

***cis*-Regulatory Control of Three Cell Fate-Specific Genes in Vulval  
Organogenesis of *C. elegans* and *C. briggsae***

Thesis by

Martha Kirouac

In partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2003

(submitted August 26<sup>th</sup>, 2002)

©2003

Martha Kirouac

All Rights Reserved

## ACKNOWLEDGEMENTS

I have been very fortunate to have Dr. Paul Sternberg as my advisor. His enthusiasm and dedication toward science are an inspiration in my life, and his personal support has meant much to me.

I would like to thank my committee members, Dr. Scott Fraser, Dr. Ellen Rothenberg, Dr. Raymond Deshaies, and Dr. Elliot Meyerowitz, for their time and guidance.

To the Sternberg lab members, past and present, I offer thanks for discussions, guidance, and friendship. A special thanks to Yen, Aidyl, Lisa, Cheryl, Gladys, Rene, and Shawn with whom I have had the joy of sharing a room, and who know my penchant for singing while I work. Thanks to Dave, Takao, Bhagwati, Byung, Rene, Nadeem and Minqin, who have helped me shape my work into a story, and to Mary, for always caring how I was and showing me the ropes.

Thanks to Mr. Bandura, who first inspired my interest in biology. I still remember the inspiration I felt with your description of why leaves change color and how corn is fertilized. To Dr. Jill Salvo, who believed in my abilities long before I did, and who pushed me to achieve, I thank you for a great many things, not the least of which is your friendship.

To all things Prufrock, I owe great thanks. Members past and present, you have been my bedrock, you have helped me maintain sanity in the face of adversity, and we've had fun to boot. You have been my family, my housemates and my friends for the past six years. To my confidants and special friends, Yen, Jen-Jen, Tom, Beth, Cheryl, Chi,

Jasper, Keith, Lisa, Big Paul and Aidyl: with each of you I have shared special moments and each of you has a place in my heart. You have made my time in graduate school more enjoyable and fuller than I could have hoped.

To those people in my life who have been steadfast in their never-ending support and love: Mom, Dad, Ian, Gabrielle, and Aunty-Lou. Thanks for your love, patience, encouragement and support.

To Shanti, thanks for things that have been and things that are to be. You have made me smile when I most needed it, and you have supported me throughout our evolving friendship.

## ABSTRACT

The great-grandprogeny of the *Caenorhabditis elegans* vulval precursor cells (VPCs) adopt one of the final vulA, B1, B2, C, D, E and F cell types in a precise spatial pattern. Formation of the pattern of vulval cell types is likely to depend upon the *cis*-regulatory regions of the transcriptional targets of these intercellular signals in vulval development. The outcome of such differential activation will result in individual cell types. *egl-17*, *zmp-1*, *cdh-3* are expressed differentially in the developing vulva cells, providing a potential readout for different signaling pathways. To understand how different signaling pathways interact to specify unique vulval cell types in a precise pattern, I have identified upstream *cis*-regulatory regions that are sufficient for their ability to confer vulval cell type-specific regulation when fused in *cis* to the basal *pes-10* promoter. In the *egl-17* promoter, I have identified a 143 base pair (bp) region that drives vulC and vulD expression, and a 102 bp region that is sufficient to drive the early expression in presumptive vulE and vulF cells. In the *zmp-1* promoter, I have identified a 300 bp region that is sufficient to drive expression in vulE, vulA and the anchor cell. In the *cdh-3* promoter, I have identified a 689 bp region sufficient to drive expression in the anchor cell and vulE, vulF, vulD and vulC, a 155 bp region sufficient to drive only anchor cell expression, and a separate 563 bp region that was also sufficient to drive expression in these vulval cells. I have identified the *C. briggsae* homologs of these three genes, and the corresponding control regions, and tested these regions in both *C. elegans* and *C. briggsae*. I find that these regions of similarity in *C. elegans* and *C. briggsae* upstream of *egl-17*, *zmp-1*, and *cdh-3* promote expression in vulval cells and the anchor cell. Using

the regions defined by the sufficiency analysis and phylogenetic footprinting, I have been able to isolate over-represented sequences that may play important roles in conferring vulval and anchor cell expression.

## TABLE OF CONTENTS

Acknowledgments	iii
Abstract	v
Table of Contents	vii
<b>Chapter 1: Transcriptional <i>cis</i>-Regulation</b>	<b>I-1</b>
Introduction	I-2
<i>cis</i> -acting regulatory elements of transcription in eukaryotes	I-2
Transcriptional regulation in <i>C. elegans</i>	I-6
Conservation of trans-acting transcriptional regulators in <i>C. elegans</i>	I-7
Vulva cell specification and intracellular signaling pathways	I-8
Genomic regulatory network analysis	I-14
<i>egl-17</i> and the FGF family	I-15
<i>zmp-1</i> and the Matrix Metalloproteinases	I-17
<i>cdh-3</i> and the Cadherins	I-18
Regulatory analysis in <i>C. elegans</i>	I-19
Dissection of co-regulated genes	I-25
Phylogenetic footprinting	I-29
Thesis overview	I-32
References	I-33
Figure 1: Vulva formation in <i>C. elegans</i>	I-49
Figure 2: Available vulval marker gene expression patterns in <i>C. elegans</i>	I-51
Figure 3: <i>egl-17::GFP</i>	I-53
Figure 4: <i>zmp-1::GFP</i>	I-55
Figure 5: <i>cdh-3::GFP</i>	I-57
<b>Chapter 2: <i>cis</i>-Regulatory Control of Cell Fate-Specific Genes in                     <i>Caenorhabditis elegans</i> Vulval Organogenesis</b>	<b>II-1</b>
Abstract	II-2

Introduction	II-3
Materials and Methods	II-4
Generation of <i>C. elegans</i> promoter GFP constructs	II-4
Generation of <i>C. elegans</i> promoter deletion GFP constructs	II-6
Sequencing of constructs	II-6
Microinjection of promoter GFP constructs into <i>C. elegans</i>	II-6
Microscopy of transgenic animals	II-7
Prediction of binding sites using Transfac database	II-7
AlignACE predictions of over-represented sequences	II-8
Results	II-8
Vulval specificity in the <i>egl-17</i> cis-regulatory region in <i>C. elegans</i>	II-9
Vulva and anchor cell specificity in the <i>zmp-1</i> cis-regulatory region in <i>C. elegans</i>	II-11
Vulva and anchor cell specificity in the <i>cdh-3</i> cis-regulatory region in <i>C. elegans</i>	II-13
Transfac putative binding site predictions in upstream sequences	II-16
AlignACE predictions of over-represented sequences	II-17
Discussion	II-21
<i>egl-17</i>	II-24
<i>zmp-1</i>	II-25
<i>cdh-3</i>	II-26
Distance of elements from translational start sites in <i>egl-17</i> , <i>zmp-1</i> , and <i>cdh-3</i>	II-28
Analysis of putative <i>trans</i> -acting factors	II-28
Analysis of over-represented sequences in regions of sufficiency	II-29
Conclusions	II-30
Acknowledgments	II-31
References	II-32
Figure 1: Marker gene expression summary	II-36
Figure 2: Initial dissection of <i>egl-17</i> , <i>zmp-1</i> , and <i>cdh-3</i> regulatory regions	II-38
Figure 3: Upstream regions that direct <i>egl-17</i> expression	II-40

Figure 4: Upstream sequences of mk84-148, mk50-51, mk96-134 and mk66-67	II-42
Figure 5: Multiple regions direct <i>zmp-1</i> expression	II-44
Figure 6: Regions that direct <i>cdh-3</i> expression	II-46
Table 1: Transfac binding site predictions for regions that confer cell-specific expression	II-48
Table 2: AlignACE predictions of over-represented sequences	II-51
Supplemental Table 1: PCR primers	II-55
Supplemental Figure 1: <i>egl-17</i> cis-regulatory deletion analysis	II-59
Supplemental Figure 2: <i>zmp-1</i> cis-regulatory deletion analysis	II-61
Supplemental Figure 3: <i>cdh-3</i> cis-regulatory deletion analysis	II-64

**Chapter 3: Three Genes, Two Species: A Comparative Analysis of Upstream  
Regulatory Sequences Sufficient to Direct Vulval Expression in  
*C. elegans* and *C. briggsae***

Abstract	III-2
Introduction	III-3
Materials and Methods	III-5
Protein prediction of EGL-17, ZMP-1 and CDH-3 homologs in <i>C. briggsae</i>	III-5
Analysis of homologous upstream sequences in <i>C. elegans</i> and <i>C. briggsae</i>	III-5
Generation of <i>egl-17</i> , <i>zmp-1</i> , and <i>cdh-3</i> <i>C. briggsae</i> promoter GFP constructs	III-6
Microinjection of promoter GFP constructs into <i>C. elegans</i>	III-7
Microinjection of promoter GFP constructs into <i>C. briggsae</i>	III-7
Microscopy of transgenic animals	III-7
Prediction of binding sites using Transfac database	III-8
AlignACE predictions of over-represented sequences	III-8
Results	III-9
<i>C. briggsae</i> homologs of <i>egl-17</i> , <i>zmp-1</i> , and <i>cdh-3</i>	III-9

Comparative sequence analysis	III-11
Analysis of <i>C. briggsae</i> upstream regions	III-14
Transfac binding site prediction in conserved regions	III-18
AlignACE predictions of over-represented sequences	III-19
Discussion	III-22
Phylogenetic footprinting	III-23
Potential for specific isolation of <i>trans</i> -acting factor binding sites by phylogenetic footprinting between <i>C. elegans</i> and <i>C. briggsae</i>	III-24
Analysis of putative <i>trans</i> -acting factors using the Transfac database	III-25
Analysis of over-represented sequences in regions of sufficiency	III-26
Implications of cross-species comparison of <i>egl-17</i> , <i>zmp-1</i> and <i>cdh-3</i>	III-26
Conclusions	III-28
Acknowledgments	III-29
References	III-30
Figure 1: EGL-17 clustalW alignment in <i>C. elegans</i> and <i>C. briggsae</i>	III-34
Figure 2: ZMP-1 clustalW alignment in <i>C. elegans</i> and <i>C. briggsae</i>	III-36
Figure 3: CDH-3 clustalW alignment in <i>C. elegans</i> and <i>C. briggsae</i>	III-38
Figure 4: Seqcomp and Family Relations predictions for <i>egl-17</i> , <i>zmp-1</i> , and <i>cdh-3</i> upstream sequences	III-40
Table 1: Summary of construct expression patterns	III-42
Figure 5: <i>egl-17</i> nucleotide sequences of important regions	III-44
Figure 6: <i>zmp-1</i> nucleotide sequences of important regions	III-46
Figure 7: <i>cdh-3</i> nucleotide sequences of mk96-134 and mk162-163	III-48
Figure 8: <i>cdh-3</i> nucleotide sequences of mk66-67 and mk164-165	III-50
Figure 9: <i>C. briggsae</i> upstream regions injected in <i>C. elegans</i>	III-52
Table 2: Transfac binding site predictions in regions of similarity between <i>C. elegans</i> and <i>C. briggsae</i>	III-54
Table 3: AlignACE predictions of over-represented sequences	III-58

<b>Chapter 4: Summary</b>	<b>IV-1</b>
Thesis Summary	IV-2
Sufficiency analysis	IV-3
Determining the necessity of regions defined by sufficiency analysis	IV-4
Phylogenetic footprinting studies of <i>cis</i> -regulatory sequences	IV-4
Practical considerations when identifying phylogenetic footprints	IV-6
Combining the results of sufficiency testing and phylogenetic footprinting studies	IV-6
Analysis of putative <i>trans</i> -acting factors	IV-7
Genomic analysis	IV-11
References	IV-12
Figure 1: Selection of nematode species for comparative genomic analysis	IV-14
Figure 2: Combined results of <i>egl-17</i> sufficiency and phylogenetic analyses	IV-16
Figure 3: Combined results of <i>zmp-1</i> sufficiency and phylogenetic analyses	IV-18
Figure 4: Combined results of <i>cdh-3</i> sufficiency and phylogenetic analyses	IV-20
Table 1: Effect of genetic background on marker expression	IV-22