the microscale need not be strongly influenced by the macroscale. For example, even though the viscosity in condensates might be increased by several orders of magnitude relative to the bulk, the local nanosecond dynamics of the proteins, on which their function depends, are essentially unaltered (Galvanetto et al. 2023). To function robustly and reliably, biology needs this hierarchy of causation, with a causal separation of time and length scales, rather than all phenomena being determined from the bottom up - from the specific and "programmed" activity of genes or other molecules.

Which brings us back to Turing. By opening and closing this contribution with Turing's idea, it is certainly not my intention to imply that it is the only or even necessarily the main way to break symmetry in the living world. But it is surely one

of the most fertile and versatile, and it is fair to say that it serves as a paradigm for how this may happen in the nonequilibrium conditions on which life depends.

Turing proposed that among the molecular ingredients of the cells of the embryo are diffusive components called morphogens ("shape-formers") that are somehow responsible for triggering development of a cell or tissue along a certain pathway.

The idea of development being controlled by—and symmetry being broken by—diffusive concentration gradients of morphogens was an old one. It was first proposed at the end of the nineteenth century by the German physiologist Theodor Boveri, who drew most of his conclusions from experiments on sea urchins. Around 1901 he proposed that the fertilized sea-urchin egg acquires two "poles"—an anterior–posterior axis—by virtue of some substance within its cytoplasm that is injected into it from the "mother" even before the egg is fertilized: that is, forced rather than spontaneous symmetry breaking.

In 1924 Hans Spemann and Hilde Mangold (1924) showed that particular cells or groups of cells in an embryo may act in this way as what they called organizers that determine the fate of those around them. Working with amphibian embryos, Mangold moved cells from a region of the embryo developing into one body part (such as a limb or head) to another region, and found that the cells would induce the same body part there instead.

In support of that idea, studies in the 1930s by Waddington and others showed that some tissues can act as "organization centers" that, if implanted at a different location on the organism—or even sometimes on that of a different species—could induce the host tissue to undertake a new developmental path. Waddington spoke rather vaguely of a substance he called an evocator that represents the "active principle" of the organization centers. He seemed to regard it as a chemical compound of some kind.

This idea appeared to be supported in tissue transplantation experiments on chick embryos in 1968 conducted by John Saunders and Mary Gasseling (1968). They studied digit formation in chick wings, and identified tissues called "zones of polarizing activity" (ZPAs), the equivalent of Spemann and Mangold's organizer, which define the polarity of the wing digits. The notion of a zone of polarizing activity was codified in the late 1960s by Lewis Wolpert's French flag model, which supposes that there are concentration thresholds of the morphogen at which distinct genes switch on to produce bands of different tissue types (Wolpert 1969) (Fig. 1.4).

So the idea of morphogens switching on genes had a pedigree by the time Turing invoked it for his patterning process. The difference is that his scheme did not rely on diffusion alone: it was a more complex, nonlinear system that also involved reactions between morphogens. In other words, it is a reaction–diffusion system, capable of breaking uniform symmetry to define not just an axis of polarity but a repeating pattern in space (Fig. 1.5).

The key attributes of the chemical system that Turing posited, in which travelling or stationary waves in concentration can arise from a combination of reaction and diffusion, had already been discussed more than 40 years earlier (Lotka 1910). In 1910 the Austrian-American ecologist and mathematician Alfred Lotka described a chemical reaction that could undergo oscillations of this kind. Lotka was not

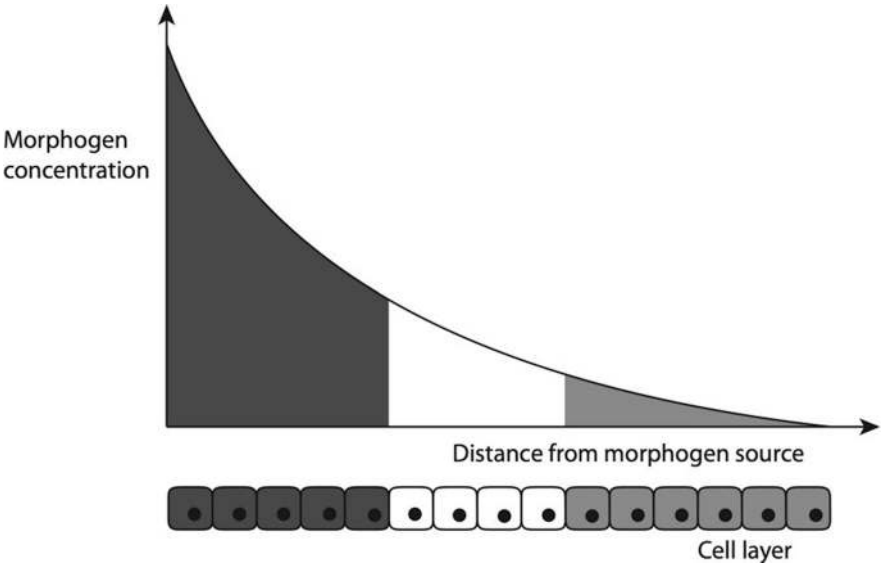


Fig. 1.4 Wolpert's "French flag" model for how continuous diffusional gradients may be converted into discrete patterns

particularly interested in chemistry per se. For him, the reaction was an analogue for the dynamics of animal populations, and the collisions and reactions that may take place between molecules were just proxies for the interactions between predatory beasts and their prey.

Lotka's 1910 paper described an autocatalytic chemical system that develops transient oscillations (Lotka 1910). Ten years later, he outlined conditions under which they could persist (Lotka 1920). Travelling wavefronts resulting from a reaction–diffusion process were also identified in the 1930s by the geneticist Ronald Fisher, who devised a mathematical model—extended by Andrei Kolmogorov and now called the Fisher–Kolmogorov equation—describing how an advantageous allele spreads in a population (Fisher 1937). This model is now understood to provide a very general prescription for a certain type of reaction–diffusion process. But at the time, with theoretical biology in its infancy, there was little appreciation that an approach devised in population genetics should have anything to say about morphogenesis.

At the same time that Turing was working out a theory that could account for pattern-forming reaction–diffusion processes, the first chemical system of this kind to receive serious attention was being discovered. In the 1950s, the Soviet biochemist Boris Belousov devised a cocktail of reagents as a simplified model of glycolysis, and he found that the mixture oscillated back and forth between two states. Belousov's results were largely dismissed until Anatoly Zhabotinsky at Moscow State University discovered a variation of Belousov's mixture that would switch

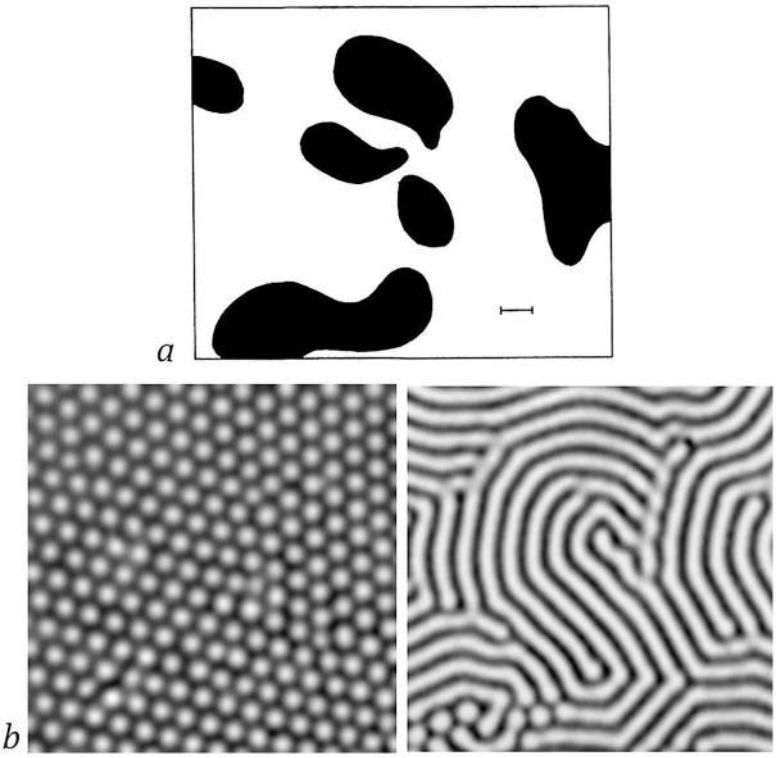


Fig. 1.5 (**a**) The patterns in two dimensions that emerge from Turing's reaction–diffusion scheme, as calculated by Turing in 1952. (**b**) The generic patterns produced by Turing's process are regular spots and stripes. (Computational results courtesy of Jacques Boissonade & Patrick De Kepper, University of Bordeaux)

back and forth between red and blue, and could exhibit symmetry breaking in the form of complex spatial patterns (Fig. 1.6).

When Zhabotinsky discussed these findings at an international conference in Prague in 1968, chemists in the West were intrigued to hear about it and began to figure out what was going on (Winfree 1972).

The qualitative essentials of Turing's model did not emerge until two decades after his paper. In 1972 the developmental biologists Hans Meinhardt and Alfred Gierer in Tübingen devised a theory of biological pattern formation caused by diffusing reagents that paralleled Turing's (Gierer and Meinhardt 1972). They were unaware of Turing's work until a referee of their paper (allegedly Max Delbrück) pointed it out. In Meinhardt and Gierer's model, stationary chemical patterns can result from two interacting ingredients, equivalent to Turing's morphogens: an autocatalytic "activator" and an "inhibitor". A theoretical model similar to this activator–inhibitor scheme was also proposed in 1972 for a predator–prey system by Segel and Jackson (1972), although it tends to be overlooked today.



Fig. 1.6 The dynamic chemical-wave patterns of the Belousov–Zhabotinsky reaction. (Image courtesy of Stephen Morris, University of Toronto)

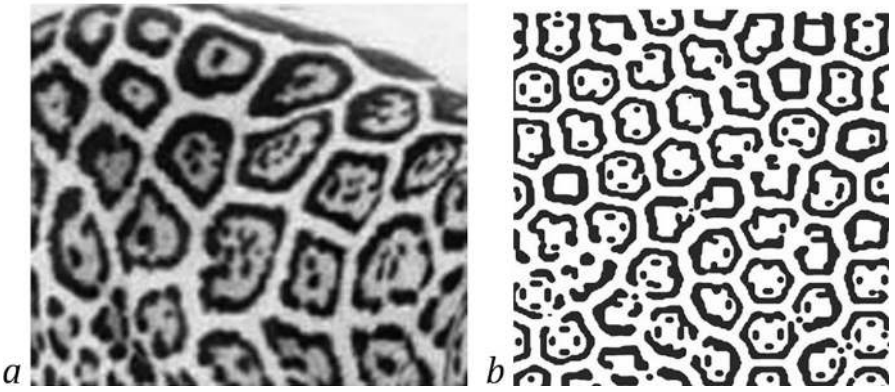


Fig. 1.7 The pigmentation patterns of jaguars (**a**) can be closely mimicked with a Turing-type reaction-diffusion scheme (**b**). From Liu et al. 2006

Waddington wrote to Turing in 1952 to suggest that his scheme might be best deployed to understand animal pigmentation patterns. That idea was developed in the 1980s by Meinhardt (1982) and Jim Murray in Oxford (Murray 1990), who showed how Turing's theory offered a possible explanation for a wide range of animal pigment patterns, from zebras to giraffes to insects. We now know that Turing schemes allow symmetry to be broken in some remarkably complex ways to potentially account for a tremendous diversity of animal pigmentation patterns (Fig. 1.7).

Turing's *travelling* waves might also produce pigmentation patterns. Meinhardt has shown that an activator–inhibitor scheme can reproduce the kinds of complex

patterns seen on some mollusk shells, which are in effect frozen traces of two-dimensional travelling waves on the rim of the growing shell (Meinhardt 1995).

But aside from these special cases of animal markings, for years it proved hard to find examples for which Turing's mechanism seemed to offer a plausible alternative to simple gradient-based models of patterning in embryogenesis. Turing himself seemed to have been conscious of the difficulties: he is said once to have remarked about the zebra, "Well the stripes are easy, but what about the horse part?"

So until recently, Turing's work had rather little impact on developmental biology. That has now changed. Over the past one or two decades, Turing-type mechanisms have been implicated in a wide range of developmental processes, from the fundamental left–right asymmetry of bilateral organisms (Chen and Schier 2002) to the formation of hair (Sick et al. 2006) and feather follicles, the parallel ridges of feather barbs (Harris et al. 2005), the grooves (rugae) of the canine palate (Economou et al. 2012), digit formation (Raspopovic et al. 2014), and the branching of lung airways (Menshykau et al. 2012).

It turns out that Turing's mechanism is in fact not generally necessary to break the symmetry of a fertilized egg. In many organisms, such as fruit flies and frogs, this is disrupted from the outset by maternal factors. But for mammalian embryos the jury is still out: there is no agreement about how the cells of the very early blastomere become distinguished. There is some indication that even at the stage of a two-cell blastomere the cells are not equivalent, but it is not clear why. One idea explored by Chen et al. (2018) is that something like a Turing-type mechanism acts to amplify intrinsic asymmetries in the individual cells themselves, for example because of the placement of organelles or the mitotic spindle.

The important unanswered question for many of these systems seems to be that, while the central elements of activation and inhibition, and reaction and diffusion, do appear to provide a useful basis for framing the problem theoretically, it is unclear to what extent the details can be reduced and simplified to the two-component scheme proposed by Turing. However, in many cases there is nothing particularly bespoke about these generic patterning modes: not only do they seem to be fundamentally the same type of symmetry-breaking process, but they typically involve the same families of morphogens, which are ubiquitous developmental proteins such as Nodal, Sonic hedgehog, Wnt, and BMP. Again, the forms that emerge are not uniquely encoded in the components that produce them but are the result of the underlying dynamics of their many-body interactions.

1.5 Turing Patterns at Large and the Creativity of Nature

Turing's discussion of the whorled leaf arrangements of the woodruff plant in his 1952 paper shows that he suspected his theory might have something to say about a biological patterning process quite different from morphogenesis of the embryo, namely phyllotaxis (Swinton 2004). His basic idea was that an activator–inhibitor system of hormones acting at the growing tip of a plant lays down spots that grow

into buds on the cylindrical stem. This seems plausible in the light of current knowledge of plant biology: phyllotaxis is regulated by proteins that ferry the hormone auxin through the outer stem up toward its apex. Existing leaf buds soak up auxin and thus act as sinks, inhibiting the formation of any new buds nearby.

Meinhardt and Koch showed how a Turing process like this operating on a cylindrical stem can produce a simple spiral phyllotaxis pattern (Koch and Meinhardt 1994), although there is no direct evidence I am aware of that such an activator–inhibitor process really operates in plants.

The feedbacks involved in replication, competition, and predation might set up Turing-type patterns in animal and plant communities. In arid regions where vegetation growth is highly moisture-limited, feedbacks can produce a range of patterns that seem to be fundamentally determined by Turing-type mechanisms (Meron 2019). Theraulaz and colleagues have shown that in the self-organized process by which a species of Mediterranean ant collects dead bodies into piles (Theraulaz et al. 2002), there is both local activation and long-ranged inhibition. The result here is a series of corpse clusters with precisely the ring geometry that Turing first explored in his paper. Mechanisms like this might underlie many other aspects of habitat formation and grouping, such as nest construction.

Turing's stripe patterns resemble not only the skin of a zebra, tiger, or angelfish but also some patterns in inanimate nature, such as the ripples in wind-blown sand. This may be no coincidence. Meinhardt suggested (Meinhardt 1995) that, at root, the formation of these sand patterns does involve a kind of activator–inhibitor system. The mounds and ridges of sand are formed by deposition of wind-blown grains. As a ridge gets bigger, it enhances its own growth by capturing more sand from the moving air. But in doing so it acts as a sink, removing sand from the wind and suppressing the formation of other ripples nearby. The balance between these two processes establishes a roughly constant mean distance between ripples.

These structures are several centimeters in size. Meanwhile, atomic-scale diffusion and reaction processes in the deposition of monolayer metal films can give rise to Turing patterns at the nanoscale, which might conceivably be controlled and fine-tuned for the spontaneous formation of quantum-wire networks (Fuseya et al. 2021).

Thus, Turing seems to have identified one of nature's general mechanisms for generating order from macroscopic uniformity and microscopic disorder. There seems little question that nature, including the living world, does use this mechanism as one way of producing its rich and often beautiful variety of forms.

From a biological perspective, the broader question is how a spontaneous process like this, which gives rise to a particular palette of shapes and patterns, interacts with natural selection. To what extent can evolution adapt and modify Turing structures? Are all such structures necessarily adaptive at all? Or are we too readily tempted, when we see order and regularity in nature, to attribute a "purpose" to it? Might some of it, at least, represent nothing more than a kind of intrinsic creative potential in the natural world?

This question surely lies beneath discussions about the nature of life from Aristotle to Darwin to Schrödinger and today. To put it another way: To what extent is life an extension of physical laws, and to what extent does it require and demand

its own autonomous operational principles? To what extent is it the result of processes that are generic, universal and timeless, and to what extent ones that are historical, path-dependent and contingent?

References

Allen M (2015) Compelled by the diagram: thinking through C.H. Waddington's epigenetic landscape. Contemporaneity 4:120–142
Andrews T (1869) On the continuity of the gaseous and liquid states of matter. Phil Trans R Soc Lond 159:575–590
ATLAS Collaboration (2012) Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC. Phys Lett B 716:1–29
Bardeen J, Cooper LN, Schrieffer JR (1957) Microscopic theory of superconductivity. Phys Rev 106:162–164
Brush SG (1983) Statistical physics and the atomic theory of matter: from Boyle and Newton to Landau and Onsager. Princeton University Press, Princeton, p 272
Chen Y, Schier AF (2002) Lefty proteins are long-range inhibitors of squint-mediated nodal signaling. Curr Biol 12:2124–2128
Chen Q, Shi J, Tao Y, Zernicka-Goetz M (2018) Tracing the origin of heterogeneity and symmetry breaking in the early mammalian embryo. Nat Commun 9:1819
Close F (2022) Elusive: how Peter Higgs solved the mystery of mass. Allen Lane, London
Curie P (1894) On symmetry in physical phenomena, symmetry of an electric field and of a magnetic field. J Phys Théor Appl 3:393–415
Di Lella L, Rubbia C (2015) In: Schopper H, Di Lella L (eds) 60 years of CERN experiments and discoveries. World Scientific, Singapore, pp 137–163
Economou AD et al (2012) Periodic stripe formation by a Turing mechanism operating at growth zones in the mammalian palate. Nat Genet 44:348–351
Fisher RA (1937) The wave of advance of advantageous genes. Ann Eugenics 7:355–369
Fuseya Y, Katsuno H, Behnia K, Kapitulnik A (2021) Nanoscale Turing patterns in a bismuth monolayer. Nat Phys 17:1031–1036
Galvanetto N et al (2023) Extreme dynamics in a biomolecular condensate. Nature. https://doi.org/10.1038/s41586-023-06329-5
Gierer A, Meinhardt H (1972) A theory of biological pattern formation. Kybernetik 12:30
Glashow S (1959) The renormalizability of vector meson interactions. Nucl Phys 10:107
Harris MP, Williamson S, Fallon JF, Meinhardt H, Prum RO (2005) Molecular evidence for an activator-inhibitor mechanism in development of embryonic feather branching. Proc Natl Acad Sci 102:11734–11739
Huxley J (1942) Evolution: the modern synthesis. Allen & Unwin, London
Ising E (1924) Beitrag zur Theorie des Ferro- und Paramagnetismus. PhD dissertation, University of Hamburg
Koch AJ, Meinhardt H (1994) Biological pattern formation: from basic mechanisms to complex structures. Rev Mod Phys 66:1481–1507
Liu RT et al (2006) Two-stage Turing model for generating pigment patterns on the leopard and the jaguar. Phys Rev E 74:011914
Lotka AJ (1910) Contribution to the theory of periodic reaction. J Phys Chem 14:271
Lotka AJ (1920) Analytical note on certain rhythmic relations in organic systems. Proc Natl Acad Sci 6:410
Meinhardt H (1982) Models of biological pattern formation. Academic, London
Meinhardt H (1995) The algorithmic beauty of sea shells. Springer, Berlin

Menshykau D, Kraemer C, Iber D (2012) Branch mode selection during early lung development. PLoS Comput Biol 8:e1002377
Meron E (2019) Vegetation pattern formation: the mechanisms behind the forms. Phys Today 72:30–36
Murray JD (1990) Mathematical biology. Springer, Berlin
Nambu Y, Jona-Lasinio G (1961) Dynamical model of elementary particles based on an analogy with superconductivity. I. Phys Rev 122:345
Needham J (1959) A history of embryology, 2nd edn. Cambridge University Press, Cambridge
Noether E (1918) Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-Physikalische Klasse 1918:235–257
Onsager L (1944) Crystal statistics. I. A two-dimensional model with an order-disorder transition. Phys Rev 65:117
Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. PLoS Biol 4:e309
Raspopovic D, Marcon L, Russo L, Sharpe J (2014) Modelng digitrs: Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. Science 345:566–570
Salam A, Ward JC (1964) Electromagnetic and weak interactions. Phys Lett 13:168–171
Saunders JW, Gasseling MT (1968) In: Fleischmajer R, Billingham RE (eds) Epithelial-mesenchymal interactions. Williams & Wilkins, Baltimore, pp 78–97
Schrödinger E (1944) What is life? Cambridge University Press, Cambridge
Segel LA, Jackson JL (1972) Dissipative structure: an explanation and an ecological example. J Theor Biol 37:545–559
Shin Y, Brangwynne CP (2017) Liquid phase condensation in cell physiology and disease. Science 357:eaaf4382
Sick S, Reinker S, Timmer J, Schlake T (2006) WNT and DKK determine hair follicle spacing through a reaction-diffusion mechanism. Science 314:1447–1450
Spemann H, Mangold H (1924) über Induktion von Embryonalagen Durch Implantation artfrem der Organisatoren. Archiv für Mikroskopische Anatomie und Entwicklungsmechanik 100:599–563. Translated & reprinted in Int. J Dev Biol 45, 13–38 (2001)
Stanley HE (1971) Introduction to phase transitions and critical phenomena. Oxford University Press, Oxford
Swinton J (2004) Watching the daisies grow: Turing and Fibonacci phyllotaxis. In: Teuscher C (ed) Alan Turing: life and legacy of a great thinker. Springer, Berlin
Theraulaz G, Bonabeau E, Nicolis SC, Deneubourg J-L (2002) Spatial patterns in ant colonies. Proc Natl Acad Sci 99:9645–9649
Thompson DW (1919) On growth and form. Cambridge University Press, Cambridge
Turing AM (1952) The chemical basis of morphogenesis. Philos Trans R Soc B 237:37–72
Van de Waals JD (1988) In: Rowlinson JS (ed) On the continuity of the gaseous and liquid states. North Holland, Amsterdam
Waddington CH (1940) Organisers and genes. Cambridge University Press, Cambridge
Watson JD, Crick FHC (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature 171:737–738
Weinberg S (1967) A model of leptons. Phys Rev Lett 19:1264–1266
Winfree AT (1972) Science 175:634–636
Wolpert L (1969) Positional information and the spatial pattern of cellular differentiation. J Theor Biol 25:1–47
Wu CS, Ambler E, Hayward RW, Hoppes DD, Hudson RP (1957) Experimental test of parity conservation in beta decay. Phys Rev 105:1413–1415

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 2
Integrating Mathematical Modeling and Experiments in the Study of Collective Neural Crest Cell Migration



Samuel W. S. Johnson, Ruth E. Baker, and Philip K. Maini

Abstract Collective cell migration underpins a wide range of biological processes, including embryonic development, wound repair, and tumor invasion, yet many of the mechanisms by which both cell–environment and cell–cell interactions give rise to coordinated group-level behaviors remain incompletely understood. In the study of collective cell migration, the neural crest is a particularly powerful model system, as it exhibits robust and highly stereotyped migratory behaviors in vivo, while remaining amenable to direct experimental perturbations and subsequent analysis. In this chapter, we review how the synthesis of experimental and theoretical studies has advanced our understanding of collective neural crest cell migration, and offer an outlook on how these approaches may be further combined to elucidate the mechanisms that underlie collective cell migration. In doing so, we position the neural crest as paradigmatic in our understanding of collective cell migration and for developing general principles that may be applicable across a broad range of developmental and pathological contexts.

Keywords Embryonic development · Cranial neural crest · Collective cell migration · Chemotaxis · Contact inhibition of locomotion · Mathematical modeling

Authors Ruth E. Baker and Philip K. Maini have equally contributed to this chapter.

S. W. S. Johnson (✉) · R. E. Baker · P. K. Maini
Mathematical Institute, Wolfson Centre for Mathematical Biology, University of Oxford, Oxford, UK
e-mail: samuel.johnson@maths.ox.ac.uk

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_2

2.1 Introduction

Collective cell migration underpins many biological processes, such as wound healing (Li et al. 2013), tumor progression (Deisboeck and Couzin 2009), and development (Schumacher 2019a). In such processes, the migration of cells is governed not only by cell–environment interactions but also by cell–cell interactions, which together give rise to emergent cell behaviors and migration as coordinated streams, chains, clumps, and other multicellular patterns (Rørth 2009). These collective behaviors arise in both epithelial and mesenchymal cell populations; however, the underlying biological mechanisms governing migration differ markedly depending on the context (Theveneau and Mayor 2013).

Epithelial collectives typically maintain stable cell–cell junctions and migrate as cohesive sheets, as observed during wound healing (Poujade et al. 2007). In such contexts, collective motion is facilitated by long-lived intercellular connections. In contrast, mesenchymal collectives are characterized by more transient cell–cell interactions and a greater degree of individual cell autonomy, yet, as in epithelial collectives, coordinated migration emerges through dynamic interactions that give rise to collective group behaviors (Theveneau and Mayor 2013). As such, mesenchymal cell collectives provide a particularly informative context in which to investigate how coordinated migration can emerge from transient and dynamic cell–cell interactions.

Studies of collective migration stem as far back as the late nineteenth and early twentieth centuries (Barfurth 1891; Eycleshymer 1907), though historically, its complex and multifaceted nature means that in-depth experimental investigations have, until recent decades, been limited by a lack of appropriate technology. However, with recent advances in experimental methods including spatial transcriptomics (Moses and Pachter 2022), single-cell RNA sequencing (Saliba et al. 2014), and high-resolution live imaging (Lemon and McDole 2020), collective migration has emerged as an increasingly prominent focus of cell biology, complementing longstanding studies of single-cell migration.

Clinically, a greater understanding of collective cell migration may inform the development of novel therapeutic strategies and interventions that target coordinated cellular behaviors, with potential applications in limiting collective tumor invasion during cancer progression (Cheung and Ewald 2016), promoting efficient epithelial migration during wound repair (Friedl and Gilmour 2009), preventing aberrant migratory behaviors during development that give rise to congenital defects (Schumacher 2019a), and informing tissue engineering (Snider and Mishina 2014) and regenerative medicine efforts (Moiseeva et al. 2025).

To elucidate the principles of collective cell migration during embryonic development, it is imperative to study a system that both robustly exhibits collective behavior and permits direct experimental interrogation of model-derived hypotheses. The neural crest is a cell population that satisfies both of these criteria and therefore serves as a powerful model paradigm for collective mesenchymal cell movement (Le Douarin and Kalcheim 1999; Bronner and Le Douarin 2012; Mayor

and Theveneau 2013). During embryogenesis, neural crest cells migrate collectively according to a range of biological mechanisms toward distal target sites, where they subsequently differentiate to contribute to the formation of many embryonic structures including the peripheral nervous system, pigment and endocrine cells, and a range of mesenchymal tissues (Le Douarin and Kalcheim 1999). Neural crest cell migration exhibits several features that are characteristic of collective cell migration in other developmental and pathological contexts, suggesting that mechanisms identified in this system may reflect more general biological principles (Theveneau and Mayor 2013). Consequently, the migration of neural crest cells has been studied extensively both in vitro and in vivo; however, despite this wealth of experimental data, many aspects of how local cell–cell interactions give rise to robust collective behavior remain poorly understood. Mathematical modeling has, therefore, been widely utilized alongside experiments to generate novel biological hypotheses and to test existing ones (Giniūnaitė et al. 2020; Szabó and Mayor 2016).

In this review, we synthesize experimental and mathematical studies of neural crest cell migration to evaluate the contribution of integrated modeling–experimental approaches to our understanding of collective mesenchymal cell migration, and conclude by discussing how closer integration of these disciplines may provide further insight into both neural crest-specific behaviors and the general principles of collective migration across many biological contexts.

2.2 Collective Cell Migration in the Neural Crest

The neural crest is a transient cell population that undergoes extensive collective migration during vertebrate embryonic development (Szabó and Mayor 2018). Following migration, neural crest cells contribute to the formation of a wide range of tissues and structures in the developing embryo (Mayor and Theveneau 2013). Neural crest cells are specified at the border of the neural plate and, following an epithelial-to-mesenchymal transition, delaminate from the neuroepithelium to migrate collectively over long distances to target sites throughout the embryo (Fig. 2.1a) (Bronner and Le Douarin 2012; Szabó and Mayor 2018). Through this process, neural crest cells generate a diverse range of derivatives, including components of the peripheral nervous system and connective tissues (Donoghue et al. 2008).

Neural crest cell migration is tightly regulated by both cell–cell and cell–environment interactions, and the factors expressed within the environment through which migration occurs (Kulesa and Gammill 2010). Disruption of neural crest cell migration can lead to a broad spectrum of downstream congenital pathologies, collectively termed neurocristopathies (Sato et al. 2019; Vega-Lopez et al. 2018). These include craniofacial malformations such as cleft palate, which arise in part from defects in cranial neural crest cell migration, proliferation, or patterning during embryogenesis (Cordero et al. 2011). Such conditions have motivated extensive investigation into the mechanisms governing neural crest development and

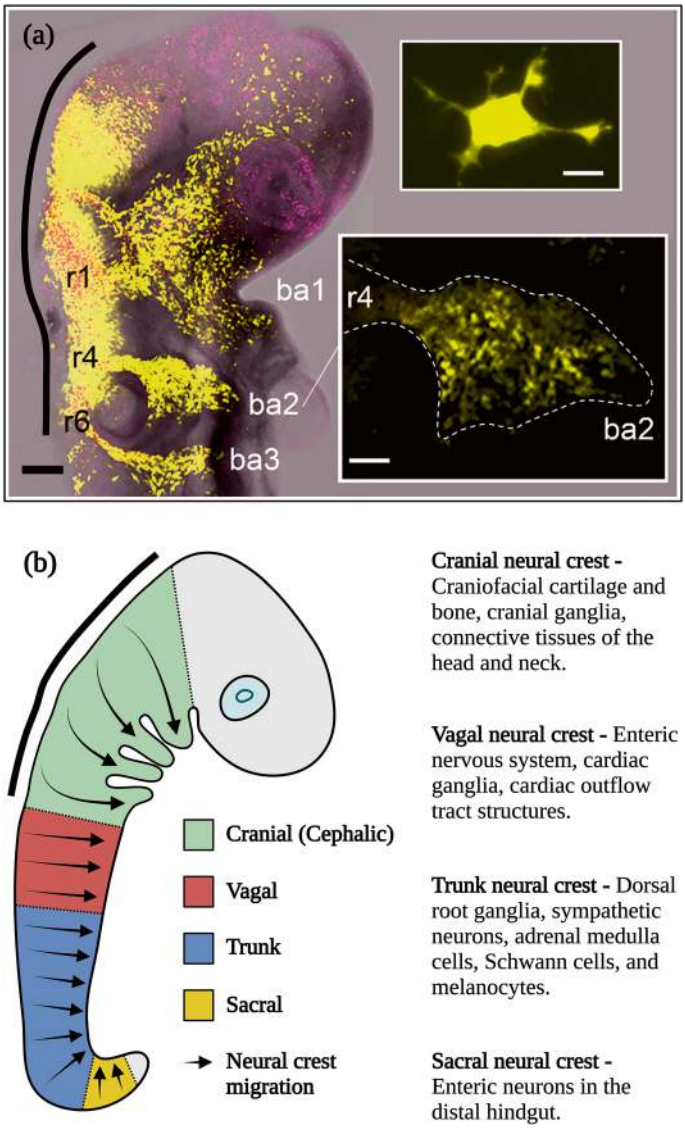


Fig. 2.1 (**a**) Neural crest cell migration in the chick head. Cranial neural crest cells (yellow) delaminate from the embryonic hindbrain and some are sculpted into streams adjacent to rhombomeres (r) 1, 4, and 6. Collective migration of these cells then occurs toward branchial arches (ba) 1, 2, and 3. A scale bar (black) is shown in the bottom left and corresponds to a distance of 200 μm. The upper inset shows a representative single migrating neural crest cell (scale bar, 10 μm; white). The lower inset highlights a segment of the neural crest migratory stream adjacent to r4 (scale bar, 100 μm; white). Adapted from Giniūnaitė et al. (2020). (**b**) Schematic of the axial subpopulations of the neural crest and their derivatives. In both panels, a black line is used to illustrate the cranial region of the embryo from which these cells emerge. Created in BioRender by Samuel W.S. Johnson: https://BioRender.com/8rmbhg0

migration to identify potential targets for therapeutic intervention aimed at mitigating or preventing neurocristopathy-associated defects.

In vivo, the neural crest is commonly divided into four axial subpopulations: cranial, vagal, trunk, and sacral, each defined by their position along the anterior-to-posterior axis of the embryo and their characteristic migratory trajectories (Fig. 2.1b) (Le Douarin and Kalcheim 1999; Bronner and Le Douarin 2012; Mayor and Theveneau 2013). Cranial (sometimes termed cephalic) neural crest cells migrate collectively from the hindbrain and midbrain into the branchial arches, where they contribute to craniofacial development in the embryo (Kulesa et al. 2010; Knight and Schilling 2006), while vagal neural crest cells migrate rostro-caudally through the developing gut mesenchyme to form the enteric nervous system, advancing as a coordinated migratory wave (Gershon et al. 1993; Hutchins et al. 2018). In addition, a spatially restricted subset of the vagal neural crest, termed the cardiac neural crest, colonizes the pharyngeal arches and conotruncal region, where it plays essential roles in great vessel patterning and cardiac septation (Stoller and Epstein 2005). Trunk neural crest cells follow stereotyped dorsolateral and ventral migratory pathways that give rise to melanocytes and components of the peripheral nervous system, respectively (Gammill and Roffers-Agarwal 2010; Vega-Lopez et al. 2017), and sacral neural crest cells migrate later in development to populate distal regions of the hindgut and pelvic ganglia (Burns and Le Douarin 1998; Wang et al. 2011), though their contribution to the enteric nervous system in mice has recently been contested (Yu et al. 2024). Disruption to the migration of any of these subpopulations can lead to developmental abnormalities, underscoring the importance of robust collective migration during neural crest development. Additionally, these subpopulations generate distinct derivatives yet share a reliance on coordinated, long-range migration and dynamic cell–cell interactions, making them well-suited for comparative studies of the mechanisms underlying collective cell migration.

2.2.1 Mechanisms of Neural Crest Cell Migration

Collective neural crest cell migration is regulated by the interplay between multiple conserved mechanisms that coordinate cell–cell cohesion and guidance across axial subpopulations. Chemotactic signaling provides directional information during migration, with neural crest cells responding to both long-range and short-range chemical cues in their environment (Shellard and Mayor 2016). Such cues bias collective movement and, in some contexts, are integrated with contact-mediated behaviors to generate persistent directional migration, as exemplified by the *chase-and-run* interactions between cranial neural crest cells and neighboring placodal cells in *Xenopus* (Theveneau et al. 2013). Chemotactic guidance has also been implicated in chick trunk neural crest cell migration, where chemical gradients help to steer cells toward specific peripheral targets (Kasemeier-Kulesa et al. 2010; de Bellard et al. 2018).

Cell–cell interactions play a central role in shaping collective dynamics through contact inhibition of locomotion (CIL), a process in which cells repolarize and change direction following contact with neighbors (Stramer and Mayor 2017). In *Xenopus* and zebrafish neural crest cells, homotypic CIL promotes dispersion within collectives, while reinforcing front–rear polarity at the population level, thereby contributing to coherent collective migration (Carmona-Fontaine et al. 2008). CIL is modulated by cell–cell adhesion and polarity signaling pathways (Bahm et al. 2017; Moore et al. 2013), and can also occur heterotypically, where interactions with other cell types enhance directional persistence during migration (Theveneau et al. 2013).

Neural crest cell guidance is further shaped by interactions with the extracellular matrix (ECM) through which migration occurs, which provides both permissive substrates and inhibitory boundaries that constrain and pattern migration (Perris and Perissinotto 2000; Tsang et al. 2010). Neural crest cells preferentially migrate on fibronectin- and laminin-rich matrices, while other non-permissive ECM components restrict invasion and hence, promote directionality, particularly in the trunk region of the embryo (Perris et al. 1989; Landolt et al. 1995; Kubota et al. 1999). Neural crest cells also actively remodel their microenvironment through proteolytic activity, which facilitates their movement through the embryonic microenvironment (Monsonego-Ornan et al. 2012).

At the level of the collective, cohesion is maintained by co-attractive interactions (CoA) that counterbalance dispersive forces arising from CIL. In several neural crest cell populations, short-range self-generated chemoattractant signaling promotes reaggregation following cell–cell separation, which stabilizes migrating groups and supports robust collective responses to external guidance cues (Carmona-Fontaine et al. 2011; Broders-Bondon et al. 2016).

Despite substantial progress in identifying the molecular and cellular mechanisms that regulate neural crest cell migration, many open questions remain regarding how these processes are integrated to generate robust, directed collective behavior in vivo. In particular, it is unclear how chemotactic cues, contact-mediated interactions, ECM guidance, and co-attractive forces are combined across spatial and temporal scales to produce persistent stream-level migration, or whether all of these mechanisms are required in all developmental contexts. Moreover, many of these interactions are inherently local, yet give rise to long-range order and guidance, which raises the question of how global migratory patterns emerge from local rules of cell–cell and cell–environment interactions.

In addition, much of the experimental evidence underpinning current models of neural crest cell migration is derived from a limited set of organisms, most notably *Xenopus* and chick, and the extent to which these mechanisms are conserved across other organisms remains incompletely resolved. Addressing these questions requires frameworks that are capable of integrating multiple interacting mechanisms, testing their sufficiency and robustness, and exploring emergent collective behaviors that

are difficult to infer from experiments alone. Mathematical and computational modeling, therefore, provides a tool for disentangling the relative contributions of individual mechanisms and for identifying minimal sets of rules capable of reproducing observed neural crest cell migratory dynamics across biological systems.

2.2.2 *Collective Neural Crest Cell Migration in Xenopus and Chick*

Experimental studies of collective neural crest cell migration that have been informed by mathematical modeling have focused predominantly on *Xenopus* and chick, owing to the relative ease with which these systems may be experimentally perturbed and studied (Kulesa et al. 2004). In both systems, embryos can be cultured externally and are sufficiently large to permit direct experimental manipulation, targeted perturbation of gene function, and high-resolution live imaging of cell migration (de Bellard and Bronner-Fraser 2005; DeSimone et al. 2005). These properties enable precise spatiotemporal control of experimental interventions, including tissue grafting and explant culture. Neural crest cell populations in both organisms also migrate along stereotyped and well-characterized pathways (Kulesa et al. 2010; Smith et al. 1997), facilitating reproducible analyses of cell–cell interactions and responses to guidance cues at both the single-cell and collective scales. The conservation of molecular signaling pathways governing cell specification and migration further suggests that insights gained from these models are broadly applicable to other vertebrates in which experimental perturbations are less feasible.

Among the neural crest cell subpopulations, cranial neural crest cells (CNCCs) have been especially informative for dissecting the mechanisms of collective migration. A notable feature of collective CNCC migration is that cells are largely mitotically quiescent during early migratory phases, prior to invasion of the branchial arches. Cell cycle mapping has shown that in chick, proliferative activity within the cranial neural crest remains low during migration and increases only after CNCCs reach their target sites (Ridenour et al. 2014) and in *Xenopus*, proliferation is infrequent throughout CNCC migration (Szabó et al. 2016). This temporal separation of migration and proliferation contrasts with other neural crest populations, such as the vagal neural crest, in which proliferation at the migratory wavefront contributes to sustained tissue colonization and a directional advance (Simpson et al. 2007a). As a result, studies of CNCC migration have provided a valuable context in which collective motility can be examined with minimal confounding effects from population growth (Szabó and Mayor 2018; Schumacher et al. 2016).

2.2.2.1 Collective CNCC Migration Through CIL and CoA

Studies of CNCC migration in *Xenopus* have provided key insights into the cell–cell interactions that underpin collective migration. The prevailing model of collective CNCC migration in *Xenopus* is the contact inhibition of locomotion–co-attraction (CIL–CoA) paradigm (Fig. 2.2a), in which short-range homotypic CoA promotes cohesion within CNCC collectives, while homotypic CIL induces repolarization upon collision, thus preventing excessive clustering (Carmona-Fontaine et al. 2008, 2011). However, as CIL and CoA are intrinsically local interactions without an inherent directional bias, they are unlikely, in many contexts, to fully account for the robust long-range guidance of CNCC streams toward distal targets in vivo. In *Xenopus*, additional directional guidance is provided by interactions with neighboring tissues. In particular, CNCCs migrate chemotactically toward placode-derived stromal cell-derived factor 1 (Sdf1), and reciprocal CNCC–placode contact triggers heterotypic CIL that displaces placodal cells, yielding a chase-and-run mechanism (Theveneau et al. 2013). In this setting, CIL and CoA can be interpreted as governing intra-collective cohesion and rearrangements, while chase-and-run provides an external steering cue through CNCC–placode interactions. A parsimonious interpretation is, therefore, that directed CNCC migration in *Xenopus* emerges from the interplay between these processes; placodal Sdf1 biases the net direction of advance, while homotypic CIL and CoA maintain a coherent, motile collective capable of sustained stream-like migration. However, while this model provides a compelling description of CNCC migration in *Xenopus*, the extent to which this guidance

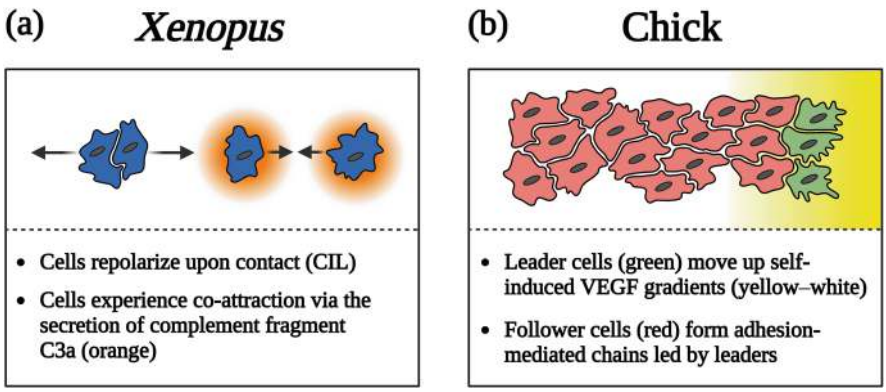


Fig. 2.2 Schematics for the mechanisms of CNCC migration in *Xenopus* and chick. (**a**) Mechanisms of collective CNCC migration in *Xenopus*. Cells repolarize upon contact, thus preventing excessive crowding. Cells also experience homotypic co-attraction via the secretion of complement fragment C3a, which binds collectives together. (**b**) Mechanisms of collective CNCC migration in chick. Cells at the leading edge of streams migrate up self-induced gradients in vascular endothelial growth factor (VEGF), while trailing cells form connected streams that move collectively behind leaders. Created in BioRender by Samuel W.S. Johnson: https://BioRender.com/afkim2a

strategy generalizes to amniote neural crest and to non-cranial neural crest subpopulations, as opposed to reflecting a context-specific mechanism arising from CNCC–placode interactions in anamniotes, remains to be established.

2.2.2.2 The Leader–Follower Hypothesis

In chick, experimental evidence suggests a different model of collective CNCC migration to that of *Xenopus*. Previous experimental studies have shown that the transcriptome of chick CNCCs varies within streams in a manner correlated with their relative stream positions (McLennan et al. 2012; Morrison et al. 2017). The first of these studies divided r4–ba2 streams into leading and trailing cell subpopulations according to their distance from r4 (McLennan et al. 2012). Using both laser-capture microdissection (LCM) and fluorescence-activated cell sorting (FACS), 84 genes implicated in CNCC migration were screened across the two subpopulations. Nineteen of these genes were found to be differentially expressed between leading and trailing CNCC populations, indicating substantial transcriptional differences between leading and trailing cell subpopulations. In these experiments, CNCCs in leading positions were found to express genes associated with chemotaxis, ECM remodeling, and ECM attachment at levels higher than cells in trailing populations. Conversely, trailing cells were found to express cadherins distinct from leading subpopulations (McLennan et al. 2012). These observations have been used to infer that collective CNCC migration in chick is facilitated by heterogeneous CNCC behaviors, wherein cells at the leading edge of streams move directedly in response to gradients in vascular endothelial growth factor (VEGF) formed in the overlying ectoderm, and follower cells use cadherin-mediated cell–cell adhesion to form loosely connected groups that are guided toward the branchial arches by leader cells at the invasive front (Fig. 2.2b). This type of leader–follower migration has been well-documented in many other biological systems, including epithelial cell migration (Vishwakarma et al. 2018) and angiogenesis (Chen et al. 2019). However, despite experimental evidence of leader–follower-like migration within chick CNCC streams, it remains unclear whether CNCC phenotypes conform to a strict leader–follower binary or instead span a continuum of phenotypic states within streams. Moreover, it is not known whether these phenotypes are stable during migration or whether CNCCs undergo frequent phenotypic transitions during migration in response to their local microenvironment and cell–cell interactions.

Experimental studies in both chick and *Xenopus* highlight that collective neural crest cell migration emerges from the interplay of local cell–cell interactions, environmental guidance cues, and, in some contexts, phenotypic heterogeneity within migrating populations. Disentangling how such local rules give rise to robust, long-range collective behaviors is challenging to address through experiments alone, motivating the use of mathematical and computational models to formalize proposed mechanisms and to rapidly generate quantitative, experimentally testable predictions.

2.3 Mathematical Modeling in the Study of Neural Crest Cell Migration

Mathematical modeling has played an increasingly important role in the study of neural crest cell migration, providing a quantitative framework for evaluating experimentally derived hypotheses and for generating new, testable predictions. In particular, models that are developed and interrogated in close dialogue with experiments have yielded significant insight into the mechanisms governing collective neural crest cell migration. In what follows, we discuss key contributions from the literature that have yielded insights in this area, with an emphasis on studies that integrate theoretical and experimental approaches. More comprehensive surveys of the neural crest modeling literature can be found in reviews by Giniūnaitė et al. (2020) and Szabó and Mayor (2016) (Fig. 2.3).

2.3.1 Agent-Based Models

Agent-based models (ABMs) have been widely used to study neural crest cell migration, owing to their ability to represent individual cells as discrete, interacting agents and to resolve cell–cell interactions and phenotypic heterogeneity at the single-cell level.

ABMs of neural crest cell migration span a range of computational formalisms, including on-lattice, off-lattice, and cellular Potts approaches, each offering different trade-offs between computational efficiency and biological detail. In contrast to tumor growth models, where very large cell numbers often necessitate coarse-grained representations of cell density, neural crest cell collectives typically comprise relatively small populations (Morrison et al. 2017; Szabó et al. 2016). This has enabled the widespread use of off-lattice and cellular Potts-based frameworks that resolve cell position, cell shape, and mechanical interactions in greater detail.

Across modeling frameworks, ABMs have demonstrated that robust collective neural crest cell migration cannot be explained by passive movement along pre-patterned ECM tracks alone. On-lattice models have shown that persistent leader–follower neural crest cell chains require directional bias in both leader and follower cells and sustained leader–follower contact to maintain cohesion (Wynn et al. 2012), while hybrid models incorporating intracellular Rac1 dynamics indicate that collective chemotaxis is shaped by interactions among chemotactic guidance, confinement, and cell–cell interactions (Freingruber et al. 2024).

Off-lattice ABMs have played a central role in elucidating the minimal interaction rules required for collective neural crest cell migration. Motivated by experimental observations in *Xenopus* and zebrafish, a computational model has demonstrated that the balance between contact inhibition of locomotion (CIL) and co-attraction (CoA) is sufficient to drive the spontaneous formation of migratory streams and persistent collective motion (Carmona-Fontaine et al. 2011). Subsequent

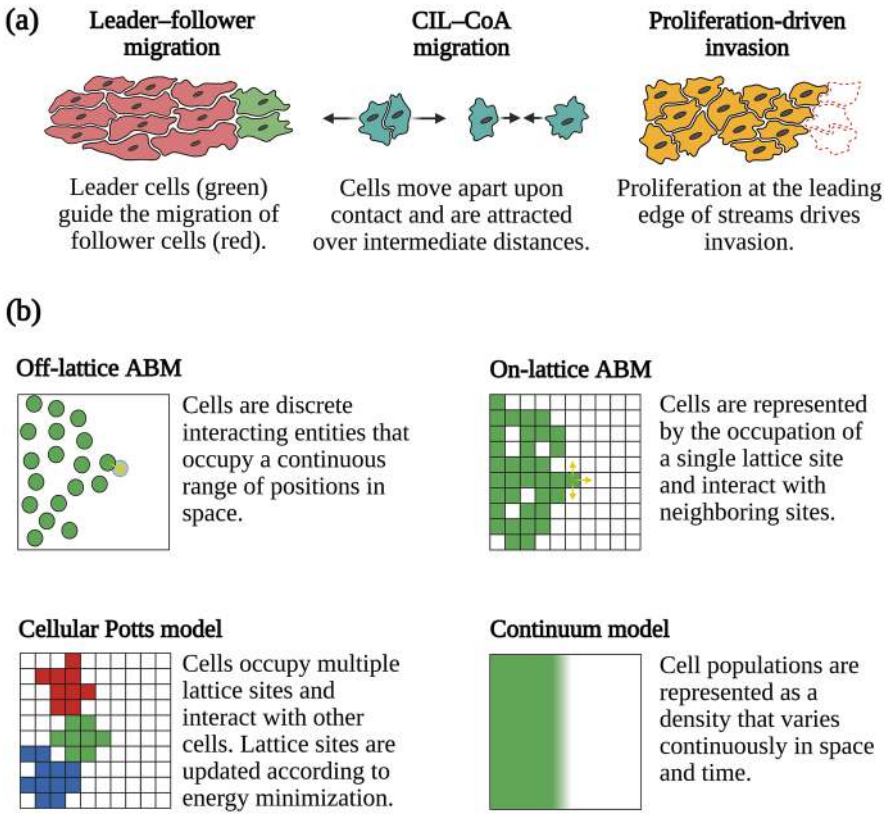


Fig. 2.3 (**a**) Schematics of leader–follower, CIL–CoA, and proliferation-driven modes of collective neural crest cell migration. (**b**) Schematics of off-lattice agent-based models (ABMs), on-lattice ABMs, cellular Potts models, and continuum models of neural crest cell migration. Created in BioRender by Samuel W.S. Johnson: https://BioRender.com/udye7f0

models incorporating explicit cell shape, elastic collisions, or spatial confinement have further shown that tissue geometry and mechanical constraints enhance directional persistence and migration efficiency (Szabó et al. 2016; Woods et al. 2014). More recent work has shown that, under conditions of sustained cell influx, CIL-like interactions alone can generate persistent directed collective migration through density-dependent alignment within channel-like geometries (Khataee et al. 2021).

ABMs have also been instrumental in dissecting how environmental guidance cues and heterogeneity shape neural crest patterning. A chemotactic model of trunk neural crest cell migration has shown that a single chemotactic guidance cue is insufficient to reproduce experimentally observed segregation patterns in vivo, whereas robust patterning emerges when cells respond to multiple, temporally distinct signals (Dyson et al. 2018). Similarly, recent modeling work suggests that collective neural crest cell behaviors can arise from the interplay between cell-generated forces, adhesion, and ECM remodeling, with external directional cues

primarily biasing or stabilizing these dynamics rather than being strictly required to initiate them (Martinson et al. 2023). Related biophysical modeling of mammalian neural crest ingression has further highlighted how adhesion, stochastic cell-shape fluctuations, and entropic packing may contribute to neural crest cell behaviors independently of simple pre-patterned guidance cues (Pasiliao et al. 2025).

An additional model resolved at the single-cell level has been used to generate a hypothesis that VEGF drives the morphological heterogeneity observed within streams of chick CNCCs (Teddy and Kulesa 2004; Johnson et al. 2025). By studying the optimal cell morphology for the detection of exogenous VEGF gradients in the cranial microenvironment, the optimal CNCC morphology was shown to vary in a manner akin to the morphological transition observed within streams in vivo. Additionally, a corresponding ABM of chick CNCC migration showed that persistent cell movement, mediated by genes such as AQP1 (McLennan et al. 2020), can yield further increases in the efficiency of migration through heterogeneous VEGF landscapes.

2.3.2 Continuum Models

Continuum models provide a complementary framework for studying neural crest cell migration by representing cell populations as continuously varying density fields rather than as discrete individuals. By focusing on the spatiotemporal evolution of cell density, these models offer greater mathematical tractability than agent-based approaches but do not explicitly resolve individual cell–cell interactions. As such, continuum models are particularly well suited for identifying population-level mechanisms that govern large-scale invasion and tissue colonization. Continuum modeling has been especially influential in studies of vagal neural crest cell migration and enteric nervous system development. Early partial differential equation (PDE) models incorporating diffusion and proliferation demonstrated that proliferation at the migratory wavefront is a key driver of sustained invasion of the gut mesenchyme (Simpson et al. 2006, 2007a). These models predicted that insufficient proliferative activity leads to incomplete colonization, a hypothesis that was subsequently validated experimentally using chick–quail tissue grafts (Simpson et al. 2007a). This modeling–experimental synthesis provided a quantitative explanation for how population growth and migration jointly contribute to robust long-range invasion, and has since been extended to connect cell-scale and population-level descriptions of invasion (Simpson et al. 2007b).

Beyond proliferation-driven invasion, continuum models have also been used to explore the role of cell–cell interactions in shaping collective neural crest cell migration. A nonlocal PDE framework incorporating repulsive interactions over finite spatial ranges has been developed to represent contact-mediated behaviors such as contact inhibition of locomotion (CIL) (Painter et al. 2015). Simulations of this model showed that nonlocal repulsion can substantially enhance invasion speed relative to a diffusion–proliferation model without contact repulsion. These

observations established a mechanistic link between microscale interactions and emergent macroscale behaviors, and thus, demonstrated how collective behaviors traditionally studied at the single-cell level can be embedded within continuum descriptions through appropriate nonlocal terms.

2.3.3 Synthesizing Mathematical and Experimental Insights in the Study of Collective Neural Crest Cell Migration

Having reviewed the broader mathematical modeling literature on neural crest cell migration, we now focus on a set of studies in which mathematical modeling and experimentation have been closely integrated, yielding insights that neither discipline could provide alone. In this body of work, iterative feedback between theory and experiments has been used to refine biological hypotheses and uncover the mechanisms governing collective neural crest cell migration.

In the study of chick CNCC migration, a family of predominantly off-lattice hybrid ABMs has been developed alongside experiments to interrogate the mechanisms underlying CNCC stream formation and guidance (Johnson et al. 2026a; McKinney et al. 2020; McLennan et al. 2012, 2015, 2017, 2020, 2023). In chick, CNCCs migrate along gradients of VEGF (McLennan et al. 2010), and the early consensus on chick CNCC migration was that VEGF-driven chemotaxis alone could facilitate collective stream-like migration through a so-called cell-induced gradient mechanism, wherein CNCCs at the leading edge of streams locally degrade VEGF, thereby creating gradients to guide both themselves and trailing cells toward migratory targets in the branchial arches. This hypothesis was then tested in a hybrid ABM in which VEGF was represented as a continuously varying scalar field within a two-dimensional domain, and CNCCs were represented as interacting, non-overlapping discs that migrate in response to gradients in VEGF. Numerical simulations of this model predicted that chemotaxis alone was insufficient to reproduce experimentally observed migration patterns in chick, as cells located behind the leading edge of streams (red circles in Fig. 2.4) invaded regions of depleted VEGF, such that VEGF gradients could not be consistently detected throughout migration, leading to atypical patterning and failed invasion of the branchial arches (McLennan et al. 2012).

These findings motivated a subsequent hypothesis that there is a division of labor within CNCC streams, wherein CNCCs in leading stream positions respond to cues distinct from trailing CNCCs. In this framework, leader cells were hypothesized to migrate directionally in response to self-induced VEGF gradients, while follower cells were hypothesized to rely primarily on cell–cell adhesion-mediated interactions to maintain cohesive, elongated streams led by leader cell VEGF chemotaxis. In the hybrid ABM framework (Fig. 2.4b), this yielded streams akin to those observed in vivo, motivating biological experiments which, through transcriptional analyses, revealed differential gene expression profiles between leader and follower

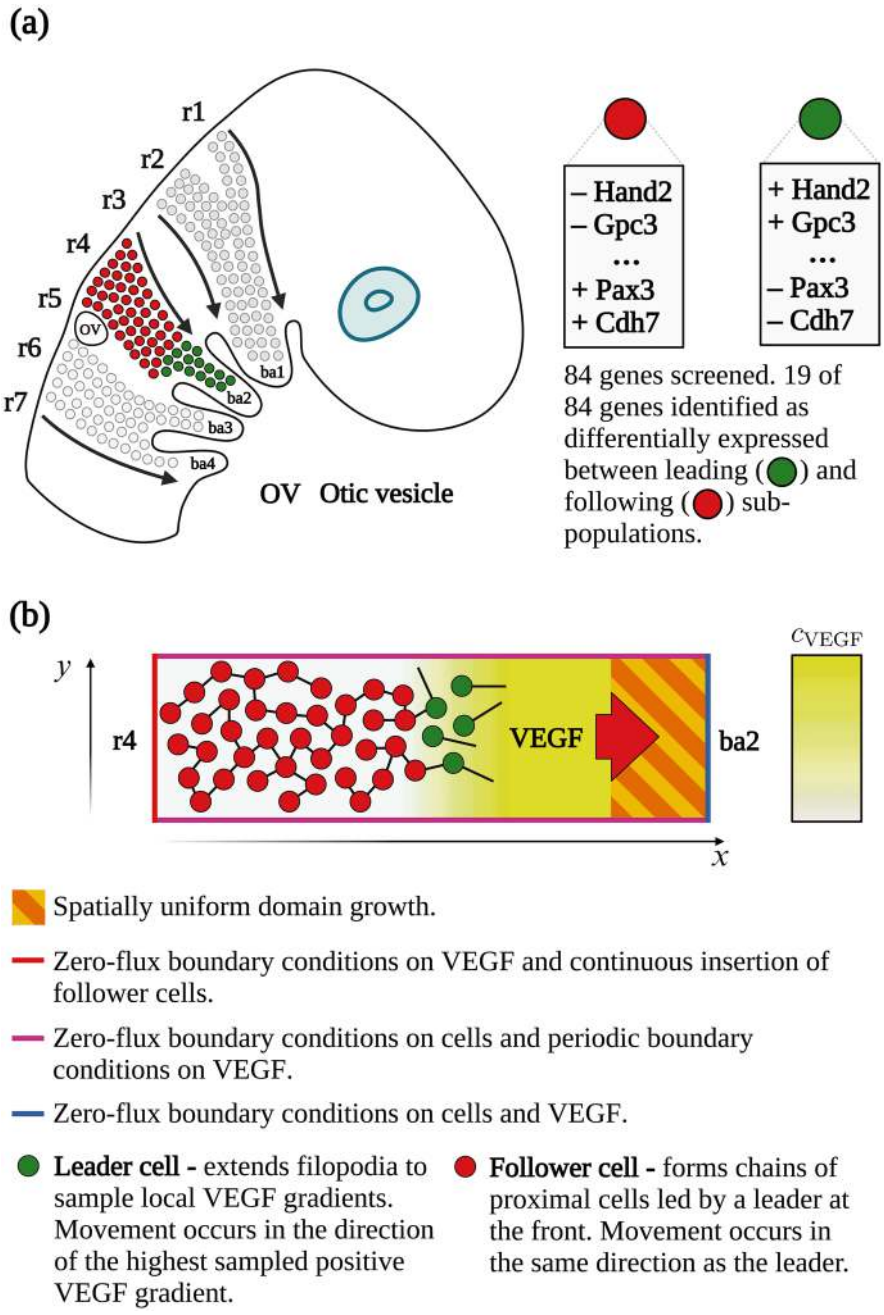


Fig. 2.4 (**a**) Schematic of the leader–follower transcriptional analysis conducted in chick (McLennan et al. 2012). Using both LCM and FACS, CNCCs migrating toward ba2 were found to express genes correlated with their positions in streams. Leading cells (green) were found to express genes associated with VEGF guidance and cell–ECM interactions, whereas trailing cells were found to express genes associated with cell–cell adhesion at higher levels than leading cells. (**b**) Schematic of the leader–follower ABM of chick CNCC migration (Johnson et al. 2026a; McKinney et al. 2020; McLennan et al. 2012, 2015, 2017, 2020, 2023). Leader cells migrate up gradients in VEGF detected by filopodia, while follower cells form cell–cell attachments to create chains of CNCCs led by a leader at the front of the stream. Created in BioRender by Samuel W.S. Johnson: https://BioRender.com/004nf6n

populations (Fig. 2.4a) (McLennan et al. 2012). This hitherto biologically undiscovered phenotypic heterogeneity agreed with predictions generated by the corresponding mathematical model, with modeling and experiments thus being integrated to formulate a leader–follower model of chick CNCC migration (Sect. 2.2.2.2).

A key limitation of the initial leader–follower hybrid ABM formulation was the tendency for leader cells to detach from trailing followers, leading to stream fragmentation, which, in turn, suggested the presence of an additional regulatory mechanism governing the speed of collectives. Subsequent experimental work identified the factor DAN, a bone morphogenetic protein antagonist expressed in the cranial microenvironment, as a key modulator of CNCC motility (McLennan et al. 2017). DAN was shown to reduce CNCC migration speed both in vitro and in vivo, and numerical simulations including an initial model representation of cell speed modulation from DAN demonstrated that DAN expression could stabilize leader–follower collectives and prevent stream breakup.

Another simplifying assumption of the original model was the imposition of zero-flux boundary conditions on cell movement, reflecting the observation that CNCC streams in vivo are confined to discrete migratory pathways separated by regions devoid of CNCCs (Kulesa et al. 2010). Computational analyses later showed that zero-flux boundary conditions on migrating CNCCs need only be imposed for approximately one-third of the migratory domain in the direction of migration to achieve maximal stream confinement (McLennan et al. 2023). This prediction was validated by subsequent experimental findings that the guidance molecules Trail and Colec12 are expressed over a similar proportion of the chick embryo in the regions adjacent to the corridor along which collective CNCC migration occurs. More recent ABMs have further clarified how the combined effects of Colec12, Trail, and DAN observed in vitro can give rise to robust confinement of CNCC trajectories in vivo (Johnson et al. 2026a).

Additional modeling work in this context has further demonstrated that relaxing the assumption of a strict leader–follower phenotypic binary can still generate robust stream formation under appropriate conditions, and that leader–follower phenotypic heterogeneity may emerge spontaneously during VEGF-driven migration (Schumacher 2019b; Johnson et al. 2026b). As such, it remains an open question as to how well the phenotypes of CNCCs in chick conform to a strict binary phenotypic division, or whether CNCC phenotypes vary in a continuous manner within streams.

Collectively, these studies underscore the strength of tightly coupled mathematical and experimental methods. Mathematical modeling has played a central role in formalizing mechanistic hypotheses, revealing inconsistencies with experimental observations, and generating testable predictions. In parallel, experiments have been essential for validating model behavior and identifying new molecular regulators in CNCC migration. It is this sustained interplay between theory and data that has enabled a coherent, multiscale understanding of chick CNCC migration to emerge.

2.4 Conclusion and Outlook

In this review, we have examined how experimental and mathematical methods have been combined to advance our understanding of collective neural crest cell migration. By drawing together work across model organisms and mathematical modeling frameworks, we have positioned the neural crest as a particularly effective model system for uncovering the mechanisms underlying collective cell migration, owing to its robust and stereotyped collective behaviors alongside its ease of experimental perturbation.

The vast body of experimental literature has shown that in vivo, migration is shaped by combinations of contact inhibition of locomotion (CIL), co-attraction (CoA), chemotactic signaling, and proliferation, with the relative importance of each mechanism varying between axial subpopulations and model organisms. In addition, these mechanisms often act within collectives that are themselves phenotypically heterogeneous. Mathematical models have been of great importance in elucidating which of these processes are most important in many of these contexts and, in addition, have enabled the formulation of precise, testable hypotheses that would be difficult to isolate experimentally in complex embryonic environments.

Agent-based models have been particularly valuable in resolving how local cell–cell interactions give rise to large-scale migratory patterning. These models have demonstrated that coherent collective migration can emerge from short-range cell–cell and cell–environment interactions alone, while also showing that long-range cues can guide collectives and that phenotypic heterogeneity is, in some contexts, necessary for collective migration to occur. Continuum models have provided complementary insight at the level of the collective, especially in systems such as the vagal neural crest where proliferation plays a dominant role in invasion. This diverse body of work, therefore, illustrates that no single modeling framework captures all aspects of neural crest cell migration and that different approaches are required to interrogate processes operating at different spatial and temporal scales.

Despite significant progress, many open questions remain in the study of collective neural crest cell migration. One important direction for future work is the tighter quantitative integration of mathematical models and experiments. While many existing mathematical models reproduce observed behaviors qualitatively, few are fully parameterized using experimental data. Recent advances in live imaging and single-cell profiling now offer opportunities to address this limitation, but doing so will require the incorporation of statistical inference and uncertainty quantification into modeling–experimental workflows.

A second challenge concerns phenotypic heterogeneity and plasticity within migrating neural crest cell populations. Experimental evidence increasingly suggests that cell behaviors and molecular states vary continuously within collectives and dynamically in time, raising questions regarding how phenotypic transitions are regulated in vivo and how they influence collective behavior. Models that allow for continuous cell states, rather than, for example, fixed leader and follower identities, are likely to be important for reconciling observations across systems and for understanding how collective migration remains robust to perturbation.

Mathematical models of neural crest cell migration should also give greater consideration to the tissues through which migration occurs. In vivo, neural crest cell migration takes place within mechanically active, three-dimensional environments that comprise complex ECM organization and that reciprocally interact with cells. Extending existing models to account for tissue mechanics and dynamic geometries will be essential for further linking cell-scale environmental interactions to emergent macroscale collective behaviors.

More broadly, the mechanistic insights derived from neural crest cell migration, and in particular those formalized and tested through mathematical modeling, provide a basis for hypothesis-driven comparisons with other systems that exhibit collective cell migration. Testing whether similar interaction rules govern other processes of collective cell migration, such as wound healing, angiogenesis, or tumor invasion, will help to distinguish the universal principles of collective cell behavior from context-specific mechanisms.

In conclusion, the integration of experimental and mathematical studies has substantially advanced our understanding of collective neural crest cell migration. Continued progress in uncovering the biological mechanisms underlying this process will likely depend on a closer quantitative coupling between theory and experiment, as well as the development of mathematical models that better capture cellular heterogeneity and cell–tissue interactions. In this regard, neural crest cell migration provides a tractable and well-characterized setting in which mechanistic hypotheses pertaining to collective cell behavior can be formulated, quantitatively tested, and compared across experimental and modeling frameworks.

Acknowledgments S.W.S.J. receives support from the Biotechnology and Biological Sciences Research Council (BBSRC) (grant number BB/T008784/1). R.E.B. is supported by a grant from the Simons Foundation (MP-SIP-00001828). All schematics were created in BioRender and are separately licensed under CC BY 4.0. During production of this work, the authors utilized an OpenAI large language model to assist with grammar, phrasing, and proofreading. The authors reviewed and edited all content as required and take full responsibility for all aspects of the work.

Conflict of Interest Statement The authors declare no conflict of interest.

References

Bahm I, Barriga EH, Frolov A, Theveneau E, Frankel P, Mayor R (2017) PDGF controls contact inhibition of locomotion by regulating N-cadherin during neural crest migration. Development 144(13):2456–2468

Barfurth D (1891) Zur Regeneration der Gewebe. Arch Mikrosk Anat 37(1):406–491

Broders-Bondon F, Paul-Gilloteaux P, Gazquez E, Heysch J, Piel M, Mayor R, Lambris JD, Dufour S (2016) Control of the collective migration of enteric neural crest cells by the complement anaphylatoxin C3a and N-cadherin. Dev Biol 414(1):85–99

Bronner ME, Le Douarin NM (2012) Development and evolution of the neural crest: an overview. Dev Biol 366(1):2–9

Burns AJ, Le Douarin NM (1998) The sacral neural crest contributes neurons and glia to the post-umbilical gut: spatiotemporal analysis of the development of the enteric nervous system. Development 125(21):4335–4347

Carmona-Fontaine C, Matthews HK, Kuriyama S, Moreno M, Dunn GA, Parsons M, Stern CD, Mayor R (2008) Contact inhibition of locomotion in vivo controls neural crest directional migration. Nature 456(7224):957–961

Carmona-Fontaine C, Theveneau E, Tzekou A, Tada M, Woods M, Page KM, Parsons M, Lambris JD, Mayor R (2011) Complement fragment C3a controls mutual cell attraction during collective cell migration. Dev Cell 21(6):1026–1037

Chen W, Xia P, Wang H, Tu J, Liang X, Zhang X, Li L (2019) The endothelial tip-stalk cell selection and shuffling during angiogenesis. J Cell Commun Signaling 13(3):291–301

Cheung KJ, Ewald AJ (2016) A collective route to metastasis: seeding by tumor cell clusters. Science 352(6282):167–169

Cordero DR, Brugmann S, Chu Y, Bajpai R, Jame M, Helms JA (2011) Cranial neural crest cells on the move: their roles in craniofacial development. Am J Med Genet A 155(2):270–279

de Bellard ME, Bronner-Fraser M (2005) Neural crest migration methods in the chicken embryo. In: Cell migration: developmental methods and protocols. Humana Press, pp 247–267

de Bellard ME, Ortega B, Sao S, Kim L, Herman J, Zuhdi N (2018) Neuregulin-1 is a chemoattractant and chemokinetic molecule for trunk neural crest cells. Dev Dyn 247(7):888–902

Deisboeck TS, Couzin ID (2009) Collective behavior in cancer cell populations. BioEssays 31(2):190–197

DeSimone DW, Davidson L, Marsden M, Alfandari D (2005) The *Xenopus* embryo as a model system for studies of cell migration. In: Cell migration: developmental methods and protocols. Humana Press, pp 235–245

Donoghue PCJ, Graham A, Kelsh RN (2008) The origin and evolution of the neural crest. BioEssays 30(6):530–541

Dyson L, Holmes A, Li A, Kulesa PM (2018) A chemotactic model of trunk neural crest cell migration. Genesis 56(9):e23239

Eycleshymer AC (1907) The closing of wounds in the larval *Necturus*. Am J Anat 7(2):317–325

Freingruber V, Painter KJ, Ptashnyk M, Schumacher LJ (2024) A biased random walk approach for modeling the collective chemotaxis of neural crest cells. J Math Biol 88(3):32

Friedl P, Gilmour D (2009) Collective cell migration in morphogenesis, regeneration and cancer. Nat Rev Mol Cell Biol 10(7):445–457

Gammill LS, Roffers-Agarwal J (2010) Division of labor during trunk neural crest development. Dev Biol 344(2):555–565

Gershon MD, Chalazonitis A, Rothman TP (1993) From neural crest to bowel: development of the enteric nervous system. J Neurobiol 24(2):199–214

Giniūnaitė R, Baker RE, Kulesa PM, Maini PK (2020) Modelling collective cell migration: neural crest as a model paradigm. J Math Biol 80(1–2):481–504

Hutchins EJ, Kunttas E, Piacentino ML, Howard AGA IV, Bronner ME, Uribe RA (2018) Migration and diversification of the vagal neural crest. Dev Biol 444:S98–S109

Johnson SWS, Parsons M, Baker RE, Maini PK (2025) An energy-based mathematical model of actin-driven protrusions in eukaryotic chemotaxis. arXiv preprint. https://arxiv.org/abs/2509.20303

Johnson SWS, Kulesa PM, Baker RE, Maini PK (2026a) Mathematical modeling predicts novel mechanisms of stream confinement from Trail/Colec12/Dan in the collective migration of cranial neural crest cells. Dev Dyn 255(1):43–65

Johnson SWS, Baker RE, Maini PK (2026b) The spontaneous emergence of leaders and followers in a mathematical model of cranial neural crest cell migration. arXiv preprint. https://arxiv.org/abs/2601.00374

Kasemeier-Kulesa JC, McLennan R, Romine MH, Kulesa PM, Lefcort F (2010) CXCR4 controls ventral migration of sympathetic precursor cells. J Neurosci 30(39):13078–13088

Khataee H, Czirok A, Neufeld Z (2021) Contact inhibition of locomotion generates collective cell migration without chemoattractants in an open domain. Phys Rev E 104(1):014405

Knight RD, Schilling TF (2006) Cranial neural crest and development of the head skeleton. Adv Exp Med Biol 589:120–133

Kubota Y, Morita T, Kusakabe M, Sakakura T, Ito K (1999) Spatial and temporal changes in chondroitin sulfate distribution in the sclerotome play an essential role in the formation of migration patterns of mouse neural crest cells. Dev Dyn 214(1):55–65

Kulesa PM, Gammill LS (2010) Neural crest migration: patterns, phases and signals. Dev Biol 344(2):566–568

Kulesa P, Ellies DL, Trainor PA (2004) Comparative analysis of neural crest cell death, migration, and function during vertebrate embryogenesis. Dev Dyn 229(1):14–29

Kulesa PM, Bailey CM, Kasemeier-Kulesa JC, McLennan R (2010) Cranial neural crest migration: new rules for an old road. Dev Biol 344(2):543–554

Landolt RM, Vaughan L, Winterhalter KH, Zimmermann DR (1995) Versican is selectively expressed in embryonic tissues that act as barriers to neural crest cell migration and axon outgrowth. Development 121(8):2303–2312

Le Douarin N, Kalcheim C (1999) The neural crest. Cambridge University Press

Lemon WC, McDole K (2020) Live-cell imaging in the era of too many microscopes. Curr Opin Cell Biol 66:34–42

Li L, He Y, Zhao M, Jiang J (2013) Collective cell migration: implications for wound healing and cancer invasion. Burns Trauma 1(1):21–26

Martinson WD, McLennan R, Teddy JM, McKinney MC, Davidson LA, Baker RE, Byrne HM, Kulesa PM, Maini PK (2023) Dynamic fibronectin assembly and remodeling by leader neural crest cells prevents jamming in collective cell migration. eLife 12:e83792

Mayor R, Theveneau E (2013) The neural crest. Development 140(11):2247–2251

McKinney MC, McLennan R, Giniūnaitė R, Baker RE, Maini PK, Othmer HG, Kulesa PM (2020) Visualizing mesoderm and neural crest cell dynamics during chick head morphogenesis. Dev Biol 461(2):184–196

McLennan R, Teddy JM, Kasemeier-Kulesa JC, Romine MH, Kulesa PM (2010) Vascular endothelial growth factor (VEGF) regulates cranial neural crest migration in vivo. Dev Biol 339(1):114–125

McLennan R, Dyson L, Prather KW, Morrison JA, Baker RE, Maini PK, Kulesa PM (2012) Multiscale mechanisms of cell migration during development: theory and experiment. Development 139(16):2935–2944

McLennan R, Schumacher LJ, Morrison JA, Teddy JM, Ridenour DA, Box AC, Semerad CL, Li H, McDowell W, Kay D et al (2015) VEGF signals induce trailblazer cell identity that drives neural crest migration. Dev Biol 407(1):12–25

McLennan R, Bailey CM, Schumacher LJ, Teddy JM, Morrison JA, Kasemeier-Kulesa JC, Wolfe LA, Gogol MM, Baker RE, Maini PK et al (2017) DAN (NBL1) promotes collective neural crest migration by restraining uncontrolled invasion. J Cell Biol 216(10):3339–3354

McLennan R, McKinney MC, Teddy JM, Morrison JA, Kasemeier-Kulesa JC, Ridenour DA, Manthe CA, Giniūnaitė R, Robinson M, Baker RE et al (2020) Neural crest cells bulldoze through the microenvironment using Aquaporin 1 to stabilize filopodia. Development 147(1):dev185231

McLennan R, Giniūnaitė R, Hildebrand K, Teddy JM, Kasemeier-Kulesa JC, Bolanos L, Baker RE, Maini PK, Kulesa PM (2023) Colec12 and Trail signaling confine cranial neural crest cell trajectories and promote collective cell migration. Dev Dyn 252(5):629–646

Moiseeva A, Nikolenko V, Oganesyan M, Nikitina A, Rizaeva N, Zharikova T, Pontes-Silva A, Zharikov Y (2025) Neural crest cells: bridging embryology and regenerative medicine. Neuroscience 579:259–266

Monsonego-Ornan E, Kosonovsky J, Bar A, Roth L, Fraggi-Rankis V, Simsa S, Kohl A, Sela-Donenfeld D (2012) Matrix metalloproteinase 9/gelatinase B is required for neural crest cell migration. Dev Biol 364(2):162–177

Moore R, Theveneau E, Pozzi S, Alexandre P, Richardson J, Merks A, Parsons M, Kashef J, Linker C, Mayor R (2013) Par3 controls neural crest migration by promoting microtubule catastrophe during contact inhibition of locomotion. Development 140(23):4763–4775

Morrison JA, McLennan R, Wolfe LA, Gogol MM, Meier S, McKinney MC, Teddy JM, Holmes L, Semerad CL, Box AC et al (2017) Single-cell transcriptome analysis of avian neural crest migration reveals signatures of invasion and molecular transitions. eLife 6:e28415
Moses L, Pachter L (2022) Museum of spatial transcriptomics. Nat Methods 19(5):534–546
Painter KJ, Bloomfield JM, Sherratt JA, Gerisch A (2015) A nonlocal model for contact attraction and repulsion in heterogeneous cell populations. Bull Math Biol 77(6):1132–1165
Pasiliao CC, Thomas EC, Yung T, Zhu M, Tao H, Sun Y, Goyal S, Hopyan S (2025) A thermodynamic perspective on mammalian neural crest ingression. Proc Natl Acad Sci 122(36):e2504185122
Perris R, Perissinotto D (2000) Role of the extracellular matrix during neural crest cell migration. Mech Dev 95(1–2):3–21
Perris R, Paulsson M, Bronner-Fraser M (1989) Molecular mechanisms of avian neural crest cell migration on fibronectin and laminin. Dev Biol 136(1):222–238
Poujade M, Grasland-Mongrain E, Hertzog A, Jouanneau J, Chavrier P, Ladoux B, Buguin A, Silberzan P (2007) Collective migration of an epithelial monolayer in response to a model wound. Proc Natl Acad Sci 104(41):15988–15993
Ridenour DA, McLennan R, Teddy JM, Semerad CL, Haug JS, Kulesa PM (2014) The neural crest cell cycle is related to phases of migration in the head. Development 141(5):1095–1103
Rørth P (2009) Collective cell migration. Annu Rev Cell Dev Biol 25(1):407–429
Saliba A-E, Westermann AJ, Gorski SA, Vogel J (2014) Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res 42(14):8845–8860
Sato TS, Handa A, Priya S, Watal P, Becker RM, Sato Y (2019) Neurocristopathies: enigmatic appearances of neural crest cell–derived abnormalities. Radiographics 39(7):2085–2102
Schumacher L (2019a) Collective cell migration in development. Adv Exp Med Biol 1146:105–116
Schumacher LJ (2019b) Neural crest migration with continuous cell states. J Theor Biol 481:84–90
Schumacher LJ, Kulesa PM, McLennan R, Baker RE, Maini PK (2016) Multidisciplinary approaches to understanding collective cell migration in developmental biology. Open Biol 6(6):160056
Shellard A, Mayor R (2016) Chemotaxis during neural crest migration. Semin Cell Dev Biol 55:111–118
Simpson MJ, Landman KA, Hughes BD, Newgreen DF (2006) Looking inside an invasion wave of cells using continuum models: proliferation is the key. J Theor Biol 243(3):343–360
Simpson MJ, Zhang DC, Mariani M, Landman KA, Newgreen DF (2007a) Cell proliferation drives neural crest cell invasion of the intestine. Dev Biol 302(2):553–568
Simpson MJ, Merrifield A, Landman KA, Hughes BD (2007b) Simulating invasion with cellular automata: connecting cell-scale and population-scale properties. Phys Rev E 76(2):021918
Smith A, Robinson V, Patel K, Wilkinson DG (1997) The EphA4 and EphB1 receptor tyrosine kinases and ephrin-B2 ligand regulate targeted migration of branchial neural crest cells. Curr Biol 7(8):561–570
Snider TN, Mishina Y (2014) Cranial neural crest cell contribution to craniofacial formation, pathology, and future directions in tissue engineering. Birth Defects Res C Embryo Today 102(3):324–332
Stoller JZ, Epstein JA (2005) Cardiac neural crest. Semin Cell Dev Biol 16:704–715
Stramer B, Mayor R (2017) Mechanisms and in vivo functions of contact inhibition of locomotion. Nat Rev Mol Cell Biol 18(1):43–55
Szabó A, Mayor R (2016) Modelling collective cell migration of neural crest. Curr Opin Cell Biol 42:22–28
Szabó A, Mayor R (2018) Mechanisms of neural crest migration. Annu Rev Genet 52:43–63
Szabó A, Melchionda M, Nastasi G, Woods ML, Campo S, Perris R, Mayor R (2016) In vivo confinement promotes collective migration of neural crest cells. J Cell Biol 213(5):543–555
Teddy JM, Kulesa PM (2004) In vivo evidence for short- and long-range cell communication in cranial neural crest cells. Development 131:6141–6151
Theveneau E, Mayor R (2013) Collective cell migration of epithelial and mesenchymal cells. Cell Mol Life Sci 70(19):3481–3492

Theveneau E, Steventon B, Scarpa E, Garcia S, Trepat X, Streit A, Mayor R (2013) Chase-and-run between adjacent cell populations promotes directional collective migration. Nat Cell Biol 15(7):763–772

Tsang KY, Cheung MCH, Chan D, Cheah KSE (2010) The developmental roles of the extracellular matrix: beyond structure to regulation. Cell Tissue Res 339(1):93–110

Vega-Lopez GA, Cerrizuela S, Aybar MJ (2017) Trunk neural crest cells: formation, migration and beyond. Int J Dev Biol 61(1–2):5–15

Vega-Lopez GA, Cerrizuela S, Tribulo C, Aybar MJ (2018) Neurocristopathies: new insights 150 years after the neural crest discovery. Dev Biol 444:S110–S143

Vishwakarma M, Di Russo J, Probst D, Schwarz US, Das T, Spatz JP (2018) Mechanical interactions among followers determine the emergence of leaders in migrating epithelial cell collectives. Nat Commun 9(1):3469

Wang X, Chan AKK, Sham MH, Burns AJ, Chan WY (2011) Analysis of the sacral neural crest cell contribution to the hindgut enteric nervous system in the mouse embryo. Gastroenterology 141(3):992–1002.e6

Woods ML, Carmona-Fontaine C, Barnes CP, Couzin ID, Mayor R, Page KM (2014) Directional collective cell migration emerges as a property of cell interactions. PLoS One 9(9):e104969

Wynn ML, Kulesa PM, Schnell S (2012) Computational modelling of cell chain migration reveals mechanisms that sustain follow-the-leader behaviour. J R Soc Interface 9(72):1576–1588

Yu Q, Liu L, Du M, Müller D, Gu Y, Gao Z, Xin X, Gu Y, He M, Marquardt T et al (2024) Sacral neural crest-independent origin of the enteric nervous system in mouse. Gastroenterology 166(6):1085–1099

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 3
Mechanisms and Evolution of Color Pattern Formation in Female-Limited Batesian Mimicry of *Papilio* Butterflies

Haruhiko Fujiwara

Abstract Complex color patterns are frequently observed across diverse insect taxa and play crucial roles in predator avoidance, including mimicry and camouflage. Several *Papilio* butterflies exhibit female-limited Batesian mimicry, a phenomenon that has attracted biological interest since the era of Darwin and Wallace. However, the underlying mechanisms and evolutionary origins of this phenomenon remained unclear for many years. In this review, I compare the genomic regions responsible for female-limited Batesian mimicry among several *Papilio* species. In each species, the responsible chromosomal locus exhibits a "supergene" architecture, although their internal structures differ. I summarize recent findings on the genetic mechanisms underlying mimetic color pattern formation and discuss the evolutionary processes that have shaped this remarkable form of mimicry.

Keywords *Papilio polytes* · *Papilio memnon* · *Papilio* butterflies · Wing color pattern · Pattern formation · Female-limited Batesian mimicry · Supergene · Gene regulation · Chromosomal inversion · Evolution

3.1 Research History of Female-Limited Batesian Mimicry

Batesian mimicry—where palatable animals imitate the appearance of distantly related, unpalatable, or toxic species—is widely observed across the animal kingdom (Mallet and Joron 1999; Ruxton et al. 2004; Kunte 2009a). For example, a bird that has previously experienced consuming a distasteful butterfly with aposematic coloration tends to avoid palatable butterflies with similar wing patterns. As a result, palatable butterflies that display Batesian mimicry reduced predation pressure. This phenomenon was first reported more than 150 years ago by the British naturalist H. W. Bates (1862).

H. Fujiwara (✉)
University of Tokyo, Tokyo, Japan
e-mail: haruh@edu.k.u-tokyo.ac.jp

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_3

A contemporary of Bates, A. R. Wallace, who explored with him in South America and later in Southeast Asia, described an exceptional example of Batesian mimicry in swallowtail butterflies in *The Malay Archipelago* (Wallace 1869). In the Great Mormon butterfly (*Papilio memnon*), mimetic females resemble a model species, whereas nonmimetic females display entirely different patterns. Males exhibit a more subdued pattern that differs from both female types. Wallace noted that the mimetic and nonmimetic females differ so dramatically that they were once considered distinct species. In *P. memnon*, the model species and mimetic females possess tails on their hindwings, while nonmimetic females and males lack them (Fig. 3.1b).

A closely related species, the Common Mormon (*Papilio polytes*), displays a similar dimorphism: females occur in both mimetic and nonmimetic forms, whereas males resemble the nonmimetic females. Unlike *P. memnon*, however, all adults of *P. polytes* possess hindwing tails. In both species, only females exhibit Batesian mimicry (Fig. 3.1b). Although the hindwing patterns of mimetic females differ slightly between the two species, the patterns of nonmimetic females differ substantially, and males of *P. memnon* display patterns distinct from all others.

Classical genetic studies demonstrated that the mimetic and nonmimetic phenotypes of females in these butterflies are governed by a single autosomal locus (termed *A* in *P. memnon* and *H* in *P. polytes* in this chapter) (Clarke et al. 1968; Clarke and Sheppard 1971, 1972). Mimetic alleles (*A* or *H*) are dominant, whereas nonmimetic alleles (*a* or *h*) are recessive. More than 90 years ago, the eminent statistician R. A. Fisher proposed that this single locus must contain multiple tightly linked genes involved in mimicry (Fisher 1930). Fisher considered that these complex adaptive traits showing intraspecific polymorphism are likely controlled by a cluster of multiple genes in a chromosomal region maintained by recombination suppression—similar to sex chromosomes. Such regions were later termed "supergenes" (Darlington and Mather 1949). Approximately 50 years after Fisher's prediction, butterfly geneticists Clark and Sheppard provided genetic evidence

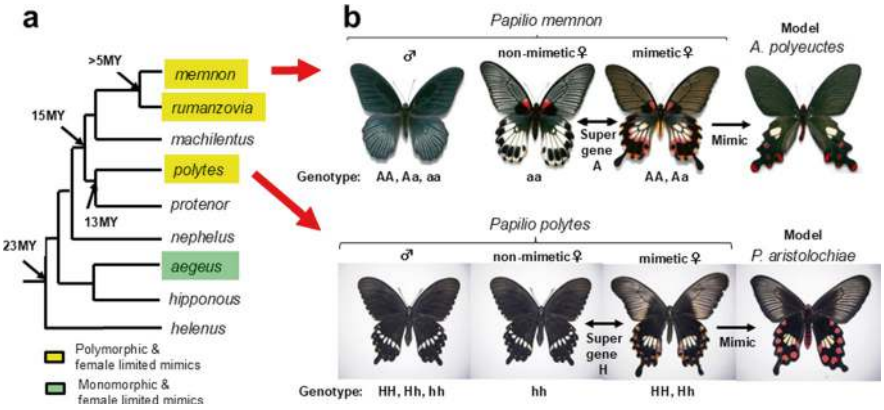


Fig. 3.1 Female-limited Batesian mimicry of *Papilio* butterflies. (**a**) Phylogeny of *Papilio* species in the Menelaides subtribe. (**b**) Photographs of *P. polytes*, *P. Memnon*, and their model species

supporting this hypothesis (Clarke et al. 1968; Clarke and Sheppard 1971, 1972). Nevertheless, the detailed structure and function of the supergene remained unclear until recently, when genomic information and technologies finally made such investigations possible.

3.2 Phylogenetic Distribution of Female-Limited Batesian Mimicry in *Papilio* Butterflies

Most swallowtail butterflies of the genus *Papilio* do not exhibit mimetic coloration, but several species display Batesian mimicry (Kunte 2009a, b; Timmermans et al. 2017). Cases include species in which both sexes mimic the same (or different) model species, as well as species in which only females mimic. Notably, no species has been reported in which mimicry is restricted to males. One possible explanation for female-limited mimicry is that females, being heavier due to their egg-laying organs and flying more slowly than males, are more susceptible to predation; thus, mimicry provides a stronger selective advantage to females (Ohsaki 2005; Komata et al. 2020; Yoshioka et al. 2025).

Female-limited mimicry occurs in two forms:

1. Monomorphic mimicry, in which all females mimic the same model species, and
2. Polymorphic mimicry, in which multiple female forms exist, only some of which mimic a model (Fig. 3.1a).

P. polytes and *P. memnon* fall into the latter category. Previous studies have shown that species exhibiting female-limited Batesian mimicry are not monophyletic but instead occur broadly and discontinuously across the *Papilio* phylogeny (Kunte 2009a; Komata et al. 2016, 2022a). This suggests that such mimicry did not originate from a single ancestral species but instead evolved independently in multiple lineages. The scattered distribution of female-limited mimicry likely reflects the involvement of diverse molecular mechanisms, making its evolutionary origins especially intriguing.

P. polytes and *P. memnon* belong to the monophyletic subtribe Menelaides (Fig. 3.1a). Within this group, some species are nonmimetic, whereas others—such as *P. rumanzovia*, a close relative of *P. memnon*—exhibit female dimorphic Batesian mimicry. *P. argeus* displays monomorphic mimicry (Kunte 2009b; Palmer and Kronforst 2020; Komata et al. 2022a). Even within these closely related species, wing pattern formation and the evolution of adaptive mimicry patterns vary remarkably.

Both *P. polytes* and *P. memnon* are widely distributed across Asia. In Japan, *P. polytes* inhabits Okinawa and other southwestern islands. Its model species is thought to be *Pachliopta aristolochiae*, a species found in the same region and sharing similar wing patterns (Fig. 3.1b). In contrast, in Japan only the nonmimetic females and males of *P. memnon* are present; mimetic females do not occur (Iijima

et al. 2018; Komata et al. 2022a). In recent years, possibly due to global warming, nonmimetic females (genotype *aa*) have been observed even in the Tokyo area, although the mimetic (*AA* or *Aa*) forms remain absent. This pattern may indicate that the mimetic allele *A* is maladapted to colder climates. The model species of *P. memnon* is considered to be *Atrophaneura polyeuctes*, which inhabits Taiwan and other regions and shares similar wing patterns (Iijima et al. 2018; Komata et al. 2022a).

3.3 Responsible Regions for Female-Limited Batesian Mimicry in Two *Papilio* Species

To elucidate the supergene responsible for female-limited Batesian mimicry in *P. polytes* and *P. memnon*, we performed whole-genome sequencing analyses. High-quality genome assemblies were essential for identifying the supergene, which spans a large chromosomal region, and for determining whether chromosomal inversions are involved. To minimize heterogeneity inherent in short-read next-generation sequencing (NGS), we extracted genomic DNA from a single *P. polytes* female larva reared continuously under laboratory conditions (Nishikawa et al. 2015). At the larval stage, it was impossible to determine whether the individual would develop into a mimetic (*HH* or *Hh*) or nonmimetic (*hh*) adult. Fortunately, the larva turned out to be heterozygous (*Hh*), enabling us to obtain sequence reads derived from both the *H* and *h* chromosomes.

Linkage analysis using RAPD polymorphisms, combined with a more detailed SNP-based genome-wide association study (GWAS), indicated that the region responsible for mimicry lies within a ~150 kb interval surrounding the doublesex (*dsx*) gene on chromosome 25 (Nishikawa et al. 2015). Strikingly, only this region exhibited a halved sequencing coverage depth relative to the rest of the genome—an indication of substantial divergence between the two homologous chromosomes. Aside from the sex chromosomes (ZW), this was the only autosomal region showing such extreme allelic differentiation (Nishikawa et al. 2015; Komata et al. 2022a). Given that linkage mapping also pointed to this region, it became evident that the genomic segment containing *dsx* on chromosome 25 constitutes the mimicry-determining region in *P. polytes*.

Critically, the sequences of the *H* and *h* allele-specific regions differ dramatically (Fig. 3.2a). Comparative nucleotide analyses revealed only ~50–70% identity across the ~150 kb region containing *dsx*, and a complete inversion spanning the region outside *dsx* was detected between *H* and *h*. This inversion likely suppressed recombination between homologous chromosomes for extended evolutionary periods, enabling the accumulation of numerous mutations and ultimately generating a highly diverged segment. We refer to this region as the highly diversified region (HDR). These findings are consistent with Fisher's classical prediction of a supergene maintained by recombination suppression and resolve a longstanding mystery of over 90 years (Kunte et al. 2014; Nishikawa et al. 2015).

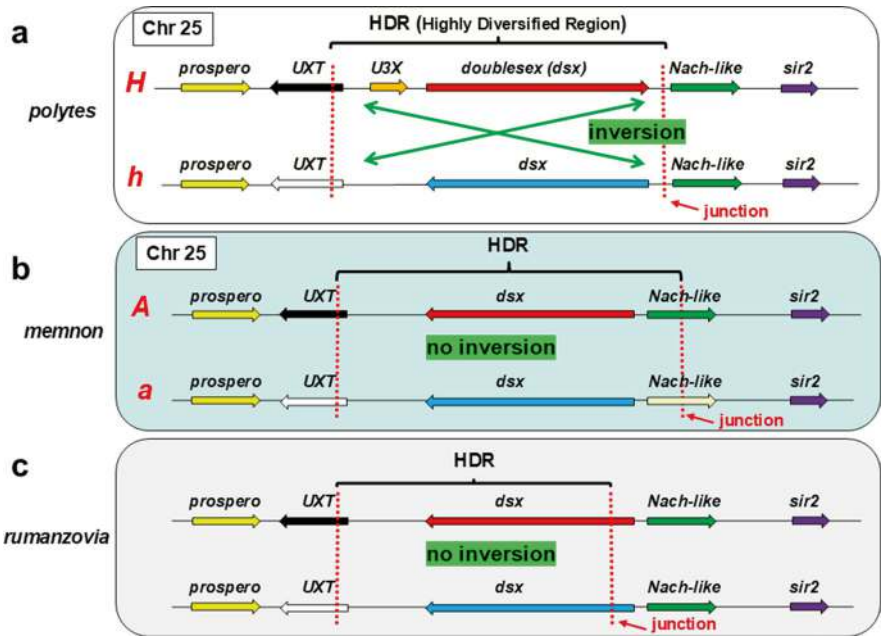


Fig. 3.2 Mimicry supergene structure in three *Papilio* species: *P. polytes* (**a**), *P. memnon* (**b**), and *P. rumanzovia* (**c**)

A few years later, genomic analyses and GWAS approaches were used to identify the mimicry region and assess the presence of a supergene structure in *P. memnon* (Iijima et al. 2018). As expected, Batesian mimicry in this species mapped to a region in chromosome 25—the corresponding location in *P. polytes*—and this region also showed pronounced divergence between the *A* (mimetic) and *a* (nonmimetic) alleles (Fig. 3.2b). Despite these similarities, a surprising finding emerged: no chromosomal inversion was detected in *P. memnon*. Thus, recombination suppression between *A* and *a* must be mediated by a mechanism other than inversion. This alternative mechanism is considered later in the section on evolutionary processes.

3.4 Structural Comparison of Batesian Mimicry Supergenes Between Two *Papilio* Butterflies

Although the mimicry supergenes of *P. polytes* and *P. memnon* occupy homologous regions on chromosome 25, their internal architectures differ markedly (Fig. 3.2). In the nonmimetic *h*-derived HDR (h-HDR) of *P. polytes*, the doublesex (*dsx*) gene is

oriented in the same direction as in other lepidopteran species. In contrast, within the mimetic *H*-derived HDR (H-HDR), the *dsx* orientation is reversed because of two flanking inversions. These inversions appear to have affected structures of the three adjacent genes inside H-HDR: *dsx*, *UXT*, and *U3X* (Nishikawa et al. 2015; Komata et al. 2022a).

Dsx is a transcription factor that regulates sexual differentiation across a wide range of animals (Baker 1989). Typically, its amino acid sequence is conserved within a species. However, 14 amino acid differences exist between the *H*-derived and *h*-derived Dsx in *P. polytes* (Nishikawa et al. 2015), and four amino acid substitutions distinguish the *A*- and *a*-derived Dsx in *P. memnon* (Iijima et al. 2018). Most of these substitutions—except one located in the dimerization domain of H-derived Dsx—occur outside functional domains (the DNA-binding and dimerization domains). Furthermore, none of the substitution sites are shared between the two species. These observations suggest that amino acid changes in Dsx are not the primary cause of mimicry.

UXT, an evolutionarily conserved ubiquitously expressed prefoldin-like chaperone, is implicated in transcriptional regulation and other cellular processes (Zhou et al. 2015; Su et al. 2016), although its role in mimicry remains unclear. Interestingly, the left boundary of the HDR in both *P. polytes* and *P. memnon* lies within the 5′-untranslated region (UTR) of *UXT* (Fig. 3.2). Although the UXT protein sequence does not differ between mimetic (A/H) and nonmimetic (a/h) alleles, transcriptional regulatory differences likely modulate its expression.

U3X, a long noncoding RNA consisting of three exons, is located upstream of the *H*-derived *dsx* transcription start site in *P. polytes* but is absent from the *h* allele. U3X likely originated during the formation of the *H*-specific inversion, although its biological function remains unknown. A long noncoding RNA upstream of *dsx* in *Daphnia* regulates *dsx* expression (Kato et al. 2018; Perez et al. 2021), raising the possibility that U3X modulates *dsx-H* transcription. Notably, no *U3X*-like sequence is present in *P. memnon*.

The inversion in *P. polytes* spans from the 5′-UTR of *UXT* on the left to downstream of *dsx* on the right (Fig. 3.2). In *P. memnon*, by contrast, no inversion exists, yet an HDR exhibiting similar levels of divergence is observed between the *A* and *a* alleles. As in *P. polytes*, the left HDR boundary lies within the 5′-UTR of *UXT*; however, the right boundary extends downstream of *dsx* into the Nach-like gene, which encodes a protein resembling a sodium channel, but whose function remains unknown (Zelle et al. 2013; Latorre-Estivalis et al. 2021). Thus, the HDRs in *P. memnon* encompass *UXT*, *dsx*, and *Nach-like* (Iijima et al. 2018; Komata et al. 2022a). The following sections explore the evolutionary implications of the presence or absence of inversions and the emergence of HDRs in these related species.

3.5 Colors and Pattern Formation of Batesian Mimic *Papilio* Butterflies

3.5.1 Comparison of Color Pigments and Gene Expression Between Model and Mimetic Wing Patterns

In Batesian mimicry, the color patterns, pigment composition, and wing shapes of mimetic and model species—despite being phylogenetically distant—appear remarkably similar. An important question is whether this similarity is achieved using homologous genes and pathways, or whether different molecular mechanisms have evolved independently. To address this, we compared the pigments used in the mimetic and nonmimetic wings of *P. polytes* with those in its model species *Pachliopta aristolochiae*, and analyzed the expression of associated pigment-synthesis genes (Nishikawa et al. 2013; Yoda et al. 2021) (Fig. 3.3).

In addition to the extensive black regions composed of melanin, these wings contain white (or pale yellow) and conspicuous red areas. Chemical analyses of scale pigments revealed that the band-shaped white patches on the hindwings of nonmimetic *P. polytes* contain papiliochrome II (PCII), a pigment composed of kynurenine and N-β-alanyl dopamine (NBAD) (Rembold and Umebachi 1985). In contrast, the pale-yellow central region characteristic of mimetic females appears to contain a chemically distinct pigment (Nishikawa et al. 2013; Yoda et al. 2021), which lacks kynurenine and is highly cross-linked—properties that hinder purification. The corresponding pale-yellow region in the model species *P. aristolochiae* is similar both in appearance and in chemical composition to that of the mimetic female of *P. polytes* (Nishikawa et al. 2013) (Fig. 3.3). These findings indicate that the white/pale-yellow patterns differ not only in shape but also in pigment biosynthetic pathways, which vary among nonmimetic females, mimetic females, and the model species.

Under ultraviolet (UV) illumination, the PCII-containing white band of nonmimetic *P. polytes* fluoresces, whereas the pale-yellow regions of mimetic females and the model species primarily reflect UV light (Yoda et al. 2021; Komata et al. 2023) (Figs. 3.3 and 3.5). Because birds and swallowtail butterflies can perceive UV wavelengths, these UV reflectance differences likely amplify the contrast between the mimetic and nonmimetic phenotypes. Consequently, the mimetic form of *P. polytes* may more effectively reduce predation by appearing more similar to the model species in the UV spectrum. Previous studies have reported mixed findings regarding mate preference: Uesugi (2000) found that male *P. polytes* prefer nonmimetic females resembling males in pattern, whereas other studies report differing tendencies (Uesugi 2000; Westerman et al. 2018, 2019). Given the prominent UV reflectance differences in hindwing patterns, these properties likely influence not only predator–prey interactions but also mate choice in *P. polytes*.

We also compared the red marginal spots of mimetic *P. polytes* with those of *P. aristolochiae* (Nishikawa et al. 2013) (Fig. 3.3a). Under UV light, the red spots of both species exhibited similar reflectance patterns. However, their pigment

a

Wing pattern / Pigment nature	White (pale yellow)			Red	
	Non-mimetic	Mimetic	Model	Mimetic	Model
					
Cross-linked	NO	YES	YES	YES	NO
Molecular weight	PII (874, 652, 430)	-	-	-	NBAD +688, +658
NBAD Pathway	YES	NO	NO	YES	NO
Kynurenine pathway	YES	YES	YES	YES	YES
UV reflection	NO	YES	YES	YES	YES
UV absorption Fluorescence	YES	NO	NO	NO	NO

b

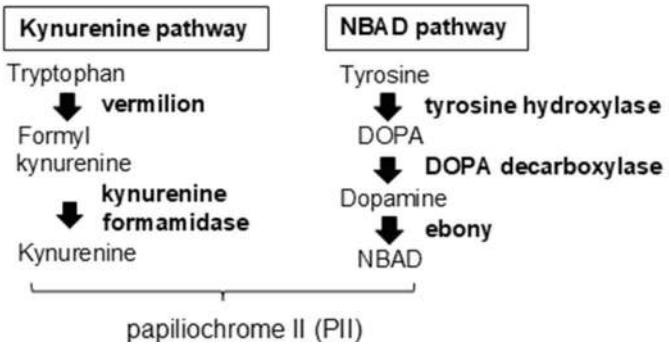


Fig. 3.3 Pigment characteristics in the hindwings of *P. polytes* and its model species. (**a**) Comparison of characteristics for white (pale yellow) and red pigments in the regions outlined in the wing photographs. (**b**) Enzymes involved in the kynurenine and NBAD pathways

chemistry differs: the red pigment of *P. aristolochiae* is not cross-linked and can be solubilized, whereas that of mimetic *P. polytes* is highly cross-linked and difficult to dissolve for chemical analysis (Umebachi 1978, 1985). Umebachi reported the latter pigment, termed papiliochrome R, which contains kynurenine and NBAD, while the pigment of *P. aristolochiae* ("red pigment A") lacks kynurenine (Umebachi 1978, 1985).

Gene expression analyses in the red spot regions during the pupal stage revealed that all three NBAD-pathway genes were expressed in both model and mimic. However, vermilion and kynurenine formamidase, key enzymes of the kynurenine

pathway, were not expressed in the red spots of *P. polytes*, consistent with pigment composition data (Nishikawa et al. 2013) (Fig. 3.3b). Recently, more detailed characterization of the red pigment in *P. memnon* revealed the presence of NBAD, sulfate groups, and an unidentified compound with a molecular mass of 388.1 (Komata et al. 2025). Collectively, these findings show that although mimetic and model species exhibit visually similar red and pale-yellow patterns, they use chemically distinct pigments and, likely, different downstream genetic pathways. This suggests that female-limited Batesian mimicry in *P. polytes* may represent a striking case of convergent evolution, in which different biochemical pathways produce similar adaptive phenotypes. Further comparative analyses will be necessary to clarify this hypothesis.

3.5.2 *Function of* dsx *in Wing Pattern Formation*

Many striking insect surface patterns—including mimicry—are found in non-model species, making functional analyses challenging. Although CRISPR/Cas9 genome editing has recently enabled mechanistic studies of complex traits, it remains technically demanding to rapidly assess gene function during wing-pattern formation. RNA interference (RNAi) is widely used to suppress gene expression, but in lepidopteran insects such as butterflies and moths, exogenous siRNAs do not readily enter cells, making efficient knockdown difficult (Terenius et al. 2011; Yamaguchi et al. 2011).

To address this limitation, we developed an electroporation-mediated method that introduces exogenous siRNAs and *piggyBac* transposons directly into epidermal cells. This technique enables rapid functional analyses during wing-pattern development (Ando and Fujiwara 2013; Fujiwara and Nishikawa 2016).

Using this system, siRNAs targeting the mimetic allele-specific *dsx-H* of *P. polytes* were introduced into the hindwings of mimetic females immediately after pupation. This knockdown (KD) caused dramatic alterations in the adult wing patterns (Nishikawa et al. 2015) (Fig. 3.4). Specifically, the pale-yellow central region characteristic of the mimetic phenotype was transformed into a nonmimetic pattern: the white band extended laterally, and the red marginal spots were reduced in size. Under UV illumination, the KD region fluoresced, indicating restoration of PCII—the pigment typical of the nonmimetic form (Yoda et al. 2021; Komata et al. 2023). Thus, *dsx-H* KD effectively converted mimetic wings into nonmimetic wings.

Similarly, knockdown of *dsx-A* in mimetic females of *P. memnon* transformed their wing patterns into those of nonmimetic females. Intriguingly, simultaneous KD of both *dsx-A* and *dsx-a* in heterozygous mimetic females (*Aa*) and in nonmimetic females (*aa*) produced wing patterns resembling those of males (Komata et al. 2022b). These results indicate that the male wing pattern represents the developmental default, whereas *dsx-a* induces the nonmimetic female pattern and *dsx-A* induces the mimetic female pattern. The same logic likely applies to *dsx-h* and *dsx-H* in *P. polytes*.

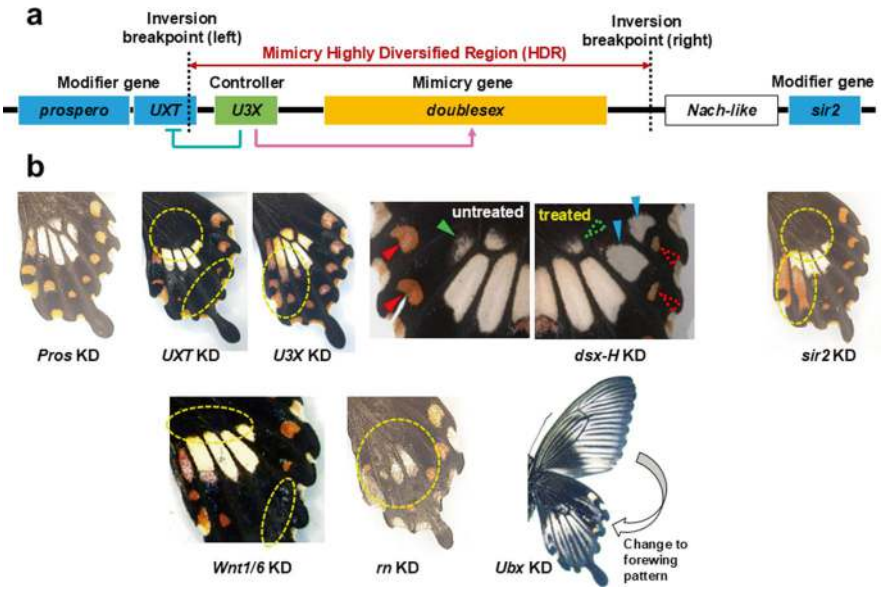


Fig. 3.4 Gene function of the mimicry supergene of *P. polytes*. (**a**) Structure of supergene. *U3X* is shown to upregulate *doublesex* and downregulate *UXT*. (**b**) Functional analyses using siRNA-based knockdown (KD) for each gene. The regions indicated by dashed outlines and arrowheads (and dashed arrowheads) represent areas showing changes in color patterns

A notable additional trait occurs in *P. memnon*: mimetic females exhibit an orange warning coloration in the abdomen. KD of *dsx-A* in the abdomen converted this orange coloration into the black coloration typical of nonmimetic females, indicating that *dsx-A* regulates mimetic traits in body regions beyond the wings (Komata et al. 2022b).

Butterfly wings consist of two forewings and two hindwings arranged bilaterally, with dorsal and ventral surfaces. Some species display four distinct patterns (forewing vs. hindwing; dorsal vs. ventral) (Nijhout 1991). It has long been known that Ultrabithorax (*Ubx*) determines hindwing identity (Ubx^+ for hindwings, Ubx^- for forewings), while apterous (*ap*) controls dorsal–ventral pattern differences (Fig. 3.4).

In *P. polytes*, mimetic females possess vertical white stripes on the forewings, which resemble those of the model species *P. aristolochiae*, whereas nonmimetic females lack these stripes (Fig. 3.1). When *dsx-H* was knocked down in the forewings of mimetic females, the affected striped regions shifted to nonmimetic black patterns (Yoda et al., unpublished), indicating that *dsx-H* contributes to mimetic pattern formation in both forewings and hindwings.

Knockdown of *Ubx* in the hindwings of mimetic females (i.e., converting them to a forewing-like Ubx^- state) resulted in patterns resembling mimetic forewings (Fig. 3.4). Furthermore, simultaneous KD of Ubx and *dsx-H* in hindwings produced patterns similar to the nonmimetic forewing phenotype (Yoda et al., unpublished).

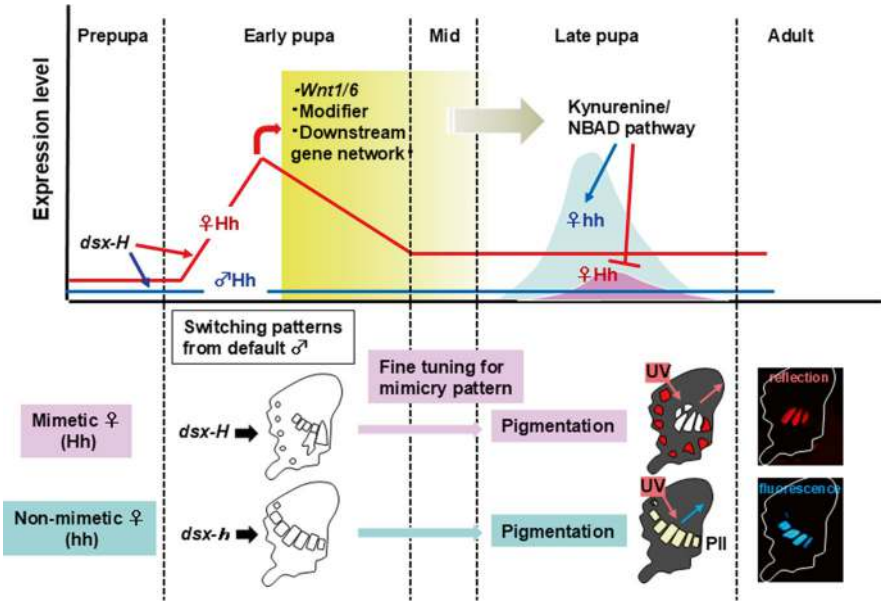


Fig. 3.5 Temporal changes of gene expression patterns in hindwings during pupal development

These findings reveal that the mimetic-patterning function of *dsx-H* operates under Ubx control, confirming a hierarchical gene-regulatory relationship.

Wing-pattern prepatterns are believed to form during the larval to prepupal stages (Nijhout 1991). Although our KD experiments were conducted immediately after pupation—later KD being technically challenging—they were capable of switching the final wing pattern from mimetic to nonmimetic. However, traits such as forewing/hindwing shapes (determined by Ubx) could not be altered at this stage. Notably, no intermediate forms were observed following *dsx-H* knockdown. This suggests that fundamental prepatterns for mimetic and nonmimetic wings, as well as forewing and hindwing identity, are established prior to pupation, and *Ubx* and *dsx-H* act post-prepattern to select among these predetermined alternatives (Fig. 3.5).

3.5.3 *Functional Involvement of Genes Inside and Outside the Mimicry HDR*

In *P. polytes*, the mimicry supergene appears to correspond closely to the HDR, because the region of marked sequence divergence between the *H* and *h* chromosomes coincides with the boundaries of the inversion. In *P. memnon*, however, although the HDR spans a region similar to that of *P. polytes*, no inversion is present. According to Fisher's classical supergene hypothesis, a mimicry supergene

should contain multiple tightly linked genes. In the present system, the HDR includes three candidates: (1) *dsx*, (2) *UXT*, and (3) *U3X* (in *P. polytes*) or *Nach-like* (in *P. memnon*).

This raises three questions:

1. Does the HDR itself constitute the entire supergene?
2. Are all genes within the HDR involved in mimicry?
3. Are genes flanking the HDR uninvolved, or might some also contribute?

As described earlier, *dsx* is the primary gene governing female-limited Batesian mimicry in both species. Therefore, we examined whether the other HDR genes also contribute to pattern formation (Komata et al. 2022a, 2023).

When *UXT* was knocked down in mimetic *P. polytes* hindwings, the central pale-yellow region—typically rounded in mimetic females—became flattened, resembling the nonmimetic phenotype. However, unlike in *dsx-H* KD, the characteristic white band did not extend laterally (Komata et al. 2023) (Fig. 3.4b). The red marginal spots were also reduced. Because these changes overlap with, but are weaker than, those induced by *dsx-H* KD, *UXT* likely functions as a modifier gene, refining the pattern produced by *dsx-H* (Fig. 3.4a).

Knockdown of *U3X* decreased the distance between the central pale-yellow region and the red marginal spots, suggesting that U3X influences the spatial arrangement of these pattern elements (Fig. 3.4b). Together, these results indicate that *dsx-H* acts as the core mimicry gene, while *UXT* and *U3X* modify its output to produce a more accurate resemblance to the model species. Recombination suppression within the HDR may have allowed these interacting genes to coevolve as a stable mimicry module (Fig. 3.4a).

Genes outside the HDR exhibit ~99% sequence identity between *H* and *h* (or *A* and *a*), suggesting that the HDR might define the functional supergene. To test this, we examined neighboring genes by KD experiments in mimetic females of *P. polytes* (Komata et al. 2023). Knockdown of Prospero, a neurogenesis-related homeobox gene located just outside of *UXT*, resulted in slight enlargement of the red marginal spots. Knockdown of Sir2, a NAD-dependent histone deacetylase located downstream of *Nach-like*, yielded more pronounced effects, flattening the upper portion of the central pale-yellow region and altering the shape of the red spots (Fig. 3.4b).

These findings demonstrate that HDR-flanking genes also contribute, albeit modestly, to mimicry pattern formation and may serve as additional modifier genes. Such results imply that regulatory elements within the HDR might influence the expression of neighboring genes, thereby extending the functional boundaries of the mimicry supergene. Consequently, the definition of the mimicry supergene may need to be reconsidered: the supergene may not correspond strictly to the HDR but instead encompass a broader set of genes functionally integrated through shared regulatory mechanisms.

3.5.4 *Female-Limited Expression of* dsx-H *and the Pleiotropic Function of Dsx*

In lepidopteran insects, females have ZW sex chromosomes and males have ZZ chromosomes, with the W chromosome acting as the primary female-determining chromosome. The doublesex (*dsx*) gene, located on an autosome, regulates sexual differentiation by producing male-specific and female-specific mRNA isoforms downstream of the sex-determining cascade. In *P. polytes* and *P. memnon*, the mimetic alleles (*H* and *A*) generate mimicry in females but do not induce mimetic traits in males, even when males carry the mimetic allele.

Initially, we hypothesized that sex-specific *dsx* isoforms differ structurally between males and females (Baker 1989). However, no amino acid differences were detected between the sex-specific forms, making this explanation unlikely. Instead, expression analyses revealed that *dsx-H* is strongly expressed in the hindwings of mimetic females (Hh) during the early pupal stage, whereas it is barely expressed in hindwings of males with the same genotype (Hh) (Komata et al. 2023)) (Fig. 3.5).

This suggests that a female-specific *cis*-regulatory element—recognizing transcription factors associated with sex determination—exists within the *H*-derived HDR, driving *dsx-H* expression only in female wings. Consequently, female-limited expression of *dsx-H* provides the mechanistic basis for female-limited mimicry.

Although Dsx-H and Dsx-A differ substantially in amino acid sequence from Dsx-h and Dsx-a, their ability to induce sexual differentiation remains intact. In *P. polytes* individuals homozygous for *H* (*HH*) and in *P. memnon* individuals homozygous for *A* (*AA*), both sexes develop normally, confirming that the "mimetic" versions of Dsx retain essential sex-differentiation functions.

Female dimorphism in *P. polytes* is often interpreted through a fitness cost hypothesis, in which the antipredator advantage of mimicry is balanced by physiological or reproductive disadvantages (Ohsaki 1995; Katoh et al. 2020; Komata et al. 2020). Our analyses comparing *P. polytes* females of genotypes *HH*, *Hh*, and *hh* revealed significant differences:

- Female lifespan: approx. 1 week (*HH*), 2 weeks (*Hh*), 4 weeks (*hh*)
- Male lifespan: approx. 2 weeks across all genotypes
- Fecundity and larval survival: reduced in females carrying the *H* allele (Komata et al. 2020)

These findings indicate that *Dsx-H* imposes certain physiological costs, although the precise mechanisms remain unknown. The coexistence of mimetic and nonmimetic females within populations likely reflects an evolutionary balance between the benefits of predator avoidance and the costs associated with *H*-allele–dependent traits.

3.5.5 Gene Networks Regulated by Mimetic Dsx Function

Dsx is a transcription factor, and in butterflies that exhibit female-limited Batesian mimicry, its downstream gene networks involved in wing and body patterning and morphogenesis, as well as sexual differentiation, are considered to have been acquired during evolution (Iijima et al. 2019; VanKuren et al. 2023, 2025). To clarify the gene network controlled by *dsx-H*, we knocked down its expression in mimetic female hindwings of *P. polytes* immediately after pupation (P0: day 0 pupa), then extracted RNA from the KD regions 1–3 days later for transcriptome analysis (Iijima et al. 2019) (Fig. 3.5).

This analysis identified numerous genes whose expression was either upregulated (i.e., normally repressed by Dsx-H) or downregulated (i.e., normally activated by Dsx-H). The Hox gene *AbdA*, which plays a key role in abdominal development in *Drosophila*, was upregulated in nonmimetic hindwings and downregulated in mimetic hindwings during the P1 stage (day 1 pupa). This suggests that suppression of *AbdA* by Dsx-H may contribute to the switch from nonmimetic to mimetic hindwing identity.

In contrast, *wnt1* and *wnt6*—encoding key secreted proteins in the Wnt signaling pathway involved in morphogenesis—were significantly downregulated in *dsx-H* KD wings. In nonmimetic hindwings, these genes are barely expressed, but in mimetic hindwings their expression is rapidly induced during the P1–P2 stages. Knockdown of *wnt1* and *wnt6* transformed mimetic females into a nonmimetic-like phenotype: the central pale-yellow region became flattened and the red spots diminished (Figs. 3.4b and 3.5). These results indicate that Dsx-H promotes mimicry partly by activating Wnt signaling components.

Functional analyses of *UXT* and *U3X* within the HDR revealed that their knockdown also reduced expression of *wnt1*, *wnt6*, and the zinc-finger transcription factor rotund (*rn*) (Komata et al. 2023). This suggests that induction of these genes constitutes a central component of the mimicry-specific gene network. Knockdown of *rn* severely disrupted mimetic patterns, converting the pale-yellow area to a UV-fluorescent state typical of nonmimetic pigment synthesis (Komata et al. 2023) (Fig. 3.4b). These findings indicate that Rn plays a major role in overall pattern organization and pigment-pathway switching.

Knockdown of *U3X* additionally decreased expression of the neighboring *dsx-H* while increasing expression of the adjacent *UXT* gene (Komata et al. 2023) (Fig. 3.4a). These observations suggest that U3X, a long noncoding RNA, helps activate *dsx-H* while suppressing *UXT*, reinforcing its position as a regulatory component within the HDR.

During the early- and mid-pupal stages, swallowtail wings lack pigmentation; pigment deposition occurs in the late pupal stages (Fig. 3.5). As described earlier, both the kynurenine pathway and the NBAD/melanin pathway contribute to generating nonmimetic pigments such as PCII (Fig. 3.3b). When *dsx-H* was knocked down in the pale-yellow central region of mimetic hindwings, expression levels of vermilion and kynurenine formamidase (kynurenine pathway) and ebony and

tyrosine hydroxylase (TH) (NBAD pathway) all increased (Yoda et al. 2021). This indicates that Dsx-H normally suppresses the enzyme suite responsible for PCII synthesis. Although the exact molecular mechanism through which Dsx-H represses these genes remains unknown, these results show that mimetic pigmentation is achieved by suppressing nonmimetic pigment pathways, coordinated through Dsx-H-mediated transcriptional regulation.

3.6 Evolutionary Process of Female-Limited Batesian Mimicry in *Papilio memnon* and *P. polytes*

One of the longstanding questions in evolutionary biology is how a complex adaptive trait such as female-limited Batesian mimicry—controlled by a supergene—originated and diversified. Even within the Menelaides subtribe, to which *P. polytes* and *P. memnon* belong, closely related species exhibit strikingly different forms of mimicry. As noted above, the mimicry supergene of *P. polytes* contains a large inversion, whereas *P. memnon* lacks such an inversion, raising the question of how recombination is suppressed in the latter (Figs. 3.1 and 3.2).

Recombination tends to be suppressed when homologous chromosomes accumulate extensive sequence divergence, which can disrupt proper chromosomal pairing during meiosis. This phenomenon is well known in sex chromosomes, where the substantial divergence between X and Y often results in recombination being restricted to small pseudoautosomal regions (Charlesworth 2016, 2023). In addition to inversions, the accumulation of transposable elements (TEs) on sex chromosomes also contributes to recombination suppression.

In *P. memnon*, approximately 1.5-fold more TEs accumulate within both *A* and *a* HDRs compared to other autosomal regions (Komata et al. 2022a). The TE accumulation likely contributes to recombination suppression by generating extensive local sequence divergence (Dolgin and Charlesworth 2008; Li et al. 2016; Kent et al. 2017). A similar pattern of TE enrichment is present in the H–h HDR of *P. polytes* (Komata et al. 2022a). Because recombination is suppressed, these TEs persist and accumulate over evolutionary time (Fig. 3.6).

Within the Menelaides subtribe, *P. rumanzovia*—a close relative of *P. memnon*—also exhibits dimorphic female-limited Batesian mimicry. Draft genome analyses revealed the presence of an HDR and mimicry supergene in this species as well, without any inversion, paralleling the situation in *P. memnon* (Palmer and Kronforst 2020; Komata et al. 2022a). SNP-based comparisons of the two chromosomes in *P. rumanzovia* showed that the left boundary of the recombination–suppression region is located within the 5′-UTR of *UXT*, consistent with *P. polytes* and *P. memnon*. However, the right boundary differs: in *P. memnon* it lies within the Nach-like gene, whereas in *P. rumanzovia* it lies within the internal region of *dsx*. Interestingly, the HDR of *P. rumanzovia* is the shortest among the three species. This conservation of the left boundary and divergence of the right boundary may provide important clues for reconstructing the evolutionary history of these supergenes.

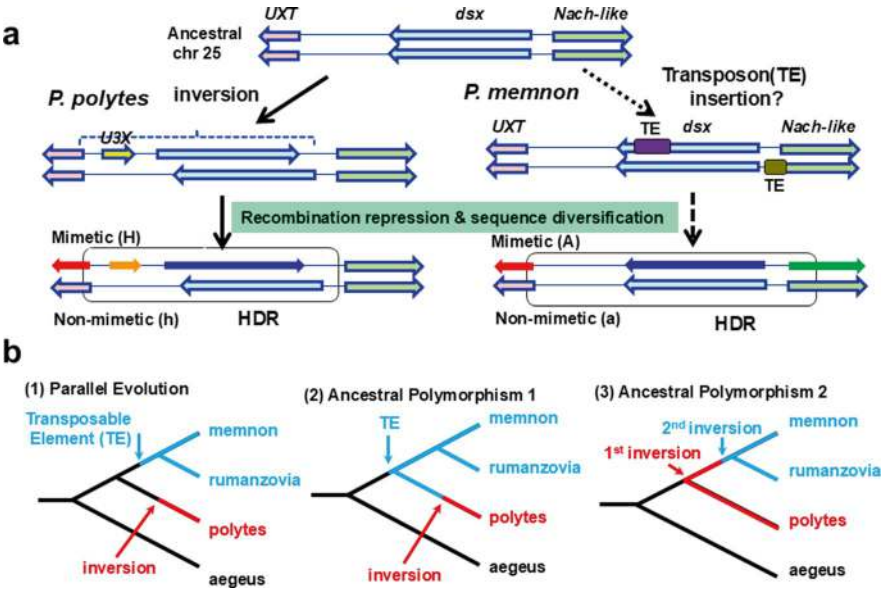


Fig. 3.6 Evolution of mimicry supergene among *Papilio* species. (**a**) Putative evolutional process of mimicry supergene in *P. polytes* and *P. memnon*. (**b**) Three proposed models for the evolution of the mimicry supergene in *Papilio* species

Based on these observations, several hypotheses can be proposed regarding the evolution of the mimicry supergene in the Menelaides subtribe (Komata et al. 2022a) (Fig. 3.6b):

1. *Ancestral polymorphism 1*: TE-driven recombination suppression predates inversions

   A common ancestor of the three species may have experienced TE accumulation in the shared region of chromosome 25, reducing recombination. Later, *P. polytes* alone may have acquired an inversion, further reinforcing recombination suppression.
2. *Ancestral polymorphism 2*: ancestral inversion followed by lineage-specific modifications

   An inversion may have originated in the common ancestor, but subsequently additional inversions or recombination events may have occurred in the *P. memnon–P. rumanzovia* lineage, eliminating or rearranging the original inversion.
3. *Parallel evolution in separate lineages*

   Given the substantial biological differences between *P. polytes* and *P. memnon*—including differences in the mimicry patterns of males and females and the presence or absence of a tail, we cannot rule out the possibility that the two species arose independently through parallel evolution.

More detailed genomic comparisons among closely related nonmimetic species, especially analyses of TE insertion patterns and recombination landscapes near the supergene, will improve our understanding of how female-limited Batesian mimicry evolved in these species.

References

Ando T, Fujiwara H (2013) Electroporation-mediated somatic transgenesis for rapid functional analysis in insects. Development 140:454–458. https://doi.org/10.1242/dev.085241

Baker BS (1989) Sex in flies: the splice of life. Nature 340:521–524

Bates HW (1862) Contributions to an insect fauna of the Amazon Valley (Lepidoptera: Heliconidae). Trans Linn Soc Lond 23:495–556

Charlesworth D (2016) The status of supergenes in the 21st century: recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. Evol Appl 9:74–90. https://doi.org/10.1111/eva.12291

Charlesworth D (2023) Why and how do Y chromosomes stop recombining? J Evol Biol 36:632–636. https://doi.org/10.1111/jeb.14137

Clarke CA, Sheppard PM (1971) Further studies on the genetics of the mimetic butterfly Papilio memnon L. Philos Trans R Soc Lond B 263:35–70

Clarke CA, Sheppard PM (1972) The genetics of the mimetic butterfly Papilio polytes L. Philos Trans R Soc Lond B 26:431–458

Clarke CA, Sheppard PM, Thornton IWB (1968) The genetics of the mimetic butterfly Papilio memnon L. Philos Trans R Soc Lond B 254:37–89

Darlington CD, Mather K (1949) Elements of genetics. George Allen and Unwin, London

Dolgin ES, Charlesworth B (2008) The effects of recombination rate on the distribution and abundance of transposable elements. Genetics 178:2169–2177

Fisher RA (1930) The Genetical theory of natural selection. Clarendon Press, Oxford

Fujiwara H, Nishikawa H (2016) Functional analysis of genes involved in color pattern formation in Lepidoptera. Curr Opin Insect Sci 17:16–23. https://doi.org/10.1016/j.cois.2016.05.015

Iijima T, Kajitani R, Komata S et al (2018) Parallel evolution of Batesian mimicry supergene in two Papilio butterflies, P. polytes and P. memnon. Sci Adv 4:eaao5416. https://doi.org/10.1126/sciadv.aao5416

Iijima T, Yoda S, Fujiwara H (2019) The mimetic wing pattern of Papilio polytes butterflies is regulated by a doublesex-orchestrated gene network. Commun Biol 2:257. https://doi.org/10.1038/s42003-019-0510-7

Kato Y, Perez CAG, Ishak NSM et al (2018) A 5′ UTR-overlapping lncRNA activates the male-determining gene doublesex1 in the crustacean Daphnia magna. Curr Biol 28:1811–1817

Katoh M, Tatsuta H, Tsuji K (2020) Mimicry genes reduce pre-adult survival rate in Papilio polytes: a possible new mechanism for maintaining female-limited polymorphism in Batesian mimicry. J Evol Biol 33:1487–1494. https://doi.org/10.1111/jeb.13686

Kent TV, Uzunović J, Wright SI (2017) Coevolution between transposable elements and recombination. Philos Trans R Soc Lond B 372:20160458

Komata S, Lin C-P, Iijima T et al (2016) Identification of doublesex alleles associated with the female-limited Batesian mimicry polymorphism in Papilio memnon. Sci Rep 6:34782. https://doi:10.1038/srep34782

Komata S, Kitamura T, Fujiwara H (2020) Batesian mimicry has evolved with deleterious effects of the pleiotropic gene doublesex. Sci Rep 10:21333

Komata S, Kajitani R, Itoh T, Fujiwara H (2022a) Genomic architecture and functional unit of mimicry supergene in female limited Batesian mimic Papilio butterflies. Philos Trans R Soc Lond Ser B Biol Sci 377:20210198. https://doi.org/10.1098/rstb.2021.0198

Komata S, Lin C-P, Fujiwara H (2022b) doublesex controls both hindwing and abdominal mimicry traits in the female-limited Batesian mimicry of Papilio memnon. Front Insect Sci 12:929518. https://doi.org/10.3389/finsc.2022.929518

Komata S, Yoda S, KonDo Y et al (2023) Functional unit of supergene in female-limited Batesian mimicry of Papilio polytes. Genetics 223:iyac177. https://doi.org/10.1093/genetics/iyac177

Komata S, Kajitani R, Yamabe T et al (2025) Whole genome analysis of toxic Papilionidae butterflies utilizing aristolochic acid, Pachliopta aristolochiae and Byasa alcinous. DNA Res. https://doi.org/10.1093/dnares/dsaf038

Kunte K (2009a) The diversity and evolution of Batesian mimicry in Papilio swallowtail butterflies. Evolution 63:2707–2716
Kunte K (2009b) Female-limited mimetic polymorphism: a review of theories and a critique of sexual selection as balancing selection. Anim Behav 78:1029–1036
Kunte K, Zhang W, Tenger-Trolander A et al (2014) doublesex is a mimicry supergene. Nature 507:229–232. https://doi.org/10.1038/nature13112
Latorre-Estivalis JM, Almeida FC, Pontes G et al (2021) Evolution of the insect PPK gene family. Genome Biol Evol 13:evab185
Li SF, Zhang GJ, Yuan JH et al (2016) Repetitive sequences and epigenetic modification: inseparable partners in plant sex chromosome evolution. Planta 243:1083–1095
Mallet J, Joron M (1999) Evolution of diversity in warning color and mimicry. Annu Rev Ecol Syst 30:201–233. https://doi.org/10.1146/annurev.ecolsys.30.1.201
Nijhout HF (1991) The analysis of wing patterns. In: Nijhout HF (ed) The development and evolution of butterfly wing patterns. Smithsonian Institution Press, Washington/London, pp 51–84
Nishikawa H, Iga M, Yamaguchi J et al (2013) Molecular basis of wing coloration in a Batesian mimic butterfly, Papilio polytes. Sci Rep 3:3184
Nishikawa H, Iijima T, Kajitani R et al (2015) A genetic mechanism for female-limited Batesian mimicry in Papilio butterfly. Nat Genet 47:405–409. https://doi.org/10.1038/ng.3241
Ohsaki N (1995) Preferential predation of female butterflies and the evolution of Batesian mimicry. Nature 378:172–175. https://doi.org/10.1038/378373a0
Ohsaki N (2005) A common mechanism explaining the evolution of female-limited and both-sex Batesian mimicry in butterflies. J Anim Ecol 74:728–734
Palmer DH, Kronforst M (2020) A shared genetic basis of mimicry across swallowtail butterflies. Nat Commun 11:6. https://doi.org/10.1038/s41467-019-13859-y
Perez CAG, Adachi S, Nong QD et al (2021) Sense-overlapping lncRNA as a decoy of translational repressor protein for dimorphic gene expression. PLoS Genet 17:e1009683
Rembold H, Umebachi Y (1985) The structure of papiliochrome II, the yellow wing pigment of Papilionidae butterflies. In: Schlossberger HG et al (eds) Progress in tryptophan and serotonin research. Walter de Gruyter, Berlin, pp 743–746
Ruxton GD, Sherratt TN, Speed MP (2004) Avoiding attack: the evolutionary ecology of Crypsis, aposematism and mimicry. Oxford University Press
Su SK, Li CY, Lei PJ et al (2016) The EZH1–SUZ12 complex positively regulates NF-κB target genes through interaction with UXT. J Cell Sci 129:2343–2353
Terenius O, Papanicolaou A, Garbutt JS et al (2011) RNA interference in Lepidoptera: an overview and design implications. J Insect Physiol 57:231–245. https://doi.org/10.1016/j.jinsphys.2010.11.006
Timmermans MNTN, Thompson MJ, Collins S, Vogler AP (2017) Independent evolution of sexual dimorphism and female-limited mimicry in swallowtails. Mol Ecol 26:1273–1284. https://doi.org/10.1111/mec.14012
Uesugi K (2000) Avoiding attack in adult butterflies. In: Osaki N (ed) The natural history of butterflies. Hokkaido University Press, Sapporo, pp 106–123. (in Japanese)
Umebachi Y (1978) Red pigments in the wings of papilionid butterflies: extraction and purification. Sci Rep Kanazawa Univ 23:119–128
Umebachi Y (1985) Papiliochrome, a new pigment group of butterfly. Zool Sci 2:163–174
VanKuren NW, Doellman MM, Sheikh SI et al (2023) Acute and long-term consequences of co-opted doublesex on color pattern development. Mol Biol Evol 40:msad196. https://doi.org/10.1093/molbev/msad196
VanKuren NW, Sheikh SI, Fu CL et al (2025) Functional genetic elements of a butterfly mimicry supergene. Proc Natl Acad Sci USA 122:e2509864122. https://doi.org/10.1073/pnas.2509864122
Wallace AR (1869) The Malay archipelago. Macmillan, London

Westerman EL, Letchinger R, Tenger-Trolander A et al (2018) Does male preference maintain female-limited polymorphism in a Batesian mimic? Behav Process 150:47–58. https://doi.org/10.1016/j.beproc.2018.02.014

Westerman EL, Antonson N, Kreutzmann S et al (2019) Behaviour before beauty: signal weighting during mate selection in Papilio polytes. Ethology 125:565–574. https://doi.org/10.1111/eth.12884

Yamaguchi J, Mizoguchi T, Fujiwara H (2011) siRNAs induce efficient RNAi in Bombyx mori embryos. Sci Rep 6:e25469. https://doi.org/10.1371/journal.pone.0025469

Yoda S, Sakakura K, Kitamura T et al (2021) Genetic switch in UV response of mimicry-related yellow coloration in Papilio polytes. Sci Adv 7:eabd6475. https://doi.org/10.1126/sciadv.abd6475

Yoshioka S, Katoh M, Suzuki T et al (2025) Mimetic females do not bear reproductive costs: challenging sexual selection in female-limited mimicry. J Evol Biol 38:1487–1495. https://doi.org/10.1093/jeb/voaf098

Zelle KM, Lu B, Pyfrom SC, Ben-Shahar Y (2013) Genetic architecture of degenerin/epithelial sodium channels in Drosophila. G3 (Bethesda) 3:441–450

Zhou Y, Ge R, Wang R et al (2015) UXT potentiates angiogenesis by attenuating Notch signaling. Development 142:774–786

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 4
A Model for the Population Dynamics of the Mimetic Butterfly *Papilio polytes* in the Sakishima Islands, Japan: Mathematical Analysis and Computer Simulations

Toshio Sekimura

Abstract This chapter reviews a mathematical model for the population dynamics of the mimetic swallowtail butterfly *Papilio polytes* in the Sakishima Islands, Japan. *P. polytes* is widely distributed across India and Southeast Asia, including Southeast of China, the Philippines, Taiwan, and the Ryukyu Islands, to which the Sakishima Islands belong. The female of *P. polytes* in the Ryukyu Islands has two different forms, the mimetic form f. *polytes* and the nonmimetic form f. *cyrus* resembling the monomorphic male in appearance. The form f. *polytes* mimics the unpalatable butterfly *Pachiliopta aristolochiae* as a mimetic model. Mimicry in the female of *P. polytes* is known to be Batesian mimicry, which is one of the most interesting biological phenomena in nature. In the Sakishima Islands, fieldwork had been conducted over a long period, and the populations of butterflies had been recorded. A mathematical model describing that population dynamics is proposed. It includes four key variables: population densities of three kinds of butterflies and their predator. The model system consists of three ordinary differential equations (ODEs). The system is analyzed mathematically, and computer simulations confirm the analytical results, the field records on population dynamics of *P. polytes*, and the positive dependence of the relative abundance (RA) of the mimic on the advantage (AI) of the mimicry in the Sakishima Islands.

Keywords *Papilio polytes* · Wing color pattern · Female-limited Batesian mimicry · The Sakishima Islands · Population dynamics · Mathematical model

T. Sekimura (✉)
Chubu University, Kasugai, Aichi, Japan

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_4

4.1 Population Dynamics of the Mimetic Butterfly *Papilio polytes* in the Sakishima Islands, Japan

We consider a mathematical model for the population dynamics of the mimetic swallowtail butterfly *Papilio polytes* in the Sakishima Islands, Japan (Fig. 4.1).

The Sakishima Islands are located at the south-west end of Japan and are part of the Ryukyu Islands, which include both the Miyako Islands and the Yaeyama Islands. The Miyako Islands include the Miyako-jima Is., Tarama-jima Is., and so on, and the Yaeyama Islands include the Ishigaki-jima Is., Hateruma-Is., Iriomote-jima Is., Taketomi-jima Is., Kohama-jima Is., and so on (Sekimura et al. 2014).

4.1.1 Female Mimicry of P. polytes *Is Batesian Mimicry*

P. polytes is a mimetic swallowtail butterfly species widely distributed across India and Southeast Asia, including Southeast of China, the Philippines, Taiwan, and the Ryukyu Islands of Japan (Clarke and Sheppard 1972). *P. polytes* exhibits the female-limited polymorphism, i.e., the female is polymorphic, whereas the male is monomorphic and exhibits a white bar on the black hind wing. In the Ryukyu Islands located in the southwest of Japan, the female of *P. polytes* has two different forms, the mimetic form f. *polytes* and the nonmimetic form f. *cyrus*, resembling the monomorphic male in appearance. The form f. *polytes* mimics the unpalatable butterfly *Pachiliopta aristolochiae* as a mimetic model, which has a large white area in the

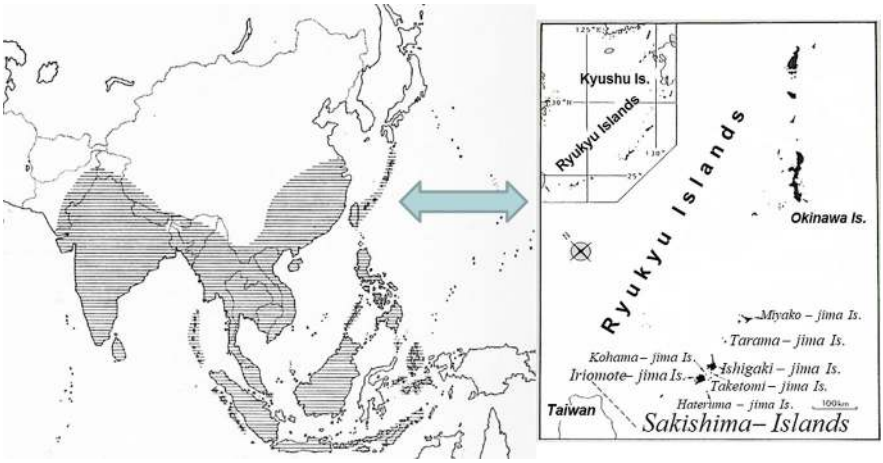


Fig. 4.1 Location of Sakishima Islands, Japan. The distribution area of the races of *P. polytes* is hatched (Clarke and Sheppard 1972)



Fig. 4.2 Female-limited polymorphism of the mimetic butterfly of *Papilio polytes* and the model butterfly *Pachliopta aristolochiae*. Top: two female forms of *P. polytes* f. *cyrus* (left, nonmimetic form) and *P. polytes* f. *polytes* (right, mimetic form). Bottom: *P. polytes* male (left) and the unpalatable model butterfly *Pachliopta aristolochiae* (right)

center, and a row of submarginal red spots on the black hind wing (Fig. 4.2). Mimicry in the female of *P. polytes* is known to be Batesian mimicry (Ohsaki 1995, Uesugi 1996).

4.1.2 General Discussions on Batesian Mimicry

Mimicry in the female of *P. polytes*, known to be Batesian mimicry, is one of the most interesting biological phenomena in nature. The literature has many general discussions about the benefits of the mimic to be protected from predation. For example, Holling (1965) found a relationship between the frequency of the mimic and the effectiveness of mimicry, and Emlen (1968) noted that the probability of a

mimic being attacked by a predator on an encounter decreases with increasing frequency of the model when models are rarer than mimics and increases conversely. Kumazawa et al. (2006) built the population dynamics equation of Batesian mimicry, in which two species of the palatable mimic and the unpalatable model interact under Lotka–Volterra-type interspecific competition. Assuming that both the benefit of mimicry for the mimic and the cost for the model are dependent on their frequency relatively to each other, they concluded that the benefit of the mimic and the cost of the model are key factors in determining the evolutionary process of Batesian mimicry. Yamauchi (1994) built a mathematical model for the population dynamics of Batesian mimicry, in which two populations of both model and mimic species were considered. The dynamic model has two components: growth at intrinsic growth rate and carrying capacity, and reduction by predation. The probability of a predator catching prey on an encounter was assumed to depend on the frequency of the mimic. He applied the dynamic model to field records of butterflies in the Ryukyu Islands, and his model had successfully explained some features, e.g., multiple dynamic equilibria between the model and the mimic in the field. However, his model did not account for realistic model–mimic systems such as polymorphism in *P. polytes* and intraspecific competition between the mimic f. *polytes* and the nonmimic f. *cyrus*.

4.1.3 Observation of Temporal Change in the Population of the Mimetic Female of P. polytes *in the Miyako-Jima Island*

4.1.3.1 Brief History of Field Observation and some Related Problems

In the Sakishima Islands (Fig. 4.1), substantial field data of *P. polytes* was recorded and accumulated since the early twentieth century. Regarding the mimetic female form *P. po*lytes f. *polytes*, field data showing the existence was recorded in June 1934 on the Hateruma-jima Island (Esaki 1934), in March 1964 in the Miyako-jima Island (Mizukoshi 1965), in March 1966 in the Iriomote-jima Island (Okada 1967), and in March 1967 in the Ishigaki-jima Island (Iwanaga and Takayama 1968). On the other hand, the model butterfly *Pachiliopta aristolochiae* was first recorded in 1968 in the Yaeyama Islands (i.e., the Iriomote-jima Is., Hateruma-jima Is., and Ishigaki-jima Is.) located in the south area of the Sakishima Islands after establishment of the mimetic form *P. polytes* f. *polytes* (Azuma 1969; Matsumoto 1969; Ohira and Terashita 1969). The above-mentioned record seems to indicate an interesting observation, i.e., two forms of the female of *P. polytes* f. *cyrus* and f. *polytes* were, to a greater or less extent, living in the Yaeyama Islands without the model *P. aristolochiae* before it was first recorded in 1968. Moreover, the mimetic form f. *polytes* has been collected every year and recorded without a break in the Yaeyama Islands since 1974, 6 years after the first record of establishment of the model *P. aristolochiae.* The continuous record of f. *polytes* suggests that the population

ratio of f. *polytes* in the islands may have remarkably increased since the establishment of the model *P. aristolochiae.* In the Miyako Islands located in the northeast about 120 km from the Yaeyama Islands, the model *P. aristolochiae* was reported to have been established in 1975, 7 years after the establishment in the Yaeyama Islands (Sunagawa et al. 1976). As mentioned above, in the Miyako-jima Island, two female forms f. *cyrus* and f. *polytes* were considered to have been coexisting before the establishment of the model *P. aristolochiae* (Mizukoshi 1965; Ohara 1972). Uesugi (1992) reported that the population ratio of the mimic f. *polytes* to the two forms (f. *cyrus* and f. *polytes*) was about 18% as of 1975 in the Miyako-jima Island. Since then, he continued to observe the ratio for the next 14 years, and the details are described in Sect. 4.1.3.2.

4.1.3.2 Observation of Temporal Change in the Population of the Mimetic Female of *P. polytes* in the Miyako-Jima Island

Uesugi (1992) observed the temporal change of the relative abundance of the mimic f. *polytes* in all the females of *P. polytes* for 14 years from 1975 to 1989 after establishment of the model *P. aristolochiae*. Here, the relative abundance (RA) denotes the population ratio of the mimic f. *polytes* to the two forms (f. *cyrus* and f. *polytes*), and RA is defined as follows: population of the mimic f. *polytes* by populations of the nonmimic f. *cyrus* and of the mimic f. *polytes* × 100(%). The observational result revealed that the RA increases with time like a sigmoidal curve, and in 1985 about 10 years after starting observation, the RA reached a saturated value (or the value of equilibrium) of about 50% (Fig. 4.3).

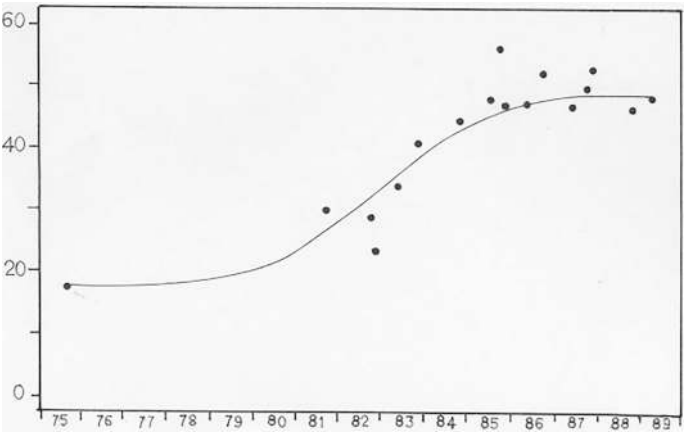


Fig. 4.3 Temporal change for 14 years of the relative abundance (RA) of the population of the mimetic female of *P. polytes* in the Miyako-jima Island (Uesugi 1992)

4.2 Fundamental Facts on the Mimicry of *P. polytes*

4.2.1 *Difference in Predation Risk between Two Forms f.* polytes *and f.* cyrus

The butterfly *P. polytes* has two female forms as described in Sect. 4.1: one is the nonmimic f. *cyrus* that resembles the monomorphic male with a white bar on the black hind wing, while the other is the mimic f. *polytes* that resembles the unpalatable butterfly species *P. aristolochiae* (Fig. 4.2). The form f. *polytes* is considered to have an advantage over f. *cyrus* with respect to protection from predation, when it lives in sympatry with the model *P. aristolochiae*. In reality, Uesugi (1996) examined the idea by learning experiments. Unexperienced birds, broweared bulbuls, *Hypsipetis amaurotis oryeri* as predators were first trained to take food from two feeders in a cavity, and then offered *P. aristolochiae* in one of the feeders. After experiencing an uncomfortable encounter with this butterfly, the birds reduced the frequency of taking regular food from the feeder where the butterfly had been placed.

On the other hand, Ohsaki (1995) paid attention to the rates of beak marks by predators on the wings of both palatable and unpalatable butterflies. By analyzing the number of beak marks on butterfly wings, he found that by comparing the species *P. polytes* and the model species *P. aristolochiae*, nonmimetic females were selectively attacked, while males, mimetic females, and models were less frequently attacked. Thus, the mimic f. *polytes* is considered to gain the benefit of reduced predation risk by living in sympatry with the model *P. aristolochiae*.

4.2.2 *Males Prefer the Nonmimic f.* cyrus *to the Mimic f.* polytes*?*

The mimetic female f. *polytes* might be less distinguishable than the nonmimetic female f. *cyrus* by the male, because the form f. *polytes* resembles a different species *P. aristolochiae*, while the form f. *cyrus* resembles the male of the same species. This means that f. *cyrus* could get more chances to mate with the male than f. *polytes*. In reality, Uesugi (1997) counted mating times of young butterflies of both f. *cyrus* and f. *polytes* just after emergence. The results showed that f. *polytes* had no approach from the male and no count, while f. *cyrus* could mate with the male. From the viewpoint of making offspring, the nonmimic f. *cyrus* has an advantage over the mimic f. *polytes*. Thus, viewing from the preference of the male to females in the species, mimicry is not always beneficial.

4.2.3 *Difference in Physiological Life Span and Egg Number between f.* cyrus *and f.* polytes

Ohsaki (2005) measured the physiological life spans of three types (the male, the nonmimic f. *cyrus*, and the mimic f. *polytes*), and found that the order of the life span length is as follows: f. *cyrus* > male > f. *polytes*. The problem is why the life span of the mimic f. *polytes* is the shortest. To explain the reason, Ohsaki proposed a hypothesis that the mimic f. *polytes* pays an additional cost for producing red colored pigments (carotenoid) by activating biochemical reaction networks in cells. The hypothesis coincides with the result that the bigger the red colored area on the wing that the butterfly has, the shorter the physiological life span that the butterfly has. Further, by using *P. polytes* caught in both the Okinawa Island and the Miyakojima Island, Kinjyou (2000) showed experimentally that the difference in physiological life span between f. *cyrus* and f. *polytes* is dependent on the feeding condition. For example, there is no difference in the life span when both f. *cyrus* and f. *polytes* are fed in a good condition, under which honey is given every day. On the other hand, when butterflies are fed in a bad condition, under which no honey or plain water is given, the life span of the nonmimic f. *cyrus* becomes significantly longer than that of the mimic f. *polytes.* This result suggests that in nature, when environmental conditions or food conditions (e.g., dry weather, and long rainy weather) become worse, the life span of f. *polytes* becomes shorter than that of f. *cyrus.* In any case, these results suggest that in natural environment, the physiological life span of f. *cyrus* is more or less longer than that of f. *polytes*. For the last few years, we measured the total number of laid-eggs of both f. *cyrus* and f. *polytes* in the laboratory, and found that in a regular feeding condition at room temperature, there was no significant difference between their laid-egg numbers (Sekimura, unpublished data). However, if we combine the result with the shortness of the physiological life span of the mimic f. *polytes*, the total laid-egg number during a life time is smaller than that of the nonmimic f. *cyrus*.

4.2.4 *The Genetic System of the Mimetic Butterfly* P. polytes

According to Clarke and Sheppard (1972), the genetic system of the mimetic butterfly species *P. polytes* is the female-limited inheritance, i.e., the male has only one phenotype while the female has two different phenotypes, i.e., the nonmimetic form f. *cyrus* and the mimetic form f. *polytes* (Fig. 4.2). It is known that the gene expressing mimicry is dominant. If we describe the allele related to the mimicry as H (dominant) and h (recessive), then the genotype of the nonmimetic form f. *cyrus* is homozygous recessive (hh), and the genotype of the mimetic form f. *polytes* is heterozygous (Hh) or homozygous dominant (HH). The genotype of the male could not be distinguished from its phenotype only. In the genetic system of *P. polytes*, the

male may act as the server or provider of all kinds of genotypes to females. In this case, it would not be useful to include the male population density directly as a variable in the current model system.

4.3 Mathematical Model and Simulation

The objective of this chapter is to present a mathematical model for the population dynamics of *P. polytes* observed in the Sakishima Islands (Fig. 4.1). Our mathematical model is constructed on the basis of field data in the islands and also experimental data in the laboratory, as well as further new facts on the mimicry of *P. polyets* (Ohsaki 2009). Three types of butterflies (i.e., two female forms of *P. polytes* and the model *P. aristolochiae*) and predator are involved in our model in terms of mimicry-related interactions.

4.3.1 Mathematical Model

The mathematical model system consists of three ordinary differential equations (ODEs) Eqs. (4.1), (4.2), and (4.3), in which the variables are the population densities of three butterflies: the unpalatable butterfly *P. aristolochiae* and two female forms of *P. polytes*. The male of *P. polytes* is not included directly in the system, but included indirectly through parameter values such as growth rates and intraspecific competition coefficients. We denote four population densities n_1, n_2, n_3, and p of the model *P. aristolochiae*, the mimic f. *polytes*, the nonmimic f. *cyrus*, and predator, respectively.

In the system Eqs. (4.1), (4.2) and (4.3), we evaluated the density effect or saturation effect by using carrying capacities K_1, K_2, K_3 for *P. aristolochiae*, f. *polytes*, and f. *cyrus*, respectively. We assume that K_2 and K_3 have the same value since both f. *polytes* and f. *cyrus* are females in the same species. With respect to growth rates of three butterflies r_1, r_2, r_3, we assume that according to discussions in Sect. 4.2, an inequality $r_2 < r_3$ should be satisfied.

In Eqs. (4.2) and (4.3), we introduced an intraspecific competition between f. *polytes* and f. *cyrus* through competition coefficients. The intraspecific competition effect includes competition for resources such as nectar and indirectly the male as noted in Sect. 4.3.2. Equations (4.1) and (4.2) represent the effect of mimicry, i.e., the negative density effect implying that the increase in the density of f. *polytes* n_2 causes the decrease in densities of both *P. aristolochiae* and f. *polytes*, n_1 and n_2. Furthermore, the parameters α, β represent the difference in the predation rate between *P. aristolochiae* and *P. polytes*, respectively, and it would be reasonable to assume the inequality $\alpha < \beta$, because *P. aristolochiae* is an unpalatable butterfly

species. In Sect. 4.2.1, we noted the result on beak marks by predators showing that nonmimetic females were selectively attacked, while males, mimetic females, and model butterflies were attacked less. This fact is evaluated mathematically in the second term of Eqs. (4.2) and (4.3) (Sekimura et al. 2014).

Taking the density-dependent effect through carrying capacity into account, we set up a model system consisting of three ordinary differential equations (ODEs), analyze it mathematically and provide computer simulations that confirm the analytical results.

The mathematical model for the population dynamics of *Papilio polytes* takes the form:

$$\frac{dn_1}{dt} = n_1 \left\{ r_1 \left(1 - \frac{n_1}{K_1} \right) - \alpha \left(\frac{n_2}{n_1 + n_2} \right) p \right\} \tag{4.1}$$

$$\frac{dn_2}{dt} = n_2 \left\{ r_2 \left(1 - \frac{n_2 + a_{23} n_3}{K_2} \right) - \beta \left(\frac{n_2}{n_1 + n_2} \right) p \right\} \tag{4.2}$$

$$\frac{dn_3}{dt} = n_3 \left\{ r_3 \left(1 - \frac{n_3 + a_{32} n_2}{K_2} \right) - \beta p \right\} \tag{4.3}$$

where a_{22} and a_{23} are competition coefficients. Eqs. (4.1), (4.2), and (4.3) include key factors to solve the problem: (1) self-density dependent growth, (2) mimetic effect, (3) interspecific competition, and (4) differences in growth rate and predation rate.

4.3.2 *Time Evolution of the Three-ODE System*

One of our objectives is to explain the results in Fig. 4.3 by mathematical analysis and computer simulations of the system equations.

Figure 4.4 shows the simulation results on temporal changes of population densities of the model (n_1), f. *polytes* (n_2), and f. *cyrus* (n_3) in Eqs. (4.1), (4.2), and (4.3). RA is the relative abundance of the mimic, i.e., the population ratio of f. *polytes* (n_2) to the two female forms of *P. polytes* (f. *cyrus* (n_3) and f. *polytes* (n_2)). AI is the advantage index of Batesian mimicry, which is defined to be the population ratio of the model (n_1) to all the related butterflies (the model and two female forms of *P. polytes*).

Numerical simulation results of temporal change in the RA explain the temporal change for 14 years in the population of the mimetic female of *P. polytes* in the Miyako-jima Island (Fig. 4.3, Uesugi 1992) (Fig. 4.5).

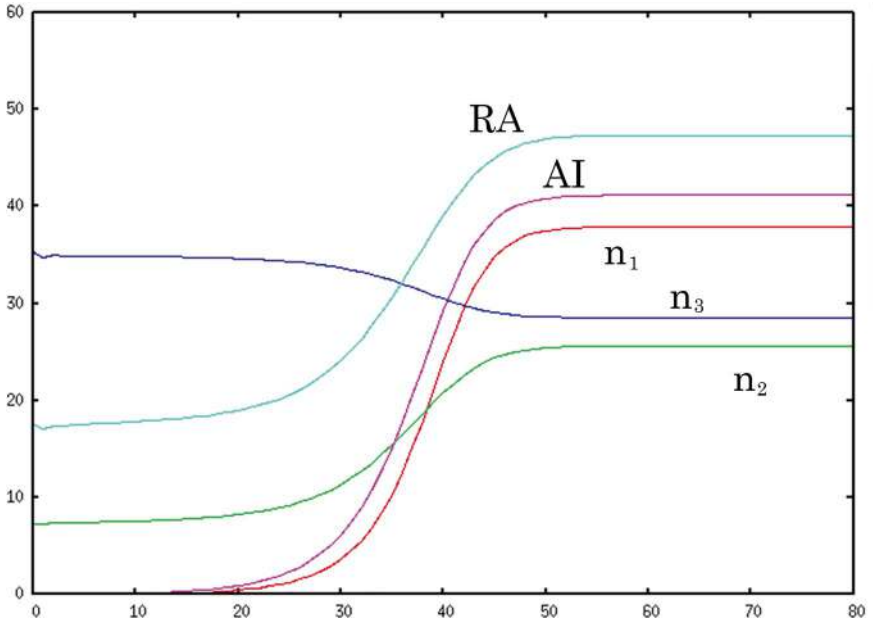


Fig. 4.4 Simulation results on temporal changes of all related quantities. The blue line denotes the population density n_3 of the nonmimic f. *cyrus*, the green line is the density n_2 of the mimic f. *polytes*, the light blue line is the relative abundance of the mimic (RA), the pink line is the advantage index (AI) of Batesian mimicry, and the orange line denotes the density n_1 of the model *P. aristolochiae*

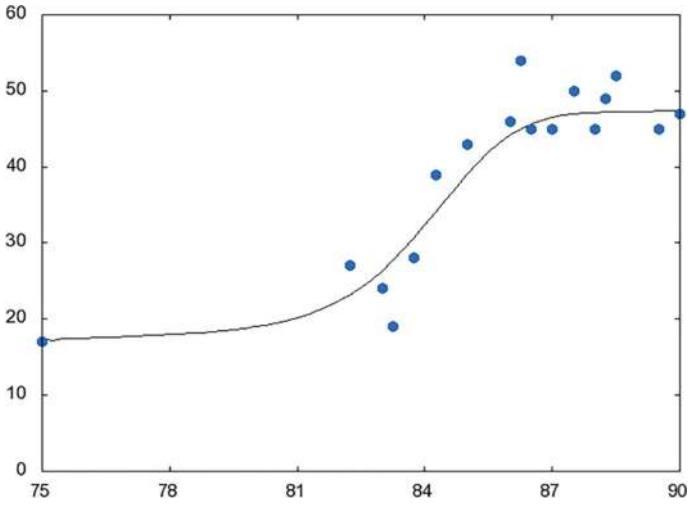


Fig. 4.5 Temporal change in the population of the mimetic female of *P. polytes* on the Miyakojima Island (blue dots, re-plotted from Fig. 4.3) and numerical simulation results of temporal change in the RA by the use of Eqs. (4.1), (4.2), and (4.3) (solid line)

4.3.3 *Simulations of Variation in the Relative Abundance RA in the Sakishima Islands*

Another objective of this chapter is to explain the variation in the relative abundance (RA) in the Sakishima Islands observed by Uesugi (1992) (plotted in Fig. 4.6).

We solve the system Eqs. (4.1), (4.2), and (4.3) numerically to simulate variation in the relative abundance RA in Sakishima Islands. Simulation results in Fig. 4.7 (solid line) show clearly the positive dependence of the mimic on the model in the Sakishima islands 14 years after establishment of the model in the Yaeyama Islands (Uesugi 1992), with which observational results of blue circles are put in the figure. The simulation results using the system Eqs. (4.1), (4.2), and (4.3) show the positive dependence of the RA on the AI. This is very important from the view point of comparison with observational results of variation in the RA in the Sakishima Islands.

4.4 Summary and Discussions

In this chapter, we presented a mathematical model for the population dynamics of the mimetic butterfly *P. polytes* in the Sakishima Islands, Japan, which is based on the best possible data currently available to us (in Sects. 4.1 and 4.2). Our mathematical model includes three major effects and two experimental results on differences in growth rate and predation rate, namely, (1) self-density effect or density saturation through carrying capacity, (2) mimetic effect, i.e., the probability that a predator attacks prey on an encounter is proportional to the relative frequency of the

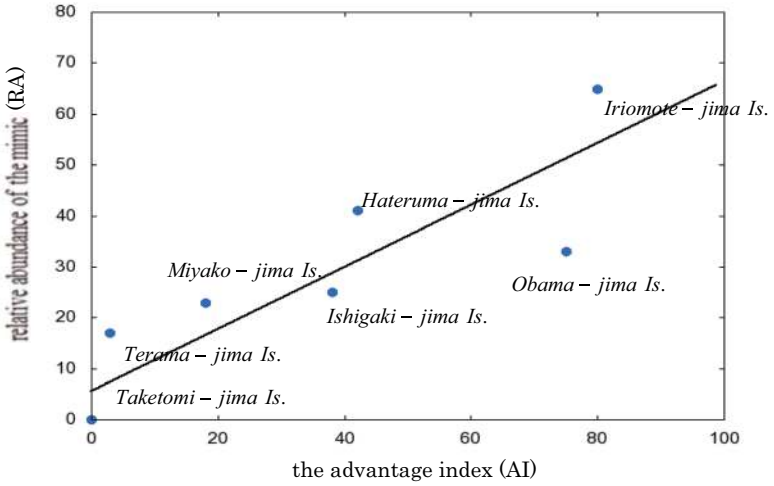


Fig. 4.6 Positive dependence of the mimic on the model in the Sakishima islands 14 years after establishment of the model in the Yaeyama Islands (Uesugi 1992)

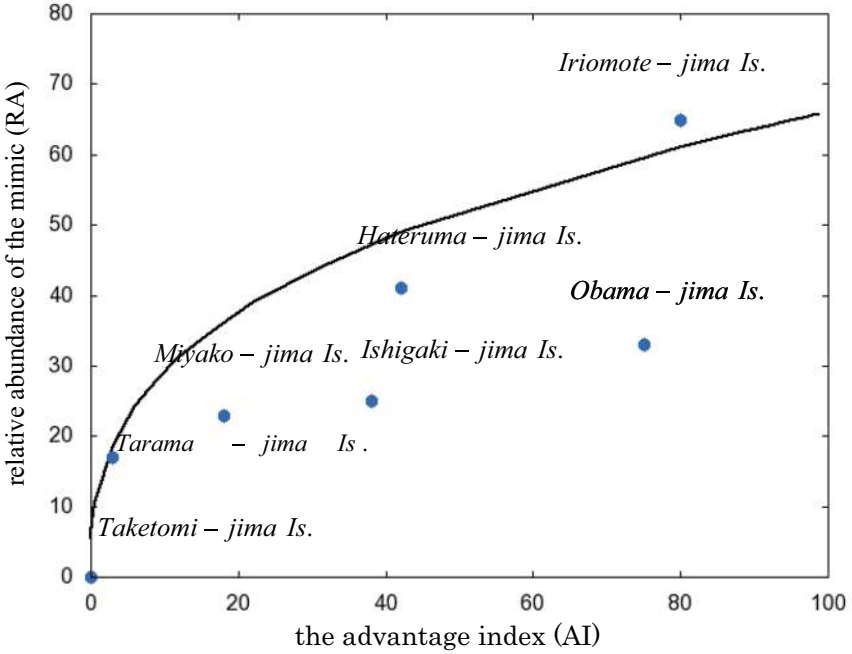


Fig. 4.7 Variation in the relative abundance RA in the Sakishima Islands and its numerical simulation results. The horizontal axis shows the advantage index AI, and the vertical axis is the relative abundance RA. Blue circles represent the RA corresponding to the AI on each island of seven islands in the Sakishima Islands (Uesugi 1992)

mimic among the model and the mimic, and (3) intraspecific competition between two female forms f. *polytes* and f. *cyrus* for resources such as nectar and, indirectly, the male, and difference in growth rate between two females f. *polytes* and f. *cyrus*.

Using mathematical analysis and computer simulations of the system equations, we clarified the logical relationship hidden behind field data on the population dynamics of the mimetic butterfly *P. polytes*. This provides the theoretical basis for the positive dependence of the RA on the carrying capacity K_1 of the model butterfly (Sekimura et al. 2014).

The horizontal axis in Figs. 4.6 and 4.7 shows the advantage index (AI) of Batesian mimicry, which is defined to be the population ratio of the model to all the related butterflies (the model and two female forms of *P. polytes*). The vertical axis is the relative abundance RA defined as the ratio of population rate of f. *polytes* to the two forms (f. *cyrus* and f. *polytes*). Blue solid circle represents the RA corresponding to the AI on each island of seven islands (i.e., the Taketomi-jima Is., Tarama-jima Is., Miyako-jima Is., Ishigaki-jima Is., Hateruma-jima Is., Obama-jima Is., and Iriomote-jima Is.) in the Sakishima Islands (from Uesugi 1992). The solid line in Fig. 4.7 represents numerical simulation results of Eqs. (4.1), (4.2), and (4.3), which clearly show the positive dependence of the RA on the AI with a

convexity. An analytical result of the system of equations provides the theoretical basis for the positive dependence of butterfly *P. polytes* in the Sakishima Islands (Sekimura et al. 2014). In particular, we discussed both temporal change in the relative abundance RA since 1975 in the Miyako-jima Island and variation in the RA in the Sakishima Islands. Our results show that one of the key factors to understand field data is the carrying capacity K_1 of the model in each island. The positive dependence of the RA on the AI originates from the result that changes of both the relative abundance RA and the advantage index AI with respect to the carrying capacity are positive. The female-limited polymorphism of the mimetic butterfly *P. polytes* in the Sakishima Islands is considered to be maintained by the negative frequency-dependent selection (Ohsaki 2009), which is another expression of the above-mentioned positive dependence of the RA on the AI. We also estimated conditions for existence of equilibrium solutions of butterfly populations by making a comparison between experimental results and mathematical analyses of our model equations. Therefore, we propose that the mathematical analysis and computer simulations in this chapter using a mathematical model provide the theoretical basis on the female-limited polymorphism of *P. polytes* in the Sakishima Islands. We have not discussed directly the origin of Batesian mimicry in this chapter, which is another interesting and unsolved problem. As an example of other types of approach to the problem, Vane-Wright et al. (1999) determined sister group relationships of the mimetic butterfly *Papilio dardanus* through the molecular phylogenetic method by using four different DNA regions (i.e., mitochondrial 16S rRNA and cytochrome *b* regions, and nuclear Elongation Factor EF1α and the ITS-1 region of the ribosomal RNA locus). We hope that our discussions in this chapter can contribute to the analysis and solution of the origin of Batesian mimicry.

References

Azuma S (1969) Yaeyama de benimon-ageha saishu sareru (The new records of Pachiliopta aristolochiae from the Yaeyama Islands). Nat Insects 4:9–10

Clarke CA, Sheppard PM (1972) The genetic of the mimetic butterfly Papilio polytes. Philos Trans R Soc Lond Ser B Biol Sci 263:431–458

Emlen JM (1968) Batesian mimicry: a preliminary theoretical investigation of quantitative aspects. Am Nat 102:235–241

Esaki T (1934) Yaeyama-shoto konchu saishuki (An account of a trip of collecting butterflies in the Yaeyama Islands). Ent World 2:537–564

Holling CS (1965) The functional response of predators to prey density and its role in mimicry and population regulation. Mem Entomol Soc Can 97:5–60

Iwanaga S, Takayama M (1968) The second female type so-called pammon-type of Papilio polytes polytes Linnaeus collected in Ishigaki and Iriomote Islands. TYO TO GA (Trans Lep Soc Jap) 19:54–55. (In Japanese)

Kinjyou A (2000) Mimetic relationship in swallowtail butterflies in the Ryukyus. Nat Insects 36(12):24–27. (In Japanese)

Kumazawa F, Asami T, Hayashi T, Yoshimura J (2006) Population dynamics of Batesian mimicry under interspecific competition. Evol Ecol Res 8:591–604

Matsumoto K (1969) Yaeyama-gunnto ni okeru benimon-ageha Menelaides aristolochiae ni tsuite (On the Menelaides(=Pachiliopta) aristolochiae in the Yaeyama Islands). Crude 4:3–6. (In Japanese)
Mizukoshi Y (1965) Dai 2 ji Okinawa chosa hokoku [IV] chorui (Reports on 2nd Okinawa research [IV] butterflies). Waseda Seibutsu 13:11–19. (In Japanese)
Ohara K (1972) Yaeyama-gunto haru no chorui (On the butterflies in the Yaenyama Islands in spring). Satsuma 21:24–31. (In Japanese)
Ohira H, Terashita M (1969) Hateruma-jima no konchu (On the insects in the Hateruma-Jima Islands). Haterumamura. Aichi Educ Col:33–35. (In Japanese)
Ohsaki N (1995) Preferential predation of female butterflies and the evolution of Batesian mimicry. Nature 378:173–175
Ohsaki N (2005) A common mechanism explaining the evolution of female-limited and both-sex Batesian mimicry in butterflies. J Anim Ecol 74:728–734
Ohsaki N (2009) Evolution of mimicry. Kaiyu-sha. (In Japanese)
Okada T (1967) Iriomote-jima no chorui (On the butterflies of Iriomote-Jima Island). Waseda Seibutsu 16:81–89. (In Japanese)
Sekimura T, Zhu M, Cook J, Maini PK, Murray JD (1999) Pattern formation of scale cells in lepidoptera by differential origin-dependent cell adhesion. Bull Math Biol 61:807–827
Sekimura T, Fujihashi Y, Takeuchi Y (2014) A model for population dynamics of the mimetic butterfly Papilio polytes in the Sakishima Islands, Japan. J Theor Biol 361:133–140
Sunagawa H, Hamakawa M, Omoto K (1976) Miyako-gunto no chorui (On the butterflies in the Miyako Islands). Larva 10:1–8. (In Japanese)
Uesugi K (1992) Polymorphism of the mimetic butterfly Papilio Polytes, L. Insectalium 22:4–10. (In Japanese)
Uesugi K (1996) The adaptive significance of Batesian mimicry in the butterfly, *Papilio Polytes* (Insecta, Papilionidae): associative learning in a predator. Ethology 102:762–775
Uesugi K (1997) Iden 51(2):68–71. (In Japanese)
Vane-Wright RI, Raheem DC, Cieslak A, Vogler AP (1999) Evolution of the mimetic African swallowtail butterfly *Papilio dardanus*: molecular data confirm relationships with *P. Phorcas* and *P. Constantinus*. Biol J Linn Soc 66:215–229
Yamauchi A (1994) A population dynamic model of the Batesian mimicry. Res Popul Ecol 53:295–315

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 5
The Development and Evolution of Butterfly Wing Patterns: Finding the Primitive Pattern



H. Frederik Nijhout and Richard Gawne

Abstract The color patterns of butterflies and moths constitute a highly evolved system of visual communication used in sexual signaling, camouflage, warning, territoriality, and mimicry. These patterns are based on extensive modifications of a ground plan of three pairs of bands. The primitive color pattern from which this patterning system evolved arose long before the evolution of the Lepidoptera. Wing patterns of the caddis flies (Trichoptera) and scorpion flies (Mecoptera) resemble those found in basal moths. We will show that these patterns are all based on variations on a set of six parallel bands, and that this basic patterning system can be found in the earliest winged insects, represented today by dragonflies (Odonata). We show by mathematical modeling that a simple reaction–diffusion system can readily generate three symmetry systems, which can be visualized as either three pairs of bands or six identical bands. We show that from this banding system, a lateral inhibition mechanism can, in turn, generate the basic spotting and banding patterns that we see in the Odonata, Mecoptera, Trichoptera, and Lepidoptera.

Keywords Color pattern · Wing pattern · Primitive pattern · Pattern evolution · Symmetry system · Lateral inhibition · Substrate depletion · Mathematical model

5.1 Introduction

We have been exploring the problem of what is the primitive color pattern of butterflies and moths—the color pattern from which all modern patterns of moths and butterflies are derived. It turns out that the answer to that question is somewhat surprising, it was to us at least, and it also has some interesting challenges for the

H. F. Nijhout (✉)
Department of Biology, Duke University, Durham, NC, USA
e-mail: hfn@duke.edu

R. Gawne
Nevada State Museum, Las Vegas, NV, USA

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_5

mathematical biology of pattern formation. In order to understand this, we need to start by briefly outlining how butterfly color patterns are organized.

The color patterns of butterflies are immensely diverse. No two species look quite alike and species can almost always be identified based on their color pattern alone. Yet, in spite of their individuality, color patterns are seldom used in the taxonomic identification of butterflies, nor are they used in phylogenetic analyses of the evolution of butterflies.

Wing patterns are adaptations for communication. They are used for sexual signaling, territorial display, aposematism, mimicry, and camouflage. Each of these is an adaptation to different habitats and ecological conditions. Butterfly color patterns can evolve rapidly under selection and appear to diversify more rapidly than other morphological characteristics (Oliver et al. 2009; Beldade et al. 2002; Wallbank et al. 2016; Mitsuho et al. 2017; Basu et al. 2023). As a result, color patterns evolve separately from other body parts and rapidly adapt to various ecological conditions. Accordingly, it is difficult to derive a phylogeny of butterflies based on color patterns alone.

There are, nevertheless, profound commonalities among butterfly wing patterns because this entire diversity of patterns is organized by a common system of homologies called the nymphalid ground plan (Schwanwitsch 1924, 1929; Süffert 1927; Nijhout 1991, 1994). The nymphalid ground plan is composed of three pairs of color bands. Each pair is referred to as a symmetry system: the basal symmetry system, the central symmetry system, and the border symmetry system (Fig. 5.1). Ideally, these bands run uninterrupted from the anterior to the posterior margin of the wing but more often than not these bands are disrupted wherever they cross a wing vein, so that the pattern becomes compartmentalized in these different wing regions, and it is thought that this compartmentalization has been responsible for enabling the tremendous evolutionary diversity of color patterns.

The nymphalid ground plan is not intended to represent the primitive pattern, it is purely a diagram of homologies that show how elements of the color pattern are arranged on the wing. Some aspects of the ground plan could be primitive, but we do not know that. That is something that is to be determined.

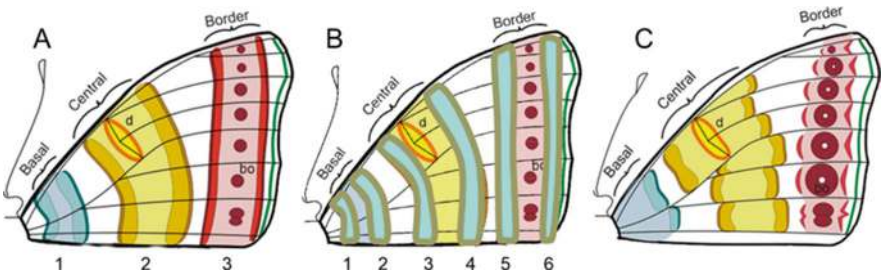


Fig. 5.1 The nymphalid ground plan is a diagram of how pattern elements are arranged on the wing that illustrates homologies among the elements of the color pattern. It is made up of three symmetry systems (**A**), each consisting of a pair of bands (**B**), labeled 1-6. In some species the wing veins compartmentalize and disrupt the bands of the ground plan (**C**) so the bands look dislocated and unconnected across the wing

5.2 The Arctiidae

The standard way of identifying the primitive or ancestral form of a trait is by outgroup comparison. If the trait is found in a closely related taxon, the outgroup, it is considered to be primitive or ancestral for the taxon of interest. So, to identify what characteristics of the color pattern are primitive, one needs to know their condition in the sister group of butterflies. Butterflies are a monophyletic clade within the moths, and some years ago we began this study by analyzing color patterns in moths of the Family Arctiidae, the so-called tiger moths (Gawne and Nijhout 2019, 2020). This is a very diverse group of day-flying moths with brilliant color patterns used primarily for aposematic signaling. We found that there is an enormous amount of individual variation in the color patterns, which turned out to be a very useful feature for the analysis of homologies in the pattern. Figure 5.2 shows examples of individual variation in three species.

In these species the pattern is composed of six bands running between the anterior and posterior margins of the wing. In some cases, these bands are present as rows of spots whose outer rings can merge. In *Utetheisa*, the black spots are surrounded by white outer rings that merge into bands. Rows of spots and their surrounding white bands can merge with adjacent rows in various ways, as shown in Fig. 5.2. Merging of the outer bands indicates that they are produced by the same developmental process. The simplest explanation is that spots and bands are contours of a morphogenetic landscape. Dark spots indicate higher values and white outer bands lower values of the morphogen. The overall pattern being due to a landscape of hills and valleys of morphogen concentration.

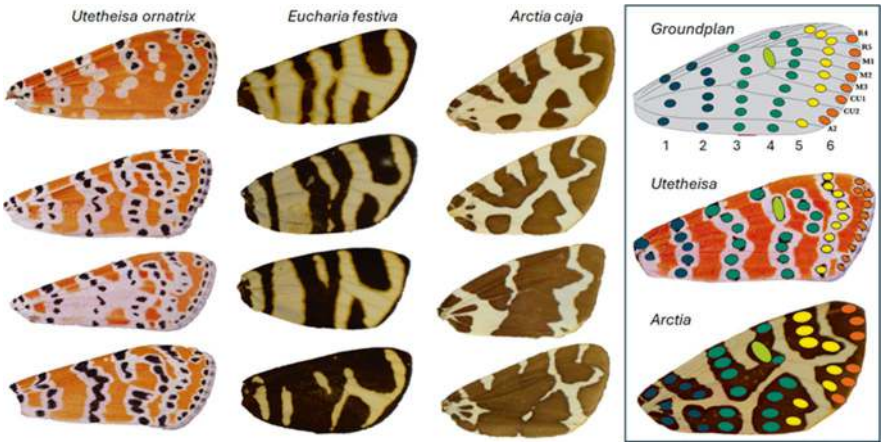


Fig. 5.2 Pattern variation and ground plan of the Arctiidae. Individual variation of the color pattern in three species shows that bands can merge in various ways to make up complex-looking patterns. In *Utetheisa* the bands are white, with black centers, on an orange background. In *Eucharia* and *Arctia* the bands are brown on a white background. The box on the right shows the arctiid ground plan consisting of six bands of spots, and how these merge in different ways in *Utetheisa* and *Arctia*

Spotted patterns like those of *Utetheisa* are uncommon in the Arctiidae. More often, their color patterns are made up of dark bands against a lighter background, which are shown in the cases of *Arctia caja* and *Eucharia festiva* (Fig. 5.2). In these species there is individual variation in the width, length, and degree of fusion of the dark-pigmented bands. A summary of the arctiid color pattern ground plan and how the patterns of *Utetheisa*, *Arctia*, and *Eucharia* are derived from it is shown in the box in Fig. 5.2.

This ground plan raises a question that has actually existed in the field for a long time: Do bands come about by fusion of spots or were spots originally bands that were broken up by the wing veins during evolution? In other words, are bands primitive or are rows of spots primitive? Fusion of spots into bands or larger areas of solid color, as shown in Fig. 5.3, is actually quite common in butterflies too. The large black areas that make up the bold dorsal color patterns of many butterflies are actually composites of fused smaller spots (e.g., see Nijhout and Wray 1986, 1988; Nijhout et al. 1990).

The tiger moths suggest that fusion of spots to form larger composite patterns is shared with butterflies and may therefore be a primitive or ancestral property for the clade to which they both belong. But there are actually two possibilities. Bands and large regions of black formed by fusion of developmentally independent spot-like pattern elements may be a derived condition. Alternatively, bands could be an ancestral primitive condition, and cases where bands are broken up by wing veins

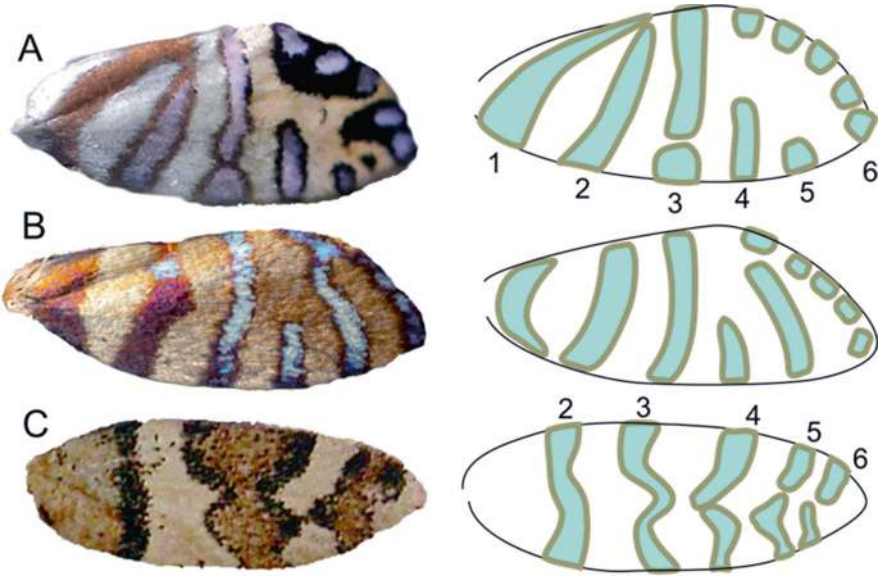


Fig. 5.3 Color patterns of the Micropterigidae. Three species of *Sabatinca* illustrate patterns made up of five or six bands. Next to each wing is a diagrammatic summary of the pattern. From top to bottom: (**A**) *S. aurella* (adapted from iNaturalist Saryu Mae (CC BY)), (**B**) *S. doroxena*, and (**C**) *S. lucilla*

that somehow impose secondary developmental barriers or compartments across the wing, resulting in discrete spots, may be the derived condition.

The Arctiidae thus share with butterflies a pattern consisting of six bands running anterior posterior on the wing, which can be broken up into spots and fuse laterally into larger regions of homogeneous color, usually black. The six bands in the Arctiidae do not appear to pair up as three symmetry systems as they do in butterflies, although they do so in other moths (e.g., the Saturniidae: Schwanwitsch 1924; Süffert 1927; Nijhout 1991).

5.3 The Micropterigidae

The outgroup to all butterflies and moths is an interesting group called the Micropterigidae, the mandibled moths. They are the first offshoot or branch in the phylogeny of the Lepidoptera, so the color pattern features they share with the other Lepidoptera should be primitive. Like the butterflies, the wing patterns of the Micropterigidae are extremely diverse, and species can be recognized based on their color pattern alone (Gibbs 2014; Gibbs and Lees 2014).

Schachat and Brown (2015, 2016) did an exceptionally thorough analysis of the color patterns of the Micropterigidae. They deduced that the wing patterns of the Micropterigidae are based on a system of either five or six bands that run across the wing. We illustrate three examples from the genus *Sabatinca* in Fig. 5.3. On the left side are the wings of three species, and on the right side are their patterns diagrammatically outlined in light blue and brown.

In two of the species illustrated, we see variations on the same pattern made up of six bands that run across the wing anterior to posterior, or only partially across the wing, with a given band sometimes reduced to a short bar, a spot, or a row of spots. The bands typically have margins of a contrasting color. In *Sabatinca lucilla*, the bands are paired, with a contrasting color between them, and appear to be three symmetry systems, much like those we find in butterflies.

Thus, the color patterns of the Micropterigidae are quite sophisticated and elaborate, with features we find throughout moths and butterflies. This implies that this patterning system must have existed sometime before the Micropterigidae arose, and it seems unlikely that these patterns represent a truly primitive system that originated in this group.

5.4 The Trichoptera and Mecoptera

Thus, since the primitive traits are assumed to be those that are shared with the phylogenetic sister group, we need to look at the phylogenetic sister group of the Lepidoptera as a whole, which are the Trichoptera (caddisflies). Trichoptera can be quite colorful animals. They do not have scales on their wings but hairs, and the

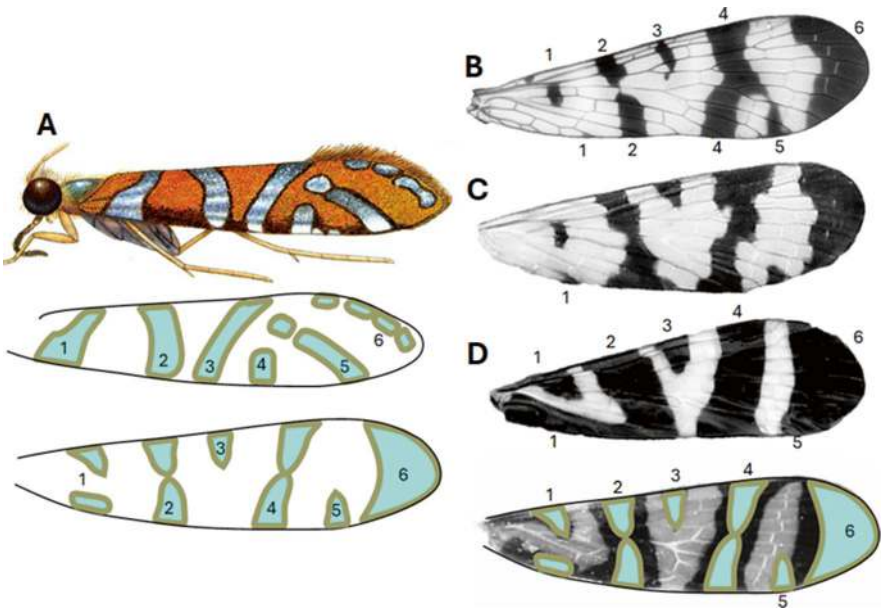


Fig. 5.4 Wing patterns of Trichoptera and Mecoptera. Left column, Trichoptera: (**A**) an undescribed *Nectopsyche* sp. (courtesy Ralph Holzenthal, University of Minnesota (CC BY-NC-ND)), with diagrammatic summary of its pattern. Right column, Mecoptera: (**B**) *Panorpa confusa*, (**C**) *Panorpa helena*, and (**D**) *Panorpa nuptialis* (with diagrammatic summary of its pattern)

pigments that make up the wing color patterns are located in the membranes of the wings, not the scales, as in Lepidoptera. Caddisflies can have complex and diverse color patterns, some of which resemble those of moths. As in the Arctiidae, there is a lot of individual variability in the patterns of different species (Fig. 5.4). Figure 5.4A shows the pattern of a still undescribed species of *Nectopsyche*. Its color pattern looks remarkably like that of some Micropterigidae. It is a pattern of six stripes, diagrammatically illustrated in Fig. 5.3 to enable comparison to the patterns of moths. The Trichoptera illustrated here also share with the Mircropterygidae the tendency of the band along the wing margin to break up into spots.

Thus, the Trichoptera already have all the pattern elements found in the Micropterigidae. There are six bands, but these are not obviously arranged in pairs that can be recognized as symmetry systems. Some bands can be incomplete and appear to emerge from the wing margin (Fig. 5.4B), as in *Sabatinca* (Fig. 5.3A, B). Likewise, the color bands of *Nectopsyche* are outlined in a contrasting color, much as we see in *Sabatinca*.

In order to understand where the color patterns that are shared by the Lepidoptera and Trichoptera arose, we will have to look at their joint sister group. A sister group to the Lepidoptera/Trichoptera clade, which also has wing color patterns, is the Mecoptera, the scorpion flies. The color patterns of selected Mecoptera are illustrated in Fig. 5.4B–D. The Mecoptera also have a system of bands that can be continuous across the wing or can appear to emerge from the anterior and posterior

wing margins. There is a lot of individual variability in the color patterns, for example bands arising from a wing margin can merge with bands originating from the opposite margin, as well as with neighboring bands on the same margin, much as we see in some Arctiidae.

The patterns of the Mecoptera look, in principle, very much like those of Trichoptera and Micropterigidae: there are six rows of bands, sometimes incomplete, sometimes partially merged, but there are nevertheless six pattern elements that run between the anterior and posterior margins of the wing.

5.5 The Odonata

Thus, the Mecoptera appear to share fundamental features of color pattern structure with the other orders we have examined. All have patterns that are fundamentally composed of six bands. But is this likely to be primitive? To address this point, we examined the dragonflies (Odonata), which are the most basal existing Order in the insect phylogeny whose members have color patterns on their wings (Grimaldi and Engel 2005). Thus, they are the ultimate outgroup in terms of wing color pattern formation.

Dragonflies can be quite colorful insects and many have very nice wing patterns made up of spots and bands. The color patterns of dragonflies are often sexually dimorphic, and within a species there is a considerable amount of individual variation in the color pattern. Some species have only a large spot or band at the tip of the wing, in others there is a band in the middle of the wing, and others have a large spot or band at the base of the wing (Fig. 5.5). Some species have all three bands, and in some species some of the bands are incomplete and appear to emerge from the anterior and posterior margins of the wing.

The forewings and the hindwings tend to have similar patterns, but because the hindwings are broader their patterns tend to be more broken and reveal the composite nature of what sometimes appear to be continuous bands on the forewing (e.g., Fig. 5.5F). In many cases, broad bands have dark margins (e.g., Fig. 5.5B–E) that can be quite broad (e.g., Fig. 5.5E), and sometimes, as in *Plathemis subornata*, the band in the middle of the wing can be broad and homogeneous, or have a variably pale core area, which can be completely colorless, so the pattern becomes two parallel bands. The broad apical and basal bands likewise can have a darker margin and a pale region that can become colorless, thus leaving a narrower band. So the Odonata have a system of three bands, each of which can be broken into two, so we have, again, a system of six bands that we see in the other Orders. In the Odonata, each band is effectively a symmetry system.

As is the case for most other color pattern elements in other Orders, it is often not possible to determine the identity of a band in a single individual. Whether two bands are band 2 and 3, or 3 and 4, or 2 and 4, for instance, can usually only be ascertained by comparative studies: by examining individual variation of the pattern and the patterns in closely related species.

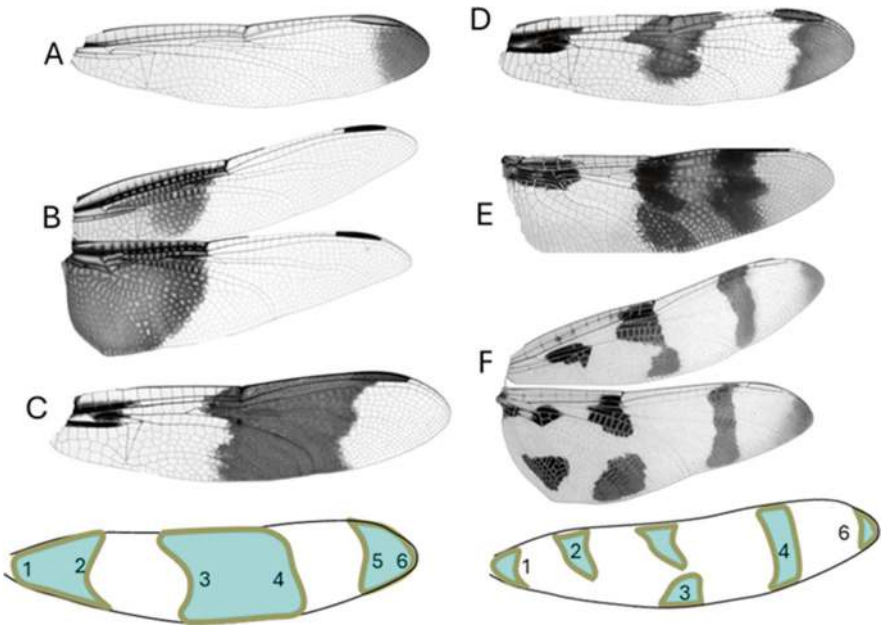


Fig. 5.5 Wing patterns of Odonata, illustrating diversity of banding patterns. (**A**) *Libellula flavida*, (B) *Libellula luctuosa*, (**C**) *Plathemis lydia* (male), (**D**) *Plathemis lydia* (female), (**E**) *Plathemis subornata* (adapted from iNaturalist, Laura Gaudette (CC BY)), and (**F**) *Celithemis eponina*. *Below* are two diagrammatic summaries: left, broad regions, each of which is assumed to be a symmetry system made up of two bands (as in E); right, pattern diagram of forewing of (**F**)

The point, however, is that in dragonflies, we have a system of six bands that we find across pretty much all the insects. In each insect lineage there has been great diversification of pattern, but what they all appear to share, in at least some species, are simple patterns made up of bands, a system of six bands, grouped in three pairs. Insofar as this pattern system can be found at the base of the insect phylogeny, the dragonflies, it is likely to constitute the primitive pattern from which all others are evolved.

5.6 Modeling Bands and Symmetry Systems

We next examine, using a mathematical model, how a system of six bands or three symmetry systems can arise. We start by using one of the mechanisms that Hans Meinhardt explored, the substrate depletion mechanism (Meinhardt 1982). It is a reaction–diffusion mechanism with two players, an activator and a substrate. The activator is synthesized at an autocatalytic rate but requires the substrate for its synthesis. The substrate is synthesized autonomously and is used up in the process that makes the activator. Both activator and substrate break down at an autonomous rate

and diffuse, with the substrate diffusing more rapidly than the activator. This system is described by the following equations:

$$\frac{\partial a}{\partial t} = ca^2 s - \rho a + D_a\left(\frac{\partial^2 a}{\partial x^2} + \frac{\partial^2 a}{\partial y^2}\right)$$

$$\frac{\partial s}{\partial t} = c_0 - ca^2 s - \nu s + D_s\left(\frac{\partial^2 s}{\partial x^2} + \frac{\partial^2 s}{\partial y^2}\right)$$

We solve these equations in a rectangular array, representing a wing, with initial concentrations everywhere set at the steady-state values of activator and substrate, except at one end of the array, the base of the wing, where the value of the activator is set at 2× steady-state. We impose zero flux boundary conditions. The progress of pattern development is shown in Fig. 5.6A and leads to a stable steady-state with wave-like patterns of activator and substrate consisting of three broad bands or ridges. If we assume that different pigments can be synthesized at particular levels of the activator, then each of the three bands forms a symmetry system.

So we can think of this as being either a set of three bands if pigments are made only above a threshold (red, for instance, in Fig. 5.6A), or we can think of them as three pairs of bands if pigments are synthesized only at a specific value (yellow, for instance, in Fig. 5.6). The latter could represent the foundation of the six bands that we see in the Lepidoptera and all the other insects.

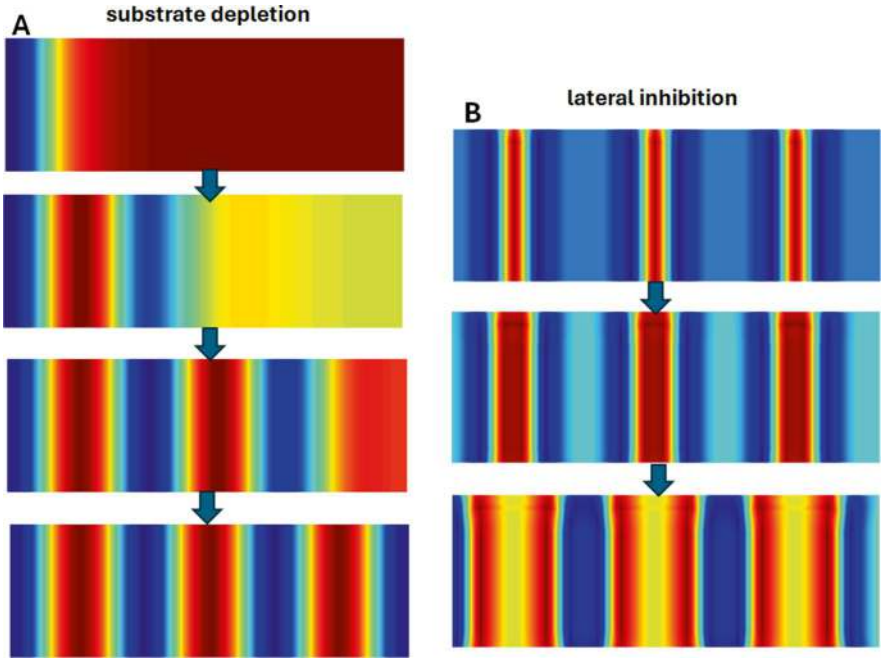


Fig. 5.6 Model results: (**A**) three symmetry systems can arise from a substrate depletion model; (**B**) a lateral inhibition process can split a band into two symmetry systems

5.7 Bands and Secondary Symmetry Systems

We noted above that the bands in the color patterns of Micropterigidae and Trichoptera often have boundaries of a contrasting color. Thus, we can think of each such band as a symmetry system. In some moths too, Süffert found that a band of a symmetry system can be subdivided into a symmetrical pair of bands (Süffert 1929). In some butterflies, too, the bands of a symmetry system can become self-symmetrical (Nijhout 1991). These features are called secondary symmetry systems.

We used a simple lateral inhibition mechanism based on the work of Swindale and Young (Swindale and Barlow 1980; Young 1984) to show how a solid band can be split into a secondary symmetry system. Figure 5.6B shows the result of such a simulation, which shows how each of three narrow bands can split into a symmetry system, yielding six bands.

5.8 Partial Bands and Fusion of Bands

Many banding patterns do not look like symmetry systems, nor do the bands always come in obvious pairs, nor do they run from the anterior to posterior margin of the wing. More often than not, the bands are incomplete and appear to emerge from the anterior or posterior margins of the wing (e.g., Fig. 5.7 shows a diagrammatic summary of the patterns discussed above).

In order to approach this, we did the following. We started with the results of the substrate depletion system that produces the set of three bands, established a threshold on those bands, and added some noise to produce a set of three noisy bands shown in Fig. 5.8. We then used these as the initial conditions for the Swindale/Young lateral inhibition program, using elliptical activator and inhibitor domains with the long axis of the ellipse parallel to the bands. This mechanism produces a pattern that develops over time and stabilizes into a system of spots or bands, determined by the relative strengths of activator and inhibitor, the relative sizes of the activator and inhibitor domains, and the aspect ratios of the activator and inhibitor ellipses.

Figure 5.9A illustrates some of the patterns that are produced by this simple lateral inhibition system. These compare well to the idealized target patterns of the Micropterigidae, Trichoptera, Mecoptera, and Odonata shown in Fig. 5.7B. Figure 5.9B illustrates a typical model landscape from which the patterns in Fig. 5.9A are derived. Selection of different threshold levels in Fig. 5.9B can broaden the patterns and lead to fusion of bands.

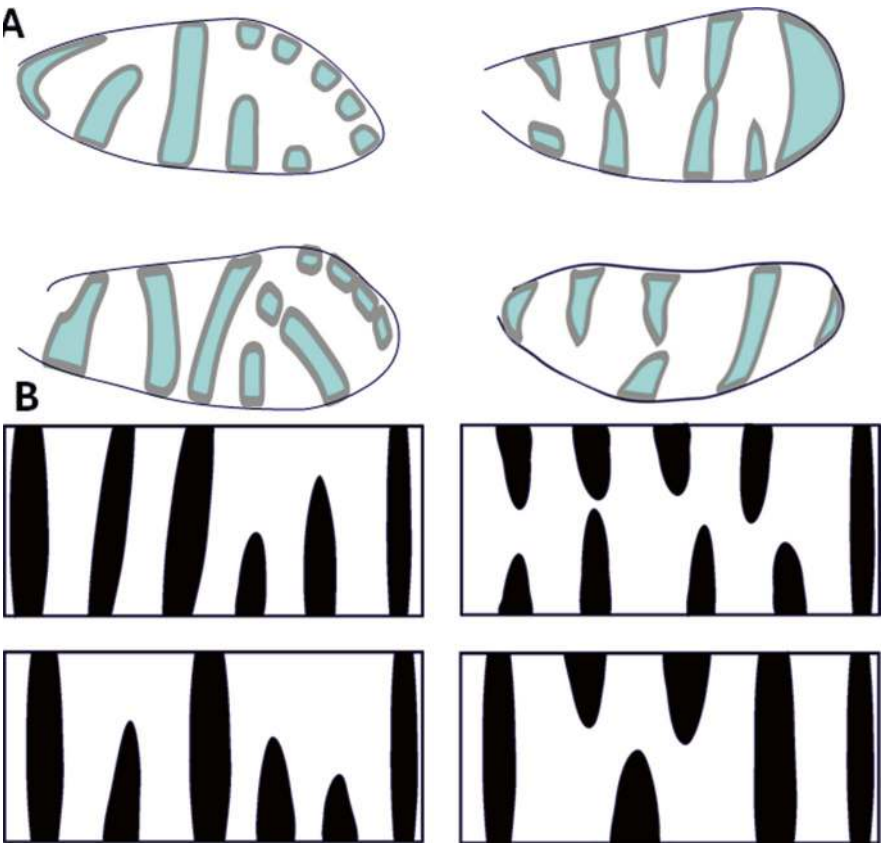


Fig. 5.7 (**A**) Diagrams of banded patterns from Figs. 5.3, 5.4, and 5.5. (**B**) Idealized patterns projected onto a rectangular grid as targets for modeling

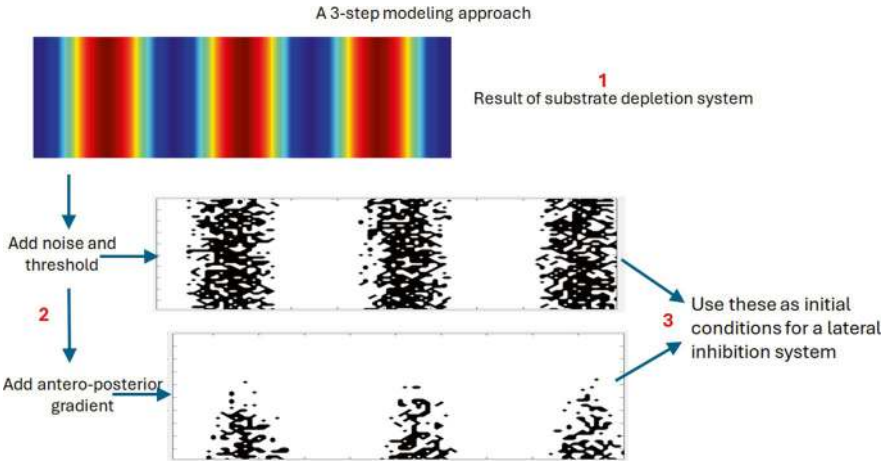


Fig. 5.8 Modeling approach for banded wing patterns

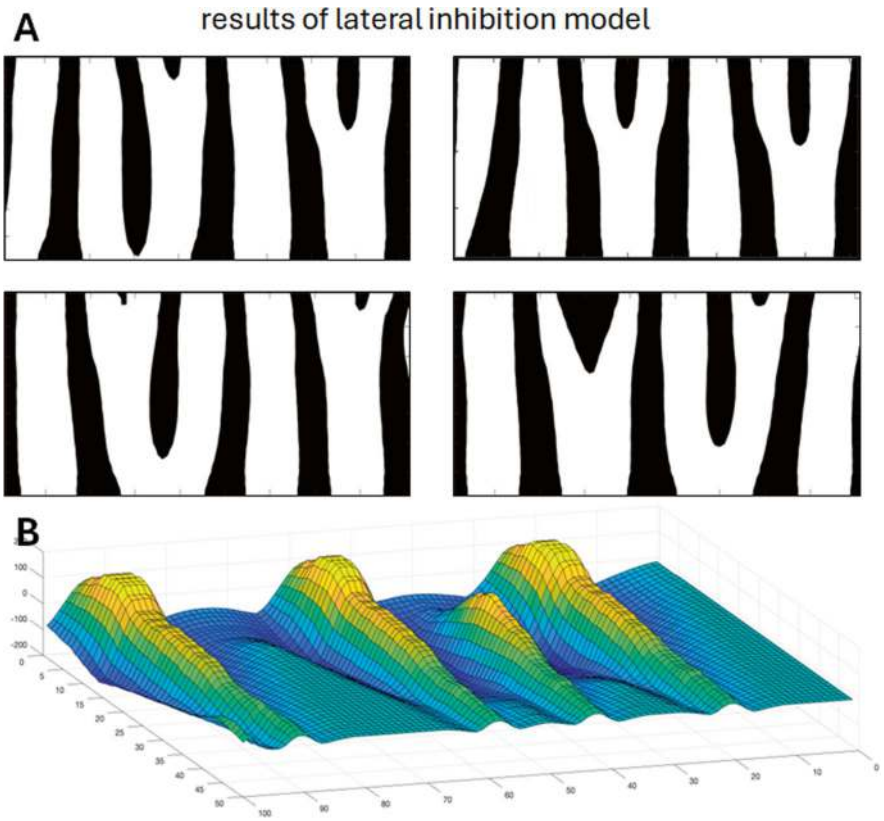


Fig. 5.9 Selected results from a lateral inhibition model approach based on Fig. 5.8: (**A**) thresholded patterns, (**B**) example of surface of lateral inhibition results from which pattern thresholds are taken

5.9 Conclusion

Our comparative analysis suggests that symmetry systems, originally described as parts of the nymphalid ground plan (Schwanwitsch 1924; Süffert 1927; Nijhout 1991), are phylogenetically ancient and can be found in dragonflies, the most basal insects that have color patterns on their wings. Symmetry systems are therefore likely to be the primitive color pattern of insects from which all other patterns are derived and thus may trace back to when the first winged insects evolved, some 350 million years ago (Grimaldi and Engel 2005).

Mathematical modeling shows that a simple reaction–diffusion system can produce a stable pattern of three bands. Each band is a long ridge of morphogen, so depending on how thresholds are chosen, each band can be wide or narrow. Multiple thresholds can specify a system of pigment bands on each side of the ridge, which is characteristic of a symmetry system.

Starting with such a banding system, we have shown that a lateral inhibition mechanism can break up a band into two self-symmetrical bands. Lateral inhibition can also produce a broad diversity of banding patterns, including partial bands that appear to emerge from the wing margin, bands that split, and new bands that intercalate between other bands.

The mathematical modeling shows that making three symmetry systems is simple, and thus provides circumstantial support for the comparative analysis that suggests that such patterns are primitive. Modeling also suggests that a simple lateral inhibition system is sufficient to produce the banding and spotting patterns we find on wings today. Finally, our findings suggest that bands, not spots, are primitive. Rows of spots in the color pattern are most likely to arise when the system that generates bands becomes bounded and compartmentalized by wing veins.

References

Basu DN, Bhaumik V, Kunte K (2023) The tempo and mode of character evolution in the assembly of mimetic communities. Proc Natl Acad Sci USA 120:e2203724120

Beldade P, Koops K, Brakefield PM (2002) Modularity, individuality, and evo-devo in butterfly wings. Proc Natl Acad Sci 99:14262–14267

Gawne R, Nijhout HF (2019) Expanding the nymphalid groundplan's domain of applicability: pattern homologies in an arctiid moth (Utetheisa ornatrix). Biol J Linn Soc 126:912–924

Gawne R, Nijhout HF (2020) The Arctiid archetype: a new lepidopteran groundplan. Front Ecol Evol 8

Gibbs GW (2014) Micropterigidae (Insecta: Lepidoptera). Fauna NZ, Landcare Research 72:5–127

Gibbs GW, Lees DC (2014) New Caledonia as an evolutionary cradle: a re-appraisal of the jaw-moth genus Sabatinca (Lepidoptera: Micropterigidae) and its significance for assessing the antiquity of the Island's fauna. Mémoires du Muséum National d'Histoire Naturelle 206:239–266

Grimaldi D, Engel MS (2005) Evolution of the insects. Cambridge University Press, New York

Meinhardt H (1982) Models of biological pattern formation. Academic, London

Mitsuho K, Haruki T, Kazuki T (2017) Rapid evolution of a Batesian mimicry trait in a butterfly responding to arrival of a new model. Sci Rep 7:6369

Nijhout HF (1991) The development and evolution of butterfly wing patterns. Smithsonian Institution Press, Washington, DC

Nijhout HF (1994) Symmetry systems and compartments in lepidopteran wings: the evolution of a patterning mechanism. Development 1994:225–233

Nijhout HF, Wray GA (1986) Homologies in the colour patterns of the genus *Charaxes* (Lepidoptera: Nymphalidae). Biol J Linn Soc 28:387–410

Nijhout HF, Wray GA (1988) Homologies in the colour patterns of the genus *Heliconius* (Lepidoptera: Nymphalidae). Biol J Linn Soc 33:345–365

Nijhout HF, Wray GA, Gillbert LE (1990) An analysis of the phenotypic effects of certain colour pattern genes in *Heliconius* (Lepidoptera: Nymphalidae). Biol J Linn Soc 40:357–372

Oliver JC, Robertson KA, Monteiro A (2009) Accommodating natural and sexual selection in butterfly wing pattern evolution. Proc Biol Sci 276:2369–2375

Schachat SR, Brown RL (2015) Color pattern on the forewing of *Micropterix* (Lepidoptera: Micropterigidae): insights into the evolution of wing pattern and wing venation in moths. PLoS One 10:e0139972

Schachat SR, Brown RL (2016) Forewing color pattern in Micropterigidae (Insecta: Lepidoptera): homologies between contrast boundaries, and a revised hypothesis for the origin of symmetry systems. BMC Evol Biol 16:116

Schwanwitsch BN (1924) On the ground-plan of wing-pattern in nymphalids and certain other families of rhopalocerous Lepidoptera. Proc Zool Soc London 34:509–528

Schwanwitsch BN (1929) Two schemes of the wing-pattern of butterflies. Z Morphol Okol Tiere 14:36–58

Süffert F (1927) Zur vergleichende Analyse der Schmetterlingszeichnung. Biologisches Zentralblatt 47:385–413

Süffert F (1929) Morphologische Erscheinungsgruppen in der Flügelzeichnung der Schmetterlinge, insbesondere die Querbindenzeichnung. Wilhelm Roux Arch. EntwMech Org 120:299–383

Swindale NV, Barlow HB (1980) A model for the formation of ocular dominance stripes. Proc R Soc London Ser B Biol Sci 208:243–264

Wallbank RW, Baxter SW, Pardo-Diaz C, Hanly JJ, Martin SH, Mallet J, Dasmahapatra KK, Salazar C, Joron M, Nadeau N, McMillan WO, Jiggins CD (2016) Evolutionary novelty in a butterfly wing pattern through enhancer shuffling. PLoS Biol 14:e1002353

Young DA (1984) A local activator–inhibitor model of vertebrate skin patterns. Math Biosci 72:51–58

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Part II
Morphology and Evolution

Chapter 6
Snail Coiling: CRISPR Editing of a Single Gene Turns Righties into Lefties



Reiko Kuroda

Abstract Most animals are symmetric externally but chiral internally. Body-plans, including chirality, are genetically determined, and thus the molecular processes of body-plan establishment are a critically important aspect of "chiromorphology," a research concept that uses chirality to link microscopic phenomena to the macroscopic morphology. We chose to investigate the body-plan of the freshwater snail *Lymnaea stagnalis* as the organism possesses several unique characteristics not found in other animal phyla. Remarkably, by CRISPR/Cas9 genome editing, we could show that snails carrying a homozygous frameshift mutation in the *lsdia1* gene produced sinistral snails over many generations, despite having an otherwise entirely dextral genetic background. Thus, we could prove that this formin-related gene alone determines handedness. Dia proteins are involved in actin polymerization and are present in all eukaryotes. In *L. stagnalis*, chirality is destined by the maternal *Lsdia1* gene already translated by the undivided egg stage. This information is transmitted within and between cells, leading to organismal chirality by controlling the expression site of *nodal*, the important gene for controlling chirality in vertebrates. Using *L. stagnalis*, we hope to elucidate the molecular mechanisms of chirality determination during development that are either universal or specific for vertebrate or invertebrate species.

Keywords Chirality · Chiromorphology · Left–right symmetry · Snail coiling · *L. stagnalis* · CRISPER/Cas9 · Actin assembly · Blastomere

R. Kuroda (✉)
Center for Radical Transformation, Chiba Institute of Technology,
Tsudanuma, Narashino, Chiba, Japan
e-mail: kuroda.reiko@chibatech.ac.jp

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_6

6.1 Introduction

Chirality (also known as handedness) is a geometric property of objects that are mirror images of each other but cannot be superimposed upon one another. In the biological world, chirality is observed at all levels of hierarchy, from molecules to cells, organs, and then to individual organisms. At the molecular level, the biological world is totally homochiral and all living organisms on Earth use molecules of a unique invariant handedness, i.e., only D(dextro)-deoxyribose in DNA and D-ribose in RNA, as well as L(levo)-amino acids in proteins/peptides. At the macroscopic level, most animals appear symmetric externally, but exhibit chirality within the body cavity, as exemplified by the arrangement of visceral organs and their twisted morphology. If the asymmetric arrangement and morphology are not maintained, the organs cannot function normally and may cause serious symptoms, as is the case in *situs ambiguus* or heterotaxy. *Situs inversus*, where the laterality of all organs is completely reversed in a mirror image, occurs in about 1 person in 10,000, but most people have no medical symptoms or complications, because the relationship among organs is not changed. As the body-plan including chirality is genetically determined by intricate molecular interactions between genes and proteins, the cause of *situs inversus* may be elucidated at the molecular level. Similarly, the molecular process by which the body-plan is established from the fertilized egg stage onward is an excellent research subject for studying "chiromorphology," which is a research concept that uses chirality to connect macroscopic and microscopic phenomena across hierarchy in both the biological and non-biological domains (Kuroda 1999).

The body-plans of most bilaterians are composed of three axes: anterior–posterior (A–P), dorsal–ventral (D–V), and left–right (L–R) axes. The A–P axis is essential in defining polarity in bilaterian animals, providing orientation to the organism. Because the front of the body encounters the environment before the rest of the body, sensory organs connected to a central nervous system such as eyes tend to develop there (Finnerty 2005). The D–V axis runs perpendicular to the A–P axis; typically, the central nervous system is positioned dorsally in the deuterostome branch of bilaterians, and ventrally in the protostome branch. These axes are genetically established early in development, but not pre-patterned in the egg, which often has a distinct set of patterning cues that function initially along the animal–vegetal (A–V) axis. The A–V axis is the primary axis of polarity in animal eggs and early embryos, extending from the nutrient-poor, active animal pole to the yolk-rich, less active vegetal pole, which is established during oogenesis. As development proceeds, the A–V axis is transformed, and additional axes are added such that ultimately the definitive A–P, D–V, and L–R axes are established. These morphogenetic events are well-understood at the molecular and cellular levels in several deuterostomes (e.g., sea urchins, ascidians, vertebrates; Munro et al. 2006, Range 2014, Durston and Zhu 2015) and ecdysozoans (e.g., arthropods, nematodes; Rose and Gönczy 2014, Ma et al. 2016), but in the Spiralia/Lophotrochozoa, the third major branch of bilaterians, few studies have documented even the basic morphogenetic events involved in this transformation (Iwata 1985; Smith et al. 1996; Maslakova

et al. 2004; Martín-Durán et al. 2016; Lyons and Henry 2014; Lyons et al. 2017). The Spiralia are a large, morphologically diverse group of protostomes (e.g., mollusks, annelids, nemerteans) that share a homologous mode of early development called spiral cleavage.

In contrast to these well-characterized A–P and D–V axes, which define a dichotomy of biological and physical properties, the L–R axis has essentially no functional distinction. During development, the A–P axis always precedes the D–V axis, while the L–R axis is established last. The left–right (LR) symmetry-breaking steps appear markedly different between vertebrates and invertebrates. For vertebrates, there is general agreement that LR asymmetry is determined by leftward fluid flow, called nodal flow, across the embryonic midline at a late development stage, and the flow is caused by motile cilia (Nonaka et al. 1998, 2002; Okada et al. 2005; Shiratori and Hamada 2006; Hirokawa et al. 2006). This flow is sensed by immotile cilia through mechanical sensing or chemical transport, and triggers *Nodal* expression only on the left side of the embryo. This is reinforced by *Lefty* and *Pitx2* expression (Hamada 2020). Before cilia form, ion fluxes (e.g., H^+/K^+ ATPases) and membrane potential differences across the L–R axis have been observed in frogs and chicks (Levin 2021). These may bias cilia orientation or gene expression, suggesting an upstream symmetry-breaking layer. It has also been shown that the cytoskeleton (comprising actin filaments and microtubules) possesses an intrinsic chirality that can drive left–right asymmetry independently of, or prior to, cilia formation, as observed in vertebrates such as frogs and chicks (Inaki et al. 2024). In invertebrates, studies have been carried out in *Drosophila*, Ascidians, *Caenorhabditis* (*C.*) *elegans*, and mollusks among others. Unconventional type I myosin and the Gα protein GPA-16 are reported to regulate the spindle orientation in *Drosophila* and *C. elegans*, respectively (Bergmann et al. 2003; Hozumi et al. 2006; Spéder et al. 2006). Other mechanisms involving cellular polarity, gap junctions, and (H^+/K^+) ATPase activity have been proposed as potential initial symmetry-breaking steps (Levin et al. 1995; Adams et al. 2002; Levin and Palmer 2007; Vandenberg and Levin 2009).

It is unclear how organisms have acquired diverse mechanisms for determining left–right symmetry during evolution (Vandenberg and Levin 2013; Blum et al. 2014; Coutelis et al. 2014, and references therein). Also poorly understood is the initial step that determines LR asymmetry and how this chirality information is transmitted to determine subsequent body formation in both vertebrates and invertebrates. It turns out that mollusks have proven to be an attractive system in which to address these fundamental questions. This chapter describes studies conducted primarily by my research group that have elucidated the remarkable chiromorphogenesis of the freshwater snail *L. stagnalis* from the fertilized egg stage to the adult organism.

6.2 The Advantages of Studying *L. stagnalis*

We have chosen to study mollusks and especially *L. stagnalis*, as they possess the following unique characteristics not found in other organisms that make for an ideal system for research on chiromorphology:

1. The body of most animals is bilateral externally but chiral internally. However, the chirality of snails is manifested not only internally but also externally in the easily observable directionality of shell coiling. Thus, it is advantageous to study snails. A correlation between the chirality of the spiral cleavage of the blastomere and the directions of shell coiling was reported in 1894 based on observations of the sinistral *Physa* (*P.*) *heterostropha* and dextral *L. columella* (Crampton 1894). By contrast, in other systems such as *Drosophila*, it is necessary to check the rotation direction of genitalia and hindgut.
2. Whereas animals within a phylum generally adopt the same chirality, gastropods are unique wherein ca. 90% of species are dextral and 10% sinistral. Furthermore, the genus *Lymnaea* is dimorphic with 98% dextral and 2% sinistral within a species. Mirror-image pairs within the species serve as perfect internal controls in the study of chiromorphogenesis.
3. *Lymnaea* are hermaphrodites and reproduce by both self-fertilization and sexual interaction with a partner. This feature facilitates the breeding of strains with specific genotypes.
4. The body chirality of animals is determined genetically. The unique hereditary attributes of *Lymnaea* were recognized in 1923 when Boycott and Diver reported the inheritance of sinistrality in *L. peregra* by breeding sinistral and dextral individuals in the laboratory (Boycott and Diver 1923). Sturtevant interpreted the results and proposed a much simpler Mendelian model of maternal inheritance, which is the first clear example ever reported of chromosome-dependent maternal inheritance (Sturtevant 1923). Stimulated by Sturtevant's hypothesis, Boycott et al. performed extensive breeding experiments on *L. peregra*, with 6000 broods and over a million snails, and published an 81-page paper with substantial supplemental data (Boycott et al. 1930). They proposed that chirality is determined by a handedness-determining gene, but their concept of a gene was that in the old Mendelian sense, as their work was carried out before the discovery of DNA as the hereditary material (Avery et al. 1944; Hershey and Chase 1952) and its double helical structure (Watson and Crick 1953).

Except for work by Meshcheryakov and Beloussov (1975), and Meshcheryakov (1990), *L. stagnalis* had not been studied as much as *L. peregra*, However, we chose to use *L. stagnalis* in our chiromorphogenesis studies as our preliminary work had shown simple inheritance for *L. stagnalis* compared with *L. peregra*. Moreover, the size of the egg masses and the number of eggs per egg mass are much larger for *L. stagnalis*.

6.3 Dominant Dextral and Recessive Sinistral Embryos within a Species Are Not Mirror Imaged at the Third Cleavage

In *Lymnaea*, significant chirality is observed at the third spiral cleavage (i.e., from the 4-cell to the 8-cell stage) as a rotation direction of micromeres with respect to the mother macromeres, as described by Crampton (1894). Many textbook

illustrations show mirror-image relationships throughout development for a pair of sinistral and dextral snails, which are drawn as mirror-images of each other. In fact, a mirror-image relationship was not observed for the dextral and sinistral pair of dimorphic *L. stagnalis* in terms of temporal and spatial cytoskeletal dynamics during the third cleavage (Shibazaki et al. 2004). For dextral embryos, a helical deformation of the blastomere was observed at the animal hemisphere in metaphase to anaphase as indicated by the white arrows in Fig. 6.1a(ii) right, and consequently, the new micromeres necessarily bud off dextrotropically from the mother macromeres (Fig. 6.1a(iii) right). The "spiral deformation (SD)" is similar to that described by Meshcheryakov and Beloussov (1975). In marked contrast, no SD occurs in sinistral embryos, and each blastomere presents a symmetrical shape (Fig. 6.1a(ii), (iii) left). The new micromeres emerge almost directly above the paired macromeres along the A–V axis. The sinistral embryos do not show any helicity until furrow ingression, when new micromeres rotate levotropically.

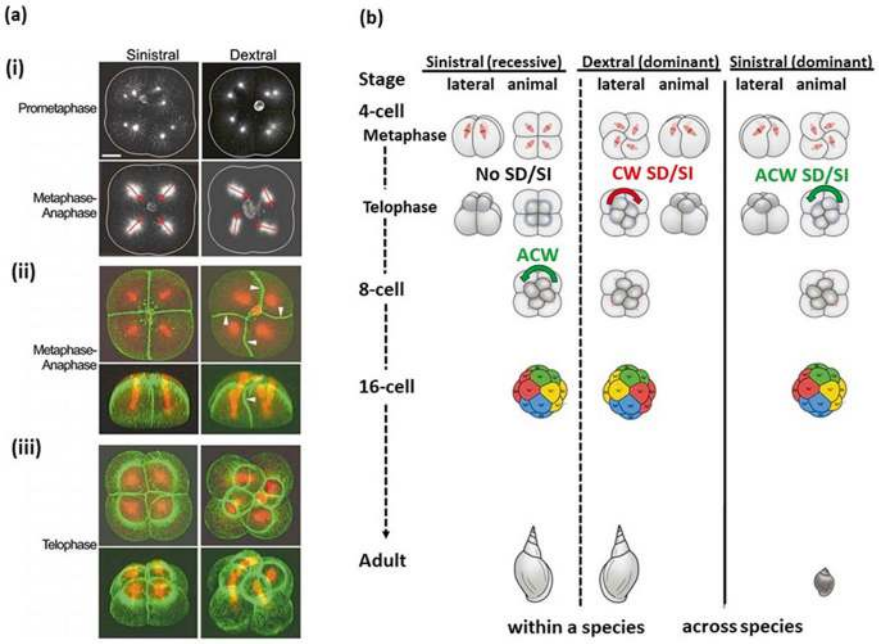


Fig. 6.1 Microscopic images and schematic diagrams illustrating cell division orientation and developmental stages in sinistral and dextral snail embryos. Panel (**a**) shows immunofluorescence images of cells stained for beta-tubulin (red) and filamentous actin (green) at the prometaphase, metaphase-anaphase, and telophase stages of the third cleavage (4–8-cell stage), comparing sinistral and dextral embryos viewed from the animal pole and lateral sides. Red arrows highlight spindle orientation. Panel (**b**) presents a schematic summary of cell cleavage patterns during the 4-cell to 16-cell stages and adult shell coiling phenotypes, comparing sinistral recessive and dextral dominant strains within a species, as well as another species with dominant sinistrality. The diagram indicates the presence or absence of spindle inclination (SI) and spiral deformation of blastomeres (SD), with arrows showing micromere rotation directionality relative to the mother macromeres. Colour-coded cells at the 16-cell stage represent cell lineages from the 4-cell stage

Spindles are oriented horizontally in a random direction at prometaphase in both the dextral and the sinistral embryos (Fig. 6.1a(i)). By metaphase, spindles are fully-grown and oriented along the A–V axis. In the dextral embryos of *L. stagnalis*, spindles orient helically at metaphase, which becomes more substantial by anaphase (Fig. 6.1a(ii) right). In sinistral embryos, however, the spindles in metaphase–anaphase are positioned radially, devoid of chirality, and almost along the A–V axis (Fig. 6.1a(ii) left). Thus, "spindle inclination (SI)" in metaphase–anaphase is a dextral specific event. The non-mirror-image cytoskeletal dynamics was also observed in sinistral and dextral pairs of *L. peregra* (Shibazaki et al. 2004). Thus, the timing of chirality generation is different for the dextral and sinistral blastomeres within a species and not enantiomorphic. Interestingly, a mirror-image relationship of cytoskeletal dynamics was observed in pairs of different species, e.g., sinistral-only monomorphic *P. acuta* and dominant dextral *L. stagnalis*. Sinistral-only monomorphic *P. acuta* exhibits substantial SD and SI levotropically. Figure 6.1b summarizes the cytoskeletal dynamics of dextral and sinistral embryos within and across species (Shibazaki et al. 2004).

Detailed analyses of cytoskeletal dynamics revealed that micromere positions in the sinistral embryos of *L. stagnalis* are somewhat heterogeneous in terms of the direction and extent of the rotation angles relative to the mother macromeres: it shows anti-clockwise (ACW) but sometimes partial ACW (only some out of four micromeres rotated ACW), neutral (vertical), and partial clockwise CW (only some out of four micromeres rotated CW), as well as very rarely clockwise (CW). The proportion of sinistral embryos that develop into sinistral adult snails even with full ACW rotation is low compared with the dextral embryos with CW rotation, 48% vs. 98%, respectively. These observations suggest different cleavage mechanisms may operate for the enantiomorphically-shaped snails within a species (with recessive and dominant chirality). Nevertheless, recessive sinistral strains of *L. stagnalis* exist in the wild and have been kept and propagated in laboratories over many years. SD and SI may provide robustness to the cleavage process.

6.4 Identification of the Handedness-Determining Gene

6.4.1 Positional Cloning

We carried out backcross breeding between dextral and sinistral inbred strains using the dextral animals as donor "fathers" and the sinistral animals as recurrent "mothers." The backcross F10 progeny (99.9% and 0.1% of the sinistral-derived and dextral-derived genome, respectively) indicated that dextrality is determined by a single gene or by genes at closely linked loci (Hosoiri et al. 2003; Kuroda 2014). Inhibition experiments using the actin depolymerization agent latrunculin A showed a concomitant absence of SD and SI, although the spindle assembly and the mitotic process themselves appeared unchanged. Inhibition of spindle assembly in dextral

embryos by the microtubule depolymerization agent nocodazole resulted in even more pronounced SD than in untreated embryos. These experiments clearly show that actin polymerization is responsible for dextral-specific SD, and SI is a consequence of SD (Shibazaki et al. 2004).

Using pure dextral and sinistral strains as well as the F10 backcrossed lines, the single handedness-determining gene locus was mapped by genetic linkage analysis, BAC cloning, and chromosome walking (Kuroda et al. 2016). Thirteen of the 15 genes in the region containing the handedness-determining gene(s) could act with maternal effect, out of which five actin-related genes were selected for study, namely *homologous to fat*, *mycbp2*, *fry*, *dia1*, and *dia2* (Fig. 6.2a). The full-length cDNA of these candidate genes of both dextral and sinistral strains was cloned, and their sequences were compared. A homozygous single-nucleotide deletion in exon 3 of *Lsdia1* was revealed in the sinistral strains (Fig. 6.2b), whereas the other genes did not show any critical mutations that affect evolutionary conserved amino acids

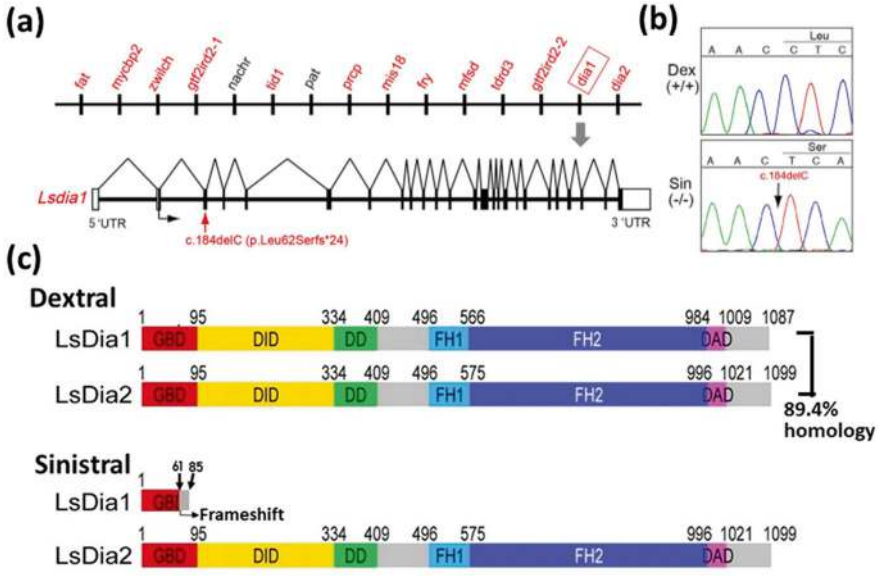


Fig. 6.2 Diagram with three panels illustrating genetic and protein structure related to snail handedness. Panel (**a**) shows a schematic of 15 candidate genes for determining snail handedness, with maternally expressed genes labeled in red. The gene structure diagram of Lsdia1 below highlights its exon-intron layout and a mutation site labeled c.184delC (p.Leu62Serfs*24). Panel (**b**) presents DNA sequencing chromatograms comparing dextral (+/+) and sinistral (-/-) genotypes at the mutation site, showing nucleotide peaks and the deletion indicated by an arrow. Panel (**c**) depicts the schematic protein domain structures of LsDia1 and LsDia2 isoforms for dextral and sinistral L. stagnalis, with coloured blocks representing domains: GBD (red), DID (yellow), DD (green), FH1 (light blue), FH2 (blue), and DAD (purple). In the sinistral snails, LsDia1 is truncated at amino acid 85 as a result of a frameshift mutation at amino acid 61, but it possesses the same LsDia2 as dextral snails. LsDia1 and LsDia2 share 89.4% homology

within the protein coding sequences. Thus, *Lsdia1* was the strongest candidate gene for handedness determination (Kuroda et al. 2016).

There are tandemly duplicated genes, *Lsdia1* (3261 bp and 1087 amino acids) and *Lsdia2* (3297 bp, 1099 amino acids), and the protein products have 89.4% amino acid homology. In the sinistral strains, both *Lsdia1* alleles carry a 184delC deletion that results in a frameshift after amino acid 62 (p.Leu62Serfs*24) and a premature termination at amino acid 85 in the *Lsdia1* mRNA (Fig. 6.2c). *Lsdia2* is expressed maternally in both dextral and sinistral snails (Kuroda et al. 2016).

A paper by Davison et al. (2016) was published just before ours and describes the same gene as associated with LR asymmetry (their *Ldia2* = our *Lsdia1*); however, their DNA sequences of these genes, results of molecular biology work, and the conclusions are different from ours.

6.4.2 CRISPR/Cas9 Genome Editing

At this stage, *Lsdia1* was still just a candidate for the handedness-determining gene. Other genes in other genomic regions may be involved. Therefore, techniques were developed to knock out a particular gene using CRISPR/Cas9 (Abe and Kuroda 2019) and ex-Ovo culture (EOC) of the manipulated embryos to healthy fertile adults (Kuroda et al. 2009; Abe and Kuroda 2019). Genome editing in mollusks is known to be more difficult than in other phyla, and at the time, no successful editing had yet been reported. A single-guide RNA was designed in the FH1 region of *Lsdia1* as that is the very rare site where the base sequences of *Lsdia1* and *Lsdia2* can be distinguished. gRNA and Cas9 mRNA were injected into the dextral strain eggs at the 1-cell stage, and the embryos were carefully raised to adult F0 snails. Adult snails were self-fertilized to produce the F1 generation, and F1 were individually bred, and their genotype was determined.

Remarkably, analysis of many thousands of offspring from each F1 lineage clearly showed that F1 snails with a homozygous knockout mutation (e.g., del5/del5, del13/del17) produced sinistral snails in the otherwise totally dextral genetic background. By contrast, snails with biallelic non-frameshift mutations (e.g., del57/del57) or with heterozygous knockouts where one of the alleles possesses no mutation (wt) or a non-frameshift mutation (e.g., wt/del5, del3/del10) produced only the dextral F2 generation (Fig. 6.3a). Clearly, the genetic chirality control introduced by the CRISPR/Cas9 at F0 is stably inherited and continues generation after generation. This finding has now been verified up to the F20 generation. Thus, *Lsdia1* is the handedness-determining gene sought for over a century. No gene editing occurred in *Lsdia2*, and off-target effects are unlikely judging from the results of several independent genome editing experiments (Abe and Kuroda 2019).

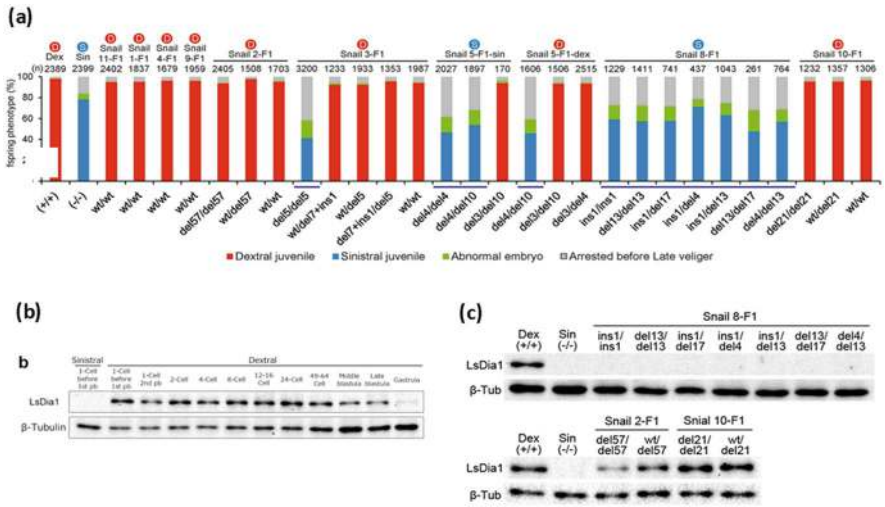


Fig. 6.3 CRISPR/Cas9 genome editing of *Lsdia1*. Panel (**a**) is a bar chart showing the distribution of offspring (F2) phenotypes for each F1 genotype (shown at the bottom of the bar chart), including wild-type, Dex(+/+) and Sin (−/−), control (snail 11), and non-edited F1 (snails 1, 4, and 9). The total number of embryos studied (n) is shown at the top of the bar graph. D in a red circle and S in a blue circle indicate dextral and sinistral snails, respectively. Four phenotype categories are colour-coded: dextral juvenile (red), sinistral juvenile (blue), abnormal embryo (green), and arrested before late veliger (grey). Panel (**b**) shows the presence/absence of LsDia1 protein during the development of the wt sinistral and dextral snails as determined by Western blotting analyses of LsDia1. Panel (**c**) shows the presence/absence of LsDia1 in the genome-edited F2 embryos for each representative genotype of F1 at the stage prior to 1st pb extrusion. (Taken from Abe and Kuroda 2019)

6.4.3 *Perfect Correlation Between the Maternal Genotype and the Presence/Absence of LsDia1 Protein*

Although the amino acid sequences of LsDia1 and LsDia2 are very similar, a specific anti-LsDia1 antibody was generated by targeting the LsDia1-specific sequence in the C-terminal region. Western blot analyses showed that LsDia1 protein was completely absent in sin (−/−), whereas the protein was present in dex (+/+) cells from oviposition to the late blastula stage (Fig. 6.3b). This protein is already present before the first polar body (pb) extrusion, indicating that the gene is translated in the oocyte, consistent with the observation of chiral twisting of embryos at the first cleavage (see Sect. 6.5.1). The evidence that *Lsdia1* is the handedness-determining gene was further strengthened by Western blot analyses of each F1 genotype of *Lsdia1* knockouts. There was a perfect correlation between the maternal genotype

and the presence/absence of the protein in the 1-cell embryos immediately after oviposition (Fig. 6.3c).

6.5 LsDia1 Dictates Cleavage Patterns and *nodal* Expression

6.5.1 *Cleavage Pattern Is Controlled by* Lsdia1 *from the First Cleavage*

To address the question of when chirality first appears during development, we investigated the cytoskeletal dynamics in detail. Embryos treated with trypsin, which reduces the tension of the vitelline membrane, exhibited substantial dynamic chirality during the first cleavage (1-cell to the 2-cell stage). The chirality is mirror-image (ACW or CW rotation) depending on the maternal *Lsdia1* genotype. The observation indicates that the single-fertilized egg carries intrinsic left or right chirality information already at that stage, which is controlled by the presence or absence of LsDia1 protein. This is the earliest observed symmetry-breaking event linked directly to body handedness in the phyla of the animal kingdom.

The third cleavage patterns in which chirality is easily observable without trypsin treatment demonstrated that the progeny of F1 with biallelic non-frameshift mutants and heterozygous knockouts (where one of the alleles is wt or possesses a non-frameshift mutation) exhibited CW rotation with SD and SI, just like the wild-type dextral strains, even with up to a 57-base pair deletion in both alleles. By contrast, the offspring of F1 snails with homozygous frameshifted *Lsdia1* knockouts exhibited cleavage patterns such as the wild-type (−/−) sinistral embryos, i.e., mostly ACW, but sometimes partial ACW, neutral, and rarely partial CW and CW, with no SD/SI (Abe and Kuroda 2019).

6.5.2 *Sites of* nodal *and* Pitx *Gene Expression Are Reversed in* Lsdia1 *Knockouts, Resulting in Mirror-Imaged Animals*

In vertebrates and in some invertebrates, chiromorphogenesis at the organismal level is achieved by asymmetric expression of *nodal* and *Pitx* genes. The Nodal pathway was found in gastropods in the sinistral snail *Biomphalaria glabrata* and the dextral snail *Lottia gigantea* (Grande and Patel 2009), and in dextral and sinistral strains of *Lymnaeids* (Kuroda et al. 2009). In *L. stagnalis*, *nodal* expression was first detected as early as cell stages 33–49 in a specific blastomere destined to develop into ectoderm, on either the right or the left side, for the sinistral and dextral *L. stagnalis*, respectively (Fig. 6.4a, left). At the late trochophore stage, *nodal* is expressed in the left or right lateral ectoderm, near the left or right side of the developing shell (Fig. 6.4b, left).

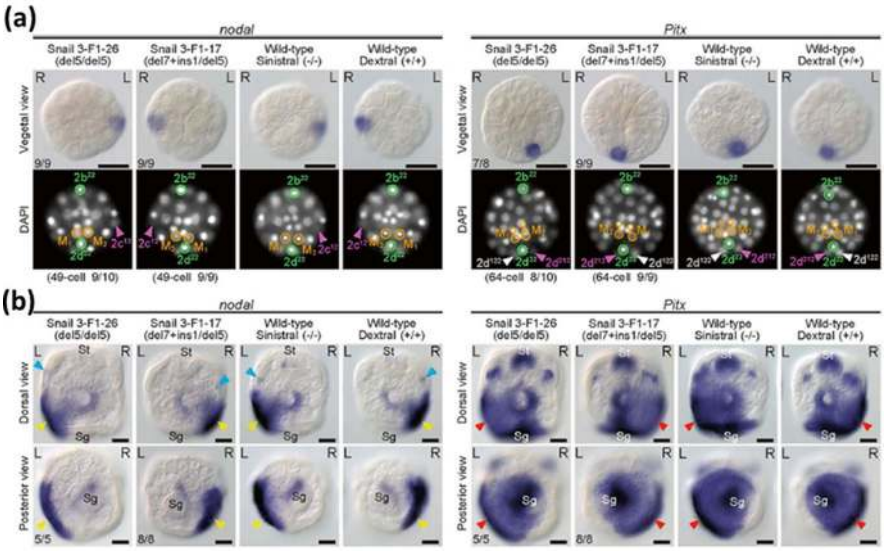


Fig. 6.4 CRISPR/Cas9-mediated gene knockouts show mirror-image expression patterns of *nodal* and *Pitx*. (**a**) Whole-mount in situ hybridization at the 49-cell stage for *nodal* and at the 64-cell stage for *Pitx* for the gene-edited F1 snails as well as wt (−/−) and (+/+). Vegetal view (top) and DAPI (4′,6-diamidino-2-phenylindole) staining to confirm the orientation and the developmental stage of the embryos (bottom). The midline goes between M1 and M2 cells and connects 2b22 and 2d22 cells. Scale bars: 50 μm. (**b**) *nodal* and *Pitx* expression patterns at the late trochophore stage. Dorsal view (top) and posterior view (bottom). The sites of asymmetric expression are indicated by the blue arrowhead (cephalic ectoderm), the yellow arrowhead (lateral ectoderm near the shell gland), and the red arrowhead (posterolateral ectoderm). Sg, shell gland. Scale bars: 50 μm. (Taken from Abe and Kuroda 2019)

Similarly, *Pitx* is expressed asymmetrically in the left or right regions of the posterior and lateral ectoderm, in the stomodeum, and in the visceral mass for sinistral and dextral *L. stagnalis*, respectively (Fig. 6.4a, right). The gene expression continued toward the late developmental stage when asymmetrical morphology was noticeable, and the shell had started to develop (Fig. 6.4b, right) (Kuroda 2014; Kuroda et al. 2016).

Homozygous knockout embryos produced by the gene-edited F1 snails exhibited *nodal* and *Pitx* gene expression patterns that were complete mirror images of the original dextral embryos, whereas expression in heterozygous knockout snails was just like the wild-type dextral embryos (Fig. 6.4a, b) (Abe and Kuroda 2019). Thus, *nodal* and *Pitx* expression patterns are also controlled by *Lsdia1*.

6.6 Mechanogenetics: Creation of Mirror-Image Animals by Micromanipulation

The third cleavage is the most prominent, with substantial CW and ACW rotation of the micromeres for the dextral and the sinistral snails, respectively. Amazingly, at this stage, if the micromeres were gently pushed with two thin glass rods under a microscope to reverse the direction of rotation, mirror-image snails, i.e., *situs inversus*, were born (Kuroda et al. 2009). After the blastomere compaction, the artificially dextralized (or artificially sinistralized) embryos at the third cleavage underwent the fourth cleavage in the opposite sense to that determined genetically, keeping the alternate rotation direction in successive spiral cleavages. This manipulation reversed the *nodal* and *Pitx* expression patterns, and as a result, these manipulated embryos grew to "dextralized" sinistral and "sinistralized" dextral snails, i.e., healthy fertile *situs inversus* organisms. Their offspring, however, adopted the original genetically-determined chirality as is expected, since their genomes had not been manipulated (Kuroda et al. 2009). Similar mechanical reversal of blastomere dynamics during the second cleavage resulted in *situs solitus*, as the blastomere arrangement that was artificially reversed at the 4-cell stage returned to its normal arrangement at the 8-cell stage. After the third cleavage, embryonic chirality could not be mechanically altered. This indicates that chirality, whether genetically determined or artificially manipulated, is permanently fixed at the 8-cell stage, and the subsequent chiromorphogenesis proceeds using the chiral information established at this stage (Kuroda et al. 2009). This handedness-determining process was also found in other snail species. For example, when we performed a similar manipulation on the monomorphic sinistral-only *P. acuta*, dextral snails were born, which are not found naturally in the wild (Abe et al. 2014).

The mechanical micromanipulation at the third cleavage did not move the micromere 1c to the other side of the midline, but it altered the *nodal* expression site (first expressed at cell $2c^{12}$) to the mirror-image site, and resulted in *situs inversus* (Kuroda et al. 2009). This indicates that the identity of A or C (blastomeres at the 4-cell stage are named as A to D) seems not yet specified at the 4-cell stage, and later cell-cell interactions play an important role in cell fate determination (Kuroda 2015). Mechanical treatment of *C. elegans* at the 6-cell stage produced chirality-reversed animals (Wood 1991). This also suggests that in the 6-cell embryo stage, the pair of anterior (AB) blastomeres on the right is equivalent to the pair on the left. In *C. elegans*, spindle orientation seems to be the key in LR determination (Bergmann et al. 2003). In *L. stagnalis*, orthologues of the cell-polarity-related protein Par6 (partitioning defective 6) and its functional partner aPKC (atypical protein kinase C) were found to bind to the mitotic spindles; however, the identical sequence and localization of LsPar6 in the dextral and the sinistral snails exclude the possibility of the gene being the primary determinant of handedness (Homma et al. 2011).

The handedness-determining pathway of *L. stagnalis* starting from the fertilized egg stage to the adults and the next generation is summarized in Fig. 6.5.

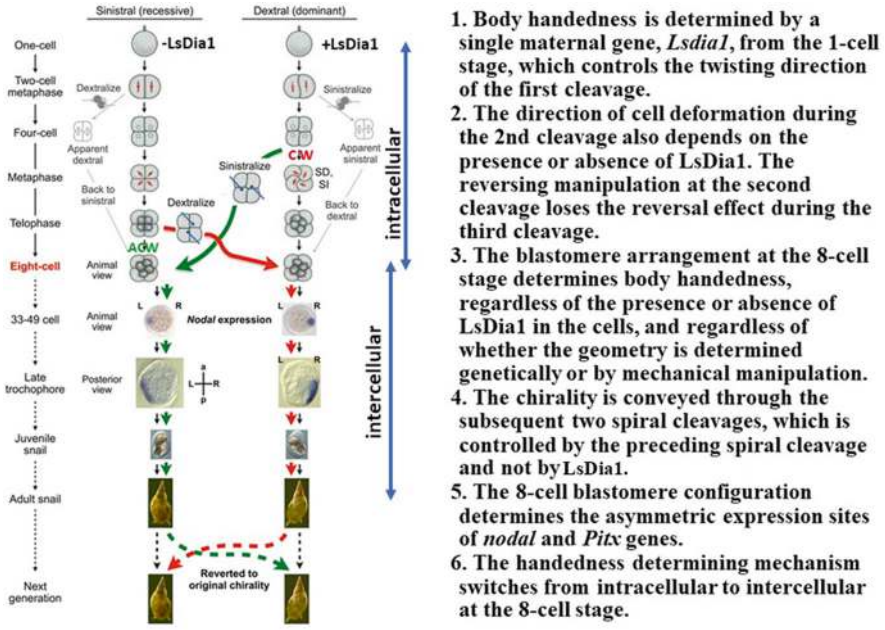


Fig. 6.5 Diagram illustrating the determination of body handedness in snails through cellular and genetic mechanisms. The left side represents sinistral (recessive) snail development without LsDia1, and the right side shows dextral (dominant) snail development with LsDia1. The process begins before the one-cell stage, progresses through two-cell metaphase, four-cell, metaphase, telophase, and reaches the eight-cell stage, with corresponding changes in cell twisting direction and blastomere arrangement. The developmental stages continue through late trochophore, juvenile snail, adult snail, and next generation. The eight-cell blastomere configuration influences asymmetric expression of *nodal* and *Pitx* genes, which determine body handedness. The diagram also notes that handedness control shifts from intracellular to intercellular mechanisms at the eight-cell stage. The thick red or green arrows indicate the process of artificial mechanical manipulation during the third cleavage and the consequent development towards the adult and then to the next generation. Red arrows show the sinistralization of dextral embryos and green arrows the dextralization of sinistral embryos. The artificially dextralized (or artificially sinistralized) embryos at the third cleavage reversed the *nodal* and *Pitx* expression patterns and grew to healthy fertile "dextralized" sinistral and "sinistralized" dextral snails, i.e., situs inversus organisms, overwriting the presence/absence of LsDia1. Their offspring adopted the original genetically-determined chirality as shown by dotted arrows (Kuroda et al. 2009). Accompanying text summarizes six key points about the genetic and cellular basis of handedness determination

6.7 Structure and Functions of the Dia Proteins

6.7.1 Structure of Dia Proteins

Diaphanous-related Formin homology proteins (Dia) control the elongation of actin filaments and are essential for cell division and the formation of cell polarity. They are widely present in eukaryotes. Dia protein is a large protein, consisting of six conserved domains: Rho GTPase-binding domain (GBD), diaphanous inhibitory

domain (DID), dimerization domain (DD), formin homology 1 domain (FH1), formin homology 2 domain (FH2), and diaphanous autoregulatory domain (DAD) (Fig. 6.2c). Dia interacts with various proteins through their respective specific domains. It operates as a dimer formed via the DD dimerization domain (Fig. 6.6a).

The structural studies of Bni1 of *S. cerevisiae* (Moseley et al. 2004), mDia1 of mouse (Rose et al. 2005; Otomo et al. 2005; Nezami et al. 2006; Shimada et al. 2004), and of human Daam1 (Lu et al. 2007; Yamashita et al. 2007) have revealed partial structures, including the DID/FH2-DAD, GBD-DID-DD/RhoC-GTP complexes, DID-DD-partial coiled coil substructure, and FH2 substructure. Structural analysis of the entire protein using cryo-electron microscopy revealed the overall architecture (Maiti et al. 2012) (Fig. 6.6b) and the mechanisms of actin polymerization (Chou and Pollard 2019). The FH1 domain is too flexible to yield a three-dimensional structure by conventional structural analyses.

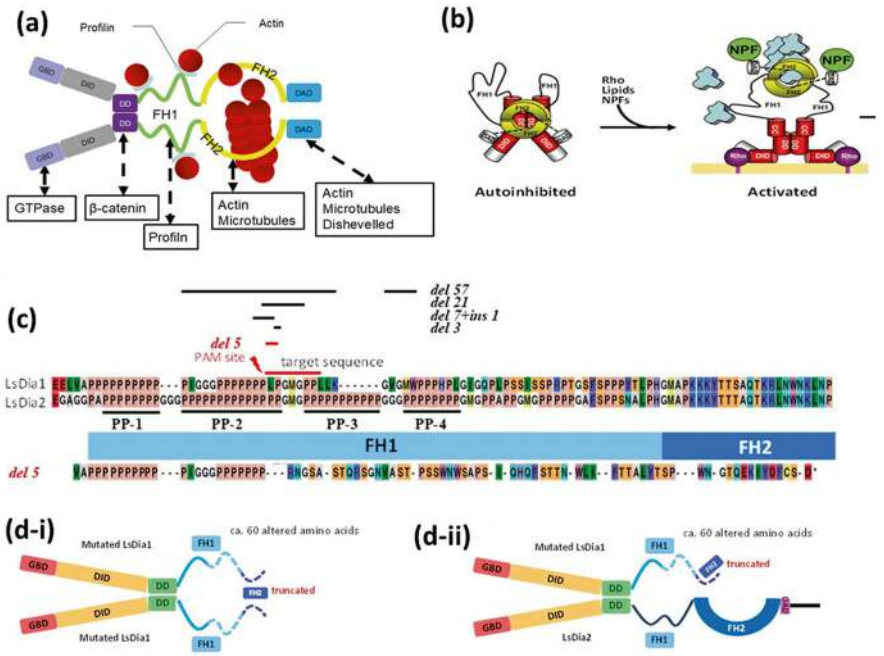


Fig. 6.6 Diagram illustrating the structure, function, and mutation effects of Diaphanous-related formin proteins. Panel (**a**) shows a schematic diagram of the domains of the mouse mDia1 protein, including GBD, DID, DD, FH1, and FH2, and their interactions with molecules such as GTPase, β-catenin, profilin, actin, and microtubules. The protein regulates cytoskeletal dynamics in its dimerized form. Panel (**b**) illustrates the protein in autoinhibited and activated states, highlighting the role of Rho, lipids, and NPFs in activation (modified from Maiti et al. 2012 with permission). Panel (**c**) presents a sequence alignment of two protein variants, LsDia1 and LsDia2. FH1 is characterized by four PP (polyproline) sequences, PP-1 to PP-4. The results of CRISPR/Cas9 genome editing are shown with the PAM (Protospacer Adjacent Motif) site. The sites of non-frameshift deletion which produced dextral snails, i.e., *del 3*, *del 7+ins1*, *del 21* and *del 57*, are shown above the sequence. The homozygous frameshift mutation, *del 5*, resulted in approximately 50 incorrect amino acid residues in the C-terminal side of the FH1 domain and 10 incorrect amino acid residues at the N-terminal side of the FH2 domain followed by a stop codon (shown as a dot) and produced fertile sinistral snails. Panel (**d**) shows hypothetical structures of a homodimer of frameshift mutant LsDia1(i) and of a heterodimer of mutated LsDia1 and LsDia2 (ii)

FH1 is characterized by a PP (poly proline) sequence or proline-rich motifs (PRMs) that capture the profilin (the G-actin–binding protein)–actin complex and transport it directly to the barbed end of F-actin (Pollard 2011), thereby greatly enhancing ATP–actin assembly at the FH2-bound filament tip (Kovar et al. 2006; Mizuno et al. 2011). Computer modeling of the FH1 domain suggests that it comprises a PP region with a rigid helical structure interspersed with flexible segments (Horan et al. 2018). FH2 consists of about 400 amino acids and forms a ring-shaped dimer that binds around the barbed end of F-actin (Fig. 6.6a, b). FH2 at the barbed end elongates the assembling actin filament in a processive manner (Pring et al. 2003; Higashida et al. 2004).

Dia proteins adopt a self-inhibitory structure through interaction between the N-terminal DID domain and the C-terminal DAD domain within the same molecule (Fig. 6.5b). This structure sterically controls the functions of the FH1 and FH2 domains, which are involved in the regulation of actin polymerization and microtubule architecture (Nezami et al. 2006; Maiti et al. 2012). The binding of small Rho-GTPase to the GBD domain, or the binding of NPFs (nucleation promoting factors) to DAD, disrupts the interaction between DID and DAD and allows the protein to adopt an active conformation (Fig. 6.6b).

6.7.2 Difference in LsDia1 and LsDia2 Structures

The amino acid sequences of LsDia1 and LsDia2 proteins are very similar and the differences are concentrated in the FH1 domain. There is also a slight difference in the C-terminal region. LsDia2 FH1 has more PP regions than LsDia1 FH1. There are four PP regions in LsDia2: PP-1, PP-2, PP-3, and PP-4 consisting of 11, 14, 11, and 8 consecutive prolines, respectively. However, in LsDia1, PP-2 and PP-4 are shorter (7 and 3 consecutive Ps instead of 11 and 8 in the PP-2 and PP-4 regions of LsDia2, respectively). There is only one region of two consecutive Ps (i.e., PP) instead of 11 consecutive Ps in the PP-3 of LsDai2. PP-1 is much the same, 10 vs. 8 consecutive Ps in the LsDia1 and LsDia2, respectively (Fig. 6.6c) (Kuroda 2022).

On the other hand, there are hardly any differences in the FH2 domains of LsDia1 and LsDia2 except for a minor change at the C-terminal side. These structural features are common to *L. stagnalis* and *L. peregra* (Kuroda et al. 2016; Kuroda 2022).

6.7.3 Roles of the Dia Domains Revealed by our CRISPR/Cas9 Genome Editing

The *Lsdia1* gene knockout studies provided insights into the functions of LsDia1 (Abe and Kuroda 2019, Kuroda 2022, 2023, to be submitted). The target sequence for CRISPR was designed for the FH1 domain and the mutations introduced in the

FH1 domain included del 3, del 21, del 57, del 7 + ins 1, del 4, del 5, del 10, del 13, del 17, and ins 1 (Fig. 6.6c).

Homozygous frameshift mutations resulted in approximately 50 incorrect amino acid residues in the C-terminal side of the FH1 domain and 10 incorrect amino acid residues at the N-terminal side of the FH2 domain followed by a stop codon and produced fertile sinistral snails. Namely, the genome editing with homozygous frameshift mutations produced LsDia1, which has a correct structure up to 35% of the FH1 domain (i.e., up to residue 517). This result indicates that all or part of the two-thirds of FH1 (C-terminal side)-FH2-DAD-C-terminal domains play a crucial role in realizing dextrality. As the truncated LsDia1 proteins have the DD domain, they may dimerize to form a homodimer of truncated LsDia1 (Fig. 6.6d, i) or a heterodimer with normal LsDia2 (Fig. 6.6d, ii). The frameshift mutants lost the DAD domain and thus the knockout products cannot form the autoinhibitory structure through the DID–DAD interaction.

In contrast, mutations that do not generate a frameshift, even with loss of 57 bases (i.e., 19 amino acid residues corresponding to the deletion of complete PP-2 and the portion before the PP-4) retained LsDia1 function and resulted in healthy dextral snails. This suggests that these regions, which have significant amino acid sequence differences between LsDia1 and LsDia2, are not directly responsible for the determination of handedness. The non-frameshift mutations may affect the polymerization rate (Kuroda 2022, 2023, to be submitted).

Sinistral snails produced by our gene editing of dextral snail embryos showed a higher ratio of abnormality than did the wild sinistral snail embryos during their development. Naturally occurring sinistral *L. stagnalis* snails have an early stop codon near the N-terminus, and thus even a short partial protein is not produced. The truncated LsDia1 with partial incorrect amino acid sequences created by the genome editing may cause toxicity.

6.7.4 Dia Proteins and SD/SI in Pond Snails

Snails that belong to the Lymnaeids, i.e., *L. stagnalis* and *L. peregra*, are dimorphic and exist in the wild as predominantly dextral (dominant) and minorly sinistral (recessive). Dextral snails have both type 1 and type 2 Dia proteins (LsDia1/LsDia2 and LpDia1/LpDia2) and exhibit CW SD/SI followed by CW micromere rotation at the third cleavage. Sinistral only snails, *P. acuta* and *Indoplanorbis* (*I.*) *exustus*, possess only type2 Dia protein and exhibit ACW SD/SI followed by ACW micromere rotation.

Namely, for different opposite coiling snail species, i.e., dominant dextral dimorphic snails and sinistral only snails, the cytoskeletal dynamics are mirror images of each other like their exterior outlook (Fig. 6.1b). Thus, it may be reasonable to assume that the directions of SD/SI and micromere rotation are coupled and managed by the same mechanism. Does this mean that LsDia1 determines the dextral cytoskeletal dynamics?

However, recessive sinistral snails of Lymnaeid show no SD/SI but micromeres predominantly rotate ACW (Kuroda et al. 2016), indicating that micromere rotation and SD/SI are independent. It is intriguing how sinistral *L. stagnalis* rotates levotropically after the non-chiral emergence of micromeres. Study on the minor recessive sinistral snail has revealed the complex nature of the cytoskeletal dynamics involved.

Type 1 and 2 *Lsdia* genes are tandemly repeated and exhibit 89.4% amino acid homology. Therefore, one of them is presumably derived from the other during evolution. These proteins probably have overlapping roles, but at the same time, they must have specialized functions. A recent molecular phylogeny study of *dia* genes in gastropods claims that duplication or loss of *dia* does not play a primary role in reversal evolution based, in part, on the overstatement that reversed *L. stagnalis* barely survive to the veliger stage (Noda et al. 2023). In fact, sinistral snails that lack *Lsdia1* exist in the wild. Sinistral snails produced by knocking out the *Lsdia1* gene in dextral *L. stagnalis* show a higher rate of abnormalities during development but are not lethal and reproduce for many generations (Kuroda et al. 2016, Abe and Kuroda 2019, Kuroda to be submitted). Clearly, further work is needed to define the roles of the type 1 and type 2 Dia proteins. Moreover, additional studies that include other snail species may lead to a greater understanding of the important and complex process of chiral evolution.

6.8 Prospects

We have demonstrated that the right-handedness of *L. stagnalis* is determined by a single formin-related gene, *Lsdia1*. Formin family proteins are actin polymerization nucleation promoters that are widely distributed in eukaryotes. The reason why the coexistence of two proteins, LsDia1 and LsDia2, which are 89.4% homologous, results in dextral snails, while the presence of LsDia2 alone results in sinistral snails is profound. We are currently examining the role of *Lsdia2*. Differences in the amino acid sequences of LsDia1 and LsDia2 are concentrated in the FH1 domain and not in other domains except for a slight difference in the C-terminal region. Genome editing revealed that partial loss of FH1 (without frameshift) makes no difference to chirality, and that all or part of the two-thirds of FH1 (C-terminal side)-FH2-DAD-C-terminal domains play a crucial role in achieving dextrality. Chirality information already established at the undivided egg stage is transmitted within and between cells and determines the chirality of individual organisms via controlling the expression site of the *nodal* gene, the important gene for chirality determination in vertebrates. Using *L. stagnalis*, we hope to elucidate the molecular mechanisms of chirality determination that are common between or specific to vertebrate and invertebrate species.

Acknowledgments This work was initially financially supported by the Japan Science and Technology Agency (JST) to R.K.: ERATO Kuroda Chiromorphology Project (1999–2004) and ERATO-SORST Kuroda Chiromorphology Team (2004–2009), and by the JSPS for Grants-in-Aid for Scientific Research, Challenging Exploratory Research (24657249) to R.K. (2012–2014). The author thanks members of the Kuroda Research Group, especially Y. Shibazaki, K. Fujikura, B. Endo, and M. Abe.

References

Abe M, Kuroda R (2019) The development of CRISPR for a mollusc establishes the formin Lsdia1 as the long-sought gene for snail dextral/sinistral coiling. Development 146:dev17597

Abe M, Takahashi H, Kuroda R (2014) Spiral cleavages determine the left-right body plan by regulating the Nodal pathway in monomorphic gastropods, Physa acuta. Int J Dev Biol 58:513–520

Adams DS, Robinson KR, Fukumoto T, Yuan S, Albertson RC, Yelick P, Kuo L, McSweeney M, Levin M (2002) Early H^+-V-ATPase-dependent proton flux is necessary for consistent left-right patterning of non-mammalian vertebrates. Development 133:1657–1671

Avery OT et al (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types. J Exp Med 79:137–157

Bergmann DC, Lee M, Robertson B, Tsou MF, Rose LS, Wood WB (2003) Embryonic handedness choice in *C. elegans* involves the Gα protein GPA-16. Development 130:5731–5740

Blum M, Fiested K, Thumberger T, Schweickert A (2014) The evolution and conservation of left-right patterning mechanisms. Development 141:1603–1613

Boycott AE, Diver C (1923) On the inheritance of sinistrality in *Limnaea peregra*. Proc R Soc Lond B Biol Sci 95:207–213

Boycott AE, Diver C, Garstang SL, Hardy AC, Turner FM (1930) The inheritance of sinistrality in *Lymnaea peregra*. Phil Trans R Soc Lond B 219:51–131

Chou SZ, Pollard TD (2019) Mechanism of actin polymerization revealed bycryo-EM structures of actin filaments with three different bound nucleotides. Proc Natl Acad Sci USA 116:4265–4274

Coutelis J-B, Gonzalez-Morales N, Geminard C, Noselli S (2014) Diversity and convergence in the mechanisms establishing L/R asymmetry in metazoa. EMBO J 15:926–937

Crampton HE (1894) Reversal of cleavage in a sinistral gastropod. Ann N Y Acad Sci 8:167–170

Davison A, McDowell GS, Holden JM, Johnson HF, Koutsovoulos GD, Liu MM, Hulpiau P, Van Roy F, Wade CM, Banerjee R, Yang F, Chiba S, Davey JW, Jackson DJ, Levin M, Blaxter ML (2016) Formin is associated with left-right asymmetry in the pond snail and the frog. Curr Biol 26:654–660

Durston AJ, Zhu K (2015) A time space translation hypothesis for vertebrate axial patterning. Semin Cell Dev Biol 42:86–93

Finnerty JR (2005) Did internal transport, rather than directed locomotion, favor the evolution of bilateral symmetry in animals? BioEssays 27:1174–1180

Grande C, Patel NH (2009) Nodal signaling is involved in l–r asymmetry in snails. Nature 457:1007–1011

Hamada H (2020) Molecular and cellular basis of left–right asymmetry in vertebrates. Proc Jpn Acad Ser B 96:273–296

Hershey AD, Chase M (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. J Gen Physiol 36:39–56

Higashida C, Miyoshi T, Fujita A, Oceguera-Yanez F, Monypenny J, Andou Y, Narumiya S, Watanabe N (2004) Actin polymerization–driven molecular movement of mDia1 in living cells. Science 303:2007–2010

Hirokawa N, Tanaka T, Okada Y, Takeda S (2006) Nodal flow and the generation of left-right asymmetry. Cell 125:33–45

Homma T, Shimizu M, Kuroda R (2011) The polarity protein Par6 is coupled to the microtubule network during molluscan early embryogenesis. Biochem Biophys Res Commun 404:173–178
Horan BG, Zerze GH, Kim YC, Vavylonis D, Mittal J (2018) Computational modeling highlights the role of the disordered Formin Homology 1 domain in profilin-actin transfer. FEBS Lett 592:1804–1816
Hosoiri Y, Harada Y, Kuroda R (2003) Construction of a backcross progeny collection of dextral and sinistral individuals of a freshwater gastropod, *Lymnaea stagnalis*. Dev Genes Evol 213:193–198
Hozumi S, Maeda R, Taniguchi K, Kanai M, Shirakabe S, Sasamura T, Spéder P, Noselli S, Aigaki T, Murakami R (2006) An unconventional myosin in Drosophila reverses the default handedness in visceral organs. Nature 440:798–802
Inaki M, Higashi T, Okuda S, Matsuno K (2024) Distinctive cellular and junctional dynamics independently regulate the rotation and elongation of the embryonic gut in *Drosophila*. PLoS Genet 20:e1011422
Iwata F (1985) Foregut formation of the Nemerteans and its role in Nemertean Systematics. Integr Comp Biol 25:23–36
Kovar DR, Harris ES, Mahaffy R, Higgs HN, Pollard TD (2006) Control of the assembly of ATP- and ADP-actin by formins and profilin. Cell 124:423–435
Kuroda R (1999) Chiromorphology. JST ERATO project
Kuroda R (2014) How a single gene twists a snail. Integr Comp Biol e54:677–687
Kuroda R (2015) A twisting story: how a single gene twists a snail? Q Rev Biophys 48:445–452
Kuroda R (2022) Exploring the structural difference in homologous proteins that are crucial in organismal body handedness determination. Sogo-Kogaku 34:38–45
Kuroda R (2023) Exploring the functions of the handedness-determining single gene Lsdia1 in freshwater snail *Lymnaea stagnalis*: influence on the normal development. Sogo-Kogaku 35:45–52
Kuroda R, Endo B, Abe M, Shimizu M (2009) Chiral blastomere arrangement dictates zygotic left–right asymmetry pathway in snails. Nature 462:790–794
Kuroda R, Fujikura K, Abe M, Hosoiri Y, Asakawa S, Shimizu M, Umeda S, Ichikawa F, Takahashi H (2016) Diaphanous gene mutation affects spiral cleavage and chirality in snails. Sci Rep 6:34809
Levin M (2021) Bioelectric signaling: reprogrammable circuits underlying embryogenesis, regeneration, and cancer. Cell 184:1971–1989
Levin M, Palmer AR (2007) Left–right patterning from the inside out: widespread evidence for intracellular control. BioEssays 29:271–287
Levin M, Johnson RL, Stern CD, Kuehn M, Tabin C (1995) A molecular pathway determining left–right asymmetry in chick embryogenesis. Cell 82:803–814
Lu J, Meng W, Poy F, Maiti S, Goode BL, Eck MJ (2007) Structure of the FH2 domain of Daam1: implications for formin regulation of actin assembly. J Mol Biol 369:1258–1269
Lyons DC, Henry JQ (2014) Ins and outs of Spiralian gastrulation. Int J Dev Biol 58:413–428
Lyons DC, Perry KJ, Henry JQ (2017) Morphogenesis along the animal-vegetal axis: fates of primary quartet micromere daughters in the gastropod Crepidula fornicate. BMC Evol Biol 17:217
Ma J, He F, Xie G, Deng W-M (2016) Maternal AP determinants in the drosophila oocyte and embryo. Wiley Interdiscip Rev Dev Biol 5:562–581
Maiti S, Michelot A, Gould C, Blanchoin L, Sokolova O, Goode BL (2012) Structure and activity of full-length formin mDia1. Cytoskeleton 69:393–405
Martín-Durán JM, Passamaneck YJ, Martindale MQ, Hejnol A (2016) The developmental basis for the recurrent evolution of deuterostomy and protostomy. Nat Publ Group 1:005
Maslakova SA, Martindale MQ, Norenburg JL (2004) Vestigial prototroch in a basal nemertean, Carinoma tremaphoros (Nemertea; Palaeonemertea). Evol Dev 6:219–226
Meshcheryakov VN (1990) The common pond snail *Lymnaea stagnalis*. In: Dettlaff TA, Vassetzky SG (eds) Animal species for developmental studies. Consultants Bureau, New York, pp 69–132
Meshcheryakov VN, Beloussov LV (1975) Asymmetrical rotations of blastomeres in early cleavage of Gastropoda. Wilhelm Roux's Arch 177:193–203

Mizuno H, Higashida C, Yuan Y, Ishizaki T, Narumiya S, Watanabe N (2011) Rotational movement of the formin mDia1 along the double helical strand of an actin filament. Science 331:80–83
Moseley JB, Sagot I, Manning AL, Xu Y, Eck M, Pellman D, Goode BL (2004) A conserved mechanism for Bni1- and mDia1-induced actin assembly and dual regulation of Bni1 by Bud6 and profilin. Mol Biol Cell 15:896–907
Munro E, Robin F, Lemaire P (2006) Cellular morphogenesis in ascidians: how to shape a simple tadpole. Curr Opin Genet Dev 16:399–405
Nezami AG, Poy F, Eck MJ (2006) Structure of the autoinhibitory switch in formin mDia1. Structure 14:257–263
Noda T, Satoh N, Gittenberger E, Asami T (2023) Left–right reversal recurrently evolved regardless of diaphanous-related formin gene duplication or loss in snails. J Mol Evol 91:721–729
Nonaka S, Tanaka Y, Okada Y, Takeda S, Harada A, Kanai Y, Kido M, Hirokawa N (1998) Randomization of left–right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. Cell 95:829–837
Nonaka S, Shiratori H, Saijoh Y, Hamada H (2002) Determination of left–right patterning of the mouse embryo by artificial nodal flow. Nature 418:96–99
Okada Y, Takeda S, Tanaka Y, Belmonte JCI, Hirokawa N (2005) Mechanism of nodal flow: a conserved symmetry. Cell 121:633–644
Otomo T, Otomo C, Tomchick DR, Machius M, Rosen MK (2005) Structural basis of Rho GTPase-mediated activation of the formin mDia1. Mol Cell 18:273–281
Pollard TD (2011) Formin tip tracking – a simple protein machine tracks the tip of a growing actin filament. Science 331:39
Pring M, Evangelista M, Boone C, Yang C, Zigmond SH (2003) Mechanism of formin-induced nucleation of actin filaments. Biochemistry 42:486–496
Range R (2014) Specification and positioning of the anterior neuroectoderm in deuterostome embryos. Genesis 52:222–234. https://doi.org/10.1002/dvg.22759
Rose L, Gönczy P (2014) Polarity establishment, asymmetric division and segregation of fate determinants in early C. elegans embryos. In: WormBook: the online review of *C. elegans* biology, pp 1–43
Rose R, Weyand M, Lammers M, Ishizaki T, Ahmadian MR, Wittinghofer A (2005) Structural and mechanistic insights into the interaction between Rho and mammalian Dia. Nature 435:513
Shibazaki Y, Shimizu M, Kuroda R (2004) Body handedness is directed by genetically determined cytoskeletal dynamics in the early embryo. Curr Biol 14:1462–1467
Shimada A, Nyitrai M, Vetter IR, Kühlmann D, Bugyi B, Narumiya S, Geeves MA, Wittinghofer A (2004) The core FH2 domain of diaphanous-related formins is an elongated actin binding protein that inhibits polymerization. Mol Cell 13:511–522
Shiratori S, Hamada H (2006) The left-right axis in the mouse: from origin to morphology. Development 133:2095–2104
Smith CM, Lans D, Weisblat DA (1996) Cellular mechanisms of epiboly in leech embryos. Development 122:1885–1894
Spéder P, Adám G, Noselli S (2006) Type ID unconventional myosin controls left–right asymmetry in Drosophila. Nature 440:803–807
Sturtevant AH (1923) Inheritance of coiling direction in *Lymnaea*. Science 58:269–270
Vandenberg LN, Levin M (2009) Perspectives and open problems in the early phases of left–right patterning. Semin Cell Dev Biol 20:456–463
Vandenberg LN, Levin M (2013) A unified model for L–R asymmetry? Comparison and synthesis of molecular models of embryonic laterality. Dev Biol 379:1–15
Watson JD, Crick FHC (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. Nature 171:727–738
Wood WB (1991) Evidence for reversal of handedness in C. elegans embryos for early cell interactions determining cell fate. Nature 349:536–368
Yamashita M, Higashi T, Suetsugu S, Sato Y, Ikeda T, Shirakawa R, Kita T, Takenawa T, Horiuchi H, Fukai S, Nureki O (2007) Crystal structure of human DAAM1 formin homology 2 domain. Genes Cells 12:1255–1265

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 7
Morphospace Approaches to Evolution: Geometric and Developmental Considerations



Sylvain Gerber

Abstract Morphospaces are spatial representations of patterns of morphological variation. They have been of both conceptual and empirical importance in biology and paleobiology. Conceptually, morphospaces have fueled critical discussions regarding the structure of variation and the roles of natural selection, evolutionary constraints, and historical contingencies underlying this structure. Empirically, morphospaces have proven powerful quantitative tools to document and analyze morphological patterns in various research areas from population-level morphometric analyses to large-scale studies of clade dynamics. In this chapter, we review the diverse kinds of morphospaces encountered in morphological studies and the implications of their geometry for the interpretations of evolutionary patterns. We also show how the consideration of organismal development, for instance through the use of pattern formation models, suggests alternative morphospace representations favoring evolutionary accessibility over morphological similarity as a criterion for organizing morphological variants.

Keywords Development · Evolutionary constraints · Morphometrics · Morphospace · Pattern formation · Morphogenesis

7.1 Introduction

In an important paper on the role of development in phenotypic evolution, Alberch argued that "Nature is not chaos, neither is it a boundless continuum of forms" (Alberch 1982). In other words, morphological variation in the natural world is characterized by order and discontinuities. Alberch then proceeded with a thought experiment on the spatial representation of this general morphological pattern. If one conceptualizes an abstract space capable of encompassing all the diversity of

S. Gerber (✉)
Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, EPHE, Sorbonne Université, UA, Paris, France
e-mail: sylvain.gerber@mnhn.fr

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_7

forms encountered in nature—each one depicted as a point in this space—one would observe a markedly nonuniform distribution of biological forms: there would be noticeable hierarchically structured clusters of forms and empty ("non-occupied") regions in the space. This structured variation constitutes a foundational basis in systematics (the classification of species into a hierarchical taxonomic system) and in discussions on the relative importance of internal and external factors in morphological evolution (e.g., Alberch 1982, 1989; Erwin 2007; Gould 2002; Maynard Smith et al. 1985; Wagner 2014; Wake and Larson 1987).

At this highest macroevolutionary level of morphological variation, the disparity of body plans renders impossible the construction of such a space, for lack of a sufficiently general and accommodating morphometric framework (Gould 1991). In studies of particular morphological structures and at lower taxonomic levels, however, the notion of morphological space, or morphospace, becomes empirically and methodologically tractable. One of the earliest concrete examples of morphospace is found in the work of Raup with the morphospatial representation of the diversity of shell shapes of mollusks and brachiopods (Raup 1966). Raup revisited the seminal work of earlier authors (e.g., Moseley 1838; Thompson 1942) to operationalize the quantitative characterization of the geometry of shell coiling using the logarithmic spiral model. His morphometric scheme allows the description of shell coiling geometry with three parameters and the depiction of possible shell shapes in a three-dimensional morphospace. Just as described by Alberch, the placing of real shell forms in this space revealed a strongly nonuniform distribution with more or less densely occupied regions and empty ones. This morphospace pattern has been used as a key example in discussions of developmental constraints and how such constraints might bias the direction of evolution through the restriction of morphological variation exposed to natural selection (e.g., Maynard Smith et al. 1985; McGhee 1999).

With the advent of morphometrics, the statistical analysis of form and morphological variation, the construction of morphospaces has become a well-established approach to analyze and visualize morphological patterns and address a variety of biological questions (Adams et al. 2004; Mitteroecker and Schaefer 2022). At the (macro)evolutionary level, morphospace analyses are, for instance, used to investigate questions about morphological disparity, the spread and spacing of taxa in morphospace (e.g., Foote 1997; Hopkins and Gerber 2021; Wills 2001) and large-scale trends and modes of trait evolution (e.g., Collyer and Adams 2013; Hopkins 2016; Hunt 2007; Stayton 2015). Implicit in these questions and their study with morphospace approaches is the idea that morphospaces possess geometric properties that adequately capture the empirical pattern and allow a meaningful representation of the biological and evolutionary concepts involved. In fact, many of the tools routinely used to explore morphospaces indicate that a metric structure is often assumed for morphospaces, regardless of the morphological descriptors used for their construction. Likewise, the visual depiction of a morphospace is often interpreted through the prism of Euclidean intuition. However, do all morphospaces encountered in biology and paleobiology actually possess metric structures? Can the visualization of morphospaces misguide biological interpretations? Are metric

topologies relevant to all research questions? The present chapter reviews the common types of morphospaces and discusses recent work in morphospace theory regarding these questions.

7.2 Affine and Euclidean Structures in Morphospaces

7.2.1 Theoretical Morphology

In mathematics, a space is defined as a set endowed with a structure establishing relationships among its elements. In biology and paleobiology, morphospaces are not usually defined in such a formal way but instead recognized from the type of morphological descriptors used to describe and distinguish morphological variants. These different kinds of morphospaces are generally visualized in the same way, however, with observed variants abstracted as points in a space of possible variants. The structure on this set is implicitly considered as some distance function—a morphological distance—that quantifies the degree of morphological distinctness among variants.

In geometric morphometrics, the most widespread framework for morphological analysis, the shape distance between two variants described by configurations of landmarks is the Procrustes metric of Kendall's shape spaces, and the geometry of these spaces is well-described and understood (Kendall et al. 1999). However, Mitteroecker and Huttegger (2009) note that for many other morphospaces, a metric is assumed but is in fact not clearly defined nor necessarily definable. Typical instances of such cases occur in the field of theoretical morphology (McGhee 1999; Eble 2000). The field has been mostly developed in paleobiology, but it shares some of its goals with the work on pattern formation in mathematical and developmental biology. In theoretical morphology, a mathematical model simulates the shape of the biological structure of interest, and the model parameters are taken as descriptors of the shape. The classic and most employed theoretical morphological model, both for its theoretical value and morphometric efficiency, is the shell coiling model of Raup mentioned in the introduction. It has given rise to numerous empirical applications and has motivated the proposal of many derived theoretical shell coiling models (see Urdy 2015 for a review). If one approximates the shell apertural margin (peristome) as a circular generating curve revolving around and translating along a coiling axis as growth proceeds isometrically, each of its points draws a conchospiral. In a cylindrical coordinate system, the path of the center of this generative curve for n revolutions is given by:

$$r(\theta) = \left(r_0 e^{k\theta}, r_0\left(e^{k\theta} - 1\right)\cot\alpha\right),$$

with r_0, $k > 0$, $\alpha \in \left(0, \pi/2\right]$, and $\theta \in [0, n2\pi]$. Three parameters are extracted from the model to describe the shell shape: The expansion rate $W = e^{k2\pi}$, corresponding to

the ratio of center radii one revolution apart; the translation rate $T = \cot \alpha$, corresponding to the ratio of the vertical over the horizontal shifts of the center one revolution apart; and the distance of the generating curve from the coiling axis $D = (r_0 - R_0)/(r_0 + R_0)$, where R_0 is the radius of the generating curve measured at $\theta = 0$. Any 3-tuple of values for these parameters generates a unique shell geometry and points to a unique location in Raup's iconic cube (Raup 1966). How do we quantify the morphological dissimilarity between two shells in this space? As ratios of lengths, the three parameters are dimensionless quantities. They also widely differ in their ranges of possible values. The parameter D is bound to the unit interval while W and T have no theoretical upper bounds. Because these parameters are incommensurate, there is no natural inner product that can be defined to equip the morphospace. As a consequence, there is no notion of distance and direction in the morphospace, and its historical portrayal as a three-dimensional space with orthogonal coordinate axes is arbitrary. Any alternative scaling and orientation of the axes is equally valid and may yet suggest a very different picture of the morphological pattern. The problem is exemplified in Fig. 7.1, which focuses on the $D - W$ plane of Raup's space, i.e., the subspace of planispiral coiling geometries ($T = 0$), which includes the shells of most cephalopods (ammonoids, nautiloids). Figure 7.1a illustrates the distribution of Paleozoic ammonoid genera (data from Saunders et al. 2004) with the scales and angle among the parameter axes following Raup's original depiction (Raup 1967). Devoid of inner product, the morphospace only exhibits an affine geometry. It is easy to intuit from the figure that changing the scales and angles between W and D can radically alter our perception of closeness of morphological variants and of the relative proportion of occupied vs. unoccupied regions of morphospaces. This perceptual bias emerges from our natural tendency to attach Euclidean properties to pictorial representation of spaces (Mitteroecker and Huttegger 2009).

Likewise, many other models in theoretical morphology, while effective at simulating particular morphological systems, share similar geometric limitations for the quantitative assessment of morphological variation due to incommensurate parameters. For instance, the parameters of Niklas' land plant morphospace include a bifurcation angle, a rotation angle, and a probability of apical bifurcation (Niklas 1999), and Kendrick's crinoid cup model is based on the cup plate's height and angle (Kendrick 2007).

One should therefore be careful with the tools used for the quantitative exploration of affine morphospaces. The only meaningful inferences that can be drawn from such spaces are those that are invariant to affine transformations (Huttegger and Mitteroecker 2011), that is, those concerning aspects of the evolutionary pattern and narrative that rest on affine properties such as parallelism, barycentric constructions, and ratios of multidimensional volumes. For instance, it has been customary to illustrate Raup's morphospace occupation patterns by means of bivariate kernel density estimates, as in Fig. 7.1a (e.g., Raup 1967; Ward 1980; Saunders et al. 2004; Smith et al. 2014). Kernel density estimation requires a metric, however, and its use for the analysis of affine morphospaces is therefore inappropriate. There are two main options to circumvent these limitations. One is to recast the model within a

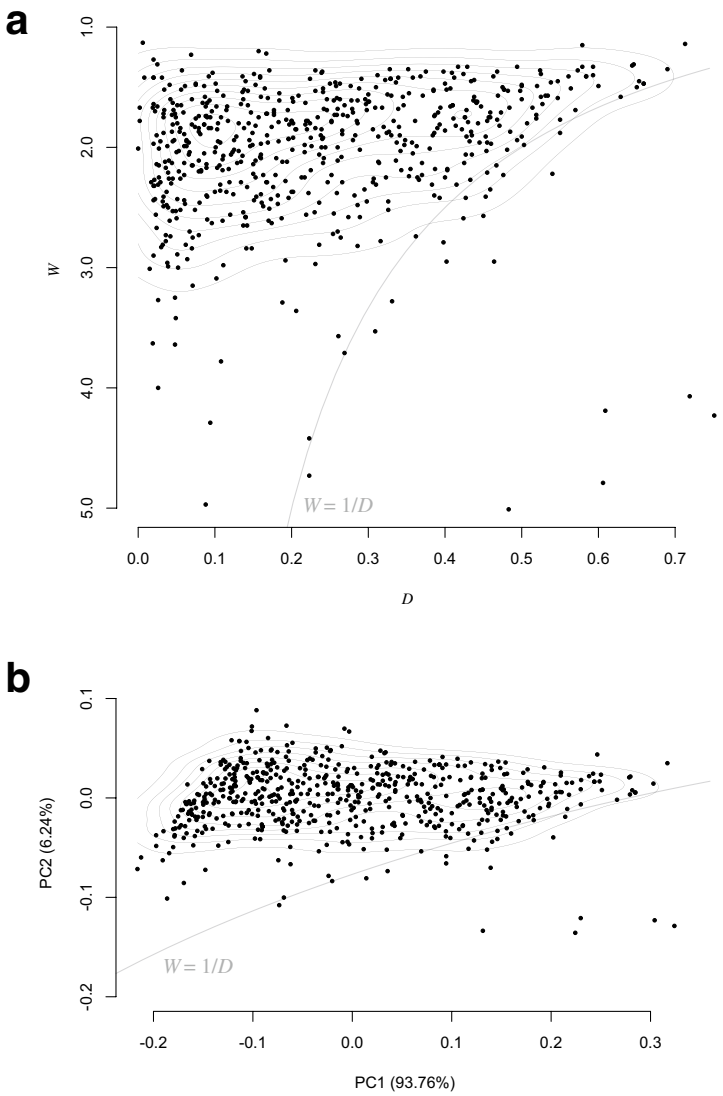


Fig. 7.1 (**a**) Distribution of 597 Paleozoic ammonoid genera in Raup's *D*-*W* morphospace (data from Saunders et al. 2004). The orthogonality of the axes and the aspect ratio follow Raup's (1967) original depiction of the space. (**b**) The same dataset displayed in Procrustes tangent shape space. Note that the visual perception of morphological similarity is different in the two morphospaces, but only the Procrustes tangent shape space possesses a metric structure that warrants a notion of morphological distance. (Modified from Gerber 2017)

morphometric framework that comes with a richer, metric geometry. Raup's model can, for instance, be fairly easily converted into a landmark-based scheme and expressed in the metric geometry of Procrustes shape spaces (Fig. 7.1b; Gerber 2017). Another option is to resort to affine invariant statistics (e.g., ratio of generalized variances as a measure of variation) and limit the interpretation of a morphological pattern to its affine invariant features (Huttegger and Mitteroecker 2011).

7.2.2 Multivariate Morphometrics and Phenomics

In multivariate morphometrics, a morphological structure is described by a set of p linear measurements. In the earliest contributions, it is stated that all these measurements ought to be of the same physical dimension and expressed in the same units (Mosimann 1970; Darroch and Mosimann 1985). This way, it is considered possible to promote a biologically driven choice of scales and angles among variables into a precise inner product (Dempster 1969). Hence, the morphospace is identified with Euclidean p-space (i.e., an affine space associated with a p-dimensional Euclidean vector space) and possesses meaningful notions of distance and direction. Unfortunately, less rigorous morphological analyses sometimes combine variables of different kinds, such as lengths, ratios, surface areas, or angles. While these analyses then tend to be carried out with the established statistical routines, the Euclidean structure of their morphospaces is no longer warranted. These analyses face the same geometric issues as those of theoretical morphology, and affine invariant analytical approaches should be preferred.

Paralleling the genomic revolution that occurred in molecular biology, there have been incentives to develop a comparable field of phenomics. Phenomics is the idea of acquiring high-dimensional phenotypic data on an organism-wide scale (Houle et al. 2010). Two types of phenotyping are distinguished: extensive and intensive phenotyping. Extensive phenotyping consists of the exhaustive description of phenomes with a large and diverse set of physiological, morphological, and behavioral phenotypic traits. With regard to the above discussion, extensive phenotyping poses important challenges. From the phenomic standpoint, a morphospace is a subspace of a phenome space. But even as a subset of the collection of phenotypic traits, the set of p morphological traits is bound to include a large diversity of incommensurable traits such as mensural, meristic, or categorical traits describing morphological attributes of different kinds. Such sets of morphological traits cannot be conducive to morphospaces exhibiting metric geometries. There have been important developments in statistics for the visualization of high-dimensional spaces, but many approaches (e.g., UMAP, t-SNE) require identifying the feature space with $\mathbb{R}^p$ and rely on Euclidean distance in the process of generating visualizations. Such a metric structure is difficult to conceive for incommensurate descriptors.

On the other hand, intensive phenotyping consists of a detailed focus on a particular phenotypic structure. In geometric morphometrics, the approach has been referred to as high-density morphometrics (e.g., Goswami et al. 2019) and follows

from the spectacular progress in imaging technologies and landmarking (semi)automation. Because high-density morphometrics essentially consists in increasing the density of the landmark coverage of the biological structure of interest, it has no impact on the metric geometry of the shape space except for its dimensionality. The associated increase in morphospace dimensionality can lead to the "curse of dimensionality" (when the number of dimensions exceeds the sample size; see Rohlf 2021 in the morphometric context), including issues of computational complexity, multicollinearity and unstable statistical estimates (calling for regularization procedures), and counterintuitive or deceptive visualizations. Some multivariate ordinations of high-dimensional morphospaces can, for instance, generate artefactual separations of morphologically similar groups (e.g., Cardini et al. 2019) and thereby misguide the description of a pattern based on its visual representation.

7.3 Discrete Character Morphospaces

Another common way that has been employed to construct morphospaces is to assign to each morphological variant a description of its morphology as a sequence of discrete character states. This approach is rooted in the field of numerical taxonomy (Sneath and Sokal 1973) and has since been widely used in macroevolutionary studies of morphological variation. As alluded to in the introduction, at high taxonomic levels, it can become difficult to find homologous morphological features common to all sampled taxa. Besides, many studies focus on articulated structures (e.g., skeletal elements) or include fragmentary materials (e.g., incompletely preserved fossil specimens) that cannot be handled by geometric morphometric methods. Discrete character morphospaces can handle these cases through presence/absence coding schemes and incorporate analytical steps and distance functions that can mitigate the effect of moderate amounts of missing data (e.g., Lehmann et al. 2019). Over the last two decades, these morphospaces have become increasingly popular through the co-option of discrete character datasets originally collated for phylogenetic purposes (Lloyd 2016; Gerber 2019). While discrete character morphospaces are often explored in the same manner as other morphospaces, their geometry is quite different and has implications for their representation (Gerber 2019).

In this discrete framework, each specimen is described by a p-tuple of integers corresponding to possible states for a set of binary and/or ordered/unordered multistate characters. Various distance functions have been advocated to accommodate different preferences or requirements (e.g., inapplicable character, unequal weighing, missing data, Euclidean likeness; Hopkins and St John 2018; Lehmann et al. 2019), but all are essentially based on the pattern of matches and mismatches among sequences. While the resulting distance matrix is sufficient for many aspects of morphospace exploration (e.g., the appraisal of morphological disparity), assessing patterns of morphospace occupation invites the visualization of the morphospaces constructed. The morphospace visualization is obtained from a multivariate ordination of the distance matrix. Most often, metric multidimensional scaling (principal

coordinate analysis) is used for this purpose. It involves the Euclidean representation of the distance matrix (Gower 1966). In the simple case of a set of p commensurate, unweighted binary characters, the morphospace is akin to a Hamming space and can be represented as a hypercube graph where the distance between two vertices equals the Hamming distance between the corresponding sequences of character states (Fig. 7.2). The metric multidimensional scaling of the square root of the Hamming distance matrix produces a Euclidean representation of the hypercube graph (Gerber 2019). The observed variants are located at, or "occupy," a subset of the vertices of the hypercube. This is similar to a principal component analysis of the covariance matrix of the observed variants' state sequences, considered as coordinates in Euclidean p-space. This multivariate ordination, however, can be prone to misinterpretation. As in other morphospaces, only observed variants (occupied vertices) are depicted and usually only the Cartesian coordinate system of the ordination is shown. However, the true morphospace, embedded in Euclidean space, is only a restricted and specific subset of $\mathbb{R}^p$. For instance, the absence of variants in any of the corners of the ordination plane in Fig. 7.2a does not call for biological explanations. These regions of the ordination space are geometric impossibilities within the discrete character scheme. Displaying edges connecting neighboring vertices can help delineate the spread of the morphospace in the ordination space (Fig. 7.2b). The subset of edges displayed also highlights that our perception of the morphological similarity of variants can be severely misguided. In the present case, all edges are of unit length, and they are either parallel or orthogonal to one another. Faithful and informative depictions of hypercube-like structures are difficult to produce (Abramson and Zanette 2003; Horiike and Fujishiro 2025) and can be further affected by suboptimal coding strategies and missing data. Missing data can introduce inconsistencies leading to overall estimates of match/mismatch that do not satisfy all the axioms of a distance function.

Another frequent issue with the exploration of discrete character morphospaces is their incorrect identification with the ordination spaces in which they are embedded (Gerber 2019). This often leads to the use of inappropriate methodologies relying on the Euclidean geometry of the ordination space and overlooking the discrete nature of the morphospace (e.g., the use of squared-change parsimony to map a phylogenetic hypothesis onto the morphospace).

7.4 Developmental Considerations

7.4.1 Morphological Distance and Accessibility

The biological structures studied and portrayed in a morphospace are not static structures. Their transformation from one variant to another in a lineage does not occur at the level of their morphological description. Even though shape changes are often visualized as deformation grids in morphometrics (Klingenberg 2013), an evolutionary change in shape does not proceed by homeomorphism. Organs and

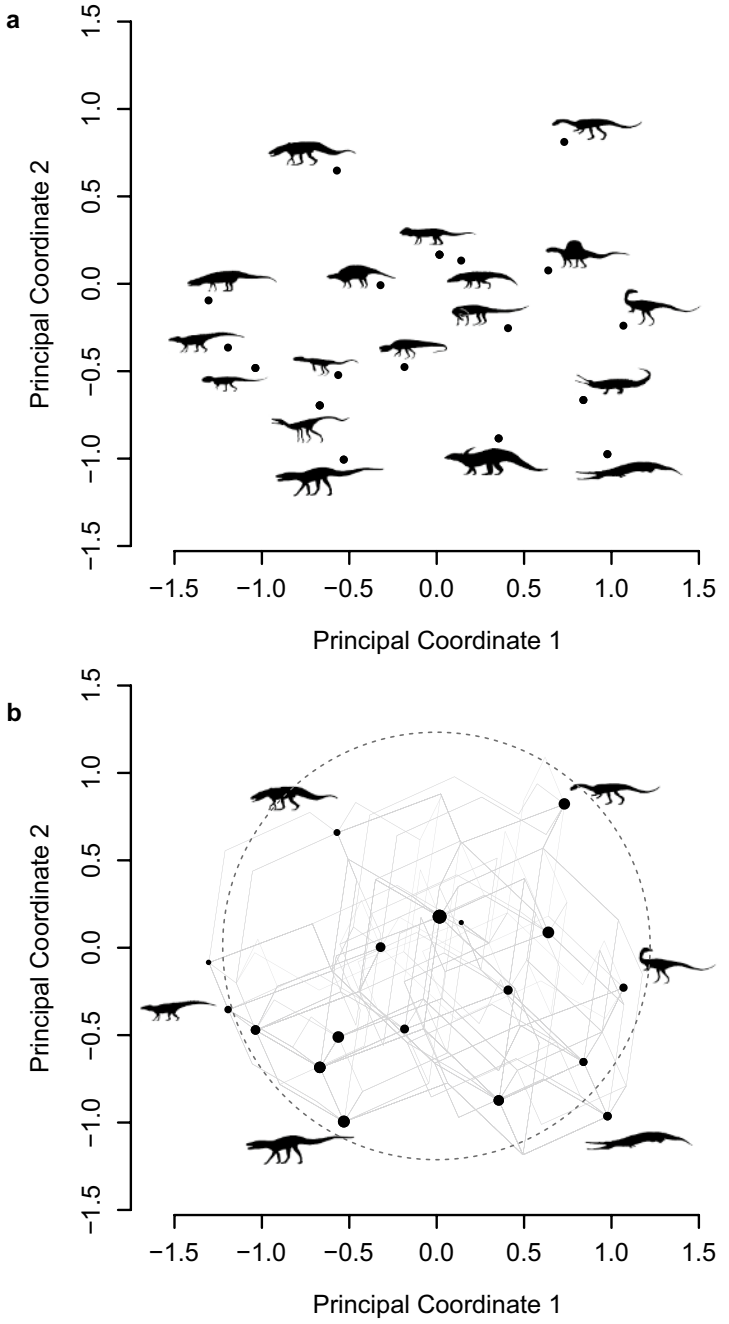


Fig. 7.2 (**a**) A discrete character space of archosaur taxa based on ten binary characters (subset of data from Brusatte et al. 2008). Its ordination is obtained by principal coordinate analysis of the morphological (Hamming) distance matrix. (**b**) The same space and ordination, but the discrete architecture of the morphospace is made visible by depicting some of the hypercube's edges. The dashed circle circumscribes 95% of the 210 vertices. The size of the vertices represented is proportional to their position along the third principal coordinate axis orthogonal to the plane of the picture. (Modified from Gerber 2019)

body parts are built anew at each generation, for each individual, and their evolutionary transformation implies changes in their underlying genetic and/or epigenetic systems. These aspects are important to keep in mind if one views morphospaces as useful constructs to improve our understanding of the evolutionary process. It appears from the above that, historically, distance functions have been favored to distinguish and compare morphological variants. This was logically motivated by the aim to capture morphological similarity and organize variants in morphospace accordingly. This is the reason why a metric structure is almost always defined or assumed for morphospaces. While this is a reasonable choice for many questions in biology, it is not the most appropriate approach to allow causal interpretations of morphological forms and evolutionary patterns.

Figure 7.3 illustrates the limits of metric topologies. The example is based on the empirical study of the evolutionary transition between two Cambrian trilobite species from a morphometric analysis of the shape of their cranidia (the central part of the trilobite head). From the comparative analysis of the developmental trajectories

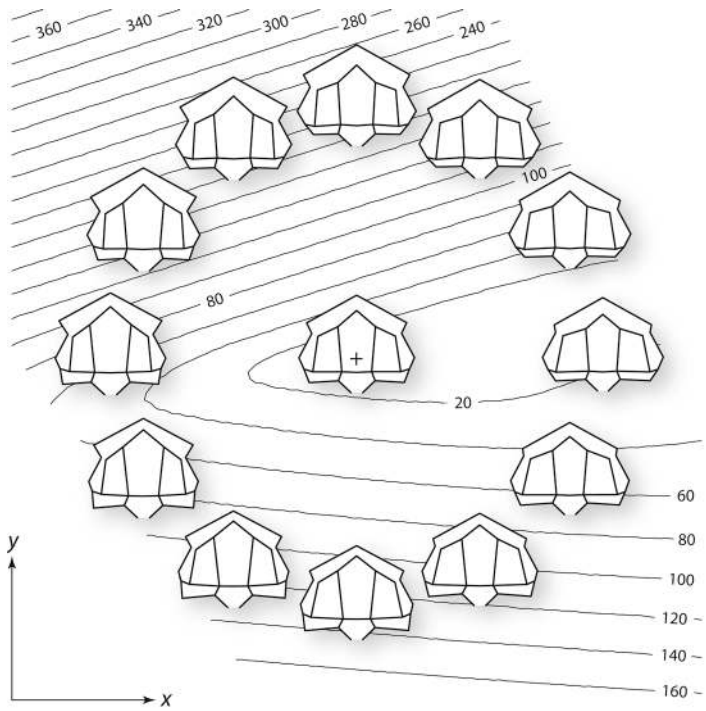


Fig. 7.3 Accessibility pattern mapped onto a metric morphospace (Procrustes shape space). The degree of evolutionary adjacency of morphological variants in the vicinity of the ancestral form is depicted by isolines of accessibility. It highlights the discordance between accessibility and morphological distance. Accessibility of variants (i.e., morphological variability) is markedly anisotropic. Isotropic accessibility would translate as circular isolines of various radii around the ancestral form. (Modified from Gerber 2014)

of the two species, the evolutionary change in development underlying morphological differences has been ascribed to mosaic heterochrony (Gerber and Hopkins 2011). Mosaic heterochrony involves shifts in the rate and/or timing of development along conserved trajectories of shape change for each of the parts (modules) constituting the overall morphological structure. Assuming that this evolutionary mode reflects the variational properties of the cranidium in this lineage, it is possible to determine the alternative morphological options that were equally available from the ancestral condition (ignoring the effect of external selection). Figure 7.3 shows the accessibility of possible shapes in the vicinity of the ancestor by means of lines of equal accessibility. It appears that not all equally accessible variants are equally distant from the ancestral shape. The evolutionary accessibility of variants or, equivalently, the variability of the ancestral morphological variant, is highly anisotropic (Gerber 2014). As a consequence, the evolutionary navigability of morphospace strongly departs from the isotropic expectation of a metric topology. This has significant consequences for the reading of morphospace patterns.

7.4.2 *Morphological Variant Representation and Accessibility Structure*

Since morphological phenotypes do not interconvert at the morphological level at which they are described, is there a more appropriate level of description? This invites us to reconsider the two defining features of mathematical spaces for biological morphospaces: How do we define the set of variants? And what appropriate topology or mathematical structure is relevant to endow this set of variants?

In the standard morphospace approaches presented above, the theoretical set of possible variants is imposed by the morphometric framework. In shape analysis of landmark data, it is the set of shapes of p not totally coincident, labeled landmarks in $k = 2$ or 3 dimensions:

$$\left\{\mathbb{R}^{p\times k} \setminus \mathcal{C}\right\} / \mathrm{Sim}(k),$$

where $\mathcal{C}$ denotes the coincidence set and Sim(k) is the group of Euclidean similarity transformations in k dimensions.

In multivariate morphometrics, it is the subset of the nonnegative orthant in $\mathbb{R}^p$ (the set of p-tuples of positive real values) defined as:

$$\left\{x \in \mathbb{R}_+^p \mid g(x) = 1\right\},$$

where the function $g : \mathbb{R}_+^p \to \mathbb{R}_+$ is a size measure (a positively homogeneous function of degree 1; Mosimann 1970).

In the discrete character scheme and for the specific case of the hypercube discussed, it is the set of all 2^p binary sequences of length p:

$$\left\{\left(s_1,\cdots,s_p\right)\middle|\, s_i \in \{0,1\}\right\}.$$

In Raup's model, it is the Cartesian product of the sets of possible values for the parameters *W*, *T*, and *D*:

$$(1,+\infty)\times[0,+\infty)\times[0,1).$$

Clearly, not all the options in a given framework make biological sense. Many of them specify morphological configurations that are biologically unrealistic or impossible. They correspond to regions of the morphospace that are both developmentally and evolutionarily inaccessible and they have therefore no explanatory value.

A radical but fruitful alternative is to consider a generative, genetic and/or epigenetic, representation of the morphological structure (e.g., Alberch 1989; Wagner and Altenberg 1996), that is, to identify a morphological variant with the particular developmental context(s) leading to its expression. Ideally, if the developmental system underlying the morphological structure is well characterized, then it also specifies its variational properties. Hence, from the morphospace perspective, the developmental representation defines the set of possible variants and determines their accessibility relations.

In practice, the term development hides a complex network of causal processes that mediate the way genetic variation is translated into phenotypic variation. It includes the dynamics of gene regulatory networks controlling patterns of gene expression, and the inductive and feedback effects of gene products on gene regulation, cell behavior, and cell–cell interactions underlying pattern formation and morphogenesis (e.g., Salazar-Ciudad 2021). Except maybe for very simple morphological structures or substructures, it appears difficult, and maybe out of reach, to obtain a detailed picture of the intricacies of the developmental structure of phenotypic traits. Development is therefore often mathematically abstracted as a Genotype–Phenotype map in theoretical work where the global structure and statistical properties of the map are instead emphasized (e.g., pleiotropy of gene effects, epistatic interaction patterns). Likewise, in many studies of pattern formation and models of morphogenesis, the aim is not necessarily to precisely identify, say, the diffusible molecule or morphogen involved, but to improve our understanding of the fundamental principles underlying self-organization in development (Nijhout 2003). Such models therefore are valuable to emulate developmental mechanisms and provide a basis for proposing developmental representations of morphological traits. The phenotype landscape framework (Rice 1998, 2004) offers a practical way to implement this idea (Gerber 2025).

A phenotype landscape is the expression of the possible variants that a phenotype can manifest at a given life stage as a function *f* of some underlying causal factors:

$$f : U \to M,$$

where M is the set of possible variants and U is the space of underlying factors or parameter space. This parameter space can include a variety of heritable or non-heritable causal processes, such as gene product concentrations and diffusion rates, cell behavior (e.g., division rate), mechanical properties of tissues (e.g., elasticity), and environmental effects (e.g., temperature). The robustness of developmental processes to genetic or environmental perturbations is expressed by the surjectivity of the developmental function f. The function thus defines an equivalence relation on U, with morphological variants as equivalence classes of tuples of values or states for the underlying factors. A variant m_i is developmentally represented in the parameter space by its preimage under f:

$$f^{-1}\left[m_i\right] = \left\{u \in U \mid f\left(u\right) = m_i\right\}.$$

The set of morphological variants is canonically identified with the set of all equivalence classes in U:

$$U / \sim_f .$$

Accessibility among variants is then estimated from the neighborhood relations of their preimages in the parameter space. Various measures have been proposed depending on the topological properties of the parameter space (e.g., Alberch 1989, Gerber 2025; Stadler et al. 2001). If the parameter space is akin to a n-dimensional manifold, the accessibility of the variant m_j from the variant m_i can be expressed as:

$$A\left(\left[m_j\right] \curvearrowleft \left[m_i\right]\right) = \frac{V_{n-1}\left(\partial f^{-1}\left[m_j\right] \cap \partial f^{-1}\left[m_i\right]\right)}{V_{n-1}\left(\partial f^{-1}\left[m_i\right]\right)}.$$

Importantly, a measure of accessibility, which represents the probability of transition from one variant to another, is not necessarily symmetric (Fontana 2002). The notion of accessibility is thus conceptually distinct from the notion of distance. The mathematical structure associated with the notion of accessibility is a pretopology (Stadler et al. 2001).

7.4.3 Morphospaces as Directed Graphs

Given this pretopological structure, a more appropriate representation of a morphospace is as a directed graph with directed edges informing the likelihood of evolutionary transitions and vertices representing developmentally possible variants. The graph representation of the morphospace captures the variability of the morphological trait considered. These variational properties can induce evolutionary constraints in the sense of opposing or modulating the effect of external selective pressures acting on the phenotype at the life stage considered (Schwenk and Wagner 2004).

The graph can be mapped onto the standard metric representation of the morphospace by assigning to the vertices the coordinates of their morphometric characterization (see the example below). Morphological variation is the subset of nodes that correspond to observed variants and it can be estimated with classical statistical measures of dispersion in the metric representation. The distinction between variation and variability, allowed by the generative representation of the morphological trait, is critical for the causal explanation of morphological patterns in terms of selection and constraint (Wagner and Altenberg 1996).

7.4.4 Pattern Formation and Morphogenetic Models

Many pattern formation and morphogenetic models can in principle lead to a developmental representation of the trait they model. For instance, Gerber (2025) explored Nijhout's (1991) model of the prepatterning of butterfly wing eyespots based on a reaction–diffusion system. The results showed that the accessibility structure induced by the dynamics of activator–inhibitor interactions restricts the set of possible positions for the eyespot within the wing cell compartment and that small changes in developmental parameter values can lead to discontinuous (w.r.t. the metric expectation) changes in morphology.

We now provide another example based on the computational model of tooth morphogenesis developed by Salazar-Ciudad and Jernvall (2010) and implemented in ToothMaker. Their model is based on 19 parameters regulating both gene network properties and cellular properties (tissue mechanics).

A sample of ten virtual teeth of the ringed seal (*Phoca hispida*) along with their associated tuples of 19 parameter values provided in Salazar-Ciudad and Jernvall (2010) serve as a description of the local topography of the phenotype landscape. For the sake of illustration and ease of computation, the parameter space is reduced by principal component analysis to a two-dimensional space accounting for 75% of the sample variation in parameter values (Fig. 7.4b). Tooth shape is captured by a configuration of nine two-dimensional landmarks placed at the tips and bases of the four tooth cusps (Fig. 7.4a). The ten landmark configurations are superimposed by partial generalized Procrustes analysis and orthogonally projected onto the tangent space (Dryden and Mardia 1998). The a priori set of morphological variants considered consists of the set of shapes lying in the plane defined by the first two principal components of the tangent space and which accounts for 80% of the sample shape variation. The developmental function f that maps parameter value tuples to shapes is obtained from multivariate quartic regression, which approximates the landscape's local topography (Fig. 7.4b). The set of variants is further discretized into a square lattice of 81 shapes covering the range of empirical shape variation.

Accessibility among variants is based on the relative sizes of their preimages under f (surface areas in parameter space). It is calculated using a modified version of the A_1 measure (Gerber 2025) to account for the multidimensionality of the phenotype and allow a nonzero probability of self-transition. As a ratio of surface areas,

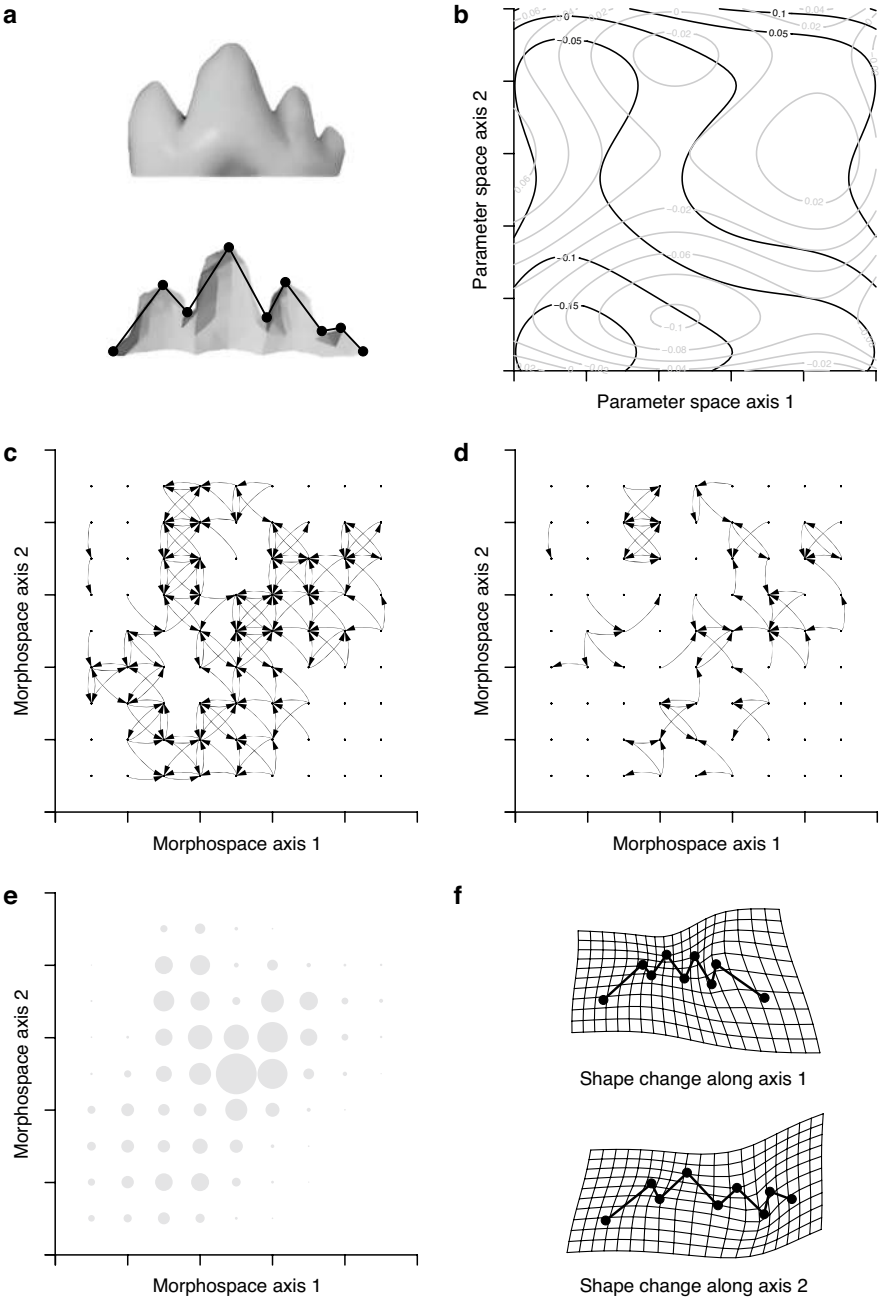


Fig. 7.4 (**a**) An example of a real ringed seal tooth and its simulation with ToothMaker. The shape of the simulated tooth is captured by nine landmarks. (**b**) The local geometry of the tooth phenotype landscape visualized as contour lines of equal morphological values for the two dimensions of the morphospace. (**c**) The accessibility structure displayed as a directed graph mapped onto the metric geometry of the morphospace, with an accessibility threshold value of 0.12. (**d**) The same with an accessibility threshold value of 0.20. (**e**) The expected null model of morphospace occupation. The size of the circles is proportional to the expected frequency of variants. (**f**) The shape changes associated with the two axes of the morphospace (at 0.2 units of Procrustes distance from the mean shape). See text for details

this measure has the advantage of being invariant to affine transformations. This is a useful property given the disparate nature of the underlying parameters, which precludes the definition of a metric topology for the parameter space.

Figure 7.4c, d shows the resulting accessibility structure as a directed graph. Vertices correspond to the 81 a priori morphological variants and they are placed at their respective locations in the morphospace. Directed edges indicate the accessibility relations among variants. Different threshold values have been set to decide whether or not a directed edge should be drawn. This facilitates the visualization of the accessibility structure but it also points to cases where some evolutionary transitions can be deemed unlikely or impossible because their associated probabilities are too small.

The graphs depicted in Fig. 7.4 show that morphological variability can be highly anisotropic (differential accessibility of neighboring variants) and heterogeneous (variant-specific accessibility pattern). Accessibility also appears often to be asymmetric, indicating unequal probabilities of transition between a given pair of variants.

The approach also allows assessing the likely frequencies of occurrence of variants (in the absence of external selection) through the limiting distribution of the Markov chain associated with the directed graph representation (Gerber 2025). This is a basis for the expected null model of morphospace occupation, prior to the action of selection at the chosen focal life stage, driven by the properties of the tooth developmental system (Fig. 7.4e). It shows substantial spatial heterogeneity in the frequencies of developmentally possible variants and, as discussed above, points to the fact that many morphometrically conceivable variants are biologically irrelevant. Hence, the fact that the bottom-right corner of the morphospace is "unoccupied" does not call for any biological explanations. This is an artefact of the metric structure of the morphometric scheme. Identifying the set of developmentally possible variants is therefore an important step toward refined interpretations of evolutionary patterns of morphological variation.

7.5 Conclusions

The concept of morphospace has proved very useful for framing the discussion and analysis of the patterns and processes of morphological evolution. Practically, morphospaces come in a variety of types reflecting the diversity of morphological phenotypes and the diversity of morphometric schemes. The expectation of a unified analytical framework for morphospace exploration comes up against the diversity of their geometries. Not all exploratory tools can be used in all morphospaces and not all types of inference can be made from morphospace patterns. The definition of the set of morphological variants and of its organizing structure should be guided by the questions one intends to address. Metric topologies have traditionally been favored, and they do provide appropriate morphospace structures for many biological questions. Organizing variants only according to their morphological similarity, however, overlooks the nature of morphological traits and their variational properties. Devising morphospace structures that reflect the genetic-, cell-, tissue-, and

organ-level processes that are causally involved in the emergence of morphological traits also appear fundamental to address key questions in evolution. Comprehensive descriptions of developmental systems are presently out of reach, but many pattern formation and morphogenetic models proposed for plants and animals open avenues for the construction of developmentally informed morphospaces.

Acknowledgments I thank Tomohiko Yamaguchi and all the committee members for the organization of the PFN 2023 "Diversity of Biological Patterns and Forms in Nature: Toward a Comprehensive Understanding." I also thank Toshio Sekimura, Shigeru Kondo, Takao Ubukata, Philip Ball, and Oliver Steinbock for discussions and Philip Maini, Yoland Savriama, Zuzana Gerber, and Maxime Varey for their critical reading of the manuscript.

References

Abramson G, Zanette DH (2003) Two-dimensional projections of a hypercube. Phys Rev E 67:057101

Adams DC, Rohlf FJ, Slice DE (2004) Geometric morphometrics: ten years of progress following the 'revolution'. Italian J Zool 71(1):5–16

Alberch P (1982) Developmental constraints in evolutionary processes. In: Bonner JT (ed) Evolution and development. Dahlem Konferenzen, Springer-Verlag, Berlin, pp 313–332

Alberch P (1989) The logic of monsters: evidence for internal constraint in development and evolution. Geobios 22:21–57

Brusatte SL, Benton MJ, Ruta M, Lloyd GT (2008) Superiority, competition, and opportunism in the evolutionary radiation of dinosaurs. Science 321:1485–1488

Cardini A, O'Higgins P, Rohlf FJ (2019) Seeing Distinct Groups Where There are None: Spurious Patterns from Between-Group PCA Evolutionary Biology 46(4):303–316 https://doi.org/10.1007/s11692-019-09487-5

Collyer M, Adams D (2013) Phenotypic trajectory analysis: comparison of shape change patterns in evolution and ecology. Hystrix 24(1):75–83

Darroch JN, Mosimann JE (1985) Canonical and principal components of shape. Biometrika 72(2):241–252

Dempster AP (1969) Elements of continuous multivariate analysis. Addison-Wesley

Dryden IL, Mardia KV (1998) Statistical shape analysis. Wiley

Eble GJ (2000) Theoretical morphology: state of the art. Paleobiology 26(3):520–528

Erwin DH (2007) Disparity: morphological pattern and developmental context. Palaeontology 50(1):57–73

Foote M (1997) The evolution of morphological diversity. Annu Rev Ecol Syst 28(1):129–152

Fontana W (2002) Modelling 'evo-devo' with RNA. BioEssays 24(12):1164–1177

Gerber S (2014) Not all roads can be taken: development induces anisotropic accessibility in morphospace. Evol Dev 16(6):373–381

Gerber S (2017) The geometry of morphospaces: lessons from the classic Raup shell coiling model. Biol Rev 92:1142–1155

Gerber S (2019) Use and misuse of discrete character data for morphospace and disparity analyses. Palaeontology 62(2):305–319

Gerber S (2025) Modeling evolutionary accessibility in phenotype space. Evol Biol 52:107–119

Gerber S, Hopkins MJ (2011) Mosaic heterochrony and evolutionary modularity: the trilobite genus *Zacanthopsis* as a case study. Evolution 65(11):3241–3252

Goswami A, Watanabe A, Felice RN, Bardua C, Fabre AC, Polly PD (2019) High-density morphometric analysis of shape and integration: the good, the bad, and the not-really-a-problem. Integr Comp Biol 59(3):669–683

Gould SJ (1991) The disparity of the Burgess Shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. Paleobiology 17(4):411–423
Gould SJ (2002) The structure of evolutionary theory. Harvard University Press
Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53(3–4):325–338
Hopkins MJ (2016) Magnitude versus direction of change and the contribution of macroevolutionary trends to morphological disparity. Biol J Linn Soc 118(1):116–130
Hopkins MJ, Gerber S (2021) Morphological disparity. In: de la Rosa N, Muller GB (eds) Evolutionary developmental biology: a reference guide. Springer International Publishing, Cham, pp 965–976
Hopkins MJ, St John K (2018) A new family of dissimilarity metrics for discrete character matrices that include inapplicable characters and its importance for disparity studies. Proc R Soc B 285(1892):20181784
Horiike Y, Fujishiro S (2025) Orthogonal projections of hypercubes. Phys Rev E 112(4):045304
Houle D, Govindaraju DR, Omholt S (2010) Phenomics: the next challenge. Nat Rev Genet 11(12):855–866
Hunt G (2007) The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. Proc Natl Acad Sci 104(47):18404–18408
Huttegger SM, Mitteroecker P (2011) Invariance and meaningfulness in phenotype spaces. Evol Biol 38(3):335–351
Kendall DG, Barden D, Carne TK, Le H (1999) Shape and shape theory. Wiley
Kendrick DC (2007) Theoretical morphology of the crinoid cup. Paleobiology 33(2):337–350
Klingenberg CP (2013) Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. Hystrix 24(1):15
Lehmann OE, Ezcurra MD, Butler RJ, Lloyd GT (2019) Biases with the Generalized Euclidean Distance measure in disparity analyses with high levels of missing data. Palaeontology 62(5):837–849
Lloyd GT (2016) Estimating morphological diversity and tempo with discrete character-taxon matrices: implementation, challenges, progress, and future directions. Biol J Linn Soc 118(1):131–151
Maynard Smith J, Burian R, Kauffman S, Alberch P, Campbell J, Goodwin B, Lande R, Raup D, Wolpert L (1985) Developmental constraints and evolution: a perspective from the Mountain Lake conference on development and evolution. Q Rev Biol 60(3):265–287
McGhee GR Jr (1999) Theoretical morphology: the concept and its applications. Columbia University Press, New York
Mitteroecker P, Huttegger SM (2009) The concept of morphospaces in evolutionary and developmental biology: mathematics and metaphors. Biol Theory 4:54–67
Mitteroecker P, Schaefer K (2022) Thirty years of geometric morphometrics: achievements, challenges, and the ongoing quest for biological meaningfulness. Am J Biol Anthropol 178:181–210
Moseley H (1838) XVII. On the geometrical forms of turbinated and discoid shells. Philos Trans R Soc Lond 128:351–370
Mosimann JE (1970) Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. J Am Stat Assoc 65(330):930–945
Niklas KJ (1999) Evolutionary walks through a land plant morphospace. J Exp Bot 50(330):39–52
Nijhout HF (1991) The development and evolution of butterfly wing patterns. Smithsonian Institution
Nijhout HF (2003) Gradients, diffusion and genes in pattern formation. In: Muller GB, Newman SA (eds) Origination of organismal form: beyond the gene in developmental and evolutionary biology. MIT Press, pp 165–181
Raup DM (1966) Geometric analysis of shell coiling: general problems. J Paleontol:1178–1190
Raup DM (1967) Geometric analysis of shell coiling: coiling in ammonoids. J Paleontol:43–65
Rice SH (1998) The evolution of canalization and the breaking of von Baer's laws: modeling the evolution of development with epistasis. Evolution 52(3):647–656
Rice SH (2004) Evolutionary theory: mathematical and conceptual foundations. Sinauer

Rohlf FJ (2021) Why clusters and other patterns can seem to be found in analyses of high-dimensional data. Evol Biol 48(1):1–16

Salazar-Ciudad I (2021) Mechanisms of pattern formation, morphogenesis, and evolution. In: de la Rosa N, Muller GB (eds) Evolutionary developmental biology: a reference guide. Springer International Publishing, Cham, pp 555–570

Salazar-Ciudad I, Jernvall J (2010) A computational model of teeth and the developmental origins of morphological variation. Nature 464(7288):583–586

Saunders WB, Work DM,Nikolaeva SV (2004) The evolutionary history of shell geometry in Palaeozoic ammonoids. Paleobiology 30(1):19–43

Smith PL, Longridge LM, Grey M, Zhang J, Liang B (2014) From near extinction to recovery: late Triassic to middle Jurassic ammonoid shell geometry. Lethaia 47:337–351

Sneath PH, Sokal RR (1973) Numerical taxonomy. W.H. Freeman & Co

Stadler BMR, Stadler PF, Wagner GP, Fontana W (2001) The topology of the possible: formal spaces underlying patterns of evolutionary change. J Theor Biol 213(2):241–274. https://doi.org/10.1006/jtbi.2001.2423

Stayton CT (2015) The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. Evolution 69(8):2140–2153

Schwenk K, Wagner GP (2004) The relativism of constraints on phenotypic evolution. In: Pigliucci M, Preston K (eds) Phenotypic integration: studying the ecology and evolution of complex phenotypes. Oxford University Press, pp 390–408

Thompson D'AW (1942) On growth and form. Cambridge University Press, Cambridge

Urdy S (2015) Theoretical modelling of the molluscan shell: what has been learned from the comparison among molluscan taxa? In: Klug C, Korn D, De Baets K, Kruta I, Mapes R (eds) Ammonoid paleobiology: from anatomy to ecology. Topics in geobiology, vol 43. Springer, Dordrecht

Wagner GP (2014) Homology, genes, and evolutionary innovation. Princeton University Press

Wagner GP, Altenberg L (1996) Perspective: complex adaptations and the evolution of evolvability. Evolution 50(3):967–976

Wake DB, Larson A (1987) Multidimensional analysis of an evolving lineage. Science 238(4823):42–48

Ward P (1980) Comparative shell shape distributions in Jurassic-Cretaceous ammonites and Jurassic-Tertiary nautilids. Paleobiology 6(1):32–43. https://doi.org/10.1017/S0094837300012471

Wills MA (2001) Morphological disparity: a primer. In: Adrain JM, Edgecombe GD, Lieberman BS (eds) Fossils, phylogeny, and form: an analytical approach. Plenum, pp 55–144

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 8
Functional Morphospace Analysis of Molluscan Shells

Takao Ubukata

Abstract A functional morphospace visualizes the form–function relationship as a contour surface of functional performance across shape variation. It enables predictions about which forms should be common or rare from a functional morphology, and to test the prediction through biometric analyses. Over the past several decades, advances in mathematical modeling, geometric morphometrics, and continuum biomechanics have substantially improved functional morphospace approaches. In particular, molluscan shell form has been a popular target for functional morphospace analysis in paleobiology. Many studies using functional morphospaces have focused on the efficiency of shell construction in ectoconchleate mollusks, hydrodynamic properties in cephalopods, postural stability in gastropods, umbonal conflict in bivalves, and so on. Collectively, such studies have sought to evaluate the role of functional properties in shaping the macroevolutionary trajectories of molluscan groups. In this chapter, I review the studies on shell morphology using functional morphospaces.

Keywords Biomechanics · Functional morphology · Molluscan shells · Morphospace · Theoretical morphology

8.1 Introduction

Since studies on the evolution of ancient life highly depend on fossil morphology, paleontologists have a long tradition of studying the relationship between the structure and function of organisms, or functional morphology (Savazzi 1999; Foote and Miller 2007). Several categories of methods and approaches are included under the umbrella of functional morphology. A common approach adopted in modern paleobiology is biomechanics, in which structure and function are studied from the

T. Ubukata (✉)
Department of Geology and Mineralogy, Kyoto University, Kitashirakawa Oiwake-cho, Kyoto, Japan
e-mail: ubukata@kueps.kyoto-u.ac.jp

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_8

mechanical point of view (Rayfield 2007; Marcé-Nogué 2020; Bishop et al. 2021). Such an approach typically focuses on an observed structure to assess how it works and how efficient it is for the function of interest (e.g., Shiino et al. 2009; Lemanis et al. 2016; Bicknell et al. 2018). An alternative approach, which has been conventionally used in paleontology, is the "paradigm method," which is based on the comparison between an observed structure and the ideal design for the supposed function (Rudwick 1961, 2018; Paul 1999). The "paradigm method" has been used to evaluate how the observed structure is optimized for the function (e.g., Tayler 1987; Oji and Okamoto 1994). In addition to the cases above, we may be interested in what kind of structures and how they are efficient for a supposed function rather than in assessing the functional properties of a particular structure. To answer this question, a large-scale comparison of morphology is needed across several taxa. Morphospace is a powerful tool for studying the relationship between structure and function at the macroevolutionary scale.

Morphospaces are geometric spaces in which a particular organic form can be plotted with reference axes chosen to represent morphological parameters (Savazzi 1999; Hammer and Haper 2024). Abstraction of key morphological traits allows complex shape variation to be summarized within a low-dimensional morphospace. When functional performance can be quantified for the shape represented in that space, the relationship between functional performance and morphological traits can be visualized as a surface on the morphospace representing a fitness parameter that mimics an "adaptive landscape" (McGhee 1999, 2007; Swan 2001). Since adaptive forms are expected to be common in the evolutionary history, the functional morphospace allows us to predict which shape is common or rare from the viewpoint of functional morphology. Such a prediction can be readily tested by paleontologists by referring to the actual fossil records. A particularly useful tool for the analysis of the distribution of the actual forms in a morphospace is a mathematical model approximating the shape of the organisms.

Raup (1966) introduced a morphospace constructed on the basis of the parameters of a mathematical model: this type of morphospace is called a theoretical morphospace (McGhee 1999). Theoretical morphospaces have been used to document the range of observed forms of organisms as a subset of forms that seem theoretically possible but may be nonexistent (Hickman 1993). Evolutionary morphology using theoretical morphospaces, called theoretical morphology (Raup 1966), has been applied to various living and fossil groups, including foraminifera (Tyszka 2006), plants (Niklas 1997), corals (Ohno et al. 2015), bryozoans (McKinney and Raup 1982), and brachiopods (McGhee 1980). In particular, molluscan shell forms have been popular targets for theoretical morphology because of their geometric regularity and the availability of the richest fossil records among animals. Mollusks show great variation in shell form (Fig. 8.1); however, the manner of shell coiling matches a logarithmic helicospiral in many cases, and therefore, their shells are suitable for functional morphospace analysis. This chapter reviews the studies on shell morphology using functional morphospaces in the relevant fields of paleobiology, focusing on the importance of the function in the macroevolution of a molluscan group, rather than on the shell shape of the species.

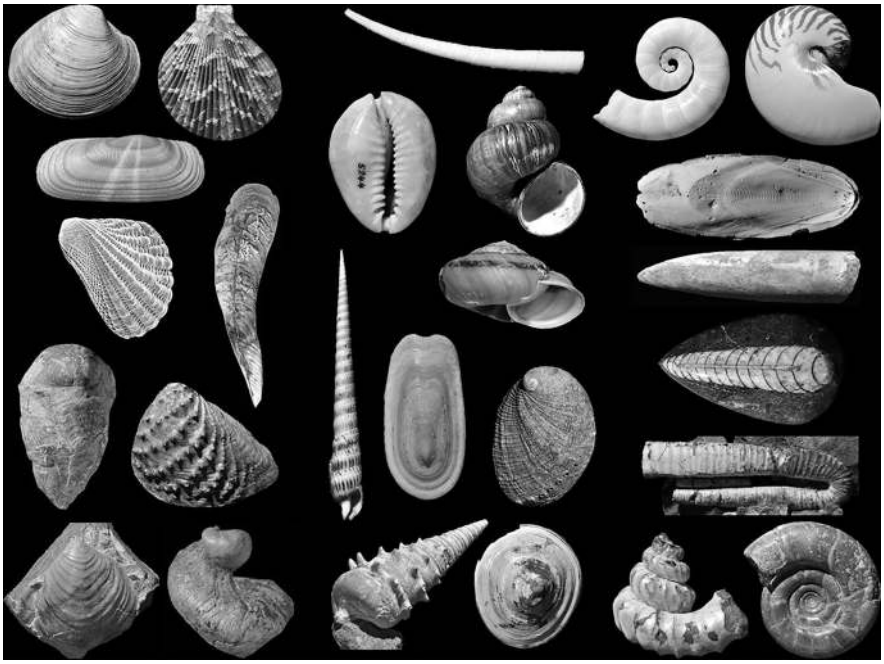


Fig. 8.1 Photographs of the shells of living and fossil molluscan species including bivalves, scaphopods, gastropods, and cephalopods, showing a wide range of morphological variation

8.2 Morphospace Analyses of Cephalopod Shells

A pioneering example of functional morphospace analysis is Raup (1967). Raup (1966) proposed a logarithmic helicospiral model describing the geometry of coiled shells (Raup 1962; Raup and Michelson 1965). The model generates shell forms as the growth trajectory of a circular- or elliptic-generating curve, which approximates the shape of the whorl cross-section. In planispiral forms such as ammonoid shells, shell growth is represented by the expansion and rotation of the generating curve around the coiling axis (Fig. 8.2A). In this case, the coiling geometry is defined by the whorl expansion rate (W) and the position of the generating curve in relation to the coiling axis (D). Figure 8.2B shows the spectrum of the computer-produced shell forms with various values of W and D, or a W–D morphospace. When W is large, the whorl radius rapidly increases with growth. When D is large, the shell has a wide umbilicus, which is a depressed central area around the coiling axis. Raup (1967) estimated the values of the parameters from more than 400 ammonoid genera and found that ammonoids are limited to a small region in the morphospace. The most common shell shapes among ammonoids were observed in the region around $W = 2.0$ and $0.3 \leq D \leq 0.4$ in the morphospace.

Raup (1967) listed several possible reasons for the biased distribution of ammonoid shell forms in the morphospace from the functional viewpoint, such as

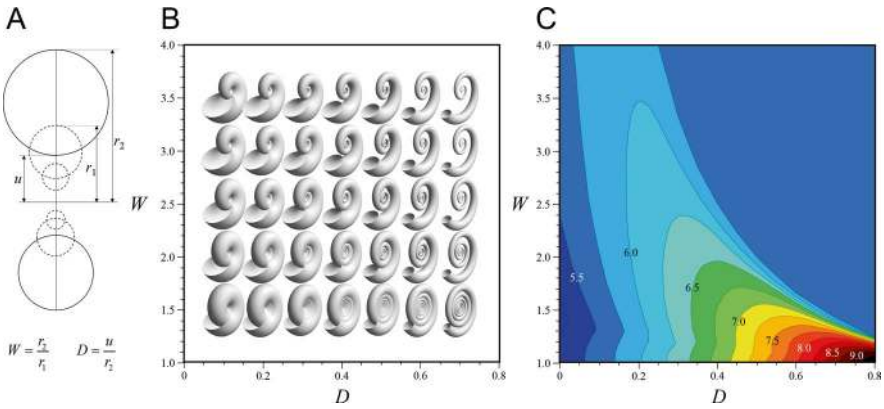


Fig. 8.2 Theoretical and functional morphospaces based on the Raup (1967) model. (**A**) Rationale for the model designed for ammonoid shell forms with definitions of the parameters. (**B**) Spectrum of the computer-generated shell shapes with a circular-generating curve and various values of *W* and *D* when the value of each parameter is fixed throughout growth. (**C**) Three-dimensional contour diagram showing the relationship between parameters (*W* and *D*) and construction efficiency when the shell thickness is 0.0772 of the radial length of the circular-generating curve

hydrostatic stability and efficiency of shell construction. He assessed the efficiency of shell construction for a given shell shape by the ratio of the volume of the internal shell space to the volume of shell material required to enclose the space, based on Raup and Chamberlain (1967). Figure 8.2C shows a functional morphospace depicting a contour plot of the construction efficiency assuming that the shell thickness is 0.0772 of the radial length of the circular-generating curve. A shell shape with small *W* and large *D* values is most efficient in terms of the necessary amount of calcium carbonate (Fig. 8.2C) but is rather rare in nature. Raup (1967) concluded that the efficiency of shell construction is not the only factor constraining the morphological diversity of ammonoid shell geometry. Raup (1967) also computed hydrostatic stability in relation to *W* and *D* as the relative distance from the center of the volume to the center of the mass with respect to the diameter of the shell. He found that hydrostatic stability is of little importance for determining the observed distribution of ammonoid shell shape.

In Raup (1967), the relative shell thickness is defined in relation to the radius of the generating curve. In this case, ammonoids with a wider umbilicus (larger *D*) inevitably have a thinner shell when the relative shell thickness is fixed to be constant. The dependency of construction efficiency on *D*, shown by Raup (1967), can be partly attributed to an artefact caused by the definition of relative shell thickness. In contrast, Parent et al. (2020) used the ratio of the square root of the surface area to the cubic root of the volume (Vogel 1981) of an ammonoid shell as a metric of construction inefficiency. They calculated the area-to-volume ratio for various shell shapes using an original mathematical model introduced by Parent et al. (2010) and indicated that construction inefficiency tends to increase with increasing whorl overlap.

In addition to construction efficiency, hydrodynamic properties are possibly important determinants of ammonoid shell form. Hydrodynamic properties depend on body chamber length as well as on shell form in ammonoids. Living cephalopods swim using thrust force produced by the hyponomic jet. In living *Nautilus*, the force acts to rotate the shell, resulting in a rocking motion during swimming. This instability is a potential problem for any other ectoconchleate cephalopods with planispiral conchs when they swim. A long body chamber seems to place the aperture higher and positions the hyponome at the same level as the center of gravity so that the hyponomic jet can act directly through the center of gravity (Jacobs and Landman 1993; Klug and Korn 2004). Therefore, the aperture orientation, presuming a neutrally buoyant condition, is a possible indicator of hydrodynamic efficiency. The line between the center of gravity and the center of buoyancy is vertical. Swan and Saunders (1987) assessed the effect of shell shape on the aperture orientation. They computed the aperture orientation for various combinations of body chamber length and parameter value (W or D) under a fixed relative shell thickness based on the method in Saunders and Shapiro (1986). Swan and Saunders found that even minor changes in W need considerable changes in body chamber length to maintain neutral buoyancy, resulting in appreciable changes in aperture orientation, whereas a change in D has little effect on aperture orientation.

Chamberlain (1981) focused on a hydrodynamic property of a given shell shape in ammonoids and evaluated the resistance of the shell in seawater using a drag coefficient. Chamberlain performed tank experiments using a series of shell models and assessed drag coefficients for various shell shapes to obtain a functional morphospace. As a result, he found that one of the positions in the morphospace realizing the minimum drag coefficient (around $W = 1.9$ and $0.35 \leq D \leq 0.42$) coincides with the peak of the most abundant ammonoid shell shapes as mentioned above. This result suggests that hydrodynamic efficiency is an important determinant for the biased distribution of ammonoid shell forms.

Recently, Hebdon et al. (2020) performed analyses of computational fluid dynamics for various shell shapes and contoured drag coefficients on a ternary morphospace with three end-member morphotypes of ammonoids called "Westermann morphospace" (Ritterbush and Bottjer 2012). They found multiple optima of the drag coefficients in the morphospace and revealed that the landscape of the functional morphospace varies with swimming velocity. Their results partly support the conventional hypothesis that the most efficient shell form depends on the mode of life of ammonoids but also suggest that the relationship between shell shape and hydrodynamic properties is more complex than previously suspected.

Functional morphospace analysis was also applied to the cephalopod septa, which are partitions subdividing the interior of the shell into chambers. Ectoconchleate cephalopods had a conch with chambered phragmocone that served as buoyancy apparatus. In some ammonoids, the periphery of the septum is folded and frilled to form a complex, branched structure (Klug and Hoffman 2015; Furui and Ubukata 2022). The complex septa had been conventionally thought to have helped reinforce the shell with its fluted structure against hydrostatic pressure (Buckland 1836; Westermann 1958). In the past couple of decades, the "Buckland

hypothesis" has been examined using a finite-element analysis of continuum mechanics (Daniel et al. 1997; Hassan et al. 2002; Lemanis et al. 2016). Daniel et al. (1997) performed finite-element stress analyses for various shapes of septal models exposed to pressure stresses. Daniel et al. (1997) described the geometry of the septum by a radially folded surface contacted inside a cylinder using a two-dimensional Fourier method. They assessed a mechanical safety factor as the ratio of breaking stress of shell material to the maximum principal stress applied to the septal surface. The safety factors were computed for many combinations of wave number and amplitude to form a functional morphospace. More complicated septal surfaces showed lower safety factors and yielded higher stress in the septa. This result seems to be a kind of counter-evidence against the "Buckland hypothesis," although it is based on an oversimplified model of the septal surface. More recently, the "Buckland hypothesis" was rejected by a finite-element analysis using the complete models taken from the computed tomography data (Lemanis et al. 2016; Lemanis 2020).

8.3 Functional Morphospace Analyses of Gastropod Shells

The efficiency of shell construction appears to be important for gastropod snails, particularly for terrestrial and freshwater species, because such environments generally lack calcium carbonate, which is necessary for shell formation. The amount of necessary shell material depends on shell shape as well as on shell thickness. The efficiency of shell construction has been studied in helicospiral gastropods. Heath (1985) analyzed the relation of the amount of overlap between successive whorls to the ratio of the area of shell material to the volume enclosed (= inefficiency of shell construction) based on a simple model of shell cross-section. Heath estimated the amount of overlap in many terrestrial gastropod species and found that the actual amount of overlap is generally greater than the optimal amount of overlap for the minimum inefficiency.

Stone (1999) computed the ratio of the volume of shell material to the volume of the space enclosed (inefficiency) using a sophisticated mathematical model designed to reproduce allometric shell growth. Stone evaluated the effect of the amount of whorl overlap on the inefficiency of shell construction by varying one of the morphological parameters obtained from some actual shells and found that the actual amounts of overlap are smaller than the optimal amount of overlap in the marine species utilized. The results obtained by Heath (1985) and Stone (1999) indicated that factors other than construction efficiency determined gastropod shell form.

Another possible cost of shell possession for gastropod snails is related to gravity. Snails attach to substrates with various topographies, such as slopes, vertical walls, or undersides, and horizontal floors. They carry their shells on their backs while creeping. Mechanical stability while creeping depends on the shell shape as well as on its posture. Postural stability can be critical, particularly in land snails, because of the absence of buoyancy provided by water in terrestrial environments.

Okajima and Chiba (2009, 2012) focused on the relationship between postural stability and the relative spire height of the shell in terrestrial pulmonate gastropods. They approximated the shell shape by a cylindrical or conical model and calculated the moment of gravitational force acting at the center of gravity about the fulcrum point for a given posture. The shell shape was represented by the height-to-diameter ratio of the model, or spire index (Cain 1977), and the posture was defined as the angle between the directions of gravitational force and the cylindrical or conical axis of the model. Okajima and Chiba (2009) computed the moment of the force for the various combinations of spire index and postural angle to obtain the contour of the postural instability. As a result, they found that conical shells with a high spire as well as planospiral ones with the lowest spire exhibit a small moment of force and are well balanced on a vertical wall. They concluded that the bimodal distribution of spire height observed in land snails (Cain 1977; Cowie 1995) can be attributed to the relationship between shell shape and postural stability.

Studies on gastropod shell form using functional morphospaces focused on either construction efficiency or postural stability, whereas a particular shell form may reflect a compromise between them. Noshita et al. (2012) analyzed the relation of shell form to the functional properties such as postural stability and construction efficiency using the Raup (1966) mathematical model and assessed how the biased distribution of gastropod shell forms in a morphospace can be explained by these functions. In the model, coiling geometry is defined by the translation rate of the generating curve along the coiling axis (T), as well as by the whorl expansion rate (W) and the relative position of the generating curve (D) (Fig. 8.3A). In gastropods, T represents the relative spire height of the shell while D represents the relative width of the umbilicus, which is a hollow space within a coiled shell (Fig. 8.4A). Noshita et al. (2012) computed the postural stability and construction efficiency for various combinations of values of the parameters W, D, and T for a given posture. Postural stability was estimated by the moment of force when the aperture represented by a generating curve is circular and parallel to the substrate. Therefore, the posture of the snail attached to the substrate is determined by the surface inclination of the substrate. Noshita et al. (2012) calculated postural stability as the average moment of force for various postures by integrating the moment over all ranges of inclination angles (Fig. 8.3C). On the other hand, the efficiency of shell construction was estimated by the ratio of the volume of internal shell space to the volume of space enclosed by the external shell surface. Figure 8.4A, B shows three-dimensional contour diagrams of postural stability and construction efficiency on the T–D morphospace when W was fixed at $10^{0.2}$. The results showed that postural stability tended to increase with decreasing T or D (Fig. 8.4B): a compact shell with a low spire and narrow umbilicus exhibited high postural stability. However, shell construction was efficient when T and D were large (Fig. 8.4C); that is, a less overlapped shell form with a high spire and wide umbilicus made a large available space for the soft body. This result indicated a functional trade-off between postural stability and construction efficiency.

In Raup's original model, the generating curve was parallel to both the whorl radius and coiling axis (Raup 1962; Raup and Michelson 1965). However, the

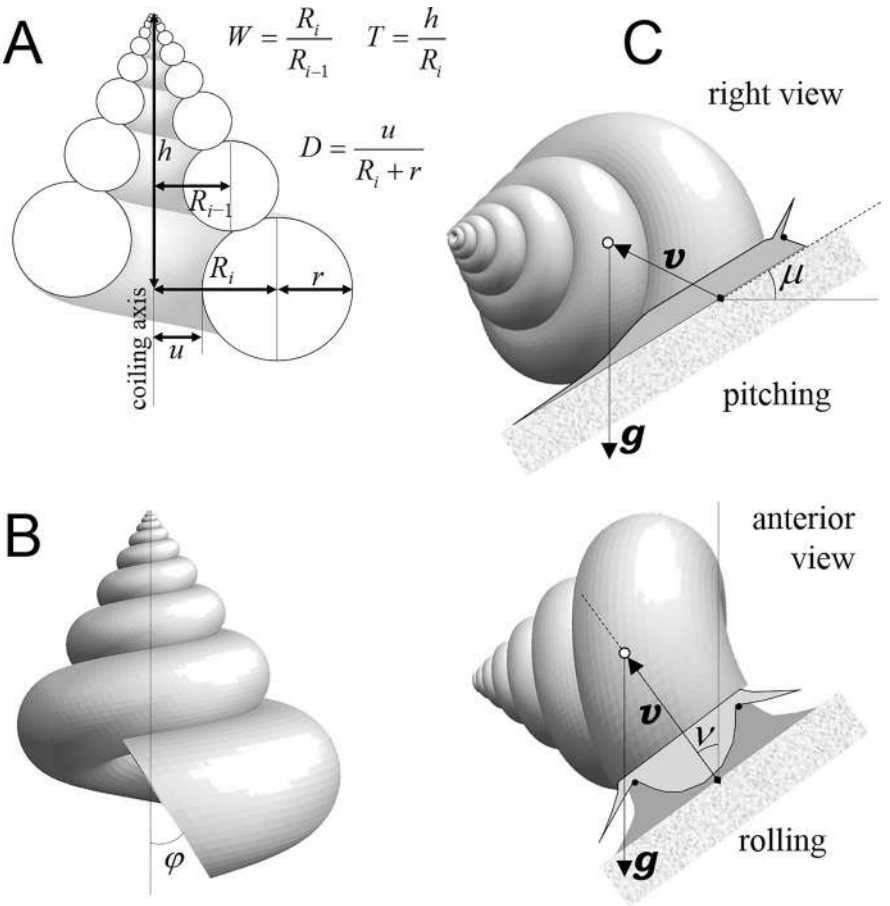


Fig. 8.3 Schematic figures illustrating gastropod shells. (**A**) Rationale for the Raup (1966) model for shell forms of gastropod snails with definitions of the parameters. (**B**) Additional parameter representing an inclination angle of the aperture (φ) adopted by Noshita et al. (2012). (**C**) Snails with different postures showing the positional vector of the center of gravity ($\boldsymbol{v}$) and the gravitational vector ($\boldsymbol{g}$)

aperture of a gastropod shell was commonly inclined with respect to the whorl radius and coiling axis, and the orientation of the aperture affects mobility and adhesion to the substrate (Linsley 1977). Noshita et al. (2012) adopted an additional parameter related to aperture inclination defined by the angle between the generating curve and coiling axis (φ) (Fig. 8.3B) and computed postural stability and construction efficiency for models possessing an inclined aperture. Figure 8.4C shows the results when the aperture is inclined toward the umbilical side (downward in Fig. 8.4D) at an angle of 60°. In this case, the optimum of postural stability shifts toward the upper right direction in the morphospace (Fig. 8.4E) compared to the case when the aperture is parallel to the coiling axis (Fig. 8.4B). The topography of

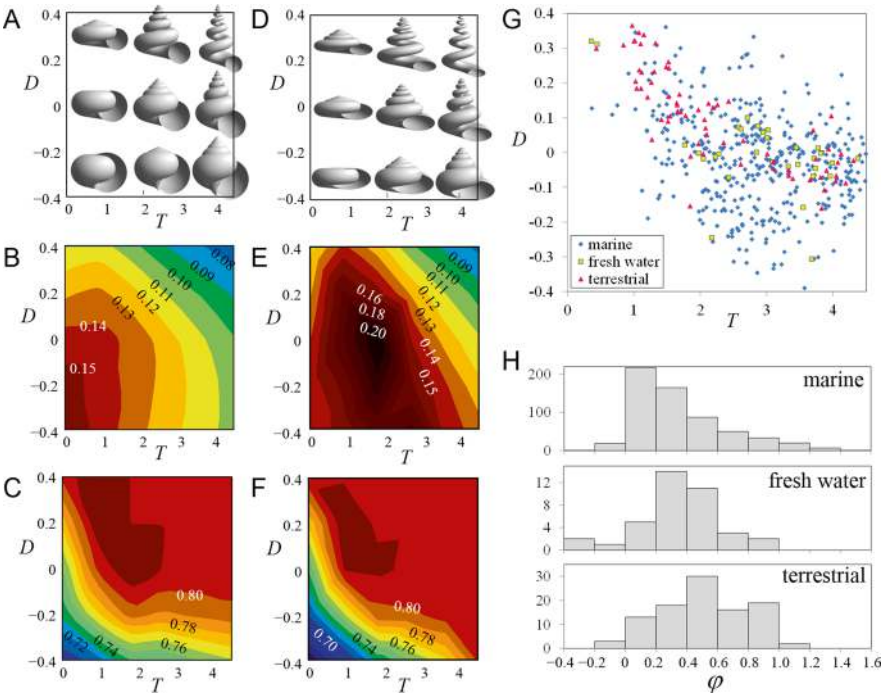


Fig. 8.4 Functional morphospace analyses of gastropod shell forms by Noshita et al. (2012). (**A**–**F**) Results of computer simulations with varying values of T and D at $W = 10^{0.2}$ when the circular aperture is parallel (**A**–**C**) or inclined ($\varphi = 60°$) (**D**–**F**) to the coiling axis. (**A**, **D**) Examples of computer-produced shells with various values of T and D. (**B**, **E**) Three-dimensional contour plots showing the relationship between the parameters (T and D) and postural stability (**B**) or construction efficiency (**E**). (**G**) Biplots of the values of T and D collected from the extant gastropod shells showing their life habitats using different color symbols. (**H**) Histograms of the estimated values of the inclination angle of the aperture (φ) for each life habitat

functional morphospace of construction efficiency is slightly affected by aperture inclination (Fig. 8.4F). The shell shapes around optimal postural stability are also efficient in shell construction in the case of aperture inclination ($\varphi = 60°$; Fig. 8.4E, F). The result indicates that aperture inclination downward relaxes the functional trade-off between postural stability and construction efficiency in gastropod snails.

The functional relationship between postural stability and construction efficiency and its relation to shell form revealed by Noshita et al. (2012) is possibly a key to understanding the biased distribution of gastropod shell shapes. Figure 8.4G shows biplots of the values of T and D collected from a total of 740 extant gastropod species. Shell shapes along the diagonal region of the morphospace from the upper-left to the lower-right corner may meet the requirements of both functions (Fig. 8.4E, F). Actual snails do not occupy the lower-left and upper-right regions of the morphospace, and land snails show more restricted distribution along the diagonal region than aquatic species (Fig. 8.4G). This suggests that a balance between

postural stability and efficiency is important, particularly for land snails, because of the lack of buoyancy and calcium carbonate in the terrestrial environment. In that case, it is predicted that aperture inclination downward is more common in land snails than in marine species. This prediction was supported by the comparison in frequency distribution of the angle of aperture inclination among different habitats (Fig. 8.4H). Land snails tend to have a larger aperture inclination than the marine species, and freshwater species show intermediate inclination between them. The results suggest that aperture inclination allows land snails to form well-balanced and efficient shells.

Noshita et al. (2012) focused on the morphological parameters T and D, whereas most gastropod snails have a small umbilicus characterized by a small D value. Araki and Noshita (2023) further explored the effect of the aperture inclination on functional properties. Along with postural instability, they also assessed the degree of shell hindrance during locomotion because shells with a smaller projected area seem to be more suitable for moving through a narrow gap. The degree of hindrance during locomotion was defined by the projected area of the shell in the direction of movement, standardized by the volume of space enclosed. Araki and Noshita constructed functional morphospaces of postural instability and hindrance during locomotion on a T–φ morphospace and revealed that the region of shells with a high spire and an aperture inclined toward the umbilical or apical side shows large scores of the instability and hindrance. They collected the values of T and φ from more than 400 extant species and found that no snail occupies those regions. The result indicates the reason why a high-spired shell with large aperture inclination is absent in nature (Vermeij 1971; Schindel 1990).

8.4 Functional Morphospace Analyses of Bivalve Shells

Bivalves have a shell with much larger whorl expansion rates (W) than do other mollusks. In bivalves, construction efficiency increases with increasing shell convexity if shell thickness is proportional to shell size. In such a case, shell construction is inefficient when an animal has such a flat shell as scallops because the surface area of the shell is large in relation to the volume enclosed. However, the relationship between shell shape and construction efficiency would be more complicated if an increase in shell thickness was allometric with respect to shell size. In bivalves, shell material continued to be accumulated on the inner surface of the existing shell during growth, and the inner surface of the shell portion produced in the earlier growth stage was heavily covered with many coats of shell layer. As a result, the early produced shell portion near the umbo becomes as thick as the shell margin formed in the adult growth stage. Therefore, shell thickness is far from proportional to shell size in most bivalves.

Okabe and Yoshimura (2017) introduced the scaling exponent of shell thickness as an additional parameter for a Raupian mathematical model to assess the effect of the allometric exponent on construction efficiency. Through functional

morphospace analyses, they found that a bivalve-like shell formed with a large expansion rate and a narrow umbilicus was the optimal design for efficient shell construction when the absolute shell thickness was constant with the shell portion. Their result suggested that very large expansion rates commonly observed in bivalves were a good match for the overcoating manner of accretionary shell growth. Okabe and Yoshimura (2017) also indicated that a pointed conical shell with a small expansion rate, such as some turritelliform gastropods, was the optimal design for efficient shell formation when shell thickness was proportional to shell size.

Unlike gastropods and cephalopods, the shell of bivalves consists of two valves that are connected by a ligament attached along the hinge axis. The ligament generates the torque for opening the two valves, such that the animal closes its valves by contracting the adductor muscles and opens them by relaxing the muscles. Bivalves face a peculiar functional problem since they have two valves. If they have an inflated shell, the projected edge of the shell or umbo of one valve may protrude beyond the commissure plane and two umbonal areas would conflict with each other. Even if the umbonal areas do not conflict, how widely the animals can open their valves depends on shell shape (Savazzi 1987). Avoidance of mutual interference of the umbones is critical for bivalves and may constrain their shell forms. Shell coiling of bivalves is not always planispiral, but sometimes helicospiral and the umbo is directed anteriorly, showing a gyrate appearance of the maximum growth line or prosogyrous shell form (Fig. 8.5A) (Trueman 1969). Anterior

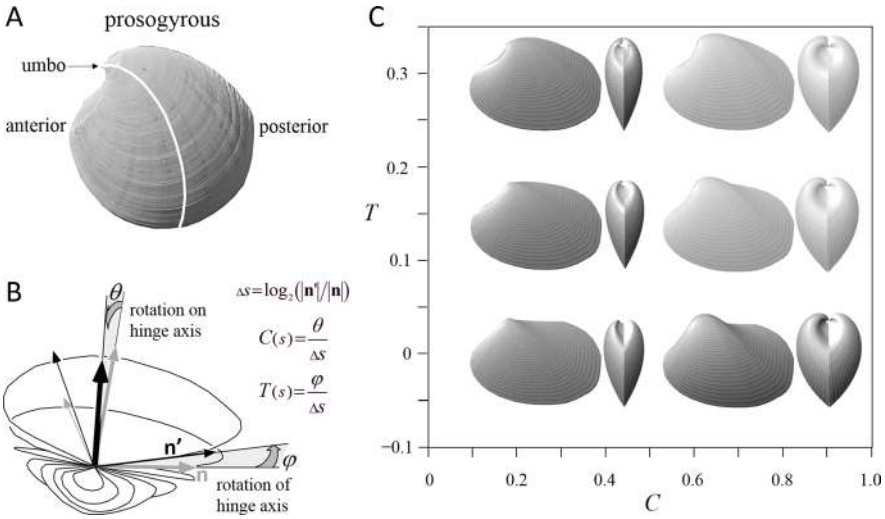


Fig. 8.5 Theoretical morphology of bivalve shells. (**A**) Example of the prosogyrous shell form showing a gyrate appearance of the maximum growth line from the umbo to the venter. (**B**) Rationale for the Ubukata (2000) model for bivalve shell forms with definitions of the parameters. (**C**) Examples of computer-generated shells with six combinations of parameter values ($C = 0.3$ or $C = 0.7$ and $T = 0$, $T = 0.15$ or $T = 0.3$) under a given outline of the generating curve. The darker shaded models depicted in the morphospace show spatial overlap between umbones

inclination of umbones and shell convexity are essential components of bivalve shell coiling.

Ubukata (2000) represented the growth of a bivalve shell by expanding and rotating a given outline of the generating curve and estimated the angle of the maximum valve opening for various shell shapes. The model was defined by the parameters, *C* and *T*, which showed valve convexity and umbonal inclination, respectively (Fig. 8.5B). The shell is inflated if *C* is large and the umbo is anteriorly directed, showing a prosogyrous appearance when *T* is positive (Fig. 8.5C). The darker shaded models depicted in Fig. 8.5C with larger values of *C* and *T* show spatial overlap between umbones and cannot be formed by actual bivalves without any additional mechanism. The relationship between shell shape and angle of the maximum valve opening depends on the ligament structure. The hinge line along which both valves touch each other is not always the same as the opening axis of the two valves. In many bivalves, the opening axis lies along the hinge line (Fig. 8.6A). However, in some bivalve groups possessing a parivincular-type ligament with a cylindrical shape, the opening axis is located along the dorsal edge of the shell and is away from the hinge line (Fig. 8.6B).

Ubukata (2000) calculated the angle of maximum valve opening for various shape models, considering the position of the opening axis. Figure 8.6A, B shows three-dimensional contour diagrams of functional morphospaces for shell opening. When the opening axis lies along the hinge line, a prosogyrous shell with a large *T* value tends to cause difficulties in valve opening (Fig. 8.6A). However, the result of the calculation in the case of the parivincular-type ligament showed that the angle of maximum valve opening does not depend on *T* because the opening axis is away from the hinge line (Fig. 8.6B). The results indicate that the parivincular-type ligament mitigates the risk of conflict between umbones attributed to formation of a prosogyrous shell with a large *T* value. This prediction was supported by morphometric data on *T* collected from 284 bivalve species. The species possessing the parivincular-type ligament tend to have larger *T* values than the others (Fig. 8.6C, D), and prominently prosogyrous shells are exclusively found in the phylogenetic group possessing the parivincular-type ligament, which is called the Heteroconchia. This suggests that the parivincular-type ligament allows evolution of a prosogyrous shell form for heteroconch bivalves without inhibiting the function of valve opening.

A prosogyrous shell form is considered to be efficient for burrowing into sediments (Stanley 1970) because a large *T* value makes the anteroventral shell edge sharp, opposing sediment penetration (Savazzi 1987; Ubukata 2001). This feature is characteristic of the Heteroconchia, and this fact may explain why efficient burrowers are found only in this group. The fossil record indicates that the Heteroconchia increased rapidly after the Cretaceous Period (Fig. 8.6E), coinciding with the increase in shell-crushing predators during the Mesozoic marine revolution (Vermeij 1977). The parivincular-type ligament seems to allow heteroconchs to develop adaptive shell shapes for burrowing through their macroevolution, and this may be why they became prominent among bivalves after the Cretaceous.

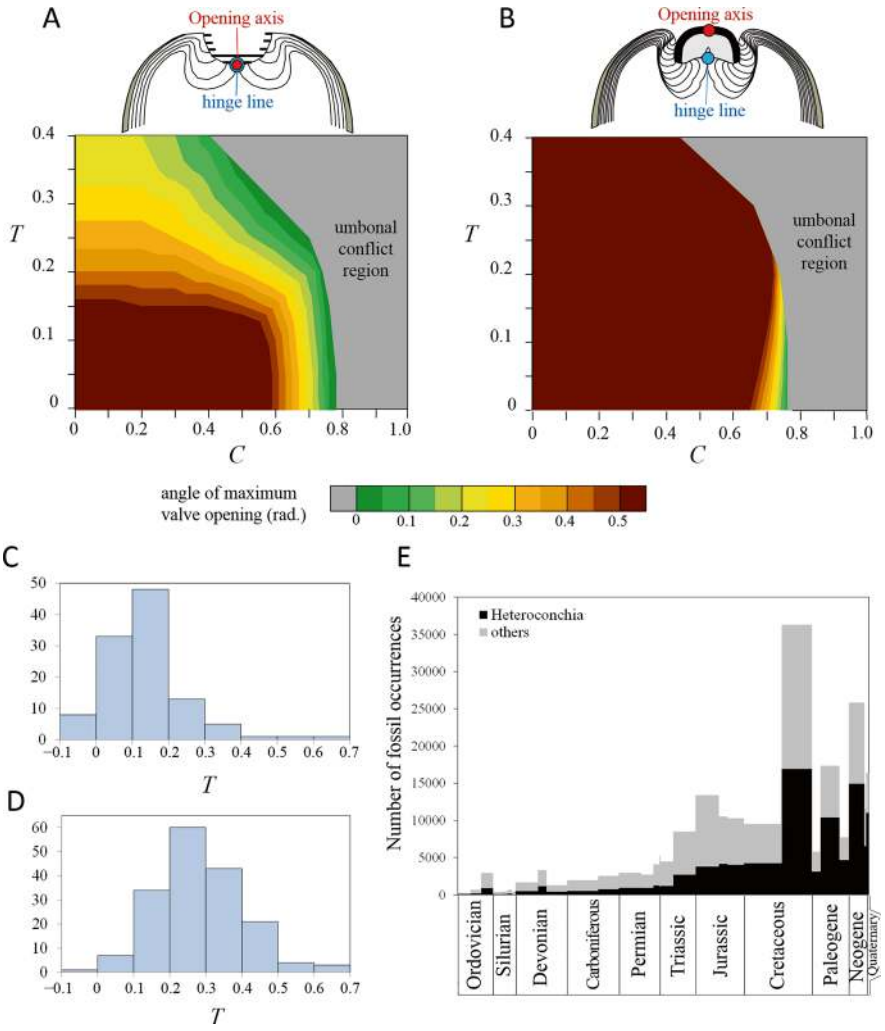


Fig. 8.6 Functional morphospace analyses of bivalve shell form by Ubukata (2000). (**A**, **B**) Three-dimensional contour diagrams showing the relationship between the parameters (*C* and *T*) and the angle of maximum valve opening (radian) with illustrations of the perpendicular cross-sections of the hinge area showing the position of the opening axis and hinge line. (**C**, **D**) Histograms of the estimated values of *T* from measurements. (**A**, **C**) Results in the case of a non-parivincular-type ligament. (**B**, **D**) Results in the case of the parivincular ligament. (**E**) Change in the number of fossil occurrences of heteroconch and other bivalves across the Phanerozoic showing the increase of the proportion of heteroconchs. The occurrence data were downloaded from the Paleobiology Database on November 11, 2025, with the following parameters: taxonomy = Bivalvia or Heteroconchia, time intervals = Early Ordovician to Quaternary

8.5 Concluding Remarks and Perspectives

The adaptive landscape concept originally introduced by Wright (1932) was extended by McGhee (1980, 2007) to include the contour diagrams depicting the functional properties on the theoretical morphospaces. Such functional morphospaces have been used to explore the effects of morphological parameters on a function of interest. A series of theoretical morphological models of molluscan shells, pioneered by Raup (1966, 1967), has advanced the evolutionary morphology of mollusks using functional morphospaces. Functional morphospace is a powerful tool for providing a working hypothesis for connections between the morphological distribution and adaptation. Modern improvement of computer performance enables the combination of continuum biomechanics with functional morphospace analysis.

The present chapter focuses on the studies of molluscan shell forms using theoretical morphospaces. However, functional morphospace analysis is potentially useful for various characteristics of a wide range of organisms, including fossils. Functional morphospaces constructed through geometric morphometric and multivariate analyses enable the exploration of the functional implications of the pattern of morphological variation (Pierce et al. 2008; O'Higgins et al. 2011). The functional morphospace concept can be extended to include discrete categories based on qualitative characters defining a design space (Hickman 1993), which applies to a broad range of taxa with distinct body plans. Evolutionary morphology based on functional morphospace analysis can be facilitated by big-data analyses using morphological databases. Integrating the conventional functional morphospace analyses with their potentially relevant techniques is an interesting subject of future research.

References

Araki A, Noshita K (2023) Theoretical morphological analysis of differential morphospace occupation patterns for terrestrial and aquatic gastropods. Evolution 77:1864–1873. https://doi.org/10.1093/evolut/qpad110

Bicknell RDC, Ledogar JA, Wroe S et al (2018) Computational biomechanical analyses demonstrate similar shell-crushing abilities in modern and ancient arthropods. Proc R Soc B 285:20181935. https://doi.org/10.1098/rspb.2018.1935

Bishop PJ, Cuff AR, Hutchinson JR (2021) How to build a dinosaur: musculoskeletal modeling and simulation of locomotor biomechanics in extinct animals. Paleobiology 47:1–38. https://doi.org/10.5061/dryad.73n5tb2v9

Buckland W (1836) Geology and mineralogy considered with reference to natural theology, treatise VI, the Bridgewater treatise on the power, wisdom, and goodness of god as manifested in the creation. W Pickering

Cain A (1977) Variation in spire index of some coiled gastropod shells, and its evolutionary significance. Philos Trans R Soc B 277:377–428

Chamberlain JA Jr (1981) Hydromechanical design of fossil cephalopods. In: House MR, Senior JR (eds) The ammonoidea. Academic, New York, pp 289–336

Cowie RH (1995) Variation in species diversity and shell shape in Hawaiian land snails: in situ speciation and ecological relationships. Evolution 49:1191–1202
Daniel TL, Helmuth BS, Saunders WB et al (1997) Septal complexity in ammonoid cephalopods increased mechanical risk and limited depth. Paleobiology 23:470–481. https://doi.org/10.1017/S0094837300019849
Foote M, Miller AI (2007) Principles of paleontology, 3rd edn. Freeman, New York
Furui H, Ubukata T (2022) Allometry between suture line length and phragmocone volume in some cretaceous ammonoids. Paleontol Res 26:55–73. https://doi.org/10.2517/PR200031
Hammer Ø, Haper DA (2024) Paleontological data analysis, 2nd edn. Wiley, Malden
Hassan MA, Westermann GE, Hewitt RA et al (2002) Finite-element analysis of simulated ammonoid septa (extinct Cephalopoda): septal and sutural complexities do not reduce strength. Paleobiology 28:113–126. https://doi.org/10.1666/0094-8373(2002)028<0113:feaosa>2.0.co;2
Heath DJ (1985) Whorl overlap and the economical construction of the gastropod shell. Biol J Linn Soc 24:165–174
Hebdon N, Ritterbush K, Choi Y (2020) Assessing the morphological impacts of ammonoid shell shape through systematic shape variation. Integr Comp Biol 60:1320–1329. https://doi.org/10.1093/icb/icaa067
Hickman CS (1993) Theoretical design space: a new program for the analysis of structural diversity. Neues Jahrb Geol Palaontol Abh 190:169–182
Jacobs D, Landman NH (1993) *Naulitus*—a poor model for the function and behavior of ammonoids? Lethaia 26:101–111
Klug C, Hoffman R (2015) Ammonoid septa and sutures. In: Klug C, Korn D, De Baets K et al (eds) Ammonoid paleobiology: form anatomy to ecology, Topics in geobiology, vol 43. Springer, Dordrecht, pp 45–90
Klug C, Korn K (2004) The origin of ammonoid locomotion. Acta Palaeontol Pol 49:235–242. http://app.pan.pl/acta49/app49−235.pdf
Lemanis R (2020) The ammonite septum is not an adaptation to deep water: re-evaluating a centuries-old idea. Proc R Soc B 287:20201919. https://doi.org/10.1098/rspb.2020.1919
Lemanis R, Zachow S, Hoffman R (2016) Comparative cephalopod shell strength and the role of septum morphology on stress distribution. PeerJ 4:e2434. https://doi.org/10.7717/peerj.2434
Linsley RM (1977) Some "laws" of gastropod shell form. Paleobiology 3:196–206
Marcé-Nogué J (2020) One step further in biomechanical models in palaeontology: a nonlinear finite element analysis review. PeerJ 10:e13890. https://doi.org/10.7717/peerj.13890
McGhee GR Jr (1980) Shell form in the biconvex articulate Brachiopoda: a geometric analysis. Paleobiology 6:57–76
McGhee GR Jr (1999) Theoretical morphology: the concept and its application. Columbia University Press, New York
McGhee GR Jr (2007) The geometry of evolution. Cambridge University Press, Cambridge
McKinney FK, Raup DM (1982) A turn in the right direction simulation of erect spiral growth in the bryozoans *Archimedes* and *Bugula*. Paleobiology 8:196–210
Niklas KJ (1997) Effects of hypothetical developmental barriers and abrupt environmental changes on adaptive walks in a computer-generated domain for early vascular land plants. Paleobiology 23:63–76. https://doi.org/10.1017/S009483730001664X
Noshita K, Asami T, Ubukata T (2012) Functional constraints on coiling geometry and aperture inclination in gastropods. Paleobiology 38:322–334. https://doi.org/10.1666/10060.1
O'Higgins P, Cobb SN, Fitton LC et al (2011) Combining geometric morphometrics and functional simulation: an emerging toolkit for virtual functional analyses. J Anat 218:3–15. https://doi.org/10.1111/j.1469-7580.2010.01301.x
Ohno R, Sentoku A, Matsumoto S et al (2015) Morphological variability in azooxanthellate scleractinian dendrophylliids governed by regular modes of asexual reproduction: a computer simulation approach. Paleontol Res 19:195–203. https://doi.org/10.2517/2015PR002

Oji T, Okamoto T (1994) Arm autotomy and arm branching pattern as anti-predatory adaptations in stalked and stalkless crinoids. Paleobiology 20:27–39
Okabe T, Yoshimura J (2017) Optimal designs of mollusk shells from bivalves to snails. Sci Rep 7:42445. https://www.nature.com/articles/srep42445
Okajima R, Chiba S (2009) Cause of bimodal distribution in the shape of a terrestrial gastropod. Evolution 63:2877–2887. https://doi.org/10.1111/j.1558-5646.2009.00780.x
Okajima R, Chiba S (2012) Adaptation from restricted geometries: the shell inclination of terrestrial gastropods. Evolution 67:429–437. https://doi.org/10.1111/j.1558-5646.2012.01772.x
Parent H, Greco AF, Bejas M (2010) Size-shape relationships in the Mesozoic planispiral ammonites. Acta Palaeontol Pol 55:85–98. https://doi.org/10.4202/app.2009.0066
Parent H, Matias B, Andres G (2020) Shell area-to-volume ratio in ammonoids. Paleontol Res 24:216–225. https://doi.org/10.2517/2019PR013
Paul CRC (1999) The paradigm method. In: Savazzi E (ed) Functional morphology of the invertebrate skeleton. Wiley, Chichester, pp 25–28
Pierce SE, Angielczyk KD, Rayfield EJ (2008) Patterns of morphospace occupation and mechanical performance in extant crocodilian skulls: a combined geometric morphometric and finite element modeling approach. J Morphol 269:840–864. https://doi.org/10.1002/jmor.10627
Raup DM (1962) Computer as aid in describing form in gastropod shells. Science 138:150–152
Raup DM (1966) Geometric analysis of shell coiling: general problems. J Paleontol 40:1178–1190
Raup DM (1967) Geometric analysis of shell coiling: coiling in ammonoids. J Paleontol 41:43–65
Raup DM, Chamberlain JA Jr (1967) Equations for volume and center of gravity in ammonoid shells. J Paleontol 41:566–574
Raup DM, Michelson A (1965) Theoretical morphology of the coiled shell. Science 147:1294–1295
Rayfield EJ (2007) Finite element analysis and understanding the biomechanics and evolution of living and fossil organisms. Annu Rev Earth Planet Sci 35:541–576. https://doi.org/10.1146/annurev.earth.35.031306.140104
Ritterbush KA, Bottjer DJ (2012) Westermann Morphospace displays ammonoid shell shape and hypothetical paleoecology. Paleobiology 38:424–446. https://doi.org/10.1666/10027.1
Rudwick MJS (1961) The feeding mechanism of the Permian brachiopod *Prorichthofenia*. Palaeontol 3:450–471
Rudwick MJS (2018) Functional morphology in paleobiology: origins of the method of 'paradigms'. J Hist Biol 51:135–178. https://doi.org/10.1007/s10739-017-9478-7
Saunders WP, Shapiro EA (1986) Calculation and simulation of ammonoid hydrostatics. Paleobiology 12:64–79
Savazzi E (1987) Geometric and functional constraints on bivalve shell morphology. Lethaia 20:293–306
Savazzi E (1999) Introduction to functional morphology. In: Savazzi E (ed) Functional morphology of the invertebrate skeleton. Wiley, Chichester, pp 3–13
Schindel DE (1990) Unoccupied morphospace and the coiled geometry of gastropods: architectural constraint or geometric covariation? In: Ross RM, Allmon WD (eds) Causes of evolution: a paleontological perspective. University of Chicago Press, Chicago, pp 270–304
Shiino Y, Kuwazuru O, Yoshikawa N (2009) Computational fluid dynamics simulations on a Devonian spiriferid *Paraspirifer bownockeri* (Brachiopoda): generating mechanism of passive feeding flows. J Theor Biol 259:132–141. https://doi.org/10.1016/j.jtbi.2009.02.018
Stanley SM (1970) Relation of shell form to life habits in the Bivalvia (Mollusca). GSA Memoir 125:1–296
Stone JR (1999) Using a mathematical model to test the null hypothesis of optimal shell construction by four marine gastropods. Mar Biol 134:397–403
Swan ARH (2001) Occupation of morphospace. In: Brigs DEG, Crowther PR (eds) Palaeobiology II. Blackwell, Oxford, pp 157–161
Swan ARH, Saunders WB (1987) Function and shape in late Paleozoic (mid-carboniferous) ammonoids. Paleobiology 13:297–311
Tayler MA (1987) How tetrapods feed in water: a functional analysis by paradigm. Zool J Linnean Soc 91:171–195

Trueman ER (1969) Ligament. In: Cox LR, Newell ND, Boyd DW et al (eds) Mollusca 6, Bivalvia. Treatise on invertebrate paleontology, part N. GSA and University of Kansas, Boulder, pp N58–N64

Tyszka J (2006) Morphospace of foraminiferal shells: results from the moving reference model. Lethaia 39:1–12. https://doi.org/10.1080/00241160600575808

Ubukata T (2000) Theoretical morphology of hinge and shell form in Bivalvia: geometric constraints derived from space conflict between umbones. Paleobiology 26:606–624. https://doi.org/10.1666/0094-8373(2000)026<0606:TMOHAS>2.0.CO;2

Ubukata T (2001) Morphological significance of the orientation of shell coiling and the outline of "the aperture" in bivalve molluscs. Neues Jahrb Geol Palaontol Abh 221:249–270. https://doi.org/10.1127/njgpa/221/2001/249

Vermeij GJ (1971) Gastropod evolution and morphological diversity in relation to shell geometry. J Zool 163:15–23

Vermeij GJ (1977) The Mesozoic marine revolution: evidence from snails, predators and grazers. Paleobiology 3:245–258

Vogel S (1981) Life in moving fluids: the physical biology of flow. Princeton University Press, Princeton

Westermann GEG (1958) The significance of septa and sutures in Jurassic ammonite systematics. Geol Mag 95:441–455

Wright S (1932) The roles of mutation, inbreeding, crossbreeding and natural selection in evolution. In: Jones DF (ed) Proceedings of the sixth international congress of genetics, vol 1. Brooklyn Botanic Gardens, New York, pp 356–366

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Part III
Modeling Biomorphs in Nature

Chapter 9
Mechanical Optimality Hidden in the Structure of Plants



Hiroyuki Shima

Abstract The natural world displays plants in a wide variety of forms. The diversity is the result of evolution, as plants—bound by the fate that "once they take root, they cannot move for the rest of their lives"—adapt to their surrounding environments. Among numerous fascinating examples, this chapter focuses specifically on the shapes and appearances of four plant species particularly familiar in Japan: (1) bamboo, (2) butterbur, (3) rice, and (4) melon. First, I will discuss the shapes of bamboo and butterbur, which are both exceptionally tall plants with hollow culms and stems. Their shapes demonstrate an optimal structure for supporting their own weight. Next, I will present a detailed study on rice, which unveiled the trade-off relationship between the tapered shape and rigidity of rice culms. Finally, I will explain how the net-like pattern on the rind of the muskmelon, highly valued by the Japanese as an expensive and delicious fruit, exhibits specific geometric properties.

Keywords Bamboo · Butterbur · Rice culm · Muskmelon · Bending stiffness · Buckling · Expansion fracture · Structural mechanics

9.1 Introduction

Many plant species surrounding us today have faced fierce competition for survival since ancient times. This fact implies that the forms of extant plant species bear traces of having evolved intelligent structures to adapt to their surroundings and survive. For example, the form of tree branches and leaves spreading toward the sky is the result of long-term evolution to capture as much sunlight as possible. Another example includes climbing plants such as morning glories and green beans, which save energy by wrapping around or twining with other plants to ascend, efficiently reaching well-lit areas while avoiding the need to develop their own sturdy stems. Furthermore, other plants develop thorns or tough leaves to deter animals from

H. Shima (✉)
Department of Environmental Sciences, University of Yamanashi, Yamanashi, Japan
e-mail: hshima@yamanashi.ac.jp

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_9

eating them, or give their seeds wing-like structures to efficiently disperse them far away via wind. Numerous examples exist of intelligent forms and structures acquired by plants to survive competition, yet relatively few cases have quantitatively described the relationship between plant organ morphology and mechanical function from the perspectives of theoretical physics or structural mechanics. Herein lies significant potential for physics and mechanics to make meaningful contributions to plant science.

This chapter explores the intriguing correlation between morphology and mechanical function, particularly focusing on the following four types of agricultural crops. The first target is "bamboo," which is one of the most particularly intriguing plant species from the perspective of structural mechanics. Unlike ordinary trees, bamboo is hollow inside and possesses nodes at regular intervals along its culms. Why did bamboo evolve such a unique shape? What survival advantages do its hollow structure and nodes confer? Similar questions arise regarding the second target, "butterbur." The stem of butterbur can reach a maximum diameter of about 10 cm yet grows vertically to heights of 1–2 m or more. Moreover, it spreads an abnormally large, heart-shaped leaf horizontally from its tip. In other words, it supports an extremely heavy leaf with just a single, slender stem. What is the structural secret behind this remarkable feat? The question does not end here; a distinct class of structural optimality is hidden in the culm of "rice," the source of staple food beloved by the Japanese. In recent years, a group of Japanese researchers systematically examined the bending stiffness and tapered shape of 12 types of rice culms, and revealed a common rule in the culm structures that is effective for avoiding culm fracture under strong wind and self-weight. The last target is "muskmelon," which are also the agricultural crops particularly favored by the Japanese. Muskmelons produced in Japan feature a uniform, fine net-like pattern covering the surface of their rind. This net pattern is actually an intriguing subject from the perspective of fracture mechanics.

9.2 Optimal Structure of Bamboo

9.2.1 Bamboo Nodes Suppress Bending Deformation of Culm

Bamboo is a group of the grass family (not a tree), and the fastest growing plant in the world (Shima et al. 2023). Some species of bamboo can grow up to 1 m in a single day. This rapid growth originates mainly from the hollow structure of bamboo culms. Because the culms are hollow, they require fewer photosynthetic products for building their own body, enabling this rapid growth. This biological characteristic is used as a metaphor for Princess Kaguya (a fictional girl who grows quite rapidly), the heroine of the oldest folk tale in Japan (Inoue and Shima 2018). From an engineering perspective, the rapid maturation of bamboo forests provides a low-cost, readily available material for use in the construction and pulp

industries. Furthermore, bamboo has remarkable capabilities for biomass accumulation, particularly carbon storage, which is beneficial to the global carbon cycle. These features have led to bamboo emerging as a green alternative to other woody materials.

While the hollow nature of bamboo culms offers various advantages, it also presents a serious drawback that compromises structural stability. This drawback is the heightened risk of culms collapsing under external forces such as wind, rain, or their own weight. To mitigate this risk, bamboo produces a series of nodes at intervals along the culm. In plant terminology, a node is a combination of an external ridge at the outer surface and an internal circular disc (called diaphragm) embedded in the hollow cavity. Since a diaphragm acts as a ring stiffener for the region near the node, sequential insertion of many diaphragms with an appropriate spacing along the culm leads to improvement in the mechanical stiffness of the whole bamboo culm against crosswind and self-weight.

The reason why the mechanical resistance increases with the number of nodes can be understood as follows. Generally, when a hollow cylindrical object member is subjected to bending deformation, the cross-section of the cylinder becomes elliptical, while it is circular in the original state (i.e., prior to deformation). For example, if you pinch both ends of a straight rubber tube and bend it sharply, the cross-section of the tube will become squashed flat. Now, imagine that the rubber tube is completely filled with sand. After filling it with sand, the tube becomes rock-hard. No matter how much force you apply to bend it, it will not deform at all. The reason is that the sand pushes outward against the tube's walls, preventing deformation of the cross-section. In other words, if you can suppress deformation of the cylindrical cross-section, you can make the entire cylinder far more resistant to bending. Bamboo employs the same mechanism. Namely, bamboo locally reinforces its culm by inserting disc-shaped stiffeners (diaphragms) into hollow sections. Furthermore, by inserting numerous stiffeners from the base to the tip of the bamboo, increasing the number of reinforcement points, it significantly enhances the overall mechanical stability of the culm.

However, it should be noted that more nodes are not necessarily better. Excessive node formation inhibits the bamboo's growth rate and increases its self-weight. Therefore, the number of nodes must be set considering the balance between stability and growth rate. This trade-off suggests that bamboo has discovered the optimal number of nodes through millions of years of evolution.

So how are the optimal number and distribution of nodes determined? Simply put, more (or fewer) nodes should form at positions subjected to strong (or weak) bending stress. Wild bamboo embodies precisely this optimal solution (Shima et al. 2016).

According to structural mechanics theory, the degree to which the nodes (diaphragms) suppress the bending deformation of a hollow cylinder when bending stress is applied to it can be expressed by the following dimensionless parameter:

$$\Omega = \left(\frac{E_{\parallel}}{E_{\perp}} \right)^{\frac{1}{4}} \cdot \left(\frac{\mathrm{l}^2 w}{r^3} \right)^{\frac{1}{2}} \tag{9.1}$$

where $E_{\parallel}$ and $E_{\perp}$ are the Young's moduli parallel to the cylinder axis and in the circumferential direction, respectively. In typical bamboos, the ratio $E_{\parallel}/E_{\perp}$ quantifies the mechanical anisotropy in the woody portion of the bamboo culms, whose value has been estimated to be on the order of 10 or less. r is the cross-sectional radius of the culm, ℓ is the separation distance between adjacent nodes along the culm, and w is the wall thickness. It follows from structural mechanics theory that the bending deformation of the cylinder is most effectively suppressed when this dimensionless parameter Ω satisfies the following inequality:

$$\Omega < \left(\frac{256}{15}\right)^{\frac{1}{4}} \simeq 2.033 \tag{9.2}$$

The right-bottom panel in Fig. 9.1 shows the measurement results of Ω as a function of the elevation from the ground; x is the relative internode number ($0 \leq x \leq 1$) that corresponds to the elevation from the ground surface ($x = 0$) to the tip of the culm ($x = 1$). The curve of Ω shows a monotonic increase with x. An important observation is that the criterion $\Omega < 2.033$ holds for every $x < 0.22$. The graph also tells us that the degree of reinforcement is greatly enhanced near the base, where Ω becomes less than 1.0. The considerable reduction in Ω is quite reasonable, since bamboo segments near the base are subjected to significant bending moments even when a cross-wind of uniform strength acts on the entire culm. This coordinated monotonicity in the $\Omega(x)$ curve, and the biased bending-moment distribution under cross-wind, indicates a plausible guiding principle of wild bamboo morphology. Namely, in the growing process, wild bamboos control their geometric parameters such as $\ell(x)$, $r(x)$, and $w(x)$ so that the resulting $\Omega(x)$ increases with x in a monotonic manner.

It should be emphasized that an assessment of the spatial variation of $\Omega(x)$ is crucially important to gaining insight into the role of nodes and internodes in bamboo mechanics. A simple observation of $\ell(x)$ alone is of no use; in reality, the combination of spatial variation in $\ell(x)$, $r(x)$, and $w(x)$ results in a monotonic increase in $\Omega(x)$, i.e., a gradual increase in the degree of reinforcement from the base to the tip, which effectively resists the applied bending moment sketched in the bottom-right panel in Fig. 9.1. We thus conclude that the wild bamboo geometry is optimized in the sense that every internode is endowed with an appropriate value of Ω.

9.2.2 Bamboo Fibers Are Optimally Distributed in Culm

In addition to the node-driven reinforcement explained above, another interesting mechanism lurks in the cross-section of the bamboo to increase the culm stability as a whole—the bamboo fiber.

The woody part of bamboo contains numerous long fibers embedded longitudinally. These fibers, known in biological terms as vascular bundles, function as vertical transport vessels for water and nutrients. Structurally, bamboo is thus considered a fiber-reinforced composite primarily composed of two types of tissues: the

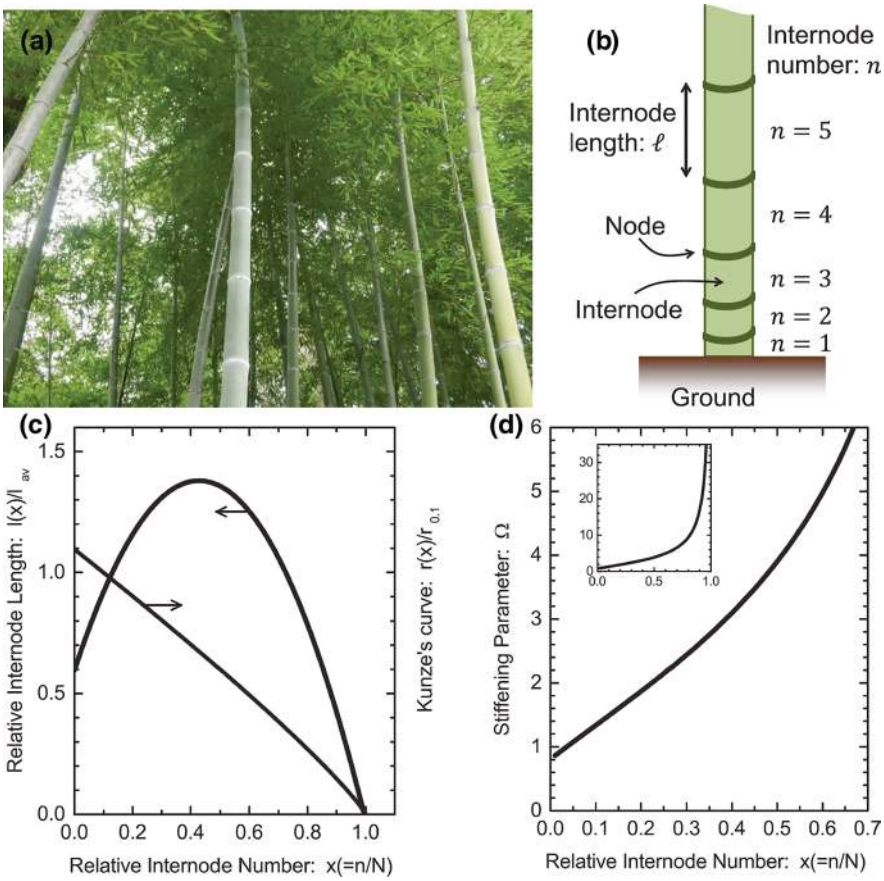


Fig. 9.1 (**a**) Wild bamboo grove. (**b**) Notation and terminology specifying each segment of a whole bamboo culm. (**c**) A parabolic curve expressing the relative internode length $\ell(x)/\ell_{av}$ as a function of the relative internode number. The normalized Kunze's curve regarding the culm radius $r(x)$ is also shown. Here, $r_{0.1}$ represents the bamboo culm radius at a relative height of 10% from the base. (**d**) The curve of $\Omega(x)$ within a limited range of x, indicating that the criterion $\Omega < 2.033$ holds for $x < 0.59$. Here, N denotes the total number of internodes in a given culm. Inset: The whole curve of $\Omega(x)$ over $0 \leq x \leq 1$. (Reprinted with permission from Shima et al. 2016, Copyright 2016, American Physical Society)

vascular bundles as reinforcing fibers and a matrix composed of parenchymatous cells. The fibers possess sufficient rigidity, with a Young's modulus on the order of several gigapascals, comparable to steel. Such rigid fibers are aligned longitudinally and embedded within a continuous matrix that contributes little to the bamboo's rigidity. Structural mechanics theory has revealed that these fibers play a crucial role in the mechanical stability of the bamboos.

The point of interest is the spatial distribution of fibers within the cross-section. If you cut a part of the cross-section of a bamboo and then observe it using a microscope, you will find that the fibers are not evenly distributed over the entire cross-section (see Fig. 9.2a). Specifically, more fibers are embedded closer to the outer

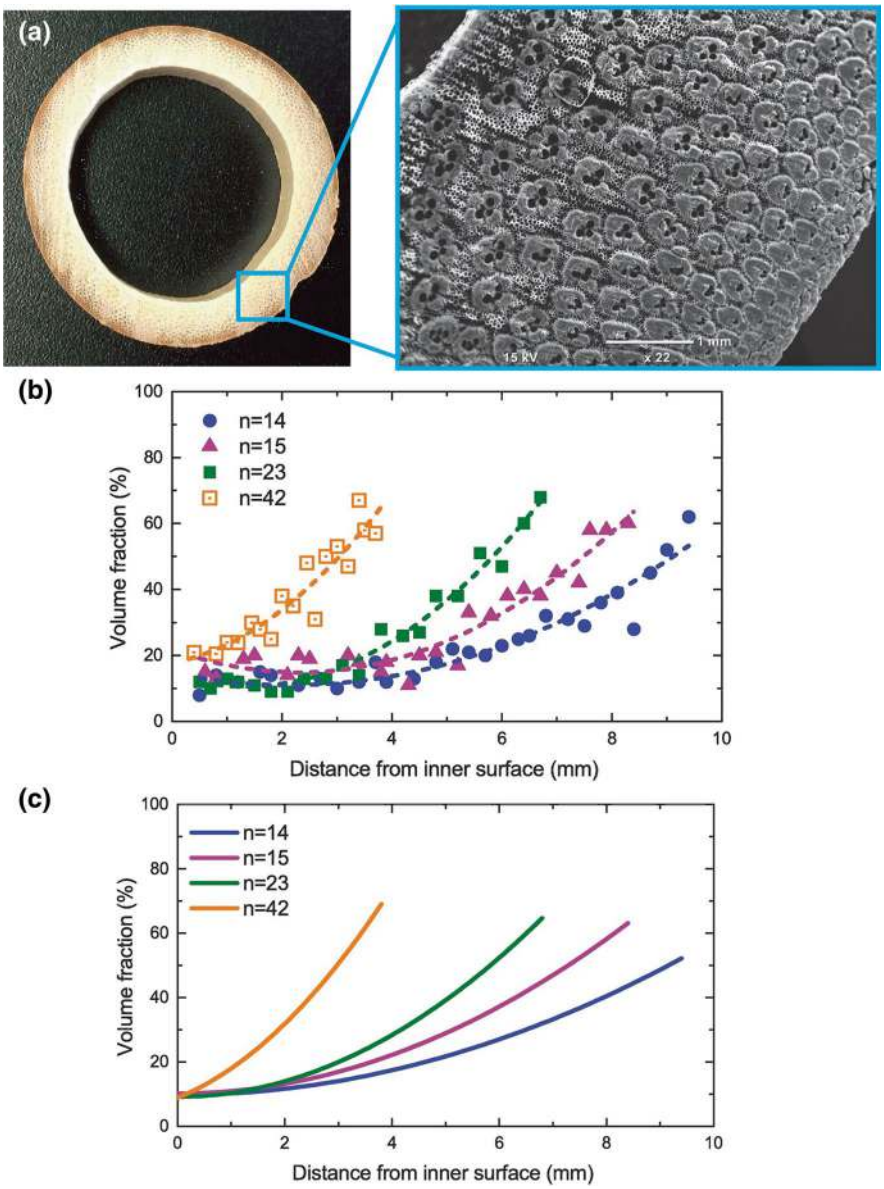


Fig. 9.2 (**a**) A photo of the cross-section of a bamboo. A wild moso bamboo (*Phyllostachys pubescens*) was sampled and its cross-section was observed with a microscope. The gradient distribution of the vascular bundles across the wall thickness is clearly observed. (**b**) Experimentally measured values of the fiber's volume fraction in wild moso bamboos as a function of the distance from the inner surface of the hollow culm. The index *n* in the legend represents the internode number from the ground. (**c**) Theoretically derived optimal distribution of fibers in elastic hollow cylinders. (Adapted from Sato et al. 2017, License CC BY)

skin. Elastic approximation theory tells us that concentrating many reinforcing fibers near the outer skin makes the entire bamboo less prone to bending. This indicates that the bamboo makes its own body stronger by efficiently arranging a limited number of reinforcing fibers inside the cross-section. More concretely, wild bamboos maximize their flexural rigidity by optimally regulating the radial gradation of their vascular bundle distribution (Sato et al. 2017). This fact provides us with a plant-mimetic design principle that enables the realization of high-stiffness and lightweight cylindrical composites.

The optimality in the fiber distribution can be described by the theoretical model given below. Suppose that a bamboo culm is modeled by an elastic hollow cylinder with radius a and wall thickness h. Over the cross-section, the reinforcing fibers are scattered in a gradient from the inner hollow core to the outer surface in a radial direction. It is then hypothesized that the volume fraction of the fibers, denoted by $V_f(r)$, obeys the expression:

$$V_f\left(r\right) = c_0 + c_1 r + c_2 r^2 \left(a - \frac{h}{2} \le r \le a + \frac{h}{2} \right). \tag{9.3}$$

In a real bamboo's cross-section, a limited (but large) number of vascular bundles are distributed optimally for maximizing the flexural rigidity of the entire culm. To explore the optimal distribution, we start from a virtual situation, wherein a given number of reinforcing fibers are uniformly scattered with equal density throughout the wall thickness; we then impose a spatial gradient on the fiber's density, while keeping the total number of fibers fixed. In the initial virtual situation, $V_f(r)$ is a constant, denoted by V_f^{av}, at arbitrary positions in the cross-section. After imposing a certain gradient, $V_f(r)$ becomes r-dependent. The effect of the gradient distribution is thus evaluated by comparing the flexural rigidity obtained from the r-dependent $V_f(r)$ with that from the constant V_f^{av}.

The flexural rigidity, D_c, of the fiber-reinforced cylindrical composites under bending is given by

$$D_c = \int_{a-\frac{h}{2}}^{a+\frac{h}{2}} dr \int_0^{2\pi} d\theta \left[E_c\left(r\right) r^3 \sin^2\theta \right], \tag{9.4}$$

where $E_c(r)$ represents the r-dependent local Young's modulus of the cylindrical composite. In the present situation, it is a linear combination of the Young's modulus of the reinforcing fiber, E_f, and that of the host matrix, E_m, expressed as:

$$E_c\left(r\right) = E_f V_f\left(r\right) + E_m \left[1 - V_f\left(r\right) \right]. \tag{9.5}$$

Substituting Eqs. (9.3) and (9.5) into Eq. (9.4) and integrating, we can express D_c as a linear combination of c_j (j = 0, 1, 2). The remaining task is, then, to find the optimal values of c_j (j = 0, 1, 2) that maximize D_c for a given set of parameters: V_f^{av}, E_f, E_m, a, and h. These parameter values generally change gradually from the base to the tip of the culm. Therefore, by cutting down actual bamboo culms and measuring the vertical variation of these parameter values, then substituting them into the equations shown above, we can theoretically derive the fiber distribution that maximizes the flexural rigidity D_c.

A direct comparison of the theoretical results for the optimal fiber distribution with the existing experimental data for wild bamboos validates the theory explained above. Figure 9.2b depicts the experimentally measured values of the fiber's volume fraction $V_f(r)$ in wild moso bamboos. In the graph, the measured values of $V_f(r)$ are plotted as a function of the distance from the inner surface of the hollow culm. The index n in the legend represents the internode number from the ground; e.g., n = 1 at the bottommost internode for a given individual bamboo culm (see Fig. 9.1b). The plot tells us that in wild moso bamboos, fibers embedded near the bottom of an entire culm distribute quadratically in the radial direction, while those near the top distribute almost linearly.

The characteristic curves of the experimentally obtained $V_f(r)$ in Fig. 9.2b are well captured using the theory that was developed in Sato et al. (2017), as shown in Fig. 9.2c. Surprisingly, the optimal fiber distribution that was theoretically deduced is in quantitative agreement with the earlier experimental data of the vascular bundle distribution in wild moso bamboo. It is clearly observed that the distribution curves theoretically deduced are in good agreement with the curves obtained through the experiment. The agreement between the theory and experiment assures the optimality in the fiber distribution toward reinforcement of elastic hollow cylinders.

The key conclusion drawn from these results is that bamboo is a nature-designed functionally graded material (FGM). FGMs are a class of advanced materials endowed with spatial variations in their physical properties and/or chemical compositions. In addition to artificial FGMs, naturally designed FGMs can be found in our daily life, ranging from bones and teeth in our body to seashells and plants observed outdoors. A possible reason for the frequent realization of FGMs in nature is that the biological cells tend to adapt to external stimuli; their self-adaptive property may have caused a natural gradient structure during the evolutionary process for enhancing the stiffness-to-weight ratio. We can say that bamboo is also one of the most typical and intriguing FGMs that occurs in nature, because the volume fraction of the fibers increases radially from the inner to the outer surface of the bamboo culm. The biased distribution of the fibers across the wall thickness is considered to enhance the flexural rigidity of the culm, as a whole, against crosswinds and bending because of gravity. A closer look at the fiber distribution in the cross-section, therefore, will lead us to an optimal design strategy for fiber-reinforced cylindrical composites inspired by the bamboo morphology.

9.3 Optimal Cross-section of Butterbur Petiole

In the previous section, we saw how bamboo culms enhance their own mechanical stability by arranging diaphragms (i.e., circular plates at nodes) and embedding vascular bundles (i.e., reinforcing fibers) within their bodies. On the other hand, another method to increase the bending resistance of hollow cylindrical structures involves devising the cross-sectional shape. From the view of structural mechanics, the difficulty of bending a hollow pipe is determined by its cross-sectional shape and the direction of the bending force. In fact, the hollow pipes used in industrial products are designed with a cross-sectional shape that is hard to bend. If plants employed the same strategy, then they could enhance resistance to the most significant mechanical stresses (such as crosswinds, stem weight, and petal weight) by selecting the optimal cross-sectional shape (elliptical, polygonal, etc.) according to their growth environment. One plant exemplifying this concept is the Japanese butterbur.

The Japanese butterbur is a well-known spring mountain vegetable that grows in the wild (Fig. 9.3a). It extends numerous leaf stalks vertically from its rhizome, bearing large, rounded heart-shaped leaves at their tips. Butterbur is characterized by its large leaves. At the tip of a thin stem, the very large leaves are spread upward. Remarkably, the weight of the large, spreading leaves at the stem tip accounts for 70% of the entire butterbur plant's weight (Fig. 9.3b). In order to support this large and heavy leaf, it is necessary to devise the structure of the stem. A close examination of the cross-section of the stem of the Japanese butterbur reveals that it has a shape similar to that of a horseshoe (Fig. 9.3c). It is thus hypothesized that this peculiar horseshoe-shaped stem cross-section is the optimal structure for supporting large leaves at the tip of the stem. This hypothesis has been verified both experimentally and theoretically by a Japanese master's degree graduate student (Horiuchi 2022).

The starting point of the consideration was defining a metric to quantify how closely the cross-sectional shape of the butterbur resembles that of a horse's hoof. Figure 9.3d illustrates this definition. Here, A_1 denotes the area of the irregularly shaped butterbur cross-section (green region). Among the outer edges of this irregular region, the convex boundary is approximated by an ellipse, and the entire region is enclosed by this ellipse. The area of the void region sandwitched by the original region's concave boundary and this ellipse is denoted by A_2. The ratio of these two areas, A_2/A_1, is defined as the "horseshoe degree" of the butterbur cross-section shape. The closer the horseshoe degree value is to zero, the closer the cross-sectional shape of the butterbur is to a perfect ellipse, indicating that the defect area indicated by A_2 is small.

A key mechanical property to progress our discussion is called the section modulus. In structural mechanics, the section modulus is a geometric property of a given cross-section used in the design of beams or flexural members. The section modulus depends on the dimensions of a cross-section, such as height and width. The greater the cross-section, the greater the section modulus and the bending resistance of a beam. The section modulus is calculated as

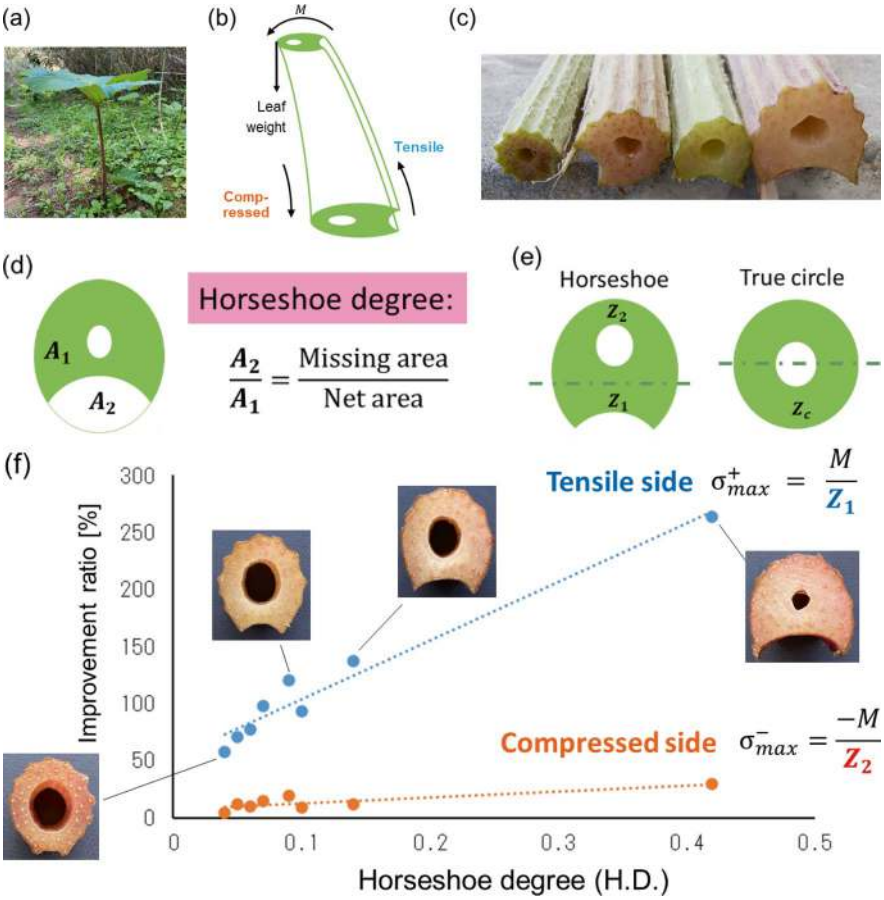


Fig. 9.3 (**a**) Wild butterbur. Its slender, elongated stem supports broad, horizontally spreading leaves. (**b**) Stress acting on the stem. Since the direction of leaf spreading aligns with the convex direction of the stem cross-section, the direction of the bending moment *M* acting at the tip always coincides with the convex direction. Consequently, compressive stress acts on the convex portions of the stem, while tensile stress acts on the concave portions. (**c**) Horseshoe-shaped cross-section of butterbur. (**d**) Definition of the horseshoe degree. (**e**) Section modulus *Z* and the neutral axis of the cross-section. (**f**) Correlation between the improvement ratio of section modulus and horseshoe degree. The left side along the horizontal axis corresponds to the stem tip, and the right side corresponds to the stem base. At the sectional concavity where tensile stress acts, the section modulus increases significantly, thereby markedly reducing tensile stress

$$Z = \frac{I}{c} \tag{9.6}$$

where I is the second moment of the area and c is the distance from the neutral axis to the most extreme fiber. It is used to determine the yield moment strength of a section

$$M_y = Z \cdot \sigma_y \tag{9.7}$$

where σ_y is the yield strength of the material.

For symmetrical cross-sections, such as circular, rectangular, and H-profiles, the section modulus of the top and bottom fibers is the same. For instance, the cross-section of the tree trunk is circular and solid. These solid circular cross-sections can evenly resist mechanical loads from all directions. On the other hand, if we have an asymmetrical complex cross-section, as observed in the butterbur's stem (Fig. 9.3e), the section modulus of the top (i.e., Z_2 in Fig. 9.3e) and the bottom fiber (Z_1) are different. This leads to different bending moment capacities for positive and negative moments. Therefore, the hypothesis arises that the horseshoe-shaped cross-section of the butterbur offers a mechanical advantage over a cylindrical cross-section.

The following three facts are important in inferring the advantages of the cross-sectional shape of butterbur. First, the leaf weight of butterbur is more than 70% of the total weight of the plant. Second, the leaf loads act only in a specific direction, bending the stem. This bending direction does not change much after growth. Lastly, because of the bending, the tensile stress always acts at the position where the cross-section is constricted. In contrast, the compressive stress always acts at the position where the cross-section is swollen. This means that there is harmony between the bending direction of the stem and the shape of the cross-section. Therefore, it is speculated that the horseshoe shape was one of the survival strategies to efficiently support the large heavy leaves.

It is reasonable to infer that the most important thing for butterbur survival is to reduce the maximum stress acting on the stem surface. As mentioned earlier, the butterbur stem is always bent in a specific direction determined by the cross-section. In such a configuration, the magnitude of the maximum stress is always greater at the tensile side than at the compressed side. Namely, the absolute value of $\sigma^{+}_{\max} = M / Z_1$ is always greater than that of $\sigma^{-}_{\max} = -M / Z_2$, where M is the bending moment applied at the tip of the stem (see Fig. 9.3b). Therefore, bending-induced fracture of the butterbur stem always occurs at the tensile side, where the stem cross-section is constricted. This means that if we increase the section modulus on the tensile side, denoted by Z_1, $\sigma^{+}_{\max}$ will decrease so that fracture risk can be reduced.

Figure 9.3f provides experimental evidence for the horseshoe effect on maximum stress reduction. Image analysis of cross-sections taken at different heights on wild butterbur confirmed that, as a rule, the horseshoe degree is greater near the stem base and smaller near the stem apex. Subsequently, the relationship between the horseshoe degree and the section modulus was calculated for each cross-section. The vertical axis of the graph in Fig. 9.3f represents the improvement ratio (i.e., increase rate) of the section modulus (denoted by Z_i with $i = 1, 2$) compared to that of a true circular section (denoted by Z_c), defined as $(Z_i - Z_c)/Z_c$. The results revealed that the horseshoe-shaped cross-section exhibits a pronounced effect on the tension side. Specifically, the section modulus increases significantly at the stem base, leading to a highly effective reduction in maximum stress. This finding indicates that the

butterbur skillfully reduces fracture risk by adopting a horseshoe-shaped cross-section.

In short, the cross-section of the butterbur stem is regulated sophisticatedly to resist the bending force at the base caused by the heavy leaf at the top. Specifically, a horseshoe-shaped cross-section increases the section modulus Z_1 on the concave side, thus reducing the maximum stress acting on this side. As seen in this section, it would be considered highly significant to apply theoretical investigations into how noncircular stem cross-sectional shapes affect stem stability to a diverse range of plant species (Shima et al. 2020).

9.4 Mechanical Trade-offs Inherent in Rice Culms

Rice is a staple food in many countries, primarily in Asia and Africa, and increasing rice yields is a major goal in agriculture. However, to increase yields, rice plants must be mechanically supported to prevent them from lodging due to their own weight or wind. Here, "lodging" refers to the permanent bending or breaking of rice stems from a vertical position toward the ground. Previous research has identified three main types of lodging in rice: (1) bending-type lodging, where the culm bends significantly and fails to recover, (2) breaking-type lodging, where the culm breaks due to excessive weight, and (3) root-break lodging, where the plant lodges because the roots break. Among these, the bending and breaking types (1) and (2) have been well studied, revealing that rice characteristics such as internode length and thickness, and the height from the ground to the panicle, are crucial for these types of lodging. However, no established method existed for comparing varieties with different characteristics, which is essential for rice breeding.

Recently, a group of Japanese researchers measured the culm shape and hardness of various rice varieties to identify those with high lodging resistance (Tsugawa et al. 2023). They discovered significant differences in culm shape and hardness between the relatively tall 5-internode group (varieties with five internodes) and the shorter 4-internode group (varieties with four internodes). Their motivation was to identify varieties with strong lodging resistance. Further analysis revealed a surprising fact: the tall 5-internode group had thicker, softer stems, while the shorter 4-internode group had thinner, harder stems. This characteristic is highly intriguing, suggesting that rice adapts its traits based on culm height to resist lodging. They named this characteristic the "thickness–hardness trade-off relationship." This term describes the trade-off where prioritizing thickness sacrifices hardness and prioritizing hardness sacrifices thickness.

To investigate this discovered trade-off relationship in greater detail, physical quantities that allow comparison between rice plants with different traits were explored. Among various indicators, the following dimensionless parameters were defined: the slenderness ratio λ_i as the morphological safety index, and the gravity-to-bending stiffness ratios, α_i and β_i, as the mechanical safety indices:

$$\lambda_i = \frac{L_i}{R_i}, \quad \alpha_i = \frac{F_i L_i^3}{E_i I_i}, \quad \beta_i = \frac{\rho_i L_i^2}{E_i I_i} \tag{9.8}$$

where L_i is the length of the ith segment, R_i is the sectional radius of gyration of the ith segment, F_i is the tip load of the ith segment, E_i is the Young's modulus (indicator of stiffness) of the ith segment, I_i is the sectional moment of inertia (indicator of thickness) of the ith segment, and ρ_i is the weight of the ith segment.

Generally, dimensionless parameters can express properties independent of spatial scales like length. Thus, λ_i, α_i, and β_i enable us to characterize the lodging resistance of rice plants with different traits in a comparable format. As a result, we could compare each safety index, as shown in Fig. 9.4a. Surprisingly, it was found that these safety indices were nearly equivalent between the 5-internode group and the 4-internode group at the node closest to the ground (see Fig. 9.4b, c).

Therefore, the Japanese research group constructed an elastic column theory to determine whether this mechanical safety index truly characterizes rice lodging resistance. The elastic column theory quantitatively characterizes the weight at which column structures, such as utility poles, bend, enabling the evaluation of column safety. Under the loading conditions shown in Fig. 9.5a, it was found that the local angle θ_i at the ith internode in Fig. 9.5b satisfies the following equation:

$$\frac{d^2\theta_i}{ds_i^2} = -(\alpha_i + \beta_i s_i)\sin\theta_i \quad (i = 1, 2, \cdots b), \tag{9.9}$$

where b is the number of internodes and s_i represents the distance along the internode curve calculated from the top of the ith internode. Deriving this equation enabled quantitative discussion of how much rice bends under relative loads α_i and

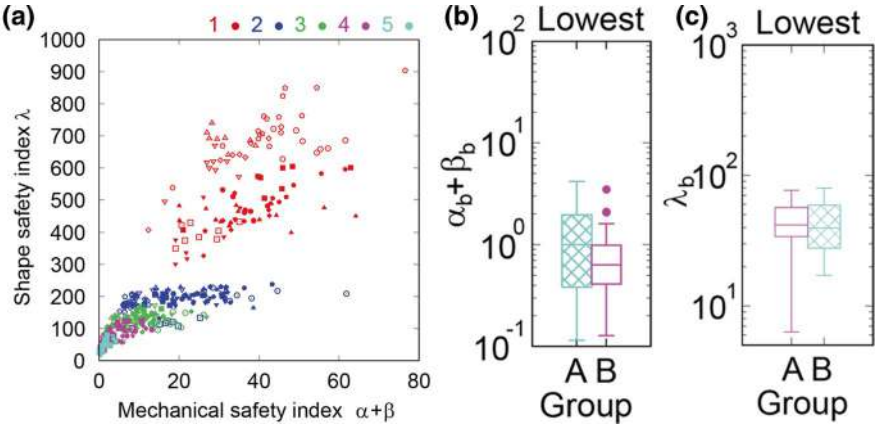


Fig. 9.4 (**a**) Quantitative analysis results showing the morphological safety index and the mechanical safety index. (**b**) Morphological safety index of the internode closest to the ground. (**c**) Mechanical safety index of the internode closest to the ground. (Adapted from Tsugawa et al. 2023, License CC BY)

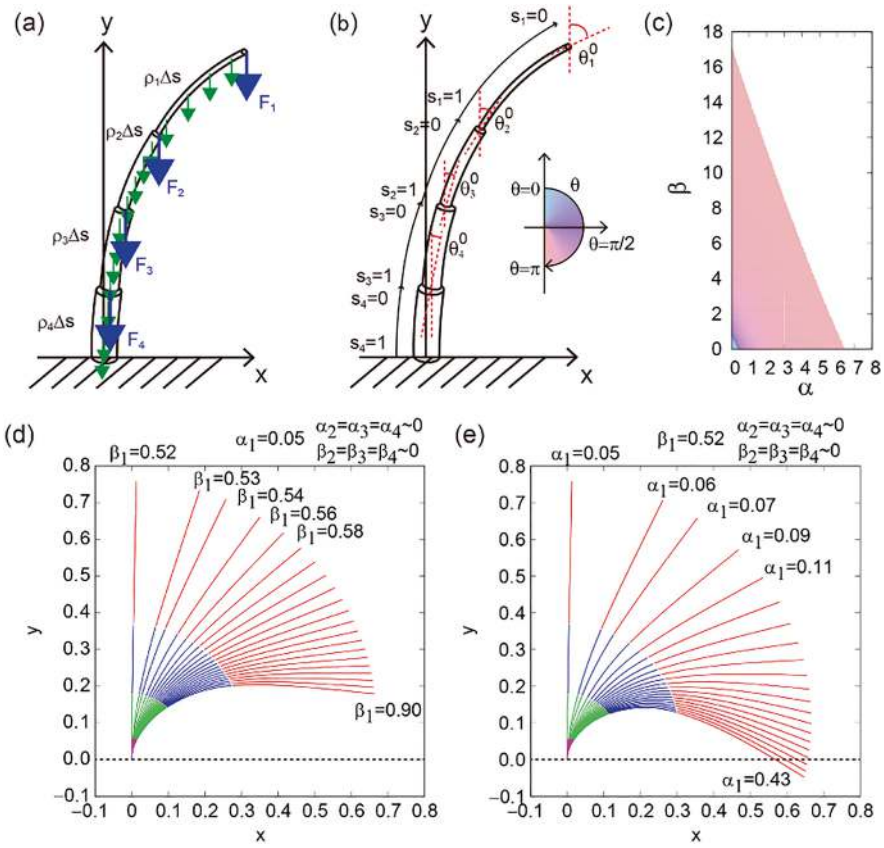


Fig. 9.5 Inference of deflected morphology of rice culms derived from the elastic column theory compared to actual data. (**a**) Force components on each internode. (**b**) Definition of angle for each internode. (**c**) Inference of deflected morphology by the elastic column theory as a function of α and β. (**d**, **e**) Resulting shapes of the theoretical model as a function of the index β_1 (**d**) and the index α_1 (**e**). (Adapted from Tsugawa et al. 2023, License CC BY)

β_i (Fig. 9.5c–e). This equation can be solved approximately, allowing investigation of which parameters most influence lodging resistance. Analysis revealed for the first time that the internode closest to the panicle, which is most closely related to panicle weight, is most effective in determining lodging resistance to self-weight, followed by internode length, stiffness, and outer diameter.

The achievement of this research enables easy comparison of lodging resistance among different varieties by calculating the morphological safety index λ_i and the mechanical safety indices α_i, β_i using actual measurement values. Furthermore, comparisons between varieties also aid in identifying the causes of lodging for each variety. Furthermore, elastic column theory revealed the critical importance of the internode closest to the panicle. Therefore, improving the characteristics of this internode is expected to lead to the development of rice varieties with enhanced

lodging resistance, contributing to solving food issues. Going forward, establishing a foundation for mechanical theories related to crops is anticipated to contribute to agricultural science, such as plant breeding.

9.5 Net-like Pattern on Muskmelon Rind

Muskmelon is a general term for melons possessing a strong, musk-like fragrance. The fruit is nearly spherical, and its rind develops a cork-like net pattern. The flesh is yellow-green, juicy, and exceptionally sweet. It is thus representative of high-grade melons. Cultivation of muskmelon is extremely difficult; strict temperature and humidity control, along with meticulous care, are essential to produce high-quality fruit. Furthermore, only one melon is allowed to grow per vine. Consequently, prices are high, with premium specimens sometimes exceeding a few hundred dollars per muskmelon.

An interesting fact is that the market value of muskmelons, particularly in Japan, strongly depends on the fineness of the net-like pattern on their rind. The surface of a muskmelon is covered in a white net-like pattern. This net is formed by suberin, an oily polymer substance produced by the plant. Since suberin is one of the main components of cork, the net pattern on melons made of suberin is also called "cork cells." This suberin net pattern is a remnant of cracks that occurred during the melon's growth. In fact, early-stage muskmelons lack this surface net. After pollination, muskmelons grow to about the size of a chicken egg in roughly 1 week, but at this stage, their surface is still smooth and netless (see Fig. 9.6a). About 2 weeks after pollination, the melon's skin suddenly hardens, entering the hardening stage, where its mechanical flexibility is lost. Even after entering this hardening stage, the inner flesh of the melon continues to expand (enlargement growth). As a result, the skin's growth rate cannot keep up with the flesh's expansion, causing fine cracks to begin appearing one after another on the skin. When cracks appear on the surface, a liquid substance called suberin is secreted from within. This suberin fills the cracks, preventing water evaporation. The suberin secreted into these gaps hardens into a cork-like substance, forming the distinctive net pattern on the surface of the muskmelon.

As explained above, the net-like pattern on muskmelons results from the rind fragmenting into fine pieces due to the expansion of the flesh. In other words, it can be viewed as a phenomenon where spherical shell-like structures undergo successive (layered) fractures under internal pressure. Similar fracture processes can be observed in many natural phenomena beyond muskmelons. For example, the cracks seen in dried starch paste (Akiba et al. 2017; Akiba and Shima 2019) or in aged molten lava (Akiba et al. 2021a, b) occur because drying/cooling creates a contracting force within the medium, and cracks form when this force becomes too great to withstand. When examining these natural crack patterns from a statistical perspective, common underlying statistical laws are often revealed. From this perspective, examining the pattern of melon netting raises the question of whether statistical laws can be identified there. However, since very few studies have

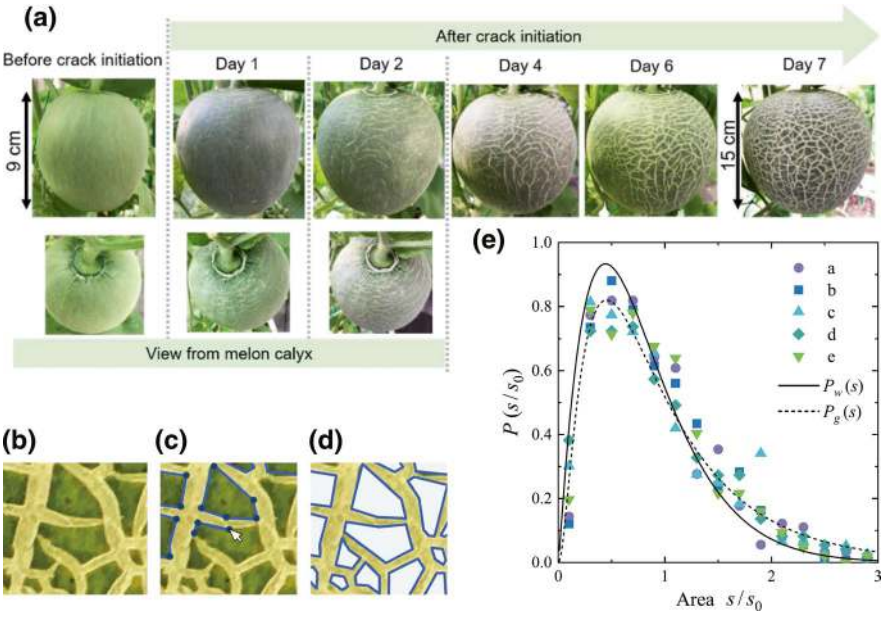


Fig. 9.6 (**a**) Changes in melon surface as it grows. After fine cracks spread across the entire surface, a net-like pattern develops over the entire surface 7 days later. (**b–d**) Area evaluation of melon rind fragments using image analysis. (**b**) Magnified photograph of a muskmelon rind. The corky mesh (light green) surrounding small rind fragments (dark green) is visible. (**c**) The computer identifies the boundaries (blue lines) of the peel fragments. (**d**) The area enclosed by the blue lines (light blue) is calculated. (**e**) Probability distribution of melon rind fragment areas. The horizontal axis s/s_0 represents the fragment area s divided by the scaling constant s_0. The vertical axis represents the percentage of total fragments with a given area s. As seen in the graph, the fragment area distributions for the five melon specimens (represented by different symbols) follow a nearly identical distribution curve (solid and dotted lines). (Adapted from Akiba et al. 2022, License CC BY)

investigated the geometric properties of melon netting, no clear conclusion could be reached.

Against this backdrop, a Japanese research group tried to explore the correlation between the geometric characteristics of the net-like pattern on melon rinds and fruit quality (Akiba et al. 2022). The research group took photographs of several muskmelons sold at a retail store (for quality assurance) and investigated the shape and size of the small green peel fragments separated by white reticular stripes (see Fig. 9.6b–d). Calculating the areas of these small peel fragments, obtaining the probability distribution curves of the areas, and scaling them, it was found that the distribution curves were on a common curve regardless of the type and size of individual melons.

Figure 9.6e shows the distribution data of fragmented areas obtained from the experiment. The horizontal axis s/s_0 represents the fragment area s divided by the

scaling constant s_0. The vertical axis of the graph represents the percentage of total fragments with a given area. As seen in the graph, the fragment area distributions for the five melon specimens follow nearly identical distribution curves (solid and dotted lines in the graph). This result implies that the seemingly random net patterns on each melon actually follow a common scientific law. Furthermore, the research team derived the specific mathematical expression for this distribution curve using fracture mechanics theory. By extending the expansion fracture theory for one-dimensional brittle rings to two-dimensional brittle spherical shells, they demonstrated that this distribution curve can be described by a Bessel function of the second kind. In other words, they proved that the law governing the muskmelon net pattern, experimentally obtained in this study, is not mere coincidence but a result grounded in mechanical theory.

It is empirically known that the appearance of this mesh pattern correlates with the quality of melons and is used as an important indicator for quality evaluation at production sites. In fact, the quality of mask melons is primarily evaluated by their appearance. Furthermore, the price difference between melons of different grades is significantly larger compared to other fruits. Therefore, from the perspective of producers and sellers, it is necessary to accurately assess the deliciousness inside without cutting the melon, relying solely on its appearance. Currently, melon sorting is performed visually by experts such as producers and market professionals. However, relying solely on such artisan skills carries risks: labor shortages may arise, and subjective judgment could influence grade determination. Furthermore, there is an inherent limit to the number of muskmelons that can be evaluated through visual inspection by humans.

Developing the findings from this research could offer a potential solution to these challenges. While this study focused solely on commercially available muskmelons (melons whose quality is guaranteed by their appearance), examining the appearance of melons discarded during cultivation or sorting processes could similarly reveal the statistical properties of the net pattern and its correlation with the deliciousness of the melon flesh. Furthermore, clarifying the relationship between the melon's growth process and the accompanying formation of the net pattern could lead to technology enabling the automatic and objective determination of melon ripeness through image recognition.

9.6 Summary

This chapter introduced several research examples that applied theoretical physics and structural mechanics theory—often focused on artificial structures—to plants that have survived the harsh natural environment of survival of the fittest, thereby formalizing the wisdom hidden within their morphology. Applying insights from physics and structural mechanics to plants does more than merely yield engineering knowledge useful for manufacturing. Shifting the focus of theory from artificial

objects to plants drives the development of the theory itself, enabling the elucidation of the mechanical rationality within more complex systems. Furthermore, compared to traditional botanical research, which primarily relies on narrow, deep experimental approaches focused on specific species, a theoretical approach based on mechanical models that extract shape characteristics common to plants enables us to touch upon the "essence of wisdom" that plants have cherished throughout their evolutionary process of changing form, including the "rational resistance mechanism to gravity that acts equally on all plants." This interdisciplinary approach transcends existing academic frameworks to fuse botany and engineering, laying the groundwork for pioneering new research domains. Moreover, through plants that have coexisted with nature since ancient times, we may even learn about next-generation engineering practices that live in harmony with nature.

References

Akiba Y, Shima H (2019) Flow velocity-dependent transition of anisotropic crack patterns in $CaCO_3$ pastes. J Phys Soc Jpn 88:024001. https://doi.org/10.7566/JPSJ.88.024001

Akiba Y, Magome J, Kobayashi H, Shima H (2017) Morphometric analysis of polygonal cracking patterns in desiccated starch slurries. Phys Rev E 96:023003. https://doi.org/10.1103/PhysRevE.96.023003

Akiba Y, Takashima A, Inoue A, Ishidaira H, Shima H (2021a) Geometric attributes of polygonal crack patterns in columnar joints, Earth and Space. Science 8:2020EA001457. https://doi.org/10.1029/2020EA001457

Akiba Y, Takashima A, Shima H (2021b) Universal fluctuation of polygonal crack geometry in solidified lava. Phys Rev E 104:025009. https://doi.org/10.1103/PhysRevE.104.025009

Akiba Y, Ishibashi A, Sato M, Shima H (2022) Empirical rule of fruit rind fragmentation in muskmelon epidermis netting. J Phys Soc Jpn 91:104801. https://doi.org/10.7566/JPSJ.91.104801

Horiuchi A (2022) Mechanical characteristics of butterbur petiole cross-section with horseshoe shape. Master's thesis, Hokkaido University

Inoue A, Shima H (2018) Survival time of Princess Kaguya in an air-tight bamboo chamber. Forma 33:1–5. https://doi.org/10.5047/forma.2018.001

Sato M, Inoue A, Shima H (2017) Bamboo-inspired optimal design of functionally graded hollow cylinders. PLoS One 12:e0175029. https://doi.org/10.1371/journal.pone.0175029

Shima H, Sato M, Inoue A (2016) Self-adaptive formation of uneven node spacings in a wild bamboo. Phys Rev E 93:022406. https://doi.org/10.1103/PhysRevE.93.022406

Shima H, Furukawa N, Kameyama Y, Inoue A, Sato M (2020) Cross-sectional performance of hollow square prisms with rounded edges. Symmetry 12:996. https://doi.org/10.3390/sym12060996

Shima H, Inoue A, Sato M (2023) Bamboo: a mechanically optimum design in nature. In: Palombini FL, Nogueira FM (eds) Bamboo science and technology. Springer, Singapore. https://doi.org/10.1007/978-981-99-0015-2_1

Tsugawa S, Shima H, Ishimoto Y, Ishikawa K (2023) Thickness-stiffness trade-off improves lodging resistance in rice. Sci Rep 13:10828. https://doi.org/10.1038/s41598-023-37992-3

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 10
Spiral Patterns and Morphology in Chemical Systems



Oliver Steinbock, Pamela Knoll, and Amrutha S.V.

Abstract Far-from-equilibrium chemical systems generate diverse self-organized spatial and temporal patterns that defy intuition and often show remarkable parallels to biological organization. This chapter surveys major examples, beginning with precipitation-driven chemical gardens that together with inorganic biomorph structures illustrate how nonequilibrium processes can produce permanent architectures. A different class of examples includes excitation waves in reaction–transport systems such as chemical waves in the Belousov–Zhabotinsky reaction, action potentials in the human heart, and certain types of surface corrosion. Many but not all of these systems generate spiral-shaped and helical patterns, which form a central focus of our discussion. Together, these phenomena reveal how simple physical and chemical principles can orchestrate rich pattern formation across scales and vastly different chemistries, linking the behavior of inorganic reactions with mechanisms familiar from living systems.

Keywords Chemical self-organization · Reaction–diffusion waves · Spiral and scroll waves · Chemical gardens · Corrosion waves · Biomorphs · Far-from-equilibrium systems

O. Steinbock (✉) · A. S.V.
Department of Chemistry and Biochemistry, Florida State University, Tallahassee, FL, USA
e-mail: osteinbock@fsu.edu

P. Knoll
UK Centre for Astrobiology, School of Physics and Astronomy, University of Edinburgh, Edinburgh, Scotland, UK

© The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4_10

10.1 Introduction

Life unfolds at macroscopic length scales but is driven by microscopic processes that originate in chemistry. It also evolves toward ever more intricate forms while the universe progresses toward an unstructured heat death. These fundamental tensions are often overlooked in modern biological research, which tends to focus on molecular assemblies, genetic control, and biochemical pathways. Yet an entire research area is built around this tension: the study of chemical self-organization far from thermodynamic equilibrium.

The origins of dissipative pattern formation extend back to the nineteenth century and, in spirit, even to the era of alchemical investigation. In those early centuries, minerals and ores were commonly believed to grow underground in a plant-like fashion and at times to emerge at the surface as new crystalline deposits. This belief was strengthened by the appearance of life-like accretions and extrusions on mine shafts and weathered rock faces. Within this historical context lies the work of Johann Rudolf Glauber (1604–1670) on what we now call chemical gardens, one of the inorganic yet distinctly organism-like systems discussed in this chapter.

Some two centuries later, spatiotemporal patterns such as oscillations and traveling waves became clearly recognizable in electrochemistry, in particular in experiments by Gustav Theodor Fechner who studied the anodic dissolution of nickel in nitric acid as well as in related work by Wilhelm Ostwald on chromium dissolution (Alkire and Kolb 2002). Electrochemical interfaces provide a natural setting for nonlinear feedback because local reaction rates depend sensitively on potential, surface coverage, and mass transport. Under galvanostatic or potentiostatic control, this interplay can drive the electrode surface into periodic or even chaotic dynamics. These regimes allow detailed studies of coupled oscillators, synchronization phenomena, and the occurrence of chimera states. Dynamically similar processes synchronize neurons in the suprachiasmatic nucleus, a region of the vertebrate brain, to form the master circadian clock.

An even more prominent example of self-organized chemical behavior is the Belousov–Zhabotinsky (BZ) reaction, which has become the canonical system for chemical oscillations and for the spontaneous emergence of macroscopic wave patterns (Pojman and Epstein 1998). The mechanisms that generate these dynamics rely on nonlinear feedback and on the coupling between reaction and transport, concepts that have become central to modern systems biology. Many of the spatial and temporal patterns observed in the BZ reaction reappear throughout biology, from single cells and microbial colonies to tissues and coordinated groups of organisms.

A notable example is the rotating spiral wave that emerges in excitable and oscillatory media. In the Belousov–Zhabotinsky reaction, these spirals appear to the unaided eye as blue, chemically oxidized bands that move across a red, reduced background. They exhibit a pitch of roughly 1 mm and rotation periods on the order of 1 min. In the language of activator and inhibitor kinetics, such waves arise when the inhibitor diffuses more slowly than the activator. The opposite regime, in which

the inhibitor diffuses more rapidly, can give rise to stationary patterns in related systems such as the CIMA reaction, including stripes, zigzags, and hexagonal arrays of dots (Horváth et al. 2009). These structures have long been discussed in a biological context, tracing back to Alan Turing's 1952 paper on "The Chemical Basis of Morphogenesis," but were discovered in chemical experiments only 40 years later (Turing 1952).

This chapter reviews several families of far-from-equilibrium chemical systems, with occasional reference to biological analogues when space permits or when conceptual connections are helpful. Specifically, we outline the following topics: (1) chemical gardens and current research directions, (2) excitation waves in different spatial dimensions, (3) corrosion waves and their technological implications, and (4) open questions related to inorganic biomorphs. Despite their diverse molecular compositions and physicochemical mechanisms, these systems share often recurring features: geometric spirals, nonlinear interactions, and origins that lie far from equilibrium.

10.2 Chemical Gardens

Our first example concerns inorganic patterns that form during precipitation reactions, i.e., chemical processes that produce solid products from solutions. In the classic version of the chemical garden experiment, small salt particles are placed into sodium silicate solutions (Barge et al. 2015). As the crystal begins to dissolve, the metal ions form insoluble hydroxides in the strongly basic medium. These colloidal metal hydroxide particles aggregate rapidly and produce a thin, continuous membrane around the crystal, effectively compartmentalizing the reaction system.

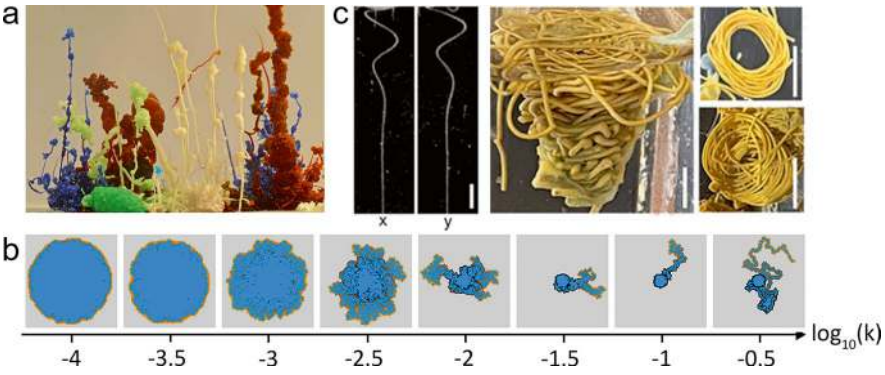


Fig. 10.1 Pattern formation in chemical gardens. (**a**) Centimeter-scale structures formed from different salt crystals in sodium silicate solution. (**b**) Simulated precipitate structures for different age biases *k*. Gold and black regions show the precipitate; gray and blue regions are the two different reagents. (**c**) Helical chemical garden tubes in an injection experiment. All scale bars: 5 mm. (Adapted with permission from Barge et al. 2015; Batista et al. 2023; Siev et al. 2024)

The freshly formed membrane is semipermeable and allows water to flow inward toward the crystal through osmotic pressure. This flow quickly ruptures the fragile inorganic membrane and causes the ejection of a buoyant jet of salt solution around which additional precipitate is templated. The net result is the formation of hollow macroscopic structures, typically straight tubes, branched tubes, and shell-like objects (Fig. 10.1a).

The chemistry that produces these precipitate tubes is extremely diverse and requires primarily a rapid precipitation reaction or a comparable solidification process. Additional requirements are less well established but likely include a conducive surface charge of the precipitate particles and favorable mechanical characteristics of the aging membrane material. Accordingly, chemical gardens have been grown from numerous inorganic salt seeds in solutions containing silicates, borates, phosphates, and other anions. The resulting tubes possess wall thicknesses on the order of a few microns and often consist of porous metal hydroxide with a thin outer layer of amorphous silica.

Today, the systematic investigation of chemical gardens relies heavily on experiments in which the salt seed is replaced by a salt solution that is steadily injected into the silicate solution. This approach was first demonstrated by our group in 2003 and made it possible to identify several distinct growth regimes (Thouvenel-Romans and Steinbock 2003). The most striking cases are the following: (1) jetting, in which an open-tipped smooth tube grows around a steady jet of solution that exits from the tip, (2) popping, in which the tube is mostly closed and expands by the periodic release of balloon-like membrane segments, and (3) budding, in which these segments remain attached to the main stalk and new growth emerges from local breach sites. Other experimental approaches employ spatial constraints, with most work focusing on the injection of salt solutions into Hele–Shaw cells or microfluidic channels.

From the perspective of pattern formation, the structures that emerge in the thin gaps of Hele–Shaw cells are of particular interest. First reported by Haudin et al. in 2014, these patterns display remarkable morphological variability while remaining reproducible with respect to their key geometric characteristics (Haudin et al. 2014). Surprisingly, this structural repertoire includes membrane segments that trace logarithmic spirals, i.e., spirals with an exponentially increasing pitch.

These unusual precipitate spirals do not form simultaneously in the Hele–Shaw cell but arise from preexisting membranes that experience a breach and are pushed outward and deformed. We speculate that during this process, new membrane material forms near the breach site while the older sections are pushed outward and bend under the friction imposed by the Hele–Shaw plates. Another intriguing result is that phyllotactic spiral patterns can develop in radially growing systems that maintain an intrinsic wavelength between neighboring spots (Facchini et al. 2025). This behavior has been demonstrated through numerical models of phase transitions and Turing structures and through experiments on chemical precipitation.

The transition of patterns from near circular precipitation fronts to irregular labyrinthine structures and eventually to filamentous conduits, which are close analogs of three-dimensional tubes, can be modeled with an unconventional cellular

automaton framework (Batista et al. 2023). In this approach, each lattice site contains one of the two reactants or the precipitate, and the injection of fresh reactant randomly replaces precipitate to generate an expanding circular front. When the model includes an age bias that preferentially removes newly formed precipitate, thin-walled filaments appear and elongate at their tips, closely reproducing the self-extending structures observed experimentally (Fig. 10.1b). The addition of a simple buoyancy rule further enables the model to capture both branched and unbranched chemical garden architectures in two and three dimensions. Taken together, these results show that temporally evolving membrane properties and biased self-healing processes play a central role in the formation and diversification of chemical garden morphologies.

Another spiral motif in unconstrained, three-dimensional chemical garden experiments is the formation of helical tubes (Fig. 10.1c) (Siev et al. 2024). Recent work by our group showed that certain chemical garden tubes create a colloidal "exhaust" above their open tip. This suspension flows upward and triggers the formation of a secondary tube that coils into a helix between the solution meniscus and its freely suspended base, which continues to produce new tube material.

10.3 Excitable Reaction–Diffusion Systems

In certain reaction–diffusion systems, spirals become highly dynamic entities. To understand them more fully, it is useful to begin with excitation waves in one spatial dimension. For systems such as neurons or the chemical Belousov–Zhabotinsky reaction, this domain possesses a stable homogeneous rest state (Murray 2003). Small local or global perturbations of this steady state lead to an unremarkable decay back into the rest state. If, however, a system-specific threshold is exceeded, the system undergoes a long excursion in phase space before returning to the steady state. The initial phase of this excitation cycle is often driven by an autocatalytic process, such as the rapid self-amplifying production of $HBrO_2$ in the BZ reaction (Epstein and Showalter 1996) or, in neurons, the sudden depolarization caused by voltage-gated ionic currents (FitzHugh 1961). If the perturbation is local, this entire process triggers a traveling pair of excitation pulses that effectively spatially unfold the temporal dynamics. Accordingly, the front of the steadily moving pulse tends to be steep, followed by a slower (often exponential) relaxation in the wave back (Dockery et al. 1988). The overall pulse shape and speed are conserved. During this relaxation, the system is locally refractory, meaning that new pulse initiations are difficult or impossible. A direct consequence of this refractory zone is that colliding pulses annihilate like forest-fire fronts, rather than undergoing interference.

In two space dimensions and for isotropic systems, a local perturbation triggers a steadily expanding circular wave (Zhabotinsky and Zaikin 1973). Regardless of whether planar or circular, if such an excitation front is disrupted, some systems will respond with the creation of a pair of counter-rotating spiral waves (Winfree 1972; Müller et al. 1985). The basic mechanism behind these intriguing vortex patterns is a slowdown of the front propagation at the highly curved spiral tip (Mikhailov and

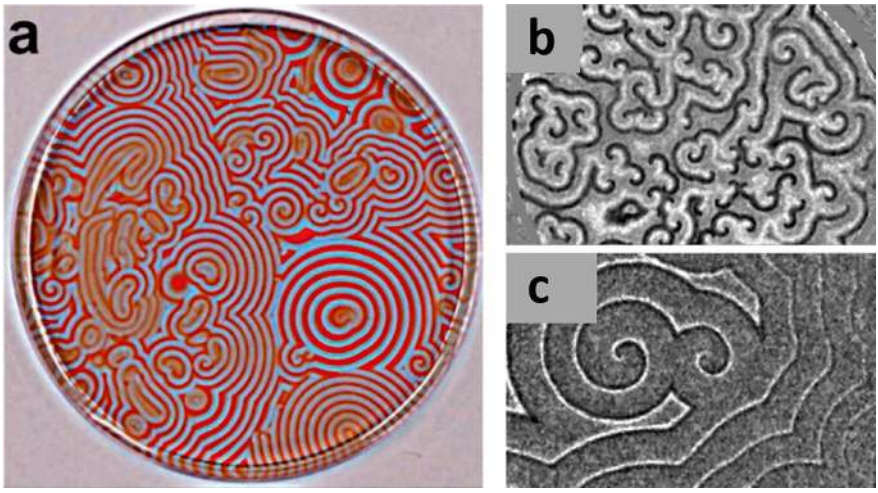


Fig. 10.2 Rotating spiral waves in excitable reaction–diffusion systems. (**a**) Chemical waves in the BZ reaction and (**b**, **c**) in the cellular slime mold *Dictyostelium discoideum*. (Adapted with permission from Knoll et al. 2024; Weijer 2019)

Zykov 1991). At a molecular level, this deceleration is due to the geometric dilution of the diffusive flux of the activator species. The resulting spirals have a system-specific rotation period and wavelength that in the BZ reaction are typically of the order of seconds to minutes and millimeters, respectively (Fig. 10.2a). Detailed studies of the spiral tip revealed that its motion can trace simple circles as well as hypo- and epicycles (Barkley 1994).

Over the past decades, spiral waves of excitation have been observed in numerous experimental systems. The earliest examples include aggregation waves in the cellular slime mold *Dictyostelium discoideum* (Fig. 10.2b) that use traveling waves of cAMP concentration to organize their transition from single amoebae to a multicellular pseudoplasmodium and subsequently spore-bearing fruiting body (Siegert and Weijer 1989). For these cells, excitation waves and coupled chemotaxis solve the problem of providing directional cues over centimeter-scale distances. We emphasize that a simple release of a passively diffusing chemoattractant could not replace this clever mechanism as it would lead to either shallow concentration gradients that fail to give directional guidance or a hopeless oversaturation of cAMP receptor near the aggregation target. Surprisingly, the spiral waves in *Dictyostelium* can be seen with the unaided eye as moving and resting cells have different heights and lead to subtle yet noticeable deformations of the thin water film that covers them.

Another important example is the CO oxidation on Pt surfaces. These studies, which contributed to Gerhard Ertl's 2007 Nobel Prize in Chemistry, are typically performed by photoelectron emission microscopy (PEEM) (Ertl 2008; Jakubith et al. 1990). By monitoring the local work function, this method images the

spatiotemporal evolution of the surface's CO coverage. PEEM pictures and movies revealed that this CO coverage is not homogeneous but patterned by rotating spiral waves and other patterns.

10.4 Vortices in Higher Dimensions

A fascinating extension of spiral wave dynamics is the emergence of their three-dimensional counterparts (Winfree 2001). For topological reasons, wave surfaces in three dimensions cannot rotate around points or small point-like regions but must instead be organized by one-dimensional curves. This idea is easiest to visualize for a vertical stack of identical two-dimensional spirals that together form a scroll-like surface. The rotation backbone of this simplest scroll wave is a straight line that is usually referred to as the filament of the vortex pattern. The phase of the constituent spirals can vary along this filament, which then induces a twisted scroll wave. In addition to such a twist, the filament can deviate from a straight line and follow smooth curves, leading to curvature of the filament itself.

The possible variations of this scroll wave motif are tightly controlled by topological constraints. Most importantly, the filament cannot end within the excitable volume. It must either terminate on the system boundary or curl back on itself to form a closed loop or knot, giving rise to a subtype known as scroll rings. Even for the simplest case of a circular filament of radius R, the filament is not static but contracts or expands depending on the system-specific filament tension α. For positive filament tension, the filament loop collapses slowly according to the simple equation: $R(t) = \sqrt{R_0^2 - 2\alpha t}$, where t and R_0 are time and the initial radius, respectively (Keener and Tyson 1992; Totz et al. 2015). For negative values of α, the scroll ring filament expands but also buckles and quickly folds into shapes showing spatiotemporal chaos (Biktashev et al. 1994). This complex state is known as negative-filament or Winfree turbulence (Alonso et al. 2003) and is not only of intellectual interest but directly related to certain cardiac arrhythmias.

The human heart has four chambers of which the walls of the ventricles are sufficiently thick to require a description as a three-dimensional system. Under normal, healthy circumstances, a natural pacemaker on one of the atria emits an excitation pulse about once per second. This action potential is relayed through the myocardium and channeled via the Purkinje fibers into the ventricles. This spatiotemporal pattern and its rhythmic repetition orchestrate our normal heart function, allowing proper pump action of the contracting organ. However, these excitation pulses can induce spiral waves that in cardiology are known as reentrant arrhythmias (Christoph et al. 2018). In tachycardia, one vortex dominates and causes very high but surprisingly regular heart rates of up to 300 beats per minute. In ventricular fibrillation, this spiral wave undergoes an instability, likely related to negative filament tension, and effectively induces many vortices that typically lead to the death of the affected person within 1 or 2 minutes (DeTal et al. 2022).

Even higher space dimensions can be encountered if the conventional assumption of a continuous space is replaced by a high-dimensional regular network. This scenario was first investigated by Zanotto and Steinbock for four dimensions (Zanotto and Steinbock 2021). Here, the dimensional pairings of domain dimensions and filament dimensions continue and hyperscroll rotation occurs around two-dimensional surfaces. Examples include planar surfaces, spheres, and tori. These surfaces again evolve under the influence of local curvature and a system-specific surface tension, and each carries a local phase whose gradients relax over time. These hyperscroll waves also show some remarkable robustness against the random removal of node connections and even allow for the appearance of hyperscroll turbulence.

10.5 Corrosion Waves

Returning to excitation waves in quasi-two-dimensional systems, a surprising example is found in corrosion science, specifically the behavior of iron alloys under thin layers of acid (Agladze and Steinbock 2000). Most of us think of rust formation as a slow process that unfolds over months and years rather than seconds, but under the appropriate conditions, very fast rust waves can form. The central idea here is the concept of surface passivation. Depending on pH and other parameters, a metal dipped into an acid might either actively corrode or quickly form a thin oxide layer that stops further reactions.

In 1925, Lillie studied such a system in the form of a long steel wire submerged in nitric acid, finding that fast corrosion waves propagated along the otherwise passive wire (Lillie 1925). Later Agladze and Steinbock (2000) also observed rotating spiral waves that appear as dark bands on the grayish steel plates. The reported wave speeds were about 5 mm/s at rotation periods of a few seconds. Wave propagation was also associated with a measurable loss in material height, equaling about 850 nm/wave. Interestingly, the spiral tips deposited thin dark lines onto the steel, indicating different reaction conditions and products. The corrosion waves on steel plates can also be easily manipulated by protecting the surface locally with adhesive tape (Agladze et al. 2001). For long channels constructed by this method, wave propagation failed if the channel width fell below ~0.3 mm, indicating sink-like boundaries for these obstacles.

Corrosion waves also exist in the context of other technologically relevant alloys such as light-weight magnesium and aluminum materials that play important roles in many civilian and military applications. An excellent example is the magnesium alloy AZ31B, which is frequently used in high-end camera and cell phone frames, helicopter parts, and other products that require or benefit from low-density materials. If this alloy is brought into contact with a moderately concentrated saline solution (e.g., 0.6 M NaCl or ocean water), it undergoes corrosion. This corrosion is strikingly visible as surface waves, both by scanning vibrating electrode technique (SVET) and optical inspection, and also associated with the production of streams

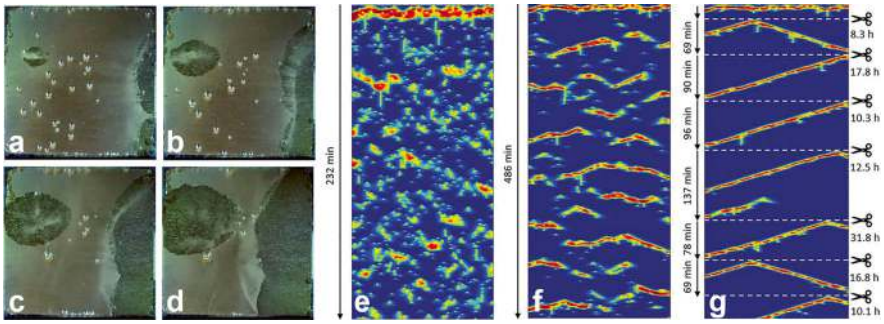


Fig. 10.3 Corrosion waves on the magnesium alloy AZ31B exposed to a 0.6 M solution of NaCl. (**a**–**d**) Experimental photos obtained at a time interval of 1440 s. Field of view: 4 cm^2. (**e**–**g**) Simulated time–space plots showing corrosion waves along the edge of a 1 cm wide AZ31B lamella. The simulations differ in the rate of background corrosion: (**e**) spots for high rates, (**f**) wave segments for medium rates, and (**g**) rare waves for slow rates. Waves are triggered and affected by IMPs in the alloy. (Adapted with permission from Batista et al. 2024a)

of small hydrogen-gas bubbles (Fig. 10.3a–d). The propagation speed of these corrosion waves is much slower than the aforementioned waves on low-carbon steel. Typical values for AZ31B under 0.6 M NaCl solution are 0.1 mm/min.

A fascinating dynamical feature of waves in the AZ31B system is that wave nucleation is tightly coupled to intermetallic particles and clusters (IMPs) in the alloy (Batista et al. 2024a). These IMPs are often crystalline (e.g., $Al_{11}Mn_4$, Al_8Mn_5, and ε-AlMn), have sizes of a few microns or less, and form during alloy production. Once exposed to the solution-facing surface of the metal, they become the main driver of the corrosion process. This results in a vicious cycle of IMP-induced wave nucleation, wave-driven material loss by corrosive dissolution, and the exposure of new IMPs that had been buried in the bulk of the alloy earlier. These dynamics can be captured in nonlinear reaction–diffusion simulations that are scaled and informed by quantitative measurements of IMP-size distributions and other parameters. Once this correspondence between model and material has been established, the framework allows quantitative predictions for the evolution of the surface roughness and the average corrosion rate. Batista et al. showed that these simulations for a single lamella spanning 1 cm of AZ31B can be performed for real corrosion times of up to several weeks, opening opportunities for new alloy protection, alloy design, and model-based recommendations of service schedules (Batista et al. 2024a).

In these pseudo-two-dimensional simulations (active 1D edge of lamella plus lamella depth), the effect of IMPs is not always wave initiation as, depending on the local dynamic state of the alloy surface, IMPs can also trigger or block waves. In addition, variation of an important model parameter, the background corrosion rate, reveals distinct regimes in the spatiotemporal dynamics of the system. These regimes include patchy noise, intermittent corrosion spots, wave segments, sustained waves traveling the entire system width, and rare waves in which the time between new pulse initiation events can approach several days (see Fig. 10.3e–g for example). This rich spectrum of nonlinear dynamics makes the IMP-based

corrosion model an intriguing research subject beyond its immediate technological impact.

Moreover, the close link between corrosion science and nonlinear dynamics in AZ31B behavior suggests new ways to mitigate corrosion damage in this system. We reported experimental evidence that the corrosion waves are subexcitable, meaning that open wave ends do not curl into spirals but instead shrink laterally as they propagate in their normal direction. These withdrawing tips generate shrinking wave segments that eventually disappear. Consequently, a local disruption or breakage of a wave front limits its potential to cause surface damage compared to a continuous circular front that affects the entire surface.

In this context, we proposed the use of local wave breakers and investigated their effectiveness in reducing the spread of corrosion waves and the associated material loss (Batista et al. 2025). This study strategically placed small unexcitable disks on the surface according to simple patterning rules. First, for a constant protected-to-total surface fraction, many small wave breakers performed better than a small number of large ones. Second, certain placement patterns provided superior protection. For instance, randomly placed wave obstacles are a poor choice, whereas wave breakers arranged on triangular lattices typically offer the best performance.

We emphasize that not all types of corrosion follow the themes of the surface-dominated iron and magnesium examples. In aluminum alloys, for example, intergranular corrosion takes the front seat. A good example is the aluminum–magnesium alloy AA5456, which is frequently used in marine applications, for car parts (frames, panels), airplane fuselages, and wing structures, as well as for pressure vessels. AA5456 has a complex grain structure that typically involves highly elongated crystalline domains (Batista et al. 2024b). The boundaries between the grains feature a multitude of β phase particles that strongly promote intergranular corrosion. This process is a very interesting target for modeling efforts and simulations that effectively need to describe front propagation in a stochastic and anisotropic three-dimensional network. Recent studies of this problem have focused on both simple percolation models and cellular automata (di Caprio et al. 2016).

10.6 Biomorphs

Yet another chemical reaction capable of producing spiral and helical structures is the biomorph system (Fig. 10.4) (Kellermeier et al. 2012). The actual experiment giving rise to these patterns is exceedingly simple to perform, involving millimolar aqueous solutions of sodium silicate and barium (or strontium) chloride. As CO_2 diffuses from the air into a thin layer of this solution, small barium carbonate crystals form. The further growth of these crystallites is strongly affected by silicate ions, resulting in micron-sized dumbbell-shaped objects. These globules then nucleate the main biomorph structures that take the shapes of helical worms, double helices, urns, funnels, and leaf-shaped sheets (Knoll and Steinbock 2018).

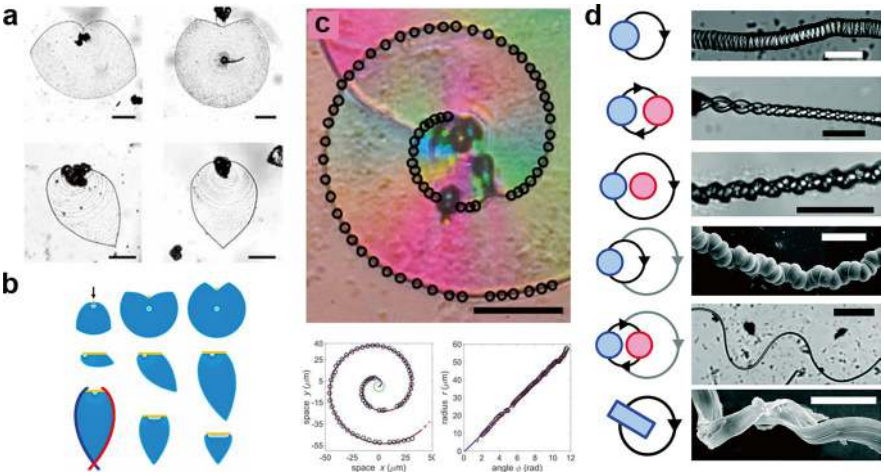


Fig. 10.4 Biomorph structures resulting from the reaction of silicate and barium ions with CO_2 or carbonate ions. (**a**) Leaf-like biomorph sheets. Scale bars: 100 μm (top row), 50 μm (bottom row). (**b**) Reaction–diffusion simulations reproducing the shapes in (**a**). (**c**) Pinned biomorph front tracing an Archimedean spiral. (**d**) Different types of helices and superhelices. Scale bars: 40 μm (top four), 200 μm (fifth panel), and 20 μm (bottom panel). (Adapted with permission from Nakouzi et al. 2016 (panel **d**), Knoll et al. 2017 (panels **a**, **b**), and Knoll et al. 2020 (panel **c**))

Biomorphs form within a few hours and reach maximal lengths of up to 1 mm. However, if the atmospheric CO_2 is replaced by a dissolved carbonate source, growth can last for weeks spawning structures extending for centimeters (Nakouzi et al. 2016). All biomorphs are hierarchically ordered and consist of witherite nanorods of very high aspect ratios that are co-aligned in the local growth direction of the biomorph structure. As witherite is birefringent, this feature gives rise to typical Maltese-cross patterns if sheets and similar structures are observed under crossed polarizers.

Spirals occur as two variants in the biomorph system. First, the leaf-like sheets at the base of the reaction container have edges that trace logarithmic spirals, i.e., curves that in polar coordinates (r, φ) obey $r = a \exp(b\,\varphi)$ (Fig. 10.4a). Two of us showed that this robust feature can be explained by the dynamics of a nonlinear reaction front that once disrupted retreats laterally at a constant speed, analogous to the excitation wave behavior in subexcitable media (Fig. 10.4b) (Knoll et al. 2017).

Second, we observed that biomorph sheets that are typically 1 μm thick can form additional layers (Knoll et al. 2020). This process occurs when a sheet front encounters another globule and effectively splits into two branches that slowly engulf the obstacle. Behind the globule, one crystallization front can jump onto its counterpart and induce spiral rotation around the obstacle. These pinned vortices last for up to three rotations and trace Archimedean spirals, i.e., spirals of constant pitch λ according to $r = \lambda\,\varphi/(2\pi)$ (Fig. 10.4c). The typical pitch values were found to be 20–40 μm.

The nucleation of these crystallization vortices is interesting as it requires a symmetry-breaking mechanism that is not understood. In addition, the existence of

logarithmic and Archimedean spirals in the same system raises the question whether pinned spirals can be stable in subexcitable media. We reported results using Barkley's reaction–diffusion model that answered this question in the affirmative (Knoll et al. 2020). However, front pinning under subexcitable conditions does require sufficiently large anchors (here the globules) and the needed size increases as the system's parameters depart further from the "normal" excitable regime.

In addition to these two types of spiral shapes, biomorphs also form single and various types of double helices (Knoll et al. 2017). These structures show a highly robust geometry. Their mean double-helix wavelength of 19 ± 3 μm and width of 9.6 ± 0.8 μm change by less than 12% across three orders of magnitude in carbonate concentration. The single helices, also referred to as worms, appear to follow a dynamically distinct mechanism that is possibly related to the aforementioned pinned biomorph spirals. Double helices exist in different forms, including superhelices, shown in Fig. 10.4d.

10.7 Conclusions

We have discussed a broad range of systems in which far-from-equilibrium conditions give rise to nonintuitive temporal rhythms and spatial structures. Although most examples in this chapter were abiotic and often inorganic, their qualitative and sometimes deeper links to living systems are apparent: the grain structures of alloys resemble the cellular textures of tissues, the waves in the BZ reaction connect to cardiac arrhythmia, and chemical gardens provide an experimental model for possible origin-of-life processes in hydrothermal vents (an interesting topic we did not address here).

In many systems, traveling waves and spirals occur that biological systems routinely exploit to relay information over macroscopic distances. This ties our research directly to current activities in chemical and otherwise unconventional computing and the question how simple chemical systems can make decisions and find optimal solutions (Steinbock and Zhu 2025). At the same time, we encountered many system-specific differences, including the intriguing absence of spiral waves in the AZ31B system, which may offer pathways toward system-level strategies for corrosion mitigation.

Discovering such similarities across extremely different and seemingly unrelated systems is a major driving force of much work in this field, which spans applied mathematics, chemistry, biology, and engineering. Given that these universalities often manifest themselves as spatiotemporal patterns, it will be interesting to see how advances in machine learning and artificial intelligence might further accelerate progress. If this promise comes to fruition, a new renaissance of a unified natural science may emerge, overcoming the artificial boundaries that scientists erected during the past century.

Acknowledgments This work was supported by the National Science Foundation under grant no. 2515818.

References

Agladze K, Steinbock O (2000) Waves and vortices of rust on the surface of corroding steel. J Phys Chem A 104:9816–9819. https://doi.org/10.1021/jp002237n

Agladze K, Thouvenel-Romans S, Steinbock O (2001) Electrochemical waves on patterned surfaces: propagation through narrow gaps and channels. J Phys Chem A 105:7356–7363. https://doi.org/10.1021/jp011294t

Alkire RC, Kolb DM (eds) (2002) Advances in electrochemical science and engineering, vol X. Wiley-VCH, Weinheim

Alonso S, Sagués F, Mikhailov AS (2003) Taming Winfree turbulence of scroll waves in excitable media. Science 299:1722–1725. https://doi.org/10.1126/science.1080207

Barge LM et al (2015) From chemical gardens to chemobrionics. Chem Rev 115:8652–8703. https://doi.org/10.1021/acs.chemrev.5b00014

Barkley D (1994) Euclidean symmetry and the dynamics of rotating spiral waves. Phys Rev Lett 72:164–167. https://doi.org/10.1103/PhysRevLett.72.164

Batista BC, Morris AZ, Steinbock O (2023) Pattern selection by material aging: modeling chemical gardens in two and three dimensions. Proc Natl Acad Sci USA 120:e2305172120. https://doi.org/10.1073/pnas.2305172120

Batista BC, Romanovskaia E, Romanovski V, Emmanuel M, Burns JT, Ma J, Kiss IZ, Scully JR, Steinbock O (2024a) Morphogenic modeling of corrosion reveals complex effects of intermetallic particles. Adv Sci 11:2404986. https://doi.org/10.1002/advs.202404986

Batista BC, Shabana MA, Jeyamohan R, Harris ZD, Burns JT, Ma J, Scully JR, Kiss IZ, Steinbock O (2024b) A facile procedure for simulating grain boundary maps. ACS Appl Eng Mater 2:2361–2368. https://doi.org/10.1021/acsaenm.4c00389

Batista BC, Romanovskaia E, Romanovski V, Scully JR, Kiss IZ, Steinbock O (2025) Wavebreakers in excitable systems and possible applications for corrosion mitigation. Chaos 35:013103. https://doi.org/10.1063/5.0237388

Biktashev VN, Holden AV, Zhang H (1994) Tension of organizing filaments of scroll waves. Philos Trans R Soc A 347:611–630. https://doi.org/10.1098/rsta.1994.0070

Christoph J et al (2018) Electromechanical vortex filaments during cardiac fibrillation. Nature 555:667–672. https://doi.org/10.1038/nature26001

DeTal N, Kaboudian A, Fenton FH (2022) Terminating spiral waves with a single designed stimulus: Teleportation as the mechanism for defibrillation. Proc Natl Acad Sci USA 119:e2117568119. https://doi.org/10.1073/pnas.2117568119

di Caprio D, Stafiej J, Luciano G, Arurault L (2016) 3D cellular automata simulations of intra- and intergranular corrosion. Corros Sci 112:438–450. https://doi.org/10.1016/j.corsci.2016.07.028

Dockery JD, Keener JP, Tyson JJ (1988) Dispersion of traveling waves in the Belousov–Zhabotinskii reaction. Physica D 30:177–191. https://doi.org/10.1016/0167-2789(88)90105-4

Epstein IR, Showalter K (1996) Nonlinear chemical dynamics: oscillations, patterns, and chaos. J Phys Chem 100:13132–13147. https://doi.org/10.1021/jp953547m

Ertl G (2008) Reactions at surfaces: from atoms to complexity (Nobel Lecture). Angew Chem Int Ed 47:3524–3535. https://doi.org/10.1002/anie.200800480

Facchini G, Budroni MA, Schuszter G, Brau F, De Wit A (2025) Phyllotactic structures in radially growing spatial symmetry breaking systems. Phys Rev Lett 135:018001. https://doi.org/10.1103/7tnn-pq49

FitzHugh R (1961) Impulses and physiological states in theoretical models of nerve membrane. Biophys J 1:445–466. https://doi.org/10.1016/S0006-3495(61)86902-6

Haudin F, Cartwright JHE, Brau F, De Wit A (2014) Spiral precipitation patterns in confined chemical gardens. Proc Natl Acad Sci USA 111:17363–17367. https://doi.org/10.1073/pnas.1409552111

Horváth J, Szalai I, De Kepper P (2009) An experimental design method leading to chemical Turing patterns. Science 324:772–775. https://doi.org/10.1126/science.1169973

Jakubith S, Rotermund HH, Engel W, von Oertzen A, Ertl G (1990) Spatiotemporal concentration patterns in a surface reaction: propagating and standing waves, rotating spirals, and turbulence. Phys Rev Lett 65:3013–3016. https://doi.org/10.1103/PhysRevLett.65.3013

Keener JP, Tyson JJ (1992) The dynamics of scroll waves in excitable media. SIAM Rev 34:1–39. https://doi.org/10.1137/1034001

Kellermeier M, Cölfen H, García-Ruiz JM (2012) Silica biomorphs: complex biomimetic hybrid materials from 'sand and chalk'. Eur J Inorg Chem 2012:5123–5144. https://doi.org/10.1002/ejic.201201029

Knoll P, Steinbock O (2018) Inorganic reactions self-organize life-like microstructures far from equilibrium. Isr J Chem 58:682–692. https://doi.org/10.1002/ijch.201700136

Knoll P, Nakouzi E, Steinbock O (2017) Mesoscopic reaction–diffusion fronts control biomorph growth. J Phys Chem C 121:26133–26138. https://doi.org/10.1021/acs.jpcc.7b09559

Knoll P, Wu N, Steinbock O (2020) Spiraling crystallization creates layered biomorphs. Phys Rev Mater 4:063402. https://doi.org/10.1103/PhysRevMaterials.4.063402

Knoll P, Ouyang B, Steinbock O (2024) Patterns lead the way to far-from-equilibrium materials. ACS Phys Chem Au 4:19–30. https://doi.org/10.1021/acsphyschemau.3c00050

Lillie RS (1925) Factors affecting transmission and recovery in the passive iron nerve model. J Gen Physiol 7:473–507. https://doi.org/10.1085/jgp.7.4.473

Mikhailov AS, Zykov VS (1991) Kinematical theory of spiral waves in excitable media: comparison with numerical simulations. Physica D 52:379–397. https://doi.org/10.1016/0167-2789(91)90134-U

Müller SC, Plesser T, Hess B (1985) The structure of the core of the spiral wave in the Belousov–Zhabotinskii reaction. Science 230:661–663. https://doi.org/10.1126/science.230.4726.661

Murray JD (2003) Mathematical biology: II. Spatial models and biomedical applications. Springer, New York

Nakouzi E, Knoll P, Hendrix KB, Steinbock O (2016) Systematic characterization of polycrystalline silica–carbonate helices. Phys Chem Chem Phys 18:23044–23052. https://doi.org/10.1039/C6CP04153A

Pojman JA, Epstein IR (1998) An introduction to nonlinear chemical dynamics: oscillations, waves, patterns, and chaos. Oxford University Press, New York

Siegert F, Weijer C (1989) Digital image processing of optical density wave propagation in *Dictyostelium discoideum* and analysis of the effects of caffeine and ammonia. J Cell Sci 93:325–335. https://doi.org/10.1242/jcs.93.2.325

Siev EA, Batista BC, Steinbock O (2024) Coiling of secondary tubes formed from the colloidal exhaust of primary chemical gardens. J Phys Chem B 128:2028–2036. https://doi.org/10.1021/acs.jpcb.3c07840

Steinbock O, Zhu W (2025) Modeling the Monty Hall decision problem with reaction kinetics. Phys Chem Chem Phys 27:15055–15061. https://doi.org/10.1039/D5CP01722J

Thouvenel-Romans S, Steinbock O (2003) Oscillatory growth of silica tubes in chemical gardens. J Am Chem Soc 125:4338–4341. https://doi.org/10.1021/ja028782a

Totz JF, Engel H, Steinbock O (2015) Spatial confinement causes lifetime enhancement and expansion of vortex rings with positive filament tension. New J Phys 17:093043. https://doi.org/10.1088/1367-2630/17/9/093043

Turing AM (1952) The chemical basis of morphogenesis. Philos Trans R Soc B 237:37–72. https://doi.org/10.1098/rstb.1952.0012

Weijer CJ (2019) Spiral waves of the chemo-attractant cAMP organise multicellular development in the social amoebae *Dictyostelium discoideum*. In: Tsuji K, Müller SC (eds) Spirals and vortices. Springer, Cham. https://doi.org/10.1007/978-3-030-05798-5_10

Winfree AT (1972) Spiral waves of chemical activity. Science 175:634–636. https://doi.org/10.1126/science.175.4022.634

Winfree AT (2001) The geometry of biological time, 2nd edn. Springer, New York

Zanotto FM, Steinbock O (2021) Hyperscroll dynamics: vortices in four-dimensional networks. Chaos 31:053132. https://doi.org/10.1063/5.0044396

Zhabotinsky AM, Zaikin AN (1973) Autowave processes in a distributed chemical system. J Theor Biol 40:45–61. https://doi.org/10.1016/0022-5193(73)90164-1

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Appendix (*Courtesy of the Yamada Science Foundation*)

Abstracts of Invited Talks

The 6th Yamada Symposium on
"Diversity of Biological Patterns and Forms in Nature:

Toward a Comprehensive Understanding"
August 1–3, 2023

Global Front, Tokyo Surugadai Campus,
Meiji University, Japan

Organizing Committee
Chair: Hiroshi Matano (MIMS)
Co-Chair: Hiraku Nishimori (MIMS)

Ken-ichi Nakamura (MIMS)
Hirokazu Ninomiya (MIMS)
Norikazu Saito (The University of Tokyo)
Toshio Sekimura (Chubu University)
Masashi Shiraishi (MIMS)
Tomohiko Yamaguchi (MIMS)
(MIMS: Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University)
Sponsored by Yamada Science Foundation

© The Editor(s) (if applicable) and The Author(s) 2026

P. K. Maini et al. (eds.), *Patterns and Forms in Nature*, Theoretical Biology,
https://doi.org/10.1007/978-981-92-0331-4


PFN2023
6th Yamada Symposium on
"Diversity of Biological Patterns and Forms in Nature:
toward a Comprehensive Understanding"
Tue. 1st - Thu. 3rd August 2023
Venue
Global Front, Meiji University, Tokyo Surugadai Campus
← More Information:
http://www.mims.meiji.ac.jp/index-e.html
Invited Speakers
Philip Ball (Science writer, London, UK)
Haruhiko Fujiwara (The University of Tokyo)
Sylvain Gerber (Muséum National d'Histoire Naturelle, France)
Shigeru Kondo (Osaka University)
Reiko Kuroda (Chubu University)
Philip K. Maini (University of Oxford, UK)
H. Frederik Nijhout (Duke University, USA)
Hiraku Nishimori (Meiji University)
Takao Ubukata (Kyoto University)
Toshio Sekimura (Chubu University)
Hiroyuki Shima (University of Yamanashi)
Oliver Steinbock (Florida State University, USA)
Organizing Committee
Chairperson
Hiroshi Matano (MIMS)
Hiraku Nishimori (MIMS)
Members
Ken-ichi Nakamura (MIMS)
Hirokazu Ninomiya (MIMS)
Norikazu Saito (The University of Tokyo)
Toshio Sekimura (Chubu University)
Masashi Shiraishi (MIMS)
Tomohiko Yamaguchi (MIMS)
Organizers and Cooperation
Meiji Institute for Advanced Study of Mathematical Sciences (MIMS)
MIMS Interdisciplinary Research Project for Mathematical Sciences and Biology
Private University Research Branding Project
Sponsored by Yamada Science Foundation
Meiji Institute for Advanced Study of Mathematical Sciences
Meiji University, Nakano Campus
Web site: http://www.mims.meiji.ac.jp/index.html
E-mail: mims@mics.meiji.ac.jp

1.1 Symmetry Breaking from Physics to Biology

Philip Ball



What is the developmental source of biological form? The question preoccupied natural philosophers from Aristotle to William Harvey to Rudolf Virchow, and it remains debated today. Early ideas assumed that biological form was either already manifest in some way within the embryo—the preformationist idea, famously illustrated by the Dutch mathematician and microscopist Nicolaas Hartsoeker's drawing of the homunculus in a sperm in the seventeenth century – or emerged via the guidance of a kind of life force or soul, as Aristotle asserted: the essentially teleological position called epigenesis. We might now regard both theories as sleights of hand designed to circumvent rather than to explain what is now seen as an example of symmetry breaking.

With the emergence of modern genetics, questions about form in biology became questions about genes: the phenotype of the developing organism was considered to be programmed by the genotype. DNA could in this view be regarded as a kind of informational homunculus: form already existed in the embryo, but in an encoded format, which Erwin Schrodinger famously called the "code-script" of life.

In his approach to the origins of embryogenesis, Alan Turing was thus going against the grain. It may be partly for this reason that his 1952 paper on morphogenesis had little impact for several decades. His approach follows the physical scientist's instinct to express the problem at hand in the simplest possible terms, and then to explore what the minimal requirements are of a model that captures the essential phenomena. Turing pointed out that the question of how an embryo develops and acquires shape and form—bilateral symmetry, say, and the budding of limbs—seemed indeed to be one of spontaneous symmetry breaking.

In this presentation I shall set out to explore how, by expressing the question in this manner, Turing aligned it with what is arguably the major theme of twentieth-century physics.

Perhaps the first intimation of the role of symmetry in physical processes was expressed in 1894 by Pierre Curie in a paper titled "On Symmetry in Physical Phenomena." Here he asked the question: What is the relationship between the physical and the symmetry properties of a system? Curie stated the fundamental principle of symmetry breaking by saying that some phenomena occur when a symmetry is *removed*: he said that "asymmetry is what creates the phenomenon."

In Curie's doctorate in the early 1890s on magnetism, he reported that ferromagnets such as iron will lose their magnetism when heated: this falls to zero at a temperature now called the Curie temperature. The onset of spontaneous magnetization by cooling through the Curie temperature occurs when random orientations of the atomic spins due to thermal motion switch to a collective state of mutual alignment.

This phase transition was studied in the 1920s using the Ising model, a lattice model developed by Wilhelm Lenz and his student Ernst Ising. It was soon appreciated that the same model could be used to represent the transition between the gaseous and liquid states of matter, as investigated in the 1870s by Johannes Diderik van der Waals. Here the presence of a particle at a point in the lattice denotes the liquid, while its absence denotes a gas. Cooling a magnet through its Curie temperature then looks equivalent to cooling a fluid through its critical point, where the uniform symmetry of the supercritical state breaks to permit two distinct phases of differing density.

This connection was arguably the first glimpse of *universality* in physics: the idea that there is a deep parallel between systems that look superficially as though they share nothing in common. It stems from the fact that many-body behavior often does not discriminate between the fine differences in the natures of those bodies: all that matters are the general features of how they interact, such as how many neighbors feel the influence of each body and the dimensionality of the system.

The 1956 proposal by Tsung Dao Lee and Chen Ning Yang that the weak interaction might break left–right symmetry (parity violation), experimentally confirmed by Chien-Shiung Wu that same year in experiments on beta decay, opened the path for the unification of the electromagnetic and weak interactions. The idea was developed in the 1960s by Sheldon Glashow and Abdus Salam, who had to introduce the symmetry breaking by hand. Their theory predicted three massive new bosons, later identified with the W and Z bosons that mediate the weak force, along with the massless photon. In the later 1960s, Steven Weinberg showed how this symmetry breaking could happen spontaneously. The theory was verified by the discovery of the W and Z bosons in CERN's Proton–Antiproton Collider in 1983.

This Nobel-winning work on electroweak symmetry breaking was intimately tied up with the theory of the Higgs field and the Higgs boson. The fundamental insight here came in the late 1950s from theoretical physicist Yoichiro Nambu, who saw an analogy between the theory of superconductivity and quantum field theory. The idea won Nambu the 2008 Nobel prize for "the discovery of the mechanism of spontaneous broken symmetry in subatomic physics." But Nambu's work raised new problems, and it was by resolving them in the early 1960s that Peter Higgs and others concluded that space is permeated by the Higgs field, which exerts a drag on the particles and so affords them mass.

Turing's work on morphogenesis takes on a different complexion when considered against this context. Symmetry breaking is, one might say, nature's way of making many from one: of developing complexity out of simplicity. In this way it offers a kind of order and form for free: a distinction between a *here* and a *there* in what was initially a homogeneous system. Biology, it is now clear, exploits this source of organization to produce structure at scales well beyond the molecular. Turing's mechanism is just one among several now known to feature in the living world at scales ranging from the cellular to the ecological.

The broader question is how the palette of shapes and patterns available from spontaneous symmetry breaking interacts with natural selection. To what extent can evolution adapt and modify Turing structures, for example? Are all such structures necessarily adaptive at all? Or are we too readily tempted, when we see order and regularity in biology, to attribute a function to it? Might some of it, at least, represent nothing more than the intrinsic creative potential in the natural world?

1.2 Modeling Collective Cell Behavior in Development and Disease

Philip K. Maini
Wolfson Centre for Mathematical Biology, Mathematical Institute, Oxford, UK
maini@maths.ox.ac.uk



Synopsis Understanding collective movement is a major and important challenge. In the context of cell movement, we present two case studies: (1) cancer and (2) neural crest (normal development). Using partial differential equation models and hybrid cell-based models, we show how modeling can increase our understanding of collective cell invasion and migration.

Collective movement occurs throughout nature across a multitude of scales, from the long-distance movement of whales, to flocking starlings, to bacterial pattern formation. Understanding what are the hallmarks of collective movement is still an

open question. With colleagues, I have been studying this problem in the context of cell movement and

I will present some results from this work in the context of disease spread and normal development. Understanding how cancer cells invade tissue is an important part of preventing the disease spreading. Several models have been proposed for this. Recently, we proposed a partial differential equation model for this phenomenon in which cells cooperate (Strobl et al. 2020). That is, instead of having a generalist population that can overcome obstacles and invade, we consider two specialist populations. In this context, we consider two barriers to invasion: normal cells and the extracellular matrix, and we consider two cancer cell types: one that produces lactic acid that can kill normal cells and one that produces matrix-degrading enzymes. We analyze the travelling wave behavior of this system via linear stability analysis and numerical simulation. We then analyze in more detail show a simplified submodel, with degenerate cross-diffusion and a continuum of steady states, and find that it poses challenging problems for travelling wave analysis.

We then study normal development in the context of the cranial neural crest. These are cells that leave the neural tube and migrate to form part of the skull. Neural crest cells give rise to many parts of the body and nearly one-third of birth defects arise from abnormal behavior of neural crest cells. Moreover, neural cells are the precursors of melanocytes, which can form melanoma, one of the most aggressive cancers. We use a hybrid cell-based model as a hypothesis-testing and hypothesis-generating tool and, in a long-standing collaboration with experimental colleagues, we have uncovered new biological insights concerning this process. Some of these results will be presented (Giniunaite et al. 2020; Johnson et al. in prep.).

1.2.0.1 References

Giniunaite R et al. (2020) Dev Dynamics, 249:270-280
Strobl MAR et al (2020) Bull Math Biol 82:15

1.3 How Organisms Convert 2D Patterns into 3D

Keisuke Matsuda[1], Hiroki Gotoh[2], Haruhiko Adachi[1], Yasuhiro Inoue[3] & Shigeru Kondo[1]*
[1]:Osaka University, Frontier Biosciences
[2]:Shizuika University, Faculty of Science
[3]:Kyoto University, Faculty of Technology
* skondo@fbs.osaka-u.ac.jp, 81-6-6879-7975





Abstract The question "How are organisms formed?" is one of the central questions in developmental biology. In the past 30 years, embryological experiments and mathematical modeling studies have revealed the mechanisms by which individual genes are expressed in specific regions, the so-called principles of pattern formation. However, those principles deal, for the most part, only with the division of fixed fields into regions. In actual morphogenetic phenomena, pattern formation and the expansion and deformation of the field proceed simultaneously. As such, it is extremely difficult to model such a situation, because the two different elements affect each other.

To circumvent these difficulties, it is useful to study phenomena in which cellular changes and field deformation occur at different times. We believe that the deformations associated with molting in exoskeletal organisms serve this purpose.

Late morphogenesis in exoskeletons occurs through molting. During molting, (1) grooves are formed in the intracellular layer in advance, and (2) large-scale deformation occurs by expanding these grooves. The process (2) is a purely physical process because it occurs in a short time. By measuring the number, depth, and direction of the grooves, the final 3D structure can be calculated from the groove pattern.

On the other hand, process (1) is a biological process in which the position and angle of the wrinkles are determined in a two-dimensional cell sheet, which is a kind of pattern formation. We first examined the beetle horn primordium. Beetle pupae have huge horns, but in larvae the horn primordia are compactly folded in the head. Reproducing the folding of the primordium from serial sections and extending it with physical calculations, the structure of the pupal horn emerged. This indicates that the three-dimensional morphology of the horns is already completed by the folding pattern of the primordium. Furthermore, by examining the wrinkle patterns corresponding to the various parts of the horns, part of the logic behind the creation of the three-dimensional form was revealed.

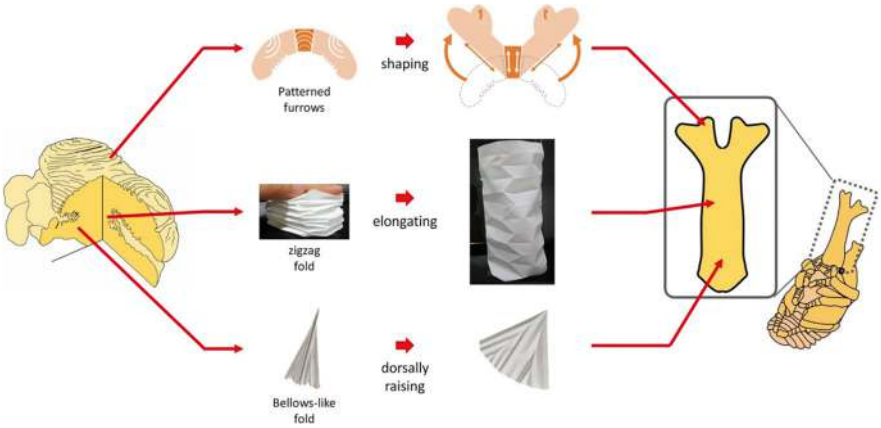
We are currently studying the folding patterns of the helmets of treehoppers, which have more complex and diverse three-dimensional structures, through which we expect to elucidate how the complex structures of insects are designed.

Sequential photos of Japanese beetle pupation. Giant horns emerge in just 100 min. Because of the extremely short time, this deformation is thought to be due to the physical unfolding of the folded cell sheet



Scientific Reports **volume 11, Article number: 1017 (2021)**

Relationship between wrinkle patterns in cell sheets of horn precursors and 3D morphology after unfolding



Zoological Letters **volume 6, Article number: 3 (2020)**
Scientific Reports volume 7(1), Article number: 13939 (2017)

1.4 Mechanisms and Evolution of Female-Limited Batesian Mimicry in *Papilio* Butterflies

Haruhiko Fujiwara
University of Tokyo (Emeritus Professor)
haruh@edu.k.u-tokyo.ac.jp



To deceive predators, palatable species often resemble color patterns and morphology of unpalatable species, which is called Batesian mimicry. Many animals employ Batesian mimicry, whose mechanisms and evolution remain largely unknown. In two closely related swallowtail butterflies, *Papilio polytes* and *P. memnon*, there are two types (mimetic and nonmimetic coloration) females, while males show nonmimetic pattern. The female-limited Batesian mimicry of these butterflies attract public interest from Darwin and Wallace's era. In both species, mimetic phenotype is superior to nonmimetic and known to be controlled by a single autosomal locus (*H* in *polytes* and *A* in *memnon*). Thus, the mimetic female genotype is HH or Hh (AA or Aa) and nonmimetic is hh (aa), while all males with any genotype are nonmimetic. Fisher RA (1930) suggested that this polymorphic female-limited mimicry is regulated by "supergene," which is a cluster of multiple genes tightly linked together.

Recent genome projects revealed many supergenes involved in various complex adaptive phenomena, but the functional unit of supergene has not been clarified in any cases. We have identified the mimicry supergene locus (*H* for *P. polytes* and *A* for *P. memnon*) in a region around *doublesex* (*dsx*) of chromosome 25 (Nishikawa et al. 2015): The region includes several genes, *dsx*, *UXT*, *U3X* (only for *polytes*), and *Nach-like*, which sequences are highly diversified between *H* and *h* (*A* and *a*) due to recombination suppression, caused by chromosomal inversion in *P. polytes* but probably by repetitive sequences without inversion in *P. memnon* (Iijima et al. 2018; Komata et al. 2022). To identify the functional unit of mimicry supergene, we performed knock-down of candidate gene function by electroporation mediated RNAi in mimetic female wings of *P. polytes*. We found that *dsx-H* was responsible for switching from the nonmimetic to mimetic traits, and that not only *UXT* within highly diversified region (HDR) but also *sir2* adjacent to HDR affected the wing

pattern (Komata et al. 2023). Furthermore, it was suggested that *U3X* regulates both *dsx-H* and *UXT* expression. These results indicate that *dsx-H* works as a mimicry gene and other flanking genes as modifier and that functional unit of mimicry supergene is expanded to a wider region than we thought before. In the symposium, I also discuss how two mimicry supergenes in *Papilio* species have evolved.

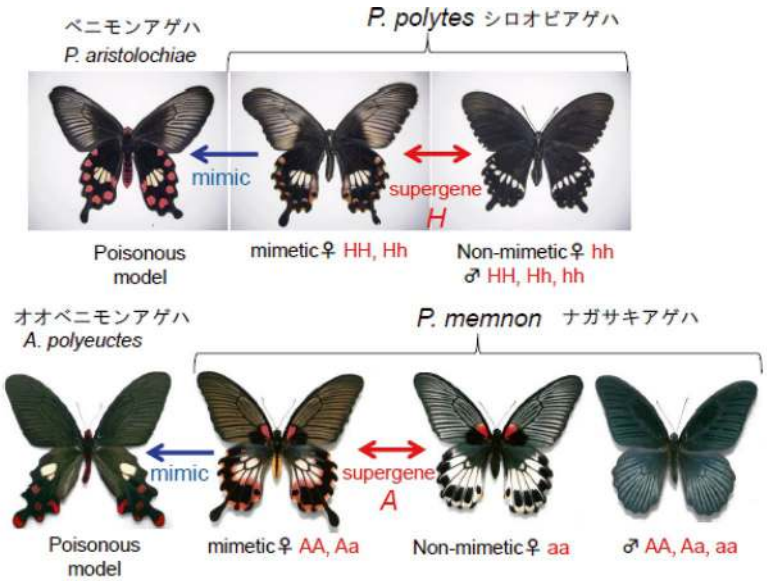
1.4.0.1 References

Iijima T et al (2018) Parallel evolution of Batesian mimicry supergene in two *Papilio* butterflies, *P. polytes* and *P. memnon*. Sci Adv 4:eeao5416

Komata S et al (2022) Genomic architecture and functional unit of mimicry supergene in female limited Batesian mimic *Papilio* butterflies. Philos Trans R Soc Lond B Biol Sci B 377(1856):20210198

Komata S et al (2023) Functional unit of supergene in female-limited Bastian mimicry of *Paplio polytes*. Genetics 223:iyac177

Nishikawa H et al (2015) A genetic mechanism for female-limited Batesian mimicry in *Papilio* butterfly. Nat Genet 47:405–409



Female limited Batesian mimicry in *P. polytes* and *memnon*

1.5 (1) A Model for Pattern Formation and (2) Population Dynamics of the Mimetic Butterfly *Papilio polytes* in the Sakishima Islands, Japan: Mathematical Analysis and Computer Simulations

Toshio Sekimura
Chubu University, Japan



1.5.0.1 (1) A Model for Pattern Formation

Butterfly wing pigmentation patterns are one of the most spectacular and vivid examples of pattern formation in biology. They have attracted much attention from experimentalists and theoreticians, who have tried to understand the underlying genetic, chemical, and physical processes that lead to patterning. In this talk, I present a brief review of this field by first considering the generation of the localized, eyespot, patterns, and then the formation of more globally controlled patterns. We present some new results applied to pattern formation on the wing of the mimetic butterfly *Papilio dardanus* and *Papilio polyte*s.



The African female-limited mimetic butterfly *Papilio dardanus*

1.5.0.2 (2) Population Dynamics of the Mimetic Butterfly *Papilio polytes* in the Sakishima Islands, Japan

I next present a mathematical model for population dynamics of the mimetic swallow butterfly *Papilio polytes* in the Sakishima Islands, Japan.

Introduction *P. polytes* is a mimetic swallowtail butterfly species widely distributed across India and Southeast Asia, including Southeast of China, the Philippines, Taiwan, and the Ryukyu Islands of Japan (Clarke and Sheppard 1972). *P. polytes* exhibits the female-limited polymorphism, i.e., the female is polymorphic, whereas the male is monomorphic and exhibits a white bar on the black hind wing. In the Ryukyu Islands located in the southwest of Japan, the female of *P. polytes* has two different forms, the mimetic form f. *polytes* and the nonmimetic form f. *cyrus* resembling the monomorphic male in appearance. The form f. *polytes* mimics the unpalatable butterfly *Pachiliopta aristolochiae* as a mimetic model, which has a large white area in the center, and a row of submarginal red spots on the black hind wing. Mimicry in the female of *P. polytes* is known to be Batesian mimicry. Batesian mimicry is one of the most interesting biological phenomena in nature.



Papilio polytes **and the numerical simulation result of fore and hind wings**

1.6 The Development and Evolution of Butterfly Wing Patterns: Finding the Primitive Pattern

Fred Nijhout
Department of Biology
Duke University
Durham, NC 27708
hfn@duke.edu



Synopsis The color patterns of butterflies and moths constitute a highly evolved system of visual communication used in sexual signaling, camouflage, warning, territoriality, and mimicry. These patterns are based on extensive modifications of a ground plan of three pairs of bands. The ancestors of this pattern can be found as far

back as the dragonflies, which are among the first winged insects, and model simulations will illustrate how this could have happened.

Outline The development and evolution of butterfly wing patterns have been extensively studied over the past 30 years. Much of our understanding of the structure of these color patterns derives from the classical work of Schwanwitch (1924) and Süffert (1927). These authors independently derived the diversity of color patterns in the butterfly Family Nymphalidae, from a common "ground plan" composed of three parallel pairs of bands that run from the anterior to posterior margin of the wing. Each pair of bands has a characteristic pigmentation pattern and is called a symmetry system because the pigments of the bands mirror each other. Subsequent work by these authors and their associate showed that the "nymphalid ground plan" (Fig. A.1) applies widely to other families of butterflies and to many moths as well.

A long-standing question in this field is whether the nymphalid ground plan somehow represents the primitive butterfly wing pattern from which all others were derived. The butterflies are phylogenetically within the moths, so if there is a primitive pattern of Lepidoptera, it needs to be sought among the moths. Recently, Schachat and Brown (2016) studied the color patterns of the Micropterigidae, the most basal group of moths, and deduced that their color patterns were based on a system of five to six parallel bands. In many moths and butterflies, the crossbands

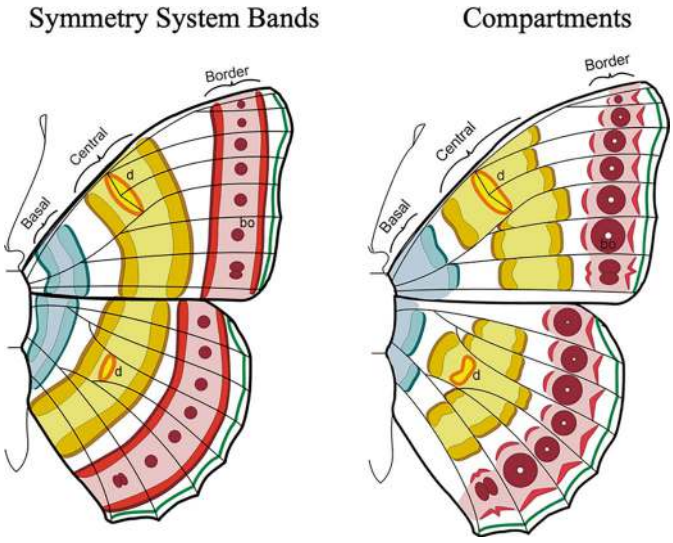


Fig. A.1 The *nymphalid ground plan* is composed of three sets of bands called the basal-, central-, and border-symmetry systems. In some butterflies and moths, the bands are uninterrupted, as in the left panel, but in most cases the bands are disrupted by the vein system (as in the right panel), which compartmentalizes the pattern into semi-independent developmental units. (After Nijhout 1991)

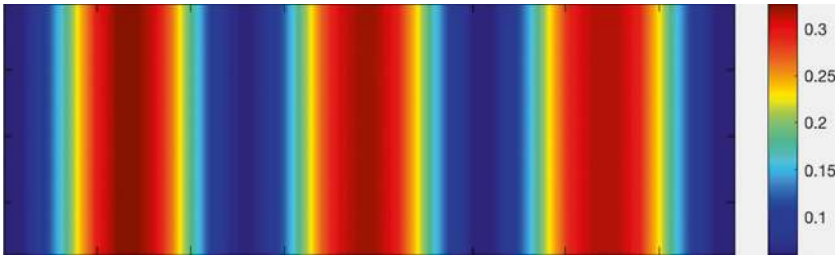


Fig. A.2 Pattern of three symmetry systems formed by a substrate-depletion model of reaction–diffusion

are interrupted by the wing veins and appear as a row of spots, and it is still an open question whether bands arose from rows of spots or whether rows of spots came from broken bands.

In my presentation I will show that the primitive color pattern from which the moth patterns evolved arose long before the evolution of the Lepidoptera. Color patterns of the caddis flies (Trichoptera) and scorpion flies (Mecoptera) all resemble those found in the basal moths. I will show that these patterns are all based on variations on a set of six parallel bands, and that this basic patterning system can be found in the earliest winged insects, represented today by the dragonflies (Odonata).

A simple reaction–diffusion system can readily generate three symmetry systems, which can be visualized as either three pairs of bands or six identical bands (e.g., the yellow stripes in Fig. A.2).

I will show that from this banding system, a lateral inhibition mechanism can, in turn, generate much of the diversity of spotting and banding patterns that we see in the dragonflies and other patterned insect wings, including the Lepidoptera.

1.6.0.1 References

Nijhout HF (1991) Development and evolution of butterfly wing patterns. Smithsonian Press, Washington, DC

Schachat SR, Brown RL (2016) Forewing color pattern in Micropterigidae (Insecta: Lepidoptera): homologies between contrast boundaries, and a revised hypothesis for the origin of symmetry systems. BMC Evol Biol 16:116

Schwanwitsch BN (1924) On the groundplan of the wing-pattern in nymphalids and certain other families of rhopalocerous Lepidoptera. Proc Zool Soc Lond B 34:509–528

Süffert F (1927) Zur vergleichende Analyse der Schmetterlingszeichnung. Biologisches Zentralblatt 47:385–413

1.7 Snail Coiling: CRISPR Editing of a Single Gene Turns Righties into Lefties

Reiko Kuroda
Frontier Research Institute, Chubu University, Japan
rkuroda@isc.chubu.ac.jp



Synopsis The establishment of left–right (l–r) body asymmetry is a key biological process, which is strictly regulated genetically. We have demonstrated unequivocally that the handedness of the freshwater snail *Lymnaea (L.) stagnalis* is determined by the single maternal diaphanous gene *Lsdia1*. The gene operates at the one-cell stage to set the chirality and dictates the entire chiromorphogenesis program across the levels of biological hierarchy. We have also revealed that mechanical micromanipulation during the third cleavage to reverse the spiral-cleavage direction creates mirror-image animals.

Abstract Most animals display external bilateral symmetry, but the structure and position of the visceral organs show asymmetries that are determined genetically. To reveal the still unknown mechanism of the establishment of left–right body asymmetry, we have chosen to work on the freshwater snail *L. stagnalis*, which has several unique characteristics: (1) Both the dextral and the sinistral strains within a species are found in the wild with 98% of the snails dextral and 2% sinistral. (2) The asymmetry is displayed both externally and internally. (3) Snail handedness depends on the direction of spiral cleavages of the embryos, as was reported as early as 1894 (Crampton 1894). (4) The snail is a hermaphrodite, and hence both self-crossing and crossing can occur. This makes the genetics more versatile and informative.

Figure A.3 shows three left–right pairs of *L. stagnalis*: (1) a pair in the wild, (2) one created by mechanical manipulation, and (3) the sinistral one is created by CRISPR/Cas9 genome editing of the dextral one. Contrary to what has been believed, temporal and spatial cytoskeletal dynamics for the left- and right-handed snails within a species are not mirror images of each other. Thus, during the third cleavage, helical spindle inclination (SI) and spiral blastomere deformation (SD) are observed only in the dominant dextral embryos at metaphase–anaphase, whereas

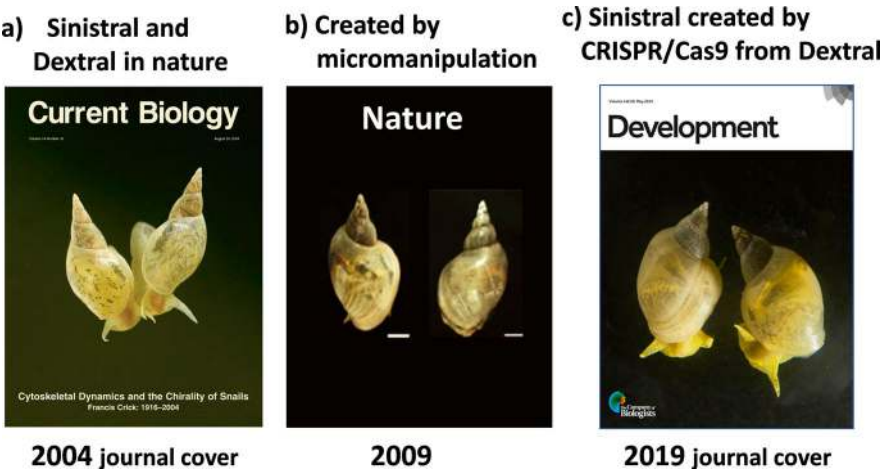


Fig. A.3 Three l–r pairs in the wild (**a**) and created (**b**, **c**)

in the recessive sinistral embryos, helicity emerges during the furrow ingression (Fig. A.4) (Shibazaki et al. 2004). We have constructed the congenic F10 strain that, on average, has inherited 99.9% of the sinistral-derived genome and only 0.1% of the dextral strain-derived genome. All the dextral embryos oviposited by F10 animals that inherited the dextrality gene(s) within the 0.1% of the dextral-derived genome exhibited dextrotropical rotation with SD/SI at the third cleavage and without exception grew to possess dextral body shape. Equally, all the sinistral embryos oviposited by F10 animals that do not inherit the dextrality-determining gene(s) showed levotropical rotation without SD/SI, and grew to possess the sinistral body shape, again without exception. These findings strongly suggested that dextrality is determined by a single gene or by genes at closely linked loci (Kuroda 2015).

We carried out positional cloning and identified that *Lsdia1* is the strongest candidate gene for the handedness determination. There are tandemly duplicated *Lsdia1* and *Lsdia2* genes whose proteins share 89.4% amino acid similarity. In the sinistral strains, both *Lsdia1* alleles had a frameshift mutation early in the coding region (c.184delC), which leads to protein truncation (Kuroda et al. 2016). Further, knocking out the *Lsdia1* gene using CRISPR/Cas9 produced sinistrally coiled offspring generation-after-generation in the otherwise totally-dextral genetic background, if biallelic frameshift mutations had occurred (Fig. A.3c) (Abe and Kuroda 2019). We could show that the gene sets the chirality at the one-cell stage, the earliest observed symmetry-breaking event linked to body-handedness in the animal kingdom. The gene dictates the entire chiromorphogenesis program across the levels of biological hierarchy wherein the early intracellular chirality is superseded by the intercellular chirality during the third cleavage, leading to asymmetric *nodal/Pitx* expressions and then to organismal body handedness (Abe and Kuroda 2019).

In parallel, we revealed that mechanical micromanipulation during the third cleavage to reverse the spiral cleavage direction creates mirror-image animals

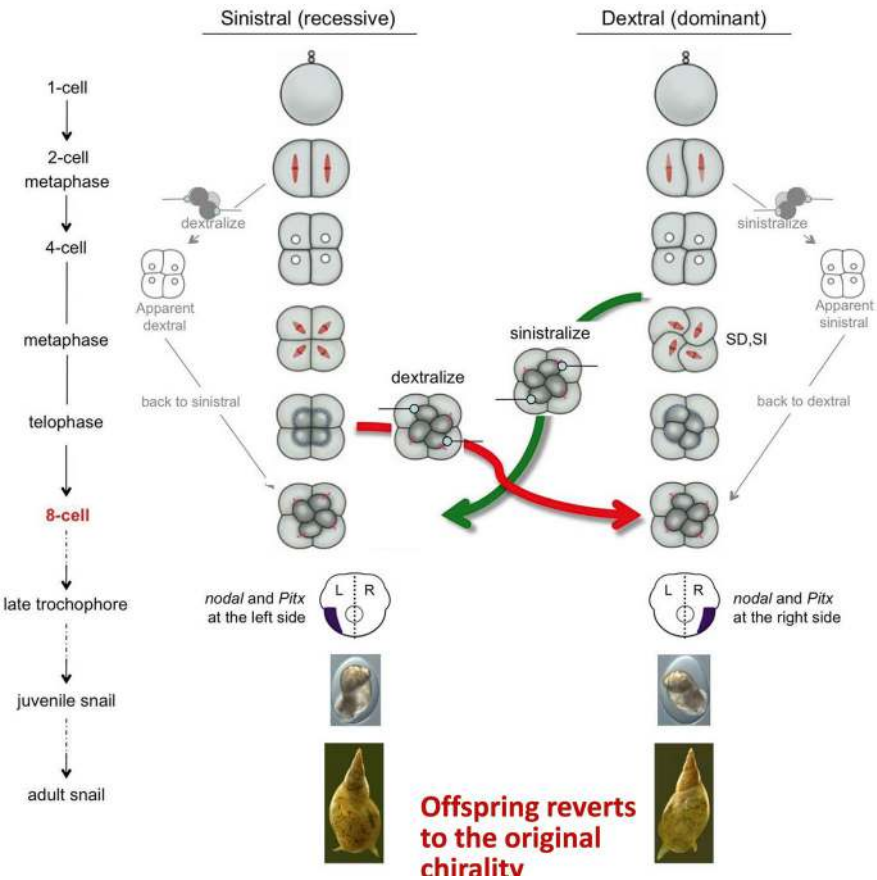


Fig. A.4 Breaking of mirror-image relationship, and l–r inversion by mechanical manipulation

(Figs. A.3b and A.4). The expression site of the *nodal/Pitx* genes was inverted by mechanical micromanipulation (Fig. A.4) (Kuroda et al. 2009).

Diaphanous genes are present in all the Eukaryotes, and therefore, our work featuring formin-controlled early onset of chirality in *L. stagnalis* may provide new insight on unifying mechanisms of l–r body plan formation in animals.

1.7.0.1 References

Abe M, Kuroda R (2019) Development 146:Dev175976
Crampton HE (1894) Ann N Y Acad Sci 8:167–170
Kuroda R (2015) Q Rev Biophys Discov 48:445–452
Kuroda R et al (2009) Nature 462:790–794
Kuroda R et al (2016) Sci Rep 6:34809
Shibazaki Y, Shimizu M, Kuroda R (2004) Curr Biol 14:1462–1467

1.8 Morphospace Approaches to Evolution: Geometric and Developmental Considerations

Dr. Sylvain Gerber
Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle,
CNRS, SU, EPHE-PSL, UA, 57 rue Cuvier, CP50, 75005 Paris, France
sylvain.gerber@mnhn.fr



Synopsis The biological notion of morphospace is examined from its conceptual and analytical dimensions. We review the diverse morphospace geometries encountered in the literature and their implications for the interpretations of morphological patterns. We also show how considerations of organismal development suggest novel mathematical structures to endow the set of morphological variants with evolutionarily meaningful notions of accessibility.

1.8.0.1 Main Text

The notion of morphospace is both of conceptual and empirical importance in biology and paleobiology. Conceptually, it has entailed critical discussions about the relative prevalence of selective forces, variational constraints, and historical contingencies in shaping patterns of morphological evolution. Empirically, morphospaces have proven powerful quantitative tools to document and analyze morphological patterns in various research areas, from comparative anatomy and systematics to functional morphology and large-scale studies of clade dynamics.

A morphospace is generally understood as the depiction of morphological variants as points in an abstract space in such a way that the relative proximity of points expresses the morphological similarity of variants. Here, we discuss whether this way of conceiving morphospace is always warranted in empirical studies or even adequate for all research questions. Two main aspects pertaining to morphospace geometry are considered: (1) the potential misinterpretation of the geometry induced by some morphometric descriptors and (2) a notion of evolutionary accessibility

among variants derived from the developmental representation of morphological traits.

Morphospace narratives often include statements about distance (between morphological variants) and direction (of morphological change), indicating that an inner-product geometry (generally the dot product of Euclidean geometry) is assumed for the morphospace. However, not all morphospaces are equipped with the metric properties that one tends to intuitively attach to the notion of "space" (Huttegger and Mitteroecker 2011; Gerber 2017). The so-called "theoretical morphospaces" are particularly prone to such nonmetric geometries (Gerber 2017). Raup's iconic shell coiling morphospace is taken as an example to illustrate this phenomenon and discuss its implications for biological inferences.

In paleobiology, the use of discrete character data does lead to metric morphospaces (assuming no missing data), but their Hamming-like geometry is sometimes overlooked (Gerber 2019). In particular, the morphospace and its multivariate ordination are often confused for one another, inviting the use of inappropriate methodologies for their analyses (Fig. A.5a).

Historically, methodological approaches to morphospace construction have typically aimed for metric topologies defined at the morphological level, organizing variants according to their morphological similarity. However, this choice may be suboptimal for addressing some important evolutionary questions, since morphological distance among variants can be a poor predictor of their evolutionary accessibility (Fig. A.5b). We show how a genetic or developmental representation of morphological characters allows the definition of accessibility among variants and enriches the explanatory power of morphospace with regards to the role of constraints and selection in morphological evolution (Gerber 2014).

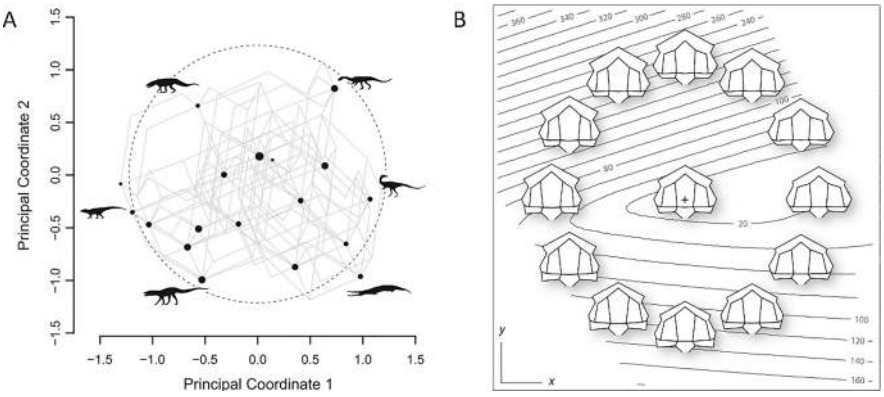


Fig. A.5 (**a**) Morphospace representation from discrete character data and (**b**) accessibility structure mapped onto a morphospace

1.8.0.2 References

Gerber S (2014) Not all roads can be taken: development induces anisotropic accessibility in morphospace. Evol Dev 16(6):373–381
Gerber S (2017) The geometry of morphospaces: lessons from the classic Raup shell coiling model. Biol Rev 92(2):1142–1155
Gerber S (2019) Use and misuse of discrete character data for morphospace and disparity analyses. Palaeontology 62(2):305–319
Huttegger SM, Mitteroecker P (2011) Invariance and meaningfulness in phenotype spaces. Evol Biol 38:335–351

1.9 Functional Morphospace Analysis of Molluscan Shells

Takao Ubukata
Kyoto University
ubukata@kueps.kyoto-u.ac.jp



Morphospace is a powerful tool for large-scale comparison of morphology across a number of species. Morphospaces are geometric spaces in which a particular organic form can be plotted with reference axes chosen to represent morphological parameters. Abstraction of essential morphological traits allows us to summarize complex morphological variation into a low dimensional morphospace. If a functional performance can be assessed for a set of organic shapes accommodated in the morphospace, the relationship between the functional performance and morphological traits can be visualized as a surface on the morphospace that mimics an adaptive landscape. The functional morphospace allows us to predict which shape is common or rare from the view point of functional morphology. Such a prediction can be readily tested by paleontologists by referring to actual fossil records. The present talk introduces a couple of case studies on evolutionary morphology of fossil and extant molluscan shells using morphospaces based on theoretical morphologic models.

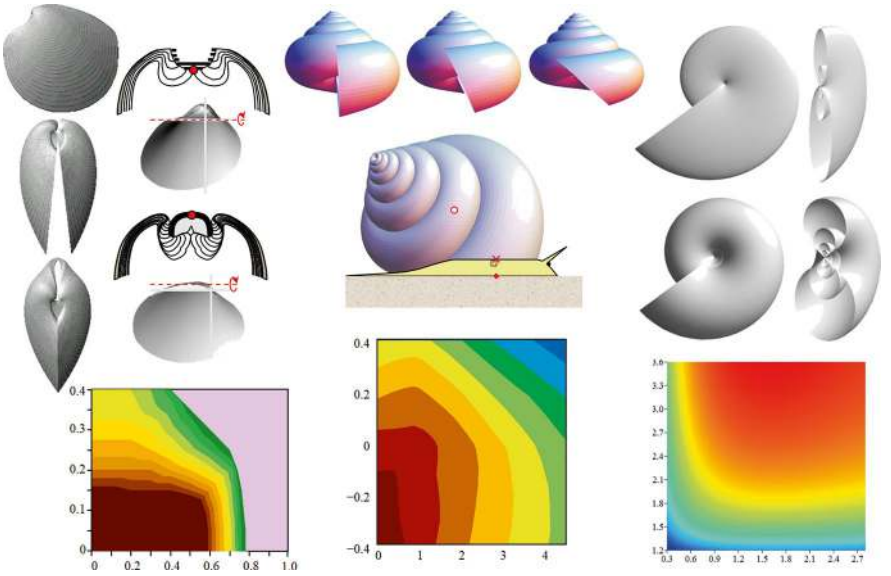
In the first topic on bivalve shell morphology, geometric constraints of bivalve shell form derived from space conflict between umbones of both valves are discussed. A functional morphospace analysis of bivalve shells using a theoretical morphologic model revealed that the space conflict between umbones restricts

morphological diversity of bivalve shells of each group depending on hinge morphology. A characteristic type of ligament evolved in a heteroconch bivalve in the late Mesozoic era may have allowed the clams to develop an adaptive shell shape for burrowing through their macroevolution.

The second topic on shell form of gastropod snails focuses on a functional trade-off between postural stability and efficiency of shell construction. The result of a functional morphospace analysis of gastropod shells revealed that mechanical stability of a snail while creeping and the amount of shell material necessary for shell construction both depend on shell shape and are hardly compatible with each other. The aperture of a snail is often inclined to the coiling axis, particularly in the species living in the environment with less buoyancy and calcium carbonate. The results suggest that the aperture inclination downward relaxes the functional trade-off between postural stability and efficiency of shell construction.

The subject of the third topic on cephalopods focuses on the functional properties for their nektonic mode of life. The functional morphospaces constructed using a theoretical model representing planispiral cephalopod shells clearly show a functional trade-off between hydrostatic and hydrodynamic efficiencies—a shell form that minimizes a form drag acting on the shell while swimming readily maximizes the density of the totally produced shell. The results of morphometric analyses indicate the following: Nautilids tend to be optimized for hydrostatics in comparison with other groups. Some ammonoid groups seem to be optimized for hydrodynamic efficiency. Tarphycerid cephalopods tend to have high scores of postural stability. The results suggest that the macroevolutionary pattern is clearly different among cephalopod phylogenies.

Morphospaces constructed through geometric morphometric and multivariate analyses appear to be potentially useful for a functional morphospace analysis. Evolutionary morphology based on morphospace analysis can be facilitated by big-data analyses using morphological databases. Combining the conventional functional morphospace analyses and their potentially relevant techniques is an interesting subject of future research.



1.10 Mechanical Optimality Hidden in the Structure of Plants

Hiroyuki Shima
University of Yamanashi, Japan
hshima@yamanashi.ac.jp



Plants of various morphology (shape) can be found in nature. These diverse shapes are believed to be the result of the evolution of plants, which cannot move, in order to adapt to the surrounding environment. In this talk, I will particularly focus on the shapes of three plant types: (1) bamboo (Shima et al. 2016, 2023; Inoue and Shima 2018; Sato et al. 2017), (2) butterbur, and (3) muskmelon (Akiba et al. 2022). First, I will demonstrate how the shapes of bamboo and butterbur, which are both tall and hollow plants, are optimal for supporting their own weight. Next, I will explain that the rind of muskmelons develops a mesh pattern with a certain geometric property, which may be used by farmers as a visual indicator of optimum sweetness of flesh; this is a consequence of a farmer seeking the optimum sweetness of the flesh.

1.10.0.1 Optimal Structure of Bamboo

Bamboo is the fastest-growing plant worldwide. This rapid maturity stems from their hollow structures, which require less material to develop, thereby promoting rapid growth. One critical disadvantage of hollowness is the propensity to easily collapse under external forces such as wind, rain, and self-weight. To overcome this limitation, bamboos generate a sequence of "circular plates" within the long cavity. The plates stiffen the culm locally; therefore, numerous plates from the bottom to the top of the bamboo enhance the stability of the culm. However, this is not a case of "the more, the better," because an excessive number of plates will inhibit the growth rate of the bamboo plant. Therefore, the number of plates has to be "neither

too many nor too few." This trade-off between stability and growth rate implies an optimum number of stiffening plates that bamboo has acquired over millions of years of evolution.

Against this backdrop, our research group has explored the optimum solution to this issue by investigating wild bamboo groves in Japan. Analysis of the measurement data showed that a single universal law explaining the interdependence between the preferred number of stiffening plates and their ideal arrangement along the culm governs the growth of bamboo (Shima et al. 2016). Specifically, the density at which fibers are embedded increases as it gets closer to the outer skin. This universal law thus provides a design criterion, endowing less material and high strength to wild bamboos, which are advantageous for survival over competing tree species.

Surprisingly, another mechanism lurks in the cross-section of the bamboo to increase its stability as a whole. It is bamboo fiber. Inside the bamboo, many long fibers are embedded in the longitudinal direction. These fibers are called "vascular sheaths" in biological terminology, and they act as straws that convey water and nutrients vertically. The structural mechanics theory revealed that this fiber plays an important role in the mechanical properties of bamboo. If a part of the cross-section of bamboo is cut and observed using an electron microscope, it can be seen that the fibers are not evenly distributed over the entire cross-section. Specifically, the closer it is to the outer skin, the more fibers are embedded. We were able to prove within the scope of the elastic approximation theory that by gathering many reinforcing fibers near the outer skin of the bamboo, the bamboo as a whole is less likely to bend (Sato et al. 2017). This result indicates that the bamboo makes its own body stronger by efficiently arranging a limited number of reinforcing fibers inside the cross-section.

1.10.0.2 Optimal Cross-section of Butterbur Petiole

There are many plant species with hollow stems and petioles, including those belonging to the family Poaceae. One of the reasons why many plant species have hollow tubular structures is thought to be the structural stabilization provided by hollowness. A tube with a hollow cross-section deforms less than a solid tube (filled) with the same cross-section. Thus, it can support its own weight by using minimal resources. In addition, by choosing the optimal cross-sectional shape (circular, polygonal, etc.) according to the growth environment, the resistance to the most influential mechanical loads (crosswind, self-weight of stems, weight of petals, etc.) increases.

From this perspective, the plant that we focused on in this study was the Japanese butterbur, a type of wild grass. A close examination of the cross-section of the stem of the Japanese butterbur reveals that it has a shape similar to that of a horse's hoof, which is rarely seen in other plants. We hypothesized that this peculiar horseshoe-shaped stem cross-section is the optimal structure for supporting large leaves at the tip of the stem. We have verified this hypothesis both experimentally and theoretically.

1.10.0.3 Geometric Law of Mesh Pattern of Melon Skin

The surface of muskmelons, which is representative of high-grade melons, has a fine mesh pattern (net). It is empirically known that the appearance of this mesh pattern correlates with the quality of melons and is used as an important indicator for quality evaluation at production sites. However, only a few studies have investigated the geometric properties of melon nets.

Our research group took photographs of several muskmelons sold at a retail store (for quality assurance) and investigated the shape and size of the small green peel fragments separated by white reticular stripes. Calculating the areas of these small peel fragments, obtaining the probability distribution curves of the areas, and normalizing them, it was found that the distribution curves were on a common curve regardless of the type and size of individual melons (Akiba et al. 2022). Furthermore, this common probability distribution curve was derived from fracture mechanics theory. This discovery suggests the possibility of nondestructive identification and evaluation of melon quality using mesh pattern imaging.

1.10.0.4 References

Akiba Y, Ishibashi A, Sato M, Shima H (2022) Empirical rule of fruit rind fragmentation in muskmelon epidermis netting. J Phys Soc Jpn 91:104801

Inoue A, Shima H (2018) Survival time of Princess Kaguya in an air-tight bamboo chamber. Forma 33:1–5. [https://forma.katachi-jp.com/frame/33.html]

Sato M, Inoue A, Shima H (2017) Bamboo-inspired optimal design of functionally graded hollow cylinders. PLoS One 12:e0175029

Shima H, Inoue A, Sato M (2023) Bamboo: a mechanically optimum design in nature. In: Bamboo science and technology. Springer. ISBN: 978-981-99-0014-5

Shima H, Sato M, Inoue A (2016) Self-adaptive formation of uneven node spacings in a wild bamboo. Phys Rev E 93:022406

1.11 Spiral Patterns and Morphology in Chemical Systems

Oliver Steinbock
Florida State University,
Department of Chemistry and Biochemistry,
Tallahassee, FL 32306-4390, USA
E-mail: osteinbock@fsu.edu



Synopsis This talk will showcase different examples of chemical systems undergoing self-organization far from the thermodynamic equilibrium. Starting with a brief look at the history of chemical waves and the work of the field's pioneers, we will discuss similarities to the growth of unusual polycrystalline solids called biomorphs and go beyond to the self-propulsion of chemical garden tubes.

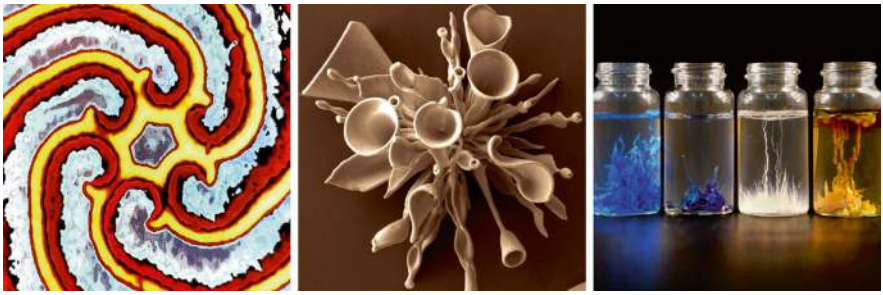
Abstract Living systems masterfully program molecular events to create macroscopic shapes and dynamics. Can the same be done with inorganic chemistry? I believe the answer is yes if reactions are kept far from the thermodynamic equilibrium (and your expectations are not too high). In my talk, I will introduce several self-organizing chemical systems including the Belousov–Zhabotinsky (BZ) reaction, chemical gardens, biomorphs, and self-propelled catalytic units capable of swarm-like cooperativity. These examples will illustrate unexpected behaviors if reactions are allowed to unfold in space and time. Perhaps not surprisingly, similar processes are used by living systems, suggesting system-level interpretations of important biological phenomena (e.g., cardiac arrhythmia) and novel approaches to materials science and engineering.

Our discussion of the BZ reaction will focus on three-dimensional patterns and discuss them in the context of pioneering work by Profs. Stefan Müller, Benno Hess, and Art Winfree. These and other prominent scientists studied two-dimensional spirals, the basic building blocks of scroll waves that provide the foundation for understanding 3D dynamics. While spiral waves rotate around point singularities, scroll waves rotate around 1D filaments that move according to their local curvature. Today tomographic methods, numerical simulations, and controlled manipulations are used to obtain deeper insights into the local pinning to heterogeneities as well as scroll behavior in gradients. Many of these studies, including the

results to be presented, are motivated by similar phenomena in the ventricles of the human heart and their link to sudden cardiac death.

Similar nonlinear wave dynamics appear to rule the growth of certain polycrystalline assemblies known as biomorphs. The smoothly curved and life-like biomorphs form when aqueous solutions of $BaCl_2$ and silicate react with CO_2. The shapes are about 100 μm in size and include helices, funnels, and urns. The entirely abiotic biomorphs can form from natural spring waters and constitute intriguing problems for the identification of Earth's earliest microfossils as well as the search for life's remnants on Mars and other planetary bodies.

Lastly, we will present novel results on chemical gardens (CG). In the classical CG experiment, these colorful, hollow precipitate tubes form when a salt crystal is placed into an alkaline sodium silicate solution. Specific aspects to be covered include a cellular automaton model of CGs, the involvement of colloidal particles in their growth, and applications as self-propelled micro-rockets. Our model considers CG growth based on a simple $A + B \rightarrow C$ reaction and a lattice for which each site is exactly one of these three species. We show that the aging of the material modulates the likelihood of mechanical changes in the position of the solid product C and that a strong preference for changes in the fresh material causes filament (2D) and tube (3D) growth. The self-propulsion of CG tubes is demonstrated for Mn-containing units in H_2O_2 solutions. The tubes are propelled forward by the rhythmic ejection of O_2 microbubbles formed in the catalytic tube cavity. We also observed a characteristic acoustic signature of this propulsion that allows detailed analyses of the bubble dynamics. These analyses reveal that the tube speed is determined by the tube diameter (which approximates the step size per bubble) and the bubble ejection frequency (which is nearly proportional to the H_2O_2 concentration). Lastly, we will discuss the swarming behavior of the CG tubes and their interactions with bubble rafts.



1.11.0.1 References

Batista BC, Morris AZ, Steinbock O (2023) Pattern selection by material aging: modeling chemical gardens in two and three dimensions. Proc Natl Acad Sci USA 120:e2305172120

Knoll P, Wu N, Steinbock O (2020) Spiraling crystallization creates layered biomorphs. Phys Rev Materials 4:063402

Zanotto FM, Steinbock O (2021) Hyperscroll dynamics: vortices in four-dimensional networks. Chaos 31:053132

1.12 Autonomous Workload Regulation Mechanism in Foraging Ants

Hiraku Nishimori*, Masashi Shiraishi, Osamu Yamanaka
Meiji Institute for Advanced Study of Mathematical Sciences,
Meiji University
* Correspondence: nishimor2@meiji.ac.jp



As the basic mechanism to enable task allocation in ant colonies without a leader, the primary role of the caste system has been discussed (Wilson 1983), through which system, specific tasks are autonomously assigned to individual workers according to their size, age, and other factors. Even among specialized ants that are involved in the same task, a hierarchical structure, namely a heterogeneity of workload, is observed.

In the present study, we introduced an experimental system to observe the statistical behavior of colonies of ants, *Camponotus japonicus*, using tiny RFID tags attached to the bodies of all ants in five colonies (Fig. A.6) (Yamanaka et al. 2019). By analyzing the "big data of ant society" focusing on foraging behavior obtained through more than 3 months of continuous measurement, we found various kinds of statistical characteristics of the ant society: (1) The emergence of the foraging workload hierarchy, namely the heterogeneity of the diligence on foraging task, that obeys a simple mathematical form of distribution; (2) Autonomous workload compensation after the separation of each colony into two different groups, the diligent half before separation, and the lazy half; and (3) The slowly decreasing correlation of rank–order of diligence with a characteristic duration of a month.

As a mathematical model to numerically reproduce the characteristic behavior of social insects, ants and bees, the response threshold model (Bonabeau et al. 1996) is widely known. To analytically treat the response model, we introduce a corresponding master equation. In conjunction with foraging behavioral data obtained through the above-mentioned long-term observation using tiny RFID tags (Yamanaka et al. 2019), we found that a master equation approach is effective in quantitatively describing the statistical features of the collective foraging-task engagement of ants. Specifically, the following two kinds of statistical characters related to the

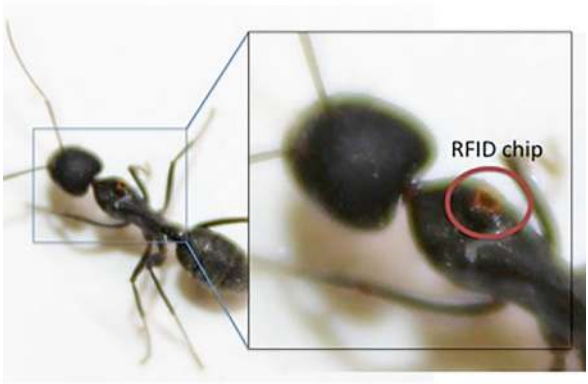


Fig. A.6 RFID tag is affixed to the center of the red circle in the enlarged image

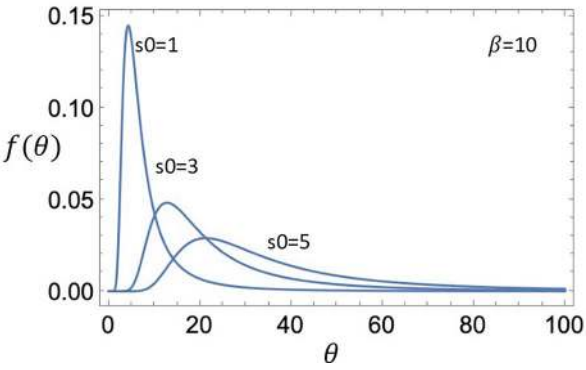


Fig. A.7 Estimated distributions of the response threshold $f(\theta)$ of the foraging task through the combinatory analysis of master equation and the experimental outcomes of the foraging activity of *Camponotus japonicus*'s colony. The shape of $f(\theta)$ varies according to the experimentally unconfirmed values of parameters, α, β, and δ. However, if β is fixed, the change of α and δ is reflected in $s0$, then all graphs overlap after rescaling

autonomous workload compensation in ant colonies were indicated: (1) Response threshold θ for starting the foraging task is not uniformly nor normally distributed but largely skewed among ants in each colony (Fig. A.7); and (2) Foraging task in a colony is carried only by a small fraction of workers in a colony.

1.12.0.1 References

Bonabeau E et al (1996) Proc R Soc Lond B 263
Charbonneau D et al (2017) PLoS One 12(9):e0184074
Wilson EO (1983) Behav Ecol Sociobio 14:47
Yamanaka O et al (2019) Sci Rep 9:1