UNIVERSITY OF CALIFORNIA, SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Identification of regulatory factors that control nervous system form, function, and regeneration in the planarian *Schmidtea mediterranea*

A dissertation submitted in partial satisfaction of the

Requirements for the degree Doctor of Philosophy

in

Biology

by

Martis William Cowles

Committee in charge:

University of California, San Diego

Professor James Posakony
Professor Deborah Yelon

San Diego State University

Professor Ricardo Zayas, Chair
Professor Robert Zeller, Co-Chair
Professor Robert Edwards

2014

The Dissertation of Martis William Cowles is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

University of California, San Diego

San Diego State University

2014

DEDICATION

I dedicate this thesis to my family. To my parents, thank you for encouraging me to always follow my dreams and supporting me along the way. To my wife Claire, thank you for always being there for me, in both good and bad times; I can't imagine this experience without you.

EPIGRAPH

Science is much more than a body of knowledge. It is a way of thinking.

This is central to its success. Science invites us to let the facts in,

even when they don't form with our preconceptions.

- Carl Sagan

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AP	Alkaline Phosphatase
bHLH	Basic Helix Loop Helix
BLAST	Basic local alignment search tool
BrdU	5-bromo-2'-deoxyuridine
cDNA	Complementary DNA
CG	Cephalic ganglia
ChIP-Seq	Chromatin immunoprecipitation sequencing
CNS	Central nervous system
COE	Collier/Olfactory-1/Early B-cell factor
dFISH	Double Fluorescent in situ hybridization
DG	Dentate gyrus
DMSO	Dimethyl sulfoxide
ES	Embryonic stem cells
FACS	Fluorescent activated cell sorting
FISH	Fluorescent in situ hybridization
GFP	Green fluorescent protein
Gy	Gray
HCl	Hydrocholoric acid
Hr	Hour
iN	Induced neuron

iPS	Induced pluripotent stem cell
PCR	Polymerase chain reaction
PH3	Phosphorylated histone H3 serine 10
PMP	Postmitotic progeny
PR	Photoreceptor
qPCR	Real-time quantitative polymerase chain reaction
Reg	Regeneration
RNA-seq	RNA sequencing
RNAi	RNA interference
S-phase	Synthesis phase
SVZ	Subventricular zone
VNC	Ventral nerve cord
WISH	Whole mount in situ hybridization

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my mentor Dr. Ricardo Zayas for his guidance throughout my graduate training. He has a contagious passion for science, which has made this part of my scientific career a true joy. His tireless dedication to his students is truly an inspiration.

I would like to thank the rest of my committee members, Dr. Robert Edwards, Dr. James Posakony, Dr. Deborah Yelon, and Dr. Robert Zeller, for their time, insight and thoughtful comments. It was a wonderful opportunity to have such great minds contributing to my research.

I would like to thank all of the past and current members of the Zayas lab for their helpful comments and thought provoking conversations. In particular, I would like to thank Jordana Henderson, Dr. Amy Hubert, and Kelly Ross for their help with editing my fellowship applications, abstracts, and manuscripts and Brianna Stanley and David Brown for providing technical assistance that was critical to the success of this project.

Chapter 1, in full, is a reprint of the material as it appears in the Journal of Development 2013. Cowles, MW.; Brown DDR.; Nisperos, SV.; Stanley BN.; Pearson BJ.; Zayas RM. The dissertation author was the primary investigator and author of this manuscript.

Chapter 2, in full, has been submitted for publication of the material. Cowles, MW.; Stanley, BN.; Omuro, KC.; Quintanilla CG.; Zayas, RM. The dissertation author was the primary investigator and author of this manuscript.

VITA

EDUCATION

Doctor of Philosophy in Biology – March 2014
University of California San Diego and San Diego State University
Joint Doctoral Program in Biology – Cell and Molecular Biology

Bachelor of Science, Biology - May 2007
San Diego State University

RESEARCH EXPERIENCE

Graduate thesis project (Ph.D. Program), San Diego State University
Advisor: **Ricardo Zayas** August 2010-present
- Identification of regulatory factors that control adult neurogenesis in planarians

Graduate thesis project (Master's Program), San Diego State University
Advisor: **Ricardo Zayas** August 2008-2010
- Analysis of neuronal migration genes in the planarian *Schmidtea mediterranea*

Research Manufacturing Associate, Illumina, La Jolla CA
- Performed high-throughput genotyping assays
- Conducted routine maintenance on TECAN liquid-handling robots

TEACHING EXPERIENCE

Teaching Assistant, San Diego State University
Cell and Molecular Biology Lab (2008-2010)
- Presented course material and demonstrated lab techniques

Supplemental Instructor, Organismal and Cell Biology
Mira Costa Community College (Spring 2003 and 2005)
- Hosted group meetings to review challenging lecture topics

Piano Instructor
Martis' Musical Instruction (Fall 2003-Spring 2007)
- Provided classical and jazz piano instruction

MENTORING EXPERIENCE

Supervision of undergraduate research, 3 students

San Diego State University (1 student Fall 2009-Summer 2011, 1 student Fall 2011-Spring 2013, 1 student Summer 2013-Present)

- Characterized axon guidance genes in the planarian *S. mediterranea*

PUBLICATIONS

Cowles, M.W., Stanley, B.N. Muth, K. Szeterlak, C., and Zayas, R.M. Semaphorin signaling regulates stem cell homeostasis and regeneration in planarians. *Submitted.*

Cowles, M.W., Stanley, B.N., Omuro, K.C., Quintanilla C.G. and Zayas, R.M. COE loss-of-function analysis reveals a genetic program underlying neuronal regeneration and stem cell regulation in planarians. *Submitted.*

Hubert A, Henderson J.M., **Cowles, M.W.**, and Zayas, R.M. A functional genomics screen identifies an Importin-a homolog as a regulator of stem cell function and tissue patterning during planarian regeneration. *In preparation.*

Cowles, M.W., Brown, D.R., Stanley, B.N., Nisperos S.V., Pearson, B.J., and Zayas, R.M. 2013. Genome-wide analysis of the basic Helix-Loop-Helix gene family in planarians identifies factors important for adult neurogenesis and neuronal regeneration. *Development* 140(23);4691-702. (Featured on cover).

Hubert A, Henderson J.M., Ross K.G., **Cowles M.W.**, Torres J, Zayas R.M. 2013. Epigenetic regulation of planarian stem cells by the SET1/MLL family of histone methyltransferases. *Epigenetics* 8 (1): 1-13.

Cowles, M.W., Hubert A, and Zayas R.M. 2012. A *Lissencephaly-1* homologue is essential for mitotic progression in the planarian *Schmidtea mediterranea*. *Developmental Dynamics* 241 (5): 901–910. (Featured on cover).

SCIENTIFIC PRESENTATIONS

Cowles, M.W., Brown, D.R., Stanley, B., Nisperos, S.V., Pearson, B.J., and Zayas, R.M. Planarians as a model to investigate factors regulating adult neurogenesis. Tissue Regeneration & Repair Gordon Research Seminar, New London NH. June 15-16, 2013. Oral Presentation.

Cowles, M.W., Brown, D.R., Stanley, B., Nisperos, S.V., Pearson, B.J., and Zayas, R.M. Planarians as a model to investigate factors regulating adult neurogenesis. Tissue Regeneration & Repair Gordon Research Conference, New London NH. June 16-21, 2013. Poster.

Cowles, M.W., Brown, D.R., Stanley, B., Nisperos, S.V., Pearson, B.J., and Zayas, R.M. Planarians as a model to investigate factors regulating adult neurogenesis and neuronal regeneration. 2nd North American Planarian Meeting, Kansas City MO, May 15-17, 2013. Oral Presentation.

Cowles, M.W., Brown, D.R., Stanley, B., Nisperos, S.V., Pearson, B.J., and Zayas, R.M. Identification of somatic neural progenitor populations in the planarian Schmidtea mediterranea. Stem Cell Interest Group. Sanford Consortium for Regenerative Medicine. La Jolla CA, March 27, 2013. Oral Presentation.

Cowles, M.W., Brown, D.R., Stanley, B., Nisperos, S.V., Pearson, B.J., and Zayas, R.M. Genome-wide analysis of the basic Helix-Loop-Helix gene family in planarians identifies factors involved in neurogenesis. Society for Developmental Biology 71st Annual Meeting, *Abs. #184*, Montreal, CA, July 19-23, 2012. Poster.

Cowles, M.W., Nisperos, S.V, Zayas, R.M. Characterization of regulatory factors that control neurogenesis in planarians. San Diego State University Student Research Symposium, *Abs. #397*, San Diego, CA. March 9-10, 2012. Oral Presentation.

Cowles, M.W., Hubert, A. and Zayas, R.M**.** Stem cell-based nervous system regeneration in planarians. 15th Annual Biosymposium, San Diego CA, October 7th, 2011. Oral Presentation.

Cowles, M.W., Hubert, A. and Zayas, R.M. A *Lissencephaly-1*-like gene is required for stem cell maintenance in the planarian *Schmidtea mediterranea.* 70th Society for Developmental Biology Annual Meeting, Chicago, IL. July 21-25, 2011. Poster.

Cowles, M.W., Hubert, A. and Zayas, R.M. A planarian ortholog of *Lissencephaly-1* is required for stem cell maintenance. San Diego State University Student Research Symposium, *Abs. #466*, San Diego, CA. March 4-5, 2011. Oral Presentation.

Cowles, M.W. and Zayas, R.M. Identification and functional analysis of neuronal migration genes in planarians. San Diego State University Student Research Symposium, *Abs. #362*, San Diego, CA. March 5-6, 2010. Oral Presentation.

Cowles, M.W. and Zayas, R.M. Identification and characterization of neuronal migration genes in planarians*.* 68th Society for Developmental Biology Annual Meeting, *Abs. L92,* San Francisco, CA. July 23-27, 2009. Poster.

Cowles, M.W. and Zayas, R.M. Identification and characterization of neuronal migration genes in planarians. West Coast Regional Developmental Biology Meeting, Pacific Grove, CA; May 1-3, 2009. Poster.

AWARDS

Inamori Foundation Fellowship (2012- Present)

Achievement Reward for College Scientists (ARCS) Fellowship (August 2011-Present)

2nd Place Winner, 25th Annual California State Student Research Competition. Fresno CA, May 6-7, 2011

President's Award, San Diego State University Student Research Symposium. San Diego CA, March 4-5, 2011

Graduate Student Travel Award, San Diego State University, provided funds to attend the 70th Society for Developmental Biology Annual Meeting, Chicago, IL. July 21-25, 2011.

Best Student Poster Award, West Coast Regional Society for Developmental Biology Meeting, Pacific Grove, CA; May 1-3, 2009.

PROFESSIONAL SOCIETY MEMBERSHIP

Society for Neuroscience, since 2010

Society for Developmental Biology, since 2009

ABSTRACT OF THE DISSERTATION

Identification of regulatory factors that control nervous system form, function, and regeneration in the planarian *Schmidtea mediterranea*

by

Martis William Cowles

Doctor of Philosophy in Biology

University of California, San Diego 2014

San Diego State University 2014

Professor Ricardo Zayas, Chair

Professor Robert Zeller, Co-Chair

Neurons are born from stem cells, migrate to their final location, form synaptic connections, and terminally differentiate by a process known as neurogenesis. Although this phenomenon is observed in adult organisms across metazoans, most animals lack the ability to repair catastrophic damages to the

central nervous system (CNS) caused by injury, disease or aging. By contrast, planarians have the amazing ability to regenerate all tissue types (including the CNS) from a population of pluripotent adult stem cells they maintain throughout their life, making these animals a powerful system to research stem cell-based regeneration *in vivo*. To investigate how adult stem cells are directed to generate new neurons during CNS regeneration, we examined the basic helix-loop-helix (bHLH) transcription factor family in planarians. Many bHLH family members regulate neurogenesis during development and are associated with nervous system diseases, yet their functions in adult stem cells and mature neurons remain unclear. We identified 44 planarian bHLH homologs, determined their tissue-specific expression in the adult animal, and examined their function using RNA interference. These analyses identified nine bHLHs expressed in stem cells and neurons that were required for CNS regeneration, including homologs of Collier/Olfactory-1/Early B-cell factor (*coe*), Single-minded (*sim*), and Hairy enhancer of split (*hesl-3*). Furthermore, we demonstrated that *coe*, *sim* and *hesl-3* mRNA were detected in lineage-committed progenitors. Our functional screen revealed that gene silencing of *coe* results in CNS regeneration defects. COE genes play conserved roles in nervous system development and are associated with CNS diseases; however, the genetic programs downstream of these genes remain largely unknown. By comparing the transcriptome profiles of control and *coe*-deficient animals, we identified over 900 differentially expressed genes, including 397 downregulated genes enriched for CNS functions. We examined downregulated genes and identified

new targets of COE in mature neurons, some of which were required for CNS regeneration. Furthermore, we found novel genes expressed in stem cell progeny that function downstream of COE and were critical for stem cell homeostasis. These findings demonstrate that COE regulates genetic programs essential for CNS homeostasis and regeneration, providing insights into how COE proteins function in the adult nervous system.

INTRODUCTION OF THE DISSERTATION

ADULT NEUROGENESIS AND CENTRAL NERVOUS SYSTEM REGENERATION

Neurogenesis is the process by which new neurons are born from stem cells, migrate to their correct location, form appropriate synaptic connections, and terminally differentiate. Once thought to only occur during development [1], adult neurogenesis has been observed throughout the animal kingdom, including humans [2, 3]. These observations overturned the long held dogma that the central nervous system (CNS) is a static organ post embryogenesis and has revolutionized the way we view nervous system maintenance and repair [4].

Mammalian adult neurogenesis is observed in two neurogenic zones in the brain, the dentate gyrus (DG) and subventricular zone (SVZ), and is essential to maintain functional plasticity of neural circuitry in the adult CNS [3]. Neural stem cells in the DG generate neural precursors that differentiate into hippocampal neurons, which function in learning and memory [5]. Remarkably, humans generate approximately 700 hippocampal neurons per day, suggesting that hippocampal neurogenesis is essential for normal brain function [6]. The connection between adult neurogenesis and learning and memory is not only observed in mammals, but is also documented in diverse species including crickets [7] and song birds [8]. In the SVZ, newly generated neuronal precursors migrate to the olfactory bulb via the rostral migratory stream, where they differentiate into GABAergic and glutaminergic interneurons and integrate with the existing neural architecture [9]. New neuron formation in the olfactory bulb is important for the spatial and temporal detection of odors [10].

Despite the ability to create new neurons under normal physiological conditions, most animals (including humans) fail to repair major damage to the CNS caused by injury, disease or aging [11, 12]. However, the creation of new neurons throughout adulthood holds promise to the idea that stem cell-based therapies could be developed to repair injury to the CNS and reestablish normal nervous system function. Thus, a major focus of the field of regenerative medicine is to elucidate how stem cells can be directed to generate specific neuronal subtypes that functionally integrate with existing neuronal structures.

In vitro analyses of stem cell function have revealed mechanisms that drive neuronal differentiation and demonstrated that cell fate plasticity is much greater than previously expected. Neural stem cells can be isolated from adult brains, cultured *in vitro*, and differentiated into a wide variety of neurons [13]. Embryonic stem (ES) cells from the inner cell mass of embryos can be isolated to generate neural precursors by co-culturing cells with stromal cells or astrocytes [14, 15] or by treating cells with sequential growth factors [16, 17] or signaling molecules [18, 19]. Remarkably, fibroblasts can be converted into induced pluripotent stem (iPS) cells by ectopically expressing specific transcription factors [20-22]. The generation of iPS cells from fibroblasts sidesteps many of the ethical implications associated with ES cells, which are harvested from discarded embryos obtained from *in vitro* fertilization clinics. Furthermore, this tool allows investigators to generate neurons from patients with CNS-specific disease pathologies such as Parkinson's or Alzheimer's disease and examine how these "disease state" neurons respond to

genetic manipulations or pharmacological treatments [23]. More recently, fibroblasts were directly converted into induced neurons (iN) by ectopic activation of the transcription factors ascl1, pou3f2, and Myt1l [24, 25]. The ability to dramatically alter developmental potential and cell fate demonstrates that cellular identity is highly plastic and can be directed toward alternative states via genetic or pharmacological intervention. In addition, these studies provide the methodology to generate a large number of neural precursors that could be used for stem cell-based therapeutic applications.

Although cell culture research has vastly improved our understanding of stem cell biology, there are major limitations and considerations when using cultured stem cells for regenerative applications. First, there is the major risk of injected ES or iPS cells to generate tumors [27]. Therefore, great care must be taken when designing stem cell-based therapies and determining their safety in humans. Second, culturing stem cells *in vitro* is highly artificial and poorly reflects the environment in which these cells will finally be placed. Stem cell behavior and differentiation is greatly influenced by their niche or microenvironment, which includes a diverse concoction of secreted factors (such as growth factors, cytokines and hormones) that regulate cell proliferation, migration, differentiation, and survival [28]. Consequently, injection of stem cells into a local environment that is not permissive for neurogenesis will fail to induce neural differentiation and survival [29]. In mammals, CNS injury induces a strong immune response that causes scar formation, which ultimately leads to additional neuronal death,

degeneration of damaged axons, a physical barrier that inhibits new axon projections across the injury site, and a "hostile" local environment that is not conducive to neurogenesis [30]. Thus, understanding what mechanisms control stem cell differentiation in a dish is only half of the puzzle; we must also investigate adult stem cell function *in vivo* to resolve how stem cells interact with their microenvironment and how these interactions modulate their proliferation, differentiation potential, maturation, and survival.

Unlike mammals, some animals, including hydra [31], planaria [32], zebrafish [33], and axolotl [34] can repair severed axons and/or regenerate new neurons following injury that integrate with existing neural architecture and contribute to functional recovery. These animals provide a unique opportunity to ask important unanswered questions in regeneration biology that cannot be easily addressed in mammalian animal models. For example: What regulatory factors function to promote or inhibit adult neurogenesis? Is neurogenesis differentially regulated in developing and adult tissues? Are there basic mechanisms that underlie nervous system regeneration in all animals? Comparative studies using animal models of CNS regeneration will be essential to elucidate how adult stem cells respond to CNS injury and contribute to its recovery.

In the following sections I will discuss neurogenesis regulation during development and focus on the roles of the basic helix-loop-helix (bHLH) transcription factor gene family in controlling this process. Next, I will introduce the amazing ability of planarians to regenerate and discuss how these animals provide a

tractable system to investigate adult stem cell function during CNS maintenance and regeneration *in vivo*. Finally, I will introduce my thesis project, which identified transcription factors important for adult neurogenesis during CNS regeneration using the planarian *Schmidtea mediterranea* as an animal model.

THE ROLE OF BASIC HELIX LOOP HELIX TRANSCRIPTION FACTORS IN NEUROGENESIS REGULATION

Highly conserved gene families regulate different stages of neural fate determination during development [35]. The phenotypic progression of neurogenesis (summarized in Figure 1) is initiated when stem cells receive signaling cues (such as FGF, Wnt or Hedgehog), which activate patterning transcription factors (homeodomain genes) [36]. In vertebrates, Otx, Gb and Hox gene families function to establish anteroventral positioning, whereas genes from the Pax, Nkx and Irx families provide positional cell fate information along the dorsoventral axis [35]. Next, patterning genes activate proneural factors of the basic helix-loop-helix (bHLH) gene family, which function to establish cells as neural progenitors by promoting early events in differentiation such as cell cycle exit and neural specification [37]. Although most bHLH genes activate neurogenesis, some bHLH genes, such as *Olig2* and *hes1*, restrict neural differentiation by inhibiting the function of other bHLH proteins and repressing proneural gene expression [38, 39]. Within a differentiating neuron, patterning and proneural protein combinations are integrated to deploy specific genetic programs (differentiation genes) that dictate

subtype specification [35, 40, 41]. Neurogenesis is completed as differentiation genes direct migration of neurons to their final location and guide formation of synaptic connections, thus facilitating maturation and terminal differentiation. Recent studies demonstrate that some transcription factors important for subtype specification continue to be expressed in mature neurons and are required to maintain neuronal identity and function; these factors are referred to as "terminal selector" genes [42, 43].

The bHLH superfamily consists of six monophyletic groups (Groups A – F) based on the presence or absence of various protein domains [44]. Group A (e.g. *achaete-scute*, *tal*, and *neuroD*) contain a single HLH motif and play major roles in ectoderm and mesoderm differentiation; group B genes (e.g. *myc* and *max*) also contain a leucine zipper domain and primarily function in cell proliferation and differentiation [37]; group C (e.g. *sim* and *arnt*) contain PAS domain(s) and regulate midline patterning, metabolism, and CNS development [45]; group D (e.g. *id* and *emc*) lack a basic domain and act to antagonize genes from group A [37]; group E (e.g. *hes* and *hey*) also contain an Orange and WRPW domain and act to repress neural differentiation [46]; group F (e.g. *collier* and *EBF*) contain a COE domain and regulate differentiation in a wide variety of cell types including neurons, myocytes, and B-cells [47, 48]. COE homologs also function as terminal selectors for cholinergic identity in *Drosophila melanogaster* and *Caenorhabditis elegans* [49, 50].

Although some bHLHs continue to be expressed in adult tissues and are associated with various human diseases including cancer [51-53], the role of these

genes outside of embryonic development are poorly understood. Furthermore, it is unknown whether these genes are redeployed following injury to regulate tissue morphogenesis or cell specification during CNS regeneration.

SCHMIDTEA MEDITERRANEA AS A MODEL OF STEM CELL-BASED TISSUE REGENERATION

Planarians are non-parasitic flatworms from the phylum Platyhelminthes, an understudied group of animals within the superphylum Lophotrochozoa [54]. These animals are among the simplest organisms to possess all three germ layers and bilateral symmetry [54]. Remarkably, planarians can regenerate entire animals from small body fragments due to a population of pluripotent adult stem cells (called neoblasts) they maintain throughout their lives [54-57]. Following amputation or injury, the stem cells proliferate, migrate to the wound site and form a regeneration blastema, where they differentiate to replace lost or damaged tissues [54]. Planarians also exhibit an extraordinary ability to alter the scale and proportion of tissues in response to metabolic needs, further demonstrating a remarkable amount of tissue plasticity even in uninjured animals [58]. The recent development of molecular tools for the planarian *Schmidtea mediterranea* have made this animal a powerful system to investigate the basic molecular mechanisms that underlie stem cell-based tissue regeneration *in vivo.* Molecular tools available for *S. mediterranea* include a sequenced genome [59], transcriptome profiles of stem cell, progeny and differentiated cell populations [60-62], whole mount *in situ* hybridization (WISH) to

examine tissue-specific gene expression [63-65], and RNA interference (RNAi) to analyze gene function [66-69].

The planarian stem cell pool makes up approximately 15-20% of all cells in the animal and shares conserved features with human embryonic stem cells [60-62]. Remarkably, implantation of a single stem cell into a lethally irradiated planarian is sufficient to reconstitute the entire stem cell pool, demonstrating that a fraction of planarian stem cells are truly pluripotent [70]. However, expression analyses of stem cell-specific genes have demonstrated that the planarian stem cell pool is heterogeneous [32, 71, 72]. Based on the transplantation studies described above, it is estimated that less than 5% of the stem cells are pluripotent, suggesting that planarians must also maintain a large population of lineage-restricted stem cells or progenitors [32, 70, 72]. In support of this hypothesis several laboratories have recently identified lineage-specific transcription factors expressed in subsets of the stem cells, which function to generate specific tissue types such as photoreceptors [73, 74] or protonephridia (the planarian excretory system) [75]. To date, the organization and complexity of the planarian stem cell hierarchy remains enigmatic.

THE MOLECULAR BASIS OF CENTRAL NERVOUS SYSTEM REGENERATION IN PLANARIANS

Planarians are among the simplest metazoans to possess a CNS, which consists of cephalic ganglia (the brain) and ventral nerve cords that extend along the length of the animal [76]. Despite having a relatively basal CNS morphology, the

planarian nervous system exhibits extensive structural complexity, with the presence of a submuscular, subepidermal and subgastrodermal nerve plexuses [77]. Expression analyses of neural specific genes have also revealed a considerable amount of molecular diversity [65, 78, 79]. Furthermore, many neural subtypes have been identified that are shared with higher organisms, such as dopaminergic, cholinergic and serotonergic neurons [80-84]. All neuronal subtypes are replaced within days after amputation. Thus, the planarian nervous system provides a unique paradigm to dissect the molecular mechanisms that guide the generation and differentiation of specific neural cell types during physiological turnover and regeneration of the adult nervous system.

In planarians, CNS patterning is maintained during normal physiological conditions by the constitutive expression of conserved secreted signaling molecules [85]. Wnt and Hedgehog signaling coordinate to establish anteroposterior polarity [86, 87], whereas BMP signaling determines the dorsoventral axis [88]. While these signaling molecules act to determine animal polarity, studies on a planarian FGF receptor-like gene, *nou-darake* (*ndk*, Japanese for "brains everywhere"), showed that FGF signaling plays a specific role in neural patterning and determination [89]. FGF signaling is vital for various stages of neural development and is conserved across metazoans [90]. This signaling pathway is activated by FGF ligands that bind FGF receptors and can be repressed by FGF-like receptors (FGFLRs), which lack a cytosolic activation domain and act to sequester FGF ligands [91, 92]. *Ndk* is predicted to encode a protein that is similar to FGFLRs and is expressed in discrete

cells located in the anterior region of the animal [89]. Gene knockdown of *ndk* induces the generation of ectopic brain-specific neurons and photoreceptors in more posterior regions of the animal. These data suggest that *ndk* inhibits FGF signaling and acts to restrict the generation of brain specific structures to the head [89]. Thus, signaling molecules such as Wnt, Hedgehog, and BMP, play roles in determining polarity, whereas FGF signaling may function more specifically in neural determination.

In planarians, injury activates a wound-induced program that functions to reestablish signaling gradients and activate genes required for regeneration [93]. Over one thousand genes are differentially expressed following head amputation [94]. Within 30 to 60 minutes following injury, intermediate early genes are activated [93]. A second wave of genes is upregulated between six to twelve hours following amputation, which include patterning factors such as Wnts [93]. Injuries resulting in loss of tissue (such as head amputation), induce an additional set of genes including follistatin, which is essential to promote stem cell proliferation required for regeneration blastema formation [93, 95, 96]. Recently, it was shown that the vast majority of wound control genes are expressed in muscle cells, demonstrating that planarian musculature plays a critical role in sensing injury and promoting appropriate regeneration responses [97].

The specific molecules or pathways activated by wound control genes or signaling molecules remain unknown. However, several transcription factors have been identified that are required for the generation of specific neuronal

subpopulations [73, 74, 98, 99]. For example, *ovo* acts as a master regulator of both pigment and photosensitive cells of the photoreceptors [74], whereas *pitx* and *lhx1/5-1* are required for specification and maintenance of serotonergic neurons [98, 99]. Unlike *pitx* and *lhx1/5-1*, which are only detected in stem cells following amputation, *ovo* is expressed in a small subpopulation of cycling stem cells in uninjured animals located just posterior to the photoreceptors [74]. Thus, *ovo*$^{+}$ cells represent the first identified neural progenitor population in uninjured planarians. Whether or not *ovo*$^{+}$ cells have the capacity to self renew or generate cell types outside of the optic system remains unclear.

The major goal of my study was to identify regulatory factors that control adult neurogenesis during CNS maintenance and regeneration. I chose to perform my studies in the planarian *S. mediterranea* for two major reasons: 1) planarians quickly and efficiently regenerate CNS structure *de novo* following injury, and 2) unlike hydra or zebrafish, which dedifferentiate cells near the wound site as a source of cells for regeneration blastema formation [100], planarian regeneration is dependent on the proliferation and differentiation of pluripotent adult stem cells.

In chapter 1 we performed a genome-wide characterization of the bHLH gene family in the planarian *S. mediterranea*. Given the highly conserved roles that these factors play in neural stem cells and progenitors during development, we hypothesized that these factors would also regulate neurogenesis and nervous system function in adult animals; furthermore, we speculated that our analyses might further reveal heterogeneity in the planarian stem cell pool. Consistent with

our hypotheses, we identified nine bHLHs expressed in stem cells and neurons that were required for CNS regeneration, including homologs to *Collier/Olfactory-1/Early B-cell factor* (*coe*), *Single-minded* (*sim*), and *Hairy enhancer of split* (*hesl-3*). We showed that *coe*, *sim* and *hesl-3* label neural progenitor populations in uninjured animals and found that coe^{+} and sim^{+} progenitors underlie formation of the regeneration blastema. These findings demonstrate that bHLH function is critical for adult neurogenesis during CNS maintenance and regeneration and identify several genes that label novel neural progenitor populations in planarians. Our data further support the hypothesis that the planarian stem cell pool is comprised of a diverse collection of lineage-committed progenitors [72].

One gene of interest identified from our screen was *coe*, which is known to play highly conserved roles in neural specification and maintenance [47, 48, 50, 101]. Interestingly, COE proteins are also proposed to function as tumor suppressors and are associated with several cancers of the CNS such as glioblastoma [53, 102, 103]; however, it remains unclear how defects in COE contribute to CNS dysfunction. In chapter 2 we compared the transcriptome profiles of control and *coe*-deficient animals, identifying novel downstream targets of COE in mature neurons and stem cell progeny. Our findings revealed a genetic program downstream of COE that is essential for stem cell homeostasis and maintenance of neuronal identity in multiple classes of adult neurons. This study provides insights into how COE proteins function in the adult CNS and demonstrates how planarians

can be used to identify and characterize regulatory molecules with roles in CNS maintenance and repair.

FIGURES

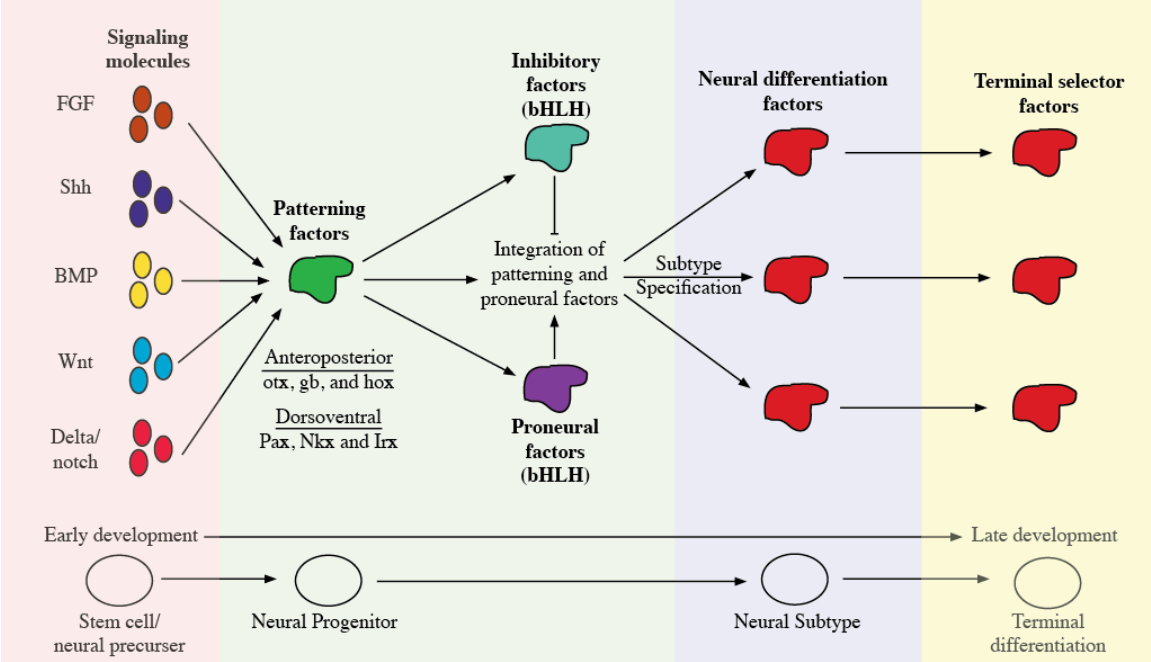


Figure 1.1: Phenotypic progression of neurogenesis during development. During early development, signaling molecules activate patterning factors; patterning proteins then activate proneural factors from the bHLH gene family. The combinatorial effects of patterning and proneural factors activate developmental programs that initiate neural subtype specification by promoting neural differentiation and maturation. Terminal selector factors function to maintain neuronal identity in mature neurons. Arrows and barbed ends indicate gene activation and repression, respectively.

REFERENCES

1. Ramón y Cajal, S., *Degeneration and Regeneration of the Nervous System*, ed. R. Day, translator. 1928, London: Oxford University Press.

2. Gage, F.H., *Neurogenesis in the adult brain.* J Neurosci, 2002. **22**(3): p. 612-3.

3. Lindsey, B.W. and V. Tropepe, *A comparative framework for understanding the biological principles of adult neurogenesis.* Prog Neurobiol, 2006. **80**(6): p. 281-307.

4. Kempermann, G., *Adult neurogenesis*. 2nd ed. 2011, New York: Oxford University Press. xiii, 601 p.

5. Kempermann, G., *New neurons for 'survival of the fittest'.* Nat Rev Neurosci, 2012. **13**(10): p. 727-36.

6. Spalding, K.L., et al., *Dynamics of hippocampal neurogenesis in adult humans.* Cell, 2013. **153**(6): p. 1219-27.

7. Cayre, M., et al., *Understanding the regulation and function of adult neurogenesis: contribution from an insect model, the house cricket.* Chem Senses, 2007. **32**(4): p. 385-95.

8. Barnea, A. and V. Pravosudov, *Birds as a model to study adult neurogenesis: bridging evolutionary, comparative and neuroethological approaches.* Eur J Neurosci, 2011. **34**(6): p. 884-907.

9. Gheusi, G., G. Lepousez, and P.M. Lledo, *Adult-born neurons in the olfactory bulb: integration and functional consequences.* Curr Top Behav Neurosci, 2013. **15**: p. 49-72.

10. Lledo, P.M. and S. Lagier, *Adjusting neurophysiological computations in the adult olfactory bulb.* Semin Cell Dev Biol, 2006. **17**(4): p. 443-53.

11. Yiu, G. and Z. He, *Glial inhibition of CNS axon regeneration.* Nat Rev Neurosci, 2006. **7**(8): p. 617-27.

12. Fawcett, J.W., *Overcoming inhibition in the damaged spinal cord.* J Neurotrauma, 2006. **23**(3-4): p. 371-83.

13. Rietze, R.L. and B.A. Reynolds, *Neural stem cell isolation and characterization.* Methods Enzymol, 2006. **419**: p. 3-23.

14. Kawasaki, H., et al., *Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity.* Neuron, 2000. **28**(1): p. 31-40.

15. Morizane, A., et al., *Generation of graftable dopaminergic neuron progenitors from mouse ES cells by a combination of coculture and neurosphere methods.* J Neurosci Res, 2006. **83**(6): p. 1015-27.

16. Reubinoff, B.E., et al., *Neural progenitors from human embryonic stem cells.* Nat Biotechnol, 2001. **19**(12): p. 1134-40.

17. Perrier, A.L., et al., *Derivation of midbrain dopamine neurons from human embryonic stem cells.* Proc Natl Acad Sci U S A, 2004. **101**(34): p. 12543-8.

18. Gerrard, L., L. Rodgers, and W. Cui, *Differentiation of human embryonic stem cells to neural lineages in adherent culture by blocking bone morphogenetic protein signaling.* Stem Cells, 2005. **23**(9): p. 1234-41.

19. Itsykson, P., et al., *Derivation of neural precursors from human embryonic stem cells in the presence of noggin.* Mol Cell Neurosci, 2005. **30**(1): p. 24-36.

20. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.

21. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. **131**(5): p. 861-72.

22. Wernig, M., et al., *In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state.* Nature, 2007. **448**(7151): p. 318-24.

23. Peitz, M., J. Jungverdorben, and O. Brustle, *Disease-specific iPS cell models in neuroscience.* Curr Mol Med, 2013. **13**(5): p. 832-41.

24. Marro, S., et al., *Direct lineage conversion of terminally differentiated hepatocytes to functional neurons.* Cell Stem Cell, 2011. **9**(4): p. 374-82.

25. Vierbuchen, T., et al., *Direct conversion of fibroblasts to functional neurons by defined factors.* Nature, 2010. **463**(7284): p. 1035-41.

26. Noble, M. and J. Dietrich, *The complex identity of brain tumors: emerging concerns regarding origin, diversity and plasticity.* Trends Neurosci, 2004. **27**(3): p. 148-54.

27. Lee, A.S., et al., *Tumorigenicity as a clinical hurdle for pluripotent stem cell therapies.* Nat Med, 2013. **19**(8): p. 998-1004.

28. Sequerra, E.B., et al., *Adult neural stem cells: plastic or restricted neuronal fates?* Development, 2013. **140**(16): p. 3303-9.

29. Lopez-Bendito, G. and P. Arlotta, *Cell replacement therapies for nervous system regeneration.* Dev Neurobiol, 2012. **72**(2): p. 145-52.

30. Kawano, H., et al., *Role of the lesion scar in the response to damage and repair of the central nervous system.* Cell Tissue Res, 2012. **349**(1): p. 169-80.

31. Kumar, A. and J.P. Brockes, *Nerve dependence in tissue, organ, and appendage regeneration.* Trends Neurosci, 2012. **35**(11): p. 691-9.

32. Rink, J.C., *Stem cell systems and regeneration in planaria.* Dev Genes Evol, 2013. **223**(1-2): p. 67-84.

33. Becker, C.G. and T. Becker, *Adult zebrafish as a model for successful central nervous system regeneration.* Restor Neurol Neurosci, 2008. **26**(2-3): p. 71-80.

34. Roy, S. and S. Gatien, *Regeneration in axolotls: a model to aim for!* Exp Gerontol, 2008. **43**(11): p. 968-73.

35. Guillemot, F., *Spatial and temporal specification of neural fates by transcription factor codes.* Development, 2007. **134**(21): p. 3771-80.

36. Jessell, T.M., *Neuronal specification in the spinal cord: inductive signals and transcriptional codes.* Nat Rev Genet, 2000. **1**(1): p. 20-9.

37. Bertrand, N., D.S. Castro, and F. Guillemot, *Proneural genes and the specification of neural cell types.* Nat Rev Neurosci, 2002. **3**(7): p. 517-30.

38. Kageyama, R., T. Ohtsuka, and T. Kobayashi, *Roles of Hes genes in neural development.* Dev Growth Differ, 2008. **50 Suppl 1**: p. S97-103.

39. Ross, S.E., M.E. Greenberg, and C.D. Stiles, *Basic helix-loop-helix factors in cortical development.* Neuron, 2003. **39**(1): p. 13-25.

40. Powell, L.M. and A.P. Jarman, *Context dependence of proneural bHLH proteins.* Curr Opin Genet Dev, 2008. **18**(5): p. 411-7.

41. Wang, J.C. and W.A. Harris, *The role of combinational coding by homeodomain and bHLH transcription factors in retinal cell fate specification.* Dev Biol, 2005. **285**(1): p. 101-15.

42. Hobert, O., *Regulatory logic of neuronal diversity: terminal selector genes and selector motifs.* Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20067-71.

43. Hobert, O., *Regulation of terminal differentiation programs in the nervous system.* Annu Rev Cell Dev Biol, 2011. **27**: p. 681-96.

44. Simionato, E., et al., *Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics.* BMC Evol Biol, 2007. **7**: p. 33.

45. Kewley, R.J., M.L. Whitelaw, and A. Chapman-Smith, *The mammalian basic helix-loop-helix/PAS family of transcriptional regulators.* Int J Biochem Cell Biol, 2004. **36**(2): p. 189-204.

46. Iso, T., L. Kedes, and Y. Hamamori, *HES and HERP families: multiple effectors of the Notch signaling pathway.* J Cell Physiol, 2003. **194**(3): p. 237-55.

47. Dubois, L. and A. Vincent, *The COE--Collier/Olf1/EBF--transcription factors: structural conservation and diversity of developmental functions.* Mech Dev, 2001. **108**(1-2): p. 3-12.

48. Jackson, D.J., et al., *Developmental expression of COE across the Metazoa supports a conserved role in neuronal cell-type specification and mesodermal development.* Dev Genes Evol, 2010. **220**(7-8): p. 221-34.

49. Eade, K.T., et al., *Developmental transcriptional networks are required to maintain neuronal subtype identity in the mature nervous system.* PLoS Genet, 2012. **8**(2): p. e1002501.

50. Kratsios, P., et al., *Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene.* Nat Neurosci, 2012. **15**(2): p. 205-14.

51. Baer, R., *TAL1, TAL2 and LYL1: a family of basic helix-loop-helix proteins implicated in T cell acute leukaemia.* Semin Cancer Biol, 1993. **4**(6): p. 341-7.

52. Leow, C.C., P. Polakis, and W.Q. Gao, *A role for Hath1, a bHLH transcription factor, in colon adenocarcinoma.* Ann N Y Acad Sci, 2005. **1059**: p. 174-83.

53. Liao, D., *Emerging roles of the EBF family of transcription factors in tumor suppression.* Mol Cancer Res, 2009. **7**(12): p. 1893-901.

54. Newmark, P.A. and A. Sánchez Alvarado, *Not your father's planarian: a classic model enters the era of functional genomics.* Nat Rev Genet, 2002. **3**(3): p. 210-9.

55. Baguñà, J., *The planarian neoblast: the rambling history of its origin and some current black boxes.* Int J Dev Biol, 2012. **56**(1-3): p. 19-37.

56. Baguñà, J., E. Saló, and C. Auladell, *Regeneration and pattern formation in planarians. III. Evidence that neoblasts are totipotent stem cells and the source of blastema cells.* Development, 1989. **107**: p. 77- 86.

57. Sánchez Alvarado, A., *Planarian regeneration: its end is its beginning.* Cell, 2006. **124**(2): p. 241-5.

58. Oviedo, N.J., P.A. Newmark, and A. Sánchez Alvarado, *Allometric scaling and proportion regulation in the freshwater planarian Schmidtea mediterranea.* Dev Dyn, 2003. **226**(2): p. 326-33.

59. Robb, S.M., E. Ross, and A. Sánchez Alvarado, *SmedGD: the Schmidtea mediterranea genome database.* Nucleic Acids Res, 2008. **36**(Database issue): p. D599-606.

60. Labbé, R.M., et al., *A Comparative Transcriptomic Analysis Reveals Conserved Features of Stem Cell Pluripotency in Planarians and Mammals.* Stem Cells, 2012.

61. Önal, P., et al., *Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells.* EMBO J, 2012. **31**(12): p. 2755-69.

62. Resch, A.M., et al., *Transcriptome analysis reveals strain-specific and conserved stemness genes in Schmidtea mediterranea.* PLoS One, 2012. **7**(4): p. e34447.

63. King, R.S. and P.A. Newmark, *In situ hybridization protocol for enhanced detection of gene expression in the planarian Schmidtea mediterranea.* BMC Dev Biol, 2013. **13**: p. 8.

64. Pearson, B.J., et al., *Formaldehyde-based whole-mount in situ hybridization method for planarians.* Dev Dyn, 2009. **238**(2): p. 443-50.

65. Umesono, Y., K. Watanabe, and K. Agata, *A planarian orthopedia homolog is specifically expressed in the branch region of both the mature and regenerating brain.* Dev Growth Differ, 1997. **39**(6): p. 723-7.

66. Rouhana, L., et al., *RNA interference by feeding in vitro-synthesized double-stranded RNA to planarians: methodology and dynamics.* Dev Dyn, 2013. **242**(6): p. 718-30.

67. Orii, H., M. Mochii, and K. Watanabe, *A simple "soaking method" for RNA interference in the planarian Dugesia japonica.* Dev Genes Evol, 2003. **213**(3): p. 138-41.

68. Sánchez Alvarado, A. and P.A. Newmark, *Double-stranded RNA specifically disrupts gene expression during planarian regeneration.* Proc Natl Acad Sci U S A, 1999. **96**(9): p. 5049-54.

69. Newmark, P.A., et al., *Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians.* Proc Natl Acad Sci U S A, 2003. **100 Suppl 1**: p. 11861-5.

70. Wagner, D.E., I.E. Wang, and P.W. Reddien, *Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration.* Science, 2011. **332**(6031): p. 811-6.

71. Elliott, S.A. and A. Sánchez Alvarado, *The history and enduring contributions of planarians to the study of animal regeneration.* WIREs Dev Biol 2012, doi: 10.1002/wdev.82 2012.

72. Reddien, P.W., *Specialized progenitors and regeneration.* Development, 2013. **140**(5): p. 951-7.

73. Lapan, S.W. and P.W. Reddien, *dlx and sp6-9 Control optic cup regeneration in a prototypic eye.* PLoS Genet, 2011. **7**(8): p. e1002226.

74. Lapan, S.W. and P.W. Reddien, *Transcriptome analysis of the planarian eye identifies ovo as a specific regulator of eye regeneration.* Cell Rep, 2012. **2**(2): p. 294-307.

75. Scimone, M.L., et al., *A regulatory program for excretory system regeneration in planarians.* Development, 2011. **138**(20): p. 4387-98.

76. Agata, K., et al., *Structure of the planarian central nervous system (CNS) revealed by neuronal cell markers.* Zoolog Sci, 1998. **15**(3): p. 433-40.

77. Baguñà, J., *The Nervous System in Planarians: Peripheral and Gastrodermal plexuses, Pharynx Innervation, and the relationshp between Central Nervous System Structure and the Acoelomate Organization.* Journal of Morphology, 1978. **15**.

78. Collins, J.J., 3rd, et al., *Genome-wide analyses reveal a role for peptide hormones in planarian germline development.* PLoS Biol, 2010. **8**(10): p. e1000509.

79. Gentile, L., F. Cebrià, and K. Bartscherer, *The planarian flatworm: an in vivo model for stem cell biology and nervous system regeneration.* Dis Model Mech, 2011. **4**(1): p. 12-9.

80. Nishimura, K., et al., *Reconstruction of dopaminergic neural network and locomotion function in planarian regenerates.* Dev Neurobiol, 2007. **67**(8): p. 1059-78.

81. Nishimura, K., et al., *Identification and distribution of tryptophan hydroxylase (TPH)-positive neurons in the planarian Dugesia japonica.* Neurosci Res, 2007. **59**(1): p. 101-6.

82. Nishimura, K., et al., *Characterization of tyramine beta-hydroxylase in planarian Dugesia japonica: cloning and expression.* Neurochem Int, 2008. **53**(6-8): p. 184-92.

83. Nishimura, K., et al., *Analysis of motor function modulated by cholinergic neurons in planarian Dugesia japonica.* Neuroscience, 2010. **168**(1): p. 18-30.

84. Nishimura, K., et al., *Identification of glutamic acid decarboxylase gene and distribution of GABAergic nervous system in the planarian Dugesia japonica.* Neuroscience, 2008. **153**(4): p. 1103-14.

85. Reddien, P.W., *Constitutive gene expression and the specification of tissue identity in adult planarian biology.* Trends Genet, 2011. **27**(7): p. 277-85.

86. Gurley, K.A., J.C. Rink, and A. Sánchez Alvarado, *Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis.* Science, 2008. **319**(5861): p. 323-7.

87. Rink, J.C., et al., *Planarian Hh signaling regulates regeneration polarity and links Hh pathway evolution to cilia.* Science, 2009. **326**(5958): p. 1406-10.

88. Reddien, P.W., et al., *BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration.* Development, 2007. **134**(22): p. 4043-51.

89. Cebrià, F., et al., *FGFR-related gene nou-darake restricts brain tissues to the head region of planarians.* Nature, 2002. **419**(6907): p. 620-4.

90. Mason, I., *Initiation to end point: the multiple roles of fibroblast growth factors in neural development.* Nat Rev Neurosci, 2007. **8**(8): p. 583-96.

91. Bottcher, R.T. and C. Niehrs, *Fibroblast growth factor signaling during early vertebrate development.* Endocr Rev, 2005. **26**(1): p. 63-77.

92. Trueb, B., *Biology of FGFRL1, the fifth fibroblast growth factor receptor.* Cell Mol Life Sci, 2011. **68**(6): p. 951-64.

93. Wenemoser, D., et al., *A molecular wound response program associated with regeneration initiation in planarians.* Genes Dev, 2012. **26**(9): p. 988-1002.

94. Sandmann, T., et al., *The head-regeneration transcriptome of the planarian Schmidtea mediterranea.* Genome Biol, 2011. **12**(8): p. R76.

95. Gavino, M.A., et al., *Tissue absence initiates regeneration through Follistatin-mediated inhibition of Activin signaling.* Elife, 2013. **2**: p. e00247.

96. Roberts-Galbraith, R.H. and P.A. Newmark, *Follistatin antagonizes activin signaling and acts with notum to direct planarian head regeneration.* Proc Natl Acad Sci U S A, 2013. **110**(4): p. 1363-8.

97. Witchley, J.N., et al., *Muscle cells provide instructions for planarian regeneration.* Cell Rep, 2013. **4**(4): p. 633-41.

98. Currie, K.W. and B.J. Pearson, *Transcription factors lhx1/5-1 and pitx are required for the maintenance and regeneration of serotonergic neurons in planarians.* Development, 2013.

99. März, M., F. Seebeck, and K. Bartscherer, *A Pitx transcription factor controls the establishment and maintenance of the serotonergic lineage in planarians.* Development, 2013. **140**(22): p. 4499-509.

100. Tanaka, E.M. and P.W. Reddien, *The cellular basis for animal regeneration.* Dev Cell, 2011. **21**(1): p. 172-85.

101. Crozatier, M. and A. Vincent, *Control of multidendritic neuron differentiation in Drosophila: the role of Collier.* Dev Biol, 2008. **315**(1): p. 232-42.

102. Maher, E.A., et al., *Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities.* Cancer Res, 2006. **66**(23): p. 11502-13.

103. Zhao, L.Y., et al., *An EBF3-mediated transcriptional program that induces cell cycle arrest and apoptosis.* Cancer Res, 2006. **66**(19): p. 9445-52.

CHAPTER 1:
GENOME-WIDE ANALYSIS OF THE BHLH GENE FAMILY IN PLANARIANS IDENTIFIED FACTORS REQUIRED FOR ADULT NEUROGENESIS AND NEURONAL REGENRATION

© 2013. Published by The Company of Biologists Ltd | Development (2013) 140, 4691-4702 doi:10.1242/dev.098616

RESEARCH ARTICLE

STEM CELLS AND REGENERATION

Genome-wide analysis of the bHLH gene family in planarians identifies factors required for adult neurogenesis and neuronal regeneration

Martis W. Cowles[1], David D. R. Brown[2,3], Sean V. Nisperos[1], Brianna N. Stanley[1], Bret J. Pearson[2,3,4] and Ricardo M. Zayas[1,*]

ABSTRACT

In contrast to most well-studied model organisms, planarians have a remarkable ability to completely regenerate a functional nervous system from a pluripotent stem cell population. Thus, planarians provide a powerful model to identify genes required for adult neurogenesis *in vivo*. We analyzed the basic helix-loop-helix (bHLH) family of transcription factors, many of which are crucial for nervous system development and have been implicated in human diseases. However, their potential roles in adult neurogenesis or central nervous system (CNS) function are not well understood. We identified 44 planarian bHLH homologs, determined their patterns of expression in the animal and assessed their functions using RNAi. We found nine bHLHs expressed in stem cells and neurons that are required for CNS regeneration. Our analyses revealed that homologs of *coe*, *hes* (*hesl-3*) and *sim* label progenitors in intact planarians, and following amputation we observed an enrichment of coe^+ and sim^+ progenitors near the wound site. RNAi knockdown of *coe*, *hesl-3* or *sim* led to defects in CNS regeneration, including failure of the cephalic ganglia to properly pattern and a loss of expression of distinct neuronal subtype markers. Together, these data indicate that *coe*, *hesl-3* and *sim* label neural progenitor cells, which serve to generate new neurons in uninjured or regenerating animals. Our study demonstrates that this model will be useful to investigate how stem cells interpret and respond to genetic and environmental cues in the CNS and to examine the role of bHLH transcription factors in adult tissue regeneration.

KEY WORDS: Basic helix-loop-helix, *Coe*, *Single-minded*, *Hes*, Neurogenesis, Lophotrochozoan, Planarians, Regeneration, *Schmidtea mediterranea*, Stem cells, Neoblasts

INTRODUCTION

The discovery that neurogenesis persists in the central nervous system (CNS) of adult animals (Gage, 2002) changed a long-held doctrine that neurons were only produced in the embryo (Ramón y Cajal, 1928; Kempermann, 2011). Although it is now well accepted that adult neurogenesis is a widespread phenomenon across diverse metazoans (Lindsey and Tropepe, 2006; Kempermann, 2012), the ability of most organisms to produce new neurons does not compensate for cells lost after injury or disease. Therefore, to examine how neural precursors could be directed to repair CNS neurons *in vivo*, comparative approaches using animal models of regeneration will help us to gain insights into the basic mechanisms needed to reestablish nervous system function after injury or the onset of neurodegenerative disease (Kempermann, 2011).

Freshwater planarians have emerged as an excellent model to examine the molecular mechanisms underlying stem cell biology and tissue replacement (Elliott and Sánchez Alvarado, 2013; King and Newmark, 2012). Following amputation, planarians are capable of restoring lost body parts from a population of adult pluripotent stem cells called neoblasts (Baguñà, 2012; Elliott and Sánchez Alvarado, 2013). Planarian stem cells share conserved pluripotency determinants with mammalian stem cells (Labbé et al., 2012; Önal et al., 2012; Solana et al., 2012) and serve to replace cells lost during normal physiological cell turnover or after amputation. In contrast to most model organisms currently studied, planarians have the remarkable ability to regenerate their CNS after injury. Thus, these animals provide an excellent opportunity to analyze mechanisms underlying stem cell regulation and CNS regeneration *in vivo*. The planarian CNS consists of bi-lobed cephalic ganglia (brain) that are connected to two longitudinal ventral nerve cords projecting posteriorly along the length of the animal. Distinct neuronal cell types have been described by histochemistry, including unipolar and bipolar neurons as well as neurosecretory cells (Bullock and Horridge, 1965; Lentz, 1968). The generation of genomic resources has led to identification of hundreds of neural markers that have been used to show that the planarian CNS is molecularly complex (Gentile et al., 2011), but little is known about how these animals regenerate their nervous system. On the basis of elegant single cell transplantation studies in lethally irradiated planarians (Wagner et al., 2011), it has been estimated that less than 5% of the planarian stem cells are truly pluripotent (Rink, 2013). Therefore, we and others hypothesize that a fraction of the heterogeneous stem cell pool may be comprised of lineage-committed or specialized progenitor cells (Reddien, 2013). To fully understand the mechanisms underlying how neuronal diversity is maintained or reestablished in planarians it will be essential to define any neural precursor populations that may exist.

Transcription factors from the basic helix-loop-helix (bHLH) gene family play vital regulatory roles throughout the different stages of neurogenesis in embryos, including neural fate commitment, subtype specification, migration and axon guidance (Bertrand et al., 2002; Guillemot, 2007). Genes of the *Drosophila achaete-scute* complex represent the prototypical proneural genes that are important for development of the peripheral and central nervous systems (Jan and Jan, 1994). Proneural genes have been identified

[1]Department of Biology, San Diego State University, San Diego, CA 92182, USA. [2]The Hospital for Sick Children, Program in Developmental and Stem Cell Biology, Toronto, ON M5G 0A4, Canada. [3]University of Toronto, Department of Molecular Genetics, Toronto, ON M5S 1A8, Canada. [4]Ontario Institute for Cancer Research, Toronto, ON M5G 1L7, Canada.

*Author for correspondence (rzayas@mail.sdsu.edu)

Received 4 May 2013; Accepted 12 September 2013

Development (2013) doi:10.1242/dev.098616

in sponges (Richards et al., 2008) and their roles are conserved from cnidarians to vertebrates (Lindsey and Tropepe, 2006; Galliot et al., 2009). However, the precise function of bHLH genes in embryos or adult neural stem cells remains poorly understood (Kintner, 2002). Here we have performed a genome-wide analysis of bHLH family genes to identify factors essential for CNS tissue renewal in adult planarians. Our screen identified nine genes that are expressed in both the stem cells and neurons and are required for normal CNS regeneration, including homologs of *collier/olfactory-1/early B-cell factor* (*coe*), *hairy/enhancer of split* (*hes-like-3*) and *single-minded* (*sim*). To characterize and follow the fate of *coe*, *hesl-3* and *sim* stem cells, we performed bromodeoxyuridine (BrdU) pulse-chase experiments and found that *coe* and *sim* are expressed in proliferating cells adjacent to the CNS, which can be traced to the brain or ventral nerve cords over the course of 2 days. During regeneration, we observed an enrichment of *coe* and *sim* progenitors near the wound site. Furthermore, RNAi knockdown of *coe*, *hesl-3* or *sim* led to defects in CNS regeneration, including failure of the cephalic ganglia to reconnect or pattern, and a loss of expression of genes unique to distinct neuronal subtypes. Together, these data suggest that *coe*, *hesl-3* and *sim* are expressed in neural progenitor cells and that these bHLH genes are required to generate new neurons in uninjured and regenerating animals. Our study demonstrates that this model will be useful to investigate how stem cells interpret and respond to genetic and environmental cues in the CNS and to examine the role of bHLH transcription factors in adult tissue regeneration.

RESULTS

Identification of bHLH family genes in planarians

We identified 44 sequences in the planarian *Schmidtea mediterranea* predicted to encode a bHLH motif and named them according to their homology, as described in the Materials and methods (supplementary material Fig. S1, Table S1; for brevity we have omitted the *Smed* prefix from the gene names). Recent transcriptional profiles generated from sorted cell populations indicate that most bHLH homologs are expressed in the planarian stem cells and their postmitotic progeny (Labbé et al., 2012; Önal et al., 2012; Resch et al., 2012). To investigate cell- and tissue-specific patterns of bHLH gene expression, we performed whole mount *in situ* hybridization (WISH). We confirmed the presence of transcripts in stem cell or their progeny by testing for the loss of gene expression throughout the parenchyma (mesenchyme) 6 days following exposure to γ-irradiation, a treatment that depletes all stem cells and postmitotic progeny (Reddien et al., 2005b; Eisenhoffer et al., 2008). Consistent with the transcriptome data, we found that 35/43 bHLH genes tested are expressed in stem cells and their progeny (supplementary material Table S1). As expected, we also observed bHLH expression in differentiated tissues, including the CNS (21 genes), epidermis (three genes), pharynx (14 genes) or intestine (nine genes) (supplementary material Fig. S2, Table S1). One gene, *neuroD-2*, was not detected by WISH.

A subset of bHLH genes is expressed in neurons

Of the 12 CNS- and stem-cell-expressed bHLH genes, we selected *atoh*, *coe*, *fer3l-1*, *hesl-3* and *sim* for detailed expression analyses because their transcripts were detected in discrete cell populations (supplementary material Fig. S2). To confirm the pattern of mRNA expression in the CNS and visualize the distribution of these cell populations in reference to the brain, we performed double-fluorescent *in situ* hybridization experiments (dFISH) using the pan-neural marker *pc2* (Collins et al., 2010). *atoh*, *coe*, *fer3l-1*, *hesl-3* and *sim* were expressed in discrete neural populations throughout the brain and in regenerating tissues (Fig. 1; supplementary material Fig. S3). In addition, *fer3l-1* (supplementary material Fig. S3D), *hesl-3* (Fig. 1D) and *sim* (Fig. 1G) were detected in cells distributed throughout the mesenchyme. We then investigated the expression of *coe*, *hesl-3* and *sim* in specific neuronal subtypes by performing dFISH with markers of cholinergic, GABAergic, octopaminergic, dopaminergic and serotonergic neurons (Umesono et al., 2011). *coe*, *hesl-3* and *sim* were each co-expressed in cholinergic neurons (Fig. 1J-L). We also detected transcripts for *coe* in GABAergic, octopaminergic, dopaminergic and serotonergic neurons and *sim* in octopaminergic and dopaminergic neurons (Fig. 1J,L). These results demonstrate that a subset of bHLH genes is expressed in molecularly distinct differentiated neurons.

coe, *hesl-3* and *sim* label cycling cells in close proximity to the CNS

Following amputation, stem cells proliferate beneath the wound site (post-blastema) and give rise to the regeneration blastema, the structure where postmitotic cells differentiate to form the missing tissues. *atoh*, *coe*, *fer3l-1*, *hesl-3* and *sim* were expressed in the newly regenerated tissues, but it was also noted that these mRNAs were present in cells located in the post-blastema (Fig. 1; supplementary material Fig. S3). Therefore, we examined whether *atoh*, *coe*, *fer3l-1*, *hesl-3* and *sim* could be detected in mitotic cells. We found that, with the exception of *atoh*, their transcripts could be visualized in a subset of anti-phosphohistone-H3 cells (Fig. 1C,F,I; supplementary material Fig. S3C,F), which we also observed in uninjured planarians (data not shown).

To distinguish between gene expression in stem cells/progeny or differentiated neurons, we examined the expression of *atoh*, *coe*, *fer3l-1*, *hesl-3* and *sim* following 6 days of γ-irradiation treatment. Compared with control animals, *atoh* expression was reduced throughout the mesenchyme, but we were unable to detect changes in expression in the head or the pre-pharyngeal area even when we used FISH (supplementary material Fig. S3G-J′). In contrast to *atoh*, we observed a reduction of *coe* cells near the brain and between the cephalic ganglia and ventral nerve cords (VNCs) (supplementary material Fig. S3K,K′). *fer3l-1* expression was broadly reduced in the mesenchyme, except for a few cells located on the dorsal surface of the cephalic ganglia and distributed throughout the mesenchyme (supplementary material Fig. S3L,L′). *hesl-3* and *sim* staining were also reduced in the mesenchyme and near the cephalic ganglia following γ-irradiation (supplementary material Fig. S3M-N′). To validate further that *coe*, *hesl-3* and *sim* were expressed in a subset of stem cells, we co-stained these genes with the stem cell marker *smedwi-1* (Reddien et al., 2005b; Eisenhoffer et al., 2008) (Fig. 1M-O). Taken together, our analyses confirmed that bHLH genes were expressed in subsets of stem cells and postmitotic progeny. In addition, we noted that cell populations that expressed *coe*, *hesl-3* and *sim* near the CNS were γ-irradiation-sensitive, further supporting potential roles of these genes in differentiation of neural precursor-like cells.

coe and *sim* are expressed in differentiating neurons

Stemming from our observations that *coe*, *hesl-3* and *sim* were expressed in stem cells and in diverse neural subtypes, we reasoned that these genes label lineage-committed progenitors and differentiating neurons. To address this possibility, we sought to label stem cells expressing *coe*, *hesl-3* or *sim* and map their relative positions in the animal over time. Although the most commonly used tools to trace cell lineages (e.g. genetic marks) (Kretzschmar

Development (2013) doi:10.1242/dev.098616

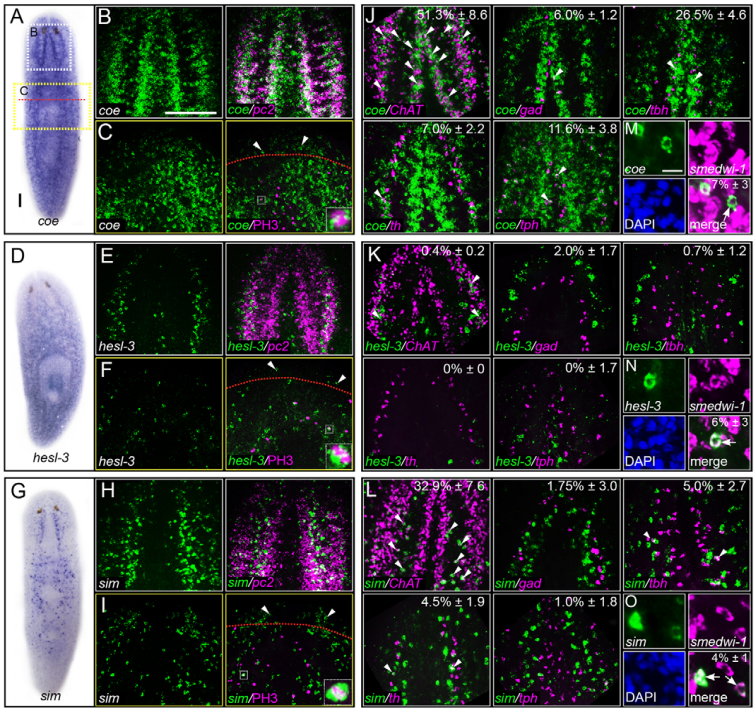


Fig. 1. *coe*, *hesl-3* and *sim* are expressed in stem cells and neurons. (A) Expression pattern of *coe*. Dashed boxes indicate zoom area of the brain or regeneration blastema shown in B and C, respectively. Dashed red line indicates site of amputation. (B) FISH to *coe* (green) and *pc2* (magenta). (C) Animals processed for FISH to *coe* and counterstained with anti-phosphohistone-H3 (ph3) to label mitotic cells in 3-day regenerates. Arrowheads denote *coe*$^+$ cells within the blastema. D-F and G-I show similar analyses for *hesl-3* and *sim*, respectively. White dashed boxes in C, F and I highlight *bHLH*/ph3-positive cells shown at high magnification within merged image insets. (J-L) Animals were processed for dFISH to *coe*, *hes-3* or *sim* and markers of cholinergic (*ChAT*, *choline acetyltransferase*), GABAergic (*gad*, *glutamic acid decarboxylase*), octopaminergic (*tbh*, *tyramine β-hydroxylase*), dopaminergic (*th*, *tyrosine hydroxylase*) and serotonergic (*tph*, *tryptophan hydroxylase*) neurons. White arrowheads point to co-labeled cells. (M-O) Representative photos of cells imaged at high magnification from animals that were processed for dFISH to *coe*, *hesl-3* or *sim* and *smedwi-1* and counterstained with DAPI. The percentage of the total number of subtype-specific neurons or *smedwi-1*$^+$ cells that co-expressed *coe*, *hesl-3* or *sim* are shown in J-O. Scale bars: 100 μm in A,B; 10 μm in M.

and Watt, 2012) are not yet available in planarians, it is possible to label S-phase neoblasts with the thymidine analog BrdU and then determine the location of label-retaining cells (Newmark and Sánchez Alvarado, 2000; Eisenhoffer et al., 2008). This approach has been used to study planarian eye (Lapan and Reddien, 2011) and intestinal (Forsthoefel et al., 2011) cell differentiation. Previous studies have estimated that the length of S/G2/M in planarians is between 12 and 16 hours (Newmark and Sánchez Alvarado, 2000). We predicted that at later time points BrdU$^-$ cells in the head marked by any of these bHLH genes would represent differentiating stem cells. Accordingly, we noted that, as expected, *coe*$^+$ and *sim*$^+$ cells were *smedwi-1*$^-$ in the anterior region of the brain (supplementary material Fig. S4). We pulsed animals with BrdU and inspected animals for BrdU$^-$ and *coe*$^+$, *hesl-3*$^+$ or *sim*$^+$ cells in the head, pre-pharyngeal and post-pharyngeal areas after a 4-, 24- or 48-hour chase period (Fig. 2). We found that most double-labeled cells were located in the head and pre-pharyngeal regions, and we focused our analyses on these areas.

Following a 4-hour chase, BrdU$^-$/*coe*$^+$ cells were located throughout the mesenchyme of the head, pre- and post-pharyngeal regions, with some cells in close proximity to the brain and VNCs (Fig. 2A,A′). Over time, BrdU$^-$/*coe*$^+$ progeny were detected in more anterior and lateral regions of the cephalic ganglia or directly adjacent to the VNCs (Fig. 2B,C). After 4 hours, BrdU$^-$/*sim*$^+$ cells were only detected in the pre- and post-pharyngeal areas (Fig. 2D,D′). Similar to *coe*$^+$ progenitors, BrdU$^-$/*sim*$^+$ cell populations could be traced to the CNS over time; by 24 hours, progenitors were observed near the posterior end of the brain, and by 48 hours, we observed BrdU$^-$/*sim*$^+$ cells at the most anterior tip of the cephalic ganglia (Fig. 2E,F). Cells progressing through S-phase that expressed *hesl-3* were observed throughout the animal (Fig. 2G,G′). In contrast to *coe*$^+$ and *sim*$^+$ progenitors, the distribution of BrdU$^-$/*hesl-3*$^+$ cells remained relatively consistent over the course of 48 hours (Fig. 2H,I). To determine whether *coe*, *hesl-3* and *sim* label unique cell populations, we performed dFISH to either *coe* or *hesl-3* with *sim*. We found that *coe*, *hesl-3* and *sim* were not co-expressed in the head; however, although rare, we did observe *coe*$^+$/*sim*$^+$ cells in the pre-pharyngeal area (supplementary material Fig. S5A-E). Consistent with the expression of these genes in *smedwi-1*$^+$ cells, these data suggest that *coe*, *hesl-3* and *sim* are expressed in lineage-committed progenitors. Moreover, we observed *coe*$^+$ and *sim*$^+$ progenitors near the CNS over time, suggesting that these genes are expressed in differentiating neurons.

In addition to determining the location of BrdU$^-$ cells marked by bHLHs, we quantified the number of double-positive cells over time (Fig. 2J-M). After 4 hours, BrdU$^-$/*coe*$^+$ cells comprised ~3.7% of BrdU$^-$ cells in the head or pre-pharyngeal area (Fig. 2K). Interestingly, by 48 hours, we observed an increase in the proportion of BrdU$^-$/*coe*$^+$ in both the head (9.5%) and pre-pharyngeal region (10%; Fig. 2K). Similar to *coe*$^+$ progenitors, the proportion of BrdU$^-$/*sim*$^+$ cells also increased over time (Fig. 2L). By contrast, the proportion of BrdU$^-$/*hesl-3*$^+$ cells remained consistent throughout the head and pre-pharyngeal area over the course of 48 hours (Fig. 2M). Intriguingly, we also noted that after 48 hours we still observed

RESEARCH ARTICLE Development (2013) doi:10.1242/dev.098616

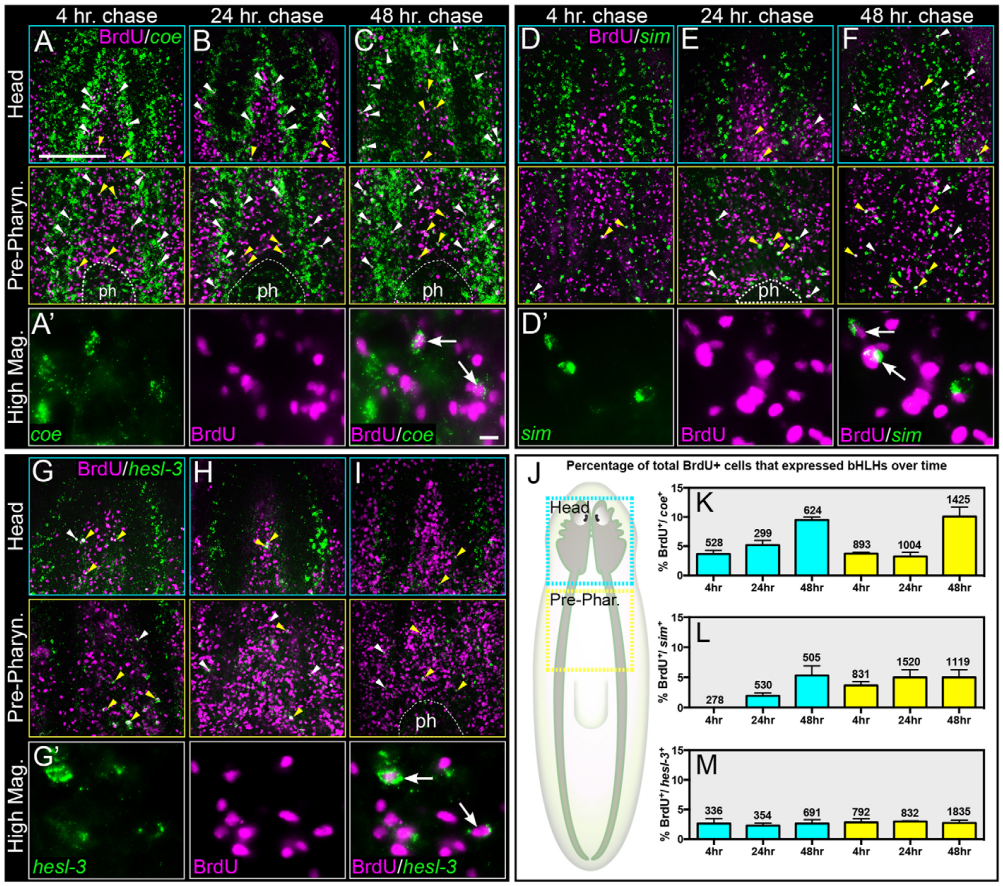


Fig. 2. Birthdating of *coe*⁺, *hesl-3*⁺ and *sim*⁺ progenitors. (A-C) Intact animals were pulsed with BrdU, followed by a 4-, 24- or 48-hour chase and stained for BrdU and *coe*. Arrowheads point to BrdU⁺/*coe*⁺ cells near the CNS (white) or the mesenchyme (yellow). (A') High magnification of animals in A; the arrows indicate BrdU⁺/*coe*⁻ cells. Similar analyses for *sim* and *hesl-3* are shown in D-F and G-I, respectively. (J) Cartoon depicting the regions counted in K-M. (K-M) Percentage of the total number of BrdU⁺ cells that are *coe*⁺, *hesl-3*⁺ or *sim*⁺ following a 4, 24- or 48-hour pulse; numbers correspond to the total number of BrdU⁺ cells counted. Blue and yellow bars indicate cell counts from the head and pre-pharyngeal regions, respectively. Scale bars: 100 μm in A; 10 μm in A'. ph, pharynx.

BrdU cells expressing *coe*, *hesl-3* or *sim* in the pre-pharyngeal area (yellow arrowheads in Fig. 2C,F,I), the same location where most *coe*, *hesl-3* and *sim* cycling cells were first detected following a 4-hour chase period (Fig. 2A,D,G), which suggests that new progenitors were generated or differentiating cells turned on expression of these genes. The increase in the proportion of BrdU cells that expressed *coe* or *sim* near the brain and VNCs, combined with the observation that these genes were expressed in diverse neuronal subtypes (Fig. 1J,L), demonstrate that at least some *coe* and *sim* progenitors differentiate into neurons. The fact that we did not observe changes in the proportion of *hesl-3* suggests a potential role of this gene in progenitor cell maintenance (Ishibashi et al., 1995; Kageyama et al., 2008).

In addition, we investigated whether these *coe* and *sim* progenitor populations contribute to the generation of the regeneration blastema following amputation. By 2 and 3 days of regeneration, BrdU /*coe* and BrdU /*sim* cells were detected in the post-blastema of animals regenerating new heads, with most BrdU /*coe* progenitors located directly adjacent to the VNCs and many BrdU /*sim* progenitors located between the VNCs (Fig. 3A-D). In *S. mediterranea*, head regeneration is completed within 7 days following amputation, and we found that the distribution of BrdU /*coe* or BrdU /*sim* cells observed in uninjured animals was reestablished by this time point (Fig. 2A,D, Fig. 3G-J). Taken together, these data are consistent with the hypothesis that the blastema is generated from a heterogeneous population of lineage-committed cells (Reddien, 2013).

Development (2013) doi:10.1242/dev.098616

FGF signaling modulates expression of *sim*⁺ and *coe*⁺ neurons and progenitors

Gene silencing of *nou-darake* (*ndk*), an FGF receptor-like gene, disrupts anterior patterning and leads to ectopic expression of brain-specific neurons outside of the head domain (Cebrià et al., 2002). We hypothesized that *ndk* silencing would cause an increase in the number of *coe* and *sim* progenitor cells. As we expected, ectopic expression of brain-specific markers (*npp-4* and *gpas*) extended from the posterior end of the brain to the anterior boundary of the pharynx following 14 days of *ndk* RNAi treatment (Fig. 4A,B). We then examined *sim* and *coe* mRNA expression in the pre-pharyngeal area of control and *ndk*(*RNAi*) animals after 14 days of RNAi; we also exposed animals from each RNAi group to a lethal dose of γ-irradiation to distinguish stem cell or progeny expression from differentiated neurons (Fig. 4C-E). In control animals, we consistently observed *coe* and *sim* cells in the pre-pharyngeal region and found that most irradiation-sensitive cells were located between the VNCs. Following *ndk* RNAi, we found there were nearly twice the number of *coe* and ~40% more *sim* cells between the VNCs (Fig. 4F-H). Irradiated *ndk*(*RNAi*) animals confirmed that the majority of the additional cells between the VNCs are stem cells or early progeny. We conclude that *coe* and *sim* progenitor generation is regulated by signals downstream of FGF signaling.

Analysis of bHLH gene function in CNS regeneration

We took advantage of the experimental ease to inhibit gene function in planarians by RNAi to investigate the role of all 44 bHLH genes in planarian tissue regeneration. To screen for defects in CNS architecture and stem cell regulation, animals were stained with the pan-neural marker anti-SYNAPSIN and the mitotic cell marker anti-phosphohistone-H3 following dsRNA treatment (supplementary material Fig. S6A). We observed a wide range of regeneration phenotypes following RNAi knockdown of 11 bHLH genes, including lesions (*mitfl-1*), defects in CNS morphology (*arnt*, *arh*, *atoh8-1*, *coe*, *da*, *max*, *mxi-1* and *sim*) and patterning (*hesl-3* and *myoD*) (Table 1; supplementary material Fig. S6B). We did not observe obvious defects in cell division following knockdown of any bHLH gene (data not shown). *mitfl-1* was primarily detected in differentiated intestinal cells (supplementary material Fig. S2), and gene knockdown caused severe regeneration abnormalities and dorsal lesions that resulted in death, a phenotype reminiscent of defects observed after the loss of intestinal integrity (Forsthoefel et al., 2012). Consistent with previous reports, gene silencing of *tfc15* (Wagner et al., 2011) resulted in no discernible phenotype, whereas *myoD* RNAi caused regeneration blastema patterning defects (Reddien et al., 2005a).

Proneural bHLHs form heterodimers with ubiquitously expressed E proteins (such as *daughterless*, *da*) to bind DNA and function to commit progenitors to the neural fate during development (Bertrand et al., 2002). We did not observe regeneration defects following RNAi of candidate proneural gene homologs, such as *acheate-scute* or *neuroD*, which we validated by real-time quantitative PCR and found that our treatment strongly silenced each gene that we tested (supplementary material Fig. S6C). It has been shown that developmental defects are exacerbated in *Drosophila* when proneural factors are co-silenced with *da* (Goulding et al., 2000; Huang et al., 2000). Therefore, we performed double-RNAi experiments of *da* and *ascl-1*, *ascl-2*, *atoh*, *neuroD-1* or *neuroD-2*. We also co-silenced predicted bHLH paralogs to test if genes may be functionally redundant (*ascl-1;ascl-2* and *hesl-1;hesl-2* RNAi). Due to the possibility that these proteins perdure, we also conducted long-term knockdown experiments (6 weeks of RNAi treatment; *ascl-1*, *ascl-2*, *hesl-1*, *hesl-2*, *sim*). Neither combinatorial nor long-term RNAi experiments revealed any additional regeneration defects. However, extended knockdown experiments increased the penetrance of *sim*(*RNAi*) animals from ~50% to 100% (n=10/10; data not shown).

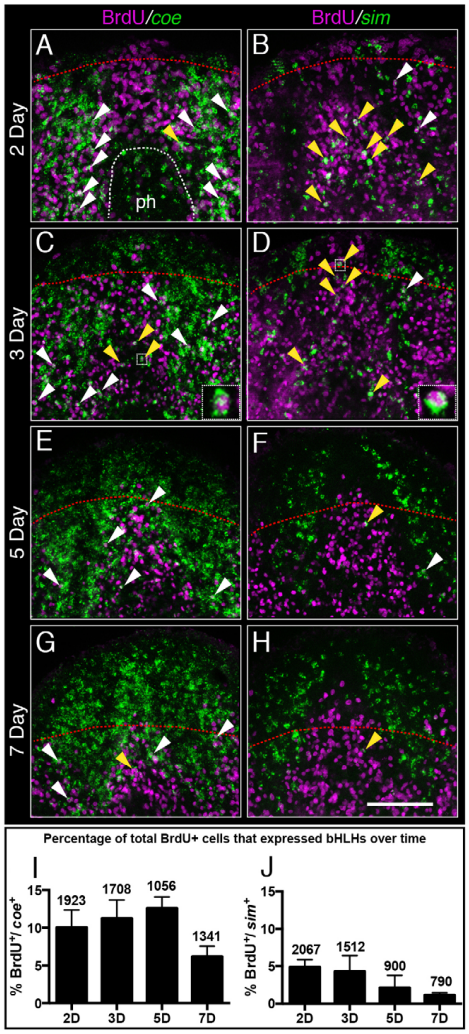


Fig. 3. Analysis of *coe*⁺ and *sim*⁺ cycling cells during regeneration. (A-H) 2, 3, 5 and 7 day regenerates were soaked in BrdU for 1 hour, chased for 4 hours, and co-labeled for *coe* or *sim* and BrdU. Red line denotes amputation site. Yellow and white and arrowheads indicate *coe*⁺/BrdU⁺ or *sim*⁺/BrdU⁺ cells that are in the mesenchyme or in close proximity to the CNS, respectively. (I,J) Percentages of total BrdU⁺ cells that were *coe*⁺ or *sim*⁺ at each time point. Numbers above each bar correspond to the total number of BrdU⁺ cells counted. ph, pharynx. Scale bar: 100 μm.

Development (2013) doi:10.1242/dev.098616

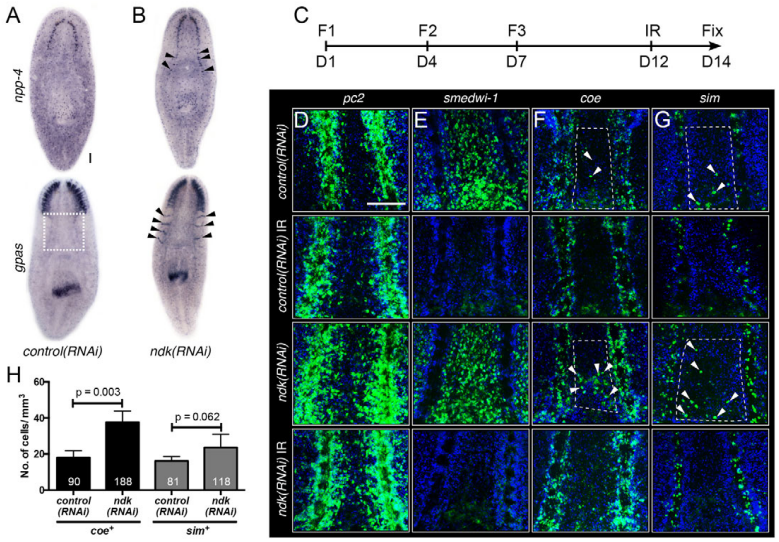


Fig. 4. Induction of ectopic neurogenesis causes an increase in the number of *coe*⁺ and *sim*⁺ neurons and progenitors. (A,B) *control*(*RNAi*) and *ndk*(*RNAi*) animals were processed for WISH to *npp-4* and *gpas* (*n*=10). Arrows point to ectopic *npp-4*⁺ or *gpas*⁺ cells. White dashed box shows region imaged in D-G. (C) Schematic showing RNAi feeding (F, feeding; D, days) and γ-irradiation (IR) schedule for animals shown in D-G. (D-G) RNAi animals were processed for FISH to *pc2*, *smedwi-1*, *coe* or *sim* and counterstained with DAPI (blue) (*n*=15). (H) Quantification of *coe*⁺ and *sim*⁺ progenitors within white boxed areas in F-G; total number of cells counted are indicated within each bar. Scale bars: 100 μm.

Furthermore, we capitalized on the robust *ndk* RNAi phenotype to test the hypothesis that bHLHs required for neurogenesis would suppress ectopic nervous system expansion. We performed combinatorial RNAi experiments using *ndk*, which has been successfully used to investigate the role of other genes in planarian body patterning and CNS regeneration (Felix and Aboobaker, 2010; Iglesias et al., 2011; Blassberg et al., 2013), and screened 15 genes (*ascl-1*, *ascl-2*, *atoh*, *atoh8-1*, *coe*, *da*, *hesl-1*, *-2*, *-3*, *hlh*, *id4*, *neuroD-1*, *neuroD-2*, *sim* and *usf*) by inspecting *bHLH;ndk*(*RNAi*) animals for changes in *gpas* and *npp-4* expression. Induction of *gpas* expression posterior to the cephalic ganglia was not suppressed by inhibiting any of the bHLH genes together with *ndk*. However, *ascl-1;ndk*(*RNAi*) animals exhibited a 60% decrease in ectopic *npp-4* cells, whereas *ndk;hesl-3*(*RNAi*) and *ndk;neuroD-1*(*RNAi*) animals exhibited a 40% decrease of ectopic *npp-4* cells when compared with *gfp;ndk*(*RNAi*) animals (supplementary material Fig. S7A-G). These data suggest that *ascl-1* and *neuroD-1* may function in neural specification, but do not cause gross morphological CNS regeneration defects following gene knockdown.

coe, *hesl-3* and *sim* are required for neuronal regeneration or maintenance

Our RNAi screen revealed that *coe*, *hesl-3* or *sim* led to clear defects in brain regeneration (Fig. 5A-D). *coe*(*RNAi*) regenerates displayed photoreceptors with abnormal morphology and smaller cephalic ganglia that failed to form anterior commissures (*n*=94/114; Fig. 5B). In *hesl-3*(*RNAi*) regenerates, the CNS was abnormally patterned, with animals regenerating a single or an ectopic eyespot and brains with abnormal morphology (*n*=20/35; Fig. 5C). *sim*(*RNAi*) animals regenerated photoreceptors with reduced pigmentation (*n*=17/65) and displayed reduced density of the brain neuropil (*n*=30/65; Fig. 5D). Gene knockdown of *arnt* or *sim* resulted in similar regeneration defects and these genes have been shown to interact with each other (Probst et al., 1997). Hence, we also tested the effect of co-silencing *sim* and *arnt*; however, *sim;arnt*(*RNAi*) did not increase the severity of the phenotype above single RNAi treatments (data not shown). The *coe*, *hesl-3* and *sim* knockdown phenotypes, together with the expression of these genes in progenitors and neurons (Figs 1-3), led us to further investigate their potential roles in neuronal regeneration and homeostasis.

Next, we examined the specific roles of *coe*, *hesl-3* and *sim* in nervous system differentiation by evaluating the effect of gene knockdown on the expression of neuronal subtype-specific genes (Fig. 5E-H). We selected the neural marker *ChAT*, which is broadly expressed in the CNS and was co-detected with *coe* , *hesl-3* and *sim* cells (Fig. 1J-L), and *cpp-1*, *npp-4* and *npy-2*, which are strongly expressed in neuropeptidergic neurons in the brain (Collins et al., 2010). Using *ChAT* staining we measured the brain area of 7-day regenerates and found that *coe* and *sim* RNAi animals regenerated smaller brains (Fig. 5I). In addition, we observed a significant reduction of *cpp-1* cells in *coe*(*RNAi*) animals (Fig. 5F,J) and of *npp-4* and *npy-2* cells in *coe*(*RNAi*) and *sim*(*RNAi*) animals (Fig. 5F,H,K,L). Furthermore, we observed *ChAT* , *npy-2* and *npp-4* cells in aberrant locations following *coe* RNAi (Fig. 5F). Although the brain area difference in *hesl-3*(*RNAi*) regenerates was not statistically significant, we did detect fewer *cpp-1* , *npp-4* and *npy-2* neurons (Fig. 5G,I-L). Due to the fact that we observed CNS patterning-like defects in *coe*(*RNAi*) and *hesl-3*(*RNAi*) animals, we tested whether these abnormalities were caused by defects in the stem cells (*smedwi-1*), progeny (*NB.32.1g*, early progeny; *agat-1*, late progeny) or midline signals (*slit* expression) (Cebrià et al., 2007), but we did not find obvious changes in the expression of these markers after *coe* or *hesl-3* RNAi (supplementary material Fig. S8A-C). These data demonstrate that *coe*, *hesl-3* and *sim* are required for expression of neuronal-specific genes and may be necessary for the replacement of neurons following injury. In combination with our expression analyses, these data suggest that *coe*, *hesl-3* and *sim* are expressed in a subset of stem cells committed to neural fates and their function is crucial for neural progenitor maintenance or differentiation.

Development (2013) doi:10.1242/dev.098616

Table 1. Summary of phenotypes observed following RNAi of bHLH homologs in *S. mediterranea*

Gene name	Gene symbol	Phenotype	Developmental role
aryl hydrocarbon receptor	*arh*	Reduced brain neuropil density (14/20)	B-cell and nervous system differentiation
aryl hydrocarbon receptor nuclear translocator	*arnt*	Delayed regeneration and reduced brain neuropil density (12/40)	Differentiation of multiple cell types
atonal homolog 8-1	*atoh8-1*	Delayed regeneration and smaller cg (19/35)	Nervous system differentiation
collier/olfactory-1/early B-cell factor	*coe*	Abnormal pr morphology, flattened morphology, and failure of cg to reconnect (94/114)	B-cell, muscle and nervous system differentiation
daughterless	*da*	Ruffled body margin edges and reduced cg neuropil (70/70)	Neurogenesis, oogenesis and sex determination
hairy and enhancer of split like-3	*hesl-3*	Abnormal pr morphology; single or third pr (20/35)	Negative regulation of Notch signaling
max-interactor-1	*mxi-1*	Abnormal pr morphology (13/40), expanded and disorganized cg (20/40)	Negative regulation of cell proliferation
microphthalmia-associated transcription factor like-1	*mitfl-1*	Lysis (8/40), smaller and disorganized cg (12/40)	Osteoclast differentiation
myc associated factor X	*max*	Lighter pr and smaller cg (15/20)	Negative regulation of gene expression
myogenic differentiation	*myoD*	Failure to regenerate (8/30), abnormal pr morphology; cyclops or bowtie-shaped pr pair (14/30)	Mesoderm specification
single-minded	*sim*	Reduced pr pigmentation (17/65) and cg neuropil density (30/65)	Axon guidance, nervous system differentiation

The number of animals showing the phenotype(s) among the total number examined is indicated in parentheses.
cg, cephalic ganglia; pr, photoreceptor.

Planarians continuously replace cells during normal tissue homeostasis. Therefore, we also assessed the roles of *coe*, *hesl-3* and *sim* in nervous system maintenance by performing extended RNAi treatments (6 weeks) on intact planarians. Knockdown of *hesl-3* and *sim* resulted in no external phenotype or alterations in CNS architecture (data not shown). By contrast, long-term *coe* RNAi resulted in a strong behavioral phenotype in which animals exhibited impaired negative phototaxis and a flattened and stretched body shape with ruffling along the body margin (Fig. 5M; supplementary material Movies 1, 2). Analysis of *ChAT* and *pc2* neurons in *coe*(*RNAi*) animals showed that the CNS appeared largely intact except for the absence of *ChAT* and *pc2* neurons located at the anterior brain commissure (Fig. 5N). This phenotype was reminiscent of the defect observed in *coe* knockdown regenerates, in which the brain fails to reconnect (Fig. 5B). Because *coe*(*RNAi*) regenerates showed a dramatic reduction of *cpp-1* brain neurons, we examined whether this cell population was also affected in uninjured *coe*(*RNAi*) animals. Strikingly, we observed an 80% reduction in the number of *cpp-1* cells in *coe*(*RNAi*) planarians (Fig. 5O,P). Furthermore, when we performed dFISH to *coe* and *cpp-1*, we found that a majority of *cpp-1* cells also expressed *coe* (81±1.3%; Fig. 5Q). Taken together, our data indicate that *coe* is required for normal function and maintenance of neural tissues and strongly suggest that *cpp-1* may be downstream of *coe*.

DISCUSSION

Although it has been demonstrated that planarians possess pluripotent stem cells (Baguñà et al., 1989; Wagner et al., 2011; Guedelhoefer and Sánchez Alvarado, 2012), several studies support the hypothesis that the stem cell population is heterogeneous (Elliott and Sánchez Alvarado, 2013; Reddien, 2013; Rink, 2013). Analyses of the planarian photoreceptor, excretory and serotonergic cells have shown that tissue-specific transcription factors are detected in the stem cells in intact (Lapan and Reddien, 2012) and regenerating tissues (Lapan and Reddien, 2011; Scimone et al., 2011; Currie and Pearson, 2013); these studies have identified the first sets of precursor cells in planarians outside of the germ cells (Newmark et al., 2008) and have generated a working model in which planarians possess diverse lineage-committed progenitors that contribute to the maintenance and regeneration of tissues (Reddien, 2013; Rink, 2013). In contrast to the well-defined excretory system and photoreceptors, the nervous system represents a formidable challenge. At the molecular level, there are potentially dozens of neuronal subtypes (Cebrià, 2007; Collins et al., 2010; Gentile et al., 2011; Umesono et al., 2011), and it is largely unknown whether the generation of neural diversity is solely dependent on the pluripotent stem cells or lineage-restricted progenitors. In our study, we investigated this question by analyzing the bHLH gene family. By combining *in situ* hybridization analyses and RNAi studies, we identified nine bHLH genes expressed in specific neural and stem cell subpopulations that were required for regeneration (Fig. 6A), which strongly suggested that these phenotypes could be due to abnormal neural differentiation and/or function.

Identification of neuronal progenitor cells in planarians

Owing to the mRNA expression in stem cells and neurons, we focused our analyses on *coe*, *hesl-3* and *sim*, which are known to serve major roles in neurogenesis in both vertebrate and invertebrate organisms (Dubois and Vincent, 2001; Kewley et al., 2004; Kageyama et al., 2008). As we expected, using BrdU, we observed *coe*, *hesl-3* and *sim* expression in cycling stem cells located in the mesenchyme of intact animals. Over the course of 48 hours, we observed an increase in the proportion of BrdU cells that expressed *coe* and *sim* and detected many of these cells in the cephalic ganglia. We hypothesize that the observed increase in the proportion of BrdU /*coe* and BrdU /*sim* cells over time is from both progenitors that maintain expression of *coe* or *sim* as they divide and begin to differentiate and additional cells that turn on expression of these genes during differentiation. Additionally, the increase in the proportion of double-labeled cells could also be accounted for by a contribution of newly generated progenitor cells. Together with the observation that the number of *coe* and *sim* cells increases following induction of ectopic neurogenesis and the requirement of *coe* and *sim* during CNS regeneration (RNAi studies discussed below), our data suggest that a subset of *coe*- and *sim*-expressing cells represent multipotent neural progenitors (Fig. 6B). We propose that these *coe* and *sim* progenitors migrate and terminally differentiate in the CNS. By contrast, we did not observe an increase

Development (2013) doi:10.1242/dev.098616

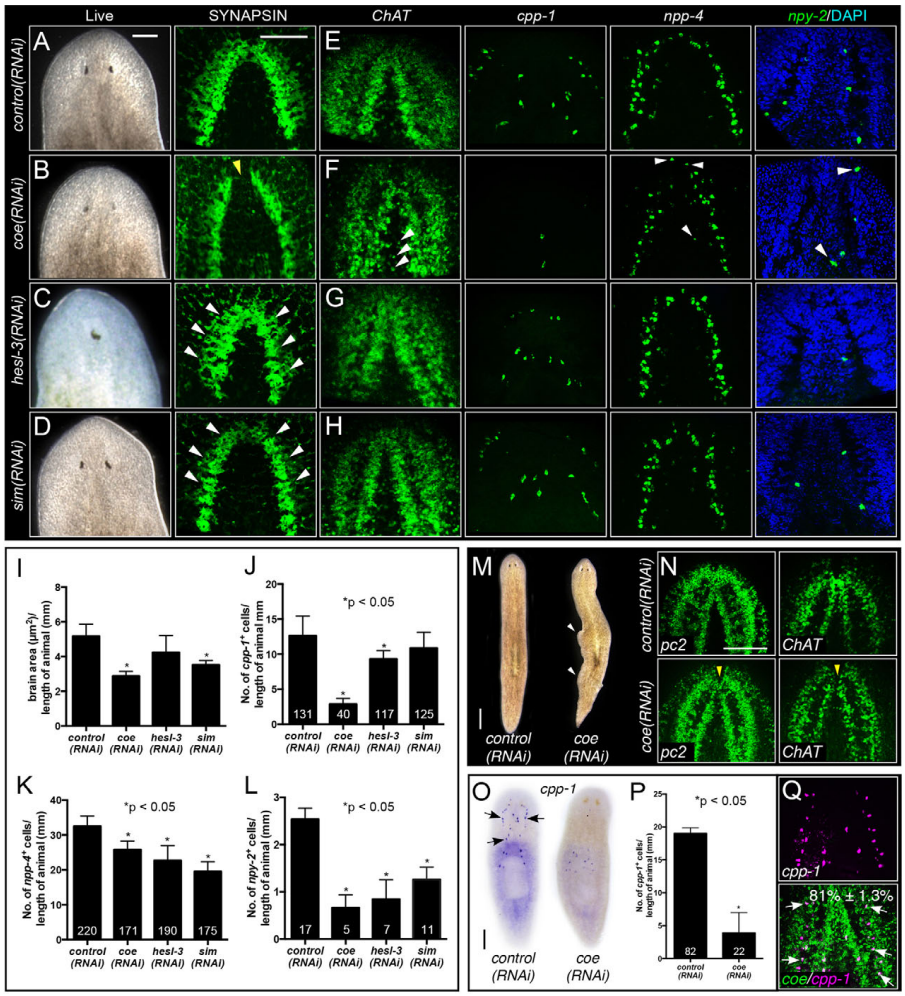


Fig. 5. *coe*, *hesl-3* and *sim* are required for CNS regeneration. (A-D) Images of live or immunostained RNAi-treated animals after 10 days of regeneration. The yellow arrowhead in B marks a commissure defect, and white arrowheads in C and D mark the abnormal brain morphology and a dramatic reduction of neuropil density observed in *hesl-3* and *sim* RNAi planarians, respectively. (E-H) Seven days following amputation, RNAi animals were processed for FISH to *ChAT*, *cpp-1*, *npp-4* or *npy-2* (n=20). Arrowheads in F denote aberrant location of *npp-4*$^+$ and *npy-2*$^+$ neurons (the latter were counterstained with DAPI to visualize the brain). (I-L) Quantification of neurons shown in E-H; the total number of cells counted are indicated within each bar. (M-O) After 24 days of RNAi treatment uninjured control and *coe*(*RNAi*) animals were imaged live (M) or processed for FISH to *pc2* or *ChAT* (N; n=10) or WISH to *cpp-1* (O; n=15). (P) Quantification of *cpp-1*$^+$ cells in O. (Q) FISH to *cpp-1* and *coe*, and quantification of *cpp-1*$^+$ cells that also expressed *coe*. Scale bars: 100 μm.

in the proportion of BrdU /*hesl-3* cells near the brain; it is possible that *hesl-3* expression is downregulated during cell-fate specification and that this gene may be regulating progenitor maintenance or the timing of neural stem cell differentiation, scenarios that are consistent with known roles of HES genes (Hatakeyama et al., 2004; Kageyama et al., 2008). Although some *coe* and *sim* cells were observed in the posterior end of the animal, neural progenitors were most prevalent in the area anterior to the pharynx and posterior to the base of the brain, the location where eye progenitors (*ovo* /*smedwi-1* cells) were also detected (Lapan and Reddien, 2012). Interestingly, our observation that the proportion of BrdU /*coe* and BrdU /*sim* cells located in the pre-pharyngeal area increased over time suggests that this area may represent a 'neurogenic zone' in planarians. Our data support a model in which

Development (2013) doi:10.1242/dev.098616

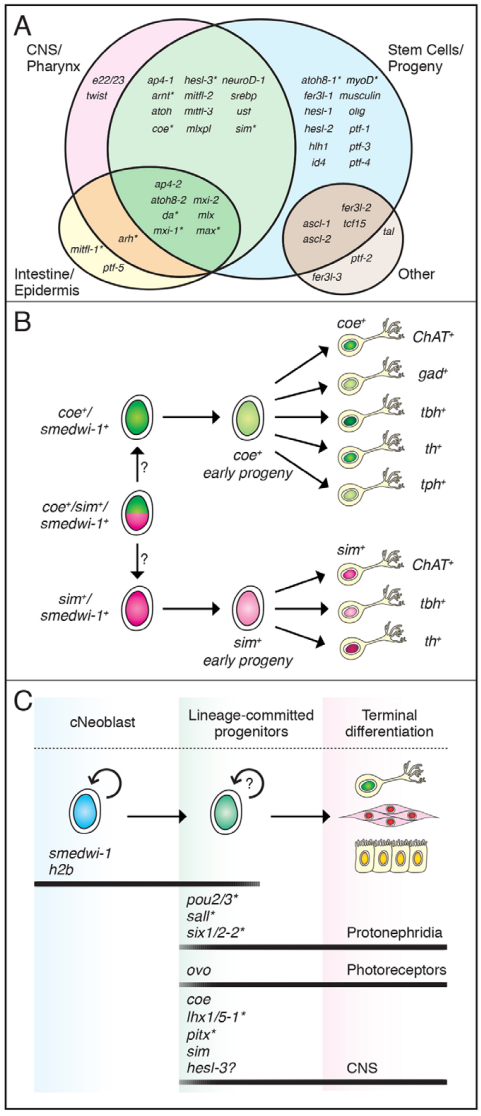


Fig. 6. Planarians possess lineage-committed neural progenitors. (A) Venn diagram summarizing genome-wide expression and functional analysis of bHLH genes in planarians. *Regeneration defects were observed following RNAi. (B) Model of *coe*+ and *sim*+ progenitor cell differentiation into specific neural subtypes. (C) Model of cell differentiation in planarians. Pluripotent adult stem cells (cNeoblasts; *smedwi-1*+ and *h2b*+) have the ability to self-renew and generate lineage-committed progenitors. Summary of identified genes marking photoreceptor (Lapan and Reddien, 2012), protonephridia (Scimone et al., 2011), serotonergic (Currie and Pearson, 2013) and novel CNS (bHLH) progenitors, respectively. *Progenitors that are only observed during regeneration.

pluripotent stem cells (cNeoblasts) maintain lineage-committed progenitors, which generate most if not all of the cells required to meet normal physiological demands in uninjured planarians (Fig. 6C).

bHLH genes with roles in planarian CNS regeneration

Planarians possess members from all of the families of proneural factors, including homologs of *acheate-scute*, *atonal*, *neuroD* and *da*, all of which are primarily expressed in the stem cells. With the exception of *atoh8-1* RNAi, which caused animals to regenerate smaller brains, we found that gene knockdown of most proneural homologs failed to cause overt regeneration defects, even after long-term or combinatorial RNAi. Nonetheless, we did find that co-silencing of *ascl-1* or *neuroD-1* together with *ndk* suppressed ectopic formation of *npp-4* neurons. We surmise that knockdown of some bHLHs may cause subtle defects in neural specification, which are difficult to detect with the use of pan-neural markers. Future functional studies using discrete nervous system markers may reveal additional roles of bHLH genes in CNS differentiation.

On the bases of gene expression patterns and RNAi phenotypes, we further explored the function of *coe*, *hesl-3* and *sim*. *coe* genes are conserved in metazoans and are known to play roles in neuronal specification, migration, axon guidance, dendritogenesis, neuronal subtype specification (Wightman et al., 1997; Dubois et al., 1998; Prasad et al., 1998; Garel et al., 2000; Pozzoli et al., 2001; Garcia-Dominguez et al., 2003; Hattori et al., 2007; Jinushi-Nakao et al., 2007; Crozatier and Vincent, 2008; Demilly et al., 2011; Kratsios et al., 2012), and cellular reprogramming (Richard et al., 2011). In planarians, *coe* knockdown led to a failure of animals to connect the cephalic ganglia. Analysis of this defect using neural subtype markers showed that neurons were found in aberrant locations. In addition, long-term silencing caused animals to exhibit abnormal locomotion and decreases of cholinergic and *pc2* neurons at the anterior commissure and brain *cpp-1* neurons. In *C. elegans*, *coe* (*unc-3*) mutants exhibit behavioral abnormalities (Wightman et al., 1997), a defect that is caused by a loss of cholinergic motoneuron properties (Kratsios et al., 2012). Our data show that *coe* is playing a conserved role in neuronal differentiation during both CNS regeneration and maintenance. *coe* homologs in humans (EBF transcription factors) have been associated with cancers of the nervous system (Liao, 2009), yet the genetic targets of *coe* homologs have not been fully characterized. Thus, further investigation of *coe* function in planarians may reveal mechanisms regulating neural progenitor populations.

hes genes are a primary target of Notch signaling and defects in *hes* genes cause premature neural differentiation and depletion of the neural progenitor pool in mice (Ishibashi et al., 1995; Kageyama et al., 2008). In planarians, *hesl-3* knockdown led animals to regenerate mispatterned brains and a reduction of *cpp-1* , *npp-4* and *npy-2* brain neurons. These data suggest *hesl-3* plays a role in neural fate regulation during CNS repair. At present, the role of Notch signaling in planarians has not been extensively characterized. Thus, analysis of *hesl* genes in stem cell regulation should be a focus of future investigations.

Finally, in flies and crustaceans, *sim* functions as a master regulator of midline cells by regulating the specification of midline progenitors (Nambu et al., 1991; Vargas-Vila et al., 2010), whereas in vertebrates, *sim* controls the differentiation (Michaud et al., 1998; Eaton and Glasgow, 2006) and migration (via plexinC1) (Xu and Fan, 2007) of certain neuroendocrine lineages. *sim* does not appear to function as a master regulator of the midline in planarians. However, *sim* RNAi caused animals to regenerate smaller brains

Development (2013) doi:10.1242/dev.098616

with fewer *npp-4* and *npy-2* neurons, suggesting a potential role in specification and/or guidance of cells from the neuroendocrine lineage. To further explore this possibility, future experiments should test the effects of *sim* knockdown on the fate of all neuropeptide-expressing neurons (Collins et al., 2010) or the expression of guidance molecules, such as plexin homologs.

Conclusions

Our work has revealed that planarians possess lineage-committed progenitors that contribute to the maintenance and regeneration of the CNS. We also identified nine bHLH genes that regulate adult neurogenesis and are required for nervous system repair. This study sets the stage to use planarians as a model to elucidate roles of bHLH genes in adult pluripotent stem cell differentiation. Furthermore, by extending our analysis of bHLH factors genome-wide, this study will serve as a resource for future investigation into bHLH evolution and function.

MATERIALS AND METHODS

Animals

Asexual *Schmidtea mediterranea* (CIW4) were maintained as previously described (Cebrià and Newmark, 2005). Animals 2-5 mm in length that were starved for 1 week were used for all experiments.

bHLH identification, phylogenetic analysis and cloning

To identify planarian bHLH genes, TBLASTN searches were performed against the *S. mediterranea* genome (Robb et al., 2008) and several transcriptomes (Zayas et al., 2005; Adamidi et al., 2011; Labbé et al., 2012; Önal et al., 2012) using bHLH protein sequences from human, mouse and fly. Putative planarian bHLH homologs were validated by performing reciprocal BLASTX against the nr database (NCBI). The bHLH superfamily consists of six monophyletic groups (Groups A-F), which are also characterized by the presence or absence of various additional protein domains (Simionato et al., 2007). Due to the large number of putative paralogs in Groups A and B, the predicted protein sequences were aligned using T-Coffee (Notredame et al., 2000) and subjected to Bayesian analyses as described previously (Currie and Pearson, 2013; Zhu and Pearson, 2013); Group C-F genes were categorized based on clear top BLASTP hits against the Swiss-Prot database (UniProt) and the presence of class-specific protein domains (supplementary material Table S1). bHLH sequences were obtained from a cDNA collection (Zayas et al., 2005) or cloned from cDNA into pJC53.2 (Collins et al., 2010) or pPR244 (Reddien et al., 2005a) using gene specific primers or 3′ RACE, respectively. bHLH sequences were deposited in GenBank. The primers used and GenBank accession numbers are listed in supplementary material Table S2.

In situ hybridization

Riboprobes were synthesized and animals were processed for *in situ* hybridization as previously described (Pearson et al., 2009). For γ-irradiation treatments, animals were exposed to 100 Gy in a JL Shepherd Mark I Cesium-137 irradiator and fixed 6 days after treatment. To visualize bHLH transcripts by multiple fluorescent *in situ* hybridization (FISH), we used horseradish peroxidase substrates as described previously (Pearson et al., 2009) or the alkaline phosphatase (AP) substrate Fast Blue (Lauter et al., 2011; Currie and Pearson, 2013). For Fast Blue staining, animals were developed in 0.25 mg/ml Fast Blue BB (Sigma F3378) and NAMP (Sigma 855) in AP staining buffer (0.1 M Tris-HCl pH 8.2, containing 50 mM $MgCl_2$, 100 mM NaCl, 0.1% Tween 20) (Hauptmann, 2001; Lauter et al., 2011).

BrdU staining

Experiments were conducted by soaking animals in BrdU for 1 hour as previously described (Cowles et al., 2012), chasing for 4, 24 or 48 hours before fixation and processing for FISH, and then processing for BrdU labeling starting with the HCl treatment.

RNA interference

For regeneration studies, we administered six feedings of bacterially expressed dsRNA over 3 weeks as previously described (Gurley et al, 2008). *gfp* was used as a control for all experiments. 24 hours following the final RNAi treatment, animals were amputated anterior to the pharynx, observed for 10 days and then processed for *in situ* hybridization or immunostaining. For long-term experiments, animals were fed 12 times over 6 weeks before amputation; uninjured animals were fixed 1 week after the final feeding. Relative gene expression after RNAi was determined by real-time quantitative PCR as described previously (Hubert et al., 2013); primers are listed in supplementary material Table S2.

Immunohistochemistry

Immunostaining with anti-SYNAPSIN (1:400, 3C10, Developmental Studies Hybridoma Bank) and anti-phosphohistone-H3 (S10) (1:1000, D2C8, Cell Signaling) were performed as previously described (Cowles et al., 2012).

Imaging

Images were acquired using a Leica DFC450 camera mounted on a Leica M205 stereomicroscope. Animals labeled with fluorescent probes were imaged with an Axiocam MRm camera mounted on a Zeiss SteREO Lumar V.12 or Axio Observer.Z1 equipped with an ApoTome, or a Hamamatsu ImagEM C9100-13 camera mounted on an Olympus IX81 microscope equipped with a Yokogawa CSU X1 spinning-disk confocal scan head.

Cell counting

Ten 1-μm optical sections were captured from selected regions and merged, and cells were hand-counted using ImageJ 1.43u software. The proportions of cells co-expressing specific neurotransmitters or *smedwi-1* and *coe*, *hesl-3* or *sim* were calculated from >100 cells counted from three to five animals. The proportion of $BrdU^+$ cells co-expressing specific genes was calculated from >300 $BrdU^+$ cells counted from three to five animals. For analysis of *ndk* RNAi animals, coe^+ and sim^+ cells were counted and normalized per mm^3; ectopic $npp\text{-}4^+$ cells were counted from the posterior end of the brain to the posterior boundary of the pharynx and normalized to the length of the animal. Mean and s.d. values were computed and statistical comparisons were performed using an unpaired Student's *t*-test. Error bars in graphs are s.d.

Acknowledgements

We thank Jordana Henderson, Amy Hubert and Kelly Ross for helpful comments on the manuscript, Kayla Muth for assistance with RNAi experiments, Claire Cowles for artwork design, and the anonymous reviewers whose constructive criticisms improved this work.

Competing interests

The authors declare no competing financial interests.

Author contributions

M.W.C. and R.M.Z. designed and interpreted the experiments and wrote the manuscript. M.W.C., D.D.R.B., S.V.N. and B.N.S. conducted the experiments and analyzed the data. B.J.P. performed phylogenetic analysis. M.W.C, B.J.P. and R.M.Z. discussed the results and edited the final version of the manuscript.

Funding

M.W.C. acknowledges support from the San Diego Chapter of the Achievement Rewards for College Scientists Foundation and the Inamori Foundation. D.D.R.B. was supported by a Canadian Institutes of Health Research Student Fellowship [GSD-121763]. This research was supported by a Natural Sciences and Engineering Research Council of Canada grant [402264-11] and Ontario Institute for Cancer Research New Investigator Award [IA-026] to B.J.P. and a California Institute for Regenerative Medicine grant [RN2-00940-1] to R.M.Z.

Supplementary material

Supplementary material available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.098616/-/DC1

References

Adamidi, C., Wang, Y., Gruen, D., Mastrobuoni, G., You, X., Tolle, D., Dodt, M., Mackowiak, S. D., Gogol-Doering, A., Oenal, P. et al. (2011). De novo assembly

and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics. *Genome Res.* **21**, 1193-1200.
Baguñà, J. (2012). The planarian neoblast: the rambling history of its origin and some current black boxes. *Int. J. Dev. Biol.* **56**, 19-37.
Baguñà, J., Saló, E. and Auladell, C. (1989). Regeneration and pattern formation in planarians. III. Evidence that neoblasts are totipotent stem cells and the source of blastema cells. *Development* **107**, 77-86.
Bertrand, N., Castro, D. S. and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nat. Rev. Neurosci.* **3**, 517-530.
Blassberg, R. A., Felix, D. A., Tejada-Romero, B. and Aboobaker, A. A. (2013). PBX/extradenticle is required to re-establish axial structures and polarity during planarian regeneration. *Development* **140**, 730-739.
Bullock, T. H. and Horridge, G. A. (1965). *Structure and Function in the Nervous Systems of Invertebrates*. San Francisco, CA, USA: W. H. Freeman.
Cebrià, F. (2007). Regenerating the central nervous system: how easy for planarians! *Dev. Genes Evol.* **217**, 733-748.
Cebrià, F. and Newmark, P. A. (2005). Planarian homologs of netrin and netrin receptor are required for proper regeneration of the central nervous system and the maintenance of nervous system architecture. *Development* **132**, 3691-3703.
Cebrià, F., Kobayashi, C., Umesono, Y., Nakazawa, M., Mineta, K., Ikeo, K., Gojobori, T., Itoh, M., Taira, M., Sánchez Alvarado, A. et al. (2002). FGFR-related gene nou-darake restricts brain tissues to the head region of planarians. *Nature* **419**, 620-624.
Cebrià, F., Guo, T., Jopek, J. and Newmark, P. A. (2007). Regeneration and maintenance of the planarian midline is regulated by a slit orthologue. *Dev. Biol.* **307**, 394-406.
Collins, J. J., 3rd, Hou, X., Romanova, E. V., Lambrus, B. G., Miller, C. M., Saberi, A., Sweedler, J. V. and Newmark, P. A. (2010). Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLoS Biol.* **8**, e1000509.
Cowles, M. W., Hubert, A. and Zayas, R. M. (2012). A Lissencephaly-1 homologue is essential for mitotic progression in the planarian Schmidtea mediterranea. *Dev. Dyn.* **241**, 901-910.
Crozatier, M. and Vincent, A. (2008). Control of multidendritic neuron differentiation in Drosophila: the role of Collier. *Dev. Biol.* **315**, 232-242.
Currie, K. W. and Pearson, B. J. (2013). Transcription factors lhx1/5-1 and pitx are required for the maintenance and regeneration of serotonergic neurons in planarians. *Development* **140**, 3577-3588.
Demilly, A., Simionato, E., Ohayon, D., Kerner, P., Garcès, A. and Vervoort, M. (2011). Coe genes are expressed in differentiating neurons in the central nervous system of protostomes. *PLoS ONE* **6**, e21213.
Dubois, L. and Vincent, A. (2001). The COE–Collier/Olf1/EBF–transcription factors: structural conservation and diversity of developmental functions. *Mech. Dev.* **108**, 3-12.
Dubois, L., Bally-Cuif, L., Crozatier, M., Moreau, J., Paquereau, L. and Vincent, A. (1998). XCoe2, a transcription factor of the Col/Olf-1/EBF family involved in the specification of primary neurons in Xenopus. *Curr. Biol.* **8**, 199-209.
Eaton, J. L. and Glasgow, E. (2006). The zebrafish bHLH PAS transcriptional regulator, single-minded 1 (sim1), is required for isotocin cell development. *Dev. Dyn.* **235**, 2071-2082.
Eisenhoffer, G. T., Kang, H. and Sánchez Alvarado, A. (2008). Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian Schmidtea mediterranea. *Cell Stem Cell* **3**, 327-339.
Elliott, S. A. and Sánchez Alvarado, A. (2013). The history and enduring contributions of planarians to the study of animal regeneration. *Wiley Interdiscip. Rev. Dev. Biol.* **2**, 301-326.
Felix, D. A. and Aboobaker, A. A. (2010). The TALE class homeobox gene Smed-prep defines the anterior compartment for head regeneration. *PLoS Genet.* **6**, e1000915.
Forsthoefel, D. J., Park, A. E. and Newmark, P. A. (2011). Stem cell-based growth, regeneration, and remodeling of the planarian intestine. *Dev. Biol.* **356**, 445-459.
Forsthoefel, D. J., James, N. P., Escobar, D. J., Stary, J. M., Vieira, A. P., Waters, F. A. and Newmark, P. A. (2012). An RNAi screen reveals intestinal regulators of branching morphogenesis, differentiation, and stem cell proliferation in planarians. *Dev. Cell* **23**, 691-704.
Gage, F. H. (2002). Neurogenesis in the adult brain. *J. Neurosci.* **22**, 612-613.
Galliot, B., Quiquand, M., Ghila, L., de Rosa, R., Miljkovic-Licina, M. and Chera, S. (2009). Origins of neurogenesis, a cnidarian view. *Dev. Biol.* **332**, 2-24.
Garcia-Dominguez, M., Poquet, C., Garel, S. and Charnay, P. (2003). Ebf gene function is required for coupling neuronal differentiation and cell cycle exit. *Development* **130**, 6013-6025.
Garel, S., Garcia-Dominguez, M. and Charnay, P. (2000). Control of the migratory pathway of facial branchiomotor neurones. *Development* **127**, 5297-5307.
Gentile, L., Cebrià, F. and Bartscherer, K. (2011). The planarian flatworm: an in vivo model for stem cell biology and nervous system regeneration. *Dis. Model. Mech.* **4**, 12-19.
Goulding, S. E., zur Lage, P. and Jarman, A. P. (2000). amos, a proneural gene for Drosophila olfactory sense organs that is regulated by lozenge. *Neuron* **25**, 69-78.
Guedelhoefer, O. C., 4th and Sánchez Alvarado, A. (2012). Amputation induces stem cell mobilization to sites of injury during planarian regeneration. *Development* **139**, 3510-3520.
Guillemot, F. (2007). Spatial and temporal specification of neural fates by transcription factor codes. *Development* **134**, 3771-3780.
Gurley, K. A., Rink, J. C. and Sánchez Alvarado, A. (2008). Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. *Science* **319**, 323-327.
Hatakeyama, J., Bessho, Y., Katoh, K., Ookawara, S., Fujioka, M., Guillemot, F. and Kageyama, R. (2004). Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation. *Development* **131**, 5539-5550.
Hattori, Y., Sugimura, K. and Uemura, T. (2007). Selective expression of Knot/Collier, a transcriptional regulator of the EBF/Olf-1 family, endows the Drosophila sensory system with neuronal class-specific elaborated dendritic patterns. *Genes Cells* **12**, 1011-1022.
Hauptmann, G. (2001). One-, two-, and three-color whole-mount in situ hybridization to Drosophila embryos. *Methods* **23**, 359-372.
Huang, M. L., Hsu, C. H. and Chien, C. T. (2000). The proneural gene amos promotes multiple dendritic neuron formation in the Drosophila peripheral nervous system. *Neuron* **25**, 57-67.
Hubert, A., Henderson, J. M., Ross, K. G., Cowles, M. W., Torres, J. and Zayas, R. M. (2013). Epigenetic regulation of planarian stem cells by the SET1/MLL family of histone methyltransferases. *Epigenetics* **8**, 79-91.
Iglesias, M., Almuedo-Castillo, M., Aboobaker, A. A. and Saló, E. (2011). Early planarian brain regeneration is independent of blastema polarity mediated by the Wnt/β-catenin pathway. *Dev. Biol.* **358**, 68-78.
Ishibashi, M., Ang, S. L., Shiota, K., Nakanishi, S., Kageyama, R. and Guillemot, F. (1995). Targeted disruption of mammalian hairy and Enhancer of split homolog-1 (HES-1) leads to up-regulation of neural helix-loop-helix factors, premature neurogenesis, and severe neural tube defects. *Genes Dev.* **9**, 3136-3148.
Jan, Y. N. and Jan, L. Y. (1994). Neuronal cell fate specification in Drosophila. *Curr. Opin. Neurobiol.* **4**, 8-13.
Jinushi-Nakao, S., Arvind, R., Amikura, R., Kinameri, E., Liu, A. W. and Moore, A. W. (2007). Knot/Collier and cut control different aspects of dendrite cytoskeleton and synergize to define final arbor shape. *Neuron* **56**, 963-978.
Kageyama, R., Ohtsuka, T. and Kobayashi, T. (2008). Roles of Hes genes in neural development. *Dev. Growth Differ.* **50** Suppl., S97-S103.
Kempermann, G. (2011). *Adult Neurogenesis*. New York, NY, USA: Oxford University Press.
Kempermann, G. (2012). New neurons for 'survival of the fittest'. *Nat. Rev. Neurosci.* **13**, 727-736.
Kewley, R. J., Whitelaw, M. L. and Chapman-Smith, A. (2004). The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *Int. J. Biochem. Cell Biol.* **36**, 189-204.
King, R. S. and Newmark, P. A. (2012). The cell biology of regeneration. *J. Cell Biol.* **196**, 553-562.
Kintner, C. (2002). Neurogenesis in embryos and in adult neural stem cells. *J. Neurosci.* **22**, 639-643.
Kratsios, P., Stolfi, A., Levine, M. and Hobert, O. (2012). Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. *Nat. Neurosci.* **15**, 205-214.
Kretzschmar, K. and Watt, F. M. (2012). Lineage tracing. *Cell* **148**, 33-45.
Labbé, R. M., Irimia, M., Currie, K. W., Lin, A., Zhu, S. J., Brown, D. D., Ross, E. J., Voisin, V., Bader, G. D., Blencowe, B. J. et al. (2012). A comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. *Stem Cells* **30**, 1734-1745.
Lapan, S. W. and Reddien, P. W. (2011). dlx and sp6-9 Control optic cup regeneration in a prototypic eye. *PLoS Genet.* **7**, e1002226.
Lapan, S. W. and Reddien, P. W. (2012). Transcriptome analysis of the planarian eye identifies ovo as a specific regulator of eye regeneration. *Cell Rep.* **2**, 294-307.
Lauter, G., Söll, I. and Hauptmann, G. (2011). Two-color fluorescent in situ hybridization in the embryonic zebrafish brain using differential detection systems. *BMC Dev. Biol.* **11**, 43.
Lentz, T. L. (1968). *Primitive Nervous Systems*. New Haven, CT, USA: Yale University Press.
Liao, D. (2009). Emerging roles of the EBF family of transcription factors in tumor suppression. *Mol. Cancer Res.* **7**, 1893-1901.
Lindsey, B. W. and Tropepe, V. (2006). A comparative framework for understanding the biological principles of adult neurogenesis. *Prog. Neurobiol.* **80**, 281-307.
Michaud, J. L., Rosenquist, T., May, N. R. and Fan, C. M. (1998). Development of neuroendocrine lineages requires the bHLH-PAS transcription factor SIM1. *Genes Dev.* **12**, 3264-3275.
Nambu, J. R., Lewis, J. O., Wharton, K. A., Jr and Crews, S. T. (1991). The Drosophila single-minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development. *Cell* **67**, 1157-1167.
Newmark, P. A. and Sánchez Alvarado, A. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Dev. Biol.* **220**, 142-153.
Newmark, P. A., Wang, Y. and Chong, T. (2008). Germ cell specification and regeneration in planarians. *Cold Spring Harb. Symp. Quant. Biol.* **73**, 573-581.
Notredame, C., Higgins, D. G. and Heringa, J. (2000). T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205-217.
Önal, P., Grün, D., Adamidi, C., Rybak, A., Solana, J., Mastrobuoni, G., Wang, Y., Rahn, H. P., Chen, W., Kempa, S. et al. (2012). Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. *EMBO J.* **31**, 2755-2769.
Pearson, B. J., Eisenhoffer, G. T., Gurley, K. A., Rink, J. C., Miller, D. E. and Sánchez Alvarado, A. (2009). Formaldehyde-based whole-mount in situ hybridization method for planarians. *Dev. Dyn.* **238**, 443-450.

Development (2013) doi:10.1242/dev.098616

Pozzoli, O., Bosetti, A., Croci, L., Consalez, G. G. and Vetter, M. L. (2001). Xebf3 is a regulator of neuronal differentiation during primary neurogenesis in Xenopus. *Dev. Biol.* **233**, 495-512.

Prasad, B. C., Ye, B., Zackhary, R., Schrader, K., Seydoux, G. and Reed, R. R. (1998). unc-3, a gene required for axonal guidance in Caenorhabditis elegans, encodes a member of the O/E family of transcription factors. *Development* **125**, 1561-1568.

Probst, M. R., Fan, C. M., Tessier-Lavigne, M. and Hankinson, O. (1997). Two murine homologs of the Drosophila single-minded protein that interact with the mouse aryl hydrocarbon receptor nuclear translocator protein. *J. Biol. Chem.* **272**, 4451-4457.

Ramón y Cajal, S. (1928). *Degeneration and Regeneration of the Nervous System*. London, UK: Oxford University Press.

Reddien, P. W. (2013). Specialized progenitors and regeneration. *Development* **140**, 951-957.

Reddien, P. W., Bermange, A. L., Murfitt, K. J., Jennings, J. R. and Sánchez Alvarado, A. (2005a). Identification of genes needed for regeneration, stem cell function, and tissue homeostasis by systematic gene perturbation in planaria. *Dev. Cell* **8**, 635-649.

Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. and Sánchez Alvarado, A. (2005b). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* **310**, 1327-1330.

Resch, A. M., Palakodeti, D., Lu, Y. C., Horowitz, M. and Graveley, B. R. (2012). Transcriptome analysis reveals strain-specific and conserved sternness genes in Schmidtea mediterranea. *PLoS ONE* **7**, e34447.

Richard, J. P., Zuryn, S., Fischer, N., Pavet, V., Vaucamps, N. and Jarriault, S. (2011). Direct in vivo cellular reprogramming involves transition through discrete, non-pluripotent steps. *Development* **138**, 1483-1492.

Richards, G. S., Simionato, E., Perron, M., Adamska, M., Vervoort, M. and Degnan, B. M. (2008). Sponge genes provide new insight into the evolutionary origin of the neurogenic circuit. *Curr. Biol.* **18**, 1156-1161.

Rink, J. C. (2013). Stem cell systems and regeneration in planaria. *Dev. Genes Evol.* **223**, 67-84.

Robb, S. M., Ross, E. and Sánchez Alvarado, A. (2008). SmedGD: the Schmidtea mediterranea genome database. *Nucleic Acids Res.* **36**, D599-D606.

Scimone, M. L., Srivastava, M., Bell, G. W. and Reddien, P. W. (2011). A regulatory program for excretory system regeneration in planarians. *Development* **138**, 4387-4398.

Simionato, E., Ledent, V., Richards, G., Thomas-Chollier, M., Kerner, P., Coornaert, D., Degnan, B. M. and Vervoort, M. (2007). Origin and diversification of the basic helix-loop-helix gene family in metazoans: insights from comparative genomics. *BMC Evol. Biol.* **7**, 33.

Solana, J., Kao, D., Mihaylova, Y., Jaber-Hijazi, F., Malla, S., Wilson, R. and Aboobaker, A. (2012). Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNAseq, RNA interference and irradiation approach. *Genome Biol.* **13**, R19.

Umesono, Y., Tasaki, J., Nishimura, K., Inoue, T. and Agata, K. (2011). Regeneration in an evolutionarily primitive brain–the planarian Dugesia japonica model. *Eur. J. Neurosci.* **34**, 863-869.

Vargas-Vila, M. A., Hannibal, R. L., Parchem, R. J., Liu, P. Z. and Patel, N. H. (2010). A prominent requirement for single-minded and the ventral midline in patterning the dorsoventral axis of the crustacean Parhyale hawaiensis. *Development* **137**, 3469-3476.

Wagner, D. E., Wang, I. E. and Reddien, P. W. (2011). Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* **332**, 811-816.

Wightman, B., Baran, R. and Garriga, G. (1997). Genes that guide growth cones along the C. elegans ventral nerve cord. *Development* **124**, 2571-2580.

Xu, C. and Fan, C. M. (2007). Allocation of paraventricular and supraoptic neurons requires Sim1 function: a role for a Sim1 downstream gene PlexinC1. *Mol. Endocrinol.* **21**, 1234-1245.

Zayas, R. M., Hernández, A., Habermann, B., Wang, Y., Stary, J. M. and Newmark, P. A. (2005). The planarian Schmidtea mediterranea as a model for epigenetic germ cell specification: analysis of ESTs from the hermaphroditic strain. *Proc. Natl. Acad. Sci. USA* **102**, 18491-18496.

Zhu, S. J. and Pearson, B. J. (2013). The Retinoblastoma pathway regulates stem cell proliferation in freshwater planarians. *Dev. Biol.* **373**, 442-452.

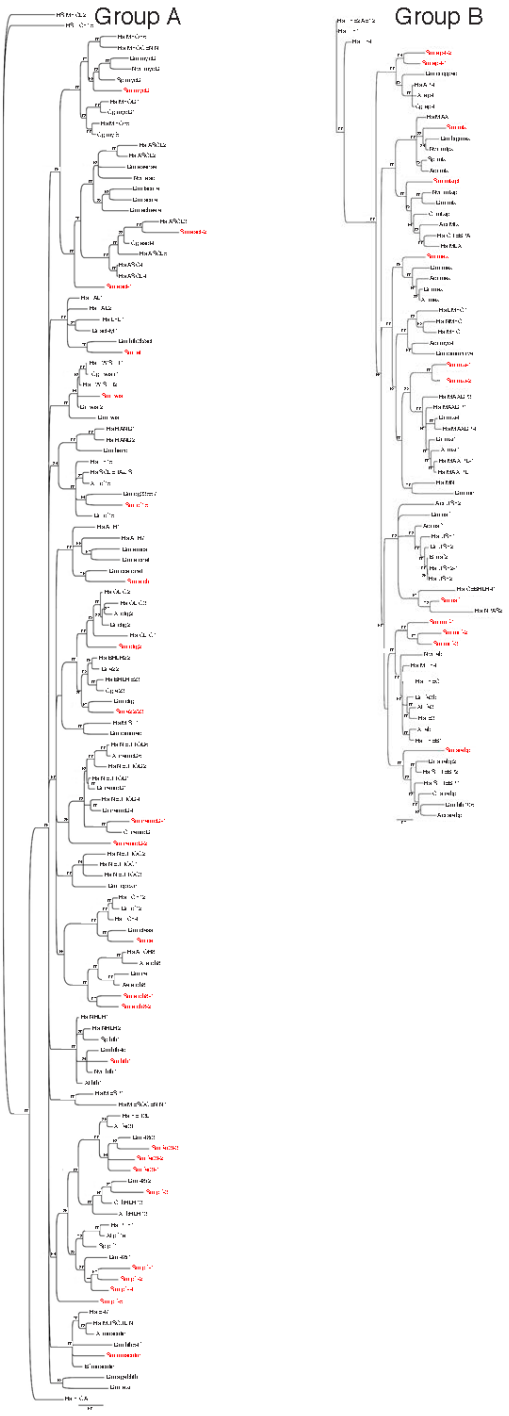


Fig. S1: Bayesian Phylogeny of Groups A and B bHLH transcription factor homologs. Protein sequences used in the phylogenies were obtained from the NCBI Entrez protein database or directly from the genome sequencing projects of included organisms. Sequences were aligned using T-Coffee (Notredame et al., 2000) and the program Geneious (www.geneious.com) was used for Bayesian analyses with the following settings: 1 million replicates, WAG substitution model, 4 heated chains, 25% burnin, and subsample frequency of 1000. Consensus tree images were saved through Geneious and then manipulated in Adobe Photoshop. Only bootstrap values over 50 are shown. Species used: Adi=*Acropora digitifera*; Ag=*Anopheles gambiae*; Bf=*Branchiostoma floridae*; Bt=*Bos taurus*; Ct=*Capitella teleta*; Dm=*Drosophila melanogaster*; Dr=*Danio rerio*; Gg=*Gallus gallus*; Hs=*Homo sapiens*; Nvit=*Nematostella vectensis*; Sp=*Strongylocentrotus purpuratus*; Sm=*Schmidtea mediterranea*; Xl=*Xenopus laevis*; Xt=*Xenopus tropicalis*. *S. mediterranea* homologs are in red fonts.

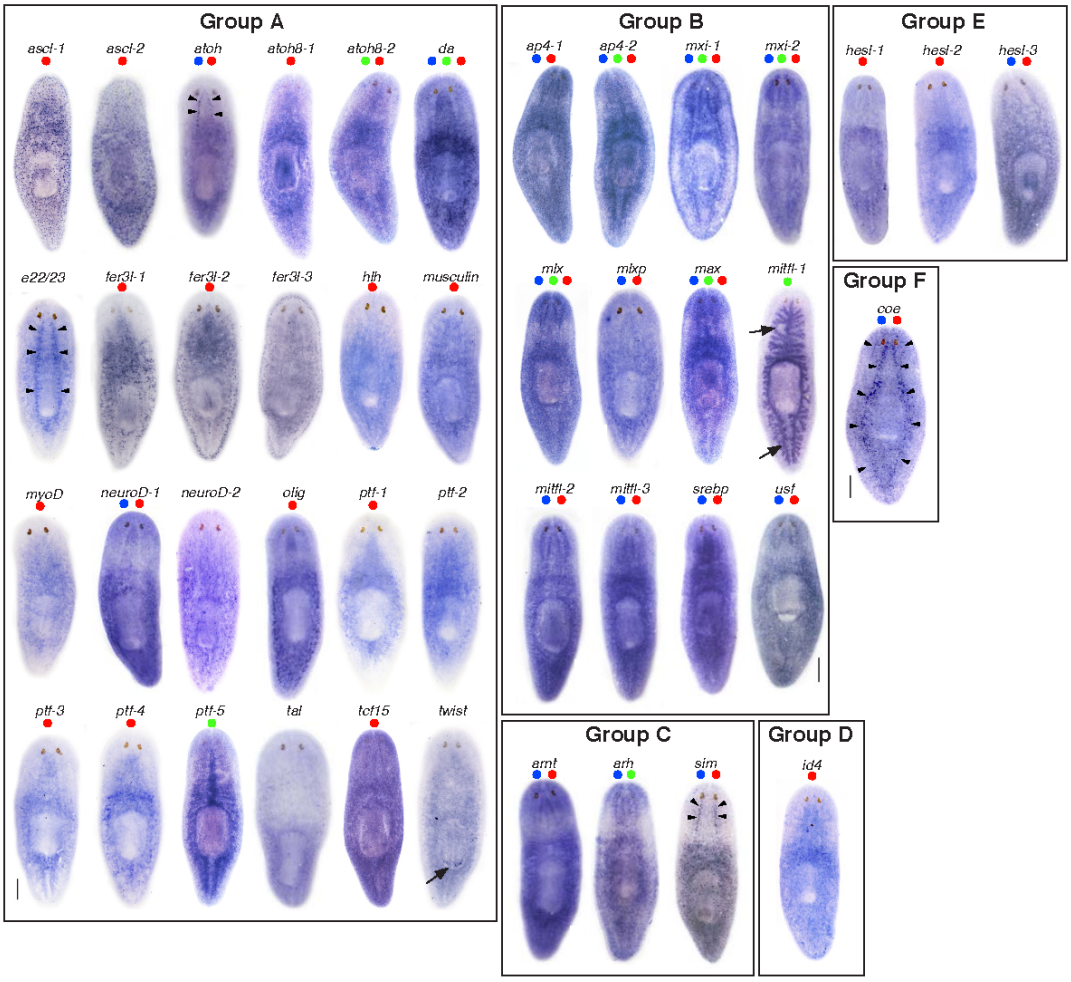


Fig. S2: Expression analysis of bHLH genes in *S. mediterranea*. Intact animals were processed for whole-mount in situ hybridization to bHLH genes. Gene names are indicated above each animal. bHLHs were expressed in a wide range of cells and tissues. *ascl-1* and *ascl-2* were expressed in a punctate pattern throughout the mesenchyme. *atoh*, *e22/23*, *sim*, and *coe* were expressed in distinct cells in the CNS (black arrowheads). *mitfl-1* and *twist* were exclusively detected in the intestine and pharynx (black arrows). Many bHLH genes, including *arnt*, *da*, *ap4-1*, *max* and *srebp*, were detected ubiquitously throughout the animal. *fer3l-1*, *fer3l-2*, and *fer3l-3* exhibited related expression patterns with *fer3l-1* expression detected in the interior mesechyme (stem cell-like) and *fer3l-2* and *fer3l-3* detected more exteriorly (similar to a post-mitotic progeny pattern). No definitive expression pattern was observed for *neuroD-2*. The expression of genes in planarian stem cells and immediate progeny is characterized by parenchymal (mesenchymal) staining ranging from punctate expression in stem cell or progeny to diffuse expression throughout the animal and γ-irradiation sensitive. As expected, most bHLHs displayed reduced expression following γ-irradiation (see Table S1). Blue and green dots above the animals denote expression in the CNS and intestine, respectively; red dots denote genes that were γ-irradiation sensitive. Genes were categorized in bHLH Groups A-F based on their homology. Scale bars = 200 μm.

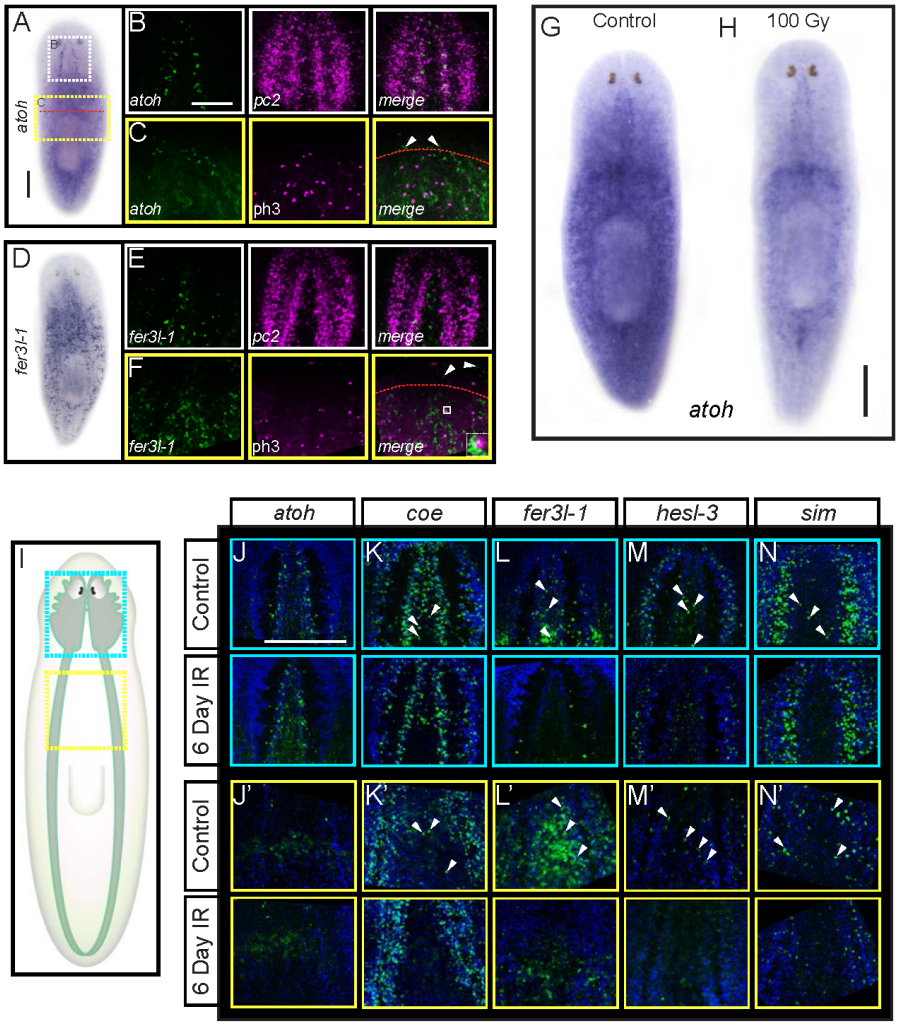


Fig. S3: bHLH genes are expressed in γ-irradiation-sensitive populations near the CNS and stem cell compartment. (A) Expression pattern of atoh. Dashed boxes indicate zoom area of the brain or regeneration blastema shown in B and C, respectively. Dashed red line indicates site of amputation. (B) FISH to *atoh* (green) and *pc2* (magenta). (C) FISH to *atoh* counterstained with anti-phosphohistone-H3 (ph3) to label mitotic cells in 3 day regenerates. Arrowheads denote *atoh*$^{+}$ cells within the blastema. (D-F) show similar analysis for *fer3l-1*. White dashed box in F highlights a bHLH/ph3-positive cells shown at high magnification within the merged image inset. (G and H) WISH to *atoh* in controls or animals 6 days post-irradiation (100 Gy). (I) Cartoon depicting the planarian CNS; blue (head) and yellow (pre-pharyngeal) boxes denote areas of the animal imaged in J-N and J'-N', respectively. (J-N') Control and γ-irradiated animals processed for fluorescent in situ hybridization to *atoh*, *coe*, *fer3l-1*, *hesl-3*, and *sim*, and counterstained with DAPI. Arrowheads denote representative cell populations lost following γ-irradiation. Scale bars = 200 µm.

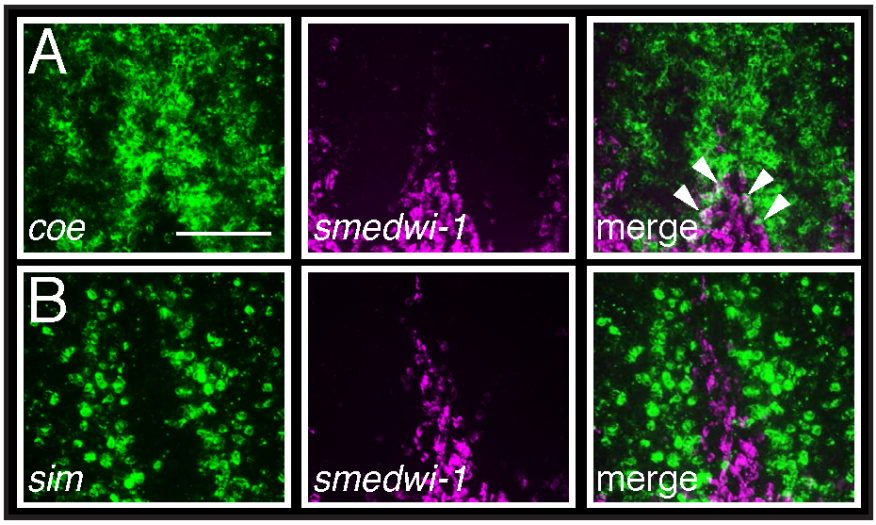


Fig. S4: *coe* and *sim* are not co-expressed with *smedwi-1* in the anterior region of the cephalic ganglia. (A-B) Representative images from the head region of animals processed for double-fluorescent in situ hybridization to *coe* or *sim* and *smedwi-1*. White arrows point to *coe*$^{+}$/*smedwi-1*$^{+}$ cells. Scale bar in A = 100 µm.

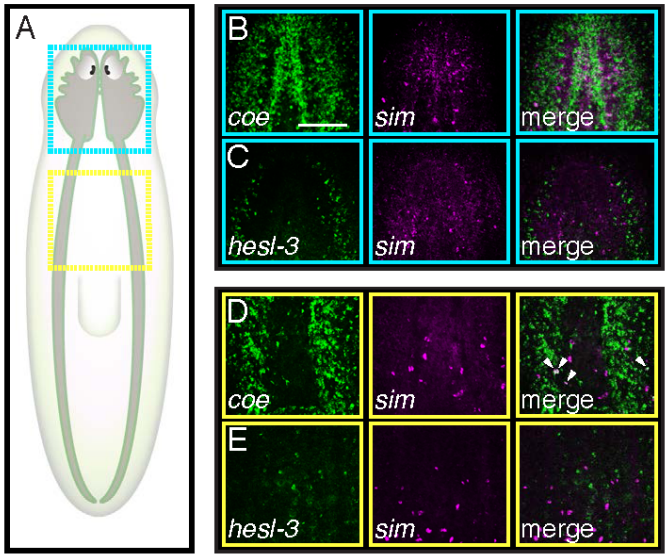


Fig. S5: *coe* and *sim* are co-expressed in cells in the pre-pharyngeal area. (A) Cartoon depicting the planarian CNS. Blue (head) and yellow (pre-pharyngeal) boxes denote the region of the animal imaged in B-D, respectively. (B-E) Images of the brain region of animals processed for double-fluorescent in situ hybridization to *coe* and *sim* or *hesl-3* and *sim*. Scale bar in B = 100 µm.

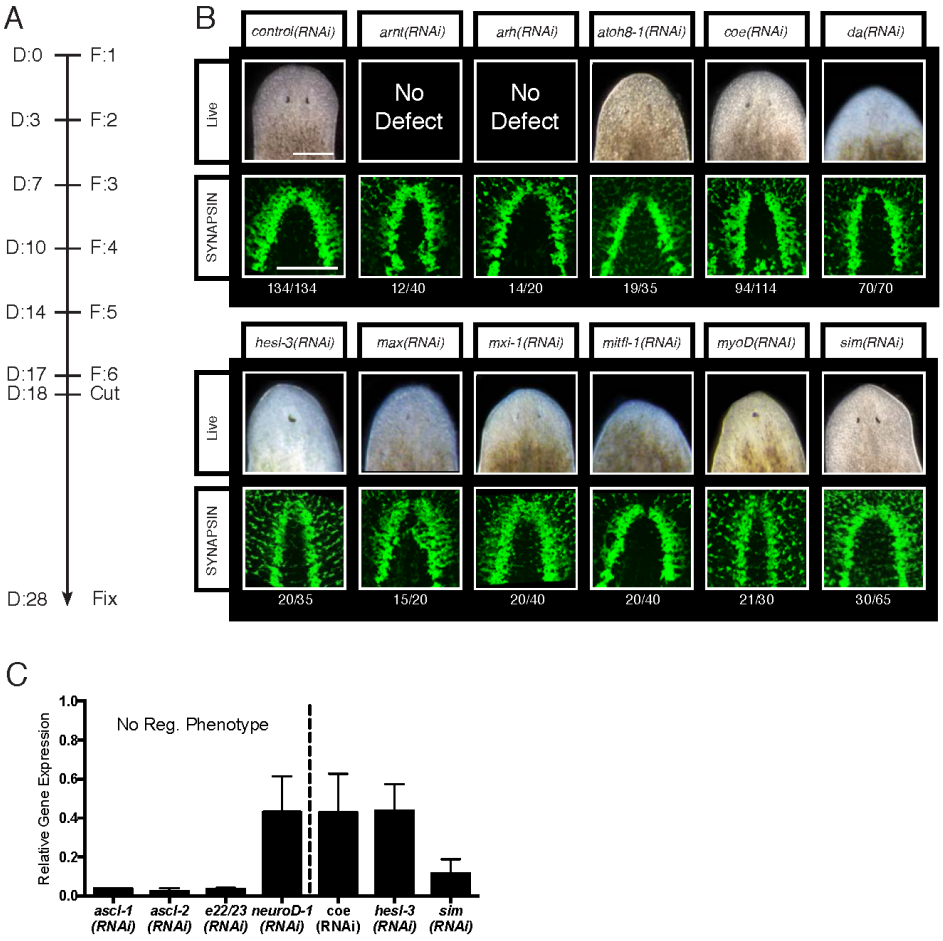


Fig. S6: bHLH RNAi screen for defects in CNS regeneration. (A) Experimental design for RNAi screen; D and F denote days and number of bacterial feedings, respectively. (B) Summary of RNAi phenotypes following bHLH knockdowns. Images shown are of 10-day regenerates. Numbers below images refer to the number of animals with an observable regeneration defect. (C) Quantitative real-time PCR measurements of relative mRNA expression after RNAi knockdown of selected bHLH genes. *ascl-1*, *ascl-2*, *e22/23* and *neuroD-1* did not result in a regeneration phenotype following RNAi. Scale bars = 200 µm

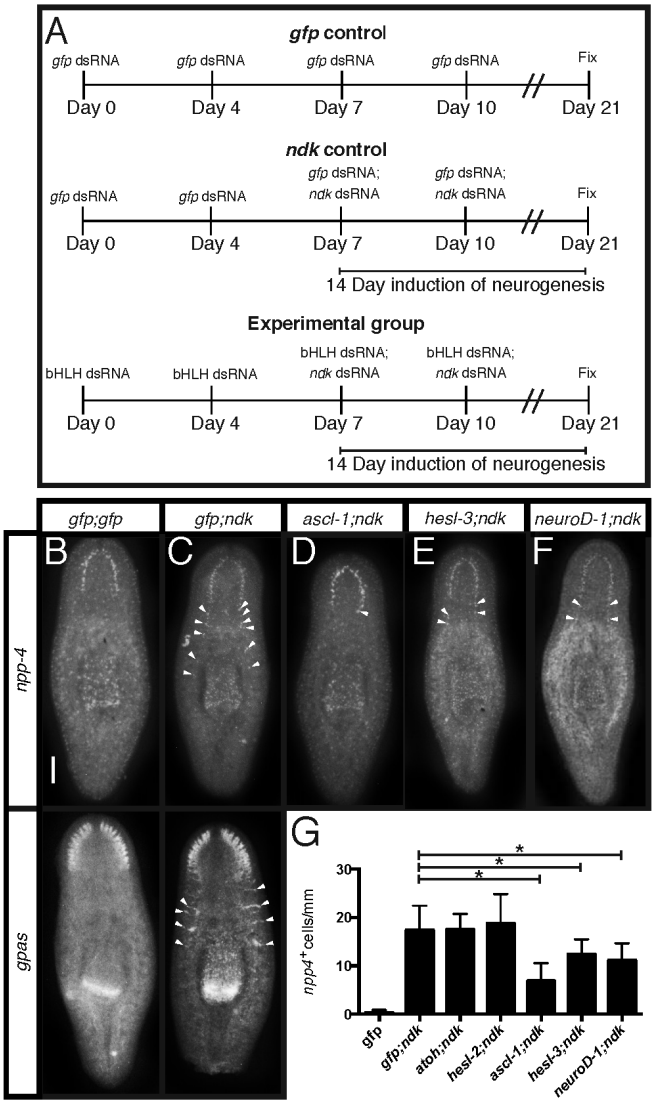


Fig. S7. *ascl-1*, *hesl-3*, and *neuroD-1* suppress ectopic formation of *npp-4*$^+$ cells when co-silenced with *ndk*. (A) Schematic of RNAi-based suppression assay. For double knockdown experiments, bacterial pellets containing dsRNA for each gene were mixed 1:1. For select bHLH and ndk co-silencing experiments, planarians were fed dsRNA four times over two weeks. The first two RNAi feedings contained *bHLH* dsRNA and the final two RNAi treatments contained both *bHLH* and *ndk* dsRNA. (B-F) *gfp;gfp(RNAi)*, *gfp;ndk(RNAi)*, *ascl-1;ndk(RNAi)*, *hesl-3;ndk(RNAi)*, and *neuroD-1;ndk(RNAi)* animals were processed for FISH to *npp-4* or *gpas*. (G) Quantification of ectopic *npp-4*$^+$ cells (arrowheads in B-F of top row; n > 9 animals per group); neurons were normalized by the length of the animal (mm). Asterisks denote a significant reduction of cells when compared with *gfp;ndk(RNAi)* animals ($p < 0.05$, Student's t-test). The expansion of *gpas* after *ndk* RNAi (arrowheads in B and C of bottom row) was not affected after *bHLH* knockdowns. Scale bar = 200 μm.

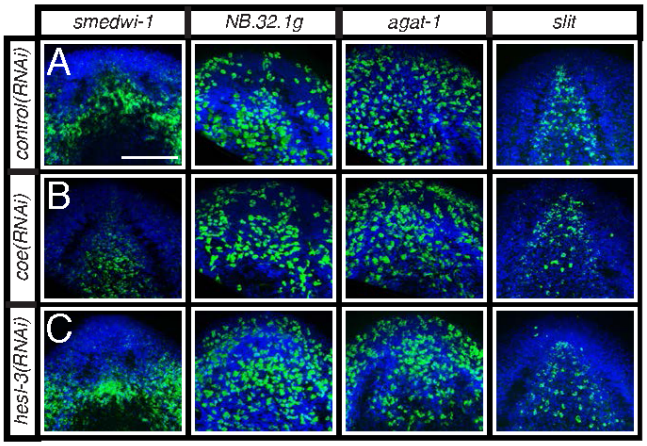


Fig. S8: *coe* and *hesl-3* RNAi phenotypes are not due to a loss of the stem cells, progeny, or midline signals. (A-C) RNAi treated animals were amputated, allowed to regenerate for 5 days, and then processed for fluorescent in situ hybridization to *smedwi-1*, *NB.32.1g*, *agat-1*, or *slit* and counterstained with DAPI. Scale bar = 100 μm.

Chapter 1, in full, is a reprint of the material as it appears in the Journal of Development 2013. Cowles, M.W., Brown, D.R., Stanley, B.N., Nisperos S.V., Pearson, B.J., and Zayas, R.M. The dissertation author was the primary investigator and author of this manuscript.

CHAPTER 2:

COE LOSS-OF-FUNCTION ANALYSIS REVEALS A GENETIC PROGRAM UNDERLYING NEURONAL REGENERATION AND STEM CELL REGULATION IN PLANARIANS

INTRODUCTION

The Collier/Olfactory-1/Early B-cell factor (COE) family of transcription factors is necessary for animal development. COE proteins possess an atypical HLH domain and a unique zinc finger DNA binding domain conserved across metazoans (1). Invertebrates encode a single homolog of COE, with roles in mesoderm and ectoderm development (2, 3), whereas vertebrates have four COE paralogs with functions in diverse cell types including B-cells and adipocytes (4). In the central nervous system (CNS), COE regulates neuronal differentiation, migration, axon guidance, and dendritogenesis during development (2, 3, 5-11) and maintains neuronal identitity throughout adulthood (12-14). COE proteins have also been proposed to function as tumor suppressors (15) and are associated with cancers such as acute lymphoblastic leukemia and glioblastoma (16-19). The specific regulatory programs regulated by these genes in adult stem cells and mature neurons remain poorly understood; therefore, it is unclear how COE dysfunction contributes to nervous system diseases.

Stem cells can be studied to determine how transcriptional regulators orchestrate developmental processes or cause disease (20). An excellent model organism to investigate stem cell regulation *in vivo* is the freshwater planarian *Schmidtea mediterranea* (21). *S. mediterranea* has the ability to regenerate all tissue types from a population of adult stem cells (called neoblasts). These cells constitute approximately 10-20% of all cells in the animal and include pluripotent stem cells (22) and lineage-committed progenitors (23-26). The planarian CNS is composed of

anterior cephalic ganglia and two ventral nerve cords that run along the length of the animal composed of molecularly diverse neuronal subtypes that are generated within days after injury or amputation (27, 28). Functional analysis of transcription factors in planarians using RNA interference (RNAi) has begun to identify regulatory molecules required for the generation and maintenance of specific neuronal subtypes in the CNS such as serotonergic and cholinergic neurons (23-26, 29-31). Thus, planarians provide a powerful system for studying basic mechanisms that underlie stem cell-based maintenance, repair, and regeneration of the adult CNS.

We previously identified a planarian *coe* homolog that is primarily expressed in neural progenitors and neurons. We showed that *coe* is required for brain regeneration; in uninjured animals fed dsRNA designed to silence *coe* expression (*coe(RNAi)* animals), we observed a strong behavioral defect and loss of expression of neural subtype-specific genes (*ChAT*, *pc-2,* and *cpp-1*) [23]. In this study we provide evidence that COE activates genetic programs in multiple classes of neurons and differentiating stem cells to ensure CNS maintenance and regeneration. To determine which genes are regulated by COE, we compared the transcriptome profiles of control and *coe(RNAi)* animals and identified downregulated genes involved in CNS function. We validated candidate targets by testing for loss of expression after *coe* knockdown and visualizing their expression in *coe*$^{+}$ cells. To examine how defects in COE contribute to CNS dysfunction, we knocked down a subset of downregulated transcripts and found additional genes required for CNS regeneration including homologs of the transcription factors NKX2 and POU4.

Finally, we mined our RNA-seq dataset for candidate COE targets expressed in stem cell progeny and identified novel *postmitotic progeny* genes required for stem cell homeostasis. Consistent with these findings, we showed that *coe* is also required for stem cell homeostasis and cell death. Combined, these data identify new targets of COE in mature neurons and differentating stem cells, providing novel insights into how COE regulates nervous system form, function, and regeneration.

METHODS

Animal husbandry

Asexual *Schmidtea mediterranea* (CIW4) were fed homogenized calf liver and maintained in Instant Ocean salts (51). Animals 2 -5 mm in length were starved for one week prior to experimentation.

RNA interference

Animals were administered six feedings of bacterially expressed dsRNA complementary to the indicated gene over three weeks as previously described (52); *gfp* dsRNA was fed as a control. For regeneration experiments, planarians were amputated pre- and post-pharyngeally 24 hours following the final dsRNA feeding.

Wholemount *in situ* hybridization and immunostaining

Animals were processed for *in situ* as described in (32); γ-irradiation treatments and detection of fluorescent *in situ* hybridization experiments were performed as described in (23). Accession numbers for the sequences used in this study are listed in Appendixes 1-2. For immunostaining with anti-PH3 (1:1000, D2C8, Cell Signaling), anti-SYNORF1 (1:400, 3C11, DSHB), or anti-VC-1 (1:10,000; kindly provided by Hidefumi Orii) animals were fixed with Carnoy's solution (53); with anti-CRMP-2 (1:50, 9393S, Cell Signaling) or anti-β-tubulin (E7, 1:1000, DSHB)

animals were fixed with formaldehyde, processed, and labeled using TSA (32) except that the reduction step was omitted.

DAVID Analysis

We determined human accession numbers for differentially expressed *Schmidtea mediterranea* transcripts by performing BLASTX against the human UniProt database (cutoff < 1 X 10^{-4}). Human accession numbers were used to assigned Gene Ontology terms and perform clustering analysis using DAVID software (1, 2) with the "Panther_BP_all" and "Panther_MF_all" gene annotation settings and Enrichment Score cutoff >1.3.

Quantitative Real-Time Polymerase Chain Reaction

RNA was extracted with Trizol reagent (Invitrogen), DNAsed using Turbo DNA-free Kit (Life Technologies) and purified using RNeasy MinElute Cleanup kit (Qiagen). cDNA was synthesized using the iScript cDNA Synthesis Kit (BioRad). qPCR was performed on a Bio-Rad CFX Connect Real-Time System using SsoAdvanced SYBR Green Supermix (Bio-Rad) with a two-step cycling protocol and annealing/extension temperature of 58.5°C. Two technical replicate PCRs were performed on each of two separate control and RNAi sample sets. The relative amount of each cDNA target was normalized to that of *Smed*-b-*tubulin* (accession no. DN305397). Two or three biological replicates were used for each experiment. Primers are listed in Table S6.

Flow Cytometry

One week following the last RNAi treatment, animals were dissociated using papain and stained with 25 mg/ml of Hoescst 33342 (Life Technologies) as previously described (3). Samples were analyzed using a FACSAria (BD Biosciences) Cell Sorter. For irradiation samples, control animals were exposed to 100 Gy of g-irradiation 6 days prior to dissociation.

Statistical analysis

Statistical analysis was performed using Student's *t*-test. Error bars in graphs are standard deviations.

RESULTS AND DISSCUSION

coe is required for maintenance of nervous system structure.

Analysis of *coe* using an optimized wholemount *in situ* hybridization protocol (WISH) (32) confirmed that *coe* mRNA was restricted to stem cells and neurons of *S. mediterranea* (Figure. 2.1A-C). We investigated the role of *coe* in head and tail regeneration by amputating *coe(RNAi)* and control animals pre- and post-pharyngeally and analyzing the trunk fragment. The brains regenerated by *coe(RNAi)* animals were smaller than those of controls and failed to connect at the anterior commissure (Figure. 2.1D) (23). Additionally, we observed ventral nerve cord (VNC) defects at posterior facing wounds (Figure 2.1E). We further analyzed the nervous system patterning defects using anti-VC-1, a marker of the photoreceptor neurons and their axons, and found that the optic chiasm also failed to form in *coe(RNAi)* animals (Figure 2.1F). These data show that *coe* is required for neuronal regeneration at both anterior and posterior facing wounds and suggest that *coe* regulates genes required for reestablishing midline patterning following amputation.

Next, we investigated the role of *coe* in tissue homeostasis by probing for markers that label the CNS, intestine, and muscle, using *ChAT*, *mat* (22), and *collagen* (33), respectively. We found a systemic loss of *ChAT*$^{+}$ expression in *coe(RNAi)* animals (Figure 2.2A) (23). By contrast, the expression of markers of the intestine and muscle were similar to those in controls (Figure 2.2B-C). To determine whether COE function is required for maintenance of nervous system architecture, we

labeled neuronal cell bodies and their projections using anti-CRMP-2 and anti-β-tubulin (Figure 2.2D-F). In *coe(RNAi)* animals, we observed a striking decrease in nerve bundles labeled by anti-CRMP-2 and anti-β-tubulin relative to that in controls; however, expression of CRMP-2 was retained in the cell bodies (Figure 2.2F). In addition, when we labeled sensory neurons using *cintillo* (34), *coe(RNAi)* animals exhibited significantly fewer *cintillo*$^{+}$ cells (Figure 2.2G). These results strongly suggest that nervous system architecture is severely reduced or lost in the absence of *coe.*

Identification of COE targets in the adult nervous system.

In *Drosophila* and *C. elegans*, COE homologs are required to maintain molecular identity in a subset of cholinergic or sensory neurons in adult animals (12-14). Although COE has been shown to drive differentiation of multiple classes of neurons during development (35), the role of COE in adult nervous system function or regeneration remains largely unknown. Thus, we took advantage of the requirement of *coe* for planarian CNS homeostasis to investigate downstream genetic programs controlled by this transcription factor. By comparing RNA-seq profiles of control and *coe(RNAi)* animals, we identified 909 differentially expressed genes; 397 were downregulated and 512 were upregulated (Appendix 3). Functional annotation using DAVID software showed that the set of downregulated genes was significantly enriched for Gene Ontology (GO) terms associated with 'ion channel', 'neuronal activities', 'nerve-nerve synaptic transmission', 'voltage-gated

ion channel', and 'cell adhesion molecule'; by contrast, the upregulated genes were enriched for GO terms associated with 'cytoskeletal protein' and 'muscle development' (Table 2.1). We selected 66 downregulated genes associated with neural functions or annotated as transcription factor homologs and performed WISH analysis to determine their tissue-specific pattern of expression (Appendix 4). As we expected, the most prominent mRNA expression pattern of genes downregulated following *coe* RNAi was in the nervous system (22/66 genes; Appendix 4) such as *ChAT* and *cpp-1*, which we had previously found to be putative downstream targets of COE (23). In addition, we observed genes that were expressed broadly in the nervous system (such as *neural cell adhesion molecule-2* (*ncam-2*), *vesicle-associated membrane protein like-1* (*vamp*), *gamma-aminobutyric acid receptor subunit beta like-1* (*gbrb-1*), *voltage-gated sodium channel alpha-1* (*scna-1*)) or in discrete neuronal subpopulations (such as *secreted peptide prohormone-19, -18, -2* (*spp-19, -18, -2*), *neuropeptide like-1* (*npl-1*), *voltage-gated sodium channel alpha-2*, (*scna-2*), and *caveolin-1* (*cav-1*)) (Figure 2.3A-J). We also identified several transcripts that labeled subsets of neurons in the brain, including *netrin-1* (36) and the transcription factors *iroquios-1* (*irx-1*) and *pou class 4 transcription factor 4 like-1* (*pou4l-1*) (Figure 2.3K-M).

Next, based on strong expression in discrete cell populations, we assessed 34/66 genes for changes in expression following *coe* RNAi knockdown. All 34 genes that we tested agreed with our RNA-seq dataset (Figure 2.3A'-L' and appendix 4). *coe* knockdown caused a systemic loss of expression of all genes evaluated

throughout the animal, with the exception of *scna-2* and *cav-1*, which were specifically lost in cells located at the midline (Figure 2.3A-N). These results suggest that *coe* function is involved in maintaining brain patterning in uninjured animals. Because many of the transcription factors were weakly expressed and difficult to detect by WISH (Figure 2.4A), we investigated expression changes in *coe(RNAi)* animals using quantitative real-time PCR (qPCR). The changes in gene expression detected by qPCR were in concordance with our RNA-seq dataset (Figure 2.4B). All of the genes identified above were detected in *coe*$^{+}$ cells with the exception of *scna-2*, *cav-1*, *irx-1*, *pou4l-1*, and *scna-2*, which we were unable to detect reliably by FISH (Figure 2.5). These experiments identified 10 previously unknown targets of *coe* in the nervous system, including genes important for maintaining neuronal subtype identity (e.g., ion channels and neuropeptides). The results obtained from RNA-seq experiments also raised the possibility that *coe* might negatively regulate mesoderm specification (Table 2.1), which is required for muscle development (3, 37). However, *coe* mRNAs were not detected in a muscle pattern (Figure 2.1) (23) nor did we detect obvious phenotypes associated with muscle differentiation (Figure 2.2). Although our experiments do not rule in or out a function of COE in mesodermal differentiation or maintenance, our data do clearly indicate that *coe* is required for expression of nervous system-specific genes in adult planarians.

***nkx2l-1* and *pou4l-1* expression is required for CNS regeneration.**

To gain insights into how the loss of COE function affects nervous system regeneration, we analyzed the role of nine downregulated genes that were expressed in neurons or predicted to encode transcription factors. Following RNAi, animals amputated pre- and post-pharyngeally were stained with anti-SYNAPSIN after 10 days of regeneration to screen for defects in CNS architecture (Table 2.2). The efficiency of inhibition of expression of selected genes was confirmed using qPCR (Figure 2.6A). We found that knockdown of *scna-2*, *pou4l-1*, and *nkx2l* caused defects in CNS regeneration. *scna-2(RNAi)* animals had little eye pigmentation and rarely developed a single photoreceptor (Figure 2.6B-C). These observations are consistent with the idea that ion channels play critical roles in CNS development and regeneration (38-40). *nkx2l(RNAi)* animals exhibited photoreceptor defects and had significantly reduced tail blastemas (Figure 2.6D and F), whereas *pou4l-1(RNAi)* animals had less photoreceptor pigment and smaller brains than controls (Figure 2.6E). We further examined the function of *nkx2l* and *pou4l-1* by staining RNAi-treated regenerates for *coe* targets, including *ChAT* and *npl*; *ChAT*$^{+}$ brains were smaller in *nkx2l(RNAi)* and *pou4l-1(RNAi)* animals compared to controls (Figure 2.6G). Interestingly, despite their smaller brains, *pou4l-1(RNAi*) animals regenerated 22% more *npl*$^{+}$ cells than controls (Figure 2.6H), suggesting that this gene plays a role in controlling neuronal fate specification.

It is noteworthy that several transcription factors that we identified in our screen are putative COE targets in *Xenopus* development, including *irx-1*, *tal*, *pou4l-*

1, and *nkx2l* (35). Of these genes, we found that expression of *pou4l-1* and *nkx2l* was important for CNS regeneration. NKX and POU orthologs play critical roles during CNS development of invertebrate and vertebrate organisms (41-43). These data demonstrate that the regulatory program downstream of COE is conserved throughout evolution and functions to control adult neurogenesis and CNS regeneration.

Novel postmitotic progeny genes are required for planarian stem cell differentiation.

We previously observed *coe* expression in a subset of stem cells and their progeny (23). To gain insights into the mechanisms by which *coe* may regulate functions of adult somatic stem cells, we compared our *coe(RNAi)* transcriptome dataset to stem cell and progenitor transcriptomes (44, 45) and identified 10 genes reported to be highly expressed in stem cell progeny: six genes were downregulated and four upregulated in *coe(RNAi)* animals compared to the controls (Figure 2.7A-B and Appendix 4). None of these 10 genes shared significant sequence homology with known genes; however, analysis of the predicted protein sequences using InterProScan (46) revealed that all had N-terminal signal peptide sequences. To confirm that these genes were expressed in differentiating cells, we exposed animals to γ-irradiation, a treatment that ablates stem cells and progenitors (47, 48), and observed a dramatic loss of expression for these transcripts (Figure 2.7C-D). The genes identified did not co-label with *h2b* (data not shown), confirming that these

transcripts are primarily expressed in differentiating stem cells. We termed these 10 genes *postmitotic progeny* (*pmp*) genes.

We next asked whether *pmp* genes are expressed in a homogeneous cell population by staining animals for downregulated (*pmp-3/pmp-4* or *pmp-4/pmp-2*) or upregulated (*pmp-7/pmp-8* or *pmp-9/pmp-8*) candidates; there was 99% and 98% overlap of expression, respectively (representative results shown in Figure 2.7E-F). Similarly, when animals were co-stained with *pmp-4* (downregulated) and *pmp-8* (upregulated), 93% of *pmp-4*$^+$ cells were also *pmp-8*$^+$ (Figure2.7G). Similar proportions of *pmp-4*$^+$ and *pmp-8*$^+$ cells expressed the late progeny marker *agat-1* (90%) (47), *coe* (11%), and *ChAT* (13-15%; Figure 2.7H-L). We rarely detected co-expression of *pmp-4* with the early progeny marker, *prog-2* (data not shown) (47). From these observations we conclude that *pmp* genes are abundantly expressed in late progenitors and that their expression is maintained in differentiating neurons.

Expression analysis suggested that *pmp* genes might be required to maintain progenitor identity or terminal differentiation. Planarian regeneration relies on the stem cell population and serves as an excellent readout to test the function of stem cell regulatory genes. When we inhibited expression of downregulated *pmp* genes using RNAi, we noted that *pmp-1(RNAi)* and *pmp-2(RNAi)* animals had significantly increased numbers of PH3$^+$ cells in trunk fragments (Figure 2.7M-N). We also observed animal death within a few days following amputation in 9/17 *pmp-1(RNAi)* head fragments, indicating the stem cells failed to restore lost cell types (Table 2.2). By contrast, we did not observe a significant effect on the mitotic index

in *pmp-5(RNAi)* animals; however, these animals had delayed photoreceptor regeneration and smaller brains than controls (Table 2.2). Our data demonstrate that *coe* is required for normal expression of *pmp* genes or functions in concert with a subset of these genes during stem cell differentiation.

COE regulates stem cell homeostasis.

COE homologs have been implicated in tumor suppression as COE promotes cell cycle exit and death (15). Based on our observations that *pmp-1(RNAi)* and *pmp-2(RNAi)* animals had more $PH3^+$ cells than controls, we examined the effect of *coe* silencing on stem cell homeostasis and cell survival. Similar to *pmp-1(RNAi)* and *pmp-2(RNAi)* animals, *coe(RNAi)* animals exhibited an increase in the number of $PH3^+$ cells compared to controls (Figure 2.8 A-C). Interestingly, we also found that the number of $TUNEL^+$ cells was diminished in *coe(RNAi)* animals (Figure 2.8D-F). We hypothesized that inhibition of *coe* expression affects cell cycle progression dynamics. To test this, we attempted to label proliferating cells using BrdU; however, *coe(RNAi)* animals were refractory to staining. As an alternative method, we labeled cycling stem cells using *h2b*, which, as we expected, labeled nearly 100% of cells in S-phase (Figure 2.8G). In *coe(RNAi)* animals, we observed a significant loss of $h2b^+$ cells between the cephalic ganglia and the ventral nerve cords (Figure 2.8H-J). We used flow cytometry to estimate the total number of stem cells and progeny in control and *coe(RNAi)* animals. We found that the number of stem cells and their progeny were not significantly different (Figure 2.8K-M). This result is consistent

with our previous observation that only a small proportion of proliferating stem cells expressed *coe* (approximately 4-7%) (23). We speculate that loss of *coe* function perturbs the ability of a subset of stem cells to exit the cell cycle or undergo normal cell death as occurs in cancer (15). It will be necessary to determine the exact identity of *coe*$^{+}$ stem cells in order to investigate whether defects in stem cell function are caused by stem cell senescence, death, or aberrant differentiation. Notwithstanding, our data support the idea that tumor suppressor genes play key roles in stem cell-based regeneration (49, 50) and further accentuate the utility of the planarian model to gain insights into how tumor suppressors regulate adult stem cells within the context of specific organs or tissues.

CONCLUSIONS

COE proteins are known to function as terminal selectors of neuronal identity in adult organisms (12-14); yet, it remains unclear how defects in COE contribute to nervous system dysfunction and disease. In this study we exploited the high rate of tissue turnover and regenerative capacity of planarians to expand our understanding of how COE functions in adult stem cells and neurons (Figure 2.9). We found that *coe* is essential to maintain nervous system architecture and drive the expression of genes important for neuronal identity (such as neurotransmitter receptors, ion channels, and neuropeptides) in multiple classes of neurons distributed throughout the CNS. By examining the function of COE targets, we identified several genes important for CNS regeneration including *scna-2*, *nkx2l-1*, and *pou4l-1*. Remarkably, five out of the seven transcription factors we analyzed in this study (*irx-1*, *nkx2l-1*, *pou4l-1*, *tal*, and *tlx*), were also identified as putative COE targets during *Xenopus* CNS development, suggesting that genetic programs regulated by COE are highly conserved during neurogenesis and redeployed during CNS turnover and regeneration. Finally, we demonstrated that COE is required for stem cell homeostasis and normal cell death, functions that may be controlled in part by COE-dependent regulation of *pmp* genes in differentiating stem cells. These data strongly suggest that the role of COE in tumor suppression is conserved in planarians (15). However, we do not know the precise mechanisms by which *coe* inhibition causes systemic defects in proliferation, cell death or *pmp* gene expression. The next step will be to perform ChIP-seq and combine it with our

expression data to elucidate which genes are direct targets of COE genome-wide. In conclusion, our study underscores the importance of COE genes in CNS maintenance and demonstrates that planarians are a powerful animal model to identify COE targets *in vivo* and examine their function in neuronal turnover and repair.

Chapter 2, in full, has been submitted for publication of the material. Cowles, MW.; Stanley, BN.; Omuro, KC.; Quintanilla CG.; Zayas, RM. The dissertation author was the primary investigator and author of this manuscript.

FIGURES

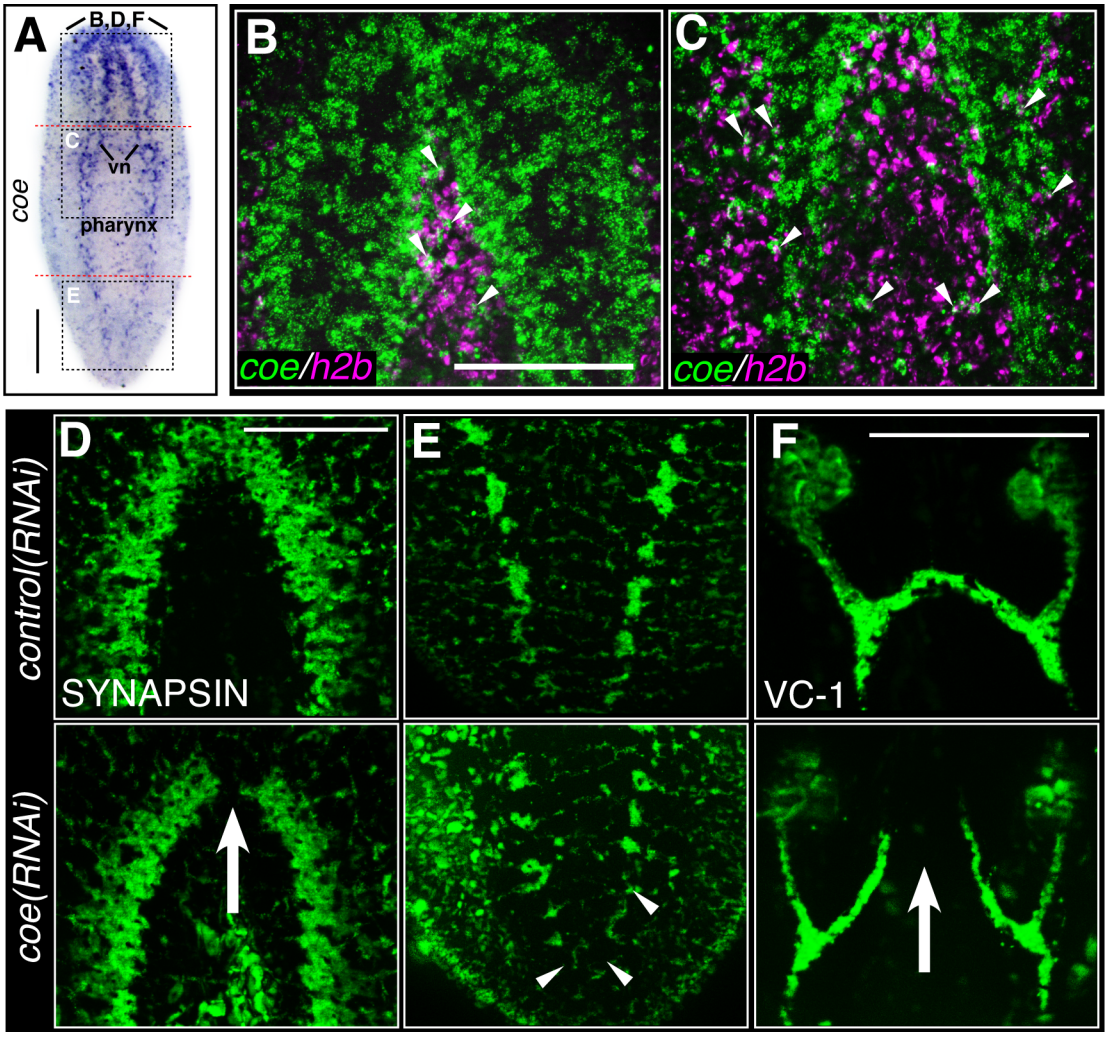


Figure. 2.1: *coe* is required for CNS regeneration and midline patterning. (A) *In situ* hybridization to *coe* in *S. mediterranea* (vn, ventral nerve cords). Dashed boxes show regions imaged in B-F; red dashed lines denote amputation sites for animals shown in D-F. (B-C) Double fluorescent in situ hybridization to *coe* and *h2b.* Regions imaged are shown in A. Arrows point to double-labeled cells (D-F) Control and *coe(RNAi)* 7-day regenerates immunostained with anti-SYNAPSIN (D, E) or labeled with anti-VC-1 (F). Arrows in D and F denote defects in anterior commissure and optic nerve patterning, respectively. Scale bars in A = 200 μm; those in B, D, F = 100 μm

Figure. 2.2: *coe* is required for nervous system maintenance.
(A-C) RNAi-treated animals were processed for fluorescent in situ hybridization to *ChAT, mat*, or *collagen*. Dashed boxes in A indicate regions imaged at higher magnification shown in the insets to the right. Dashed box in C denotes region shown at higher magnification in inset. (D) Head or tail images from an animal stained with anti-CRMP-2 and *ChAT*. CRMP-2 is expressed in axon projections and neuronal cell bodies. (E) High magnification image of region denoted by white box in F shows CRMP-2 is detected in *ChAT*$^{+}$ cell bodies (Arrowhead). Nuclei were labeled in blue using DAPI. (F) Uninjured control and *coe(RNAi)* planarians immunolabeled with anti-CRMP-2 and anti-β-TUBULIN. (G) Control and *coe(RNAi)* animals processed for WISH to *cintillo*. Quantification of *cintillo*$^{+}$ cells Shown in top right corner. Scale bar in A = 200 μm; Scale bar in D = 100μm; Scale bar in E = 50 μm.

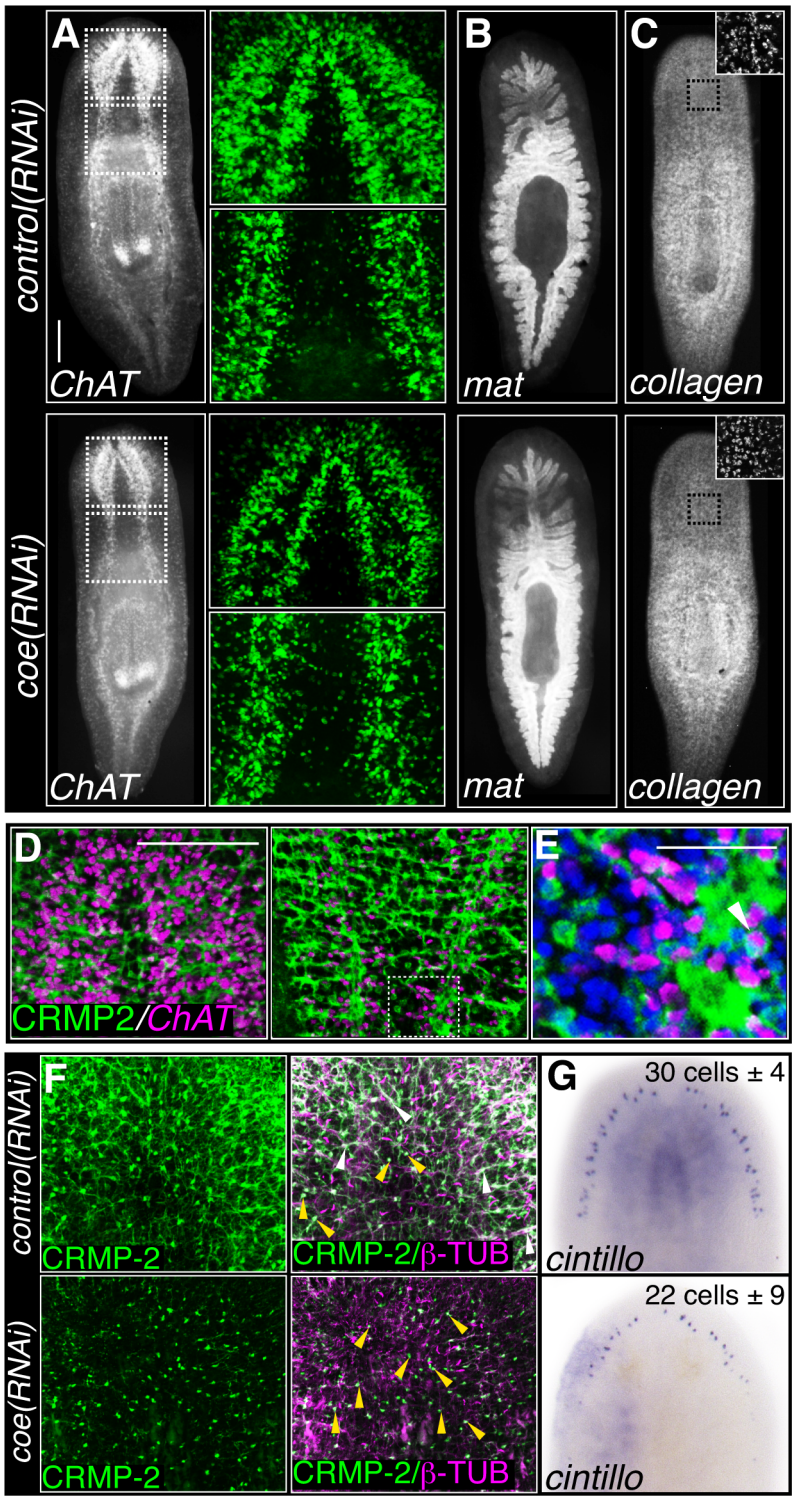
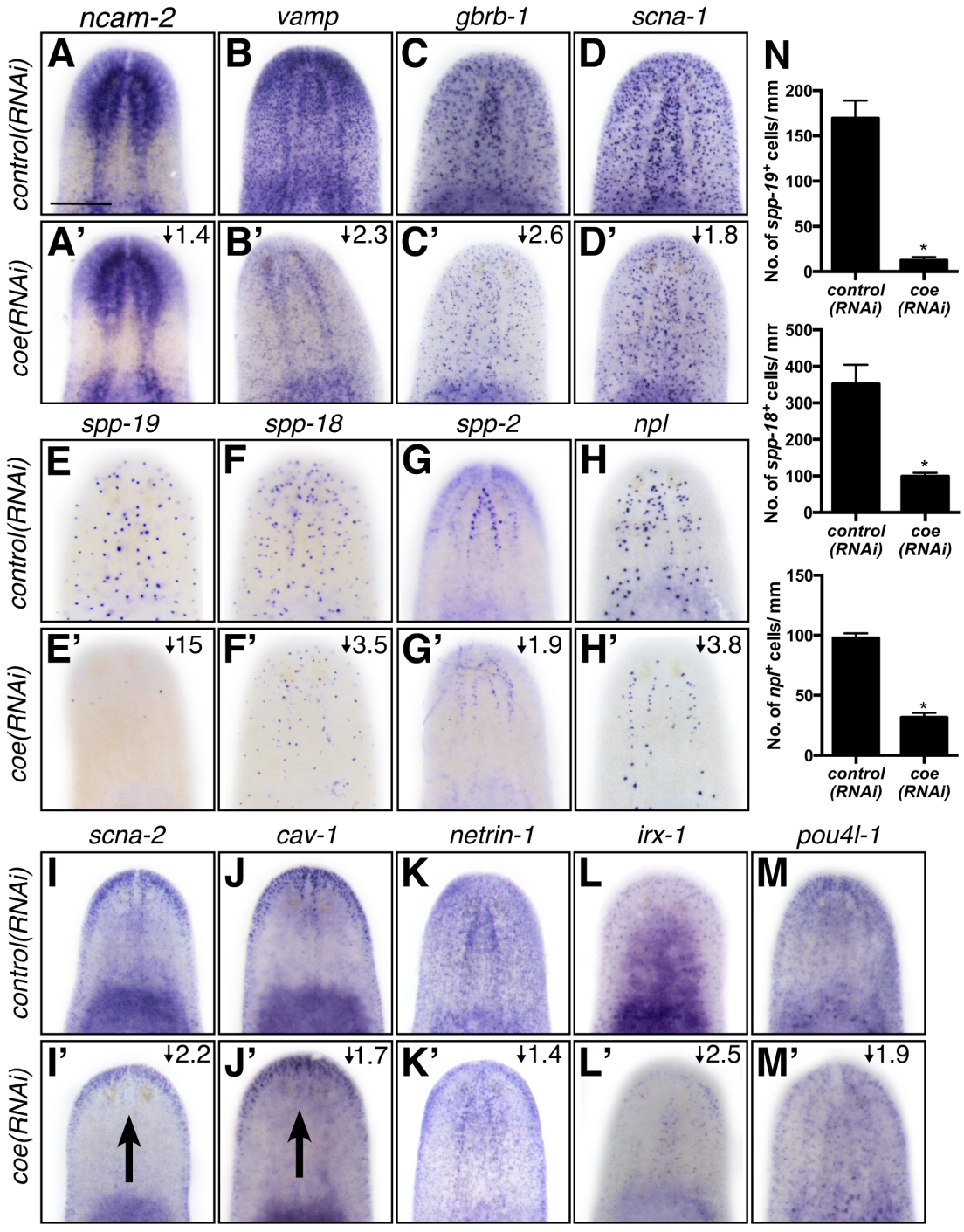

A
B
C
control(RNAi)
ChAT
mat
collagen
coe(RNAi)
ChAT
mat
collagen
D
CRMP2/ChAT
E
F
control(RNAi)
CRMP-2
CRMP-2/β-TUB
G
30 cells ± 4
cintillo
coe(RNAi)
CRMP-2
CRMP-2/β-TUB
22 cells ± 9
cintillo

Figure. 2.3: Identification of downstream targets of COE in the central nervous system.
(A-M) Control and *coe(RNAi)* treated animals were evaluated by WISH for genes denoted above each panel. Numbers in top right corner indicate fold changes in mRNA expression between control and *coe(RNAi)* animals. Arrows in I and J point to loss of expression at the midline. (N) Quantification of *spp19*$^{+}$, *spp-18*$^{+}$, and *npl*$^{+}$ cells from animals shown in E, F, and H. Scale bars = 100 μm


ncam-2
vamp
gbrb-1
scna-1
control(RNAi)
coe(RNAi)
A
B
C
D
A'
↓1.4
B'
↓2.3
C'
↓2.6
D'
↓1.8
N
No. of spp-19+ cells/ mm
200
150
100
50
0
control (RNAi)
coe (RNAi)
spp-19
spp-18
spp-2
npl
E
F
G
H
E'
↓15
F'
↓3.5
G'
↓1.9
H'
↓3.8
No. of spp-18+ cells/ mm
500
400
300
200
100
0
No. of npl+ cells/ mm
150
100
50
0
scna-2
cav-1
netrin-1
irx-1
pou4l-1
I
J
K
L
M
I'
↓2.2
J'
↓1.7
K'
↓1.4
L'
↓2.5
M'
↓1.9

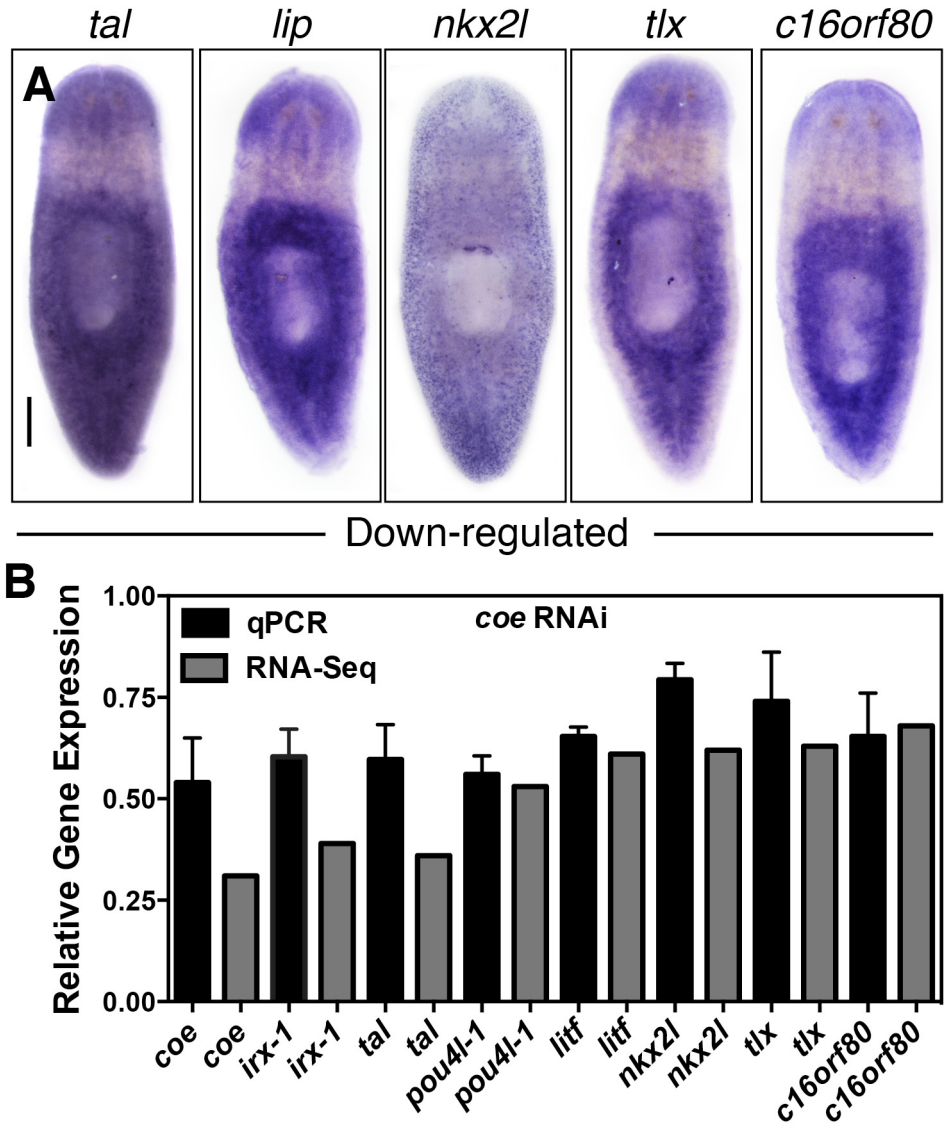


Figure. 2.4: Expression analysis of transcription factors differentially regulated following coe gene silencing.
(A) WISH to *tal*, *lip*, *nkx2l*, *tlx*, and *c16orf80* demonstrated that these transcription factors are expressed at low levels throughout the animal. (B) Quantification of changes in gene expression of selected transcription factors using RNA-sequencing or quantitative real-time PCR. Scale bar = 200 µm.

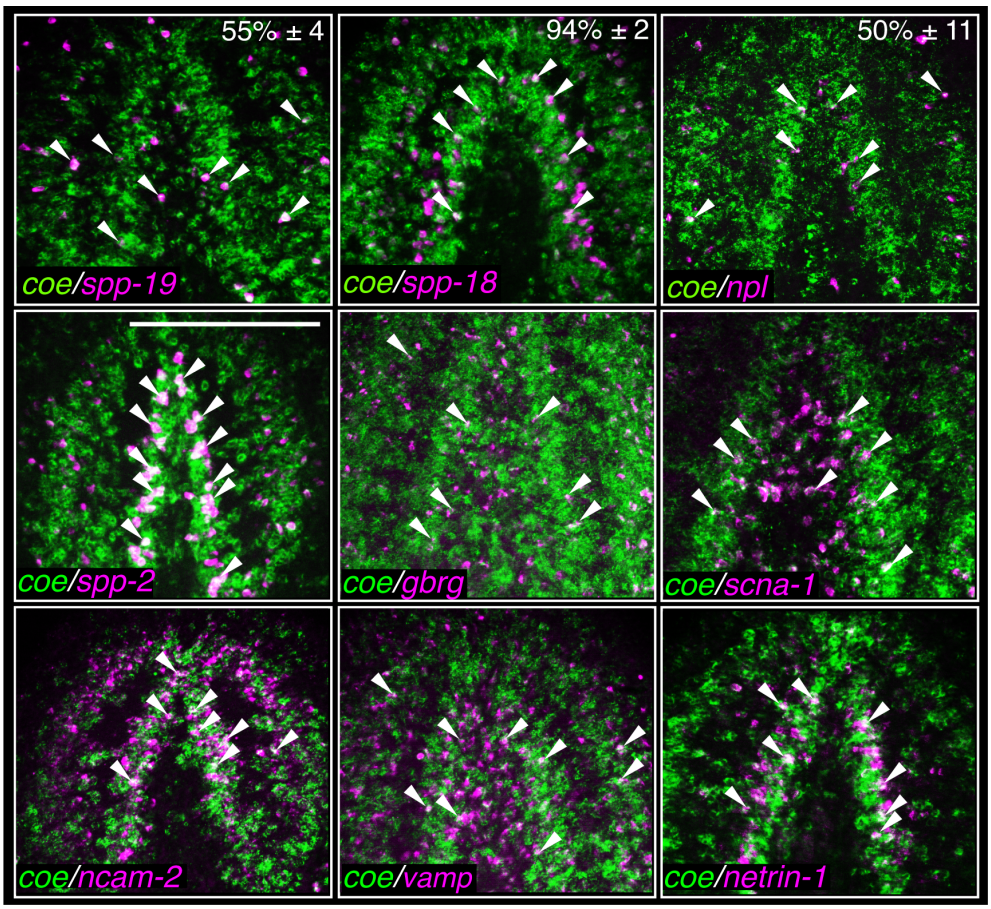


Figure 2.5: Double-labeling of candidate downstream targets with *coe.* Fluorescent *in situ* hybridization analysis of *coe* and *spp-19*, *spp18*, *npl, spp-2, gbrb, scna-1, ncam-2, vamp,* and *netrin-1*. Numbers indicate the proportion of cells that were also *coe*$^{+}$. Arrowheads denote double-labeled cells. Scale bar = 200 μm.

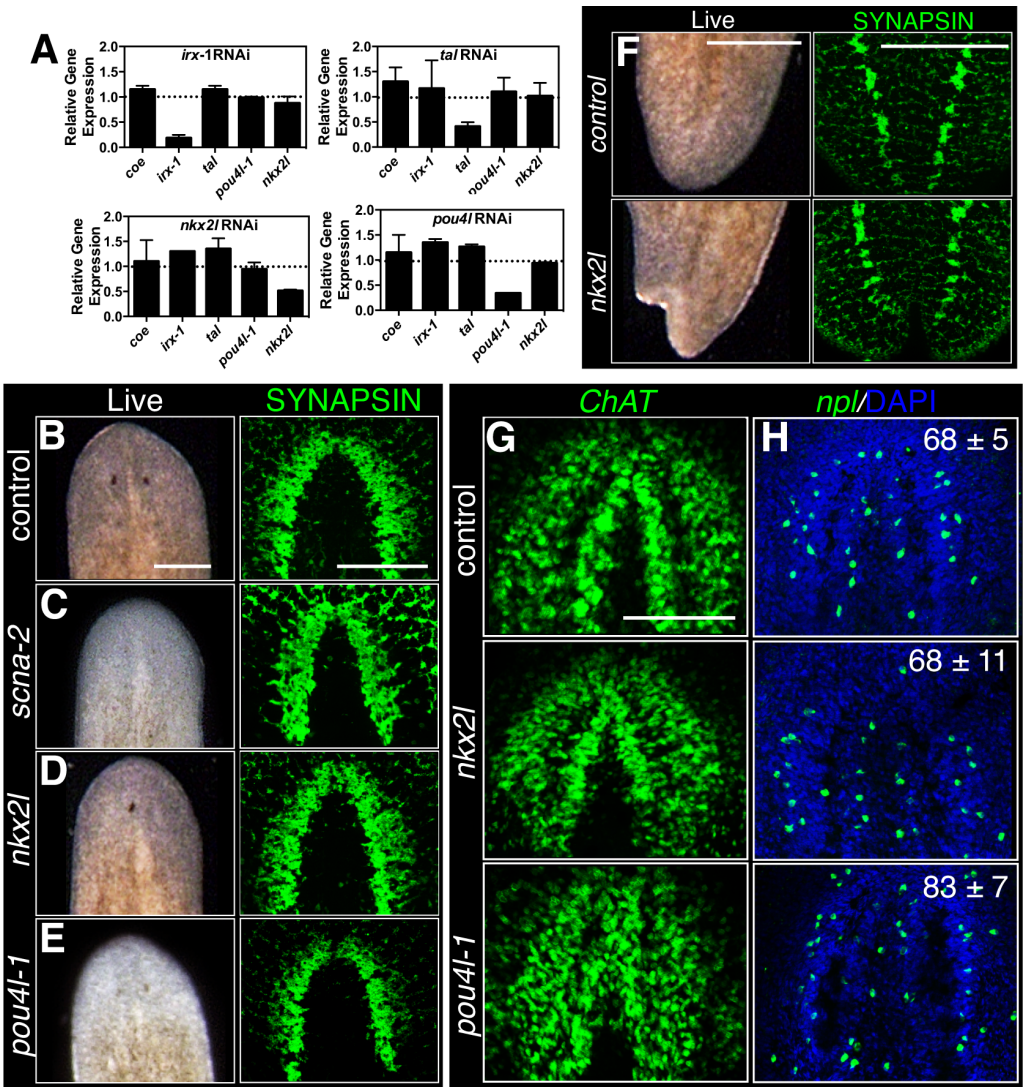


Figure 2.6: Characterization of downstream targets of COE involved in tissue regeneration.

(A) Quantification of relative mRNA levels of selected transcription factors following RNAi treatment. (B-H) Animals were fed dsRNA targeting the indicated gene. Ten-day regenerates were imaged live and immunostained with anti-SYNAPSIN, or processed for fluorescent *in* situ hybridization to *ChAT* or *npl*. Panel F shows posterior wound. Counterstain with DAPI in panel H was used to visualize brain morphology. The numbers of *npl*$^{+}$ cells in the brains are given in panel H. Scale bars = 100 μ

Figure 2.7: COE regulates the expression of novel progenitor genes.
(A-D) *control(RNAi)*, *coe(RNAi)*, or γ-irradiation-treated animals were processed for WISH to evaluate expression of *pmp* genes denoted above each panel. Numbers in top right corner indicate mRNA expression fold changes in *coe(RNAi)* animals relative to controls. (E-F) Double-fluorescent *in situ* hybridization using riboprobe pairs for *pmp* genes that were downregulated (*pmp-4/pmp-2*) or upregulated (*pmp-9/pmp-8*), or both (*pmp-4/pmp-8*), and for *coe/pmp-4*, coe/pmp-8, *ChAT/pmp-4, or ChAT/pmp-8*. Numbers in E-L show the proportion of double-positive cells. Arrowheads in I-L point to double-labeled cells. (M) RNAi-treated regenerates immunostained with anti-PH3. (N) Quantification of $PH3^+$ cells from animals shown in J. *P < 0.05. Scale bars = 100 μm.

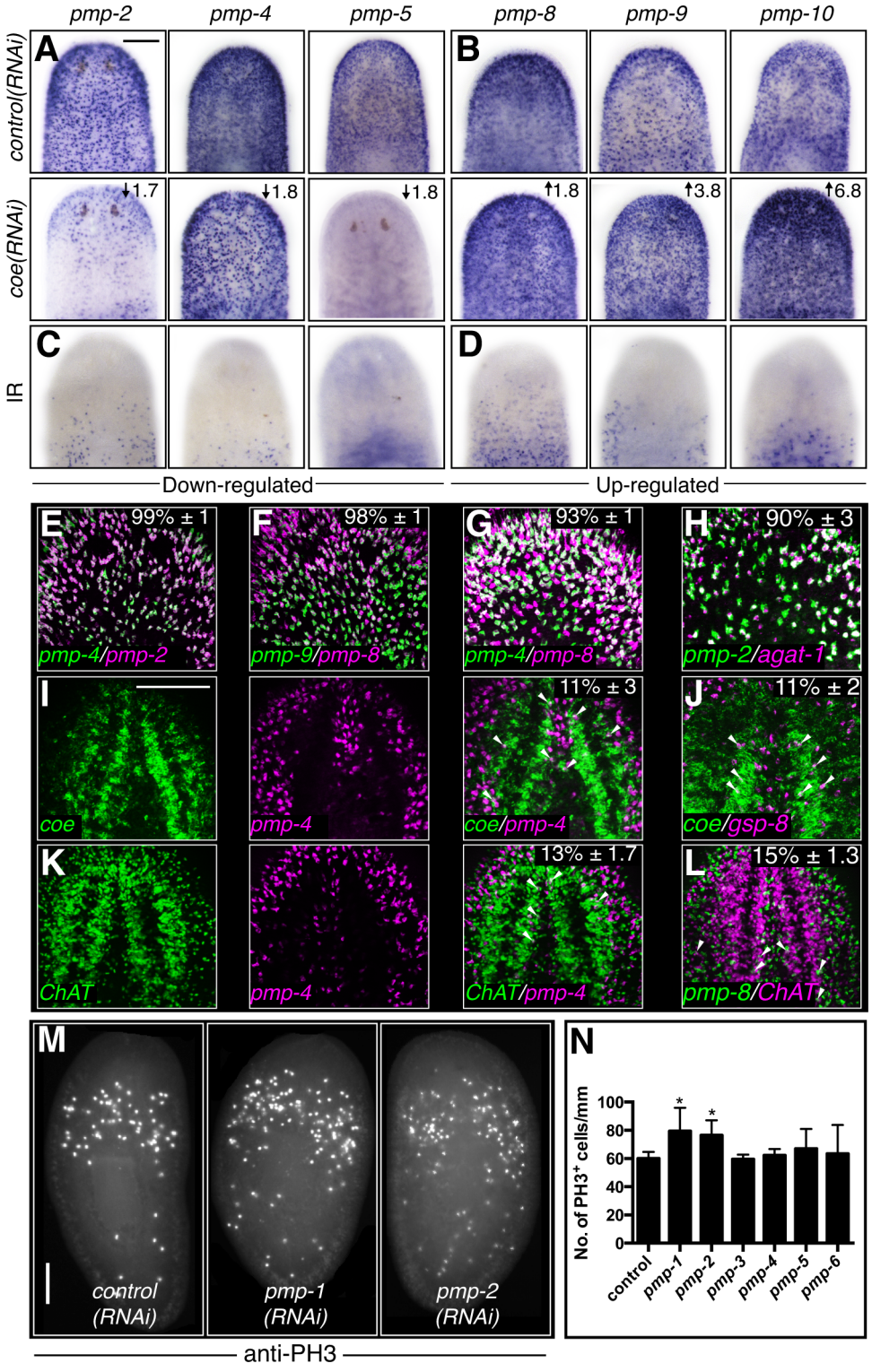
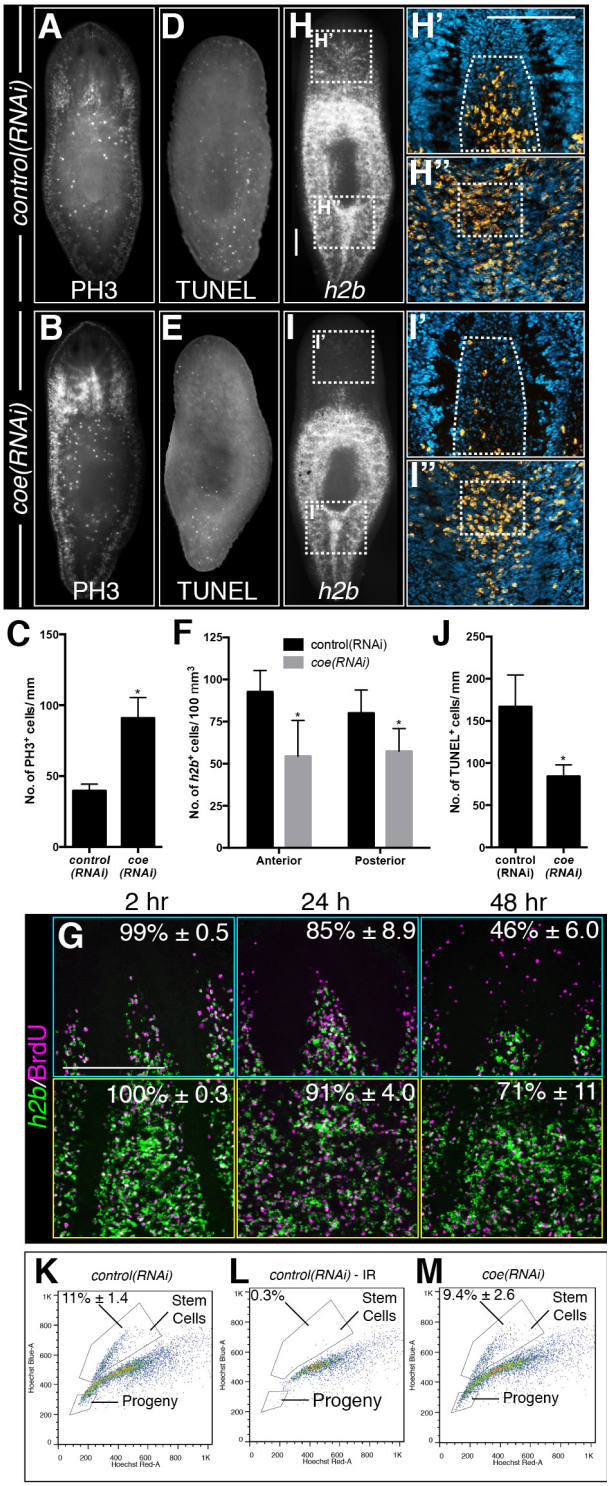

pmp-2
pmp-4
pmp-5
pmp-8
pmp-9
pmp-10
control(RNAi)
coe(RNAi)
IR
A
B
C
D
↓1.7
↓1.8
↓1.8
↑1.8
↑3.8
↑6.8
Down-regulated
Up-regulated
E
99% ± 1
pmp-4/pmp-2
F
98% ± 1
pmp-9/pmp-8
G
93% ± 1
pmp-4/pmp-8
H
90% ± 3
pmp-2/agat-1
I
coe
pmp-4
11% ± 3
coe/pmp-4
J
11% ± 2
coe/gsp-8
K
ChAT
pmp-4
13% ± 1.7
ChAT/pmp-4
L
15% ± 1.3
pmp-8/ChAT
M
control (RNAi)
pmp-1 (RNAi)
pmp-2 (RNAi)
anti-PH3
N
No. of PH3+ cells/mm
100
80
60
40
20
0
control
pmp-1
pmp-2
pmp-3
pmp-4
pmp-5
pmp-6

Figure 2.8: COE regulates stem cell homeostasis.
(A-F) RNAi treated animals were stained with anti-PH3 or processed for TUNEL. Quantification of $PH3^+$ and $TUNEL^+$ cells are shown in C and F, respectively. (G) Planarians were soaked with BrdU, chased for two, 24, or 48 hours, labeled with anti-BrdU as described previously (11) and processed for fluorescent in situ hybridization to *h2b*. Percentages in F show the number $h2b^+$ cells that were also labeled with BrdU. (H-J) RNAi animals were processed for FISH to *h2b*.White boxes in K and L denote areas imaged at higher magnification shown in H'-I". (H'-I") Dashed regions show representative fields used for cell counts summarized in M (n = 5). (K-M) Control and *coe(RNAi)* animals were dissociated and cell suspensions were labeled with Hoechst dye and analyzed using flow cytometry. *control(RNAi)* IR planarians were exposed to 100 Gy γ-irradiation six days prior to dissociation to ablate both the stem cells and their progeny. Percentages are averaged from two independent experiments; standard deviation is shown. Scale bars = 100 μm.


control(RNAi)
coe(RNAi)
A
B
D
E
H
I
H'
H''
I'
I''
PH3
TUNEL
h2b
C
No. of PH3+ cells/ mm
control (RNAi)
coe (RNAi)
F
control(RNAi)
coe(RNAi)
No. of h2b+ cells/ 100 mm3
Anterior
Posterior
J
No. of TUNEL+ cells/ mm
2 hr
24 h
48 hr
G
99% ± 0.5
85% ± 8.9
46% ± 6.0
100% ± 0.3
91% ± 4.0
71% ± 11
h2b/BrdU
K
control(RNAi)
11% ± 1.4
Stem Cells
Progeny
L
control(RNAi) - IR
0.3%
M
coe(RNAi)
9.4% ± 2.6
Hoechst Blue-A
Hoechst Red-A

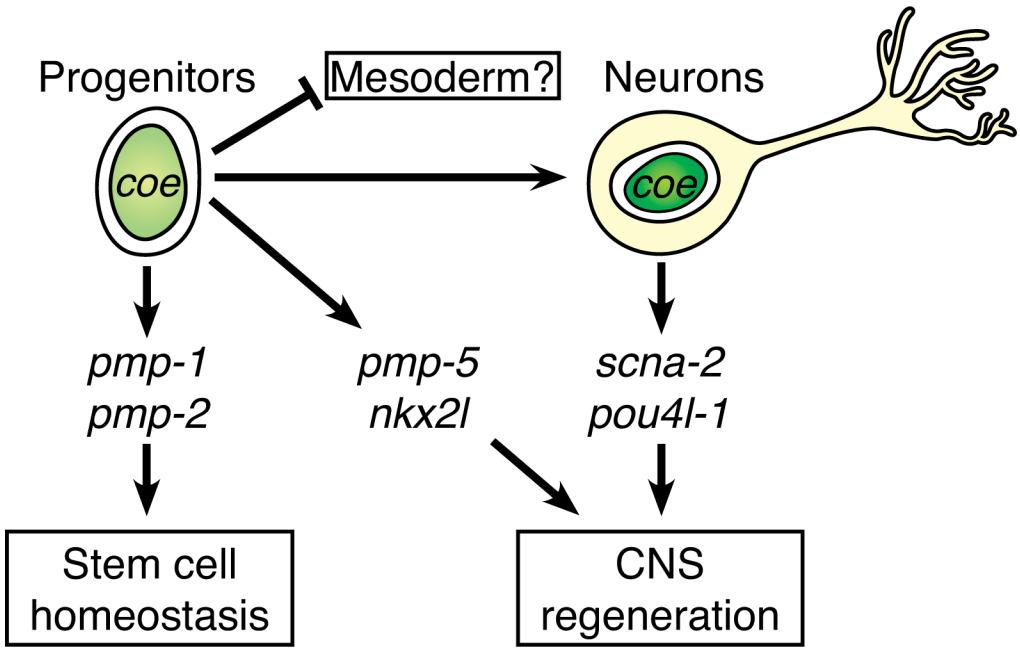


Figure 2.9: Model of COE regulation in planarians.
coe is expressed in lineage-committed progenitors (23), *pmp*$^{+}$ progenitor cells, and diverse classes of mature neurons. To gain insights into how loss of COE function contributes to defects in stem cell and nervous system function, we analyzed the function of genes that were downregulated in *coe(RNAi)* animals relative to controls. These analyses identified genes critical for CNS regeneration (*pmp-5*, *nkx2l*, *scna-2*, and *pou4l-1*) and for stem cell homeostasis (*pmp-1* and *pmp-2*). In *coe(RNAi)* animals, we also observed 597 upregulated genes enriched for GO terms associated with muscle development, suggesting that COE may also function to repress the expression of mesoderm-specific genes.

TABLES

Table 2.1: Functional cluster analysis of differentially expressed genes in coe-deficient animals using DAVID software.

Functional Cluster	Enrichment score
Ion channel	7.63
Neuronal activities	4.66
Microtubule binding motor protein	2.44
Nerve-nerve synaptic transmission	2.41
Voltage-gated ion channel	2.28
Cell adhesion molecule	1.4

Functional Cluster	Enrichment score
Cytoskeletal protein	5.28
Muscle development	4.9

Notes:

1. Downregulated and upregulated genes shown in blue and red, respectively.

Table 2.2: Functional analysis of COE targets in CNS tissue regeneration.

Gene Name	RNAi Phenotype
gbrg	No phenotype observed
irx-1	No phenotype observed
nkx2l	Photoreceptor defects (Bowtie;9/35), abnormal CNS morphology (35/35), and blasetma patterning (7/35)
pmp-1	Head lesions (9/19), increased mitosis
pmp-2	Increased mitosis
pmp-3	No phenotype observed
pmp-4	No phenotype observed
pmp-5	Delayed photoreceptor regeneration (12/17), blastema patterning defects at posterior facing wounds (3/17)
pmp-6	No phenotype observed
pou4l-1	Lighter photoreceptors (5/40) and smaller cephalig ganglia (14/20)
scna-1	No phenotype observed
scna-2	Delayed regeneration photorecptor formation and smaller brains (10/27); regenerated a single visible photoreceptor (1/27)
scna-3	No phenotype observed
spc-1	No phenotype observed
vamp-1	Head and trunk fragments failed to move in response to light (10/10); Regeneration normal (10/10)

Notes:

1. Parentheses show the number of animals exhibiting the phenotype out of the number of animals tested. At least 10 animals tested per experimental group.

2. RNAi animals are from two or more independent experiments.

REFERENCES

1. Daburon V*, et al.* (2008) The metazoan history of the COE transcription factors. Selection of a variant HLH motif by mandatory inclusion of a duplicated exon in vertebrates. *BMC Evol Biol* 8:131.

2. Demilly A*, et al.* (2011) Coe genes are expressed in differentiating neurons in the central nervous system of protostomes. *PLoS One* 6(6):e21213.

3. Jackson DJ*, et al.* (2010) Developmental expression of COE across the Metazoa supports a conserved role in neuronal cell-type specification and mesodermal development. *Dev Genes Evol* 220(7-8):221-234.

4. Liberg D, Sigvardsson M, & Akerblad P (2002) The EBF/Olf/Collier family of transcription factors: regulators of differentiation in cells originating from all three embryonal germ layers. *Mol Cell Biol* 22(24):8389-8397.

5. Crozatier M & Vincent A (2008) Control of multidendritic neuron differentiation in Drosophila: the role of Collier. *Dev Biol* 315(1):232-242.

6. Dubois L*, et al.* (1998) XCoe2, a transcription factor of the Col/Olf-1/EBF family involved in the specification of primary neurons in Xenopus. *Curr Biol* 8(4):199-209.

7. Garel S, Garcia-Dominguez M, & Charnay P (2000) Control of the migratory pathway of facial branchiomotor neurones. *Development* 127(24):5297-5307.

8. Jinushi-Nakao S*, et al.* (2007) Knot/Collier and cut control different aspects of dendrite cytoskeleton and synergize to define final arbor shape. *Neuron* 56(6):963-978.

9. Pozzoli O, Bosetti A, Croci L, Consalez GG, & Vetter ML (2001) Xebf3 is a regulator of neuronal differentiation during primary neurogenesis in Xenopus. *Dev Biol* 233(2):495-512.

10. Prasad BC*, et al.* (1998) unc-3, a gene required for axonal guidance in Caenorhabditis elegans, encodes a member of the O/E family of transcription factors. *Development* 125(8):1561-1568.

11. Wightman B, Baran R, & Garriga G (1997) Genes that guide growth cones along the C. elegans ventral nerve cord. *Development* 124(13):2571-2580.

12. Kratsios P, Stolfi A, Levine M, & Hobert O (2012) Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. *Nat Neurosci* 15(2):205-214.

13. Eade KT, Fancher HA, Ridyard MS, & Allan DW (2012) Developmental transcriptional networks are required to maintain neuronal subtype identity in the mature nervous system. *PLoS Genet* 8(2):e1002501.

14. Hattori Y, *et al.* (2013) Sensory-Neuron Subtype-Specific Transcriptional Programs Controlling Dendrite Morphogenesis: Genome-wide Analysis of Abrupt and Knot/Collier. *Dev Cell*.

15. Zhao LY, *et al.* (2006) An EBF3-mediated transcriptional program that induces cell cycle arrest and apoptosis. *Cancer Res* 66(19):9445-9452.

16. Liao D (2009) Emerging roles of the EBF family of transcription factors in tumor suppression. *Mol Cancer Res* 7(12):1893-1901.

17. Maher EA, *et al.* (2006) Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. *Cancer Res* 66(23):11502-11513.

18. Mullighan CG, *et al.* (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 446(7137):758-764.

19. Zardo G, *et al.* (2002) Integrated genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors. *Nat Genet* 32(3):453-458.

20. Zhu Z & Huangfu D (2013) Human pluripotent stem cells: an emerging model in developmental biology. *Development* 140(4):705-717.

21. Elliott SA & Sánchez Alvarado A (2012) The history and enduring contributions of planarians to the study of animal regeneration. *WIREs Dev Biol 2012, doi: 10.1002/wdev.82*.

22. Wagner DE, Wang IE, & Reddien PW (2011) Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* 332(6031):811-816.

23. Cowles MW, *et al.* (2013) Genome-wide analysis of the bHLH gene family in planarians identifies factors required for adult neurogenesis and neuronal regeneration. *Development* 140(23):4691-4702.

24. Currie KW & Pearson BJ (2013) Transcription factors lhx1/5-1 and pitx are required for the maintenance and regeneration of serotonergic neurons in planarians. *Development*.

25. Lapan SW & Reddien PW (2012) Transcriptome analysis of the planarian eye identifies ovo as a specific regulator of eye regeneration. *Cell Rep* 2(2):294-307.

26. März M, Seebeck F, & Bartscherer K (2013) A Pitx transcription factor controls the establishment and maintenance of the serotonergic lineage in planarians. *Development* 140(22):4499-4509.

27. Collins JJ, 3rd*, et al.* (2010) Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLoS Biol* 8(10):e1000509.

28. Gentile L, Cebrià F, & Bartscherer K (2011) The planarian flatworm: an in vivo model for stem cell biology and nervous system regeneration. *Dis Model Mech* 4(1):12-19.

29. Lapan SW & Reddien PW (2011) dlx and sp6-9 Control optic cup regeneration in a prototypic eye. *PLoS Genet* 7(8):e1002226.

30. Wenemoser D, Lapan SW, Wilkinson AW, Bell GW, & Reddien PW (2012) A molecular wound response program associated with regeneration initiation in planarians. *Genes Dev* 26(9):988-1002.

31. Sandmann T, Vogg MC, Owlarn S, Boutros M, & Bartscherer K (2011) The head-regeneration transcriptome of the planarian Schmidtea mediterranea. *Genome Biol* 12(8):R76.

32. King RS & Newmark PA (2013) In situ hybridization protocol for enhanced detection of gene expression in the planarian Schmidtea mediterranea. *BMC Dev Biol* 13:8.

33. Witchley JN, Mayer M, Wagner DE, Owen JH, & Reddien PW (2013) Muscle cells provide instructions for planarian regeneration. *Cell Rep* 4(4):633-641.

34. Oviedo NJ, Newmark PA, & Sánchez Alvarado A (2003) Allometric scaling and proportion regulation in the freshwater planarian Schmidtea mediterranea. *Dev Dyn* 226(2):326-333.

35. Green YS & Vetter ML (2011) EBF factors drive expression of multiple classes of target genes governing neuronal development. *Neural Dev* 6:19.

36. Cebrià F & Newmark PA (2005) Planarian homologs of netrin and netrin receptor are required for proper regeneration of the central nervous system and the maintenance of nervous system architecture. *Development* 132(16):3691-3703.

37. Crozatier M & Vincent A (1999) Requirement for the Drosophila COE transcription factor Collier in formation of an embryonic muscle: transcriptional response to notch signalling. *Development* 126(7):1495-1504.

38. Tseng AS, Beane WS, Lemire JM, Masi A, & Levin M (2010) Induction of vertebrate regeneration by a transient sodium current. *J Neurosci* 30(39):13192-13200.

39. Beane WS, Morokuma J, Adams DS, & Levin M (2011) A chemical genetics approach reveals H,K-ATPase-mediated membrane voltage is required for planarian head regeneration. *Chem Biol* 18(1):77-89.

40. Beane WS, Morokuma J, Lemire JM, & Levin M (2013) Bioelectric signaling regulates head and organ size during planarian regeneration. *Development* 140(2):313-322.

41. Latchman DS (1999) POU family transcription factors in the nervous system. *J Cell Physiol* 179(2):126-133.

42. Urbach R & Technau G (2008) Dorsoventral Patterning of the Brain: A Comparative Approach. *Brain Development in Drosophila melanogaster,* Advances in Experimental Medicine and Biology, ed Technau G (Springer New York), Vol 628, pp 42-56.

43. McMahon AP (2000) Neural patterning: The role of Nkx genes in the ventral spinal cord. *Genes Dev* 14(18):2261-2264.

44. Labbé RM*, et al.* (2012) A Comparative Transcriptomic Analysis Reveals Conserved Features of Stem Cell Pluripotency in Planarians and Mammals. *Stem Cells*.

45. Önal P*, et al.* (2012) Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. *EMBO J* 31(12):2755-2769.

46. Quevillon E*, et al.* (2005) InterProScan: protein domains identifier. *Nucleic Acids Res* 33(Web Server issue):W116-120.

47. Eisenhoffer GT, Kang H, & Sánchez Alvarado A (2008) Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian Schmidtea mediterranea. *Cell Stem Cell* 3(3):327-339.

48. Wagner DE, Ho JJ, & Reddien PW (2012) Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis. *Cell Stem Cell* 10(3):299-311.

49. Pomerantz JH & Blau HM (2013) Tumor suppressors: enhancers or suppressors of regeneration? *Development* 140(12):2502-2512.

50. Pearson BJ & Sánchez Alvarado A (2008) Regeneration, stem cells, and the evolution of tumor suppression. *Cold Spring Harb Symp Quant Biol* 73:565-572.

51. Roberts-Galbraith RH & Newmark PA (2013) Follistatin antagonizes activin signaling and acts with notum to direct planarian head regeneration. *Proc Natl Acad Sci U S A* 110(4):1363-1368.

52. Gurley KA, Rink JC, & Sánchez Alvarado A (2008) Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. *Science* 319(5861):323-327.

53. Umesono Y, Watanabe K, & Agata K (1997) A planarian orthopedia homolog is specifically expressed in the branch region of both the mature and regenerating brain. *Dev Growth Differ* 39(6):723-

CONCLUSION OF THE DISSERTATION

The discovery that adult neurogenesis is widely observed throughout the animal kingdom[1, 2] overturned the long held dogma that the post-embryonic CNS is immutable and incapable of repairing damage [3]. However, new neuron formation is limited in most adult organisms and fails to compensate for catastrophic damages caused by aging, injury, or disease [4, 5]. The development of methods to isolate stem cells (adult stem cells, ES, and iPS) and direct their differentiation into specific neuronal lineages generated great hopes and promises that stem cell-based therapies would be used in the near future to repair CNS damage in humans. Unfortunately, limitations in our understanding of how stem cells behave *in vivo* following injury have thwarted our ability to develop safe and efficient regenerative therapies. The mechanistic study of CNS regeneration has been arduous and problematic due to the limited regenerative capacity of most animal models currently studied. By contrast, planarians have the amazing capacity to regenerate their entire CNS *de novo* from a population of pluripotent adult stem cells they maintain throughout their life, making these animals an ideal system to study basic mechanisms that underlie stem cell-based CNS regeneration *in vivo*.

Members of the bHLH transcription factor family play essential roles in regulating neurogenesis during development, but their functions in the adult nervous system remain poorly understood. In chapter 1 we utilized the regenerative capacity of planarians to identify bHLH factors with roles in adult neurogenesis during CNS regeneration. Of the 44 bHLH factors in planarians we identified nine genes (including *coe*, *sim* and *hesl-3*) important for CNS repair based on exclusive

expression in stem cells and neurons and a requirement for brain regeneration. Our findings demonstrate that highly conserved bHLH factors play critical roles in CNS regeneration, strongly suggesting that these factors are important for adult neurogenesis.

In planarians, prototypical proneural factors such as *acheate-scute* and *atonal* were expressed in stem cells and progenitors but did not result in overt CNS regeneration defects, suggesting that these factors may not exhibit proneural activity in the adult CNS. The fact that these factors may function differently in developing or adult contexts was not entirely surprising since these proteins are likely functioning in concert with unique sets of factors (regulatory factors, growth factors, hormones etc.) when in different developmental scenarios. A related finding was observed in rodents where Ascl-1 regulates cortical interneuron specification in the developing subventricular zone (SVZ), whereas in the adult SVZ Ascl-1 regulates oligodendrocyte specification [6]. We cannot rule out the possibility that silencing of proneural factors caused subtle defects in CNS regeneration that were undetectable by the markers used in this study. Consequently, future experiments using additional CNS markers will be required to confirm whether these genes play proneural roles. Furthermore, analysis of bHLH function during planarian embryogenesis should provide useful insights into how the role of these genes may be altered in different developmental contexts.

We also found that *coe*, *sim* and *hesl-3* label novel neural progenitor populations in uninjured animals and that *coe*$^{+}$ and *sim*$^{+}$ stem cells underlie

formation of the regeneration blastema. These findings further support the hypothesis that the planarian stem cell pool is comprised of lineage-committed progenitors [7]. It will be interesting to investigate if these cells represent transiently amplifying or self-renewing cell populations. Of particular interest are *hesl-3*$^{+}$ stem cells, given the conserved role of these genes in inhibiting neural differentiation and promoting stem cell self-renewal [8]. Without transgenic tools or the ability to culture stem cells or progenitors *in vitro*, assessing self-renewal capacity in planarians is not yet possible. However, our lab has developed methods to label cycling stem cells by soaking animals in F-ara-Edu, a thymidine analog similar to BrdU ([9];unpublished data); therefore, it may be possible to label transiently amplifying cell populations by performing sequential pulse experiments using BrdU and F-ara-Edu.

The identification of neural progenitors in planarians raises another fascinating question: How do pluripotent stem cells generate and maintain progenitor diversity following injury? Lineage-restricted progenitor populations identified thus far in planarians are expressed in distinct spatial domains in uninjured animals. For example, *ovo*$^{+}$ progenitors are located just posterior to the photoreceptors [10], whereas *coe*$^{+}$ progenitors are mostly found between the ventral nerve cords anterior to the pharynx [11]. Therefore, it will be possible to perform amputation experiments on RNAi-treated animals that remove specific progenitor populations and subsequently, assay their ability to replace these cell types. A similar strategy was used to demonstrate that the transcription factor

doublesex/male-abnormal-3 domain (dmd-1) was required for germ cell specification [12]. For this experiment *dmd-1(RNAi)* animals were amputated just posterior to the photoreceptors, an amputation that excluded the reproductive structures from the head fragment; head fragments were then observed for their ability regenerate germline structures *de novo* [12].

Our bHLH screen identified a COE homolog that was expressed in neural progenitors and neurons and was required for brain regeneration. Unlike most bHLH factors, COE proteins bind DNA via a unique zinc-finger domain. Thus, these genes may represent a unique family of HLH transcription factors [13]. COE proteins are well studied during development and play highly conserved roles in neurogenesis [14-16]. COE factors continue to be expressed in the adult CNS [17] and are associated with human cancers of the nervous system [18, 19], yet very little is known about how COE proteins function in the post-embryonic nervous system. In chapter 2 we compared the transcriptome profiles of *coe*-deficient and control animals to identify putative COE targets in differentiating and mature neurons. By examining the expression and function of genes differentially regulated following COE loss, we uncovered three major findings: 1) COE is essential to maintain CNS architecture and drive expression of "neuronal identity" genes (such as ion channels, neurotransmitter receptors and neuropeptides) in multiple classes of neurons distributed throughout the CNS, 2) transcription factors downstream of COE are conserved in planarians and play an essential role in CNS regeneration, and 3) *coe* is required for stem cell homeostasis and normal cell death.

We identified COE targets in mature neurons by validating changes in gene expression in *coe(RNAi)* animals using WISH and performing double-labeling experiments with *coe* using dFISH. However, due to weak or diffuse gene expression many candidate targets were difficult to resolve by *in situ* hybridization and could not be examined in our assays. Included in this list were upregulated genes enriched for mesodermal function. Future investigations aimed at identifying COE DNA-binding sites genome-wide will be required to determine if changes in gene expression in *coe(RNAi)* animals are a direct or indirect consequence of COE loss. COE proteins bind a unique palindromic DNA sequence [17, 20, 21], which can easily be identified *in silico*; however, the consensus sequence is only 10 nucleotides in length and thus, appears by chance in nearly every promoter sequence throughout the planarian genome (unpublished preliminary data). Therefore, the next logical step is to develop an antibody specific to COE that can be used for ChIP-seq analysis. Combining ChIP-seq analysis with our RNA-seq expression dataset will be essential to validate our findings, identify additional COE targets genome-wide, and determine if COE may also function to repress mesodermal expression or fate.

We found five transcription factors downstream of COE that were also putative targets of COE in *Xenopus* development. Candidate targets of COE were identified in *Xenopus* by ectopically expressing EBF3 during development and measuring changes in gene expression using microarray analysis [22]. Our findings suggest that the regulatory program downstream of COE is conserved and redeployed during CNS regeneration. Gene silencing of two of these factors, *nkx2l*

and *pou4l-1*, resulted in CNS regeneration defects. Because orthologs of NKX2 and POU4 play highly conserved roles in specifying and patterning neurons during vertebrate development [23-25], we hypothesize that future studies into the function of these genes will reveal basic mechanisms shared by metazoans that regulate adult neuron formation during CNS regeneration.

We showed that *coe* was required to maintain stem cell homeostasis and cell death, a phenotype that was reminiscent to tumor suppressor-like defects observed when COE is misexpressed in cancer cells *in vitro* [19]. These findings underscore the importance of tumor suppressor genes in tissue regeneration [26] and demonstrate that planarians are a suitable animal model to investigate this phenomenon. Interestingly, unlike COE defects in mammals, which are associated with tumor formation, we did not observe abnormal growths in *coe(RNAi)* animals. If fact, our data suggest that the number of cycling stem cells was slightly reduced in *coe*-deficient animals. We hypothesize that changes in stem cell number are a result of defects in *coe*$^{+}$ stem cells; however, whether stem cell defects are a result of aberrant cell differentiation, cell death, or senescence remains unclear. To investigate these possibilities, additional studies will be needed to examine stem cell dynamics and cell death after different time points following *coe* RNAi treatment.

Our RNA-seq dataset suggests that COE may control stem cell homeostasis through regulating expression of novel *pmp* genes in late progenitors. We showed that *pmp-1* and *pmp-2* are downregulated in *coe(RNAi)* animals and result in an increase in the number of PH3^{+} cells when silenced, consistent with the *coe(RNAi)*

phenotype. The expression profile and RNAi phenotype exhibited by *pmp-1* or *pmp-2* is similar to that of the tumor suppressor p53 [27]. p53 is a transcriptional activator that detects cellular stress and activates downstream pathways to induce repair, growth arrest, autophagy, apoptosis, or senescence [28]. In planarians, p53 is expressed throughout early stem cell progeny and gene knockdown results in a temporary increase in the number of mitotic cells followed by a dramatic collapse of the stem cell population, demonstrating an essential role in regulating stem cell proliferation and self-renewal [27]. It is intriguing that *p53*, *pmp-1,* and *pmp-2* gene silencing affect stem cell proliferation, yet these genes are expressed in postmitotic progenitors. We hypothesize that similar to *p53*, *pmp-1* and *pmp-2* are required to negatively regulate cell cycle progression; thus, silencing of these genes causes cells to reenter the cell cycle, which may ultimately result in stem cell senescence or death. Alternatively, an increase in mitosis could also be a result of aberrant cell differentiation. It will be interesting to investigate the cellular roles of *pmp* genes and further characterize how defects in these genes affect stem cell homeostasis and differentiation.

Planarians are masters of regeneration. These animals have the extraordinary ability to regenerate CNS structures *de novo*, a phenomenon conferred by a population of pluripotent adult stem cells they maintain throughout their lives. My studies revealed that many factors with highly conserved roles in neurogenesis during development are redeployed following injury to repair the CNS. These observations further demonstrate that *S. mediterranea* is a tractable

system to investigate basic mechanisms that underlie CNS regeneration *in vivo* and underscore the importance of using “simple” animal models (relative to mammals) to investigate complex developmental mechanisms. Looking forward, it will be fascinating to take basic principles learned from planarian regeneration and investigate how they apply to mammalian biology.

REFERENCES

1. Gage, F.H., *Neurogenesis in the adult brain.* J Neurosci, 2002. **22**(3): p. 612-3.

2. Lindsey, B.W. and V. Tropepe, *A comparative framework for understanding the biological principles of adult neurogenesis.* Prog Neurobiol, 2006. **80**(6): p. 281-307.

3. Kempermann, G., *Adult neurogenesis*. 2nd ed. 2011, New York: Oxford University Press. xiii, 601 p.

4. Fawcett, J.W., *Overcoming inhibition in the damaged spinal cord.* J Neurotrauma, 2006. **23**(3-4): p. 371-83.

5. Yiu, G. and Z. He, *Glial inhibition of CNS axon regeneration.* Nat Rev Neurosci, 2006. **7**(8): p. 617-27.

6. Jessberger, S., et al., *Directed differentiation of hippocampal stem/progenitor cells in the adult brain.* Nat Neurosci, 2008. **11**(8): p. 888-93.

7. Reddien, P.W., *Specialized progenitors and regeneration.* Development, 2013. **140**(5): p. 951-7.

8. Hatakeyama, J., et al., *Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation.* Development, 2004. **131**(22): p. 5539-50.

9. Neef, A.B. and N.W. Luedtke, *Dynamic metabolic labeling of DNA in vivo with arabinosyl nucleosides.* Proc Natl Acad Sci U S A, 2011. **108**(51): p. 20404-9.

10. Lapan, S.W. and P.W. Reddien, *Transcriptome analysis of the planarian eye identifies ovo as a specific regulator of eye regeneration.* Cell Rep, 2012. **2**(2): p. 294-307.

11. Cowles, M.W., et al., *Genome-wide analysis of the bHLH gene family in planarians identifies factors required for adult neurogenesis and neuronal regeneration.* Development, 2013. **140**(23): p. 4691-702.

12. Chong, T., et al., *A sex-specific transcription factor controls male identity in a simultaneous hermaphrodite.* Nat Commun, 2013. **4**: p. 1814.

13. Crozatier, M. and A. Vincent, *Requirement for the Drosophila COE transcription factor Collier in formation of an embryonic muscle: transcriptional response to notch signalling.* Development, 1999. **126**(7): p. 1495-504.

14. Jackson, D.J., et al., *Developmental expression of COE across the Metazoa supports a conserved role in neuronal cell-type specification and mesodermal development.* Dev Genes Evol, 2010. **220**(7-8): p. 221-34.

15. Crozatier, M. and A. Vincent, *Control of multidendritic neuron differentiation in Drosophila: the role of Collier.* Dev Biol, 2008. **315**(1): p. 232-42.

16. Dubois, L. and A. Vincent, *The COE--Collier/Olf1/EBF--transcription factors: structural conservation and diversity of developmental functions.* Mech Dev, 2001. **108**(1-2): p. 3-12.

17. Kratsios, P., et al., *Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene.* Nat Neurosci, 2012. **15**(2): p. 205-14.

18. Liao, D., *Emerging roles of the EBF family of transcription factors in tumor suppression.* Mol Cancer Res, 2009. **7**(12): p. 1893-901.

19. Zhao, L.Y., et al., *An EBF3-mediated transcriptional program that induces cell cycle arrest and apoptosis.* Cancer Res, 2006. **66**(19): p. 9445-52.

20. Wang, M.M. and R.R. Reed, *Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast.* Nature, 1993. **364**(6433): p. 121-6.

21. Wang, S.S., R.Y. Tsai, and R.R. Reed, *The characterization of the Olf-1/EBF-like HLH transcription factor family: implications in olfactory gene regulation and neuronal development.* J Neurosci, 1997. **17**(11): p. 4149-58.

22. Green, Y.S. and M.L. Vetter, *EBF factors drive expression of multiple classes of target genes governing neuronal development.* Neural Dev, 2011. **6**: p. 19.

23. Cornell, R.A. and T.V. Ohlen, *Vnd/nkx, ind/gsh, and msh/msx: conserved regulators of dorsoventral neural patterning?* Curr Opin Neurobiol, 2000. **10**(1): p. 63-71.

24. Latchman, D.S., *POU family transcription factors in the nervous system.* J Cell Physiol, 1999. **179**(2): p. 126-33.

25. Technau, G.M., *Advances in Experimental Medicine and Biology. Brain development in Drosophila melanogaster. Preface.* Adv Exp Med Biol, 2008. **628**: p. v-vi.

26. Pomerantz, J.H. and H.M. Blau, *Tumor suppressors: enhancers or suppressors of regeneration?* Development, 2013. **140**(12): p. 2502-12.

27. Pearson, B.J. and A. Sánchez Alvarado, *A planarian p53 homolog regulates proliferation and self-renewal in adult stem cell lineages.* Development, 2010. **137**(2): p. 213-21.

28. Jiang, P., W. Du, and X. Yang, *p53 and regulation of tumor metabolism.* J Carcinog, 2013. **12**: p. 21.

APPENDIX

Appendix 1: Accession numbers and names of genes studied in Chapter 2.

Gene ID	Gene name	Accession #
CUFF.1657.1	*CUFF.1657.1*	DN294040
CUFF.231395.1	*CUFF.231395.1*	DN308442
CUFF.238332.1	*CUFF.238332.1*	DN308214
isotig14061	*isotig14061*	DN308508
isotig14071	*isotig14071*	DN304540
isotig19062	*isotig19062*	HO005688
isotig19669	*isotig19669*	HO005946
isotig19703	*isotig19703*	DN307282
isotig21980	*isotig21980*	HO005671
isotig24035	*isotig24035*	HO006564
isotig24454	*isotig24454*	HO006197
isotig24719	*isotig24719*	HO007741
isotig25033	*isotig25033*	DN307458
isotig00709	*Smed-ankyrin repeat protein like-1*	DN304041
isotig23253	*Smed-ankyrin repeat protein-2*	DN293104
isotig21105	*Smed-BTB/POZ domain-containing protein like (btbdl)*	KJ187182
isotig16785	*Smed-c16orf80 (c16orf80)*	DN305596
isotig18792	*Smed-caveolin-1 (cav-1)*	KJ187183
isotig14125	*Smed-cerebral peptide prohormone like-1*	DAA33901
CUFF.121024.1	*Smed-choline acetyltransferase (ChAT)*	FG310880
isotig16398	*Smed-cytochrome p450*	HO005329
isotig24461	*Smed-dual specificity protein phosphatase-1 (dusp-1)*	DN303035
isotig24887	*Smed-dynein heavy chain like*	HO008100
isotig17285	*Smed-dynein light chain axonemal like-1*	DN302772
contig07577	*Smed-E3 ubiquitin-protein ligase mindbomb like-1 (mib-1)*	DN305440
isotig18858	*Smed-fas apoptotic inhibitory molecule*	HO006454
isotig21111	*Smed-gamma irradiation insensitive population-2 (gip-2)*	DN305478
isotig25538	*Smed-gamma-aminobutyric acid receptor subunit beta like (gbrb1)*	DN306571
CUFF.134339.1	*Smed-gamma-aminobutyric acid receptor subunit gamma like (gbrg)*	DN303054
CUFF.25097.1	*Smed-gli pathogenesis related-2 (glipr-2)*	HO005517
isotig16695	*Smed-glyine receptor, alpha (glra)*	DN305020
isotig23606	*Smed-hemicentrin-1*	DN306833
isotig22081	*Smed-iroquois-1 (irx-1)*	DN307336
isotig01231	*Smed-leishmanolysin-like peptidase (LMLN)*	DN302820
isotig09651	*Smed-Lipopolysaccharide-induced tumor necrosis factor (litaf)*	DN306897
isotig19133	*Smed-multidrug and toxin extrusion protein like*	HO005808
CUFF.275890.2	*Smed-musashi*	DN311207
isotig25789	*Smed-netrin-1*	AAY23350
isotig24656	*Smed-neural cell adhesion molecule-2 (ncam-2)*	DN304033
isotig17884	*Smed-neuropeptide y prohormone-3*	DAA33898
isotig06243	*Smed-neurotrypsin-like*	HO005364
isotig08944	*Smed-nidogen2 like*	DN302848
isotig11314	*Smed-nkx2 like-1 (nkx2l-1)*	HO006644
isotig22131	*Smed-notch-1*	KJ187184
isotig15054	*Smed-outer dense fiber protein 3*	DN303365
CUFF.221711.1	*Smed-peptidase inhibitor 16*	DN305218
isotig14448	*Smed-postmitotic progeny-1 (pmp-1)*	AFJ24810
isotig07364	*Smed-postmitotic progeny-10 (pmp-10)*	KJ187185
isotig01740	*Smed-postmitotic progeny-2 (pmp-2)*	HO005560
CUFF.166860.1	*Smed-postmitotic progeny-3 (pmp-3)*	DN307774
CUFF.117103.1	*Smed-postmitotic progeny-4 (pmp-4)*	HO007198
isotig07887	*Smed-postmitotic progeny-5 (pmp-5)*	DN306723

Appendix 1: Accession numbers and names of genes studied in Chapter 2

isotig16173	*Smed-postmitotic progeny-6 (pmp-6)*	DN307817
CUFF.198121.1	*Smed-postmitotic progeny-7 (pmp-7)*	DN310346
isotig18309	*Smed-postmitotic progeny-8 (pmp-8)*	HO007117
contig17595	*Smed-postmitotic progeny-9 (pmp-9)*	HO004941
CUFF.216200.1	*Smed-potassium channel subfamily K (kcnka)*	KJ187186
isotig19235	*Smed-potassium voltage-gated channel, Shab-related-like*	KJ187187
isotig21311	*Smed-pou class 4 transcription factor 3 like-1 (pou4l-1)*	KJ187188
isotig11603	*Smed-protein tyrosine non-receptor type like-1*	DN305232
CUFF.190549.1	*Smed-RAS-like, estrogen-regulated, growth inhibitor (RERG)*	KJ187189
CUFF.229666.1	*Smed-secreted peptide prohormone 18 (spp18)*	ADC84438
CUFF.247824.3	*Smed-secreted peptide prohormone 19 (spp19)*	ADC84440
isotig11493	*Smed-secreted peptide prohormone-2 (spp2)*	DAA33921
isotig05428	*Smed-signal peptide containing-1 (spc-1)*	DN308176
isotig19500	*Smed-Sodium channel protein-1 (scna-1)*	DN303323
isotig19431	*Smed-splicing factor 3b subunit 4*	DN307929
isotig11373	*Smed-T cell acute leukemia (tal)*	AGZ94920
isotig08511	*Smed-T-cell leukemia homeobox protein (tlx)*	KJ187190
isotig13360	*Smed-tetraspanin like*	HO005410
isotig13551	*Smed-tetratricopeptide repeat protein 30 like*	DN305472
isotig06925	*Smed-vesicle-associated membrane protein like-1 (vamp-1)*	DN307612
CUFF.97636.1	*Smed-voltage-gated sodium channel (scna-1)*	KJ187191
CUFF.260325.1	*Smed-voltage-gated sodium channel (scna-2)*	KJ187192
isotig25405	*Smed-voltage-gated sodium channel (scna-3)*	KJ187193
isotig17556	*Smed-WD repeat-containing protein-1 (wdr-1)*	KJ187194

Appendix 2: List of primers used in Chapter 2.

Gene ID	Cloning Primer Forward	Cloning Primer Reverse	qPCR Primer Forward	qPCR Primer Reverse
isotig21105	CGGATCAGTGATGGCTTGTGAAATG	AGAGGAGGTTCTAGAGGCAC	NA	NA
isotig18792	CGGAATGGCGGGAAAGATTGCTG	TTACAAGAAGGCTAAGGTTTGC	NA	NA
isotig22081	NA	NA	GCTCGCAATTCTCACACAAA	TTTGCAATCCAACCGATTTT
isotig09651	NA	NA	ATCCTCCTCCTCCTGCGTAT	CTGGATATGGCGGATATGGT
isotig11314	NA	NA	AACCATCCAACCGAATCATC	GCTGCTGCAACATCTGGATA
isotig22131	CGGCCAGAACAAAGGTCGCTGTTA	CTTGGTGGCTTAAATGCAGTG	NA	NA
isotig07364	CGGAAAACCGATGTGAAATTGGAG	TAAACTATCAGTGGTATTTTTCTTTTCGTTACAG	NA	NA
CUFF.216200.1	CGGATATTCTGTTGGGTTTGTTGAGC	TAAACTATTCTCAATAGCAGAACGTTTGGA	NA	NA
isotig19235	CGGCCTTTATCCTGCTCTCAGTCTTG	TAAACTATACCGTTTTGTTGGCACTTTC	NA	NA
isotig21311	CGGTGCGACCTAGTTTGAATATTTTAGTG	TAAACTATAAAACGGACGGATAATGTCG	GGCAATGGCCTCATTTACAC	GCTAGCGGCACTGCTAACTT
CUFF.190549.1	CGGGAACAACCGATTCTCCGACG	AAAGTTCTTGACCGCCTC	NA	NA
isotig11373	NA	NA	TGTAGCACCGAGGAAATCGT	GGCAATATTTGTCGGAGGTC
isotig08511	CGGTGTCAGCATCGTTTTCCATT	GTTTGTCGCCTCCATTTTGT	CAAGCAGCCAGTCGGTATTT	TTGGATATGGCGGAAAGAGT
CUFF.97636.1	CGGTTGATGAAGAATTGGGAGTGG	TAAACTATGGATCTTCGTCGAATTTTGG	NA	NA
CUFF.260325.1	CGGATGCGATTCCGTCAATTTTC	GCATGATTTCCATCCATCCT	NA	NA
isotig25405	CGGGAATATCAGCGGAGATTATACCG	TAAACTATAGGAATCGGTTTCTGTGGAG	NA	NA
isotig17556	CGGTTATTTACGGGCTGGGTTGC	CAACAAACACAACGAAAC	NA	NA

XhoI site added to all Forward Cloning primers (CCGCTCGAG)

NotI site added to all Reverse Cloning primers (ATAAGAATGCGGCCGC)

Appendix 3: List of differentially expressed genes in *coe*-deficient animals.

ID	FC	logCPM	FDR	Blast Hit Acc.	E-value	Uniprot Acc.
CUFF.247824.3	-15.104	2.5163279	2.00E-56	ADC84440	2.5494E-53	No Hit
isotig22173	-8.2915	0.9832116	2.31E-21	YP_004290409	2.2869E-09	No Hit
isotig18390	-6.4521	1.5249937	3.79E-23	#N/A	#N/A	No Hit
CUFF.317539.1	-4.5234	0.4638613	8.59E-10	XP_002571957	0.00086039	O14980
isotig19235	-4.123	0.6965006	4.42E-10	EFZ17440	2.0133E-97	Q92953
isotig05428	-3.7812	1.2939794	5.70E-12	#N/A	#N/A	No Hit
isotig20240	-3.7047	2.8013765	1.21E-20	#N/A	#N/A	No Hit
isotig14448	-3.6046	6.6850396	3.71E-32	AFJ24810	6.1997E-07	No Hit
isotig12939	-3.4974	0.1665633	3.18E-06	EKC29408	4.5256E-66	P11908
isotig03352	-3.486	7.1335683	6.01E-31	#N/A	#N/A	No Hit
isotig24719	-3.4159	7.5106983	3.02E-30	CCD82334	2.1719E-06	No Hit
isotig26190	-3.4112	0.0825483	1.01E-05	#N/A	#N/A	No Hit
isotig21980	-3.3547	5.1262748	2.16E-26	#N/A	#N/A	No Hit
isotig23681	-3.2531	3.535882	2.40E-20	AGZ94924	0	Q9H4W6
isotig23253	-3.2053	0.2914742	8.31E-06	EJY57493	1.9221E-34	Q12955
isotig19062	-3.1524	6.2717708	1.23E-25	#N/A	#N/A	No Hit
isotig21267	-3.1464	6.4783375	1.01E-25	GAA55911	7.1464E-07	No Hit
CUFF.229666.1	-3.1443	5.1466809	6.44E-24	ADC84438	2.5865E-47	No Hit
isotig09640	-3.1111	6.1468269	6.24E-25	#N/A	#N/A	No Hit
isotig24892	-2.9998	0.9345229	2.61E-07	AFJ51623	2.7446E-06	P62079
isotig14896	-2.8843	3.0149971	1.36E-14	#N/A	#N/A	No Hit
isotig25899	-2.8426	0.4223356	4.20E-05	EKC21550	1.308E-70	P35499
isotig18689	-2.8375	3.4335577	9.44E-16	XP_002734182	1.7736E-65	Q96EM0
isotig11838	-2.8328	1.8368048	9.60E-10	#N/A	#N/A	No Hit
isotig11373	-2.7547	0.0192984	0.0006487	EKC21640	9.63E-50	P17542
isotig15923	-2.6867	3.0859106	4.86E-13	#N/A	#N/A	No Hit
isotig15092	-2.6377	3.4315979	1.53E-13	#N/A	#N/A	No Hit
CUFF.260325.1	-2.6304	0.1139664	0.0008258	XP_001601254	6.6811E-64	Q99250
isotig00709	-2.5847	1.9218181	1.56E-08	XP_002952050	5.1358E-08	Q92527
isotig25556	-2.5458	1.6973999	1.40E-07	#N/A	#N/A	No Hit
CUFF.134339.1	-2.5413	1.7065867	1.35E-07	ELU17488	3.5934E-69	Q99928
isotig22081	-2.538	1.2198641	3.67E-06	GAA49245	3.7548E-30	P78415
CUFF.288056.1	-2.5224	2.5679694	5.96E-10	#N/A	#N/A	No Hit
CUFF.269036.1	-2.5216	7.9468719	1.71E-17	AFJ24802	2.5893E-43	No Hit
isotig24917	-2.4957	4.4339014	2.99E-14	XP_003731181	8.9821E-76	No Hit
contig21557	-2.4777	0.5165833	0.0003436	#N/A	#N/A	No Hit
CUFF.238332.1	-2.4506	0.0145085	0.0039518	#N/A	#N/A	No Hit
isotig12571	-2.4141	0.3597987	0.0011466	XP_003450309	2.2918E-13	No Hit
isotig13360	-2.4095	2.824355	1.03E-09	CAX73483	3.4211E-05	No Hit
isotig25378	-2.4055	1.8769643	3.14E-07	#N/A	#N/A	No Hit
isotig24454	-2.3841	3.6674863	1.43E-11	#N/A	#N/A	No Hit
CUFF.293166.1	-2.36	7.0513247	7.88E-15	AFJ24802	6.2291E-64	No Hit
contig07577	-2.3566	0.7478498	0.0003196	XP_003532904	5.5645E-08	Q86YT6
isotig00685	-2.3467	3.8538822	1.75E-11	XP_001649473	9.922E-09	P16157
isotig24035	-2.3467	1.2855871	2.66E-05	#N/A	#N/A	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig24035	-2.3467	1.2855871	2.66E-05	#N/A	#N/A	No Hit
CUFF.1657.1	-2.3458	5.642807	7.77E-14	#N/A	#N/A	No Hit
isotig06925	-2.3002	0.5568896	0.0012607	XP_422640	1.3739E-34	P63027
CUFF.318822.1	-2.3002	2.0009871	7.08E-07	GAA37002	2.6605E-10	No Hit
CUFF.175557.3	-2.2935	3.0314122	2.83E-09	XP_001649473	3.0008E-33	Q12955
isotig07515	-2.2891	0.2273799	0.0045479	GAA41547	1.4339E-17	No Hit
CUFF.329572.1	-2.2459	0.6933065	0.0010356	GAA55481	2.3022E-10	No Hit
CUFF.246623.1	-2.2358	7.4421798	3.89E-13	#N/A	#N/A	No Hit
isotig17006	-2.233	3.4304676	1.59E-09	#N/A	#N/A	No Hit
isotig19703	-2.2254	0.9182886	0.0005053	#N/A	#N/A	No Hit
CUFF.25097.1	-2.2028	0.9614569	0.0004779	AFJ24739	1.1485E-50	Q6UWM5
CUFF.216200.1	-2.197	0.3392335	0.005685	GAA52537	7.2483E-37	P57789
isotig15570	-2.1919	0.4637042	0.0039595	ELU15220	6.6188E-11	A7E2U8
isotig25538	-2.1667	1.3316039	0.0001609	ELU17487	1.1022E-50	P18505
isotig19500	-2.1303	2.5019996	9.96E-07	AAC63049	9.348E-106	Q15858
CUFF.231395.1	-2.1276	0.7652679	0.0021927	#N/A	#N/A	No Hit
CUFF.203310.1	-2.1258	0.9758051	0.001045	CCD77361	4.467E-10	No Hit
isotig23247	-2.0993	4.4061772	1.71E-09	#N/A	#N/A	No Hit
isotig21997	-2.0964	1.2690986	0.0003961	CCD75549	2.8057E-72	Q13308
isotig22215	-2.0937	2.7729355	6.39E-07	#N/A	#N/A	No Hit
isotig12291	-2.0871	4.0244156	7.22E-09	NP_001096513	5.4347E-91	Q96CU9
isotig01250	-2.0867	0.8130877	0.0028185	#N/A	#N/A	No Hit
isotig00933	-2.0825	0.0483838	0.029585	ADY47987	4.837E-48	Q96QA6
isotig14071	-2.0792	3.4335892	6.79E-08	#N/A	#N/A	No Hit
isotig16945	-2.0777	1.9421769	3.57E-05	GAA55616	1.6054E-42	O15162
isotig19476	-2.0752	1.5912272	0.0001536	XP_002411948	1.8202E-43	Q9BYG0
CUFF.251115.1	-2.0649	1.8952132	4.99E-05	#N/A	#N/A	No Hit
isotig14125	-2.0626	1.9528173	4.29E-05	DAA33901	4.7537E-63	No Hit
isotig26027	-2.0502	0.9511213	0.0023992	ABA60382	7.014E-77	Q15822
isotig18853	-2.0458	3.1452058	3.88E-07	XP_003393881	1.6756E-08	O43657
isotig22677	-2.0428	3.6382903	7.34E-08	EKC26077	1.287E-31	No Hit
CUFF.312619.1	-2.0382	0.9542489	0.0032147	#N/A	#N/A	No Hit
isotig09214	-2.0312	4.1583172	2.16E-08	AEQ00955	2.5638E-44	Q12913
isotig01740	-2.0226	7.5588138	4.42E-10	#N/A	#N/A	No Hit
CUFF.106573.1	-2.0109	0.1034703	0.0357027	#N/A	#N/A	No Hit
isotig19189	-2.0001	0.6951339	0.0088526	GAA53959	1.3326E-25	No Hit
isotig18858	-1.9915	3.3935748	5.53E-07	ELT99248	1.7423E-49	Q9NVQ4
isotig11603	-1.9889	0.5519627	0.0144611	XP_002582135	0	Q06124
isotig13551	-1.9851	1.0302241	0.0034617	XP_001894316	1.0323E-75	Q8N4P2
isotig23728	-1.9829	6.9261516	2.05E-09	#N/A	#N/A	No Hit
CUFF.319906.1	-1.9788	1.5168035	0.0007381	ELU05583	6.631E-57	A6NFQ2
CUFF.286082.1	-1.9774	1.990525	0.0001103	EKC37364	1.027E-81	P19440
isotig19669	-1.9629	0.1070482	0.0459201	#N/A	#N/A	No Hit
isotig19431	-1.9589	2.5329202	2.61E-05	GAA56093	2.584E-138	Q15427
isotig25881	-1.9512	0.4115805	0.0254939	CCD82648	1.206E-28	No Hit
isotig16707	-1.9442	0.7720185	0.0110396	AAW27824	1.8545E-07	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig26114	-1.935	1.0904301	0.0048456	GAA52537	1.0656E-69	P57789
CUFF.113640.1	-1.9342	2.49729	4.48E-05	#N/A	#N/A	No Hit
CUFF.183849.1	-1.9234	2.2674381	0.0001052	EKC31091	3.2745E-38	P56539
isotig24630	-1.9196	0.7246074	0.0154229	#N/A	#N/A	No Hit
CUFF.250015.1	-1.9188	2.0018451	0.0002641	XP_002577171	7.0703E-16	No Hit
isotig06243	-1.9134	9.1844218	1.00E-08	XP_003454280	1.9367E-36	P56730
isotig15054	-1.9111	3.1351292	7.26E-06	CAX74044	1.351E-31	Q96PU9
CUFF.266184.1	-1.905	3.0538942	1.04E-05	#N/A	#N/A	No Hit
CUFF.286815.1	-1.904	0.4818969	0.0307634	#N/A	#N/A	No Hit
CUFF.221711.1	-1.902	0.5544561	0.0262514	AAI18027	3.2595E-30	Q6UXB8
isotig16398	-1.8941	4.2651614	6.39E-07	XP_003439757	4.8261E-59	Q6VVX0
isotig14061	-1.8889	2.6690891	5.36E-05	EKC35100	3.3311E-14	F8WEG0
CUFF.271873.1	-1.8863	1.6196392	0.0016109	XP_002811794	2.7271E-38	O75897
isotig25033	-1.8835	2.9332877	2.56E-05	#N/A	#N/A	No Hit
isotig11493	-1.8832	2.282354	0.0002132	DAA33921	4.8539E-61	No Hit
isotig21311	-1.8802	3.1491555	1.37E-05	CAA49382	1.667E-121	Q15319
isotig19131	-1.8663	0.7875272	0.0194368	ELU09984	4.3888E-07	No Hit
isotig19133	-1.8648	2.856897	4.64E-05	XP_004149091	3.7352E-13	Q86VL8
CUFF.296605.1	-1.8623	0.3944558	0.0474522	#N/A	#N/A	No Hit
CUFF.299172.2	-1.862	1.3644033	0.0040662	#N/A	#N/A	No Hit
CUFF.166860.1	-1.8548	5.1019902	4.91E-07	#N/A	#N/A	No Hit
CUFF.244578.1	-1.8514	2.9227259	4.80E-05	XP_001604478	1.5786E-42	Q96CU9
isotig13119	-1.848	2.3070645	0.0003521	CCD75294	5.115E-110	Q9P2U8
isotig21111	-1.8398	2.555595	0.0001815	#N/A	#N/A	No Hit
isotig01231	-1.8367	6.6326405	2.26E-07	XP_689096	2.1108E-77	Q96KR4
isotig09183	-1.8363	0.4693811	0.049969	XP_002111114	1.7826E-09	Q8N5Y8
isotig23810	-1.8338	0.65315	0.0333384	#N/A	#N/A	No Hit
CUFF.121024.1	-1.8294	1.34737	0.0076575	BAG16388	1.3191E-55	P28329
isotig25405	-1.8277	1.3924216	0.005894	ABX47011	1.1344E-68	Q15858
CUFF.328331.1	-1.8275	0.8254823	0.0238047	GAA55481	7.8727E-18	No Hit
isotig13790	-1.8259	2.0762417	0.0009091	EKC42074	7.4792E-35	P57721
isotig19918	-1.8222	1.7401328	0.00266	XP_002111960	1.2211E-39	A1A4V9
CUFF.25077.2	-1.8161	1.7915368	0.00231	AFJ24739	7.2644E-66	Q6UXB8
CUFF.117103.1	-1.8135	7.7953044	3.11E-07	#N/A	#N/A	No Hit
CUFF.239306.1	-1.8106	2.6453794	0.0002641	AAW24939	2.0375E-18	P63316
CUFF.196931.2	-1.8042	3.5970469	2.40E-05	EJY57493	2.6509E-28	O15084
isotig09393	-1.8014	1.8855945	0.0027748	AAX25744	1.3844E-16	No Hit
isotig08135	-1.8004	1.8090917	0.0030075	XP_001955987	3.7064E-07	No Hit
CUFF.251724.1	-1.7998	7.4875824	5.08E-07	AFJ24810	9.9725E-07	No Hit
CUFF.97636.1	-1.7981	3.1464569	8.25E-05	AAC63049	8.496E-111	Q99250
isotig22222	-1.7886	2.2963911	0.0009236	CAG02703	0.00048063	No Hit
isotig16932	-1.7761	3.4033602	7.01E-05	XP_002571413	8.4224E-22	No Hit
CUFF.80923.1	-1.7757	1.0664127	0.0233446	AAA40116	3.0513E-09	Q16874
CUFF.251475.1	-1.775	2.1521378	0.0015459	GAA51571	1.6751E-12	No Hit
isotig04494	-1.7729	1.9262031	0.0029312	#N/A	#N/A	No Hit
isotig07887	-1.7695	5.7496917	2.88E-06	#N/A	#N/A	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig05367	-1.763	5.2410554	5.30E-06	XP_001199228	2.47E-110	P07711
CUFF.277702.1	-1.7588	3.3499503	0.000112	ELT89283	9.0017E-13	No Hit
isotig22162	-1.7518	1.3143185	0.0169777	#N/A	#N/A	No Hit
isotig18792	-1.7497	1.5153857	0.0109731	EKC31091	6.8021E-37	P56539
isotig02435	-1.7476	1.2233076	0.0214304	AFJ24739	4.7016E-57	Q6UXB8
CUFF.251485.1	-1.7472	4.3990109	2.34E-05	EKC34055	2.9923E-10	P43146
CUFF.291167.1	-1.7469	2.1261921	0.0027046	BAE78814	2.5352E-84	No Hit
CUFF.196043.1	-1.7464	0.8954823	0.0402723	ELT96089	1.4322E-42	Q9GZQ6
isotig05736	-1.746	3.3946168	0.0001442	XP_002409162	2.809E-102	B7Z9G9
isotig22415	-1.7447	3.1944095	0.0002175	EKC24880	1.2839E-29	Q96M69
isotig18767	-1.741	0.9192114	0.0423339	XP_001496842	0.00011284	Q9Y2G3
CUFF.297113.1	-1.7373	1.0288572	0.033947	ELT90372	1.6586E-47	Q8TCU5
isotig18667	-1.734	2.1012853	0.0031878	#N/A	#N/A	No Hit
CUFF.203314.1	-1.7271	1.4232814	0.0169909	GAA51400	9.0673E-44	Q96J84
isotig26364	-1.7265	1.024915	0.0372671	ADZ13686	4.2956E-14	No Hit
CUFF.96037.1	-1.7211	1.7721244	0.0083884	ELU00356	7.1411E-07	Q5H9R7
isotig21677	-1.7209	1.0895611	0.0360358	AAR11265	2.0626E-57	Q13332
isotig01328	-1.714	5.1496292	2.25E-05	#N/A	#N/A	No Hit
CUFF.269374.1	-1.714	2.3535762	0.0023966	GAA37002	0.00011221	No Hit
isotig20401	-1.7116	2.032686	0.0053963	CAX74741	4.2006E-24	P57078
isotig20913	-1.711	1.9701384	0.006105	#N/A	#N/A	No Hit
isotig16173	-1.7079	6.0984552	1.36E-05	#N/A	#N/A	No Hit
isotig17472	-1.706	1.2233912	0.031163	#N/A	#N/A	No Hit
isotig09858	-1.7056	1.9574209	0.0070421	EKC41007	1.024E-50	A8MV24
isotig22114	-1.7055	1.4310184	0.0210208	ELU04556	1.0214E-62	Q9Y4C4
isotig22086	-1.7048	2.8519633	0.0009236	GAA37002	1.8292E-10	No Hit
CUFF.300184.1	-1.7019	1.6863011	0.0131998	BAE78814	5.5818E-15	No Hit
CUFF.251976.1	-1.6988	1.2442926	0.0320421	GAA53244	1.6537E-92	Q8WXX0
isotig12744	-1.6978	1.2110108	0.034322	#N/A	#N/A	No Hit
CUFF.263502.1	-1.69	2.1389088	0.0057466	XP_003223316	1.807E-21	Q9BQS2
CUFF.260845.2	-1.6889	1.0717189	0.0496086	EKC21427	3.5544E-13	Q9P0L9
isotig09488	-1.685	2.2281658	0.0051573	ELU09120	2.3123E-77	Q53EV4
isotig22025	-1.683	3.7703495	0.000282	AFJ24792	4.266E-18	Q9Y4K3
CUFF.296163.1	-1.6812	1.1772012	0.0421399	#N/A	#N/A	No Hit
isotig16538	-1.6786	3.4140833	0.0005632	ADF47429	4.7747E-39	O00463
isotig22843	-1.6771	3.2345432	0.0007381	ELU00727	7.6671E-18	No Hit
CUFF.160648.1	-1.6754	1.5875712	0.0210631	#N/A	#N/A	No Hit
isotig21882	-1.6751	2.7673054	0.0019854	#N/A	#N/A	No Hit
isotig17777	-1.6703	2.7166774	0.0023026	#N/A	#N/A	No Hit
isotig15606	-1.6695	5.0813084	7.85E-05	AFJ24792	3.4055E-21	Q9Y4K3
isotig18507	-1.6694	4.4395885	0.0001547	#N/A	#N/A	No Hit
isotig09765	-1.6631	2.7068568	0.0026483	#N/A	#N/A	No Hit
CUFF.181829.1	-1.6598	2.4081042	0.0050742	XP_002941054	0.00031531	Q71SY5
CUFF.230880.1	-1.6594	2.5838157	0.0036174	#N/A	#N/A	No Hit
CUFF.175533.1	-1.6586	2.9243041	0.0019193	#N/A	#N/A	No Hit
CUFF.67844.1	-1.6584	1.7268023	0.0192265	GAA56288	2.531E-56	Q6ZR08

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig19334	-1.6565	1.4138507	0.0349681	XP_002575539	4.3981E-29	Q6W5P4
isotig18315	-1.6552	1.4455232	0.035319	#N/A	#N/A	No Hit
CUFF.230625.1	-1.6546	3.3230224	0.0010109	GAA42275	2.681E-113	Q9HC56
isotig24658	-1.6497	2.2921046	0.0073864	ELU16968	9.5921E-71	Q8IZ02
CUFF.241197.1	-1.6483	2.6640217	0.0034658	GAA31584	6.6582E-54	Q5JQ13
isotig24961	-1.6472	5.2091006	0.0001226	ELU01393	1.302E-118	Q8TC71
CUFF.63887.1	-1.6458	3.4897349	0.0009114	GAA49681	8.1065E-25	No Hit
isotig17492	-1.6429	1.5590475	0.0331877	#N/A	#N/A	No Hit
CUFF.215172.1	-1.6352	1.730512	0.0253037	BAF57622	1.105E-173	O00571
CUFF.279881.1	-1.6351	2.7759144	0.0040403	#N/A	#N/A	No Hit
isotig09651	-1.6346	4.3831245	0.0003544	CAG06054	9.576E-12	Q99732
isotig18808	-1.6339	2.4124846	0.0077309	XP_003441411	7.5887E-19	Q9Y345
isotig20131	-1.6289	2.8678246	0.0036174	DAA33914	1.5375E-22	No Hit
CUFF.95594.1	-1.6271	1.7582296	0.0264082	#N/A	#N/A	No Hit
CUFF.293397.1	-1.6267	2.1000735	0.0147553	CCD58525	3.0804E-12	Q8WY19
isotig17729	-1.6241	2.2433071	0.012616	XP_001635956	9.9829E-90	Q5VZ52
isotig11314	-1.6173	1.6052842	0.0376034	AAV85465	2.7246E-22	Q8TAU0
CUFF.224152.1	-1.6157	2.4698508	0.0088819	#N/A	#N/A	No Hit
isotig02864	-1.6155	2.292363	0.0120694	#N/A	#N/A	No Hit
isotig17285	-1.6154	1.8009991	0.0299992	EDL04633	6.7717E-29	O96015
isotig05085	-1.6122	3.131468	0.0031615	#N/A	#N/A	No Hit
isotig09309	-1.6106	3.3923652	0.0021392	XP_002592650	5.1188E-88	Q3ZCV2
CUFF.30816.1	-1.6089	1.7884743	0.0307865	XP_002715067	1.8884E-17	Q9H0A6
isotig17657	-1.6079	1.8740209	0.0278163	XP_002192223	3.6631E-11	Q6ZU64
isotig21105	-1.6073	2.4428891	0.0110396	EKC24917	5.3623E-52	C9JJ37
CUFF.235417.1	-1.607	2.2865847	0.0136822	ELU04638	7.8299E-20	P48065
isotig20439	-1.6058	1.6762212	0.038302	GAA51498	5.1688E-16	O00591
isotig19747	-1.6057	2.1743039	0.0184881	CAX70453	1.6928E-08	C9JVW0
isotig21568	-1.6037	4.2836168	0.0007438	AAH49362	6.768E-44	S4R410
isotig23606	-1.6033	3.6439935	0.0017347	XP_001632823	8.6471E-48	Q96RW7
isotig24889	-1.6027	2.9280032	0.0049373	#N/A	#N/A	No Hit
isotig26426	-1.6025	1.7876272	0.0338348	AGA95402	3.0034E-49	Q9C005
isotig23678	-1.6019	3.5626058	0.0020288	XP_001649474	4.8387E-41	P16157
isotig20671	-1.6017	1.7944277	0.0329416	CCD76613	5.5101E-05	No Hit
isotig24027	-1.5961	2.059358	0.0233446	XP_002732835	7.4622E-34	B7ZC32
CUFF.246214.2	-1.595	1.595752	0.0495358	ELU00857	7.107E-140	Q8TDX7
isotig23987	-1.5931	1.6155368	0.0493081	AEJ87271	1.47E-136	O43497
isotig19845	-1.5927	2.8056015	0.0073268	ELT92118	2.7831E-05	No Hit
isotig16695	-1.5897	3.9185154	0.0016841	GAA51498	2.388E-157	P23416
isotig02449	-1.5896	6.1455234	0.0003322	AFJ24739	3.4469E-73	Q6UXB8
isotig08511	-1.5895	2.0940146	0.0230962	EKC40719	1.2402E-41	O43711
isotig18446	-1.5893	3.2349043	0.0038798	#N/A	#N/A	No Hit
isotig18882	-1.5858	2.2785354	0.0191979	XP_002572261	3.0105E-08	No Hit
CUFF.85189.1	-1.5848	2.5346499	0.0128766	XP_002733392	8.1571E-29	Q7Z442
CUFF.91045.1	-1.5825	2.0000007	0.0300127	BAC06343	0	Q16099
isotig05250	-1.5797	4.5161465	0.0011189	ELU15121	5.96E-109	E9PHT4

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig22602	-1.5792	2.4219812	0.0170704	EKC37252	2.7195E-91	Q5I0G3
isotig12171	-1.5782	3.6351492	0.0029858	XP_781718	8.0549E-08	O75317
contig18172	-1.5772	2.6651842	0.012704	#N/A	#N/A	No Hit
isotig25184	-1.5762	2.0510772	0.0300127	#N/A	#N/A	No Hit
isotig15969	-1.576	6.6156282	0.0003987	BAJ10272	2.062E-110	Q02094
CUFF.290145.1	-1.5753	5.7416229	0.0005626	#N/A	#N/A	No Hit
CUFF.55595.1	-1.5752	4.1189182	0.0017146	EKC20777	1.397E-125	K7EK91
isotig14769	-1.5736	3.1296048	0.0062011	#N/A	#N/A	No Hit
isotig22138	-1.573	3.4470319	0.004183	XP_002582126	1.538E-19	No Hit
isotig25374	-1.572	2.8859956	0.0093901	#N/A	#N/A	No Hit
isotig13972	-1.5711	6.9262114	0.0004165	GAA28183	0	Q9NYC9
isotig16037	-1.5701	4.4193503	0.0015421	EKC24012	4.1278E-33	P62820
CUFF.55597.1	-1.5668	4.5131229	0.0014714	ELT87666	0	K7EK91
isotig25008	-1.566	3.4053162	0.0048611	#N/A	#N/A	No Hit
isotig19917	-1.5643	1.8969757	0.0447004	ELU04033	1.4449E-12	No Hit
CUFF.230158.1	-1.56	2.1196849	0.0337479	#N/A	#N/A	No Hit
CUFF.96196.1	-1.5593	2.1571997	0.033448	GAA52351	1.8815E-51	Q7L2J0
CUFF.300419.1	-1.5583	6.1117445	0.0007335	EKC20777	0	Q9UFH2
isotig11537	-1.5575	4.1025119	0.0026733	XP_973373	4.836E-127	P06280
isotig05213	-1.5564	4.8139429	0.0014539	#N/A	#N/A	No Hit
isotig23586	-1.5532	2.5899787	0.0192265	ELT99958	2.514E-168	Q12756
isotig14076	-1.553	9.7026277	0.0005202	ELQ76425	7.6788E-44	No Hit
isotig14731	-1.5523	4.4521462	0.002206	ADF47430	6.6092E-18	Q13114
isotig06704	-1.5508	3.4511264	0.0064082	#N/A	#N/A	No Hit
isotig15748	-1.548	4.2384679	0.0027802	#N/A	#N/A	No Hit
CUFF.185731.1	-1.547	2.2249229	0.0329103	#N/A	#N/A	No Hit
isotig03617	-1.5467	3.3506696	0.0075187	XP_002609250	2.3931E-11	Q5CZ79
isotig25473	-1.5431	2.1532543	0.038302	XP_002593300	1.9504E-22	Q9Y4F1
CUFF.96335.1	-1.5412	3.3513433	0.0085845	#N/A	#N/A	No Hit
CUFF.199361.1	-1.5403	3.7139185	0.005667	ELU15408	5.9451E-32	Q9Y6K8
isotig08944	-1.5394	9.1912082	0.0007501	EKC26122	7.634E-15	Q14112
CUFF.284876.1	-1.5387	2.7290472	0.0193559	CCD60929	2.171E-113	Q6ZR08
isotig12836	-1.5382	6.5034177	0.0010649	XP_002163647	7.837E-124	P07098
CUFF.182587.3	-1.5365	9.8656422	0.0007911	XP_002733296	1.171E-05	Q08629
CUFF.294826.1	-1.5338	2.7926865	0.0188622	XP_004227474	6.284E-119	Q8TD57
isotig18714	-1.5328	2.644582	0.0238047	#N/A	#N/A	No Hit
CUFF.115652.1	-1.5326	3.931412	0.0053319	#N/A	#N/A	No Hit
isotig20500	-1.5306	2.213008	0.0418685	ABO52851	1.2393E-73	Q14721
CUFF.279323.1	-1.5295	3.4334975	0.0094362	GAA55041	1.7174E-13	Q63HQ2
CUFF.111276.1	-1.5282	2.4735826	0.0311622	EFW45779	1.344E-27	Q96NX5
isotig22295	-1.5271	2.8502704	0.0193559	EKC18822	8.2559E-14	No Hit
isotig22555	-1.5265	3.3091999	0.011149	ELT88007	5.8871E-35	Q8N9S7
isotig21282	-1.525	5.1378841	0.002514	EKC41830	0	Q9P2D7
isotig16211	-1.5249	8.4829743	0.0011234	#N/A	#N/A	No Hit
isotig02172	-1.5233	4.6685671	0.0035468	XP_003726783	2.1494E-25	Q6UXB8
CUFF.184531.1	-1.5228	7.6899845	0.0012887	#N/A	#N/A	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig05907	-1.5188	2.3131287	0.0443349	#N/A	#N/A	No Hit
isotig16575	-1.5184	4.7970292	0.0036037	#N/A	#N/A	No Hit
isotig13597	-1.5173	2.6165077	0.0303062	ELT88007	2.6523E-74	B1ARL5
isotig15901	-1.517	4.0394912	0.00643	#N/A	#N/A	No Hit
CUFF.307784.1	-1.5149	3.4535147	0.0121244	XP_001364847	9.2142E-83	Q96HU8
isotig25462	-1.5148	2.8309816	0.0238047	EKC37639	7.6737E-13	No Hit
isotig16445	-1.5143	3.4169823	0.0128856	GAA57409	4.8416E-82	Q12908
CUFF.190549.1	-1.5139	3.4677345	0.012186	ADD20562	2.7552E-41	Q96A58
isotig16592	-1.5131	2.9149648	0.0225483	ELT97237	2.3887E-35	Q9UBV2
CUFF.169158.1	-1.513	2.5151191	0.0351632	XP_002745652	2.458E-71	Q13563
isotig20408	-1.5096	3.9247632	0.0085207	#N/A	#N/A	No Hit
isotig13275	-1.5096	8.0349248	0.0017166	ABX80072	1.8581E-19	Q8TEU8
isotig23460	-1.5079	5.4233819	0.0032215	EKC37609	8.5335E-24	Q9Y6K8
isotig21178	-1.5075	2.9275429	0.0234169	XP_002579723	1.9369E-14	Q8NBI5
CUFF.317283.1	-1.5068	3.1439135	0.0193486	BAE78814	3.046E-150	No Hit
isotig24481	-1.5049	3.8172008	0.0100718	NP_998717	0	P52209
CUFF.284217.1	-1.5033	3.3187588	0.0171045	GAA55616	1.0712E-84	O15162
CUFF.207497.1	-1.5027	5.4765169	0.0035468	#N/A	#N/A	No Hit
isotig23691	-1.5023	2.8512085	0.0278163	ELU02812	3.1738E-65	Q8N7U6
CUFF.59202.1	-1.5014	2.7345146	0.0329012	XP_003140688	5.8313E-43	P49585
isotig22789	-1.5007	2.8008802	0.0300127	EKC42879	2.3053E-37	Q9H5I5
CUFF.219098.1	-1.4992	5.6030507	0.0036174	CAX74321	1.7629E-26	Q6UXB8
CUFF.178052.2	-1.4988	2.7616435	0.0327415	EKC39718	3.7527E-14	Q9BQS2
isotig15788	-1.4937	2.8833313	0.0308171	XP_002576723	1.966E-108	Q8IXY8
CUFF.234050.1	-1.4922	3.6945308	0.0142698	EKC23244	9.3666E-91	No Hit
isotig22423	-1.4919	3.82955	0.0128015	XP_002579107	1.645E-130	Q9NZM6
isotig21906	-1.4909	2.5734563	0.0468834	XP_003381271	1.6301E-33	Q9GZV3
isotig25541	-1.4902	3.6474172	0.0152934	ELT98433	1.3372E-82	O43603
isotig11796	-1.49	5.8103658	0.0041134	ELU07223	5.648E-130	Q9BXF9
CUFF.300810.1	-1.4895	3.4596057	0.0188473	GAA55980	1.0757E-29	P30626
isotig21482	-1.4887	2.8219385	0.0348378	GAA49259	4.1158E-29	No Hit
CUFF.233422.2	-1.4878	5.3450161	0.0051791	XP_414335	1.425E-174	Q8NE62
isotig26142	-1.4874	2.6161909	0.0449453	XP_002005999	6.4608E-24	O15269
CUFF.142856.1	-1.4871	4.976445	0.0064842	ELT95590	7.8988E-06	No Hit
isotig05896	-1.4861	2.9061182	0.0328566	#N/A	#N/A	No Hit
CUFF.173585.1	-1.486	3.126683	0.0261192	XP_004057985	2.6994E-78	Q4G0P3
isotig03910	-1.4848	5.949018	0.0044662	EKC19365	1.0534E-95	Q96PU9
CUFF.126082.1	-1.4846	4.7156769	0.007705	XP_002734164	0	Q5W041
CUFF.192982.1	-1.4832	7.2650734	0.0035672	AFJ24736	2.9116E-23	No Hit
CUFF.60300.1	-1.4815	4.4155786	0.0099804	EKC29175	1.3593E-76	Q9Y365
CUFF.74088.2	-1.4804	3.6803011	0.0172523	XP_002611479	5.0335E-15	A6NNP5
isotig21003	-1.48	7.2528917	0.0038216	XP_796478	1.942E-171	Q8NE62
isotig24472	-1.4795	3.020637	0.0319492	XP_002579204	1.532E-32	Q6ZVS7
isotig23800	-1.4791	3.768143	0.016939	AAX25214	1.0208E-27	No Hit
isotig19465	-1.4787	6.8846771	0.0041597	EKC30360	7.705E-08	P07602
isotig20326	-1.477	4.0295086	0.0142625	GAA56776	1.9307E-05	I6L8F1

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig19057	-1.4757	2.97244	0.0356943	XP_002736123	3.158E-81	Q53EV4
isotig20961	-1.4748	4.8386838	0.0088829	EJY57493	6.5193E-67	P16157
CUFF.198142.1	-1.4721	2.7807489	0.0456901	ELT96523	2.251E-143	Q15822
isotig20354	-1.4703	2.9369617	0.0404313	EKC21494	1.8409E-18	Q6ZUG5
CUFF.99394.1	-1.4701	3.628771	0.0217279	EKC24412	1.447E-112	Q9H0C1
isotig17884	-1.4699	3.4110838	0.0256823	DAA33898	1.4258E-30	No Hit
isotig09831	-1.4693	3.3225255	0.0286577	GAA49482	3.367E-44	No Hit
isotig24461	-1.4688	4.1602644	0.0151013	GAA27180	9.3632E-46	O95147
CUFF.265888.1	-1.4671	3.215161	0.0328199	GAA48622	0	Q8TD57
isotig25930	-1.465	3.9684554	0.0188048	#N/A	#N/A	No Hit
isotig22907	-1.4622	3.6530107	0.0241727	AAA63593	0	Q8TD57
isotig16785	-1.4619	4.5012448	0.0140072	EHJ64454	3.778E-126	Q9Y6A4
isotig23850	-1.4609	3.7849743	0.0226782	AFK83801	2.394E-27	P62158
CUFF.78240.1	-1.4586	4.5955093	0.0137238	AFJ24739	4.3843E-19	Q6UXB8
isotig13168	-1.4557	3.5850458	0.0284165	EKC22176	1.4501E-58	Q8N9Z9
CUFF.300194.1	-1.4553	3.8770019	0.0230034	GAA31597	2.3176E-22	P30085
isotig20488	-1.455	7.6689883	0.0066139	P81906	3.3425E-10	O95925
CUFF.271149.1	-1.4546	3.9783764	0.0219288	GAA57475	2.019E-129	E7EMB3
isotig21098	-1.4541	4.1960569	0.0192265	EKC26516	1.809E-131	Q8NHU2
isotig15178	-1.4537	4.67743	0.0149024	#N/A	#N/A	No Hit
isotig25789	-1.452	3.4706844	0.0329153	AAY23350	0	O95631
isotig18167	-1.45	4.0902915	0.0209899	#N/A	#N/A	No Hit
isotig10526	-1.4446	4.653477	0.0178042	GAA51305	1.0382E-68	Q86XN7
isotig21167	-1.4441	4.2143183	0.0224775	EKC38147	1.4718E-37	Q9HCF6
isotig13262	-1.4402	4.7564501	0.0182184	AFJ24739	1.643E-67	Q6UWM5
CUFF.223066.1	-1.4371	10.913541	0.0089821	ACR27085	0	Q8NDH3
isotig10982	-1.4371	3.3436914	0.0459201	ELU12161	4.7949E-13	A7E2S9
isotig07891	-1.4369	3.4087074	0.0445495	XP_001947549	1.1422E-72	Q8WZA2
CUFF.131271.1	-1.4362	3.9638719	0.0299992	XP_003047604	3.6996E-06	No Hit
CUFF.209532.1	-1.436	4.0330906	0.0289109	DAA33904	4.3497E-76	Q96JM3
isotig13203	-1.4354	6.6967287	0.0117082	#N/A	#N/A	No Hit
CUFF.190116.1	-1.4341	4.3183151	0.0253524	BAC06342	0	Q9ULK0
isotig24905	-1.433	5.9253629	0.01426	EKC18085	3.322E-164	Q9UIF3
isotig17932	-1.4322	3.951452	0.0324739	#N/A	#N/A	No Hit
isotig22559	-1.4311	3.5368977	0.0421399	ELT87666	4.5581E-78	Q9UFH2
isotig04651	-1.4305	5.4262356	0.0170704	XP_003453162	1.425E-144	Q86U10
isotig22131	-1.4293	3.9566722	0.0338521	GAA33950	1.283E-73	P46531
isotig15756	-1.4291	6.3362228	0.014183	#N/A	#N/A	No Hit
isotig13671	-1.4279	4.692034	0.0233238	NP_001120461	8.511E-100	Q8TD08
isotig06883	-1.4279	5.6503229	0.0169621	EKC31089	8.625E-34	P56539
CUFF.235152.1	-1.4271	4.2187548	0.0304836	XP_002603579	4.2189E-24	Q8N6F8
isotig24887	-1.4249	7.3290782	0.0135305	#N/A	#N/A	Q8TE73
isotig11139	-1.4246	4.8813143	0.0230962	ELU06846	7.776E-100	Q5THR3
isotig06092	-1.4231	4.4894164	0.0281605	#N/A	#N/A	No Hit
CUFF.114122.1	-1.423	5.48276	0.0193559	ELU07223	2.2746E-89	Q9BXF9
isotig13195	-1.422	8.6771482	0.0131998	ACS72290	4.7705E-18	P10646

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig21608	-1.4194	5.0500439	0.0236757	EMC82052	3.2561E-18	Q969S8
isotig07941	-1.4181	4.3771728	0.0328487	GAA51253	0	Q9H5I5
CUFF.317085.1	-1.417	4.1421119	0.036711	XP_790066	8.6207E-18	No Hit
isotig15852	-1.4157	4.2181567	0.0360106	#N/A	#N/A	No Hit
isotig13698	-1.4132	4.177882	0.0388698	ELU15468	6.187E-101	A5D8W1
CUFF.30620.1	-1.4099	7.5752945	0.0181823	#N/A	#N/A	No Hit
isotig08877	-1.4073	7.2098911	0.0195243	XP_002431644	1.603E-143	Q9UBN7
isotig19804	-1.4057	4.369747	0.0396842	Q8AXW7	2.5955E-14	P39877
isotig01275	-1.4052	7.1046823	0.0206502	#N/A	#N/A	No Hit
CUFF.275890.2	-1.4048	4.3207005	0.0410437	BAG15904	5.7717E-78	No Hit
CUFF.280080.1	-1.4046	4.0013146	0.048889	XP_001897507	2.0335E-06	O15484
CUFF.247548.1	-1.4044	4.8037991	0.0338023	XP_002593070	7.0074E-42	Q7RTY0
isotig23416	-1.4039	6.8953162	0.0214665	#N/A	#N/A	Q9P2D7
CUFF.267267.1	-1.4029	4.6423659	0.036758	AEZ03834	9.338E-09	Q8N9I0
isotig15032	-1.3966	4.4126599	0.0459201	XP_002594588	1.7348E-31	Q99932
isotig09935	-1.3917	4.6747354	0.0447004	ELT97299	3.3911E-39	O94759
isotig15574	-1.3884	5.6159162	0.0355672	EKC22192	8.053E-113	Q01581
CUFF.208811.1	-1.3872	5.1396103	0.041143	EKC35700	1.1286E-94	Q9ULI1
CUFF.241477.2	-1.3809	5.9738975	0.0377722	GAA56448	0	O14815
isotig11064	-1.3807	5.8798356	0.0388982	XP_004010837	1.4715E-08	P61916
isotig17423	-1.3806	5.4193676	0.0428697	#N/A	#N/A	No Hit
CUFF.293645.1	-1.3768	5.5406488	0.0445495	#N/A	#N/A	No Hit
isotig10248	-1.3765	6.2275984	0.0394659	ADY41222	0	Q12791
isotig11774	-1.373	5.9046966	0.0447004	#N/A	#N/A	No Hit
isotig24656	-1.3713	6.5535107	0.0417248	BAE94188	0	O15394
CUFF.172880.1	-1.371	7.8078608	0.0376773	AAW27755	6.693E-144	P68371
isotig06340	-1.3707	5.6087334	0.0496086	#N/A	#N/A	No Hit
CUFF.205607.1	-1.3697	10.622967	0.0360358	XP_002161860	0	P68371
isotig17556	-1.3655	6.4188135	0.0474522	NP_001106545	0	Q8N1V2
CUFF.234078.1	-1.3607	7.7595638	0.0462994	#N/A	#N/A	No Hit
isotig15863	-1.3599	7.5352495	0.047655	EFX70245	1.039E-109	P84085
isotig22396	1.3624	6.5972873	0.0493395	#N/A	#N/A	Q8WZ42
isotig13998	1.36323	8.0157603	0.0435492	CAZ33895	0	Q5HY54
isotig17598	1.36438	6.3983107	0.0489204	#N/A	#N/A	No Hit
isotig24540	1.36443	6.2278574	0.0496086	GAA27843	1.122E-171	Q9NTJ3
isotig22683	1.36781	6.8895159	0.0429394	GAA54374	1.173E-116	P98160
isotig22065	1.37106	5.8451967	0.0471912	AAP06498	7.7236E-98	P51911
isotig05482	1.37202	7.9218807	0.0367767	XP_002122451	1.6668E-12	Q14CN2
isotig14014	1.37203	7.633198	0.0372671	GAA32777	1.9179E-57	P59282
CUFF.301775.2	1.37499	5.9855879	0.0421399	#N/A	#N/A	No Hit
CUFF.288264.1	1.37692	5.3185329	0.0475644	GAA52616	3.2854E-36	Q8N1W1
CUFF.270280.1	1.37882	6.9631585	0.0347167	#N/A	#N/A	No Hit
isotig07317	1.37933	6.7564426	0.035095	GAA53233	1.3386E-88	O43491
CUFF.263049.2	1.37956	9.8202519	0.0304752	AAL29934	0	P22897
isotig17419	1.38343	5.2468532	0.0434588	AAH46638	3.8392E-36	Q92614
isotig20294	1.38392	5.8496425	0.03668	#N/A	#N/A	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig22285	1.38405	5.2312535	0.0431153	XP_002574137	1.594E-108	O75150
CUFF.280753.1	1.38485	4.7529702	0.0486983	XP_003750780	1.9881E-11	Q4VXY6
isotig11850	1.38664	5.8850057	0.0348378	AAA92786	1.5515E-72	P55081
isotig23027	1.38675	6.4687049	0.0316943	CCD78816	2.2958E-20	P12270
CUFF.313057.1	1.38737	6.3669678	0.0316943	XP_002571464	7.0449E-66	Q14315
CUFF.136069.1	1.38848	8.0614749	0.0269276	XP_002577959	5.014E-139	O00429
isotig05727	1.38867	6.8488544	0.0293115	NP_001073190	6.2659E-74	O00429
isotig17163	1.39222	5.9840997	0.030836	XP_005146890	7.5233E-62	Q9Y3Z3
isotig18177	1.39439	6.3635328	0.0280079	XP_002595092	1.8121E-12	Q8WXH0
CUFF.152534.1	1.39623	4.5115073	0.0445523	AAQ63200	6.4277E-23	P11277
CUFF.180544.1	1.39684	6.7690263	0.0253274	BAF57623	8.5044E-42	Q92499
isotig22398	1.39731	4.6350909	0.0410437	#N/A	#N/A	Q8WZ42
isotig15244	1.39731	5.500888	0.031173	XP_004923370	1.5674E-13	Q9BXD5
isotig00533	1.39828	4.1605829	0.0498339	#N/A	#N/A	No Hit
isotig14196	1.39836	5.3852628	0.0312813	#N/A	#N/A	No Hit
isotig13332	1.40005	6.4105828	0.0249563	CCD59036	5.213E-124	Q14315
isotig14426	1.4006	5.5302224	0.0293115	AFJ24739	4.8037E-45	Q6UWM5
CUFF.308588.1	1.40081	5.8487599	0.0271019	#N/A	#N/A	No Hit
isotig10645	1.40085	8.760518	0.0203438	XP_004558647	2.8798E-63	Q5T1M5
isotig24918	1.40106	4.8503897	0.035385	AFJ24791	1.6832E-23	Q13114
CUFF.305172.1	1.40712	7.8661359	0.0188622	XP_004073302	1.1539E-24	Q09666
isotig20169	1.40918	4.1222813	0.0428501	#N/A	#N/A	No Hit
isotig17925	1.4105	6.3626145	0.0203438	#N/A	#N/A	No Hit
isotig04051	1.41212	7.415326	0.0175209	XP_003748261	0	E9PEB9
isotig20557	1.41224	6.4452756	0.0195243	ELT89509	0	O60841
isotig02925	1.41289	6.4416316	0.0192263	XP_003389753	1.4754E-44	Q12913
CUFF.223068.1	1.41471	5.664946	0.0218888	XP_002570049	3.6002E-39	Q96T23
CUFF.287671.1	1.41689	4.6119182	0.0299452	Q5BJL5	0	Q9Y2G9
isotig23801	1.41693	4.7478334	0.0279674	#N/A	#N/A	No Hit
CUFF.308785.1	1.41721	4.4076351	0.0321887	#N/A	#N/A	No Hit
isotig24912	1.41853	5.6107784	0.0207964	XP_003285810	9.2639E-07	No Hit
isotig17961	1.41956	8.1433754	0.0142698	#N/A	#N/A	No Hit
isotig18502	1.41978	4.0350638	0.0376773	#N/A	#N/A	No Hit
CUFF.305170.2	1.42069	6.6443735	0.016105	XP_687696	2.24E-12	Q09666
isotig14389	1.42194	6.9713109	0.0149715	#N/A	#N/A	No Hit
CUFF.285306.1	1.42222	4.3595563	0.0307371	EKC27657	2.2186E-56	Q03001
isotig21328	1.42308	3.7182209	0.043516	XP_785018	2.1264E-16	H0YGG5
CUFF.83504.1	1.42343	4.5840025	0.0266421	EOR00235	9.0388E-38	O00429
CUFF.60615.1	1.4237	4.4961589	0.0288005	CCD59345	2.6869E-10	P19338
isotig11073	1.42538	3.4942443	0.049969	CCC91647	1.1925E-08	Q96BK5
isotig14246	1.42736	6.7308076	0.0138898	GAA48156	2.2104E-98	Q9H4G4
isotig03872	1.42746	7.2130091	0.0129901	EKC26094	7.9185E-31	Q04721
isotig15108	1.42775	4.4322505	0.0270575	XP_001783052	1.808E-44	Q9Y6E0
isotig24748	1.42985	5.9781903	0.0149354	XP_001607390	7.0962E-37	O75385
CUFF.118601.2	1.43083	7.0157447	0.0124062	#N/A	#N/A	No Hit
isotig15088	1.43189	6.9592317	0.0121503	#N/A	#N/A	Q5VXJ5

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig11901	1.43236	6.7348669	0.0124062	#N/A	#N/A	No Hit
isotig21165	1.43299	4.0941367	0.0290427	#N/A	#N/A	No Hit
CUFF.193169.1	1.43308	4.3683854	0.0243676	AGA83299	0	P15104
isotig17994	1.43457	4.9208986	0.0193559	NP_001086240	8.1906E-64	P48594
isotig13253	1.43607	3.9342917	0.0304565	XP_003977803	2.3557E-69	Q8NCC3
isotig10685	1.43768	6.9498199	0.0107108	#N/A	#N/A	No Hit
CUFF.66600.1	1.43799	3.3146583	0.047655	#N/A	#N/A	No Hit
contig13546	1.43873	3.906078	0.0305176	#N/A	#N/A	No Hit
CUFF.257855.1	1.43934	4.0801443	0.026493	#N/A	#N/A	No Hit
isotig10741	1.43944	4.4790867	0.0211977	#N/A	#N/A	No Hit
isotig11388	1.43995	4.1064249	0.0264087	XP_005092472	6.181E-27	Q8IWJ2
isotig13333	1.43995	8.5033734	0.0088908	#N/A	#N/A	Q5HY54
isotig10716	1.44029	5.0610572	0.0164067	#N/A	#N/A	O75475
isotig15836	1.44066	5.4207489	0.014183	#N/A	#N/A	No Hit
isotig08462	1.44372	9.0465156	0.0079656	ELU16128	3.5441E-32	M0R387
CUFF.251381.1	1.44417	6.7191625	0.0095193	EKC26094	1.8824E-64	Q16820
isotig23264	1.44515	3.3275979	0.0394307	XP_004370785	2.6993E-82	P11940
isotig02574	1.44578	6.0145282	0.0106256	ELT95888	1.8548E-43	Q12913
isotig09042	1.44756	4.0193487	0.0251653	XP_002596721	1.002E-24	Q7L7X3
CUFF.244268.1	1.44934	6.3921472	0.0091358	CAZ36123	1.039E-152	Q13813
CUFF.67130.1	1.45043	3.8786401	0.0253037	#N/A	#N/A	Q15075
CUFF.326272.1	1.45309	3.540087	0.0307634	#N/A	#N/A	No Hit
isotig23476	1.45313	4.228423	0.0198854	GAA36649	2.8904E-23	O15083
CUFF.313133.1	1.45463	5.5421058	0.0102606	EGW07575	2.4409E-45	Q86YT6
isotig10078	1.45468	9.5711994	0.0060675	Q05870	0	P12883
CUFF.36729.1	1.45551	3.9036087	0.0226782	CCD75970	1.7171E-17	No Hit
isotig15937	1.45676	7.4362979	0.0064842	#N/A	#N/A	No Hit
isotig13202	1.45796	3.294638	0.0355088	#N/A	#N/A	No Hit
isotig21440	1.46042	4.0019073	0.0198854	#N/A	#N/A	No Hit
CUFF.140805.1	1.46156	3.2168252	0.0360106	XP_001604740	1.3914E-32	Q92614
isotig22780	1.46244	7.3418323	0.0057673	CCD58825	2.9318E-11	Q9Y520
isotig10715	1.463	5.0492924	0.0104169	#N/A	#N/A	O75475
CUFF.231807.1	1.46314	3.4874963	0.0289202	#N/A	#N/A	No Hit
isotig23618	1.46447	5.0442197	0.0099175	XP_005091510	3.7562E-48	Q96DT5
CUFF.199217.1	1.46449	3.0135289	0.0412445	XP_003447800	1.518E-160	P11586
isotig18248	1.46469	3.7413998	0.0223278	#N/A	#N/A	No Hit
CUFF.300840.1	1.46491	4.1820799	0.0163188	#N/A	#N/A	No Hit
CUFF.259498.1	1.47049	4.2510654	0.0143875	XP_971393	5.3625E-05	No Hit
isotig24729	1.47068	2.7494281	0.0499293	EFX81466	1.1708E-11	No Hit
CUFF.247418.1	1.47207	3.1797244	0.0316943	#N/A	#N/A	Q9NZW4
isotig22863	1.47348	3.217281	0.0306119	#N/A	#N/A	Q08378
CUFF.282987.1	1.47691	3.0016511	0.0344313	#N/A	#N/A	No Hit
CUFF.56987.1	1.47804	3.9414855	0.0147553	#N/A	#N/A	Q8WZ42
isotig17442	1.48064	4.0471761	0.0139956	#N/A	#N/A	Q02224
isotig20515	1.4829	2.7114682	0.0434323	#N/A	#N/A	No Hit
isotig19061	1.48703	8.71511	0.0028997	BAA34954	0	P13533

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig23244	1.48715	2.9337542	0.0316943	#N/A	#N/A	Q7Z7B0
isotig07425	1.48758	7.158493	0.0032946	ELU08680	8.0732E-21	Q7Z5P9
isotig22207	1.4916	6.1106135	0.0036724	XP_002193085	1.452E-133	O94804
isotig26441	1.49266	2.5279255	0.0468834	#N/A	#N/A	No Hit
CUFF.169829.2	1.49583	3.1710098	0.0216765	NP_001006351	2.3378E-59	A2VEC9
isotig22198	1.49598	7.1865754	0.0026784	XP_004020529	2.8195E-73	Q86YT6
CUFF.244364.1	1.4962	6.9554392	0.0027748	CCD76118	0	Q13813
CUFF.307565.1	1.49712	3.7858936	0.012186	AAX28513	1.728E-25	Q08379
isotig22539	1.49839	3.9287153	0.0104547	CCD75805	1.219E-13	Q7L7X3
CUFF.287376.1	1.49932	2.6451052	0.0378844	#N/A	#N/A	No Hit
CUFF.150920.3	1.50345	4.0394868	0.0087132	#N/A	#N/A	Q8NCU4
isotig20629	1.50362	2.4418971	0.0445495	#N/A	#N/A	No Hit
isotig17939	1.50389	5.9940487	0.002939	XP_002581967	1.5145E-58	Q2HIZ1
isotig11081	1.50766	4.2895848	0.0064082	AAW26183	1.4944E-35	P48788
CUFF.67084.1	1.51024	2.7973749	0.026529	#N/A	#N/A	Q15643
isotig02176	1.51191	7.8235023	0.0016543	#N/A	#N/A	No Hit
isotig26285	1.51559	4.4856969	0.0045985	#N/A	#N/A	No Hit
isotig11418	1.51758	2.8866652	0.0210688	EKC27939	5.5699E-24	P48507
CUFF.72496.1	1.51773	3.6934991	0.0092581	XP_002577000	1.1313E-09	H0YM25
CUFF.34027.1	1.51792	4.4728773	0.0045003	#N/A	#N/A	Q9Y388
CUFF.170637.1	1.51857	5.6584883	0.0023265	#N/A	#N/A	No Hit
isotig23929	1.52094	4.7304454	0.003633	#N/A	#N/A	No Hit
isotig13472	1.52106	2.1474836	0.0491598	EGT58221	2.5942E-12	O95140
CUFF.162944.1	1.52339	3.7656276	0.0075579	ELT98778	9.0714E-61	J3KNX9
isotig13479	1.52579	5.8052015	0.0018492	EMC79134	2.2501E-22	Q5W0W3
CUFF.275869.2	1.52837	3.7963458	0.006867	CCD78771	6.5036E-29	Q15075
CUFF.183206.1	1.53043	8.7281025	0.0009621	#N/A	#N/A	No Hit
isotig22579	1.5338	2.2432977	0.0376034	#N/A	#N/A	No Hit
isotig22989	1.53431	5.6286035	0.0015758	EFN69033	2.5618E-26	Q6IE36
CUFF.197733.2	1.53588	2.7112148	0.0205526	XP_002802251	4.0973E-12	Q5TZA2
isotig23475	1.53878	3.6146887	0.0064842	#N/A	#N/A	Q02224
isotig02041	1.53882	9.9049406	0.0007446	XP_002579884	0	Q13813
isotig17630	1.53906	5.553408	0.0014435	XP_002577236	8.3643E-41	Q9NYW8
CUFF.89745.1	1.54042	2.7565288	0.020479	#N/A	#N/A	Q9UGV2
isotig26413	1.54057	2.2761708	0.0339945	EFN69299	1.3657E-97	P15121
isotig03702	1.54135	6.7199391	0.0009407	XP_004080924	6.5424E-15	A2VEC9
isotig22806	1.54324	6.9773651	0.0008518	#N/A	#N/A	Q8N1M1
isotig11528	1.54509	2.4391849	0.0274436	#N/A	#N/A	P49454
CUFF.228641.1	1.54848	8.7826263	0.0006089	#N/A	#N/A	No Hit
CUFF.152949.1	1.55058	5.0432604	0.0014524	#N/A	#N/A	No Hit
CUFF.243802.1	1.55314	2.5724338	0.0192265	#N/A	#N/A	No Hit
CUFF.72500.1	1.55333	2.6410676	0.0177736	#N/A	#N/A	No Hit
isotig24878	1.55659	7.350195	0.000572	#N/A	#N/A	No Hit
CUFF.228663.2	1.55703	1.9350058	0.0482572	#N/A	#N/A	No Hit
isotig17949	1.55796	3.7911899	0.0045716	#N/A	#N/A	No Hit
CUFF.315928.2	1.55827	7.4577396	0.0005429	#N/A	#N/A	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

CUFF.223997.1	1.55915	5.5212638	0.0009254	CCD78748	0.00011431	No Hit
isotig19583	1.55939	5.6155027	0.0008775	#N/A	#N/A	No Hit
isotig22350	1.56061	6.425512	0.000642	#N/A	#N/A	Q3V6T2
CUFF.326355.1	1.56141	7.3765196	0.0005053	BAA34954	0	P13533
isotig20586	1.56153	6.5083877	0.000601	XP_002166145	8.453E-39	No Hit
CUFF.199313.1	1.56232	3.7665313	0.0036256	GAA54964	2.1253E-22	Q92665
isotig11510	1.56239	3.4168168	0.0055483	#N/A	#N/A	No Hit
CUFF.205499.1	1.56273	3.4087076	0.0051882	#N/A	#N/A	No Hit
isotig09075	1.56474	3.5599558	0.0041787	EMC85203	2.3725E-71	P20813
CUFF.290550.1	1.56684	3.2880895	0.0061041	XP_003211385	7.4826E-09	Q8WX70
isotig25580	1.56852	2.4768861	0.0188622	#N/A	#N/A	Q15075
CUFF.218509.2	1.57069	3.0565134	0.0074732	XP_002574173	9.81E-173	Q14683
contig23292	1.57185	4.1366862	0.0019796	EFN75477	9.055E-63	P14625
isotig22671	1.57205	7.230895	0.000388	EKC22191	2.7072E-47	M0QZD8
CUFF.281031.1	1.5727	3.0368151	0.0073864	#N/A	#N/A	P15924
isotig21273	1.57284	3.1715566	0.0064082	#N/A	#N/A	Q3V6T2
isotig23711	1.57368	2.2160197	0.0252602	#N/A	#N/A	No Hit
isotig03652	1.57733	7.175471	0.0003436	#N/A	#N/A	No Hit
CUFF.215778.1	1.57734	3.5923815	0.0031958	AAA40023	4.4879E-18	P14625
CUFF.174311.1	1.57758	8.7536662	0.0002843	BAA34954	0	P13533
isotig21507	1.57826	6.4236452	0.0003987	ELU00280	7.8258E-36	Q03001
isotig20056	1.5824	2.7219416	0.0102606	XP_001518489	1.6865E-32	Q6VVX0
CUFF.318580.1	1.58382	1.8756959	0.0354675	CAX73979	2.6404E-64	Q13568
CUFF.174294.1	1.58785	2.9104131	0.0070856	AFJ24774	8.252E-149	Q9UQK1
isotig23397	1.58873	3.5742325	0.0024764	XP_001971413	7.124E-119	Q12955
isotig25665	1.58997	3.4802662	0.0029521	#N/A	#N/A	Q13111
CUFF.183184.1	1.5911	8.4704446	0.0002023	#N/A	#N/A	No Hit
isotig24604	1.59289	2.0922064	0.0262822	#N/A	#N/A	No Hit
CUFF.227805.1	1.59411	5.7894571	0.0003442	#N/A	#N/A	No Hit
isotig08698	1.59461	5.6977763	0.0003568	ELU00280	7.6218E-57	Q03001
isotig16172	1.59529	4.8747951	0.0005767	EJT45113	1.9555E-34	P00441
CUFF.8307.1	1.59674	2.1496273	0.020326	#N/A	#N/A	P30622
isotig13048	1.59763	7.687057	0.0001847	BAA34954	0	P12883
isotig07616	1.6009	4.3237414	0.0008929	EKC27073	1.6914E-21	Q8IZT6
isotig22643	1.60216	4.9632174	0.0004711	#N/A	#N/A	Q02224
CUFF.221082.1	1.60325	2.6562019	0.0083884	GAA54774	1.1297E-42	Q9BXS9
isotig22846	1.60398	3.2680223	0.0031615	GAA36806	5.127E-169	Q14683
CUFF.67152.1	1.60617	3.7746019	0.001344	#N/A	#N/A	Q9NZW4
isotig10613	1.60912	7.0443806	0.0001536	ABN79674	2.93E-132	P13533
CUFF.38157.1	1.61247	3.0120683	0.003975	#N/A	#N/A	No Hit
isotig23000	1.61357	2.6765626	0.0067338	#N/A	#N/A	Q8IY85
CUFF.319206.1	1.61366	5.0605175	0.0003228	XP_002579987	2.7259E-32	Q9NR61
isotig13494	1.6139	4.305602	0.0006561	CCD59345	6.3304E-10	P19338
CUFF.31336.1	1.61418	2.4973453	0.0092095	AFJ24792	1.9352E-27	C9J1S8
isotig22957	1.61439	5.18315	0.0003058	#N/A	#N/A	No Hit
isotig08463	1.61454	5.6748129	0.0002117	#N/A	#N/A	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

CUFF.262789.1	1.61494	4.5526188	0.0005152	ELU01156	2.1367E-34	Q3V6T2
isotig11995	1.61607	1.9285466	0.0245272	XP_004225628	7.739E-93	Q92878
isotig24752	1.61816	2.2687586	0.0129901	#N/A	#N/A	No Hit
isotig22964	1.61943	6.9514012	0.0001155	BAA34954	1.45E-124	P12883
isotig20722	1.62173	3.5998289	0.0013176	#N/A	#N/A	No Hit
isotig20291	1.6257	2.7067634	0.0050742	XP_002738175	0	Q71U36
isotig17815	1.62778	6.4275638	0.0001088	#N/A	#N/A	No Hit
isotig07618	1.62821	2.0788893	0.0155525	EKC27073	1.7915E-22	Q8IZT6
isotig13493	1.62879	1.6796124	0.0288891	GAA57060	0	P51795
CUFF.106298.3	1.63039	1.9902081	0.017571	XP_784891	4.6609E-17	Q14203
isotig05356	1.63185	7.4387924	7.49E-05	#N/A	#N/A	Q99996
CUFF.205495.1	1.63278	1.5637322	0.0338023	#N/A	#N/A	No Hit
CUFF.251135.1	1.63346	3.3960931	0.0014296	#N/A	#N/A	No Hit
isotig23798	1.63515	1.4663788	0.0411394	#N/A	#N/A	No Hit
CUFF.84964.1	1.6355	1.3532165	0.0491103	#N/A	#N/A	No Hit
isotig21250	1.63608	6.5684194	8.29E-05	#N/A	#N/A	H0YM25
CUFF.257746.1	1.63677	3.4785487	0.0011943	XP_004773644	5.0677E-29	Q5TB80
isotig10092	1.63902	2.2171153	0.0094362	XP_002733861	0	P48147
contig21268	1.64181	4.2716209	0.0003735	#N/A	#N/A	No Hit
CUFF.131046.1	1.64257	7.3943639	5.52E-05	#N/A	#N/A	Q13439
isotig16278	1.64368	4.7701081	0.0002127	#N/A	#N/A	P13533
isotig24397	1.64504	2.9641887	0.0025977	#N/A	#N/A	Q02224
isotig21883	1.64655	1.3490898	0.0445495	AFL03408	3.8023E-08	Q13114
CUFF.110549.1	1.65	6.7164954	5.28E-05	EKC20356	3.6071E-32	P04114
CUFF.165459.1	1.65355	3.9744201	0.0004268	EKC38542	3.7074E-55	Q3V6T2
CUFF.325669.1	1.65373	2.100694	0.0117082	XP_002570049	5.09E-23	Q96T23
CUFF.265038.1	1.65697	1.8020917	0.0182184	#N/A	#N/A	No Hit
CUFF.177923.1	1.65945	3.216548	0.0011217	#N/A	#N/A	No Hit
isotig26294	1.65993	3.1594065	0.0013551	#N/A	#N/A	No Hit
CUFF.290548.1	1.66238	2.4903518	0.0044483	#N/A	#N/A	Q7Z7A1
CUFF.301052.1	1.66465	4.0653934	0.000263	#N/A	#N/A	No Hit
isotig03465	1.66723	6.0321102	4.54E-05	XP_003703682	3.8429E-57	Q13813
isotig13968	1.66732	1.3282939	0.0385342	#N/A	#N/A	P11532
isotig08423	1.66844	6.3265563	3.78E-05	#N/A	#N/A	Q9NQX4
isotig17213	1.67386	2.9129968	0.0015799	XP_003723292	4.8036E-19	Q92805
CUFF.139826.1	1.67508	1.3110266	0.0343055	#N/A	#N/A	No Hit
CUFF.193587.2	1.67612	1.3981533	0.0291483	GAA49319	1.0491E-19	No Hit
CUFF.205389.1	1.67893	6.1014313	3.04E-05	ELU00280	4.6986E-48	P15924
CUFF.262493.1	1.67901	1.6330324	0.0192311	XP_005111426	8.0214E-85	No Hit
isotig21175	1.68074	1.823518	0.013085	GAA48356	1.0371E-07	No Hit
isotig06220	1.681	5.800333	3.59E-05	AAD28718	0	Q9Y2K3
isotig01971	1.68267	6.3069971	2.57E-05	GAA49076	2.7071E-28	P49747
isotig26420	1.68274	6.8004459	2.12E-05	AAB95253	1.046E-107	Q9UKX2
isotig08235	1.68423	6.430072	2.37E-05	#N/A	#N/A	Q3V6T2
CUFF.177872.1	1.68442	2.7288285	0.0017576	EKC40008	2.0419E-35	No Hit
isotig13971	1.68569	6.0261471	2.78E-05	#N/A	#N/A	Q9UKX2

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig09246	1.69582	4.3685799	8.70E-05	#N/A	#N/A	No Hit
CUFF.147604.1	1.69771	1.7763349	0.0117433	GAA52149	2.043E-18	Q8IUD2
CUFF.279329.1	1.69784	1.5117225	0.0202494	AEI69697	2.832E-133	P11161
isotig09943	1.70233	2.2788258	0.0035372	#N/A	#N/A	No Hit
isotig26362	1.70692	2.2565207	0.0036789	AFB74718	7.3125E-76	Q8IW35
isotig03280	1.71016	7.5894695	7.53E-06	XP_002575931	0	P35580
isotig03855	1.71132	6.911404	8.61E-06	XP_002161429	1.1287E-49	Q9HDC9
CUFF.67086.1	1.71336	1.550191	0.0145237	#N/A	#N/A	P49454
CUFF.134762.1	1.7242	1.8348081	0.0073864	GAA51421	3.2546E-08	Q96T58
CUFF.324123.1	1.72457	4.6377734	3.04E-05	#N/A	#N/A	Q8IY85
CUFF.283819.1	1.72469	2.6809785	0.0010157	#N/A	#N/A	Q8TC20
CUFF.219677.1	1.7322	5.1660729	1.42E-05	CAG23924	1.28E-30	Q504Z1
isotig09593	1.73487	3.9916618	6.29E-05	ACN93794	1.4632E-52	Q7Z449
isotig23261	1.73588	2.7918675	0.0007297	XP_005090907	1.5355E-10	Q8IWJ2
CUFF.252096.4	1.73796	4.0856281	4.76E-05	XP_004539305	1.4193E-13	P35579
isotig20867	1.74025	1.0334308	0.0338571	#N/A	#N/A	No Hit
isotig09860	1.74378	3.1267941	0.0002855	ADF47424	5.561E-24	No Hit
CUFF.67956.3	1.75252	5.0838434	8.82E-06	GAA31575	1.3641E-41	P12270
CUFF.303378.1	1.75404	4.8063721	1.15E-05	#N/A	#N/A	No Hit
isotig21254	1.75663	2.6460455	0.0006645	#N/A	#N/A	No Hit
isotig02586	1.76175	1.8143993	0.0044998	ELT95888	4.2907E-45	Q12913
isotig25888	1.76614	5.9797404	2.66E-06	#N/A	#N/A	No Hit
contig32948	1.76896	3.3119533	0.0001086	#N/A	#N/A	No Hit
isotig06379	1.7697	6.5159228	1.79E-06	#N/A	#N/A	No Hit
isotig24928	1.7749	2.4435787	0.0007925	#N/A	#N/A	No Hit
isotig16429	1.77492	5.7054898	2.70E-06	#N/A	#N/A	No Hit
isotig22617	1.78055	3.8051032	3.11E-05	XP_002409116	6.4029E-27	Q8TBE0
CUFF.26911.1	1.78515	1.2889171	0.0128202	#N/A	#N/A	No Hit
isotig18805	1.78617	3.4728469	4.99E-05	#N/A	#N/A	No Hit
isotig23965	1.7868	0.7870533	0.0354745	#N/A	#N/A	No Hit
CUFF.281929.1	1.78717	2.2655716	0.0009776	#N/A	#N/A	Q86VS8
CUFF.309553.1	1.79011	1.0807189	0.0198854	#N/A	#N/A	No Hit
isotig25750	1.79015	2.8708492	0.0002023	#N/A	#N/A	No Hit
CUFF.253162.2	1.79555	1.6231621	0.0050742	#N/A	#N/A	K7EQA3
CUFF.189336.2	1.8009	5.8848479	1.04E-06	GAA54503	0	P35579
isotig21955	1.80436	1.9070696	0.0023104	XP_002615229	1.1368E-28	No Hit
isotig21587	1.80439	6.1881615	7.89E-07	#N/A	#N/A	Q3V6T2
CUFF.56778.1	1.80535	1.9340841	0.0016719	#N/A	#N/A	Q8WZ42
isotig19734	1.80882	1.3628904	0.0091358	#N/A	#N/A	No Hit
CUFF.20803.1	1.80953	1.6116867	0.0040473	XP_003200508	8.4444E-05	No Hit
CUFF.281630.3	1.81024	3.9002315	1.26E-05	#N/A	#N/A	No Hit
isotig22760	1.81111	1.4397636	0.0064411	#N/A	#N/A	No Hit
CUFF.221090.1	1.81117	3.1008	6.85E-05	GAA54774	3.0578E-29	P40879
CUFF.120578.1	1.81383	5.0226016	1.84E-06	#N/A	#N/A	Q15075
isotig22813	1.81462	5.2607512	1.37E-06	#N/A	#N/A	No Hit
CUFF.151949.1	1.81889	4.502578	3.04E-06	#N/A	#N/A	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig24285	1.81929	0.7914346	0.0290665	#N/A	#N/A	Q5SW02
CUFF.71072.1	1.82396	2.5479938	0.00027	EFN65250	2.9096E-15	P30622
CUFF.201806.1	1.8271	1.1531014	0.0120522	#N/A	#N/A	No Hit
CUFF.279907.1	1.82803	1.0479504	0.0155363	#N/A	#N/A	No Hit
isotig08981	1.83325	2.900886	7.78E-05	#N/A	#N/A	No Hit
isotig06172	1.83517	7.9161769	1.53E-07	GAA28498	2.582E-112	Q5TF21
isotig03468	1.83551	6.5630542	2.48E-07	ELT96713	0	Q01082
CUFF.252093.1	1.83567	3.1054268	4.77E-05	XP_003972149	1.8853E-14	P35579
isotig21412	1.8393	4.0742846	3.71E-06	#N/A	#N/A	Q13464
CUFF.258811.1	1.83989	3.4542632	2.08E-05	XP_005102451	3.1572E-34	Q9UQ35
isotig11186	1.84533	7.7753328	1.16E-07	XP_002165026	1.757E-113	P38571
isotig09007	1.84664	9.5941952	8.50E-08	EFZ17192	7.3344E-09	P07602
isotig22319	1.84941	4.1245975	2.88E-06	#N/A	#N/A	No Hit
isotig16272	1.85161	4.2819762	2.32E-06	AAA29064	1.245E-13	J3QRK5
isotig24467	1.85838	5.3628002	3.47E-07	#N/A	#N/A	P35663
CUFF.4604.1	1.85924	2.3356894	0.000282	GAA54535	2.5612E-13	S4R2Y4
isotig04873	1.86021	2.5036405	0.0001427	NP_568058	1.0161E-06	Q86T82
isotig18904	1.86383	1.4298744	0.0035469	XP_005109533	1.9781E-23	P56539
isotig05070	1.86616	5.8833344	1.52E-07	#N/A	#N/A	No Hit
isotig24324	1.86621	1.7017217	0.0017528	#N/A	#N/A	Q9BYW2
CUFF.56977.1	1.8688	1.042827	0.0102606	#N/A	#N/A	No Hit
isotig15601	1.87003	6.6352554	8.50E-08	#N/A	#N/A	Q14789
CUFF.277439.2	1.87362	0.8577095	0.0163954	BAK55646	1.4031E-27	Q5JZH0
CUFF.242706.1	1.87414	0.8697661	0.0160033	GAA52149	6.0171E-55	Q8IUD2
isotig17526	1.87677	4.2285525	1.10E-06	#N/A	#N/A	No Hit
CUFF.310363.1	1.8774	1.8260355	0.0008518	#N/A	#N/A	No Hit
CUFF.159057.1	1.87746	2.2607674	0.0002408	ELU05355	1.6852E-46	Q8N434
isotig06221	1.87937	10.553343	2.80E-08	BAA34955	0	P13533
CUFF.168610.1	1.88176	2.7016175	6.20E-05	EGT53696	1.8251E-45	P83111
isotig11992	1.88654	0.4256134	0.03595	ELU12706	6.222E-67	Q9UNT1
CUFF.198121.1	1.88683	4.3762781	6.15E-07	#N/A	#N/A	No Hit
CUFF.321144.1	1.88884	6.2325168	6.13E-08	BAA34955	1.769E-170	P12883
CUFF.324350.1	1.89093	4.5780286	4.28E-07	XP_004917420	1.7099E-25	Q5T200
isotig17856	1.89348	2.2261066	0.0002032	#N/A	#N/A	No Hit
isotig17837	1.89365	5.3332235	1.31E-07	#N/A	#N/A	Q15811
isotig20260	1.89668	4.5428666	3.71E-07	#N/A	#N/A	No Hit
isotig00279	1.89734	3.5322333	3.37E-06	#N/A	#N/A	No Hit
CUFF.309615.1	1.89858	2.012058	0.0003769	AFB74713	1.806E-167	Q7Z7A1
isotig09547	1.8989	4.3883533	5.11E-07	#N/A	#N/A	Q3V6T2
isotig08914	1.90019	4.406676	4.91E-07	XP_002573469	1.531E-149	P17844
isotig20367	1.90324	4.8681151	1.83E-07	#N/A	#N/A	P07197
CUFF.64217.1	1.90415	5.3528931	8.50E-08	#N/A	#N/A	No Hit
CUFF.144554.1	1.90513	3.5123711	3.28E-06	#N/A	#N/A	P12270
isotig02844	1.90633	10.746949	1.15E-08	#N/A	#N/A	Q9Y493
isotig13519	1.91079	3.4285857	3.82E-06	#N/A	#N/A	Q13439
CUFF.56981.1	1.91172	1.2570296	0.0034596	XP_002434229	4.7191E-10	Q86TC9

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig24592	1.91184	1.4145355	0.0026566	#N/A	#N/A	No Hit
isotig24642	1.91253	2.0796765	0.0002891	#N/A	#N/A	Q8WVS4
CUFF.258215.1	1.9244	0.1851635	0.0495358	XP_002642935	2.4968E-81	A3KN83
isotig10913	1.92751	6.120938	2.04E-08	BAF57620	1.5808E-19	Q8IX12
CUFF.279130.1	1.93115	0.5995775	0.0201933	#N/A	#N/A	No Hit
isotig19936	1.9312	0.3417708	0.0345956	EKC23192	1.3983E-14	Q16572
CUFF.56983.1	1.93477	1.4551461	0.0014335	#N/A	#N/A	No Hit
CUFF.6400.1	1.93503	3.9463583	5.92E-07	#N/A	#N/A	H0YJ97
isotig02171	1.93524	10.543043	4.54E-09	EFX65941	1.6685E-48	M0R387
CUFF.85020.1	1.93767	1.4806332	0.0014054	#N/A	#N/A	Q03001
isotig22912	1.95105	3.166517	2.95E-06	EGT32873	2.9844E-13	No Hit
CUFF.290258.3	1.95177	3.3700014	2.01E-06	#N/A	#N/A	No Hit
isotig19938	1.95351	2.1879291	9.41E-05	AFB74713	5.864E-156	Q7Z7A1
CUFF.291313.1	1.95684	1.6723826	0.0006249	#N/A	#N/A	No Hit
CUFF.234203.1	1.96191	5.2126419	1.92E-08	AFJ24739	1.3706E-41	Q6UXB8
isotig15894	1.96194	6.8211702	4.23E-09	AAL78671	1.478E-97	P12883
CUFF.137789.1	1.9707	1.7653148	0.0003929	#N/A	#N/A	No Hit
isotig17307	1.97606	5.4691129	8.71E-09	#N/A	#N/A	No Hit
CUFF.327136.1	1.97615	0.9426038	0.0045893	#N/A	#N/A	No Hit
CUFF.330021.1	1.98309	1.4528231	0.0009621	#N/A	#N/A	Q5VXJ5
isotig11488	1.98773	0.4094974	0.0209146	ELT91800	2.7877E-81	Q96IJ6
CUFF.225747.1	1.98934	2.1865013	5.17E-05	#N/A	#N/A	No Hit
isotig25815	1.9904	1.2974137	0.0014032	#N/A	#N/A	No Hit
isotig25475	1.99219	0.5010479	0.0147553	ELU09383	1.468E-82	Q8WVC0
CUFF.89589.1	1.99627	1.9343154	0.0001445	AFB74718	4.02E-106	Q8IW35
CUFF.90716.1	1.99772	0.3358344	0.0216765	GAA55401	1.5735E-33	Q9HC10
CUFF.163316.1	1.99871	5.6024284	3.71E-09	#N/A	#N/A	No Hit
isotig00510	1.9989	0.339906	0.0225888	#N/A	#N/A	No Hit
isotig22976	2.00624	5.772155	2.34E-09	#N/A	#N/A	Q3V6T2
isotig03228	2.00746	2.7349445	5.30E-06	#N/A	#N/A	No Hit
CUFF.293440.1	2.00914	4.9402065	8.39E-09	#N/A	#N/A	Q9UKX3
CUFF.197852.1	2.01296	1.4598062	0.0006089	#N/A	#N/A	Q8IUD2
isotig25322	2.0185	4.8928345	6.56E-09	CAX73132	1.2237E-08	Q7L4I2
CUFF.252114.1	2.0246	2.2785333	2.31E-05	#N/A	#N/A	No Hit
isotig18309	2.02639	3.8757849	7.31E-08	#N/A	#N/A	No Hit
CUFF.268433.1	2.02721	1.6234308	0.0002691	#N/A	#N/A	No Hit
CUFF.327377.1	2.02738	3.6063609	1.40E-07	#N/A	#N/A	No Hit
CUFF.277441.1	2.03812	0.6935659	0.0060675	#N/A	#N/A	M0QY59
CUFF.75786.2	2.04326	4.6091462	6.47E-09	#N/A	#N/A	Q15075
CUFF.71076.1	2.04358	1.2615709	0.0008074	#N/A	#N/A	No Hit
isotig14078	2.05597	5.6060621	7.72E-10	#N/A	#N/A	No Hit
CUFF.24637.3	2.05676	2.8697289	9.09E-07	#N/A	#N/A	No Hit
CUFF.215906.1	2.06057	5.0671645	1.59E-09	#N/A	#N/A	Q66K74
CUFF.251863.1	2.0613	5.7460623	5.53E-10	XP_005094049	8.406E-100	Q14554
isotig21992	2.06575	2.6130092	2.40E-06	GAA56620	7.0713E-78	Q9UPN3
CUFF.40255.1	2.06957	0.0039678	0.0355088	#N/A	#N/A	No Hit

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig13046	2.07293	3.3071157	1.45E-07	#N/A	#N/A	No Hit
CUFF.283609.1	2.07418	0.745389	0.0040662	#N/A	#N/A	No Hit
CUFF.286191.1	2.08235	4.4853315	3.03E-09	#N/A	#N/A	Q02224
isotig25920	2.08778	1.794351	5.60E-05	GAA54535	2.9885E-17	S4R2Y4
isotig26365	2.09589	8.5683053	2.73E-11	#N/A	#N/A	No Hit
isotig07850	2.09848	10.753312	1.98E-11	#N/A	#N/A	No Hit
isotig22799	2.09996	4.9160144	7.13E-10	#N/A	#N/A	No Hit
CUFF.6406.7	2.10433	4.4289125	1.82E-09	XP_001636649	2.5033E-06	Q9BXU7
isotig17794	2.10751	2.6959804	6.82E-07	#N/A	#N/A	No Hit
isotig08834	2.11278	10.099524	1.27E-11	#N/A	#N/A	M0QZD8
isotig01373	2.11683	6.2883611	4.78E-11	XP_005094049	1.872E-126	Q14554
isotig18220	2.1318	2.5055446	1.13E-06	CCD81469	2.2071E-13	Q8N9T8
isotig19902	2.13385	10.83065	6.03E-12	GAA55911	5.0687E-07	No Hit
CUFF.239140.1	2.14096	1.7232557	3.78E-05	GAA48492	4.4708E-25	Q13439
isotig19717	2.14306	9.9469072	4.79E-12	CCD82334	4.7853E-05	No Hit
CUFF.24629.1	2.14926	2.8897083	1.44E-07	#N/A	#N/A	No Hit
isotig11827	2.15097	9.3831923	3.91E-12	CCD82334	0.00049992	No Hit
CUFF.315133.1	2.15218	7.0359641	7.74E-12	ELU04621	1.4183E-95	M0QZD8
isotig21731	2.15433	4.2934816	7.90E-10	XP_003210747	4.3528E-70	Q5BKV1
CUFF.286787.1	2.15608	0.1547898	0.0148386	#N/A	#N/A	Q3V6T2
isotig18053	2.16347	9.7256568	2.43E-12	#N/A	#N/A	No Hit
isotig22100	2.1665	8.8860685	2.47E-12	#N/A	#N/A	No Hit
isotig08028	2.18043	4.4020724	2.42E-10	EKC20683	3.1281E-14	Q9NX58
isotig01872	2.18312	9.066496	1.39E-12	XP_002575931	0	P35579
CUFF.267901.1	2.19689	3.0623641	2.72E-08	#N/A	#N/A	No Hit
isotig18673	2.19824	3.6573661	1.72E-09	#N/A	#N/A	No Hit
isotig20799	2.2003	2.1007951	2.48E-06	#N/A	#N/A	No Hit
isotig17450	2.2125	3.2259154	8.49E-09	#N/A	#N/A	No Hit
CUFF.251864.2	2.21468	3.3975233	3.68E-09	#N/A	#N/A	No Hit
isotig25153	2.2208	1.1376585	0.000228	AFJ24791	2.7225E-24	Q13077
CUFF.24617.1	2.23154	2.8488827	2.52E-08	#N/A	#N/A	No Hit
CUFF.24641.1	2.23301	2.8523775	3.31E-08	#N/A	#N/A	No Hit
isotig18161	2.23316	6.2538828	1.12E-12	AFJ24739	2.2847E-59	Q6UWM5
CUFF.194270.1	2.23383	2.0641694	1.68E-06	#N/A	#N/A	No Hit
isotig15003	2.24667	6.2762762	8.20E-13	#N/A	#N/A	No Hit
isotig12578	2.25704	6.9343506	2.84E-13	#N/A	#N/A	A2VEC9
isotig20787	2.26449	4.027223	8.46E-11	#N/A	#N/A	Q9H2G4
CUFF.170291.1	2.2775	0.238422	0.0058872	NP_001086383	3.0215E-24	Q9NRN7
CUFF.282346.1	2.28085	7.9451702	6.82E-14	#N/A	#N/A	No Hit
CUFF.282973.1	2.28977	0.7719008	0.0005643	#N/A	#N/A	No Hit
isotig20446	2.30114	5.7318179	2.65E-13	AGM37974	1.2879E-14	M0QZD8
CUFF.310043.1	2.30147	5.9688166	1.81E-13	AFJ24739	4.3249E-59	Q6UXB8
isotig05880	2.31489	7.1158307	3.57E-14	#N/A	#N/A	Q9NZW4
isotig06551	2.32591	4.0953014	1.27E-11	#N/A	#N/A	No Hit
isotig03590	2.32592	3.9166585	2.58E-11	#N/A	#N/A	P46821
CUFF.6405.6	2.32961	3.0378805	2.30E-09	#N/A	#N/A	Q15643

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

isotig22786	2.34484	1.2478714	3.87E-05	GAA48580	1.1371E-72	Q8IZC6
isotig03593	2.35606	2.3170111	5.90E-08	#N/A	#N/A	No Hit
isotig08824	2.36079	1.0332463	9.12E-05	#N/A	#N/A	No Hit
isotig23066	2.36581	2.6957435	5.08E-09	#N/A	#N/A	No Hit
CUFF.92860.1	2.37012	4.2807753	1.93E-12	#N/A	#N/A	No Hit
CUFF.24631.1	2.38222	4.2240673	1.55E-12	#N/A	#N/A	No Hit
CUFF.194264.1	2.3863	0.7343126	0.0002872	#N/A	#N/A	No Hit
CUFF.75785.1	2.39028	0.7384185	0.0002434	#N/A	#N/A	No Hit
CUFF.197846.1	2.3951	2.7959256	1.72E-09	XP_002168418	4.8917E-12	Q69YH5
isotig21410	2.3969	2.6516621	3.49E-09	#N/A	#N/A	No Hit
isotig16544	2.39894	0.6077885	0.0005394	XP_002572573	0	P50579
CUFF.24635.1	2.39956	3.1876011	1.44E-10	#N/A	#N/A	No Hit
isotig14320	2.42837	5.7103657	6.25E-15	#N/A	#N/A	No Hit
CUFF.145261.1	2.42966	6.4787097	1.54E-15	AFJ24834	1.1051E-10	Q8IUA0
isotig10825	2.43305	6.0709851	2.44E-15	#N/A	#N/A	No Hit
CUFF.56973.1	2.4349	0.5883518	0.0003521	#N/A	#N/A	No Hit
isotig22544	2.44241	6.8068382	6.79E-16	#N/A	#N/A	P46821
isotig22613	2.457	1.3541322	3.82E-06	#N/A	#N/A	Q96NI6
CUFF.324342.1	2.46328	0.7437147	0.0001251	#N/A	#N/A	No Hit
CUFF.322288.1	2.46738	0.0064429	0.0031175	ADF47415	1.175E-148	Q9UKV8
isotig02073	2.47792	0.7953442	8.11E-05	XP_001770798	4.0028E-16	Q13153
CUFF.194979.3	2.4906	3.4243984	5.80E-12	XP_002607617	4.2639E-45	Q9BV43
CUFF.315000.1	2.51965	6.4312806	7.80E-17	#N/A	#N/A	No Hit
CUFF.56969.1	2.53219	1.4413025	9.59E-07	#N/A	#N/A	No Hit
CUFF.285000.1	2.55992	1.4147788	5.82E-07	ABC25064	5.0164E-23	Q96PD5
isotig24834	2.61066	4.8560978	2.73E-16	#N/A	#N/A	No Hit
isotig19303	2.6155	1.302891	1.08E-06	CAF98974	5.518E-14	P39060
isotig26279	2.64007	5.752134	6.57E-18	#N/A	#N/A	Q15431
isotig22185	2.67527	1.3792966	3.47E-07	GAA35055	8.4533E-20	P52569
CUFF.302188.1	2.6827	1.6119476	4.50E-08	#N/A	#N/A	Q8IWJ2
isotig06070	2.71301	5.566699	9.41E-19	#N/A	#N/A	No Hit
isotig21655	2.71949	2.4691838	2.15E-11	EKC32300	4.9937E-16	Q9NZK7
isotig06505	2.73081	2.9607174	4.72E-13	#N/A	#N/A	No Hit
isotig07559	2.73099	5.8179482	3.44E-19	#N/A	#N/A	No Hit
isotig21210	2.77517	4.6361218	5.36E-18	#N/A	#N/A	No Hit
CUFF.24633.1	2.79295	2.4854315	6.68E-12	#N/A	#N/A	No Hit
isotig20782	2.80631	5.9885261	2.12E-20	#N/A	#N/A	No Hit
isotig16628	2.82738	8.3792559	4.00E-22	#N/A	#N/A	No Hit
CUFF.48626.1	2.8415	0.52773	2.30E-05	GAA34633	5.7591E-35	O43246
contig05717	2.90932	8.2294429	2.66E-23	#N/A	#N/A	No Hit
CUFF.219287.1	2.94435	0.9192513	4.52E-07	CAO79607	0	P68371
CUFF.31680.1	2.97651	7.5764626	4.62E-24	#N/A	#N/A	No Hit
CUFF.198450.1	3.005	4.7140562	4.81E-21	#N/A	#N/A	No Hit
isotig15480	3.03889	2.8575495	3.69E-15	XP_003095974	1.016E-11	P48307
CUFF.283799.1	3.06586	5.7962513	8.41E-24	#N/A	#N/A	No Hit
isotig22930	3.06657	4.6176003	1.59E-21	#N/A	#N/A	P46100

Appendix 3: List of differentially expressed genes in *coe*-deficient animals

contig17595	3.08065	5.7438491	7.22E-24	#N/A	#N/A	No Hit
isotig15920	3.1273	3.3774227	4.09E-18	XP_004068241	1.9844E-32	Q7Z449
CUFF.251417.1	3.15134	7.9373004	1.14E-26	#N/A	#N/A	No Hit
CUFF.252662.1	3.16506	5.999282	2.45E-25	#N/A	#N/A	No Hit
isotig24412	3.2022	2.0868229	6.26E-13	#N/A	#N/A	No Hit
isotig22163	3.36884	2.6506238	7.69E-17	#N/A	#N/A	No Hit
CUFF.176168.1	3.38684	2.0997384	3.57E-14	#N/A	#N/A	No Hit
isotig26074	3.39779	3.0612215	4.10E-19	XP_002610338	3.312E-111	P07098
CUFF.75804.1	3.41217	1.3353843	9.27E-11	#N/A	#N/A	Q15075
isotig14785	3.5655	6.1909469	6.01E-31	#N/A	#N/A	P54108
isotig14225	3.93185	4.2071217	1.38E-29	#N/A	#N/A	No Hit
isotig07757	3.93676	4.8594598	3.10E-32	#N/A	#N/A	No Hit
isotig18351	3.95401	2.2088902	3.83E-18	AFJ24821	7.1829E-34	P10646
CUFF.263942.1	4.2396	4.6464789	8.49E-35	#N/A	#N/A	No Hit
CUFF.235322.1	4.84596	3.0318732	9.07E-30	#N/A	#N/A	No Hit
isotig01778	4.8468	7.6533286	1.18E-48	#N/A	#N/A	No Hit
isotig13967	5.40407	5.3626874	5.60E-49	#N/A	#N/A	No Hit
isotig04500	5.76175	6.6783395	4.45E-57	AAL29937	6.8819E-25	P22897
isotig17750	5.8915	7.2301759	3.19E-59	#N/A	#N/A	No Hit
isotig14434	5.93077	4.3889058	1.11E-47	#N/A	#N/A	No Hit
CUFF.259889.2	5.9321	4.4911338	6.70E-49	#N/A	#N/A	No Hit
isotig04497	6.18779	6.2968925	4.15E-60	AAL29937	4.0236E-25	P11226
isotig24919	6.24534	4.7452733	1.15E-52	#N/A	#N/A	No Hit
CUFF.143111.1	6.35259	6.5215888	3.61E-62	#N/A	#N/A	No Hit
isotig07364	6.81147	4.1318197	3.47E-52	#N/A	#N/A	No Hit
CUFF.283797.1	6.81191	3.2244176	4.63E-43	#N/A	#N/A	No Hit
isotig13000	7.08194	5.1569476	3.95E-62	#N/A	#N/A	No Hit
isotig15350	7.25693	5.2602296	1.49E-64	#N/A	#N/A	No Hit
isotig08591	8.39594	6.3770874	6.57E-79	AAL29940	8.8279E-16	Q9BWP8
isotig14690	8.85341	2.609106	3.02E-43	#N/A	#N/A	No Hit
CUFF.249094.1	9.62097	2.6355756	9.99E-47	#N/A	#N/A	No Hit
isotig12661	10.031	5.8247141	2.69E-86	ADW66116	1.1761E-89	P22897
isotig21593	10.2879	2.8948492	9.76E-52	#N/A	#N/A	No Hit
CUFF.280399.1	10.9173	2.3944872	8.42E-46	#N/A	#N/A	No Hit
isotig16862	13.1252	4.4402816	3.18E-86	#N/A	#N/A	No Hit
isotig22305	26.5095	5.1992526	6.17E-134	ACO82054	1.231E-123	P38571
isotig20350	62.2418	6.1412231	3.92E-194	#N/A	#N/A	No Hit

FC = Fold change
Acc. = Accession number
FDR = False discovery rate
Blast hit acc. = Top human Blast Hit against Human UniProt database
Human Accession numbers used for DAVID analysis

Appendix 4: WISH validation of "downregulated" and "*pmp*" genes.

Gene Name	Discrete	Val.	dFISH	Neural
Smed-secreted peptide prohormone 19 (spp19)	Yes	Yes	Yes	Yes
Smed-potassium voltage-gated channel, Shab-related-like	Yes	Yes	NA	Yes
Smed-signal peptide containing-1 (spc-1)	Yes	Yes	Yes	Yes
isotig24719	No	No	NA	No
isotig21980	Yes	Yes	No	No
Smed-ankyrin repeat protein-2	No	NA	NA	No
isotig19062	No	Yes	NA	No
Smed-secreted peptide prohormone 18 (spp18)	Yes	Yes	Yes	Yes
Smed-T cell acute leukemia (tal)	No	NA	NA	No
Smed-voltage-gated sodium channel (scna-2)	Yes	Yes	NA	Yes
Smed-ankyrin repeat protein like-1	No	NA	NA	No
Smed-gamma-aminobutyric acid receptor subunit gamma like (gbrg)	Yes	Yes	Yes	Yes
Smed-iroquois-1 (irx-1)	Yes	Yes	NA	Yes
CUFF.238332.1	Yes	Yes	NA	No
Smed-tetraspanin like	No	NA	NA	No
isotig24454	No	NA	NA	No
Smed-E3 ubiquitin-protein ligase mindbomb like-1 (mib-1)	No	Yes	NA	No
isotig24035	No	Yes	NA	No
CUFF.1657.1	No	Yes	NA	No
Smed-vesicle-associated membrane protein like-1 (vamp-1)	Yes	Yes	Yes	Yes
isotig19703	No	NA	NA	No
Smed-gli pathogenesis related-2 (glipr-2)	Yes	Yes	Yes	No
Smed-potassium channel subfamily K (lkcnka)	No	NA	NA	No
Smed-gamma-aminobutyric acid receptor subunit beta like (gbrb1)	No	NA	NA	No
Smed-Sodium channel protein-1 (scna-1)	No	NA	NA	No
CUFF.231395.1	Yes	Yes	NA	Yes
isotig14071	No	NA	NA	No
Smed-cerebral peptide prohormone like-1	Yes	Yes	Yes	Yes
Smed-glyine receptor, alpha (glra)	Yes	Yes	NA	Yes
Smed-fas apoptotic inhibitory molecule	No	NA	NA	No
Smed-protein tyrosine non-receptor type like-1	No	NA	NA	Yes
Smed-tetratricopeptide repeat protein 30 like	No	NA	NA	No
isotig19669	No	NA	NA	No
Smed-splicing factor 3b subunit 4	Yes	Yes	NA	Yes
Smed-neurotrypsin-like	No	NA	NA	No
Smed-outer dense fiber protein 3	No	NA	NA	No
Smed-peptidase inhibitor 16	Yes	Yes	NA	No
Smed-cytochrome p450	No	NA	NA	No
isotig14061	No	NA	NA	No
isotig25033	No	NA	NA	No
Smed-secreted peptide prohormone-2 (spp2)	Yes	Yes	Yes	Yes
Smed-pou class 4 transcription factor 3 like-1 (pou4l-1)	Yes	Yes	NA	Yes
Smed-multidrug and toxin extrusion protein like	Yes	Yes	Yes	No
Smed-gamma irradiation insensitive population-2 (gip-2)	Yes	Yes	No	No
Smed-leishmanolysin-like peptidase (LMLN)	No	NA	NA	No
Smed-choline acetyltransferase (ChAT)	Yes	Yes	Yes	Yes
Smed-voltage-gated sodium channel (scna-3)	Yes	Yes	NA	Yes
Smed-voltage-gated sodium channel (scna-1)	Yes	Yes	Yes	Yes
Smed-caveolin-1	Yes	Yes	NA	Yes
Smed-Lipopolysaccharide-induced tumor necrosis factor (litaf)	No	NA	NA	Yes
Smed-nkx2.3 like-1 (nkx2.3)	Yes	Yes	NA	No
Smed-dynein light chain axonemal like-1	No	NA	NA	No

Appendix 4: WISH validation of "downregulated" and "*pmp*" genes .

Smed-BTB/POZ domain-containing protein like	No	NA	NA	No
Smed-hemicentrin-1	Yes	Yes	Yes	No
Smed-T-cell leukemia homeobox protein (tlx)	No	NA	NA	No
Smed-nidogen2 like	Yes	Yes	no	No
Smed-RAS-like, estrogen-regulated, growth inhibitor	Yes	Yes	NA	No
Smed-neuropeptide y prohormone-3	Yes	Yes	NA	Yes
Smed-dual specificity protein phosphatase-1 (dusp-1)	No	No	NA	No
Smed-c16orf80 (c16orf80)	No	NA	NA	No
Smed-netrin-1	Yes	Yes	Yes	Yes
Smed-notch-1	Yes	Yes	NA	No
Smed-dynein heavy chain like	Yes	Yes	Yes	No
Smed-musashi	Yes	Yes	NA	Yes
Smed-neural cell adhesion molecule-2 (ncam-2)	Yes	Yes	Yes	Yes
Smed-WD repeat-containing protein-1 (wdr-1)	No	Yes	NA	No

Identification of differentially expressed genes in postmitotic progenitors following *coe* gene silencing

Gene Name	BPKG ID	SC	Progeny	Diff.
Smed-post mitotic progeny-1 (pmp-1)	NA	NA	NA	NA
Smed-post mitotic progeny-2 (pmp-2)	BPKG20360	5.3	30.6	5.97
Smed-post mitotic progeny-3 (pmp-3)	BPKG5592	2.35	7.45	2.23
Smed-post mitotic progeny-4 (pmp-4)	BPKG21509	61.6	164.19	11.68
Smed-post mitotic progeny-5 (pmp-5)	BPKG20307	3.56	15.26	2.66
Smed-post mitotic progeny-6 (pmp-6)	BPKG15931	2.56	10	2.81
Smed-post mitotic progeny-7 (pmp-7)	BPKG22234	NA	NA	NA
Smed-post mitotic progeny-8 (pmp-8)	NA	NA	NA	NA
Smed-post mitotic progeny-9 (pmp-9)	BPKG864	1502	888.94	231.72
Smed-post mitotic progeny-10 (pmp-10)	BPKG14365	1.53	14.34	3.86

Notes:
Numbers in "SC Expression", "Progeny Expression", and "Diff. Tissues"
are RPKMs (Reads Per Kilobase Mapped), a relative measure of gene expression
used in RNA-seq analysis. Data was collected from Labbe et. al 2012

Discrete = Discreate Expression
Val. = WISH Validation
dFISH = double-labeling with *coe*
Neural = Neural expression pattern
BPKG ID = Corresponding ID number in BPKG transcriptome (Labbe et. al 2012)
SC = Stem cells or X1
Progeny = Postmitotic progenitors or X2
Diff. = Differentiated Tissues or Xins