Insect Molecular Genetics

An Introduction to Principles and Applications

Third Edition

Marjorie A. Hoy

Insect Molecular Genetics

An Introduction to Principles and Applications

Third Edition

Marjorie A. Hoy

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
32 Jamestown Road, London NW1 7BY, UK
225 Wyman Street, Waltham, MA 02451, USA
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

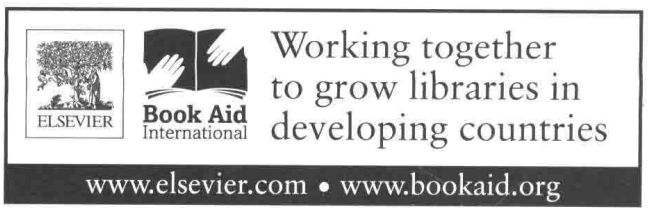
Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-415874-0

Typeset by MPS Limited, Chennai, India
www.adi-mps.com

Printed and bound in United States of America

13 14 15 16 10 9 8 7 6 5 4 3



Insect Molecular Genetics

An Introduction to Principles and Applications

Preface to the Third Edition

Once again, I must acknowledge the very rapid pace of advances in arthropod molecular genetics. Because ticks and mites are increasingly included in molecular genetic studies, especially now that the genomes of a tick, the twospotted spider mite, *Varroa* mite, and a predatory mite have been sequenced, I use the term arthropod often in this edition. Since the year 2000, the number of arthropod genomes that have been sequenced has increased dramatically, primarily due to the decreasing costs of sequencing using Next-Generation methods. The cover of edition three contains photos of many of the arthropods for which complete genome-sequence data are available. A proposal was made in 2011 to sequence 5000 insect genomes in the next few years and, as a result, the cover of this book would soon be unable to "host" all the photographs of arthropods with completely sequenced genomes. The enormous amount of genomic data being developed will affect all aspects of the study of arthropods, including physiology, phylogenetics, ecology, evolution, behavior, and applied pest-management methods. The release of transgenic insects into the environment is receiving increased focus, now that releases into the environment of transgenic pink bollworms, predatory mites, and mosquitoes have occurred.

A very diverse group of scientists use molecular genetic tools to solve problems that can be answered in no other manner, so it is critical that all entomologists (and acarologists) become familiar with the terminology, concepts, and tools of molecular genetics–even if they do not wish to do such research themselves. Once again, I am writing this book for students and advanced scientists who have a limited background in insect molecular genetics. I hope to provide a portal into the concepts, methods, and applications of these powerful tools. Because the field has expanded so much, I cannot cite all relevant papers and have tended to include reviews and partial lists to give the reader a way to access the literature. I apologize to those scientists who have published excellent papers on diverse topics of relevance, but space considerations limited the examples that can be cited.

After discussing what material should be covered in this edition with several colleagues, I have once again provided simple protocols in the methods chapters (5–9) in an effort to explain what is being done when DNA or RNA are being extracted, genes are being cloned and sequenced, genomic or transcriptomic libraries are developed, and the polymerase chain reaction is being conducted. Although numerous commercial kits are available and widely used, the recommendations were that it is instructive for novices to gain an understanding as to what is being achieved (and how) when carrying out these procedures. The precise details of kit methodologies are often considered proprietary information, but the protocols provided here may make it possible for students to understand the concepts sufficiently that they can learn to "trouble shoot" if problems occur in their own projects.

All chapters have been revised and updated. Three areas have changed dramatically since the second edition: 1) The development of genetically modified insects for potential use in pest-management programs has advanced to the point that several releases have been made. 2) The development of low-cost Next-Generation sequencing methods permit the sequencing of arthropod genomes to be commonplace, although the explosion in genome (and transcriptome) sequencing data threatens to overwhelm our ability to analyze and use the data. Perhaps the next major advance will be the availability of new tools for bioinformatics analyses so that even novices can effectively analyze genomic data. It now costs US$10,000–$30,000 to sequence a genome, but the cost of analyzing the data may be ten times that. Even newer sequencing methods just developed will reduce the time and cost to sequence genomes. 3) Finally, the use of molecular methods has allowed immense increases in our knowledge of the diversity and roles that microbial symbionts play in the biology of their arthropod hosts. This expansion in knowledge is reflected in the publication of several books devoted to insect symbiosis. It has been difficult to choose among the many interesting examples of symbiosis, but I have included examples that illustrate the impact that symbionts have on the physiology, behavior, sex determination, reproductive compatibility, and pest status of their hosts.

I thank the students and others who have provided constructive feedback on the earlier editions of this book and Ke Wu, Niklaus Hostettler, Jennifer Gillett-Kaufman, and Aaron Pomerantz, who assisted me in the preparation of this edition. I especially thank Geoffrey Attardo, Susan Brown, Lyle Buss, Chun-che Chang, Michael Clark, Cameron Currie, Marian Goldsmith, Thom Kaufman, Jason Meyer, James Newman, Jeffrey Stuart, and Yoshi Tomoyasu for their excellent photographs of the arthropod species illustrated on the cover of this book for which genome-sequence data are available.

My very best wishes for a productive and interesting read. I believe that molecular-genetic tools will help to resolve many critical issues in arthropod biology, evolution, and pest management.

Marjorie A. Hoy
Eminent Scholar and Davies,
Fischer and Eckes Professor of Biological Control,
Department of Entomology and Nematology,
University of Florida, Gainesville
September 2012

Preface to the Second Edition

Amazing progress has been made in insect molecular genetics since the first edition appeared in 1994. Transformation of insects other than *Drosophila melanogaster* has become an almost routine project. The *Drosophila* Genome Project was completed in 1999 and produced many surprises and promises a fruitful future for mining genes and developing an understanding of genome structure, function and evolution. The mining of this treasure trove of data will require some years of work, but the possibility exists that we ultimately will be able to understand how this insect develops. Insect biology will become synthetic again with the use of genomics, transcriptomics, and proteomics approaches. The complete sequencing of other complex eukaryotic genomes, including that of *Caenorhabditis elegans* and *Homo sapiens*, opened additional doors to compare genome organization, evolution, and gene function.

Molecular methods and technology have changed rapidly in the past few years, with a plethora of new kits available for extracting and purifying DNA and RNA, for cloning, sequencing, and amplifying DNA and RNA by the polymerase chain reaction (PCR). Gene chip or microarray methods offer new tools for learning about gene function. All the improvements in these molecular toolkits make molecular methods ever more accessible to the entomological community.

The same basic organization with three major sections has been maintained in this edition, but the chapters have been reorganized in Section I, and all have been updated with recent references. References were included that provide an entry into the recent literature; where possible, review articles are cited. I regret I could not include references to all the new molecular studies on insects, but there are just too many! That signals that molecular entomology is maturing.

This book is dedicated to entomologists just beginning their research careers; I hope this book helps you to start exciting and productive projects that employ these valuable molecular tools. For those of you with no background in molecular genetics, the book should be read from start to finish. Key concepts are

highlighted in the Overview and reading it both before and after reading each chapter may be helpful. The diagrams, especially those illustrating molecular methods, should be evaluated while reading the text. In many cases, the concepts involved are most readily obtained if the text and diagrams are read together.

Finally, I thank all the people who so kindly provided feedback on the first edition and those who made thoughtful suggestions on earlier drafts of this one, including A. Jeyaprakash, Lucy Skelley, Juan Alvarez, and Alison Walker. I sincerely thank those who kindly provided new illustrations, Bruce Webb, Carol Lauzon, and Felix Guererro. Thanks to Pam Howell, Mike Sanford, and Pat Hope for their assistance in getting this manuscript and its illustrations completed.

Marjorie A. Hoy
Department of Entomology and Nematology
University of Florida, Gainesville

Preface to the First Edition

The development of recombinant DNA techniques during the past 20 years has resulted in exciting advances in the detailed study of specific genes at the molecular level as well as breakthroughs in molecular, cellular, and developmental biology. Of the molecular genetics studies conducted on insects, most have been directed to *Drosophila melanogaster*. Relatively few data have been generated by molecular biological methods from analyses of other insects. Yet, the application of molecular genetics to insects other than *Drosophila* has the potential to revolutionize insect population and organismal biology.

Why have molecular genetic techniques been used so little by entomologists? There may be a number of reasons. Recombinant DNA techniques are most readily carried out by people trained in biochemistry and relatively few entomologists are so trained. The techniques have been, until recently, relatively complex and difficult, so that strong technical skills were required. Also, most entomologists have been slow to ask whether these techniques were appropriate for studies of population or organismal biology because much of the published literature has focused on fundamental issues of *Drosophila* gene structure, regulation and function, developmental regulation, and evolution.

Goals

My goal is to introduce entomologists to the concepts of molecular genetics without assuming that they have received previous training in molecular biology. This book is not intended to substitute for formal training in biochemistry or molecular genetics. If novice readers wish to develop molecular genetics skills, they must obtain additional training in genetics and biochemistry. However, the book will provide an introduction to terminology, as well as an overview of principles, techniques, and possible applications of molecular genetics to problems of interest to entomologists.

In preference to using examples from the *Drosophila* literature, I have used examples in which other arthropods have been studied. However, without doubt, *Drosophila* is the premier model for insect molecular genetics study. One fond hope is that this book will be a bridge for entomologists seeking to apply the exciting methods developed for *Drosophila* and that it will introduce *Drosophila* workers to some of the problems and issues of interest to entomologists seeking to solve applied problems. Perhaps this book will help to break down the barriers between entomologists and *Drosophila* workers isolated from each other by perspective and technical jargon. If this book helps to achieve these goals, it will have served its purpose.

Organization

The book was designed for a one-semester course in insect molecular genetics for upper division undergraduates or beginning graduate students. The initial portion of the book reviews basic information about DNA, RNA, and other important molecules (Chapters 1–4). Readers with a recent course in genetics could skip this section. Chapter 5 describes the genetic systems found in insects and an overview of development sufficient to understand subsequent techniques such as *P*-element–mediated transformation and sex determination. Chapters 6–9 provide introductions to useful techniques, including cloning, library construction, sequencing, the polymerase chain reaction, and P element-mediated transformation of *Drosophila*. Most molecular biologists reading this book could skip this section as well. Chapters 6–9 are not intended as a laboratory manual, but an outline of laboratory protocols is provided in order to furnish the novice with a sense of the complexity or simplicity of the procedures and some of the issues to consider in problem solving. Throughout the book, references are provided for the reader interested in pursuing specific topics and techniques, although they are not exhaustive. Despite the value of providing an historical overview, I have not always provided references to the first publication on a subject. Rather, review articles or recent publications that include references to earlier work are cited.

Finally, in the third section of the book (Chapters 10–16), I have attempted to demonstrate how molecular genetic techniques can solve a diverse array of basic and applied problems. Part III is intended to introduce readers to the exciting molecular research that is revolutionizing insect biology, ecology, systematics, behavior, physiology, development, sex determination, and pest management. Each chapter in this section could be read by itself, assuming that reader understands the appropriate concepts or information presented in Parts I and II.

Each chapter begins with an overview or brief summary of the material being covered. The overview should be read both before and after reading each chapter to review the concepts covered. The overview is followed by a brief introduction covering the history or rationale for the topic. References at the end of the chapter are provided for further reading. Where possible, books or reviews are cited to provide an entry into the literature. Recent references are provided, but no attempt has been made to review all the literature on a specific topic. Simple protocols may be given to provide the flavor of specific techniques, although these are not intended to be complete. References to handbooks or techniques books are also provided at the end of appropriate chapters. When a term that may be unfamiliar is first introduced, it is underlined and a brief definition or description is given in the Glossary at the end of the book. Finally, in Appendix I, a time line of some significant advances in genetics, molecular biology, and insect molecular genetics provides a perspective of the pace with which dramatic advances have been, and continue to be, made.

Progress is rapid in molecular genetics, and this book can only provide an introduction to the principles of insect molecular genetics and some of its applications. It is impossible to provide a complete review of the insect molecular genetics literature in a book of this size. The literature cited includes references to through 1993 and focuses on genetics. It is not intended to be an introduction to all aspects of "molecular entomology," which has been defined as "... a blend of insect science, molecular biology, and biochemistry." The dividing line between molecular entomology and insect molecular genetics is sometimes difficult to resolve.

Shortly before this book went to the publisher, two related books were published: *Molecular Approaches to Fundamental and Applied Entomology*, edited by J. Oakeshott and M. J. Whitten, and *Insect Molecular Science*, edited by J. M. Crampton and P. Eggleston. Both multiauthored books cover some of the topics included here, but assume the reader is familiar with molecular genetic techniques and terminology; they would be daunting for the novice.

Acknowledgments

This book would not have been written without the encouragement of many people. As most authors remark, it was more difficult to write than expected, in part due to a move from the University of California at Berkeley to the University of Florida in Gainesville about half-way through the endeavor. Certainly, this book would not have been written without the support of the Rockefeller Foundation who provided me with 5 weeks of precious time at the Bellagio Study and Conference Center in Italy. I gained a crucial grasp on the scope of the project and gathered my courage there while on sabbatical from the University of California, Berkeley. A fellowship spent at Cold Spring Harbor Laboratories in the summer of 1985 in the "Molecular Cloning of Eukaryotic Genes" course provided my initial training in molecular genetic techniques. I thank the instructors for their patience. Likewise, G. M. Rubin kindly allowed me to participate in his cloning course at the University of California, Berkeley.

Many people have contributed information, advice, and valuable time reviewing this book. I especially thank Mary Bownes, Michael Caprio, Gary Carvalho, Howell Daly, Owain Edwards, Marilyn Houck, James Hoy, A. Jeyaprakash, Srini Kambhampati, Carolyn Kane, James Presnail, Veronica Rodriguez, Mark Tanouye, and Tom Walker, who either reviewed drafts of individual chapters or the entire book. Richard Beeman, Owain Edwards, Glenn Hall, A. Jeyaprakash, Ed Lewis, Jim Presnail, and A. Zacharopoulou provided photographs or illustrations. Lois Caprio and Denise Johanowicz assisted with many of the logistical issues. Finally, I especially thank my husband Jim, who repeatedly motivated me to clarify and elucidate the principles and applications presented here and tolerated my preoccupation with this project over many months, and John Capinera, who patiently waited for me to emerge from my compulsion. Despite the best efforts of the reviewers, errors probably persist and are my responsibility.

Marjorie A. Hoy
University of California, Berkeley
University of Florida, Gainesville

Contents

Preface to the Third Edition xvii
Preface to the Second Edition xxi
Preface to the First Edition xxiii
Acknowledgments xxvii

PART I: GENES AND GENOME ORGANIZATION IN EUKARYOTES 1

Chapter 1 DNA, Gene Structure, and DNA Replication 3
1.1 Overview 3
1.2 DNA is the Hereditary Material: A Brief History 4
1.3 The Central Dogma 6
1.4 The "RNA World" Came First? 8
1.5 The Molecular Structure of DNA 8
1.6 The Molecular Structure of RNA 11
1.7 The Double Helix 13
1.8 Complementary Base Pairing is Fundamental 14
1.9 DNA Exists in Several Forms 14
1.10 Genes 16
1.11 The Genetic Code for Protein-Coding Genes is a Triplet and is Degenerate 16
1.12 Gene Organization 18
1.13 Efficient DNA Replication is Essential 21
1.14 DNA Replication is Semiconservative 21
1.15 Replication Begins at Replication Origins 21
1.16 DNA Replication Occurs Only in the 5′ to 3′ Direction 22
1.17 Replication of DNA Requires an RNA Primer 23
1.18 Ligation of Replicated DNA Fragments 23
1.19 DNA Replication during Mitosis in Eukaryotes 25
1.20 Telomeres at the End: A Solution to the Loss of DNA during Replication 28
1.21 DNA Replication Fidelity and DNA Repair 28
1.22 Mutations in the Genome 29
1.23 Common Genetic Terminology 32
1.24 Independent Assortment and Recombination during Sexual Reproduction 33

General References 33
References Cited 34

Chapter 2 Transcription, Translation, and Regulation of Eukaryotic DNA 37

2.1 Overview 37
2.2 Introduction 38
2.3 RNA Synthesis is Gene Transcription 41
2.4 Transcription Involves Binding, Initiation, Elongation, and Termination 43
2.5 RNA Transcripts of Protein-Coding Genes 45
2.6 RNA of Protein-Coding Genes Must Be Modified and Processed in Eukaryotes 45
2.7 Splicing Out the Introns 47
2.8 Translation Involves Protein Synthesis 48
2.9 RNA Surveillance: Damage Control 52
2.10 Import and Export from the Nucleus 53
2.11 Transport of Proteins within the Cytoplasm 54
2.12 mRNA Stability 55
2.13 Chaperones and the Proteosome 56
2.14 RNA Silencing or Interference and miRNAs 56
2.15 Gene Regulation in Eukaryotes 58
2.16 Insulators and Boundaries 63
2.17 Chromosome or Gene Imprinting in Insects 64
2.18 Eukaryotic Genomes and Evolution 66
References Cited 67

Chapter 3 Nuclear and Extranuclear DNA in Insects 71

3.1 Overview 71
3.2 Introduction 72
3.3 C-Value Paradox: Is it Real? 73
3.4 Repetitive DNA is Common in Insects 74
3.5 Composition of Insect DNA 76
3.6 Chromosomes are DNA Plus Proteins 76
3.7 Packaging Long, Thin DNA Molecules into Tiny Spaces 77
3.8 Structure of the Nucleus 79
3.9 Euchromatin and Heterochromatin 80
3.10 Centromeres 81
3.11 Telomeres 82
3.12 Chromosomes during Mitosis and Meiosis 84
3.12.1 Mitosis 84
3.12.2 Meiosis 87
3.13 Chromosome Damage 90
3.14 Polyteny 92
3.15 Chromosomal Puffing 93
3.16 B Chromosomes 94
3.17 Sex Chromosomes 94

3.18 Extranuclear Inheritance in Mitochondrial Genes 95
3.19 Transposable Elements are Ubiquitous Agents that Alter Genomes 99
References Cited 99

Chapter 4 Genetic Systems, Genome Evolution, and Genetic Control of Embryonic Development in Insects 103

4.1 Overview 104
4.2 Introduction 105
4.3 Genetic Systems in Insects 105
4.4 Endopolyploidy is Common in Somatic Tissues of Arthropods 107
4.5 Genetics of Insects Other than *D. melanogaster* 107
4.6 Dynamic Insect Genomes 108
4.6.1 Horizontal Gene Transfer from Microorganisms to Insects 109
4.7 B Chromosomes 110
4.8 Unique-Sequence DNA in the Nucleus 111
4.9 Middle-Repetitive DNA in the Nucleus 111
4.9.1 Heat-Shock Genes 112
4.9.2 Histone Genes 113
4.9.3 Immune-Response Genes 113
4.9.4 Ribosomal Genes 115
4.9.5 Silk Genes 115
4.9.6 Transfer RNA Genes 116
4.9.7 Vitellogenin Genes 116
4.9.8 Transposable Elements 117
4.10 Highly Repetitive DNA 122
4.11 Producing Large Amounts of Protein in a Short Time: Gene Amplification and Gene Duplication 123
4.11.1 Chorion Genes in *Drosophila* and Moths 124
4.11.2 Insecticide Resistance 127
4.12 Multiple Genomes in or on Insects: What is the "Biological Individual"? 127
4.12.1 Multiple Symbionts 128
4.12.2 *Wolbachia* and *Cardinium* 132
4.12.3 Polydnaviruses in Parasitoids 139
4.12.4 Gut Symbionts in Arthropods 141
4.12.5 Symbiosis in Fungus-Growing Attine Ants 143
4.12.6 Southern Pine Beetle Symbionts 146
4.12.7 Aphid Symbionts 146
4.12.8 Tsetse Fly Symbionts 148
4.13 Insect Development 149
4.13.1 Oocyte Formation in *D. melanogaster* 149
4.13.2 Embryogenesis in *D. melanogaster* 151
4.13.3 Postembryonic Development 153
4.14 Dissecting Development with *D. melanogaster* Mutants 154
4.14.1 Maternal-Effect Genes 155

4.14.2 Zygotic-Segmentation Genes 157
4.14.3 Insect Metamorphosis 161
4.15 Interactions During Development 162
4.16 Similarities and Differences in Development in Other Insects 162
4.17 Evo-Devo and the Revolution in Developmental Studies 164
References Cited 166

PART II: MOLECULAR GENETIC TECHNIQUES 181

Chapter 5 Some Basic Tools: How to Cut, Paste, Copy, Measure, Visualize, and Clone DNA 183

5.1 Overview 183
5.2 Introduction to a Basic Molecular Biology Experiment 184
5.2.1 The Pros and Cons of Kits 185
5.2.2 A Simple Cloning Experiment 186
5.3 Extracting DNA from Insects 187
5.3.1 DNA Extraction Resulting in Loss of the Specimens 187
5.3.2 DNA Extraction That Does Not Require Destroying the Specimens 188
5.3.3 Assessing the Quality of Extracted DNA 189
5.4 Precipitating Nucleic Acids 190
5.5 Shearing DNA 192
5.6 Cutting DNA with Restriction Endonucleases 192
5.7 Joining DNA Molecules 195
5.8 Growth, Maintenance, and Storage of *E. coli* 196
5.9 Plasmids for Cloning in *E. coli* 196
5.10 Transforming *E. coli* with Plasmids 201
5.11 Purifying Plasmid DNA from *E. coli* 202
5.12 Electrophoresis in Agarose or Acrylamide Gels 204
5.13 Detecting, Viewing, and Photographing Nucleic Acids in Gels 206
5.14 Identifying Specific DNA by Southern Blot Analysis 206
5.15 Labeling DNA or RNA Probes 209
5.16 Removing DNA from Agarose Gels after Electrophoresis 210
5.17 Restriction-Site Mapping 211
General References 212
References Cited 213

Chapter 6 Some Additional Tools for the Molecular Biologist 215

6.1 Overview 215
6.2 Introduction 216
6.3 The Perfect Genomic Library 218
6.3.1 Lambda (λ) Phage as a Vector 219
6.3.2 Cloning with Cosmids 226
6.3.3 Cloning in the Filamentous Phage M13 228
6.3.4 Phagemids 230
6.3.5 BACs 230
6.4 cDNA Cloning 230