Integrated Genomics
A Discovery-based Laboratory Course

Guy A. Caldwell, Shelli N. Williams and Kim A. Caldwell
Department of Biological Sciences
The University of Alabama
Tuscaloosa, Alabama

John Wiley & Sons, Ltd
Integrated Genomics
Integrated Genomics
A Discovery-based Laboratory Course

Guy A. Caldwell, Shelli N. Williams and Kim A. Caldwell
Department of Biological Sciences
The University of Alabama
Tuscaloosa, Alabama
Guy A. Caldwell would like to dedicate this book to his mentor and friend, Jeffrey M. Becker, PhD

Shelli N. Williams would like to dedicate this to her mother, who taught her to be strong

Kim A. Caldwell would like to dedicate this to her beloved bunny rabbits, The Heft and Rustamov
Contents

Preface xi
Author biographies xiii
Acknowledgments xvii
List of figures xix

1 Introduction to basic laboratory genetics 1
   1.1 Transferring and handling C. elegans 5
   1.2 Introduction to laboratory genetics 8

2 Gene expression analysis using transgenic animals 15
   2.1 Transgenic gene expression analysis in C. elegans: lacZ staining 17
   2.2 Transgenic gene expression analysis in C. elegans: GFP analysis 22

3 Creation and testing of transgenic yeast for use in protein–protein interaction screening 27
   3.1 Small-scale transformation of S. cerevisiae 31
   3.2 Transformation of S. cerevisiae to test for non-specific interaction 36
   3.3 Assaying for protein–protein interaction by reporter gene expression 42

4 Yeast two-hybrid screening 47
   4.1 Protein–protein interaction screening of a C. elegans cDNA library 47
   4.2 Assaying for protein–protein interaction by reporter gene expression 54
5 Isolation and identification of interacting proteins 61
  5.1 Preparation of electrocompetent E. coli 61
  5.2 Isolation of DNA from yeast and electroporation of E. coli 65
  5.3 Small-scale isolation of plasmid DNA from E. coli: the mini-prep 71
  5.4 Sequencing of two-hybrid library plasmid DNA vectors 76

6 Using bioinformatics in modern science 83
  6.1 DNA sequence chromatogram 84
  6.2 BLASTing your sequence 85
  6.3 Evaluating sequence results and choosing an RNAi target 86
  6.4 Bioinformatics practice questions 89

7 Generation of an RNAi vector 93
  7.1 Small-scale isolation of genomic DNA from C. elegans 93
  7.2 PCR amplification of target gene sequence from C. elegans genomic DNA 97
  7.3 Preparations for cloning to generate RNAi vector 105
    7.3.1 Agarose gel electrophoresis 107
    7.3.2 Removal of dNTPs from PCR reaction 110
    7.3.3 Restriction enzyme digestion of PCR product and C. elegans RNAi vector 111
  7.4 Gel purification of DNA and ligation of vector and PCR-amplified DNA 117
    7.4.1 Preparative agarose gel electrophoresis 118
    7.4.2 Gel purification of DNA from agarose gel 119
    7.4.3 Ligation of vector and PCR-amplified DNA 121
  7.5 Transformation of ligation reactions 122
  7.6 PCR screening of transformation colonies 126
  7.7 Small-scale isolation of plasmid DNA from E. coli: the mini-prep 130
  7.8 Verifying successful ligation by restriction digestion 133

8 RNA-mediated interference by bacterial feeding 137
  8.1 Preparation of RNAi-feeding bacteria for transformation 139
  8.2 Media preparation for RNAi feeding 143
  8.3 Transformation of RNAi-feeding strain HT115(DE3) 147
  8.4 RNA interference by bacterial feeding of C. elegans 150
  8.5 Analyzing effects of dsRNAi 156
    8.5.1 Assaying for sterility (Ste) or embryonic lethality (Emb) 158
    8.5.2 Assaying for growth effect 159
    8.5.3 Assaying for morphological effects 159
    8.5.4 Assaying for general neuromuscular effects 161
    8.5.5 Assaying for specific neuronal effects 163
    8.5.6 Assaying for dauer formation 165
Appendix I  Recombinational cloning  
AI.1 Isolation of genomic DNA from *C. elegans*  
AI.2 PCR amplification of target gene sequence from *C. elegans* genomic DNA  
AI.3 Agarose gel electrophoresis and clean-up of PCR reaction  
AI.4 Entry vector cloning  
AI.5 Small-scale isolation of plasmid DNA from *E. coli*: the mini-prep  
AI.6 Destination vector cloning  
AI.7 Small-scale isolation of plasmid DNA from *E. coli*: the mini-prep

Appendix II  Recipes and media preparation  
Solution recipes  
Media preparation

Appendix III  Sterile techniques and worm protocols  
Sterile techniques  
Worm protocols

Appendix IV  Mutant *C. elegans* phenotypes

Appendix V  Vector maps

Subject index
Science is about discovery. The creation of knowledge from where once there was none. The process of discovery is the passion that drives research, is the fuel that has built our society, and spurred great technological advancement and achievement for the human race. Arguably the greatest monument to our kind has been the unraveling of our complete DNA sequence, the human genome. The wealth of underlying secrets and mysteries that scientists may reveal from this scroll of life has only begun to be tapped. The now mature fields of genomics and proteomics are merely conduits for future discovery and expansion of knowledge through the many other realms of cellular and molecular biology and their growing interdependence on computer science, physical science, and biological engineering. Through all this is that underlying thread of life, DNA, and its universal role as the primary code for how cells, microbes, viruses, animals, plants, and humans reproduce and function. Could there be a greater quest or adventure than to investigate the underlying mechanisms of nature?

College and university life is about discovery too. As students progress through their most formative years of self-growth and independence, capturing the essence of the learning experience is often lost in the practical laboratory environment. Although any accomplished researcher can appreciate a finely devised and written experimental protocol, it is clear to even the novice that actual discovery is not possible in the context of the ‘cookbook recipes’ common to laboratory-based classroom experience. We strive to change this widely held paradigm.

*Integrated Genomics* is an outgrowth of a highly successful laboratory course taught for the past six years at The University of Alabama. This course was originally devised to meet the formative expectations of the Howard Hughes Medical Institute Undergraduate Biological Sciences Education Program Grant awarded to The University of Alabama
and has been modified with subsequent support from the National Science Foundation CAREER program. Both of these fine institutions share the common goal of fostering inquiry-based pedagogy and modernizing biological science education. This text provides the instructor with a broad and innovative strategy through which they may present a true discovery-based experience for their students. *Integrated Genomics* represents a continuum of flexible yet organized exercises, wherein concepts of basic microbiology, genetics, molecular biology, genomics, proteomics, and bioinformatics are enmeshed within a framework of experiments for which the results are not predetermined. The application of simple model organisms such as bacteria, yeast and worms in this course not only keeps the overall cost of implementing the class down, but equally introduces students to the unity of biological function and the power of comparative genomic analysis. Unequivocally, the burden of discovery-based pedagogy is a challenge, but one that is welcomed by the more ambitious of teachers, where the rewards for both the instructor and students are immeasurable. In this regard, the experimental design of *Integrated Genomics* is inherently adaptable and enables the instructor to 'plug and play' with their own favorite genes and proteins. We sincerely hope that this philosophy will be embraced and extended by inventive professors, and we welcome their feedback.

Discovering the process of discovery is what this book is all about. Moreover, it is a process that is addictive – and simply a lot of fun! Enjoy!

You can find further information about *Integrated Genomic* instructional materials and experimental reagents, including all strains and plasmids at www.wiley.com/go/caldwell

Guy A. Caldwell  
Shelli N. Williams  
Kim A. Caldwell
Guy A. Caldwell, PhD, is an Associate Professor in the Department of Biological Sciences at The University of Alabama, where since 1999 he has held an undergraduate professorial appointment from the Howard Hughes Medical Institute. He holds an adjunct appointment at the University of Alabama at Birmingham, as an Assistant Research Professor of Neurology. In 2001, Dr Caldwell was named a Basil O’Connor Scholar of The March of Dimes Birth Defects Foundation for his research into the molecular basis of childhood birth defects of the brain. Dr Caldwell is a recipient of grants from The March of Dimes, National Institutes of Health, Dystonia Medical Research Foundation, Parkinson’s Disease Foundation, National Parkinson Foundation, and the Bachmann-Strauss Dystonia & Parkinson Foundation. In 2003, The Caldwell Laboratory was selected as 1 of only 11 groups worldwide to represent the research goals of The Michael J. Fox Foundation for Parkinson’s Research in their Protein Degradation Grant Initiative. For his combined teaching and research efforts, Dr Caldwell was also chosen as the recipient of a 2003 CAREER Award from the National Science Foundation. In 2005, he was named Alabama Professor of the Year by the Carnegie Foundation for the Advancement of Teaching and Council for Advancement and Support of Education. Dr Caldwell, a native of the New York City area, received his undergraduate degree in Biology from Washington & Lee University in 1986 and his PhD in Cell, Molecular & Developmental Biology from The University of Tennessee in 1993. Following receipt of his doctorate, he moved to Columbia University in New York where he was twice named the recipient of fellowships from the National Institute of Neurological Disease and Stroke. He is the author of two editions of a widely adopted textbook, Biotechnology: A Laboratory Course, published worldwide in three languages. He teaches courses in Integrated Genomics, Neuronal Signaling, General Biology, and an acclaimed seminar on the societal impact of the Human Genome Project.
Shelli N. Williams, PhD, is a research scientist at a private forensic company based in Virginia. Following her early graduation magna cum laude from undergraduate studies, Dr. Williams began her graduate work in the laboratory of Drs. Guy and Kim Caldwell at The University of Alabama, where she earned her doctorate from The University of Alabama in 2006. Dr. Williams served as an adjunct faculty member in New College, an interdisciplinary department at The University of Alabama, where she was the instructor for a seminar course demonstrating how the nature of the laboratory experience plays an essential role in the understanding and advancement of science. She has experience teaching introductory biology courses to both majors and non-majors students and has been a repeated guest lecturer in a cross-disciplinary bioethics class. As a PhD candidate, Dr. Williams served as a teaching assistant for Integrated Genomics, a discovery-based genomics course funded by the Howard Hughes Medical Institute. Dr. Williams was named the recipient of two university-wide Graduate Council Fellowships, as well as receiving recognition as an Isabella Hummel Graham Scholar honoring outstanding female students throughout the university. She also received a competitive Worthington Biochemical Travel Award from the American Society of Cell Biology, placing her among the highest honored student researchers at their 2003 conference. Subsequent graduate work establishing Caenorhabditis elegans as a model for epilepsy was highlighted in news releases by the Howard Hughes Medical Institute. In recognition of her accomplishments, Dr. Williams was awarded the 2005 Joab Langston Thomas Award, the top honor for PhD students in Biological Sciences at The University of Alabama.

Kim A. Caldwell, PhD, is an Assistant Professor in the Department of Biological Sciences at The University of Alabama. Dr. Caldwell is a Faculty Affiliate of The University of Alabama Center for Green Manufacturing and she is an Adjunct Research Assistant Professor in the Department of Neurology at the University of Alabama at Birmingham Medical School. Dr. Caldwell, a native of the Buffalo area, received her undergraduate degree in Recombinant Gene Technology from The State University of New York at Fredonia and her MS and PhD degrees in Biotechnology and Cell, Molecular & Developmental Biology, respectively, from The University of Tennessee. While at Tennessee, Dr. Caldwell was a four-time recipient of the Oak Ridge National Lab–UT Science Alliance Teaching/Research Award and the Chancellor’s Award for Extraordinary Professional Promise. Following receipt of her doctorate, she held postdoctoral research appointments at The Rockefeller University and Columbia University in New York, during which time she was named the recipient of a Revson Fellowship and a National Research Service Award from the National Institute of Child Health and Human Development. Her research has been published in many outstanding peer-reviewed journals, including Nature, Proceedings of the National Academy of Sciences, Journal of
Dr Caldwell serves as Director of the Howard Hughes Medical Institute Rural Science Scholars program at Alabama. Additionally, she has designed and taught courses in General Biology, a seminar on the societal impact of the Human Genome Project, and a course entitled 'The Language of Research', which she teaches jointly for Howard Hughes Research Interns at both Stillman College and The University of Alabama. For her teaching efforts, in 2005 Dr Caldwell was selected as a Education Fellow in the Life Sciences of the National Academy of Sciences.
Acknowledgments

The authors would like to acknowledge the significant contributions of several individuals and organizations that enabled this text to move from a concept to a reality.

Significant financial support has come from a Howard Hughes Medical Institute Undergraduate Biological Sciences Education Program Grant awarded to The University of Alabama and a National Science Foundation CAREER Award to Guy A. Caldwell.

Many thanks to Dr. Stevan Marcus and Peirong Yang of the M.D. Anderson Cancer Center for generously sharing their advice and yeast two-hybrid reagents. Thanks to Dr. Mingxia Huang of the University of Colorado Health Science Center for plasmid vector pACT2.2 and Dr. Andy Fire, Stanford University, for additional worm plasmids used in the development of this book. We wish to thank Dr. Martin Chalfie of Columbia University and Theresa Steirmagle of the Caenorhabditis Genetics Center (University of Minnesota) for providing some of the Caenorhabditis elegans strains used herein and Drs. Lisa Timmons (University of Kansas) and Lynn Boyd (University of Alabama in Huntsville) for helpful discussions on RNAi feeding methodologies. Thanks to Cody Locke for filming the worm videos associated with this text on-line. We also greatly appreciate the support of Dr. Erich Schwarz and his colleagues at Wormbase for advice on bioinformatics and for the continued development of such an outstanding resource.

We are most grateful to all of our past teaching assistants and students, who have generously contributed their thoughtful suggestions toward improvement of this course over the years. Notably, special thanks go to Jafa Armagost for her generous and careful assistance in the final editing of this text. Likewise, special thanks go to all the members of ‘The Worm Shack’ (Caldwell Laboratory) at The University of Alabama for their cumulative advice, assistance and tolerance (!) through the writing of this text. We wish
to sincerely thank our editor at Wiley, Nicky McGirr, for believing in this idea and patiently waiting for it to materialize. Finally, our single greatest thanks go to Dr Martha J. Powell, Chair of The University of Alabama Department of Biological Sciences, who has been our unwavering champion and inspirational leader at every turn in supporting the development and implementation of *Integrated Genomics*. 
List of figures

1.1 *Caenorhabditis elegans* lifecycle at 25 °C 2
1.2 Anatomical features of a *Caenorhabditis elegans* hermaphrodite 2
1.3 Progression through embryogenesis 2
1.4 Developmental staging of *Caenorhabditis elegans* hermaphrodites from L1 larva to adult 3
1.5 *Caenorhabditis elegans* adult male 9

2.1 Comparison of reporter gene expression results using either *lacZ* or GFP driven by a specific promoter fusion 16
2.2 Example of the method for making an agarose pad for use in mounting worms for microscopy 21

3.1 Schematic diagram depicting the yeast two-hybrid system 28
3.2 Transformed yeasts containing both bait and prey vectors, Panel A shows patches of yeast grown, then transferred to filter paper for assaying Begal activity as a readout for protein-protein interaction, as depicted in Panel B 44

4.1 Plasmid map of the yeast two-hybrid library vector pACT2.2 50
4.2 Images depicting set-up and results of yeast two-hybrid candidates of putative interactors 55

5.1 Priming and extension of a new DNA strand from a DNA template by DNA polymerase 76
5.2 Representation of an adenine nucleotide: dioxy vs. dideoxy 77

6.1 Example of a typical chromatographic read-out following automated DNA sequencing 84

7.1 Diagram of vector L4440 94
7.2 Example of primer design for PCR amplification of RNAi target sequence 102
7.3 Separation of DNA fragments in an agarose gel during electrophoresis 108
## List of figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.4</td>
<td>Schematic representation of cloning a fragment using a single restriction enzyme</td>
</tr>
<tr>
<td>7.5</td>
<td>Schematic representation of cloning a fragment using two different restriction enzymes</td>
</tr>
<tr>
<td>8.1</td>
<td>Induction of dsRNA production within bacterial cells HT115(DE3)</td>
</tr>
<tr>
<td>8.2</td>
<td>Examples of three different types of <em>Caenorhabditis elegans</em> mutant phenotypes</td>
</tr>
<tr>
<td>AII.1</td>
<td>Examples of well-seeded 60-mm worm plates. Note that both have bacterial lawns located in the center of the plate: (A) a small bacterial lawn for mating experiments; (B) a normal sized lawn for standard worm maintenance</td>
</tr>
<tr>
<td>AIII.1</td>
<td>Photograph depicting a worm pick</td>
</tr>
<tr>
<td>AIII.2</td>
<td><em>Caenorhabditis elegans</em> lifecycle at 20°C</td>
</tr>
<tr>
<td>AIII.3</td>
<td><em>Caenorhabditis elegans</em> lifecycle at 15°C</td>
</tr>
<tr>
<td>AV.1</td>
<td>Detailed diagram of vector pLexA</td>
</tr>
<tr>
<td>AV.2</td>
<td>Detailed diagram of vector pACT2.2</td>
</tr>
<tr>
<td>AV.3</td>
<td>Detailed diagram of vector L4440</td>
</tr>
</tbody>
</table>
Introduction to basic laboratory genetics

*Caenorhabditis elegans* emerged as a model system in the 1960s, when it was first championed by professor Sydney Brenner and colleagues at the Medical Research Council in London, England. *Caenorhabditis elegans* is an ideal model system for many reasons. It is relatively small (only 1 mm at adulthood), transparent, hermaphroditic, and easy to grow in the lab (it can be treated as a microorganism). Fortunately for researchers, it shares many characteristics with more complex organisms, including, but not limited to, neurons, basic neurotransmitters, hormones, and numerous basic developmental processes.

*Caenorhabditis elegans* is one of the three lab organisms used throughout this course. ‘The worm’, as it is fondly known, serves as a whole animal system in which to study the expression of a specific gene (Chapter 2) and the functional consequences of altering its activity (Chapter 8). Prior to embarking on these analyses, it is first necessary to learn basic worm husbandry and handling. *Caenorhabditis elegans* are typically grown in Petri dishes containing media seeded with a bacterial food source (*Escherichia coli*). The media consists of agar and nutrients for both the bacteria and worms.

Before beginning the actual handling of worms, it is important to understand the lifecycle (Figure 1.1) and basic anatomy (Figure 1.2).

*Caenorhabditis elegans* development and lifecycle

**Embryo**

Following the union of a sperm and oocyte within the gonad of the hermaphrodite, development of the embryo begins. This stage lasts for approximately 12 h, during which
2 Experiment 1

![Diagram of Caenorhabditis elegans lifecycle at 25°C](image)

**Figure 1.1** *Caenorhabditis elegans* lifecycle at 25°C

![Anatomical features of a Caenorhabditis elegans hermaphrodite](image)

**Figure 1.2** Anatomical features of a *Caenorhabditis elegans* hermaphrodite

During larval development, *C. elegans* increases in size and complexity. The first three larval stages, L1, L2, and L3, are distinguishable morphologically on a dissecting...
microscope only by differences in size. Examining the worms at higher power magnification will illuminate developmental differences, such as the start of gonadogenesis in the L2 stage and spermatogenesis in the L3 stage. The last larval stage, L4, is easily recognizable in worms because a clear spot in the middle of the ventral side of the animals becomes visible. This clearing is caused by development and formation of the vulval opening (Figure 1.4). Caenorhabditis elegans molt between each larval stage, shedding the old cuticle during these events. With some patience and a microscope of sufficiently strong magnification, it is possible to observe a worm crawling out of the old layer with hardly a pause in normal activity.

**Adult**

Adult *C. elegans* are recognizable not only by their larger size, but also by the presence of fully formed gonadal arms and a completely developed vulva. Young adults are distinguishable from older adults by the absence of eggs (Figure 1.4). Adulthood is reached approximately 48 h after initial fertilization of the oocyte by the sperm. Adult worms can live for weeks provided that they have a continuous food source. They typically produce