

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Functional analysis of RING and U-Box E3 ubiquitin ligases involved in stem cell regulation and regeneration in the planarian *Schmidtea mediterranea*

Permalink

<https://escholarship.org/uc/item/49s2333h>

Author

Allen, John Matthew

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO
SAN DIEGO STATE UNIVERSITY

Functional analysis of RING and U-Box E3 ubiquitin ligases involved in stem cell regulation
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

John M. Allen

Committee in Charge:

University of California San Diego

Professor Eric Bennett

Professor David Traver

San Diego State University

Professor Ricardo M. Zayas, Chair

Professor Nicholas J. Shikuma

Professor Robert W. Zeller

2021

The Dissertation of John M Allen is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

San Diego State University

2021

Dedication

This work is dedicated to my father. A scientist in his own right, he taught me that once you have been at it long enough science is mostly yelling at emails. That and cell counting is so easy even a child could to it, demonstrated empirically.

Epigraph

There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity from so simple a beginning endless forms most beautified and most wonderful have been, and are being, evolved.

Charles Darwin

Table of Contents

Dissertation Approval Page.....	iii
Dedication.....	iv
Epigraph.....	v
Table of Contents.....	vi
List of Abbreviations	viii
List of Figures.....	x
List of Tables.....	xi
Acknowledgements.....	xii
Vita.....	xiv
Abstract of the Dissertation.....	xvii
Introduction.....	1
Stem cells in Development and Homeostasis.....	1
Principles of Regeneration and Model Systems.....	9
Regeneration in Invertebrates: Model Systems.....	13
Ubiquitin Signaling in Stem Cells and Regeneration.....	22
NTC Function in Cellular Processes.....	25
Histone Modifications and Epigenetic Regulation of Development by Polycomb Repressive Complexes.....	28
References.....	41
Chapter 1: Dissecting the function of Cullin-RING ubiquitin ligase complex genes in planarian regeneration.....	54
Supplemental Figures.....	63
Supplemental Tables.....	69
Chapter 2: Genetic screen of RING and U-box E3 ligases in planarians identifies critical spliceosomal and epigenetic regulators of stem cell differentiation and specification.....	73
Abstract.....	74
Introduction.....	75
Methods.....	78
Results.....	82

Discussion.....	95
Acknowledgements.....	101
Figures.....	102
Tables.....	117
References.....	145
Conclusion of the Dissertation.....	149
Cullin RING ligase complexes in planarian regeneration.....	150
Spliceosomal and epigenetic ubiquitin ligases are critical regulators of planarian biology.....	152
References.....	162

List of Abbreviations

ASC	Adult Stem Cell
AP	Alkaline Phosphatase
ATR	<i>ataxia telangiectasia</i> mutated and Rad3-related kinase
bp	Base Pair
BLAST	Basic Local Alignment Search Tool
CCR4-Not	Carbon Catabolite Repressed 4—Negative On Tata-less
cDNA	complementary DNA
cPRC1	canonical PRC1
CRL	Cullin RING Ligase
DNA	Deoxyribonucleic Acid
DDR	DNA Damage Repair
DSB	Double Strand Break
ESC	Embryonic Stem Cell
FISH	Fluorescent <i>In Situ</i> Hybridization
gDNA	genomic DNA
GFP	Green Fluorescent Protein
GO	Gene Ontology
<i>hh</i>	hedgehog
iPSC	induced Pluripotent Stem Cell
KD	Knock Down
KO	Knock Out
HSC	Hematopoietic Stem Cell
mRNA	messenger RNA
NTC	NineTeen Complex
PcG	Polycomb Group
pH3	phospho-Histone H3
PTM	Post-Translational Modification

PRC1	Polycomb Repressive Complex 1
PRC2	Polycomb Repressive Complex 2
pre-mRNA	precursor mRNA
PRP19	Pre-mRNA Processing Factor 19
RING	Really Interesting New Gene
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
SAM	Sterile Alpha Motif
SCF	Skp1/Cullin-1/F-box
SCS	Single Cell Sequencing
<i>shh</i>	sonic hedgehog
Skp1	S-phase kinase-associated protein 1
snRNA	small nuclear RNA
snRNP	small nuclear RNP
ssDNA	single stranded DNA
TdT	Terminal deoxynucleotidyl Transferase
vPRC1	variant PRC1
WISH	Whole-Mount <i>In Situ</i> Hybridization

List of Figures

Supplemental Figure S1: Relationships of <i>Schmidtea mediterranea</i> Cullin proteins to those of other species were inferred using the Neighbor-Joining method.....	62
Supplemental Figure S2: Expression analysis of cullins and <i>skp1</i> in <i>Schmidtea mediterranea</i>	64
Supplemental Figure S3: Reported expression of f-box genes in <i>cul1+</i> cells.....	66
Figure 2.1: RNAi screen of RING and U-Box E3 ligases identifies regulators of stem cells and regeneration.....	101
Supplemental Figure S2.1: Whole mount <i>in situ</i> hybridization patterns for genes that showed phenotypes in RNAi screen.....	103
Figure 2.2: Inhibition of <i>prpf19</i> disrupts neoblast function but is not required for stem cell maintenance.....	104
Supplemental Figure S2.2:.....	106
Figure 2.3: <i>prpf19</i> -associated factors and downstream targets recapitulate <i>prpf19(RNAi)</i> phenotypes.....	107
Supplemental Figure S2.3: WISH to NTC core elements <i>cdc5l</i> , <i>pflg1</i> , and <i>spf27</i> , and spliceosomal RNP members <i>prpf3</i> and <i>prpf8</i>	109
Figure 2.4: Inhibition of cPRC1 function disrupts pharyngeal patterning and histone ubiquitylation.....	110
Supplemental Figure S2.4.....	111
Figure 2.5: Loss of PRC1 function causes changes to gene expression levels and spatial patterns.....	112
Supplemental Figure S2.5:.....	114

List of Tables

Supplemental Table S1: Cullin gene present in the planarian <i>Schmidtea mediterranea</i>	68
Supplemental Table S2: Identification and analysis of F-box genes present in the planarian <i>Schmidtea mediterranea</i>	69
Supplemental Table S3: Accession numbers and primers.....	70
Table 2.1: RING and U-box E3 ubiquitin ligases with phenotypes after inhibition in <i>S. mediterranea</i>	115
Table 2.2: Summary of RNA-seq results. Totals of differentially expressed genes at FDR cutoff value < 0.1.....	116
Supplemental Table S2.1 List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases.....	117
Supplemental Table S2.2: List of significantly differentially expressed genes following <i>phc(RNAi)</i> or <i>rnf2(RNAi)</i> at FDR cutoff value of ≤ 0.1	131

Acknowledgements

There is a long list of people that contributed directly to this work, my development as a scientist, or both. Thank you to my professors in graduate school, for teaching me how to think scientifically and critically. Thank you to Drs. Bernstein and Zeller for taking me into their labs for rotation projects. Thank you to the Grainger, Glembotski, Luallen, Shikuma, and Zeller labs for the use of equipment and reagents. Thank you to the support staff in the biology department, Gina, Medora, Cecilia, Jamie, and especially Patti.

Thank you to the ARCS foundation for the amazing and generous personal support. To Helga, Rachel, Holly, and Dr. Atkins for your dedication to running such an amazing organization, and to the donors the Rubin H. Fleet Fund that support my fellowship.

Thank you to my lab members past and present, Dr. Ross for her early training and collegiality, Dr. Tunes for our candid collaborations, both inside and outside of lab, the undergraduates that I have had the pleasure of working with, Celeste, Carolina, and Elizabeth and to Madison, thank you for all the hard work and long hours, it has been a privilege to watch you grow and develop.

To my committee members, Drs. Bennett, Traver, Shikuma, and Zeller, thank you for all that you have done to help me through the years. From instruction as professors, support through letters, your time, and guidance on my project.

And finally, I would like to thank Dr. Zayas for his guidance and mentorship through my graduate training. Thank you for all the opportunities and encouragement that you have given me.

The introduction, in part, is a reprint of the material as it appears in eLS 2016. Allen, JM. Ross, KG. Zayas, RM. The dissertation author was the primary author of this journal article.

Chapter 1, in full, is a reprint of the material as it appears in Developmental Biology 2018. Strand, NS.; Allen, JM.; Ghulam, M.; Taylor, MR.; Munday, RK.; Carrillo, M.; Movsesyan, A.; Zayas, RM. The dissertation author was a primary investigator and author of this manuscript.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Allen, JM.; Balagtas M.; Barajas, E.; Cano, C.; Iberkleid, I.; Zayas, RM. The dissertation author was the primary investigator and author of this manuscript.

Vita

Education

Doctor of Philosophy, Cell and Molecular Biology, September 2021
San Diego State University and University of California, San Diego Joint Doctoral Program
San Diego, CA
Advisor: Dr. Ricardo Zayas
Thesis Title: Functional analysis of RING E3 ubiquitin ligases involved in stem cell regulation and regeneration in the planarian flatworm *Schmidtea mediterranea*.

Embryology: Concepts and Techniques in Modern Developmental Biology, June-July 2018
Marine Biological Laboratory, Woods Hole, MA
Directors: Dr. Rich Schneider, Dr. Dave Sherwood

Bachelor of Science in Molecular Biology, June 2009,
Harvey Mudd College, Claremont, CA
Advisor: Dr. Robert Drewell

Professional Presentations

Aquatic Cell Biology Meeting. San Diego, CA. February 2020. PRC1 function in patterning and regeneration. **John M. Allen**. (Presentation)

Guest Lecturer. Biology 100. San Diego State University, San Diego, CA. March 2019, October 2018. Stem Cells and Regeneration. **John M. Allen**. (Presentation)

ASCB/EMBO 2018 Meeting. San Diego, CA. December 2018. Investigating RING E3 Ubiquitin Ligase Function in Stem Cell Regulation and Regeneration. **John M. Allen**, Ionit Iberkleid and Ricardo M. Zayas. (Poster)

Guest Presentation. EpiCypher Inc. Durham, NC. July 2018. Reverse genetic screen uncovers PRC1 components as critical epigenetic regulators of regeneration in the planarian flatworm. **John M. Allen**, Madison Balagtas, Ricardo Zayas. (Presentation)

10th Annual Student Research Symposium. San Diego State University, San Diego, CA. March 2017. Using the Planarian flatworm to investigate RING E3 ligase function in stem cell regulation and regeneration. **John M. Allen** and Ricardo M. Zayas. (Presentation)

4th European Meeting on Planarian Biology. Sant Feliu de Guíxols, Catalunya. September 2016. A screen to identify RING E3 ubiquitin ligases involved in stem cell regulation *in vivo*. **John M. Allen**, Ionit Iberkleid, Celeste Romero, and Ricardo M. Zayas. (Presentation)

49th Annual Drosophila Research Conference. San Diego, CA. April 2008. Molecular dissection of the IAB5 *cis*-regulatory module in *Drosophila*. Sara E. Goetz, **John M. Allen**, Robert A. Drewell. (Poster)

Publications

Stacy D. Ochoa, Michael R. Dores, **John M. Allen**, Tuan Tran, Maryan Osman, Nidia P. Vázquez Castellanos, JoAnn Trejo and Ricardo M. Zayas. June 2019. A modular laboratory course using planarians to study genes involved in tissue regeneration. In: *Biochemistry and Molecular Biology Education*.

Nicholas S. Strand, **John M. Allen**, Ricardo M. Zayas. March 2019. Post-Translational Regulation of Planarian Regeneration. In: *Seminars in Cell and Developmental Biology*.

Luiza G. Tunes, **John M. Allen**, Ricardo M. Zayas and Rubens L. do Monte-Neto. November 2018. Planarians as models to investigate the bioactivity of gold(I) complexes *in vivo*. In: *Scientific Reports*.

Nicolas S. Strand*, **John M. Allen***, Mahjoobah Ghulam, Matthew R. Taylor, Roma K. Munday, Melissa Carrillo, Artem Movsesyan and Ricardo M. Zayas. January 2018. Dissection the function of Cullin-RING ubiquitin ligase complex genes in planarian regeneration. In: *Developmental Biology* 433(2).
*Equal Authorship

Allen, John M; Ross, Kelly G; and Zayas, Ricardo M. May 2016. Regeneration in Invertebrates: Model Systems. In: eLS. John Wiley & Sons, Ltd: Chichester.

Ho MC, Johnsen H, Goetz SG, Schiller BJ, Bae E, Tran DA, Shur AS, **Allen JM**, Rau C, Bender W, Celniker SE, Drewell RA. November 2009. Functional Evolution of *cis*-Regulatory Modules at a Homeotic Gene in *Drosophila*. In: *PLOS Genetics* 5(11).

Ho MC, Goetz SE, Schiller BJ, **Allen JM**, Drewell RA. 2008. Between transcription and translation: Re-defining RNA and regulation. In *Fly* 2(3).

Honors and Awards

Achievement Award for College Scientists Fellowship, ARCS Foundation, San Diego Chapter, Fall 2017 – current

University Graduate Fellowship, Graduate and Research Affairs, San Diego State University, Fall 2017 - current

Instructionally Related Activities Travel Funding Award, College of Sciences, San Diego State University, October 2011

Dean's List, Harvey Mudd College Fall 2007, Spring 2008, Fall 2008

DePietro Scholarship Award, Harvey Mudd College 2005-2009

Professional Memberships

American Association for the Advancement of Science, 2019-current

American Society for Cell Biology, 2018-current

Society for Developmental Biology, 2016-current

Teaching and Mentoring Experience

Teaching Assistant, Biology 496L: Functional Genomics Lab, San Diego State University (Spring 2017)

Volunteer, San Diego Festival of Science and Engineering, San Diego State University (Spring 2017)

Teaching Assistant, Biology 366L: Biochemistry, Cell and Molecular Biology Lab I, San Diego State University (Spring 2016)

Guest Discussion Leader, Biology 589: Stem Cell and Regenerative Biology, San Diego State University (Fall 2015)

Teaching Assistant, Biology 203L: Cell and Molecular Biology Lab, San Diego State University (Fall 2014)

Professional Experience

Research Associate, Nucleis Inc., Jun. 2013 – Aug. 2014
Supervisor: Steve Danso

Manufacturing Technician, Sangart Inc., Aug. 2011 – Jun. 2013
Supervisor: Robert Burrus

Manufacturing Technician II, SAFC Pharma Carlsbad, Aug. 2009 – Jul. 2011
Supervisor: Mary Tomasini

Student Researcher/Teaching Assistant, Harvey Mudd College, Fall 2007 – Spring 2009,
Advisor: Dr. Robert Drewell

ABSTRACT OF THE DISSERTATION

Functional Analysis of RING E3 Ubiquitin Ligases Involved in Stem Cell Regulation and
Regeneration in the Planarian Flatworm *Schmidtea mediterranea*

by

John M. Allen

Doctor of Philosophy in Biology

University of California San Diego, 2021

San Diego State University, 2021

Professor Ricardo Zayas, Chair

Regeneration is a widely distributed but not universal phenomenon in metazoans that involves the regrowth and repair of lost or damaged body parts that are damaged or lost. The dynamic process of regeneration requires the integration of wound response and

patterning signals to establish a response that can regrow, repattern, and functionally integrate missing body parts. Ubiquitin is a small polypeptide with broad functions in cell biology including protein degradation, subcellular localization, and transcription. The specificity of ubiquitin signaling is controlled by the E3 ligases, a large protein family, that are understudied in the context of regeneration. The E3 ligases often act complexes including CRLs which utilize a *cullin* factor as an organizing scaffold and a substrate recognition factor, an example of which are the *f-box* genes as part of the SCF complex. We used the planarian, *Schmidtea mediterranea*, as a model organism to identify and investigate E3 ligases that regulate stem cells and regeneration. We used RNAi to perturb gene function for 103 RING/U-boxes, six *cullins* and 30 *f-boxes* and found phenotypes for 31 of these genes. We examined *prpf19* and *rnf2* in greater depth and for *prpf19* found, using marker genes and TUNEL, that the basis of the phenotype was not stem cell loss as expected but rather a loss of progeny cells and an increase in apoptosis. *rnf2* ubiquitylates H2A and functions within the epigenetic complex PRC1 to repress transcription. While *rnf2(RNAi)* demonstrated a mild phenotype, inhibition of PRC1 factor *phc* gave a striking phenotype of regional tissue misspecification. To understand the transcriptional targets of *rnf2* and *phc* we used RNA-seq to understand and found surprisingly that *phc* and *rnf2* largely regulated different target genes, explaining the differences in observed phenotypes. Using WISH, we found striking spatial shifts in expression for *phc* target genes after *phc(RNAi)*. These findings demonstrate key roles for E3 ligases in regeneration and stem cells and uncovered a role for cPRC1 in specifying regional tissue identity in planarians.

Introduction of the dissertation

Stem cells in development and homeostasis

The complexity of cell types observed in multicellular organisms is achieved through the coordinated behavior of cells through the process of development. During development cells progressively differentiate into functionally and phenotypically distinct fates. The information that provides this coordinated behavior is encoded by the DNA of an organism and is completely retained in nearly every somatic cell of an organism¹. Differentiated cell states are enabled by a cell type-specific gene expression program and can be preserved through repeated cell divisions. The genetic information that specifies developmental processes must be expressed and then interact with environmental factors that shape and influence development. A degree of plasticity and buffering is then introduced into development to allow acclimatization to changing environmental conditions while still forming a functional body plan. In some situations, including that of sea urchin larva separated at the two-cell stage, the developing embryo can compensate for even severe perturbations and still develop a normal morphology^{2,3}. The action of gene products can be influenced at numerous levels by various factors, by controlling the temporal and spatial expression of the genes themselves⁴, through the modification, suppression and localization of the mRNA^{5,6}, and by chemical modification of proteins⁷, including controlling rates of degradation^{8,9}.

To specify differentiated cell types and to maintain cellular identity throughout the lifespan of an organism requires that certain sets of genes be activated in some cells and silenced in others. At the same time cellular specification is occurring this activity must be

balanced against the proliferative demands of an organism, both during embryogenesis to form the tissues and organs of the developing organism and throughout the lifespan of an organism to support homeostatic turnover of tissues and regenerative processes. The ability to proliferate and self-renew, while also maintaining the ability to differentiate into other, more specialized cell types, is the distinguishing feature of stem cells^{10,11}. Stem cells can be categorized using degree of potency (e.g., totipotent, pluripotent, multipotent, unipotent) and by derivation (e.g., embryonic, adult somatic, induced). Embryonic stem cells (ESCs) are developed early during embryogenesis and before the formation of the primitive germ layers can give rise to any tissues, including extra-embryonic, of the developing organisms and are termed totipotent. As development progresses stem cell potency decreases with the pluripotent cells of the inner cell mass being able to differentiate into any tissue of the embryo proper. Cellular ontology is progressive and sequential through development and spatiotemporally controlled. The fate of a cell during development is dependent upon prior cell states that influence the activation or inhibition of ensuing genetic pathways¹². Mammalian induced pluripotent stem cells (iPSCs) can also be generated *in vitro* from somatic cells through the transduction of four transcription factors (termed Yamanaka Factors), *Oct3/4*, *Sox2*, *c-Myc*, and *Klf4*^{13,14}. Introduction of these factors into differentiated cells causes changes in cell morphology, including a reduction in cell size, loss of somatic cell markers, and an increase in proliferation^{15,16}. Other genes that are enriched in human ESCs were found to be able to substitute for certain Yamanaka factors and induce pluripotency¹⁷. Other evidence exists that lineage specifying factors, in some circumstances, can unexpectedly facilitate reprogramming and replace

reprogramming factors, which suggests that a balance of mutually exclusive lineage specifiers is sufficient to maintain or induce stem cell potency^{18,19}.

The high degree of potentiality for both ESC and iPSCs offers enormous potential for regenerative therapies to treat degenerative diseases and ageing-related disorders. iPSCs offer advantages of being derived from a patient's somatic cells, potentially obviating some issues regarding immuno-rejection, and circumventing any ethical concerns related to the harvesting of ESCs²⁰. The use of iPSCs may not be feasible to treat conditions with an underlying genetic etiology as the underlying genetic defect would still be present in the patient iPSCs. Application of gene therapies would be easier to perform *in vitro* and could be screened for efficacy before reintroduction into a patient. Outside of reintroduction therapies, iPSCs offer additional potential clinical applications in the testing and screening of drug compounds on certain types of cell or in a particular genetic background²¹.

The value of iPSCs as a therapeutic tool is contingent upon the development of a deep understanding of how to reestablish gene regulatory networks that specify cellular identity. Reprogramming of cells by Yamanaka factor induction to a de-differentiated state is initially unstable and will reactivate somatic genes if the transforming factors are removed. Only a small fraction of somatic cells that express Yamanaka factors give rise to iPSC colonies and cells undergoing de-differentiation pass through a number of intermediate cell states that are proliferative but do not develop pluripotency without continued Yamanaka factor exposure^{22,23}. Excess transgenic expression of reprogramming factors can cause cells to adopt a novel cell state that does not exhibit typical ESC-type morphologies but demonstrates pluripotency²⁴. Profiling of intermediate-state cells indicates that partially reprogrammed cells fail to transcriptionally activate pluripotency

genes and that the action of chromatin remodelers are necessary to convert these cells to a pluripotent state²⁵. The early transcriptional dynamics of iPSCs are restricted to pre-existing, accessible chromatin, and that progression from this state requires a concerted change in the somatic epigenome to activate host cell pluripotent factors²⁶. The preservation of the somatic epigenome and limited chromatin remodeling observed early during reprogramming demonstrates the stability of epigenetic regulation of somatic cell identity¹², and illustrates the necessity of understanding the mechanisms of epigenetic regulation of stem cells and progeny. Once the transition to a pluripotent state has occurred, comparison of global methylation profiles indicates that iPSCs are epigenetically more similar to ESCs than their tissue of origin²⁷, and that the methylation profile of iPSCs becomes more like that of ESCs through continued passaging²⁸. The epigenomes of ESCs and iPSCs are not identical with labs reporting that, iPSCs have a preference to preferentially differentiate towards the lineage from which they were originally derived^{29,30}, disease-associated gene imprints that were maintained during iPSC generation^{31,32}, and regions of the chromosome near centrosome and telomeres that are particularly resistant to reprogramming³³.

If cells that are derived from iPSCs are to be reintroduced to a patient the cellular state must be stable, both to ensure continued functioning and prevent any reversion towards an undifferentiated, proliferative state that could result in unchecked cell division. Effective iPSCs therapies will require an understanding of how residual epigenetic marks can affect the potency and stability of derived iPSCs and their re-differentiated progeny. Studying the mechanisms by which cellular identity is established and maintained,

especially in an adult organismal context, will be critical to ensure safe and effective therapies developed from iPSC technology.

Outside of embryogenesis, stem cells are maintained throughout the lifetime of most organisms, termed adult somatic stem cells (ASCs), these cells, in contrast to ESCs typically display limited potency and proliferation. ASCs have been identified for many of the major mammalian tissue and organ systems, including the blood cells, the liver, mammary glands, skin epithelium, intestinal cells, testis, and neurons. Organisms utilize ASCs to replenish and renew tissues, especially epithelial layers subjected to environmental insults, with some tissues like the epidermis, testis, and intestinal epithelia, being constantly self-renewing. Other organ systems, like the liver, pancreas, and lung, undergo little turn over during homeostasis but can mobilize resident ASCs to proliferate in response to damage³⁴. Hematopoietic stem cells (HSCs) reside in the bone marrow³⁵ and are multipotent stem cells³⁶ that sit atop of a hierarchy of progenitor cells that become progressively lineage restricted and give rise to mature blood cells, including red blood cells, megakaryocytes, myeloid cells, and lymphocytes³⁷. As one of the initial ASCs that was characterized³⁸, the HSC established many of the features thought to define ASCs, HSCs are relatively rare and do not divide frequently (quiescence), when division occurs it is asymmetric, creating an actively dividing daughter progenitor and new, quiescent HSC, and that the daughter progenitor progresses down a progressively lineage-restricted, unidirectional differentiation hierarchy³⁹. While many ASCs share traits with HSCs, characterization of ASCs populations in other mammalian tissues and in other model organisms demonstrates a diversity of stem cell types.

ASCs were initially hypothesized to be slow-cycling cells, identified by the retention of labeled DNA, and this characteristic was thought to preserve their proliferative potential and minimize DNA errors induced by replication. While label retention has been a useful parameter to use to screen for potential ASCs and has led to the identification of several ASCs, quiescence as a defining characteristic of stem cells and its utility in identifying novel ASCs, is nondiscriminatory and limited as the majority of mammalian cells are nondividing³⁴. Despite the limitations of using quiescence as a characteristic of ASCs, DNA label retention has been used to identify or confirm the locations of ASCs in their niche. Skeletal muscle has the capacity to regenerate from injury, an ability dependent on mononucleated satellite cells. Satellite cells were first identified based on morphology and were correctly hypothesized, without functional evidence, to be myoblasts that did not differentiate during development and could “recapitulate” embryonic development when muscle damage occurred⁴⁰. Experiments using radioactively labeled thymidine established the satellite cell as mitotically dormant during homeostasis, the cellular source for regenerated muscle fibers, and capable of asymmetric division to generate both satellite cells and differentiated muscle cells⁴¹⁻⁴³. Quiescence as a property of ASCs has also been useful in identifying populations of ASCs that contribute to the maintenance of hair follicles in the epidermis⁴⁴⁻⁴⁶.

Other work from studying epidermal specification challenges the requirement that ASCs must be both quiescent and set strict hierarchical structure. Maintenance of the epidermis and hair follicle involves several, distinct populations of both stem cells and progenitor cells, some of which are highly proliferative under homeostatic conditions, with the interfollicular epidermis being predominantly maintained by a population of

committed progenitor cells that stochastically⁴⁷⁻⁴⁹. During normal homeostatic maintenance heterogeneous stem cell populations independently maintain compartments of tissues, outside of a hierarchical structure^{49,50}. When wounding occurs stem cells contribute substantially to repair⁴⁸ and progeny from multiple distinct stem cell populations acquire lineage plasticity⁴⁹ to contribute to tissue repair. These characteristics of epidermal tissue regulation, multiple stem cell populations, progenitor plasticity, and stochastic determination of progenitor fate, do not agree with a classically designed stem cell hierarchies and might be common in spatially restricted epithelia⁵⁰, including mammary glands^{51,52} and prostate epithelium⁵³. Another example of a tissue system that relies on a system of self-renewal disparate from the HSC paradigm is the liver which normally has low rates of hepatocyte turnover during homeostasis but is capable of remarkable compensatory growth after partial hepatectomy or acute chemically induced injury⁵⁴. The homeostatic turn-over of hepatocytes is likely replenished by pre-existing hepatocytes and substantial contributions of new hepatocytes during regenerative events come from mature hepatocytes that have re-entered the cell cycle^{55,56}.

Differing organisms and even differing tissue systems within an organism have divergent biological properties and are subject to different challenges and likely rely on multiple strategies for self-renewal or regeneration. Some tissues, such as circulating blood cells, rely on a dedicated, quiescent, hierarchical stem cell whereas others rely on distributed sources for cells during renewal, including differentiated cells in certain contexts. A complete understanding of ASC function will require the study of multiple stem cell populations in a variety of contexts, as homeostatic maintenance and regenerative

events can rely on different cellular populations and have different underlying regulatory networks.

Principles of regeneration and model systems

Regeneration is a fundamental capacity of biological organisms: “If there were no regeneration there could be no life. If everything regenerated there would be no death. All organisms exist between these two extremes.”⁵⁷ When considered broadly, regeneration is observed at any considered scale of biological organization. In the broadest sense at the population level, reproduction is necessary to ensure the intergenerational maintenance of a species. At a subcellular level, biological systems exist in a state of flux, with organelles and molecules constantly being turned over and recycled. At an organismal level regeneration is the restoration of a body part that is damaged or lost. Within an organism regeneration can occur at multiple hierarchical levels, cellular, tissue, organ, structure, and even whole-body regeneration. When considered as a phenomenon separate from homeostatic turnover, reparative regeneration is the ability of an organisms to regrow a lost or missing body part and represents a postembryonic recapitulation of developmental processes. Often regenerative events occur as the result of a traumatic injury but should also be considered in the contexts of asexual reproduction and metamorphosis.

The regenerative potential of organisms varies widely with little consistent correlation between the phylogenic position of a species and its regenerative capabilities⁵⁸. The ability to regenerate is fairly common and occurs in most phyla that have been interrogated but the degree of regenerative potency and distribution of regenerative ability varies widely, even between related species. As an example, annelid worms show wide variation in regenerative ability, with both complete regeneration and a lack of regeneration reported for sister clades⁵⁹. Even within an organism variability is observed in regenerative potential between organ systems, with the mammalian liver showing

considerable regenerative ability following hepatectomy while the heart is notably one of the least regenerative organs and responds to injury insults primarily through the formation of scar tissue⁶⁰. This extensive variation in regenerative abilities is not easily explained, but the presence of whole-body regeneration in all basal metazoan lineages and in several lophotrochozoan and deuterostome phyla argues for an deep ancestral origin for regeneration that has been lost and potentially regained throughout metazoan evolution⁶¹.

Understanding the principles of regeneration, and how and why regenerative potential varies drastically between species, will require the study of multiple animal models, especially to understand regeneration in an evolutionary context. The wholesale regrowth of lost appendages is a striking event that was described even in ancient literature by Aristotle^{62,63} and features in myths, including that of Prometheus regenerating his liver every night and that of the Lernaean Hydra regenerating lost heads. Modern experimental investigation of animal regeneration begins with Abraham Trembley in the eighteenth century who wrote describing the regenerative abilities and his surgical manipulations of a small freshwater cnidarian polyp⁶⁴, commonly known as Hydra for the explicit comparisons Trembley made with the Lernaean Hydra^{58,65}. These observations by Trembley disrupted contemporary philosophies on preformation in biology and were followed by studies on regeneration in, among others, earthworms⁶⁶ and salamander appendages⁶⁷, beginning a new field of experimental inquiry into understanding regeneration^{65,68}. This experimental approach to understanding regeneration continues apace using modern imaging and molecular techniques to discover the cellular sources of regenerated tissues and the molecular signals that transduce wound signals and respecify cellular identity. A variety of models, both vertebrate and invertebrate, are employed to

study the process of regeneration, based on phenotypic position, degree of regenerative potential, tractability to laboratory culturing, and amenability to experimental techniques.

Vertebrate models of regeneration draw considerable interest due to their anatomical and evolutionary similarity to humans and thus serve an important role to study the process of regeneration in a context that potentially has more immediate clinical applications. The zebrafish, *Danio rerio*, was established as model for vertebrate embryogenesis but has been adapted to study the developmental process of regeneration. The zebrafish as a model is tractable to both forward and reverse genetic techniques and has large brood sizes with short generation times. One of the first genetic studies of regeneration in zebrafish was a genetic screen for regulators of caudal fin regeneration that utilized temperature-sensitive mutant to recover genes that were embryonic lethal⁶⁹. Cell labelling and tracking has been implemented to understand the identity and source of cells that contribute to the restoration of lost tissues and form the regeneration blastema, with the findings that, like urodele limb regeneration, the cells that form the blastema are dedifferentiated cells that retain their cellular “memory” and in the regenerated structure contribute only to the same lineage as their original derivation^{70,71}. Zebrafish is also a useful vertebrate model for exploring regeneration in organ systems that show extremely limited degrees of regeneration in mammals and includes cardiac⁷² and spinal cord regeneration^{73,74}. Amphibian models, including *Xenopus*, newts, and axolotls, have the ability to regenerate missing limbs to varying degrees. Transgenic fate mapping of cell lineages in axolotl limb regeneration indicates that, like zebrafish, upon injury cells dedifferentiate to proliferate and create lineage-restricted progenitor cells that form the regeneration blastema⁷⁵. Mammals show a reduced degree of regenerative potential when

compared to fish and amphibians, but instances of mammalian regeneration include, male deer antlers⁷⁶, digit tips in young individuals⁷⁷, ear pinna⁷⁸, and full skin regeneration in the African spiny mouse⁷⁹.

Invertebrate models of regeneration have diverse capabilities, which includes, unlike vertebrate models, complete regeneration of any missing structures in some species. These models are essential to study modes of regeneration that are not present in vertebrate models, offer insights into the evolution of regenerative processes, and offer advantages in laboratory culturing and application of experimental techniques.

Regeneration in Invertebrates: Model Systems

John M Allen, *San Diego State University, San Diego, California, USA*

Kelly G Ross, *San Diego State University, San Diego, California, USA*

Ricardo M Zayas, *San Diego State University, San Diego, California, USA*

Introductory article

Article Contents

- Introduction
- Current and Emergent Invertebrate Models of Regeneration
- Mechanisms Underlying Regeneration
- Concluding Remarks
- Acknowledgements

Online posting date: 16th May 2016

Based in part on the previous version of this eLS article 'Regeneration Model Systems: Invertebrate Epimorphic' (2003) by Jaume Baguña.

The ability of animals to replace lost tissues or body parts has captured the imagination of humans for centuries. Animals with simple body parts, such as hydra and planarians, are able to re-form complete organisms from small tissue fragments. Echinoderms such as sea urchins, starfish and sea cucumbers display a broad and diverse range of regenerative capacities that allows for the regeneration of lost organs, limbs or in some cases entire organisms. In contrast, arthropods (insects and crustaceans), which have been the most amenable invertebrates to genetic manipulation, are more limited in regenerative potential but can still faithfully regenerate complex structures of the limb. The diverse modes and capacities for regeneration in invertebrates and the advent of molecular tools to inhibit gene function and study genome-wide changes in gene expression associated with tissue repair provide outstanding opportunities for scientists to decode the cellular and molecular underpinnings of regeneration. Here we provide an overview of prominent invertebrate organisms that are interesting models to investigate stem cell biology, cellular reprogramming and regeneration.

Introduction

If there were no regeneration there could be no life. If everything regenerated there would be no death. All organisms exist between these two extremes.

R.J. GOSS, *PRINCIPLES OF REGENERATION* (1969)

eLS subject area: Developmental Biology

How to cite:

Allen, John M; Ross, Kelly G; and Zayas, Ricardo M (May 2016) Regeneration in Invertebrates: Model Systems. In: eLS. John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0001095.pub2

Regeneration refers to the ability that some organisms possess to replace lost body parts by remodelling remaining tissues or through cell proliferation. This term covers a wide range of abilities, from being able to regrow a part of an amputated appendage to the ability of some organisms to completely regenerate an entire adult form from only a bit of isolated tissue. Regeneration has been documented following asexual reproduction, predation and experimental manipulation (cutting and grafting), in divergent organisms such as jellyfishes, flatworms, annelids, insects, crustaceans, starfish and sea squirts. Over a century ago, Thomas Hunt Morgan proposed terms for two distinct 'general ways' or modes of regeneration that are not mutually exclusive. The first mode he called epimorphosis, 'regeneration in which a proliferation of material precedes the formation of the new part', and the other, morphallaxis, where 'a part is transformed directly into a new organism' without cell proliferation (Morgan, 1901).

The classical example of morphallaxis is that of hydra, which, when bisected, will re-form itself to compensate for the lost tissue, forming a complete adult but remaining at about one half its original size. However, as we discuss in the following section, most data on the cellular basis of regeneration do not support a mechanism of regeneration that occurs strictly via morphallaxis, something that Morgan anticipated in his *Regeneration* monograph. Morphallaxis may be rare in nature and represent a case of an evolutionarily derived trait (i.e. a trait that is different from the common ancestor). Thus, it may be more appropriate to focus the discussion on the cellular basis of tissue remodelling (e.g. cell death, cellular reprogramming) rather than using the often-confused term of morphallaxis.

Epimorphic regeneration is characterised by cell proliferation that forms an undifferentiated structure called a blastema, which later differentiates into the lost tissues. The source of cells in a blastema is the subject of some disagreement and varies between species. Most known instances of regeneration occur primarily through an epimorphic mechanism and there are wide differences in the capacities for epimorphic regeneration in different animal groups. Salamanders and cockroaches are both able to regenerate distal portions of amputated limbs, but neither an isolated salamander hand nor the tip of the cockroach leg can regenerate the rest of the salamander or the cockroach. However, a vast array of animals, notably planarians, annelid worms and starfish, are able to regenerate in both scenarios. The most classic example is the

freshwater planarian, which can regenerate a complete organism from a very small tissue fragment.

Not mentioned in Morgan's original studies on regeneration is a newly discovered mode of regeneration seen in the ephyrae (juvenile form) of the moon jellyfish, *Aurelia aurita*, which has been coined 'symmetrization' (Abrams *et al.*, 2015). Upon injury, *A. aurita* ephyrae do not 'regenerate' lost parts by epimorphosis or morphallaxis. The juveniles will instead reorganize remaining parts to regain the radial symmetry they require for coordinated movement and survival. In this way, although the original form is not altered, functions can be preserved.

Although generating categories is useful in studying biological processes, nature certainly does not have a preference, and modes of regeneration are not necessarily exclusive. The diverse modes to repair tissues that invertebrates have evolved offer known and yet to be discovered opportunities to investigate how tissues respond to injury and re-form to confer organismal function. In the remainder of this article, we highlight examples of classic and emerging invertebrate models of regeneration. We describe what

is currently known about the cellular and molecular mechanisms that govern the remarkable regenerative capacities observed in invertebrates and how this knowledge impacts our ability to address fundamental questions about cellular reprogramming in tissue repair. See also: [Regeneration: Principles](#); [Regeneration of Vertebrate Appendages](#); [Reproduction and Life Cycles in Invertebrates](#)

Current and Emergent Invertebrate Models of Regeneration

The ability to regenerate is a trait that is dispersed throughout animal phyla and cannot be predicted solely on phylogenetic relationships (Figure 1a). Hydra, flatworms, annelid worms, insect and crustacean appendages, various echinoderms (such as sea cucumbers) and ascidians (sea squirts) are amongst the invertebrate model organisms in which regeneration has been

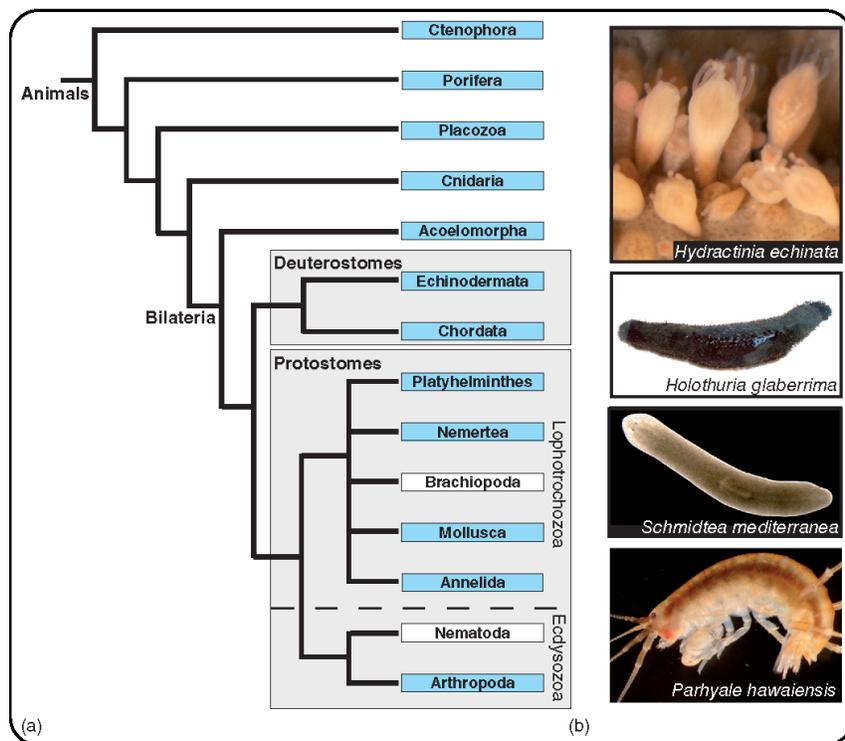


Figure 1 (a) A simplified illustration of a phylogenetic tree denoting diverse animal phyla in which regeneration has been observed (blue boxes). For additional information see: Dunn CW *et al.* (2014) *Annu ev Ecol Syst* 45: 371–395 and Sánchez Alvarado A and Tsonis PA. (2006) *Nat Rev Genet* 7: 873–884. (b) Examples of current invertebrate models of regeneration from pre-bilaterians (Cnidaria: the colonial hydroid *Hydractinia echinata*), deuterostomes (Echinodermata: the sea cucumber *Holothuria glaberima*), and protostomes (Platyhelminthes: the freshwater planarian *Schmidtea mediterranea*; Arthropoda: the crustacean *Parhyale hawaiiensis*). Photos credits: *H. echinata*, Uri Frank; *H. glaberima*, Vladimir Mashanov and José García-Arriarás; *P. hawaiiensis*, Alvina Lai.

investigated in detail. These organisms differ greatly in the ranges of regenerative ability they exhibit, but all share some common molecular underpinnings. While sharing some commonalities, each model organism offers a set of advantages and challenges for researchers, and all have important contributions to make towards a comprehensive understanding of regeneration.

The modern study of regeneration is thought to have been initiated by Abraham Trembley's studies on hydra (Lenhoff and Lenhoff, 1986), a species within the Cnidaria phylum. Cnidaria encompasses the aquatic species of jellyfish, corals, hydra and hydractinia, and are characterised by a specialised cell type, cnidocytes, which are used for capturing prey. These animals display radial symmetry (in contrast to Bilateria) and are generally considered to comprise two germ layers (diploblastic), although recent characterisation of a layer of muscle derived from a mesoderm-like source has led some to consider the Cnidarians as triploblastic (Seipel and Schmid, 2005). Hydra is generally considered to be the canonical example of morphallactic regeneration. Their remarkable ability to regenerate is likely related to their unusual cellular life cycle in which cells are constantly proliferating and being displaced towards the extremities of the animal. As cells are displaced down the body axis of the organism they undergo repatterning and acquire new cellular structures and function depending on their relative location. This ability of hydra cells to form new structures throughout their lifetime confers a remarkable degree of cellular plasticity that is typically not seen in animals and allows hydra to regenerate even when fully disassociated: when single hydra cells are mixed together, they rapidly aggregate and reorganize to form a new animal within days. Recent studies suggest that head regeneration in *Hydra vulgaris* does involve a contribution from proliferating cells (Chera *et al.*, 2009). Furthermore, cell proliferation accompanies head regeneration in the closely related cnidarians *Hydractinia echinata* and *Nematostella vectensis* (sea anemone) (Bradshaw *et al.*, 2015; Passamanek and Martindale, 2012). In hydractinia, stem cell migration and proliferation underlies head regeneration and blastema formation, which is similar to forms of epimorphic regeneration seen in other animals (Bradshaw *et al.*, 2015). In the same species, no blastema is observed during aboral pole (foot) regeneration and polyps transform into the lost body parts. These data indicate that distinct modes of regeneration can exist within the same organism. Cnidarians offer several advantages as a model organism including, ease of culturing in a laboratory and the amenability to molecular manipulations, including RNAi (RNA interference) and transgenesis. The large degree of cellular plasticity seen in cnidarians is fascinating but complicates their use as model organisms because understanding their stem cell biology is difficult. Drawing homologous comparisons of stem cells between cnidarian species, let alone with other animal species, remains challenging and the regulation and role of stem cells in cnidarians might be distinct from how stem cells are conceptualized in Bilateria. The study of cellular renewal in cnidarians offers insight into understanding the origins of tissue-specification and how the role of stem cells first arose. As one of the longest studied models of regeneration, hydra and the rest of Cnidaria continue to be important organisms for studying the diverse methods of regeneration that are observed in nature. **See also: [Regeneration in Hydra](#)**

Planarians, the common freshwater flatworms found in rivers, ponds and lakes all over the world, are excellent model organisms because they are easy to collect and culture in the laboratory (**Figure 1b**). They have a simple body structure (no segments, no coelom, no circulatory or respiratory systems and a blind gut) and have well-defined, polarised body axes (anteroposterior and dorsoventral). Their greatest advantage as a model organism with which to study regeneration has been, and continues to be, the truly astounding regenerative ability they possess. A planarian that is bisected will regenerate two full adult worms and a planarian that is cut into 10 segments will likewise result in 10 full adults being regenerated; this ability has earned planarians the reputation of being 'immortal under the edge of a knife' (Newmark and Sanchez Alvarado, 2002). A large population of pluripotent stem cells that are present in the adult worm confers this regenerative ability in planarians. As one of the first organisms in which regeneration was formally studied, the initial theories of axial gradients setting organismal polarity were developed from experiments in planarians. Today, planarian regeneration is fairly well understood at the tissue and cellular levels and is increasingly so at the molecular level. **See also: [An Introduction to Planarians and Their Stem Cells](#)**

Annelid worms are more highly organised than planarians; they are segmented, have a complete digestive system, extensive respiratory and circulatory systems, and contain a well-developed coelom. Some species demonstrate extensive regenerative ability, both terminal and intercalary, whereas many others show only limited capabilities. Blastema cells originate from dedifferentiation of differentiated cells and not from pluripotent stem cells, which do not seem to be present in annelids. In both anterior and posterior terminal regeneration, the missing terminal structures are the first to be formed with other missing segments filled in via intercalary regeneration. The missing intermediate segments are regrown, as in embryonic development, in an anteroposterior sequence. Innervation is necessary for regenerative growth and the nervous system is involved in establishing the final polarity of the regenerate. However, the cellular basis of regeneration in annelids remains poorly understood and studies at the genetic or molecular level have been scarce, likely because studies in annelids have been distributed over many species. One of the interesting questions that remain to be answered is an evolutionary explanation for the variation in regenerative potential (or lack thereof) between even closely related species. Other avenues of study involve the ecological significance of annelid regeneration, especially the role it plays as a source of renewable biomass for grazing predators. **See also: [Annelida \(Segmented Worms\); Evolutionary Aspects of Annelid Regeneration](#)**

In contrast to planarians and some annelid worms, arthropods, including insects and crustaceans, cannot regenerate a whole animal from an isolated body fragment, but are still able to regenerate lost appendages. In some arthropods, injured limbs are shed in a reflexive response called autotomy where the limb is cast off at a predetermined point proximal to the injury. Regeneration is dependent upon the moulting cycle and is closely related to levels of circulating ecdysteroid hormones that control the cycle of cuticle synthesis. Regeneration is only seen in the adults of organisms where the adult is still capable of moulting. After transection or autotomy, a haemolymph clot closes the wound

and forms an incubation chamber in which the regenerate limb bud forms from the epidermis of the stump. Epidermal cells, especially those close to the wound, begin to divide and form the epidermal blastema. At the next moult, a miniature copy of the lost appendage emerges and after three or four moults it reaches normal size. The cellular nature of the blastema cells remains unresolved as to whether they are pluripotent or retain a 'memory' of the cell type from which they dedifferentiated. Cell lineage tracing experiments in *Parhyale hawaiiensis* (beach hopper or sand flea) used transgenic cell markers to demonstrate that regenerated cell types are derived from distinct, locally present progenitor cells and not a pool of pluripotent stem cells (Konstantinides and Averof, 2014). This mechanism is similar to how blastemas are thought to form during vertebrate limb regeneration, which is generally underpinned by lineage-specific stem cells or progenitors (Kragl *et al.*, 2009; Sandoval-Guzmán *et al.*, 2014). Insect model organisms that can regenerate lost appendages and for which there exist well-developed molecular techniques include cricket nymphs, cockroaches and *Tribolium* flour beetles. These insect species offer powerful models for understanding appendage regeneration and proximal–distal patterning during the regeneration process. The legs of arthropods are especially suited to surgical manipulations and grafting experiments where segments can be grafted into discontinuous conformations (e.g. a more proximal leg segment grafted onto a more distal stump or where the dorsal–ventral (D–V) orientation of the graft is reversed), which often results in unusual regenerates that feature intercalary or even supernumerary regeneration. A complete understanding of regeneration in arthropods will require combining the research in crustaceans, focusing on the role of moulting and hormonal control of regeneration, with that done in insects, which have focused on understanding the molecular mechanisms of patterning control. **See also:** [Regeneration in Crustaceans and Insects](#)

Deuterostomia is a phylogenetic class that is defined by shared developmental features and comprises several invertebrate phyla along with all the vertebrates. This makes invertebrate deuterostomes attractive and important options for study because of their phylogenetic relationship to humans. Invertebrate groups within this clade include echinoderms that range in their regenerative ability to form an entire animal from a body fragment (e.g. starfish) to a restricted capacity to regenerate specific tissues (e.g. sea cucumbers, which eject parts of their digestive system in response to predators and then regenerate), and ascidians, which can regenerate certain lost tissues. Despite the high degree of regenerative potential in echinoderms (Carnevali, 2006), information regarding the origin of the regenerating cells and identification of the structures and factors involved in their morphogenesis has yet to be subjected to systematic analyses using either loss- or gain-of-function techniques, or using modern genome- and proteome-wide profiling approaches. Sea urchins, the best-studied echinoderms for which many molecular tools are available, have not been the subject of systematic studies on regeneration. A major challenge for echinoderms is the difficulty of culturing them in the laboratory, which would facilitate the application of genetic tools. However, the application of genomic tools such as transcriptome profiling that are being applied to echinoderms such as sea cucumbers and brittle stars should

help to yield insights into molecular mechanisms underlying regenerative processes (Czarkwiani *et al.*, 2013; Mashanov *et al.*, 2014; Ortíz-Pineda *et al.*, 2009; Rojas-Cartagena *et al.*, 2007). Ascidians are chordates and are the closest evolutionarily related extant invertebrate group to vertebrates. The regenerative capacity of adult ascidians ranges from complete regeneration from a small cell mass (in species that can reproduce by budding) to the replacement of minor body parts. Regeneration of body parts in solitary ascidians (*Ciona intestinalis*) has been demonstrated to be dependent on a stem cell population that is located near the branchial sac (Jeffery, 2015a). This regenerative potential decreases with age, providing a model with which to understand the role of cellular aging on regeneration and how senescence (programmed cell death) progresses (Jeffery, 2015b). The main advantages of ascidians as model organisms to study regeneration is their close phylogenetic relationship with vertebrates, their vast morphogenetic plasticity and amiability to molecular manipulations. **See also:** [Echinodermata](#); [Chordata \(Chordates\)](#); [Urochordata \(Tunicates\)](#); [Regeneration in Echinoderms and Ascidians](#)

It has been somewhat unfortunate that a majority of invertebrate species that demonstrate high levels of regenerative ability (planarians, annelids and echinoderms) have not been amenable to genetic or transgenic manipulation, whereas organisms that are genetically tractable and for which transgenic methodologies exist, generally show a low degree of regenerative potential (*Drosophila melanogaster*, *Caenorhabditis elegans*). However, recently developed molecular techniques have greatly increased the range of experimental set-ups and the variety of organisms that can be genetically manipulated, including several highly regenerative species. One technique is RNAi, where injection or feeding of double-stranded RNA that targets a particular sequence can induce the cell's endogenous mRNA (messenger ribonucleic acid) decay machinery to efficiently reduce, or knock down, the targeted gene's expression levels. A limitation of RNAi is that the mediated knock down is often incomplete or transient in nature. In some situations where a complete gene knock out is physiologically lethal at an early stage in the life history (e.g. embryogenesis) or early regenerative process (e.g. wound healing) and does not allow later roles in regeneration to be studied, an incomplete knock down by RNAi becomes an advantage, in that it allows for a non-lethal perturbation to assay gene function. RNAi has been implemented in many model organisms; in the planarian *Schmidtea mediterranea*, RNAi has allowed for large-scale screening of candidate genes for roles in wound response, blastema formation and differentiation. More recently, the CRISPR/Cas gene editing system uses a form of acquired immunity to viruses found in bacteria and has been adapted into a powerful and precise gene manipulation system that can target endogenous gene loci for knockout, knockdown, and direct sequence manipulation. CRISPR/Cas combines an approximately 20 base guide ribonucleic acid (gRNA) with a Cas nuclease (an enzyme that can cut a DNA strand). The gRNA will complex with Cas to guide the nuclease to a particular DNA sequence that is complementary to the gRNA. The ease and versatility of designing gRNAs represent a major improvement with respect to ease of use over existing designer nucleases

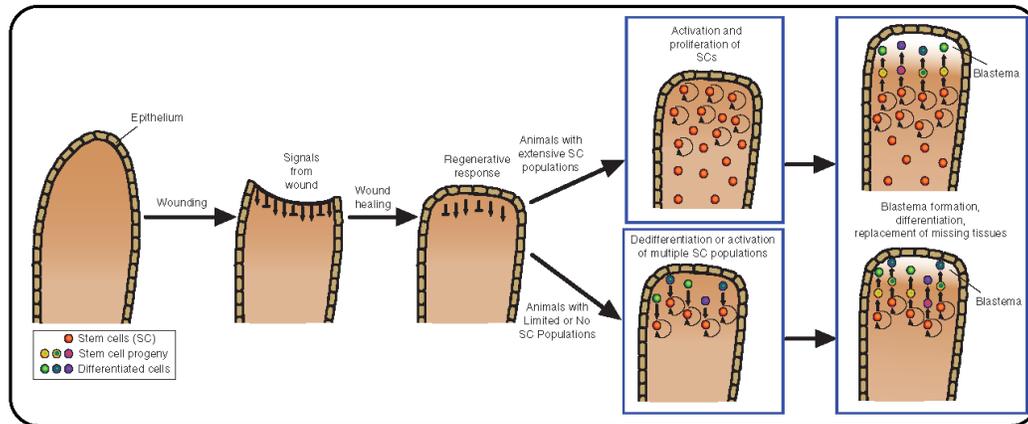


Figure 2 Steps in regeneration and blastema formation in animals. Wounding or tissue loss triggers a wound healing response regardless of an animal's regenerative capacity. Following wound healing, signals from the wound site and pre-existing tissues trigger a regenerative response that involves remodelling of tissues and a proliferative and migratory response of resident stem cells (top). Another mechanism involves proliferation of stem cells and the dedifferentiation or transdifferentiation of cells adjacent to the wound (bottom). In either scenario, proliferating cells ultimately generate the missing tissues that are integrated with pre-existing tissues and then remodelled to achieve the appropriate pattern or scale. Adapted from: Sanchez Alvarado A and Tsonis PA. (2006). © US National Library of Medicine National Institutes of Health.

systems like zinc-finger nucleases or TALEN proteins. Application of CRISPR/Cas has allowed for genetic manipulation of endogenous sequence and transgenesis in a variety of model organisms used for understanding regeneration including, crickets, *Tribolium* beetles, and ascidians, and as is in progress, for the crustacean *P. hawaiiensis*.

Mechanisms Underlying Regeneration

The above model organisms coupled with the discussed molecular techniques have led to the beginning of understanding the process of regeneration at a molecular and cellular level. Questions regarding the source of regenerating cells have been resolved in several instances (e.g. planarians) but remain unknown in others (e.g. annelids, arthropods). After proliferation, it is still necessary for the regenerate to pattern the tissue correctly if the function and form of the lost body fragment are to be regained. Much molecular work has been performed to understand the mechanism of axial gradients and segmental patterning. By understanding how regeneration occurs in these invertebrate models, we hope to fully comprehend the general mechanism of regeneration and how it relates to injury recovery and aging. To know why certain organisms can regenerate and how they do it may allow us to better understand why other organisms, especially humans, cannot. Many of the specific molecular factors involved in tissue regeneration are conserved between species, including species that lack regenerative potential. A striking example of this conservation is the Wnt/ β -catenin pathway, which appears to be implicated in the axial patterning

across Bilateria and, remarkably, the specific use of Wnt to specify posterior identity and Wnt inhibition to specify anterior seems to be largely constant within Bilateria (Petersen and Reddien, 2009a).

Regeneration is triggered by an injury, which elicits a tissue repair programme that can be categorised into four steps: wound healing, formation of a regeneration blastema, differentiation and tissue remodelling (Figure 2). Immediately following acute tissue loss, a crucial step to survival and regeneration is the proper closure and healing of the wound. Closing of a major wound is aided by the contraction of muscle bundles around the site of injury followed by the recruitment of epidermal cells that heal the wound. Gene expression programmes that turn on reparative cellular responses are also turned on by wounds, but much less is known about upstream signals that initiate these programmes. In epimorphosis, proliferating cells form a regeneration blastema once the wound site is healed. The blastema is composed of an accumulation of undifferentiated cells that will eventually give rise to the missing structures. The cells in a blastema are locally derived, either from a population of stem cells or from dedifferentiation of somatic cells near the wound site or both. Blastemas grow fast and become clearly visible as a non-pigmented region in a matter of days. After a sustained growth period, differentiated structures appear in the regenerating region. In principle, regeneration should be complete after several weeks, with regenerates having re-established normal body proportions.

Thus, it is clear that achieving complete regeneration generally requires *de novo* cellular growth. However, attaining the correct shape and proportion is dependent on cellular remodelling of the remaining tissues, which can be accomplished by the removal of excess or unwanted cells (via apoptosis), by changing (reprogramming) the identities, by altering the localization of

pre-existing cells, or by a combination of the aforementioned processes. An elegant example of tissue remodelling takes place during intestinal regeneration in planarians. Following an amputation that includes removal of major intestinal branches, the pre-existing intestinal branches are remodelled into the characteristic triclad (three main intestinal branches) planarian intestine morphology. Remodelling is accompanied by the addition of new intestinal cells derived from the planarian stem cell pool. The planarian intestine provides a robust paradigm to study how epimorphic regeneration and tissue remodelling are integrated to achieve appropriate morphology during tissue regeneration (Forsthoefer *et al.*, 2011).

Following wound response and blastema formation, an organism must begin to re-establish the identity and patterning of tissues that were lost. In order for regeneration to proceed properly, cells must differentiate into the correct cell types and form the appropriate higher-order structures. This requires a cell to receive information that allows it to continue through a positionally correct regenerative pathway. In this sense, regeneration is substantively similar in principle to embryonic development where a single cell must proliferate and eventually generate all adult tissues. Many theories have been postulated on how positional information is determined along an organismal axis but one that is increasingly accumulating evidence is that of positional identity being determined by gradients of morphological substances along the body axis. T.H. Morgan began to develop this concept of formative 'stuffs' that are more or less abundant in different parts of the body in his experiments on earthworm regeneration as early as 1897 (Wolpert, 1991). Morgan used this concept of gradients to elegantly explain the experimental outcome where amputated trunk segments of worms (containing only the medial portions of the organism without heads or tails) will regrow a head in the direction of the anterior cut and a tail along the posterior one. That is, the head would regrow in the direction of the former head and likewise will the tail. This maintenance of anterior–posterior (A–P) polarity through regeneration indicated the presence of some factor that could account for this cellular 'memory' of A–P direction. Further experiments by Morgan demonstrated that very thin transverse medial segments would often regenerate abnormally with heads regrowing at both wound sites. This result fits nicely into the positional gradients theory by demonstrating that abnormal regeneration can occur when a differential gradient cannot be established (due to the segment being too thin).

While Morgan formulated (and later abandoned) the concept of morphological gradients providing the necessary positional information to drive proper regeneration, the precise nature of these formative 'stuffs' remained unknown. Recent experiments performed with molecular biology tools have led to the elucidation of several of the factors and processes in invertebrates that are also shared with those in vertebrates. Many of the factors and pathways identified have also been implicated in the establishment of polarity in development. The Wnt/ β -catenin signalling pathway is responsible for the establishment of the A–P axis during development across the metazoans and is a key player in resetting tissue identity during regeneration.

Wnt and β -catenin are both members of a signal transduction pathway that canonically involves a secreted glycoprotein (*wnt*,

an amalgamation of the *wingless* and *int* homologues in the fly and the mouse, respectively) that can bind to a cell surface marker called Frizzled (Fz). This binding allows the activation of another protein within the cell (Dishevelled) that breaks up a complex of proteins that normally act to degrade β -catenin. Since β -catenin is not degraded when Fz is activated by Wnt signalling, β -catenin can translocate to the cell nucleus where it acts as a transcriptional regulator. In the planarian flatworm, Wnt proteins are present in concentrations that are highest in the tail region and decrease in concentration towards the head. Likewise, Wnt antagonists (molecules that bind or modify Wnt thus preventing it from binding Fz and activating β -catenin) are found in high concentrations in the head with a decreasing concentration towards the tail. A striking set of experiments in planarian flatworms demonstrate the importance of Wnt signalling on establishing A–P tissue identity in regeneration (Figure 3). When levels of Wnt signalling pathway members, *wnt1* and β -catenin, were reduced using RNAi, heads were seen to regenerate in place of tails in posterior-facing wounds. Inhibition of Wnt antagonist (e.g. *notum*) or of APC, a protein involved in the degradation of β -catenin, causes a reciprocal effect of promoting the regeneration of tails in anterior-facing wound sites (Gurley *et al.*, 2008; Petersen and Reddien, 2008, 2009b, 2011). When β -catenin is knocked down in uninjured animals, ectopic heads begin to form along the periphery indicating that the β -catenin signalling pathway is maintained in the adult and is necessary to direct A–P axial polarity during homeostasis. In addition, recent studies examined the role of Wnt signalling in regeneration of defective planarian species that fail to regenerate heads from tail fragments. Remarkably, RNAi knockdown of β -catenin restored regeneration of functional heads in tail pieces! These studies suggest that these 'regeneration-deficient' animals can in fact regenerate if the correct cues can be set. RNAi knock-down experiments in cricket (*Gryllus bimaculatus*) leg regeneration indicate that canonical Wnt signalling is involved in proximal–distal patterning and that an orthologue of β -catenin, *Gb'armadillo*, is necessary for regeneration to occur (Nakamura *et al.*, 2007). These experiments across divergent model organisms strongly indicate a common, evolutionarily conserved basis for patterning networks in regeneration.

Similar to the establishment and maintenance of the A–P axis, the D–V axis of a regenerating planarian is set by a morphological chemical gradient. The morphogen that controls D–V patterning is bone morphogenic protein (BMP), a conserved group of growth factors that promote the formation of bone and cartilage in vertebrates. In planarians, *bmp* is expressed in a gradient with high concentrations found on the dorsal side of the animal. Conversely, a BMP signalling pathway antagonist, *noggin*, is expressed in an opposite gradient with high concentrations found along the ventral surfaces. These gradients work to specify tissue identity during regeneration along the D–V axis. Inhibition of *bmp* using RNAi causes the dorsal surface of the planarian to duplicate ventral organs, with ectopic nerve cords, cephalic ganglia (primitive brain structure) and ciliated epithelial cells during regeneration (Molina *et al.*, 2007; Reddien *et al.*, 2007).

In contrast to hydra and planarians, most other invertebrates that are capable of regeneration can only replace certain tissues or body parts of the adult organism. An excellent example of this

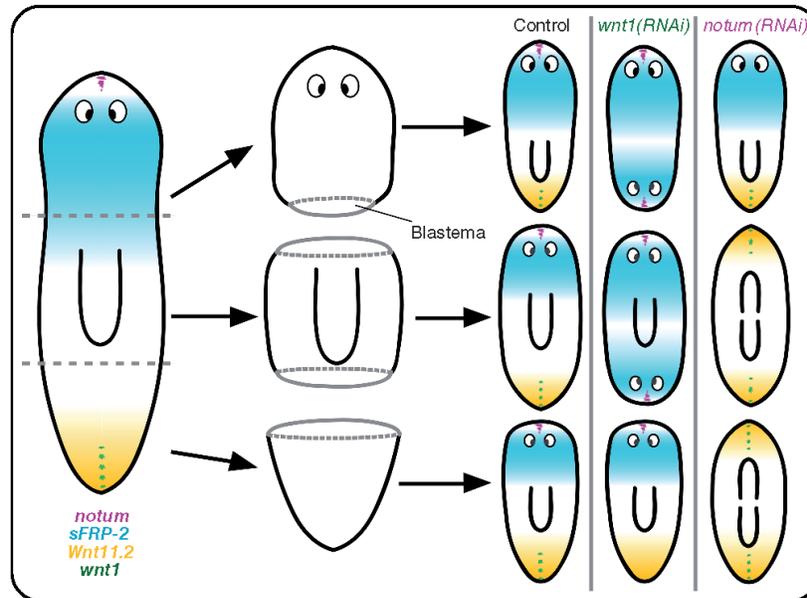


Figure 3 Illustration of a freshwater planarian cut at different body levels (grey dashed lines) to depict the role of canonical Wnt signalling in anterior to posterior body patterning during regeneration. Genes that express inhibitors of Wnt signalling (e.g. *notum* and *sFRP-2*) are expressed in discrete cells located in the anterior pole or in a decreasing gradient along the A–P axis. Certain Wnt genes are expressed in the posterior pole of the animal (e.g. *wnt1* and *wnt11-2*). These patterns of expression are quickly re-established in regenerating animals. Remarkably, decreasing Wnt signalling activity using RNAi leads to incorrect A–P patterning (e.g. the formation of ectopic heads or tails). Adapted from: Adell T, Cebria F and Salo E. (2010) © US National Library of Medicine National Institutes of Health.

restricted regeneration is the ability of many arthropods to regrow lost appendages. Experimentally induced grafts of cockroach nymph limbs have been used to investigate how proximal–distal patterning is set in regenerating limbs. A cockroach leg is made up of a number of segments, arranged along the proximodistal axis in the sequence, coxa, femur, tibia and tarsus. When a distally amputated tibia is grafted onto a host tibia stump cut at a more proximal level, localized growth occurs at the junction between graft and host, and the missing central regions are intercalated. Using cell-specific markers, it has been shown that both proximal and distal pieces contribute to regeneration. This mode of regeneration fits with a cell–cell interaction model where cell fate is determined by the interaction with immediate neighbours (through membrane-bound surface markers). When this interaction is disturbed through injury, normally non-adjacent cells are brought into contact with each other, which signals these cells to begin proliferating. This proliferation will continue until the proper sequence of cell–cell interactions is re-established, thus allowing for the intercalation of missing segments (Figure 4). The general conclusion from these grafting experiments is that from wherever the juxtaposed pieces originate, the structures regenerated are those that would normally lie in between them.

Importantly, it should be noted that the gradient model and cell–cell interaction model of tissue patterning are not mutually exclusive and could work together to ensure recapitulation

of proper tissue differentiation patterning. The gradient model works especially well to create polarity along the body axes while cell–cell interactions can create sharp boundaries between body segments or create mosaic patterns of cell differentiation where appropriate.

Concluding Remarks

Regeneration remains a fascinating phenomenon with many unresolved mechanistic questions remaining to be answered. Invertebrate model organisms offer an amenable system in which to conduct experiments. Combined with an expanding molecular tool kit, including RNAi and CRISPR/Cas, a wide and increasing set of experiments can now be performed which will begin to answer many outstanding questions. The evolutionary basis for regeneration, the selective pressures that caused its loss or gain in diverse phyla, its relation to and distinctiveness from embryonic development, and basic questions regarding stem cell biology and homeostasis are complex fields of study that are only beginning to be explored. Remarkably, the factors identified as being central players in regulating axis polarity during regeneration are the same factors that are involved in patterning during development in vertebrates, including humans. New emphasis in understanding stem cell biology, especially the molecular factors that control the

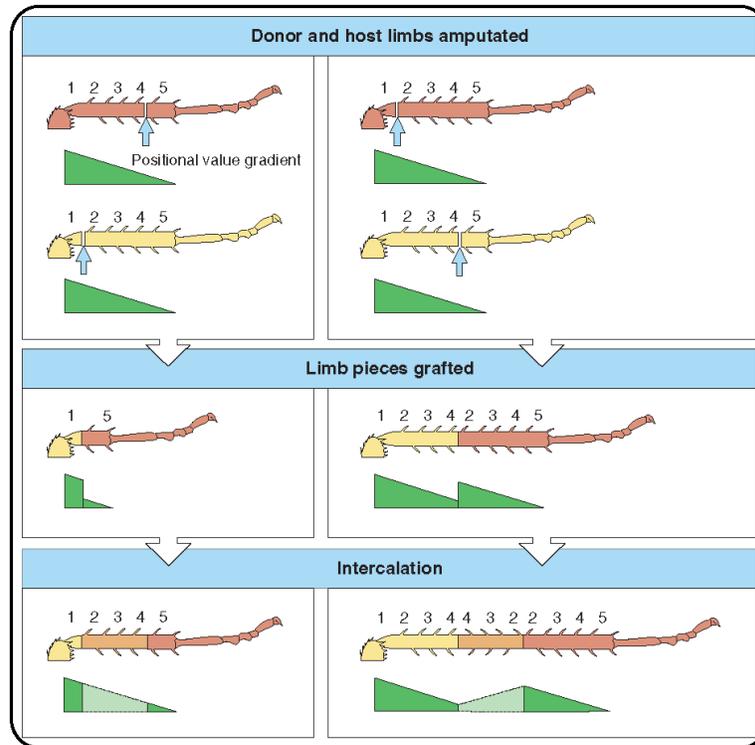


Figure 4 Intercalation of positional values by growth in the regenerating cockroach leg. Left panels: a distally amputated tibia (positional value 5) grafted to a proximally amputated host (positional value 1) induces, regardless of the proximodistal orientation of the graft, the intercalation of the positional values 2–4. A normal tibia is regenerated. Right panels: a proximally amputated tibia (positional value 1) grafted to a distally amputated host (positional value 4) regenerates a longer than normal tibia with reversed polarity as judged by the orientation of surface bristles. The reversed orientation of regeneration is due to the reversal in positional value gradient. The proposed gradient in positional value is shown after each figure. Reproduced with permission from Wolpert and Tickle (2010) © Oxford University Press.

maintenance of pluripotency, provides continuing value for the flatworm as a model organism. Cell-lineage tracing experiments using newly developed CRISPR/Cas technology are resolving the source and nature of blastemal cells in various organisms, of which in some it appears that the cells will re-differentiate into the same cell type from which they were derived, a mechanism also observed in vertebrate limb regeneration. Invertebrate model systems have proved to be valuable in discovering and understanding cellular and molecular mechanisms fundamental for organismal development and function. Regeneration is no different in this regard and the wide diversity of animal life continues to provide numerous avenues to explore.

Acknowledgements

We thank Aziz Aboobaker for helpful discussions and comments on the manuscript.

References

- Abrams MJ, Basinger T, Yuan W, *et al.* (2015) Self-repairing symmetry in jellyfish through mechanically driven reorganization. *Proceedings of the National Academy of Sciences of the United States of America* **112**: E3365–E3373.
- Bradshaw B, Thompson K and Frank U (2015) Distinct mechanisms underlie oral vs aboral regeneration in the cnidarian *Hydractinia echinata*. *Elife* **4**: e05506.
- Carnevali MC (2006) Regeneration in Echinoderms: repair, regrowth, cloning. *Invertebrate Survival Journal* **3**: 64–76.
- Chera S, Ghila L, Dobretz K, *et al.* (2009) Apoptotic cells provide an unexpected source of Wnt3 signaling to drive hydra head regeneration. *Developmental Cell* **17**: 279–289.
- Czarkwiani A, Dylus DV and Oliveri P (2013) Expression of skeletal genes during arm regeneration in the brittle star *Amphiura filiformis*. *Gene Expression Patterns* **13**: 464–472.

- Forsthoefel DJ, Park AE and Newmark PA (2011) Stem cell-based growth, regeneration, and remodeling of the planarian intestine. *Developmental Biology* **356**: 445–459.
- Gurley KA, Rink JC and Sanchez Alvarado A (2008) Beta-catenin defines head versus tail identity during planarian regeneration and homeostasis. *Science* **319**: 323–327.
- Jeffery WR (2015a) Distal regeneration involves the age dependent activity of branchial sac stem cells in the ascidian. *Regeneration (Oxford)* **2**: 1–18.
- Jeffery WR (2015b) Regeneration, stem cells, and aging in the tunicate ciona: insights from the oral siphon. *International Review of Cell and Molecular Biology* **319**: 255–282.
- Konstantinides N and Averof M (2014) A common cellular basis for muscle regeneration in arthropods and vertebrates. *Science* **343**: 788–791.
- Kragl M, Knapp D, Nacu E, *et al.* (2009) Cells keep a memory of their tissue origin during axolotl limb regeneration. *Nature* **460**: 60–65.
- Lenhoff SG and Lenhoff HM (1986) *Hydra and the Birth of Experimental Biology, 1744: Abraham Trembley's Memoirs Concerning the Natural History of a Type of Freshwater Polyp with Arms Shaped Like Horns*. Pacific Grove, California: Boxwood Press.
- Mashanov VS, Zueva OR and García-Arriarán JE (2014) Transcriptional changes during regeneration of the central nervous system in an echinoderm. *BMC Genomics* **15**: 357.
- Molina MD, Saló E and Cebrià F (2007) The BMP pathway is essential for re-specification and maintenance of the dorsoventral axis in regenerating and intact planarians. *Developmental Biology* **311**: 79–94.
- Morgan TH (1901) *Regeneration*. New York: The Macmillan Company.
- Nakamura T, Mito T, Tanaka Y, *et al.* (2007) Involvement of canonical Wnt/Wingless signaling in the determination of the positional values within the leg segment of the cricket *Gryllus bimaculatus*. *Development, Growth & Differentiation* **49**: 79–88.
- Newmark PA and Sanchez Alvarado A (2002) Not your father's planarian: a classic model enters the era of functional genomics. *Nature Reviews Genetics* **3**: 210–219.
- Ortiz-Pineda PA, Ramírez-Gómez F, Perez-Ortiz J, *et al.* (2009) Gene expression profiling of intestinal regeneration in the sea cucumber. *BMC Genomics* **10**: 262.
- Passamaneck YJ and Martindale MQ (2012) Cell proliferation is necessary for the regeneration of oral structures in the anthozoan cnidarian *Nematostella vectensis*. *BMC Developmental Biology* **12**: 34.
- Petersen CP and Reddien PW (2008) Smed-betacatenin-1 is required for anteroposterior blastema polarity in planarian regeneration. *Science* **319**: 327–330.
- Petersen CP and Reddien PW (2009a) Wnt signaling and the polarity of the primary body axis. *Cell* **139**: 1056–1068.
- Petersen CP and Reddien PW (2009b) A wound-induced Wnt expression program controls planarian regeneration polarity. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 17061–17066.
- Petersen CP and Reddien PW (2011) Polarized notum activation at wounds inhibits Wnt function to promote planarian head regeneration. *Science* **332**: 852–855.
- Reddien PW, Bernange AL, Kicza AM, *et al.* (2007) BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration. *Development* **134**: 4043–4051.
- Rojas-Cartagena C, Ortíz-Pineda P, Ramírez-Gómez F, *et al.* (2007) Distinct profiles of expressed sequence tags during intestinal regeneration in the sea cucumber *Holothuria glaberrima*. *Physiological Genomics* **31**: 203–215.
- Sandoval-Guzmán T, Wang H, Khattak S, *et al.* (2014) Fundamental differences in dedifferentiation and stem cell recruitment during skeletal muscle regeneration in two salamander species. *Cell Stem Cell* **14**: 174–187.
- Seipel K and Schmid V (2005) Evolution of striated muscle: jellyfish and the origin of triploblasty. *Developmental Biology* **282**: 14–26.
- Wolpert L (1991) Morgan's ambivalence: a history of gradients and regeneration. In: Dinsmore CE (ed) *A History of Regeneration Research: Milestones in the Evolution of a Science*, pp. 201–217. Cambridge: Cambridge University Press.

Further Reading

- Adell T, Cebrià F and Saló E (2010) Gradients in planarian regeneration and homeostasis. *Cold Spring Harbor Perspectives in Biology* **2**: a000505.
- Dunn CW, Giribet G, Edgecombe GD, *et al.* (2014) Animal phylogeny and its evolutionary implications. *Annual Review of Ecology, Evolution, and Systematics* **45**: 371–395.
- Elliott SA and Sánchez Alvarado A (2013) The history and enduring contributions of planarians to the study of animal regeneration. *Wiley Interdisciplinary Reviews: Developmental Biology* **2**: 301–326.
- Ferretti P and Géraudie J (1998) *Cellular and Molecular Basis of Regeneration: From Invertebrates to Humans*. Chichester, England: John Wiley & Sons.
- Gilbert SF (2013) Postembryonic development: metamorphosis, regeneration, and aging. *Developmental Biology*. Sunderland, MA: Sinauer Associates, Inc..
- Goff R (1969) *Principles of Regeneration*. New York: Academic Press.
- Gold DA and Jacobs DK (2013) Stem cell dynamics in Cnidaria: are there unifying principles? *Development Genes and Evolution* **223**: 53–66.
- Jeffery WR (2015c) Closing the wounds: one hundred and twenty five years of regenerative biology in the ascidian *Ciona intestinalis*. *Genesis* **53**: 48–65.
- Jeffery WR (2015d) The tunicate: a model system for understanding the relationship between regeneration and aging. *Invertebrate Reproduction and Development* **59**: 17–22.
- King RS and Newmark PA (2012) The cell biology of regeneration. *Journal of Cell Biology* **196**: 553–562.
- Poss KD (2010) Advances in understanding tissue regenerative capacity and mechanisms in animals. *Nature Reviews Genetics* **11**: 710–722.
- Sánchez Alvarado A and Tsonis PA (2006) Bridging the regeneration gap: genetic insights from diverse animal models. *Nature Reviews Genetics* **7**: 873–884.
- Tanaka EM and Reddien PW (2011) The cellular basis for animal regeneration. *Developmental Cell* **21**: 172–185.
- Wolpert L and Tickle C (2010) Limb and organ regeneration. *Principles of Development*. Oxford, England: Oxford University Press.

Ubiquitin signaling in stem cells and regeneration

Ubiquitin is a small (≈ 8.5 kD⁸⁰) polypeptide that acts as a post-translational modifier of proteins (including itself) and is expressed ubiquitously in tissues throughout the eukaryotes. Ubiquitylation functions in a wide variety of cellular and regulatory processes, including transcription, cell cycle regulation, translational fidelity, protein turnover, and degradation^{9,81-84}. As a critical signaling molecule, ubiquitin is associated with several diseases, including cancer progression and neurodegenerative diseases^{85,86}. Furthermore, ubiquitylation has been implicated in development, embryonic stem cell fate determination, sea cucumber intestinal regeneration, and in planarian regeneration⁸⁷⁻⁹².

The conjugation of ubiquitin onto a target substrate is typically achieved through a tripartite enzymatic cascade that begins with an E1 ubiquitin-activating enzyme that binds ubiquitin via a thioester bond⁹³. The activated ubiquitin is then transferred to an E2 conjugating enzyme which usually interacts with an E3 ligase to transfer ubiquitin onto a target substrate through an isopeptide bond, most commonly on a lysine residue⁹⁴. Once attached ubiquitylation can occur multiple times, either directly on the substrate again (multi-monoubiquitylation) or on ubiquitin itself, forming ubiquitin chains⁹⁵. The nature of ubiquitin chains can become quite complex as ubiquitin can be elongated at the same residue, usually lysine, each time (homotypic), different residues (heterotypic), or even multiple times at the same ubiquitin forming branched chains⁹⁵. It has been proposed that these ubiquitin chains can be read in the cell as a sort of ubiquitin code, and for some types of chains, notably Lys48-linked chains marking proteins for degradation, the signaling is well understood, while other types of chains remain uncharacterized^{95,96}.

The residue specificity of this cascade is supposed to be largely attributable to the E3 ligases⁹⁷. The E3 ligases are a large and diverse protein family comprising over 600 members in humans⁹⁸ and can be classified into two major classes, the HECTs and RINGs, that differ in their catalytic E2-binding domains⁹⁴. Many of the members of this protein family remain poorly characterized, and in many cases the individual target substrates remain unknown⁹⁹. Understanding the roles of E3 ligases and their target substrates will be essential in understanding ubiquitin biology and the diseases that are caused by ubiquitylation dysregulation. The RINGs, or Really Interesting New Genes, are the largest class of E3 ligases (RING E3s), containing 300 predicted members in humans⁹⁸. The RINGs are defined by the presence of a RING-domain¹⁰⁰, which is a conserved pattern of cysteine and histidine residues that bind two atoms of zinc^{101,102}. Unlike HECT-domain ligases, the RING E3s do not bind ubiquitin directly, but rather bind activated E2 and either directly or through a protein complex bind a target substrate, bringing the two elements into proximity with each other¹⁰³. The RING E3s might also have further activating activity towards bound E2 as an additional regulatory step to prevent spurious ubiquitylation events^{103,104}. RING E3s can act as monomers, homo- or hetero-typic dimers, or as members of larger multi-subunit complexes, an example of which is the Cullin-RING-Ligase (CRL) superfamily¹⁰³. The CRLs contain a Cullin protein that acts as a scaffold to coordinate the remaining members of the complex. This complex will generally contain a RING-domain protein to bind activated E2, various adaptor and activator members, and a substrate recognition receptor¹⁰⁵. The substrate recognition proteins include the F-box, SOCS-box, BTB, and DCAF protein families, and tend to each associate with a particular Cullin. Within each CRL family, differential association with members of the recognition receptor family

allow for modularity in substrate targeting. The RING E3s and CRLs constitute the majority of known ubiquitin ligases, target a diverse array of substrates, and are critical regulatory factors in nearly every aspect of cell biology. A current challenge in understanding RING E3 biology is that substrates for most RING-type E3:E2 pairs remain unknown¹⁰³. As most previous studies on RING E3s have been performed in cell culture or yeast models, understanding of the roles and targets of RING E3s in an in vivo, whole organismal context remains underdeveloped. In particular, understanding the role of ubiquitin signaling in stem cell regulation is insufficiently addressed in ex vivo experiments, given the criticality of surrounding tissues in regulating stem cell niches¹⁰⁶.

NTC function in cellular processes

The Prp19 complex or NineTeen Complex (NTC) is a multifunctional protein complex that is named after its founding member, pre-mRNA processing factor 19 (*PRP19*), and is highly conserved throughout eukaryotes. NTC is a large complex that in *Saccharomyces* consists of eight core proteins and up to 19 associated proteins, which expands to more than 30 associated proteins in higher eukaryotes¹⁰⁷. Human NTC displays heterogeneous complex formation with at least three distinct NTC-like complexes biochemically defined, with a core complex that is comprised of PRP19, CDCL5, PRL1, and SPF27 based on stringent purification conditions¹⁰⁸. NTC functions in diverse cellular processes including splicing, DNA damage repair (DDR), transcriptional regulation¹⁰⁹, protein degradation, and lipid biogenesis^{107,110}.

PRP19 was first identified as *PSO4* in a screen in *Saccharomyces cerevisiae* for mutants that conferred sensitivity to DNA damage from X-rays and psoralen¹¹¹ and contains three recognized motifs, a predicted coil-coil domain, a C-terminal WD40 domain, and an N-terminal U-box¹¹². The U-box is a domain that has E3 ligase activity and is structurally similar to the RING finger domain but lacks the zinc-chelating residues that define the RING protein family¹¹³. While PRP19 has been demonstrated to interact with elements of the proteasome¹¹⁴ and has been genetically shown to be involved in the proteolytic regulation of the cell cycle¹¹⁵ and notch signaling factors¹¹⁶, the direct involvement of PRP19 in ubiquitylating a target substrate for degradation has not been established.

A role for PRP19 in maintaining genomic integrity was suggested from the initial yeast screens that identified *Pso4* mutants as being sensitive to DNA damage. In human cell lines PRP19 was reported using pull down assays to associate with terminal deoxynucleotidyl transferase (TdT) which is involved in repairing DNA double strand breaks (DSB)¹¹⁷. Depletion of *PRP19* delays resumption of DNA replication after chemically induced stalling and is necessary for the timely progression through the cell cycle¹¹⁸. Functional studies have demonstrated that PRP19 can sense single stranded DNA (ssDNA) by binding to RPA-coated ssDNA and that this binding of PRP19 facilitates the accumulation of factors necessary to activate master DNA damage regulator *ataxia telangiectasia* mutated and Rad3-related (ATR) kinase. The ability of PRP19 to activate the DNA damage response is dependent upon its E3 ligase activity¹¹⁹.

The best described role for NTC and PRP19 is its function as a regulator of RNA splicing in the nucleus. Most genes in eukaryotes are first expressed as precursor mRNA (pre-mRNA) that contains both intronic and exonic sequence elements that must be processed to remove introns and ligate the appropriate exons to form mature mRNAs for nuclear export and subsequent translation. This process is catalyzed by the spliceosome, a large macromolecular complex that consists of five small nuclear RNP (snRNPs) and several non-RNP factors and is highly dynamic in its conformation and composition. The snRNPs are composed of a uridine-rich snRNA (U1, U2, U4, U5, or U6) and associated proteins, including NTC¹²⁰. NTC has a critical role in mediating the dynamic interactions of the snRNPs through its E3 ligases activity where it ubiquitylates U4 RNP factor PRP3 with nonproteolytic K63-linked chains. The ubiquitylation of PRP3 increases the affinity of PRP3 for U5 snRNP component PRP8, thus stabilizing the formation of the U4/U6.U5 tri-snRNP.

The E3 ligase activity of PRP19 is counteracted by the combined action of USP4 and the substrate targeting factor, Sart3, to de-ubiquitylate PRP3 and allow for spliceosome disassembly after pre-mRNA processing and the recycling of snRNP¹²¹.

The E3 ligase activity of PRP19 and its function through NTC has a critical regulatory role in many cellular processes but the role for this complex in a developmental context is largely unexplored. Initial identification of *PRP19* was done in yeast screens and much subsequent work has been performed in cell culture models. More recent work has placed PRP19 and spliceosome function as essential regulators of germline stem cells in both *Caenorhabditis elegans* and *Drosophila melanogaster*^{116,122} as well as a role for *prpf19* in a screen for regulators of head regeneration in *S. mediterranea*¹²³.

Histone modifications and epigenetic regulation of development by polycomb repressive complexes

Multicellular organisms utilize a common set of genetic instructions encoded by DNA to specify diverse cell types. These developmental decisions that determine lineage commitment are the effect of coordinated gene expression networks that activate and silence the genes appropriate for differentiated cell identity and function. These gene programs must adapt to both developmental and environmental signals while concurrently be stable through cell divisions and environmental insults to a degree necessary to prevent deleterious neoplasia. The DNA of a eukaryotic cells is organized in a spatial manner beyond a linear code sequence and the nature of this higher order organization of the DNA is one mechanism by which the cell is able to regulate differential gene programs during differentiation. This organization is comprised of a complex of DNA and proteins called chromatin, the basic structural unit of which is the nucleosome of ≈ 146 base pairs (bp) of DNA that is wound around an octamer of histone proteins comprised of a tetramer of H3-H4 combined with two H2A-H2B dimers¹²⁴. The histone protein family contains five major classes, H2A, H2B, H3, H4, and linker histones H1/H5, that are small (between 11 kDa and 21 kDa), rich in basic amino acids, and highly conserved throughout eukaryotes (100 of 102 amino acids are conserved between bovine and pea histone H3)¹²⁵. The core histone proteins are structurally similar and feature a globular core from which terminal “tail” domains extend. The histone proteins, especially the unstructured N-terminal tail regions, are extensively modified using several different biochemical groups, including methylation, acetylation, phosphorylation, and ubiquitylation¹²⁶. These histone modifications are strongly associated with gene transcriptional states and have been proposed to work

together to form a “histone code” that can be read out by the cell¹²⁷. Interestingly, the same modification can be associated with differential transcriptional outputs based on which histone protein is modified (i.e. monoubiquitylation of H2B-Lys120 and H2A-Lys119 are associated with transcriptional activation and silencing respectively^{128,129}) or even which amino acid residue that is modified on a particular histone protein (i.e. methylation of lysine 4 and lysine 27 on histone H3 are associated with transcriptional activation and silencing respectively^{130,131}). Each post-translational modification (PTM) is associated with associated factors that are necessary to create these modifications, “writers”, those that can recognize and bind to these marks, “readers”, and often a set of enzymes necessary to remove these modification, “erasers”¹³². The identification of specific domains that can recognize modified histones to recruit other factors suggests a role in recruiting trans-acting factors to regulate transcription, condensation, and DNA repair¹³³. The precise mechanism of action for histone modifications in epigenetic gene regulations remains in most cases controversial. Unlike some other epigenetic regulation, like DNA methylation with hemi-methylated DNA methyltransferases, a mechanism of heritability for histone modifications through the DNA synthesis and the cell cycles remains unclear, although evidence exists that retention of parental histones after replication can reestablish histone PTMs on newly synthesized DNA¹³⁴. The difficulty in translating histone PTMs into readable outcomes and that most histone PTMs are only associated with a transcriptional state have led to challenges to the causality of histone PTMs in regulating genes and to the notion that histone PTMs can be considered a “code” to the same degree to that of the genetic code of DNA sequence¹³⁵. Independent of the precise nature of histone PTMs, the

complexes associated with the various marks are essential regulators of chromatin and have critical roles during development.

The Polycomb group (PcG) proteins are a family of complexes that can remodel chromatin and can epigenetically silence genes. The PcGs were named for the homeotic transformations of posterior legs towards a more anterior leg identity that were observed in PcG mutants in *D. melanogaster*¹³⁶ and are comprised of two major complexes, Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2), that act as the major silencing complexes during development. The two complexes have significant overlap in their genomic targets and can work synergistically to effect gene silencing but do have distinct regulatory targets and catalyze the addition of different histone modifications.

PRC2 consists of three core subunits, SUZ12, EED, and EZH1/2, and in mammals can be subdivided into two main forms, PRC2.1 and PRC2.2. All the core components are essential in embryonic development as homozygous mutants in mice causes embryonic lethality while heterozygous mutations cause congenital defects. The SET domain of EZH1/2 catalyzes the addition of up to three methyl groups onto histone H3K27¹³⁷. The other core units enhance this methyltransferase activity of EZH1/2 and have roles in stabilizing the complex, mediating interactions with other factors, and targeting the complex to genomic sites¹³⁸. In the canonical model of PcG recruitment to target loci PRC2 was thought to be necessary for the subsequent binding of PRC1 and as a result generated much study on the mechanisms of PRC2 targeting. PRC2-target genes have strong overlap with CpG islands and that GC-rich genomic elements are sufficient to recruit PRC2¹³⁹, with

other proposed mechanisms of sequence-specific transcription factors or long noncoding RNAs¹³⁸.

The composition of PRC1 is more variable than PRC2 and can broadly be divided into canonical PRC1 (cPRC1) and variant PRC1 (vPRC1) and in vertebrates has an expanded number of orthologs for each component. cPRC1 is comprised of four core subunits each present in stoichiometric amounts, a RING E3 ligase (RING1 or RNF2), a Polycomb group Ring finger (PCGF), a chromobox factor (CBX), and a Polyhomeotic homolog (PHC). vPRC1 is defined by the lack of the CBX and PHC subunits and instead contain either a RYBP or YAF2 subunit which compete for the same interaction pocket as CBX with RING1/RNF2 and are thus mutually exclusive subunits¹⁴⁰. In contrast to PRC2 where the only expansion of core subunits between *Drosophila* and mammals is the duplication of E(z) to EZH1 and EZH2, PRC1 subunits have undergone substantial expansion with paralogs for every core subunit existing in mammals. The core of PRC1 consists of a heterodimer of an E3 ligase RNF2 (or its paralog RING1) and a PCGF paralog. In mammals there are six PCGF paralogs and proteomic profiling of PRC1 indicates that which PCGF is present determines the composition of the remainder of the complex¹⁴¹. Broadly, PCGF2/4 form cPRC1 complexes and the other PCGFs define at least three distinct forms of vPRC1. These vPRC1s have distinct interacting factors and have distinct but overlapping functions and genomic locations¹⁴¹. The role of each complex in repressing genes is complicated by redundancy and crosstalk between forms of PRC1 and between PRC2 function but it is likely that the vPRC1 complexes work synergistically to efficiently silence genes¹⁴². The precise role of the PRC1s in mammals remains a topic of contention with some groups even reporting a role for vPRC1 in transcriptional activation¹⁴³ while

other groups do not find evidence for PRC1 function as a transcriptional activator¹⁴². Invertebrates are believed to only have a single PCGF gene (*Psc* in *Drosophila*) that is 4-5 times larger than vertebrate PCGFs and mediates both cPRC1 and vPRC1 functions¹⁴⁴. The conclusion that invertebrates have only a single copy of each cPRC1 core component is based on sparse data from limited model organisms (mostly *Drosophila* as *C. elegans* lacks PRC1) and sequencing of genomic data from new model organisms indicates that some invertebrates contain multiple copies of the PCGF gene. These multiple copies are often assumed to indicate paralogous expansion within a lineage¹⁴⁵, but the precise evolutionary lineage of the PCGF genes remains unresolved as Ecdysozoa may instead be a derived lineage with ancestral PCGF gene loss.

To form the core structure of PRC1, a PCGF subunit forms a heterodimer with either RING1 or RNF2. RING1 and RNF2 (*Ring1A* and *Ring1B* in mouse) are RING E3 ubiquitin ligases and have E3 ligase activity that targets histone H2A for mono-ubiquitylation at lysine 119 (H2AK119ub1)¹⁴⁶. The *in vitro* ubiquityl ligase activity of RNF2 is stimulated by the addition of other PRC1 components, especially RING1 and PCGF4 (BMI1)¹²⁹. The role of ubiquitylation of histone H2A in PRC1 function and mediating gene repression is controversial and appears to be dependent on the context and genomic target. The ability of PRC1 to compact *in vitro* is not impacted by the removal of N-terminal histone tails from nucleosome templates¹⁴⁷. These *in vitro* experiments used trypsinization treatment to remove tails, as opposed to a genetic approach, so it is unclear if the lysine 119 residue targeted by PRC1 is affected by the enzymatic treatment, but the conditions used during the compaction assays likely did not support ubiquityl ligase activity. The chromatin compaction activity of PRC1 is thus separable from its E3 ligase activity, at least in an *in*

in vitro context, suggesting that the H2AK119ub1 is either dispensable to PRC1 repressive function or that it has roles in other aspects of chromatin biology such as recruiting other factors or the maintenance of repressed domains. Reinforcing these *in vitro* results, fruit flies with a point mutation in *Sce* (*Drosophila* RNF2 homolog) that ablates the E3 ubiquityl ligases activity of PRC1 do not show phenotypes characteristic of Polycomb group mutants and fully maintain repression of PRC1 target genes¹⁴⁸. The same point mutation when expressed in ESCs was found to be sufficient to compact chromatin and maintain target gene repression¹⁴⁹ as well as rescue early embryonic mouse development in a *Ring1b* knock out (KO)¹⁵⁰. Importantly, catalytically inactive mutant PRC1 flies and mice do exhibit developmental defects and lethality which suggests that PRC1 ligase activity is needed for long term maintenance of chromatin states or has specific functions in certain cell types. Support for lineage-specific requirements for PRC1 catalytic activity is demonstrated by experiments where the loss of this activity in epidermal progenitor cells leads to an expansion of Merkel cells¹⁵¹. The myriad compositions of PRC1 and the interaction with PRC2 activity further complicates analysis. Early experiments on *Ring1B* function were not necessarily performed in a *Ring1A* null background and had H2AK119ub1 present at low levels. Experiments on *Ring1B* performed in a *Ring1A*^{-/-} background indicate that H2A ubiquitylation is not necessary for chromatin compaction at *Hox* loci but is indispensable for continued repression of target genes and necessary to maintain ESC potency⁸¹. Conditional mutation systems to completely ablate PRC1 E3 ligase activity demonstrate a central role for H2AK119ub1 in maintaining PRC1-mediated gene repression, recruitment of PRC2 to target loci, and the formation of PcG chromatin domain formation^{152,153}.

The complex interplay between PRC1 and PRC2 in maintaining gene repression and promoting lineage commitment is only beginning to become appreciated. In the best understood relationship PRC2 is recruited to target loci where it catalyzes the di- and tri-methylation of H3K27. This histone modification is recognized and bound by the CBX subunit of cPRC1 which is necessary to compact chromatin and maintain efficient gene repression. In this hierarchical model the E3 ligase function of PRC1 was thought to be dispensable as in a *Ring1b*^{-/-} ESC line H3K27me3 levels at *Hox* gene loci is not decreased and the observed chromatin decompaction can be rescued by the introduction of a mutant form of *Ring1b* that lacks *in vitro* ligase activity^{129,149}. Conversely, *Eed*^{-/-} ESCs demonstrate de-repression of PcG target genes and a significant reduction of PRC1 components Ring1b and Cbx2 bound at target loci¹⁵⁴. These studies helped establish the canonical model for PcG activity where PRC2 is first recruited to a target locus and through its catalytic action establishes domains marked with H3K27me3 for repression. The H3K27me3 mark is bound by the Cbx subunit of cPRC1 which, together with PRC2, forms higher order chromatin structures. The sterile alpha motif (SAM) domain of the Polyhomeotic (Ph in *Drosophila*, PHC in humans) subunit of cPRC1 facilitates the polymerization of PRC1 and is required for the subnuclear organization of PRC1 to mediate gene repression^{155,156}.

This canonical model of PcG recruitment is challenged by vPRC1s that lack a Cbx subunit and cannot bind H3K27me3. These vPRC1s are recruited to chromatin and establish domains of H2AK119ub1 independently of PRC2¹⁴¹. The ability of vPRC1 to be recruited to chromatin separately from PRC2 explains the observation that the genome-wide distributions of PRC1 and PRC2 do not fully overlap. A new model of PcG recruitment was proposed when it was demonstrated that PRC2 preferentially associates with and can

be recruited by nucleosomes that contained H2AK119ub1 and that the ubiquityl modification promoted the catalytic activity of PRC2^{157,158}. This established a feed forward mechanism where PRC2 or vPRC1 is recruited to a target locus and through their catalytic activity provide a recognition motif that allows the recruitment of the other complex. The presence of both epigenetic marks on a nucleosome provides a stable mechanism to recruit and maintain PcG localization. The precise interplay between the various forms of the two complexes and the mechanisms of recruitment to specific targets in a highly context specific manner^{159,160}.

While the function of the PcGs was discovered and first appreciated in an embryological context, additional critical roles for the complexes have emerged in stem cell regulation, regeneration, and cancer biology. As expected, given their role in embryogenesis the PcGs are necessary in both regulating stem cell maintenance and differentiation. PRC2 subunits have been reported as being required for the self-renewal and pluripotency in mouse ESCs¹⁶¹, but this result appears to be the result of specific culture conditions as other studies reported PRC2 function as dispensable for the expression of pluripotency factors¹⁶². The finding that PRC2 activity is not required for ESC self-renewal remains unresolved but a limited role in self-renewal is consistent with PRC2 mutant mice dying during and after implantation rather than during early embryogenesis^{163,164}. The phenotypes of PRC1 single KO mutants (excepting *RNF2*, which has gastrulation arrest¹⁶⁵) likewise manifest relatively late in development and the role of PRC1 in promoting pluripotency is complicated by the numerous variant forms and redundancy between the elements. This redundancy is supported by experiments in mice where single *Phc* knockout mice have phenotypes of homeotic transformations and

perinatal lethality, that become a much stronger phenotype of embryonic lethality when two paralogs are knocked out simultaneously¹⁶⁶. Similar synergistic effects have been observed in mice that are doubly deficient for cPRC1 components *Mel18* and *Bmi1* (PCGF2 and PCGF4) that exhibit embryonic lethality earlier than knockouts for either gene alone¹⁶⁷. While initially unclear due to redundancy, experiments on *Ring1b*-KO and catalytically inactivated *Ring1b* that were performed in a *Ring1a* genetic null background demonstrated a critical role for the enzymatic function of PRC1 in repressing Polycomb targets and maintaining ESCs⁸¹.

The overlap and redundancies in PcG-mediated gene regulation make studying the individual roles of each member element challenging but various biochemical and genetic approaches are uncovering the role for PRC1 in specific conformations and developmental contexts. Changes in the subunits that comprise PRC1 has been shown to be an important factor in regulating pluripotency and differentiation toward specific lineages. *Cbx7* has been shown to be highly expressed in ESCs and to negatively regulate the expression of other *Cbx* factors, *Cbx2*, *Cbx4*, and *Cbx8*. MicroRNAs miR-125 and miR-181 regulate *Cbx7* and mediate the switch of Polycomb orthologs during differentiation¹⁶⁸. This switching of Polycomb ortholog expression allows for changes in PRC1 conformation to regulate different developmental pathways, including *cbx4* in thymic epithelial cells and *cbx2* in the negative regulation of neurite development^{169,170}. In a similar manner, incorporation of vPRC1 component PCGF1 was found to be necessary for the positive regulation of the expression of endoderm- and mesoderm-associated transcription factors¹⁷¹. PcG protein RNF2 is necessary to repress certain genes in order to direct the development of specific lineages, examples include the repression of *tbx* transcription factors in zebrafish cardiac

development¹⁷² and in mice the repression proneural gene *neurogenin1* in a temporal manner to restrict the neurogenic competence of neural progenitor cells to allow the development of astrocytes¹⁷³. *Ring1b* has further roles in the developing mouse neocortex through regulating the timing of *Fez2f* expression and the termination of the production of subcerebral projection neurons from neural precursor cells¹⁷⁴. The use of a synthetic RING1B inhibitor on human ESC lines caused an increase in the expression of neuroectodermal marker genes in differentiating embryoid bodies and a reduction in endodermal and mesodermal marker genes demonstrating a role for PRC1 in directing early fate decisions during embryonic development¹⁷⁵.

PcG proteins are important in lineage commitment in contexts outside of embryogenesis and *hox* gene regulation. PRC1 has been shown to function in oocyte development as *Ring1/Rnf2* doubly deficient mice have defects in oocyte maturation and their progeny fail to develop past the two-cell stage due to a failure to activate the zygotic genome¹⁷⁶. PcGs are among the most up-regulated genes during vertebrate diapause in African turquoise killifish and maintain epigenetic marks at key developmental genes including the repression of metabolism and muscle genes by CBX7¹⁷⁷. The long-term stability but also reversibility of epigenetic marks through PcG action allows for the switching or suspension of developmental trajectories in response to environmental conditions while minimizing any long-term trade-offs in adult growth, fertility, and life span. PRC1 functions throughout the lifespan of organisms to regulate the effective long-term repression of target genes to support the maintenance of ASCs and during regenerative events. In intestinal stem cells PRC1 preserves cellular identity by repressing non-lineage-specific transcription factors and sustaining Wnt/ β -Catenin transcriptional

activity¹⁷⁸. PRC1 is also involved in adult hematopoiesis as *phc1*-deficient (*rae28*^{-/-}) HSCs were able to support the survival of lethally irradiated mice but were unable to increase HSC numbers long-term and through transplantation experiments¹⁷⁹. A role for PRC1 in liver regeneration was suggested by the observation that mouse *Cbx2* (*M33*) translocates to the nucleus following partial hepatectomy¹⁸⁰. Further post-embryonic developmental roles for PcGs are demonstrated in insect studies where the tissue transdetermination of *Drosophila* imaginal discs regeneration in was found to depend on the suppression of PcGs through the JNK signaling pathway¹⁸¹ and in *Tribolium castaneum* (flour beetle) where inhibition of PRC1 gene *Polycomb* (*Cbx* homolog) or PRC2 gene *Enhancer of zeste* (*E(z)*) was found to cause homeotic transformations during metamorphosis and a failure of tissue re-differentiation during leg regeneration¹⁸². The PcGs have extensive roles outside of the traditional contexts of embryogenesis that are beginning to be appreciated and contribute to the specification and maintenance of cellular identity throughout the lifespan of an organism.

As critical regulators of cellular identity and in maintaining stem cell plasticity it is not surprising that the PcGs are associated with several cancers. The multiple roles and myriad conformations of the PcG complexes mean that the action of the PcGs in cancer progression can be context and tissue dependent, with even a single PcG having both oncogenic and tumor suppressor activities. A wide variety of cancer types have been shown to involve the action of the PcGs (see ¹⁸³ and ¹⁸⁴ and references therein for summary tables) and many PcGs were initially discovered outside of *Drosophila* development as oncogenes. The context specificity of PcGs in cancer is demonstrated by the action of PRC1 gene *CBX7* which was found to be highly expressed in prostate cancers and have oncogenic properties

by promoting the growth of prostate cells through the repression of the Ink4a/Arf locus¹⁸⁵ while *CBX7* was demonstrated to have tumor-suppressor activities in thyroid cancer where the loss of *CBX7* expression correlates with a highly malignant phenotype¹⁸⁶. The PCGF gene *bmi1* (*Pcgf4*) was first characterized as an oncogene that promoted lymphomagenesis by collaborating with *c-myc* to regulate the INK4a/ARF locus^{187,188}. The *mel18* paralog of *bmi1*, despite substituting for *bmi1* in cPRC1 in a redundant manner, is thought to mostly act in cancer cells as a tumor-suppressor¹⁸⁹. The expression patterns of *bmi1* and *mel18* are generally negatively correlated and may compete for integration into PRC¹⁸³. Interestingly, the opposing roles for *mel18* and *bmi1* in cancer are mirrored in ESCs where Mel-18 is down-regulated and Bmi1 upregulated during differentiation¹⁹⁰. The complicated nature of PcGs in cancer etiology underpins the necessity of studying the complexes in a variety of complexes, especially to understand the potential for off-target effects for any potential developed therapeutics.

The introduction, in part, is a reprint of the material as it appears in eLS 2016. Allen, JM. Ross, KG. Zayas, RM. The dissertation author was the primary author of this journal article.

References

- 1 Gurdon, J. B. The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles. *Development* **10**, 622-640 (1962).
- 2 Driesch, H. Entwicklungsmechanische Studien. I. Der Werth der beiden ersten Furchungszellen in der Echinodermenentwicklung. *Z. wiss. Zool.* **53**, 163-178 (1892).
- 3 Sander, K. Shaking a concept: Hans Driesch and the varied fates of sea urchin blastomeres. *Roux Arch Dev Biol* **201**, 265, doi:10.1007/BF00592106 (1992).
- 4 Levine, M. Transcriptional enhancers in animal development and evolution. *Curr Biol* **20**, R754-763, doi:10.1016/j.cub.2010.06.070 (2010).
- 5 Ho, M. C., Goetz, S. E., Schiller, B. J., Allen, J. M. & Drewell, R. A. Between transcription and translation: Re-defining RNA and regulation. *Fly (Austin)* **2**, 152-155, doi:10.4161/fly.6382 (2008).
- 6 Ye, J. & Blelloch, R. Regulation of pluripotency by RNA binding proteins. *Cell Stem Cell* **15**, 271-280, doi:10.1016/j.stem.2014.08.010 (2014).
- 7 Strand, N. S., Allen, J. M. & Zayas, R. M. Post-translational regulation of planarian regeneration. *Semin Cell Dev Biol* **87**, 58-68, doi:10.1016/j.semcd.2018.04.009 (2019).
- 8 Ciechanover, A., Heller, H., Elias, S., Haas, A. L. & Hershko, A. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proc Natl Acad Sci U S A* **77**, 1365-1368, doi:10.1073/pnas.77.3.1365 (1980).
- 9 Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67**, 425-479, doi:10.1146/annurev.biochem.67.1.425 (1998).
- 10 Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156, doi:10.1038/292154a0 (1981).
- 11 Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-1147, doi:10.1126/science.282.5391.1145 (1998).
- 12 Smith, Z. D., Sindhu, C. & Meissner, A. Molecular features of cellular reprogramming and development. *Nat Rev Mol Cell Biol* **17**, 139-154, doi:10.1038/nrm.2016.6 (2016).
- 13 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).
- 14 Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872, doi:10.1016/j.cell.2007.11.019 (2007).

- 15 Stadtfeld, M., Maherali, N., Breault, D. T. & Hochedlinger, K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* **2**, 230-240, doi:10.1016/j.stem.2008.02.001 (2008).
- 16 Brambrink, T. *et al.* Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2**, 151-159, doi:10.1016/j.stem.2008.01.004 (2008).
- 17 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-1920, doi:10.1126/science.1151526 (2007).
- 18 Shu, J. *et al.* Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* **153**, 963-975, doi:10.1016/j.cell.2013.05.001 (2013).
- 19 Montserrat, N. *et al.* Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell* **13**, 341-350, doi:10.1016/j.stem.2013.06.019 (2013).
- 20 Kastenberg, Z. J. & Odorico, J. S. Alternative sources of pluripotency: science, ethics, and stem cells. *Transplant Rev (Orlando)* **22**, 215-222, doi:10.1016/j.trre.2008.04.002 (2008).
- 21 Cherry, A. B. & Daley, G. Q. Reprogramming cellular identity for regenerative medicine. *Cell* **148**, 1110-1122, doi:10.1016/j.cell.2012.02.031 (2012).
- 22 Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151**, 1617-1632, doi:10.1016/j.cell.2012.11.039 (2012).
- 23 Hussein, S. M. *et al.* Genome-wide characterization of the routes to pluripotency. *Nature* **516**, 198-206, doi:10.1038/nature14046 (2014).
- 24 Tonge, P. D. *et al.* Divergent reprogramming routes lead to alternative stem-cell states. *Nature* **516**, 192-197, doi:10.1038/nature14047 (2014).
- 25 Sridharan, R. *et al.* Role of the murine reprogramming factors in the induction of pluripotency. *Cell* **136**, 364-377, doi:10.1016/j.cell.2009.01.001 (2009).
- 26 Koche, R. P. *et al.* Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* **8**, 96-105, doi:10.1016/j.stem.2010.12.001 (2011).
- 27 Doi, A. *et al.* Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet* **41**, 1350-1353, doi:10.1038/ng.471 (2009).
- 28 Nishino, K. *et al.* DNA methylation dynamics in human induced pluripotent stem cells over time. *PLoS Genet* **7**, e1002085, doi:10.1371/journal.pgen.1002085 (2011).
- 29 Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285-290, doi:10.1038/nature09342 (2010).

- 30 Kim, K. *et al.* Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat Biotechnol* **29**, 1117-1119, doi:10.1038/nbt.2052 (2011).
- 31 Chamberlain, S. J. *et al.* Induced pluripotent stem cell models of the genomic imprinting disorders Angelman and Prader-Willi syndromes. *Proc Natl Acad Sci U S A* **107**, 17668-17673, doi:10.1073/pnas.1004487107 (2010).
- 32 Yang, J. *et al.* Induced pluripotent stem cells can be used to model the genomic imprinting disorder Prader-Willi syndrome. *J Biol Chem* **285**, 40303-40311, doi:10.1074/jbc.M110.183392 (2010).
- 33 Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68-73, doi:10.1038/nature09798 (2011).
- 34 Clevers, H. & Watt, F. M. Defining Adult Stem Cells by Function, not by Phenotype. *Annu Rev Biochem* **87**, 1015-1027, doi:10.1146/annurev-biochem-062917-012341 (2018).
- 35 Till, J. E. & McCulloch, E. A. A Direct Measurement of the Radiation Sensitivity of Normal Mouse Bone Marrow Cells¹. *Radiation Research* **175**, 145-149, doi:10.1667/rrxx28.1 (2011).
- 36 Metcalf, D. Studies on colony formation in vitro by mouse bone marrow cells. II. Action of colony stimulating factor. *J Cell Physiol* **76**, 89-99, doi:10.1002/jcp.1040760113 (1970).
- 37 Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631-644, doi:10.1016/j.cell.2008.01.025 (2008).
- 38 Becker, A. J., Mc, C. E. & Till, J. E. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature* **197**, 452-454, doi:10.1038/197452a0 (1963).
- 39 Clevers, H. STEM CELLS. What is an adult stem cell? *Science* **350**, 1319-1320, doi:10.1126/science.aad7016 (2015).
- 40 Mauro, A. Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol* **9**, 493-495, doi:10.1083/jcb.9.2.493 (1961).
- 41 Reznik, M. Thymidine-3H uptake by satellite cells of regenerating skeletal muscle. *J Cell Biol* **40**, 568-571, doi:10.1083/jcb.40.2.568 (1969).
- 42 Moss, F. P. & Leblond, C. P. Satellite cells as the source of nuclei in muscles of growing rats. *Anat Rec* **170**, 421-435, doi:10.1002/ar.1091700405 (1971).
- 43 Schultz, E., Gibson, M. C. & Champion, T. Satellite cells are mitotically quiescent in mature mouse muscle: an EM and radioautographic study. *J Exp Zool* **206**, 451-456, doi:10.1002/jez.1402060314 (1978).
- 44 Jaks, V. *et al.* Lgr5 marks cycling, yet long-lived, hair follicle stem cells. *Nat Genet* **40**, 1291-1299, doi:10.1038/ng.239 (2008).

- 45 Jensen, K. B. *et al.* Lrig1 expression defines a distinct multipotent stem cell population in mammalian epidermis. *Cell Stem Cell* **4**, 427-439, doi:10.1016/j.stem.2009.04.014 (2009).
- 46 Cotsarelis, G., Sun, T.-T. & Lavker, R. M. Label-retaining cells reside in the bulge area of pilosebaceous unit: Implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* **61**, 1329-1337, doi:10.1016/0092-8674(90)90696-c (1990).
- 47 Clayton, E. *et al.* A single type of progenitor cell maintains normal epidermis. *Nature* **446**, 185-189, doi:10.1038/nature05574 (2007).
- 48 Mascre, G. *et al.* Distinct contribution of stem and progenitor cells to epidermal maintenance. *Nature* **489**, 257-262, doi:10.1038/nature11393 (2012).
- 49 Page, M. E., Lombard, P., Ng, F., Gottgens, B. & Jensen, K. B. The epidermis comprises autonomous compartments maintained by distinct stem cell populations. *Cell Stem Cell* **13**, 471-482, doi:10.1016/j.stem.2013.07.010 (2013).
- 50 Schepeler, T., Page, M. E. & Jensen, K. B. Heterogeneity and plasticity of epidermal stem cells. *Development* **141**, 2559-2567, doi:10.1242/dev.104588 (2014).
- 51 Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189-193, doi:10.1038/nature10573 (2011).
- 52 Rios, A. C., Fu, N. Y., Lindeman, G. J. & Visvader, J. E. In situ identification of bipotent stem cells in the mammary gland. *Nature* **506**, 322-327, doi:10.1038/nature12948 (2014).
- 53 Ousset, M. *et al.* Multipotent and unipotent progenitors contribute to prostate postnatal development. *Nat Cell Biol* **14**, 1131-1138, doi:10.1038/ncb2600 (2012).
- 54 Li, W., Li, L. & Hui, L. Cell Plasticity in Liver Regeneration. *Trends Cell Biol* **30**, 329-338, doi:10.1016/j.tcb.2020.01.007 (2020).
- 55 Chen, F. *et al.* Broad Distribution of Hepatocyte Proliferation in Liver Homeostasis and Regeneration. *Cell Stem Cell* **26**, 27-33 e24, doi:10.1016/j.stem.2019.11.001 (2020).
- 56 Matsumoto, T., Wakefield, L., Tarlow, B. D. & Grompe, M. In Vivo Lineage Tracing of Polyploid Hepatocytes Reveals Extensive Proliferation during Liver Regeneration. *Cell Stem Cell* **26**, 34-47 e33, doi:10.1016/j.stem.2019.11.014 (2020).
- 57 Goss, R. J. *Principles of regeneration*. (Academic Press, 1969).
- 58 Dinsmore, C. E. Regeneration: Principles. *eLS*, doi:10.1038/npg.els.0001112 (2001).
- 59 Bely, A. E., Zattara, E. E. & Sikes, J. M. Regeneration in spiralian: evolutionary patterns and developmental processes. *Int J Dev Biol* **58**, 623-634, doi:10.1387/ijdb.140142ab (2014).
- 60 Baddour, J. A., Sousounis, K. & Tsonis, P. A. Organ repair and regeneration: an overview. *Birth Defects Res C Embryo Today* **96**, 1-29, doi:10.1002/bdrc.21006 (2012).

- 61 Bely, A. E. & Nyberg, K. G. Evolution of animal regeneration: re-emergence of a field. *Trends Ecol Evol* **25**, 161-170, doi:10.1016/j.tree.2009.08.005 (2010).
- 62 Aristotle & Peck, A. L. *Historia animalium*. (Heinemann; Harvard University Press, 1965).
- 63 Elliott, S. A. & Sanchez Alvarado, A. The history and enduring contributions of planarians to the study of animal regeneration. *Wiley Interdiscip Rev Dev Biol* **2**, 301-326, doi:10.1002/wdev.82 (2013).
- 64 Trembley, A. *Mémoires pour servir à l'histoire d'un genre de ployes d'eau douce, à bras en forme de cornes*. (Chez Durand, 1744).
- 65 Birnbaum, K. D. & Sanchez Alvarado, A. Slicing across kingdoms: regeneration in plants and animals. *Cell* **132**, 697-710, doi:10.1016/j.cell.2008.01.040 (2008).
- 66 Bonnet, C. in *Nineteenth Century Collections Online: Science, Technology, and Medicine: 1780-1925, Part II* 1 online resource (2 volumes) (Chez Durand, libraire ..., A Paris, 1745).
- 67 Spallanzani, L., Maty, M. & Pre-1801 Imprint Collection (Library of Congress). *An essay on animal reproductions*. (T. Becket and P.A. de Hondt, 1769).
- 68 Elliott, S. A. *Studies of conserved cell-cell signaling pathways in the planarian, Schmidtea mediterranea*, Department of Neurobiology and Anatomy, University of Utah, (2016).
- 69 Johnson, S. L. & Weston, J. A. Temperature-sensitive mutations that cause stage-specific defects in Zebrafish fin regeneration. *Genetics* **141**, 1583-1595 (1995).
- 70 Knopf, F. *et al.* Bone regenerates via dedifferentiation of osteoblasts in the zebrafish fin. *Dev Cell* **20**, 713-724, doi:10.1016/j.devcel.2011.04.014 (2011).
- 71 Tu, S. & Johnson, S. L. Fate restriction in the growing and regenerating zebrafish fin. *Dev Cell* **20**, 725-732, doi:10.1016/j.devcel.2011.04.013 (2011).
- 72 Poss, K. D., Wilson, L. G. & Keating, M. T. Heart regeneration in zebrafish. *Science* **298**, 2188-2190, doi:10.1126/science.1077857 (2002).
- 73 Becker, T., Wullimann, M. F., Becker, C. G., Bernhardt, R. R. & Schachner, M. Axonal regrowth after spinal cord transection in adult zebrafish. *The Journal of Comparative Neurology* **377**, 577-595, doi:10.1002/(sici)1096-9861(19970127)377:4<577::Aid-cne8>3.0.Co;2-# (1997).
- 74 Reimer, M. M. *et al.* Motor neuron regeneration in adult zebrafish. *J Neurosci* **28**, 8510-8516, doi:10.1523/JNEUROSCI.1189-08.2008 (2008).
- 75 Kragl, M. *et al.* Cells keep a memory of their tissue origin during axolotl limb regeneration. *Nature* **460**, 60-65, doi:10.1038/nature08152 (2009).
- 76 Goss, R. J. Future directions in antler research. *Anat Rec* **241**, 291-302, doi:10.1002/ar.1092410302 (1995).
- 77 Illingworth, C. M. Trapped fingers and amputated finger tips in children. *Journal of Pediatric Surgery* **9**, 853-858, doi:10.1016/s0022-3468(74)80220-4 (1974).

- 78 Seifert, A. W. & Muneoka, K. The blastema and epimorphic regeneration in mammals. *Dev Biol* **433**, 190-199, doi:10.1016/j.ydbio.2017.08.007 (2018).
- 79 Seifert, A. W. *et al.* Skin shedding and tissue regeneration in African spiny mice (*Acomys*). *Nature* **489**, 561-565, doi:10.1038/nature11499 (2012).
- 80 Goldstein, G. *et al.* Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc Natl Acad Sci U S A* **72**, 11-15, doi:10.1073/pnas.72.1.11 (1975).
- 81 Endoh, M. *et al.* Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. *PLoS Genet* **8**, e1002774, doi:10.1371/journal.pgen.1002774 (2012).
- 82 Nakayama, K. I. & Nakayama, K. Regulation of the cell cycle by SCF-type ubiquitin ligases. *Semin Cell Dev Biol* **16**, 323-333, doi:10.1016/j.semcdb.2005.02.010 (2005).
- 83 Higgins, R. *et al.* The Unfolded Protein Response Triggers Site-Specific Regulatory Ubiquitylation of 40S Ribosomal Proteins. *Mol Cell* **59**, 35-49, doi:10.1016/j.molcel.2015.04.026 (2015).
- 84 Glickman, M. H. & Ciechanover, A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev* **82**, 373-428, doi:10.1152/physrev.00027.2001 (2002).
- 85 Carlucci, A. & D'Angiolella, V. It is not all about BRCA: Cullin-Ring ubiquitin Ligases in ovarian cancer. *Br J Cancer* **112**, 9-13, doi:10.1038/bjc.2014.594 (2015).
- 86 Zheng, Q. *et al.* Dysregulation of Ubiquitin-Proteasome System in Neurodegenerative Diseases. *Front Aging Neurosci* **8**, 303, doi:10.3389/fnagi.2016.00303 (2016).
- 87 Yang, W. *et al.* The histone H2A deubiquitinase Usp16 regulates embryonic stem cell gene expression and lineage commitment. *Nat Commun* **5**, 3818, doi:10.1038/ncomms4818 (2014).
- 88 Xu, H. *et al.* WWP2 promotes degradation of transcription factor OCT4 in human embryonic stem cells. *Cell Res* **19**, 561-573, doi:10.1038/cr.2009.31 (2009).
- 89 Luo, X. *et al.* Wwp2 targets SRG3, a scaffold protein of the SWI/SNF-like BAF complex, for ubiquitination and degradation. *Biochem Biophys Res Commun* **443**, 1048-1053, doi:10.1016/j.bbrc.2013.12.089 (2014).
- 90 Pasten, C., Ortiz-Pineda, P. A. & Garcia-Arraras, J. E. Ubiquitin-proteasome system components are upregulated during intestinal regeneration. *Genesis* **50**, 350-365, doi:10.1002/dvg.20803 (2012).
- 91 Pasten, C., Rosa, R., Ortiz, S., Gonzalez, S. & Garcia-Arraras, J. E. Characterization of proteolytic activities during intestinal regeneration of the sea cucumber, *Holothuria glaberrima*. *Int J Dev Biol* **56**, 681-691, doi:10.1387/ijdb.113473cp (2012).

- 92 Henderson, J. M. *et al.* Identification of HECT E3 ubiquitin ligase family genes involved in stem cell regulation and regeneration in planarians. *Dev Biol* **404**, 21-34, doi:10.1016/j.ydbio.2015.04.021 (2015).
- 93 Schulman, B. A. & Harper, J. W. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* **10**, 319-331, doi:10.1038/nrm2673 (2009).
- 94 Metzger, M. B., Hristova, V. A. & Weissman, A. M. HECT and RING finger families of E3 ubiquitin ligases at a glance. *J Cell Sci* **125**, 531-537, doi:10.1242/jcs.091777 (2012).
- 95 Komander, D. & Rape, M. The ubiquitin code. *Annu Rev Biochem* **81**, 203-229, doi:10.1146/annurev-biochem-060310-170328 (2012).
- 96 Williamson, A., Werner, A. & Rape, M. The Colossus of ubiquitylation: decrypting a cellular code. *Mol Cell* **49**, 591-600, doi:10.1016/j.molcel.2013.01.028 (2013).
- 97 David, Y. *et al.* E3 ligases determine ubiquitination site and conjugate type by enforcing specificity on E2 enzymes. *J Biol Chem* **286**, 44104-44115, doi:10.1074/jbc.M111.234559 (2011).
- 98 Li, W. *et al.* Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS One* **3**, e1487, doi:10.1371/journal.pone.0001487 (2008).
- 99 Iconomou, M. & Saunders, D. N. Systematic approaches to identify E3 ligase substrates. *Biochem J* **473**, 4083-4101, doi:10.1042/BCJ20160719 (2016).
- 100 Lovering, R. *et al.* Identification and preliminary characterization of a protein motif related to the zinc finger. *Proc Natl Acad Sci U S A* **90**, 2112-2116, doi:10.1073/pnas.90.6.2112 (1993).
- 101 Freemont, P. S. The RING finger. A novel protein sequence motif related to the zinc finger. *Ann N Y Acad Sci* **684**, 174-192, doi:10.1111/j.1749-6632.1993.tb32280.x (1993).
- 102 Borden, K. L. B. & Freemont, P. S. The RING finger domain: a recent example of a sequence—structure family. *Current Opinion in Structural Biology* **6**, 395-401, doi:10.1016/s0959-440x(96)80060-1 (1996).
- 103 Metzger, M. B., Pruneda, J. N., Klevit, R. E. & Weissman, A. M. RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim Biophys Acta* **1843**, 47-60, doi:10.1016/j.bbamcr.2013.05.026 (2014).
- 104 Yin, Q. *et al.* E2 interaction and dimerization in the crystal structure of TRAF6. *Nat Struct Mol Biol* **16**, 658-666, doi:10.1038/nsmb.1605 (2009).
- 105 Pintard, L., Willems, A. & Peter, M. Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *EMBO J* **23**, 1681-1687, doi:10.1038/sj.emboj.7600186 (2004).

- 106 Scadden, D. T. The stem-cell niche as an entity of action. *Nature* **441**, 1075-1079, doi:10.1038/nature04957 (2006).
- 107 Chanarat, S. & Strasser, K. Splicing and beyond: the many faces of the Prp19 complex. *Biochim Biophys Acta* **1833**, 2126-2134, doi:10.1016/j.bbamcr.2013.05.023 (2013).
- 108 Grote, M. *et al.* Molecular architecture of the human Prp19/CDC5L complex. *Mol Cell Biol* **30**, 2105-2119, doi:10.1128/MCB.01505-09 (2010).
- 109 Minocha, R. *et al.* Mud2 functions in transcription by recruiting the Prp19 and TREX complexes to transcribed genes. *Nucleic Acids Res* **46**, 9749-9763, doi:10.1093/nar/gky640 (2018).
- 110 Cho, S. Y. *et al.* Identification of mouse Prp19p as a lipid droplet-associated protein and its possible involvement in the biogenesis of lipid droplets. *J Biol Chem* **282**, 2456-2465, doi:10.1074/jbc.M608042200 (2007).
- 111 Henriques, J. A. P. & Moustacchi, E. Isolation and Characterization of *pso* Mutants Sensitive to Photo-Addition of Psoralen Derivatives in *Saccharomyces cerevisiae* *Genetics* **95**, 273-288 (1980).
- 112 Ohi, M. D. *et al.* Structural and functional analysis of essential pre-mRNA splicing factor Prp19p. *Mol Cell Biol* **25**, 451-460, doi:10.1128/MCB.25.1.451-460.2005 (2005).
- 113 Aravind, L. & Koonin, E. V. The U box is a modified RING finger - a common domain in ubiquitination. *Curr Biol* **10**, 132-134, doi:10.1016/S0960-9822(00)00398-5 (2000).
- 114 Sihm, C. R., Cho, S. Y., Lee, J. H., Lee, T. R. & Kim, S. H. Mouse homologue of yeast Prp19 interacts with mouse SUG1, the regulatory subunit of 26S proteasome. *Biochem Biophys Res Commun* **356**, 175-180, doi:10.1016/j.bbrc.2007.02.134 (2007).
- 115 Huang, R., Xue, R., Qu, D., Yin, J. & Shen, X.-Z. Prp19 Arrests Cell Cycle via Cdc5L in Hepatocellular Carcinoma Cells. *International Journal of Molecular Sciences* **18**, 778, doi:10.3390/ijms18040778 (2017).
- 116 Gutnik, S. *et al.* PRP-19, a conserved pre-mRNA processing factor and E3 ubiquitin ligase, inhibits the nuclear accumulation of GLP-1/Notch intracellular domain. *Biol Open* **7**, doi:10.1242/bio.034066 (2018).
- 117 Mahajan, K. N. & Mitchell, B. S. Role of human Pso4 in mammalian DNA repair and association with terminal deoxynucleotidyl transferase. *Proc Natl Acad Sci U S A* **100**, 10746-10751, doi:10.1073/pnas.1631060100 (2003).
- 118 Abbas, M., Shanmugam, I., Bsaili, M., Hromas, R. & Shaheen, M. The role of the human psoralen 4 (hPso4) protein complex in replication stress and homologous recombination. *J Biol Chem* **289**, 14009-14019, doi:10.1074/jbc.M113.520056 (2014).
- 119 Marechal, A. *et al.* PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR activation via a ubiquitin-mediated circuitry. *Mol Cell* **53**, 235-246, doi:10.1016/j.molcel.2013.11.002 (2014).

- 120 Will, C. L. & Luhrmann, R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol* **3**, doi:10.1101/cshperspect.a003707 (2011).
- 121 Song, E. J. *et al.* The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev* **24**, 1434-1447, doi:10.1101/gad.1925010 (2010).
- 122 Chen, X. *et al.* Precursor RNA processing 3 is required for male fertility, and germline stem cell self-renewal and differentiation via regulating spliceosome function in *Drosophila* testes. *Sci Rep* **9**, 9988, doi:10.1038/s41598-019-46419-x (2019).
- 123 Roberts-Galbraith, R. H., Brubacher, J. L. & Newmark, P. A. A functional genomics screen in planarians reveals regulators of whole-brain regeneration. *Elife* **5**, doi:10.7554/eLife.17002 (2016).
- 124 Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251-260, doi:10.1038/38444 (1997).
- 125 Lehninger, A. L., Nelson, D. L. & Cox, M. M. *Lehninger principles of biochemistry*. 4th edn, (W.H. Freeman, 2005).
- 126 Allis, C. D., Jenuwein, T. & Reinberg, D. *Epigenetics*. (Cold Spring Harbor Laboratory Press, 2007).
- 127 Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41-45, doi:10.1038/47412 (2000).
- 128 Zhu, B. *et al.* Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol Cell* **20**, 601-611, doi:10.1016/j.molcel.2005.09.025 (2005).
- 129 Cao, R., Tsukada, Y. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol Cell* **20**, 845-854, doi:10.1016/j.molcel.2005.12.002 (2005).
- 130 Strahl, B. D., Ohba, R., Cook, R. G. & Allis, C. D. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. *Proc Natl Acad Sci U S A* **96**, 14967-14972, doi:10.1073/pnas.96.26.14967 (1999).
- 131 Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* **16**, 2893-2905, doi:10.1101/gad.1035902 (2002).
- 132 Gardner, K. E., Allis, C. D. & Strahl, B. D. Operating on chromatin, a colorful language where context matters. *J Mol Biol* **409**, 36-46, doi:10.1016/j.jmb.2011.01.040 (2011).
- 133 Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705, doi:10.1016/j.cell.2007.02.005 (2007).

- 134 Reveron-Gomez, N. *et al.* Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. *Mol Cell*, doi:10.1016/j.molcel.2018.08.010 (2018).
- 135 Henikoff, S. & Shilatifard, A. Histone modification: cause or cog? *Trends Genet* **27**, 389-396, doi:10.1016/j.tig.2011.06.006 (2011).
- 136 Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565-570, doi:10.1038/276565a0 (1978).
- 137 Shen, X. *et al.* EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol Cell* **32**, 491-502, doi:10.1016/j.molcel.2008.10.016 (2008).
- 138 Di Croce, L. & Helin, K. Transcriptional regulation by Polycomb group proteins. *Nat Struct Mol Biol* **20**, 1147-1155, doi:10.1038/nsmb.2669 (2013).
- 139 Mendenhall, E. M. *et al.* GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet* **6**, e1001244, doi:10.1371/journal.pgen.1001244 (2010).
- 140 Wang, R. *et al.* Polycomb group targeting through different binding partners of RING1B C-terminal domain. *Structure* **18**, 966-975, doi:10.1016/j.str.2010.04.013 (2010).
- 141 Gao, Z. *et al.* PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol Cell* **45**, 344-356, doi:10.1016/j.molcel.2012.01.002 (2012).
- 142 Fursova, N. A. *et al.* Synergy between Variant PRC1 Complexes Defines Polycomb-Mediated Gene Repression. *Mol Cell* **74**, 1020-1036 e1028, doi:10.1016/j.molcel.2019.03.024 (2019).
- 143 Gao, Z. *et al.* An AUTS2-Polycomb complex activates gene expression in the CNS. *Nature* **516**, 349-354, doi:10.1038/nature13921 (2014).
- 144 Whitcomb, S. J., Basu, A., Allis, C. D. & Bernstein, E. Polycomb Group proteins: an evolutionary perspective. *Trends Genet* **23**, 494-502, doi:10.1016/j.tig.2007.08.006 (2007).
- 145 Sowpati, D. T., Ramamoorthy, S. & Mishra, R. K. Expansion of the polycomb system and evolution of complexity. *Mech Dev* **138 Pt 2**, 97-112, doi:10.1016/j.mod.2015.07.013 (2015).
- 146 Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873-878, doi:10.1038/nature02985 (2004).
- 147 Francis, N. J., Kingston, R. E. & Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574-1577, doi:10.1126/science.1100576 (2004).
- 148 Pengelly, A. R., Kalb, R., Finkl, K. & Muller, J. Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation. *Genes Dev* **29**, 1487-1492, doi:10.1101/gad.265439.115 (2015).

- 149 Eskeland, R. *et al.* Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol Cell* **38**, 452-464, doi:10.1016/j.molcel.2010.02.032 (2010).
- 150 Illingworth, R. S. *et al.* The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development. *Genes Dev* **29**, 1897-1902, doi:10.1101/gad.268151.115 (2015).
- 151 Cohen, I. *et al.* PRC1 Fine-tunes Gene Repression and Activation to Safeguard Skin Development and Stem Cell Specification. *Cell Stem Cell* **22**, 726-739 e727, doi:10.1016/j.stem.2018.04.005 (2018).
- 152 Tamburri, S. *et al.* Histone H2AK119 Mono-Ubiquitination Is Essential for Polycomb-Mediated Transcriptional Repression. *Mol Cell*, doi:10.1016/j.molcel.2019.11.021 (2019).
- 153 Blackledge, N. P. *et al.* PRC1 Catalytic Activity Is Central to Polycomb System Function. *Mol Cell* **77**, 857-874 e859, doi:10.1016/j.molcel.2019.12.001 (2020).
- 154 Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-353, doi:10.1038/nature04733 (2006).
- 155 Robinson, A. K. *et al.* The growth-suppressive function of the polycomb group protein polyhomeotic is mediated by polymerization of its sterile alpha motif (SAM) domain. *J Biol Chem* **287**, 8702-8713, doi:10.1074/jbc.M111.336115 (2012).
- 156 Kim, C. A. in *Polycomb Group Proteins* 5-31 (2017).
- 157 Kalb, R. *et al.* Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat Struct Mol Biol* **21**, 569-571, doi:10.1038/nsmb.2833 (2014).
- 158 Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445-1459, doi:10.1016/j.cell.2014.05.004 (2014).
- 159 Schwartz, Y. B. & Pirrotta, V. Ruled by ubiquitylation: a new order for polycomb recruitment. *Cell Rep* **8**, 321-325, doi:10.1016/j.celrep.2014.07.001 (2014).
- 160 Cohen, I., Bar, C. & Ezhkova, E. Activity of PRC1 and Histone H2AK119 Monoubiquitination: Revising Popular Misconceptions. *Bioessays*, e1900192, doi:10.1002/bies.201900192 (2020).
- 161 O'Carroll, D. *et al.* The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol* **21**, 4330-4336, doi:10.1128/MCB.21.13.4330-4336.2001 (2001).
- 162 Chamberlain, S. J., Yee, D. & Magnuson, T. Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells* **26**, 1496-1505, doi:10.1634/stemcells.2008-0102 (2008).
- 163 Faust, C., Lawson, K. A., Schork, N. J., Thiel, B. & Magnuson, T. The Polycomb-group gene *eed* is required for normal morphogenetic movements during gastrulation in the mouse embryo. *Development* **125**, 4495-4506 (1998).

- 164 Pasini, D., Bracken, A. P., Jensen, M. R., Lazzerini Denchi, E. & Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J* **23**, 4061-4071, doi:10.1038/sj.emboj.7600402 (2004).
- 165 Voncken, J. W. *et al.* Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc Natl Acad Sci U S A* **100**, 2468-2473, doi:10.1073/pnas.0434312100 (2003).
- 166 Isono, K. *et al.* Mammalian polyhomeotic homologues Phc2 and Phc1 act in synergy to mediate polycomb repression of Hox genes. *Mol Cell Biol* **25**, 6694-6706, doi:10.1128/MCB.25.15.6694-6706.2005 (2005).
- 167 Akasaka, T. *et al.* Mice doubly deficient for the Polycomb Group genes *Mel18* and *Bmi1* reveal synergy and requirement for maintenance but not initiation of Hox gene expression. *Development* **128**, 1587-1597 (2001).
- 168 O'Loughlen, A. *et al.* MicroRNA regulation of *Cbx7* mediates a switch of Polycomb orthologs during ESC differentiation. *Cell Stem Cell* **10**, 33-46, doi:10.1016/j.stem.2011.12.004 (2012).
- 169 Liu, B. *et al.* *Cbx4* regulates the proliferation of thymic epithelial cells and thymus function. *Development* **140**, 780-788, doi:10.1242/dev.085035 (2013).
- 170 Gu, X. *et al.* CBX2 Inhibits Neurite Development by Regulating Neuron-Specific Genes Expression. *Front Mol Neurosci* **11**, 46, doi:10.3389/fnmol.2018.00046 (2018).
- 171 Yan, Y. *et al.* Loss of Polycomb Group Protein *Pcgf1* Severely Compromises Proper Differentiation of Embryonic Stem Cells. *Sci Rep* **7**, 46276, doi:10.1038/srep46276 (2017).
- 172 Chrispijn, N. D. *et al.* Loss of the Polycomb group protein *Rnf2* results in derepression of *tbx*-transcription factors and defects in embryonic and cardiac development. *Sci Rep* **9**, 4327, doi:10.1038/s41598-019-40867-1 (2019).
- 173 Hirabayashi, Y. *et al.* Polycomb limits the neurogenic competence of neural precursor cells to promote astrogenic fate transition. *Neuron* **63**, 600-613, doi:10.1016/j.neuron.2009.08.021 (2009).
- 174 Morimoto-Suzuki, N. *et al.* The polycomb component *Ring1B* regulates the timed termination of subcerebral projection neuron production during mouse neocortical development. *Development* **141**, 4343-4353, doi:10.1242/dev.112276 (2014).
- 175 Desai, D., Khanna, A. & Pethe, P. Inhibition of *RING1B* alters lineage specificity in human embryonic stem cells. *Cell Biol Int*, doi:10.1002/cbin.11325 (2020).
- 176 Posfai, E. *et al.* Polycomb function during oogenesis is required for mouse embryonic development. *Genes Dev* **26**, 920-932, doi:10.1101/gad.188094.112 (2012).
- 177 Hu, C. K. *et al.* Vertebrate diapause preserves organisms long term through Polycomb complex members. *Science* **367**, 870-874, doi:10.1126/science.aaw2601 (2020).

- 178 Chiacchiera, F. *et al.* Polycomb Complex PRC1 Preserves Intestinal Stem Cell Identity by Sustaining Wnt/beta-Catenin Transcriptional Activity. *Cell Stem Cell* **18**, 91-103, doi:10.1016/j.stem.2015.09.019 (2016).
- 179 Kim, J. Y. *et al.* Defective long-term repopulating ability in hematopoietic stem cells lacking the Polycomb-group gene *rae28*. *Eur J Haematol* **73**, 75-84, doi:10.1111/j.1600-0609.2004.00268.x (2004).
- 180 Noguchi, K., Shiurba, R. & Higashinakagawa, T. Nuclear translocation of mouse polycomb m33 protein in regenerating liver. *Biochem Biophys Res Commun* **291**, 508-515, doi:10.1006/bbrc.2002.6480 (2002).
- 181 Lee, N., Maurange, C., Ringrose, L. & Paro, R. Suppression of Polycomb group proteins by JNK signalling induces transdetermination in Drosophila imaginal discs. *Nature* **438**, 234-237, doi:10.1038/nature04120 (2005).
- 182 Chou, J. *et al.* Roles of Polycomb group proteins Enhancer of zeste (E(z)) and Polycomb (Pc) during metamorphosis and larval leg regeneration in the flour beetle *Tribolium castaneum*. *Developmental Biology* **450**, 34-46, doi:https://doi.org/10.1016/j.ydbio.2019.03.002 (2019).
- 183 Koppens, M. & van Lohuizen, M. Context-dependent actions of Polycomb repressors in cancer. *Oncogene* **35**, 1341-1352, doi:10.1038/onc.2015.195 (2016).
- 184 Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780, doi:10.1126/science.aad9780 (2016).
- 185 Bernard, D. *et al.* CBX7 controls the growth of normal and tumor-derived prostate cells by repressing the *Ink4a/Arf* locus. *Oncogene* **24**, 5543-5551, doi:10.1038/sj.onc.1208735 (2005).
- 186 Pallante, P. *et al.* Loss of the CBX7 gene expression correlates with a highly malignant phenotype in thyroid cancer. *Cancer Res* **68**, 6770-6778, doi:10.1158/0008-5472.CAN-08-0695 (2008).
- 187 Haupt, Y., Alexander, W. S., Barri, G., Peter Klinken, S. & Adams, J. M. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in Eμ-myc transgenic mice. *Cell* **65**, 753-763, doi:10.1016/0092-8674(91)90383-a (1991).
- 188 Jacobs, J. J. *et al.* Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. *Genes Dev* **13**, 2678-2690, doi:10.1101/gad.13.20.2678 (1999).
- 189 Guo, W. J. *et al.* Mel-18 acts as a tumor suppressor by repressing Bmi-1 expression and down-regulating Akt activity in breast cancer cells. *Cancer Res* **67**, 5083-5089, doi:10.1158/0008-5472.CAN-06-4368 (2007).
- 190 Morey, L. *et al.* Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. *Cell Stem Cell* **10**, 47-62, doi:10.1016/j.stem.2011.12.006 (2012).

Chapter 1:

Dissecting the function of cullin-RING ubiquitin ligase complex genes in planarian regeneration



Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

Short communication

Dissecting the function of Cullin-RING ubiquitin ligase complex genes in planarian regeneration



Nicholas S. Strand¹, John M. Allen¹, Mahjoobah Ghulam, Matthew R. Taylor, Roma K. Munday, Melissa Carrillo, Artem Movsesyan, Ricardo M. Zayas*

Department of Biology, San Diego State University, San Diego, CA 92182, USA

ARTICLE INFO

Keywords:

Planarian
Cullin
F-box
Regeneration
Stem cells
E3 ubiquitin ligase

ABSTRACT

The ubiquitin system plays a role in nearly every aspect of eukaryotic cell biology. The enzymes responsible for transferring ubiquitin onto specific substrates are the E3 ubiquitin ligases, a large and diverse family of proteins, for which biological roles and target substrates remain largely undefined. Studies using model organisms indicate that ubiquitin signaling mediates key steps in developmental processes and tissue regeneration. Here, we used the freshwater planarian, *Schmidtea mediterranea*, to investigate the role of Cullin-RING ubiquitin ligase (CRL) complexes in stem cell regulation during regeneration. We identified six *S. mediterranea* cullin genes, and used RNAi to uncover roles for homologs of Cullin-1, -3 and -4 in planarian regeneration. The *cullin-1* RNAi phenotype included defects in blastema formation, organ regeneration, lesions, and lysis. To further investigate the function of *cullin-1*-mediated cellular processes in planarians, we examined genes encoding the adaptor protein Skp1 and F-box substrate-recognition proteins that are predicted to partner with Cullin-1. RNAi against *skp1* resulted in phenotypes similar to *cullin-1* RNAi, and an RNAi screen of the F-box genes identified 19 genes that recapitulated aspects of *cullin-1* RNAi, including ones that in mammals are involved in stem cell regulation and cancer biology. Our data provides evidence that CRLs play discrete roles in regenerative processes and provide a platform to investigate how CRLs regulate stem cells *in vivo*.

1. Introduction

Planarians have emerged as an important model organism to examine gene function in stem cell-based tissue regeneration (Elliott and Sánchez Alvarado, 2012; Roberts-Galbraith and Newmark, 2015; Ross et al., 2017). These animals can restore lost or damaged tissues from a population of adult pluripotent stem cells, termed neoblasts (Baguña, 2012; Rink, 2013; Wagner et al., 2011; Zhu and Pearson, 2016). Recent studies have provided insights into the molecular mechanisms that regulate regeneration in planarians at the genetic level (Elliott and Sánchez Alvarado, 2012; Roberts-Galbraith and Newmark, 2015; Wurtzel et al., 2015). However, the dynamic regulation of proteins during regeneration remains an open area of investigation. An essential cellular pathway in protein regulation is the ubiquitin system, in which cells utilize the highly conserved small ubiquitin polypeptide as a post-translational modification of other proteins, which can lead to degradation of target proteins (Ciechanover et al., 1984, 1980; Finley et al., 1984; Hershko et al., 1980). The ubiquitin-system plays a crucial role in diverse cellular processes, including DNA

repair, transcription, synaptic plasticity, and regulation of the cell cycle, wherein ubiquitin-mediated proteolysis is a key regulatory step (Bennett and Harper, 2008; Dhananjayan et al., 2005; Finley et al., 2004; Glickman and Ciechanover, 2002; Hershko and Ciechanover, 1998; Hershko et al., 2000; Kawabe and Brose, 2011; Nakayama and Nakayama, 2005; Pickart, 2004; Varshavsky, 2005).

Ubiquitin is directed onto specific substrate proteins by E3 ubiquitin ligases (Ardley and Robinson, 2005; Dikic and Robertson, 2012; Glickman and Ciechanover, 2002; Hershko and Ciechanover, 1998), a large class of enzymes (over 600 predicted genes in humans; Li et al., 2008) for which many of the target specificity and function remain poorly understood. Identification of the biological roles of the E3s has been facilitated by siRNA screens using human cells *in vitro*, and by genetic screens in model organisms, such as *Drosophila* and *C. elegans* (Williamson et al., 2013). Specific roles for ubiquitin ligases have been demonstrated in embryonic stem cell fate determination (Werner et al., 2017; Xu et al., 2009), eye development (Ou et al., 2003), and neural development (Boix-Perales et al., 2007; Bury et al., 2008; J. Chen et al., 2012; D'Arcia et al., 2010; Hoeck et al., 2010; Sobieszczak et al., 2010;

* Correspondence to: Department of Biology, San Diego State University, 5500 Campanile Dr., San Diego, CA 92182-4614, USA.

E-mail address: rzayas@mail.sdsu.edu (R.M. Zayas).

¹ Equal authorship.

<http://doi.org/10.1016/j.ydbio.2017.10.011>

Received 6 June 2017; Received in revised form 25 September 2017; Accepted 11 October 2017

Available online 25 December 2017

0012-1606/ © 2017 Elsevier Inc. All rights reserved.

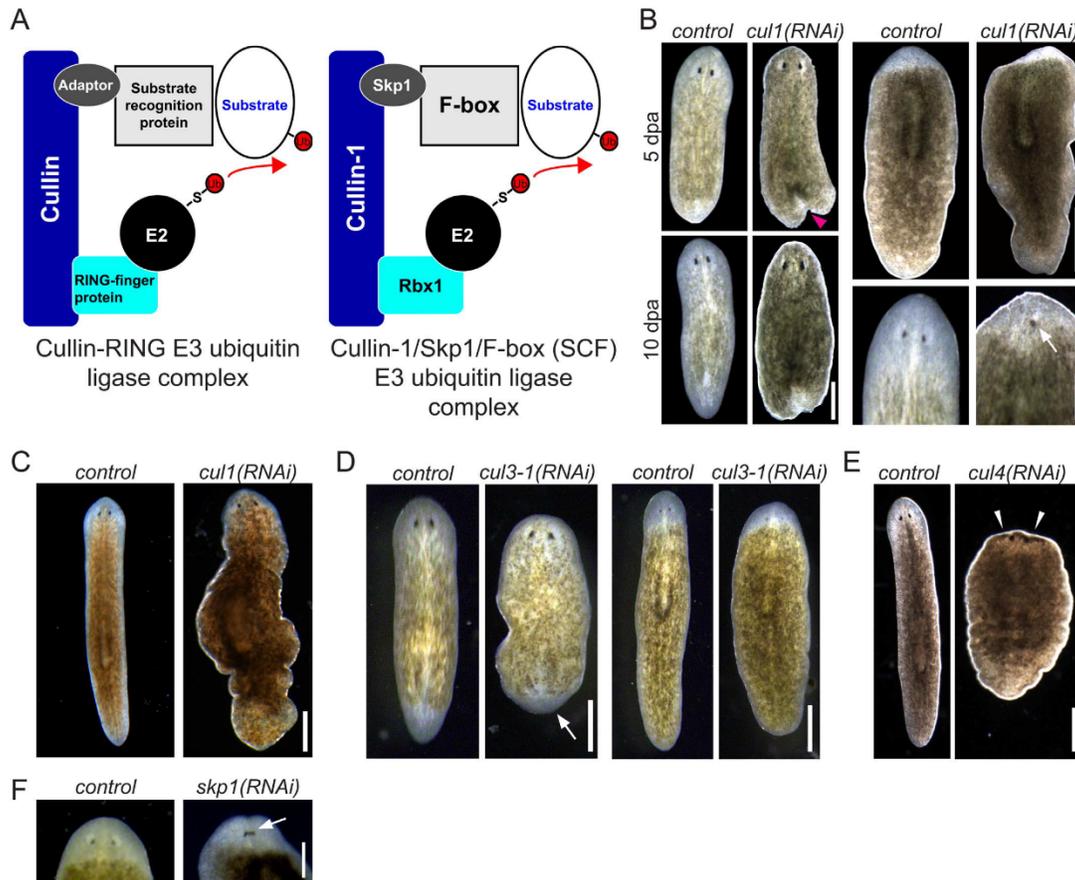


Fig. 1. Planarians possess multiple Cullin genes with distinct roles in regeneration and tissue homeostasis. **A)** Diagram summarizing the general organization of Cullin-RING E3 ubiquitin ligase complexes or the prototypical Skp1/Cullin-1/F-box (SCF) E3 ubiquitin ligase complex. E3 ubiquitin ligases transfer ubiquitin to the substrate by forming an isopeptide bond. **B)** Animals were fed dsRNA 6 times over 3 weeks against *gfp* (controls; n = 22) or *cul1* (n = 22), amputated pre-pharyngeally and allowed to regenerate for 10 days. Magenta arrowhead in head regenerate at 5 dpa indicates indented blastema in *cul1*(RNAi) planarian. White arrow marks abnormal regeneration in a single eye spot in the head of a *cul1*(RNAi) trunk regenerate. **C)** Animals were fed dsRNA over 6–8 weeks against *gfp* (controls; n = 38) or *cul1* (n = 38). All *cul1*(RNAi) worms showed loss of mobility, formed lesions, and subsequently lysed. **D)** Animals were fed dsRNA 5 times over 3 weeks against *gfp* (control; n = 30) or *cul3-1* (n = 21), amputated pre-pharyngeally, and allowed to regenerate for 10 days. *cul3-1*(RNAi) planarians showed delayed regeneration and differentiation when compared to the controls. White arrow marks the small blastema observed in head regenerates. **E)** Animals were fed dsRNA 5 times over 3 weeks against *gfp* (control; n = 30) or *cul4* (n = 21), amputated pre-pharyngeally, and allowed to regenerate for 10 days. **F)** Animals were fed dsRNA 6 times over 3 weeks against *gfp* (control; n = 50) or *skp1* (n = 45), amputated pre-pharyngeally, and allowed to regenerate for 10 days. White arrow denotes abnormal eye regeneration. Scale bars = 500 μ m.

Voigt and Papanopulu, 2006; Zhao et al., 2008; Zhu et al., 2005). Ubiquitin system components regulate regeneration in nematodes, flies, and mice, and are specifically upregulated during regeneration in sea cucumbers and axolotls (Hindi and Kumar, 2016; Pasten et al., 2012; Rao et al., 2009; Tian and Wu, 2013).

We are utilizing planarians as a model system to investigate the roles of E3 ubiquitin ligases in tissue regeneration. Previously, our lab demonstrated that members of the HECT E3 ligase gene family, which directly catalyze ubiquitin transfer onto a substrate via a ubiquitin-HECT complex intermediate (Metzger et al., 2012), are required for diverse aspects of regeneration in the planarian *Schmidtea mediterranea* (Henderson et al., 2015). In contrast to the HECT family, most E3 enzymes do not directly bind and transfer ubiquitin but rather coordinate the transfer of ubiquitin from an E2 onto a substrate, often through multimeric complexes, including the Cullin-RING ligase (CRL)

family (Sarikas et al., 2011). Cullin proteins act as molecular scaffolds that organize the binding of other elements to form an E3 complex that requires a substrate recognition subunit (Fig. 1A). These recognition subunits confer target specificity for ubiquitylation and their differential utilization allows modularity within CRL classes, thereby enabling function in multiple aspects of cellular biology.

In this study, we analyzed CRL function in tissue regeneration by inhibiting genes encoding major components of these complexes in planarians. First, we identified and performed RNAi against *cullin* genes present in *S. mediterranea* and found that homologs of Cullin-1, -3 and -4 are involved in regulating tissue homeostasis and regeneration. Robust *cullin-1* RNAi phenotypes included lesions, lysis, and defects in blastema formation, organ regeneration, and homeostatic tissue maintenance. Cullin-1 is a core component of the canonical CRL, the SCF (Skp1/Cullin-1/F-box)-E3 ubiquitin ligase complex (Fig. 1A)

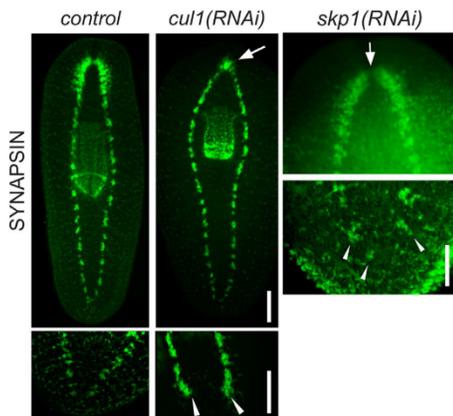


Fig. 2. *cullin-1* and *skp1* are required for normal nervous system regeneration. Animals were fed dsRNA 6 times over 3 weeks against *gfp*, *cull1* or *skp1*, amputated pre-pharyngeally, allowed to regenerate for 10 days, and stained with anti-SYNAPSIN to visualize nervous system regeneration. White arrow marks small mass of neuropil adjoining the ventral nerve cords or lack of anterior commissure formation in trunk regenerates and the arrowheads mark the ventral nerve cord stumps of head regenerates in *cull1(RNAi)* and *skp1(RNAi)* planarians. Scale bars: 250 μ m for whole worm and 100 μ m for tail inset for *cull1* RNAi; 250 μ m for *skp1* RNAi.

(Sarikas et al., 2011). By combining *in situ* hybridizations and RNAi screens targeting SCF-complex member genes that encode homologs of Skp1 and F-box proteins, we gained insight into the function of *cullin-1*-mediated cellular processes. We found that *skp1* is ubiquitously expressed while *f-box* genes are expressed in diverse tissue-types. Knockdown of *skp1* recapitulated *cullin-1* RNAi phenotypes and knockdown of various *f-box* genes recapitulated aspects of the *cullin-1* RNAi phenotype. Our study provides evidence that perturbation of elements within multimeric E3 ubiquitin ligase complexes can identify discrete functions for E3 ubiquitin ligases in regenerative processes and demonstrates conserved biological roles for a subset of the *f-box* genes, which are involved in stem cell regulation and cancer biology in other organisms.

2. Results

2.1. Cullin homologs are broadly expressed in stem cells and differentiated tissues of *S. mediterranea*

To examine the role of CRLs in planarian regeneration, we first searched the *S. mediterranea* genome and transcriptomes using human Cullin protein sequences and found six genes predicted to encode homologs of Cullin-1 through -5: *Smed-cullin-1*, *Smed-cullin-2*, *Smed-cullin-3-1*, *Smed-cullin-3-2*, *Smed-cullin-4* and *Smed-cullin-5* (see Table S1; hereon the *S. mediterranea* homologs will be referred to as *cull1*, etc.). In addition, we performed phylogenetic analysis to establish the relationship of the *S. mediterranea* *cullin* genes to well-characterized homologs in other species (Fig. S1). These genes are predicted to have broad expression in multiple *S. mediterranea* cell types by single-cell RNA-seq data, including neoblasts and differentiated tissues (Fig. S2A) (Wurtzel et al., 2015). Consistent with the single-cell expression data, whole-mount *in situ* hybridization (WISH) showed that *cullin* genes are broadly expressed in planarians (Fig. S2A-B), suggesting roles for these genes in multiple tissues.

2.2. Cullin gene knockdown leads to pleiotropic defects in regeneration and tissue homeostasis

Amputation of planarian tissues provides a simple paradigm to assess the role of genes in stem cell-based tissue renewal. We examined the function of individual *cullin* genes using RNAi, amputating the worms pre-pharyngeally and allowing the animals to regenerate for up to 10 days (a time-point by which the worms have fully regenerated and patterned their tissues). *cull1(RNAi)* worms had defects in blastema formation, cell differentiation (e.g., delayed optic pigment cup regeneration), and possible disruption of midline patterning signaling (e.g., single observable eye spot) compared to control worms (Fig. 1B). Extended *cull1(RNAi)* treatment resulted in defects in uninjured worms, including slow inching-like locomotion, lesions, and eventual lysis (Fig. 1C). *cul3-1(RNAi)* worms exhibited defects in blastema formation and photoreceptor regeneration (Fig. 1D), and *cul4(RNAi)* worms displayed ventral curling, lesions, and subsequent lysis of the worms after amputation (Fig. 1E), as well as during homeostasis ($n = 21$; not shown). These data identify roles for multiple *cullin* genes in planarian regeneration, and indicate that *cull1* and *cul4* are essential for planarian survival.

Motivated by our interest in the robust *cull1* phenotypes, we sought to investigate further the potential association of this gene to the canonical SCF complex (Fig. 1A). Skp1 mediates the interaction between Cullin-1 and F-box proteins (Ardley and Robinson, 2005; Sarikas et al., 2011). We identified a planarian homolog of Skp1, *Smed-skp1* (*skp1*) (dd_Smed_v6_1337_0_1; *Rattus norvegicus*, Q6PEC4, BLASTX = $5.4e^{-91}$), and found that, similar to *cull1*, *skp1* has broad expression in the animal (Fig. S2A-B). Furthermore, we reasoned that *skp1* RNAi should phenocopy *cull1* inhibition. Indeed, *skp1(RNAi)* planarians showed a delay in blastema formation and photoreceptor regeneration by 10 days post amputation (dpa) (Fig. 1F). We further examined their function in tissue differentiation. *cull1* and *skp1* are expressed in neurons (Fig. S2A-B); thus, we hypothesized that these genes regulate neuronal regeneration. To visualize the central nervous system, we labeled *control(RNAi)*, *cull1(RNAi)*, and *skp1(RNAi)* worms with anti-SYNAPSIN (Fig. 2). During anterior regeneration, a drastic impairment in brain formation, with absent or narrowed cephalic ganglia, was observed in both *cull1(RNAi)* and *skp1(RNAi)* worms. Similarly, the ventral nerve cords did not regenerate in *cull1(RNAi)* and *skp1(RNAi)* head regenerates (Fig. 2). These data support the hypothesis that *cull1* and *skp1* function together in putative SCF complexes of *S. mediterranea*. Therefore, we sought to dissect the roles of specific SCF complexes by examining the function and expression of genes homologous to the F-box substrate recognition partners.

2.3. Analysis of F-box genes in *S. mediterranea*

To screen for specific roles of SCF complexes in planarians, we identified F-box-encoding genes in *S. mediterranea*. F-box proteins are defined by an F-box domain, which is necessary for SKP1-F-box protein binding (Bai et al., 1996). BLAST searches were performed using F-box domains from diverse species (see Materials and Methods) and identified 35 *S. mediterranea* *f-box* genes (Table S2), all of which contain an F-box domain and were classified by homology to one of three categories that are defined by the presence of other protein-protein interacting domains: F-box “only” (FBX), F-box leucine-rich repeat containing (FXL), or F-box WD40-repeat containing (FBW) (Table S2).

We hypothesized that F-box proteins, which confer SCF target specificity, mediate discrete roles for *cullin-1* in planarian tissues and that these functions can be genetically dissected using RNAi. Thus, we performed an RNAi screen against 30 *f-box* genes to determine if these genes can phenocopy aspects of the *cull1* and *skp1* phenotypes. Worms were treated with dsRNA and observed for defects in homeostasis and mobility before being amputated pre-pharyngeally and observed

Table 1
Functional analysis of F-box encoding genes in *Schmidtea mediterranea*.

Smed Gene ID	Phenotype	Nervous system patterning phenotype
<i>Smed-btrep/fbw1a</i>	Lesions/lysis (13/13)	Not analyzed
<i>Smed-ect2-like</i>	Delayed regeneration/patterning (14/27)	Not analyzed
<i>Smed-fbw-3</i>	Delayed regeneration/patterning (10/43), loss of mobility (10/43)	None observed
<i>Smed-fbw7-like-1</i>	Delayed regeneration/patterning (12/26)	Cephalic ganglia anterior commissure defect (7/14)
<i>Smed-fbw7-like-2</i>	Delayed regeneration/patterning (28/48)	Cephalic ganglia anterior commissure defect (7/14) and decreased neuropil density (7/14)
<i>Smed-fbx-10</i>	Delayed regeneration/patterning (25/31)	None observed
<i>Smed-fbx-11</i>	Delayed regeneration/patterning (27/42)	None observed
<i>Smed-fbx-2</i>	Delayed regeneration/patterning (15/15)	None observed
<i>Smed-fbx-4</i>	Delayed regeneration/patterning (13/30)	Cephalic ganglia anterior commissure defect (7/16)
<i>Smed-fbx36</i>	Delayed regeneration/patterning (8/27)	None observed
<i>Smed-fbx38</i>	Delayed regeneration/patterning (64/86)	Ventral nerve cord defect (4/5), reduced or absent cephalic ganglia (14/22)
<i>Smed-fbx8</i>	Delayed regeneration/patterning (12/43)	None observed
<i>Smed-fxl-2</i>	Delayed regeneration/patterning (10/29)	None observed
<i>Smed-fxl-3</i>	Delayed regeneration/patterning (8/22)	Cephalic ganglia anterior commissure defect (8/14) and decreased neuropil density (12/14)
<i>Smed-fxl13</i>	Delayed regeneration/patterning (10/32)	None observed
<i>Smed-fxl16</i>	Delayed regeneration/patterning (12/20)	None observed
<i>Smed-fxl2-1</i>	Delayed regeneration/patterning (24/47)	Reduced neuropil density of cephalic ganglia (7/26)
<i>Smed-fxl20</i>	Delayed regeneration/patterning (24/32)	Reduced neuropil density of cephalic ganglia (8/11)
<i>Smed-morgue/ubc2</i>	Delayed regeneration/patterning (27/36)	Cephalic ganglia anterior commissure defect (9/13), decreased neuropil density (4/13)

through regeneration for 10 days. We found that RNAi knockdown of 19 *f-box* genes led to defects in homeostasis or regeneration, like delayed regeneration (e.g., small blastemas), defects in blastema patterning (abnormal eye regeneration), mobility defects, lesions, and lysis (Table 1 and S2), which are similar to the phenotypes observed in *cull1* and *skp1* RNAi planarians. Worms fed RNAi against *btrep* displayed body shape defects with tissue outgrowths, lesions, ventral curling, and lysis (Fig. 3A). These phenotypes were similar to those observed in *cull1*(RNAi) worms, suggesting that a planarian β -TrCP-containing SCF complex (i.e., SCF $^{\beta$ -TrCP) regulates cell signaling and survival. Additional RNAi phenotypes included loss of mobility (*fbw-3*) and blastema defects (*fbw7-like-1*, *fbw7-like-2*, *fbx38*, *fxl-3*, *fxl2-1*, and *fxl20*) (Table 1 and S2, Fig. 3B).

To further examine how *f-box* genes are involved in stem cell function or organ regeneration/patterning, we repeated RNAi experiments for 17 *f-box* genes and stained the worms with markers for the nervous system or mitosis. Anti-SYNAPSIN staining revealed defects in nervous system regeneration in eight of the RNAi treatment groups, such as a failure of the cephalic ganglia to connect at the anterior commissure, small or absent cephalic ganglia, and decreased neuropil density (Table 1, Fig. 3B). We also stained control and F-box RNAi planarians with anti-phospho-Histone H3 (PH3) to label mitotic cells. We found that nine genes displayed a decrease in the number of PH3⁺ cells, while one gene, *fxl2-1*, showed an increase in PH3⁺ cells at 10 dpa (Fig. 3C and S3A).

To determine if the expression of F-box encoding genes correlates with the cell- or tissue-specific phenotypes in planarians, we performed WISH against the 19 genes that showed phenotypes during screening. Six of the eight genes involved in nervous system regeneration had strong expression in the central nervous system (*fbw7-1*, *fbw7-2*, *fbx38*, *fxl2-1*, *fxl20*, and *morgue*) (Fig. 3D). We also observed tissue-specific gene expression patterns, such as *fbx-10* and *fxl13*, which displayed neural and pharyngeal expression patterns (Fig. 3D). Importantly, single-cell RNA-seq data support that *f-box* genes are expressed in diverse tissue types; furthermore, all 19 *f-box* genes producing RNAi phenotypes are predicted to have expression in *cull1*-expressing cells (Fig. S3B). Combined with the results from the RNAi screen, these data further indicate that *f-box* expression may be dictating tissue-specific roles of SCF complexes, including neoblast proliferation and differentiation, blastema formation, and organ patterning.

3. Discussion

In this study, we examined the function of Cullin-RING ligase (CRL) genes in planarian regeneration. Our work revealed definitive roles for three Cullin genes, *cull1*, *cull3-1* and *cull4*, in planarian regeneration and survival. Our results are consistent with extensive studies implicating these genes in stem cell biology and cell cycle regulation (Nakayama and Nakayama, 2005; Werner et al., 2017). We did not examine the function of the Anaphase Promoting Complex (APC) Cullin-domain subunit (Apc2) since previous studies indicate the APC has a conserved role in cell division (RNAi knockdown of a planarian CDC23 homolog, a component of the APC, blocks mitosis) (Azimzadeh et al., 2012; Reddien et al., 2005a). We had two major objectives in this work: to determine if multimeric E3 complexes could be studied using RNAi methodology in planarians, and further, to dissect the specific functions of E3 complex genes in stem cell regulation and tissue regeneration. Based on the results of screening the *cullin* genes, we focused on analyzing the highly conserved Skp1/Cullin-1/F-box (SCF) complex components. Knocking down *cull1* or *skp1* led to homeostasis and regeneration-specific phenotypes, like defects in blastema formation and nervous system regeneration. Due to the expected pleiotropic nature of the phenotypes observed in *cull1* and *skp1* RNAi worms, we subsequently analyzed F-box protein-encoding genes and found that knockdown of 19 genes, all of which are predicted to have expression in *cull1*⁺ cells, phenocopied aspects of the *cull1/skp1* RNAi phenotypes. Moreover, our results implicate a subset of *f-box* genes in cell cycle regulation.

Because of their well-appreciated roles in stem cell biology and cancer (Maneix and Catic, 2016; Strikoudis et al., 2014; Werner et al., 2017), F-box proteins are important candidates for mechanistic evaluation and drug design. For example, SCF $^{\beta$ -TrCP ubiquitylates substrates with key roles in signal transduction pathways that underlie many aspects cell division, development, and tumorigenesis (Fuchs et al., 2004; Willems et al., 2004). SCF $^{\beta$ -TrCP is involved in regulating β -Catenin stability (Stamos and Weis, 2013) and it is intriguing that *btrep* inhibition in *S. mediterranea* did not result in the obvious patterning defects characteristic of RNAi knockdown against canonical Wnt signaling components (Almuedo-Castillo et al., 2012). This could be due to the strength of the phenotype following *btrep* RNAi, which included tissue outgrowths, ventral curling, and lysis. Although these phenotypes are consistent with a function of SCF $^{\beta$ -TrCP in stem cell biology (Werner et al., 2017), it would be interesting to resolve a

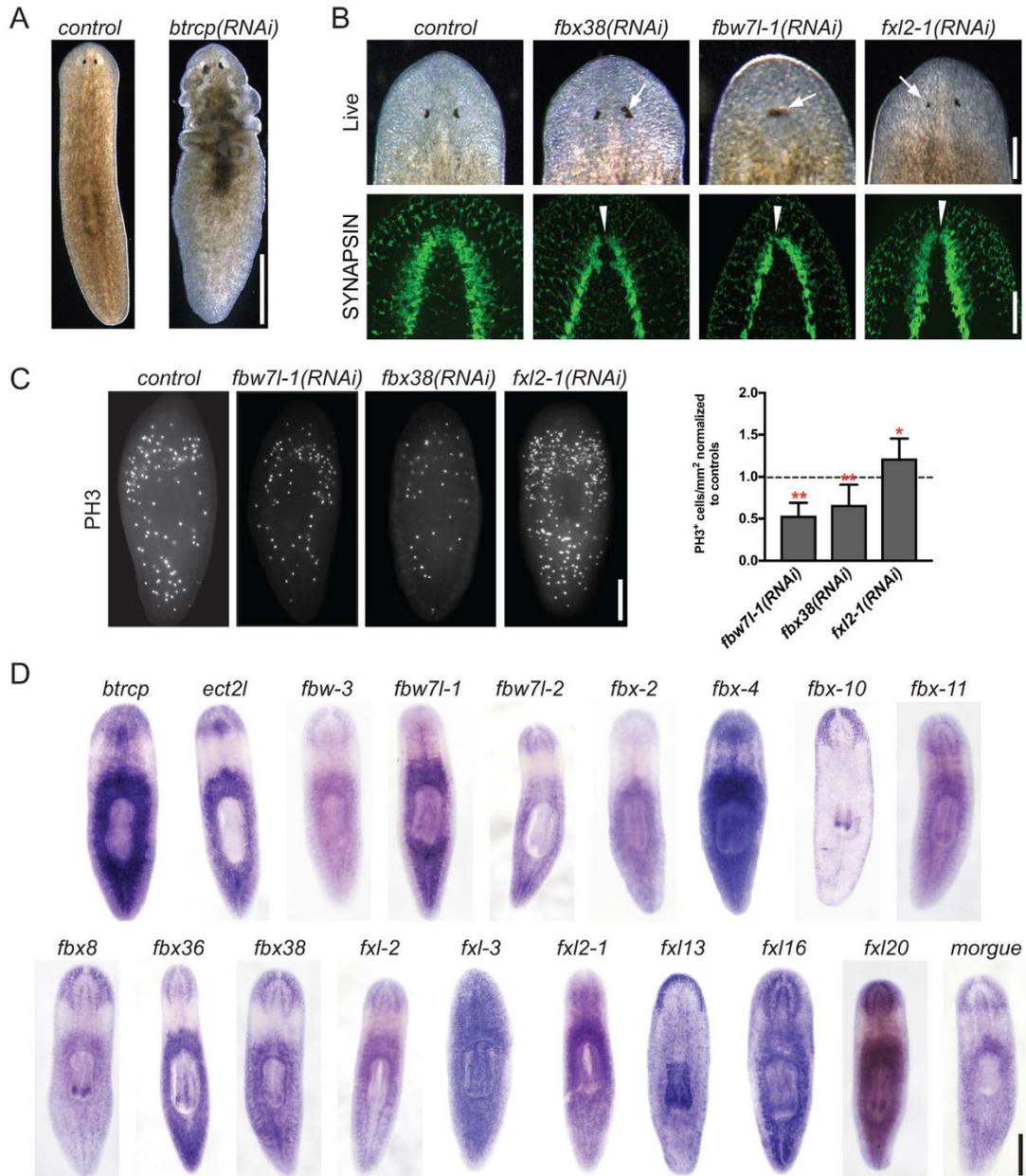


Fig. 3. RNAi and expression analysis of F-box genes in *S. mediterranea*. **A**) Animals were fed dsRNA 6 times over 3 weeks against *gfp* (control; n = 13) or *btrcp* (n = 13). **B**) Animals were fed dsRNA 6 times over 3 weeks against *gfp* (control; n = 26), *fbx38* (n = 22), *fbw71-like-1* (*fbw71-1*; n = 14), or *fxl2-1* (n = 26), amputated pre-pharyngeally, and allowed to regenerate for 10 days. Trunk-regenerate heads of live animals were imaged. Subsequently, animals were sacrificed, fixed and stained with anti-SYNAPSIN to visualize the nervous system. The white arrows in the live animals denote asymmetric or improperly patterned eyes in F-box gene RNAi worms; arrowheads in SYNAPSIN images mark failed or reduced anterior commissures observed in RNAi planarians. **C**) Animals were fed dsRNA 6 times over 3 weeks against *gfp* (control; n = 10 for each group), *fbx38* (n = 10), *fbw71-like-1* (*fbw71-1*; n = 10), or *fxl2-1* (n = 10), amputated pre-pharyngeally, allowed to regenerate for 10 days, and then stained with anti-phospho-Histone H3 (PH3) to visualize mitotic neoblasts. Graph shows the mean \pm s.d. of values normalized to *gfp* controls. *P < 0.05 or **P < 0.01, Student's *t*-test. **D**) Whole-mount *in situ* hybridization to *f-box* genes that showed a phenotype following RNAi. Abbreviations: *ect2l*, *ect2-like*; *fbw71-1*, *fbw71-like-1*; *fbw71-2*, *fbw71-like-2*. Scale bars: A = 500 μ m, B = 250 μ m for live images, 100 μ m for SYNAPSIN staining, C = 250 μ m, D = 500 μ m.

potential role for *btrcp* in canonical Wnt signaling by analyzing the *btrcp(RNAi)* tissue outgrowths or performing Western blot to *Smed- β -catenin-1* following RNAi (Stuckemann et al., 2017; Sureda-Gomez et al., 2016).

FBXW7 proteins are known tumor suppressors and regulators of stem cell differentiation (Takeishi and Nakayama, 2014). In *S. mediterranea*, two *fbxw7*-like genes (*fbw7-like-1* and *fbw7-like-2*) had roles in eye regeneration, blastema formation, mitosis, and nervous system regeneration (Fig. 3). Interestingly, loss of *Fbxw7* in mice leads to the accumulation of neural stem cells and loss of differentiated neurons (Hoeck et al., 2010; Matsumoto et al., 2011). *Fbxw7* is also a known regulator of the Notch signaling pathway (Matsumoto et al., 2011), which has recently been implicated in midline formation in the regenerating planarian (Sasidharan et al., 2017). The nervous system patterning and eye regeneration phenotypes observed in *cul1*, *skp1* and *fbw7-like-1* RNAi worms are all consistent with possible defects in midline formation (Fig. 2B, 3B, S3B). Intriguing preliminary results suggest that *cul1(RNAi)* worms fail to express *slit1* (Cebrià et al., 2007) at the midline of head blastemas (data not shown). However, given the pleiotropic nature of the *cul1* phenotype, future studies on the role of *fbw7-like-1* and *-2* in planarian regeneration are a logical next step to investigate if these genes are directly regulating neurogenesis and body patterning, and what proteins are being ubiquitinated by putative SCF^{Fbxw7-1} or SCF^{Fbxw7-2} complexes.

Another F-box that our data implicated in planarian regeneration is *fxl2-1*, a homolog of the tumor suppressor *FBXL2* (B.B. Chen et al., 2012). *FBXL2* targets Cyclin D3 to arrest mitotic activity in human and mouse lung epithelial cells (Chen et al., 2011). *fxl2-1* RNAi in planarians led to a significant increase in PH3⁺ cells (Fig. 3C), suggesting that *fxl2-1* may have a conserved function in regulating cell division. Few gene knockdowns lead to hyperproliferation phenotypes in planarians (e.g., *Smed-p53*, *-smg-1*, and *-huwe1*) (González-Estévez et al., 2012; Henderson et al., 2015; Pearson and Sánchez Alvarado, 2010), yet this work has identified a putative SCF^{Fbxw7-1} complex as a candidate suppressor of mitosis. Further analysis of the *fxl2-1* RNAi phenotype using proteomic approaches (discussed below) could uncover other factors that regulate stem cell populations.

fbx38(RNAi) worms phenocopied several *cul1* and *skp1* RNAi phenotypes, like defects in blastema formation and nervous system regeneration, and a reduction in mitotic activity (Table 1 and S2, and Fig. 3B-C). Mutations in *FBX38* have been identified in patients with bipolar disorder and spinal muscular atrophy (Feng and Zhu, 2010; Sumner et al., 2013). Additional studies of *fbx38* in *S. mediterranea* could inform how mutations in *FBX38* cause nervous system disorders in humans.

Previous work from our laboratory revealed roles for HECT E3 ubiquitin ligases in planarian regeneration and tissue patterning (Henderson et al., 2015). Here, we demonstrate that it is possible to dissect the function of CRL complexes within the context of adult tissue regeneration, accentuating the prospect of using planarians as a model to investigate how ubiquitylation regulates regenerative processes. We found that knockdown of *cul1* or *skp1* led to several phenotypes, and that *f-box* gene knockdowns recapitulated aspects of the *cul1(RNAi)* and *skp1(RNAi)* phenotypes, suggesting we can isolate the effects of specific SCF complexes *in vivo*. Further study of *f-box* genes, and of other substrate recognition proteins in CRL complexes, will define the specific roles of these complexes in stem cell regulation. Importantly, F-box domains, either alone or in combination with other domains, are involved in protein-protein interactions that could also function in other cellular processes outside of ubiquitylation (Hermand, 2006). Future studies should expand analysis of specific *f-box* genes on the neoblast population and determine the impact of RNAi knockdowns on protein ubiquitylation in planarians, which can be measured by Western blot analysis (Henderson, 2013). Once a direct link to ubiquitin signaling is established, it should be possible to apply quantitative techniques, such as DiGly proteomics (Bennett et al.,

2010; Kim et al., 2011), to identify molecular targets of SCF ubiquitin ligases involved in planarian stem cell regulation.

4. Materials and methods

4.1. Planarian Husbandry

A clonal, asexual strain of *Schmidtea mediterranea* (CIW4) was maintained as described previously (Cebrià and Newmark, 2005). Animals ranging in length from 3 to 6 mm were starved for at least one week prior to all experiments.

4.2. Informatics

TBLASTN analysis was performed using human CULLIN proteins or F-box protein domains from *S. cerevisiae*, *S. pombe*, *C. elegans*, *D. melanogaster*, *X. laevis*, *M. musculus* and *H. sapiens* against planarian transcriptomes (Adamidi et al., 2011; Brandl et al., 2016). The returned sequences were subsequently analyzed with the NCBI Conserved Domain Database Search tool (nucleotide sequences), SMART (longest ORF translation), and InterProScan (longest ORF translation) to confirm the presence of an F-box domain. Transcripts with an E-value $\leq 1e^{-03}$ from the listed domain predicting programs were considered an F-box domain-containing transcript (Table S2).

4.3. Phylogenetic analysis

S. mediterranea *cullin* transcripts (Table S1) were translated in Virtual Ribosome (Wernersson, 2006) and aligned to annotated Cullin protein sequences from *Caenorhabditis elegans*, *Drosophila melanogaster*, *Homo sapiens*, and *Mus musculus*, by ClustalW within MEGA7 software (Kumar et al., 2016) using default settings. Phylogenetic relationships were inferred using the Neighbor-Joining method in MEGA7. Evolutionary distances were computed using the Jones-Taylor-Thornton model with uniform mutation rates and pairwise deletion of gaps. Phylogenetic relationships were tested by Bootstrap resampling with 1000 replications and displayed rooted at the midpoint.

4.4. Cloning

Three Cullin genes and 13 F-box containing genes were obtained from a library of sequenced cDNA clones in pBluescript II SK (+) (Zayas et al., 2005) (Table S3). The remaining three Cullin and 22 F-box genes were directionally cloned into the pJC53.2 vector using forward and reverse primers equipped with XhoI and NotII restriction sites (Collins et al., 2010) or with primers containing overhangs homologous to the pPR-T4P vector (J. Rink) using ligation-independent cloning (Aslanidis and de Jong, 1990) (Table S3). Inserts in the pBluescript II SK (+) vector were shuttled into either pPR242 or pPR244 using the Gateway system (Invitrogen) (Reddien et al., 2005b).

4.5. Whole-mount *in situ* hybridization

Animals were sacrificed and processed using a 5% N-Acetyl Cysteine solution prior to fixation in 4% formaldehyde. Antisense RNA probes were produced as previously described (Pearson et al., 2009). Briefly, DNA templates were PCR amplified from cDNA clones in pJC53.2, pBS, pPR244/242, and pPR-T4P plasmid vectors. Antisense riboprobes labeled with digoxigenin were synthesized at 37 °C. Probes were purified via ethanol precipitation and whole-mount *in situ* hybridization was performed in an InSituPro VS liquid handling robot (Intavis). Samples were incubated with an anti-digoxigenin AP antibody (Roche, 1:2000) and developed with NBT/BCIP as previously described (Pearson et al., 2009).

4.6. RNA interference

Plasmid templates in pJC53.2, pPR244/242, or pPR-T4P were transformed into the RNase-free cell line HT115 (DE3) for double-stranded RNA production by using IPTG inducible promoters (Collins et al., 2010; Reddien et al., 2005b). Bacterially-expressed *gfp* dsRNA was used as a control in all experiments. Animals were fed a dsRNA bacterial pellet mixed with an approximate 3:1 ratio of liver:water paste. Five to six RNAi feedings were performed over a period of three weeks and animals were cut pre-pharyngeally 24–48 h after the last feed. Planarians were fixed following 10 days of regeneration. All experiments were performed in duplicate at a minimum.

4.7. Immunostaining

Fixation and immunostaining with anti-phospho-Histone H3 (S10) (1:2000, Cell Signaling) and anti-Synapsin (1:400, Hybridoma bank) were performed as described in Cowles et al. (2013). Anti-phospho-Histone H3 (S10) was visualized with Cy3-Tyramide following incubation with goat anti-rabbit-HRP secondary antibodies (1:2000); anti-Synapsin was visualized with goat anti-mouse Alexa Fluor 488 (Thermo Fisher Scientific, A11029; 1:1000, Waltham, MA).

4.8. Image acquisition and processing

Fluorescent labeled images were acquired via an AxioCam MRm camera mounted on a Zeiss Axio Observer. Z1 equipped with ApoTome or a Zeiss SteREO Lumar V.12 for whole body images. Live RNAi and in situ images were taken with a Leica DFC290 or DFC450 camera on a Leica M205 microscope. All images are of the dorsal worm with the anterior at the top. Images were processed for brightness and the figures organized using the Adobe Creative Suite.

4.9. Statistics

Quantification of anti-phospho-Histone H3⁺ cells was done by manually counting cells on ImageJ. Cell counts were normalized to the area. Student's *t*-tests were performed and graphs made on GraphPad Prism Version 6 (GraphPad Software, San Diego, CA). All graphs show mean with error bars displaying standard deviation (mean \pm s.d.). The number of animals per group is indicated in the appropriate figure legend.

Acknowledgements

We thank Maria Gutierrez for assisting with BLAST searches; Erica Campau and Andrew Kelm for assisting with RNAi experiments and imaging; Elizabeth Waters for help with phylogenetic analysis; Jordana Henderson, Ionit Iberkleid and Kelly Ross, for helpful discussions and comments on the manuscript. This work was funded by the California Institute for Regenerative Medicine RN2-00940-1 and NSF IOS-1350302.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.ydbio.2017.10.011>.

References

Adamidi, C., Wang, Y., Gruen, D., Mastrobuoni, G., You, X., Tolle, D., Dodt, M., Mackowiak, S.D., Gogol-Doering, A., Oenal, P., Rybak, A., Ross, E., Sanchez Alvarado, A., Kempa, S., Dieterich, C., Rajewsky, N., Chen, W., 2011. De novo assembly and validation of planaria transcriptome by massive parallel sequencing and shotgun proteomics. *Genome Res* 21, 1193–1200.

Almuedo-Castillo, M., Sureda-Gomez, M., Adell, T., 2012. Wnt signaling in planarians: new answers to old questions. *Int. J. Dev. Biol.* 56, 53–65.

Ardley, H.C., Robinson, P.A., 2005. E3 ubiquitin ligases. *Essays Biochem.* 41, 15–30.

Aslanidis, C., de Jong, P.J., 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* 18, 6069–6074.

Azhizadeh, J., Wong, M.L., Downhour, D.M., Sanchez Alvarado, A., Marshall, W.F., 2012. Centrosome loss in the evolution of planarians. *Science* 335, 461–463.

Baguña, J., 2012. The planarian neoblast: the rambling history of its origin and some current black boxes. *Int. J. Dev. Biol.* 56, 19–37.

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebel, M., Harper, J.W., Elledge, S.J., 1996. SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* 88, 263–274.

Bennett, E.J., Harper, J.W., 2008. DNA damage: ubiquitin marks the spot. *Nat. Struct. Mol. Biol.* 15, 20–22.

Bennett, E.J., Rush, J., Gygi, S.P., Harper, J.W., 2010. Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. *Cell* 143, 951–965.

Boix-Perales, H., Horan, I., Wise, H., Lin, H.R., Chuang, L.C., Yew, P.R., Philpott, A., 2007. The E3 ubiquitin ligase skip2 regulates neural differentiation independent from the cell cycle. *Neural Dev.* 2, 27.

Brandl, H., Moon, H., Vila-Farre, M., Liu, S.Y., Henry, I., Rink, J.C., 2016. PlanMine—a mineable resource of planarian biology and biodiversity. *Nucleic Acids Res.* 44, D764–D773.

Bury, F.J., Moers, V., Yan, J., Soungui, J., Quan, X.J., De Geest, N., Kricha, S., Hassan, B.A., Bellefroid, E.J., 2008. Xenopus BTBD6 and its Drosophila homologue lute are required for neuronal development. *Dev. Dyn.* 237, 3352–3360.

Cebrià, F., Guo, T., Jopek, J., Newmark, P.A., 2007. Regeneration and maintenance of the planarian midline is regulated by a slit orthologue. *Dev. Biol.* 307, 394–406.

Cebrià, F., 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.

Chen, B.B., Glasser, J.R., Coon, T.A., Mallampalli, R.K., 2011. FBXL2 is a ubiquitin E3 ligase subunit that triggers mitotic arrest. *Cell Cycle* 10, 3487–3494.

Chen, B.B., Glasser, J.R., Coon, T.A., Mallampalli, R.K., 2012a. F-box protein FBXL2 exerts human lung tumor suppressor-like activity by ubiquitin-mediated degradation of cyclin D3 resulting in cell cycle arrest. *Oncogene* 31, 2566–2579.

Chen, J., Lai, F., Niswander, L., 2012b. The ubiquitin ligase mLin41 temporally promotes neural progenitor cell maintenance through FGF signaling. *Genes Dev.* 26, 803–815.

Ciechanover, A., Finley, D., Varshavsky, A., 1984. The ubiquitin-mediated proteolytic pathway and mechanisms of energy-dependent intracellular protein degradation. *J. Cell Biochem.* 24, 27–53.

Ciechanover, A., Heller, H., Elias, S., Haas, A.L., Hershko, A., 1980. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proc. Natl. Acad. Sci. USA* 77, 1365–1368.

Collins, J.J., 3rd, Hou, X., Romanova, E.V., Lambrus, B.G., Miller, C.M., Saberi, A., Sweedler, J.V., Newmark, P.A., 2010. Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLoS Biol.* 8, e1000509.

Cowles, M.W., Brown, D.D., Nisperos, S.V., Stanley, B.N., Pearson, B.J., Zayas, R.M., 2013. Genome-wide analysis of the bHLH gene family in planarians identifies factors required for adult neurogenesis and neuronal regeneration. *Development* 140, 4691–4702.

D'Arca, D., Zhao, X., Xu, W., Ramirez-Martinez, N.C., Iavarone, A., Lasorella, A., 2010. Huwe1 ubiquitin ligase is essential to synchronize neuronal and glial differentiation in the developing cerebellum. *Proc. Natl. Acad. Sci. USA* 107, 5875–5880.

Dhananjayan, S.C., Ismail, A., Nawaz, Z., 2005. Ubiquitin and control of transcription. *Essays Biochem.* 41, 69–80.

Dikic, I., Robertson, M., 2012. Ubiquitin ligases and beyond. *BMC Biol.* 10, 22.

Elliott, S.A., Sanchez Alvarado, A., 2012. The history and enduring contributions of planarians to the study of animal regeneration. *WIREs Dev. Biol.* 2, 301–326.

Feng, T., Zhu, X., 2010. Genome-wide searching of rare genetic variants in WTCCC data. *Hum. Genet.* 128, 269–280.

Finley, D., Ciechanover, A., Varshavsky, A., 1984. Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85. *Cell* 37, 43–55.

Finley, D., Ciechanover, A., Varshavsky, A., 2004. Ubiquitin as a central cellular regulator. *Cell* 116 (S29–32), 22, (following S32).

Fuchs, S.Y., Spiegelman, V.S., Kumar, K.G., 2004. The many faces of beta-TrCP E3 ubiquitin ligases: reflections in the magic mirror of cancer. *Oncogene* 23, 2028–2036.

Glickman, M.H., Ciechanover, A., 2002. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* 82, 373–428.

González-Estévez, C., Félix, D.A., Smith, M.D., Paps, J., Morley, S.J., James, V., Sharp, T.V., Aboobaker, A.A., 2012. SMG-1 and mTORC1 act antagonistically to regulate response to injury and growth in planarians. *PLoS Genet.* 8, e1002619.

Henderson, J., 2013. Analysis of HECT-Domain E3 Ubiquitin Ligases During Tissue Regeneration in the Planarian *Schmidtea mediterranea*. San Diego State University, San Diego, CA.

Henderson, J.M., Nisperos, S.V., Weeks, J., Ghulam, M., Marin, I., Zayas, R.M., 2015. Identification of HECT E3 ubiquitin ligase family genes involved in stem cell regulation and regeneration in planarians. *Dev. Biol.* 404, 21–34.

Hermant, D., 2006. F-box proteins: more than baits for the SCF? *Cell Div.* 1, 30.

Hershko, A., Ciechanover, A., 1998. The ubiquitin system. *Annu. Rev. Biochem.* 67, 425–479.

Hershko, A., Ciechanover, A., Heller, H., Haas, A.L., Rose, I.A., 1980. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc. Natl. Acad. Sci. USA* 77, 1783–1786.

Hershko, A., Ciechanover, A., Varshavsky, A., 2000. Basic medical research award. The ubiquitin system. *Nat. Med.* 6, 1073–1081.

- Hindi, S.M., Kumar, A., 2016. TRAF6 regulates satellite stem cell self-renewal and function during regenerative myogenesis. *J. Clin. Invest.* 126, 151–168.
- Hoeck, J.D., Jandke, A., Blake, S.M., Nye, E., Spencer-Dene, B., Brandner, S., Behrens, A., 2010. Fbw7 controls neural stem cell differentiation and progenitor apoptosis via Notch and c-Jun. *Nat. Neurosci.* 13, 1365–1372.
- Kawabe, H., Brose, N., 2011. The role of ubiquitylation in nerve cell development. *Nat. Rev. Neurosci.* 12, 251–268.
- Kim, W., Bennett, E.J., Hurlin, E.L., Guo, A., Li, J., Possemato, A., Sowa, M.E., Rad, R., Rush, J., Comb, M.J., Harper, J.W., Gygi, S.P., 2011. Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* 44, 325–340.
- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874.
- Li, W., Bengtson, M.H., Ulbrich, A., Matsuda, A., Reddy, V.A., Orth, A., Chanda, S.K., Batalov, S., Joazeiro, C.A., 2008. Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS One* 3, e1487.
- Maneix, L., Catic, A., 2016. Touch and go: nuclear proteolysis in the regulation of metabolic genes and cancer. *FEBS Lett.* 590, 908–923.
- Matsumoto, A., Onoyama, I., Sunabori, T., Kageyama, R., Okano, H., Nakayama, K.I., 2011. Fbw7-dependent degradation of Notch is required for control of "stemness" and neuronal-gial differentiation in neural stem cells. *J. Biol. Chem.* 286, 13754–13764.
- Metzger, M.B., Hristova, V.A., Weissman, A.M., 2012. HECT and RING finger families of E3 ubiquitin ligases at a glance. *J. Cell Sci.* 125, 531–537.
- Nakayama, K.I., Nakayama, K., 2005. Regulation of the cell cycle by SCF-type ubiquitin ligases. *Semin. Cell Dev. Biol.* 16, 323–333.
- Ou, C.Y., Pi, H., Chien, C.T., 2003. Control of protein degradation by E3 ubiquitin ligases in *Drosophila* eye development. *Trends Genet.* 19, 382–389.
- Pasten, C., Oriz-Pineda, P.A., Garcia-Ararras, J.E., 2012. Ubiquitin-proteasome system components are upregulated during intestinal regeneration. *Genesis* 50, 350–365.
- Pearson, B.J., Eisenhoffer, G.T., Gurley, K.A., Rink, J.C., Miller, D.E., Sanchez Alvarado, A., 2009. Formaldehyde-based whole-mount in situ hybridization method for planarians. *Dev. Dyn.* 238, 443–450.
- Pearson, B.J., Sánchez Alvarado, A., 2010. A planarian p53 homolog regulates proliferation and self-renewal in adult stem cell lineages. *Development* 137, 213–221.
- Pickart, C.M., 2004. Back to the future with ubiquitin. *Cell* 116, 181–190.
- Rao, N., Jhamb, D., Milner, D.J., Li, B., Song, F., Wang, M., Voss, S.R., Palakal, M., King, M.W., Saranjami, B., Nye, H.L., Cameron, J.A., Stocum, D.L., 2009. Proteomic analysis of blastema formation in regenerating axolotl limbs. *BMC Biol.* 7, 83.
- Reddien, P.W., Bermange, A.L., Murfitt, K.J., Jemings, J.R., 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., Bermange, A.L., Murfitt, K.J., Jemings, J.R., Sánchez Alvarado, A., 2005b. Identification of genes needed for regeneration, stem cell function, and tissue homeostasis by systematic gene perturbation in planaria. *Dev. Cell* 8, 635–649.
- Rink, J.C., 2013. Stem cell systems and regeneration in planaria. *Dev. Genes Evol.* 223, 67–84.
- Roberts-Galbraith, R.H., Newmark, P.A., 2015. On the organ trail: insights into organ regeneration in the planarian. *Curr. Opin. Genet. Dev.* 32, 37–46.
- Ross, K.G., Currie, K.W., Pearson, B.J., Zayas, R.M., 2017. Nervous system development and regeneration in freshwater planarians. *WIREs Dev. Biol.* 6, e266.
- Sarikas, A., Hartmann, T., Pan, Z.Q., 2011. The cullin protein family. *Genome Biol.* 12, 220.
- Sasidharan, V., Marepally, S., Elliott, S.A., Baid, S., Lakshmanan, V., Nayyar, N., Bansal, D., Sanchez Alvarado, A., Vemula, P.K., Palakodeti, D., 2017. The miR-124 family of microRNAs is crucial for regeneration of the brain and visual system in the planarian *Schmidtea mediterranea*. *Development* 144, 3211–3223.
- Sobieszczuk, D.F., Poliakov, A., Xu, Q., Wilkinson, D.G., 2010. A feedback loop mediated by degradation of an inhibitor is required to initiate neuronal differentiation. *Genes Dev.* 24, 206–218.
- Stamos, J.L., Weis, W.L., 2013. The beta-catenin destruction complex. *Cold Spring Harb. Perspect. Biol.* 5, a007898.
- Strikoudis, A., Guillamot, M., Aifantis, I., 2014. Regulation of stem cell function by protein ubiquitylation. *EMBO Rep.* 15, 365–382.
- Stuckemann, T., Cleland, J.P., Werner, S., Thi-Kim Vu, H., Bayersdorf, R., Liu, S.Y., Friedrich, B., Julicher, F., Rink, J.C., 2017. Antagonistic self-organizing patterning systems control maintenance and regeneration of the anteroposterior axis in planarians. *Dev. Cell* 40 (248–263), e244.
- Sumner, C.J., d'Ydewalle, C., Wooley, J., Fawcett, K.A., Hernandez, D., Gardiner, A.R., Kalmar, B., Baloh, R.H., Gonzalez, M., Zuchner, S., Stanescu, H.C., Kleta, R., Mankodi, A., Cornblath, D.R., Boylan, K.B., Reilly, M.M., Greensmith, L., Singleton, A.B., Harms, M.B., Rossor, A.M., Houlden, H., 2013. A dominant mutation in FBXO38 causes distal spinal muscular atrophy with calf predominance. *Am. J. Hum. Genet.* 93, 976–983.
- Sureda-Gomez, M., Martín-Duran, J.M., Adell, T., 2016. Localization of planarian beta-CATENIN-1 reveals multiple roles during anterior-posterior regeneration and organogenesis. *Development* 143, 4149–4160.
- Takeishi, S., Nakayama, K.I., 2014. Role of Fbxw7 in the maintenance of normal stem cells and cancer-initiating cells. *Br. J. Cancer* 111, 1054–1059.
- Tian, X., Wu, C., 2013. The role of ubiquitin-mediated pathways in regulating synaptic development, axonal degeneration and regeneration: insights from fly and worm. *J. Physiol.* 591, 3133–3143.
- Varshavsky, A., 2005. Regulated protein degradation. *Trends Biochem. Sci.* 30, 283–286.
- Voigt, J., Papalopulu, N., 2006. A dominant-negative form of the E3 ubiquitin ligase Cullin-1 disrupts the correct allocation of cell fate in the neural crest lineage. *Development* 133, 559–568.
- Wagner, D.E., Wang, I.E., Reddien, P.W., 2011. Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* 332, 811–816.
- Werner, A., Manford, A.G., Rape, M., 2017. Ubiquitin-dependent regulation of stem cell biology. *Trends Cell Biol.*
- Wernersson, R., 2006. Virtual Ribosome—a comprehensive DNA translation tool with support for integration of sequence feature annotation. *Nucleic Acids Res.* 34, W385–W388.
- Willems, A.R., Schwab, M., Tyers, M., 2004. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim. Biophys. Acta* 1695, 133–170.
- Williamson, A., Werner, A., Rape, M., 2013. The Colossus of ubiquitylation: decrypting a cellular code. *Mol. Cell* 49, 591–600.
- Wurtzel, O., Cote, L.E., Poirier, A., Satiya, R., Regev, A., Reddien, P.W., 2015. A generic and cell-type-specific wound response precedes regeneration in planarians. *Dev. Cell* 35, 632–645.
- Xu, H., Wang, W., Li, C., Yu, H., Yang, A., Wang, B., Jin, Y., 2009. WWP2 promotes degradation of transcription factor OCT4 in human embryonic stem cells. *Cell Res.* 19, 561–573.
- Zayas, R.M., Hernández, A., Habermann, B., Wang, Y., Stary, J.M., 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.
- Zhao, X., Heng, J.I., Guardavaccaro, D., Jiang, R., Pagano, M., Guillemot, F., Iavarone, A., Lasorella, A., 2008. The HECT-domain ubiquitin ligase Huwe1 controls neural differentiation and proliferation by destabilizing the N-Myc oncoprotein. *Nat. Cell Biol.* 10, 643–653.
- Zhu, S., Perez, R., Pan, M., Lee, T., 2005. Requirement of Cul3 for axonal arborization and dendritic elaboration in *Drosophila* mushroom body neurons. *J. Neurosci.* 25, 4189–4197.
- Zhu, S.J., Pearson, B.J., 2016. Neoblast from the past: new insights into planarian stem cell lineages. *Curr. Opin. Genet. Dev.* 40, 74–80.

Fig. S1. Relationships of *Schmidtea mediterranea* Cullin proteins to those of other species were inferred using the Neighbor-Joining method. The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the Jones-Taylor-Thornton matrix-based method and are in the units of the number of amino acid substitutions per site. Sequences were selected from *Caenorhabditis elegans* (CAEEL), *Drosophila melanogaster* (DROME), *Homo sapiens* (HUMAN), *Mus musculus* (MOUSE), and *Schmidtea mediterranea* (SMED). The accession numbers for sequences included in the analysis are: *C. elegans* Cullin-1 (Q17389), Cullin-2 (Q17390), Cullin-3 (Q17391), Cullin-4 (Q17392), Cullin-5 (Q23639), Cullin-6 (Q21346); *D. melanogaster* Cullin-1 (Q24311), Cullin-2 (Q9V9R2), Cullin-3, isoform C (Q9V475), Cullin-3, isoform F (Q8IP45), Cullin-4 (Q5BI50), Cullin-5 (Q9VAQ0); *H. sapiens* Cullin-1 (Q13616), Cullin-2 (Q13617), Cullin-3 (Q13618), Cullin-4A (Q13619), Cullin-4B (Q13620), Cullin-5 (Q93034), Cullin-7 (Q14999), Cullin-9 (Q8IWT3); *M. musculus* Cullin-1 (Q9WTX6), Cullin-2 (Q9D4H8), Cullin-3 (Q9JLV5), Cullin-4A (Q3TCH7), Cullin-4B (A2A432), Cullin-5 (Q9D5V5), Cullin-7 (Q8VE73), Cullin-9 (Q80TT8).

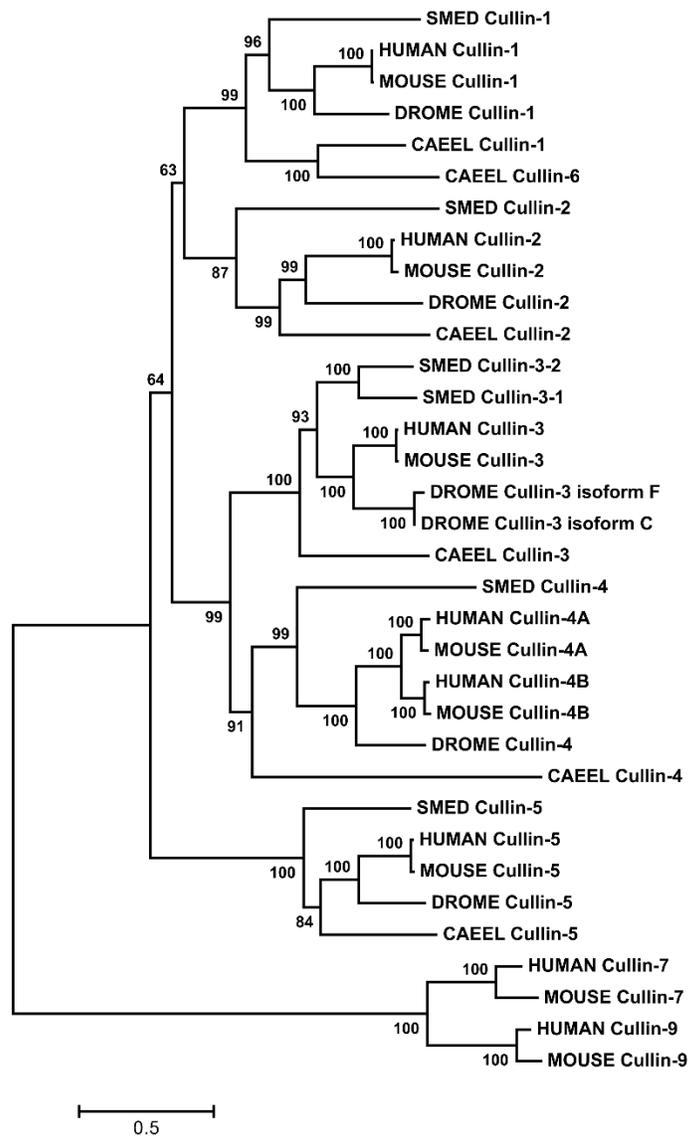
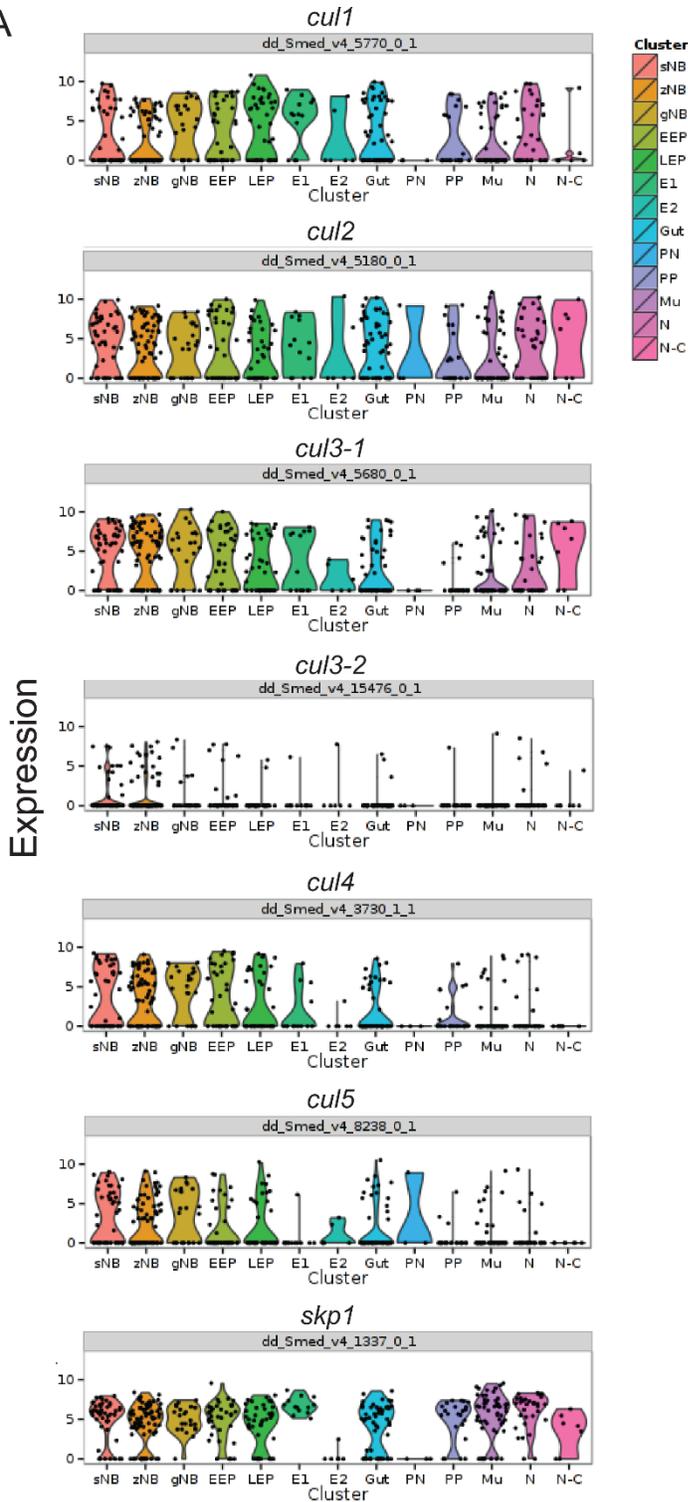
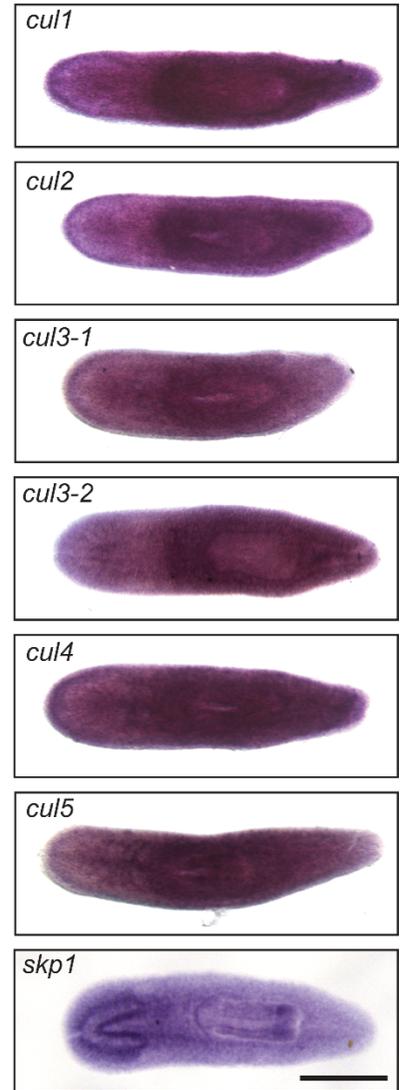


Fig. S2. Expression analysis of cullins and *skp1* in *Schmidtea mediterranea*. A) Single-cell expression profile of cullin genes and *skp1* in *S. mediterranea* from <https://radiant.wi.mit.edu/app/> (Wurtzel et al., 2015). B) Whole-mount in situ hybridization to cullin genes and *skp1*. Scale bar = 500 μm .

A



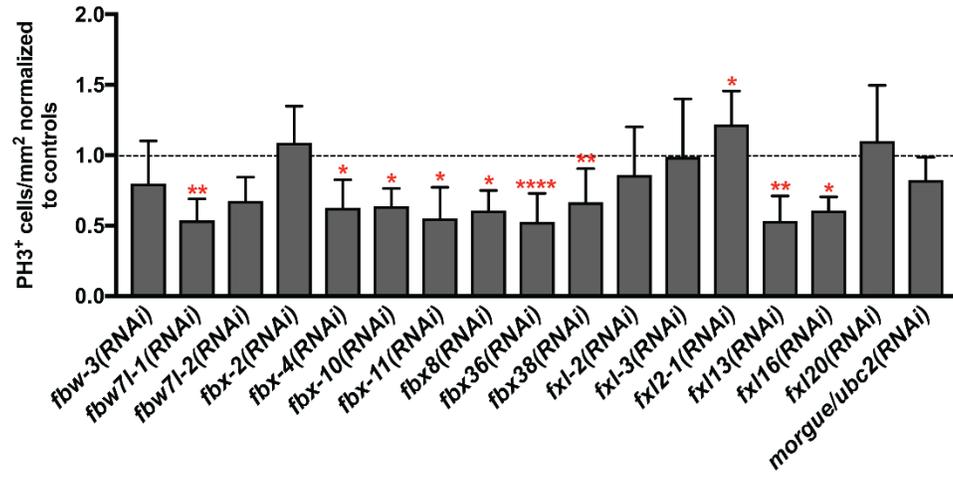
B



Smed-cullins expression in different cell types. **sNB**: sigma neoblast; **zNB**: zeta neoblast; **gNB**: gamma neoblast; **EEP**: early epidermal progenitors; **LEP**: late epidermal progenitors; **E1**: epidermis I; **E2**: epidermis 2; **Gut**: intestine; **PN**: protonephridia; **PP**: parapharyngeal; **Mu**: muscle; **N**: neural; **N-C**: neural-ciliated.

Fig. S3. Reported expression of f-box genes in *cul1*+ cells. A) Quantification of mitotic cells following f-box RNAi. Animals were fed dsRNA 6 times over 3 weeks against *gfp* or f-box genes indicated below each bar, amputated pre-pharyngeally, allowed to regenerate for 10 days, and then stained with anti-phospho-Histone H3 (PH3) to visualize mitotic neoblasts. Experiments were performed twice for each gene (N = 4–5 worms for each experiment). Graph shows the mean±s.d. of values normalized to *gfp* controls. *P<0.05, **P<0.01, or ***P<0.0001, Student's t-test. B) Co-expression of fbox genes in *cul1*+ cells obtained from <https://radiant.wi.mit.edu/app/> (Wurtzel et al., 2015). For each individual plot, *cul1* expression is on the y-axis; f-box gene expression is on the x-axis. f-box gene IDs are indicated above each plot. Legend abbreviations: sNB: sigma neoblast; zNB: zeta neoblast; gNB: gamma neoblast; EEP: early epidermal progenitors; LEP: late epidermal progenitors; E1: epidermis I; E2: epidermis 2; Gut: intestine; PN: protonephridia; PP: parapharyngeal; Mu: muscle; N: neural; N-C: neural-ciliated.

A



B

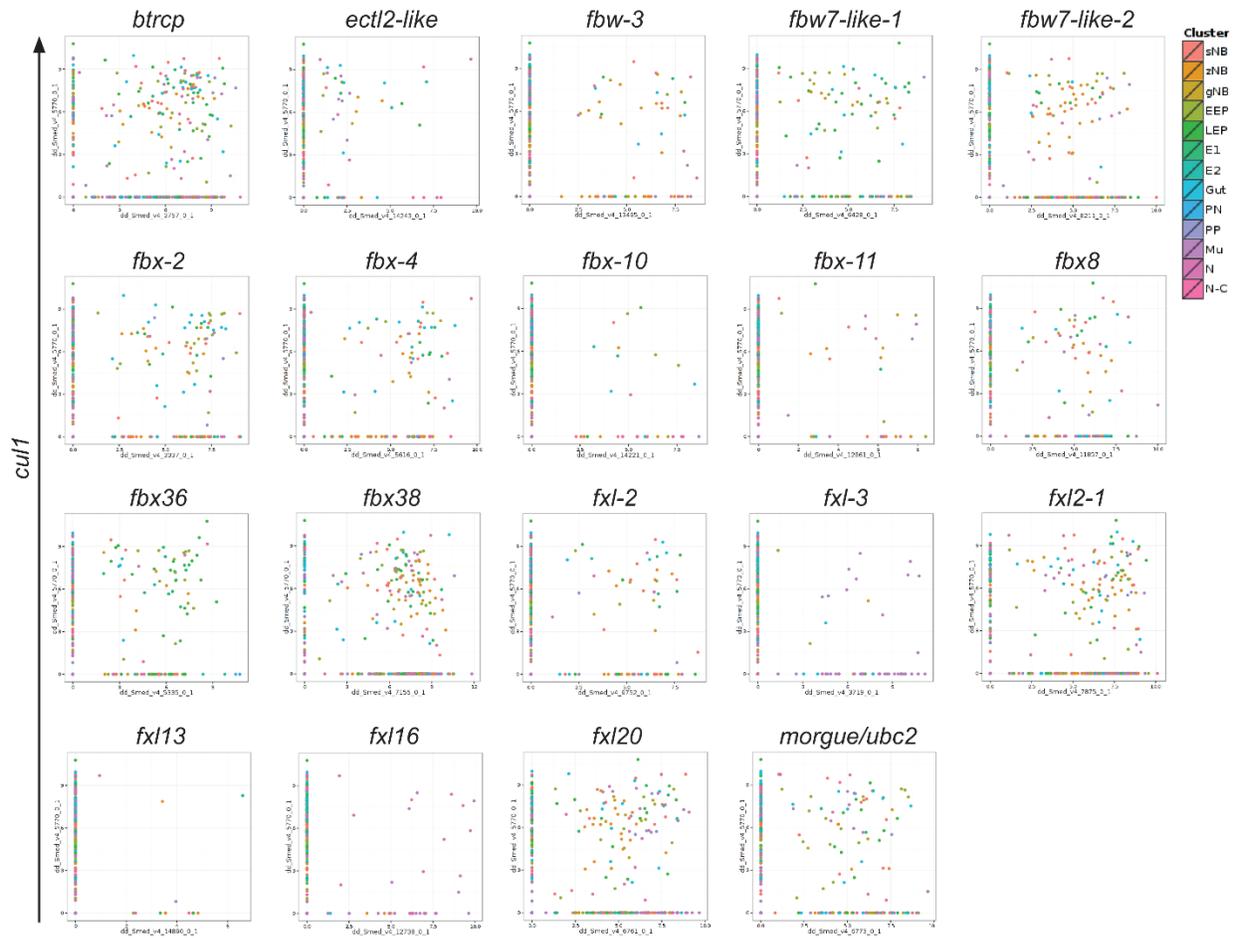


Table S1. Cullin genes present in the planarian *Schmidtea mediterranea*.

<i>S. mediterranea</i> Gene ID	dd_Smed_v6 ID	Top BLAST hit sequence ID	E-value
<i>Smed-cullin-1</i>	dd_Smed_v6_5770_0_1	CUL1_MOUSE name: Full=Cullin-1 Short=CUL-1	0
<i>Smed-cullin-2</i>	dd_Smed_v6_5180_0_1	CUL2_MOUSE name: Full=Cullin-2 Short=CUL-2	2.6e ⁻¹³⁰
<i>Smed-cullin-3-1</i>	dd_Smed_v6_5680_0_1	CUL3B_XENLA name: Full=Cullin-3-B Short=CUL-3-B	0
<i>Smed-cullin-3-2</i>	dd_Smed_v6_15476_0_1	CUL3_RAT name: Full=Cullin-3	0
<i>Smed-cullin-4</i>	dd_Smed_v6_3730_1_1	CUL4A_HUMAN name: Full=Cullin-4A Short=CUL-4A	1.16e ⁻¹⁶⁴
<i>Smed-cullin-5</i>	dd_Smed_v6_8238_0_1	CUL5_HUMAN name: Full=Cullin-5 Short=CUL-5 name: Full=Vasopressin- activated calcium-mobilizing receptor 1 Short=VACM-1	0

Table S2. Identification and analysis of F-box genes present in the planarian *Schmidtea mediterranea*.

<i>Schmidtea mediterranea</i> Gene ID	dd_Smed_v6 ID	Top BLAST hit sequence ID	E-value	Morphological phenotype	Nervous system patterning phenotype
<i>Smed-btrcp/fbw1a</i>	dd_Smed_v6_3757_0_1	FBW1A_MOUSE ame: Full=F-box WD repeat-containing 1A/Full=Beta-transducin repeat-containing protein	0	Lesions/lysis (13/13)	
<i>Smed-ect2-like</i>	dd_Smed_v6_14243_0_1	ECT2L_HUMAN ame: Full=Epithelial cell-transforming sequence 2 oncogene-like	1.25E-74	Absent PR (9/27), delayed PR regeneration (5/27)	
<i>Smed-fbw-1</i>	dd_Smed_v6_14791_0_1	FBW1B_HUMAN ame: Full=F-box WD repeat-containing 11	5.60E-16		
<i>Smed-fbw-3</i>	dd_Smed_v6_13495_0_1	FBXW7_MOUSE ame: Full=F-box WD repeat-containing 7	2.45E-25	Absent PR (3/43), asymmetric PR regeneration (7/43), immobile (10/43)	None Observed
<i>Smed-fbw10</i>	dd_Smed_v6_13271_0_1	FBW10_HUMAN ame: Full=F-box WD repeat-containing 10	1.55E-91	None Observed	
<i>Smed-fbw7-like-1</i>	dd_Smed_v6_6428_0_1	F-box WD repeat-containing 7-like	5.36E-123	Asymmetric PR regeneration (9/26), reduced blastema formation (3/26)	CG anterior commissure defect (7/14)
<i>Smed-fbw7-like-2</i>	dd_Smed_v6_8211_0_1	F-box WD repeat-containing 7-like isoform X2	1.70E-104	Delayed PR regeneration (18/30); delayed tail regeneration (5/18); ectopic PR (5/18)	CG anterior commissure defect (7/14), decreased neuropil density of CG (7/14)
<i>Smed-fbx-1</i>	dd_Smed_v6_11811_0_1	FBX40_MOUSE ame: Full=F-box only 40	5.43E-14	None Observed	
<i>Smed-fbx-10</i>	dd_Smed_v6_14221_0_1	FBX10_MOUSE ame: Full=F-box only 10	2.12E-29	Absent PR (7/31), asymmetric/delayed PR regeneration (18/31)	None Observed
<i>Smed-fbx-11</i>	dd_Smed_v6_12861_0_1	FXL21_MOUSE ame: Full=F-box LRR-repeat 21	4.30E-06	Asymmetric/delayed PR regeneration (27/42)	None Observed
<i>Smed-fbx-2</i>	dd_Smed_v6_3337_0_1	FBX4_MOUSE ame: Full=F-box only 4	2.97E-24	Absent PR (4/15), delayed PR regeneration (11/15)	None Observed

Table S3. Accession numbers and primers.

S. mediterranea gene	PlanMine dd_Smed_v6 ID	Acc. # (cDNA clones)	EST Clone ID	Primers used for restriction enzyme-based or ligation-independent cloning		Primers for nested PCR cloning	
				Forward Primer (XhoI)	Reverse Primer (Not I)	Outer Reverse Primer	Inner Reverse Primer (NotI)
Smed-cullin-1	dd_Smed_v6_5770_0_1	DN289356, DN313951	PL010001001A05, PL06017A1B11				
Smed-cullin-2	dd_Smed_v6_5180_0_1	XXXXXXXX					
Smed-cullin-3-1	dd_Smed_v6_5680_0_1	HO005153	PL08001B2G08				
Smed-cullin-3-2	dd_Smed_v6_15476_0_1	XXXXXXXX					
Smed-cullin-4	dd_Smed_v6_3730_1_1	DN309863, HO005151	PL06005B2E10, PL08001B2G08				
Smed-cullin-5	dd_Smed_v6_8238_0_1	XXXXXXXX				CAAGTCTCCATGGCCAAAC	ATAAGAAATGCGGCCGCGACTTTGAGAGACAACC
Smed-ekp1	dd_Smed_v6_1337_0_1	DN305227	PL05008A1A12				
Smed-htp2/bwfa	dd_Smed_v6_3757_0_1	DN312044	PL06011B2D09				
Smed-htp2	dd_Smed_v6_14243_0_1	XXXXXXXX		CCGCTCGAGAGTGCAGGCCAACTCTTTG	ATAAGAAATGCGGCCGCGAGATAGATAGTGCCTTTCCAG		
Smed-fbw-1	dd_Smed_v6_14791_0_1	XXXXXXXX		CATTACCATCCCGATTGAACAGACTTTGGA	CCAATTCTACCCGTCCCAATTTCTGCGCATA		
Smed-fbw10	dd_Smed_v6_13271_0_1	XXXXXXXX				AGAAACAGCAACCAACAAAC	ATAAGAAATGCGGCCGCGAAGCAACGGTCTCTATTTC
Smed-fbw3	dd_Smed_v6_13495_0_1	XXXXXXXX				TATCCCGCACTGCAATACG	ATAAGAAATGCGGCCGCGTGAACAGCACAAGACTCA
Smed-fbw7-1	dd_Smed_v6_6428_0_1	DN311839	PL06011A2B12				
Smed-fbw7-2	dd_Smed_v6_8211_0_1	XXXXXXXX				TCTGTAAAAACATCCCAATCC	ATAAGAAATGCGGCCGCGATCGTAATCCCTACTTC
Smed-fbx-1	dd_Smed_v6_11811_0_1	XXXXXXXX				AACCTGATGAAGCCAAACG	ATAAGAAATGCGGCCGCGAGCAATCGCTGGAAGTG
Smed-fbx-10	dd_Smed_v6_14221_0_1	XXXXXXXX				CCGCTCGAGAGGACGACGAGAAACAT	ATAAGAAATGCGGCCGCGCTGGAATGCAATGATTCG
Smed-fbx-11	dd_Smed_v6_12861_0_1	XXXXXXXX				AAGTGCAATTTGCTTTCG	ATAAGAAATGCGGCCGCGCAACAGGAGACCTTCAATC
Smed-fbx16	dd_Smed_v6_12714_0_1	DN313935	PL06011B2H11				
Smed-fbx-2	dd_Smed_v6_3337_0_1	DN315733	PL06021B2C12				
Smed-fbx-3	dd_Smed_v6_12894_0_1	XXXXXXXX				CCGAATACCAGTTCAGTTTATG	ATAAGAAATGCGGCCGCGCAATCAACAGCATAGGAGA
Smed-fbx32	dd_Smed_v6_11888_0_1	HO006413	PL08009B2B01				
Smed-fbx36	dd_Smed_v6_5335_0_1	DN310288	PL06020A1D06				
Smed-fbx38	dd_Smed_v6_7155_0_1	XXXXXXXX		CCGCTCGAGTGAAGCAGAAATCAAAAG	ATAAGAAATGCGGCCGCGCCGCAACCAATCATTACC		
Smed-fbx-4	dd_Smed_v6_5616_0_1	DN309031	PL06002X1F01				
Smed-fbx41	dd_Smed_v6_36984_0_1	XXXXXXXX		CATTACCATCCCGTGTCCAGCTCTGTGATAGCG	CCAATTCTACCCGGAACCAAGCTGCTTTTTCG		
Smed-fbx45	dd_Smed_v6_11757_0_1	XXXXXXXX		CATTACCATCCCGCTGGTCTTCAATCCGTGG	CCAATTCTACCCGTATACCAACTGCTGCCCGG		
Smed-fbx-5	dd_Smed_v6_13046_0_1	XXXXXXXX		CATTACCATCCCGAACCAACAGCTGTTTCSA	CCAATTCTACCCGCTCAGGATCCCAATTTGCG		
Smed-fbx-6	dd_Smed_v6_15599_0_1	XXXXXXXX				ACAGCATTTGAGCTTTGCG	ATAAGAAATGCGGCCGCGATTCAACAATGCTCTCAA
Smed-fbx-7	dd_Smed_v6_2677_0_1	DN308025	PL05015B1C05				
Smed-fbx8	dd_Smed_v6_11857_0_1	XXXXXXXX		CCGCTCGAGCCATTATTTGCTGCGATGTC	ATAAGAAATGCGGCCGCGAGAAAGCAAGTGACCCAAAC		
Smed-fbx-9	dd_Smed_v6_3570_0_1	DN307842	PL05015A1C11			ATGATATCTCGCGGCTCTTG	ATAAGAAATGCGGCCGCGAATACCGCAACCAAGAA
Smed-fxl-1	dd_Smed_v6_10134_0_1	XXXXXXXX		CCGCTCGAGATAGACGTGGATCAACTGG	ATAAGAAATGCGGCCGCGTATACGTTCCGGCATTGG		
Smed-fxl3	dd_Smed_v6_14890_0_1	XXXXXXXX				CGTAAACAATGAGGCCAAG	ATAAGAAATGCGGCCGCGTTTGAATGCTGAGATTTC
Smed-fxl16	dd_Smed_v6_12738_0_1	HO007899	PL08005A1E04				
Smed-fxl-2	dd_Smed_v6_9752_0_1	XXXXXXXX					
Smed-fxl20	dd_Smed_v6_8761_0_1	DN304091	PL05004B2F01			ACCAATTTATGTTACGGGAAAG	ATAAGAAATGCGGCCGCGAATCGGGCTGACCAAG
Smed-fxl-2-1	dd_Smed_v6_7875_0_1	DN314474	PL06018A2G06			ATTTGACAAACCCCACTTG	ATAAGAAATGCGGCCGCGCACTTGCAATTCACCAAGC
Smed-fxl-3	dd_Smed_v6_3719_1_1	DN292526	PL030011A10D02				
Smed-fxl-4-1	dd_Smed_v6_9358_0_1	XXXXXXXX				GCAGCACAACTAATTAACAGC	ATAAGAAATGCGGCCGCGAATTCACATTTTGCACACC
Smed-fxl7-1	dd_Smed_v6_10647_0_1	XXXXXXXX				GCTCCGAGATTTCCTCAACC	ATAAGAAATGCGGCCGCGGATCTGCTGTTGATTG
Smed-fxl7-2	dd_Smed_v6_15551_0_1	XXXXXXXX		CCGCTCGAGCTTCGTTTACCAACAGC	ATAAGAAATGCGGCCGCGACAATCTCCGAAAGATGCG		
Smed-morgue/ubc	dd_Smed_v6_6773_0_1	DN300983	PL04021A1D08				

Chapter 1, in full, is a reprint of the material as it appears in *Developmental Biology* 2018. Strand, NS.; Allen, JM.; Ghulam, M.; Taylor, MR.; Munday, RK.; Carrillo, M.; Movsesyan, A.; Zayas, RM. The dissertation author was a primary investigator and author of this manuscript.

CHAPTER 2:

Genetic screen of RING and U-box E3 ligases in planarians identifies critical spliceosomal and epigenetic regulators of stem cell differentiation and specification.

John M. Allen, Madison Balagtas, Elizabeth Barajas, Carolina Cano, Ionit Iberkleid, Ricardo M. Zayas

Abstract

Planarians are capable of regenerating from nearly any injury in a dynamic process that must integrate wound response and patterning signals to direct proper specification and functional integration of missing tissues. Ubiquitylation is a small protein that post-transcriptionally modifies other proteins and regulates many cellular pathways including protein degradation, mRNA splicing, and transcription. The specificity of ubiquitin signaling is conferred by the action of the E3 ligases, a large protein family whose roles in regeneration remain largely unexplored. Here we screen RING and U-box E3 ligases for function in planarians using RNAi and uncovered roles for nine genes and further explored the phenotypes of spliceosomal factor *prpf19* and epigenetic regulator *rnf2*. We examined other cofactors of *rnf2* in PRC1 and observed a striking phenotype of regional tissue misspecification. To uncover the transcriptional targets of PRC1 we performed RNA-seq and found that *rnf2* and *phc* were largely regulating separate genes despite being predicted to function in the same complex. We found using *in situ* hybridization that *rnf2* regulated levels of expression within a tissue type while *phc* was necessary for the spatial restriction of genes to the proper cell types. Collectively, this work reveals E3 ligases that regulate stem cells and regeneration, uncovers roles for RNA splicing in progenitor specification, and finds differential gene targets for PRC1 factors in invertebrates.

Introduction

A deep understanding of the networks and signaling pathways that direct the maintenance and differentiation of adult stem cells is essential for the development of regenerative therapies. The freshwater planarian, *Schmidtea mediterranea*, is an important model for the study of the molecular mechanisms that underpin stem cell-based regeneration. These worms maintain a large population of adult stem cells, a subset of which have been demonstrated to be pluripotent^{1,2}. This population of stem cells is utilized during homeostasis to continually turn-over and renew planarian tissues and is also mobilized in response to injury to contribute to regenerating tissues. As such, they offer an amenable model to study stem cell biology in a whole-organism *in vivo* context.

Extensive work has been performed to understand the molecular basis of planarian regeneration³, yet most studies have primarily examined transcriptional changes^{4,5}. Comparatively fewer studies have focused on proteomic regulation in planarian stem cells^{6,7} or the post-translational regulation of proteins important for stem cell function⁸. An essential post-translation regulator of proteins is the addition of a small, highly conserved polypeptide called ubiquitin which modifies protein function in a variety of cellular contexts, including transcription, cell cycle regulation, translational fidelity, protein turnover, and degradation⁹⁻¹².

Ubiquitin-dependent signaling events have emerged as essential regulators of stem cell functions, including self-renewal and differentiation¹³. The transfer of free ubiquitin onto a target substrate typically occurs through a tripartite enzymatic cascade that terminates with the E3 ubiquitin ligases. The E3 ligases can be grouped into two major

classes, the HECT (Homologous to the E6-AP Carboxyl Terminus) and more prevalent RING (Really Interesting New Gene) class. Of the approximately 617 genes encoding putative E3 ligases identified in the human genome, 309 were predicted to contain a RING finger (RNF) or the related U-box domain; a further 270 E3 genes are RNF-dependent through complex associations¹⁴. The RNFs are defined by a zinc-finger domain that has an evolutionarily conserved arrangement of cysteine and histidine residues that coordinate two zinc ions and bind an E2-ubiquitin conjugate¹⁵. The U-box domain forms a similar structure to the RING domain and can bind conjugated E2 but does not coordinate zinc¹⁶. Substrate recognition and binding is achieved by additional domains within the RNF protein or by association with other proteins as part of a multi-protein complex. Previous work on E3 ligase function in planarians has implicated HECT E3 and Cullin-Ring complex member ligases as essential regulators of regeneration and stem cells^{8,17}.

Here we performed a functional analysis on a subset of the large RING and U-box domain-containing gene family that are expressed in the planarian stem cells or progeny. We found several to be essential for homeostatic maintenance, regeneration, and tissue patterning. These genes included spliceosomal factor *prpf19* and epigenetic factors *rnf2* and *bre1*, which are known to ubiquitylate of histones. We found that *prpf19* was required for worm survival but not required for stem cell maintenance, suggesting a role in promoting cell differentiation. Polycomb Repressive Complex 1 (PRC1) component gene *rnf2* was required for global monoubiquitylation of Histone H2A (H2Aub1) and promoting proper regeneration. When we inhibited other homologous PRC1 genes, *phc* and *cbx*, we did not observe a global reduction in H2Aub1 levels but did observe consistent, specific, functional defects in the organization of tissue near the base of the planarian pharynx. In

summary, this functional screen of E3 ligases identified essential regulators of stem cell biology and regeneration including spliceosomal and epigenetic factors, and led to the discovery of uncovered differential phenotypes and transcriptional targets for PRC1 factors.

Methods

Planarian care

A clonal line of asexual *S. mediterranea* (CIW4) was used in all experiments and kept in 1X Montjuïc salts (1.6 mM NaCl, 1.0 mM CaCl₂, 1.0 mM MgSO₄, 0.1 mM MgCl₂, 0.1 mM KCl, 1.2 mM NaHCO₃, pH 7.0) in food-grade plastic containers at 20°C¹⁸. Animals selected for experiments were 3-6 mm in length and starved for one week prior to experimentation.

Gene identification

To find RING and U-box domain containing genes in *S. mediterranea* we filtered the Dresden transcriptome¹⁹ using InterPro Domain IDs, IPR001841 (Zinc finger, RING-type) and IPR003613 (U box domain). This list was filtered to include only the longest gene contig for each hit and was used as query sequences for a BLAST search to a curated list of human RING and U-box genes¹⁴ at an expected value cut-off of 10⁻³. We additionally filtered the Dresden transcriptome for contigs annotated with IPR013083 (Zinc finger, RING/FYVE/PHD-type). This list was filtered to remove duplicate entries and a BLAST search was performed against our list of human RING and U-box genes and as the IPR013083 family contains non-RING and U-box genes, only genes that had predicted homology to a human gene at an expected cut-off of 10⁻³ were appended to our initial list.

RNA interference

During initial screening animals were fed double stranded RNA (dsRNA) mixed with a ≈3:1 mixture of liver:water paste twice per week for eight feeds and were amputated pre-pharyngeally on day 28 of treatment to observe regeneration. dsRNA used during the initial

screening was synthesized using an *in vitro* transcription reaction or expressed bacterially; secondary screenings and RNAi for sample collection was performed using bacterially expressed dsRNA. *In vitro* dsRNA was synthesized as previously described²⁰ and the entire reaction mixture was separated into eight aliquots, mixed with liver paste, and stored until feeding. Bacterially expressed dsRNA was prepared by growing cultures of *E. coli* strain HT115 transformed with the pPR-T4P plasmid²¹ containing the gene of interest and inducing dsRNA expression using IPTG. Bacteria pellets were purified using centrifugation and mixed with liver paste for administration to animals.

TUNEL staining

Terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate nick end labeling (TUNEL) was performed to assay levels of apoptotic cells. Animals were incubated in 5% *n*-acetyl cysteine (diluted in PBS) for 5 minutes and fixed in 4% formaldehyde (diluted in PBS-Tx [PBS with 0.3% Triton X-100]) for 15 minutes. Samples were then permeabilized in 1% SDS (diluted in PBS) and bleached overnight in 6% H₂O₂ (diluted in PBS-Tx). Samples were then rinsed and stained using the Apoptag Kit (Millipore-Sigma) as previously described²².

In situ hybridization

Antisense probes for *in situ* hybridization were synthesized as previously described²³ from DNA templates amplified from pBS II SK(+) (Stratagene) or pPR-T4P²¹ plasmid vectors incorporating either digoxigenin or FITC labeled UTPs. Animals for whole-mount *in situ* hybridization were processed and hybridized as outlined previously²⁴. Briefly samples were sacrificed in 5% N-Acetyl Cysteine, fixed in 4% formaldehyde, and

bleached in formamide/hydrogen peroxide bleaching solution. Samples were pre-hybridized for two hours and then hybridized with probe overnight at 56°C. Samples were incubated with an appropriate antibody, depending on the probe-label and subsequent development strategy. For chromogenic development samples were incubated with an anti-digoxigenin-AP antibody (Roche, 1:2000) and developed with NBT/BCIP in AP buffer. Fluorescent development was performed using Fast Blue salts or Tyramide Signal Amplification (TSA) after incubation with anti-digoxigenin-AP or anti-FITC-POD (Roche, 1:300) antibodies respectively following previously described protocols²⁵.

Immunohistochemistry

Animals were incubated in ice cold 2% hydrochloric acid for 5 minutes and fixed for 2 hours in Carnoy's solution at 4°C. Samples were washed in methanol for 1 hour at 4°C and bleached overnight in 6% H₂O₂ diluted in methanol at room temperature. Animals were washed out of methanol and into PBS-Tx and blocked in 1% Bovine serum albumin (BSA) diluted in PBS-Tx for 4 hours at room temperature. Samples were incubated with anti-phosphohistone H3 (Ser 10) (Cell Signaling, 1:1000) diluted in 1% BSA/PBS-Tx overnight at 4°C. Antibody was developed using TSA as previously described²⁴.

Protein extraction and western blotting

Protein samples from RNAi worms were homogenized in Trizol and extracted from the organic phase following the manufacturer-provided protocol with a modified solubilization buffer (4M Urea, 0.5% SDS) and added sonication step of 10 one second pulses that were performed to increase protein recovery²⁶. Samples were loaded onto AnyKD TGX gels (BioRad), transferred using the semidry method to a 0.45m PVDF

membrane, and blocked in 5% nonfat milk/TBS-Tw (Tris-buffered saline with 0.1% Tween-20). Antibodies to monoubiquityl-Histone H2A (Cell Signaling 8240) and monoubiquityl-Histone H2B (Cell Signaling 5546) were diluted in 5% Bovine serum albumin in TBS-Tw at 1:2000 and 1:1000 respectively and incubated overnight at 4°C. Washes were performed with TBS-Tw and anti-rabbit-HRP (Cell Signaling 7074) was diluted in 5% nonfat milk/TBS-Tw at 1:2500 and incubated 1 hour at room temperature. Signal was developed using BioRad Clarity Western ECL Substrate (BioRad 1705061).

RNA sequencing

Worms from three independent control and experimental RNAi groups per time point were homogenized in Trizol and RNA was extracted and purified following manufacturer protocol. RNA was treated with the Turbo DNA-free kit and column purified using the RNeasy MinElute Cleanup kit. Three independent biological groups were collected at each time point assayed for both control and experimental (*rnf2* or *phc*) RNAi treatments. Samples were sequenced on an Illumina HiSeq to a read depth of at least 15 million 150bp paired-end reads. Reads were pseudo aligned to the Dresden transcriptome using Kallisto²⁷ and differential gene expression analysis was performed using the R Bioconductor package²⁸ and DESeq2²⁹ with an FDR cutoff value of ≤ 0.1 applied. To perform Gene Ontology (GO) analysis differentially expressed transcripts from the d28 *rnf2(RNAi)* data set were compared to the human proteome using BLASTX (cutoff e-value $< 1e^{-3}$). Human UniProt IDs were used as input for annotation and overrepresentation analysis (<http://geneontology.org/>) using Fisher's Exact test with an FDR multiple comparisons correction cutoff of ≤ 0.05 applied.

Results

Identification of RING and U-Box E3 ubiquitin ligase genes in *S. mediterranea*

The RING and U-Box protein domains have been identified as having a key role in mediating the ubiquitylation of a target substrate. To identify genes in *S. mediterranea* that are predicted to encode a RING or U-box domain we filtered a reference planarian transcriptome¹⁹ using InterPro domain annotations and generated a list of 393 transcripts. We took this list of putative RING and U-box domain-containing gene transcripts and performed BLAST analysis using a curated list of human E3 ubiquitin ligases¹⁴ and found 376 planarian genes that were predicted to have homology with a human RING or U-box gene (Supplemental Table S2.1) and 17 planarian transcripts that did not have predicted significant homology to a human RING or U-box gene.

A functional screen reveals genes with roles in planarian stem cell regulation and regeneration

To identify RING and U-box E3 ubiquitin ligases that regulate planarian stem cells and regeneration we selected a subset of genes on our list based on their enriched expression in FACS-isolated planarian stem cells or stem cell progeny using a previously generated transcriptomic data set⁴. We then performed RNA interference (RNAi) to perturb the function of 93 of these genes. RNAi treatments were performed over four weeks and the worms were amputated at day 28 of treatment to assess effects on regeneration (Figure 2.1A); we found that RNAi of nine genes produced phenotypes related to stem cell function in homeostasis and during regeneration (Table 2.1). Phenotypes

observed during homeostasis included head regression, epidermal lesions, ventral curling, and lysis (Figure 2.1B); other genes displayed abnormalities and delays during regeneration when inhibited (Figure 2.1C). Head regression was observed during homeostasis after depletion of *prpf19*, *march5*, *traf2-like*, *not4*, and *bre1*; lesions were observed during homeostasis after depletion of *march5*, *ran*, and *rnf8-like* inhibition; ventral curling was observed during homeostasis after depletion of *prpf19* and *not4*. The genes *prpf19*, *march5*, and *ran* were essential for worm survival and depletion of these genes caused worm lysis. Knockdown of *rnf8-like*, *bre1*, *rnf2*, and *ring1* caused defective regeneration which was typically manifested as a delayed appearance of visible eyespots when compared to *control(RNAi)* treatments. We chose to further characterize a subset of these phenotypes based on their strong expression in stem cells or predicted roles as epigenetic regulators of stem cells during developmental processes.

Spliceosomal factor *prpf19* is required for worm survival and stem cell function

We chose *prpf19* for further characterization based on its enriched expression in planarian stem cells and identification during our initial screen as being necessary for worm survival. Other aspects of the RNAi phenotype, including head regression and ventral curling, are typically associated with a depletion or loss of stem cells in planarians. These phenotypes are consistent with an earlier report for *prpf19* as being upregulated during and necessary for head regeneration in planarians³⁰. *prpf19* is known to function as a member of NineTeen Complex (NTC), with a well described role in regulating mRNA splicing that is conserved in *Saccharomyces cerevisiae*. Consistent with a role in an essential

cellular process we found broad expression of this gene using whole-mount mRNA *in situ* hybridization (WISH) (Supplemental Figure S2.1A). We confirmed that at least a subset of this expression is in the neoblasts by performing double fluorescent *in situ* hybridization (FISH) for *prpf19* and neoblast marker *piwi-1* and observed overlapping expression for these two genes (Supplemental Figure S2.2B). As *prpf19* is known to function as an E3 ligase we assayed by western blotting the effect of *prpf19* inhibition on ubiquitylated proteins and found no obvious difference with control samples (Supplemental Figure S2.2C), suggesting that *prpf19(RNAi)* does not disrupt general proteasome function but rather or has only a minor effect on global ubiquitylation or is targeting specific proteins that are not resolvable on a total ubiquitin blot.

To investigate if the *prpf19(RNAi)* phenotypes observed were the result of stem cell depletion, we performed WISH to planarian stem cell marker genes, *tgs-1*, *piwi-1*, and *h2b*, on *prpf19(RNAi)* and *control(RNAi)*-treated worms and we found, surprisingly, that all marker genes analyzed demonstrated robust expression, even in worms where the phenotype had significantly progressed (Figure 2.2A, Supplemental Figure S2.2B). Because *prpf19* was found to be expressed in additional cell types besides stem cells (Supplemental Figure S2.2B), we examined the effect of *prpf19* inhibition on epidermal progeny by performing WISH analysis using *prog-1* and *agat-1* as markers for early and late epidermal progeny, respectively, and found that, consistent with the epidermal lesions observed during the progression of the *prpf19* phenotype, staining for epidermal lineage markers is reduced in *prpf19(RNAi)* worms. Taken together, these results suggest that *prpf19* function is not required for the maintenance and survival of planarian stem cells but may affect the

differentiation of stem cells into progenitors or the maintenance of certain progenitor populations.

Inhibition of *prpf19* causes defects in stem cell proliferation and increases apoptosis

Despite being dispensable for stem cell maintenance, the strong expression of *prpf19* in stem cells and robust phenotypes that resulted from *prpf19* inhibition suggested a role for *prpf19* in regulating stem cell dynamics. To examine the effect of *prpf19(RNAi)* on cell proliferation we stained *control(RNAi)* and *prpf19(RNAi)* treated worms with anti-phospho-Histone H3 (pH3) to mark mitotic cells across several time points of treatment. We found that at late time points (day 18) when the phenotype is beginning to manifest, there was a significant decrease in the number of mitotic cells, normalized for worm size, in *prpf19(RNAi)* worms (Figure 2.2B and 2.2C). This decrease in the number of mitotic cells was not correlated with a decrease in expression of stem cell marker genes (Figure 2A, Supplemental Figure S2.2A), which further suggests that *prpf19(RNAi)* treatment is causing existent stem cells to not differentiate or differentiate at a lower rate to reflect this reduction in mitotic rates without depletion of the stem cell pool.

To better understand the severe phenotypes observed in *prpf19(RNAi)* worms, including epidermal lesioning and worm lysis, we assayed the worms for apoptotic cells. Not surprisingly, we found an increase in apoptotic cells in *prpf19(RNAi)*-treated worms compared to control worms at the time point prior to observing phenotypes and a marked increase was observed as the *prpf19(RNAi)* phenotype progressed (Figure 2.2D and 2.2E). Together with the observed loss of epidermal progenitor markers (Figure 2.2A), these data suggest that the phenotypes observed after *prpf19* depletion are not caused by a loss of

stem cells but rather by dysfunction in the proper homeostatic replacement of differentiated tissues. As the stem cells are unable to replenish tissues the worm begins to experience an increase in apoptosis and a decrease in proliferation as epidermal integrity becomes compromised. There could also be a role for *prpf19* as an antiapoptotic factor in differentiated tissues resulting in increased apoptosis in *prpf19(RNAi)* worms. Roles for *prpf19* in regulating differentiation and in regulating apoptosis are not exclusive and could both be contributing to the phenotype.

NTC components and targets are necessary for regeneration and homeostasis

NTC is a large protein complex that has a variety of cellular roles but has its best-described role in regulating pre-mRNA splicing. Named after its founding member, *prpf19*, the complex is conserved between human and yeast. In the spliceosome NTC, through its PRPF19 subunit, functions as an E3 ligase (Figure 2.3A). To examine if the observed effects of *prpf19* inhibition were being mediated through disruption of spliceosomal assembly and function we used RNAi knock down three homologs of core NTC component members, *cdc5l*, *prlg1*, and *spf27*, and found that these genes were also necessary for worm survival and regeneration (Figure 2.3B and Figure 2.3D). *cdc5l* and *prlg1* are essential for NTC function in yeast and presented very severe phenotypes of head regression, ventral curling and lysis that mirrored the observations for *prpf19(RNAi)*. *spf27(RNAi)* presented a milder phenotype than other NTC genes that were examined and showed delayed or absent regeneration in 28/37 head fragments and 33/37 trunk fragments with an additional 3 trunk fragments having more severe phenotypes of ventral curling or lysis. Likewise, we

reasoned that if the *prpf19* phenotype was the result of its ubiquityl ligase activity in NTC and the subsequent stabilization of *prpf3* interaction with *prpf8*, then disruption of this interaction by inhibiting either partner should have a similar phenotype to *prpf19(RNAi)*; indeed, we found that *prpf3(RNAi)* and *prpf8(RNAi)* worms exhibited severe phenotypes like *prpf19(RNAi)* and included head regression, ventral curling, epidermal lesions, and lysis (Figure 2.3C). WISH analysis of NTC genes, *prpf3*, and *prpf8*, both demonstrated broad parenchymal expression patterns similar to *prpf19*, with *prpf8* demonstrating a noticeable stem cell expression pattern (Supplemental Figure S2.3A). The similar phenotypes and expression patterns observed for NTC genes and downstream factors to *prpf19* suggests that the phenotype for *prpf19(RNAi)* is mediated through its role in NTC and that the NTC and spliceosome function are critical for stem cell regulation during homeostasis and regeneration.

Histone-modifying ubiquitin ligases are essential for regeneration and homeostasis

Histone proteins package DNA to form the nucleoprotein structure known as chromatin, the organization of which affects the transcriptional state of genes. Chromatin organization can be modulated by the post-translational modification of histones including ubiquitylation of histones H2A and H2B. Ubiquitylation of histone H2B is associated with transcriptional activation and is mediated by the E3 ligase complex RNF20/40 (Bre1 in yeast). We found that planarians have a single homolog for this complex designated *Smed-bre1* and that knockdown of *bre1* caused the worms to exhibit head regression and lesions prior to day 28 of treatment in 33/53 worm assayed (Figure 2.1B). When amputated most

bre1(RNAi) worms failed to regenerate and many lysed with 31/53 head fragments and 21/53 trunk fragments lysing by the end of the observation period (day 14 post amputation). To investigate if *bre1(RNAi)* was affecting global levels of monoubiquityl-histone H2B (H2Bu1) we performed western blotting analysis using a H2Bub1-specific antibody and found reduced levels of H2Bu1 in whole worm homogenates as soon as 14 days after beginning RNAi treatment (Supplemental Figure S2.4A).

In contrast to histone H2B ubiquitylation, monoubiquitylation of histone H2A is associated with transcriptional repression and occurs in a variety of cellular contexts, including developmental processes, stem cell regulation, and the DNA damage response. Histone H2A is targeted for ubiquitylation by RING1 and RNF2, which act as RING E3 ligases within Polycomb Repressive Complex 1 (PRC1). PRC1 is a repressive epigenetic complex that is active during development and acts to compact chromatin, stably silence genes, and monoubiquitylate histone H2A (Figure 2.4A). We identified two candidate homologs of RING1 and RNF2 and found that depletion of each gene caused delayed or absent regeneration when compared to controls. These phenotypes were most evident in the trunk fragments where 37/58 *rnf2(RNAi)* and 19/29 *ring1(RNAi)* worms exhibited a phenotype of delayed regeneration (measured by the appearance of dark eyespots) compared to 7/54 and 2/30 *control(RNAi)* worms assayed at the same regeneration time point (7 days post amputation). Of the 37/58 *rnf2(RNAi)* trunks and 19/29 *ring1(RNAi)* trunks with regeneration defects, 13/37 and 4/19 failed to form regeneration blastemas, respectively, whereas all *control(RNAi)* worms formed normal-sized blastemas (Figure 2.1C). No obvious phenotypes were observed during homeostatic maintenance even during long-term RNAi treatment (> 16 feeds over 8 weeks). In other organisms *rnf2* is

responsible for the bulk of H2Aub1 deposition and we confirmed that *rnf2(RNAi)* reduced bulk levels of H2Aub1 by western blot analysis on homogenates using a H2Aub1 specific antibody (Figure 2.4B). In contrast, we found that *ring1(RNAi)* did not have an appreciable effect on global H2Aub1 levels (Supplemental Figure S2.4B), which is consistent with *rnf2* being the E3 ligase subunit within PRC1 that is responsible for most of the catalytic activity on H2A³¹.

Inhibition of canonical PRC1 subunit *phc* affects the patterning of the planarian pharyngeal domain

To better understand the role of this major developmental complex in planarian biology and regeneration we examined the functional requirements for other subunits of PRC1. In vertebrates the composition of PRC1 is variable and the complex is defined by which of the six mammalian paralogs of PCGF is present. PCGF2 and PCGF4 define the canonical mammalian PRC1 complex which also includes one each of several chromobox (CBX) and Polyhomeotic paralogs (PHC) (Figure 2.4A). We identified planarian homologs for these other core PRC1 genes and found one homolog each for *cbx* and *phc*, and two homologs for *pcgf*. To investigate if the phenotypes for *rnf2(RNAi)* and *ring1(RNAi)* were mediated through their function in canonical PRC1 we used RNAi to deplete *cbx*, *phc*, *pcgf2*, and *pcgf3*. In contrast to the impaired regeneration observed after *rnf2* or *ring1* knockdown, RNAi for *phc* or *cbx* demonstrated a complex homeostasis phenotype that included the abnormal appearance of a dorsal lesion anterior to the pharynx (Figure 2.4C). In some cases, we observed the pharynx protruding from the lesioned region and

extending ectopically from the dorsal surface of the worm. As the phenotypes progressed, these RNAi worms began to exhibit defects along the body axis, showing crimped tails that were unable to affix to the dish and epidermal lesions. We also assayed the effect of inhibition of the cPRC1 genes on H2Aub1 levels and found that inhibition of *phc* or *cbx* did not have a noticeable effect on bulk H2Aub1 levels (Supplemental Figure S2.4B), which suggests that cPRC1 is not responsible for the majority of H2Aub1 deposition, similar to findings in vertebrate models³². Both *phc* and *cbx* had similar mRNA expression patterns as assayed by WISH suggesting that they can function in the same complex (Supplemental Figure S2.4C). This expression pattern overlapped with the diffuse parenchymal expression pattern for *rnf2* and *ring1* (Supplemental Figure S2.1A) but had stronger specific expression near the planarian brain and intestinal branches, the latter of which are areas known to be enriched in stem cells.

While similar, the penetrance of the *phc(RNAi)* phenotype was higher than for *cbx(RNAi)* and we chose to further examine the *phc(RNAi)* phenotype using known markers of tissue patterning. The appearance of a dorsal lesion and mis-localization of the pharynx to the dorsal surface in *phc(RNAi)* animals suggests that planarian PRC1 may be involved in the specification of specific tissues related to the pharynx or in regulating genes that provide axial positioning cues to stem cell progeny during homeostatic tissue turnover. We examined dorsal-ventral patterning factor *bmp-4* and anterior-posterior factor *ndl-3* expression after *phc(RNAi)* and found that there was no change in the expression pattern of these factors relative to controls (Supplemental Figure S2.4D). We then further examined genes that mark specific tissues related to the pharynx, including the pharynx marker, *laminin*, and the gene *NB.22.1E*, which labels marginal adhesive gland cells, the ventral

mouth opening, and a population of cells near the base of the pharynx. Following *phc* inhibition we observed that *laminin* expression was reduced to a single condensed spot of expression near where the dorsal lesion was observed and a few scattered cells near the midline of the animal (Figure 2.4D). Likewise, we observed the specific disappearance of the *NB.22.1E⁺* population of cells at the base of the pharynx following *phc(RNAi)*, while expression along the body margin and ventral mouth opening was unaffected (Figure 2.4D). This data establishes a role for PRC1 factors in maintaining tissue identity in a non-embryological context.

RNA-seq analysis of gene expression after PRC1 subunit inhibition reveals transcriptional targets of PRC1

To gain insights into which genes are dysregulated after PRC1 inhibition and to understand the transcriptional basis for the observed phenotypes for *rnf2* and *phc* we performed RNA-seq on *rnf2(RNAi)* or *phc(RNAi)* worms to determine differentially expressed genes. We chose timepoints for RNA extraction based on the phenotypic progression, qPCR analysis to confirm a robust reduction in target RNAi transcript levels, and, for *rnf2(RNAi)*, western blot analysis to ensure the RNAi treatment was reducing levels of H2Aub1. Based on these parameters we chose to extract RNA after 11 days of *phc(RNAi)* treatment and 14 and 28 days after *rnf2(RNAi)* (Supplemental Figure S2.5A).

Between both *rnf2(RNAi)* time points, we identified a combined 264 differentially regulated genes with 126 downregulated and 138 upregulated genes (Figure 2.5A, Supplemental Figure S2.5B). Not surprisingly, a longer RNAi treatment period for *rnf2*

resulted in an increase in the number of differentially expressed genes with 247 differentially expressed after 28 days of treatment compared to only 29 differentially expressed genes after 14 days of treatment (Table 2.2). There was substantial overlap between the *rnf2(RNAi)* data sets with 12 of 29 genes in the day 14 data set represented in the larger day 28 data set (Supplemental Figure S2.5C). After 11 days of *phc(RNAi)* treatment we found that 49 genes that were differentially expressed; 20 were downregulated and 29 upregulated (Figure 2.5C). Consistent with a repressive role in transcriptional regulation more genes were found to be significantly upregulated when either *phc* or *rnf2* was inhibited. Importantly, *rnf2* and *phc*, were each found to be significantly downregulated when targeted for RNAi.

Surprisingly, despite being predicted to function in a shared complex, only a single gene was found to be in common between the *phc(RNAi)* and *rnf2(RNAi)* data sets. This lack of overlap between the data sets suggests that *phc* and *rnf2* regulate different processes and pathways *in vivo* and this difference explains the disparate phenotypes that were observed after RNAi treatment. One common upregulated gene between the data sets was *cbx*, which is itself a chromatin binding element within cPRC1, and was moreover the most significantly upregulated gene in the *phc* knockdown data set.

To analyze the RNA-seq data set further we performed gene ontology (GO) analysis on both upregulated and downregulated genes from the d28 *rnf2(RNAi)* data set. Genes downregulated during *rnf2(RNAi)* were found to be significantly enriched for GO biological process terms related metabolic and catabolic processes (Supplemental Figure S2.5D). Among the GO terms significantly enriched in genes upregulated following *rnf2* depletion

were several related to cellular stress (Supplemental Figure S2.5E), especially low oxygen conditions, including, “response to hypoxia” (GO:0001666), “cellular response to decreased oxygen levels” (GO:0036294), “ATF6-mediated unfolded protein response” (GO:0036500), “regulation of transcription from RNA polymerase II promoter in response to stress” (GO:0043618), “chaperone cofactor-dependent protein refolding” (GO:0051085), “protein folding in endoplasmic reticulum” (GO:0034975), and “protein refolding” (GO:0042026). These GO terms suggest that the activity of *rnf2* represses cellular responses to stress during normal homeostatic conditions and that epigenetic mechanisms facilitate the switch between homeostasis and cellular stress responses.

To investigate the spatial expression changes of differentially expressed genes from our RNA-seq data sets, we selected a subset to examine using WISH following *phc(RNAi)* or *rnf2(RNAi)*. For *rnf2* we selected 33 differentially expressed genes that were predicted to be involved in the extracellular matrix, stress response factors, cell signaling, and regulation of chromatin or transcription and assayed their expression after *rnf2* depletion. In general, *rnf2(RNAi)* caused a subtle effect on gene expression levels. In some instances, a robust change in expression occurred after *rnf2(RNAi)*, as seen clearly for *smcd-colec10* and *smcd-colec11*; expression of these genes is nearly undetectable in *control(RNAi)* worms as compared to *rnf2(RNAi)* worms developed for the same length of time (Figure 2.5B). Taken together, the GO and *in situ* analyses indicate that *rnf2* functions in broad cellular processes and that it maintains gene expression in differentiated tissues at appropriate levels.

In contrast to the mild effect on gene expression observed in *rnf2(RNAi)* animals, assaying mRNA expression of putative PHC target genes revealed striking changes in both

expression levels and spatial patterning after *phc(RNAi)*. We examined 11 genes using *in situ* hybridization, including genes involved in cell-adhesion, cell signaling, transcription, and chromatin regulation. For 7 of these 11 genes, strong ectopic expression was observed in *phc(RNAi)* worms in the region of the worm where the dorsal lesion forms (Figure 2.5D). Genes shown to be ectopically expressed in this region included the cell adhesion factor *icam5*, the Cut homeobox transcription factor *onecut1*, and an orphan nuclear receptor (*roar*). We also found that several chromatin regulators that were misexpressed in the region near the pharynx, including *cbx*, *pc-like*, *smc4*, and *kat6a*. Additionally, we found that the extra cellular matrix protein, *egflam*, which is normally expressed in the nervous system and pharynx tip, was significantly down regulated throughout the worm.

The ectopic expression of specific factors and disruption of *NB.22.1E* and laminin expression at the site of tissue defects in *phc(RNAi)* worms indicates that *phc* function is required to maintain the proper specification and integrity of tissues in this region.

Discussion

Planarians are an effective *in vivo* model for the screening of gene function in stem cells and regeneration. The use of a whole organism, in contrast to cell-based models, allows for the study of how specific factors regulate stem cells in a dynamic regenerative context as they differentiate into complex tissues and structures. Regeneration after injury requires the integration of signaling pathways to effectively recover from injury, initiate regeneration responses, and to direct proper cellular differentiation to recreate missing tissues; the post-translational modification of proteins by ubiquitin is an important regulatory step in many of these pathways but remains understudied in a regenerative context.

To address the role of ubiquitin signaling in stem cell regulation and regeneration in a whole organism *in vivo* context we performed a functional screen of RING and U-box class of E3 ubiquitin ligases that are expressed in stem cells and progeny in *S. mediterranea*. This functional screen returned nine genes that demonstrated phenotypes related to stem cell function or regeneration, building on previous studies from our lab on the HECT¹⁷ and Cullin⁸ classes of E3 ligases.

One gene that we chose to investigate in further detail was the U-box factor, *prpf19*, which is the founding member of the large protein complex NTC. First characterized in yeast, the best described role for NTC is in the spliceosome, where the E3 ligase function of Prpf19 is essential in the formation of snRNP conformations. We found that depletion of *prpf19* caused a strong homeostasis phenotype that included head regression, lesioning, ventral curling, and lysis, morphological effects that are often caused by stem cell depletion.

We depleted other NTC member genes in this study and observed similar phenotypes to *prpf19(RNAi)*, which suggests that the *prpf19(RNAi)* phenotype is mediated through its role in NTC. Surprisingly, we found that the stem cell population was maintained in *prpf19(RNAi)* worms, suggesting an alternative mechanism of dysregulation. Our results are consistent with a previous study that observed a similar phenotype upon depletion of *prpf19*, which showed an effect on head regeneration without disruption of the stem cells³⁰.

Post-transcriptional processing of RNAs is emerging as a major regulator of planarian stem cells and differentiation. The PIWI homolog *smedwi-2* was identified as being dispensable for stem cell maintenance but necessary for proper differentiation³³, and *smedwi-3* was shown to regulate planarian stem cell mRNA transcripts through two distinct activities³⁴. Similarly, the CCR4-NOT complex regulates the post-translational degradation of mRNAs and has been shown to have a critical role in planarian stem cell biology. The phenotype of CCR4-NOT complex member gene *smed-not1* was reported to have a similar phenotype to that of *prpf19*, where the animals maintained proliferative stem cells in *not1(RNAi)* worms despite presenting a phenotype that suggests loss of tissue renewal³⁵. In this study we found that an additional CCR4-NOT subunit, *not4*, is critical for worm homeostasis and causes head regression and ventral curling upon inhibition (Figure 2.1B). This phenotype is consistent with that of *not1(RNAi)*, but in the future it will be necessary to examine the stem cell population using marker genes in *not4(RNAi)* worms to resolve if the phenotype is mediated through a similar mechanism. Regulation of mRNAs in planarian stem cells by several pathways, including piRNAs, de-adenylation, or splicing is a crucial process for homeostasis and regeneration while being dispensable for stem cell

maintenance. Together these studies implicate post-transcriptional regulation of mRNAs in planarian stem cells as a critical process for regulating differentiation.

Epigenetic regulation of gene expression is essential during development and throughout organismal life span to maintain cellular identity. Understanding how epigenetic factors regulate adult stem cells has implications for how we understand these factors as drivers of cancers and how epigenetic “memory” can affect the dedifferentiation and subsequent reprogramming of induced pluripotent stem cells. In our RNAi screen we uncovered planarian homologs of histone-targeting RING E3 ubiquitin ligases that affected worm homeostasis and regeneration and used western blotting to confirm that inhibition of *bre1* and *rnf2* reduced levels of monoubiquityl-histone H2B and H2A, respectively. This work demonstrates that both activating and repressive signals provided through histone modifiers are essential for the proper specification of stem cells and establishment of cellular identity during regenerative events.

PRC1 is a major repressive complex that works during development to ubiquitylate histone H2A, compact chromatin, and silence target gene expression. PRC1 function was first discovered and remains perhaps best characterized as a repressor of the HOX genes during development³⁶. The core of the complex is defined by a RING and PCGF protein that together form either canonical or variant PRC1 depending on which other factors are present³⁷. The RING subunit acts as an E3 ligase that targets histone H2A, and in vertebrates is either RING1 or RNF2³⁸. In contrast to other invertebrates, we found that planarians have two genes predicted to be homologs of RING1 and RNF2. While these are likely to be lineage-specific paralogs instead of direct homologs of each vertebrate genes,

we do find that as in vertebrates, the *S. med rnf2* paralog acts as the major ligase and is responsible for the bulk of histone H2A ubiquitylation. We did not observe a noticeable difference in global levels of H2Aub1 levels after *Smed-ring1* inhibition (Supplemental Figure S2.4B). However, as both genes demonstrated similar regeneration-specific phenotypes suggests that they may share common targets or pathways. In contrast to *rnf2(RNAi)*, when we inhibited other PRC1 core elements, *phc* and *cbx*, we did not see a reduction in bulk H2Aub1 levels by western blotting. From work in mammalian cell lines, it was determined that variant PRC1 activity is responsible for the majority of H2A ubiquitylation with a minimal contribution from canonical PRC1 complexes^{32,39}. Invertebrates were not thought to contain vPRC1 but more recent phylogenetic analysis that included a greater variety of invertebrate model organism indicates that vPRC1 likely evolved as early as cnidarians⁴⁰. Our western blot results showing the conserved catalytic activity of specific planarian RING protein and the identification of two *S. med pcgf* genes strongly support the presence of vPRC1 in *S. mediterranea*. Moreover, our data suggests that cPRC1 has a minor contribution to overall H2Aub1 levels *in vivo*.

To gain insight into the unexpected discrepancy in the phenotypes after *rnf2* or *phc* depletion and to understand which genes are regulated by each factor we performed RNA-sequencing after *rnf2* or *phc* RNAi. Consistent with predicted roles as subunit of a complex that represses transcription, we identified more upregulated than downregulated genes after depletion of either *rnf2* or *phc*. There was only a single gene in common between these data sets, supporting a transcriptional link to the different *rnf2(RNAi)* and *phc(RNAi)*. Interestingly the gene that was in common between the data sets was cPRC1 gene *cbx*; this gene was upregulated in both data sets and was the most significantly upregulated gene

after *phc(RNAi)*. It is possible that PRC1 autoregulates its own expression in planarians and that the disruption to chromatin homeostasis induced by PRC1 inhibition induces a compensatory response involving other chromatin factors.

Using GO analysis, we found that *rnf2* regulates genes related to the cellular stress response; when we examined the expression of candidate genes from our RNA-seq data set using WISH we observed that gene expression changes in *rnf2(RNAi)* animals occurred mainly within the endogenous expression pattern. In short, these data support a role for RNF2, and potentially H2A ubiquitylation, in tuning transcription levels within a particular cell type, especially for pathways that are adaptive and responsive to stressful stimuli.

In contrast, we saw dramatic shifts in the spatial expression of specific genes after *phc(RNAi)*, including several genes that showed ectopic expression near the base of the pharynx where the *phc(RNAi)* phenotype presented. In this affected region, we observed both up (*intercellular adhesion molecule 5*) and down regulation (*pikachurin*) of genes that encode extracellular matrix and intercellular adhesion molecules, suggesting that their dysregulation is likely linked to the formation of the dorsal lesion seen after *phc(RNAi)*. The ectopically expressed genes also included regulators of cellular specification, including nuclear receptors, transcription factors, and chromatin modifiers. One gene we identified as being misexpressed after *phc* depletion was that encoding the nuclear factor *onecut1*, a CUT and homeobox domain-containing transcription factor that promotes hepatocyte proliferation, remodels chromatin accessibility, and promotes tumor growth in colorectal cancers⁴¹⁻⁴³. Based on its role in regulating transcription and tissue identity in other animal models, we suspect it may be contributing to the change in patterning near the pharynx.

Future investigation into this factor would elucidate if *onecut1* misexpression is driving regional tissue misspecification and if inhibition of *onecut1* is suppressive towards the *phc(RNAi)* phenotype.

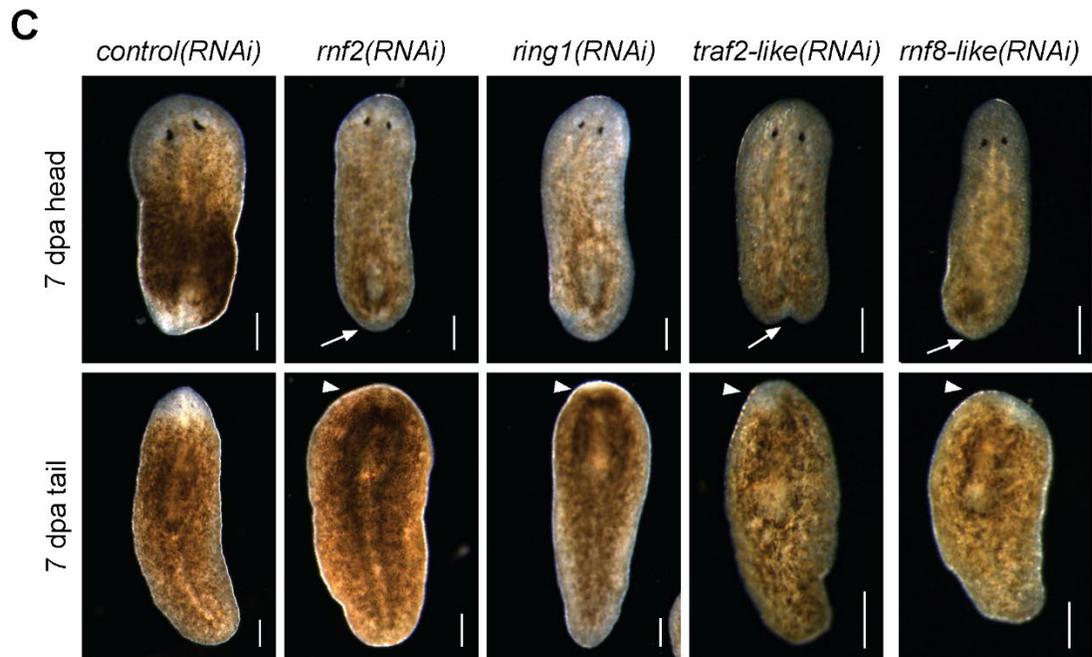
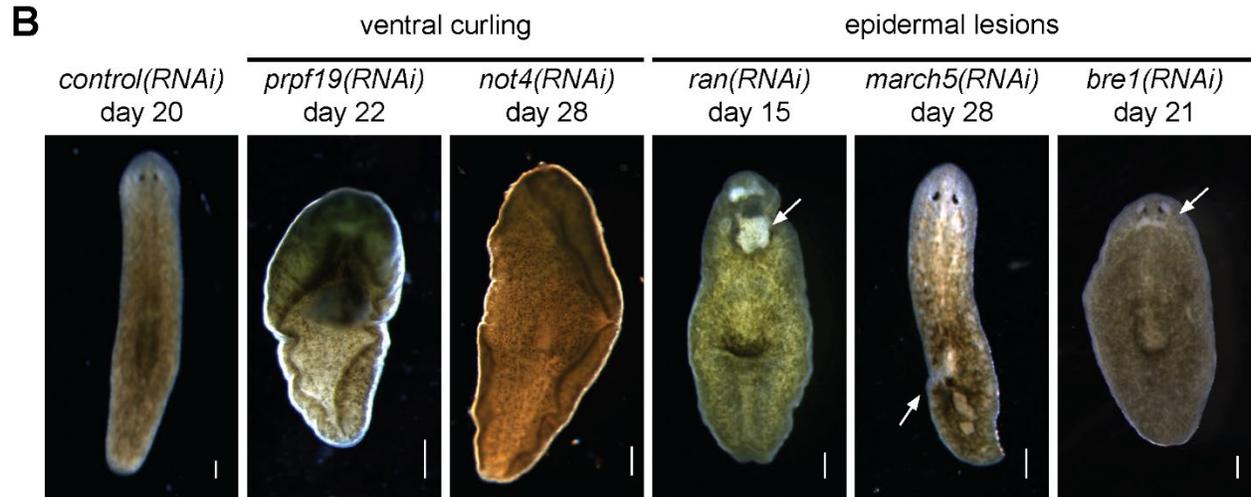
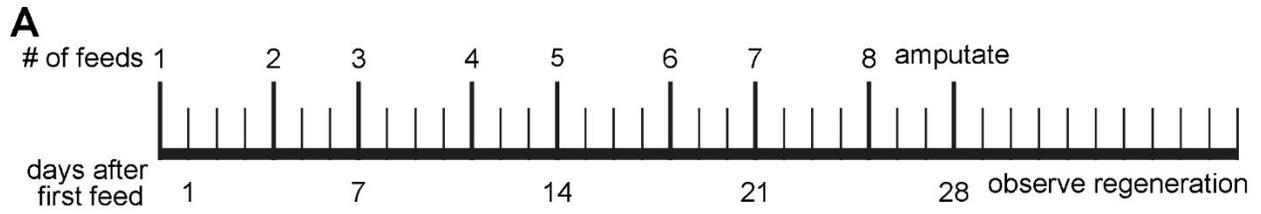
Epigenetic factors work through the modification of chromatin, a deeper understanding of the role of epigenetics in the regulation of regeneration will require understanding the genomic elements that are being regulated by each factor. Application of assays to measure where in the planarian genome certain histone marks are localized using ChIP-seq or to assay changes in chromatin accessibility more generally using ATAC-seq, would inform how genes are regulated epigenetically to promote a robust regenerative response during injury and tissue re-specification and remodeling.

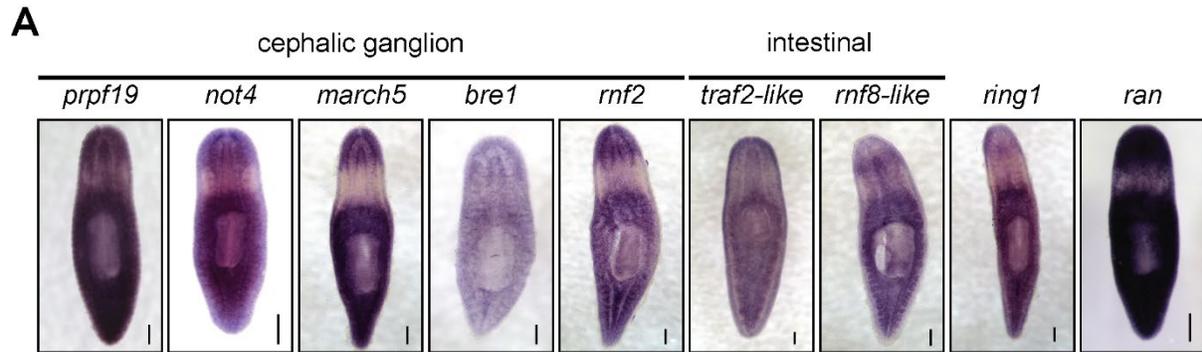
Acknowledgements

We thank Catherine May for assistance in setting up the RNA-seq analysis and providing template R scripts; Celeste Romero for help with RNAi experiments; Kelly Ross and Elizabeth Duncan for helpful discussion and feedback on the manuscript.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Allen, JM.; Balagtas M.; Barajas, E.; Cano, C.; Iberkleid, I.; Zayas, RM. The dissertation author was the primary investigator and author of this manuscript

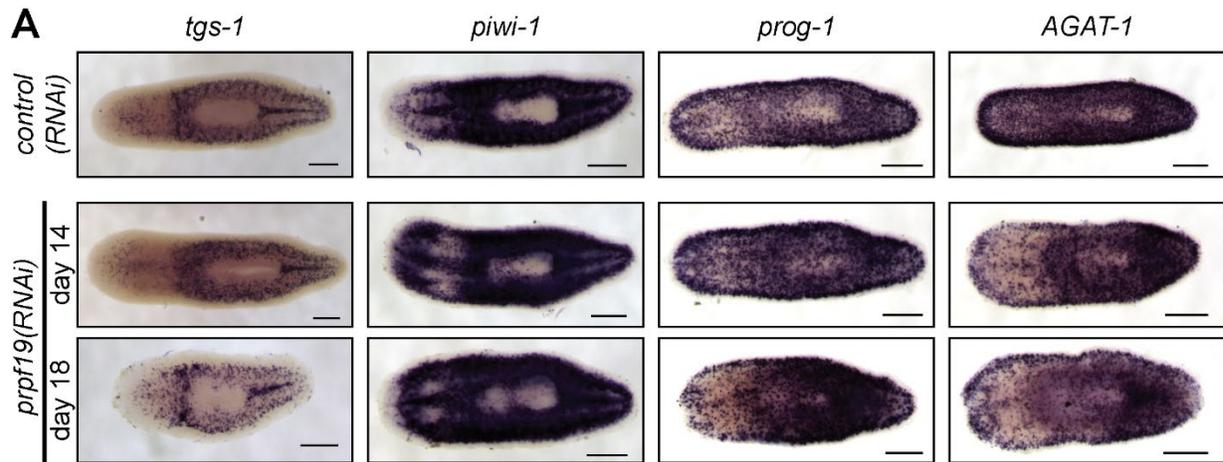
Figure 2.1. RNAi screen of RING and U-Box E3 ligases identifies regulators of stem cells and regeneration. (A) Feeding and amputation schedule for RNAi screening. Worms were fed twice per week for a total of eight feeds and amputated pre-pharyngeally on day 28. (B) Knockdown of the indicated genes displayed phenotypes including ventral curling and lesions (white arrow). Animals are shown after RNAi feedings prior to amputation. (C) Knockdown of indicated genes that demonstrated phenotypes of delayed or absent regeneration after amputation shown by smaller than normal or absent blastemas (white arrow) and missing or faint eyespots (white arrowhead) when compared to *control(RNAi)* worm at the same regeneration time point. Scale bars = 200 μ m



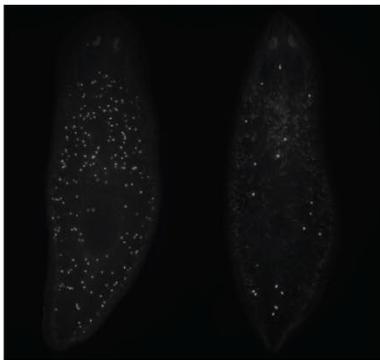


Supplemental Figure S2.1. (A) Whole mount *in situ* hybridization patterns for genes that showed phenotypes in RNAi screen. All genes examined were expressed throughout the parenchyma with some genes displaying enriched expression near the cephalic ganglion or near the intestine of the worm. Scale bars = 200 μ m.

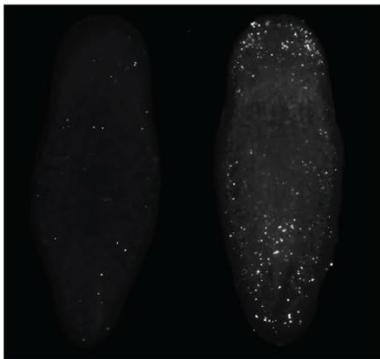
Figure 2.2. Inhibition of *prpf19* disrupts neoblast function but is not required for stem cell maintenance. (A) WISH to stem cell markers *tgs-1* and *piwi-1* and early and late epidermal progeny markers *prog-1* and *AGAT-1*, respectively, in *control(RNAi)* (upper panels) and *prpf19(RNAi)* animals at 14 (middle panels) and 18 (bottom panels) days after first RNAi feeding. (B) Representative image of animals fixed 14 days after first RNAi feeding for *control(RNAi)* (left) or *prpf19(RNAi)* (right) and immunostained for mitotic marker phospho-Histone H3. (C) Quantification of phospho-Histone H3⁺ cells per mm² of worms fixed at 11, 14, and 18 days after first RNAi feed (N = 9 – 13 per time point, ***p-value < 0.001). (D) Representative image of animals fixed 14 days after first RNAi feeding for *control(RNAi)* (left) or *prpf19(RNAi)* (right) and processed for TUNEL staining. (E) Quantification of TUNEL⁺ cells per mm² of worms fixed at 10, 14, and 17 days after first RNAi feed (N = 5 – 6 per time point, *p-value < 0.05). Scale bars = 200 μm.



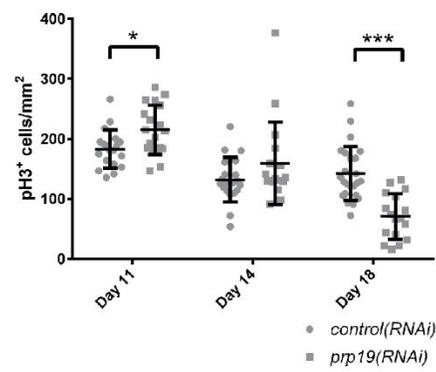
B control(RNAi) *prpf19*(RNAi)



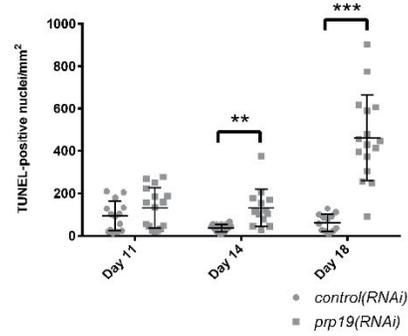
D control(RNAi) *prpf19*(RNAi)

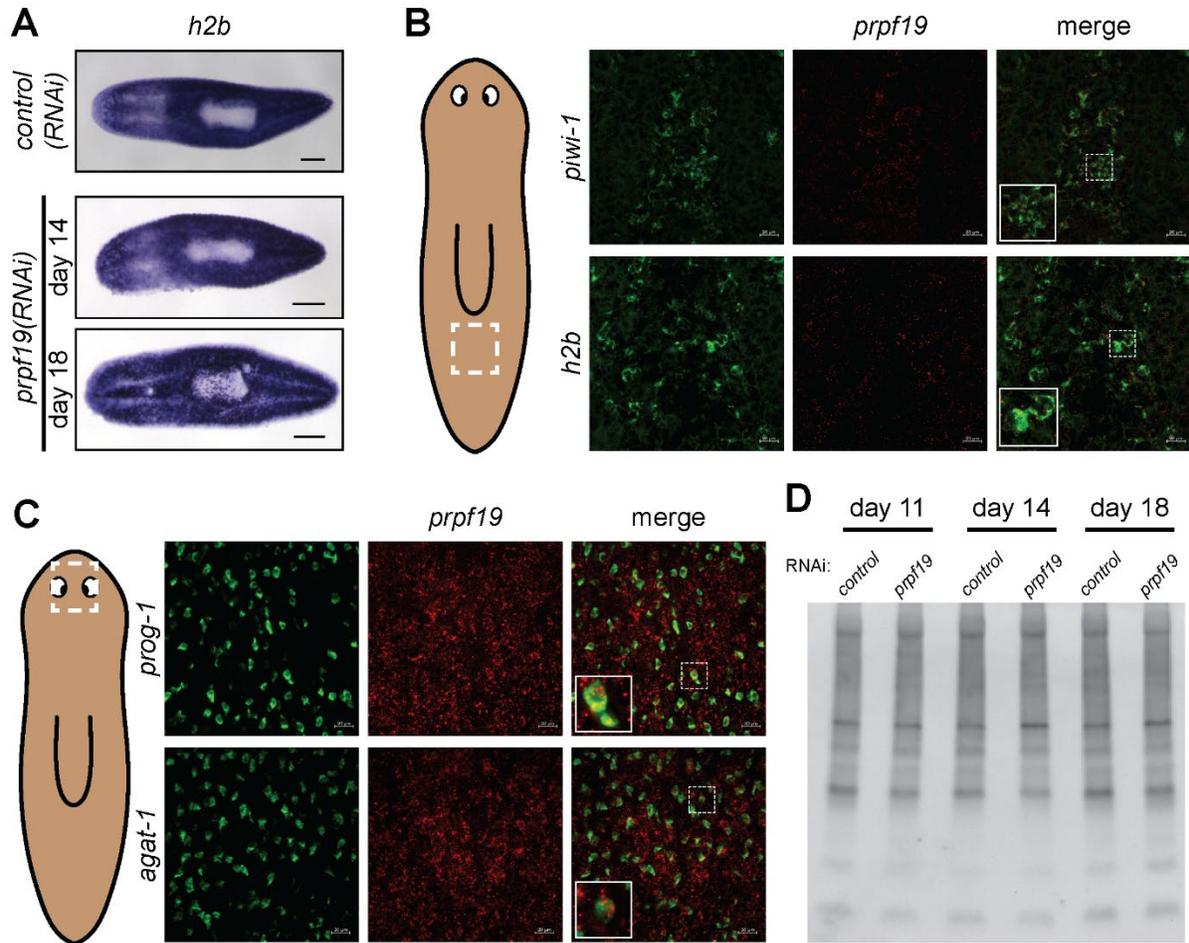


C



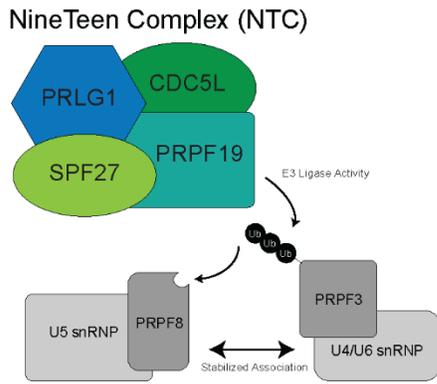
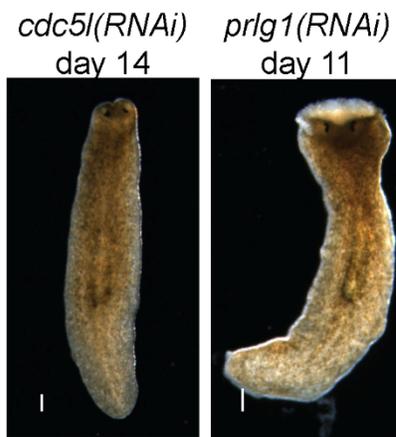
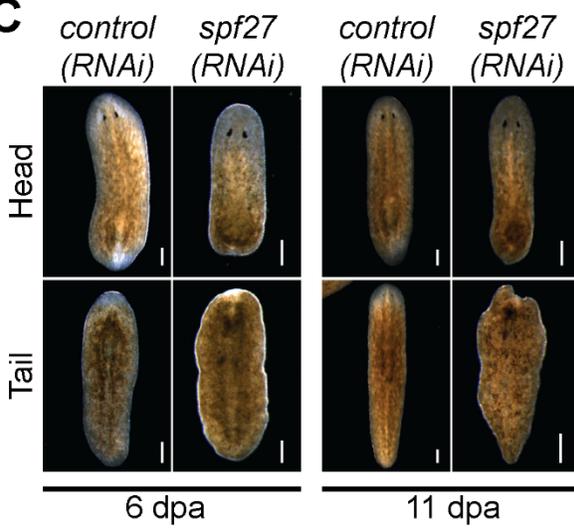
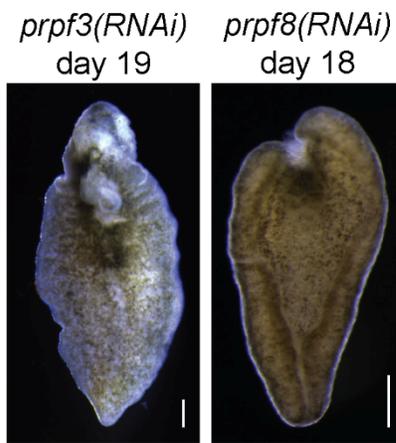
E

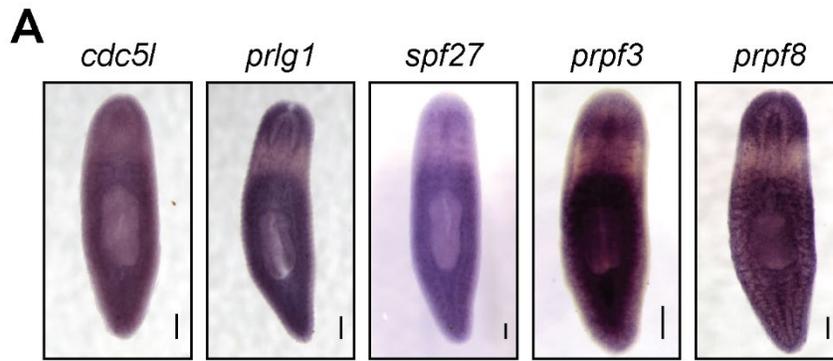




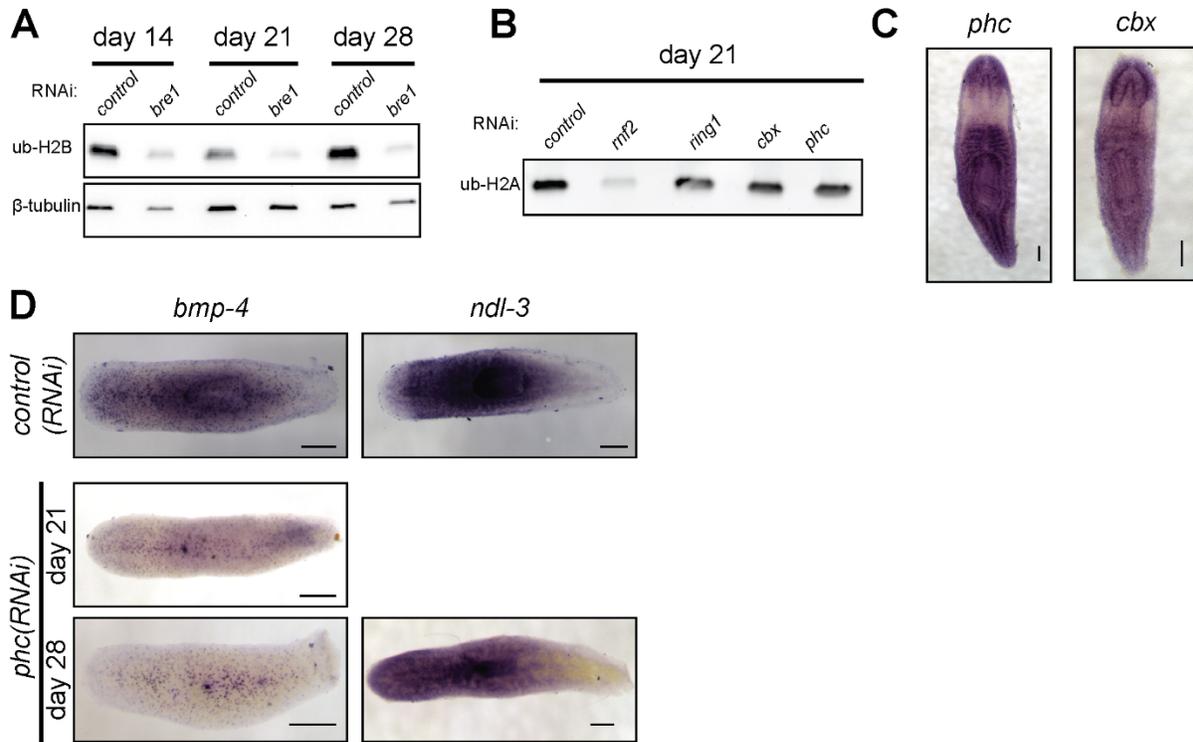
Supplemental Figure S2.2. (A) WISH to stem cell marker *H2B* in *control*(RNAi) animals (upper panel) and *prpf19*(RNAi) animals at 14 (middle panel) and 18 (bottom panel) days after first RNAi feeding. (B) Double fluorescence *in situ* hybridization showing co-expression of *prpf19* (red) with marker genes (green) for neoblasts (*piwi-1*), early epidermal progeny marker (*prog-1*), and late epidermal progeny marker (*AGAT-1*). (C) Western blot probed using anti-ubiquitin antibody for whole worm homogenate protein extracts from *control*(RNAi) or *prpf19*(RNAi) treated animals collected at days 11, 14, and 18 of treatment. Scale bars = 200 μ m (A), 20 μ m (B) (C).

Figure 2.3. *prpf19*-associated factors and downstream targets recapitulate *prpf19(RNAi)* phenotypes. (A) Prpf19 acts as an E3 ligase in NTC, interacting with core complex members PLRG1, CDC51, and SPF27 to modify U4/U6 snRNP subunit PRPF3 with nonproteolytic K63-linked ubiquitin chains. This ubiquityl mark stabilizes the interaction of PRPF3 with U5 snRNP subunit PRPF8 to allow the stable formation of the U4/U6.U5 tri-snRNP and the catalytic activity of the spliceosome. (B) Knockdown of indicated NTC core components *cdc51* and *prlg1* displaying head regression. (C) Knockdown of NTC core component SPF27 caused a reduced and delayed regenerative response in amputated worms. At 6 days post amputation *spf27(RNAi)* worms have blastemas that are reduced in size relative to *control(RNAi)* worms at the same time point. At day 11 post amputation the regenerative response in *control(RNAi)* worms is largely concluded with large blastemas and visible reformed eyespots present in trunk fragments. In comparison *spf27(RNAi)* worms have smaller blastemas and have not regenerated eyespots. (D) Inhibition of Prpf19 target *prpf3* and ubiquityl-Prpf3 binding factor *prpf8* demonstrate phenotypes similar to *prpf19(RNAi)* and includes head regression, lesions, and ventral curling. Scale bars = 200 μm .

A**B****C****D**

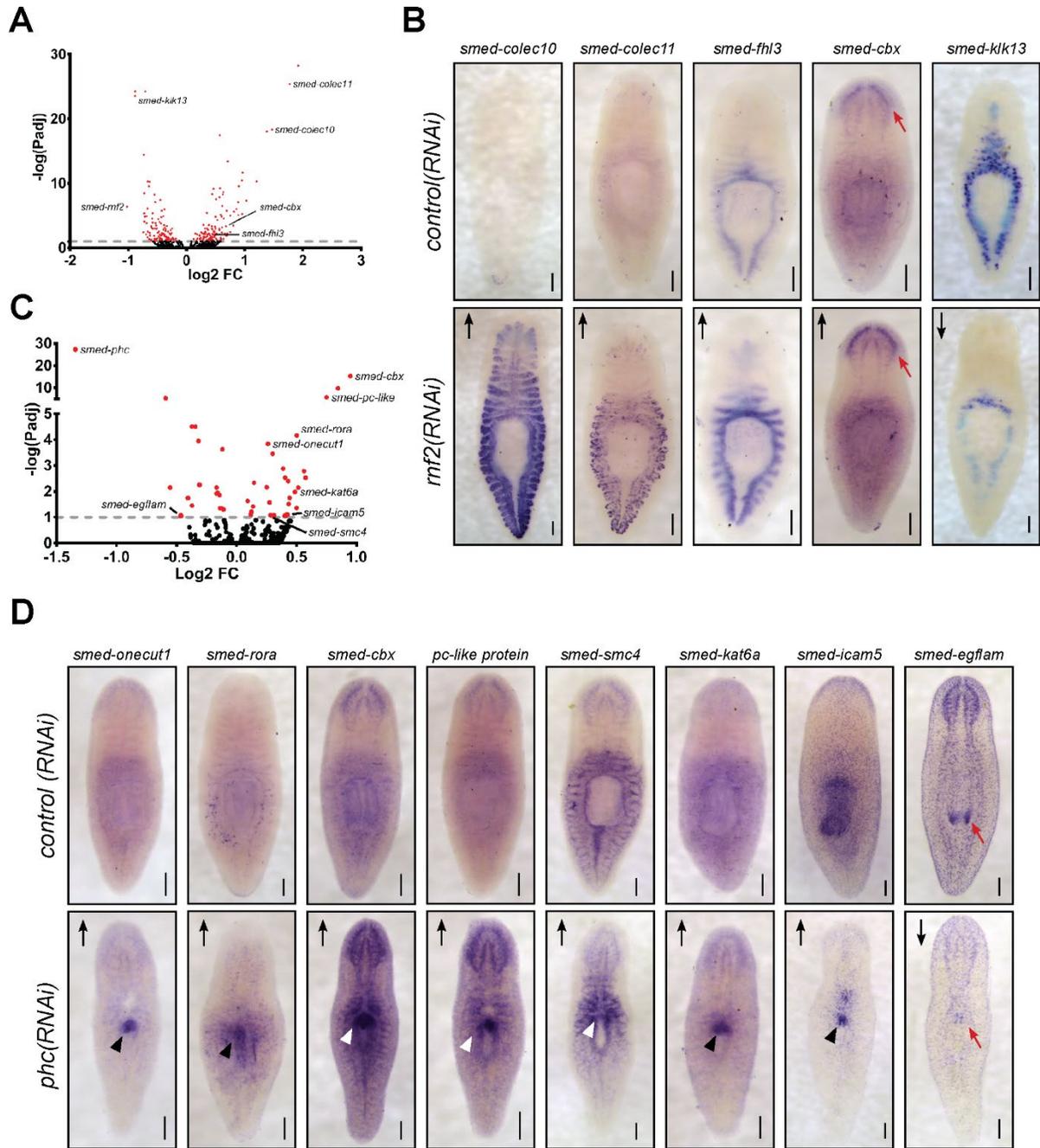


Supplemental Figure S2.3. (A) WISH to NTC core elements *cdc5l*, *pflg1*, and *spf27*, and spliceosomal RNP members *prpf3* and *prpf8*. Scale bars = 200 μ m.



Supplemental Figure S2.4. (A) Western blot analysis showing reduction in ubiquityl-Histone H2B levels following inhibition of *bre1* at 14, 21 and 28 days of RNAi treatment. (B) Western blot analysis shows no reduction in H2Aub1 levels following inhibition of cPRC1 genes *ring1*, *cbx*, and *phc* following 21 days of RNAi treatment. (C) WISH analysis showing expression patterns for cPRC1 genes *phc*, and *cbx*. (D) WISH marker gene analysis for D-V marker *bmp4* and A-P marker *ndl-3* showing no change in expression domains following *phc* inhibition. Scale bars = 200 μm.

Figure 2.5. Loss of PRC1 function causes changes to gene expression levels and spatial patterns (A) Volcano plot of differentially expressed genes between *rnf2(RNAi)* and *gfp(RNAi)* worms after 28 days of treatment. (B) WISH analysis of selected genes indicated to be differentially expressed after *rnf2(RNAi)* by RNA-seq. (C) Volcano plot of differentially expressed genes between *phc(RNAi)* and *gfp(RNAi)* after 11 days of treatment. (D) WISH analysis of selected genes indicated to be differentially expressed after *phc(RNAi)*. Arrows indicate up or down regulated expression by RNA-seq. Red arrows highlight regions with changed expression after RNAi in the brain (B) and mouth (D) regions of the worm. Arrowheads indicate regions of ectopic gene expression after RNAi treatment. Scale bars = 200 μ m.



Supplemental Figure S2.5. (A) Feeding and sampling schedule for RNA-seq experiments. Worms were fed twice a week and sampled for RNA extraction at the days indicated by arrowheads. (B) Volcano plot of differentially expressed genes between *rnf2(RNAi)* and *gfp(RNAi)* worms after 14 days of treatment. (C) Venn diagram showing overlap between day 28 and day 14 sample sets for *rnf2(RNAi)*. (D) GO analysis of downregulated genes in *rnf2(RNAi)* worms after 28 days of treatment. (E) GO analysis of upregulated genes in *rnf2(RNAi)* worms after 28 days of treatment

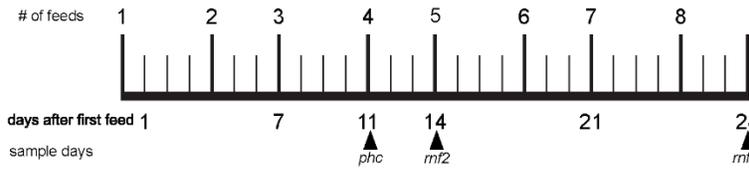
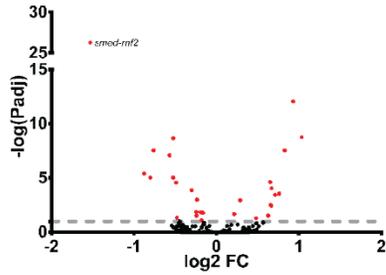
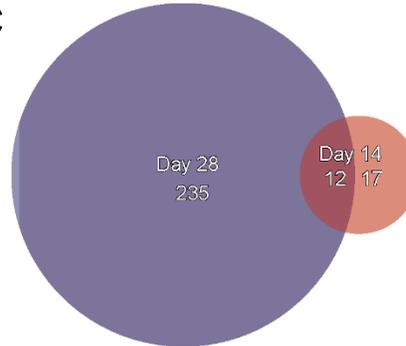
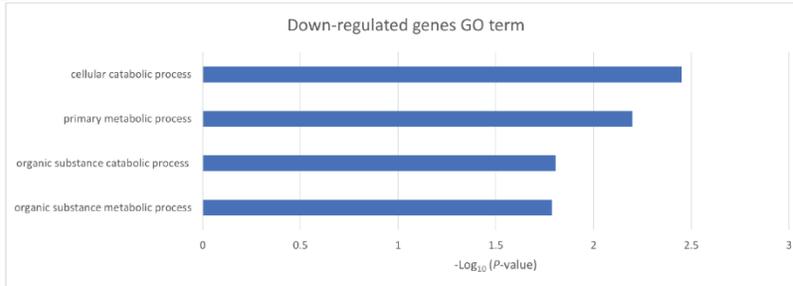
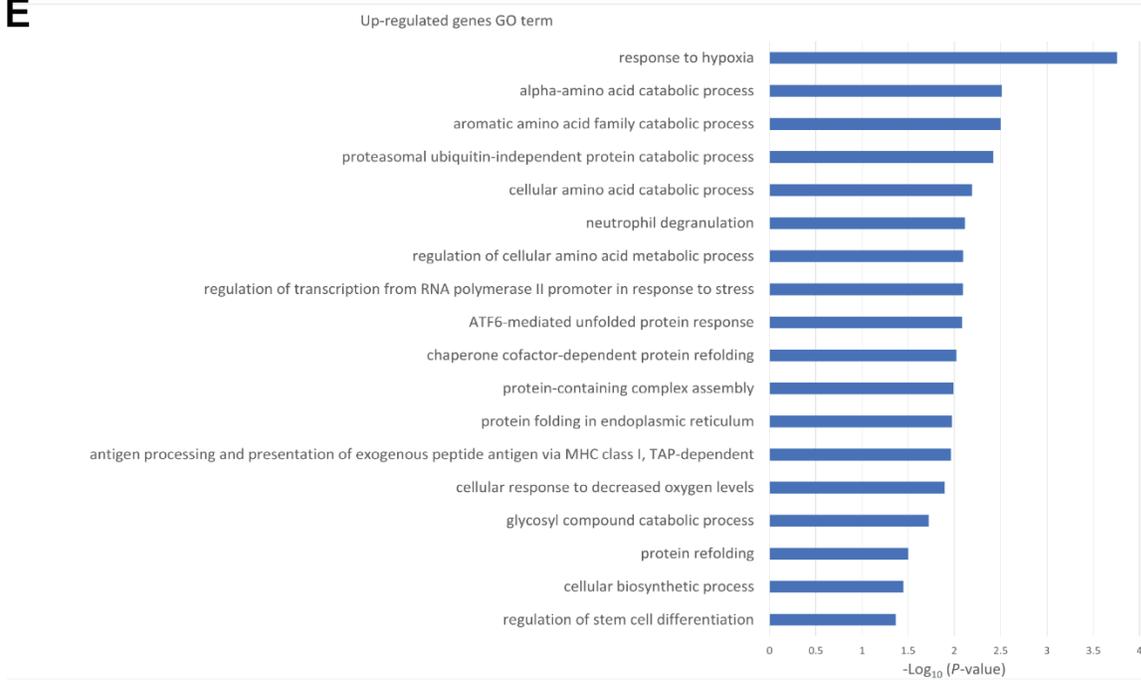
A**B****C****D****E**

Table 2.1: RING and U-box E3 ubiquitin ligases with phenotypes after inhibition in *S. mediterranea*. HR: Head regression. VC: Ventral Curling. DR: Delayed regeneration.

Gene Name	Dresden Contig ID	Human RING or U-box homolog	E Value	Phenotypes Observed
<i>Smed-prpf19</i>	dd_Smed_v6_1276_0_1	PRPF19	0.0	HR, VC, Lysis
<i>Smed-march5</i>	dd_Smed_v6_4602_0_1	MARCH5	8e-93	HR, Lesions, Lysis
<i>Smed-traf2-like</i>	dd_Smed_v6_3837_0_1	TRAF2	2e-69	HR
<i>Smed-ran</i>	dd_Smed_v6_330_0_1	CBLB*	3e-94	Lesions, Lysis
<i>Smed-not4</i>	dd_Smed_v6_4767_0_1	CNOT4	5e-87	HR, VC
<i>Smed-rnf8-like</i>	dd_Smed_v6_1137_0_5	RNF8	4e-05	DR, Lesions
<i>Smed-bre1</i>	dd_Smed_v6_4070_0_1	RNF40	1e-88	HR, DR
<i>Smed-rnf2</i>	dd_Smed_v6_8989_0_1	RNF2	6e-46	DR
<i>Smed-ring1</i>	dd_Smed_v6_12141_0_1	RING1	6e-33	DR

*Top Human BLAST hit: RAN, e-value: 9e-116

Table 2.2: Summary of RNA-seq results. Totals of differentially expressed genes at FDR cutoff value < 0.1.

Gene Name	Days of RNAi treatment	Downregulated Genes	Upregulated genes	Total
<i>Smed-phc</i>	11	20	29	49
<i>Smed-rnf2</i>	14	16	13	29
<i>Smed-rnf2</i>	28	113	134	247

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_1075_1_0_1	gi 52487176 ref NP_001005207.1	640	0	tripartite motif-containing 37	TRIM37
dd_Smed_v6_1276_0_1	sp Q9UMS4 ref NP_055317	571	0	PRP19/PSO4 pre-mRNA processing factor 19 homolog	PRPF19
dd_Smed_v6_4224_0_1	gi 187761373 ref NP_005735.2	549	0	ariadne homolog	ARIH1
dd_Smed_v6_12234_0_1	gi 21361137 ref NP_002495.2	476	7.10E-154	nuclear transcription factor, X-box binding 1	NFX1
dd_Smed_v6_9115_0_1	gi 57863277 ref NP_001009921.1	498	3.50E-153	vacuolar protein sorting 8 homolog (S. cerevisiae)	VPS8
dd_Smed_v6_4217_0_1	sp Q13356 ref NP_055152	440	3.91E-151	peptidylprolyl isomerase (cyclophilin)-like 2	PPIL2
dd_Smed_v6_9989_0_1	gi 190341104 ref NP_055978.4	457	1.16E-149	tripartite motif-containing 9	TRIM9
dd_Smed_v6_2194_0_1	gi 27436925 ref NP_115807.1	440	1.03E-148	synovial apoptosis inhibitor 1, synoviolin	SYVN1
dd_Smed_v6_6226_0_1	gi 55749557 ref NP_001006611.1	409	5.52E-143	seven in absentia homolog 1 (Drosophila)	SIAH1
dd_Smed_v6_4795_0_1	gi 21361137 ref NP_002495.2	442	1.91E-139	nuclear transcription factor, X-box binding 1	NFX1
dd_Smed_v6_6104_0_1	gi 30348954 ref NP_065825.1	447	2.18E-137	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_4887_0_1	gi 21361732 ref NP_060790.2	389	1.40E-136	ring finger protein 121	RNF121
dd_Smed_v6_4558_0_1	sp Q14139 ref NP_004779	434	2.83E-136	ubiquitination factor E4A (UFD2 homolog, yeast)	UBE4A
dd_Smed_v6_6780_0_1	gi 52426745 ref NP_005179.2	387	8.96E-126	Cas-Br-M ecotropic retroviral transforming sequence	CBL
dd_Smed_v6_3481_0_1	gi 114199475 ref NP_055211.2	390	2.22E-123	vacuolar protein sorting 41 homolog (S. cerevisiae)	VPS41
dd_Smed_v6_7882_0_1	gi 45594312 ref NP_115647.2	374	5.59E-123	TNF receptor-associated factor 7	TRAF7
dd_Smed_v6_7475_0_1	sp O95155 ref NP_006039	392	1.14E-119	ubiquitination factor E4B (UFD2 homolog, yeast)	UBE4B
dd_Smed_v6_9945_0_1	gi 291190787 ref NP_055872.4	417	1.51E-118	MYC binding protein 2	MYCBP2
dd_Smed_v6_6166_0_1	gi 160948610 ref NP_742067.3	394	2.98E-113	zinc finger protein 650	ZNF650
dd_Smed_v6_5173_0_1	gi 5032071 ref NP_005776.1	314	8.09E-108	ring finger protein 41	RNF41
dd_Smed_v6_5473_0_1	gi 45387949 ref NP_588609.1	335	9.03E-108	ring finger and SPRY domain containing 1	RSPRY1
dd_Smed_v6_330_0_1	gi 54112420 ref NP_733762.2	321	5.09E-96	Cas-Br-M ecotropic retroviral transforming sequence b	CBLB
dd_Smed_v6_4070_0_1	gi 7662230 ref NP_055586.1	305	2.49E-90	ring finger protein 40	RNF40
dd_Smed_v6_5649_1_1	gi 37588869 ref NP_071347.2	316	3.66E-90	ring finger protein 123	RNF123
dd_Smed_v6_5198_0_1	gi 113417068 ref XP_001128827.1	308	2.94E-87	mitogen-activated protein kinase kinase 1	MAP3K1
dd_Smed_v6_4767_0_1	gi 56550059 ref NP_001008226.1	283	3.44E-87	CCR4-NOT transcription complex, subunit 4	CNOT4
dd_Smed_v6_1293_0_1	gi 5454168 ref NP_006453.1	272	5.85E-87	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_2121_0_1	sp Q9UNE7 ref NP_005852	253	1.70E-84	STIP1 homology and U-box containing protein 1	STUB1
dd_Smed_v6_14099_0_1	gi 71043932 ref NP_872327.2	266	1.04E-82	zinc finger, SWIM-type containing 2	ZSWIM2
dd_Smed_v6_1704_0_1	gi 30348954 ref NP_065825.1	276	3.59E-79	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_5467_0_1	gi 24025688 ref NP_699202.1	263	2.03E-77	ligand of numb-protein X 2	LNX2
dd_Smed_v6_7536_0_1	gi 5902158 ref NP_008909.1	233	1.01E-76	ring finger protein 113A	RNF113A
dd_Smed_v6_1317_0_1	gi 30348954 ref NP_065825.1	251	2.54E-73	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_7689_0_1	gi 21071001 ref NP_001135.3	238	9.64E-72	autocrine motility factor receptor	AMFR
dd_Smed_v6_4440_0_1	gi 5454168 ref NP_006453.1	228	6.73E-70	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_23131_0_1	sp O95155 ref NP_006039	233	1.31E-69	ubiquitination factor E4B (UFD2 homolog, yeast)	UBE4B
dd_Smed_v6_2627_0_1	gi 38788243 ref NP_037394.2	229	1.60E-67	myosin regulatory light chain interacting protein	MYLIP
dd_Smed_v6_1136_0_1	gi 7657508 ref NP_055063.1	194	2.54E-66	ring-box 1	RBX1
dd_Smed_v6_12150_0_1	gi 13569903 ref NP_112198.1	207	1.05E-65	ring finger protein 32	RNF32
dd_Smed_v6_4148_0_1	gi 22027612 ref NP_066961.2	210	1.29E-63	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_6165_0_1	gi 33300635 ref NP_006773.2	199	5.71E-62	zinc finger protein-like 1	ZFPL1
dd_Smed_v6_3477_0_1	gi 31742478 ref NP_006306.2	192	7.33E-62	polycomb group ring finger 3	PCGF3
dd_Smed_v6_5532_0_1	gi 58331204 ref NP_065775.1	206	1.15E-60	ring finger protein 150	RNF150
dd_Smed_v6_4025_0_1	gi 109150431 ref NP_060469.4	216	3.80E-60	ring finger protein 31	RNF31
dd_Smed_v6_13517_0_1	gi 22027612 ref NP_066961.2	201	9.65E-60	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_3937_0_1	gi 231573214 ref NP_056380.2	220	2.45E-59	zinc finger protein 294	ZNF294
dd_Smed_v6_8994_0_1	gi 24025688 ref NP_699202.1	205	1.89E-57	ligand of numb-protein X 2	LNX2
dd_Smed_v6_5004_0_1	gi 31542783 ref NP_689480.2	174	1.73E-54	ring finger protein 185	RNF185
dd_Smed_v6_7487_0_1	gi 109150431 ref NP_060469.4	196	4.49E-54	ring finger protein 31	RNF31
dd_Smed_v6_6678_0_1	gi 17978477 ref NP_068375.3	192	1.11E-53	vacuolar protein sorting 11 homolog (S. cerevisiae)	VPS11
dd_Smed_v6_4392_0_1	gi 22027616 ref NP_003291.2	186	1.55E-53	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_10351_1	gi 54112420 ref NP_733762.2	196	1.74E-52	Cas-Br-M ecotropic retroviral transforming sequence b	CBLB
dd_Smed_v6_10841_0_1	gi 53729361 ref NP_001005373.1	189	3.06E-52	leucine rich repeat and sterile alpha motif containing 1	LRSAM1
dd_Smed_v6_6206_0_1	gi 62865649 ref NP_001017368.1	174	9.75E-52	ring finger and FYVE-like domain containing 1	RFFL

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_12753_0_1	gi 6005964 ref NP_009075.1	170	1.20E-51	polycomb group ring finger 2	PCGF2
dd_Smed_v6_11736_0_1	gi 239048907 ref NP_060693.2	175	3.21E-50	checkpoint with forkhead and ring finger domains	CHFR
dd_Smed_v6_4083_0_2	gi 188497705 ref NP_006759.3	175	4.33E-48	BRCA1 associated protein	BRAP
dd_Smed_v6_8989_0_1	gi 6005747 ref NP_009143.1	163	9.38E-48	ring finger protein 2	RNF2
dd_Smed_v6_3780_0_1	gi 33620769 ref NP_008841.2	177	1.59E-46	retinoblastoma binding protein 6	RBBP6
dd_Smed_v6_1110_0_1	gi 22027616 ref NP_003291.2	167	3.78E-46	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_7918_0_1	gi 237858654 ref NP_112216.3	149	1.33E-45	ring finger protein 170	RNF170
dd_Smed_v6_1802_0_1	gi 5454168 ref NP_006453.1	158	7.95E-45	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_13573_0_1	gi 71143112 ref NP_060594.3	162	7.56E-44	ring finger and WD repeat domain 3	RFWD3
dd_Smed_v6_5342_0_1	gi 5454168 ref NP_006453.1	152	6.17E-43	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_9317_0_1	gi 109150431 ref NP_060469.4	152	4.30E-41	ring finger protein 31	RNF31
dd_Smed_v6_2436_0_1	gi 14150005 ref NP_115644.1	136	8.38E-41	zinc and ring finger 1	ZNRF1
dd_Smed_v6_2662_0_1	gi 5454168 ref NP_006453.1	145	3.68E-40	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_802_0_1	gi 237858654 ref NP_112216.3	141	3.97E-40	ring finger protein 170	RNF170
dd_Smed_v6_168_0_1	gi 22027612 ref NP_066961.2	150	1.78E-39	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_1923_0_1	gi 14149702 ref NP_056343.1	144	2.23E-39	ring finger protein 167	RNF167
dd_Smed_v6_2819_0_1	gi 5454168 ref NP_006453.1	147	6.78E-39	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_3201_0_1	gi 22027616 ref NP_003291.2	143	1.29E-38	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_12318_0_1	gi 33620769 ref NP_008841.2	144	9.51E-38	retinoblastoma binding protein 6	RBBP6
dd_Smed_v6_1139_0_1	gi 77404348 ref NP_001029082.1	140	1.81E-37	TNF receptor-associated factor 5	TRAF5
dd_Smed_v6_3487_0_1	gi 7657522 ref NP_055060.1	118	3.32E-37	ring finger protein 7	RNF7
dd_Smed_v6_8046_0_1	gi 4759254 ref NP_004611.1	138	7.89E-37	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_7396_0_1	gi 58331204 ref NP_065775.1	135	7.61E-36	ring finger protein 150	RNF150
dd_Smed_v6_14782_0_1	gi 32454739 ref NP_006449.2	138	9.73E-36	tripartite motif-containing 3	TRIM3
dd_Smed_v6_6291_0_1	gi 5454168 ref NP_006453.1	132	1.01E-35	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_6913_0_1	gi 7657520 ref NP_055187.1	124	2.01E-35	ring finger protein 11	RNF11
dd_Smed_v6_12141_0_1	gi 51479192 ref NP_002922.2	123	1.01E-34	ring finger protein 1	RING1

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_1158_0_0_1	gi 4757762 ref NP_004281.1	90.9	3.54E-34	ring finger protein 14	RNF14
dd_Smed_v6_2669_0_1	gi 22027616 ref NP_003291.2	129	6.68E-34	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_1755_1_0_1	gi 22027612 ref NP_066961.2	130	8.36E-34	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_1858_0_1	gi 6005964 ref NP_009075.1	126	4.37E-33	polycomb group ring finger 2	PCGF2
dd_Smed_v6_1042_4_0_1	gi 4759254 ref NP_004611.1	126	8.74E-33	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_4074_0_1	gi 109150431 ref NP_060469.4	128	2.33E-32	ring finger protein 31	RNF31
dd_Smed_v6_8395_0_1	gi 30794216 ref NP_112223.1	127	3.48E-32	tripartite motif-containing 56	TRIM56
dd_Smed_v6_9091_0_1	gi 44917608 ref NP_056061.1	122	9.79E-32	mahogunin, ring finger 1	MGRN1
dd_Smed_v6_3097_0_1	gi 77404348 ref NP_001029082.1	121	1.08E-30	TNF receptor-associated factor 5	TRAF5
dd_Smed_v6_7426_0_1	gi 209180481 ref NP_079090.2	118	1.43E-30	Cas-Br-M ecotropic retroviral transforming sequence-like 1	CBLL1
dd_Smed_v6_6485_0_1	gi 22027616 ref NP_003291.2	119	2.49E-30	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_5260_0_1	gi 21389515 ref NP_653327.1	119	2.67E-29	ring finger protein 145	RNF145
dd_Smed_v6_5084_0_1	gi 33636758 ref NP_112225.2	108	5.32E-29	ring finger protein 146	RNF146
dd_Smed_v6_6800_0_1	gi 282394030 ref NP_543151.2	119	6.86E-29	mindbomb homolog 2 (Drosophila)	MIB2
dd_Smed_v6_1147_8_0_1	gi 4506343 ref NP_000309.1	107	7.12E-29	peroxisomal membrane protein 3, 35kDa (Zellweger syndrome)	PXMP3
dd_Smed_v6_6660_0_1	gi 194248079 ref NP_056086.2	118	8.52E-29	tripartite motif-containing 2	TRIM2
dd_Smed_v6_3952_0_1	gi 22027612 ref NP_066961.2	114	9.43E-29	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_4187_0_1	gi 31542783 ref NP_689480.2	102	1.09E-27	ring finger protein 185	RNF185
dd_Smed_v6_974_0_1	gi 22027612 ref NP_066961.2	109	3.41E-27	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_7291_0_1	gi 4759254 ref NP_004611.1	107	1.81E-26	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_8939_0_1	gi 7657520 ref NP_055187.1	93.2	2.65E-26	ring finger protein 11	RNF11
dd_Smed_v6_6250_0_1	gi 194248079 ref NP_056086.2	109	9.44E-26	tripartite motif-containing 2	TRIM2
dd_Smed_v6_1180_0_1	gi 37622892 ref NP_060346.2	103	2.44E-25	ring finger protein 126	RNF126
dd_Smed_v6_4623_0_1	gi 22027612 ref NP_066961.2	103	6.25E-25	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_2844_0_1	gi 37577175 ref NP_073618.3	100	6.59E-25	ring finger protein 38	RNF38
dd_Smed_v6_1439_0_1	gi 4759254 ref NP_004611.1	100	1.77E-24	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_2034_1_0_1	gi 4759254 ref NP_004611.1	100	2.34E-24	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_5507_0_1	gi 5031825 ref NP_005658.1	102	2.35E-24	ring finger protein 103	RNF103

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_1803_0_1	gi 32528299 ref NP_001158.2	100	6.76E-24	baculoviral IAP repeat-containing 4	BIRC4
dd_Smed_v6_3706_0_1	gi 22027616 ref NP_003291.2	98.6	1.71E-23	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_7122_0_1	gi 74027249 ref NP_056990.3	99.4	2.83E-23	tripartite motif-containing 33	TRIM33
dd_Smed_v6_5432_0_1	gi 53729361 ref NP_001005373.1	99	5.86E-23	leucine rich repeat and sterile alpha motif containing 1	LRSAM1
dd_Smed_v6_1934_0_1	gi 223278368 ref NP_004201.3	96.7	7.28E-23	neuralized homolog (Drosophila)	NEURL
dd_Smed_v6_5889_0_1	gi 4505225 ref NP_002422.1	91.3	1.17E-22	menage a trois homolog 1, cyclin H assembly factor	MNAT1
dd_Smed_v6_1254_8_0_1	gi 221139764 ref NP_065952.2	98.6	1.81E-22	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_3486_0_1	gi 5454168 ref NP_006453.1	88.6	8.64E-21	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_1110_0_0_1	gi 22027612 ref NP_066961.2	87.8	1.06E-20	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_3677_0_1	gi 4759254 ref NP_004611.1	89	1.31E-20	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_6768_0_1	gi 22027616 ref NP_003291.2	89.4	2.61E-20	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_1574_6_9	gi 22027616 ref NP_003291.2	89.7	5.30E-20	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_8109_0_1	gi 22027616 ref NP_003291.2	86.3	1.54E-19	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_1430_2_0_1	gi 22027612 ref NP_066961.2	84.3	1.64E-19	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_1022_0_0_1	gi 5454168 ref NP_006453.1	84.7	2.71E-19	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_4626_0_1	gi 32528299 ref NP_001158.2	84.7	2.91E-19	baculoviral IAP repeat-containing 4	BIRC4
dd_Smed_v6_9114_0_1	gi 30348954 ref NP_065825.1	84.3	6.22E-19	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_5514_0_1	gi 7662486 ref NP_055716.1	81.6	6.58E-19	ring finger protein 44	RNF44
dd_Smed_v6_1306_1_0_1	gi 22027612 ref NP_066961.2	83.2	1.18E-18	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_1431_8_0_1	gi 45594312 ref NP_115647.2	82.4	8.43E-18	TNF receptor-associated factor 7	TRAF7
dd_Smed_v6_1852_0_1	gi 37577175 ref NP_073618.3	81.6	8.61E-18	ring finger protein 38	RNF38
dd_Smed_v6_1074_1_0_2	gi 14149702 ref NP_056343.1	77.8	2.13E-17	ring finger protein 167	RNF167
dd_Smed_v6_6905_0_1	gi 109134327 ref NP_057209.3	77	2.74E-17	PTD016 protein	LOC51136
dd_Smed_v6_5782_0_1	gi 5454168 ref NP_006453.1	78.2	2.93E-17	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_5165_0_1	gi 205830432 ref NP_004638.2	80.9	3.15E-17	D4, zinc and double PHD fingers family 1	DPF1
dd_Smed_v6_2817_0_1	gi 4759254 ref NP_004611.1	77	7.17E-17	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_1564_2_0_1	gi 22027612 ref NP_066961.2	74.7	4.69E-16	TNF receptor-associated factor 2	TRAF2

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_5715_0_1	gi 4759254 ref NP_004611.1	74.3	8.95E-16	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_3878_0_1	gi 34452681 ref NP_055683.3	74.7	1.18E-15	ring finger protein 10	RNF10
dd_Smed_v6_6288_0_1	gi 22027612 ref NP_066961.2	73.2	1.51E-15	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_4880_0_1	gi 30348954 ref NP_065825.1	73.9	1.69E-15	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_3281_0_1	gi 187761373 ref NP_005735.2	74.7	1.75E-15	ariadne homolog	ARIH1
dd_Smed_v6_2026_0_1	gi 22027616 ref NP_003291.2	73.9	4.27E-15	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_6297_0_1	gi 5032071 ref NP_005776.1	67.8	8.52E-15	ring finger protein 41	RNF41
dd_Smed_v6_9512_0_1	gi 34878787 ref NP_071898.2	68.9	1.17E-14	ring finger protein 25	RNF25
dd_Smed_v6_15562_0_1	gi 40807469 ref NP_005870.2	69.3	1.54E-14	TRAF interacting protein	TRAIP
dd_Smed_v6_2712_3_9	gi 5454168 ref NP_006453.1	66.2	6.16E-14	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_14376_0_1	gi 14149702 ref NP_056343.1	64.7	1.14E-13	ring finger protein 167	RNF167
dd_Smed_v6_9478_0_1	gi 4759254 ref NP_004611.1	61.6	1.22E-13	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_81309_0_1	gi 14149702 ref NP_056343.1	61.6	1.49E-13	ring finger protein 167	RNF167
dd_Smed_v6_2015_0_1	gi 22027616 ref NP_003291.2	67.4	1.64E-13	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_5911_0_1	gi 30348954 ref NP_065825.1	65.5	9.78E-13	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_10019_0_1	gi 4504867 ref NP_003949.1	61.6	1.41E-12	ring finger protein 8	RNF8
dd_Smed_v6_2599_0_1	gi 77404348 ref NP_001029082.1	63.9	1.83E-12	TNF receptor-associated factor 5	TRAF5
dd_Smed_v6_20691_0_1	gi 44680139 ref NP_203127.3	57.8	2.03E-12	baculoviral IAP repeat-containing 8	BIRC8
dd_Smed_v6_8155_1_1	gi 32528299 ref NP_001158.2	62.8	2.24E-12	baculoviral IAP repeat-containing 4	BIRC4
dd_Smed_v6_3101_0_1	gi 157266328 ref NP_000456.2	60.8	2.43E-12	BRCA1 associated RING domain 1	BARD1
dd_Smed_v6_12678_0_1	gi 5454168 ref NP_006453.1	62.4	2.48E-12	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_27885_0_1	gi 4759254 ref NP_004611.1	58.9	2.92E-12	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_1308_0_1	gi 22027612 ref NP_066961.2	64.7	3.06E-12	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_27323_0_1	gi 5454168 ref NP_006453.1	62	3.42E-12	RanBP-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_11354_0_1	gi 4759254 ref NP_004611.1	63.2	3.85E-12	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_10569_0_1	gi 22027612 ref NP_066961.2	62.4	4.53E-12	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_9703_0_1	gi 4759254 ref NP_004611.1	62.4	6.36E-12	TNF receptor-associated factor 6	TRAF6

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_17046_0_1	gi 4759254 ref NP_004611.1	61.2	6.84E-12	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_19217_0_1	gi 77404348 ref NP_001029082.1	62	7.14E-12	TNF receptor-associated factor 5	TRAF5
dd_Smed_v6_45208_0_1	gi 40807469 ref NP_005870.2	57.8	8.56E-12	TRAF interacting protein	TRAIP
dd_Smed_v6_1395_0_1	gi 37595537 ref NP_079402.2	60.5	8.99E-12	ring finger protein 34	RNF34
dd_Smed_v6_14661_0_1	gi 4759254 ref NP_004611.1	61.2	1.14E-11	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_18178_0_1	gi 4759254 ref NP_004611.1	60.5	2.12E-11	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_5428_0_1	gi 47419909 ref NP_003843.3	62.4	3.83E-11	tripartite motif-containing 24	TRIM24
dd_Smed_v6_29967_0_1	gi 4759254 ref NP_004611.1	58.5	3.83E-11	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_5329_0_1	gi 4759254 ref NP_004611.1	58.9	3.84E-11	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_3437_0_1	gi 30348954 ref NP_065825.1	60.8	4.47E-11	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_3875_0_1	gi 22027616 ref NP_003291.2	56.6	8.26E-11	TNF receptor-associated factor 3	TRAF3
dd_Smed_v6_4827_0_1	gi 4759254 ref NP_004611.1	57.8	1.13E-10	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_5906_0_1	gi 34452684 ref NP_057204.2	55.8	1.22E-10	ring finger protein 12	RNF12
dd_Smed_v6_3416_0_1	gi 30348954 ref NP_065825.1	58.5	1.53E-10	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_3868_0_1	gi 37675277 ref NP_932351.1	55.8	1.59E-10	ring finger protein 148	RNF148
dd_Smed_v6_9026_0_1	gi 34452681 ref NP_055683.3	58.5	1.60E-10	ring finger protein 10	RNF10
dd_Smed_v6_37969_0_1	gi 221139764 ref NP_065952.2	50.8	2.36E-10	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_7519_0_1	gi 4759254 ref NP_004611.1	56.6	3.03E-10	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_42939_0_1	gi 40805104 ref NP_005793.2	51.2	3.14E-10	topoisomerase I binding, arginine/serine-rich	TOPORS
dd_Smed_v6_6290_0_1	gi 112421127 ref NP_056246.3	57.4	3.18E-10	tripartite motif-containing 58	TRIM58
dd_Smed_v6_1097_0_1	gi 157266328 ref NP_000456.2	58.9	3.25E-10	BRCA1 associated RING domain 1	BARD1
dd_Smed_v6_2039_0_1	gi 27436877 ref NP_775107.1	57	3.38E-10	tripartite motif-containing 59	TRIM59
dd_Smed_v6_4921_0_1	gi 157266328 ref NP_000456.2	55.5	3.43E-10	BRCA1 associated RING domain 1	BARD1
dd_Smed_v6_13757_0_1	gi 40805104 ref NP_005793.2	56.2	3.68E-10	topoisomerase I binding, arginine/serine-rich	TOPORS
dd_Smed_v6_13238_0_1	gi 33620769 ref NP_008841.2	50.1	1.11E-09	retinoblastoma binding protein 6	RBBP6
dd_Smed_v6_11059_0_1	gi 4759254 ref NP_004611.1	49.7	1.34E-09	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_3562_0_1	gi 4759254 ref NP_004611.1	54.7	1.39E-09	TNF receptor-associated factor 6	TRAF6

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_2453_0_1	gi 282394030 ref NP_543151.2	53.9	1.50E-09	mindbomb homolog 2 (Drosophila)	MIB2
dd_Smed_v6_19170_0_3	gi 4502139 ref NP_001156.1	48.5	1.67E-09	baculoviral IAP repeat-containing 3	BIRC3
dd_Smed_v6_5689_0_2	gi 24025688 ref NP_699202.1	51.6	1.73E-09	ligand of numb-protein X 2	LNx2
dd_Smed_v6_1243_0_3	gi 24025688 ref NP_699202.1	47.8	2.01E-09	ligand of numb-protein X 2	LNx2
dd_Smed_v6_28638_0_1	gi 22027612 ref NP_066961.2	53.5	2.14E-09	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_7660_0_1	gi 87241872 ref NP_777563.2	52.4	2.21E-09	ring finger protein 151	RNF151
dd_Smed_v6_7016_0_1	gi 30348954 ref NP_065825.1	55.5	2.38E-09	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_6964_0_3	gi 30348954 ref NP_065825.1	49.3	2.53E-09	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_10104_0_1	gi 282394030 ref NP_543151.2	47.8	2.70E-09	mindbomb homolog 2 (Drosophila)	MIB2
dd_Smed_v6_47022_0_1	gi 4759254 ref NP_004611.1	46.2	2.73E-09	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_9749_0_1	gi 24025688 ref NP_699202.1	46.6	2.83E-09	ligand of numb-protein X 2	LNx2
dd_Smed_v6_12718_0_1	gi 4759254 ref NP_004611.1	53.5	2.92E-09	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_27379_0_1	gi 30348954 ref NP_065825.1	47.4	3.26E-09	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_25112_0_1	gi 22027612 ref NP_066961.2	52.4	3.28E-09	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_10539_0_1	gi 22027624 ref NP_665694.1	50.1	4.16E-09	TNF receptor-associated factor 4	TRAF4
dd_Smed_v6_11860_0_1	gi 55953112 ref NP_001007279.1	53.1	4.35E-09	tripartite motif-containing 13	TRIM13
dd_Smed_v6_75625_0_1	gi 7662486 ref NP_055716.1	47	4.56E-09	ring finger protein 44	RNF44
dd_Smed_v6_12171_0_2	gi 57529737 ref NP_055824.1	48.9	4.83E-09	PDZ domain containing RING finger 3	PDZRN3
dd_Smed_v6_6852_0_1	gi 24025688 ref NP_699202.1	49.7	5.15E-09	ligand of numb-protein X 2	LNx2
dd_Smed_v6_30148_0_1	gi 221139764 ref NP_065952.2	46.2	5.21E-09	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_6556_0_1	gi 33620769 ref NP_008841.2	53.5	5.72E-09	retinoblastoma binding protein 6	RBBP6
dd_Smed_v6_5871_0_1	gi 282394030 ref NP_543151.2	52.4	9.20E-09	mindbomb homolog 2 (Drosophila)	MIB2
dd_Smed_v6_14842_0_1	gi 22027624 ref NP_665694.1	49.3	9.61E-09	TNF receptor-associated factor 4	TRAF4
dd_Smed_v6_1271_0_1	gi 58743365 ref NP_443148.1	53.5	1.08E-08	ring finger protein 157	RNF157
dd_Smed_v6_645_0_1	gi 282394030 ref NP_543151.2	53.1	1.21E-08	mindbomb homolog 2 (Drosophila)	MIB2
dd_Smed_v6_79682_0_1	gi 221139764 ref NP_065952.2	44.7	1.30E-08	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_18055_0_1	gi 40805104 ref NP_005793.2	51.2	1.34E-08	topoisomerase I binding, arginine/serine-rich	TOPORS

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_4529_0_1	gi 27436925 ref NP_115807.1	51.2	1.39E-08	synovial apoptosis inhibitor 1, synoviolin	SYVN1
dd_Smed_v6_3468_0_1	gi 7657508 ref NP_055063.1	47.4	1.41E-08	ring-box 1	RBX1
dd_Smed_v6_593_1_15	gi 282394030 ref NP_54315.1.2	49.3	1.44E-08	mindbomb homolog 2 (Drosophila)	MIB2
dd_Smed_v6_12016_0_1	gi 40805104 ref NP_005793.2	52	1.45E-08	topoisomerase I binding, arginine/serine-rich	TOPORS
dd_Smed_v6_72617_0_1	gi 57165361 ref NP_689683.2	43.5	1.54E-08	ring finger protein 165	RNF165
dd_Smed_v6_7801_0_1	sp O94941 ref NP_055763	50.4	1.72E-08	U-box domain containing 5	UBOX5
dd_Smed_v6_15776_0_1	gi 221139764 ref NP_06595.2.2	47.4	1.96E-08	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_5723_0_1	gi 24025688 ref NP_699202.1	44.3	2.03E-08	ligand of numb-protein X 2	LNX2
dd_Smed_v6_10411_0_1	gi 4505715 ref NP_002608.1	50.1	2.27E-08	peroxisome biogenesis factor 10	PEX10
dd_Smed_v6_48453_0_1	gi 21389515 ref NP_653327.1	50.4	2.44E-08	ring finger protein 145	RNF145
dd_Smed_v6_26821_0_1	gi 29788758 ref NP_060904.2	44.3	2.73E-08	ring finger protein 130	RNF130
dd_Smed_v6_12211_0_1	gi 194248079 ref NP_05608.6.2	50.8	2.79E-08	tripartite motif-containing 2	TRIM2
dd_Smed_v6_57645_0_1	gi 37675277 ref NP_932351.1	44.3	3.38E-08	ring finger protein 148	RNF148
dd_Smed_v6_16504_0_1	gi 11545910 ref NP_071444.1	47.4	3.81E-08	baculoviral IAP repeat-containing 7 (livin)	BIRC7
dd_Smed_v6_7538_0_1	gi 30348954 ref NP_065825.1	49.3	3.86E-08	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_10917_0_1	gi 4759254 ref NP_004611.1	49.7	5.12E-08	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_2867_0_1	gi 282394030 ref NP_54315.1.2	49.7	5.46E-08	mindbomb homolog 2 (Drosophila)	MIB2
dd_Smed_v6_3693_0_1	gi 282394030 ref NP_54315.1.2	50.4	8.15E-08	mindbomb homolog 2 (Drosophila)	MIB2
dd_Smed_v6_88252_0_1	gi 7662486 ref NP_055716.1	41.6	9.42E-08	ring finger protein 44	RNF44
dd_Smed_v6_7124_0_1	gi 5032071 ref NP_005776.1	47.8	1.03E-07	ring finger protein 41	RNF41
dd_Smed_v6_490_0_20	gi 24025688 ref NP_699202.1	43.1	1.07E-07	ligand of numb-protein X 2	LNX2
dd_Smed_v6_56572_0_1	gi 51479192 ref NP_002922.2	42.7	1.09E-07	ring finger protein 1	RING1
dd_Smed_v6_5841_0_1	gi 30348954 ref NP_065825.1	48.5	1.11E-07	mindbomb homolog 1 (Drosophila)	MIB1
dd_Smed_v6_84672_0_1	gi 221139764 ref NP_06595.2.2	44.3	1.19E-07	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_8691_0_1	gi 221139764 ref NP_06595.2.2	47	1.25E-07	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_1187_0_1	gi 4502141 ref NP_001157.1	47.4	1.78E-07	baculoviral IAP repeat-containing 2	BIRC2
dd_Smed_v6_94197_0_1	gi 40805104 ref NP_005793.2	42.4	1.98E-07	topoisomerase I binding, arginine/serine-rich	TOPORS

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_14467_0_1	gi 57529737 ref NP_055824.1	48.1	2.16E-07	PDZ domain containing RING finger 3	PDZRN3
dd_Smed_v6_91490_0_1	gi 4504867 ref NP_003949.1	40.8	4.20E-07	ring finger protein 8	RNF8
dd_Smed_v6_2170_0_1	gi 22027612 ref NP_066961.2	47	4.23E-07	TNF receptor-associated factor 2	TRAF2
dd_Smed_v6_9146_0_1	gi 32528299 ref NP_001158.2	45.1	5.79E-07	baculoviral IAP repeat-containing 4	BIRC4
dd_Smed_v6_5665_0_1	gi 4759254 ref NP_004611.1	46.6	6.44E-07	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_1137_0_5	gi 4504867 ref NP_003949.1	46.6	6.46E-07	ring finger protein 8	RNF8
dd_Smed_v6_6465_0_1	gi 38679905 ref NP_775818.2	46.6	6.85E-07	tripartite motif-containing 65	TRIM65
dd_Smed_v6_97741_0_1	gi 55749557 ref NP_001006611.1	40	9.00E-07	seven in absentia homolog 1 (Drosophila)	SIAH1
dd_Smed_v6_21033_0_1	gi 221139764 ref NP_065952.2	43.9	1.12E-06	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_5673_0_1	gi 4759254 ref NP_004611.1	45.8	1.23E-06	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_23413_0_1	gi 73747840 ref NP_001027026.1	45.1	1.26E-06	LON peptidase N-terminal domain and ring finger 3	LONRF3
dd_Smed_v6_4082_0_1	gi 4502141 ref NP_001157.1	43.1	2.26E-06	baculoviral IAP repeat-containing 2	BIRC2
dd_Smed_v6_23201_0_1	gi 4502141 ref NP_001157.1	43.5	2.28E-06	baculoviral IAP repeat-containing 2	BIRC2
dd_Smed_v6_2104_0_1	gi 4502141 ref NP_001157.1	42.7	2.33E-06	baculoviral IAP repeat-containing 2	BIRC2
dd_Smed_v6_2937_0_1	gi 4759254 ref NP_004611.1	43.5	2.37E-06	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_17728_0_1	gi 4502139 ref NP_001156.1	43.9	2.44E-06	baculoviral IAP repeat-containing 3	BIRC3
dd_Smed_v6_4528_0_1	gi 4759254 ref NP_004611.1	45.4	2.68E-06	TNF receptor-associated factor 6	TRAF6
dd_Smed_v6_44806_0_1	gi 77404348 ref NP_001029082.1	37.7	2.70E-06	TNF receptor-associated factor 5	TRAF5
dd_Smed_v6_84492_0_1	gi 11545910 ref NP_071444.1	37.4	3.05E-06	baculoviral IAP repeat-containing 7 (livin)	BIRC7
dd_Smed_v6_83991_0_1	gi 4505715 ref NP_002608.1	37.4	3.08E-06	peroxisome biogenesis factor 10	PEX10
dd_Smed_v6_17389_0_1	gi 221139764 ref NP_065952.2	42.4	9.45E-06	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_52018_0_1	gi 221139764 ref NP_065952.2	39.7	1.07E-05	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_19551_0_1	gi 87241872 ref NP_777563.2	41.2	1.40E-05	ring finger protein 151	RNF151
dd_Smed_v6_67480_0_1	gi 284447287 ref NP_775918.2	38.5	1.45E-05	ring finger protein 149	RNF149
dd_Smed_v6_47736_0_1	gi 4502141 ref NP_001157.1	40	1.58E-05	baculoviral IAP repeat-containing 2	BIRC2
dd_Smed_v6_15545_0_1	gi 24025688 ref NP_699202.1	42	1.77E-05	ligand of numb-protein X 2	LNX2
dd_Smed_v6_17342_0_1	gi 51988887 ref NP_065921.2	39.7	1.79E-05	SH3 domain containing ring finger 1	SH3RF1

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 1: Planarian RING_UBOX blast to human RING

Smed contig ID	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_42127_0_1	gi 21314654 ref NP_009149.2	37.4	3.69E-05	ring finger protein 139	RNF139
dd_Smed_v6_35439_0_1	gi 21071052 ref NP_003062.2	35	1.85E-04	helicase-like transcription factor	HLTF
dd_Smed_v6_73473_0_1	gi 221139764 ref NP_065952.2	35.4	2.23E-04	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_28174_0_1	gi 149408115 ref NP_075564.3	35.4	3.17E-04	chromosome 16 open reading frame 28	C16orf28
dd_Smed_v6_4845_0_1	gi 24307991 ref NP_055904.1	38.5	3.59E-04	p53-associated parkin-like cytoplasmic protein	PARC
dd_Smed_v6_18067_0_1	gi 24025688 ref NP_699202.1	32	4.15E-04	ligand of numb-protein X 2	LNX2
dd_Smed_v6_11314_0_1	gi 58743365 ref NP_443148.1	36.6	4.42E-04	ring finger protein 157	RNF157
dd_Smed_v6_14096_0_1	gi 221139764 ref NP_065952.2	35.8	8.98E-04	CTD-binding SR-like protein rA9	KIAA1542
dd_Smed_v6_18686_0_1	gi 24025688 ref NP_699202.1	30.8	0.001	ligand of numb-protein X 2	LNX2
dd_Smed_v6_12817_0_1					
dd_Smed_v6_13414_0_1					
dd_Smed_v6_1692_0_1					
dd_Smed_v6_22203_0_1					
dd_Smed_v6_24063_0_1					
dd_Smed_v6_2712_2_1					
dd_Smed_v6_33403_0_1					
dd_Smed_v6_38544_0_1					
dd_Smed_v6_40805_0_1					
dd_Smed_v6_41377_0_1					
dd_Smed_v6_4972_0_1					
dd_Smed_v6_50305_0_1					
dd_Smed_v6_5961_0_1					
dd_Smed_v6_5986_0_1					
dd_Smed_v6_7071_0_1					
dd_Smed_v6_76286_0_1					
dd_Smed_v6_7796_0_1					

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 2: Additional IPR 013083 Contigs

Smed contig	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_4584_0_1	gi 33589846 ref NP_005876.2	608	0	E3 ubiquitin-protein ligase MARCH6	MARCH6
dd_Smed_v6_8386_0_1	gi 190341104 ref NP_055978.4	474	2.01E-157	tripartite motif containing 9	TRIM9
dd_Smed_v6_4755_0_1	gi 19913361 ref NP_579891.1	455	2.15E-149	Wolf-Hirschhorn syndrome candidate 1 protein isoform 5	
dd_Smed_v6_12595_0_1	gi 190341104 ref NP_055978.4	396	2.79E-127	tripartite motif containing 9	TRIM9
dd_Smed_v6_4550_0_1	gi 17978485 ref NP_065908.1	384	9.29E-120	VPS18 core subunit of CORVET and HOPS complexes	VPS18
dd_Smed_v6_4236_0_1	gi 27597061 ref NP_056070.1	394	2.51E-113	ubiquitin protein ligase E3 component n-recogin 2	UBR2
dd_Smed_v6_11074_0_1	gi 19913361 ref NP_579891.1	343	7.04E-108	Wolf-Hirschhorn syndrome candidate 1 protein isoform 5	
dd_Smed_v6_4602_0_1	gi 8923415 ref NP_060294.1	282	1.39E-94	membrane associated ring-CH-type finger 5	MARCH5
dd_Smed_v6_10189_0_1	gi 19913361 ref NP_579891.1	266	1.11E-79	Wolf-Hirschhorn syndrome candidate 1 protein isoform 5	
dd_Smed_v6_3837_0_1	gi 22027612 ref NP_066961.2	233	3.21E-71	TNF receptor associated factor 2	TRAF2
dd_Smed_v6_4997_0_1	gi 8923613 ref NP_060393.1	202	4.85E-63	membrane associated ring-CH-type finger 1	MARCH1
dd_Smed_v6_1991_0_1	gi 22027616 ref NP_003291.2	199	4.61E-59	TNF receptor associated factor 3	TRAF3
dd_Smed_v6_8933_0_1	gi 22027612 ref NP_066961.2	178	2.38E-49	TNF receptor associated factor 2	TRAF2
dd_Smed_v6_5825_0_1	gi 205830432 ref NP_004638.2	163	3.10E-47	double PHD fingers 1	DPF1
dd_Smed_v6_9815_0_1	gi 19913361 ref NP_579891.1	156	2.27E-41	Wolf-Hirschhorn syndrome candidate 1 protein isoform 5	
dd_Smed_v6_6521_0_1	gi 19913361 ref NP_579891.1	154	1.44E-40	Wolf-Hirschhorn syndrome candidate 1 protein isoform 5	
dd_Smed_v6_6787_0_1	gi 5454168 ref NP_006453.1	137	1.81E-40	RANBP2-type and C3HC4-type zinc finger containing 1	RBCK1
dd_Smed_v6_2916_0_1	gi 22027612 ref NP_066961.2	122	2.11E-32	TNF receptor associated factor 2	TRAF2
dd_Smed_v6_7110_0_1	gi 22749145 ref NP_689766.1	117	2.64E-31	ring finger protein 217	RNF217
dd_Smed_v6_4420_0_1	gi 22027612 ref NP_066961.2	109	3.09E-27	TNF receptor associated factor 2	TRAF2
dd_Smed_v6_4893_0_1	gi 22027616 ref NP_003291.2	110	1.01E-26	TNF receptor associated factor 3	TRAF3
dd_Smed_v6_8166_0_1	gi 4505721 ref NP_000277.1	103	2.34E-26	peroxisomal biogenesis factor 12	PEX12
dd_Smed_v6_5797_0_1	gi 205830432 ref NP_004638.2	107	3.14E-26	double PHD fingers 1	DPF1
dd_Smed_v6_2331_0_1	gi 20336207 ref NP_612115.1	112	3.23E-26	transcriptional regulator ATRX isoform 3	
dd_Smed_v6_9090_0_1	gi 20336207 ref NP_612115.1	102	2.39E-23	transcriptional regulator ATRX isoform 3	
dd_Smed_v6_6332_0_1	gi 20336207 ref NP_612115.1	95.9	5.79E-22	transcriptional regulator ATRX isoform 3	
dd_Smed_v6_5740_0_1	gi 22027612 ref NP_066961.2	92.8	3.95E-21	TNF receptor associated factor 2	TRAF2

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 2: Additional IPR 013083 Contigs

Smed contig	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_5015_7_0_1	gi 30425370 ref NP_848545.1	82.8	1.05E-20	membrane associated ring-CH-type finger 3	MARCH3
dd_Smed_v6_5804_0_1	gi 205830432 ref NP_004638.2	86.7	3.62E-20	double PHD fingers 1	DPF1
dd_Smed_v6_129_2_1	gi 22027612 ref NP_066961.2	87	3.65E-19	TNF receptor associated factor 2	TRAF2
dd_Smed_v6_2080_1_0_1	gi 77404348 ref NP_001029082.1	74.7	2.81E-17	TNF receptor associated factor 5	TRAF5
dd_Smed_v6_6283_0_1	gi 19913361 ref NP_579891.1	80.5	1.30E-16	Wolf-Hirschhorn syndrome candidate 1 protein isoform 5	
dd_Smed_v6_7950_0_1	gi 22027616 ref NP_003291.2	76.6	1.74E-16	TNF receptor associated factor 3	TRAF3
dd_Smed_v6_1344_6_0_1	gi 30348954 ref NP_065825.1	76.3	1.46E-15	mindbomb E3 ubiquitin protein ligase 1	MIB1
dd_Smed_v6_3650_0_1	gi 77404348 ref NP_001029082.1	72.4	3.18E-15	TNF receptor associated factor 5	TRAF5
dd_Smed_v6_2171_3_0_1	gi 30425370 ref NP_848545.1	67.4	6.43E-15	membrane associated ring-CH-type finger 3	MARCH3
dd_Smed_v6_1459_1_0_1	gi 77404348 ref NP_001029082.1	52	1.38E-14	TNF receptor associated factor 5	TRAF5
dd_Smed_v6_3038_0_1	gi 4759254 ref NP_004611.1	64.3	1.08E-12	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_1124_5_0_1	gi 205830432 ref NP_004638.2	63.2	6.57E-12	double PHD fingers 1	DPF1
dd_Smed_v6_1461_4_0_1	gi 4759254 ref NP_004611.1	61.2	8.67E-12	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_7654_0_1	gi 12383066 ref NP_073737.1	63.5	1.73E-11	membrane associated ring-CH-type finger 7	MARCH7
dd_Smed_v6_6054_0_1	gi 4759254 ref NP_004611.1	60.8	1.91E-11	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_1031_1_0_1	gi 47419909 ref NP_003843.3	62.8	2.12E-11	tripartite motif containing 24	TRIM24
dd_Smed_v6_8604_0_1	gi 47419909 ref NP_003843.3	60.5	9.43E-11	tripartite motif containing 24	TRIM24
dd_Smed_v6_7394_0_1	gi 221139764 ref NP_065952.2	61.2	1.31E-10	PHD and ring finger domains 1	PHRF1
dd_Smed_v6_3334_0_1	gi 221139764 ref NP_065952.2	59.7	2.68E-10	PHD and ring finger domains 1	PHRF1
dd_Smed_v6_3984_0_1	gi 221139764 ref NP_065952.2	57.4	7.67E-10	PHD and ring finger domains 1	PHRF1
dd_Smed_v6_1997_0_1	gi 188497705 ref NP_006759.3	56.6	1.24E-09	BRCA1 associated protein	BRAP
dd_Smed_v6_1113_0_0_1	gi 221139764 ref NP_065952.2	54.3	5.85E-09	PHD and ring finger domains 1	PHRF1
dd_Smed_v6_8004_0_1	gi 4759254 ref NP_004611.1	52.4	8.53E-09	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_2391_0_1	gi 4759254 ref NP_004611.1	53.1	1.32E-08	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_2071_3_0_1	gi 4502139 ref NP_001156.1	45.1	2.17E-08	baculoviral IAP repeat containing 3	BIRC3
dd_Smed_v6_3775_0_1	gi 188497705 ref NP_006759.3	51.2	3.25E-08	BRCA1 associated protein	BRAP
dd_Smed_v6_2277_0_1	gi 4759254 ref NP_004611.1	48.9	4.81E-08	TNF receptor associated factor 6	TRAF6

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 2: Additional IPR 013083 Contigs

Smed contig	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_15039_0_1	gi 4759254 ref NP_004611.1	49.3	7.06E-08	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_8799_0_1	gi 47419909 ref NP_003843.3	50.1	9.72E-08	tripartite motif containing 24	TRIM24
dd_Smed_v6_3468_1_1	gi 50409810 ref NP_001002244.1	41.6	9.88E-08	anaphase promoting complex subunit 11	ANAPC11
dd_Smed_v6_13174_0_1	gi 22027612 ref NP_066961.2	47.4	2.98E-07	TNF receptor associated factor 2	TRAF2
dd_Smed_v6_4933_0_1	gi 221139764 ref NP_065952.2	49.7	3.85E-07	PHD and ring finger domains 1	PHRF1
dd_Smed_v6_17286_0_1	gi 4759254 ref NP_004611.1	46.6	3.91E-07	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_94260_0_1	gi 77404348 ref NP_001029082.1	42.4	4.37E-07	TNF receptor associated factor 5	TRAF5
dd_Smed_v6_53188_0_1	gi 22027616 ref NP_003291.2	39.7	4.70E-07	TNF receptor associated factor 3	TRAF3
dd_Smed_v6_14617_0_1	gi 77404348 ref NP_001029082.1	45.4	5.64E-07	TNF receptor associated factor 5	TRAF5
dd_Smed_v6_55800_0_1	gi 4759254 ref NP_004611.1	40	6.22E-07	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_18711_0_1	gi 22027616 ref NP_003291.2	40	6.90E-07	TNF receptor associated factor 3	TRAF3
dd_Smed_v6_12667_0_1	gi 44680139 ref NP_203127.3	42.7	9.18E-07	baculoviral IAP repeat containing 8	BIRC8
dd_Smed_v6_10803_0_1	gi 4502139 ref NP_001156.1	44.3	9.48E-07	baculoviral IAP repeat containing 3	BIRC3
dd_Smed_v6_1163_0_1	gi 4759254 ref NP_004611.1	45.4	1.07E-06	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_4609_0_1	gi 37588869 ref NP_071347.2	46.2	1.22E-06	ring finger protein 123	RNF123
dd_Smed_v6_4379_0_1	gi 82659109 ref NP_065816.2	44.3	1.99E-06	ubiquitin protein ligase E3 component n-recognin 4	UBR4
dd_Smed_v6_9528_0_1	gi 21361543 ref NP_057567.3	45.4	2.44E-06	PHD finger protein 7	PHF7
dd_Smed_v6_16468_0_1	gi 22027612 ref NP_066961.2	43.5	2.86E-06	TNF receptor associated factor 2	TRAF2
dd_Smed_v6_67936_0_1	gi 29788758 ref NP_060904.2	38.5	2.96E-06	ring finger protein 130	RNF130
dd_Smed_v6_4338_0_1	gi 44680139 ref NP_203127.3	42.7	2.99E-06	baculoviral IAP repeat containing 8	BIRC8
dd_Smed_v6_8447_0_1	gi 4502139 ref NP_001156.1	42.7	3.15E-06	baculoviral IAP repeat containing 3	BIRC3
dd_Smed_v6_7881_0_1	gi 19913361 ref NP_579891.1	44.7	3.81E-06	Wolf-Hirschhorn syndrome candidate 1 protein isoform 5	
dd_Smed_v6_18434_0_1	gi 4502141 ref NP_001157.1	39.3	5.03E-06	baculoviral IAP repeat containing 2	BIRC2
dd_Smed_v6_7294_0_2	gi 115430235 ref NP_001041666.1	43.9	6.94E-06	ubiquitin like with PHD and ring finger domains 1	UHRF1
dd_Smed_v6_9022_0_1	gi 6552299 ref NP_009225.1	44.3	1.29E-05	BRCA1 DNA repair associated	BRCA1
dd_Smed_v6_16183_0_1	gi 57529737 ref NP_055824.1	41.6	1.48E-05	PDZ domain containing ring finger 3	PDZRN3
dd_Smed_v6_66336_0_1	gi 221139764 ref NP_065952.2	39.7	1.99E-05	PHD and ring finger domains 1	PHRF1

Supplemental Table S2.1. List of planarian contigs with RING or U-box domain and top BLAST hit to human E3 ligases. Continued.

Tab 2: Additional IPR 013083 Contigs

Smed contig	Human Blast Hit ID	Bitscore	Evalue	Gene Name	Gene Symbol
dd_Smed_v6_1333_1_0_1	gi 148528975 ref NP_940863.3	35.8	2.16E-05	LON peptidase N-terminal domain and ring finger 2	LONRF2
dd_Smed_v6_1090_6_0_1	gi 22027616 ref NP_003291.2	32.3	2.18E-05	TNF receptor associated factor 3	TRAF3
dd_Smed_v6_3655_0_1	gi 188497705 ref NP_006759.3	40.8	2.29E-05	BRCA1 associated protein	BRAP
dd_Smed_v6_1948_7_0_1	gi 24025688 ref NP_699202.1	34.3	3.84E-05	ligand of numb-protein X 2	LNK2
dd_Smed_v6_8460_1_0_1	gi 21630277 ref NP_660215.1	35	4.00E-05	tripartite motif containing 11	TRIM11
dd_Smed_v6_4496_0_0_1	gi 14042925 ref NP_114404.1	34.3	7.43E-05	ring finger protein 26	RNF26
dd_Smed_v6_3398_0_1	gi 45594312 ref NP_115647.2	42.4	9.52E-05	TNF receptor associated factor 7	TRAF7
dd_Smed_v6_5662_0_1	gi 19913361 ref NP_579891.1	38.1	9.86E-05	Wolf-Hirschhorn syndrome candidate 1 protein isoform 5	
dd_Smed_v6_2321_9_0_1	gi 54792146 ref NP_258411.2	33.5	1.36E-04	tripartite motif containing 47	TRIM47
dd_Smed_v6_1604_6_0_4	gi 4759254 ref NP_004611.1	35.8	1.64E-04	TNF receptor associated factor 6	TRAF6
dd_Smed_v6_1277_0_0_1	gi 37588869 ref NP_071347.2	36.2	3.31E-04	ring finger protein 123	RNF123
dd_Smed_v6_4144_0_1	gi 57529737 ref NP_055824.1	33.1	3.90E-04	PDZ domain containing ring finger 3	PDZRN3
dd_Smed_v6_1018_5_0_1	gi 6005747 ref NP_009143.1	37	7.39E-04	ring finger protein 2	RNF2
dd_Smed_v6_2491_0_0_5	gi 134288906 ref NP_001004342.3	32	7.90E-04	tripartite motif containing 67	TRIM67
dd_Smed_v6_1226_1_0_1	gi 209180481 ref NP_079090.2	37	0.001	Cbl proto-oncogene like 1	CBLL1

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 .

Tab 1: Significantly differentially expressed genes after 11 days of *phc(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_12731_0_1	0.95	4.73E-16	E3 SUMO-protein ligase CBX4	O00257.3	6.00E-09
dd_Smed_v6_3632_0_1	0.84	1.62E-10			
dd_Smed_v6_9909_0_1	0.75	1.84E-06			
dd_Smed_v6_2394_0_1	0.58	0.002957566	Cytochrome P450 2J2	P51589.2	1.00E-36
dd_Smed_v6_15545_0_1	0.56	0.001631982	Ligand of Numb protein X 2	Q8N448.1	0.001
dd_Smed_v6_1692_0_5	0.51	0.007081075	28S ribosomal protein S5, mitochondrial	P82675.2	8.00E-45
dd_Smed_v6_42763_0_1	0.50	6.86E-05	Nuclear receptor ROR-alpha	P35398.2	7.00E-24
dd_Smed_v6_385_0_2	0.50	0.043329257			
dd_Smed_v6_11245_0_2	0.49	0.010360497	Histone acetyltransferase KAT6A	Q92794.2	2.00E-17
dd_Smed_v6_9559_0_2	0.44	0.017636534	WD repeat-containing protein 81	Q562E7.2	3.00E-68
dd_Smed_v6_8078_0_1	0.44	0.030639653	Eukaryotic translation initiation factor 4 gamma 1	Q04637.4	8.00E-32
dd_Smed_v6_1636_0_1	0.43	0.003994618	Zonadhesin	Q9Y493.5	1.00E-33
dd_Smed_v6_2869_0_1	0.42	0.07832413	Intercellular adhesion molecule 5	Q9UMF0.3	0.006
dd_Smed_v6_12353_0_1	0.41	0.087471986			
dd_Smed_v6_416_0_1	0.40	0.002957566			
dd_Smed_v6_1571_0_1	0.40	0.087925756			
dd_Smed_v6_1692_0_2	0.39	0.001319671	28S ribosomal protein S5, mitochondrial	P82675.2	8.00E-45
dd_Smed_v6_5368_0_13	0.31	0.083753657	Serine/threonine-protein kinase 3	Q13188.2	2.00E-17
dd_Smed_v6_636_0_1	0.30	0.00034698	von Willebrand factor	P04275.4	3.00E-07
dd_Smed_v6_5865_0_1	0.28	0.083087152	Structural maintenance of chromosomes protein 4	Q9NTJ3.2	1.00E-107
dd_Smed_v6_5235_1_1	0.27	0.026777116			
dd_Smed_v6_7877_0_2	0.26	0.000144449	Hepatocyte nuclear factor 6	Q9UBC0.1	4.00E-71
dd_Smed_v6_11691_0_16	0.25	0.006920318			
dd_Smed_v6_16566_0_2	0.14	0.004607497	Protocadherin gamma-A7	Q9Y5G6.1	1.00E-69
dd_Smed_v6_11691_0_20	0.14	0.038185236			
dd_Smed_v6_56732_0_1	0.13	0.064328855			
dd_Smed_v6_3286_0_23	0.12	0.059585366			
dd_Smed_v6_12323_0_4	0.12	0.07832413	Copine-9	Q8IYJ1.3	1.00E-46
dd_Smed_v6_10407_0_28	0.09	0.023440614			
dd_Smed_v6_14400_0_3	-0.11	0.049532048	Rho GTPase-activating protein 20	Q9P2F6.2	2.00E-07
dd_Smed_v6_2822_0_17	-0.12	0.000232145	Aurora kinase C	Q9UQB9.1	4.00E-27
dd_Smed_v6_6366_0_3	-0.13	0.044272715	Beta-1,3-glucosyltransferase	Q6Y288.2	2.00E-54
dd_Smed_v6_18182_0_1	-0.14	0.044272715			
dd_Smed_v6_20393_0_2	-0.14	0.013877322			

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 1: Significantly differentially expressed genes after 11 days of *phc(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_8834_0_1	-0.16	0.011279 334	Squamous cell carcinoma antigen recognized by T-cells 3	Q15020.1	0.017
dd_Smed_v6_8062_0_14	-0.17	0.007081 075	Acyl-CoA synthetase short-chain family member 3	Q9H6R3.1	1.00E-11
dd_Smed_v6_8268_0_2	-0.17	0.012072 179			
dd_Smed_v6_13044_0_5	-0.31	0.005545 916			
dd_Smed_v6_12353_0_3	-0.32	0.000112 255			
dd_Smed_v6_12353_0_4	-0.34	3.15E-05			
dd_Smed_v6_22230_0_2	-0.37	3.07E-05			
dd_Smed_v6_1760_0_1	-0.37	0.035307 152	Transformer-2 protein homolog alpha	Q13595.1	2.00E-17
dd_Smed_v6_6686_0_3	-0.38	2.96E-05	Dynein assembly factor 3, axonemal	Q8N9W5.4	1.00E-84
dd_Smed_v6_2072_0_32	-0.40	0.017926 325	Ankyrin-3	Q12955.3	0
dd_Smed_v6_10946_0_2	-0.46	0.083753 657	Pikachurin	Q63HQ2.2	1.00E-48
dd_Smed_v6_5620_0_8	-0.47	0.091657 674	Tensin-1	Q9HBL0.2	5.00E-37
dd_Smed_v6_5144_0_3	-0.55	0.007081 075	Endosome/lysosome-associated apoptosis and autophagy regulator family member 2	A8MWY0.2	3.00E-126
dd_Smed_v6_8936_0_3	-0.59	5.44E-06			
dd_Smed_v6_8875_0_1	-1.34	4.93E-28	Polyhomeotic-like protein 3	Q8NDX5.1	7.00E-14

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 2: Significantly differentially expressed genes after 14 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_8989_0_1	-1.53	5.29E-27	E3 ubiquitin-protein ligase RING2	Q99496.1	6.00E-46
dd_Smed_v6_2184_0_1	-0.88	3.96E-06			
dd_Smed_v6_9699_0_2	-0.80	9.30E-06	Band 3 anion transport protein	P02730.3	6.00E-45
dd_Smed_v6_10147_0_5	-0.76	2.81E-08	Monocarboxylate transporter 14	Q7RTX9.1	7.00E-17
dd_Smed_v6_3066_0_1	-0.57	8.30E-08	Transmembrane protein 41A	Q96HV5.1	2.00E-66
dd_Smed_v6_3286_0_12	-0.52	9.30E-06			
dd_Smed_v6_583_0_1	-0.52	2.15E-09			
dd_Smed_v6_777_0_1	-0.49	2.65E-05			
dd_Smed_v6_1127_0_1	-0.48	0.046195564	Glycine N-methyltransferase	Q14749.3	2.00E-116
dd_Smed_v6_2483_0_1	-0.30	0.000133265			
dd_Smed_v6_9205_0_6	-0.24	0.013554565	Guanine nucleotide exchange factor subunit RIC1	Q4ADV7.2	9.00E-159
dd_Smed_v6_10186_0_3	-0.24	0.029148468	Bardet-Biedl syndrome 2 protein	Q9BXC9.1	6.00E-171
dd_Smed_v6_6500_0_5	-0.23	0.001001943	DNA polymerase eta	Q9Y253.1	8.00E-52
dd_Smed_v6_12745_0_3	-0.19	0.014304774	RNA pseudouridylylase domain-containing protein 2	Q8IZ73.2	1.00E-63
dd_Smed_v6_6366_0_8	-0.18	0.080357031	Beta-1,3-glucosyltransferase	Q6Y288.2	6.00E-30
dd_Smed_v6_7997_1_5	-0.16	0.016522882			
dd_Smed_v6_16660_0_2	0.21	0.021373083	Jerky protein homolog-like	Q9Y4A0.2	0.002
dd_Smed_v6_15487_0_2	0.29	0.001117104			
dd_Smed_v6_758_1_1	0.48	0.050075382	Endoplasmic	P14625.1	0
dd_Smed_v6_7448_0_1	0.63	0.028916557	Mitofusin-2	O95140.3	2.00E-15
dd_Smed_v6_2080_0_1	0.65	2.32E-05	Mitofusin-2	O95140.3	8.00E-23
dd_Smed_v6_79_0_1	0.66	0.003011286	Pulmonary surfactant-associated protein A1	Q8IWL2.2	1.00E-06
dd_Smed_v6_58_0_1	0.67	0.003543346	Collectin-11	Q9BWP8.1	7.00E-12
dd_Smed_v6_6390_0_2	0.67	9.01E-05	Tubulin polyglutamylase TTLL5	Q6EMB2.3	2.00E-157
dd_Smed_v6_909_0_1	0.71	0.000363279			
dd_Smed_v6_2777_0_5	0.76	0.000280758			
dd_Smed_v6_10791_0_1	0.83	2.81E-08	FAD-dependent oxidoreductase domain-containing protein 1	Q96CU9.2	5.00E-107
dd_Smed_v6_297_0_1	0.93	8.72E-13	Cytochrome c oxidase subunit 2	P00403.1	1.00E-30
dd_Smed_v6_10_1_1	1.03	1.73E-09	Collectin-10	Q9Y6Z7.2	1.00E-07

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 3: Significantly differentially expressed genes after 28 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_79_0_1	1.92	6.10E-29	Pulmonary surfactant-associated protein A1	Q8IWL2.2	1.00E-06
dd_Smed_v6_58_0_1	1.78	3.79E-26	Collectin-11	Q9BWP8.1	7.00E-12
dd_Smed_v6_10_1_1	1.47	4.70E-19	Collectin-10	Q9Y6Z7.2	1.00E-07
dd_Smed_v6_2777_0_5	1.38	8.85E-19			
dd_Smed_v6_33_0_1	1.21	5.15E-11	Pulmonary surfactant-associated protein A2	Q8IWL1.1	4.00E-06
dd_Smed_v6_7448_0_1	1.03	4.44E-08	Mitofusin-2	O95140.3	2.00E-15
dd_Smed_v6_297_0_1	0.97	2.08E-12	Cytochrome c oxidase subunit 2	P00403.1	1.00E-30
dd_Smed_v6_606_0_3	0.96	5.32E-06	Alpha-N-acetylgalactosaminidase	P17050.2	1.00E-75
dd_Smed_v6_2080_0_1	0.95	3.73E-11	Mitofusin-2	O95140.3	8.00E-23
dd_Smed_v6_8_0_2	0.91	4.25E-05			
dd_Smed_v6_909_0_1	0.90	2.62E-10			
dd_Smed_v6_10364_0_1	0.89	8.48E-06	Dehydrogenase/reductase SDR family member 11	Q6UWP2.1	1.00E-76
dd_Smed_v6_10504_0_1	0.89	8.17E-08			
dd_Smed_v6_3194_0_1	0.86	9.29E-07	Serine/threonine-protein kinase 24	Q9Y6E0.1	1.00E-33
dd_Smed_v6_28214_0_1	0.81	2.24E-06	Muscarinic acetylcholine receptor M1	P11229.2	1.00E-29
dd_Smed_v6_9168_0_2	0.80	0.000326886	DNA polymerase alpha catalytic subunit	P09884.2	0
dd_Smed_v6_4595_0_1	0.75	0.003588929	Chromatin-remodeling ATPase INO80	Q9ULG1.2	0
dd_Smed_v6_585_0_1	0.75	1.01E-05			
dd_Smed_v6_122_1_1	0.71	3.80E-14	Lysosomal acid lipase/cholesteryl ester hydrolase	P38571.2	4.00E-111
dd_Smed_v6_2850_0_3	0.70	0.005579583			
dd_Smed_v6_2777_0_1	0.70	0.011293493	Mitofusin-2	O95140.3	2.00E-24
dd_Smed_v6_6738_0_1	0.69	1.66E-06	Ethanolamine-phosphate phospho-lyase	Q8TBG4.1	0
dd_Smed_v6_13191_0_1	0.69	0.012723791			
dd_Smed_v6_12731_0_1	0.68	0.000396239	E3 SUMO-protein ligase CBX4	O00257.3	6.00E-09
dd_Smed_v6_3638_0_1	0.66	0.005943063	Krueppel-like factor 13	Q9Y2Y9.1	2.00E-34
dd_Smed_v6_916_0_2	0.65	3.27E-05	Fumarylacetoacetase	P16930.2	0
dd_Smed_v6_8_0_1	0.64	2.85E-05			
dd_Smed_v6_6819_0_2	0.64	0.036840685	Alpha-1,4-N-acetylglucosaminyltransferase	Q9UNA3.1	3.00E-04
dd_Smed_v6_3177_0_1	0.63	2.31E-09	Tolloid-like protein 2	Q9Y6L7.1	2.00E-28
dd_Smed_v6_271_0_1	0.63	5.23E-09	Fatty acid-binding protein, brain	O15540.3	8.00E-30
dd_Smed_v6_7637_0_2	0.61	3.36E-07	Inositol 1,4,5-trisphosphate receptor type 1	Q14643.3	0
dd_Smed_v6_19_0_3	0.61	0.000647842			
dd_Smed_v6_5406_0_1	0.60	0.004289082			
dd_Smed_v6_3650_0_1	0.60	1.44E-05	TNF receptor-associated factor 5	O00463.2	2.00E-13

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 3: Significantly differentially expressed genes after 28 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_42_0_1	0.60	0.03944555 6			
dd_Smed_v6_12877_0_1	0.59	0.02618344 3			
dd_Smed_v6_388_0_1	0.59	7.07E-06	Dynamin-1-like protein	O00429.2	3.00E-134
dd_Smed_v6_496_0_1	0.59	0.08212226 9			
dd_Smed_v6_9441_0_2	0.58	0.04233087	Haloacid dehalogenase-like hydrolase domain-containing 5	Q9BXW7.1	7.00E-56
dd_Smed_v6_668_0_1	0.57	5.56E-10	Hsc70-interacting protein	P50502.2	5.00E-52
dd_Smed_v6_258_0_1	0.57	3.21E-18	NADH-ubiquinone oxidoreductase chain 1	P03886.1	3.00E-42
dd_Smed_v6_421_0_1	0.57	1.38E-08	10 kDa heat shock protein, mitochondrial	P61604.2	1.00E-38
dd_Smed_v6_2777_0_2	0.56	0.00094328	Mitofusin-2	O95140.3	2.00E-24
dd_Smed_v6_8766_0_1	0.56	0.00805830 1	2-amino-3-carboxymuconate-6-semialdehyde decarboxylase	Q8TDX5.1	1.00E-136
dd_Smed_v6_19_0_2	0.56	0.00176150 7			
dd_Smed_v6_12062_0_1	0.55	0.08001038 9	Protein ATP1B4	Q9UN42.1	6.00E-14
dd_Smed_v6_753_0_1	0.53	4.71E-05	Cytochrome c oxidase subunit 1	P00395.1	0
dd_Smed_v6_5156_0_1	0.52	0.05302258 6	Histone H2A type 1-A	Q96QV6.3	4.00E-43
dd_Smed_v6_3169_0_1	0.51	0.00561757 9			
dd_Smed_v6_601_0_1	0.51	0.03585080 5	Histone H1.8	Q8IZA3.1	6.00E-09
dd_Smed_v6_526_0_1	0.51	2.78E-08	Adenosylhomocysteinase	P23526.4	0
dd_Smed_v6_5472_0_1	0.51	0.00925418 9			
dd_Smed_v6_5660_0_1	0.50	0.00105009	Sodium/hydrogen exchanger 1	P19634.2	5.00E-59
dd_Smed_v6_602_0_1	0.50	1.57E-05	NPC intracellular cholesterol transporter 2	P61916.1	1.00E-33
dd_Smed_v6_315_0_1	0.49	0.00045916 7			
dd_Smed_v6_10696_0_3	0.48	0.02050981 4	NFX1-type zinc finger-containing protein 1	Q9P2E3.2	0
dd_Smed_v6_3794_0_1	0.48	0.01161746 4	Pantothenate kinase 3	Q9H999.1	1.00E-08
dd_Smed_v6_758_1_1	0.48	0.00248307 4	Endoplasmic	P14625.1	0
dd_Smed_v6_5726_0_1	0.48	0.00382147 9	D-amino-acid oxidase	P14920.3	4.00E-56
dd_Smed_v6_8422_0_1	0.47	0.09241566 4	Plasma alpha-L-fucosidase	Q9BTY2.2	2.00E-176
dd_Smed_v6_3387_0_1	0.47	0.00291590 2	Proteasome subunit alpha type-6	P60900.1	6.00E-99
dd_Smed_v6_56_0_1	0.47	6.27E-10	Heat shock protein HSP 90-alpha	P07900.5	0
dd_Smed_v6_4596_0_1	0.46	0.07350400 9			
dd_Smed_v6_5862_0_1	0.46	0.08493773 7	Ribonucleoside-diphosphate reductase subunit M2 B	Q7LG56.1	2.00E-31
dd_Smed_v6_5347_0_2	0.46	0.05025305 9	Nuclear pore complex protein Nup98-Nup96	P52948.4	3.00E-60
dd_Smed_v6_1794_0_1	0.45	0.00744220 6	Catechol O-methyltransferase domain-containing protein 1	Q86VU5.1	3.00E-49
dd_Smed_v6_132_0_1	0.45	0.09855989 4			
dd_Smed_v6_19_0_1	0.45	0.00137198			

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 3: Significantly differentially expressed genes after 28 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_511_0_1	0.45	4.84E-09	Endoplasmic reticulum chaperone BiP	P11021.2	0
dd_Smed_v6_11756_0_1	0.45	0.0938081	Mitogen-activated protein kinase kinase kinase 15	Q6ZN16.2	8.00E-23
dd_Smed_v6_6337_0_1	0.45	0.021388878	Phenylalanine-4-hydroxylase	P00439.1	0
dd_Smed_v6_1251_0_1	0.44	0.007262118	Four and a half LIM domains protein 3	Q13643.4	2.00E-17
dd_Smed_v6_785_0_1	0.44	0.000396239	Lysosomal protective protein	P10619.2	5.00E-145
dd_Smed_v6_678_0_1	0.44	0.000998659	PUTATIVE PSEUDOGENE: RecName: Putative heat shock protein HSP 90-beta 4	Q58FF6.1	2.00E-120
dd_Smed_v6_862_0_1	0.43	0.044539806			
dd_Smed_v6_5447_0_1	0.43	0.017101234	Solute carrier family 2, facilitated glucose transporter member 1	P11166.2	2.00E-115
dd_Smed_v6_3051_0_1	0.43	0.028823572			
dd_Smed_v6_6553_0_1	0.43	0.061743947			
dd_Smed_v6_1536_0_1	0.42	0.060484763	Peptidyl-prolyl cis-trans isomerase FKBP2	P26885.2	7.00E-50
dd_Smed_v6_1706_0_1	0.42	0.031632137	Prosaposin	P07602.2	6.00E-07
dd_Smed_v6_380_0_1	0.42	0.067089111			
dd_Smed_v6_2947_0_1	0.42	0.027865274	Succinate--CoA ligase [GDP-forming] subunit beta, mitochondrial	Q96199.2	3.00E-153
dd_Smed_v6_3508_0_1	0.41	0.032069337	Solute carrier family 28 member 3	Q9HAS3.1	7.00E-110
dd_Smed_v6_15598_0_4	0.41	0.042321041	1-acyl-sn-glycerol-3-phosphate acyltransferase alpha	Q99943.2	3.00E-41
dd_Smed_v6_3419_0_1	0.41	0.013842064	Kynurenine/alpha-aminoadipate aminotransferase, mitochondrial	Q8N5Z0.2	2.00E-13
dd_Smed_v6_4792_0_1	0.40	0.080184996	Neurogenic locus notch homolog protein 4	Q99466.2	1.00E-10
dd_Smed_v6_861_0_2	0.40	0.008423505			
dd_Smed_v6_1767_0_1	0.40	0.031403847	Bcl-2 homologous antagonist/killer	Q16611.1	2.00E-11
dd_Smed_v6_1557_0_1	0.39	0.03302288	Proteasome subunit alpha type-3	P25788.2	9.00E-112
dd_Smed_v6_1235_0_1	0.38	0.042321041	Proteasome subunit alpha type-7	O14818.1	6.00E-109
dd_Smed_v6_1921_0_1	0.38	0.000257708	Peptidyl-prolyl cis-trans isomerase FKBP4	Q02790.3	2.00E-116
dd_Smed_v6_1_0_1	0.38	0.005654931			
dd_Smed_v6_4738_0_1	0.38	0.039445556	Pyruvate kinase PKM	P14618.4	0
dd_Smed_v6_2582_1_1	0.38	0.09572883	Methyltransferase-like protein 27	Q8N6F8.2	1.00E-13
dd_Smed_v6_1771_0_1	0.37	0.046968995	Proteasome subunit beta type-1	P20618.2	4.00E-79
dd_Smed_v6_9501_0_3	0.37	0.001724746			
dd_Smed_v6_219_0_1	0.36	0.007879812	Dipeptidyl peptidase 1	P53634.2	8.00E-159
dd_Smed_v6_1584_0_1	0.35	0.000918976	Beta-ureidopropionase	Q9UBR1.1	2.00E-174
dd_Smed_v6_320_0_1	0.35	0.013473592	Heat shock cognate 71 kDa protein	P11142.1	0
dd_Smed_v6_1711_0_1	0.35	0.077169277	Glutamine--tRNA ligase	P47897.1	0

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 3: Significantly differentially expressed genes after 28 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_1170_0_1	0.35	3.48E-06	Haloacid dehalogenase-like hydrolase domain-containing 5	Q9BXW7.1	7.00E-90
dd_Smed_v6_220_0_4	0.35	0.053022586	Calreticulin	P27797.1	2.00E-151
dd_Smed_v6_2087_0_1	0.34	0.055106009	Phosphatidylserine decarboxylase proenzyme, mitochondrial	Q9UG56.4	5.00E-68
dd_Smed_v6_4058_0_1	0.34	0.058858474	Phosphoglucomutase-2	Q96G03.4	6.00E-176
dd_Smed_v6_4449_0_1	0.33	0.097153019	Cytochrome P450 2A6	P11509.4	2.00E-80
dd_Smed_v6_423_0_1	0.33	0.021388878	Cystatin-A	P01040.1	2.00E-16
dd_Smed_v6_16045_0_2	0.32	0.004961376			
dd_Smed_v6_921_0_1	0.32	0.003348694	LIM domain and actin-binding protein 1	Q9UHB6.1	6.00E-19
dd_Smed_v6_1599_0_1	0.32	0.000998659	Coronin-1B	Q9BR76.1	6.00E-72
dd_Smed_v6_278_0_1	0.32	0.006153615			
dd_Smed_v6_22388_0_2	0.31	0.004760306			
dd_Smed_v6_1986_0_1	0.30	0.018900282	Stress-induced-phosphoprotein 1	P31948.1	0
dd_Smed_v6_23320_0_1	0.30	0.007576787			
dd_Smed_v6_1087_0_1	0.30	0.028428297	Heat shock 70 kDa protein 4L	O95757.3	0
dd_Smed_v6_4296_0_2	0.30	0.005632566	RNA-binding protein with multiple splicing	Q93062.1	2.00E-23
dd_Smed_v6_929_0_1	0.29	0.019053795	Advillin	O75366.3	1.00E-59
dd_Smed_v6_18_0_1	0.29	0.000248555			
dd_Smed_v6_462_0_1	0.29	0.005943063			
dd_Smed_v6_1757_0_1	0.29	0.041595022			
dd_Smed_v6_10941_0_58	0.27	0.00473448			
dd_Smed_v6_727_0_1	0.25	0.066750474	Adenylyl cyclase-associated protein 2	P40123.1	3.00E-130
dd_Smed_v6_17904_0_2	0.25	0.003906654			
dd_Smed_v6_662_0_1	0.25	0.026946461	Phospholipid hydroperoxide glutathione peroxidase	P36969.3	9.00E-60
dd_Smed_v6_180_0_1	0.24	0.063244478	Ubiquitin-conjugating enzyme E2 L3	P68036.1	3.00E-39
dd_Smed_v6_8526_0_10	0.23	0.035890317	Tigger transposable element-derived protein 1	Q96MW7.1	9.00E-08
dd_Smed_v6_21185_0_2	0.23	0.012973447	Serine/threonine-protein phosphatase with EF-hands 1	O14829.1	1.00E-61
dd_Smed_v6_8311_0_1	0.23	0.011758265	Tyrosyl-DNA phosphodiesterase 1	Q9NUW8.2	1.00E-112
dd_Smed_v6_14153_0_1	0.19	0.030257297			
dd_Smed_v6_8834_0_3	0.18	0.008058301	Squamous cell carcinoma antigen recognized by T-cells 3	Q15020.1	3.00E-06
dd_Smed_v6_26861_0_2	0.18	0.041152772	Potassium voltage-gated channel subfamily C member 2	Q96PR1.1	1.00E-78
dd_Smed_v6_8975_0_2	0.18	0.062790677	Serine/threonine-protein kinase 4	Q13043.2	1.00E-17
dd_Smed_v6_9768_0_2	0.17	0.061459058	Omega-amidase NIT2	Q9NQR4.1	7.00E-75
dd_Smed_v6_6427_0_1	0.15	0.006187349	Exocyst complex component 1	Q9NV70.4	1.00E-89
dd_Smed_v6_1205_0_2	0.12	0.038562324	Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5	Q9UPN3.4	0

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 3: Significantly differentially expressed genes after 28 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_5368_0_1	-0.13	0.06708515 2	Serine/threonine-protein kinase 3	Q13188.2	4.00E-17
dd_Smed_v6_11507_0_2	-0.18	0.03236370 1	Nucleolar pre-ribosomal-associated protein 1	O60287.4	3.00E-20
dd_Smed_v6_27418_0_2	-0.19	0.07463177 2			
dd_Smed_v6_17809_0_1	-0.21	0.04898133 9	Zinc finger CCHC domain-containing protein 13	Q8WW36.1	0.026
dd_Smed_v6_12660_2_1	-0.22	0.04125260 4	Tigger transposable element-derived protein 1	Q96MW7.1	7.00E-134
dd_Smed_v6_906_0_1	-0.23	0.03928789 2	Staphylococcal nuclease domain-containing protein 1	Q7KZF4.1	0
dd_Smed_v6_2141_0_2	-0.23	0.02760610 6	Actin-related protein 10	Q9NZ32.1	9.00E-74
dd_Smed_v6_351_0_1	-0.24	0.00099865 9	Neurotrypsin	P56730.2	7.00E-34
dd_Smed_v6_16239_0_3	-0.26	0.00514159 8			
dd_Smed_v6_536_0_1	-0.27	0.04135255 9	Phosphoenolpyruvate carboxykinase, cytosolic [GTP]	P35558.3	0
dd_Smed_v6_48_0_1	-0.27	0.06177133 5	Procathepsin L	P07711.2	6.00E-113
dd_Smed_v6_553_0_1	-0.27	0.00176970 7	Vigilin	Q00341.2	0
dd_Smed_v6_1574_6_12 2	-0.27	0.00874152 3			
dd_Smed_v6_2284_0_1	-0.29	0.05885847 4	Ubiquitin carboxyl-terminal hydrolase 4	Q13107.3	8.00E-135
dd_Smed_v6_17717_0_1	-0.29	0.00565493 1	Tensin-2	Q63HR2.2	2.00E-12
dd_Smed_v6_47_0_1	-0.29	0.00565493 1			
dd_Smed_v6_854_1_1	-0.31	0.02666081 1	CD63 antigen	P08962.2	2.00E-04
dd_Smed_v6_1873_0_2	-0.31	0.03944555 6	Protein transport protein Sec16A	O15027.4	2.00E-17
dd_Smed_v6_1694_0_1	-0.32	0.00311989 4	Hepatocyte nuclear factor 4	Q14541.3	2.00E-111
dd_Smed_v6_1833_0_1	-0.32	0.02050981 4	Glycine dehydrogenase (decarboxylating), mitochondrial	P23378.2	0
dd_Smed_v6_131_0_2	-0.32	0.00311989 4			
dd_Smed_v6_1874_0_1	-0.32	0.05739648 4	Sialin	Q9NRA2.2	2.00E-111
dd_Smed_v6_27_0_1	-0.33	0.08493773 7			
dd_Smed_v6_542_0_1	-0.33	0.03542315 2	NADPH--cytochrome P450 reductase	P16435.2	0
dd_Smed_v6_1032_0_1	-0.34	0.00037738 4	Cyclic AMP-responsive element-binding protein 3-like protein 1	Q96BA8.1	3.00E-34
dd_Smed_v6_825_0_1	-0.34	0.08212226 9			
dd_Smed_v6_2209_0_1	-0.35	0.01743243 4	Whirlin	Q9P202.4	0.001
dd_Smed_v6_415_0_1	-0.35	0.01927060 2	X-box-binding protein 1	P17861.2	6.00E-12
dd_Smed_v6_1577_0_1	-0.35	0.04898133 9			
dd_Smed_v6_14132_0_1 0	-0.36	0.00135005 5	Zinc finger BED domain-containing protein 5	Q49AG3.2	7.00E-04
dd_Smed_v6_366_0_1	-0.36	0.01714408	6-phosphogluconate dehydrogenase, decarboxylating	P52209.3	0
dd_Smed_v6_870_0_1	-0.36	0.00234504 9	FK506-binding protein 15	Q5T1M5.2	3.00E-48
dd_Smed_v6_938_0_1	-0.37	0.01437655 3	Signal recognition particle receptor subunit beta	Q9Y5M8.3	1.00E-42

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 3: Significantly differentially expressed genes after 28 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_2649_0_1	-0.37	0.01057584 7	Fibrillin-3	Q75N90.3	4.00E-136
dd_Smed_v6_14342_0_4	-0.38	0.00068937 4	Histone-lysine N-methyltransferase SETMAR	Q53H47.2	0.002
dd_Smed_v6_1489_0_1	-0.38	1.05E-06	Macrophage mannose receptor 1	P22897.1	7.00E-17
dd_Smed_v6_2541_0_1	-0.38	0.00802096 3	Carboxypeptidase A2	P48052.3	3.00E-72
dd_Smed_v6_413_0_1	-0.38	0.01297344 7	Lysophospholipid acyltransferase 2	Q6ZWT7.2	1.00E-76
dd_Smed_v6_10141_0_4	-0.38	0.00026590 7	Transmembrane protein 39A	Q9NV64.1	3.00E-42
dd_Smed_v6_855_0_1	-0.39	0.00146891 9	Binder of sperm protein homolog 1	Q075Z2.1	2.00E-04
dd_Smed_v6_1155_0_1	-0.40	1.16E-05	Aggrecan core protein	P16112.3	1.00E-14
dd_Smed_v6_487_0_1	-0.40	0.00565493 1	Glutamate--cysteine ligase regulatory subunit	P48507.1	1.00E-26
dd_Smed_v6_643_0_1	-0.41	0.00744220 6	N-acetyl-D-glucosamine kinase	Q9UJ70.4	1.00E-52
dd_Smed_v6_970_0_1	-0.41	0.00565493 1	MKI67 FHA domain-interacting nucleolar phosphoprotein	Q9BYG3.1	9.00E-08
dd_Smed_v6_87_0_2	-0.42	0.04268316 1			
dd_Smed_v6_1585_0_1	-0.42	0.00872765 1	Cystathionine beta-synthase-like protein	P0DN79.1	0
dd_Smed_v6_131_0_1	-0.42	1.66E-06			
dd_Smed_v6_214_1_1	-0.42	0.00527630 9	Ganglioside GM2 activator	P17900.4	4.00E-19
dd_Smed_v6_606_0_1	-0.42	0.04233087	Alpha-N-acetylgalactosaminidase	P17050.2	2.00E-112
dd_Smed_v6_735_0_1	-0.43	1.86E-05			
dd_Smed_v6_13734_0_1 0	-0.44	0.00015259 1			
dd_Smed_v6_3150_0_1	-0.44	0.09857392 9	Probable tRNA N6-adenosine threonylcarbamoyltransferase	Q9NPF4.1	1.00E-180
dd_Smed_v6_975_0_1	-0.44	0.01167364 7	Kallikrein-13	Q9UKR3.1	1.00E-36
dd_Smed_v6_4455_0_1	-0.44	0.03293163 5	CDK5 regulatory subunit-associated protein 3	Q96JB5.2	9.00E-58
dd_Smed_v6_6020_0_1	-0.44	0.07716927 7	Fibroblast growth factor receptor 1	P11362.3	1.00E-53
dd_Smed_v6_7816_0_1	-0.45	0.04696899 5	60 kDa lysophospholipase	Q86U10.3	5.00E-116
dd_Smed_v6_257_0_1	-0.45	3.42E-08	Gastric triacylglycerol lipase	P07098.1	5.00E-117
dd_Smed_v6_8785_0_1	-0.45	0.05252119 2	Synaptotagmin-14	Q8NB59.2	1.00E-40
dd_Smed_v6_154_0_1	-0.45	0.00874152 3			
dd_Smed_v6_2336_0_1	-0.46	0.02655412 5	Down syndrome cell adhesion molecule-like protein 1	Q8TD84.2	4.00E-05
dd_Smed_v6_3294_1_1	-0.46	0.03854312	Leucine-rich repeat and WD repeat-containing protein 1	Q9UFC0.2	7.00E-52
dd_Smed_v6_2591_0_1	-0.47	9.88E-05	Macrophage mannose receptor 1	P22897.1	4.00E-22
dd_Smed_v6_2320_0_1	-0.47	0.00293198 9	Leishmanolysin-like peptidase	Q96KR4.2	2.00E-56
dd_Smed_v6_583_0_1	-0.48	0.01665759			
dd_Smed_v6_328_0_1	-0.48	0.00434259 7			
dd_Smed_v6_8514_0_5	-0.49	3.45E-05			

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 3: Significantly differentially expressed genes after 28 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_2559_0_1	-0.49	0.00428908 2	Leucine-rich repeat-containing protein 72	A6NJI9.2	4.00E-07
dd_Smed_v6_1446_0_1	-0.51	0.03232096 2	Tissue factor pathway inhibitor	P10646.1	2.00E-13
dd_Smed_v6_6850_0_1	-0.51	0.08212226 9	Heat shock cognate 71 kDa protein	P11142.1	0
dd_Smed_v6_9200_0_1	-0.51	0.00092399 3	Macrophage mannose receptor 1	P22897.1	9.00E-27
dd_Smed_v6_5305_0_3	-0.52	0.00032507 2	FAS-associated factor 1	Q9UNN5.2	3.00E-49
dd_Smed_v6_3603_0_1	-0.53	0.00110007 4	Peroxidase homolog	Q92626.2	0
dd_Smed_v6_5760_0_1	-0.53	0.04115277 2			
dd_Smed_v6_4633_0_1	-0.53	0.04794676 8			
dd_Smed_v6_2361_0_1	-0.53	0.00013968 8	Multidrug and toxin extrusion protein 1	Q96FL8.1	2.00E-90
dd_Smed_v6_2970_0_1	-0.54	8.22E-08	Histidine ammonia-lyase	P42357.1	2.00E-99
dd_Smed_v6_9186_0_1	-0.54	0.06708515 2	Mothers against decapentaplegic homolog 7	O15105.1	1.00E-05
dd_Smed_v6_1893_0_1	-0.55	5.77E-09	Protein disulfide-isomerase A2	Q13087.2	1.00E-33
dd_Smed_v6_7505_0_1	-0.55	0.03700093 4			
dd_Smed_v6_7577_0_1	-0.55	0.00094328	cAMP-specific 3',5'-cyclic phosphodiesterase 4D	Q08499.2	5.00E-171
dd_Smed_v6_2193_0_1	-0.56	0.01297344 7			
dd_Smed_v6_6823_0_1	-0.57	0.08736369 8			
dd_Smed_v6_72_0_1	-0.58	0.00105472 6			
dd_Smed_v6_279_0_1	-0.58	4.03E-06			
dd_Smed_v6_3260_0_1	-0.58	0.00011437 9			
dd_Smed_v6_6923_0_1	-0.58	0.09857392 9	Protein Smaug homolog 2	Q5PRF9.1	4.00E-41
dd_Smed_v6_28214_0_2	-0.59	2.86E-05	Muscarinic acetylcholine receptor M1	P11229.2	1.00E-29
dd_Smed_v6_463_0_1	-0.59	0.00078821 7			
dd_Smed_v6_5347_0_1	-0.60	0.00042539 6	Nuclear pore complex protein Nup98-Nup96	P52948.4	4.00E-60
dd_Smed_v6_1694_0_3	-0.61	0.06708911 1	Hepatocyte nuclear factor 4-gamma	Q14541.3	3.00E-93
dd_Smed_v6_6938_0_1	-0.61	0.04574696 1			
dd_Smed_v6_465_0_1	-0.62	0.01175826 5			
dd_Smed_v6_2068_0_1	-0.63	0.00478631 9	Nucleoside diphosphate-linked moiety X motif 8	Q8WV74.2	5.00E-24
dd_Smed_v6_3266_0_1	-0.63	5.88E-11	Macrophage mannose receptor 1	P22897.1	7.00E-30
dd_Smed_v6_10171_0_2	-0.64	0.04024170 7	Gamma-aminobutyric acid receptor subunit gamma-2	P18507.2	4.00E-13
dd_Smed_v6_2169_0_1	-0.64	2.69E-10			
dd_Smed_v6_5978_0_2	-0.66	0.00311989 4	Nuclear factor of activated T-cells 5	O94916.1	4.00E-81
dd_Smed_v6_254_0_1	-0.66	5.15E-11	Bone morphogenetic protein 1	P13497.2	1.00E-22
dd_Smed_v6_66_0_2	-0.67	8.16E-07	Kallikrein-7	P49862.1	1.00E-13
dd_Smed_v6_4733_0_1	-0.67	0.02055703 6			

Supplemental Table S2.2. List of significantly differentially expressed genes following *phc(RNAi)* or *rnf2(RNAi)* at FDR cutoff value of ≤ 0.1 . Continued.

Tab 3: Significantly differentially expressed genes after 28 days of *rnf2(RNAi)*.

Contig ID	log2FoldChange	padj	Top Human Uniprot Hit	Accession	e-value
dd_Smed_v6_3728_0_1	-0.68	0.000139688	Kallikrein-13	Q9UKR3.1	2.00E-26
dd_Smed_v6_750_0_1	-0.69	1.33E-05			
dd_Smed_v6_11905_0_7	-0.70	0.008057368	Partitioning defective 3 homolog	Q8TEW0.2	2.00E-11
dd_Smed_v6_634_0_1	-0.71	5.72E-25			
dd_Smed_v6_663_0_1	-0.72	1.40E-05	Prosaposin	P07602.2	2.00E-05
dd_Smed_v6_4570_0_1	-0.73	8.14E-05	Kallikrein-13	Q9UKR3.1	5.00E-32
dd_Smed_v6_827_0_1	-0.73	3.53E-09			
dd_Smed_v6_3066_0_1	-0.73	4.03E-06	Transmembrane protein 41A	Q96HV5.1	2.00E-66
dd_Smed_v6_238_1_1	-0.74	3.75E-15	Zonadhesin	Q9Y493.5	2.00E-24
dd_Smed_v6_6816_0_1	-0.74	0.003722563	Myoferlin	Q9NZM1.1	0
dd_Smed_v6_122_0_1	-0.88	5.72E-25	Teneurin-2	Q9NT68.3	7.00E-05
dd_Smed_v6_66_0_1	-0.88	2.88E-24	Kallikrein-13	Q9UKR3.1	3.00E-19
dd_Smed_v6_8989_0_1	-1.02	4.27E-07	E3 ubiquitin-protein ligase RING2	Q99496.1	6.00E-46

References

- 1 Baguna, J., Salo, E. & Auladell, C. Regeneration and pattern formation in planarians III. Evidence that neoblasts are totipotent stem cells and the source of blastema cells. *Development* **107**, 77-86 (1989).
- 2 Wagner, D. E., Wang, I. E. & Reddien, P. W. Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* **332**, 811-816, doi:10.1126/science.1203983 (2011).
- 3 Reddien, P. W. The Cellular and Molecular Basis for Planarian Regeneration. *Cell* **175**, 327-345, doi:10.1016/j.cell.2018.09.021 (2018).
- 4 Labbe, R. M. *et al.* A comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. *Stem Cells* **30**, 1734-1745, doi:10.1002/stem.1144 (2012).
- 5 Onal, P. *et al.* Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. *EMBO J* **31**, 2755-2769, doi:10.1038/emboj.2012.110 (2012).
- 6 Boser, A. *et al.* SILAC proteomics of planarians identifies Ncoa5 as a conserved component of pluripotent stem cells. *Cell Rep* **5**, 1142-1155, doi:10.1016/j.celrep.2013.10.035 (2013).
- 7 Fernandez-Taboada, E., Rodriguez-Esteban, G., Salo, E. & Abril, J. F. A proteomics approach to decipher the molecular nature of planarian stem cells. *BMC Genomics* **12**, 133, doi:10.1186/1471-2164-12-133 (2011).
- 8 Strand, N. S. *et al.* Dissecting the function of Cullin-RING ubiquitin ligase complex genes in planarian regeneration. *Dev Biol* **433**, 210-217, doi:10.1016/j.ydbio.2017.10.011 (2018).
- 9 Ciechanover, A., Finley, D. & Varshavsky, A. Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell* **37**, 57-66, doi:10.1016/0092-8674(84)90300-3 (1984).
- 10 Endoh, M. *et al.* Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. *PLoS Genet* **8**, e1002774, doi:10.1371/journal.pgen.1002774 (2012).
- 11 Higgins, R. *et al.* The Unfolded Protein Response Triggers Site-Specific Regulatory Ubiquitylation of 40S Ribosomal Proteins. *Mol Cell* **59**, 35-49, doi:10.1016/j.molcel.2015.04.026 (2015).
- 12 Nakayama, K. I. & Nakayama, K. Regulation of the cell cycle by SCF-type ubiquitin ligases. *Semin Cell Dev Biol* **16**, 323-333, doi:10.1016/j.semcdb.2005.02.010 (2005).
- 13 Werner, A., Manford, A. G. & Rape, M. Ubiquitin-Dependent Regulation of Stem Cell Biology. *Trends Cell Biol* **27**, 568-579, doi:10.1016/j.tcb.2017.04.002 (2017).

- 14 Li, W. *et al.* Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS One* **3**, e1487, doi:10.1371/journal.pone.0001487 (2008).
- 15 Lorick, K. L. *et al.* RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc Natl Acad Sci U S A* **96**, 11364-11369, doi:10.1073/pnas.96.20.11364 (1999).
- 16 Aravind, L. & Koonin, E. V. The U box is a modified RING finger - a common domain in ubiquitination. *Curr Biol* **10**, 132-134, doi:10.1016/S0960-9822(00)00398-5 (2000).
- 17 Henderson, J. M. *et al.* Identification of HECT E3 ubiquitin ligase family genes involved in stem cell regulation and regeneration in planarians. *Dev Biol* **404**, 21-34, doi:10.1016/j.ydbio.2015.04.021 (2015).
- 18 Merryman, M. S., Alvarado, A. S. & Jenkin, J. C. Culturing Planarians in the Laboratory. *Methods Mol Biol* **1774**, 241-258, doi:10.1007/978-1-4939-7802-1_5 (2018).
- 19 Brandl, H. *et al.* PlanMine--a mineable resource of planarian biology and biodiversity. *Nucleic Acids Res* **44**, D764-773, doi:10.1093/nar/gkv1148 (2016).
- 20 Rouhana, L. *et al.* RNA interference by feeding in vitro-synthesized double-stranded RNA to planarians: methodology and dynamics. *Dev Dyn* **242**, 718-730, doi:10.1002/dvdy.23950 (2013).
- 21 Liu, S. Y. *et al.* Reactivating head regrowth in a regeneration-deficient planarian species. *Nature* **500**, 81-84, doi:10.1038/nature12414 (2013).
- 22 Pellettieri, J. *et al.* Cell death and tissue remodeling in planarian regeneration. *Dev Biol* **338**, 76-85, doi:10.1016/j.ydbio.2009.09.015 (2010).
- 23 Pearson, B. J. *et al.* Formaldehyde-based whole-mount in situ hybridization method for planarians. *Dev Dyn* **238**, 443-450, doi:10.1002/dvdy.21849 (2009).
- 24 King, R. S. & Newmark, P. A. *In situ* hybridization protocol for enhanced detection of gene expression in the planarian *Schmidtea mediterranea*. *BMC Developmental Biology* **13** (2013).
- 25 Brown, D. D. R. & Pearson, B. J. in *In Situ Hybridization Methods Neuromethods* Ch. Chapter 7, 127-150 (2015).
- 26 Simoes, A. E. *et al.* Efficient recovery of proteins from multiple source samples after TRIzol((R)) or TRIzol((R))LS RNA extraction and long-term storage. *BMC Genomics* **14**, 181, doi:10.1186/1471-2164-14-181 (2013).
- 27 Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525-527, doi:10.1038/nbt.3519 (2016).
- 28 Huber, W. *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* **12**, 115-121, doi:10.1038/nmeth.3252 (2015).

- 29 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
- 30 Roberts-Galbraith, R. H., Brubacher, J. L. & Newmark, P. A. A functional genomics screen in planarians reveals regulators of whole-brain regeneration. *Elife* **5**, doi:10.7554/eLife.17002 (2016).
- 31 Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873-878, doi:10.1038/nature02985 (2004).
- 32 Fursova, N. A. *et al.* Synergy between Variant PRC1 Complexes Defines Polycomb-Mediated Gene Repression. *Mol Cell* **74**, 1020-1036 e1028, doi:10.1016/j.molcel.2019.03.024 (2019).
- 33 Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. & Sanchez Alvarado, A. SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* **310**, 1327-1330, doi:10.1126/science.1116110 (2005).
- 34 Kim, I. V. *et al.* Planarians recruit piRNAs for mRNA turnover in adult stem cells. *Genes Dev* **33**, 1575-1590, doi:10.1101/gad.322776.118 (2019).
- 35 Solana, J. *et al.* The CCR4-NOT complex mediates deadenylation and degradation of stem cell mRNAs and promotes planarian stem cell differentiation. *PLoS Genet* **9**, e1004003, doi:10.1371/journal.pgen.1004003 (2013).
- 36 Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565-570, doi:10.1038/276565a0 (1978).
- 37 Conway, E. M. & Bracken, A. P. in *Polycomb Group Proteins* 57-80 (2017).
- 38 Cao, R., Tsukada, Y. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol Cell* **20**, 845-854, doi:10.1016/j.molcel.2005.12.002 (2005).
- 39 Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445-1459, doi:10.1016/j.cell.2014.05.004 (2014).
- 40 Gahan, J. M., Rentzsch, F. & Schnitzler, C. E. The genetic basis for PRC1 complex diversity emerged early in animal evolution. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.2005136117 (2020).
- 41 Peng, Y. C., Lv, T. H., Du, Z. K., Cun, X. N. & Yang, K. M. Liver Macrophages Stimulate the Expression of Hepatocyte Nuclear Factor-6 and Promote Hepatocyte Proliferation at the Early Stage of Liver Regeneration. *Bull Exp Biol Med* **170**, 40-45, doi:10.1007/s10517-020-05000-7 (2020).
- 42 van der Raadt, J., van Gestel, S. H. C., Nadif Kasri, N. & Albers, C. A. ONECUT transcription factors induce neuronal characteristics and remodel chromatin accessibility. *Nucleic Acids Res* **47**, 5587-5602, doi:10.1093/nar/gkz273 (2019).

- 43 Jiang, K. *et al.* HNF6 promotes tumor growth in colorectal cancer and enhances liver metastasis in mouse model. *J Cell Physiol* **234**, 3675-3684, doi:10.1002/jcp.27140 (2019).

Conclusion of the dissertation

The phenomenon of regeneration has captured the fascination of naturalists and scientists for centuries. Regeneration is a process that is both distinct from and akin to embryological development, with many of the underlying regulatory pathways in regeneration representing a recapitulation of those developmental patterning. There are two major differences between regeneration and embryogenesis. The first is that the ability to regenerate is dispersed widely but not uniformly throughout the animal phyla, with significant variation, even among closely related species, in the degree of regenerative ability, which contrasts with the near universal process in metazoans of embryonic development^{1,2}. The other major difference is that embryogenesis typically proceeds from an established starting point, the zygote, while regeneration must occur adaptively in response to environmental insults or life cycle events. Regeneration therefore requires a degree of responsive plasticity to integrate wound response signals and then direct the proper specification and integration of regenerated tissues.

The post-translational modification of proteins is commonly utilized in signal transduction pathways and offers a responsive and adaptive mechanism by which an organism can regulate regeneration. One such post-translational modifier of proteins is the small polypeptide ubiquitin, which has broad functions in cellular biology including protein degradation, cellular trafficking, and transcriptional regulation³⁻⁵. The terminal enzymatic step of the ubiquitylation cascade depends on the action of the E3 ubiquitin ligases, a large family of proteins that give specificity to the process of ubiquitylation⁶. Understanding the roles of the E3 ligases during regeneration is essential to understand of how ubiquitin signaling is regulating regenerative processes. We used the freshwater planarian as a

model to investigate the role of the E3 ligases in stem cell biology and regeneration because of its ability to completely recover from nearly any injury using a large pool of adult pluripotent stem cells. We focused our study on the RING and U-box classes of E3 ubiquitin ligases and extended our work to include the investigation of several complexes that are associated with E3 ligase function, including the CRLs, NTC, and PRC1.

Cullin RING ligase complexes in planarian regeneration

The CRL complexes are defined by the association of a RING factor with a partner Cullin protein. The Cullin act as a molecular scaffold that coordinates the assembly and organization of a particular CRL complex by forming two modules, a catalytic unit formed with a RING factor and a substrate targeting module that is comprised of an adaptor protein that links the CRL to a substrate-recognition protein⁷. A particular cullin protein can form several distinct CRL complexes through the modular association with different substrate-recognition proteins, allowing each class of CRL complexes to potentially have differential targets. To investigate the roles of CRL complexes in regulating regeneration we identified six cullin homologs in *S. mediterranea* and used RNAi to perturb gene function and uncovered roles for *cullin-1*, *-3*, and *-4*. We found that *cullin-1*, which is a core component of the SCF complex, presented phenotypes during homeostasis and regeneration. These phenotypes included defects in blastema formation and patterning during regeneration and impaired movement, lesions, and eventual lysis during homeostasis. These varied phenotypes, coupled with the broad expression of *cullin-1*, suggested that the SCF complex functions in many aspects of planarian biology. To dissect these functions in more specific contexts we reasoned that, because each F-box substrate recognition subunit likely interacts with only a subset of the total factors targeted by SCF,

we could use RNAi against individual *f-box* genes to perturb only a subset of SCF function. We identified and classified 35 *f-box* genes in *S. mediterranea* and knocked down 30 of them using RNAi and found that knockdown of 19 phenocopied aspects of the *cullin-1(RNAi)*. WISH analysis determined that the 19 *f-box* genes were expressed in patterns that were more restricted than that of *cullin-1*. Targeting the *f-box* subunits within SCF offers a mechanism by which the pleiotropic functions of SCF can be dissected for study or clinical interventions.

CRL complexes and the F-boxes more specifically have roles in regulating cell cycle dynamics and in stem cell biology, often through promoting the targeted degradation of regulatory and signaling factors. We identified nine *f-box* genes that had a significant effect on proliferation in planarians, suggesting a critical role for *f-box* genes and the SCF complex in regulating cell proliferation. We recovered known tumor suppressor genes in our *f-box* screen, including homologs of *FBXW7* and *FBXL2*^{8,9}. In humans and mice, *FBXL2* targets Cyclin D3 to arrest mitotic activity¹⁰, and we found that RNAi of the planarian homolog, *fxl2-1*, led to a significant increase in pH3⁺ cells, consistent with a function as a regulator of the cell cycle and as a tumor suppressor¹⁰. We demonstrated that planarians can be used as an effective *in vivo* model to dissect the function of ubiquitin E3 ligases complexes and potentially uncover regulators of stem cell biology and cell proliferation. These uncovered regulators could become potential targets for anti-cancer therapies and that targeting substrate recognition subunits like the *f-boxes* instead of core CRL subunits would offer treatment options that were more specific in affecting only the morbid aspects of a dysregulated CRL and potentially limit any side effects that would occur from disrupting the general functioning of a CRL complex.

Spliceosomal and epigenetic ubiquitin ligases are critical regulators of planarian biology

We continued our investigation of the role of ubiquitin signaling in regulating regeneration and stem cells by performing an *in vivo* functional screen of the largest class of E3 ubiquitin ligases, the RING and U-boxes, using RNAi to disrupt gene function. We generated a list of 393 contigs that were annotated as containing a RING or U-box domain and leveraged an existing transcriptomic data set to bin our list of E3 ligases into expression classes of stem cell, stem cells and progeny, and differentiated cells. To uncover genes important in stem cell regulation we focused our screening efforts on transcripts predicted to be expressed in the stem cell and stem cell progeny classes. We screened 103 E3 ligases for function during homeostasis and regeneration and found nine that exhibited phenotypes related to stem cell function.

We chose to further examine a couple of E3 ligases that emerged from our screen in greater depth and included the spliceosomal gene *prpf19* and the epigenetic factor *rnf2*. The U-box gene *prpf19* is the founding member of the NTC and regulates the assembly of the spliceosome by ubiquitylating U4 RNP protein PRP3 with nonproteolytic K63-linked chains. This action of *prpf19* is conserved with yeast and necessary for the proper processing of pre-mRNA into mature mRNA for gene translation^{11,12}. Prpf19/NTC also has roles in the DDR repair pathway where it acts as a sensor of DNA damage and coordinates DNA repair through its ubiquitylation of RPA-ssDNA complexes and subsequent recruitment of ATR^{13,14}.

When inhibited, *prpf19* presented a robust phenotype of head regression, ventral curling, lesioning, and lysis that are all typically associated with a loss of stem cells in *S.*

mediterranea. These phenotypes were also observed when other core factors of NTC were inhibited which suggested that the *prpf19(RNAi)* phenotypes are mediated through its role in NTC. When we examined the effect of *prpf19(RNAi)* on the expression of stem cell marker genes *piwi-1*, *tgs-1*, and *h2b* we surprisingly found that inhibition of *prpf19* did not lead to any loss of expression for those markers. This finding of the *prpf19(RNAi)* phenotype not being mediated by a loss of stem cells is consistent with a previous study that identified *prpf19* as being upregulated during head regeneration¹⁵.

The broad expression pattern observed in *prpf19* WISH suggested NTC had roles in diverse cell types, and we found that *prpf19(RNAi)* did cause a reduction in staining for early and late epidermal progenitor populations suggesting that differentiation of stem cells rather than their survival is affected by *prpf19* inhibition. This reduction in progenitor density was accompanied by a decrease in pH3⁺ proliferative cells, which, to not cause a concomitant reduction in stem cell numbers, must be balanced by reduction in stem cell differentiation rates. Through its DDR activity Prpf19 has anti-apoptotic properties¹⁶, a result that we confirmed in *S. mediterranea* as *prpf19(RNAi)* caused a significant increase in the number of TUNEL⁺ cells. Taken together, these data suggest that the phenotypes observed after *prpf19* inhibition do not result for a loss of stem cells but rather a dysregulation of homeostatic tissue replacement. The phenotypes observed in *prpf19(RNAi)* worms could be the result of a failure of the stem cells to differentiate properly or that *prpf19* is necessary as an anti-apoptotic factor for worm survival.

The post-transcriptional processing and regulation of RNAs is becoming established as a major regulator of stem cell differentiation in planarian biology. Previous work identified planarian PIWI factors *smewi-2* and *smewi-3* as regulating planarian stem cells

with *smedwi-2* being dispensable for stem cell maintenance but necessary to specify progeny cells in a manner reminiscent of the *prpf19(RNAi)* phenotype^{17,18}. The CCR4-NOT complex is a post-transcriptional regulator of mRNAs degradation by promoting the deadenylation of poly(A) tails^{19,20}. We identified *not4* in our E3 ligase screen as necessary for worm survival, building on a previous report that identified *not1* as an essential CCR4-NOT factor in planarians²¹. Intriguingly, *not1(RNAi)* demonstrated a phenotype similar to that of *prpf19(RNAi)* or *smedwi-2(RNAi)* where a phenotype that is typical of stem cell depletion is observed despite stem cells being maintained. A reduction in the density of epidermal progenitors was also observed in *not1(RNAi)* worms like the results we observed in *prpf19(RNAi)* treatments. It would be interesting to examine the *not4(RNAi)* phenotype in greater detail to establish if the phenotype involves a depletion of the stem cell population or shares mechanism with *not1*, where stem cells are maintained but their ability to differentiate is impaired. The phenotypes reported in this work and others point to the post-transcriptional processing of mRNA as a major mechanism by which the proper differentiation of planarian stem cells is directed and that impairment of differentiation presents phenotypes that are analogous to those caused by a loss of stem cells.

The epigenetic regulation of genes is an essential developmental process that is necessary to determine and maintain cellular identity. The modification of histones is one mechanism of epigenetic regulation and includes ubiquitylation. The addition of ubiquitin onto a histone can be an activating or repressive mark depending on the context. In our screen of RING E3 ligases we recovered two genes that are homologs of factors that target histones for ubiquitylation. We found that the planarian homolog of *bre1*, which targets histone H2B for ubiquitylation and is associated with transcriptional activation²², when

inhibited exhibited head regression and epidermal lesions, failed to regenerate, and reduced levels of bulk ubiquityl histone H2B. Ubiquitylation of histone H2A is associated with gene repression and this ubiquitylation is catalyzed by the RING *rnf2* acting within PRC1^{23,24}. Inhibition of *rnf2* reduced levels of ubH2A and had a phenotype of delayed or impaired regeneration with a moderate penetrance.

PRC1 is a major developmental repressive complex that was first identified as a regulator of HOX genes²⁵. The canonical complex is comprised of four core subunits, a RING and PCGF that form a dimer that associates with the chromatin binding subunit CBX and a PHC subunit that is necessary to form higher-order chromatin structures²⁶. In vertebrates, variant forms of PRC1 are responsible for the preponderance of H2A ubiquitylation²⁷, and we found that inhibition of cPRC1 factors *cbx* and *phc* did not affect levels of ub-H2A. It had previously been thought that invertebrates contained only cPRC1, but this conclusion was based on limited evidence from only a few model species and more recent phylogenetic analysis indicates that variant forms of PRC1 evolved as early as cnidarians.²⁸ Our analysis of ub-H2A levels following PRC1 inhibition suggest that in planarian the *cbx* and *phc* subunits are dispensable for the ubiquityl ligase function of *rnf2* in cPRC1 or that vPRC1 conformations exist in *S. mediterranea* that are responsible for the bulk H2A ubiquitylation.

In contrast to the relatively mild effects on worm regeneration that was observed in *rnf2(RNAi)* treatments, inhibition of *phc*, and to a lesser degree *cbx*, had a dramatic and robust phenotype on worm patterning that involved the formation of a lesion on the dorsal surface of the worm just anterior to the base of the pharynx. In some instances, it was observed that the pharynx emerged from this lesion and was ectopically located on the dorsal surface. This phenotype suggested a dysregulation of patterning and tissue

specification that we examined further by observing the effect *phc(RNAi)* had on marker gene expression. We found that *phc* inhibition led to a loss of laminin expression in the pharynx and a loss of *NB.22.1E* expression in a population of cells near the base of the pharynx.

To identify the transcriptional targets of PRC1 in planarians and to understand the basis of the discrepancies in phenotypes observed when different core elements of PRC1 were perturbed, we performed RNA-seq after *rnf2(RNAi)* and *phc(RNAi)*. Our RNA-seq results agree with PRC1 being a transcriptional repressor as more genes were upregulated than downregulated after PRC1 inhibition. The data sets shared only a single common factor between them, which indicates that *rnf2* and *phc* are largely regulating separate transcriptional networks and that this difference is a probable basis for the phenotypic contrasts observed after inhibition of each gene. This discrepancy in transcriptional targets and phenotype for *rnf2* and *phc* is somewhat surprising given the deep conservation of this complex and of H2A ubiquitylation in animals. It is possible, however unlikely, that in planarians *rnf2* and *phc* do not function in the same complex, and thus, would regulate different genes. A more likely explanation is that the *rnf2* phenotype, especially the loss of ub-H2A, is mediated through vPRC1 while the phenotype for *phc* is mediated through cPRC1 and that, similar to vertebrate models, cPRC1 has a minimal role in ubiquitylating H2A. As RNF2 is essential for the formation of both variant and canonical PRC1 complexes *rnf2(RNAi)* should encompass the phenotypes observed in *phc(RNAi)* but unexpectedly do not. This could be the result of experimental methodologies, as we use the feeding of dsRNA to the worms to induce the RNAi pathway and cause a reduction in gene transcript levels, and this approach results in gene knock down rather than a genetic knock out

condition. While we assayed knock down efficiency using qPCR and saw a robust reduction in transcript levels early in treatment for both genes, there could exist differences in protein perdurance that might allow enough residual RNF2 to exist in our treatments that allowed the formation and function of cPRC1 with PHC. This would imply that vPRC1 function is more sensitive to a reduction in *rnf2* levels and be enough to cause a loss of ubH2A. It is also important to note that we found two genes that are predicted to be homologs of mammalian RNF2 and RING1, both of which function in mammalian PRC1 as E3 ligases. We note that these planarian genes are probably not direct homologs for each of the vertebrate PRC1 RINGs, but rather likely reflect independent duplication events in both lineages. Based on our examination of ubH2A levels after RNAi treatments we concluded that *Smed-rnf2* is the major E3 ligase that ubiquitylates H2A and that the contribution of *Smed-ring1* has a minor, if any, effect on ubH2A. Knockdown of *ring1* had a phenotype that was similar to that *rnf2*, with delayed or impaired regeneration seen with incomplete penetrance. These genes could be compensating for each other in cPRC1 when the other is knocked down, preventing the manifestation of the same phenotypes seen in *phc(RNAi)*. We did perform double knockdown experiments where we targeted both *rnf2* and *ring1* for inhibition and did not observe any phenotypic effects in addition to those observed from single knock down experiments. It remains a possibility that the incomplete knock down nature of our RNAi experiments allows enough gene transcript to persist to allow adequate formation of cPRC1.

Intriguingly, the gene that was shared between the *rnf2* and *phc* RNA-seq data sets was the cPRC1 chromatin binding factor *cbx*. This gene was upregulated in both data sets and was the most significantly upregulated gene after *phc* inhibition. This suggests that

PRC1 could auto-regulate its own activity or that disruption to PRC1 function induces a compensatory response to attempt to repair disrupted chromatin states. The presence of other chromatin regulators and modifiers in both data sets supports the hypothesis that the disruption of an epigenetic factor can induce a cascade of epigenetic changes.

We performed GO analysis on the genes that were upregulated following *rnf2* inhibition and found terms that were related to the cellular stress response, and included, response to hypoxia, cellular response to decreased oxygen levels, ATF6-mediated unfolded protein response, regulation of transcription from RNA polymerase II promoter in response to stress, chaperone cofactor-dependent protein refolding, protein folding in endoplasmic reticulum, and protein refolding. This analysis suggests that *rnf2* in planarians is required to keep cellular stress response genes repressed during homeostasis. The plastic and responsive nature of epigenetic regulation makes it an attractive candidate for a regulator of cellular stress responses, as this response must be induced by disruptive stimuli and must be reversible after an environmental insult has subsided for the cell return to homeostasis. A hyperactive stress response may be disadvantageous for a cell, especially from an energetics perspective, but is not likely to have a majorly deleterious effect. Our examination of differentially expressed genes from the *rnf2(RNAi)* RNA-seq data using WISH demonstrated that *rnf2* is likely modulating target gene expression levels within tissues that normally expressed a given gene rather than suppressing ectopic expression. This is consistent with a role for *rnf2* in regulating intracellular response pathways, including stress responses. Taken together the GO analysis and examination of target genes via WISH argues that the action of *rnf2*, and potentially H2A ubiquitylation, works to tune transcriptional levels within a cell type, particularly related to stress

response pathways. A role for *rnf2* in adjusting the expression of stress response pathway genes is consistent with the relatively mild phenotypes that were observed after *rnf2* inhibition as an inappropriately elevated stress response could act as an impairment to proper regeneration.

In contrast to the subdued changes in gene expression observed by WISH after *rnf2(RNAi)*, when we examined genes regulated by PHC we saw drastic spatial shifts in expression. Strikingly, these spatial shifts were concentrated in the region near the base of the pharynx that was phenotypically most affected by *phc* inhibition. The transcripts that we found ectopically expressed near the base of the pharynx included factors that regulate cellular specification, including nuclear receptors, transcription factors, and chromatin modifiers. The observed changes in expression of extracellular matrix and intercellular adhesion molecules, both up (*intercellular adhesion molecule 5*) and down regulated (*pikachurin*), after *phc* inhibition are likely involved in the formation of the lesion anterior to the pharynx. Thus, the action of *phc* is necessary to maintain proper tissue identity by repressing factors that could cause the misspecification of stem cells.

Several of the genes from our *phc* RNA-seq data regulate gene transcription and have identified roles in development. Nuclear receptor ROR-alpha ($ROR\alpha$) belongs to the orphan class of nuclear receptors that act as ligand-dependent transcription factors²⁹. $ROR\alpha$ has described roles during development in other organisms that includes the regulation of *sonic hedgehog* (*shh*) signaling to specify Purkinje cells in cerebellar development³⁰. Planarian *hedgehog* (*hh*) has been identified as being expressed in a population of ventral medial neurons and to be required for the normal production of neural progenitor cells³¹. It would be interesting to determine if *rora* has a conserved role

in directing neuronal specification in planarians and if this action is mediated through *hh* signaling. The nuclear factor Onecut1 (also Hepatocyte Nuclear Factor 6) is a CUT and homeobox domain-containing transcription factor that promotes hepatocyte proliferation, remodels chromatin accessibility, and promotes tumor growth in colorectal cancers³²⁻³⁴. The role of this gene in regulating transcription and chromatin accessibility in other models suggests that the ectopic expression we observe for this factor in *phc(RNAi)* planarians could be transforming tissue identity near the base of the pharynx. Further work will be needed to determine what the contribution of *onecut1* is to the *phc(RNAi)* phenotype, which would include using RNAi to determine if inhibition of *onecut* is suppressive towards the *phc(RNAi)* phenotype. Examining the role and transcriptional targets of *onecut1* in planarians using RNAi and RNA-seq would help elucidate how *onecut1* mis-expression is influencing cell differentiation in *phc* knockdown worms and potentially uncover which cell types are contributing the phenotype.

Regeneration is a dynamic process that involved the integration of wound signals and positional information to re-specify lost body parts. The regulation of protein function by the post-translational modification of ubiquitylation is an important but understudied phenomenon in regenerative biology. Here, we screened the function of a large class of ubiquityl E3 ligases and uncovered roles for these genes in regulating planarian biology. The identification of *prpf19* as factor dispensable for stem cell maintenance but necessary for worm survival and progenitor specification points to a key role in the processing of RNA as a regulator of stem cell differentiation in planarians. Epigenetic histone modifiers are an attractive mechanism for regulating regeneration as the marks are plastic and can be adapted to respond to a variety of situations. We found that the epigenetic repressor PRC1

had distinct effects depending on which subunit, *rnf2* or *phc*, was perturbed. To understand this difference, we performed RNA-seq after inhibition to uncover transcriptional targets and found little overlap in the data sets, suggesting these factors regulate independent processes and providing support for the existence of vPRC1 and cPRC1 complexes in invertebrates. We found that *rnf2*, and potentially its E3 ligase activity, was regulating stress-response factors, and we found that *phc* was necessary to repress the expression of several genes, including chromatin regulators and transcription factors, to properly pattern the region of the worm anterior to the pharynx. Continuing work will further dissect the function of PRC1, especially with the use of assays like ChIP-seq or ATAC-seq, to understand how chromatin is adapted to support a robust regenerative response. Especially interesting follow up work will leverage advancements in ChIP methodologies that have lower background and input requirements. These new methods, including Cut&Tag³⁵, will allow chromatin assays to be performed on limited cell populations like those of the regeneration blastema. The application of these new methods will allow us to understand how chromatin marks are shifting during cell differentiation in regeneration.

Despite the critical role of epigenetic factors in development, to date comparatively little work has been performed on studying epigenetic factors in planarians and the work presented here represents a major advancement of our understanding of ubiquitin signaling in planarian regeneration.

References

- 1 Sanchez Alvarado, A. & Tsonis, P. A. Bridging the regeneration gap: genetic insights from diverse animal models. *Nat Rev Genet* **7**, 873-884, doi:10.1038/nrg1923 (2006).
- 2 Allen, J. M., Ross, K. G. & Zayas, R. M. in *eLS* 1-9 (2016).
- 3 de Dieuleveult, M. & Miotto, B. Ubiquitin Dynamics in Stem Cell Biology: Current Challenges and Perspectives. *Bioessays* **42**, e1900129, doi:10.1002/bies.201900129 (2020).
- 4 Endoh, M. *et al.* Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. *PLoS Genet* **8**, e1002774, doi:10.1371/journal.pgen.1002774 (2012).
- 5 Ciechanover, A., Finley, D. & Varshavsky, A. Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell* **37**, 57-66, doi:10.1016/0092-8674(84)90300-3 (1984).
- 6 David, Y. *et al.* E3 ligases determine ubiquitination site and conjugate type by enforcing specificity on E2 enzymes. *J Biol Chem* **286**, 44104-44115, doi:10.1074/jbc.M111.234559 (2011).
- 7 Zheng, N. *et al.* Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**, 703-709, doi:10.1038/416703a (2002).
- 8 Chen, B. B., Glasser, J. R., Coon, T. A. & Mallampalli, R. K. F-box protein FBXL2 exerts human lung tumor suppressor-like activity by ubiquitin-mediated degradation of cyclin D3 resulting in cell cycle arrest. *Oncogene* **31**, 2566-2579, doi:10.1038/onc.2011.432 (2012).
- 9 Takeishi, S. & Nakayama, K. I. Role of Fbxw7 in the maintenance of normal stem cells and cancer-initiating cells. *Br J Cancer* **111**, 1054-1059, doi:10.1038/bjc.2014.259 (2014).
- 10 Chen, B. B., Glasser, J. R., Coon, T. A. & Mallampalli, R. K. FBXL2 is a ubiquitin E3 ligase subunit that triggers mitotic arrest. *Cell Cycle* **10**, 3487-3494, doi:10.4161/cc.10.20.17742 (2011).
- 11 Chanarat, S. & Strasser, K. Splicing and beyond: the many faces of the Prp19 complex. *Biochim Biophys Acta* **1833**, 2126-2134, doi:10.1016/j.bbamcr.2013.05.023 (2013).
- 12 Song, E. J. *et al.* The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev* **24**, 1434-1447, doi:10.1101/gad.1925010 (2010).
- 13 Mahajan, K. hPso4/hPrp19: a critical component of DNA repair and DNA damage checkpoint complexes. *Oncogene* **35**, 2279-2286, doi:10.1038/onc.2015.321 (2016).
- 14 Marechal, A. *et al.* PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR activation via a ubiquitin-mediated circuitry. *Mol Cell* **53**, 235-246, doi:10.1016/j.molcel.2013.11.002 (2014).

- 15 Roberts-Galbraith, R. H., Brubacher, J. L. & Newmark, P. A. A functional genomics screen in planarians reveals regulators of whole-brain regeneration. *Elife* **5**, doi:10.7554/eLife.17002 (2016).
- 16 Dellago, H. *et al.* ATM-dependent phosphorylation of SNEVhPrp19/hPso4 is involved in extending cellular life span and suppression of apoptosis. *Aging (Albany NY)* **4**, 290-304, doi:10.18632/aging.100452 (2012).
- 17 Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. & Sanchez Alvarado, A. SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* **310**, 1327-1330, doi:10.1126/science.1116110 (2005).
- 18 Eskeland, R. *et al.* Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol Cell* **38**, 452-464, doi:10.1016/j.molcel.2010.02.032 (2010).
- 19 Tucker, M. *et al.* The Transcription Factor Associated Ccr4 and Caf1 Proteins Are Components of the Major Cytoplasmic mRNA Deadenylation Complex in *Saccharomyces cerevisiae*. *Cell* **104**, 377-386, doi:10.1016/s0092-8674(01)00225-2 (2001).
- 20 Collart, M. A. & Panasenko, O. O. The Ccr4--not complex. *Gene* **492**, 42-53, doi:10.1016/j.gene.2011.09.033 (2012).
- 21 Solana, J. *et al.* The CCR4-NOT complex mediates deadenylation and degradation of stem cell mRNAs and promotes planarian stem cell differentiation. *PLoS Genet* **9**, e1004003, doi:10.1371/journal.pgen.1004003 (2013).
- 22 Hwang, W. W. *et al.* A Conserved RING Finger Protein Required for Histone H2B Monoubiquitination and Cell Size Control. *Molecular Cell* **11**, 261-266, doi:10.1016/s1097-2765(02)00826-2 (2003).
- 23 Cao, R., Tsukada, Y. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol Cell* **20**, 845-854, doi:10.1016/j.molcel.2005.12.002 (2005).
- 24 Cohen, I., Bar, C. & Ezhkova, E. Activity of PRC1 and Histone H2AK119 Monoubiquitination: Revising Popular Misconceptions. *Bioessays*, e1900192, doi:10.1002/bies.201900192 (2020).
- 25 Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565-570, doi:10.1038/276565a0 (1978).
- 26 Conway, E. M. & Bracken, A. P. in *Polycomb Group Proteins* 57-80 (2017).
- 27 Fursova, N. A. *et al.* Synergy between Variant PRC1 Complexes Defines Polycomb-Mediated Gene Repression. *Mol Cell* **74**, 1020-1036 e1028, doi:10.1016/j.molcel.2019.03.024 (2019).
- 28 Gahan, J. M., Rentzsch, F. & Schnitzler, C. E. The genetic basis for PRC1 complex diversity emerged early in animal evolution. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.2005136117 (2020).

- 29 Solt, L. A. & Burris, T. P. Action of RORs and their ligands in (patho)physiology. *Trends Endocrinol Metab* **23**, 619-627, doi:10.1016/j.tem.2012.05.012 (2012).
- 30 Jetten, A. M. Retinoid-related orphan receptors (RORs): critical roles in development, immunity, circadian rhythm, and cellular metabolism. *Nucl Recept Signal* **7**, e003, doi:10.1621/nrs.07003 (2009).
- 31 Currie, K. W., Molinaro, A. M. & Pearson, B. J. Neuronal sources of hedgehog modulate neurogenesis in the adult planarian brain. *Elife* **5**, doi:10.7554/eLife.19735 (2016).
- 32 Peng, Y. C., Lv, T. H., Du, Z. K., Cun, X. N. & Yang, K. M. Liver Macrophages Stimulate the Expression of Hepatocyte Nuclear Factor-6 and Promote Hepatocyte Proliferation at the Early Stage of Liver Regeneration. *Bull Exp Biol Med* **170**, 40-45, doi:10.1007/s10517-020-05000-7 (2020).
- 33 van der Raadt, J., van Gestel, S. H. C., Nadif Kasri, N. & Albers, C. A. ONECUT transcription factors induce neuronal characteristics and remodel chromatin accessibility. *Nucleic Acids Res* **47**, 5587-5602, doi:10.1093/nar/gkz273 (2019).
- 34 Jiang, K. *et al.* HNF6 promotes tumor growth in colorectal cancer and enhances liver metastasis in mouse model. *J Cell Physiol* **234**, 3675-3684, doi:10.1002/jcp.27140 (2019).
- 35 Kaya-Okur, H. S. *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* **10**, 1930, doi:10.1038/s41467-019-09982-5 (2019).