Advances in Evolutionary Developmental Biology

Edited by J. Todd Streelman

WILEY Blackwell

ADVANCES IN EVOLUTIONARY DEVELOPMENTAL BIOLOGY

Edited by

J. Todd Streelman
School of Biology, and Petit Institute for Bioengineering and Bioscience
Georgia Institute of Technology

WILEY Blackwell

Cover Design: Wiley
Cover Image: Courtesy of the author

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data

Advances in evolutionary developmental biology / edited by: J. Todd Streelman.
pages cm
Includes bibliographical references and index.
ISBN 978-1-118-13111-4 (cloth : alk. paper) 1. Developmental biology.
2. Evolution (Biology) I. Streelman, J. Todd (Jeffrey Todd).
QH491.A24 2013
571.8–dc23

2013038051

Printed in Singapore.

10 9 8 7 6 5 4 3 2 1

ADVANCES IN EVOLUTIONARY DEVELOPMENTAL BIOLOGY

PREFACE

My thoughts to work on this project began some time ago, while discussing biology at Morningside Elementary School in Atlanta. I asked a group of students to show me, using their hands, how *small* they were at their *smallest point* in life. Every child indicated that they were the smallest at birth; most had pictures at home of this day. They had no idea, perhaps couldn't even conceive, that they'd been growing and developing for months prior to being born. So I thought to organize a book around basic, fundamental processes in biology where the truth is almost too remarkable to believe: How do brains develop and evolve? How are boys and girls made differently? How does social information influence development? How do creatures regenerate body parts? How did human language evolve?

The idea for me is that most scientists are just as creative and full of wonder as are second graders, and that the most captivating questions then remain so now. My particular prism, through which to view these questions, is the discipline of evolutionary developmental biology because evo-devo links proximal data to ultimate explanation, via mechanism.

This book is organized into chapters that pass the second-grade litmus test—each addresses fundamental biological mysteries that open the eyes of the child in us all. I thank the authors who have written and illustrated with such care and beauty; my students and postdocs who provided encouragement and comment on this endeavor throughout; and of course those friends and colleagues who reviewed and improved chapter content.

J. Todd Streelman

CONTRIBUTORS

R. Craig Albertson, Department of Biology, University of Massachusetts, Amherst, MA

Gro V. Amdam, School of Life Sciences, Arizona State University, Tempe, AZ, and Department of Chemistry, Biotechnology and Food Science, University of Life Sciences, Aas, Norway

Franck Bourrat, Institut Alfred Fessard, CNRS, Gif sur Yvette, France

Cheng-Ming Chuong, Department of Pathology, Keck School of Medicine, University of Southern California, Los Angeles, CA

Adam G. Dolezal, School of Life Sciences, Arizona State University, Tempe, AZ

Kevin B. Flores, School of Life Sciences, Arizona State University, Tempe, AZ

Hélène Hinaux, Institut Alfred Fessard, CNRS, Gif sur Yvette, France

Jean-Stéphane Joly, Institut Alfred Fessard, CNRS, Gif sur Yvette, France

Genevieve Konopka, Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, TX

Jessica A. Lehoczky, Department of Genetics, Harvard Medical School, Boston, MA

Ang Li, Department of Pathology, Keck School of Medicine, University of Southern California, Los Angeles, CA

Hang Lu, School of Chemical and Biomolecular Engineering, and Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA

Judith E. Mank, Department of Genetics, Evolution and Environment, University College London, London, UK

Daniel J. Miller, Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, TX

Kevin J. Parsons, Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK

Sylvie Rétaux, Institut Alfred Fessard, CNRS, Gif sur Yvette, France

J. Todd Streelman, School of Biology, and Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA

Clifford J. Tabin, Department of Genetics, Harvard Medical School, Boston, MA

Kirsten S. Traynor, School of Life Sciences, Arizona State University, Tempe, AZ

Tobias Uller, Department of Zoology, University of Oxford, Oxford, UK

Randall Widelitz, Department of Pathology, Keck School of Medicine, University of Southern California, Los Angeles, CA

Adam S. Wilkins, Institute of Theoretical Biology, Humboldt University of Berlin, Berlin, Germany

Ping Wu, Department of Pathology, Keck School of Medicine, University of Southern California, Los Angeles, CA

Jun Yin, College of Life Science, Inner Mongolia Agricultural University, Huhhot, China

Mei Zhan, School of Chemical and Biomolecular Engineering, and Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA

CONTENTS

1

"THE GENETIC TOOL-KIT": THE LIFE-HISTORY OF AN IMPORTANT METAPHOR

Adam S. Wilkins

Institute of Theoretical Biology, Humboldt University of Berlin, Berlin, Germany

INTRODUCTION

The expression "the genetic tool-kit" denotes a central idea in evolutionary developmental biology: that there is a relatively small set of key regulatory genes for conserved developmental patterning functions—for tissues, organs, and body axial patterning—throughout the diverse phyla of bilaterally symmetrical animals, despite the tremendous morphological and developmental differences among those phyla. These genes comprise the basic "tools," denoted by the term, that are essential for the developmental construction of these animals. This chapter will attempt to describe the historical background, genesis, significance—and limitations—of the term. Yet, before launching into the topic itself, it might be helpful to begin with a look at how new terms, in general, first promote and enlarge understanding and then, with time, often come to constrain it.

We understand things only in relationship to other things, whether objects or ideas. Though preverbal infants undoubtedly have some mental capacity for making comparisons, language is essential for making clear such conceptual linkages. The two essential linguistic forms for doing so are similes ("X is like Y") and metaphors ("X *is* a Y"). Both are indispensable instruments (itself a metaphor) for making sense of the world but, in general, metaphors are more powerful, hence more effective than similes. Every simile raises an implicit question about the degree of resemblance ("How much is X *really* like Y?"), hence raising a doubt about its aptness while, in

Advances in Evolutionary Developmental Biology, First Edition. Edited by J. Todd Streelman.

contrast, metaphors stress the essential identity between the term of the metaphor and the object/process to which it refers. A good metaphor captures something fundamental and thereby sharpens understanding. This is as true of metaphors used in science as those employed in ordinary speech. Within evolutionary biology, in particular, metaphors have been a particularly important aid in understanding. Indeed, the field was built upon a metaphor, one coined by Charles Darwin, that of "natural selection." Though it was based on a simile to artificial (human-directed) selection as practiced by the plant and animal breeders, it is a metaphor. Evolutionary biology is, in fact, rife with metaphors. Some examples are "evolutionary *tinkering*," "the *selfish* gene," "the adaptive *landscape*," "phylogenetic *trees*," "genetic *drift*," "inbreeding *depression*," "hybrid *vigor*," "life *history* studies," "developmental *plasticity*," "*canalization*," "evolutionary *entropy*."

Within evolutionary developmental biology specifically, there has been no metaphor more important than the genetic tool-kit. It captures the essence of the phenomenon that launched the contemporary field. It was based on the discovery that there is a rather limited set of regulator genes that play conserved functional roles in organizing the "body plans" (another metaphor) of the bilaterian animal phyla (the bilaterally symmetric animals which comprise the great majority of animal species). Though coined relatively late, with respect to the findings it sums up, the term crystallized understanding of the phenomenon it denotes.

In the first part of this article, the history and importance of this concept, and the origins of its designating term, will be set out. I will then discuss why a continued focus on the concept and use of the term may retard the further maturation of evolutionary developmental biology. Such a history of rise-and-fall of the usefulness of an expression is not atypical for metaphors in science. At first, they can be invaluable aids to understanding but, with repeated use, they inevitably lose their freshness and as new findings appear, they begin to seem less apt. The long-standing characterization of clichés, that "a cliché is a dead metaphor," sums up this trajectory.

HISTORICAL BACKGROUND TO THE TERM

Though the expression "the genetic tool-kit" did not achieve currency until the late 1990s, it had a century-old antecedent in a prior idea. This is the concept that, in animal development and evolution, "some genes are more important than others." That was, in essence, if not in those words (the word "gene" had not yet been coined), the key implication of a classic work, published in 1894, titled *Materials for the Study of Variation, Treated with Especial Regard to the Discontinuity in the Origin of Species*. The author was William Bateson (Figure 1.1), who would later give the science of genetics its name and who was one of its foremost practitioners in the first decades of the 20th century. The thesis of Bateson's book was that it is hereditary variations of major observable phenotypic effect that are the source material for evolutionary change (Bateson 1894). It was thus a direct challenge to Darwin's belief that all evolution proceeds via the accumulation of variations of small effect, as the subtitle itself rather aggressively suggests. As pointed out by Gould (1992), Darwin's view had been, in a sense, a democratic one: it implicitly assigned equal potential importance to *all* genes affecting morphology in the evolution of traits and to all mutations that create changes of small phenotypic effect in those genes. Bateson intended to refute that idea by placing known mutations of major phenotypic effect at the center of evolutionary thought.

Figure 1.1. William Bateson (1861–1925), who coined the term "genetics" and was a pioneer of the new science of Mendelian genetics in the early 20th century. He promoted the concept of there being special genes for particular morphological features, the precursor to the idea of the "genetic tool-kit."

Central to the evidence he marshaled was the existence of what we now term "homeotic mutants," those mutants that cause the development of one body part in a new location, replacing another body part, such as the mutants found in a number of insects that replace part or all of an antenna with a leg. Bateson named the phenomenon "homeosis," but the expression "homeotic mutant" would come much later. To both Darwin and his 20th-century intellectual descendants who created NeoDarwinian evolutionary theory (reviewed in Mayr and Provine 1980; Wilkins 2008), the view championed by Bateson was anathema. They regarded such mutations as "sports," that is freaks, that would inevitably quickly die out in a state of Nature, leaving no descendants, and which, therefore, could not have any evolutionary impact.

Although thus downgraded by most 20th-century evolutionary biologists, homeotic mutations continued to exert a fascination for a small group of geneticists interested in both developmental mechanisms and evolution. Among these, none was more important than Edward ("Ed") B. Lewis (Figure 1.2), who worked on a group of homeotic mutations in *Drosophila* that transformed different segments of the thorax and abdomen of the fly into phenotypic facsimiles of other segments. A feature that particularly intrigued Lewis was that these mutations were all closely linked within one small region of the fruit fly's third chromosome. Naming this cluster of genes for the prototypic, original mutation—which partially transformed the metathorax into a second mesothorax, creating a double-mesothorax fly—Lewis called it "the bithorax complex" (Lewis 1964, 1978). Though the importance of Lewis' work was, by the late 1970s, universally recognized for its significance within developmental genetics, it is of historical interest that he began this work in the 1940s to solve an evolutionary (not a developmental) question: do new genes arise by duplication of old ones with one of the duplicates taking on a new function? (see Lewis 1996). This idea had been suggested by Calvin Bridges, based on his finding of duplicate bands in *Drosophila* salivary gland chromosomes (Bridges 1935). It had, however, been previously suggested, as a purely theoretical speculation, by J. B. S Haldane (1932) in his book, *The Causes of Evolution*, while its full significance would only become apparent in the 1970s, with Susumo Ohno's book, *Evolution by Gene Duplication* (Ohno, 1970).

Figure 1.2. Edward B. Lewis (1918–2004), the discoverer of the bithorax complex in *Drosophila* and a pioneer in the exploration of homeotic genes.

Fittingly, given their historical importance, it was genes of the bithorax complex that provided the initial evidence in support of the existence of genes of special general importance for animal "body plans" and which provided the first observations that laid the foundations for the concept of the "genetic tool-kit."

FROM "HOMEOTIC GENES" (AND "HOMEOBOXES") TO THE GENERAL IDEA OF KEY REGULATORY GENES WITH CONSERVED DEVELOPMENTAL FUNCTIONS

The experimental breakthrough from Lewis' more abstract and generic ideas about developmental control to the beginnings of modern evolutionary developmental biology was provided by the cloning of the genes of the bithorax complex (Bender et al. 1983). The sequence analysis of these genes, along with those of the so-called Antennapedia complex (Laughan and Scott 1984; McGinniss et al. 1984), led to the realization that while the homeotic genes of both these *Drosophila* gene clusters—which were soon recognized as split parts of an original single gene complex—were no longer directly recognizable duplicates of one another, they shared a 180 bp coding region homologous to DNA binding regions of certain prokaryotic regulator genes, evidence of ancient gene duplication events, as predicted by Lewis. The genes were named Hox genes, and their encoded 60 amino acid polypeptides were termed "homeodomains." Most homeobox sequences between different Hox genes themselves differ, and some of the sequence differences were soon shown to have major functional significance, but the homeobox family of sequences clearly marked a set of genes that had special importance for development in *Drosophila*. Indeed, the homeobox, as revealed in these first discoveries, was hailed in one commentary as a potential "Rosetta Stone" of biological development (Slack 1984), a special sequence that would help decode the complex "language" of DNA instructions underlying development.

Such hopes were both encouraged and dampened by the discovery that homeobox sequences were not peculiar to *Drosophila* but were found in abundance in other animals, by means of so-called "zoo blots" which detected them by means of DNA homology (McGinniss et al. 1984). The fact that they were numerous and probably universally used in the development of complex animals encouraged the view that these genes were of general and widespread significance. Yet, their sheer abundance within genomes and the fact that many were associated with genes whose expression pattern indicated no relationship to the phenomena of segment identity or segment specification, the functions originally associated with the genes of the Hox complex(es), reduced the chances that the homeobox-containing genes would provide some kind of Rosetta Stone for development. What remained was the original realization that genes containing the homeobox were all involved in development and that they specified positive regulators of gene transcription.

THE GENETIC TOOL-KIT: THE SEMINAL FINDINGS THAT LED TO ITS COINAGE AND THE KEY IDEA

It was, however, two reports on Hox genes in 1989, from two separate labs (Duboule and Dolle 1989; Graham et al. 1989), that regalvanized the whole field of homeobox gene research and which provided the cornerstone of the idea of the "genetic tool-kit." These papers reported that not only did mice possess Hox genes homologous to those of the fruit fly but, as in the fruit fly, these existed as clusters of such genes, in fact four such clusters in the genome, compared to *Drosophila*'s single (though split) cluster. Furthermore, there was a clear relationship between the mammalian Hox gene clusters and those of *Drosophila*, in terms of both the order of the homologous genes along the chromosome and the respective regions of the embryos of the two animals in which they were expressed. Thus, for example, the most anteriorly expressed genes of the Antennapedia complex in the fruit fly had matching homologues in the most anteriorly expressed genes of the four Hox gene clusters in the mouse.

These findings could hardly have been more exciting or surprising. They indicated that beneath the dramatic surface differences of development exhibited by an insect and a mammal, the two animals shared some underlying form of genetic specification of the antero-posterior (a-p) body axis. Nor were the similarities of Hox clusters in fruit fly and mouse some kind of fluke or amazing coincidence: within five years or so, Hox gene clusters with the same general features, were found throughout the bilaterian animal phyla. The findings blew apart the conventional wisdom of preceding decades, that different morphologies of structures of comparable function between different animal groups must reflect comparable differences in their genetic specification (Mayr 1963; Salvini-Plawen and Mayr 1977).

These discoveries, from the labs of Krumlauf and Duboule, set off a gold rush in many laboratories to see if other important regulator genes first identified in one organism, usually *Drosophila*, might have similar expression patterns in other, often quite different, animals. Such resemblances, if found, would provide preliminary indications of similar roles in both invertebrate animals and the far more complex vertebrates. The prospectors were not disappointed: gold was soon struck.

Within the space of half a dozen years, a panoply of genes identified in *Drosophila* as essential for the development of various specific organs and tissues in the fruit fly, and

all encoding transcription factors, were discovered to be expressed in the developing rudiments of the comparable organs and tissues of mammals. For instance, to take the most famous example, a gene (*Pax6*) crucial for eye formation in *Drosophila*, which was first identified by Thomas Hunt Morgan and his colleagues in the 1920s, as the gene *eyeless*, is required comparably for mouse and human eye development. Or, to take a second example, the fruit fly gene *tinman*, required for mesoderm and specifically heart development, was found to be represented by a small family of homologous genes in vertebrates which are similarly required for mesoderm and heart development. Furthermore, these genes were not only expressed in comparable organ sites but, as shown by knockout mutations in mice, found to be required for development of the respective murine organs and tissues. Other genetic findings involved replacement of *Drosophila* genes by their mouse homologues, with the discovery of retained developmental function. Still other findings involved overexpression of particular key regulator genes with the resulting production of small versions of the respective organs that they controlled. The most dramatic of these involved the ectopic production of eyes in *Drosophila* following overexpression of the mouse version of the *Pax6* gene (Halder et al. 1995). The ensemble of findings of conserved gene functions for development throughout the bilateria have been reviewed in Carroll et al. (2001) and Wilkins (2002).

It was not, however, just transcription factors and their roles that were found to be conserved. All the major signal transduction families were found to be similarly conserved across the bilateria, and often their roles in particular developmental processes were similarly conserved. These included the Wnt, Fgf, Bmp, and Shh signaling pathways. Some of these uses might involve convergent gene/module recruitment processes, but for the most part, it appears that there has been true conservation of these usages, or "deep homology" to use the phrase of Shubin et al. (1997).

The evolutionary implications were both clear and startling: this shared genetic machinery must go back to the ancestor of all bilaterian animals, dubbed the "Urbilaterian," which must have lived before the start of the Cambrian period (based on the discovery of a few such in the late Proterozoic), more than 540 million years ago. Evidently, their essential roles had been retained even as the organs and tissues that they helped specify evolved, often dramatically, into very different forms. Thus, beneath the tremendous visible phenotypic diversity of animals lay a previously wholly unexpected set of shared genetic "instructions."

It is this idea that is encapsulated in the term "the genetic tool-kit," referring specifically to the genes required for the development of bilaterian animal forms. (There had been several earlier usages of this term but in different contexts and with different referents.) Interestingly, however, it is not clear who coined the term. I have asked several major figures in the field and no one seems to know who used the term first. One of the earliest published uses of the term, though simply stated, as if self-explanatory, is in Knoll and Carroll (1999). In that same year, Michael Akam was using the term in his lectures at the University of Cambridge (M. Akam, pers. comm.). Furthermore, in subsequent published papers, the term is almost always used without explanation or careful definition, as if it needs no such clarification. It seems probable that the publication that gave the term its real currency was the book by Sean Carroll (Figure 1.3) and two colleagues, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (2001).

It is also interesting that no one has tried to put a number to the components of the genetic tool-kit. But if one adds up all the conserved transcription factor and signal transduction genes, one is probably looking at a set of a few hundred genes, perhaps no more

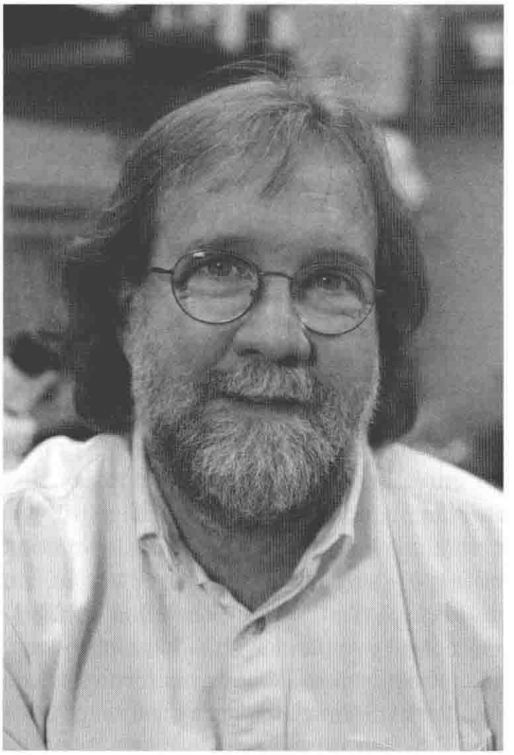

Figure 1.3. Sean A. Carroll (1960 to present), a key investigator in and writer on evolutionary developmental biology. He is the probable originator of the term, "the genetic tool-kit." Photograph courtesy of Jamie Carroll.

than 200–300, not counting paralogues, and perhaps considerably less, yet whatever the precise number, this set is clearly only a small fraction of the total number of genes in animal genomes, which range, roughly, from 12,000 to 25,000. (That range of gene numbers seems small in comparison to earlier estimates (Fields et al. 1994), but the estimated sum of "tool-kit" genes remains a small part of the total.)

In effect, it appears that the idea was so well established by the late 1990s that the function of the term seems to have been to crystallize the concept rather than initiate a new way of thinking. In this, it differs from other comparable metaphors, whose origination helped give clear definition to the respective ideas designated. Phrases such as "the adaptive landscape," "the selfish gene," "the genetic program" are all in the latter category. Yet, the genetic tool-kit (for animal development) provided a wonderful shorthand description for what was, and remains, the most surprising phenomenon in the field of evolutionary developmental biology and, indeed, one of the most unexpected in 20th-century evolutionary biology as a whole.

In using the term, it should be noted that elements of the tool-kit need not be single genes. They can be small groups of conserved interacting genes. Such co-operating tool-kit elements were termed, tongue-in-cheek, a "junta" by Claude Desplan (Desplan 1997), in deliberate contrast to the earlier concept of (single) "master genes" (Halder et al. 1995), the latter presumably individually governing a particular developmental process. The need to move away from the master gene concept became evident with discoveries that it is not *Pax6* alone that governs eye development but a group of interacting genes of which *Pax6* is but one (reviewed in Pappu and Mardon, 2004 and Treisman 1999). Formulations preferable to junta, for such conserved interacting cohorts of genes, are "module" or "kernel"