
THE REGULATORY GENOME

THE REGULATORY GENOME

**Gene Regulatory Networks
in Development and Evolution**

Eric H. Davidson

Division of Biology
California Institute of Technology
Pasadena, California



ELSEVIER
ACADEMIC
PRESS

Amsterdam Boston Heidelberg London New York Oxford
Paris San Diego San Francisco Sydney Tokyo

Academic Press is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WC1X 8RR, UK

This book is printed on acid-free paper. ∞

Copyright © 2006, by Eric H. Davidson

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone: (+44) 1865 843830, fax: (+44) 1865 853333, E-mail: permissions@elsevier.com. You may also complete your request on-line via the Elsevier homepage (<http://elsevier.com>), by selecting "Support & Contact" then "Copyright and Permission" and then "Obtaining Permissions."

Library of Congress Cataloging-in-Publication Data

Application submitted

British Library Cataloguing-in-Publication Data

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

ISBN 13: 978-0-12-088563-3

ISBN 10: 0-12-088563-8

For information on all Academic Press publications
visit our Web site at www.books.elsevier.com

Printed in Canada

06 07 08 09 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER

BOOK AID
International

Sabre Foundation

CONTENTS

<i>Preface</i>	<i>ix</i>
Chapter 1	
The “Regulatory Genome” for Animal Development	1
The Framework	1
The Regulatory Apparatus Encoded in the DNA	3
Genomes, Genes, and Genomic “Space”	3
Overview of Regulatory Architecture	8
Gene Regulatory Networks	11
The Regulatory Demands of Development	16
Readout and Generation of Regulatory Information in Developmental Specification	16
Spatial Gene Expression	18
From Regional Specification to Terminal Differentiation	21
Evolution, Development, and the Regulatory Genome	27
Chapter 2	
<i>cis</i>-Regulatory Modules, and the Structure/Function Basis of Regulatory Logic	31
General Operating Principles	31
Modularity, A General Property of Genomic <i>cis</i>-Regulatory Control Units	33
Inside the <i>cis</i>-Regulatory Module: Logic Processing and Input/Output Relations	47
<i>cis</i> -Regulatory Logic Processors: Examples	48
Target Site Occupancy and Transcriptional Output	57
Generalization: The Combinatorial <i>cis</i> -Regulatory Logic Code	59
<i>cis</i>-Regulatory Design	61
Integration of Spatial Inputs at the <i>cis</i> -Regulatory Level	61
Repression and the Diversity of <i>cis</i> -Regulatory Design	68
<i>cis</i> -Regulatory Design and the Creation of Spatial Complexity in Development	73

Control by Sequence-Specific Interactions Among Distant <i>cis</i>-Regulatory Elements	78
Development and BTA Specificity	79
Control Functions of Sequence-Specific, Intra- and Intergenic Genomic Looping	80
Chapter 3	
Development as a Process of Regulatory State Specification	87
Developmental Processes of the Pregastrular Embryo	87
Vectorial Embryogenesis	88
Maternal Anisotropy and External Cues for Axial Polarity	90
Spatial Regulatory Complexity and Interblastomere Signaling Territories	94
	100
Comparative View of Embryonic Processes, and the Organization of the Underlying Gene Regulatory Networks	102
Shallow Gene Regulatory Networks of Type 1 Embryos	103
More Complex Forms of Embryonic Process	109
Regulatory State and Network Circuitry in Building Adult Body Parts	114
Steps in the Postembryonic Developmental Process	114
The Progenitor Field Concept	115
Subdivision of the Progenitor Field Regulatory State	121
Body Plan Development and the Organism	122
Chapter 4	
Gene Regulatory Networks for Development: What They Are, How They Work, and What They Mean	125
General Structural Properties of Developmental Gene Regulatory Networks	126
Overall Network Structure	127
Gene Regulatory Networks from Different Perspectives:	
Subcircuits to Process Diagrams	135
Qualitative Diversity of Network Components	136
Extensions	140
Gene Regulatory Networks for Embryonic Development	144
The Sea Urchin Endomesoderm Network	144
Regulatory Gene Network for Specification of Mesoderm in the <i>Xenopus</i> Embryo	159
Gene Regulatory Network Controlling Dorsal-Ventral Territorial Specification in the <i>Drosophila</i> Embryo	166

Networks that Control Construction of Components of Adult Body Parts	175
Network Organization in the Specification of Pancreatic β -Cells	176
Cell Fate Exclusion by microRNAs in a Gene Regulatory Controlling Terminal Differentiation in <i>C. elegans</i> Taste Network Neurons	180
Concluding Remarks	184
Chapter 5	
Gene Regulatory Networks: The Roots of Causality and Diversity in Animal Evolution	187
Evolutionary Implications of the Structure and Function of Gene Regulatory Networks	188
The Parts of Gene Regulatory Networks, and the Qualities of Evolutionary Change	188
Linnaean Framework for the Evolutionary Consequences of Alterations in Gene Regulatory Networks	195
Bilaterian Kernels, Predicted and Real	196
Caveats and a Caution	197
Regional Specification in the Bilaterian Nervous System	200
The Bilaterian Gut	212
A Possible Pan-Bilaterian Kernel Underlying Specification of the Heart Progenitor Field	218
Distribution of Kernels	221
Subphyletic Evolution of Bilaterian Body Plans: <i>box</i> Genes as Sources of Labile Network Input/Output Switches	223
Diverse Functional Consequences of <i>box</i> Gene Expression	223
Multiple Examples: Regional Repression and Installation of Patterning Output, and the Evolutionary Diversification of Body Parts	224
<i>box</i> Gene Subnetworks and Spatial Evolution	233
Metazoan Origins and Gene Networks Before Kernels	235
“Deconstructing” Bilaterian Gene Regulatory Networks in Time	235
Concluding Reflection: The Principle of Animal Development and Evolution	239
References	241
Figure Citations	275
Index	279

PREFACE

This book came about in the following way. In the winter of 2004, Jasna Markovac, the publisher of my monograph “Genomic Regulatory Systems” (2001), asked me to think about doing a second edition of that work. But no sooner had I opened that door and begun to survey the landscape than, as in an old legend of magic, everything seemed to have been transformed to another landscape. The conceptual peaks that were the landmarks of “Genomic Regulatory Systems” were still there, but now new mountains towered over them, and the whole domain appeared to be much more brightly illuminated, and to extend farther toward the horizon. It was beyond possibility to “update” a vista so changed.

“Genomic Regulatory Systems” sought as its main objective to crystallize the incontrovertible evidence that causality in development resides ultimately in *cis*-regulatory control of spatial gene expression. But development is the output of regulatory systems comprising large numbers of regulatory genes. Though in that work I often referred to the gene networks that would someday represent developmental programs, the few examples were anecdotal, and their general properties remained entirely obscure. Then in 2002 we published the first real scale gene regulatory network (including about 50 genes) explanatory of a major piece of development, specification of the endomesoderm of the sea urchin embryo. As the ancients used to say, the scales fell from our eyes. A whole field of developmental gene regulatory networks has now sprung forth, that encompasses many different animal systems. The structure/function properties that emerge from the architecture of these networks are a large part of what has transformed the conceptual terrain of this large area of bioscience. There follow more new things: a different way to think comparatively about various forms of development; a different way to think about the process of evolution; hence the subtitle of this book, “Gene Networks in Development and Evolution”. Exploration of these new pathways toward scientific explanation of the developmental and evolutionary phenomena of biology is the central object of this volume. As with its predecessor, the approach I have taken here is that of demonstration by example: the points to be made rest upon powerful, exemplary experimental demonstrations, detailed, for those who desire experimental substance, in the figure captions. However, in no way have I attempted to be encyclopedic. So, willy nilly, there is much that could equally well have been included but was not, and my apologies en masse to the authors of these works.

I have not shied away from what are sometimes pejoratively been termed “big ideas,” nor taken the view of an anonymous reviewer of a paper of mine who recently amused me with the complaint “But the original ideas in this paper are speculative!” This book includes many diagrams in which concepts are set forth

in specific form, just so they can be subjected to precise tests of falsification, and just so they can be used in precise ways to generate predictions I may not have thought of. One such idea, which underlies everything in this book, is the concept of genomically encoded information processing. To return to my metaphor above, this is like the geological basis of the landscape. In my view, *cis*-regulatory information processing, and information processing at the gene regulatory network circuit level, are the real secret of animal development. Probably the appearance of genomic regulatory systems capable of information processing is what made animal evolution possible.

This book begins with an overview of the regulatory genome and the concept of information processing in gene regulation (Chapter 1). It proceeds to an in-depth analysis of modular *cis*-regulatory designs for generation of spatial patterns of gene expression, and consideration of how they generate regulatory output (Chapter 2); thence to a comparative treatment of developmental pathways in terms of transient regulatory states (Chapter 3); to gene regulatory network theory and the character of diverse real developmental regulatory networks (Chapter 4); and finally to the application of network structure/function relations to some unsolved problems of animal evolution (Chapter 5). The image of a genomically encoded information processing system that throughout the life cycle responds conditionally to incident regulatory inputs can never lie far from the surface of any of these subject areas.

Science is made by scientists, whose creations deeply affect each others' progress. For me there have been certain scientists in each period of my own progress whose work and ideas have particularly illuminated the world: among them I must mention as of particular importance in this present period, and for what is included herein, Mike Levine, Ellen Rothenberg, Doug Erwin, Sorin Istrail, Bill McGinnis, and Lee Hood. This book would not have whatever worth it does were it not for the generosity of these people and also of Paola Oliveri and Joel Smith, postdoctoral colleagues in my laboratory, in reading, criticizing, and improving drafts of various parts, and in some cases all, of the manuscript. I have been extremely fortunate to have had the very expert services of a superb illustrator, Tania Dugatkin. In my own domain Deanna Thomas has provided invaluable assistance with figures, references, and everything else; and my graduate student Pei-yun Lee not only helped with technical research but also with figure attributions. Nor would this project have ever reached fruition were it not for the continued encouragement of Jasna Markovac, and of the careful, obsessive work of the production manager Paul Gottehrer at Academic Press/Elsevier. I also wish to say that since so much of what follows is linked to our expanding experimental invasion of gene regulatory networks, the support we have had for that research has been indirectly essential for this book as well: mainly this support has come from the National Institute of Child Health and Human Development and from the Genomes to Life Program of DOE, but also from NIGMS, NIRR, NIHGRI, NSF, NASA, Caltech's Beckman Institute, and Applied Biosystems, Inc.

Finally, I would like to dedicate this book to the person who has worked most closely with me on it, good days and bad, and that is Jane Rigg. She has been my editor, judge, administrator, research aide, and advisor throughout, as also on three other books I have written in the more than 35 years that we have worked together. Only my first book, "Gene Activity in Early Development" (1968) preceded the Jane Rigg era, but that was a very long time ago indeed.

Eric Davidson
April 2006