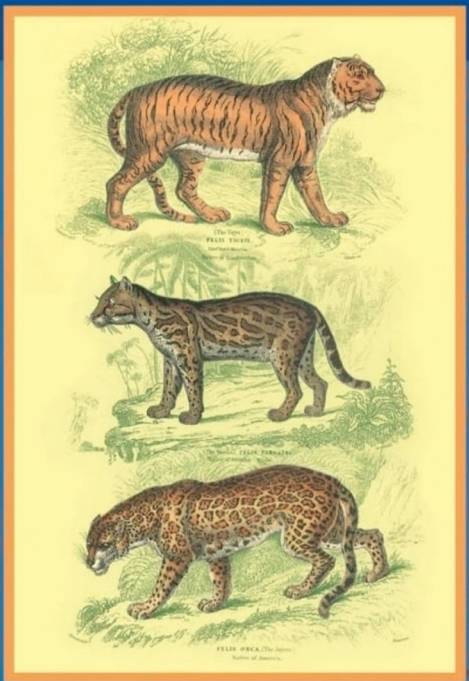
SEAN B. CARROLL JENNIFER K. GRENIER SCOTT D. WEATHERBEE

FROM DNA TO DIVERSITY

MOLECULAR GENETICS AND THE EVOLUTION OF ANIMAL DESIGN SECOND EDITION



Blackwell Publishing

From DNA to Diversity

Molecular Genetics
and the Evolution of Animal Design

2nd edition

Sean B. Carroll
Jennifer K. Grenier
Scott D. Weatherbee

Howard Hughes Medical Institute and
University of Wisconsin
Sloan Kettering Institute, New York

Blackwell
Publishing

BLACKWELL PUBLISHING

350 Main Street, Malden, MA 02148-5020, USA

108 Cowley Road, Oxford OX4 1JF, UK

550 Swanston Street, Carlton, Victoria 3053, Australia

First edition published 2001 by Blackwell Publishing Ltd

Second edition published 2005

Library of Congress Cataloging-in-Publication Data

Carroll, Sean B.

From DNA to diversity: molecular genetics and the evolution of animal design / Sean B.

Carroll, Jennifer K. Grenier, Scott D. Weatherbee.—2nd ed.

p.; cm.

Includes bibliographical references and index.

ISBN 1-4051-1950-0 (pbk. : alk. paper)

1. Evolutionary genetics.2. Molecular genetics. 3. Biological diversity. 4. Morphology (Animals)

[DNLM: 1. Animal Population Groups—genetics. 2. Variation (Genetics) 3. Animal Population Groups—anatomy & histology. 4. Evolution, Molecular. 5. Gene Expression.

QH 408 C319f 2005] I. Grenier, Jennifer K. II. Weatherbee, Scott D. III. Title.

QH390.C37 2005

572′38—dc22

2003027991

A catalogue record for this title is available from the British Library.

To our families

"From DNA to Diversity is written for a general audience, including undergraduates, with an interest in developmental and evolutionary biology, and it is a joy to read. Using striking examples, the authors summarize the current state of thinking on the interconnectedness between developmental genetics and evolutionary diversification." Axel Meyer, University of Konstanz; *Nature*

"This book helps to fill a gap in the teaching of evolutionary theory that arose because developmental biology was not a direct participant in the evolutionary synthesis. … This is an outstanding account of the latest findings in molecular developmental biology." James W. Valentine, Professor Emeritus, University of California, Berkeley

"The authors have done an excellent job of distilling the large and complex literature on molecular genetics that is pertinent to understanding how gene networks evolve … The writing is consistently clear, concise, and engaging."

Gregory A. Wray, Duke University; *Science*

"Carroll, Weatherbee, and Grenier have produced a wonderful and exciting introduction to the field of evolutionary developmental biology … Newcomers and aficionados will find this a compelling read."

Martin J. Cohn, University of Florida; *Evolution and Development*

"… this is one book that everybody should read who wants to know why 'evo-devo' is such a hot topic right now."

Manfred Laubichler, Arizona State University

"From DNA to Diversity can be, and should be read by College and University students as well as scientists out of the field, who want to be informed of what is new and promising in biology."

Jean Deutsch, Universite Phillippe et Marie Curie, Paris; *BioEssays*

Preface

The Earth is now populated by between 1 million and perhaps as many as 20 million animal species, which represent probably less than 1% of all animal species that have ever existed. An even more remarkable fact is that all of this diversity—aardvarks and ostriches, butterflies and pythons, dinosaurs, and earthworms—descended from a common bilaterally symmetrical ancestor that lived in Precambrian seas more than 540 million years ago. Traditionally approached through paleontology, systematics, and comparative anatomy, the story of animal evolution has, until recently, been sorely missing one huge chapter —namely, genetics.

Animals diverge from common ancestors through changes in their DNA. The major question, then, is, Which changes in DNA account for morphological diversity? The answer to this question has eluded us for the half-century since the Modern Synthesis was proposed and the structure of DNA was discovered. Although many reasons exist to explain this omission, foremost among them is that biology first had to address another central genetic mystery—that is, which genes out of the thousands in any species control morphology?

One of the most important biological discoveries of the past two decades is that most animals, no matter how divergent in form, share specific families of genes that regulate major aspects of body pattern. The discovery of this common genetic “toolkit” for animal development has had two major implications for researchers. First, it has enabled biologists to uncover widely conserved molecular, cellular, and developmental processes whose existence was concealed by previously incomparable anatomies. Second, it has focused the study of the genetic basis of animal diversity on how the number, regulation, and function of genes within the toolkit have changed over the course of animal evolution.

The genetic picture of morphological diversity presented in this book is highly influenced by the legacy of previous successes of genetic logic. The mysteries of enzyme induction in bacteria and bacteriophage life cycles were, through formal genetic logic and molecular biology, ultimately reduced to elegant genetic switches that determined the on/off state of groups of genes. This success laid the foundation for understanding the regulation of genes in different cell types of multicellular organisms and, in turn, the regulation of genes in space and over time during the development of individual organisms. Similarly, recent advances in understanding how the toolkit operates in the design of just a few model

species has laid the foundation for studies of the evolution of a wide variety of animal structures and patterns.

The presentation in this book lies at the intersection of evolutionary biology with embryology and genetics. Comprehensive treatment of any of these long-established, fast-growing disciplines can be found in full textbooks dedicated to each. Because our goal is to elucidate general principles about the genetic basis of morphological change, we will focus on those genes, developmental processes, and taxa that are best known and best illustrate these principles. The book is organized into two parts. The first part (Chapters 1–3) focuses on the history of animals and on animal developmental genetics and regulatory mechanisms. We first examine some of the major trends in animal design and evolution illustrated in the fossil record and by modern forms (Chapter 1). Next, we take an inventory of the genetic toolkit for the development of model species (Chapter 2). Finally, we analyze the regulation and function of these genes in the complex hierarchies that govern animal development (Chapter 3). This crucial background knowledge of the major transitions in animal evolution and the genetic logic of animal design sets the stage for the analysis of mechanisms of morphological evolution.

The second part of the book examines the genetic mechanisms underlying the evolution of animals at different morphological levels. We take a case study approach by focusing on the best-understood examples of the evolution of the genetic toolkit, the diversity of body plans and body parts, and novel structures. In the final chapter, we discuss why and how changes in gene regulation have played a primary role in the evolution of diversity across the morphological spectrum—from small-scale differences within or between species, to the large-scale differences that distinguish higher taxa.

We have provided selected references for further reading at the end of each chapter. By no means should these citations (or this book) be taken as the primary or exclusive references on a topic. For both brevity and to circumvent questions of priority in ideas or evidence, we have avoided attributions to specific authors in the text.

One of the inspirations for our approach was Mark Ptashne's classic *A Genetic Switch*, in which many of the basic physiological and molecular principles of gene regulation were illuminated by focusing on the bacteriophage λ. In the preface, Ptashne stated that "one of the charms of molecular biology is that the answers it provides to fundamental questions for the most part can be easily visualized." Few fields in biology can rival the aesthetic appeal of the new comparative embryology. Indeed, the visualization of members of the genetic toolkit in action during the development of different species has already become a surrogate for analyzing final forms. For those who find conceptual beauty in the logic and

molecular anatomy of genetic switches, the genetic switches controling animal anatomy may be even more appealing. Not only do they control the striking patterns of gene expression within developing embryos, but as we shall see, they are also key to understanding how the wonderful, but presently dwindling, diversity of animal forms has evolved.

CHANGES IN THE SECOND EDITION

The revision and expansion of *From DNA to Diversity* for this second edition is driven by advances on many fronts. Increased understanding of developmental mechanisms, systematic exploration and comparisons of animal genomes, and inquiries into new models of morphological evolution have provided a wealth of case studies from which we have selected new material. Much of the new coverage in this edition is found in the second part of the book, which has been expanded to five chapters (Chapters 4–8) from four in the first edition. Information and references have been updated throughout the book. Again, we stress that these citations are selective and that neither they nor this book should be taken as the primary or exclusive reference on a topic.

The book's overall organization remains the same, with the first three chapters devoted to the history of animals, developmental genetics, and genetic regulatory mechanisms. The second part of the book (Chapters 4–8) examines the evolution of animals at different morphological levels. The explosion in genome sequence data has provided an enormous increase in the quantity and quality of information concerning the evolution of the genetic toolkit for animal development. Many animal genomes, including our own, have been sequenced since the publication of the first edition. Some of the major insights from genome studies have been added to Chapter 4.

The growth of evolutionary developmental biology has provided new insights into the diversification of specific body plans and the origins of animal novelties. Chapters 5 and 6 have been revised and expanded to incorporate new findings ranging from mechanisms of segmentation in spiders, to the evolution of the cephalopod body plan, and the origin of the turtle shell.

There has also been an increasing focus on models of variation within species and of divergence of traits. Some of the simplest models of phenotypic variation and evolution involve the color patterns of mammals, birds, and insects. In several cases, the identity of genetic differences responsible for variation between

populations is now known. We have added a new chapter (Chapter 7) that focuses on models of variation and divergence among closely related species.

ACKNOWLEDGMENTS

We thank all of our colleagues who have shared their discoveries with us and allowed us to use their illustrations in this book. This edition would not have been possible without the constructive input of reviewers and users of the first edition; we thank you for your thoughtful comments and suggestions. We also thank Jamie Carroll for design of the cover, preparation of the text, and coordination of the editing and permissions process; Leanne Olds for additional artwork; Steve Paddock for help with images; and Nancy Whilton for her unflagging support for this book from the very beginning. We thank the many current and past members of the Carroll laboratory for their insights and hard work in developmental and evolutionary biology, without which this endeavor would not have been possible, and Lee Niswander for her support and encouragement. The authors' work has been supported by the Howard Hughes Medical Institute, National Science Foundation, National Institutes of Health, Sloan Kettering Institute, Shaw Scientist's Program of the Milwaukee Foundation, Human Frontiers Science Program, European Molecular Biology Organization, and the University of Wisconsin.

FIGURE AND TABLE ACKNOWLEDGMENTS

The authors and publisher gratefully acknowledge the permission granted to reproduce the copyright material in this book.

Every effort has been made to trace copyright holders and to obtain their permission for the use of copyright material. The publisher apologizes for any errors or omissions, and would be grateful to be notified of any corrections that should be incorporated in future reprints or editions of this book.

CHAPTER 1

A Brief History of Animals

... an understanding of regulation must lie at the center of any rapprochement between molecular and evolutionary biology; for a synthesis of the two biologies will surely take place, if it occurs at all, on the common field of development.

—*Stephen Jay Gould*

Ontogeny and Phylogeny (*1977*)

The central focus of this book is to identify the genetic mechanisms underlying the evolution of animal design, particularly with regard to the patterning of animal body plans and body parts. To approach this mystery, new discoveries and ideas from developmental genetics must be integrated into the larger framework of the evolutionary history of animal life. This history is reconstructed from many fields of

study—in particular, paleontology, systematics, and comparative biology. In this chapter, we present a brief overview of animal evolution from these three perspectives. This discussion provides a historical foundation for the consideration of the mechanistic questions that are addressed in subsequent chapters.

First, we discuss the origin of animals and the radiation of the major animal phyla based on evidence gleaned from the fossil record. Most living phyla have ancient origins, and the fundamental differences between them evolved long ago. Two milestones in early animal history that are of special interest are the evolution of bilaterally symmetrical animals and the explosive radiation of these forms in the Cambrian period more than 500 million years ago.

Second, we examine the phylogenetic relationships among animals. Understanding the direction of evolutionary change in morphological, developmental, or genetic traits and the ability to make inferences about animal ancestors requires knowledge of the structure of the animal evolutionary tree. While traditionally based upon morphological comparisons, new phylogenies based on DNA and protein sequences have revealed unexpected relationships among anatomically disparate animals, refuting long-held notions about which phyla are more closely related.

Third, we consider the comparative anatomy of selected phyla with the aim of identifying some of the major trends in the evolutionary diversification of individual phyla. In particular, we focus on the modular organization of the body plans and body parts of larger animals—the vertebrates, arthropods, and annelids. Much of the large-scale morphological diversity within these phyla (for example, between different classes) involves differences in the number and pattern of modular elements (segments, appendages, and so on). The recognition of the modular organization of these animals is an important conceptual link to understanding the genetic logic controlling their development and the mechanisms underlying the evolution of diversity.

ANIMAL ORIGINS AND THE FOSSIL RECORD

The fossil record is our primary window into the history of life. It provides many kinds of information that cannot be inferred from living animals. Fossils give us pictures of extinct forms that may be ancestors of modern animals, provide minimal estimates of the time of origin or divergence of particular groups, reveal

episodes of extinctions and radiations, and, in favorable circumstances, offer detailed accounts of the evolution of important structures.

The search for the origins of modern animals begins with an assessment of the **Cambrian** fossil record. It has been known since before Darwin's time that animal diversity increased dramatically during this period, which spans an age from roughly 545 to 490 million years ago (Ma). Molluscs, arthropods, **annelids**, chordates, **echinoderms**, and representatives of most other modern phyla make their first appearance in Cambrian fossil deposits (Fig. 1.1). The emergence of large, complex animal forms and their radiation over a 10 to 25 million year interval in the Early–Middle Cambrian is often referred to as the "Cambrian Explosion."

The appearance of these animals in the Cambrian fossil record gives us only a minimum estimate of their time of origin. The crucial question about the Cambrian Explosion is whether it marks the origin of animals or the origin of modern phyla. Did most phyla first arise in this short period, or did they predate their preservation in the Cambrian fossil record? Although the Precambrian animal fossil record is relatively scarce, several kinds of fossil evidence indicate that the origins of most modern phyla predate the Cambrian. First, the fossil record of some modern groups clearly begins before this period. For example, body fossils of both **cnidarians** and sponges predate the Cambrian (Fig. 1.2). Both of these groups are **diploblastic** animals, composed of two tissue layers. The cnidarians have a radically symmetrical body design that distinguishes them from sponges and from a much larger number of modern phyla that are **triploblastic**—that is, composed of three tissue layers—and have bilaterally symmetrical body designs (the **Bilateria**). Second, Precambrian deposits contain evidence in the form of **trace fossils**, the record of the meanderings and burrowings of animals in sediments, which indicate the existence of some bilaterian forms (Figs 1.2 & 1.3d) well before the Cambrian Explosion. A third piece of potential evidence for earlier animal origins is the **Ediacaran** fauna (575–544 Ma), named for the Australian locale in which they were first discovered.

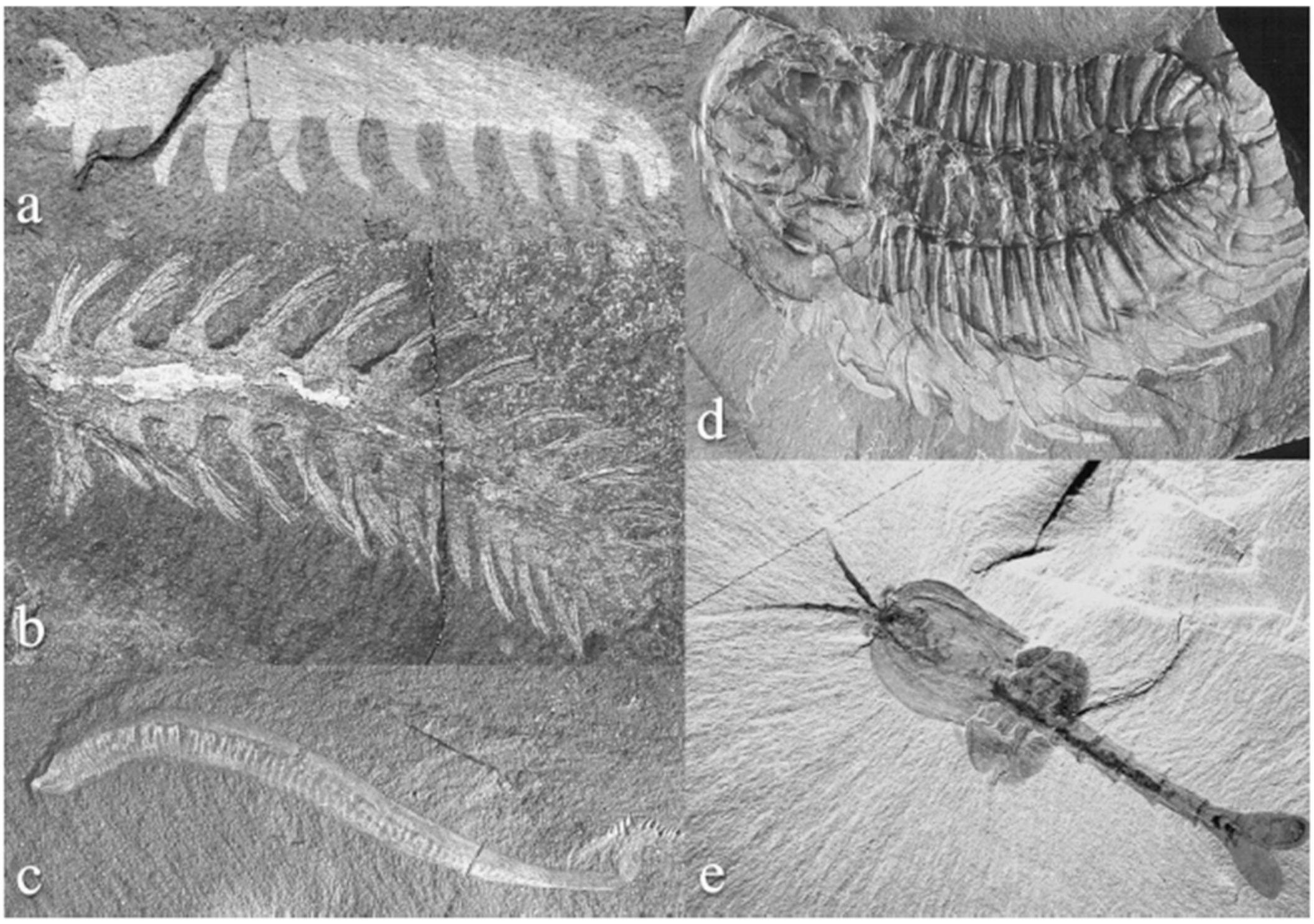
The biological interpretation of Ediacaran fossils and their relationships, if any, to modern animals remains controversial. Several distinct body plans have been identified, including radially symmetrical types and a number of frond-like and tube-like forms (Fig. 1.3). None of these bear any clear-cut similarity to modern animals, so they have been difficult to place on the tree of animal evolution. Some of the Ediacaran fossils could represent diploblastic forms related to cnidarians or sponges. Others could be primitive bilaterians that possess some, but not all, features of modern bilaterians.

The difficulties in placing Ediacarans in the scheme of animal evolution have led to the proposal that they represent an extinct experiment in multicellular life.

On the other hand, perhaps their lack of resemblance to modern groups is exactly what should be expected of primitive animals. It is possible that the Ediacaran fauna include both extinct types of diploblastic animals and primitive ancestors of modern bilaterians. The fossil record indicates that some Ediacaran forms persisted into the Cambrian, but then died out as bilaterians, sponges, cnidarians, and ctenophores flourished.

Figure 1.1 Cambrian animal fossils Representatives of many modern phyla are found in Cambrian deposits and are made up of repeating units. (**a**) *Aysheaia pedunculata*, an onychophoran; (**b**) *Burgessochaeta setigera*, a polychaete annelid; (**c**) *Pikaia gracilens*, a chordate; (**d**) *Olenoides serratus*, a trilobitomorph arthropod; (**e**) *Waptia fieldensis*, a crustacean-type arthropod.

Source: Photographs from Briggs DEG, Erwin DH, Collier FJ. *Fossils of the Burgess shale*. Washington, DC: Smithsonian Institution Press, 1994; reprinted by permission from the Smithsonian Institution Press.

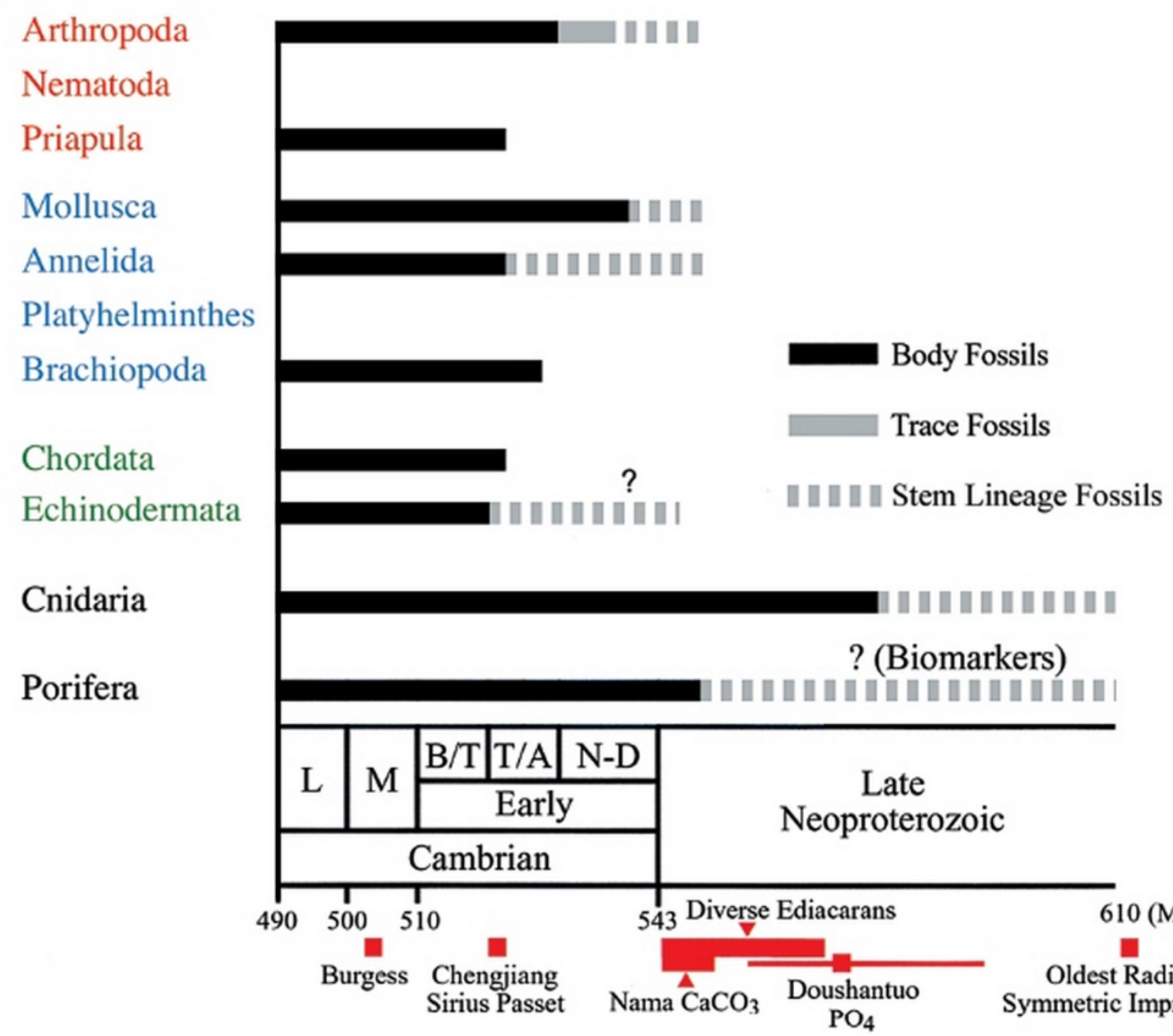


Given the uncertainty of the relationship of the Ediacarans to modern phyla and the paucity of body fossils prior to the Cambrian, it is difficult to pinpoint the origins of modern animals based on the fossil evidence. Consequently, biologists have turned to other methods to try to identify when major animal groups diverged. Using the evolution of protein and ribosomal RNA sequences between

species to calibrate **molecular clocks**, estimates of the time of divergence of most animal phyla have been made that range from approximately 650 Ma to more than 1000 Ma. While these estimates remain controversial, even the most conservative estimate suggests a period of more than 100 million years before the beginning of the Cambrian in which most bilaterian phyla had arisen but led a paleontologically cryptic existence.

Figure 1.2 The early fossil record of animals The appearance of various animal phyla in the fossil record are indicated, relative to the Cambrian and Proterozoic periods. The ages of fossils from particular localities are shown in red at the bottom. Note that the cnidarian and poriferan records clearly predate the Cambrian. Other phyla first appear in the Cambrian, although early members may exist that predate the Cambrian by a considerable period. L, Late; M, Middle; B/T, Botomian plus Toyonian; T/A, Tommotian plus Atdabanian; N-D, Nemakit-Daldynian.

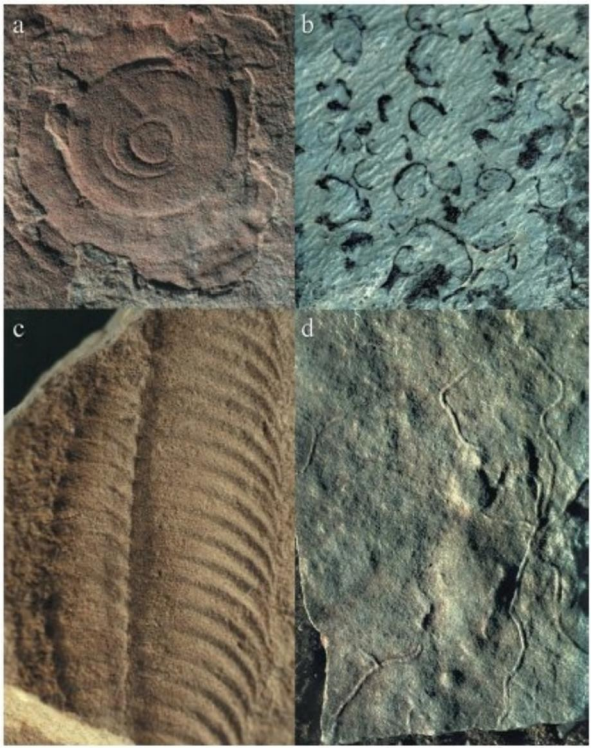
Source: Adapted from Knoll AH, Carroll SB. *Science* 1999;284:2129–2137.



It is widely believed that primitive bilaterians may have been very small and their size limited by atmospheric and oceanic oxygen levels. This fact would help to explain their slim fossil record before the Cambrian (Fig. 1.2). In the last few years, evidence has also been gathered that suggests a possible mass extinction at the boundary between the **Proterozoic** and Cambrian. Whatever the cause of such an event, it may have hastened the extinction of Ediacaran forms and opened up the ecological opportunity for bilaterians to radiate. Environmental and ecological changes may have removed constraints on bilaterians, permitting the evolution of larger animals. In addition, competitive interactions among bilaterians may have facilitated the evolution of skeletonized taxa, more sophisticated predatory and defense behaviors, and the variety of anatomical innovations that unfolded in the Cambrian.

Figure 1.3 Pre-Cambrian animal fossils and traces (**a**) *Ediacaria*, a radially symmetrical form from deposits in Australia. (**b**) Calcified fossils in limestone from Namibia. (**c**) *Pteridinium*, a frond-like ediacaran fossil form built of repeating units. (**d**) Trace fossils made in sediments by bilaterian animals.

Source: Knoll AH, Carroll SB. *Science* 1999;284:2129–2137.



THE ANIMAL TREE

There are about 35 living animal phyla. To understand the origin and evolution of

any feature found in one or more of these groups, it is necessary to have a picture of the phylogenetic relationships among animals. Ideally, the fossil record would present a complete, ordered, unambiguous picture of the branching pattern of the animal tree. Unfortunately, it does not. As the divergence of most bilaterian phyla appears to have predated the emergence of recognizable members of modern phyla in the fossil record, we must make our inferences from later, more derived forms.

Constructing an accurate picture of **metazoan** relationships has been challenging, and many alternative schemes of animal phylogeny have been proposed and scrutinized over recent decades and continue to be evaluated. Most approaches have relied on anatomical and embryological comparisons. In general, phylogenies are determined according to shared characters that are presumed to be derived and therefore reflect a close relationship. For example, all animal phyla are thought to be more closely related to each other than to any other nonanimal phylum, because of similarities in animal multicellularity, cell structure and morphology, and cell signaling. Members of the most closely related protist group, the choanoflagellates, share a similar cell architecture with sponges but are not multicellular. What is most difficult to determine is whether apparent similarities between animals (for example, segmentation in arthropods and annelids) are due to common ancestry, are superficial, or evolved independently. Also, different tree topologies can emerge when different characters are used or when the same characters are weighted differently.

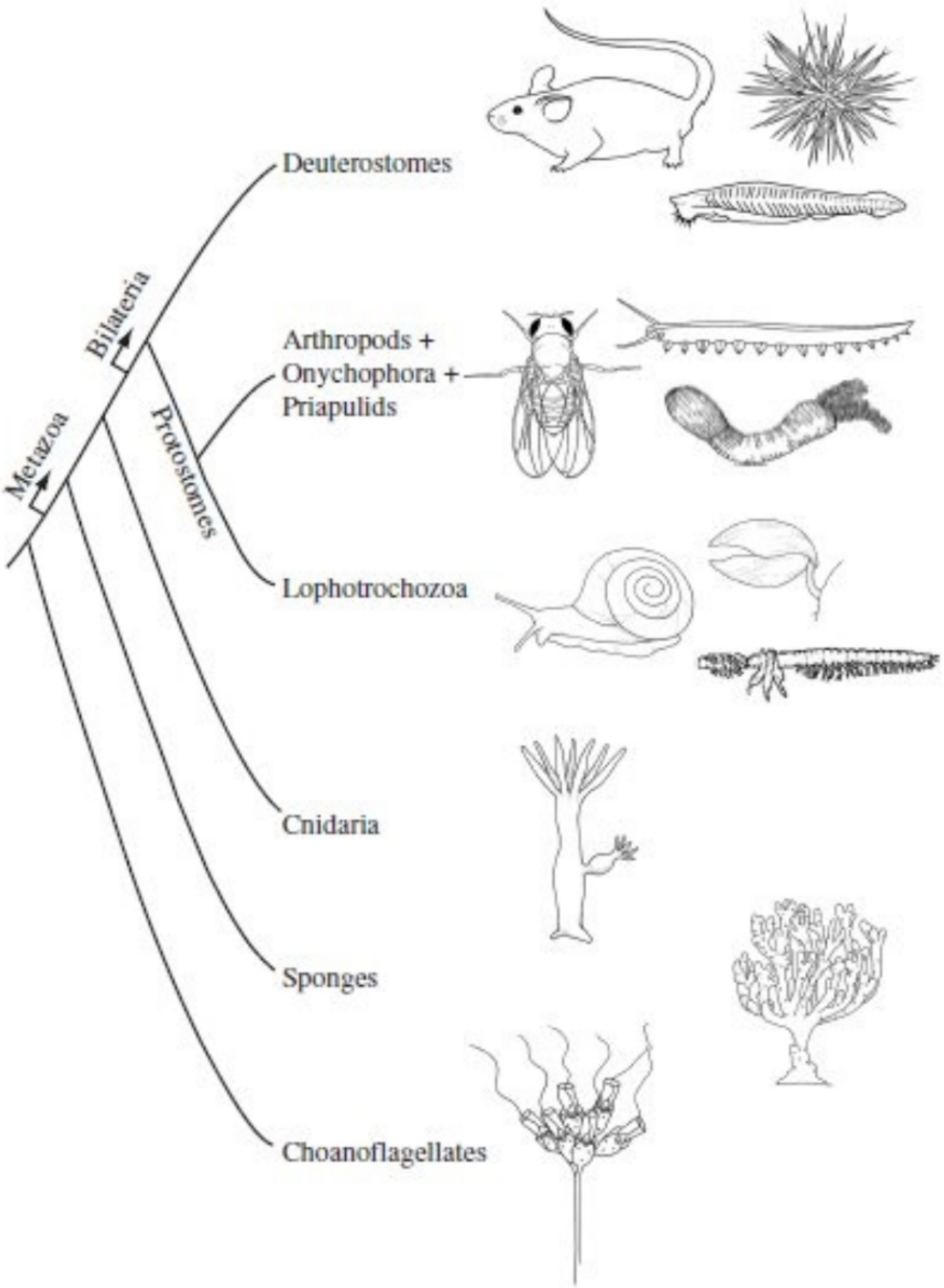
One way to circumvent the reliance on morphological comparisons is to use molecular genetic characters to construct animal phylogenies. As taxa diverge, the sequences of DNA, RNA, and protein molecules diverge as well; the relative degree of divergence can therefore be used to infer phylogenetic relationships. In addition, the presence or absence of particular genes, or the linkage of a group of genes on chromosomes, can be used to construct **phylogenetic trees**. New methods based on molecular sequences have been combined with morphology-based approaches to both prune and strengthen the animal tree.

We now recognize shared morphological, developmental, and genetic traits that suggest that the Bilateria can be organized into three great **clades** (a set of species descended from a common ancestor) (Fig. 1.4):

- The **deuterostomes**, including chordates, echinoderms, ascidians, and hemichordates. The deuterostomes are named for a shared feature of early embryonic development in which the mouth forms from a site separate from the **blastopore**, an opening in the early embryo.
- Two groups of **protostomes**, in which the mouth develops from the blastopore. The protostomes are divided into the **lophotrochozoans**,

including annelids, molluscs, and brachiopods, many of which share a trochophore larval stage in their life cycle, and a clade consisting of the arthropods, onychophora, and priapulids.

Figure 1.4 Metazoan phylogeny The current picture of metazoan phylogeny showing representatives of three major bilaterians clades—the deuterostomes, the Lophotrochozoa, and the arthropod + onychophora + priapulid clade.



Within these great clades, the branching order has been less well resolved, such that it is unclear which phyla are more closely related. It is worth noting that the recent assignment of arthropods and annelids to two different protostome clades and the assignment of pseudocoelomate phyla among different clades are major changes from previous portraits of the animal tree. The phylogenetic

placement of the nematodes, including the model organism *Caenorhabditis elegans*, remains controversial, because their rapid molecular clock complicates analysis. Some phylogenies place the nematodes close to the arthropod + onychophora + priapulid clade and others more basally near the common ancestor of all bilaterian phyla.

The anatomical and developmental features of the Bilateria are very distinct from those of the basal metazoans (cnidara, ctenophores, and porifera). The evolutionary links between basal metazoans and the bilaterians are difficult to perceive. Indeed, as we will see in Chapter 4, major differences exist between the genetic toolkit of these two groups, and the differences are much more substantial than those between most bilaterians. Because of the long divergence time since the radiation of these groups, the phylogenetic relationships between cnidarians, sponges, and ctenophores and the last common ancestor of the Bilateria are uncertain. Many extinct animal lineages, as yet unknown from the fossil record, may have branched off of the metazoan tree between the last common ancestor of all animals and of the Bilateria (Fig. 1.4).

The gaps in the fossil record; the great differences in anatomy, development, and genome content between radially symmetrical animals and bilaterians; and the cryptic early history of bilaterians, make inferences about the morphological transformations involved in the origin of animal body plans very speculative. Paleontologists have introduced the concept of **disparity** to refer to differences among body plans and use the term **diversity** to refer to the number of species within a group. The genetic and developmental bases of the morphological diversification of a *particular* body plan within a phylum are far more accessible than is the origin of *different* body plans. Therefore, we will focus primarily on evolutionary trends within a few select phyla, such as the arthropods and chordates, making the implicit assumption that the same sort of genetic mechanisms involved in the evolution of large-scale morphological diversity within phyla also gave rise to fundamental differences in body plans.

GENERAL FEATURES OF ANIMAL DESIGN AND DIVERSITY

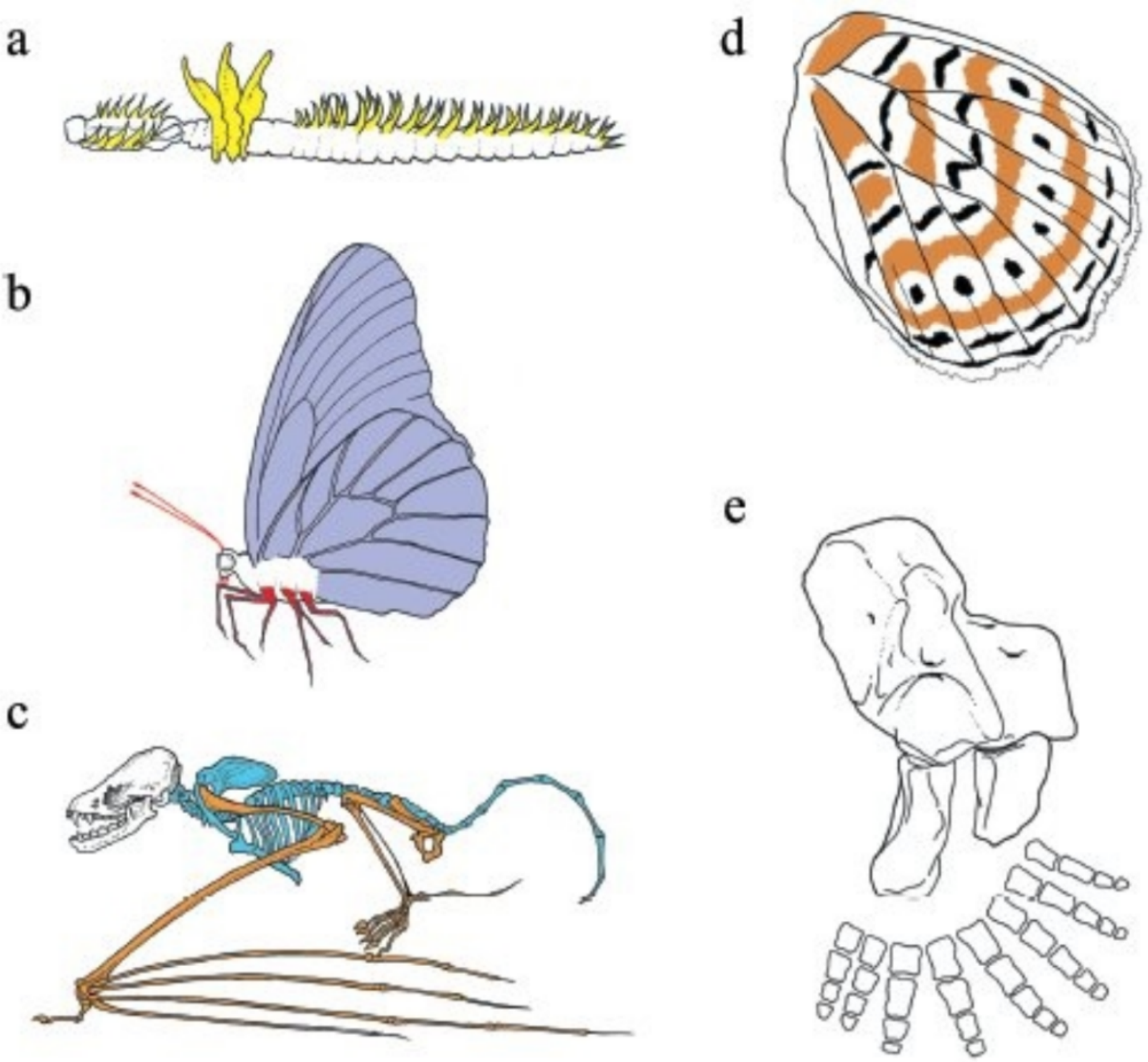
One of the most outstanding features of animal design, particularly of larger

bilaterians, is their construction from repeating structures (or modules). The segments of arthropods and annelids and the vertebrae (and associated processes) of vertebrates are the basic units of body plan organization in these phyla (Fig. 1.5a–c). Similarly, many body parts such as the insect wing (Fig. 1.5d) and the **tetrapod** hand (Fig. 1.5e) are composed of repeated structures.

An important trend in the morphological evolution of animals has been the **individualization** of modular elements. For example, among the arthropods, we observe a large number of different segment types in crustaceans and insects. This diversity far exceeds that found in the **onychophora**, a phylum closely related to the arthropods. Thus the evolution of the onychophoran/arthropod clade has been marked by increased diversity of segment types from the more uniform patterns found in earlier forms. Similarly, in some mammals, teeth are differentiated into molars, premolars, canines, and incisors, whereas in the ancestral condition exhibited by most reptiles, the teeth are of uniform shape. Because the diversification of the number, morphology, and function of these repeated units characterizes many of the large-scale differences that distinguish related taxa, understanding how repeated structures form and become individualized is a prerequisite for understanding the developmental basis of large-scale morphological evolution.

Figure 1.5 The modularity of body plans and body parts The body plans of many major phyla, including the annelids (**a**), arthropods (**b**), and chordates (**c**), are composed of many repeating parts. Some of these parts are similar or identical in appearance to other parts; others are individuated. Sets of serially homologous structures are shaded a unique color. Body parts, such as a butterfly wing (**d**), or a fossil tetrapod limb from the amphibian fossil *Acanthostega* (**e**), are also composed of repeating structures or patterns, some of which are differentiated from others. For example, *Acanthostega* has eight digits, but like its modern descendants, only five distinct types of digits can be distinguished.

Source: Parts a–c from Weatherbee SD, Carroll SB. Selector genes and limb identity in arthropods and vertebrates. *Cell* 1999; 97: 283–286; part e from Michael Coates.



The modular organization of animal bodies and body parts has long been recognized by comparative biologists. William Bateson, in his classic treatise *Materials for the Study of Variation* (1894), identified several kinds of organization found among animals. More importantly, he was the first to bring a Darwinian perspective to the question of how different body patterns may have evolved. Bateson focused particularly on the repetition of parts, cataloguing a large number of rare, but naturally occurring, variants that differed from the norms within various species with regard to either the number or individualization of characters. He suggested that these variations within species could provide insight into the evolution of the large-scale morphological discontinuities between species. For example, variations in the number of body segments within onychophora and centipede species, and of vertebrae in humans and pythons, suggested to

Bateson that such discontinuities arose at some frequency in populations and therefore represented plausible steps in the morphological diversification of species.

The question of whether evolution may progress in large, discrete steps remains controversial (we will address this issue in Chapter 8). Nevertheless, these sorts of variants and the organizational concepts espoused by Bateson have been enormously helpful in understanding the genetics and developmental logic underlying the **modularity** of animal design. In fact, they led to the discovery of genes that play key roles in morphological evolution, albeit not in the fashion Bateson first imagined.

Four fundamental kinds of large-scale, evolutionary differences in morphology are most prevalent in modularly organized animals and are the most significant in terms of adaptation:

1. *Changes in the number of repeated parts* Bateson referred to this type of change as **meristic variation** when describing differences within species. Differences in segment number and vertebral number are some of the most obvious characteristics that distinguish classes of arthropods and various classes and orders of vertebrates, respectively (Fig. 1.6).

2. *Diversification of serially homologous parts* A series of reiterated parts are termed **serially homologous**. The individualization of repeated parts in an animal reflects the diversification of serially homologous structures. For example, arthropod appendages are serially homologous structures. In the course of arthropod evolution, ancestrally similar appendages have evolved into antennae, various mouthparts, walking legs, and genital structures. In vertebrates, serially homologous vertebrae have evolved into distinct cervical, thoracic, lumbar, and sacral vertebral types.

3. *The diversification of homologous parts* One of the most prevalent trends in animal evolution is the morphological diversification of **homologous** parts between lineages. The same structures in different lineages are termed "homologous" when they share a common history, even if they no longer serve the same function. For example, all tetrapod forelimbs are homologous (Fig. 1.7). Despite their differing appearances and functions, bird wings, bat wings, and human forelimbs have all conserved the basic architecture of the tetrapod forelimbs.

4. *The evolution of* ***novelties*** New characters or "novelties" may evolve from a preexisting structure or arise *de novo* and become adapted to a new purpose. The evolution of feathers, fur, teeth, antlers, and butterfly wing eyespots are examples of such morphological novelties.

Figure 1.6 Meristic differences among arthropods and among vertebrates Among arthropods such as this trilobite (**a**), crustacean (**b**), centipede (**c**), and insect (**d**), the number of body segments differs, as does the diversity of segment morphology. Among vertebrates, the number of vertebrae and associated processes differs considerably between a fish (**e**), frog (**f**), python (**g**), and chimpanzee (**h**).

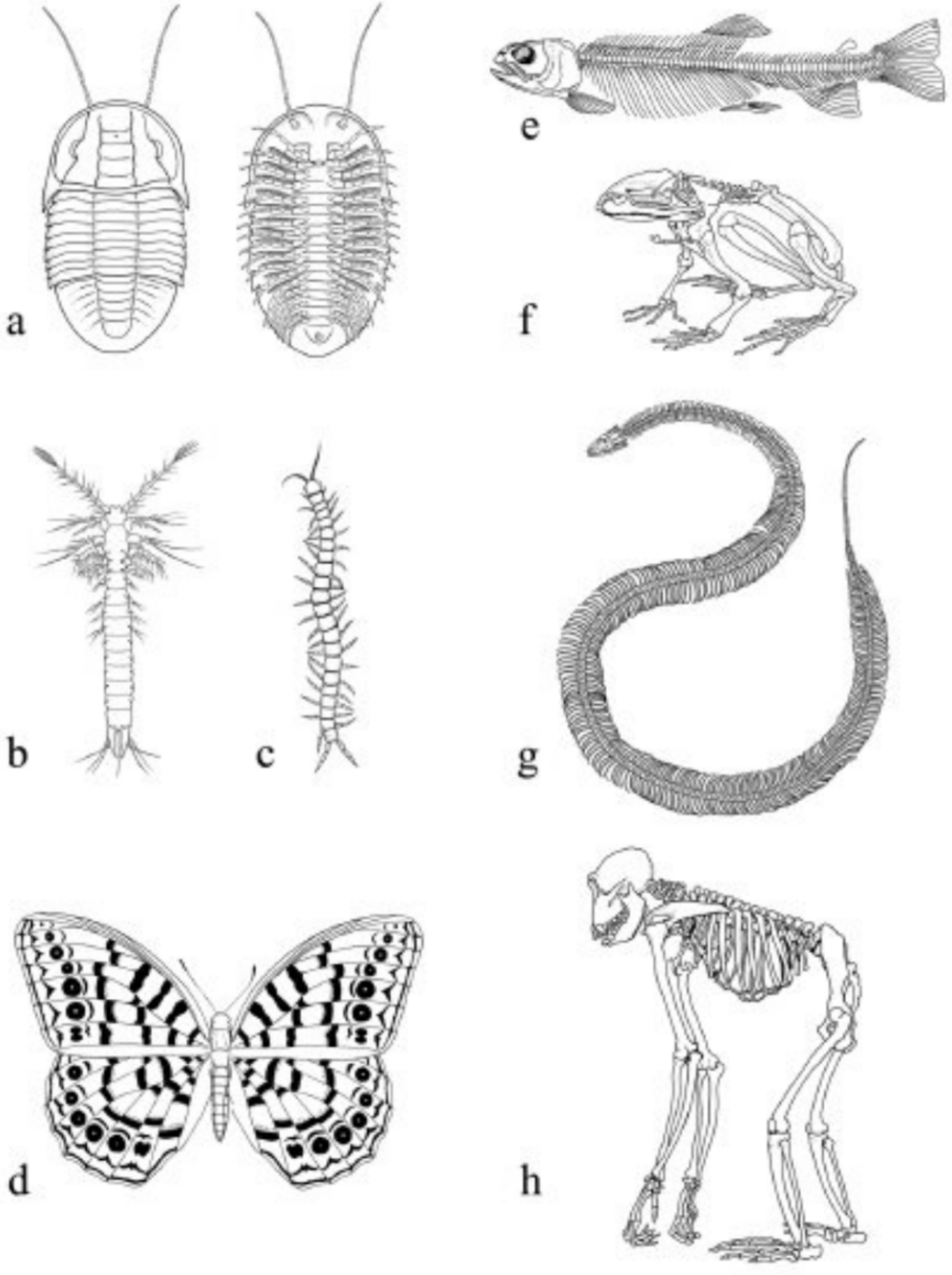
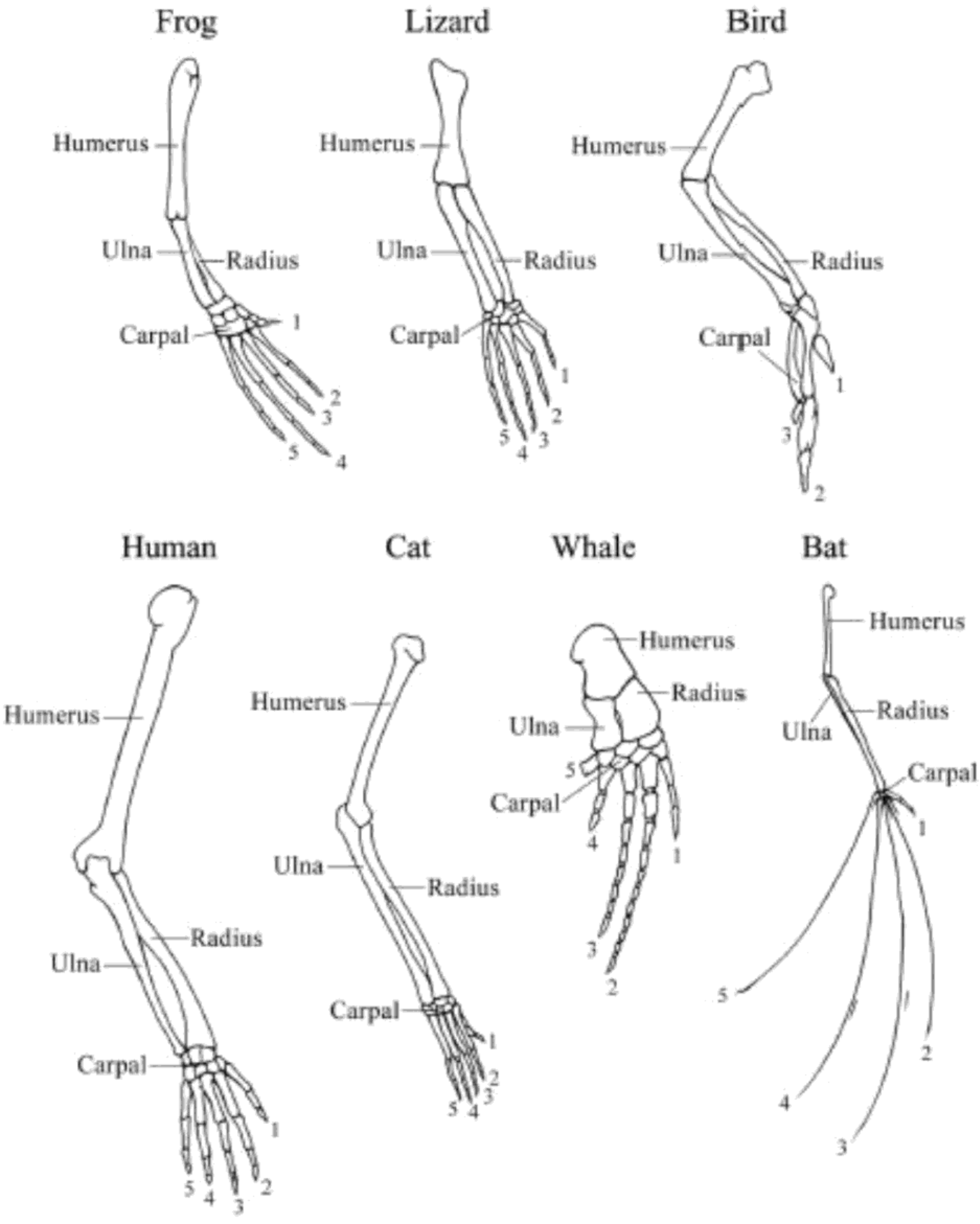


Figure 1.7 The diversification of homologous parts All vertebrate forelimbs are homologous structures whose anatomy has undergone considerable diversification in the evolution and adaptation of these various vertebrate lineages. Not to scale.

Source: Redrawn from Ridley M. *Evolution*, 2nd edn. Malden, MA: Blackwell Science, 1996.



Considering that modularly organized animals are among the most diverse groups (in terms of both the number and morphology of species), could there be a correlation between body design and evolutionary diversity? One possible explanation for this relationship is that modular organization allows one part of the

animal to change without necessarily affecting other parts. The evolution of genetic mechanisms that control the individualization of parts would allow for the uncoupling of developmental processes in one part of the body from the developmental processes in another part of the body. In this fashion, for example, vertebrate forelimbs can evolve into wings while hindlimbs remain walking legs. Dissociation of the forelimb and hindlimb developmental programs allows further modifications to occur selectively in either structure, such as the development of feathers in the forelimb of birds and scales in the hindlimb.

EVOLUTION AND DEVELOPMENT: DNA AND DIVERSITY

To understand the major trends in animal diversity and the various kinds of morphological evolution, we must first understand how animal form is generated. Morphology is the product of development, the process through which a single fertilized egg cell gives rise to an entire organism. The physical basis of animal diversity has been viewed since Darwin's time as the outcome of development. Until very recently, however, the developmental principles underlying animal design remained unknown. Although experimental embryologists of the late 1800s and the first half of the 1900s had identified many fascinating phenomena concerning the organization of embryos and the formation of particular structures, the mechanisms responsible for these properties were beyond their reach.

With better understanding of the nature of genes and the process of gene regulation, development has been increasingly viewed as a process orchestrated by the products of genes. Thus the puzzles of embryology, such as how cells come to know their position and identity within a developing animal, have become rephrased in genetic terms. Given that the DNA of (most) all cells in an animal is identical, how do different cells acquire the unique morphologies and functional properties required in the diverse organs and tissues of the body? We now understand that this process occurs through the selective expression of distinct subsets of the many thousands of genes in any animal's genome in different cells. How genes are turned on and off in different cells over the course of animal development is an exquisitely orchestrated regulatory program whose features are only now coming into detailed view.

If morphological diversity is all about development, and development results from genetic regulatory programs, then is the evolution of diversity directly related to the evolution of genetic regulatory programs? Simply put, yes. But to understand how diversity evolves, we must first understand the genetic regulatory mechanisms that operate in development. In other words, what is the genetic toolkit of development and how does it operate to build animals? In the next two chapters, we will examine some of the general features of the genetic and regulatory logic of animal development.

SELECTED READINGS

The Early Fossil Record

Chen JY, Oliveri P, Gao F, et al. Precambrian animal life: probable developmental and adult cnidarian forms from Southwest China. *Dev Biol* 2002; 248, 182–196.

Chen JY, Oliveri P, Li CW, et al. Precambrian animal diversity: putative phosphatized embryos from the Doushantuo Formation of China. *Proc Natl Acad Sci USA* 2000; 97: 4457–4462.

Conway MS. The fossil record and the early evolution of the Metazoa. *Nature* 1993; 361: 219–225.

_____. *The crucible of creation: the Burgess shale and the rise of animals*. Oxford, UK: Oxford University Press, 1998.

Fortey R. The Cambrian evolutionary “explosion”: decoupling cladogenesis from morphological disparity. *Biol J Linnean Soc* 1996; 5713–5133.

Glaessner MF. *The dawn of animal life: a biohistorical study*. Cambridge, UK: Cambridge University Press, 1984.

Gould SJ. *Wonderful life: the Burgess shale and the nature of history*. New York: Norton, 1989.

Knoll AH, Carroll SB. Early animal evolution: emerging views from comparative biology and geology. *Science* 1999; 284: 2129–2137.

Valentine JW, Jablonski D, Erwin DH. Fossils, molecules and embryos: new perspectives on the Cambrian explosion. *Development* 1999; 126: 851–859.

Molecular Clocks and Animal Evolution

Ayala FJ, Rzhetsky A, Ayala FJ. Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. *Proc Natl Acad Sci USA* 1998; 95: 606–611.

Wray G, Levinton J, Shapiro L. Molecular evidence for deep precambrian divergences among metazoan phyla. *Science* 1996; 274: 568–573.

Animal Phylogeny

Adoutte A, Balavoine G, Lartillot N, de Rosa R. Animal evolution: the end of the intermediate taxa? *Trend Genet* 1999; 15: 104–108.

Aguinaldo AM. Evidence for a clade of nematodes, arthropods, and other moulting animals. *Nature* 1997; 387: 489–493.

Blair JE, Ikeo K, Gojobori T, Hedges SB. The evolutionary position of nematodes. *BMC Genomics* 2002; 2: 7.

King N, Carroll SB. A receptor tyrosine kinase from choanoflagellates: molecular insights into early animal evolution. *Proc Natl Acad Sci USA* 2001; 98: 15032–15037.

Maley LE, Marshall CR. The coming of age of molecular systematics. *Science* 1998; 279: 505–506.

Snell EA, Furlong RF, Holland PW. Hsp70 sequences indicate that choanoflagellates are closely related to animals. *Curr Biol* 2001; 11: 967–970.

Animal Design and Evolution: Modularity and Homology

Bateson W. *Materials for the study of variation*. London: Macmillan, 1894.

Gilbert S, Opitz J, Raff R. Resynthesizing evolutionary and developmental biology. *Dev Biol* 1996; 173: 357–372.

Raff R. *The shape of life*. Chicago: University of Chicago Press, 1996.

Wagner GP. Homologues, natural kinds and the evolution of modularity. *Am Zoologist* 1996; 36: 36–43.

_____. The origin of morphological characters and the biological basis of homology. *Evolution* 1989; 43: 1157–1171.

CHAPTER 2

The Genetic Toolkit for Development

The only way in which we may hope to get at the truth is by the organization of systematic experiments in breeding, a class of research that calls perhaps for more patience and more resources than any other form of biological inquiry. Sooner or later such investigation will be undertaken and then we shall begin to know.

— W. Bateson Material for the Study of Variation (1894)

… if the mystery that surrounds embryology is ever to come within our comprehension, we must … have recourse to other means than description of the passing show.

— T.H. Morgan Experimental Embryology (1927)

The foremost challenge for embryology has been to identify the genes and proteins that control the development of animals from an egg into an adult. Early embryologists discovered that localized regions of embryos and tissues possess properties that have long-range effects on the formation and patterning of the primary body axes and appendages. Based on these discoveries, they postulated the existence of substances responsible for these activities. However, the search for such molecules proved fruitless until the relatively recent advent of genetic and molecular biological technologies. The most successful approach to understanding normal development has involved the isolation of single gene mutations that have discrete and often large-scale effects on body pattern.

In this chapter, we take an inventory of the essential genetic toolkit for animal development. We concentrate on genes first discovered in insects, where systematic screens for developmental genes were pioneered. Importantly, however, it turns out that related genes are present in many other animals. We describe how members of the genetic toolkit were identified and what kinds of gene products they encode. In addition, we illustrate the general correlation between these genes' patterns of expression with the development of the morphological features they affect. Finally, we briefly survey their distribution and function in other animals.

Only a small fraction of all genes in any given animal constitute the toolkit that is devoted to the formation and patterning of the body plan and body parts. Two classes of gene products with the most global effects on development are of special interest: families of proteins called transcription factors that regulate the expression of many other genes during development, and members of signaling pathways that mediate short- and long-range interactions between cells. The expression of specific transcription factors and signaling proteins marks the location of many classically defined regions within the embryo. These proteins control the formation, identity, and patterning of most major features of animal design and diversity.

BEFORE THE TOOLKIT—ORGANIZERS, FIELDS, AND MORPHOGENS

Long before any genes or proteins affecting animal development were characterized, embryologists sought to identify the basic principles governing animal design. In their search, they focused on the large-scale organization of the primary body axes, the differentiation of various **germ layers** (**ectoderm**, **mesoderm**, and **endoderm**), and the polarity of structures such as appendages and insect segments. By manipulating embryos and embryonic tissues, primarily by transplantation and ablation, researchers discovered many important properties of developing embryos and tissues. Much of the fascination of embryology stems from the remarkable activities of discrete regions within developing embryos in organizing the formation of body axes and body parts. Furthermore, these classical concepts of embryonic organization present a very useful framework for considering how that organization can change during evolution. We will briefly review some of these experiments and ideas before addressing their genetic and molecular manifestations.

The first demonstration of **organizers**—regions of embryos or tissues that have long-range effects on the fate of surrounding tissues—was achieved by Mangold and Spemann in 1924. They transplanted the lip of the blastopore, the invagination where mesoderm and endoderm move inside the amphibian embryo, of a newt gastrula into another newt embryo and found that the transplanted tissue could induce a second complete body axis (Fig. 2.1a). The additional embryo induced was partly derived from the transplanted graft and partly derived from the host. The equivalent of the "Spemann organizer" in amphibians has been found in chick and mouse embryos, and it is now recognized to be a structure characteristic of all chordate embryos.

Other organizers with long-range effects on surrounding tissues have been identified in the developing vertebrate limb bud. Transplantation of a discrete patch of posterior tissue to an ectopic anterior site induces the formation of limb structures (digits, tendons, muscles) with mirror-image polarity to the normal anteroposterior order (Fig. 2.1b). By contrast, transplantation or removal of anterior tissue has no effect on limb development, suggesting that this posterior region of the limb bud, dubbed the **zone of polarizing activity** (ZPA), organizes anteroposterior (that is, the thumb-to-pinkie axis) polarity and limb formation.

Another organizer operates from the most distal tip of the limb bud, the **apical ectodermal ridge** (AER). Removal of this region truncates the limb and deletes distal elements (digits), whereas transplantation of the AER to an early limb bud can induce outgrowth of a duplicate limb (Fig. 2.1b).

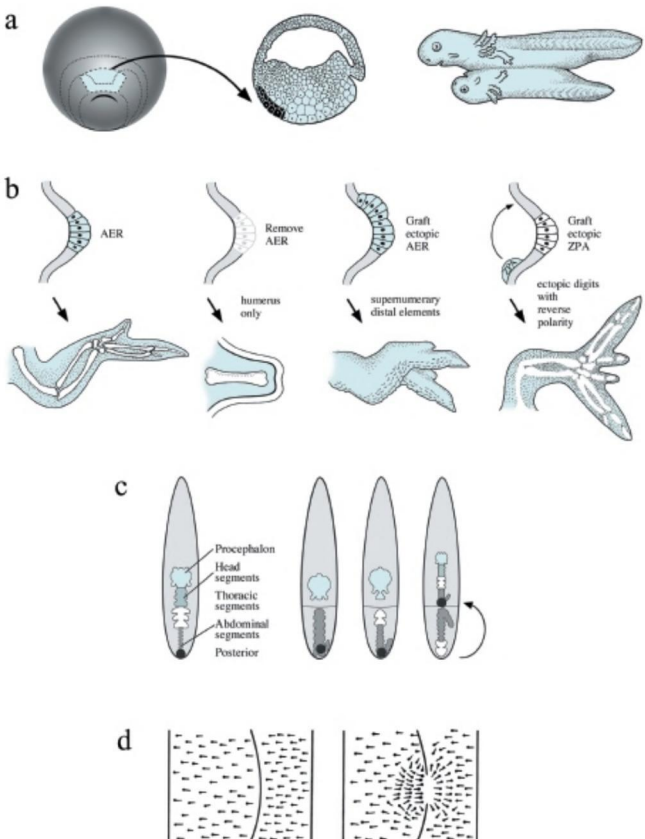
One explanation for the long-range polarizing and inductive effects of the Spemann organizer, ZPA, and AER is that these tissues are sources of inducer molecules, or **morphogens**—that is, substances whose concentrations vary within a tissue and to which surrounding cells and tissues respond in a

concentration-dependent manner. The response to a morphogen depends, then, on the distance of the responding tissue from the source. For example, if the ZPA is a source of a morphogen, then diffusion of this substance can establish a **gradient** of inducer concentration. Induction of different digit types depends on the morphogen concentration, with low levels of morphogen inducing anterior digits (thumb) and high levels inducing posterior digits (pinkie) (Fig. 2.1b).

Organizers have been demonstrated and morphogens postulated in insects as well as vertebrates. Ligature and cytoplasmic transplantation experiments first suggested that the anteroposterior axis of certain insect embryos is influenced by two organizing centers, one at each pole of the egg (Fig. 2.1c), that behave as sources of morphogens. Similarly, the polarity of cells within insect segments appears to be organized by signals that produce a graded pattern (Fig. 2.1d).

Figure 2.1 Organizers in animal embryos Transplantation and ablation experiments have been used to investigate the long-range organizing activities of embryonic tissues. (**a**) The Spemann organizer. The dorsal blastopore lip of an early amphibian embryo can induce a second embryonic axis and embryo when transplanted to the ventral region of a recipient embryo. (**b**) Limb organizers. The apical ectodermal ridge (AER) is required for formation of distal limb elements. Removal leads to loss of structures; transplantation to specific ectopic sites induces extra elements. The zone of polarizing activity (ZPA) organizes the anteroposterior pattern; transplantation to an ectopic site induces extra digits with reverse polarity. (**c**) Insect egg organizer. Ligation of the insect *Euscelis* embryo (marked by the gray line) early in development deletes the thorax and abdomen; later ligations leave more segments intact. However, transplantation of the posterior pole cytoplasm (marked by the black dot) into the anterior of a ligated embryo induces the formation of a complete embryo. This result demonstrates that the posterior cytoplasm has organizer activity. (**d**) Within insect segments, epithelial polarity is organized by signaling sources. Ablation of a segment boundary (indicated by the interruption of the black line) reorganizes segment polarity (indicated by the orientation of small black hairs).

Source: Parts a–c redrawn from Gilbert S. *Developmental biology*, 5th edn. Sunderland: Sinauer Associates, 1997; part redrawn d from Lawrence PA. *The making of a fly*. Oxford, UK: Blackwell Science, 1992.



One difficulty with this picture of morphogen-producing organizers arises when

we attempt to explain the boundaries of their range of influence. All of the cells in a growing embryo are in contact with other cells, so how is it that some parts respond and others do not? One explanation involves the concept of the **morphogenetic field.** Early embryologists demonstrated that some parts of developing animals, such as the forelimb field, could be transplanted to another site and still differentiate properly—that is, into a forelimb. In addition, if undetermined cells were introduced into the field, they could become incorporated into the limb. These transplantable, self-regulating fields are discrete physical units or modules of embryonic development. They form bounded domains within which specific programs of morphogenesis occur. The term "primary field" applies to the entire embryo before the axes are determined; the limbs, eyes, and other organs are termed "**secondary fields**," or organ **primordia.**

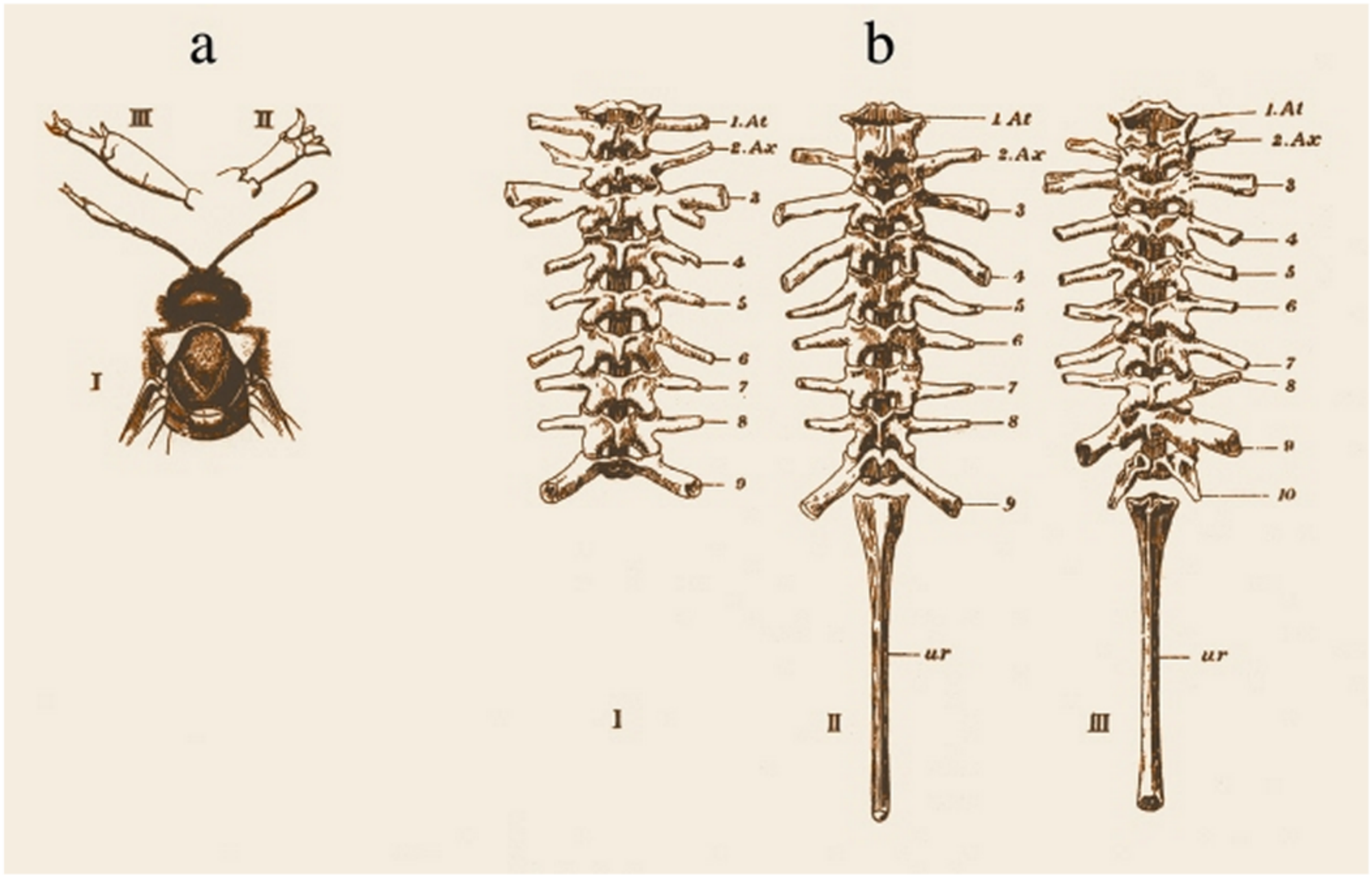
Secondary fields may be further subdivided into "tertiary fields," defined by physical or developmental boundaries. **Compartments** are one special type of subdivision. First demonstrated within the wing **imaginal disc** of the fruit fly *Drosophila melanogaster*, compartments are composed of populations of cells that do not intermix with cells outside the compartment.

Further progress in understanding the nature of organizers, morphogens, and fields stalled after their discovery and description in the first half of the 1900s. The impasse was ultimately broken by the discovery of genes whose products governed the activity of organizers, behaved as morphogens, and controlled the formation and identity of embryonic fields. These genes make up the "toolkit" for animal development.

THE GENETIC TOOLKIT

Animal genomes contain thousands of genes. Many of these genes encode proteins that function in essential processes in all cells in the body (for example, metabolism, biosynthesis of macromolecules) and are often referred to as "**housekeeping genes**." Other genes encode proteins that carry out specialized functions in particular cells or tissues within the body (for example, oxygen transport, immune defense) or, to extend the housekeeping metaphor, in specific "rooms" in the "house." But here we are interested in a different set of genes, those whose products govern the construction of the house—the toolkit that determines the overall body plan and the number, identity, and pattern of body parts.

Toolkit genes have generally first been identified based on the catastrophes or monstrosities that arise when they are mutated. Two sources of toolkit gene

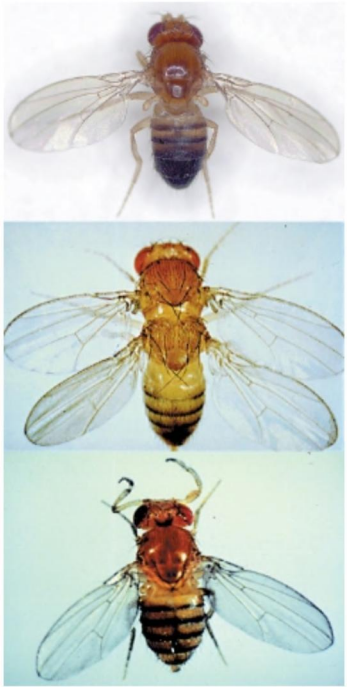


Intriguing as Bateson's specimens were, most were one-of-a kind museum pieces in which only one member of a bilateral pair of structures was affected. To carry out a thorough investigation of the phenomenon of homeosis and its genetics, researchers required mutants that would breed true in subsequent generations. In 1915, Calvin Bridges isolated a spontaneous mutation in *Drosophila*, dubbed *bithorax*, in which part of the haltere (the posterior flight appendage in flies) was transformed into wing tissue. The haltere and wing are serially homologous appendages, so the *bithorax* mutation causes the partial transformation of the identity of a structure on the third thoracic segment (the haltere) into its serial homolog found on the second thoracic segment (the wing). A more complete transformation of the entire haltere into a wing can occur if additional mutations are combined with *bithorax*, producing a four-winged fly (Fig. 2.3).

In the following decades, several more homeotic mutants were identified in *Drosophila*, and in other insects as well. All of these homeotic mutations transform the identities of segments and their associated structures into those of other segments. For example, certain *Antennapedia* mutations cause the transformation of antennae into legs (Fig. 2.3), which are also serial homologs. The direction of the homeotic transformations depends on whether a mutation causes a loss of homeotic gene function where the gene normally acts, or a gain

of homeotic gene function in places where the homeotic gene does not normally act. For example, *Ultrabithorax* (*Ubx*) acts in the **haltere** to promote haltere development and repress wing development. Loss-of-function mutations in *Ubx* transform the haltere into a wing. Dominant mutations that cause *Ubx* to gain function in the wing transform that structure into a haltere. Similarly, the antenna-to-leg transformations of *Antennapedia* mutants reflect a dominant gain of *Antennapedia* gene function in the antenna.

Figure 2.3 Homeotic mutants of *Drosophila melanogaster* (**top**) Normal fly with one pair of wings on T2 and halteres on T3. (**middle**) Triple mutant for three mutations in the *Ultrabithorax* gene abolishes *Ubx* function in the posterior thorax and causes the appearance of an extra set of wings (transformation of T3 → T2 identity). (**bottom**) *Antennapedia* mutant in which the antennae are transformed into legs.



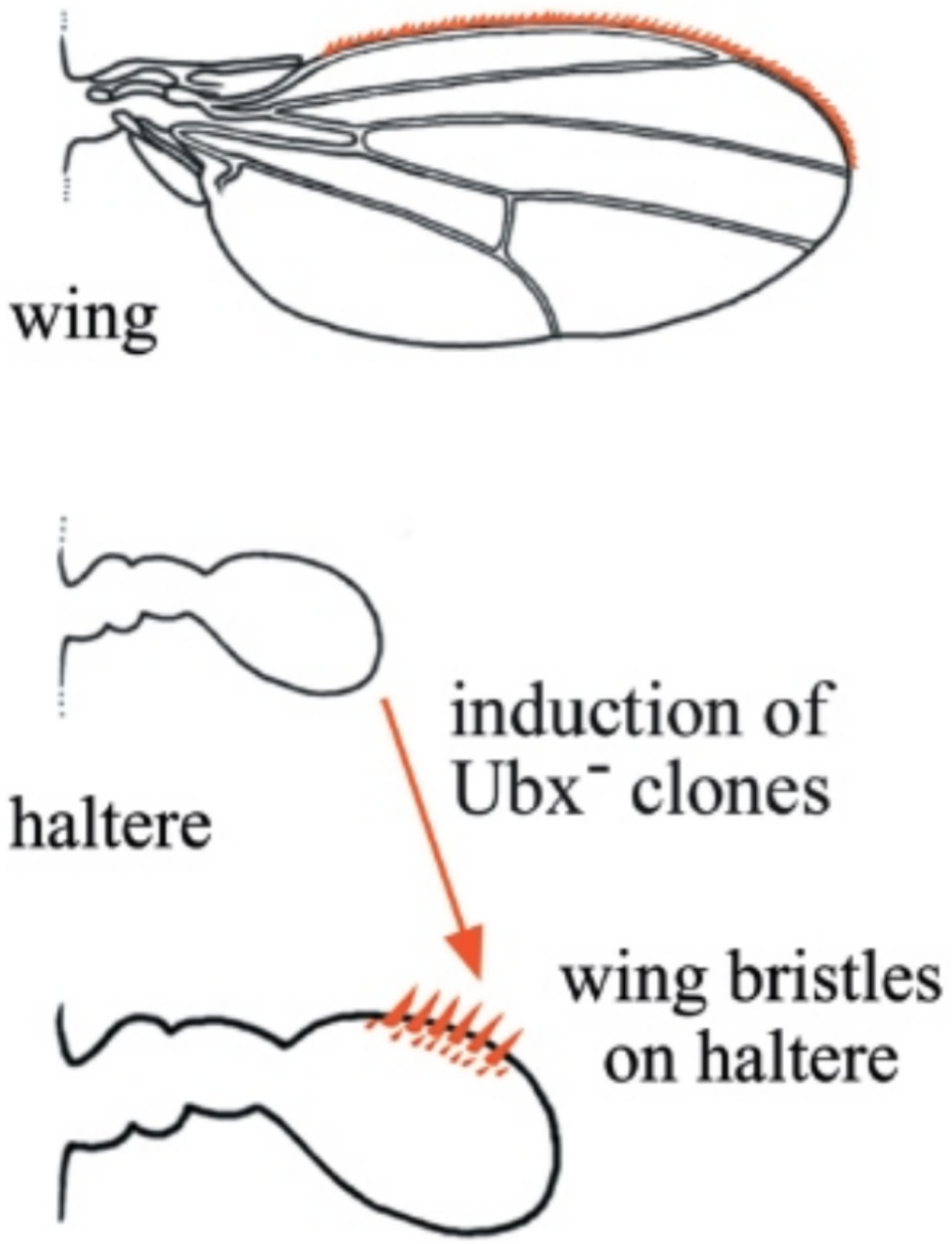
The fascination with homeotic mutants stems from two issues. First, it is startling that a single gene mutation could change entire developmental pathways so dramatically in a complex animal. Second, it is curious that the structure formed in the mutant is a well-developed likeness of another body part.

More detailed understanding of homeotic gene function was made possible by some particularly ingenious methods for analyzing the effects of mutations on the behavior of a group of cells in otherwise normal (or “wild-type”) tissues. That is, rather than being limited to examining the effect of homeotic mutations on whole

animals, the behavior of clones of mutant cells could be observed within otherwise normal animals (Fig. 2.4). This technique was used to determine that the effects of homeotic mutations generally remain limited to cells with mutant genotypes; such behavior is termed **cell autonomous.** Thus a patch of cells in the haltere that lacks *Ubx* function forms wing tissue, even when it is surrounded by normal haltere cells (Fig. 2.4). This finding suggested that homeotic genes act within cells to select their developmental fate. Homeotic genes, and other genes with analogous functions in controlling cell fate, are therefore known as **selector genes.**

Figure 2.4 Cell autonomy of homeotic mutations The *Drosophila* wing and haltere have different pattern elements, such as the occurrence of sensory bristles at the leading edge of the wing (red). Clones lacking *Ubx* function in the haltere form wing structures (for example, the sensory bristles shown in red) in positions corresponding to those of the wing.

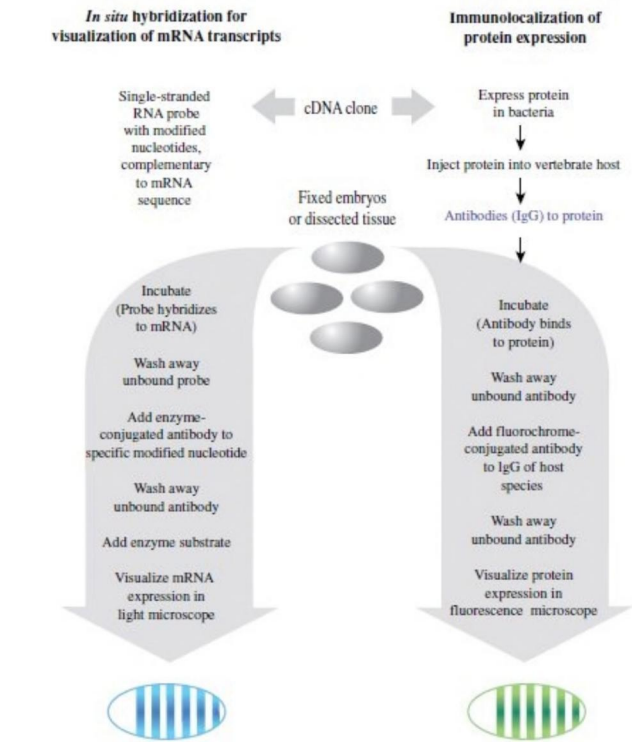
Source: Redrawn from Lawrence PA. *The making of a fly*. Oxford, UK: Blackwell Scientific, 1992.



Although homeotic genes were first identified through spontaneous mutations affecting adult flies, they are required throughout most of *Drosophila* development to determine segmental identity. Systematic screening for homeotic genes led to the identification of eight linked genes, collectively referred to as ***Hox* genes**, that affect the specification of particular segment identities in the developing *Drosophila* embryo, larva, and adult. In addition to *Ultrabithorax* (*Ubx*) and *Antennapedia* (*Antp*), they include *labial* (*lab*), *proboscipedia* (*pb*), *Deformed (Dfd)*, *Sex combs reduced* (*Scr*), *abdominal-A* (*abd-A*), and *Abdominal-B* (*Abd-B*). Generally, the complete loss of any *Hox* gene function causes transformations of segmental identity and is lethal in early development. The spontaneous homeotic mutants found in viable adults are caused by partial loss of gene function or are dominant such that in heterozygotes normal gene function is provided by the wild-type allele.

One of the most intriguing features of these *Hox* genes is that they are linked in

Figure 2.6 Methods for visualizing gene expression in developing animals
The two most common means of visualizing where a gene is transcribed and its protein product is synthesized are (**left**) *in situ* hybridization of complementary RNA probe to mRNA and (**right**) immunolocalization of protein expression. The procedures for each method are indicated. Gene expression patterns are visualized as the product of enzymatic reactions (**left**) or with fluorescently labeled compounds (**right**).



The relationship between the structure of *Hox* gene complexes and the phenotypes of *Hox* mutants was illuminated by the molecular characterization of both the Bithorax Complex and the Antennapedia Complex. Cloning of the *Hox* genes provided the means to uncover when and where each of the eight genes is expressed during development. The ability to visualize *Hox* and other gene expression patterns during development was crucial to understanding the

correlation between gene function and phenotypes. Localization of *Hox* genes' RNA transcripts by *in situ* hybridization or of Hox proteins via immunological methods (Fig. 2.6) revealed that all *Hox* genes are expressed in spatially restricted, sometimes overlapping domains within the embryo. These genes are also expressed in subsets of the developing larval imaginal discs, which proliferate during larval development and differentiate during the pupal stages to give rise to the adult fly.

The patterns of *Hox* gene expression generally correlate with the regions of the animal affected by homeotic mutations. For example, the *Ubx* gene is expressed within the posterior thoracic and most anterior abdominal segments of the embryo (Fig. 2.7a). The development of these segments is altered in *Ubx* mutants. In larvae, *Ubx* is expressed in the developing haltere, but not in the developing wing (Fig. 2.7b–e). This expression correlates with the requirement for *Ubx* to promote haltere development and to suppress wing identity.

The boundaries of *Hox* gene expression in the *Drosophila* embryo are not segmental, but usually begin in the posterior part of one segment and extend to (or beyond) the anterior portion of the next-most-posterior segment, a unit dubbed a **parasegment**. In the various imaginal tissues of the developing adult, homeotic genes are often expressed in segmental domains. For example, flies have three pairs of legs, with one pair extending from each of the three thoracic segments. Each pair of adult legs has a distinctive morphology. Indeed, genetic analysis has shown that the morphology of the first legs is largely influenced by the *Scr* gene, the second legs by the *Antp* gene, and the third legs by the *Ubx* gene. These respective genetic requirements correlate with the respective patterns of homeotic gene expression in the developing imaginal legs.

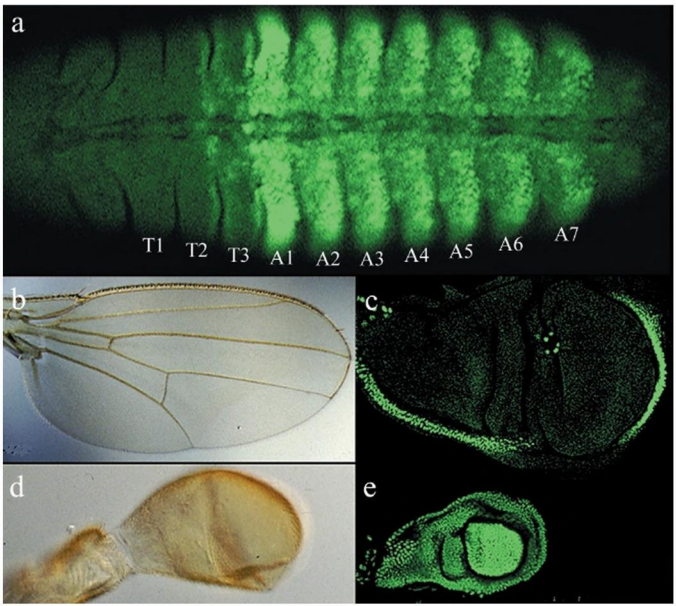
It is crucial to understand the distinction between *Hox* gene function in determining the *identity* of a field, as opposed to a requirement for *Hox* gene function in the *formation* of the field. The antennae, mouthparts, and walking limbs of flies all develop from serially homologous limb fields. In the absence of homeotic genes, each limb field develops, but with antennal identity. Therefore, *Hox* genes specify the particular identity, but are not required for the formation of the limb fields. The expression and function of *Hox* genes are not limited to body segments and their appendages. These genes act as **region-specific selectors** in all three germ layers (ectoderm, mesoderm, and endoderm) and in diverse structures and tissues.

The homeobox

The large effects of *Hox* genes on the developmental fates of entire segments and structures made the nature of the proteins encoded by these genes of special

interest. The close genetic linkage and similar function of *Drosophila Hox* genes suggested that they might have evolved through the tandem duplication of one or more ancestral *Hox* genes. This idea led to the discovery that the DNA sequences of the *Hox* genes of the Bithorax and Antennapedia Complexes were similar enough to hybridize to each other. This similarity was traced to a 180 base-pair (bp) stretch of DNA, dubbed the **homeobox**, that encodes a 60 amino acid protein domain (the **homeodomain**); the sequence of the homeodomain is very similar among the homeotic proteins (Fig. 2.8). The structure of the homeodomain resembles the DNA-binding domain of many prokaryotic regulatory proteins, suggesting that homeotic gene products exert their effects by controlling gene expression during development and that the homeodomain binds to DNA in a sequence-specific manner.

Figure 2.7 ***Hox*** **gene expression** *Hox* gene expression is restricted to regions of the body and to particular structures. For example, (**a**) the Ubx protein (shown in green) is expressed in the posterior thoracic (T) and seven anterior abdominal (A) segments of the embryo. (**b**) The adult wing. (**c**) Ubx is not expressed in the cells of the wing imaginal disc. (**d**) The adult haltere. (**e**) Ubx is expressed in the cells of the haltere imaginal disc.



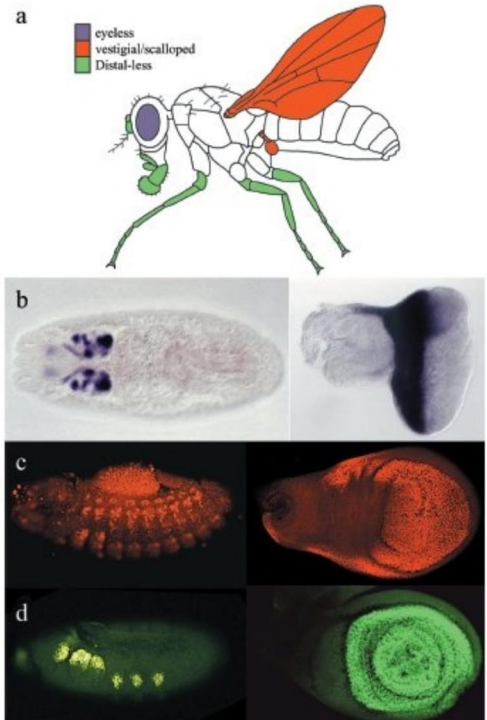
The homeobox gene family is large and diverse. In fact, the homeodomain motif is found in approximately 20 other distinct families of homeobox-containing genes, all of which encode DNA-binding proteins.

Figure 2.8 Homeodomains of *Drosophila Hox* genes Each of the eight *Drosophila Hox* genes encodes proteins containing a highly conserved 60 amino acid DNA-binding domain, the homeodomain, composed of three alpha helices. The third helix is most conserved in sequence. Conserved residues are shaded in yellow; divergent residues are shaded in red; those shared among subsets of proteins are shaded in blue or green.

rise to the flight appendages (see Fig. 2.10c). As is the case with the other field-specific selector genes, expression of *vg* with *sd* in developing eyes, legs, antenna, or genitalia can induce the formation of wing tissue. The Vg and Sd proteins form a complex that binds to DNA, indicating that their selector function is mediated by regulation of gene expression.

Figure 2.10 Field-specific selector genes (**a**) Development of parts of the *Drosophila* adult depend upon the function of the *ey* (eyes), *vg* (flight appendages), and *Dll* (limbs) selector genes. (**b–d**) These genes are expressed in both the embryonic primordia (**left**) and larval imaginal discs (**right**), which will give rise to these structures.

Source: Photomicrographs courtesy of Georg Halder and Grace Panganiban.



The formation of the *Drosophila* heart depends on still another selector gene, dubbed *tinman* (*tin*). Mutants lacking *tin* function lack a heart. The *tin* gene is expressed in the developing mesoderm and in all cells that will form the cardiac tissue of the fly. It is a member of a distinct homeobox family, and thus also a DNA-binding protein that acts by controlling gene expression.

Compartment selector genes

Several genes have been identified in *Drosophila* that act within certain developing fields to subdivide them into separate cell populations, or compartments. The *engrailed* (*en*) gene acts in the posterior part of all segments of the embryo; it is expressed continuously such that the posterior portions of all structures that develop from these segments also express *en* (Fig. 2.11a). The function of the *engrailed* gene is best understood in the embryo and in the developing wing, where it acts to determine posterior identity. Mutations in this gene cause posterior cells to develop as anterior cells but with reversed segmental polarity, resulting in mirror-image duplications of anterior tissue. The *engrailed* gene encodes member of a distinct class of homeodomain-containing transcription factors.

A second compartmental selector gene, *apterous* (*ap*), subdivides the developing wing imaginal disc into dorsal and ventral compartments (Fig. 2.11b–e). Complete loss of *apterous* function blocks wing development, whereas loss of *apterous* function within a subset of dorsal cells transforms their identity to ventral fate. The Apterous protein belongs to yet another class of homeodomain-containing transcription factors.

Cell-type-specific selector genes

Another class of selector genes operates within developing fields to control the differentiation of particular cell types. The formation of neuroblasts and other neural precursor cells in *Drosophila* requires the action of members of the Achaete-Scute Complex (AS-C), a gene complex that contains four genes. Loss of AS-C gene function in the embryo prevents formation of the nervous system; loss or reduction of individual AS-C gene functions in particular body regions in the imaginal tissues of the developing adult fly causes loss of particular sensory bristles.

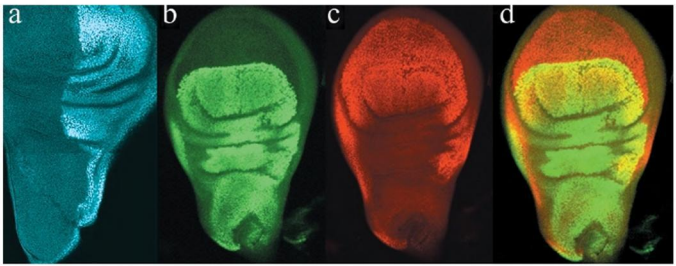
All four AS-C genes encode structurally related transcription factors. The genes are expressed in dynamic and complex patterns that foreshadow the formation of central and peripheral nervous system elements in the larva and adult. The development of neural precursors is initiated within clusters of cells that express AS-C genes, from which a single precursor segregates, divides, and gives rise to neurons and associated cells (Fig. 2.12). A similar process involving a distantly related group of transcription factors specifies muscle development in *Drosophila*. The *twist*, *nautilus,* and *Dmef-2* genes control the development and differentiation of muscle cells.

Formation of the body axes

Systematic searches for developmental genes in Drosophila

Many of the selector genes described in the previous section were first identified on the basis of the adult phenotypes of spontaneous mutants in *Drosophila*. Most of those mutations, however, did not completely disrupt the gene's function during development. Complete loss of function of many selector genes is lethal at earlier stages of development. Therefore, to find genes that control other aspects of embryo organization and patterning, genetic screens had to be designed that could identify recessive lethal mutations.

Figure 2.11 Compartmental selector genes (**a**) The Engrailed protein (shown in blue) is expressed in all cells in the posterior compartment of the wing imaginal disc. (**b**) The Apterous protein (green) is expressed in all cells in the dorsal compartment of the wing imaginal disc, and subdivides the fields of (**c**) vestigial-expressing cells (red) into (**d**) dorsal (yellow; overlap) and ventral (red) populations. (**e**) The territories marked by expression of the proteins in parts b–d in the larval imaginal disc correspond to future regions of the adult wing.



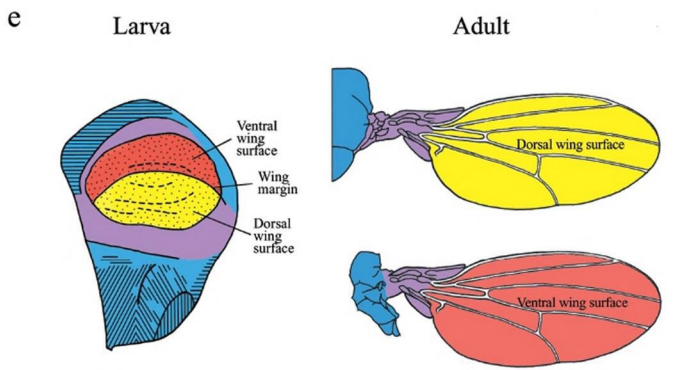


Figure 2.12 **A cell-type-specific selector gene** (**a**) The Achaete protein is expressed in clusters of proneural cells (shown in greater detail in **b**) that foreshadow the pattern of neural precursors. (**c**) Single precursor cells within each cluster will segregate and give rise to neuroblasts.