Organizer Regeneration and Patterning of Stem Cell Lineages in Planarians

by

Isaac M. Oderberg

B.A., Molecular and Cell Biology (2010)
University of California Berkeley, Berkeley, CA

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

Signature of Author..... Signature redacted

Department of Biology
May 10, 2017

Certified By..... Signature redacted

Peter W. Reddien
Professor of Biology
Thesis Supervisor

Accepted By..... Signature redacted .....

Amy E. Keating
Professor of Biology
Chair, Committee for Graduate Students

Organizer Regeneration and Patterning of Stem Cell Lineages in Planarians

By

Isaac M. Oderberg

Submitted to the Department of Biology on May 10th, 2017 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biology

Abstract

Planarians are freshwater flatworms capable of whole-body regeneration. Like development, regeneration requires the establishment of tissue patterns and the specification of appropriate cell types. However, regeneration has the additional challenges of performing these tasks in the absence of developmental cues, and in the presence of pre-existing, differentiated tissues. Planarian head regeneration involves the anterior pole, which is a cluster of cells in the tip of head required for proper head patterning. We used transplantation to show that the head tip region, containing the anterior pole, has organizing activity. We sought to establish how the anterior pole is placed during regeneration. Anterior pole progenitors are specified medially and accumulate to form a cluster at the DV median plane. Pole progenitors are specified at the pre-existing midline, and coalesce to the DV boundary in the blastema. These findings demonstrate that during its formation, the anterior pole integrates positional information from the pre-existing tissues. This process places the anterior pole in the appropriate location to organize patterning during head regeneration.

Regeneration in planarians involves two important components. Neoblasts are pluripotent stem cells that are the source of all new cells during regeneration. Position control genes (PCGs) are developmental signaling genes that are constitutively expressed in adult planarian muscle, and are thought to provide the molecular instructions for regeneration. How neoblasts respond to PCGs to produce regionally appropriate cells types remains unknown. To better understand this process, we characterized the planarian epidermis. Bulk RNA-sequencing of mature epidermal cells and subsequent *in situ* validation identified numerous spatial patterns. To understand when these patterns arose during differentiation, we performed single-cell sequencing (SCS) of epidermal progenitors. Positional identities present in the mature epidermis were also present in progenitors, with dorsal-ventral identity present in spatially distant neoblasts. Epidermal neoblasts were able change their DV identity upon inhibition of Bmp signaling. This demonstrates that neoblasts can respond to changes in their signaling environment, linking positional information from muscle to the generation of regionally-appropriate cells types.

Thesis Supervisor: Peter W. Reddien
Title: Professor of Biology

Acknowledgments

First and foremost, I would like to thank Peter for giving me the privilege of working in his lab. Peter has fostered a rigorous and supportive lab environment, is an excellent mentor, and of course, is a talented scientist with a perception for the details that matter and an almost uncanny scientific intuition. I have always left our discussions about science more excited about my work than when I entered.

I would like to thank the members of the Reddien lab, past and present, for creating an amazing space to do science. I want to start by thanking those students who came before me: Dan Wagner, Danielle Wenemoser, Sylvain Lapan, Irving Wang, Mike Gavino, and Jared Owen. They took me in when I was just started and answered all of my questions. I also want to thank the postdocs (and staff scientist) for their mentorship: Lucila Scimone, Jana Hersch, Gisele Chen, Mansi Srivastava, and Josien van Wolkswinkel. The lab has changed greatly over the years, but the spirit of exploration and collaborative work has only grown, and I have been so grateful to be a member of this group. I would in particular like to thank Omri Wurtzel for our collaboration on the epidermis project and Jack Li for our work together on the anterior pole.

I would like to thank the members of my thesis committee, Terry Orr-Weaver and Dennis Kim. I have valued their guidance with my Ph.D. projects over the years, their advice on how to proceed after graduate school, and their commitment to seeing me succeed. I would also like to thank Angela DePace, for participating in my thesis defense and stimulating critical discussion.

I also like to thank the many communities that I have had the pleasure to be a part of during my time here at MIT. First, my classmates in the Biology Department (Biograd 2010) for helping me transition to a completely new environment. I would like to thank the members of the Sidney-Pacific graduate residence, for countless experiences during the years I lived there. I would also like to thank the members of Graduate Hillel, as well as the Hillel staff, for helping me find my Jewish home away from home here at MIT. Lastly, I would like to thank the community closest to my heart, the Biomansion. The Biomansion was founded in 2011 as two floors of an apartment complex in Porter Square, but grew into a close-knit community of friends going through life together. We have numerous Saturday lunches, D&D sessions, grill nights and more. We have celebrated the good times together and supported each other through tough times, and I sincerely hope that everyone gets jobs in the area.

I would like to thank Amelie Raz. The seasons we have spent together have been the best, and I have valued her support greatly as I have approached the end of my Ph.D.

Finally, I would like to thank the members of my family. I would like to thank my siblings, Zoe and Sidney, for being there for me whenever I have needed them. I would like to thank my parent, David and Ethel, who have supported me and given me the opportunity to follow my dreams.

Table of Contents

Chapter 1

Introduction

Foreword

Regeneration is the replacement of tissues lost by injury or damage. Regeneration has many of the same challenges as development: fate specification of region-appropriate cells and arrangement of cells into patterned, functional tissues. However, regeneration is unique in that it initiates from injuries that are inherently random, and is able to replace missing tissues and restore form without the patterning events that define early embryos. The goal of my thesis work is to understand how pattern is initiated in regeneration, and how cells interpret patterning information to make appropriate fate choices, without access to developmental cues. The introductory chapter of this thesis will discuss models of tissue development and regeneration to place my work in context. I will discuss mechanisms for pattern initiation and cell fate specification in embryos using vertebrate axial and limb development as examples. I will provide a broad overview of animal regeneration and then focus on two well-studied systems for understanding regeneration, oral regeneration in *Hydra* and limb regeneration in amphibians, in order to explore to what extent embryonic patterning processes are reused during regeneration. Finally, I will explain why planarians are an excellent model system for studying regeneration, and discuss recent findings about how patterning and stem cell specification operate in planarians. How patterns are initiated and interpreted is a fundamental question in regeneration, and my work seeks to elucidate how planarians address these challenges and are able to restore form robustly following a multitude of injuries.

I. Pattern Formation in Development

Pattern Initiation and Symmetry Breaking

Embryology and theoretical models of patterning

A central question in development is how the information encoded in a group of genetically identical cells is interpreted reliably to generate a complete organism. To understand this process, scientists have performed a number of experiments manipulating embryos. These experiments indicated that the developmental trajectory of early embryos was not, in most cases, completely determined, but rather capable of regulation after perturbation. If the blastomeres of a sea urchin embryo are separated at the two or four cell stage, each resulting individual blastomere is capable of generating a complete individual [1,2]. When the German embryologist Hans Spemann bisected an amphibian embryo along the DV axis, both halves developed into a complete, albeit smaller, animal. He noted, however, that if he bisected the embryo so that only one half contained the dorsal blastopore lip, only that piece would develop completely, and the other half remained a "belly-piece" that lacked all axial organs (Figure 1A) [3]. The property of embryos to adapt to these manipulations is termed self-regulation. Self-regulation grants robustness to development, and enables it to produce a complete and functional animal despite genetic or environmental changes.

In 1924, Hans Spemann and Hilde Mangold performed a remarkable experiment in which they transplanted the dorsal blastopore lip of one amphibian embryo onto another and this resulted in the generation of an ectopic body axis (Figure 1B). Part of the ectopic axis was made from host tissues, demonstrating that the donor graft had

"induced" a change in the host. Spemann coined the term "organizer" to describe the dorsal blastopore lip as a group of cells containing powerful inductive properties [4], and received the Nobel Prize in medicine or physiology in 1932 for this discovery.

How could a small group of cells exert such an influence on their neighbors, and in such way to establish a stable pattern capable of self-regulation? In 1952, the British mathematician Alan Turing proposed that "a system of chemical substances, called morphogens, reacting together and diffusing through a tissue, adequate to account for the main phenomena of morphogenesis" (Figure 1C-D) [5]. Turing was primarily concerned with how morphogens could interact with each other to generate stable patterns from relatively homogenous starting conditions. Later, Turing's model was improved upon by Gierer and Meinhardt, whose model included a short-range activator and a long range inhibitor, both of which are diffusible (Figure 1C-D) [6]. In 1968, Lewis Wolpert proposed that cells interpret morphogen concentrations as positional information [7]. In Wolpert's "French Flag model," cells determine their position in the animal by interpreting the morphogen concentration at their position, and adopt the appropriate cell fate (Figure 1E). This system is robust to removal or addition of tissue, as well as variations in size, is thus self-regulating [7].

It would be decades before the physical underpinnings for these conceptual advances would be identified. Research in the molecular era revealed that organizers do indeed secrete morphogens, which are responsible for providing instructions to cells. Cells respond to morphogens in a concentration-dependent manner to drive gene expression that dictates the fate of the cell. Molecular interactions between morphogens were found to account for the property of self-regulation. While in principle, Turing's

proposition that random fluctuations in a reaction-diffusion system could lead to stable patterns, in reality embryos are often already polarized even as a zygote. Two examples of how embryos become polarized, and thus possess a template for patterning, will be examined here: maternal determinants and the point of sperm entry.

Maternal determinants

In *Drosophila* development, the oocyte is asymmetric along the anterior-posterior axis with regards to the presence of *bicoid* (*bcd*) mRNA at the anterior end (Figure 1F). *bcd* was identified in a genetic screen in *Drosophila* for recessive female-sterile mutations, and thus was classified as a maternal-effect gene [8]. *bcd* mRNA is transcribed in nurse cells before being transported into the developing oocyte. Both the transport into the oocyte, as well as the localization of *bcd* within the oocyte to the anterior end, require microtubules [9,10].

BCD is present in a gradient from the presumptive anterior end of the embryo, and drops to background levels approximately halfway between the anterior and posterior termini [11]. Progeny of *bcd* mutant mothers fail to specify properly the anterior half of the animal, indicating a requirement for *bcd* in the establishment of positional identity [8]. Moreover, transplantation of the anterior tip cytoplasm from wild-type embryos is sufficient to rescue AP pattern throughout the embryo [8], demonstrating that the protein acts at a distance from the anterior end. BCD is a transcription factor that elicits concentration-dependent specification by regulating expression and translation of its target genes [12,13]. In this way, a maternally-deposited mRNA in the oocyte enables the specification of the anterior-posterior axis in *Drosophila*.

Furthermore, the discovery of *bcd* represented one of the first examples of a morphogen as described by Turing decades earlier.

The point of sperm entry

In *Xenopus* development, the egg is radially symmetric about the animal-vegetal axis. However, the frog embryo will ultimately have three defined body axes: an anterior-posterior axis, a dorsal-ventral axis, and a medial-lateral (left-right) axis, and thus this radial symmetry is broken during development. The first visible sign of deviation from radial symmetry is the dorsal crescent [14], which is an unpigmented region of the egg that marks the presumptive dorsal side of the embryo. Like other amphibian embryos, if a *Xenopus* embryo is bisected into a dorsal and a ventral half, only the piece containing the dorsal crescent will be able to generate a complete tadpole with dorsal axial structures [15]. This suggests that dorsal determination is associated with the dorsal crescent. The generation of the dorsal crescent, and thus the breaking of radial symmetry, is facilitated by sperm entry (Figure 1G) [16].

Sperm entry initiates a process known as cortical rotation, during which the egg cortex rotates about 30° with respect to the yolky cytoplasm (Figure 1H). The plus ends of sperm aster microtubules extend into the vegetal cortex, directing rotation away from the point of entry. Thus, cortical rotation reveals the unpigmented dorsal crescent opposite to the point of sperm entry [17]. If microtubules are destabilized, by either inhibitors or UV irradiation, the embryo develops as a belly-piece, indicating that microtubule activity is required for later events of axial specification [18]. Cortical microtubules, in addition to mediating cortical rotation, transport membrane vesicles from the vegetal pole to the dorsal side of the embryo [19]. These vesicles contain

dorsal determinants that activate the canonical Wnt signaling pathway, and thus the nuclear localization of β-catenin, on the presumptive dorsal side of the embryo [20,21]. After the midblastula stage, the expression of *Xenopus* Nodal-related (Xnr) genes in the endoderm helps mediate induction of the mesoderm. The nuclear localization of β-catenin in the dorsal endoderm creates a gradient of Xnr activity, which is highest dorsally. The combination of high Xnr activity, in conjunction with other mesodermal signals, is responsible for the induction of the Spemann-Mangold organizer on the dorsal side [22]. In this way, a physical event, sperm entry, is utilized to establish the signaling centers that have organizing capacity (Figure 1H). These centers are then utilized for the specification of the vertebrate axes, as described below.

Vertebrate Axis Specification

The gastrulation organizer

While the Spemann-Mangold organizer (also called the gastrulation organizer) was first discovered in amphibians [4], similar structures have been characterized in chick [23,24], mouse [25], and zebrafish [26,27], indicating that the presence of a gastrulation organizer is shared among vertebrates [28]. Gastrulation organizers are temporary structures, existing in a defined location and time during development. By secreting morphogens, they are responsible for patterning ectodermal and mesodermal derivatives. Rather than providing instructions to a few cells, the gastrulation organizer signals to a field of cells, which alter their behavior in response to morphogen concentration to generate the appropriate tissue patterns.

Dorsal-ventral patterning

The molecular characterization of the Spemann-Mangold organizer, and its spatially opposed ventral center, revealed much about how the organizer robustly confers patterning. Ultimately, the organizer's role in regulating dorsal-ventral pattern operates through the action of bone morphogenic proteins (Bmps) (Figure 2A). Bmps are a group of extracellular signaling molecules that belong to the TGF-β superfamily. Bmp ligands bind to cognate receptors, leading to receptor dimerization and phosphorylation of Smad1/5/8, the intracellular effector of the pathway. Together with Smad4, Smad1/5/8 translocates to the nucleus and activates target genes [29]. On the ventral side of the embryo Bmp activity is high and drives expression of ventral target genes. Conversely, on the dorsal side, low Bmp activity drives expression of dorsal target genes [30].

The Spemann-Mangold organizer expresses the secreted Bmp antagonists Chordin, Noggin, and Follistatin, as well as the secreted Wnt antagonists Frzb-1, sFRP-2, crescent, and Dickkopf-1 [30]. Chordin in particular is important for self-regulation; it binds to Bmp ligands and creates a diffusible complex incapable of binding to Bmp receptors. The Bmp ligands Admp and Bmp2 are also expressed on the dorsal side, but because of the high concentration of Chordin dorsally, they only signal ventrally [31]. The ventral center also expresses a number of secreted proteins: the ligands Bmp4 and Bmp7, which activate Bmp signaling ventrally; Xolloid-related, a metalloproteinase that cleaves Chordin to release Bmp ligands; and the sFRP-like protein Sizzled, which functions as a feedback inhibitor of Bmp signaling by inhibiting Xolloid-related (Figure 2A) [30].

How does this exquisite protein-protein interaction network confer the self-regulatory properties observed in embryos? Even though the dorsal and ventral centers

secrete both Bmp ligands and their antagonists, the network acts to maintain a gradient of Bmp activity that is robust to perturbation and operates over multiple size scales. The key lies in the reciprocal transcriptional control of gene expression from the two centers (Figure 2B). If Bmp levels increase, transcription of antagonists such as Sizzled are increased and lead to a reduction in Bmp activity. If Bmp levels drop, Admp expression increases, and this leads to an increase in Bmp activity [15]. This network also offers an explanation for the ability of dorsal halves to reconstitute an entire organism [3]. In a dorsal half, Bmp4 and Bmp7 from the ventral side will have been removed, and so Admp expression will increase. Chordin will bind Admp and diffuse to the ventral side of the animal, where Xolloid-related is highest. Xolloid-related will cleave Chordin, releasing Admp to signal through Bmp receptors, which will drive expression of Bmp4 and Bmp7, and generate a new ventral center. The interactions between the dorsal and ventral centers will then scale the Bmp gradient to the size of the new animal, resulting in perfectly proportioned, albeit smaller, animal.

Anterior-posterior patterning

Whereas DV patterning for many animals uses Bmp signaling, the AP axis is patterned using Wnt signaling, where the anterior domain is Wnt-low and the posterior is Wnt-high [32]. In the canonical pathway, extracellular Wnt ligands bind to Frizzleds, which are seven-transmembrane receptors. β-catenin, the intracellular effector of the pathway, is normally sequestered and destroyed by a complex composed of GSK-3β, APC, and Axin proteins. Ligand binding stabilizes β-catenin and promotes its translocation to the nucleus where it regulates target gene transcription with Tcf/Lef transcription factors.

Like Bmp signaling, Wnt signaling is modulated by a number of secreted antagonists, such as sFRPs or Dickkopf-like proteins [32].

The Spemann-Mangold organizer expresses a number of these antagonists, and as such, it also marks the presumptive anterior of the animal [31]. Administration of Wnt inhibitors (along with TGF-β inhibitors) to the ventral side of the animal can induce ectopic heads, consistent with antagonism of Wnt signaling being required for head formation. Overexpression of secreted Wnt antagonists anteriorizes embryos and conversely, overexpression of Wnts or chemical activation of the pathway cause repression of anterior fates and promotion of posterior fates. Finally, an endogenous gradient of β-catenin signaling is present along the AP axis and is required for neuroectoderm patterning in a concentration dependent manner [33].

Hox genes are highly conserved homeodomain-containing transcription factors that are regionally expressed during development along the AP axis [34]. During primary axis formation, the gradient of Wnt expression from the posterior and Wnt antagonists in the anterior is utilized to initiate region-appropriate Hox gene expression [33,35,36]. Hox genes mark positional identity along the AP axis for paraxial mesoderm, lateral plate mesoderm, neuroectoderm, neural crest, and endoderm. During subsequent events of development, Hox genes drive gene expression programs required for the specification of tissue identity at particular AP positions [34].

Neural induction

In order to understand how cells interpret patterning information to generate region-appropriate cell types, it is again necessary to refer to the pioneering work of Hans

Spemann and Hilde Mangold. Their organizer transplantation experiments were performed using two different newt species, one pigmented and the other unpigmented, allowing them to trace the fate of the transplanted tissues. When they examined the secondary axes generated by transplantation, they observed that the donor piece gave rise to the ectopic notochord, whereas the ectopic central nervous system was derived from the host. That the mesodermal organizer could signal to the ectoderm and convert it to a neural fate was termed "induction" [4]. This observation was corroborated by later studies using lineage tracing [37].

During development, ectoderm will give rise to neural tissue or epidermal tissue [30]. Up until gastrulation, ectoderm is competent to become either cell type; after gastrulation, cells are committed to a particular fate [38]. Experimental manipulations revealed the relationship between this fate choice and Bmp signaling. Reduction of Chordin, Noggin, and Follistatin simultaneously resulted in a dramatic failure in the specification of neural tissues and expansion of epidermal tissues [39]. Simultaneous inhibition of Bmp2/Bmp4/Bmp7 and Admp resulted in the opposite phenotype: the entire ectoderm was converted into neural tissue at the expense of epidermal tissue [15]. Thus high Bmp activity promotes epidermal fates, whereas low Bmp activity promotes neural fates. In this way, the Bmp gradient established by dorsal and ventral centers of the embryo is interpreted by the ectoderm to specify cell fates along the dorsal-ventral axis (Figure 2C).

Vertebrate Limb Development

During early development, a group of cells capable of responding to a set of common positional signals to pattern an axis is referred to as a "primary morphogenetic field" [7].

In the case described above, cells of the mesoderm and ectoderm respond to a set of cues established by the dorsal and ventral signaling centers. This principle is reused during development in the context of "secondary morphogenetic fields," which are used to pattern the organ systems of the animal [40]. In fact, the term "morphogenetic field" was coined by the American embryologist Ross Harrison in his study of amphibian forelimb development. He found that a circular region of lateral plate mesoderm (LPM) would induce the formation of ectopic forelimbs when transplanted. When he cut this region in half, each half was capable of creating an entire limb [41]. Thus, an exploration of limb development represents another opportunity to discuss how a field of cells are patterned by organizing centers.

Limb bud initiation

Visible limb development begins with the formation of the limb bud, which is an outgrowth of mesenchymal cells from the lateral plate mesoderm (LPM) that runs along the flank of the embryo, and is covered by the surrounding ectoderm (SE) (Figure 3A). The distal tip of the limb bud is referred to as the apical ectodermal ridge (AER) and the distal posterior mesenchyme is referred to as the zone of polarizing activity (ZPA) (Figure 3A). Molecular and cellular interactions between intermediate mesoderm (IE), the AER, the ZPA and the SE pattern the limb along the anterior-posterior (AP), dorsal-ventral (DV) and proximal-distal (PD) axes [42]. Additionally, the signaling pathways controlling each axis interact with each other in order to coordinate the growth of the limb along all three axes [43]. The discussion here will focus on the initiation of the forelimb.

The first molecular signature for the forelimb is the expression of the T-box transcription factor Tbx5 in the LPM [44]. How Tbx5 expression was regulated to be present only at the appropriate AP position remained elusive for many years. If an impermeable barrier was inserted between the somites and the LPM, forelimb initiation was abrogated [45]. Additionally, transplantation experiments revealed that the cells fated to produce the limb bud were determined many steps prior to limb bud outgrowth [44,46]. Although Wnt signaling and retinoic acid (RA) signaling are required for the transcription of Tbx5, the AP position of Tbx5 is determined by a combinatorial Hox code established earlier in development (Figure 3B) [47,48].

Tbx5 expression is absolutely required for forelimb initiation, as Tbx5 knockouts fail to develop limbs [49]. Tbx5 expression drives activation of Wnt2b, which signals through β-catenin to activate Fgf10 [50]. While the LPM is initially epithelial, Tbx5 and Fgf10 expression drive an epithelial-to-mesenchymal transition (EMT) that is required for limb bud outgrowth [51]. Fgf10 signals to the surrounding ectoderm, where it induces expression of Wnt3a, which in turn activates Fgf8 in the surrounding ectodermal space (Figure 3C) [50]. The onset of expression of Wnt3a and Fgf8 in the ectoderm coincides with the formation of the AER [50], which acts as a signaling center for the duration of limb development.

Discrete signaling centers regulate axial patterning

Once the AER has formed, a circuit between Fgf8 expression in the AER and Fgf10 expression in the LPM drives limb bud outgrowth (Figure 3D) [50]. In addition to driving growth and proliferation, signals from the AER also have critical roles in PD patterning. Along the proximal-distal axis, the limb has a segmented skeletal structure, containing a

proximal stylopod, intermediate zeugopod, and distal autopod. Recent studies have led to the proposition of the "two-signal model" for PD patterning [52]. Initially, cells of the limb mesenchyme are exposed to RA signaling from the intermediate mesoderm, which drives expression of Meis homeobox transcription factors that promote proximal cell fates and stylopod identity [53,54]. As the limb grows, FGF signals from the AER promote distal cell fates and autopod identity (Figure 3E) [55]. The domain containing the zeugopod is then specified by interactions between the proximal and distal domains [54].

The ZPA, which is located at the distal posterior edge of the limb mesenchyme, was identified in a transplantation experiment in which it was grafted to the anterior of a host limb bud. The resulting limb possessed mirror symmetry along the AP axis, demonstrating that the ZPA was critical for AP polarity [56]. Later, molecular studies revealed that the morphogen sonic hedgehog (Shh) was responsible for this activity [57]. Initially, the AP axis of the presumptive limb bud is regulated by a circuit between the expression of the transcription factors Gli3 (anterior) and Hand2 (posterior) [58], with Hand2 ultimately inducing Shh expression in the ZPA (Figure 3F) [59]. Similarly to its role specifying the location of the limb bud, Hox gene expression is used to generate the initial AP polarity within the limb[60].

For the DV axis, signals from the dorsal and ventral ectoderm have been implicated in patterning. The dorsal ectoderm is defined by Wnt7a expression, and elimination of Wnt7a results in ventralized limbs [61]. Conversely, engrailed-1 (En-1) is expressed ventrally and its absence is marked by dorsalized limbs [62]. Wnt7a induces expression of Lmx1b in the dorsal mesenchyme, where it defines a dorsal compartment

[63], and is both necessary and sufficient to promote dorsal limb identity [64,65]. Interestingly, there is evidence of molecular interaction between the AER, the ZPA, and the dorsal ectoderm to regulate morphogenesis of the limb in three dimensions (Figure 3G) [66]. Once the molecular coordinates of the limb have been established, they can be interpreted to guide appropriate specification of cell fate.

Cell fate specification during limb development

During limb outgrowth, signaling from the AER maintains a population of multipotent progenitors in the distal region of the limb bud. These cells will undergo coordinated differentiation to produce the limb connective tissues, including cartilage, tendons, and dermis [67–69]. The differentiation of these multipotent cells is regulated spatially; cartilage precursors (chrondocytes) will coalesce to the center of the bud to from a chondrogenic core, which will provide the template for later bone development. The soft connective tissue types (tendon, muscle connective tissue and perichondrium) will be specified in the periphery around this core [67,70]. Cells migrating from other regions in the embryo will ultimately form muscle, blood vessels and nerves [71].

The divergence in cell fate specification from a common cell pool is achieved by Wnt signaling from the surrounding ectoderm and FGF signaling from the AER. The distal mesenchymal cells that receive both signals remain proliferative and undifferentiated. As the limb bud grows in all dimensions, a population of cells in the center, receiving neither Wnt nor FGF signals activates expression of Sox9, which promotes chondrogenic differentiation. Cells located close to the surrounding ectoderm, but not at the distal tip, receive a Wnt signal alone, which promotes proliferation and specifies cells towards soft connective tissue fates [72]. Myogenic precursors will

migrate from the somite into the limb bud and form a pattern based on the guide connective tissue [73], a process that is dependent on the PD gradient of RA and FGF signaling [74]. TGF-β signaling from muscle and chrondrocytes induces expression of the transcription factor Scleraxis (Scx) in neighboring soft connective tissue progenitors, which is required for tendon cell fate [75,76]. Soft connective tissue progenitors can also differentiate into muscle connective tissues (MCT), which are cells that surround and ensheath muscle fibers. While the exact molecular instructions guiding MCT differentiation are unclear, it appears that Wnt signaling is important [73,77,78]. Neurons extend processes from the spinal cord into the developing limb bud to innervate the skeletal muscles, which requires signaling driven by the expression of En-1 [79]. In this way, the signaling environment set up early in development promotes the cell fate specification of the diverse tissues that make up the limb.

Development proceeds from an invariant starting point, utilizing cues such a maternal determinants or the point of sperm entry to generate asymmetry along an axis. These cues lead to the establishment of organizing regions that secrete morphogens to generate long range gradients. A field of cells can interpret these gradients as positional information and make decisions regarding fate specification. These mechanisms are used throughout multiple stages of animal development to create proper form and pattern. Central to understanding regeneration is knowing how and when these mechanisms are reused to for patterning in adults.

II. Pattern Formation in Regeneration

Animal Regeneration

Regeneration is the replacement of lost body parts and tissues to restore proper form and function. Regeneration occurs as part of natural homeostasis, in the replacement of cells lost from cell death, aging, and daily wear and tear. For many animals, it can also occur after an injury, such as an amputation or internal organ damage. Central questions in the field of regeneration are [80]: (1) what are the signals required to initiate and terminate the regenerative response following injury? (2) what are the cellular sources for new material? (3) how are the proliferation and patterning of new tissues regulated? For many animals, the process of regeneration shares many interesting commonalities with development, including the generation of region-appropriate cell types, and the spatial arrangement of the cell types to produce functional organs. Regeneration also has the additional challenge of integrating newly formed tissues with pre-existing tissues.

The capacity for regeneration varies broadly across the animal kingdom. Mammals and birds have relatively limited capacity for regeneration [81], whereas lizards and juvenile frogs can regenerate their tails [82,83]. Some arthropods can regenerate limbs or antennae [84,85], and zebrafish, urodele amphibians, and some molluscs can regenerate limbs, internal organs, and even their retinas [80]. Some animals are even capable of whole-body regeneration. Annelids can replace entire body segments [86], starfish can regenerate their bodies from a single limb [87], and planarians are can restore form from a tiny fragment [88]. Hydra can also regenerate

from any fragment after injury, and in a most impressive feat of regeneration, form a new animal after being dissociated into a single cell suspension and reaggregated [89].

Replacement of tissues after injury in many animals proceeds by the formation of a regeneration blastema. A blastema is an outgrowth of cells that appears after injuries, and will ultimately differentiate to produce the lost tissues. Animals that use blastema formation during regeneration include planarians [90,91], polychaete annelids [92], crustaceans [93], echinoderms [94,95], as well as zebrafish, urodele amphibians, and lizards [96]. In principle, there are many ways missing cells can be generated. Some animals use a population of stem cells, which are cells that can both self-renew and produce one or more differentiated cell types, to replace missing tissues [97]. Additionally, some animals utilize dedifferentiation, a process by which a mature cell loses some of its differentiated characteristics, proliferates to generate more cells, and subsequently differentiates into more cells of the same or different type [98]. Yet another way is transdifferentiation, in which cells change into a cell type of a different lineage without going through a stem-like state [98,99].

Which cells are responsible for providing the molecular instructions during regeneration? Organizers are temporary structures that are crucial for patterning during development; a recent hypothesis proposed is that adult tissues can gain organizer activity in order to guide regeneration [100]. In particular, regeneration blastemas orchestrate similar cellular behaviors as developmental organizers, including proliferation, migration, and differentiation into appropriate cells types [81]. Additionally, regeneration blastemas, when transplanted, have the capacity to instruct surrounding tissues to differentiate and forms ectopic structures [101]. However, early and late

blastemas appear to have different potentials, as late blastemas are only capable of self-repair and do not recruit surrounding tissues when transplanted [102]. This suggests that the organizing capacity of regenerative tissues may be transient. Discussed here will be two case studies in which regenerative tissues can gain organizing capacity: oral regeneration in *Hydra* and limb regeneration in amphibians. The focus of the discussion will be on the initiation of pattern in the blastema and the specification of the appropriate cell types.

Hydrozoan Regeneration

Hydra are freshwater polyps that are polarized along their main body axis, with the oral end containing the hypostome (mouth) and tentacles, the aboral end containing the foot, and the body column (also called the gastric region) containing the digestive cavity between them (Figure 4A). As members of the phylum Cnidaria, they possess two germ layers (ectoderm and endoderm), and are radially symmetric along their primary body axis. *Hydra* have two tissue layers, an outer layer of ectodermal myoepithelial cells and an inner layer of endodermal myoepithelial cells (Figure 4C). Each tissue layer is a single cell thick and is separated by the ECM-rich mesoglea [103]. *Hydra* are capable of regenerating their oral ends after surgical removal, a feature that has fascinated scientists for over 270 years [104].

Oral and aboral organizers

While it was Hans Spemann who coined the terms "organizer" and "induction," Ethel Browne had observed these phenomena 15 years earlier through her experimentation on *Hydra*. She found that transplantation of an intact or regenerating head tip onto the

side of host *Hydra* could induce an ectopic axis, and that the outgrowth was composed primarily of host tissue [105]. Thus, these regions meet the criteria for the classical definition of an organizer. Later experiments by other researchers revealed the presence of a foot organizer at the aboral end of the animal [106,107].

Transplantation experiments were performed to investigate the potential for grafts to induce ectopic structures. A transplantation series where the position of the donor fragment from an intact *Hydra* varied along the oral-aboral axis, but the host site remained the same, indicated that the potential to generate an ectopic oral end was graded and highest orally. A complementary series, where the donor remained the same but the transplantation site varied, revealed the original oral end was inhibitory for the formation an ectopic axis [108]. The mysterious factors controlling these properties were termed Head Activation (HA) and Head Inhibition (HI). The possibility of a graded morphogen promoting head identity, which controlled its own inhibitor, was reminiscent of a Turing-like mechanism for guiding oral-aboral patterning [6].

There is evidence that the molecular instructions for patterning along the oral-aboral axis, as well as the molecular control of the head organizer, relies on Wnt signaling. Many Wnt ligands, including Wnt3, are expressed exclusively in the hypostome in intact animals, and appear early at the presumptive oral end during regeneration [109]. β-catenin localization is nuclear primarily at the oral end, suggesting that Wnt ligands in the organizer likely signal through canonical Wnt signaling [110]. Consistent with this hypothesis is the observation that pharmacological activation of Wnt signaling induces ectopic oral structures all throughout the body column [111]. If the HA signal are Wnt ligands, could the HI signals be Wnt inhibitors? Analysis of the promotor

of Wnt3 indicated binding sites for TCFs that promote transcription, as well as binding sites for an unidentified repressor. A repressive signal in the body column would help restrict Wnt3 expression to the hypostome [110]. The Dickkopf gene Dkk1/2/4-C is expressed in a graded fashion along the oral-aboral axis, with its highest expression just below the hypostome [112,113], however it does not indicate how Wnt3 is regulated at the transcriptional level. Thus, the molecular signal for HI remains unknown.

Given that *Hydra* can regenerate their oral ends, how is the head organizer regenerated *de novo*? After midgastric bisection, there is asymmetric activation of MAPK/CREB signaling preferentially at oral-facing wounds, which is required for initiating injury-induced apoptosis (Figure 4B) [114,115]. Apoptosis occurs in interstitial cell derivatives, and these cells secrete Wnt3, even as they are dying [116]. Subsequently, Wnt3 is transcriptionally activated in endodermal epithelial cells, which begin the cascade leading to the formation of a new head organizer [109,117]. The asymmetry in CREB activation likely relies on the polarity of the tissue fragment, although how this is regulated is unclear. To understand the requirement for tissue polarity in head organizer generation, *Hydra* were dissociated, reaggregated, and the formation of head organizers was assessed. Numerous head organizers formed *de novo* at random locations in the aggregate, with each organizer inducing its own axis in the surrounding tissue, and ultimately budding off to produce a complete animal. Wnt activity was initially present in all epithelial cells, before being confined to a few sites, indicating active regulation occurs to restrict high Wnt activity to the location of future organizers [117]. Head organizer formation can be initiated without tissue polarity, but requires Wnt signaling to be refined. Taken together, these findings highlight the

importance of Wnt signaling in head organizer maintenance and regeneration, and the role that injury-induced signals play in the *de novo* generation of pattern in the absence of developmental cues.

Lineage restricted stem cells

The cells of the *Hydra* are specialized along the oral-aboral axis, such that the ectoderm at the oral end is specialized into tentacle battery cells and ectodermal hypostome cells, whereas at the aboral end these cells are specialized into foot mucous cells and peduncle ectodermal cells. The endoderm lineage is composed of tentacle and hypostome endodermal cells at the oral end, foot endodermal cells at the aboral end, and digestive cells in between. *Hydra* have a number of other specialized cell types required for their normal biology, including germ cells, secretory gland cells, nematocytes, and two types of neurons [118].

The production of new cells, either in response to injury or during homeostasis, is achieved by the use of three defined stem cell lineages: the ectodermal epithelial stem cells, the endodermal epithelial stem cells, and the interstitial stem cells (Figure 4D). The epithelial stem cells in the gastric region divide every three to four days, and are progressively displaced towards the oral and aboral ends, where they will eventually differentiate and replace the cells of the tentacles and the foot. Similarly, interstitial stem cells are located in the gastric region, and migrate away as they proliferate and differentiate. They cycle much faster, every 24-30 hours, and reside in the interstitial space between the epithelial cells [118]. Recent work has uncovered some of the molecular signatures of the three stem cell compartments [119], however there are

presently few functional studies characterizing the roles of these genes in regulating differentiation into particular lineages.

A series of elegant experiments revealed the behavior of the interstitial stem cells, some by taking advantage of the fact that *Hydra* are capable of regeneration following dissociation and reaggregation [89]. Host *Hydra* were treated with nitrogen mustard (NM), which ablates interstitial cells, and dissociated into a single cell suspension. Donor *Hydra* were treated with tritiated-thymidine, and dissociated. These two populations were combined with the NM-treated cells in great excess, aggregated, and allowed to regenerate. *Hydra* with clonal growth descended from a single tritiated interstitial cell were analyzed. These experiments demonstrated that single interstitial cells are multipotent and could give rise to the nerve cells, gland cells, nematocytes and the gametes, but not cells from the epithelial layers [120,121]. Because proliferation of interstitial cells takes place in the gastric region, the oral and aboral regions are made up entirely of terminally differentiated cells, and as such, regeneration does not occur if these regions are isolated [122,123]. Importantly, the spatial pattern of differentiated interstitial cells in *Hydra* (such as the nerve cells or nematocytes) appears to be generated mostly by migration, rather a fate choice dependent upon position along the oral-aboral axis [124], although the cues regulating this migration are poorly understood. In principle, animals can regulate both stem cell specification and progenitor migration to regenerate correct tissue patterns. In animals such as *Hydra*, production of region-appropriate cell types is decoupled from specialization, and relies on migration alone.

Amphibian Limb Regeneration

Salamanders and other urodele amphibians are another classical model for regeneration. Similar to other vertebrates, they can replace tissues such as blood, bone muscle and epithelium as part of normal homeostasis. They are used as a model because they are also capable of regenerating entire limbs after amputation (Figure 4E) [101]. The vertebrate limb contains tissues from multiple germ layers: epidermis and peripheral neurons from ectoderm; muscle, bone, dermis, and blood vessels from mesoderm. Regeneration occurs after amputation at any proximal-distal location, generates the correct cell types, and patterns those cells to create functional organs. Regeneration in axolotls, one amphibian capable of regeneration, proceeds via blastema formation. The observation that when transplanted, blastemas retain their identity and instruct surrounding tissues, suggests that blastemas may act as transient adult tissue organizers [101].

Within hours of limb amputation, a simple epithelium migrates to cover the exposed wound. Beneath this wound epidermis, cells undergo a period of sustained proliferation to generate the mesenchymal cells that comprise the blastema. With time, the cells of the blastema will continue to proliferate, as well as differentiate into the appropriate cells types. Morphogenesis occurs as these cell types arrange themselves into functional tissues [125]. In addition, the arm varies along the proximal-distal, anterior-posterior, and dorsal-ventral axes. How is the patterning of this tissue controlled during regeneration, and to what degree does this patterning rely on developmental mechanisms? Additionally, what is the source of the diversity of cell types created during regeneration?

Signaling pathways and patterning

Many of the early observations about patterning during limb regeneration came from surgeries and transplantation experiments. Amputation of the upper arm will result in regeneration of the upper arm, lower arm, and the hand. Amputation at the wrist will only produce a hand. Thus, blastema cells are able to remember their location along the PD axis, and regenerate only those positional values distal to the wound, a concept termed positional memory [7].

Various transplantation experiments were further used to challenge the system to regenerate after unusual manipulations. First, an axolotl limb was amputated at the wrist, and the arm was grafted onto the flank of the same host. The upper arm was then amputated, to produce an animal with two regenerating arms, one inverted along the PD axis. Both wound faces only regenerated the tissue distal to the positional value of the wound face, an observation that led to the "Rule of Distal Transformation" [126]. Transplantation of a wrist blastema to an upper arm stump results not only in the formation of a hand, but regeneration of the missing positional values between the blastema and the stump, a process known as intercalary regeneration. Transplantation of an upper arm blastema onto a lower arm stump results in an inversion of the positional identities in between the blastema and the stump. These findings indicate communication between the blastema and the stump decides which positional values need to be filled in. Finally, grafting a right arm blastema onto a left arm stump, induces the formation of two supernumerary limbs. This suggests that the juxtaposition of anterior and posterior tissue could influence limb outgrowth and PD patterning [127].

The molecular instructions for proper regeneration of the proximal-distal axis were initially characterized using retinoic acid (RA). RA treatment would force cells to ignore the rule of distal transformation and produce proximal tissue at any wound face along the PD axis [128,129]. Analysis of chimeric retinoic acid receptors (RAR) identified the receptor RARδ2, which conferred proximal identity when ectopically activated [130]. The proximalizing activity of RA likely signals through the homeodomain transcription factor *Meis* (Figure 4F). *Meis* transcripts are more abundant in upper limb versus lower limb blastemas. MEIS protein is only localized to the nucleus proximally, but can be driven into the nucleus in more distal blastemas through RA administration [131]. Knockdown of MEIS1 and MEIS2 abrogrates the proximalization of RA-treated blastemas, and overexpression of *meis1a* or *meis2a* is sufficient to proximalize distal blastemas [127]. *Meis* expression proximally drives the expression of *prod1*, which is a GPI-linked cell surface protein that is required for proximal homing of cells during regeneration [132]. Finally, in addition to the work on RA, a recent study characterized the regeneration of position-specific *Hox* genes as occurring in a proximal to distal fashion [133]. Taken together, these observations highlight the similarities between proximal-distal patterning in regeneration and homeostasis (Figure 4F).

Another powerful assay for assessing limb patterning during regeneration is the accessory limb model (ALM). In this system, a nerve is rerouted from the arm into an upper arm anterior wound. This results in a bump that grows for 19 days, before it regresses and is reabsorbed. When combined with a graft that generates an AP positional conflict, these outgrowths often regenerate into a complete limb [134,135]. The ALM was used to characterize the mechanism by which the AP axis of the limb is

patterned. Activation of SHH in anterior-only blastemas (but not posterior-only blastemas) was sufficient to induce the growth of a complete limb. Conversely, activation of FGF8 in posterior-only blastemas (but not anterior-only blastemas), was able to drive complete limb growth. During normal regeneration, blastema formation activates the expression of FGF8 and Gremlin in the anterior and SHH in the posterior (Figure 4G). Early on, the FGF8 domain requires SHH for its maintenance; later expression of SHH requires FGF8. These two signaling centers reinforce and maintain each other as the limb grows and are utilized to direct AP patterning of tissues [136], again utilizing mechanisms similar to development.

Which cell types serve as the carriers of positional information during regeneration? Studies indicate that positional identity along the PD axis is tissue-specific. Lineage tracing studies (discussed in more detail below) indicated that positional identity markers were expressed primarily in connective-tissue-derived blastema cells [131]. Further work demonstrated that connective tissue derivatives and not muscle derivatives obey the rule of distal transformation upon transplantation [137]. This suggests that, similar to development, connective tissue utilizes extracellular information to generate a template for tissue pattern, which in turn is used by muscle and nerve cells to regenerate a complete limb.

Generation of new cells in the blastema

The recent development of fluorescent lineage tracing tools in axolotls has allowed for the dissection of new cell production during limb regeneration. Grafting of specific tissues from labeled donors into unlabeled hosts, followed by amputation, was used to understand how each tissue contributed to the composition of the blastema.

Transplanted dermis was able to make dermis, cartilage, and tendons but not muscle. Likewise, cartilage was also unable to make muscle. Muscle was found to make neither cartilage nor epidermis. The tissue-specificity of this technique relied on surgical precision, so to more cleanly generate labeled cell-types, specific embryonic regions were transplanted into unlabeled embryos and allowed to develop. This technique was used to label muscle cells, connective tissue, and Schwann cells. Amputation in these mosaic animals replicated the finding that during regeneration, cells respected their tissue type of origin [131]. These findings mirror the lineage restriction that occurs during development (Figure 4H). Dermis, skeleton, and other connective tissue cell types are derived from lateral plate mesoderm. Presomitic mesoderm gives rise to muscle and cells from the neural crest give rise to Schwann cells.

While clarifying questions regarding the lineage restriction of these cell types, this study did not address whether the generation of new cells was achieved by amplification of a resident lineage-restricted stem cell, or the dedifferentiation of a postmitotic cell. A comparison between a newt (*Notophthalmus viridescens*) and a salamander (*Ambystoma mexicanum*), revealed a divergent mechanism for the generation of new muscle cells. Newts primarily rely on dedifferentiation of mature muscle fibers, which fragment, proliferate, and give rise to new muscle in the blastema. In contrast, mature muscle in axolotls does not contribute to the blastema; resident stem cells are activated [138]. Another study, also in axolotls, sought to understand how the diversity of connective tissue cell types participated in digit regeneration by characterizing the behavior of chrondocytes, pericytes, periskeletal cells, and dermal fibroblasts. Given that the blastema must ultimately generate all of the mature cell types

of the limb, a reasonable assumption would be that the blastema cells must be composed of lineage-restricted progenitors in proportions commensurate with their final form. Surprisingly, examination of chrondocytes and pericytes revealed that they do proliferate in response to injury, but they do not migrate or contribute to the blastema. An early wave of periskeletal cells and dermal fibroblasts proliferated, migrated into the blastema and became chrondrocytes. A later wave of dermal fibroblasts entered the blastema and became the other soft connective tissues [139]. These observations are consistent with a model in which certain mature connective cell types are induced to dedifferentiate to contribute to the cells of the blastema. While not every cell type has been characterized, it seems likely that a combination of dedifferentiation and stem cell activation is utilized for amphibian limb regeneration.

Regeneration in both *Hydra* and amphibians relies on the establishment of organizing regions. These signaling centers activate developmental pathways to create an environment that can guide stem cells or differentiated cells to replace the appropriate missing tissues. The specification of cell fate and the organization of cells into functional tissues are challenges faced by all regenerative organisms. To understand further how these challenges are met, we next turn to a discussion of regeneration in the freshwater planarian, *Schmidtea mediterranea*.

III. Planarians as a Model for Studying Regeneration

The Planarian *Schmidtea mediterranea*

Planarians are free-living flatworms and their remarkable capacity for regeneration has fascinated scientists for over a hundred years (Figure 5A) [140,141]. Many surgical and grafting experiments have been performed to understand the principles of planarian regeneration [88]. With the advent of modern molecular techniques and their application to planarians, the process of regeneration has been interrogated with mechanistic detail. Gene function can be perturbed by RNAi [142,143] and gene expression can be visualized by *in situ* hybridization [144–146]. Planarians are members of the relatively understudied superphylum Lophotrochozoa, and as such, comparisons between them and deuterostomes can reveal insights about the last common ancestor of all Bilaterians.

Planarian anatomy

Planarians are bilaterally symmetric animals and contain derivatives of all three germ layers (ectoderm, mesoderm, endoderm). They possess three distinct body axes, an anterior-posterior (AP) axis, a dorsal-ventral (DV) axis, and a medial-lateral (ML) axis (Figure 5B), as well as a variety of differentiated tissue types. The planarian epidermis is a single-cell-layer thick epithelium that contains motile cilia on the ventral side, which are used for locomotion. The nervous system consists of the central nervous system (CNS), which contains a bi-lobed cephalic ganglia and ventral nerve cords, and peripheral nervous system (PNS) with neurons distributed thoughout the body. Connected to the cephalic ganglia are two dorsally located photoreceptors, which the

animal can use for phototaxis. Planarians eat using a muscular pharynx that is extruded through a ventrally-located mouth, which is the only opening to the outside. Food ingested from the pharynx is pushed into the intestine, which has one primary anterior branch and two primary posterior branches. Planarians utilize a number of different types of muscle, including body-wall muscle, intestinal muscle and pharyngeal muscle. In between the muscle layer and the gut is the parenchyma, which contains other mesenchymal differentiated tissue types [88]. Planaria use protonephridia for waste excretion and osmoregulation. The parenchyma is also the location of the neoblasts (adult stem cells), which are the only dividing cells in the animal [147].

Regenerative potential

Planarians are able to replace any missing tissue type and correctly restore form following injury. Remarkably, even small body fragments are capable of producing properly proportioned, albeit smaller, animals in a matter of days. Regeneration proceeds by two mechanisms: Epimorphosis, which involves the formation of a regeneration blastema at amputation planes, and morphollaxis, which is the remodeling of pre-existing tissue. The blastema is an unpigmented outgrowth at the site of the injury that will eventually differentiate into tissues removed by the injury. Recent work in the field has used planarians to explore fundamental questions in regeneration: How do animals respond to injuries? What is the source for new material in regeneration? How do animals adopt the correct form following injury?

Wound Response

Cellular responses to injury

A number of characteristic cellular responses to injury have been described in planarians. After an amputation one of the first processes is sealing of the wound by epidermal migration. Within 30 minutes, the dorsal and ventral epidermal cells have detached from the basement membrane and spread out to completely cover the wound face [148,149], a process that is facilitated by a muscular contraction to bring the wound edges closer together [150,151].

For all types of injuries, there is a systemic burst of proliferation from within the neoblast compartment that peaks at six hours following injury. In cases in which tissue has been removed, there is a second wave of proliferation close to the wound site that rises at about 48 hours post injury and persists for several days [152]. These proliferation events are thought to provide the cellular material for replacement of missing or damaged cells. Following the first wave of proliferation, neoblasts accumulate at wounds [152], a process that likely occurs by migration [153].

In addition to new cell formation, existing tissues are remodeled to adjust to the new size of the animal, and this is achieved by cell death (apoptosis). There is a peak in apoptotic cell numbers at the wound site for all wounds at four hours post injury, and in cases of missing tissue, a systemic peak in apoptotic cell numbers at three days following injury [154]. Interestingly, even tissues uninjured by the amputation itself (such as the pharynx), display an increase in apoptosis at this time that scaled with the amount of missing tissue, suggesting that apoptosis may be an important component of rescaling organ size following injury.

Transcriptional responses to injury

An unbiased microarray analysis of regenerating planarian fragments identified a program of gene expression that was induced after injury [155]. These genes were grouped by their kinetics and location into four categories: W1, W2, W3, and W4. W1, W2, and W3 genes were expressed in differentiated cells, and W4 genes were expressed in neoblasts. W1 genes are homologous to "immediate early genes" in other organisms, and their induction does not require translation. W2 genes were expressed sub-epidermally at the wound site, and often encoded regulators or ligands of developmental signaling pathways in other organisms. W3 genes were expressed in the epidermis and were not necessarily restricted to the wound site. W4 genes were expressed in neoblasts and many were transcription factors, chromatin remodeling proteins or cell cycle factors. Many of these genes were expressed in different injury types, suggesting a generic response to wounds.

A subsequent study, using single-cell RNAseq analyses of regenerating animals, revealed multiple phases of gene expression following wounding in a cell type specific manner [156]. It was found that there was indeed a generic response to all injuries, followed by a more specific response in cases requiring regeneration. Although expression of some genes occurs in all cell types following injury, most of the genes exhibited cell-type specificity and were expressed exclusively in neoblasts, epidermis, or muscle. After the generic phase has ended in tail fragment regenerating a head, gene expression falls into one of three categories: (1) expression of muscle-associated tissue patterning factors [157], (2) genes associated with specialized neoblasts [158,159], and (3) markers of differentiated anterior tissues [156]. These results link wound-induced gene expression to the proper regeneration of missing tissues.

Stem Cell Specification

Neoblasts are the source for new material in regeneration

Many animals replace missing tissues using resident stem cells, and in planarians these stem cells are referred to as neoblasts. Neoblasts were originally defined by their morphology [160,161], and by the fact that they are the only dividing cells in the animal [147]. Multiple lines of evidence suggest that regeneration in planarians requires neoblasts. Neoblasts have a characteristic response to wounds, which can vary between injuries, and they migrate to the site of regeneration [152]. Lethal irradiation, which ablates neoblasts, abrogates regeneration and ultimately leads to the death of the animal [153,162,163]. Bromodeoxyuridine (BrdU), a nucleotide analogue, is incorporated first into neoblasts, and then into their post-mitotic progeny [91,164]. Isolation of the head tip or pharynx, which do not contain neoblasts, does not result in regeneration [88]. Sequencing has identified a transcriptional signature present in all neoblasts that includes genes utilized in other systems for stem cell maintenance [164–169]. Interestingly, neoblasts also express *smedwi-1* and *smedwi-2*, genes that encode PIWI proteins, which are normally associated with germ cells (Figure 5C) [163]. Finally, the neoblast population was shown to contain pluripotent stem cells in an elegant experiment: A single cell transplanted into an irradiated host was able to both self-renew and differentiate into cells spanning all germ layers (Figure 5D). In some cases the cell was sufficient to restore regeneration and completely rescue the host [170].

Neoblast specialization

The identification of neoblasts based on morphology, location, and even gene expression led to conclusion that neoblasts represent a single homogenous stem cell population distributed throughout the animal [171]. Recent work has identified heterogeneity within the neoblast compartment that cannot be attributed to cell cycle state or location, but rather cell type. Studies involving the eye [172,173] and the protonephridia [174] were among the first to identify transcription factors required for the formation of new cells, which were expressed in and defined a subset of neoblasts. Further work characterized transcription factors associated with anterior pole [175–177], gut [169,178], muscle[179], pharynx [180], epidermis [159,164], and numerous neuronal lineages [155,158,179,181]. These findings led to the "specialized neoblast" model, in which neoblasts express a transcription factor for a particular cell type, exit the cell cycle, differentiate, and integrate into the mature organ (Figure 5D).

Patterning of New Tissue

Muscle cells contain patterning information

How are neoblasts instructed to differentiate into the appropriate missing tissues during regeneration? How are neoblast progeny directed to form functional organs? During development, a core group of signaling pathways, such as Wnt, Bmp, or Fgf, are used to control specification of cell fate and tissue patterning. Interestingly, planarian homologues of these pathway components are expressed constitutively in adults, and in many cases in a regionalized manner (Figure 5E) [182]. RNAi screens have indicated that many of these genes are required for the proper regeneration of axial pattern.

Perturbation of Wnt signaling results in defects in the regeneration of the anterior-posterior axis [183–188], and perturbation of Bmp signaling causes aberrant dorsal-ventral patterning [189–194].

Position control genes (PCGs) are defined as genes that meet the following two criteria: (1) they display regionalized expression along one or more body axes and (2) either have a patterning defect upon RNAi or are in a signaling pathway required for normal patterning [157]. A number of observations indicate that PCGs are the carriers of positional information and provide the molecular instructions for neoblasts during regeneration and homeostasis. PCGs are expressed in the same subepidermal cell population and are irradiation-insensitive, suggesting that their site of expression is a differentiated cell type. Further characterization revealed that PCGs are expressed primarily in the body-wall musculature, and are capable of dynamically altering their regionalized expression during regeneration [157]. The strongest evidence that PCGs provide instructions for regeneration is that their inhibition by RNAi results in duplications or elimination of axial structures. Inhibition of *wnt1* results in the formation of ectopic posterior heads, whereas inhibition of *notum*, a Wnt antagonist, results in the formation of anterior tails [186,188]. Two PCG circuits were identified that control the localization of the eyes and the pharynx [195,196], and notably these transformations occur even in the absence of injury. These findings suggest that PCGs are used to guide tissue patterning in both regeneration and homeostasis. One prediction of the PCG model is that neoblasts sense their position in a signaling gradient and adjust their behavior accordingly; up until this point, this has not been directly observed (Figure 5F).

Head regeneration

One of the most striking aspects of planarian regeneration is the production of a new head following head amputation in a matter of days. The decision to regenerate a head or tail has been the subject of intense study. Recent work has demonstrated that wound-induced expression of *notum* is responsible for the decision to regenerate a head only at anterior-facing wounds [186–188]. *notum* encodes a well-conserved secreted hydrolase that is an antagonist of Wnt signaling [188]. After the decision has been made, head regeneration proceeds next by the expression of PCGs that are restricted to the anterior [157,186,187]. As the pattern of PCGs are being re-established, the differentiated tissue architecture of the head is regenerated [195,197,198], and ultimately function is restored [88].

The anterior pole is a cluster of cells that is uniquely positioned in the animal; it is located at the anterior tip of the head, on the midline, and adjacent to the dorsal-ventral boundary (DVB). It is defined by the expression of the *notum* [189] and the Activin inhibitor *follistatin* [200,201], as well as the transcription factors *foxD* [176] and *zic-1* [177,178]. The anterior pole is thought to be a specialized form of body wall muscle, as it expresses the muscle marker *collagen* [178]. Like other muscle cells near the tip of the head, it expresses *sFRP-1*, *prep*, *ndk*, and *ndl-4*, all of which are anterior PCGs [176]. Recent work has demonstrated that *foxD* [176] and *zic-1* [177,178] are required for the *de novo* formation of the anterior pole during regeneration. In summary, both genes are reciprocally required for each other's expression and for the expression of *notum* and *follistatin* in the anterior pole of uninjured animals. Both *foxD* and *zic-1* are expressed in neoblasts early in regeneration, consistent with an autonomous

requirement for these genes to specialize neoblasts into anterior pole progenitors. Interestingly, the ablation of the anterior pole through RNAi of either *foxD* or *zic-1* results in dramatic head regeneration defects. These animals fail to re-express anterior PCGs in head blastemas, as well as PCGs associated with the midline. The tissues that normally form in head regeneration are either absent or collapsed at the midline, indicating a requirement for the anterior pole in AP and ML patterning of the blastema (Figure 1C) [176–178,202].

Like many other organisms, planarians have distinct responses to injury. Transcriptional events induced by injury are utilized to reestablish the signaling environment present in intact adults. Neoblasts also respond to wounds and will produce any missing tissue removed by injury, and replace all tissues during homeostasis. The wealth of knowledge about planarian patterning and stem cell biology makes them an attractive candidate for the study of how pattern initiation and stem cell specification during regeneration can proceed in the absence of developmental cues.

IV. Content Overview

During development, embryonic pattern-initiating processes are utilized to set up organizers, regions that exert a patterning influence on neighboring cells. How complex tissue patterns are created in adults in the absence of developmental cues is a central question in regeneration. Chaper 2 of this thesis demonstrates, using transplantation, the organizing capacity of the anterior head tip, which contains the anterior pole. These findings motivated using the anterior pole as a case study in how adult tissue organizers are placed during regeneration. The anterior pole relies on three landmarks at wounds for its localization: an anterior orientation, the pre-existing midline, and the dorsal-ventral boundary in the blastema. Patterning information influenced the specification of anterior pole progenitors and their migration. This process, utilizing positional information in existing tissue at unpredictably shaped wounds, can influence patterning of new tissue in a manner that facilitates integration with pre-existing tissue in regeneration.

Complex spatial patterns exist within tissues and must be regenerated. How neoblasts utilize positional information to generate the diversity of cell types and place them appropriately is poorly understood. Chaper 3 of this thesis addresses this question through a molecular characterization of the planarian epidermis. Bulk RNA-sequencing of the mature epidermis revealed a number of interesting spatial patterns. In order to understand when these patterns arose, single-cell RNA sequencing was performed on cells from throughout the epidermal lineage. Epidermal progenitors appeared to be pre-patterned near their mature counterparts. Epidermal neoblasts had distinct dorsal and ventral identities, and these were controlled by Bmp signaling. This finding provides the

first evidence of neoblasts directly reading out their position in a signaling gradient and adjusting their gene expression programs to promote appropriate regional identity.

Figure 1

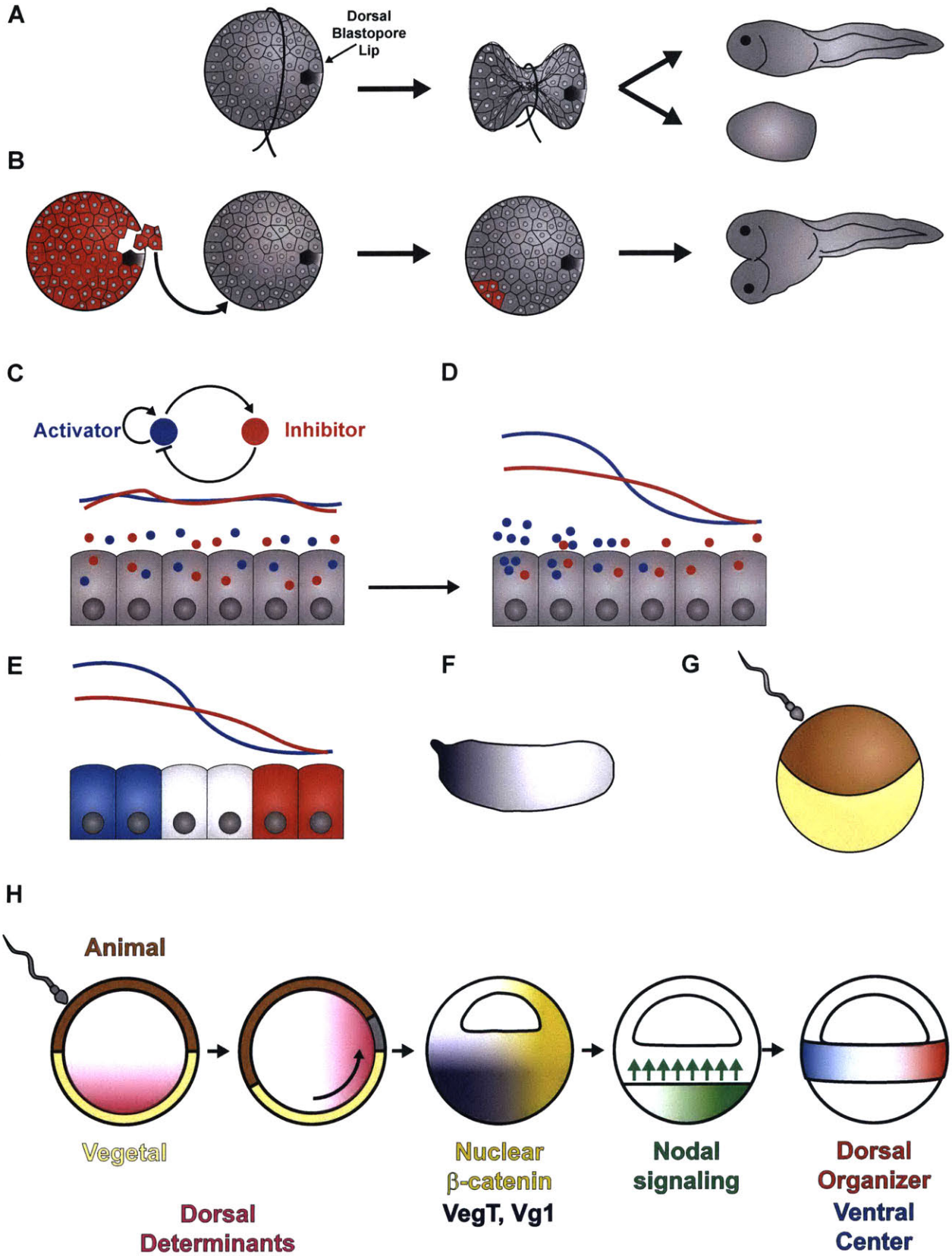

A
Dorsal
Blastopore
Lip
B
C
Activator
Inhibitor
D
E
F
G
H
Animal
Vegetal
Dorsal
Determinants
Nuclear
β-catenin
VegT, Vg1
Nodal
signaling
Dorsal
Organizer
Ventral
Center

Figure 1. Pattern formation and symmetry breaking in development.

(A) Bisection of an amphibian embryo so that only one half contained the dorsal blastopore lip results into two different pieces: A dorsal half that develops normally, and the ventral half developed into a "belly-piece," which lacks all axial organs. **(B)** Transplantation of the dorsal blastopore lip onto the ventral side of another embryo induces the formation of a secondary axis. **(C-D)** Turing proposed that a system of diffusible substances termed "morphogens" could interact in a reaction-diffusion system to create a patterned state from an initial homogenous starting state. In this example, an activator (blue) promotes its own expression and the expression of its inhibitor (red). **(E)** In this diagram, cells adopt a fate based on their position in a morphogen gradient. Cells in the highest blue region adopt a blue fate, cells in the lowest blue region adopt a red fate, and cells in the middle adopt a white fate. **(F)** Maternally-deposited *bicoid* (purple) is utilized for AP axial patterning in *Drosophila*. **(G)** The initial symmetry-breaking event in *Xenopus* is triggered by sperm entry. **(H)** Sperm entry in *Xenopus* triggers cortical rotation. This brings dorsal determinants (pink) from the vegetal pole to a point opposite sperm entry. This promotes nuclear β-catenin (yellow) and thus Wnt signaling, which is required, along with VegT and Vg1 (purple) to set up a gradient of Nodal signaling (green). Nodal signaling induces the mesoderm, and the highest point of signaling becomes the dorsal organizer (red) and the lowest point becomes the ventral center (blue).

Figure 2

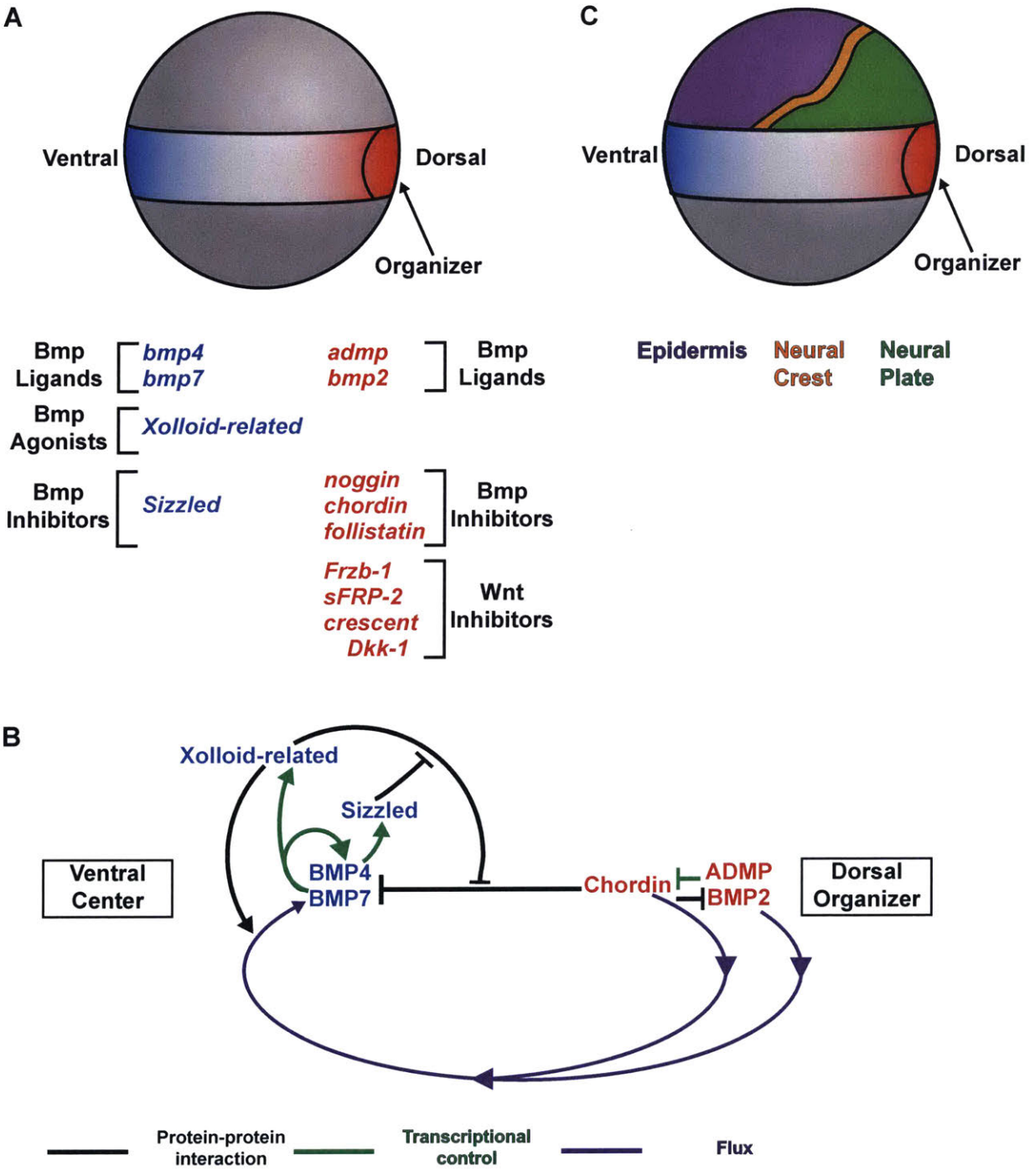

A
Ventral
Dorsal
Organizer
C
Ventral
Dorsal
Organizer
Bmp Ligands
bmp4
bmp7
admp
bmp2
Bmp Ligands
Bmp Agonists
Xolloid-related
Bmp Inhibitors
Sizzled
noggin
chordin
follistatin
Bmp Inhibitors
Frzb-1
sFRP-2
crescent
Dkk-1
Wnt Inhibitors
Epidermis
Neural Crest
Neural Plate
B
Xolloid-related
Sizzled
Ventral Center
BMP4
BMP7
Chordin
ADMP
BMP2
Dorsal Organizer
Protein-protein interaction
Transcriptional control
Flux

Figure 2. Vertebrate Axis Specification.

(A) Dorsal-ventral patterning is controlled by signaling molecules secreted from the dorsal organizer (red) and the ventral center (blue). **(B)** Properties in development such as robustness and scaling are granted by interactions between the components secreted from the dorsal organizer and ventral center. Genes in red are expressed from the dorsal organizer, and genes in blue are expressed from the ventral center. Protein-protein interactions are represented by black lines, transcriptional control is represented by green lines, and flux across the embryo is represented by purple lines. **(C)** The gradient of Bmp signaling is interpreted by the ectoderm to guide fate specification. Cells closest to the Bmp source adopt epidermal fates (purple), whereas cells closest to the dorsal organizer adopt neuronal fates (green). Cells at the boundary between these two become the neural crest (orange).

Figure 3

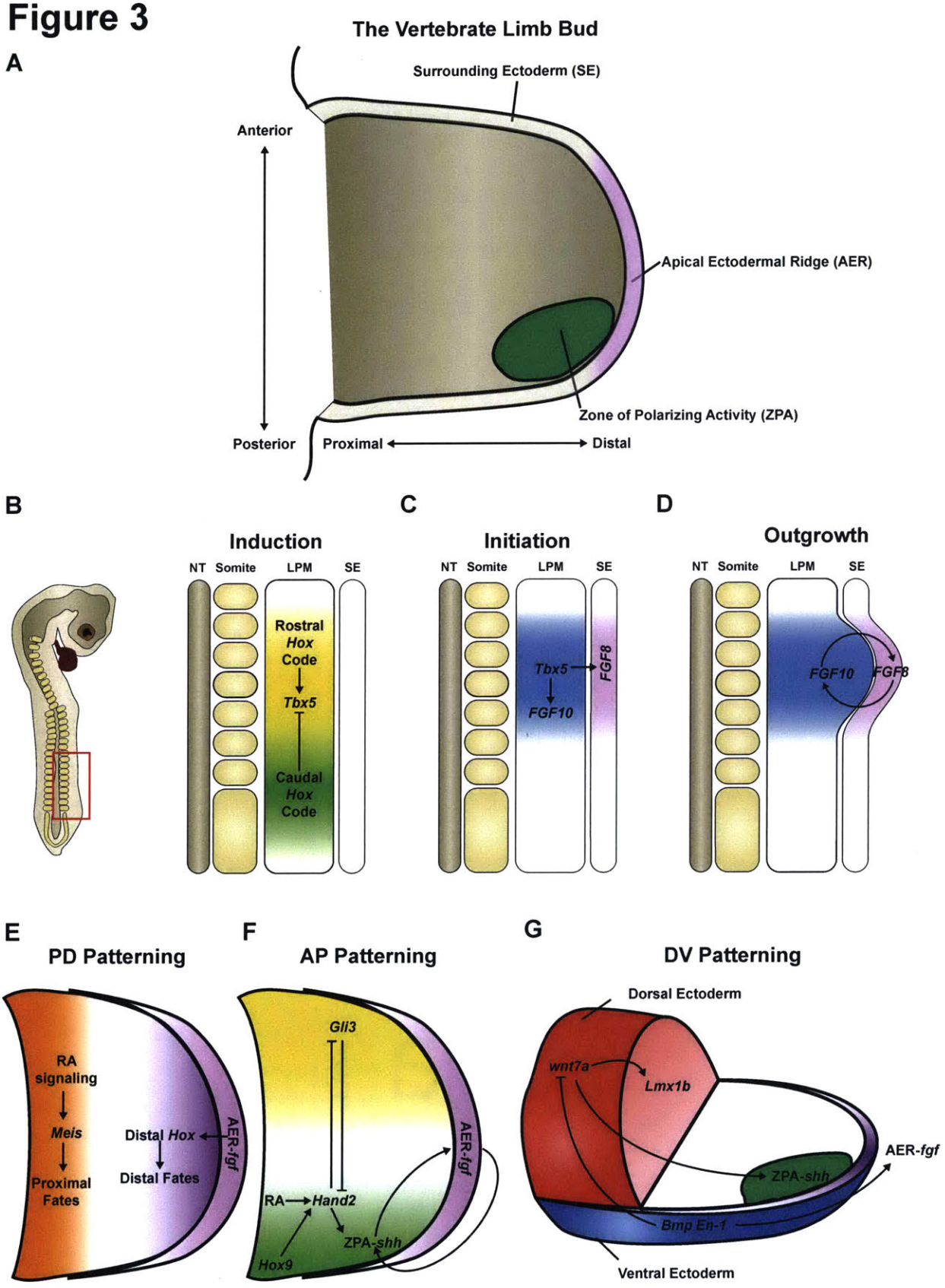

A
The Vertebrate Limb Bud
Surrounding Ectoderm (SE)
Anterior
Apical Ectodermal Ridge (AER)
Zone of Polarizing Activity (ZPA)
Posterior
Proximal
Distal
B
C
D
Induction
Initiation
Outgrowth
NT
Somite
LPM
SE
Rostral Hox Code
Tbx5
Caudal Hox Code
Tbx5
FGF8
FGF10
FGF10
FGF8
E
F
G
PD Patterning
AP Patterning
DV Patterning
RA signaling
Meis
Proximal Fates
Distal Hox
Distal Fates
AER-fgf
Gli3
RA
Hand2
Hox9
ZPA-shh
AER-fgf
Dorsal Ectoderm
wnt7a
Lmx1b
AER-fgf
ZPA-shh
Bmp En-1
Ventral Ectoderm

Figure 3. Vertebrate Limb Development.

(A) The vertebrate limb bud is an outgrowth of lateral plate mesoderm (LPM) surrounded by an ectodermal sheath. The distal ectoderm is the apical ectodermal ridge (AER, purple), and the more proximal ectoderm on both the dorsal and ventral sides is the surrounding ectoderm (SE). The zone of polarizing activity (ZPA, green) is located in the posterior distal mesenchyme. **(B)** During vertebrate embryonic development, the formation of the limb bud begins when the *Hox* code is interpreted in the LPM to activate expression of *Tbx5* at the appropriate AP position for forelimb induction. This utilizes activating signals from the rostral *Hox* code (yellow) and inhibiting signals from the caudal *Hox* code (green). **(C)** The limb bud is initiated when *Tbx5* (blue) drives expression of *FGF10* in the LPM and *FGF8* in the SE (purple). *FGF8* marks the site of the future AER. **(D)** Limb bud outgrowth is driven by a positive feedback loop between *FGF10* (blue) and *FGF8* (purple). NT = Neural Tube. LPM = Lateral Plate Mesoderm. SE = Surrounding Ectoderm. **(E)** Proximally, the limb bud is patterned by retinoic acid (RA, orange) signaling that drives the expression of *meis* transcription factors that promote posterior fates. Distally, a gradient of FGF signaling from the apical ectodermal ridge (AER, purple) acts to induce expression of *Hox* genes that promote proximal fates. **(F)** *Gli3* in the anterior (yellow) and a combination of RA, *Hox9* and *Hand2* in the posterior (green), induce the zone of polarizing activity (ZPA) in the posterior distal mesenchyme, which secretes *shh*. **(G)** The limb bud is patterned along the DV axis by *wnt7a* (red) from the dorsal ectoderm and Bmp and *En-1* signals (blue) from the ventral ectoderm.

Figure 4

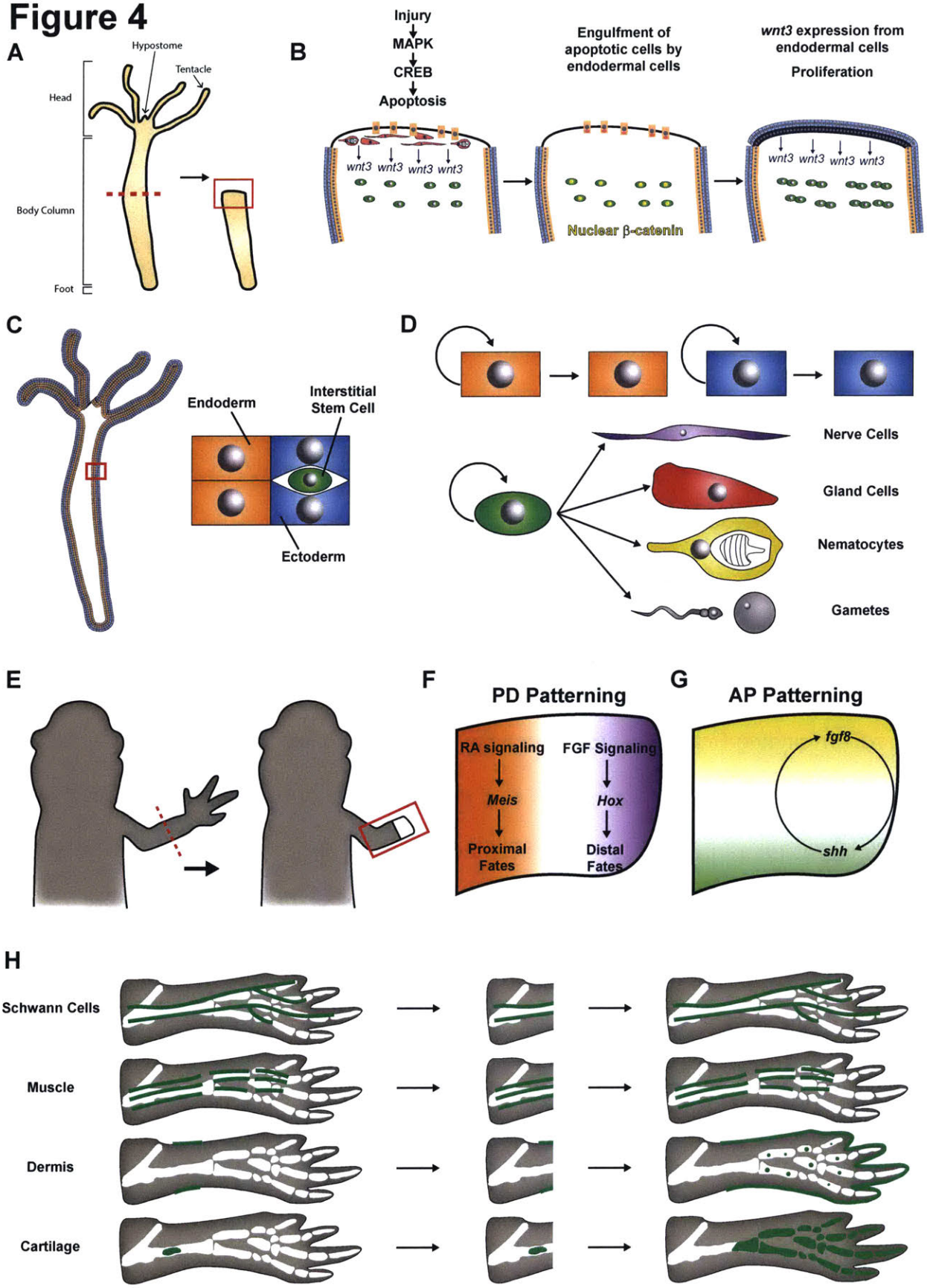

A
Hypostome
Tentacle
Head
Body Column
Foot
B
Injury
MAPK
CREB
Apoptosis
wnt3 wnt3 wnt3 wnt3
Engulfment of apoptotic cells by endodermal cells
Nuclear β-catenin
wnt3 expression from endodermal cells
Proliferation
wnt3 wnt3 wnt3 wnt3
C
Endoderm
Interstitial Stem Cell
Ectoderm
D
Nerve Cells
Gland Cells
Nematocytes
Gametes
E
F
PD Patterning
RA signaling
FGF Signaling
Meis
Hox
Proximal Fates
Distal Fates
G
AP Patterning
fgf8
shh
H
Schwann Cells
Muscle
Dermis
Cartilage

Figure 4. Pattern formation in regeneration.

(A) *Hydra* are composed along the oral-aboral axis of a head, body column and foot. The head has specialized structures like the hypostome and the tentacles. *Hydra* can perform oral regeneration following bisection. **(B)** During regeneration, injury signals drive apoptosis in interstitial cell derivatives preferentially at oral-facing wounds after gastric amputation. These cells release Wnt3. Wnt3 activates β-catenin in interstitial cells, and endodermal epithelial cells engulf apoptotic cells. Interstitial cells undergo rounds of synchronized cell divisions. Endodermal epithelial cells now secrete Wnt3 that promotes the formation of a new oral organizer. **(C)** *Hydra* are diploblastic and composed of an outer ectodermal later (blue) and an inner endodermal layer (orange). Interstitial stem cells (green) reside in between the ectodermal epithelial cells. **(D)** Ectodermal epithelial cells give rise to more ectodermal epithelial cells and their derivatives, and endodermal epithelial cells give rise to more endodermal epithelial cells and their derivatives. Interstitial stem cells give rise to nerve cells (purple), gland cells (red), nematocytes (yellow), and gametes (gray). **(E)** Amphibians can regenerate following limb amputation. **(F)** During amphibian limb regeneration, proximal-distal patterning utilizes mechanisms similar to development. **(G)** A circuit of FGF8 and SHH is utilized to pattern the anterior-posterior axis of regenerating amphibian limbs. **(H)** Cell types respect their embryonic origins during regeneration. Schwann cells can only regenerate Schwann cells, and muscle can only regenerate muscle. Dermis and cartilage, both derivatives of the lateral plate mesoderm, can differentiate into each other.

Figure 5

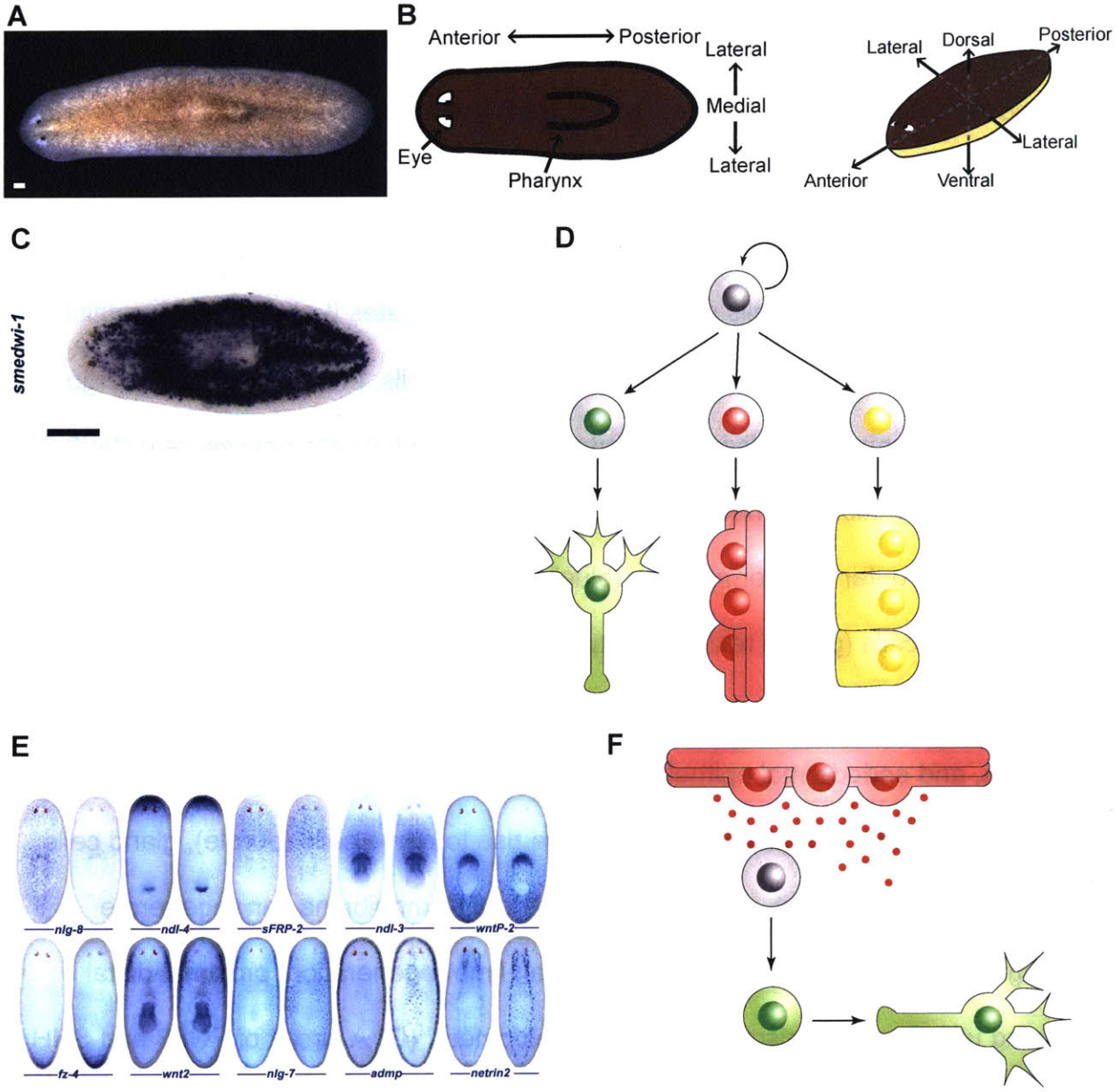

A
B
Anterior
Posterior
Lateral
Medial
Lateral
Eye
Pharynx
Lateral
Dorsal
Posterior
Lateral
Anterior
Ventral
C
smedwi-1
D
E
nlg-8
ndl-4
sFRP-2
ndl-3
wntP-2
fz-4
wnt2
nlg-7
admp
netrin2
F

Figure 5. Planarians as a model to study regeneration.

(A) Live images of an intact planarian. **(B)** Planarians have three distinct body axes: an anterior-posterior axis, a dorsal-ventral axis, and a medial-lateral axis. Also labeled are the eyes and pharynx. **(C)** Colorimetric *in situ* hybridization for *smedwi-1*, which marks neoblasts. **(D)** Neoblasts can specialize to adopt different cell fates, and then exit the cell cycle and differentiate. Some neoblasts are pluripotent, and as such, can both self-renew and generate any cell type in the animal. **(E)** Colorimetric *in situ* hybridization for a number of signaling molecules expressed in a regional manner. These molecules are primarily expressed in body-wall muscle. Images taken from [157], used with permission. **(F)** The PGC model predicts that neoblasts can sense their position in a signaling gradient established by muscle, and respond by differentiating into tissue appropriate for that region.

References

1. Driesch, H. (1891). The potency of the first two cleavage cells in the development of echinoderms. Gt. Exp. Biol. Prentice-Hall, Englewoods Cliffs, 210–4.

2. Böveri, T. (1907). Zellenstuden VI: Die Entwicklung dispermer Seeigeleier. Ein Beitrag. zur Befruchtungslehre und zur Theorie des Kernes. Jena Zeitschr Naturw. *43*, 1–292.

3. Spemann, H. (1918). Über die Determination der ersten Organanlagen des Amphibienembryo I--VI. Dev. Genes Evol. *43*, 448–555.

4. Spemann, H., and Mangold, H. (1924). Uber Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. Arch. Fur Mikroskopische Anat. Und Entwicklungsmechanik. *100*, 599–638.

5. Turing, A.M. (1952). The chemical basis of morphogenesis. Bull. Math. Biol. *52*, 153–97.

6. Gierer, A., and Meinhardt, H. (1972). A theory of biological pattern formation. Kybernetik. *12*, 30–9.

7. Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. J. Theor. Biol. *25*, 1–47.

8. Frohnhöfer, H.G., and Nüsslein-Volhard, C. (1986). Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid.* Nature. *324*, 120–5.

9. Becalska, A.N., and Gavis, E.R. (2009). Lighting up mRNA localization in

Drosophila oogenesis. Development. *136*, 2493–503.

10. Kugler, J.M., and Lasko, P. (2009). Localization, anchoring and translational control of *oskar*, *gurken*, *bicoid* and *nanos* mRNA during Drosophila oogenesis. Fly (Austin). *3*, 15–28.

11. Driever, W., and Nüsslein-Volhard, C. (1988). A Gradient of *bicoid* Protein in Drosophila Embryos. Cell. *54*, 83–93.

12. Struhl, G., Struhl, K., and Macdonald, P.M. (1989). The gradient morphogen *bicoid* is a concentration-dependent transcriptional activator. Cell. *57*, 1259–73.

13. Driever, W., and Nüsslein-Volhard, C. (1989). The *bicoid* protein is a positive regulator of hunchback transcription in the early *Drosophila* embryo. Nature. *337*, 138–43.

14. Bracket, J. (1977). An old enigma: The gray crescent of amphibian eggs. Curr. Top. Dev. Biol. *11*, 133–86.

15. Reversade, B., and De Robertis, E.M. (2005). Regulation of ADMP and BMP2/4/7 at opposite embryonic poles generates a self-regulating morphogenetic field. Cell. *123*, 1147–60.

16. Elinson, R.P., and Manes, M.E. (1978). Morphology of the site of sperm entry on the frog egg. Dev. Biol. *63*, 67–75.

17. Gerhart, J., Danilchik, M., Doniach, T., Roberts, S., Rowning, B., and Stewart, R. (1989). Cortical rotation of the *Xenopus* egg: consequences for the anteroposterior pattern of embryonic dorsal development. Development. *107*, 37–

51.

18. Houston, D.W. (2012). Cortical rotation and messenger RNA localization in *Xenopus* axis formation. Wiley Interdiscip. Rev. Dev. Biol. *1*, 371–88.

19. Rowning, B.A., Wells, J., Wu, M., Gerhart, J.C., Moon, R.T., and Larabell, C.A. (1997). Microtubule-mediated transport of organelles and localization of β-catenin to the future dorsal side of *Xenopus* eggs. Proc. Natl. Acad. Sci. U. S. A. *94*, 1224–9.

20. Larabell, C.A., Torres, M., Rowning, B. a, Yost, C., Miller, J.R., Wu, M., Kimelman, D., and Moon, R.T. (1997). Establishment of the dorso-ventral axis in *Xenopus* embryos is presaged by early asymmetries in β-catenin that are modulated by the Wnt signaling pathway. J. Cell Biol. *136*, 1123–36.

21. Schneider, S., Steinbeisser, H., Warga, R.M., and Hausen, P. (1996). β-Catenin Translocation Into Nuclei Demarcates the Dorsalizing Centers in Frog and Fish Embryos. Mech. Dev. *57*, 191–8.

22. Agius, E., Oelgeschläger, M., Wessely, O., Kemp, C., and De Robertis, E.M. (2000). Endodermal Nodal-related signals and mesoderm induction in *Xenopus*. Development. *127*, 1173–83.

23. Waddington, C.H. (1933). Induction by the primitive streak and its derivatives in the chick. J. Exp. Biol. *10*, 38–46.

24. Waddington, C.H. (1932). Experiments on the development of chick and duck embryos, cultivated in vitro. Philos. Trans. R. Soc. London. Ser. B, Contain. Pap.

a Biol. Character. *221*, 179–230.

25. Beddington, R.S. (1994). Induction of a second neural axis by the mouse node. Development. *120*, 613–20.

26. Oppenheimer, J.M. (1936). Structures Developed in Amphibians by Implantation of Living Fish Organizer. Proc. Soc. Exp. Biol. Med. *34*, 461–3.

27. Shih, J., and Fraser, S.E. (1996). Characterizing the zebrafish organizer: microsurgical analysis at the early-shield stage. Development. *122*, 1313–22.

28. Joubin, K., and Stern, C.D. (2001). Formation and maintenance of the organizer among the vertebrates. Int. J. Dev. Biol. *45*, 165–75.

29. Feng, X., and Derynck, R. (2005). Specificity and Versatility in Tgf-β Signaling Through Smads. Annu. Rev. Cell Dev. Biol. *21*, 659–93.

30. De Robertis, E.M., and Kuroda, H. (2004). Dorsal-Ventral Patterning and Neural Induction in *Xenopus* Embryos. Annu Rev Cell Dev Biol. *20*, 285–308.

31. De Robertis, E.M., Larraín, J., Oelgeschläger, M., and Wessely, O. (2000). The establishment of Spemann's organizer and patterning of the vertebrate embryo. Nat. Rev. Genet. *1*, 171–81.

32. Petersen, C.P., and Reddien, P.W. (2009). Wnt signaling and the polarity of the primary body axis. Cell. *139*, 1056–68.

33. Kiecker, C., and Niehrs, C. (2001). A morphogen gradient of Wnt/β-catenin signalling regulates anteroposterior neural patterning in *Xenopus*. Development. *128*, 4189–201.

34. Alexander, T., Nolte, C., and Krumlauf, R. (2009). Hox genes and segmentation of the hindbrain and axial skeleton. Annu. Rev. Cell Dev. Biol. *25*, 431–56.

35. Forlani, S., Lawson, K. a, and Deschamps, J. (2003). Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. Development. *130*, 3807–19.

36. Ikeya, M., and Takada, S. (2001). Wnt-3a is required for somite specification along the anteroposterior axis of the mouse embryo and for regulation of *cdx-1* expression. Mech. Dev. *103*, 27–33.

37. Gimlich, R.L., and Cooke, J. (1983). Cell lineage and the induction of second nervous systems in amphibian development. Nature. *306*, 471–3.

38. Willier, B.H., Weiss, P.A., and Hamburger, V. Analysis of development. Saunders; 1955.

39. Khokha, M.K., Yeh, J., Grammer, T.C., and Harland, R.M. (2005). Depletion of three BMP antagonists from Spemann's organizer leads to a catastrophic loss of dorsal structures. Dev. Cell. *8*, 401–11.

40. De Robertis, E.M. (2009). Spemann's organizer and the self-regulation of embryonic fields. Mech. Dev. *126*, 925–41.

41. Harrison, R.G. (1918). Experiments on the development of the fore limb of Amblystoma, a self-differentiating equipotential system. J. Exp. Zool. *25*, 413–61.

42. Tickle, C. (2015). How the embryo makes a limb: determination, polarity and identity. J. Anat. *227*, 418–30.

43. Sheeba, C.J., Andrade, R.P., and Palmeirim, I. (2016). Getting a handle on embryo limb development: Molecular interactions driving limb outgrowth and patterning. Semin. Cell Dev. Biol. *49*, 92–101.

44. Saito, D., Yonei-Tamura, S., Kano, K., Ide, H., and Tamura, K. (2002). Specification and determination of limb identity: evidence for inhibitory regulation of *Tbx* gene expression. Development. *129*, 211–20.

45. Stephens, T.D., and McNulty, T.R. (1981). Evidence for a metameric pattern in the development of the chick humerus. J. Embryol. Exp. Morphol. *61*, 191–205.

46. Chaube, S. (1959). On axiation and symmetry in transplanted wing of the chick. J. Exp. Zool. *140*, 29–77.

47. Nishimoto, S., Minguillon, C., Wood, S., and Logan, M.P.O. (2014). A Combination of Activation and Repression by a Colinear Hox Code Controls Forelimb-Restricted Expression of *Tbx5* and Reveals Hox Protein Specificity. PLoS Genet. *10*, 1–13.

48. Nishimoto, S., Wilde, S.M., Wood, S., and Logan, M.P.O. (2015). RA Acts in a Coherent Feed-Forward Mechanism with *Tbx5* to Control Limb Bud Induction and Initiation. Cell Rep. *12*, 879–91.

49. Agarwal, P., Wylie, J.N., Galceran, J., Arkhitko, O., Li, C., Deng, C., Grosschedl, R., and Bruneau, B.G. (2003). *Tbx5* is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. Development. *130*, 623–33.

50. Kawakami, Y., Capdevila, J., Büscher, D., Itoh, T., Esteban, C.R., and Belmonte, J.C.I. (2001). WNT signals control FGF-dependent limb initiation and AER induction in the chick embryo. Cell. *104*, 891–900.

51. Gros, J., and Tabin, C.J. (2014). Vertebrate limb bud formation is initiated by localized epithelial-to-mesenchymal transition. Science. *343*, 1253–6.

52. Tabin, C., and Wolpert, L. (2007). Rethinking the proximodistal axis of the vertebrate limb in the molecular era. Genes Dev. *21*, 1433–42.

53. Mercader, N., Leonardo, E., Azpiazu, N., Serrano, A., Morata, G., Martínez, C., and Torres, M. (1999). Conserved regulation of proximodistal limb axis development by Meis1/Hth. Nature. *402*, 425–9.

54. Mercader, N., Leonardo, E., Piedra, M.E., Martínez-A, C., Ros, M.A., and Torres, M. (2000). Opposing RA and FGF signals control proximodistal vertebrate limb development through regulation of Meis genes. Development. *127*, 3961–70.

55. Mariani, F. V., Ahn, C.P., and Martin, G.R. (2008). Genetic evidence that FGFs have an instructive role in limb proximal-distal patterning. Nature. *453*, 401–5.

56. Tickle, C. (2006). Making digit patterns in the vertebrate limb. Nat. Rev. Mol. Cell Biol. *7*, 45–53.

57. Riddle, R.D., Johnson, R.L., Laufer, E., and Tabin, C. (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. Cell. *75*, 1401–16.

58. Te Welscher, P., Fernandez-Teran, M., Ros, M.A., and Zeller, R. (2002). Mutual genetic antagonism involving GLI3 and dHAND prepatterns the vertebrate limb

bud mesenchyme prior to SHH signaling. Genes Dev. *16*, 421–6.

59. Galli, A., Robay, D., Osterwalder, M., Bao, X., Bénazet, J.D., Tariq, M., Paro, R., Mackem, S., and Zeller, R. (2010). Distinct roles of Hand2 in initiating polarity and posterior *Shh* expression during the onset of mouse limb bud development. PLoS Genet. *6*.

60. Tarchini, B., Duboule, D., and Kmita, M. (2006). Regulatory constraints in the evolution of the tetrapod limb anterior-posterior polarity. Nature. *443*, 985–8.

61. Parr, B.A., and McMahon, A.P. (1995). Dorsalizing signal *Wnt-7a* required for normal polarity of D-V and A-P axes of mouse limb. Nature. *374*, 350–3.

62. Loomis, C.A., Harris, E., Michaud, J., Wurst, W., Hanks, M., and Joyner, A.L. (1996). The mouse *Engrailed-1* gene and ventral limb patterning. Nature. *382*, 360–3.

63. Arques, C.G., Doohan, R., Sharpe, J., and Torres, M. (2007). Cell tracing reveals a dorsoventral lineage restriction plane in the mouse limb bud mesenchyme. Development. *134*, 3713–22.

64. Riddle, R.D., Ensini, M., Nelson, C., Tsuchida, T., Jessell, T.M., and Tabin, C. (1995). Induction of the LIM homeobox gene *Lmx1* by WNT6a establishes dorsoventral pattern in the vertebrate limb. Cell. *83*, 631–40.

65. Vogel, A., Rodriguez, C., Warnken, W., Izpisúa Belmonte, J.C., and Izpisua Belmonte, J.C. (1995). Dorsal cell fate specified by chick *Lmx1* during vertebrate limb development. Nature. *378*, 716–20.

66. Tzchori, I., Day, T.F., Carolan, P.J., Zhao, Y., Wassif, C.A., Li, L., Lewandoski, M., Gorivodsky, M., Love, P.E., Porter, F.D., et al. (2009). LIM homeobox transcription factors integrate signaling events that control three-dimensional limb patterning and growth. Development. *136*, 1375–85.

67. Searls, L. (1965). Isolation of Mucopoilysacclharide from the Precartilaginous Embryonic Chick Limb Bud. Exp. Biol. Med., 1172–6.

68. Stark, R.J., and Searls, R.L. (1973). A description of chick wing bud development and a model of limb morphogenesis. Dev. Biol. *33*, 138–53.

69. Pearse, R. V., Scherz, P.J., Campbell, J.K., and Tabin, C.J. (2007). A cellular lineage analysis of the chick limb bud. Dev. Biol. *310*, 388–400.

70. Thorogood, P. V, and Hinchliffe, J.R. (1975). An analysis of the condensation process during chondrogenesis in the embryonic chick hind limb. J. Embryol. Exp. Morphol. *33*, 581–606.

71. Mariani, F. V. (2010). Proximal to distal patterning during limb development and regeneration: a review of converging disciplines. Regen. Med. *5*, 451–62.

72. ten Berge, D., Brugmann, S.A., Helms, J.A., and Nusse, R. (2008). Wnt and FGF signals interact to coordinate growth with cell fate specification during limb development. Development. *135*, 3247–57.

73. Kardon, G., Harfe, B.D., and Tabin, C.J. (2003). A *Tcf4*-positive mesodermal population provides a prepattern for vertebrate limb muscle patterning. Dev. Cell. *5*, 937–44.

74. Mok, G.F., Cardenas, R., Anderton, H., Campbell, K.H.S., and Sweetman, D. (2014). Interactions between FGF18 and retinoic acid regulate differentiation of chick embryo limb myoblasts. Dev. Biol. *396*, 214–23.

75. Pryce, B.A., Watson, S.S., Murchison, N.D., Staverosky, J.A., Dünker, N., and Schweitzer, R. (2009). Recruitment and maintenance of tendon progenitors by TGFβ signaling are essential for tendon formation. Development. *136*, 1351–61.

76. Havis, E., Bonnin, M.M.-A., De Lima, J.J.E., Charvet, B., Milet, C., and Duprez, D. (2016). TGFβ and FGF promote tendon progenitor fate and act downstream of muscle contraction to regulate tendon differentiation during chick limb development. Development. *143*, 3839–51.

77. Hasson, P., DeLaurier, A., Bennett, M., Grigorieva, E., Naiche, L.A., Papaioannou, V.E., Mohun, T.J., and Logan, M.P.O. (2010). *Tbx4* and *Tbx5* Acting in Connective Tissue Are Required for Limb Muscle and Tendon Patterning. Dev. Cell. *18*, 148–56.

78. Hasson, P. (2011). "Soft" Tissue Patterning: Muscles and Tendons of the Limb Take Their Form. Dev. Dyn. *240*, 1100–7.

79. Huettl, R.E., Luxenhofer, G., Bianchi, E., Haupt, C., Joshi, R., Prochiantz, A., and Huber, A.B. (2015). *Engrailed 1* mediates correct formation of limb innervation through two distinct mechanisms. PLoS One. *10*, 1–16.

80. Poss, K.D. (2010). Advances in understanding tissue regenerative capacity and mechanisms in animals. Nat. Rev. Genet. *11*, 710–22.

81. Bely, A.E., and Nyberg, K.G. (2010). Evolution of animal regeneration: re-emergence of a field. Trends Ecol. Evol. *25*, 161–70.

82. Lin, G., and Slack, J.M.W. (2008). Requirement for Wnt and FGF signaling in *Xenopus* tadpole tail regeneration. Dev. Biol. *316*, 323–35.

83. Nain, Z., Islam, M.A., Chowdhury, S.H., and Afroza, S. (2016). Current Understanding on Tail Regeneration in Green Anoles (*Anolis carolinensis*).

84. Lailvaux, S.P., Reaney, L.T., and Backwell, P.R.Y. (2009). Dishonest signalling of fighting ability and multiple performance traits in the fiddler crab *Uca mjoebergi*. Funct. Ecol. *23*, 359–66.

85. Mito, T., and Noji, S. (2008). The Two-Spotted Cricket *Gryllus bimaculatus*: An Emerging Model for Developmental and Regeneration Studies. Cold Spring Harb. Protoc. *3*, 1–15.

86. Bely, A.E., Zattara, E.E., and Sikes, J.M. (2014). Regeneration in spiralians: evolutionary patterns and developmental processes. Int. J. Dev. Biol. *58*, 623–34.

87. Ben Khadra, Y., Ferrario, C., Benedetto, C. Di, Said, K., Bonasoro, F., Candia Carnevali, M.D., and Sugni, M. (2015). Re-growth, morphogenesis, and differentiation during starfish arm regeneration. Wound Repair Regen. *23*, 623–34.

88. Reddien, P.W., and Sánchez Alvarado, A. (2004). Fundamentals of Planarian Regeneration. Annu. Rev. Cell Dev. Biol. *20*, 725–57.

89. Gierer, A., Berking, S., Bode, H., David, C.N., Flick, K., Hansmann, G., Schaller,

C.H., and Trenkner, E. (1972). Regeneration of hydra from reaggregated cells. Nat. New Biol. *239*, 98–101.

90. Saló, E., and Baguñà, J. (1984). Regeneration and pattern formation in planarians. J. Embryol Exp. Morph. *83*, 63–80.

91. Newmark, P.A., and Sánchez Alvarado, A. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. Dev. Biol. *220*, 142–53.

92. Paulus, T., and Müller, M. (2006). Cell proliferation dynamics and morphological differentiation during regeneration in *Dorvillea bermudensis* (Polychaeta, Dorvilleidae). J. Morphol. *267*, 393–403.

93. Konstantinides, N., and Averof, M. (2014). A common cellular basis for muscle regeneration in arthropods and vertebrates. Science. *343*, 788–91.

94. Candia Carnevali, M.D., Bonasoro, F., and Biale, A. (1997). Pattern of bromodeoxyuridine incorporation in the advanced stages of arm regeneration in the feather star *Antedon mediterranea*. Cell Tissue Res. *289*, 363–74.

95. Kondo, M., and Akasaka, K. (2010). Regeneration in crinoids. Dev. Growth Differ. *52*, 57–68.

96. Brockes, J.P., and Kumar, A. (2008). Comparative aspects of animal regeneration. Annu. Rev. Cell Dev. Biol. *24*, 525–49.

97. Weissman, I.L., Anderson, D.J., and Gage, F. (2001). Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. Annu. Rev. Cell Dev. Biol. *17*, 387–403.

98. Jopling, C., Sleep, E., Raya, M., Martí, M., Raya, A., and Izpisúa Belmonte, J.C. (2010). Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. Nature. *464*, 606–9.

99. Selman, K., and Kafatos, F.C. (1974). Transdifferentiation in the labial gland of silk moths: is DNA required for cellular metamorphosis? Cell Differ. *3*, 81–94.

100. Vogg, M.C., Wenger, Y., and Galliot, B. How Somatic Adult Tissues Develop Organizer Activity. Curr. Top. Dev. Biol., vol. 116. 1st ed., Elsevier Inc.; 2016, p. 391–414.

101. Nye, H.L.D., Cameron, J.A., Chernoff, E.A.G., and Stocum, D.L. (2003). Regeneration of the urodele limb: A review. Dev. Dyn. *226*, 280–94.

102. McCusker, C.D., and Gardiner, D.M. (2013). Positional Information Is Reprogrammed in Blastema Cells of the Regenerating Limb of the Axolotl (*Ambystoma mexicanum*). PLoS One. *8*, 1–14.

103. Tanaka, E.M., and Reddien, P.W. (2011). The cellular basis for animal regeneration. Dev. Cell. *21*, 172–85.

104. Galliot, B. (2012). Hydra, a fruitful model system for 270 years. Int. J. Dev. Biol. *56*, 411–23.

105. Browne, E.N. (1909). The production of new hydranths in hydra by the insertion of small grafts. J. Exp. Zool. *7*, 1–23.

106. Hicklin, J., and Wolpert, L. (1971). Positional information and pattern regulation in regeneration of hydra. Symp. Soc. Exp. Biol. *25*, 391–415.

107. Kadu, V., S, S.G., and Ghaskadbi, S. (2012). Induction of secondary axis in hydra revisited: New insights into pattern formation. Int J Mol Cell Med. *1*, 11–20.

108. Shimizu, H. (2012). Transplantation analysis of developmental mechanisms in *Hydra*. Int. J. Dev. Biol. *56*, 463–72.

109. Lengfeld, T., Watanabe, H., Simakov, O., Lindgens, D., Gee, L., Law, L., Schmidt, H.A., Özbek, S., Bode, H., and Holstein, T.W. (2009). Multiple Wnts are involved in *Hydra* organizer formation and regeneration. Dev. Biol. *330*, 186–99.

110. Nakamura, Y., Tsiairis, C.D., Özbek, S., and Holstein, T.W. (2011). Autoregulatory and repressive inputs localize *Hydra Wnt3* to the head organizer. Proc. Natl. Acad. Sci. U. S. A. *108*, 9137–42.

111. Broun, M., Gee, L., Reinhardt, B., and Bode, H.R. (2005). Formation of the head organizer in hydra involves the canonical Wnt pathway. Development. *132*, 2907–16.

112. Augustin, R., Franke, A., Khalturin, K., Kiko, R., Siebert, S., Hemmrich, G., and Bosch, T.C.G. (2006). Dickkopf related genes are components of the positional value gradient in Hydra. Dev. Biol. *296*, 62–70.

113. Guder, C., Pinho, S., Nacak, T.G., Schmidt, H.A., Hobmayer, B., Niehrs, C., and Holstein, T.W. (2006). An ancient Wnt-Dickkopf antagonism in *Hydra*. Development. *133*, 901–11.

114. Chera, S., Ghila, L., Wenger, Y., and Galliot, B. (2011). Injury-induced activation of the MAPK/CREB pathway triggers apoptosis-induced compensatory

proliferation in hydra head regeneration. Dev. Growth Differ. *53*, 186–201.

115. Kaloulis, K., Chera, S., Hassel, M., Gauchat, D., and Galliot, B. (2004). Reactivation of developmental programs: the cAMP-response element-binding protein pathway is involved in hydra head regeneration. Proc. Natl. Acad. Sci. U. S. A. *101*, 2363–8.

116. Chera, S., Ghila, L., Dobretz, K., Wenger, Y., Bauer, C., Buzgariu, W., Martinou, J.C., and Galliot, B. (2009). Apoptotic Cells Provide an Unexpected Source of Wnt3 Signaling to Drive *Hydra* Head Regeneration. Dev. Cell. *17*, 279–89.

117. Hobmayer, B., Rentzsch, F., Kuhn, K., Happel, C.M., von Laue, C.C., Snyder, P., Rothbächer, U., and Holstein, T.W. (2000). WNT signalling molecules act in axis formation in the diploblastic metazoan *Hydra*. Nature. *407*, 186–9.

118. Galliot, B. (2013). Regeneration in *Hydra*. eLS.

119. Hemmrich, G., Khalturin, K., Boehm, A.M., Puchert, M., Anton-Erxleben, F., Wittlieb, J., Klostermeier, U.C., Rosenstiel, P., Oberg, H.H., Domazet-Lošo, T., et al. (2012). Molecular Signatures of The Three Stem Cell Lineages in *Hydra* and the Emergence of Stem Cell Function at the Base of Multicellularity. Mol. Biol. Evol. *29*, 3267–80.

120. David, C.N., and Murphy, S. (1977). Characterization of Interstitial Stem Cells in Hydra by Cloning. Dev. Biol. *58*, 372–83.

121. Bosch, T.C.G., and David, C.N. (1987). Stem Cells of *Hydra magnipapillata* Can Differentiate into Somatic Cells and Germ Line Cells. Dev. Biol. *121*, 182–91.

122. Steele, R.E. (2002). Developmental signaling in *Hydra*: what does it take to build a "simple" animal? Dev. Biol. *248*, 199–219.

123. Hobmayer, B., Jenewein, M., Eder, D., Eder, M.K., Glasauer, S., Gufler, S., Hartl, M., and Salvenmoser, A.W. (2012). Stemness in *Hydra* - a current perspective. Int. J. Dev. Biol. *56*, 509–17.

124. David, C.N. (2012). Interstitial stem cells in *Hydra*: Multipotency and decision-making. Int. J. Dev. Biol. *56*, 489–97.

125. Tanaka, E.M. (2016). The Molecular and Cellular Choreography of Appendage Regeneration. Cell. *165*, 1598–608.

126. Butler, E.G. (1955). Regeneration of the urodele forelimb after reversal of its proximo-distal axis. J. Morphol. *96*, 265–81.

127. Nacu, E., and Tanaka, E.M. (2011). Limb Regeneration: A New Development? Annu. Rev. Cell Dev. Biol. *27*, 409–40.

128. Maden, M. (1982). Vitamin A and pattern formation in the regenerating limb. Nature. *295*, 672–5.

129. Maden, M. (1983). The effect of vitamin A on the regenerating amphibian limb. J Embryol Exp Morphol. *77*, 273–95.

130. Pecorino, L.T., Entwistle, A., and Brockes, J.P. (1996). Activation of a single retinoic acid receptor isoform mediates proximodistal respecification. Curr. Biol. *6*, 563–9.

131. Kragl, M., Knapp, D., Nacu, E., Khattak, S., Maden, M., Epperlein, H.H., and

Tanaka, E.M. (2009). Cells keep a memory of their tissue origin during axolotl limb regeneration. Nature. *460*, 60–5.

132. Shaikh, N., Gates, P.B., and Brockes, J.P. (2011). The Meis homeoprotein regulates the axolotl Prod 1 promoter during limb regeneration. Gene. *484*, 69–74.

133. Roensch, K., Tazaki, A., Chara, O., and Tanaka, E.M. (2013). Progressive specification rather than intercalation of segments during limb regeneration. Science. *342*, 1375–9.

134. Endo, T., Bryant, S. V., and Gardiner, D.M. (2004). A stepwise model system for limb regeneration. Dev. Biol. *270*, 135–45.

135. Satoh, A., Gardiner, D.M., Bryant, S. V., and Endo, T. (2007). Nerve-induced ectopic limb blastemas in the axolotl are equivalent to amputation-induced blastemas. Dev. Biol. *312*, 231–44.

136. Nacu, E., Gromberg, E., Oliveira, C.R., Drechsel, D., and Tanaka, E.M. (2016). FGF8 and SHH substitute for anterior–posterior tissue interactions to induce limb regeneration. Nature. *1*, 1–16.

137. Nacu, E., Glausch, M., Le, H.Q., Damanik, F.F.R., Schuez, M., Knapp, D., Khattak, S., Richter, T., Tanaka, E.M., Alvares, L.E., et al. (2013). Connective tissue cells, but not muscle cells, are involved in establishing the proximo-distal outcome of limb regeneration in the axolotl. Development. *140*, 513–8.

138. Sandoval-Guzmán, T., Wang, H., Khattak, S., Schuez, M., Roensch, K., Nacu, E.,

Tazaki, A., Joven, A., Tanaka, E.M., and Simon, A. (2014). Fundamental differences in dedifferentiation and stem cell recruitment during skeletal muscle regeneration in two salamander species. Cell Stem Cell. *14*, 174–87.

139. Currie, J.D., Kawaguchi, A., Traspas, R.M., Schuez, M., Chara, O., and Tanaka, E.M. (2016). Live Imaging of Axolotl Digit Regeneration Reveals Spatiotemporal Choreography of Diverse Connective Tissue Progenitor Pools. Dev. Cell. *39*, 411–23.

140. Harriet Randolph. (1897). Observations and experiments on regeneration in planarians. Arch. Fur Entwickelungsmechanik Der Org. *5*, 352–72.

141. Morgan, T.H. (1900). Regeneration in Planarians. Arch. Fur Entwickelungsmechanik Der Org. *10*, 58–119.

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

143. Rouhana, L., Weiss, J.A., Forsthoefel, D.J., Lee, H., King, R.S., Inoue, T., Shibata, N., Agata, K., and Newmark, P.A. (2013). RNA interference by feeding in vitro synthesized double-stranded RNA to planarians: methodology and dynamics. Dev. Dyn. *242*, 1–43.

144. Umesono, Y. (1997). A planarian *orthopedia* homolog is specifically expressed in the branch region of both the mature and regenerating brain. Dev. Growth ....

145. Pearson, B.J., Eisenhoffer, G.T., Gurley, K.A., Rink, J.C., Miller, D.E., and

Sánchez Alvarado, A. (2009). A Formaldehyde-based Whole-Mount In Situ Hybridization Method for Planarians. Dev. Dyn. *238*, 443–50.

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

147. Baguna, J., Saló, E., and Auladell, C. (1989). Regeneration and pattern formation in planarians. III . Evidence that neoblasts are totipotent stem cells and the source of blastema cells. Development. *86*, 77–86.

148. Baguñà, J., Salo, E., Romero, R., Garcia-fernàndez, J., and Bueno, D. (1994). Regeneration and pattern formation in planarians: cells, molecules and genes. Zoolog. Sci. *11*, 781–95.

149. Sánchez Alvarado, A., and Newmark, P.A. (1998). The use of planarians to dissect the molecular basis of metazoan regeneration. Wound Repair Regen. *6*, 413–20.

150. Chandebois, R. (1980). The dynamics of wound closure and its role in the programming of planarian regeneration. II—Distalization. Dev. Growth Differ. *22*, 693–704.

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

152. Wenemoser, D., and Reddien, P.W. (2010). Planarian regeneration involves distinct stem cell responses to wounds and tissue absence. Dev. Biol. *344*, 979–

91.

153. Guedelhoefer, O.C., and Sánchez Alvarado, A. (2012). Amputation induces stem cell mobilization to sites of injury during planarian regeneration. Development. *139*, 3510–20.

154. Pellettieri, J., Fitzgerald, P., Watanabe, S., Mancuso, J., Green, D.R., Sánchez Alvarado, A., and Sánchez, A. (2010). Cell death and tissue remodeling in planarian regeneration. Dev. Biol. *338*, 76–85.

155. Wenemoser, D., Lapan, S.W., Wilkinson, A.W., Bell, G.W., and Reddien, P.W. (2012). A molecular wound response program associated with regeneration initiation in planarians. Genes Dev. *26*, 988–1002.

156. Wurtzel, O., Cote, L.E., Poirier, A., Satija, R., Regev, A., and Reddien, P.W. (2015). A Generic and Cell-Type-Specific Wound Response Precedes Regeneration in Planarians. Dev. Cell. *35*, 632–45.

157. Witchley, J.N., Mayer, M., Wagner, D.E., Owen, J.H., and Reddien, P.W. (2013). Muscle Cells Provide Instructions for Planarian Regeneration. Cell Rep. *4*, 1–9.

158. Scimone, M.L., Kravarik, K.M., Lapan, S.W., and Reddien, P.W. (2014). Neoblast specialization in regeneration of the planarian *Schmidtea mediterranea*. Stem Cell Reports. *3*, 339–52.

159. van Wolfswinkel, J.C., Wagner, D.E., and Reddien, P.W. (2014). Single-Cell Analysis Reveals Functionally Distinct Classes within the Planarian Stem Cell Compartment. Cell Stem Cell. *15*, 1–14.

160. Dubois, F. Contribution a l'etude de la migration des cellules de regeneration chez les Planaires dulcicoles. Bulletin Biologique de la France et de le Belgique, 1949.

161. Wolff, E. Recent researches on the regeneration of planaria. Regen. 20th growth Symp. Ronald Press. New York, 1962, p. 53–84.

162. Wolff, E. (1948). Sur la migration des cellules de régénération chez les planaires. Rev. Suisse Zool.

163. Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sánchez Alvarado, A. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science. *310*, 1327–30.

164. Eisenhoffer, G.T., Kang, H., and Sánchez Alvarado, A. (2008). Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea*. Cell Stem Cell. *3*, 327–39.

165. Galloni, M. (2012). Global irradiation effects, stem cell genes and rare transcripts in the planarian transcriptome. Int. J. Dev. Biol. *56*, 103–16.

166. Labbé, R.M., Irimia, M., Currie, K.W., Lin, A., Zhu, S.J., Brown, D.D.R., Ross, E.J., Voisin, V., Bader, G.D., Blencowe, B.J., et al. (2012). A Comparative Transcriptomic Analysis Reveals Conserved Features of Stem Cell Pluripotency in Planarians and Mammals. Stem Cells. *30*, 1734–45.

167. Önal, P., Grün, D., Adamidi, C., Rybak, A., Solana, J., Mastrobuoni, G., Wang, Y., Rahn, H.-P., Chen, W., Kempa, S., et al. (2012). Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. EMBO

J. *31*, 2755–69.

168. Resch, A.M., Palakodeti, D., Lu, Y.-C., Horowitz, M., and Graveley, B.R. (2012). Transcriptome Analysis Reveals Strain-Specific and Conserved Stemness Genes in Schmidtea mediterranea. PLoS One. *7*, e34447.

169. Wagner, D.E., Ho, J.J., and Reddien, P.W. (2012). Genetic Regulators of a Pluripotent Adult Stem Cell System in Planarians Identified by RNAi and Clonal Analysis. Stem Cell. *10*, 299–311.

170. Wagner, D.E., Wang, I.E., and Reddien, P.W. (2011). Clonogenic Neoblasts Are Pluripotent Adult Stem Cells That Underlie Planarian Regeneration. Science. *332*, 811–6.

171. Reddien, P.W. (2013). Specialized progenitors and regeneration. Development. *140*, 951–7.

172. Lapan, S.W., and Reddien, P.W. (2011). *dlx* and *sp6-9* Control Optic Cup Regeneration in a Prototypic Eye. PLoS Genet. *7*, e1002226.

173. Lapan, S.W., and Reddien, P.W. (2012). Transcriptome Analysis of the Planarian Eye Identifies *ovo* as a Specific Regulator of Eye Regeneration. Cell Rep. *2*, 1–14.

174. Scimone, M.L., Srivastava, M., Bell, G.W., and Reddien, P.W. (2011). A regulatory program for excretory system regeneration in planarians. Development. *138*, 4387–98.

175. Scimone, M.L., Lapan, S.W., and Reddien, P.W. (2014). A *forkhead* Transcription

Factor Is Wound-Induced at the Planarian Midline and Required for Anterior Pole Regeneration. PLoS Genet. *10*, e1003999.

176. Vásquez-Doorman, C., and Petersen, C.P. (2014). *zic-1* Expression in Planarian Neoblasts after Injury Controls Anterior Pole Regeneration. PLoS Genet. *10*, e1004452.

177. Vogg, M.C., Owlarn, S., Pérez Rico, Y.A., Xie, J., Suzuki, Y., Gentile, L., Wu, W., and Bartscherer, K. (2014). Stem cell-dependent formation of a functional anterior regeneration pole in planarians requires Zic and Forkhead transcription factors. Dev. Biol. *390*, 136–48.

178. Forsthoefel, D.J., James, N.P., Escobar, D.J., Stary, J.M., Vieira, A.P., Waters, F. a, and Newmark, P. a. (2012). An RNAi screen reveals intestinal regulators of branching morphogenesis, differentiation, and stem cell proliferation in planarians. Dev. Cell. *23*, 691–704.

179. Cowles, M.W., Brown, D.D.R., Nisperos, S. V, Stanley, B.N., Pearson, B.J., and Zayas, R.M. (2013). Genome-wide analysis of the bHLH gene family in planarians identifies factors required for adult neurogenesis and neuronal regeneration. Development, 1–12.

180. Adler, C., Seidel, C., McKinney, S., and Alvarado, A. (2014). Selective amputation of the pharynx identifies a FoxA-dependent regeneration program in planaria. eLife, 1–22.

181. März, M., Seebeck, F., and Bartscherer, K. (2013). A Pitx transcription factor controls the establishment and maintenance of the serotonergic lineage in

planarians. Development. *140*, 4499–509.

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

183. Gurley, K.A., Rink, J.C., and Sánchez Alvarado, A. (2008). β-Catenin Defines Head Versus Tail Identity During Planarian Regeneration and Homeostasis. Science. *319*, 323–7.

184. Iglesias, M., Gomez-Skarmeta, J.L.J., Saló, E., and Adell, T. (2008). Silencing of *Smed-βcatenin1* generates radial-like hypercephalized planarians. Development. *135*, 1215–21.

185. Petersen, C.P., and Reddien, P.W. (2008). *Smed-βcatenin-1* is required for anteroposterior blastema polarity in planarian regeneration. Science. *319*, 327–30.

186. Petersen, C.P., and Reddien, P.W. (2009). A wound-induced Wnt expression program controls planarian regeneration polarity. Proc. Natl. Acad. Sci. U. S. A. *106*, 17061–6.

187. Gurley, K.A., Elliott, S.A., Simakov, O., Schmidt, H.A., Holstein, T.W., and Sánchez Alvarado, A. (2010). Expression of secreted Wnt pathway components reveals unexpected complexity of the planarian amputation response. Dev. Biol. *347*, 24–39.

188. Petersen, C.P., and Reddien, P.W. (2011). Polarized *notum* activation at wounds inhibits Wnt function to promote planarian head regeneration. Science. *332*, 852–

5.

189. Orii, H., and Watanabe, K. (2007). Bone morphogenetic protein is required for dorso-ventral patterning in the planarian *Dugesia japonica*. Dev. Growth Differ., 345–9.

190. Molina, M.D., 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. Dev. Biol. *311*, 79–94.

191. Reddien, P.W., Bermange, A.L., Kicza, A.M., and Sánchez Alvarado, A. (2007). BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration. Development. *134*, 4043–51.

192. Molina, M.D., Saló, E., and Cebrià, F. (2009). Expression pattern of the expanded noggin gene family in the planarian *Schmidtea mediterranea*. Gene Expr. Patterns. *9*, 246–53.

193. Gaviño, M.A., and Reddien, P.W. (2011). A Bmp/Admp Regulatory Circuit Controls Maintenance and Regeneration of Dorsal-Ventral Polarity in Planarians. Curr. Biol. *21*, 294–9.

194. Molina, M.D., Neto, A., Maeso, I., Gómez-Skarmeta, J.L., Saló, E., Cebrià, F., Go, L., and Salo, E. (2011). Noggin and noggin-like genes control dorsoventral axis regeneration in planarians. Curr. Biol. *21*, 300–5.

195. Scimone, M.L., Cote, L.E., Rogers, T., and Reddien, P.W. (2016). Two FGFRL-Wnt circuits organize the planarian anteroposterior axis. eLife. *5*, e12845.

196. Lander, R., and Petersen, C. (2016). Wnt, Ptk7, and FGFRL expression gradients control trunk positional identity in planarian regeneration. eLife. *5*, 1689–99.

197. Kobayashi, C., Umesono, Y., Sa, A., and Cebria, F. (2002). FGFR-related gene nou-darake restricts brain tissues to the head region of planarians. Nature. *419*.

198. Hill, E.M., and Petersen, C.P. (2015). Wnt/Notum spatial feedback inhibition controls neoblast differentiation to regulate reversible growth of the planarian brain. Development. *142*, 4217–29.

199. Gaviño, M.A., Wenemoser, D., Wang, I.E., and Reddien, P.W. (2013). Tissue absence initiates regeneration through Follistatin-mediated inhibition of Activin signaling. eLife. *2*, e00247.

200. Roberts-Galbraith, R.H., and Newmark, P.A. (2013). Follistatin antagonizes Activin signaling and acts with Notum to direct planarian head regeneration. Proc. Natl. Acad. Sci. U. S. A. *110*, 1363–8.

201. Owlarn, S., and Bartscherer, K. (2016). Go ahead, grow a head! A planarian's guide to anterior regeneration. Regeneration. *3*, 139–55.

Chapter 2

Landmarks in existing tissue at wounds are utilized to generate pattern in regenerating tissue

Isaac M. Oderberg, Dayan J. Li, M. Lucila Scimone, Michael A. Gaviño, and Peter W. Reddien

Experiments shown in Figures 1A-C and Figure S1 were performed by IMO and DJL. Experiments shown in Figures 1D-F, Figures 2-4, and Figures S2-S4 were performed by IMO. All authors contributed to the design of experiments or editing of the manuscript.

Published as:

Oderberg, I. M., Li, D. J., Scimone, M. L., Gavino, M. A., & Reddien, P. W. (2017). Landmarks in existing tissue at wounds are utilized to generate pattern in regenerating tissue. Curr. Biol., *27*, 733-42.

Abstract

Regeneration in many organisms involves the formation of a blastema, which differentiates and organizes into the appropriate missing tissues. How blastema pattern is generated and integrated with pre-existing tissues is a central question in the field of regeneration. Planarians are free-living flatworms capable of rapidly regenerating from small body fragments [1]. A cell cluster at the anterior tip of planarian head blastemas (the anterior pole) is required for establishment of anterior-posterior (AP) and midline blastema pattern [2–4]. Transplantation of the head tip into tails induced host tissues to grow patterned head-like outgrowths containing a midline. Given the important patterning role of the anterior pole, understanding how it becomes localized during regeneration would help explain how wounds establish pattern in new tissue. Anterior pole progenitors were specified at the pre-existing midline of regenerating fragments, even when this location deviated from the medial-lateral (ML) median plane of the wound face. Anterior pole progenitors were specified broadly on the dorsal-ventral (DV) axis, and subsequently formed a cluster at the DV boundary of the animal. We propose that three landmarks of pre-existing tissue at wounds set the location of anterior pole formation – a polarized AP axis, the pre-existing midline, and the dorsal-ventral median plane. Subsequently, blastema pattern is organized around the anterior pole. This process, utilizing positional information in existing tissue at unpredictably shaped wounds, can influence patterning of new tissue in a manner that facilitates integration with pre-existing tissue in regeneration.

Introduction

Pattern formation during animal development can be initiated by symmetry-breaking mechanisms including asymmetric maternal factors in oocytes and the location of sperm entry [5,6]. Similar to development, animal regeneration requires mechanisms to establish tissue pattern, in tissue outgrowths called blastemas. However, tissue pattern establishment in blastemas must occur in the absence of embryonic pattern-initiating processes. Furthermore, regeneration has the additional challenge of integrating new tissue with existing tissue in the context of unpredictable injuries. An elegant solution to this challenge would be if pattern-initiating processes relied on cues present at the wound face. Planarians are a classic regenerative model system and are capable of regenerating from a large array of injuries, making them well suited to address the origins of pattern in blastemas. Planarian regeneration involves both blastema formation and the remodeling of pre-existing tissue [1], and requires proliferative cells (neoblasts) that include pluripotent stem cells [7].

Planarian head regeneration involves the formation of a cluster of specialized muscle cells at the anterior head tip called the anterior pole [2]. The anterior pole expresses *notum* [8], *follistatin* [9,10], and the transcription factors *foxD* [2,10] and *zic-1* [3,4]. *foxD* [2] and *zic-1* [3,4] are required for anterior pole formation through the specialization of neoblasts into anterior pole progenitors. *foxD* or *zic-1* RNAi blocks pole regeneration and results in aberrant AP and ML patterning gene expression and absent or medially collapsed differentiated head tissues. These findings indicate a requirement for the anterior pole in AP and ML blastema patterning [2–4]. The planarian anterior pole has some similarities to other discrete regions of cells in developing embryos, such

as the amphibian Spemann-Mangold organizer, that regulate patterning of neighboring tissue [2,3,11]. Therefore, the position of the regenerated anterior pole is likely critical to proper patterning, but how this positioning is controlled is poorly understood.

Results and Discussion

Tissue fragments containing the anterior pole can induce patterned outgrowths in host tissue following transplantation

Transplantation was used to test the inductive capacities of the planarian head tip, which contains the anterior pole ("pole fragments") (Figure 1A, Figure S1A-G). Donor animals were lethally irradiated to ablate all neoblasts [12]; any resultant outgrowth involving new cells would therefore be produced by host tissues in response to the transplant. As a control, we also transplanted equally sized head tip regions that were offset from the midline and lacking the anterior pole ("flank fragments"). Juxtaposition of different body regions causes outgrowths in planarians [13–17], and accordingly, both pole and flank transplants generated outgrowths (Figure 1A).

Out of 86 "pole fragment" transplantations, 22 animals exhibited outgrowths, 18 of which possessed two eyes and were mobile (Figure 1A). Out of 66 "flank fragment" transplantations, 29 animals exhibited outgrowths, only six of which possessed eyes; the rest were immobile tissue spikes (Figure 1A). To assess patterning in these outgrowths, we performed fluorescent *in situ* hybridization (FISH) with RNA probes for *foxD/notum* (anterior pole), *sFRP-1* (anterior head tip), *slit* (midline), *ndl-2* (pre-pharyngeal region), *opsin* (photoreceptor neurons), and *laminB* (DV boundary). Pole transplant outgrowths with eyes (n = 8/8) and without eyes (n=3/4) displayed AP

patterning and a midline (Figure 1B, S1H). These outgrowths displayed entirely dorsal identity and little dorsal-ventral boundary (n=4/4) (Figure S1I). Flank transplant outgrowths with eyes also had AP patterning and a midline (n = 6/6), whereas the spike outgrowths expressed only *ndl-2* and no other anterior or midline markers (n = 8/8) (Figure 1B, S1H). Transplanted flank regions thus had less efficient, but not absent, potential to induce head pattern. Together with prior RNAi experiments that ablated the anterior pole, these data indicate an important role for cells at the anterior head tip in organizing head pattern. This motivated study of pole formation mechanisms to understand the logic by which head blastema pattern is formed and integrated with the pattern of pre-existing tissue.

Anterior pole progenitors are specified medially and in a broad DV domain

To understand pole positioning during regeneration, we first characterized anterior pole progenitor specification. Anterior pole progenitors are specified by expression of transcription factors such as *foxD* and *zic-1* and appear medially at anterior-facing wounds [3,4,18]. The distribution of pole progenitors along the DV axis is poorly understood, so we imaged and quantified anterior pole progenitor specification in blastemas viewed *en face* (head-on) (Figure 1C, Figure S2A).

Pole progenitors were detected with *foxD* and *notum* expression. Because *notum* is also expressed in the brain [8,19], we examined only *foxD*/*notum* double-positive cells, which were present by 48 hours post amputation (hpa). These cells were centered at the ML median plane (midpoint between the right and left wound sides) from 48 hpa

to 72 hpa (Figures 1C-D). The position of pole progenitors was quantified relative to the DV median plane (the midpoint between the dorsal and ventral sides) estimated using *laminB* expression (which marks the dorsal-ventral boundary [20]). *foxD+/notum+* cells were initially dispersed on the DV axis (Figure 1C, Figure 1E). The distributions of these progenitors grew narrower and closer to the DV median plane with time, accumulating in a cluster by 72 hpa (Figure 1C, Figure 1E). Similarly, the average distance of anterior pole progenitors to the DV median plane grew smaller with time (Figure 1F). *zic-1+/notum+* cells and *follistatin+/notum+* cells displayed similar dynamics (Figures S2B-H).

We hypothesized that the *foxD+/notum+* cells include early stage anterior pole progenitors, which are the result of neoblast specialization. Irradiation depletes neoblasts (Figure S2J) and, consistent with previous reports [2,3], blocked formation of anterior pole progenitors and the pole (Figures S2K). To support the idea that the *foxD+/notum+* cells observed were the result of neoblast specialization, animals were labeled with RNA probes to *foxD* and *smedwi-1*, which is a marker for neoblasts [26]. *foxD+/smedwi-1+* cells were present from 24 to 72 hpa (Figure 2A) and were biased medially (Figure 2B), similar to *foxD+/notum+* cells. However, the distributions of *foxD+/smedwi-1+* cells along the DV axis did not become significantly narrower over time (Figure 2C) and the distance between the *foxD+/smedwi-1+* cells and the DV median plane did not become smaller over time (Figure S2I). These results suggest that anterior pole progenitors are specialized from neoblasts at the ML median plane, but broadly along the DV axis, only accumulating at the DV median plane as post-mitotic progenitors.

Anterior Pole Cells Coalesce at the DV Median Plane

Neoblast-derived progenitors have been observed to move from a specification location to their final destination for several tissue types, notably for the epidermis [21,22] and the eye [23,24]. To assess possible movement of dispersed pole progenitors along the DV axis into the anterior pole, we utilized bromodeoxyuridine (BrdU) labeling. BrdU is incorporated into neoblasts and can be used to trace the behavior of a cell cohort [25]. Animals were amputated, immediately BrdU-pulsed, and analyzed in a regeneration time course (Figure 2D). The average distance of *foxD*+/BrdU+ cells to the DV median plane was smaller at 60 and 72 hpa than at 48 hpa (Figure 2E), even as the number of *foxD*+/BrdU+ cells per animal remained relatively constant (Figure 2F). These results are consistent with net movement of pole progenitors from being broadly dispersed on the DV axis into a coalesced anterior pole.

Anterior Pole Progenitor Specification Occurs at the Old Midline

Anterior pole progenitor specification could in principle occur either at the ML median plane of the wound, such as might be the case if the blastema involves *de novo* pattern organization, or at the midline of pre-existing tissue. To distinguish between these possibilities, animals were subjected to either transverse, oblique, or sagittal amputations, and the expression of *notum*, *laminB*, and *slit* (which marks the midline [26]) was examined. In transversely amputated animals, the anterior pole formed at the pre-existing midline (yellow arrowheads), which was coincident with the ML median plane of the anterior-facing wound (white arrowheads) (Figure 3A, Figure S3A).

Strikingly, in the oblique and sagittal amputations, the anterior pole progenitors formed at the pre-existing midline, even though this point was far from the ML median plane of the wound. (Figure 3A, Figure S3A). Although the pole was shifted away from the ML median plane, with time these animals regenerated with the correct shape (Figure S3B). *foxD+/smedwi-1+* cells (Figure 3B) were also shifted away from the ML median plane of the wound in oblique and sagittal fragments (Figure 3C). Instead, the pole-specialized neoblasts were present at the pre-existing *slit+* midline, suggesting that the anterior pole progenitor specification zone is biased to be within the pre-existing midline (Figure 3D).

If pole progenitors are specified at a pre-existing midline, what happens to planarian fragments lacking a midline? To address this question, we generated fragments that had little-to-no pre-existing midline with parasagittal amputation (Figure S3D). As expected, the anterior pole in the parasagittal thick fragments regenerated at the pre-existing midline (Figure S3E). The thin fragments, which had no pre-existing midline, displayed new *slit* expression, followed by an anterior pole (Figure S3F). This result is consistent with the possibility that anterior pole formation happens at the midline, and that body fragments lacking a midline undergo a process of *de novo* midline formation prior to anterior pole formation.

The Plane of Symmetry in Asymmetric Fragments is Centered on the Anterior Pole

The asymmetric formation of the anterior pole in amputated fragments lacking initial ML symmetry (Figure 3A) is consistent with the fact that obliquely amputated planarian body fragments produce head blastemas offset from the center of the wound face

[1,27]. We assessed the expression patterns of multiple patterning genes (*sFRP-1, ndl-2, wnt2, slit,* and *nlg7*) in these regenerating, asymmetric fragments at seven days post amputation. The plane of spatial expression symmetry for each of these patterning molecules was centered at the regenerating anterior pole (Figure 3E, Figure S3C). The plane of symmetry during regeneration for many differentiated tissues, including the cephalic ganglia (*PC2*+), mechanosensory neurons (*cintillo*+), photoreceptor neurons (*opsin*+), GABAergic neurons (*gad*+), ciliated epidermis (*rootletin*+), secretory cells (*mag-1*+), and intestine (*madt*+), was also aligned with the location of the anterior pole (Figure 3E, Figure S3C). These results demonstrate for many tissues and gene expression domains that symmetry is centered around the pole, even if this position is offset from the ML midpoint of the wound face. These results are consistent with the role of the anterior pole in facilitating organization of new tissue pattern.

Midline Patterning Molecules Affect the Medial Zone of Pole Progenitor Specification

The appearance of pole progenitors at the prior (*slit*+) midline raised the possibility that the environment of the prior midline is permissive for pole progenitor specification at anterior-facing wounds. To test this possibility, animals were subjected to RNAi of *wnt5* and *slit*, which regulate the planarian ML axis [26,28]. *wnt5* negatively regulates the expression domain of *slit.* After 8 dsRNA feedings, *wnt5(RNAi)* animals displayed ectopic lateral eyes, and *slit(RNAi)* animals displayed ectopic medial eyes, confirming RNAi had perturbed medial-lateral pattern (Figure 3F). Furthermore, the *slit* expression domain was wider in *wnt5(RNAi)* animals and *slit* expression in *slit(RNAi)* animals was

reduced and narrower than in controls (Figures S3G-H). Some *wnt5(RNAi)* animals had wider anterior poles than did control animals, whereas some of the *slit(RNAi)* animals had narrower poles than did control animals (Figures 3G-H). When we examined the distributions of pole-specialized neoblasts in these animals, we found that the distributions were wider in *wnt5(RNAi)* animals (Figures S3I-J). We conclude that the zone marked by the expression of *slit,* and defined by antagonistic roles for *wnt5* and *slit,* regulates the ML zone competent for pole progenitor specification at anterior-facing wounds.

Anterior Pole Progenitors Accumulate at the DV Median Plane of the Blastema

We next sought to understand how the anterior pole is placed at a particular location along the DV axis. Of the many surgeries applied to planarians, few create asymmetry along this shortest of planarian axes. We developed a DV oblique cut involving transverse amputation and dorsal tissue removal, creating a wound with more dorsal than ventral tissue removed (Figure 4A). Despite the severity of this wound, these animals eventually regenerated relatively normally (Figure S4A).

DV oblique animals formed a line of *laminB*+ cells in the blastema that was shifted dorsally from the preexisting *laminB*+ plane, consistent with a dorsal shift in the boundary formed between the dorsal and ventral epidermis during wound closure (Figure 4A-B). Early after DV oblique injury, pole progenitors were distributed along the DV axis; at later timepoints, the anterior pole coalesced far dorsal to the DV median plane of pre-existing tissue (Figure 4B-D). *laminB* expression in the blastema was used

to estimate the DV boundary in the blastema at 72 hpa. Pole progenitors at 72 hpa were coincident with this new *laminB* expression plane (Figure 4E), suggesting that anterior pole progenitors accumulate near the approximate location of contact between the dorsal and ventral sides formed early during regeneration (DV boundary). At 7 days post injury, the expression gap in *ndl-2*, *sFRP-1* expression, and the location of the brain as visualized by *PC2* expression, were all shifted dorsally and retained their relative position to the anterior pole (Figure S4B).

Perturbation of Bmp Signaling Affects the DV Positioning of Pole Progenitors

To determine if the pre-existing DV axis has a role in positioning anterior pole progenitors, animals were subjected to RNAi of *bmp4* and *smad1*, which normally promote dorsal tissue identity for the planarian DV axis [29–31]. As previously reported, the *laminB* expression domain was thickened with ectopic dorsal patches in both *bmp4* RNAi and *smad1* RNAi animals (Figure 4F, Figure S4C) [29,32]. RNAi of *bmp4* and *smad1* causes ventralization; cells responding to Bmp activity levels for localization should be shifted dorsally when that signal is reduced. Indeed, anterior pole progenitors were shifted dorsally at 48 hpa relative to *laminB*+ cells in pre-existing tissue in both *bmp4* and *smad1* RNAi animals (Figure 4F). This result was confirmed by measuring the position of the anterior pole progenitors relative to the pre-existing DV median plane, and by examining their distributions (Figure 4G). Quantification of anterior pole progenitors at 72 hpa revealed that this dorsal bias persisted (Figure 4H). These results indicate that the correct DV localization of anterior pole progenitors requires Bmp signaling.

Conclusions

One of the central challenges of regeneration is having a system that is capable of responding to injuries with different wound site architectures to produce a correctly patterned animal. Regenerative tissue outgrowths (blastemas) likely initiate pattern formation by a mechanism distinct from what occurs at the beginning of embryogenesis given the drastically different starting conditions for these processes. Given prior RNAi data [2–4] and the head tip transplantation results, the anterior pole likely has an important role in determining how the planarian head blastema organizes its pattern. Our findings suggest that the position of anterior pole regeneration relies on three cues: an anterior-facing wound [2–4], the pre-existing midline, and the DV boundary in the blastema (Figure 4I). The positioning of the pole at the midline during regeneration can allow ML pattern of new tissue to align with ML pattern of pre-existing tissue. Coordinating AP axis information with a DV axial plane to set a point of tissue organization and growth has parallels to other biological systems, such as in *Drosophila* imaginal discs [33,34]. Positioning of an organizer at the DV median plane in regeneration could facilitate a vector of growth and pattern formation on the AP axis perpendicular to the old DV axis. Planarians use positional information actively as adults for maintenance and restoration of axial pattern [35]. We conclude that the process of anterior pole formation during regeneration integrates axial cues at wounds, providing a mechanism to coordinate patterning and growth during head regeneration in a manner coherent with pre-existing tissue pattern (Figure 4J).

Figure 1

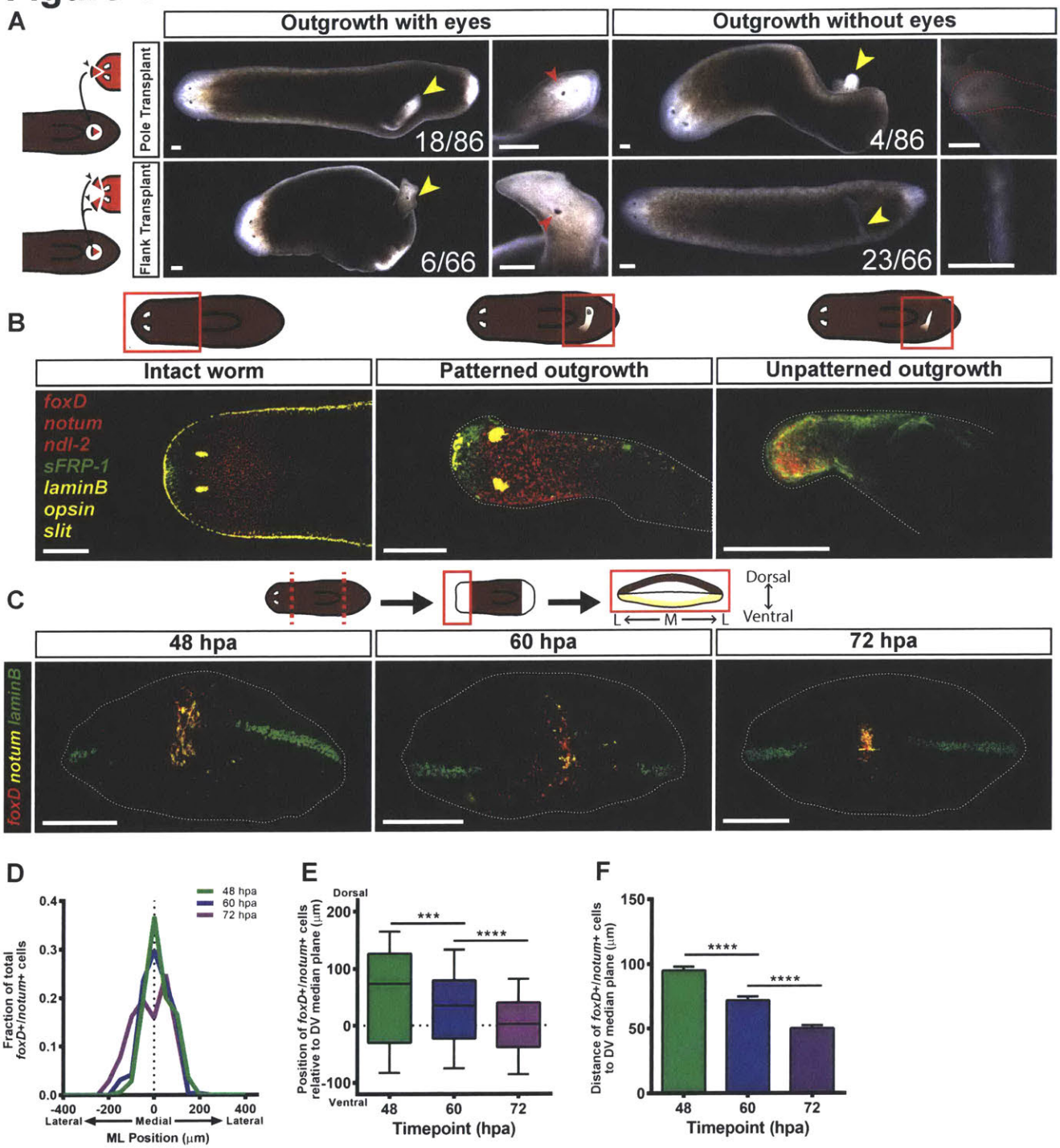

A
Outgrowth with eyes
Outgrowth without eyes
Pole Transplant
Flank Transplant
18/86
4/86
6/66
23/66
B
Intact worm
Patterned outgrowth
Unpatterned outgrowth
foxD
notum
ndl-2
sFRP-1
laminB
opsin
slit
C
Dorsal
Ventral
L
M
L
48 hpa
60 hpa
72 hpa
foxD notum laminB
D
48 hpa
60 hpa
72 hpa
Fraction of total foxD+/notum+ cells
0.4
0.3
0.2
0.1
0.0
-400
-200
0
200
400
Lateral
Medial
Lateral
ML Position (μm)
E
Dorsal
Ventral
Position of foxD+/notum+ cells relative to DV median plane (μm)
200
100
0
-100
***
****
48
60
72
Timepoint (hpa)
F
Distance of foxD+/notum+ cells to DV median plane (μm)
150
100
50
0
****
****
48
60
72
Timepoint (hpa)

Figure 1. Anterior pole progenitors are specified medially and in a broad DV domain. (A) Live images of animals 14 days after pole or flank transplantation. Yellow arrowheads mark the site of an outgrowth. Red arrowheads mark an ectopic eye. **(B)** Triple FISH for *foxD*/*notum*/*ndl-2*, *sFRP-1*, and *slit*/*laminB*/*opsin* at 14 days following transplantation. Merged images are displayed. For single-channel images see Figure S1H. For **(A-B)** dorsal view, anterior is to the left. Scale bars, 200 μm. **(C)** Triple FISH for *foxD*, *notum* and *laminB* at 48, 60, and 72 hpa. Medial *foxD*+/*notum*+ cells coalesce to the DV median plane with time. Maximal intensity projections shown. *en face* view of anterior blastema, dorsal is up. Scale bars, 200 μm. **(D)** The distribution of the positions of *foxD*+/*notum*+ cells on the ML axis is centered roughly around the ML median plane of the fragment (dotted line). 48 hpa, n=18 worms; 60 hpa, n=24 worms; 72 hpa, n=18 worms. **(E)** The distribution of the positions of *foxD*+/*notum*+ cells on the DV axis becomes tighter and closer to the DV median plane over time. Data were compared with a Mann-Whitney test. 48 vs 60 hpa, ***p=0.0003, n>289 cells; 60 vs 72 hpa, ****p<0.0001, n>289 cells. **(F)** The distance between *foxD*+/*notum*+ cells and the DV median plane grows smaller over time. Data were compared using a Student's t-test. 48 vs 60 hpa, ****p<0.0001, n>289 cells; 60 vs 72 hpa, ****p<0.0001, n>289 cells.

See also Figure S1 and Figure S2.

Figure 2

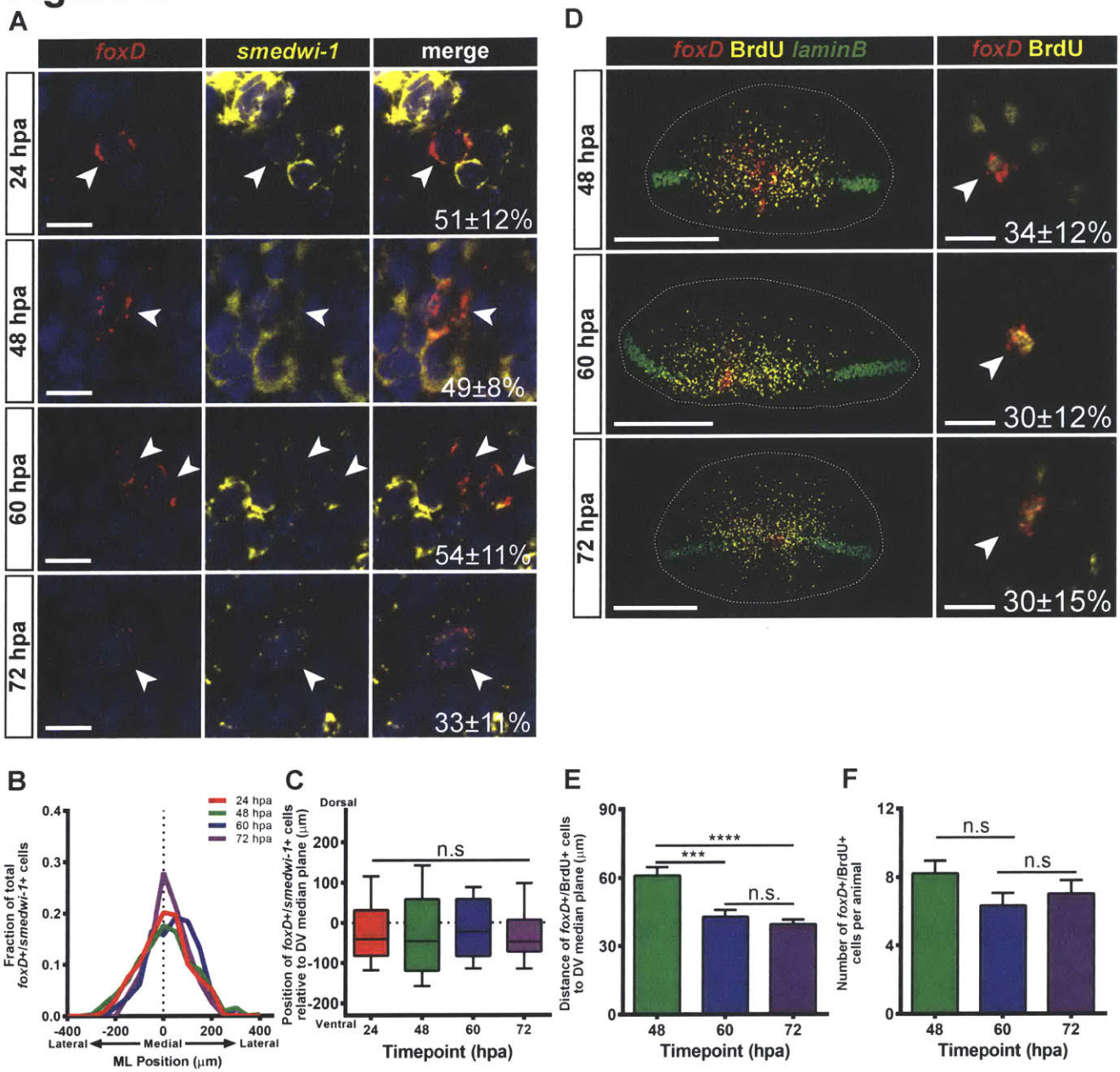

A
foxD
smedwi-1
merge
24 hpa
48 hpa
60 hpa
72 hpa
51±12%
49±8%
54±11%
33±11%
D
foxD BrdU laminB
foxD BrdU
34±12%
30±12%
30±15%
B
24 hpa
48 hpa
60 hpa
72 hpa
Fraction of total foxD+/smedwi-1+ cells
Lateral
Medial
Lateral
ML Position (μm)
C
Dorsal
Ventral
Position of foxD+/smedwi-1+ cells relative to DV median plane (μm)
n.s
Timepoint (hpa)
E
Distance of foxD+/BrdU+ cells to DV median plane (μm)
****
***
n.s.
Timepoint (hpa)
F
Number of foxD+/BrdU+ cells per animal
n.s
n.s
Timepoint (hpa)

Figure 2. Anterior pole progenitors are specified medially and in a broad DV domain.

(A) Double FISH for *foxD* and *smedwi-1* at 24, 48, 60, and 72 hpa. Nuclear signal (DAPI) is shown in blue. A proportion of *foxD*+ cells were also *smedwi-1*+ at all timepoints examined (24 hpa, 51±21%, n = 810 cells; 48 hpa, 49±8%, n = 763 cells; 60 hpa, 54±11%, n = 211 cells; 72 hpa, 33±11%, n = 242 cells). **(B)** The distribution of the positions of *foxD*+/*smedwi-1*+ cells on the ML axis is centered roughly around the ML median plane of the fragment (dotted line). 24 hpa, n=29 worms; 48 hpa, n=19 worms; 60 hpa, n=8 worms; 72 hpa, n=8 worms. **(C)** *foxD*+/*smedwi-1*+ cells are spread along the DV axis. DV median plane, dotted line. Data compared by Kruskal-Wallis test. n.s. n>75 cells. **(D)** Double FISH for *foxD* and *laminB*, and immunolabeling for BrdU at 48, 60, and 72 hpa. Left panel: *foxD*+/BrdU+ cells are scattered along the DV axis, and coalesce to a cluster with time. Right panel: White arrowhead indicates *foxD*+/BrdU+ cell. Percent of *foxD*+ cells which were BrdU+, lower right. **(E)** The distance between *foxD*+/BrdU+ cells and the DV median plane is smaller at 60 and 72 hpa than it is at 48 hpa. Data were compared by Student's t-test. 48 hpa vs 60 hpa, ***p=0.0007, n>114 cells; 48 hpa vs 72 hpa, ****p<0.0001, n>189 cells; 60 hpa vs 72 hpa, n.s., n>114 cells. **(F)** The number of *foxD*+/BrdU+ cells per animal is similar at all timepoints examined. Data were compared by Student's t-test. 48 hpa vs 60 hpa, n.s., n>18 animals; 60 hpa vs 72 hpa, n.s., n>18 animals. All images are an *en face* view of anterior blastema, dorsal is up. Left panels of **(D)**, maximal intensity projections shown. Scale bars, 200 μm. For **(A)** and right panels of **(D)** single confocal slices shown. Scale bars, 10 μm. For **(E-F)** data are represented as mean ± SEM. See also Figure S2.

Figure 3

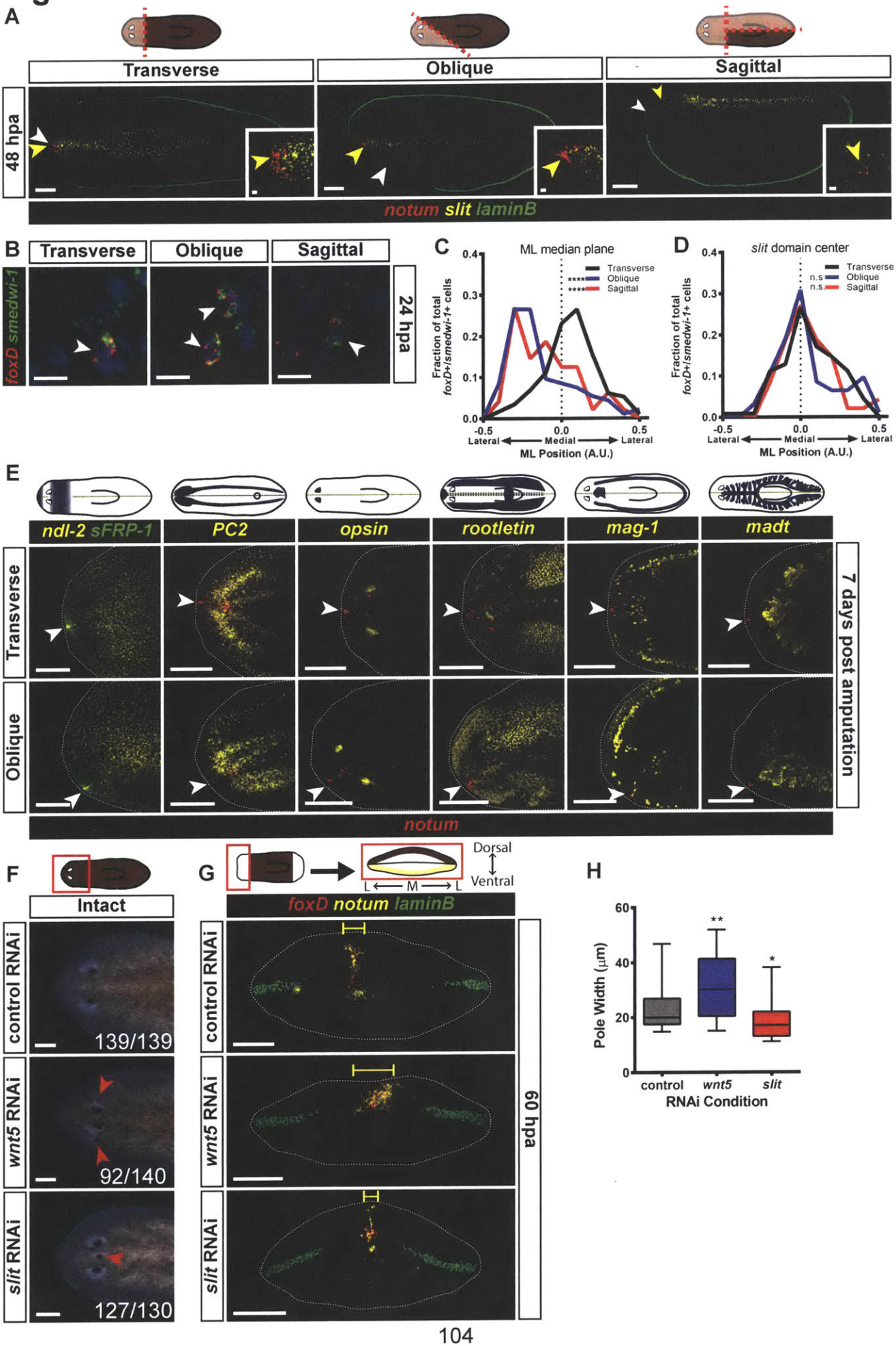

A
Transverse
Oblique
Sagittal
48 hpa
notum slit laminB
B
Transverse
Oblique
Sagittal
foxD smedwi-1
24 hpa
C
ML median plane
Transverse
Oblique
Sagittal
****
****
Fraction of total foxD+/smedwi-1+ cells
0.4
0.3
0.2
0.1
0.0
-0.5
0.0
0.5
Lateral
Medial
Lateral
ML Position (A.U.)
D
slit domain center
Transverse
Oblique
Sagittal
n.s
n.s
Fraction of total foxD+/smedwi-1+ cells
ML Position (A.U.)
E
ndl-2 sFRP-1
PC2
opsin
rootletin
mag-1
madt
Transverse
Oblique
7 days post amputation
notum
F
Intact
control RNAi
wnt5 RNAi
slit RNAi
139/139
92/140
127/130
G
Dorsal
Ventral
L
M
L
foxD notum laminB
control RNAi
wnt5 RNAi
slit RNAi
60 hpa
H
Pole Width (μm)
60
40
20
0
**
*
control
wnt5
slit
RNAi Condition

Figure 3. Anterior pole formation occurs at the prior midline.

(A) Triple FISH for *notum*, *slit* and *laminB* at 48 hpa for transverse, oblique or sagittal amputations. *notum*+ cells appear near the *slit* domain in asymmetric wounds. White arrowheads, middle of anterior-facing wound; yellow arrowheads, middle of *slit* domain. Cartoon demonstrates surgery. Transverse 48 hpa, n=5/5; Oblique 48 hpa, n=10/10; Sagittal 48 hpa n=9/9. **(B)** Triple FISH for *foxD*, *smedwi-1* and *laminB* at 24 hpa for transverse, oblique or sagittal amputations. Nuclear signal (DAPI) is shown in blue. *foxD*+/*smedwi-1*+ cells were quantified. **(C-D)** Distribution of *foxD*+/*smedwi-1*+ cells along the ML axis at 24 hpa for transverse, oblique or sagittal amputations. **(C)** Reference is the ML median plane of the wound (dotted line). Transverse amputation distributions were centered on this point, oblique and sagittal distributions were offset. Data were compared by a Kolmogorov-Smirnov Test. Transverse vs Oblique ****$p<0.0001$ $n>94$ cells; Transverse vs Sagittal ****$p<0.0001$ $n>48$ cells. **(D)** Reference is the *slit* domain center (dotted line). Transverse, oblique and sagittal amputation distributions were centered on this point. Data were compared by a Kolmogorov-Smirnov Test. Transverse vs Oblique n.s. $n>94$ cells; Transverse vs Sagittal n.s. $n>48$ cells. **(E)** Triple FISH for *notum*, and *ndl-2* and *sFRP-1*, *PC2*, *opsin*, *rootletin*, *mag-1* or *madt* at 7 dpa for transverse and oblique amputations. White arrowheads, anterior pole. For clarity, ventral images (*PC2*) were vertically flipped. The plane of symmetry for gene expression was centered at the position of the anterior pole. **(F)** Live images at 28 days following inhibition of control gene, *wnt5* or *slit*. *wnt5(RNAi)* animals displayed ectopic lateral eyes and *slit(RNAi)* animals displayed ectopic medial eyes. Number of animals

with phenotype, lower right. Red arrowheads, ectopic eyes. **(G)** Triple FISH for *foxD*, *notum* and *laminB* at 60 hpa following inhibition of control gene, *wnt5* or *slit*. *foxD+/notum+* cells occupied a wider area in *wnt5(RNAi)* animals, and a narrower area in *slit(RNAi)* animals, compared to control RNAi animals. Yellow bracket, pole width. **(H)** Each animal was assigned a "pole width" score, which is the average distance of a cell to the center of its pole for that animal. *wnt5(RNAi)* animals had wider anterior poles and *slit(RNAi)* animals had narrower anterior poles. Data were compared by a Mann-Whitney test. control RNAi vs *wnt5* RNAi, **p=0.0025, n>41 animals; control vs *slit* RNAi, *p=0.0137, n>44 animals. For **(A)**, **(E)**, and **(F)** maximal intensity projections shown. Dorsal view, anterior is to the left. Scale bars, 200 μm. Inset scale bars, 20 μm. For **(B)** single confocal slices shown. *en face* view of anterior blastema, dorsal is up. Scale bars, 10 μm. For **(G)** maximal intensity projections shown. *en face* view of anterior blastema, dorsal is up. Scale bars, 200 μm. See also Figure S3.

Figure 4

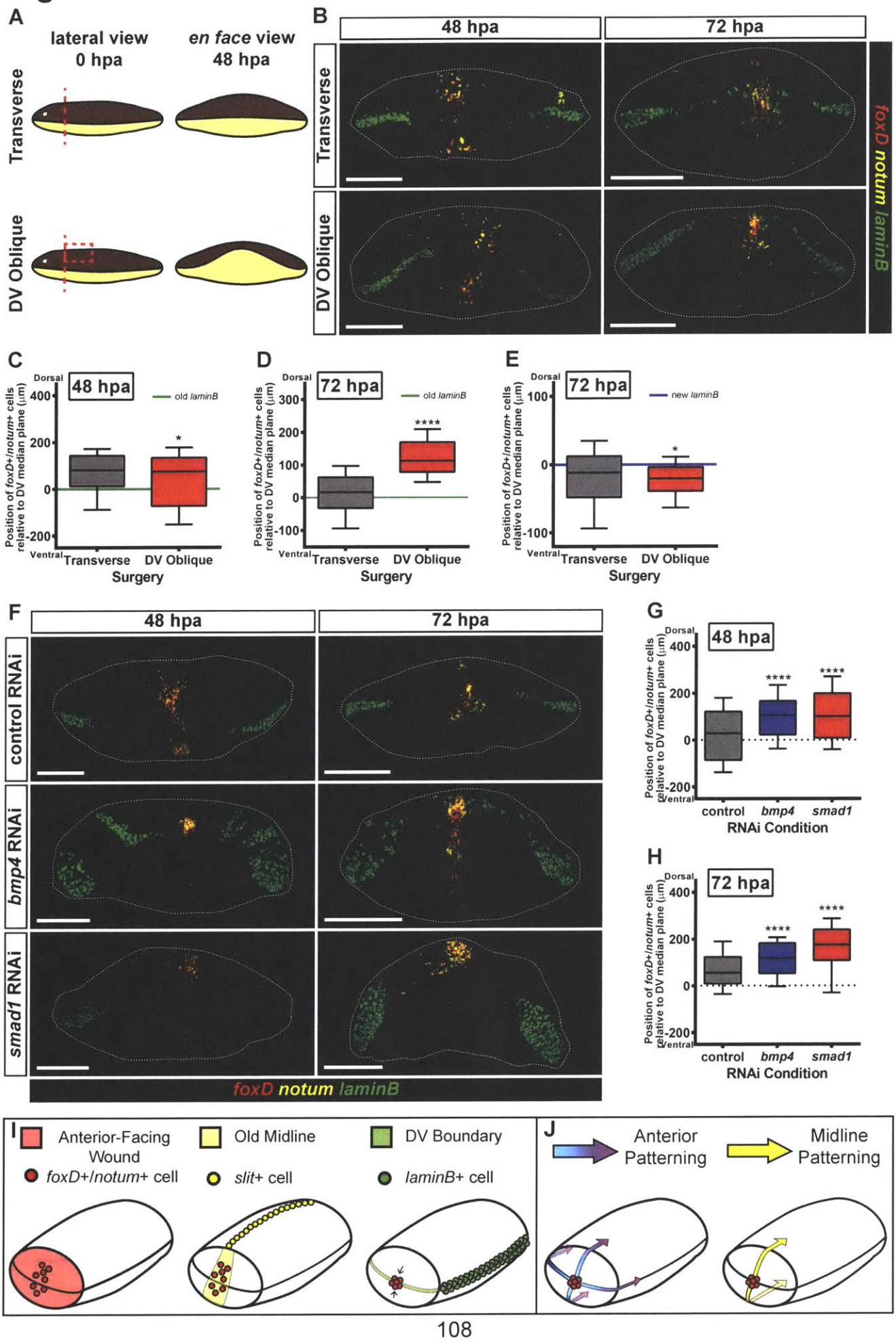

A
lateral view
0 hpa
en face view
48 hpa
Transverse
DV Oblique
B
48 hpa
72 hpa
Transverse
DV Oblique
foxD notum laminB
C
48 hpa
old laminB
Position of foxD+/notum+ cells relative to DV median plane (µm)
Dorsal
Ventral
400
200
0
-200
*
Transverse
DV Oblique
Surgery
D
72 hpa
old laminB
300
200
100
0
-100
****
E
72 hpa
new laminB
100
0
-100
*
F
48 hpa
72 hpa
control RNAi
bmp4 RNAi
smad1 RNAi
foxD notum laminB
G
48 hpa
****
control
bmp4
smad1
RNAi Condition
H
72 hpa
I
Anterior-Facing Wound
foxD+/notum+ cell
Old Midline
slit+ cell
DV Boundary
laminB+ cell
J
Anterior Patterning
Midline Patterning

Figure 4. The final location of the anterior pole is influenced by the DV axis.

(A) Left panel: Cartoon of a transverse and DV oblique amputation. To generate DV oblique fragments, animals had a transverse amputation followed by a precise removal of dorsal tissue. Lateral view, dorsal is up. Right panel: In order to close the wound following a DV oblique cut, ventral tissue must migrate further to make contact with dorsal tissue and seal the wound. *en face* view, dorsal is up. **(B)** Triple FISH for *foxD*, *notum* and *laminB* at 48 and 72 hpa following transverse or DV oblique amputation. *foxD+/notum+* cells accumulate at the DV boundary (*laminB+*) over time. **(C-E)** The distribution of *foxD+/notum+* cells on the DV axis. The transverse and DV oblique conditions were compared by a Mann-Whitney test. **(C)** At 48 hpa, *foxD+/notum+* cells were broader in DV oblique animals, but not dorsally shifted. *$p<0.0417$, $n>228$ cells. **(D)** At 72 hpa, *foxD+/notum+* cells were shifted dorsally in DV oblique animals. ****$p<0.0001$, $n>297$ cells. **(E)** At 72 hpa, *foxD+/notum+* cells were close to the DV boundary. *$p<0.0152$, $n>230$ cells. **(F)** Triple FISH for *foxD*, *notum* and *laminB* at 48 and 72 hpa following inhibition of control gene, *bmp4* or *smad1*. *foxD+/notum+* cells were shifted dorsally in *bmp4(RNAi)* and *smad1(RNAi)* as compared to control RNAi animals. **(G-H)** The distribution of *foxD+/notum+* cells on the DV axis. Data were compared by a Mann-Whitney test. **(G)** At 48 hpa, *foxD+/notum+* cells were dorsally shifted following inhibition of *bmp4* or *smad1*. control RNAi vs *bmp4* RNAi, ****$p<0.0001$, $n>346$ cells; control RNAi vs *smad1* RNAi, ****$p<0.0001$, $n>354$ cells. **(H)** At 72 hpa, *foxD+/notum+* cells were dorsally shifted following inhibition of *bmp4* or *smad1*. control RNAi vs *bmp4* RNAi, ****$p<0.0001$, $n>295$ cells; control RNAi vs *smad1* RNAi, ****$p<0.0001$, $n>295$ cells. For **(B)** and **(F)** maximal intensity projections shown. *en face* view of anterior

blastema, dorsal is up. Scale bars, 200 μm. For **(C-D)** and **(G-H)** the reference point for these coordinates is the DV median plane, which is estimated using old *laminB* expression (green or dotted line). For **(E)** the reference point for these coordinates is the DV boundary in the blastema, which was estimated using new *laminB* expression (blue line). **(I)** Anterior pole formation relies on three landmarks at wounds in order to integrate the pattern of the new and pre-existing tissues: an anterior-facing wound, the prior midline, and the boundary between the dorsal and ventral sides of the animal. **(J)** Once the anterior pole is formed, it acts to help pattern the AP and ML axes of the regenerating head. See also Figure S4.

Figure S1

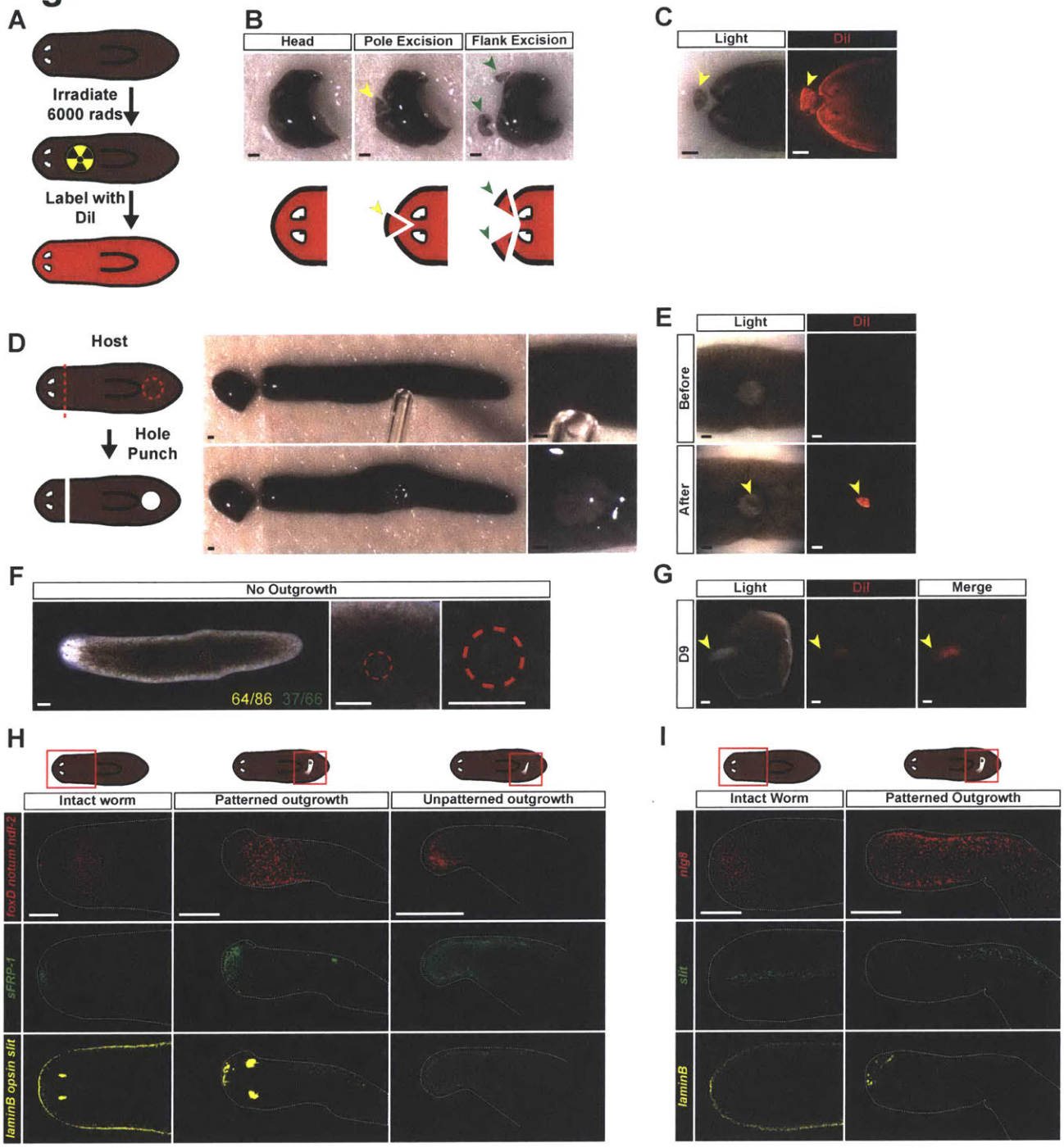

A
Irradiate
6000 rads
Label with
DiI
B
Head
Pole Excision
Flank Excision
C
Light
DiI
D
Host
Hole
Punch
E
Light
DiI
Before
After
F
No Outgrowth
64/86
37/66
G
Light
DiI
Merge
D9
H
Intact worm
Patterned outgrowth
Unpatterned outgrowth
foxD notum ndl-2
sFRP-1
laminB opsin slit
I
Intact Worm
Patterned Outgrowth
nlg8
slit
laminB

Figure S1. Pole transplantation procedure and outcomes. Related to Figure 1.

(A) Donors were lethally irradiated (6000 rads) 24 hours prior to transplantation. Animals were incubated in a 1:500 dilution of DiI (2 μg/mL) in planaria H2O overnight. **(B)** The donors' heads were amputated. From a single donor, a pole piece was excised, and transplanted, and then two flank pieces were excised, and transplanted. **(C)** The donor heads and pole fragments were positive for DiI signal. **(D)** Hosts were placed on a filter paper over ice, and a glass capillary was used to create a "hole punch" wound in the host. **(E)** The pole fragment or flank fragment was transplanted into the "hole punch" wound, and the animals was placed in a recovery chamber. **(F)** A majority of transplantations resulted in no outgrowth. n=64/86 pole transplants and n=37/66 flank transplants resulted in no outgrowth. A red dotted circle marks the site of failed outgrowth. **(G)** Outgrowths were positive for DiI signal, indicating that the outgrowths contained donor tissue. **(B-G)** Dorsal view, anterior is to the left. Scale bars, 200 μm. **(H)** Triple FISH for *foxD/notum/ndl-2*, *sFRP-1*, and *slit/laminB/opsin* at 14 days following transplantation. Single-channel images are displayed. For merged images see Figure 1B. **(I)** Triple FISH for *nlg8*, *slit*, and *laminB* at 14 days following transplantation. Single-channel images are displayed. Pole transplant outgrowths have entirely dorsal identity (n=4/4). **(H-I)** Maximal intensity projections shown. Dorsal view, anterior is to the left. Scale bars, 200 μm.

Figure S2

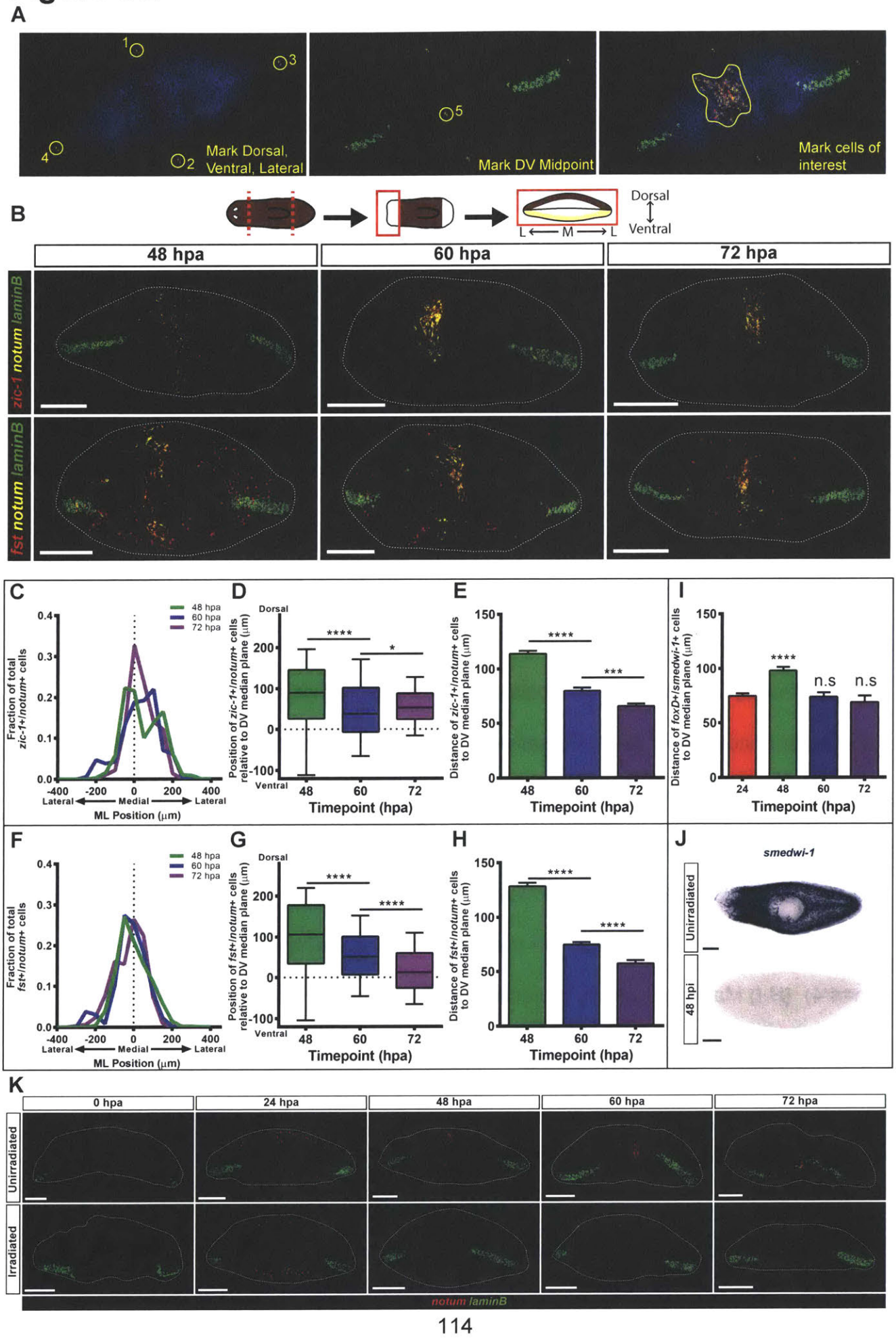

A
Mark Dorsal, Ventral, Lateral
Mark DV Midpoint
Mark cells of interest
B
Dorsal
Ventral
L M L
48 hpa
60 hpa
72 hpa
zic-1 notum laminB
fst notum laminB
C
48 hpa
60 hpa
72 hpa
Fraction of total zic-1+/notum+ cells
ML Position (μm)
Lateral
Medial
D
Dorsal
Ventral
Position of zic-1+/notum+ cells relative to DV median plane (μm)
Timepoint (hpa)
E
Distance of zic-1+/notum+ cells to DV median plane (μm)
Timepoint (hpa)
I
Distance of foxD+/smedwi-1+ cells to DV median plane (μm)
n.s
Timepoint (hpa)
F
Fraction of total fst+/notum+ cells
G
Position of fst+/notum+ cells relative to DV median plane (μm)
H
Distance of fst+/notum+ cells to DV median plane (μm)
J
smedwi-1
Unirradiated
48 hpi
K
0 hpa
24 hpa
48 hpa
60 hpa
72 hpa
Unirradiated
Irradiated
notum laminB

Figure S2. Anterior pole progenitors are specified medially and in a broad DV domain. Related to Figure 1 and Figure 2.

(A) During image quantification, multiple attributes of each image were marked in the following order. The locations of the dorsal-most point, ventral-most point, and the two lateral-most points were marked. The location of the estimated DV median plane was marked using *laminB* signal as a guide. The positions of the cells of interest were marked. **(B)** Triple FISH for *zic-1* or *follistatin*, and *notum* and *laminB* at 48, 60, and 72 hpa. *zic-1+/notum+* and *follistatin+/notum+* cells appear around 48 hpa at the ML median plane of the anterior-facing wound, and accumulate over time near the DV median plane. (**C)** The distribution of the positions of *zic-1+/notum+* cells on the ML axis is centered on the ML median plane of the fragment (dotted line). 48 hpa, n=16 worms; 60 hpa, n=15 worms; 72 hpa, n=14 worms. **(D)** The distribution of the positions of *zic-1+/notum+* cells on the DV axis becomes tighter over time. Data were compared by a Mann-Whitney test. 48 vs 60 hpa, ****p<0.0001, n>473 cells; 60 vs 72 hpa, *p=0.0283, n>364 cells. **(E)** The distance between *zic-1+/notum+* cells and the DV median plane grows smaller over time. Data were compared using a Student's t-test. 48 vs 60 hpa, ****p<0.0001, n>473 cells; 60 vs 72 hpa, ***p=0.0002, n>364 cells. **(F)** The distribution of the positions of *follistatin+/notum+* cells on the ML axis is centered roughly around the ML median plane of the fragment (dotted line). 48 hpa, n=16 worms; 60 hpa, n=16 worms; 72 hpa, n=12 worms. **(G)** The distribution of the positions of *follistatin+/notum+* cells on the DV axis becomes tighter and closer to the DV median plane over time. Data were compared with a Mann-Whitney test. 48 vs 60 hpa, ****p<0.0001, n>479 cells; 60 vs 72 hpa, ****p<0.0001, n>298 cells. **(H)** The distance between *follistatin+/notum+* cells

and the DV median plane grows smaller over time. Data were compared using a Student's t-test. 48 vs 60 hpa, ****p<0.0001, n>479 cells; 60 vs 72 hpa, ****p<0.0001, n>298 cells. **(I)** The distance between *foxD+/smedwi-1+* cells and the DV median plane at 60 and 72 hpa is comparable to the distance at 24 hpa. At 48 hpa this distance is increased. Data are represented as mean ± SEM and analyzed using a Student's t-test. 48 hour, 60 hour, and 72 hour timepoints were compared to 24 hpa. 48 hpa ****p<0.0001, n>311 cells, 60 hpa n.s., n>114 cells, 72 hpa n.s., n>75 cells. **(J)** Colorimetric *in situ* hybridization for the expression of the gene *smedwi-1*, which marks neoblasts. 48 hours post irradiation (hpi), animals were completely depleted of neoblasts. **(K)** Double FISH for *notum* and *laminB* at 0, 24, 48, 60, and 72 hpa following lethal irradiation (6000 rads). At 0 hpa, anterior pole expression of *notum* has been completely eliminated by transverse amputation. At 24 hpa, wound-induced *notum* expression still occurs in animals that have been lethally irradiated. From 48 to 72 hpa, anterior pole progenitors do not appear and no anterior pole is formed in animals which have been lethally irradiated. For FISH images in **(B)** and **(K)** *en face* view of the anterior blastema, dorsal is up. Scale bars, 200 μm. For colorimetric images in **(J)**, ventral view, anterior is to the left. Scale bars, 200 μm. For **(D)** and **(G)**, reference point is the DV median plane, which is estimated using *laminB* expression (dotted line). For **(E)** and **(H)** data are represented as mean ± SEM.

Figure S3

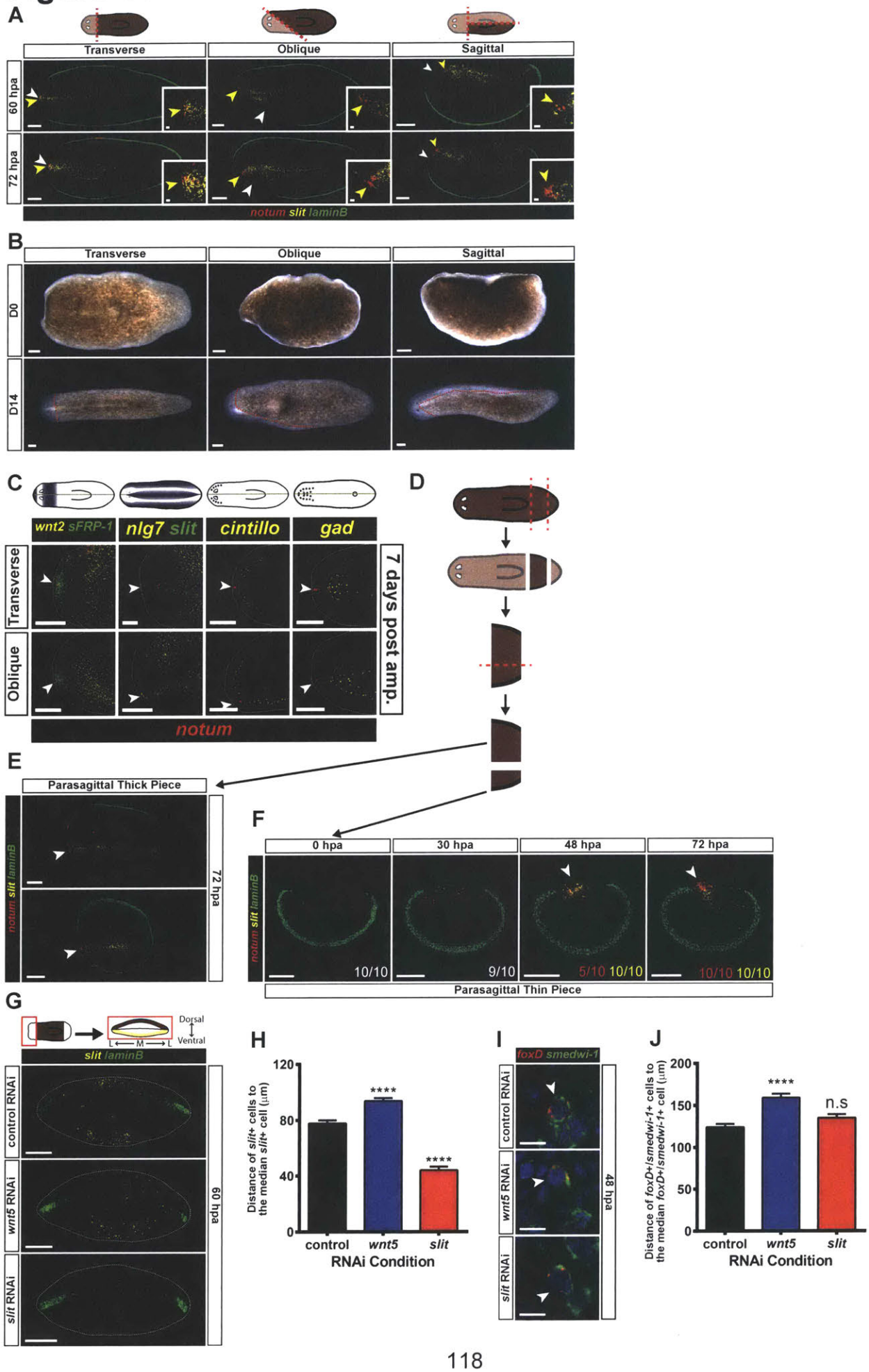

A
Transverse
Oblique
Sagittal
60 hpa
72 hpa
notum slit laminB
B
Transverse
Oblique
Sagittal
D0
D14
C
wnt2 sFRP-1
nlg7 slit
cintillo
gad
Transverse
Oblique
7 days post amp.
notum
D
E
Parasagittal Thick Piece
notum slit laminB
72 hpa
F
0 hpa
30 hpa
48 hpa
72 hpa
notum slit laminB
10/10
9/10
5/10 10/10
10/10 10/10
Parasagittal Thin Piece
G
Dorsal
Ventral
L ← M → L
slit laminB
control RNAi
wnt5 RNAi
slit RNAi
60 hpa
H
Distance of slit+ cells to the median slit+ cell (µm)
120
80
40
0
****
****
control
wnt5
slit
RNAi Condition
I
foxD smedwi-1
control RNAi
wnt5 RNAi
slit RNAi
48 hpa
J
Distance of foxD+/smedwi-1+ cells to the median foxD+/smedwi-1+ cell (µm)
200
150
100
50
0
****
n.s
control
wnt5
slit
RNAi Condition

Figure S3. Surgeries and *wnt5/slit* RNAi affect the ML axis. Related to Figure 3.

(A) Triple FISH for *notum*, *slit* and *laminB* at 60 and 72 hpa for transverse, oblique or sagittal amputations. *notum*+ cells appear near the *slit* domain in asymmetric wounds. White arrowheads, middle of anterior-facing wound; yellow arrowheads, middle of *slit* domain. Cartoon demonstrates surgery. Transverse 60 hpa, n=6/6; Oblique 60 hpa, n=11/11; Sagittal 60 hpa, n=10/10; Transverse 72 hpa, n=8/8; Oblique 72 hpa, n=12/12; Sagittal 72 hpa, n=20/20. **(B)** Live images at 0 and 14 days following transverse, oblique or sagittal amputation. Animals are capable of successfully regenerating from these amputations. **(C)** Triple FISH for *notum*, and *wnt2* and *sFRP-1*, *slit* and *nlg7*, *cintillo*, or *gad* at 7 dpa for transverse and oblique amputations. White arrowheads, anterior pole. For clarity, ventral images (*nlg7/slit* and *gad*) were vertically flipped. The plane of symmetry for gene expression was centered at the position of the anterior pole. **(D)** Cartoon depiction of the generation of thick and thin parasagittal pieces. **(E)** Triple FISH for *notum*, *slit* and *laminB* at 72 hpa for thick parasagittal pieces. White arrowheads, anterior pole. The anterior pole regenerates at the pre-existing midline. Two representative examples are shown. **(F)** Triple FISH for *notum*, *slit* and *laminB* at 0, 30, 48 and 72 hpa for thin parasagittal pieces. Immediately after the surgery, there is very little *slit* expression (10/10 animals). By 30 hpa, *slit* and wound-induced *notum* are expressed at the wound face (9/10 animals). At 48 hpa, some of the animals had begun to form a new anterior pole (5/10 animals), and all the animals had *slit* expression (10/10 animals). At 72 hpa all of the animals had formed a new anterior pole (10/10 animals), and all animals had *slit* expression extending into the fragment (10/10

animals). **(G)** Double FISH for *slit* and *laminB* at 60 hpa following inhibition of control gene, *wnt5* or *slit*. *slit*+ cells occupy a wider area in *wnt5(RNAi)*, and a narrow area in *slit(RNAi)*, compared to control RNAi animals. **(H)** The distance between *slit*+ cells and the median cell is larger in *wnt5(RNAi)* animals and smaller in *slit(RNAi)* animals, as compared to control RNAi animals. Data are represented as mean ± SEM and analyzed using a Student's t-test. control vs *wnt5*(RNAi), ****p<0.0001, n>621 cells; control vs *slit*(RNAi), ****p<0.0001, n>454 cells. **(I)** Double FISH for *foxD* and *smedwi-1* at 48 hpa following inhibition of control gene, *wnt5* or *slit*. Nuclear signal (DAPI) is shown in blue. *foxD*+/*smedwi-1*+ cells could be found and quantified. **(J)** The distance between *foxD/smedwi-1*+ cells and the median cell is larger in *wnt5(RNAi)* animals as compared to control RNAi animals. Data are represented as mean ± SEM and analyzed using a Student's t-test. control vs *wnt5*(RNAi), ****p<0.0001, n>510 cells; control vs *slit*(RNAi), n.s., n>496 cells. For **(A-C) and (E-F)** maximal intensity projections shown. Dorsal view, anterior is to the left. Scale bars, 200 μm. Insets (where present) are a zoomed in view of the anterior-facing wound. Scale bars, 20 μm. For **(G)** images shown are maximal intensity projections. Images are an *en face* view of the anterior blastema. Dorsal is up. Scale bars, 200 μm. For **(I)** images are a single confocal slice. Images are an *en face* view of the anterior blastema. Dorsal is up. Scale bars, 10 μm.

Figure S4

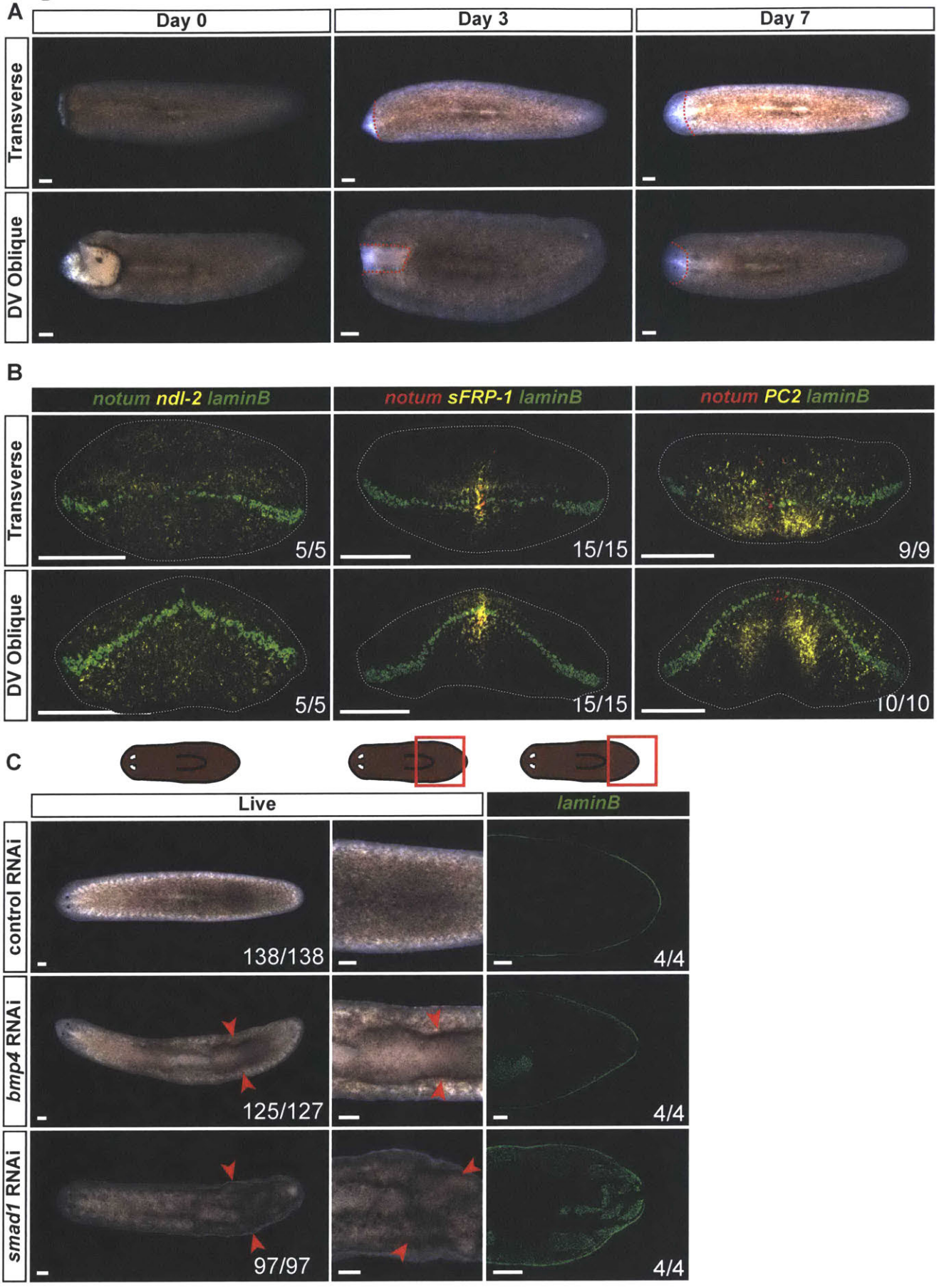

A
Day 0
Day 3
Day 7
Transverse
DV Oblique
B
notum ndl-2 laminB
notum sFRP-1 laminB
notum PC2 laminB
Transverse
DV Oblique
5/5
15/15
9/9
5/5
15/15
10/10
C
Live
laminB
control RNAi
bmp4 RNAi
smad1 RNAi
138/138
4/4
125/127
4/4
97/97
4/4

Figure S4. DV oblique cuts and Bmp pathway inhibition affect the DV axis. Related to Figure 4.

(A) Live images at 0, 3 and 7 days following transverse or DV oblique amputation. Animals are capable of restoring form after these injuries. Red dotted line, old tissue/blastema boundary. **(B)** Triple FISH for *notum* and *ndl2*, *notum* and *sFRP-1, or notum* and *PC2, and laminB* at 7 dpa for transverse and DV oblique amputations. Gene expression domains and tissue architecture appear to maintain their spatial relationship with the anterior pole. Number of animals, lower right. **(C)** Left columns: Live images at 28 days following inhibition of control gene, *bmp4* or *smad1. bmp4(RNAi)* and *smad1(RNAi)* animals display ridges on their dorsal side. Red arrowheads, dorsal ridges. Right Column: Single FISH for *laminB* at 28 days following inhibition of control gene, *bmp4* or *smad1.* Patches of ectopic *laminB* expression can be found on the dorsal sides of *bmp4(RNAi)* and *smad1(RNAi)* animals. Number of animals, lower right. For **(A)** and **(C),** dorsal view, anterior is to the left. For **(B),** *en face* view of the anterior blastema, dorsal is up. For all images, maximal intensity projections shown. Scale bars, 200 μm.

Tables

Table S1. List of genes used in this study. Related to Figures 1-4.

Gene Symbol	Reference	Pubmed ID	NCBI Accession Number	Ref. No.
bmp4	Molina 2007	17905225	EF633689	[30]
cintillo	Oviedo 2003	12557210	AY067542	[36]
follistatin	Roberts-Galbraith 2013	23297191	KC161222	[10]
foxD	Vogg 2014	24704339	KC577557	[4]
gad	Nishimura 2008	18440152	AB332029	[37]
laminB	Tazaki 2002	12203092	AY067086	[20]
mag-1	Zayas 2010	20865784	HM803280	[38]
madt	Wenemoser 2010	20599901	EG413862	[39]
ndl-2	Scimone 2016	27063937	KT983961	[40]
nlg7	Molina 2009	19174194	FJ471489	[41]
nlg8	Molina 2009	19174194	FJ471490	[41]
notum	Petersen 2011	21566195	JF725701	[8]
opsin	Sanchez Alvarado 1999	10220416	AF112361	[42]
PC2	Collins 2010	20967238	BK007043	[43]
rootletin	Glazer 2010	19852954	AY068190	[44]
sFRP-1	Petersen 2008	18063755	EU296635	[45]
slit	Cebria 2007	17553481	DQ336176	[26]
smad1	Molina 2007	17905225	EF633692	[30]
smedwi-1	Reddien 2005	16311336	DQ186985	[46]
wnt2	Petersen 2008	18063755	EU296634	[45]
wnt5	Adell 2009	19211673	FJ463749	[47]
zic-1	Vogg 2014	24704339	KF751216	[4]

Materials and Methods

Animals and radiation treatment

Asexual *Schmidtea mediterranea* strain (CIW4) were maintained in 1x Montjuic planarian water at 20°C as previously described [48]. Animals were starved 7–14 days prior experiments were used. Irradiated animals were exposed to a 6,000 rads dose of radiation using a dual Gammacell-40 ^{137}cesium source and amputated two days after irradiation. Animals were selected to be size-matched within an experiment, for all experiments.

Double-stranded RNA synthesis for RNAi experiments

Double stranded RNA (dsRNA) was synthesized as previously described [49]. Briefly, PCR templates of sequences for the forward and reverse of the target genes were prepared with a 5' flanking T7 promoter (TAATACGACTCACTATAGGG). The forward and reverse templates (16 μl) were mixed, in separate reactions, with 1.6 μl of 100 mM rNTPs (Promega); 0.6 μl of 1M dithiothreitol (DTT; Promega); 4 μl of T7 polymerase; and 24 μl of 5x Transcription optimized buffer (Promega). Reactions were incubated for 2 h at 37°C and then supplemented with RNase-free DNase for 45 minutes. RNA was purified by ethanol precipitation, and finally resuspended in 24 μl of milliQ H2O. RNA was analyzed on 1% agarose gel. RNA for forward and reverse strands were combined and annealed by heating the reactions in a thermo-cycler to 90°C and lowering gradually the temperature to 20°C.

RNAi

Animals were starved for at least 7 days prior to the first feeding. The food mixture was as follows: 26 μl of 100% homogenized beef liver, 12 μl of dsRNA, and 2 μl of red food coloring. Animals were fed twice a week.

Transplantation and Microsurgery

Transverse trunk fragments were generated by amputation beneath the auricles and immediately posterior to the pharynx. Oblique fragments were generated by a single diagonal cut made from the right side of the animal adjacent to the right eye to the left side of the animal below the pharynx. Sagittal fragments were made by amputating the head and performing a single sagittal cut. Transplantation of pole and flank pieces was as follows. Large (10-12 mm long) animals were selected for the procedure. Donors were lethally irradiated (6000 rads) 24 hours prior to transplantation and then were incubated in a 1:500 dilution of DiI (2 μg/mL) in planarian water overnight (Figure S1A). Thirty minutes prior to transplantation the DiI solution was removed and the donors were transferred to fresh plates. Hosts were anesthetized in 0.2% chloretone diluted in planarian water for 5 minutes, and briefly rinsed in Holtfreter's before being place on a filter paper on ice moistened with Holtfreter's. Host heads were amputated to prevent the hosts from moving following the transplantation procedure, as has been previously described [14]. A glass capillary tube of 0.75 mm interior diameter and 0.7 mm exterior diameter (FHC, Bowdoin, ME, USA) were used to remove host tissue from the future transplantation site posterior to the pharynx (Figure S1D) [50]. Donors were placed on a filter paper on top of a petri-dish lid over ice moistened with Holtfreter's and either a pole

fragment or flank fragment was excised (Figure S1B). Because donors were labeled with DiI, the donor fragments were easily visualized under a fluorescent microscope (Figure S1C). The pole or flank fragment was moved with a pipette from the donor to the host, and a pair of surgical scalpels were used to force the pole or flank fragment into the host (Figure S1E). The host was then placed in a recovery chamber as described previously [50]. Twenty-four hours after transplantation, hosts were moved from the recovery chamber into a petri dish with planaria water. Live animals were examined under a fluorescent microscope to look for DiI signal (Figure S1G).

Fixations

Animals were fixed and labeled as previously described [51]. In brief, animals were killed in 5% N-acetyl-cysteine in PBS for 5 minutes and then fixed in 4% formaldehyde for 20 minutes at room temperature. Fixative was removed and worms were rinsed 1X with PBSTx (PBS + 0.3% Triton X-100). PBSTx was replaced with preheated Reduction solution for 10 minutes at 37°C. Animals were dehydrated by a methanol series and stored in methanol at -20°C.

Whole-mount *in situ* hybridizations

RNA probes were synthesized and nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) colorimetric whole-mount in situ hybridizations (ISH) were performed as described [52]. Fluorescence in situ hybridizations (FISH) were performed as described [51] with minor modifications. Briefly, animals were killed in 5% NAC and

treated with proteinase K (4 μg/ml). Following overnight hybridizations, samples were washed twice in pre-hyb buffer, 1:1 pre-hyb-2X SSC, 2X SSC, 0.2X SSC, PBSTx. Subsequently, blocking was performed in 0.5% Roche Western Blocking Reagent (10% solution, Roche) and 5% inactivated horse serum PBSTx solution when anti-DIG or anti-DNP antibodies were used, and in 1% Roche Western Blocking Reagent PBSTx solution when an anti-FITC antibody was used. Post-antibody binding washes and tyramide development were performed as described [51]. Peroxidase inactivation with 1% sodium azide was done for 90 minutes at RT. Most specimens were counterstained with DAPI overnight (Sigma, 1 μg/ml in PBSTx).

BrdU labeling and immunofluorescence

BrdU (Sigma) was administered by soaking fragments starting 1 hour following amputation for 2 hours in planaria H2O containing 25 mg/ml BrdU and 3% DMSO; animals were then chased in Instant Ocean dissolved at 5 g/L until they were fixed. For BrdU labeling, samples were pretreated by incubation in 2N HCl (+0.5% triton-X-100) for 45 minutes followed by brief neutralization in 0.1M sodium borate. BrdU was detected with mouse anti-BrdU (BD Biosciences) followed by goat anti-Mouse IgG HRP Conjugate (Life Technologies), blocked in 0.6% BSA, 1% Western Blocking Reagent (Roche), and 5mM thymidine in PBSTx(0.3%). The BrdU signal was developed as previously described [52] using the Cy5 tyramide.

***en face* mounting**

Following staining, animals were placed in a petri dish with PBSTx and the anterior blastemas were cut off using a sharp scalpel. The anterior blastemas were placed on a slide with the cut edge down. A coverslip was placed over the same and Vectashield was allowed to flow in from the side.

Microscopy and image analysis

Fluorescent images were taken with a Zeiss LSM700 Confocal Microscope. Light images were taken with a Zeiss Discovery Microscope. For *en face* images, the marking of cell positions was performed using Fiji/ImageJ. A text file containing the positions of all the cells, and 5 reference positions (Figure S2A), was exported from Fiji/ImageJ. Reference point 1 is the dorsal most point, reference point 2 is the ventral most point, reference points 3 and 4 are the right and left sides, respectively, and reference point 5 is the estimated DV boundary in the blastema. This file was imported into MATLAB, which was used to perform the following transformations:

Coordinate Transformation I: The mean of reference points 1 and 2 was subtracted from all coordinates. This moves the distribution so that rotation will be around this point.

Coordinate Rotation: Matrix multiplication was used to rotate the coordinates so that reference point 1 was directly vertical.

Coordinate Transformation II: The x-value mean of reference points 3 and 4 was used to subtract from the X coordinates, and the y-value of reference point 5 was used to subtract from the Y coordinates. This places the origin of the data at the middle between the right and left sides of the fragment, and at the DV boundary.

The transformed coordinates, as well as measurements of distances between the coordinates and various landmarks were exported into an Excel document for further analysis. The histogram function in Excel was used on the X coordinates to generate the histograms for Figures 1D, 2B, 3C-D, S2C, S2F. The Y coordinates were plotted to generate Figures 1E, 2C, 4C-E, 4G-H, S2D, and S2G. The absolute values of the Y coordinates were used to generate the distances from the DV boundary in Figures 1F, 2E, S2E, S2H, and S2I.

Graphs and statistical analysis

All graphs and statistical analysis were done using the Prism software package (GraphPad Inc., La Jolla, CA). Comparisons between the means of two populations were done by a Student's t-test. Comparisons between two distributions were done using a Mann-Whitney test. Comparisons between two rows in a contingency table were done using Fisher's exact test. Significance was defined as $p < 0.05$.

Acknowledgements

We thank the members of the Reddien Lab for comments and discussion. D.J.L is supported by the National Institute of General Medical Sciences (T32GM007753) and by the Paul and Daisy Soros Fellowship for New Americans. P.W.R. is an Investigator of the Howard Hughes Medical Institute and an associate member of the Broad Institute of Harvard and MIT. We acknowledge support from the NIH (R01GM080639).

References

1. Reddien, P.W., and Sánchez Alvarado, A. (2004). Fundamentals of Planarian Regeneration. Annu. Rev. Cell Dev. Biol. *20*, 725–57.

2. Scimone, M.L., Lapan, S.W., and Reddien, P.W. (2014). A *forkhead* Transcription Factor Is Wound-Induced at the Planarian Midline and Required for Anterior Pole Regeneration. PLoS Genet. *10*, e1003999.

3. Vásquez-Doorman, C., and Petersen, C.P. (2014). *zic-1* Expression in Planarian Neoblasts after Injury Controls Anterior Pole Regeneration. PLoS Genet. *10*, e1004452.

4. Vogg, M.C., Owlarn, S., Pérez Rico, Y.A., Xie, J., Suzuki, Y., Gentile, L., Wu, W., and Bartscherer, K. (2014). Stem cell-dependent formation of a functional anterior regeneration pole in planarians requires Zic and Forkhead transcription factors. Dev. Biol. *390*, 136–48.

5. Driever, W., and Nüsslein-Volhard, C. (1988). A Gradient of *bicoid* Protein in Drosophila Embryos. Cell. *54*, 83–93.

6. De Robertis, E.M., Larraín, J., Oelgeschläger, M., and Wessely, O. (2000). The establishment of Spemann's organizer and patterning of the vertebrate embryo. Nat. Rev. Genet. *1*, 171–81.

7. Wagner, D.E., Wang, I.E., and Reddien, P.W. (2011). Clonogenic Neoblasts Are Pluripotent Adult Stem Cells That Underlie Planarian Regeneration. Science. *332*, 811–6.

8. Petersen, C.P., and Reddien, P.W. (2011). Polarized *notum* activation at wounds inhibits Wnt function to promote planarian head regeneration. Science. *332*, 852–5.

9. Gaviño, M.A., Wenemoser, D., Wang, I.E., and Reddien, P.W. (2013). Tissue absence initiates regeneration through Follistatin-mediated inhibition of Activin signaling. eLife. *2*, e00247.

10. Roberts-Galbraith, R.H., and Newmark, P.A. (2013). Follistatin antagonizes Activin signaling and acts with Notum to direct planarian head regeneration. Proc. Natl. Acad. Sci. U. S. A. *110*, 1363–8.

11. Spemann, H., and Mangold, H. (1924). Uber Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. Arch. Fur Mikroskopische Anat. Und Entwicklungsmechanik. *100*, 599–638.

12. Dubois, F. Contribution a l'etude de la migration des cellules de regeneration chez les Planaires dulcicoles. Bulletin Biologique de la France et de le Belgique, 1949.

13. Sugino, H. (1937). Transplantation experiments in *Planaria gonocephala*. Jour. Zool. *7*, 373–439.

14. Santos, F. V. (1929). Studies on transplantation in planaria. Biol. Bull. *57*, 188–97.

15. Santos, F. V. (1931). Studies on transplantation in planaria. Physiol. Zool. *4*, 111–64.

16. Chandebois, R. (1985). Intercalary regeneration and level interactions in the fresh-water planarian Dugesia lugubris. Roux's Arch. Biol. *53*, 390–6.

17. Kato, K., Orii, H., Watanabe, K., and Agata, K. (1999). The role of dorsoventral interaction in the onset of planarian regeneration. Development. *126*, 1031–40.

18. Scimone, M.L., Kravarik, K.M., Lapan, S.W., and Reddien, P.W. (2014). Neoblast specialization in regeneration of the planarian *Schmidtea mediterranea*. Stem Cell Reports. *3*, 339–52.

19. Hill, E.M., and Petersen, C.P. (2015). Wnt/Notum spatial feedback inhibition controls neoblast differentiation to regulate reversible growth of the planarian brain. Development. *142*, 4217–29.

20. Tazaki, A., Kato, K., Orii, H., Agata, K., and Watanabe, K. (2002). The body margin of the planarian *Dugesia japonica*: characterization by the expression of an intermediate filament gene. Dev. Genes Evol. *212*, 365–73.

21. Eisenhoffer, G.T., Kang, H., and Sánchez Alvarado, A. (2008). Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea*. Cell Stem Cell. *3*, 327–39.

22. van Wolfswinkel, J.C., Wagner, D.E., and Reddien, P.W. (2014). Single-Cell Analysis Reveals Functionally Distinct Classes within the Planarian Stem Cell Compartment. Cell Stem Cell. *15*, 1–14.

23. Lapan, S.W., and Reddien, P.W. (2011). *dlx* and *sp6-9* Control Optic Cup Regeneration in a Prototypic Eye. PLoS Genet. *7*, e1002226.

24. Lapan, S.W., and Reddien, P.W. (2012). Transcriptome Analysis of the Planarian Eye Identifies *ovo* as a Specific Regulator of Eye Regeneration. Cell Rep. *2*, 1–

14.

25. Newmark, P.A., and Sánchez Alvarado, A. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. Dev. Biol. *220*, 142–53.

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

27. Morgan, T.H. (1900). Regeneration in Planarians. Arch. Fur Entwickelungsmechanik Der Org. *10*, 58–119.

28. Gurley, K.A., Elliott, S.A., Simakov, O., Schmidt, H.A., Holstein, T.W., and Sánchez Alvarado, A. (2010). Expression of secreted Wnt pathway components reveals unexpected complexity of the planarian amputation response. Dev. Biol. *347*, 24–39.

29. Reddien, P.W., Bermange, A.L., Kicza, A.M., and Sánchez Alvarado, A. (2007). BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration. Development. *134*, 4043–51.

30. Molina, M.D., 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. Dev. Biol. *311*, 79–94.

31. Orii, H., and Watanabe, K. (2007). Bone morphogenetic protein is required for dorso-ventral patterning in the planarian *Dugesia japonica*. Dev. Growth Differ., 345–9.

32. Gaviño, M.A., and Reddien, P.W. (2011). A Bmp/Admp Regulatory Circuit Controls Maintenance and Regeneration of Dorsal-Ventral Polarity in Planarians. Curr. Biol. *21*, 294–9.

33. Diaz-Benjumea, F.J., and Cohen, S.M. (1993). Interaction between dorsal and ventral cells in the imaginal disc directs wing development in Drosophila. Cell. *75*, 741–52.

34. Diaz-Benjumea, F.J., and Cohen, S.M. (1995). Serrate signals through Notch to establish a Wingless-dependent organizer at the dorsal/ventral compartment boundary of the *Drosophila* wing. Development. *121*, 4215–25.

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

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

37. Nishimura, K., Kitamura, Y., Umesono, Y., Takeuchi, K., Takata, K., Taniguchi, T., and Agata, K. (2008). Identification of glutamic acid decarboxylase gene and distribution of GABAergic nervous system in the planarian *Dugesia japonica*. Neuroscience. *153*, 1103–14.

38. Zayas, R.M., Cebrià, F., Guo, T., Feng, J., and Newmark, P.A. (2010). The use of lectins as markers for differentiated secretory cells in planarians. Dev. Dyn. *239*, 2888–97.

39. Wenemoser, D., and Reddien, P.W. (2010). Planarian regeneration involves distinct stem cell responses to wounds and tissue absence. Dev. Biol. *344*, 979–91.

40. Scimone, M.L., Cote, L.E., Rogers, T., and Reddien, P.W. (2016). Two FGFRL-Wnt circuits organize the planarian anteroposterior axis. eLife. *5*, e12845.

41. Molina, M.D., Saló, E., and Cebrià, F. (2009). Expression pattern of the expanded noggin gene family in the planarian *Schmidtea mediterranea*. Gene Expr. Patterns. *9*, 246–53.

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

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

44. Glazer, A.M., Wilkinson, A.W., Backer, C.B., Lapan, S.W., Gutzman, J.H., Cheeseman, I.M., and Reddien, P.W. (2010). The Zn finger protein Iguana impacts Hedgehog signaling by promoting ciliogenesis. Dev. Biol. *337*, 148–56.

45. Petersen, C.P., and Reddien, P.W. (2008). *Smed-βcatenin-1* is required for anteroposterior blastema polarity in planarian regeneration. Science. *319*, 327–30.

46. Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sánchez Alvarado, A. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science. *310*, 1327–30.

47. Adell, T., Salò, E., Boutros, M., and Bartscherer, K. (2009). Smed-Evi/Wntless is required for β-catenin-dependent and -independent processes during planarian regeneration. Development. *136*, 905–10.

48. Sánchez Alvarado, A., Newmark, P.A., Robb, S.M.C., and Juste, R. (2002). The *Schmidtea mediterranea* database as a molecular resource for studying platyhelminthes, stem cells and regeneration. Development. *129*, 5659–65.

49. Rouhana, L., Weiss, J.A., Forsthoefel, D.J., Lee, H., King, R.S., Inoue, T., Shibata, N., Agata, K., and Newmark, P.A. (2013). RNA interference by feeding in vitro synthesized double-stranded RNA to planarians: methodology and dynamics. Dev. Dyn. *242*, 1–43.

50. Guedelhoefer, O.C., and Sánchez Alvarado, A. (2012). Amputation induces stem cell mobilization to sites of injury during planarian regeneration. Development. *139*, 3510–20.

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

52. Pearson, B.J., Eisenhoffer, G.T., Gurley, K.A., Rink, J.C., Miller, D.E., and Sánchez Alvarado, A. (2009). A Formaldehyde-based Whole-Mount In Situ Hybridization Method for Planarians. Dev. Dyn. *238*, 443–50.

Chapter 3

Planarian epidermal stem cells respond to positional cues to promote cell type diversity

Omri Wurtzel[1], Isaac M. Oderberg[1], Peter W. Reddien

[1]These authors contributed equally to this work.

Experiments shown in Figures 1-7 and Figures S1-S7 were performed by OW and IMO. All authors contributed to the design of experiments or editing of the manuscript.

Published as:

Wurtzel, O., Oderberg, I.M., and Reddien, P.W. (2017). Planarian Epidermal Stem Cells Respond to Positional Cues to Promote Cell-Type Diversity. Dev. Cell. *40*, 491–504.e5.

Abstract

Successful regeneration requires that progenitors of different lineages form the appropriate missing cell types. However, simply generating lineages is not enough. Cells produced by a particular lineage often have distinct functions depending on their position within the organism. How this occurs in regeneration is largely unexplored. In planarian regeneration, new cells arise from a proliferative cell population (neoblasts). We used the planarian epidermal lineage to study how the location of adult progenitor cells results in their acquisition of distinct functional identities. Single-cell RNA sequencing of epidermal progenitors revealed the emergence of distinct spatial identities as early in the lineage as the epidermal neoblasts, with further pre-patterning occurring in their post-mitotic migratory progeny. Establishment of dorsal-ventral epidermal identities and functions, in response to BMP signaling, required neoblasts. Our work identified positional signals that activate regionalized transcriptional programs in the stem cell population and subsequently promote cell type diversity in the epidermis.

Introduction

A major challenge of adult regeneration and tissue turnover is the production of region-appropriate cell types in the absence of embryonic patterning mechanisms [1]. Progenitors for regeneration, such as stem cells or dedifferentiated cells, must be regulated to choose which cell types to make, and these cell types must be appropriate for their location [2]. Furthermore, cells of the same lineage and cell type often have specialized functions depending on their location [3], which requires additional control over their differentiation [4,5]. Therefore, mechanisms governing lineage choice and the regional specialization of cell function are of central importance in regeneration. Here, we focus on the questions of how and when region-appropriate specialization occurs within a lineage.

Planarians are free-living flatworms that use adult stem cells to maintain tissues and to regenerate [6]. The only proliferating cell population in planarians, neoblasts, contain pluripotent stem cells [7]. Many neoblasts are specialized towards particular cell types including cells of the protonephridia [8], intestine [9], pharynx [10,11], nervous system [11,12], eye [13,14], and anterior pole [15]. The location of a neoblast [16] impacts its identity: For example, eye-specialized neoblasts are not found in the posterior of the animal [14] and intestinal neoblasts are often in proximity to the planarian gut [7]. Therefore, spatial information likely affects the identity of neoblasts and their progeny [16].

Conversely, it is unknown how the spatial distribution of neoblasts and progenitors within a lineage generates a diversity of cellular identities and functions. The

planarian epidermis presents an ideal system for studying this question: First, multiple cellular identities with specialized functions are found in the epidermis in specific body locations [17,18], and these cells appear to emerge from a single specialized neoblast lineage (ζneoblasts; Fig 1A) [19]. Second, there are well-established assays for evaluating planarian epidermal integrity and function [19–21], and the status of the lineage from neoblasts [19] all the way to mature cells [20]. Finally, the epidermal lineage has well-characterized differentiation stages (Fig 1A) that are both spatially and temporally distinct [19,20,22].

The epidermal lineage is derived from mitotic ζneoblasts that are distributed throughout the mesenchyme [19]. A ζneoblast exits the cell-cycle and progresses through a series of defined stages during its differentiation (Fig 1A) [20,22], a process that takes at least seven days [19]. During differentiation, cells migrate from the mesenchyme outwards to the epidermis, and ultimately integrate into the mature epidermis, a single-layered epithelial sheet [20]. In this process progenitors display a stereotyped sequence of gene expression changes, collectively described as maturation (Fig 1A) [20]. This apparently homogeneous population of progenitors gives rise to spatially distinct mature epidermal cell types [17].

We devised a strategy to determine at which stage in the epidermal lineage cells express genes associated with distinct spatial identities. We developed a method for isolating dorsal and ventral epidermis that facilitated the detection of eight spatial mature epidermal identities. We asked whether these identities emerge in the differentiated cells, their immediate progenitors, or even as early as in the spatially and temporally distant ζneoblasts that will produce these cells. The emergence of distinct

expression patterns was analyzed by single-cell RNA sequencing (SCS) of 303 cells spanning every step of epidermal maturation, combined with in situ hybridization (ISH) of over 125 epidermal genes. We found that epidermal neoblasts (ζneoblasts) and progenitors from all stages of epidermal maturation express genes according to their location within the animal. Analysis of ζneoblasts and progenitors across the dorsal-ventral (DV) axis revealed divergent transcriptional programs that correlated with the emergence of distinct dorsal or ventral epidermal identities. Inhibition of *bmp4*, a dorsalizing factor [23–25] that is constitutively expressed in dorsal muscle cells [26], resulted in the rapid dorsal emergence of ventral ζneoblasts and their progeny. This lineage ventralization, following *bmp4* inhibition, did not occur in the absence of ζneoblasts even when their progeny were present. This indicates that ζneoblasts respond to positional signals in their environment and activate region-appropriate transcriptional programs. These findings demonstrate that single-cell RNA sequencing, from distinct body regions, is a powerful method for revealing the spatial identity of progenitors for regeneration and tissue maintenance. Our results demonstrate that distinct regional identities are detectable within the neoblast population, and that signals from differentiated cells can be read by neoblasts to promote spatial cell-type diversity.

Results

Epidermal genes are expressed in at least eight different spatial patterns

Multiple cellular identities, distinguished by distinct gene expression domains, make the planarian epidermis. Our goal was to find how and when, during the formation of new epidermal cells, distinct epidermal identities emerge. Previous work identified three epidermal identities: ciliated epidermis [17], non-ciliated epidermis, and dorsal-ventral (DV) boundary epidermis [18]. However, other epidermal identities might exist. Because our work required a broad classification of epidermal identities, we first characterized mature epidermal identities through analysis of epidermal gene expression. We implemented an epidermal-enrichment strategy based on ammonium thiocyanate treatment [27], which allowed for the collection of the epidermis (Fig S1A; Methods). Isolated epidermis was used for RNA extractions from six biological replicates from the dorsal or ventral epidermal surfaces, separately (Fig S1A-B; Methods), which was followed by preparation of RNA sequencing libraries (Methods). Comparison of the epidermis-enriched RNAseq libraries with libraries prepared from RNA extracted from whole worms identified 3,315 genes that were overexpressed in the epidermis (Fig 1C-E; fold-change ≥ 2; FDR < 1E-4; Methods), including 393 and 233 genes enriched in either the dorsal or ventral epidermis, respectively (Fig S1D; Table S1).

We characterized the diverse patterns of epidermal gene expression by selecting 125 epidermis-enriched genes (Fig S1D-E; Table S1) and performing whole-mount in situ hybridizations with RNA probes (WISH). WISH analysis revealed that 90% (113/125) of the genes were in fact enriched in the epidermis (Fig 1B, S1F; Table S1), which we then classified to eight distinct expression patterns (Fig 1B-C; Fig S1F-G;

Table S1), representing diverse mature epidermal identities (Fig 1C). We subsequently used representatives of these patterns to study when, during epidermal differentiation, spatial gene expression domains emerge.

Reconstructing the epidermal lineage by single-cell sequencing

The expression of epidermal genes can emerge before differentiation is complete [20]. For example, *vim-1* was expressed in epidermal cells that were not yet integrated into the epidermis (Fig S1G), in contrast to PRSS12 and *laminB* (Fig S1G), which were expressed only in the mature epidermis. However, it is unclear how early in differentiation mature epidermal gene expression can initiate. To determine the stage of epidermal differentiation at which epidermal identities emerge, we characterized the transcriptomes of epidermal progenitors with single cell resolution. We dissected different body regions, including dorsal, ventral, or lateral tissues (Fig 1D; Methods) and, following maceration of the fragments, we isolated cells by fluorescence-activated cell sorting (FACS; Methods). Then, we selected cells for sequencing by quantitative PCR (qPCR) for the expression of an epidermal progenitor marker, *agat-1*, which we found to be expressed in all post-mitotic epidermal progenitors both by SCS [28] and FISH (Fig S2A-B; Supplementary Methods). In addition, we isolated lateral tissues and selected cells expressing an epidermal boundary marker (*laminB*; Fig 1B, S1G; [18]). In total, we collected 205 epidermal progenitors (*agat-1*+) and 6 epidermal boundary cells (*laminB*+). We used these cells for SCS [29], and the data that was generated was analyzed in combination with published epidermal lineage SCS data (98 cells; [28]). SCS gene expression data was visualized by t-distributed stochastic neighbor

embedding (tSNE) [30], and the data was clustered using the Seurat package [31] (Fig 1E; Table S2; Methods). The clustering analysis recapitulated known epidermal lineage cell states [20], here defined as: "ζNeoblast", "Early Stage", "Late Stage", "Later Stage", and 3 mature cell types (Fig 1E-F; S2C-F). Gene expression comparison between clusters identified 1,452 genes enriched in particular clusters (Table S3; FDR < 0.001, fold-change > 4; Methods). We selected highly specific gene markers to each maturation state based on this analysis (AUC range 0.81-1; Fig 1F-G; S2C-D), and used them for FISH (Fig 1H). These data represent comprehensive gene expression profiles of every stage of planarian epidermal differentiation.

Many mature epidermal genes are expressed in Early Stage progenitors

The SCS data we collected spanned every stage of epidermal maturation, which allowed determination of whether genes expressed in the mature epidermis (Table S1), were also expressed in progenitors and at what stage. We found that 24% of these genes (Fig 2A) were expressed in at least 30% of the Early Stage or Late Stage progenitors (Table S1; Supplementary Methods). Interestingly, some of these genes encode proteins that are associated with mature epidermal functions, such as ciliogenesis (Fig 2B, S3A). This raised the possibility that progenitor populations display distinct epidermal identities despite having at least four days to complete their maturation [19]. We tested these results by FISH on whole animals (Fig 2B-D; S3A-C) or by FISH on FACS-sorted cells (cell FISH; Fig S3D; Methods). Expression of tested genes encoding cilia components (*e.g.* BBS1 and *ift88-like*) was detectable in Early Stage progenitors and mature epidermis (Fig 2B-D, S3A-D), but not in dividing

neoblasts (Fig S3E). SCS, however, predicted that 23% of the centriole-associated genes [32] were also expressed in neoblasts (Table S4). This observation is consistent with the recent report by Duncan et al. on the expression of cilia components in neoblasts [33]. Furthermore, transcripts for 75% (20/27) of centriole-associated genes [32], which are expressed in ciliated cells [28,32] were detectable in Early Stage or Late Stage progenitors (Table S4), despite the fact that cilia were restricted to the mature epidermis (Fig S3F). Therefore, cilia appeared to be only assembled in mature epidermal cells despite transcription of genes encoding cilia components in progenitors. The correlation of co-expression of multiple structural cilia genes in single cells was much higher in Later Stage progenitors compared to Early Stage or Late Stage progenitors (Average Pearson correlation r=0.08, 0.1, 0.43 for Early Stage, Late Stage, and Later Stage, respectively; Supplementary Methods). However, expression of cilia-encoding genes was unlikely non-specific even in Early Stage progenitors, because expression of genes encoding cilia components was completely undetectable in non-ciliated cell types, such as muscle or gut (Methods; [28]). These results suggest that epidermal cells can adopt a distinct functional identity (becoming ciliated or not) early during differentiation, rather than acquiring an identity only after integration into the mature epidermis.

RNAi identifies transcription factors required for spatially-restricted epidermal identity

To identify transcription factors (TFs) that are required for the emergence of distinct spatial epidermal identities, we searched for TFs that were expressed specifically in the

DV epidermal boundary cells using our SCS data (Supplementary Methods). We found eight such TFs (Fig 2E) and could validate the DV-boundary epidermis expression for seven of eight by FISH (Fig 2F, S3G). We inhibited their expression by RNAi (Fig 2G; Methods). Then, we performed FISH on regenerating heads or tails for a DV epidermal marker (*laminB*). RNAi of three genes, *Post-2a*, *Post-2b*, and BARHL1, resulted in a striking reduction of epidermal DV boundary expression (*laminB*; Fig 2G), and in addition BARHL1 RNAi also resulted in lesions around the pharyngeal cavity (Fig S3H), a region in which BARHL1 is expressed (Fig S3I). In contrast, these three TFs did not affect other mature epidermal cell types, as tested by FISH, indicating their specific requirement for activation of DV-boundary epidermis gene expression (Fig S3J). These results demonstrate that certain TFs specifically drive gene expression of spatially restricted epidermal subpopulations.

Spatially-restricted epidermal identities emerge in progenitors

Mature epidermal identities have distinct spatial distributions across the different body axes (Fig 1B-C). These distinct patterns of gene expression might (1) emerge in progenitors, or (2) might only be present after cells have been incorporated into the mature epidermis. We selected epidermal TFs, which are expressed in a spatially restricted manner in the mature epidermis (Fig 1B-C, 2F, S3G), and analyzed their expression in epidermal progenitors. In the mature epidermis, *ovo-2* (Fig 1B-C, Fig 3A-B, S4A-B) was expressed in the dorsal midline and the lateral edges. Imaging a plane of internal tissues (Fig 3A-B, Fig S4B; Methods) showed spatially restricted expression in immature epidermal progenitors (*agat-3*+) that was highly reminiscent of the pattern of

ovo-2 expression in the mature epidermis (Fig 1B-C, 3A). Inhibition of *wnt5*, which causes expansion of the planarian midline [34,35], resulted in expansion of the *ovo-2* expression domain (Student's unpaired t-test $p < 0.02$; Fig S4C). In a similar fashion to *ovo-2*, *tlx-1*, which in the epidermis is specifically expressed in the DV boundary (Fig 2E-F, S3G), was expressed in epidermal progenitors (*agat-3*+) only next to the DV boundary and not in epidermal progenitors further away (Fig 3C-D). Furthermore, *foxJ1-4,* which is expressed in ciliated epidermis (Fig S3K), was abundantly expressed in *agat-3*+ cells near ciliated epidermal regions, such as on the ventral surface (Fig 3E), but not in the vicinity of non-ciliated areas, such as the dorsal midline region (Fig 3E). Finally, we examined the spatially restricted expression of the gene DCLK2, which encodes a doublecortin protein kinase rather than a TF, and which is expressed in the ventral epidermis (Fig 1B-C). DCLK2 was expressed in multiple ventral epidermal progenitors, but not in any dorsal progenitors (Fig 3F; Methods). By contrast, the expression of DYRK4 (Fig 1B-C) was not detectable in epidermal progenitors (Fig S4D), despite having a spatially restricted expression pattern in mature cells. These results demonstrate that gene expression in immature epidermal progenitors is pre-patterned in a manner that reflects some of the spatial identities that are found in the mature epidermis. This suggests that migratory epidermal progenitors detect their position in the animal to activate regionally-appropriate transcriptional programs that will reflect the pattern of gene expression in the mature epidermis.

Dorsal and ventral gene expression distinguishes Early Stage post-mitotic progenitors

The results above demonstrated that some genes are expressed in epidermal progenitors based on their location in the animal (Fig 3, S4B). Because the dorsal and ventral epidermis display strikingly different expression patterns (Fig 1B-C), we sought to elucidate how early during maturation dorsal and ventral identities emerge. Approaches for identifying genes that are expressed in a spatially restricted pattern within a lineage are challenging because they require comparing cells from a similar maturation stage, but from different locations. The single cells we sequenced were isolated from physically separated dorsal or ventral regions in the animal (Fig 4A), and through SCS clustering (Fig 1E-F; Methods), they were assigned to a maturation stage (Fig 1E, 4A; Table S2). Therefore, this approach allows for the systematic identification of genes with spatially restricted expression in epidermal progenitors. To find the earliest differences in gene expression across epidermal populations we compared gene expression of Early Stage progenitors from dorsal and ventral regions (Fig 4A). In total, we found 23 genes that were significantly enriched (Fig 4B; Fold-change > 4; FDR < 0.1; Power > 0.4; Table S5) in dorsal or ventral progenitors. This finding was corroborated by the observation that 19 of these 23 genes were biased in bulk RNA sequencing from dorsal or ventral animal fragments (Fig 4B). Remarkably, the expression of two genes strongly predicted whether a cell is dorsal (*PRDM1-1*; AUC = 0.85; Fig S5A; Table S5) or ventral (*kal1*; AUC = 0.9; Fig S5B; Table S5), a finding that is consistent with the whole-mount expression patterns for these genes (Fig 4C). FISH analysis validated that *prog-2+/PRDM1-1+* cells were restricted dorsally and that *prog-2+/kal1+* were exclusively ventral (Fig 4D). Importantly, whereas *PRDM1-1* expression appeared exclusively in dorsal epidermal progenitors, *kal1* was expressed in other cell

types in ventral tissues, including neural and muscle cells (Fig S5C-D), suggesting it is a broader ventral marker. The expression of *PRDM1-1* and *kal1* strongly correlated with other DV-biased identities: the majority of cells expressing *ovo-2* (log_2(CPM) > 5), which is expressed only dorsally (Fig 1B-C, 3A-B), also expressed *PRDM1-1* in SCS data (Fig S5E; overlap = 67%; percentile rank of fraction overlap = 0.98; Supplementary Methods), but none expressed *kal1* (Fig S5E). Similarly, the expression of DCLK2 overlapped with the expression of *kal1*(Fig S5F; overlap = 35%; percentile rank of overlap = 0.84), a gene that is expressed specifically in the ventral epidermis (Fig 1B-C, 3F), but not with the expression of *PRDM1-1* (0%). These results indicate that the DV-bias observed in Early Stage gene expression is correlated with subsequent spatially-restricted epidermal identities.

Epidermal neoblasts express dorsal and ventral markers

Epidermal progenitors are the product of mitotic ζneoblasts [19]. The expression of DV-biased genes in epidermal progenitors raised the question whether DV identities exist as early in the lineage as in ζneoblasts. We tested this hypothesis by analyzing the expression of DV-biased epidermal progenitor genes in neoblasts. *PRDM1-1* and *kal1* expression was highly specific to ζneoblasts in SCS data (Fig S6A), and was detectable in 68% of the cells (Fig 5A). In addition, we used cell FISH to quantify the fraction of dividing neoblasts expressing *PRDM1-1* or *kal1*, and found that 6.7% and 9.7% of the neoblasts expressed *PRDM1-1* or *kal1*, respectively (Fig 5B). Similarly, we quantified the proportion of ζneoblasts in the dividing neoblast pool by cell FISH and found that they comprised 31% of dividing neoblasts (Fig S6B). Therefore, we estimated that

~53% of ζneoblasts express either *PRDM1-1* or *kal1* using cell FISH. Strikingly, the expression of *PRDM1-1* and *kal1* in SCS was mutually exclusive in ζneoblasts (Fig 5A), a finding we corroborated with cell FISH on dividing neoblasts (Fig S6C; number of cells counted = 1061).

We examined the spatial distribution of *PRDM1-1*+ ζneoblasts by FISH, and detected them only on the dorsal side (Fig 5C). Conversely, *kal1*+ ζneoblasts were only found ventrally (Fig 5C, S6D). We quantified the positions of *kal1*+ neoblasts along the DV axis by imaging transverse sections (Fig 5D, S6E), which were labeled for major anatomical structures (intestine, *madt* [36]; and DV-boundary epidermis, *laminB*). *kal1*+ cells were always ventral to the intestine or next to the DV-boundary epidermis (Fig 5D-E). We divided each transverse section into 10 regions along the DV axis and counted the number of *kal1*+ neoblasts in each region (Fig 5E; Supplementary Methods). The majority (79%) of the *kal1*+ neoblasts were found in the two regions above the ventral epidermis (Fig 5E), and notably *kal1*+ neoblasts were not found in any of the dorsal regions. Utilizing ζneoblast markers, we found that *kal1*+ ζneoblasts were closer to the ventral epidermis, and that more internal ζneoblasts did not express *kal1* (Fig 5F, S6F). These findings show that *kal1* expression in neoblasts is correlated with the position of the cells across the DV axis. To confirm these results with another method, we isolated groups of 500 dividing neoblasts by FACS from either dorsal or ventral planarian fragments. We performed qPCR analysis on eight dorsal and six ventral neoblast groups (Methods) and found that in dividing neoblasts, *PRDM1-1* was significantly overexpressed on the dorsal side, and by contrast, *kal1* was significantly overexpressed on the ventral side (Fig 5F). These findings demonstrate that gene expression in

neoblasts within a lineage reflects the position of those neoblasts in the animal. This positional identity is likely subsequently propagated to the spatially divergent epidermal progenitors, which ultimately generate regionally-appropriate mature epidermal cells.

The BMP signaling gradient affects the DV identity of ζneoblasts and epidermal progenitors

The dorsal and ventral identities identified in ζneoblasts raised the possibility that ζneoblasts respond to an extracellular signal to activate position-specific transcriptional programs. In planarians, many genes regulating adult patterning are expressed in muscle cells [26], but the connection between patterning gene expression in muscle and neoblast states is poorly understood. *bmp4* expression from dorsal muscle cells [26] regulates the polarization of the DV axis [23–25], and thus it is an attractive candidate for regulating the emergence of DV identities in ζneoblasts. Importantly, the expression of *kal1* was spatially opposed to *bmp4* expression (Fig 5D). *bmp4* inhibition leads to progressive ventralization of the animals [23–25], resulting in ventral features appearing dorsally. We assessed if BMP signaling actively represses the expression of ventral epidermal genes in dorsal epidermal lineage cells. We inhibited *bmp4* and examined *kal1* expression in two scenarios (Fig 6A; Methods). First, in intact animals, 10 days following *bmp4* inhibition, a time-point at which animals appeared indistinguishable from controls (Fig 6A) and; second, in regenerating animals, 10 days following amputation, and a total of 20 days after initiating *bmp4* inhibition. In this condition, the animals displayed major morphological defects reflecting ventralization of the regenerating tissue (Fig 6A). Following RNAi, *kal1*+ cells were found dorsally in *bmp4* RNAi animals

(Fig 6B), but never dorsally in controls (Fig 6B). Importantly, *kal1*+ Early Stage progenitors (*prog-2*+) were present dorsally in both intact and regenerating *bmp4* RNAi animals (Fig 6B). Furthermore, *kal1*+ ζneoblasts were detectable in *bmp4* RNAi animals dorsally (Fig 6C) 10 days from the beginning of the experiment, suggesting that proximity to a BMP source represses the ventral ζneoblast identity.

Neoblasts are required for responding to changes in BMP signaling

The outcome of *bmp4* RNAi on epidermal lineage cells at their various stages, ventralization, could result from a response only in neoblasts, which produce the lineage, or a response in neoblasts and post-mitotic cells of the lineage. To distinguish between these possibilities, we tested whether ventralization of the epidermal lineage following *bmp4* RNAi required neoblasts. Animals were split into two groups (Fig 6D), with the first group lethally irradiated on day zero and the second unirradiated. Lethal irradiation ablates all neoblasts by 24 hours post-irradiation but post-mitotic epidermal progenitors persist for up to 6 days as they transit towards differentiation [22]. At 24 hours post-irradiation, half of each group was injected with *bmp4* or control dsRNA. At four, six, or eight days following injection animals were fixed and analyzed for the expression of *kal1* by FISH (Fig 6E-I). Numerous ectopic *kal1*+ progenitors were present in unirradiated *bmp4* RNAi animals, at all time points examined (Fig 6E-F, S7A). Moreover, *kal1* expression in the mature dorsal epidermis was detected eight days following injection (Fig 6G). By contrast, we did not detect *kal1* expression in progenitors or mature epidermal cells on the dorsal side of animals that were irradiated prior to *bmp4* inhibition (Fig 6E-I). Next, we examined the expression of *rootletin*, which

is normally spatially restricted to dorsal epidermis and expressed ubiquitously in ventral epidermis (Fig S7B). Dorsal expression of *rootletin* was expanded and resembled the ventral side in 3/8 unirradiated *bmp4* injected animals (Fig 6G). This transformation did not occur in animals that were irradiated prior to the injection (100%; Fig 6G). Finally, we quantified the number of dorsal *kal1*+ cells, and confirmed our observations that ectopic *kal1* expression was only present in unirradiated animals following *bmp4* inhibition (Fig 6H-I). These findings suggest that within the epidermal lineage, it is primarily ζneoblasts that respond to BMP signaling. Finally, we examined the spatial distribution of the ectopic *kal1*+ cells following *bmp4* RNAi by imaging transverse sections of unirradiated *bmp4* and control RNAi animals (Fig 6J). In all affected animals, we found that ectopic *kal1*+ cells appeared only near the dorsal epidermis, and not in the interior of the animal (Fig 6J). Since we observed (Fig 5F, S6F) that *kal1*+ ζneoblasts are generally closer to the epidermal surface than *kal1*- ζneoblasts, this result indicates that *kal1* expression is dependent not only on DV position, but also on proximity to the periphery of the animal. Therefore, we suggest that ζneoblasts sense the BMP levels in their environment and obtain a DV spatial identity, promoting cell type diversity to form region-appropriate epidermis.

Discussion

The body plan of even relatively simple animals is constructed of diverse, spatially arranged cell types. Adult planarians can replace all cell types during tissue turnover or regeneration through neoblast differentiation and integration of their progeny into organs and tissues [6]. Therefore, the molecular signals that specify neoblasts and their

descendants towards the appropriate cell types are important for understanding tissue turnover and regeneration. Equally important is understanding how cells within a particular lineage acquire unique functions based on their position in the animal [3]. These functions are often essential for basic activities. For example, the production of spatially arranged ciliated epidermis is essential for normal planarian motility [17]. A major challenge in identifying regulators of this process is that cell population-scale analyses often utilize cells that are heterogeneous with regards to cell type, differentiation state, and location within the animal. Therefore, single-cell analyses are highly advantageous for studying where, when, and how specialization emerges within a cell population.

Expression of spatially restricted epidermal identities in the epidermal lineage

We used the planarian epidermis to study the location, timing, and mechanism of spatially-restricted epidermal identity acquisition during differentiation (Fig 7). We first characterized the different spatially-restricted identities that are found in the epidermis and then analyzed, using SCS and FISH, the transcriptomes of epidermal progenitors from throughout the lineage. We identified transcriptional programs that were dependent on the position of the cell within the animal (Fig 7A-B). Importantly, position-specific gene expression for different identities emerged at different stages of maturation. For example, the spatially-restricted expression of DYRK4 (Fig 1B-C, S4D) was detectable only in the mature epidermis. The expression of *ovo-2*, on the other hand, could be found both in the mature epidermis and in a pattern similar to its ultimate epidermal pattern in migratory, mesenchymal, Late Stage progenitors (Fig 1B-C, Fig 3E). The expression of genes encoding cilia components was observed by FISH in Early Stage

progenitors (Fig 2B-C, Table S4), and SCS suggested that some cilia components were already expressed in neoblasts (Table S4). In fact, Duncan et al. reported recently that some cilia components, such as *cfap53* and *rsph6A*, are expressed in neoblasts, suggesting that cilia specification might occur in mitotic cells [33]. Importantly, DV-restricted gene expression was found even at the top of the epidermal lineage, in ζneoblasts (Fig 7C), which have at least eight days and multiple stages though which to transit through before they have fully matured and migrated to their final destination [19].

The epidermis is a model for generation of regionally appropriate cell types

Planarian epidermal progenitors have been studied in multiple reports, with findings focusing on phases of epidermal maturation [19,20,22,37]. Less emphasis has been given to the distinct cellular identities individual cells can acquire, and the spatiotemporal regulation of these distinct identities.

The emergence of distinct gene expression domains in the planarian epidermis (e.g. *ovo-2* and DYRK4 at the dorsal epidermis; *laminB* in the DV boundary; DCLK2 at the ventral epidermis; Fig 1B-C) is very likely regulated by multiple mechanisms. Our analysis demonstrated that patterning molecules, such as BMP [23–25] and Wnt5 [34,35] were essential for the normal formation of some of these patterns. Interestingly, the major source of these factors is muscle cells at distinct regions of the animal [26]. We hypothesize that progenitors for regeneration and tissue-maintenance sense the levels of patterning factors at their location, and respond by activating a region-appropriate transcriptional program. This hypothesis is consistent with several observations made on regional gene expression in planarians: First, patterning

molecules are constitutively expressed in planarians [2]; second, RNAi of these molecules leads to patterning defects in intact and regenerating animals [2], and finally, with the rapid establishment of regionally-appropriate gene expression of patterning factors following injuries, prior to the generation of regionally-appropriate cell types [26,28,35,38]. Importantly, it has been unclear at what stage of differentiation cells respond to these signals, and in particular whether it is mature cell types, post-mitotic progenitors, and/or neoblasts that respond.

DV gene expression distinguishes populations of epidermal progenitors and ζneoblasts

To identify the cells that respond to positional cues, we first searched for the earliest discernable transcriptional differences between spatially-distinct epidermal lineage cells. Using SCS on samples obtained from dorsal or ventral regions, we identified 23 genes that were potentially expressed in a regionally-restricted manner at a very early stage of differentiation. Remarkably, the expression of two genes was sufficient to identify the DV-location of the vast majority of Early Progenitors (*PRDM1-1* for dorsal cells and *kal1* for ventral cells). Furthermore, these genes were expressed in spatially-divergent epidermal neoblast populations, as observed by FISH, SCS, and qPCR analysis. These results demonstrated that ζneoblasts are heterogeneous in a manner that is explained best by their position within the animal (Fig 7C), which suggests that ζneoblasts can respond to positional cues.

Neoblasts respond to BMP levels by acquiring a DV-positional identity

One positional cue, *bmp4*, is expressed from dorsal muscle cells [26] in a gradient that is strongest at the midline. It promotes acquisition of dorsal tissue identities, and its

inhibition causes progressive ventralization of the animal [23–25]. Strikingly, following a short *bmp4* RNAi treatment, expression of a ventral gene (*kal1*) appeared dorsally in epidermal neoblasts and their progeny. Regions in the dorsal epidermis, normally devoid of cilia, co-expressed *kal1* and *rootletin* in these *bmp4(RNAi)* animals, indicating that some ectopic *kal1*+ cells adopted a ventral functional identity. Following ablation of neoblasts by lethal irradiation, however, *bmp4* inhibition did not result in dorsal expression of *kal1* or in ectopic *rootletin* expression. Therefore, our data suggest that neoblasts are the only cells in the epidermal lineage capable of adopting ectopic ventral identities in response to BMP inhibition. Subsequently, these ventral identities were were then propagated to neoblast progeny. Furthermore, this suggests that mature epidermal identities are not plastic even in response to changes in the BMP signaling environment.

Since *bmp4* is expressed largely exclusively in dorsal muscle cells [26] we suggest that neoblasts read a BMP gradient established by dorsal muscle, and respond by activating a DV-appropriate transcriptional program, which subsequently generates the correct types of epidermis (Fig 7B-C). Planarian regeneration requires two components: collectively pluripotent neoblasts [16,39] that are the cellular source for all new tissues, and a set of positional signals [2] that are expressed from muscle cells [26], and provide the patterning information required for regeneration. Our results support a model in which neoblasts, within a lineage, respond directly to the muscle-derived positional signals to produce the required regionally-appropriate cell types. This model helps explain how planarians maintain a complex body plan during tissue turnover and in response to injury.

Figure 1

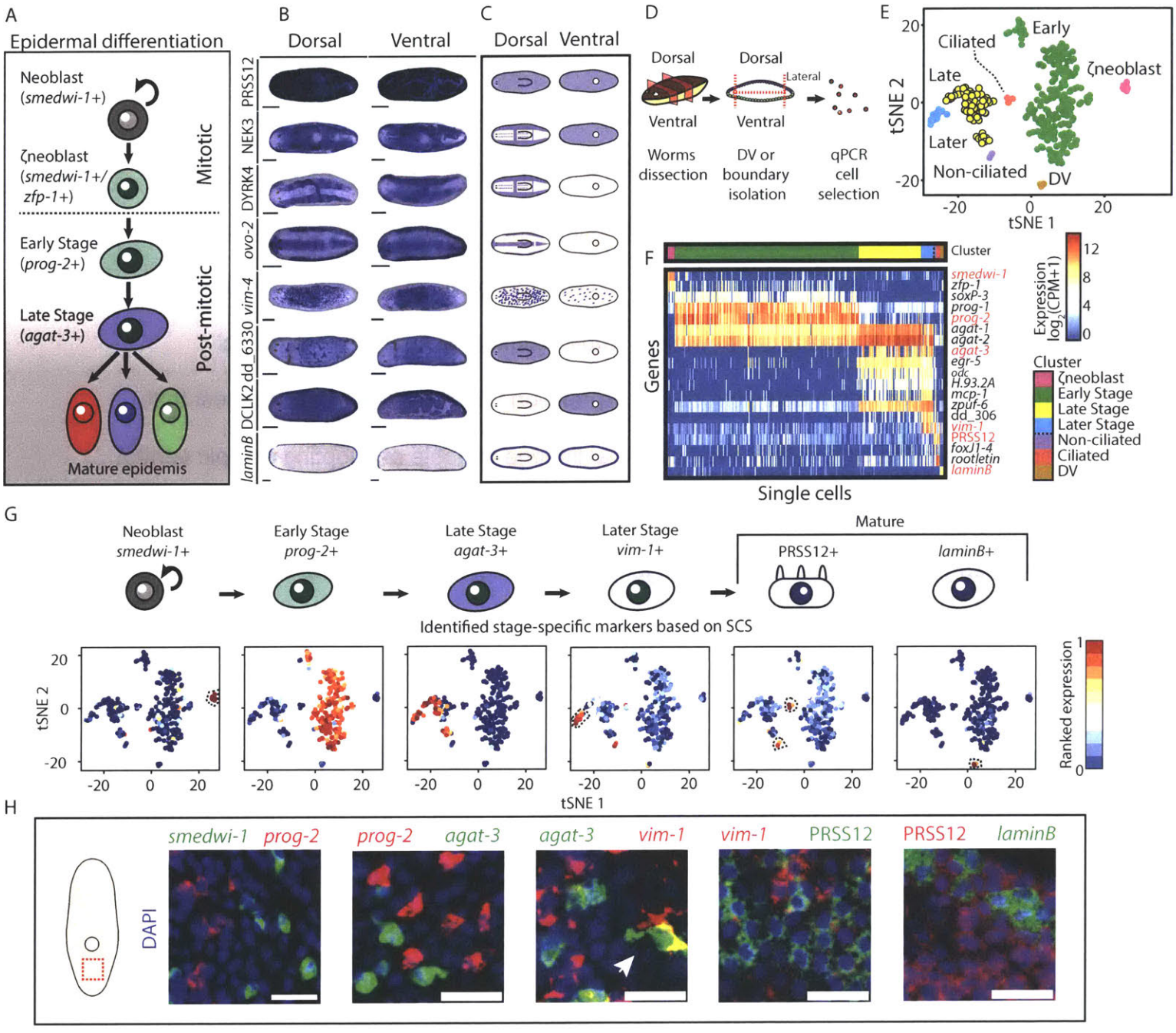

A
Epidermal differentiation
Neoblast
(smedwi-1+)
ζneoblast
(smedwi-1+/
zfp-1+)
Mitotic
Early Stage
(prog-2+)
Late Stage
(agat-3+)
Post-mitotic
Mature epidemis
B
Dorsal
Ventral
PRSS12
NEK3
DYRK4
ovo-2
vim-4
dd_6330
DCLK2
laminB
C
Dorsal
Ventral
D
Dorsal
Ventral
Dorsal
Lateral
Ventral
Worms dissection
DV or boundary isolation
qPCR cell selection
E
Early
Ciliated
ζneoblast
Late
Later
Non-ciliated
DV
tSNE 1
tSNE 2
F
Cluster
Genes
smedwi-1
zfp-1
soxP-3
prog-1
prog-2
agat-1
agat-2
agat-3
egr-5
odc
H.93.2A
mcp-1
zpuf-6
dd_306
vim-1
PRSS12
foxJ1-4
rootletin
laminB
Single cells
Expression log2(CPM+1)
Cluster
ζneoblast
Early Stage
Late Stage
Later Stage
Non-ciliated
Ciliated
DV
G
Neoblast
smedwi-1+
Early Stage
prog-2+
Late Stage
agat-3+
Later Stage
vim-1+
Mature
PRSS12+
laminB+
Identified stage-specific markers based on SCS
tSNE 2
tSNE 1
Ranked expression
H
DAPI
smedwi-1
prog-2
prog-2
agat-3
agat-3
vim-1
vim-1
PRSS12
PRSS12
laminB

Figure 1. Mature epidermal identities revealed by RNA-seq. (A) A model for planarian epidermal maturation. Epidermal progenitors have spatiotemporally distinct stages [20,22]. Stage-specific gene expression markers distinguish different maturation stages. Epidermal progenitors generate multiple epidermal cell types characterized by different spatial distribution and gene expression. (B) Different epidermal gene expression patterns were detected by WISH. Shown are eight representative expression patterns (blue) in dorsal (left column) and ventral (right column) views. As animals are semitransparent, following bleaching [40], assessing domains of expression requires examination from multiple viewpoints (scale bar = 100 μm). (C) Cartoon of the expression patterns reported in Fig 1B. Blue color represents region of epidermal expression. (D) Analysis of the epidermal lineage by SCS. Isolation of epidermal progenitors; pre-pharyngeal fragments were dissected to isolate dorsal, ventral, or lateral fragments, separately (Methods). Tissues were disassociated and individual cells were sorted to plates by FACS. *agat-1* expression [22] was measured by qPCR and cells expressing *agat-1* were sequenced and used for analysis, in combination with previously published SCS of the epidermal lineage [28]. (E) t-distributed stochastic neighbor embedding (tSNE) representation of epidermal cells (dots) is shown (Methods). Colors represent different clusters, with labels for cluster identities based on gene expression markers for epidermal maturation stages [20]. (F) Gene expression in single cells (columns) of previously described epidermal genes (rows). Gene expression is represented by color (blue to red, low to high expression; counts per million, CPM; Hierarchal clustering was used to order cells within a cluster). Red labels denote genes that were subsequently used for FISH experiments, because of their power to

discriminate between clusters, and their low background expression in FISH (Table S3). (G) Upper panel: Cluster-specific gene expression markers that were used for evaluating the distribution of epidermal lineage progenitors in situ. Markers were selected based on their fold-enrichment and power to discriminate between clusters. Lower panel: tSNE plot of cells (dots) colored by the ranked expression of the gene markers we selected across all sequenced cells (blue to red, low to high gene expression, respectively; dotted black line highlights cell clusters). (H) The selected gene expression markers were assessed by co-expression FISH analysis (scale bar = 20 μm). The markers had a nearly complete lack of co-expression, therefore permitting the analysis of distinct maturation stages.

Figure 2

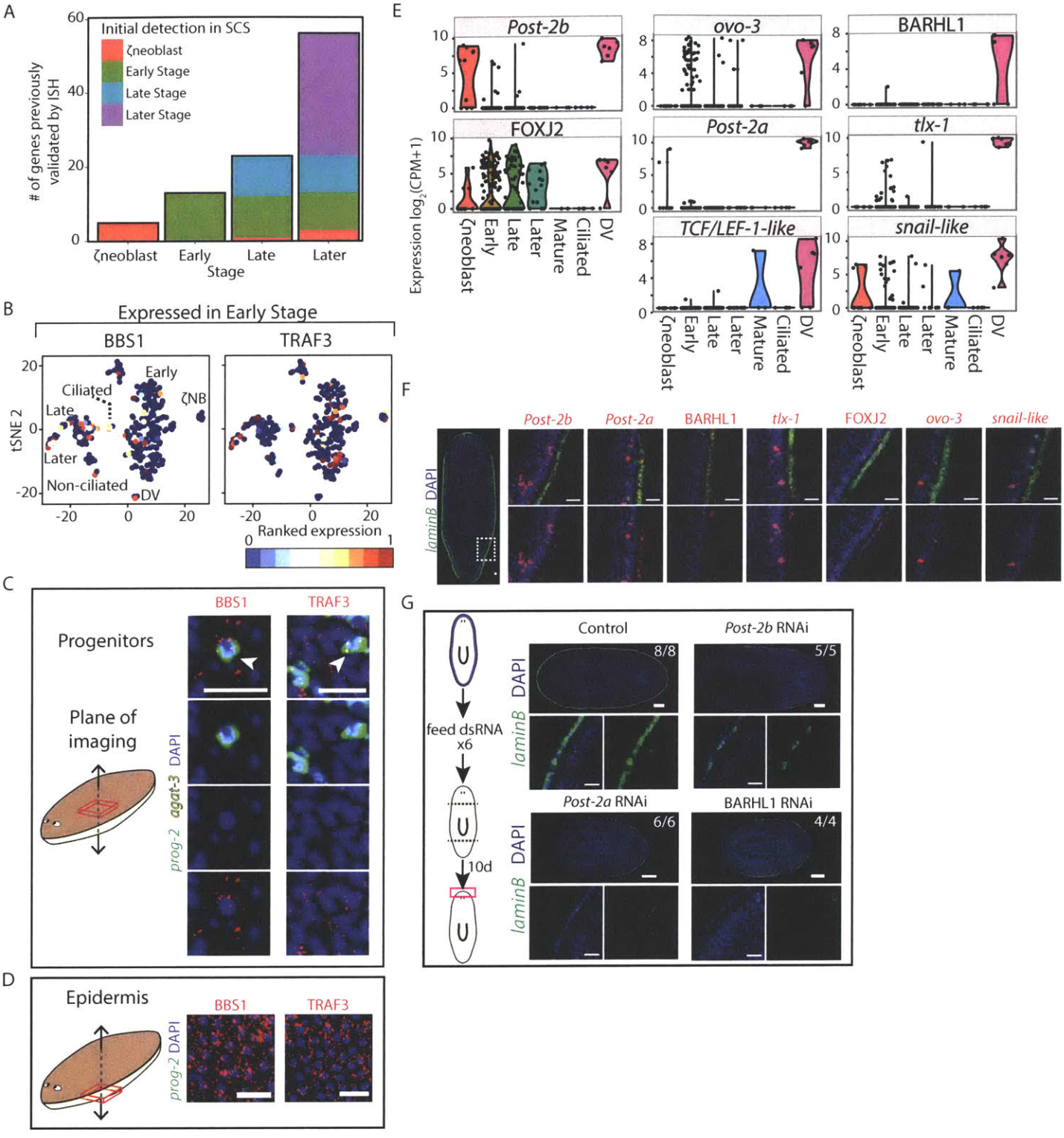

A
Initial detection in SCS
ζneoblast
Early Stage
Late Stage
Later Stage
# of genes previously validated by ISH
ζneoblast
Early
Late
Later
Stage
B
Expressed in Early Stage
BBS1
TRAF3
Early
Ciliated
ζNB
Late
Later
Non-ciliated
DV
tSNE 2
Ranked expression
C
Progenitors
Plane of imaging
BBS1
TRAF3
DAPI
agat-3
prog-2
D
Epidermis
BBS1
TRAF3
prog-2 DAPI
E
Post-2b
ovo-3
BARHL1
FOXJ2
Post-2a
tlx-1
TCF/LEF-1-like
snail-like
Expression log2(CPM+1)
ζneoblast
Early
Late
Later
Mature
Ciliated
DV
F
Post-2b
Post-2a
BARHL1
tlx-1
FOXJ2
ovo-3
snail-like
laminB DAPI
G
feed dsRNA x6
10d
Control
Post-2b RNAi
Post-2a RNAi
BARHL1 RNAi
8/8
5/5
6/6
4/4
DAPI
laminB

Figure 2. Mature epidermal gene expression is abundant in epidermal progenitors. (A) Expression of validated epidermal genes in SCS data across epidermal clusters. Shown are genes that were expressed only in the epidermal lineage. Color coding corresponds to the maturation stage in which the gene was initially expressed. (B) tSNE plots of two epidermal genes validated by WISH (BBS1 and TRAF3), that were found to be expressed in Early Stage progenitor cells (expression low to high, blue to red, respectively). (C-D) FISH validation of epidermal gene expression in Early Stage and Late Stage progenitors (C) and in mature epidermis (D), taken in different focal planes in the same animals (see schematic) by confocal microscopy (scale bar = 20 μm). White arrows indicate co-expression. Early Stage (*prog-2*) and Late Stage (*agat-3*) progenitor markers are not expressed in the mature epidermis. (E) Violin plots of the eight TFs that were enriched in the DV boundary epidermal cells are shown. Black dots represent expression of a single cell. (F) FISH validation of DV-boundary epidermis TFs, including co-expression with *laminB*. The expression of the TFs is shown in higher magnification at a DV region corresponding to the white-dashed box in the overview (left) image. See also Fig S3G. Scale bar = 20 μm. (G) Inhibition of three of the DV epidermal boundary TFs led to a reduction in DV epidermal boundary gene expression in regenerating animals 10 days following amputation, as observed by FISH for *laminB* expression. Top panels: The expression of *laminB* is greatly reduced in the perturbed animals. Scale bar = 100 μm. Bottom panels: Higher magnification of the DV boundary in the different RNAi conditions. Scale bar = 20 μm.

Figure 3

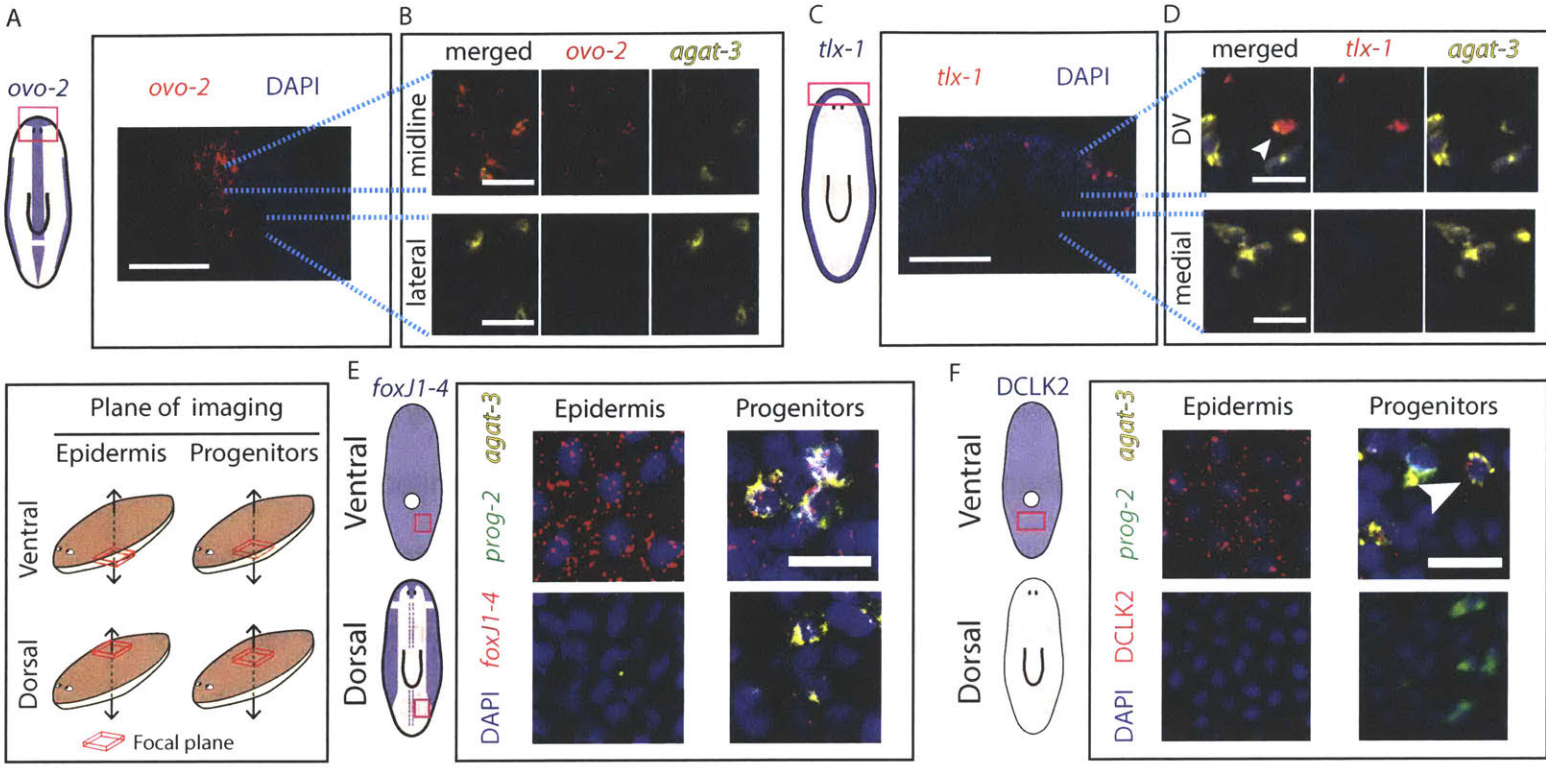

A
ovo-2
ovo-2
DAPI
B
merged
ovo-2
agat-3
midline
lateral
C
tlx-1
tlx-1
DAPI
D
merged
tlx-1
agat-3
DV
medial
Plane of imaging
Epidermis
Progenitors
Ventral
Dorsal
Focal plane
E
foxJ1-4
Ventral
Dorsal
agat-3
prog-2
foxJ1-4
DAPI
Epidermis
Progenitors
F
DCLK2
Ventral
Dorsal
agat-3
prog-2
DCLK2
DAPI
Epidermis
Progenitors

Figure 3. Expression of spatially-restricted mature epidermal identities is pre-patterned in epidermal progenitors. (A) *ovo-2* expression in the head is restricted to the midline (Scale bar = 100 μm; schematic shows spatial expression restriction in blue). (B) In the focal plane of the progenitors, *ovo-2* expression is limited to progenitors in the midline (upper panel), and is not found in progenitors immediately lateral to the midline (lower panel). Blue dotted lines indicate higher magnification area. Scale bar = 20 μm. (C) *tlx-1* expression in mature epidermis is found in DV epidermal cells (left schematic), however, *tlx-1* expression was also detectable in epidermal progenitors in proximity to the DV boundary, but not detected in progenitors in other parts of the animal (scale bar = 20 μm). (E-F) Shown are confocal images at different focal planes of dorsal or ventral regions of the animal. See key (left panel). Scale bar = 20 μm. (E) *foxJ1-4*, an essential gene for ciliogenesis [21] is expressed in ventral epidermis (top-left), and the lateral flanks the dorsal epidermis, but not in the dorsal-medial part (bottom-left; See schematic). The expression of *foxJ1-4* is present in epidermal progenitors at the ventral region (top-right), but not in the dorsal region (bottom-right). (F) The expression of DCLK2 is detectable in ventral but not dorsal epidermis (Fig 1B; left schematic). Similarly, ventral epidermal progenitors, but not dorsal progenitors already displayed DCLK2 expression, demonstrating that epidermal progenitors activate region-appropriate transcriptional program according to their DV position.

Figure 4

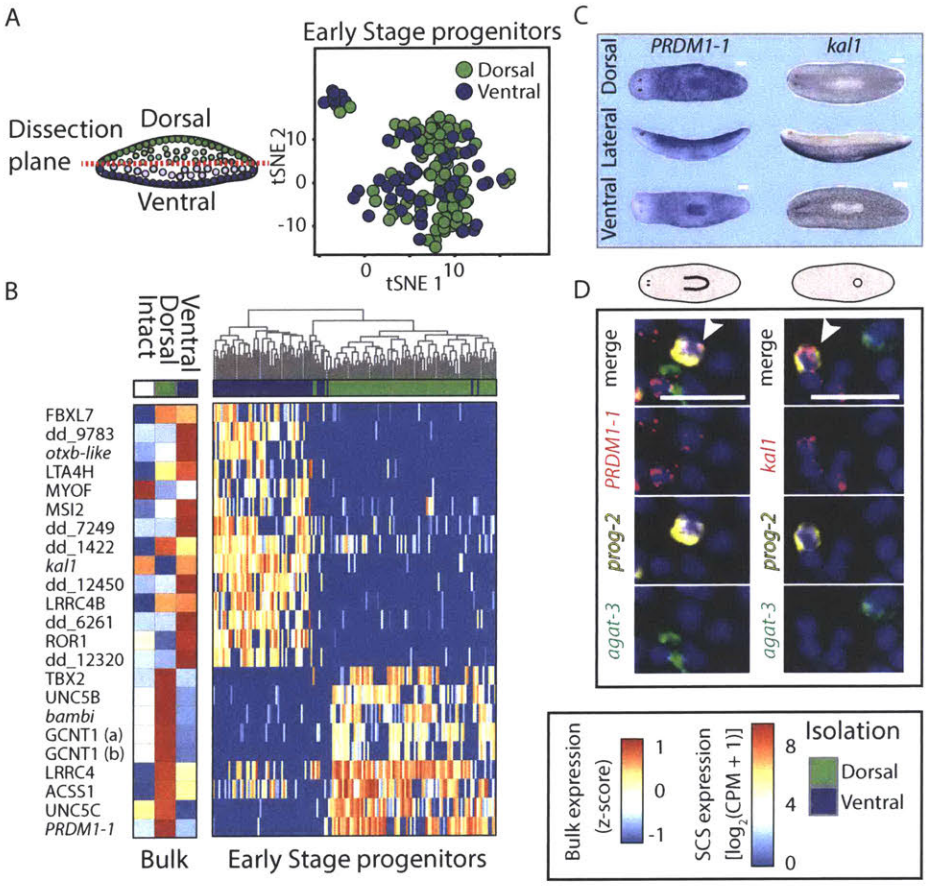

A
Early Stage progenitors
Dorsal
Ventral
Dissection plane
Dorsal
Ventral
tSNE 2
tSNE 1
B
Ventral
Dorsal
Intact
FBXL7
dd_9783
otxb-like
LTA4H
MYOF
MSI2
dd_7249
dd_1422
kal1
dd_12450
LRRC4B
dd_6261
ROR1
dd_12320
TBX2
UNC5B
bambi
GCNT1 (a)
GCNT1 (b)
LRRC4
ACSS1
UNC5C
PRDM1-1
Bulk
Early Stage progenitors
C
PRDM1-1
kal1
Dorsal
Lateral
Ventral
D
merge
PRDM1-1
prog-2
agat-3
kal1
Bulk expression (z-score)
SCS expression [log2(CPM + 1)]
Isolation
Dorsal
Ventral

Figure 4. Early Stage progenitors activate divergent transcriptional programs based on their DV location. (A) Left: Tissues were dissected from dorsal and ventral regions and epidermal progenitors were isolated by FACS and qPCR (Fig 1D; Methods). Early Stage progenitors were identified computationally (Methods), and gene expression of dorsal and ventral Early Stage progenitor samples was compared. Right: tSNE plot of Early Stage progenitors from dorsal and ventral regions (green and blue dots, dorsal and ventral cells, respectively). (B) Right: Differentially expressed genes (rows) in dorsal and ventral Early Stage progenitors (columns) demonstrate differences between cells that differ in their position across the boundary. Dendrogram was generated by hierarchical clustering of the gene expression. Many of the genes that have a biased DV-expression in the progenitors also display biased expression in the mature epidermis-enriched samples (left heat map; shown is the average expression across replicates from the same sample type). (C) WISH analyses of the dorsal-specific gene marker (*PRDM1-1*) and the ventral-specific gene marker (*kal1*) demonstrate distinct regionalized expression (scale bar = 100 μm). (D) Expression of *PRDM1-1* and *kal1* is found in Early Stage progenitors and is restricted to the dorsal or ventral regions, respectively. Shown is co-expression of either *PRDM1-1* or *kal1* with an Early Stage marker (*prog-2*; white arrow).

Figure 5

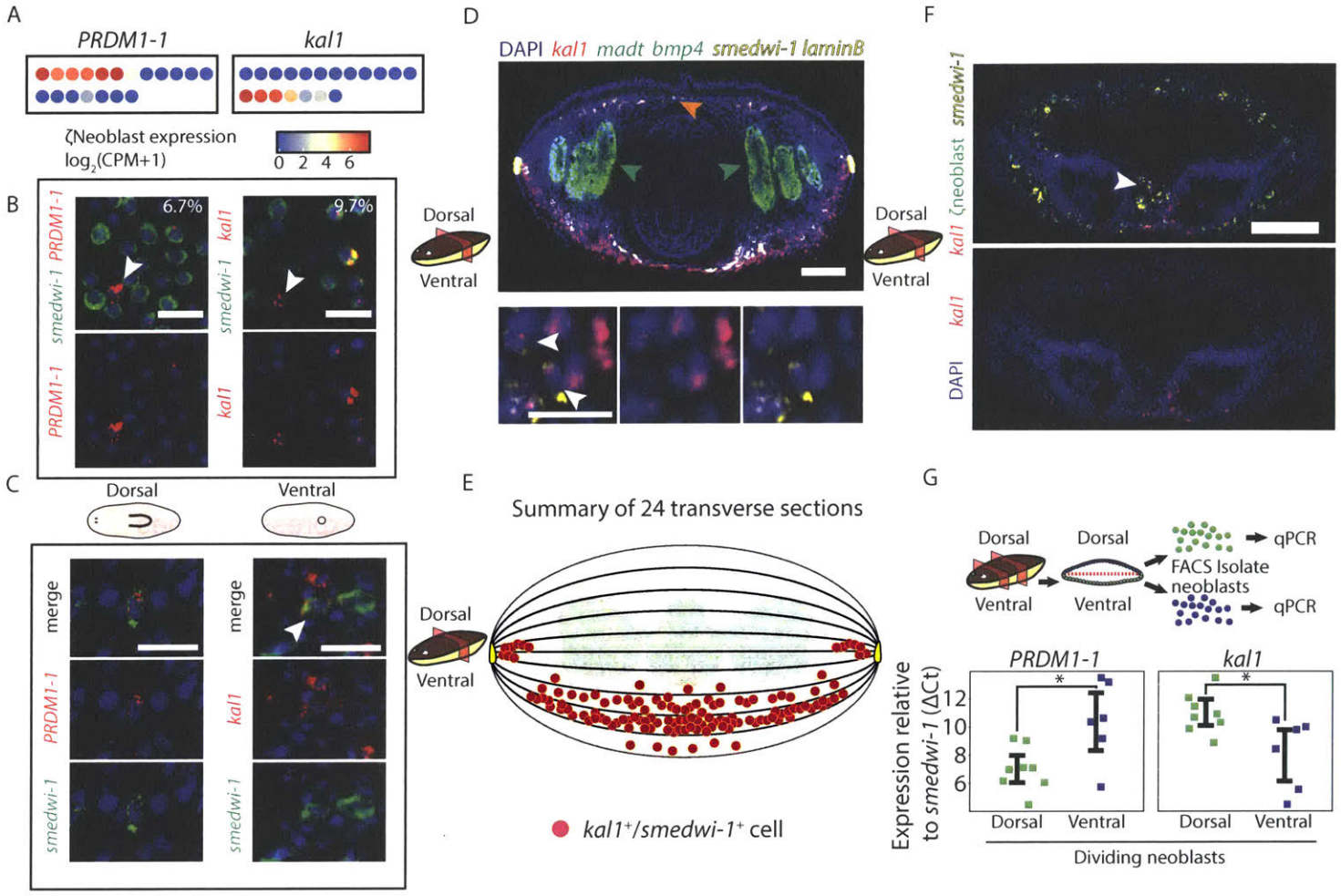

A
PRDM1-1
kal1
ζNeoblast expression
log2(CPM+1)
0 2 4 6
B
6.7%
9.7%
PRDM1-1
smedwi-1
kal1
C
Dorsal
Ventral
merge
D
DAPI kal1 madt bmp4 smedwi-1 laminB
Dorsal
Ventral
E
Summary of 24 transverse sections
kal1+/smedwi-1+ cell
F
kal1 ζneoblast smedwi-1
kal1
DAPI
G
qPCR
FACS Isolate neoblasts
Expression relative to smedwi-1 (ΔCt)
Dividing neoblasts

Figure 5. Spatially opposed expression of *PRDM1-1* and *kal1* in neoblasts (A) ζneoblasts express *PRDM1-1* or *kal1* in mutually exclusive single ζneoblasts (dots; blue to red, low to high gene expression, respectively). (B) The fraction of neoblasts expressing dorsal (*PRDM1-1*) or ventral (*kal1*) markers was estimated by co-expression analysis of FACS-isolated dividing neoblasts, and co-staining with a neoblast marker (*smedwi-1*). (C) FISH showing *PRDM1-1* expression in dorsal neoblasts (left), and *kal1* is found only in ventral neoblasts (right). *kal1* was not expressed in dorsal neoblasts in at least 10 animals; *PRDM1-1* was not expressed in ventral neoblasts in at least 10 animals. (D) Shown is a transverse section (see cartoon) labeled for *kal1* (red), madt (green; intestinal marker; green arrows), *bmp4* (green; orange arrow), *smedwi-1* (yellow; neoblast marker), and *laminB* (yellow; DV-boundary epidermis). Scale bar = 100 μm. *kal1+/smedwi-1+* cells were counted (bottom panel, white arrows) and their position in the animal was documented (Supplementary Methods). Scale bar = 20 μm. See also Fig S6E. (E) Summary of the location of *kal1+/smedwi-1+* cells in 24 transverse sections (magenta circles). Sections were divided to 10 regions (bounded by black lines) and the number of *kal1+/smedwi-1+* cells was counted in each region (Supplementary Methods). The medial-lateral position of the cells was not measured and the medial-lateral spread is for clarity, except for cells next to the DV-boundary epidermis (yellow cartoon), which represent the approximate distance of these cells from the boundary. All *kal1+/smedwi-1+* cells were ventral to the intestine (green cartoon), except for those next to the DV-boundary epidermis. (F) FISH on animals labeled for ζneoblasts markers (pool of *soxP-3* and *zfp-1*) and *kal1* shows that *kal1-* ζneoblasts extends more internally than *kal1* expression (white arrow), suggesting that

ventral ζneoblasts express *kal1* when they are only near the epidermal surface. Top panel shows all channels; bottom panel shows *kal1* expression and DAPI. Scale Bar = 100 μm. See Figure S6F for single channel images. (G) Top panel: Schematic of dividing neoblast isolation by FACS (Supplementary Methods). Pre-pharyngeal segments were isolated and were dissected into dorsal and ventral fragments. The fragments were macerated, separately, and cells (green and blue dots, dorsal and ventral cells, respectively) were isolated by FACS into groups of 500 cells (Methods; eight dorsal samples and six ventral samples). The samples were processed (Methods) and; bottom panel: tested by qPCR for the expression of *PRDM1-1* and *kal1*, with their expression normalized by *smedwi-1* expression (green and blue squares represent 500-cell samples, dorsal or ventral, respectively). Significance of difference in expression was tested by Student's t-test and corrected for multiple hypotheses testing using Bonferroni correction.

Figure 6

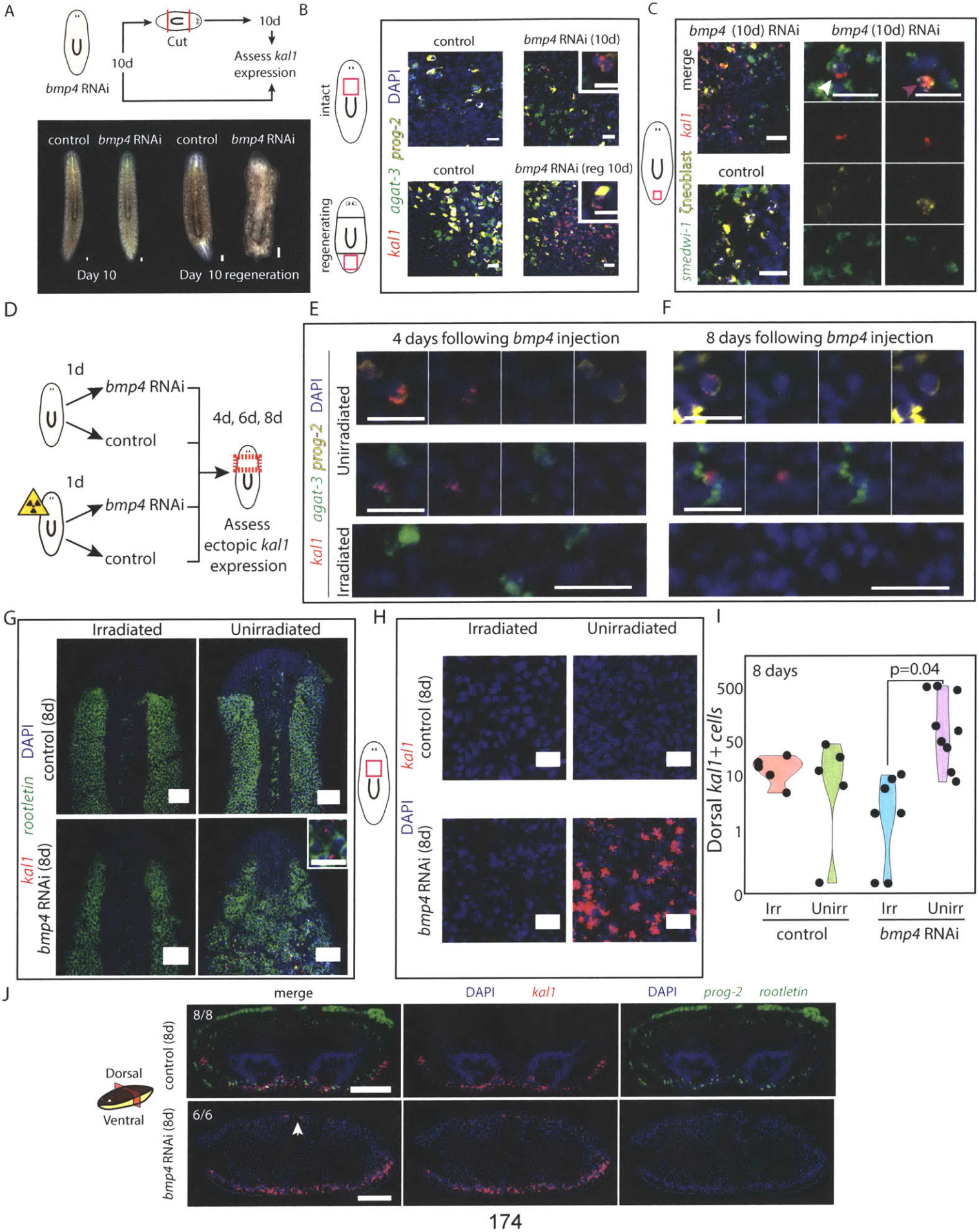

A
bmp4 RNAi
10d
Cut
10d
Assess kal1 expression
control
bmp4 RNAi
control
bmp4 RNAi
Day 10
Day 10 regeneration
B
intact
regenerating
control
bmp4 RNAi (10d)
control
bmp4 RNAi (reg 10d)
DAPI
prog-2
agat-3
kal1
C
bmp4 (10d) RNAi
bmp4 (10d) RNAi
merge
kal1
control
neoblast
smedwi-1
D
1d
bmp4 RNAi
control
1d
bmp4 RNAi
control
4d, 6d, 8d
Assess ectopic kal1 expression
E
4 days following bmp4 injection
F
8 days following bmp4 injection
DAPI
prog-2
agat-3
kal1
Unirradiated
Irradiated
G
Irradiated
Unirradiated
control (8d)
bmp4 RNAi (8d)
DAPI
rootletin
kal1
H
Irradiated
Unirradiated
control (8d)
bmp4 RNAi (8d)
DAPI
kal1
I
8 days
p=0.04
Dorsal kal1+ cells
500
50
10
1
0
Irr
Unirr
Irr
Unirr
control
bmp4 RNAi
J
merge
DAPI
kal1
DAPI
prog-2
rootletin
Dorsal
Ventral
control (8d)
bmp4 RNAi (8d)
8/8
6/6

Figure 6. Planarian ζneoblast gene expression is regulated by extracellular positional signaling. (A) *bmp4* was inhibited by RNAi (Methods), and 10 days following the first feeding of three, animals were either fixed or cut and allowed to regenerate for 10 days (upper panel). Animals did not have any visible defects prior to cutting (bottom panel), but following cutting and regeneration, they displayed major morphological defects, as previously reported [23–25]. (B) *kal1* expression was observed 10 days following *bmp4* RNAi on the dorsal surface, but never in control animals. Shown are animals before cutting (top) and animals at 10 day following regeneration (labeled reg 10d). Insets show a dorsal cell in *bmp4* RNAi animal that co-expresses an Early Stage marker (*prog-2*) and *kal1*, demonstrating that reduction in BMP expression is sufficient for the appearance of ventral markers dorsally. Inset scale bar = 10 μm. (C) *kal1* expression is observed in dorsal ζneoblasts, suggesting that ζneoblasts sense *bmp4* expression in their environment and specify their positional gene expression accordingly. At day 10 following *bmp4* inhibition animals displayed *kal1* expression dorsally (left panel). Middle panel: In *bmp4* RNAi animals, we detected dorsal *kal1*+ cells that also express *smedwi-1* (a pan-neoblast marker), and ζneoblast markers (combination of *soxP-3* and *zfp-1* [19]), indicating that they were indeed epidermal neoblasts (white arrow). Right panel: In addition, we identified *kal1*+/ζneoblast+ cells that did not express *smedwi-1*, suggesting that they are more differentiated epidermal progenitors. Scale bar = 20 μm. (D) Animals were divided into two cohorts, one of which was lethally irradiated on day zero; the other was unperturbed. Half of the animals each cohort received a single injection of either *bmp4* dsRNA or a control dsRNA. Animals were fixed and analyzed at day four, day six and

day eight following injection. (E-F) At days four and eight following *bmp4* dsRNA injections, epidermal progenitors on the dorsal side (*prog-2*+ or *agat-3*+ cells, top and middle panels, respectively), which expressed *kal1*, were detectable. Conversely, *kal1*+ dorsal epidermal progenitors were not detected in irradiated animals. At day eight following irradiation, epidermal progenitors are completely ablated. See also Fig S7A. Scale bar = 20 μm. (G) The dorsal expression of *rootletin* in the irradiated *bmp4* animals is indistinguishable from control animals. Conversely, in three of eight unirradiated *bmp4* RNAi animals, the dorsal *rootletin* expression resembled the ventral surface of the animal. See Fig S7B. Scale bar = 100 μm. Inset shows *rootletin*+ cells expressing *kal1*. Inset scale bar = 20 μm. (H) Expression of *kal1* in dorsal pre-pharyngeal regions. Only unirradiated *bmp4* RNAi had dorsal expression. Scale bar = 20 μm. (I) Quantification of the number of *kal1*+ cells on the pre-pharyngeal dorsal region of each animal (black dot) analyzed. Significance was assessed using Student's t-test. Groups labeled Irr and Unirr represent irradiated and unirradiated groups, respectively. (J) Transverse section showing the spatial distribution of *kal1*+ cells in control and *bmp4* RNAi animals. *kal1*+ cells were never found in the middle of the animal, despite ectopic expression in all *bmp4* RNAi sections analyzed. Scale bar = 100 μm.

Figure 7

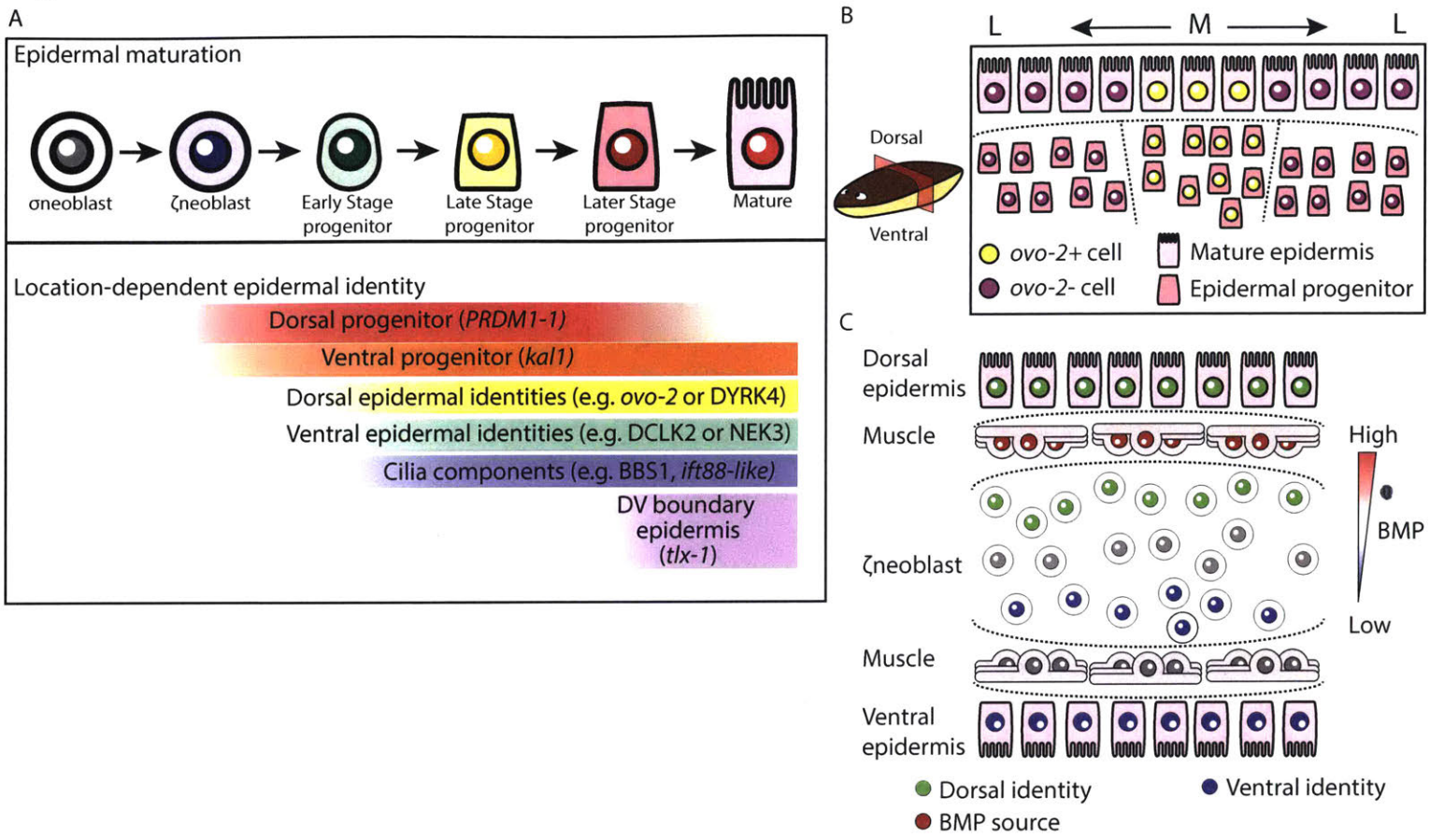

A
Epidermal maturation
σneoblast
ζneoblast
Early Stage progenitor
Late Stage progenitor
Later Stage progenitor
Mature
Location-dependent epidermal identity
Dorsal progenitor (PRDM1-1)
Ventral progenitor (kal1)
Dorsal epidermal identities (e.g. ovo-2 or DYRK4)
Ventral epidermal identities (e.g. DCLK2 or NEK3)
Cilia components (e.g. BBS1, ift88-like)
DV boundary epidermis (tlx-1)
B
L
M
L
Dorsal
Ventral
ovo-2+ cell
ovo-2- cell
Mature epidermis
Epidermal progenitor
C
Dorsal epidermis
Muscle
ζneoblast
Muscle
Ventral epidermis
High
BMP
Low
Dorsal identity
Ventral identity
BMP source

Figure 7. Model for planarian epidermal differentiation. (A) Top panel: Planarian epidermal cells mature in a spatiotemporally defined sequence, in which progenitors go through several transitions defined by cellular morphology and spatial distribution. Bottom panel: Genes that are associated with distinct epidermal identities are expressed in all stages of the epidermal maturation, in a spatially-defined manner. Some of these genes are already expressed in ζneoblasts, suggesting that ζneoblasts sense their position in the animal. (B) Some spatially-restricted epidermal identities are pre-patterned in epidermal progenitors. The model presents a view of a transverse cross-section. Dorsal is up. Medial mature epidermal cells (M; Large shapes) express *ovo-2* (yellow nuclei), whereas lateral cells do not (L; purple nuclei). FISH analysis of epidermal progenitors determined that epidermal progenitors near the midline already express *ovo-2* (yellow), whereas lateral epidermal progenitors do not (purple). (C) The specification of ζneoblasts to ventral (*kal1+*) identities is repressed by *bmp4* expression from dorsal muscle cells (red elongated cells). ζneoblasts, which are in proximity to the BMP source, can express dorsal markers (*PRDM1-1+*; green circles), which is correlated with the emergence of additional dorsal epidermal identities. By contrast, cells that are far from the BMP source can specialize to ventral progenitors (*kal1+*; blue circles), which are associated with ventral epidermal identities. Inhibition of BMP signaling results in the emergence of *kal1+* ζneoblasts dorsally.

Figure S1

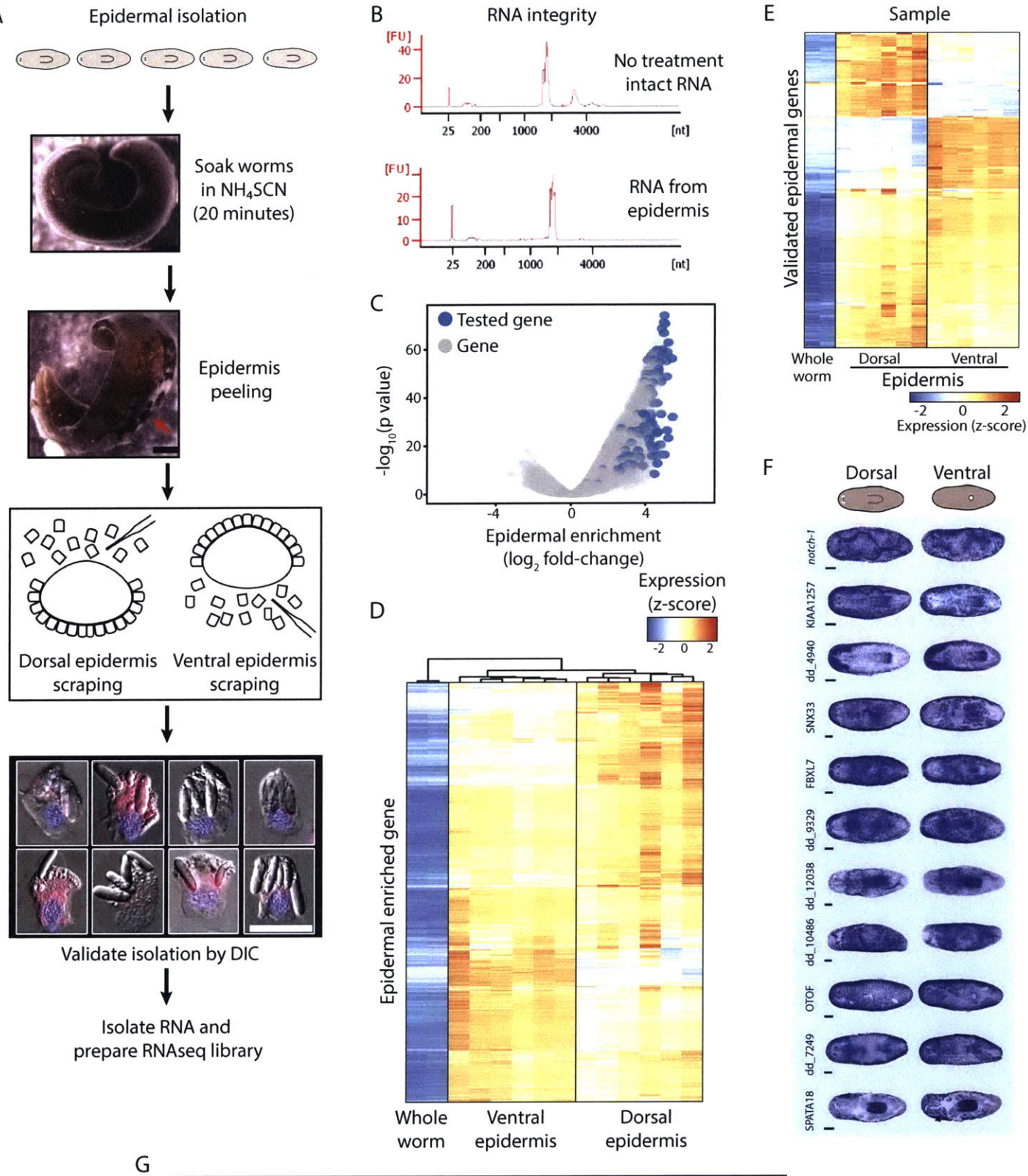

A
Epidermal isolation
Soak worms in NH_4SCN (20 minutes)
Epidermis peeling
Dorsal epidermis scraping
Ventral epidermis scraping
Validate isolation by DIC
Isolate RNA and prepare RNAseq library
B
RNA integrity
[FU]
No treatment intact RNA
RNA from epidermis
[nt]
C
Tested gene
Gene
$-\log_{10}$(p value)
Epidermal enrichment ($\log_2$ fold-change)
D
Expression (z-score)
Epidermal enriched gene
Whole worm
Ventral epidermis
Dorsal epidermis
E
Sample
Validated epidermal genes
Whole worm
Dorsal
Ventral
Epidermis
Expression (z-score)
F
Dorsal
Ventral
notch-1
KIAA1257
dd_4940
SNX33
FBXL7
dd_9329
dd_12038
dd_10486
OTOF
dd_7249
SPATA18
G
DAPI
vim-1
PRSS12
laminB
Dorsal
Ventral

Figure S1 related to Figure 1. Epidermal expression patterns identified by epidermal RNAseq. (A) A method for epidermal RNA extraction (Methods). Worms were soaked in ammonium thiocyanate [27] in batches of 5 animals for 20 minutes, which resulted in epidermal peeling (red arrow). Dorsal or ventral epidermis were scraped separately and the cells were visualized by differential interference contrast (DIC) microscopy. RNA was then extracted for RNAseq library preparation (Methods). (B) The quality of epidermal RNA obtained by ammonium thiocyanate treatment and scraping was compared to RNA from whole animals. Shown is output of RNA fragment analysis by BioAnalyzer 2100 (Agilent Technologies). X-axis is the RNA fragment length and y-axis is its fluorescence intensity. Control whole-worm RNA (top) and epidermal RNA sample (down) had comparable fragment size distributions indicating of comparable RNA integrity. (C) Shown is gene expression in epidermis-enriched samples compared to controls, as determined by RNASeq. A volcano plot of planarian genes (gray dots) showing the average epidermal $\log_2$ fold-enrichment over whole worm control expression (x-axis) and the significance (p value) of the differential expression (y-axis). Highly enriched genes (blue dots) were selected for validations by in situ hybridization. p values were calculated using edgeR [41] (Methods). (D) Heat map of all genes that were differentially expressed in the epidermis enriched libraries compared to the whole worm controls; samples (columns) and genes (rows). Expression is shown as row z-score (blue to red, low to high, respectively). Dendrogram of samples (top) is based on hierarchical clustering of the epidermal-enriched genes (Table S1). (E) Heat map of epidermis-enriched genes tested by ISH. Columns correspond to samples and rows correspond to genes tested by ISH (Table S1; row z-score; blue to red, low to

high). (F) Epidermal expression patterns. Shown are representative epidermal gene expression patterns. Genes were selected for WISH based on the epidermis-enriched RNAseq gene expression analysis (Table S1; see also Fig 1B-C). Note that some genes were also expressed in other tissues (*e.g.* dd_4940 in the brain, and SPATA18 in the pharynx). Scale bar = 100 μm. (G) Top panel: Transverse section showing the expression of two distinct mature epidermal gene expression patterns (PRSS12 and *laminB*), and expression of *vim-1*, which is abundant in mature cells, but is also found in epidermal progenitors at sub-epithelial layer (white arrowhead). Bottom panel: higher magnification. Scale bar = 20 μm.

Figure S2

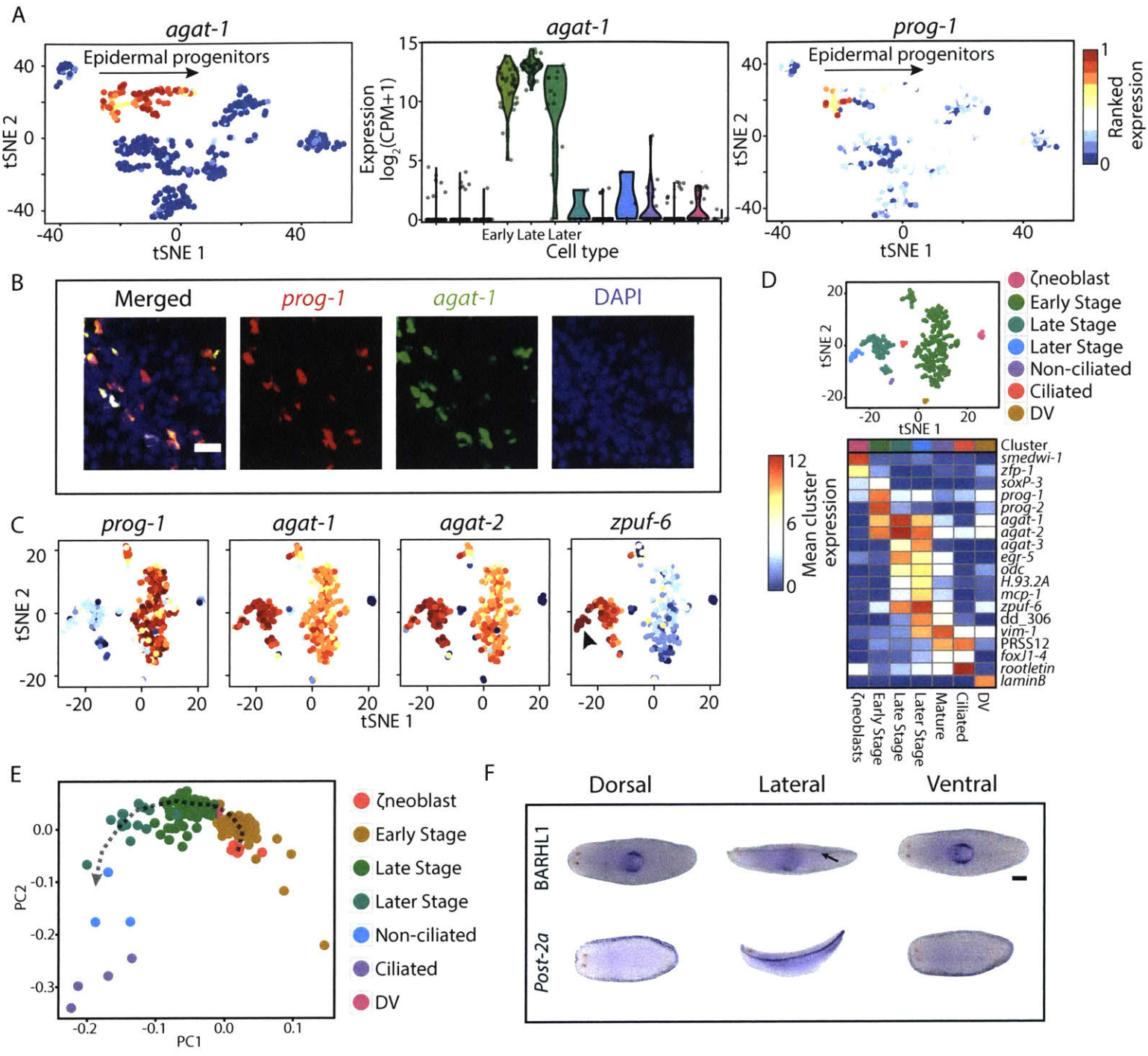

A
agat-1
Epidermal progenitors
tSNE 2
tSNE 1
agat-1
Expression log2(CPM+1)
Early Late Later
Cell type
prog-1
Epidermal progenitors
Ranked expression
B
Merged
prog-1
agat-1
DAPI
C
prog-1
agat-1
agat-2
zpuf-6
tSNE 2
tSNE 1
D
ζneoblast
Early Stage
Late Stage
Later Stage
Non-ciliated
Ciliated
DV
Mean cluster expression
Cluster
smedwi-1
zfp-1
soxP-3
prog-1
prog-2
agat-1
agat-2
agat-3
egr-5
odc
H.93.2A
mcp-1
zpuf-6
dd_306
vim-1
PRSS12
foxJ1-4
rootletin
laminB
ζneoblasts
Early Stage
Late Stage
Later Stage
Mature
Ciliated
DV
E
PC1
PC2
ζneoblast
Early Stage
Late Stage
Later Stage
Non-ciliated
Ciliated
DV
F
Dorsal
Lateral
Ventral
BARHL1
Post-2a

Figure S2 related to Figure 1. Clustering and classification of epidermal progenitors. (A) Left panel: a tSNE plot showing a single cell resolution expression of *agat-1*. Cells are dots colored from blue to red, low to high expression, respectively. Data obtained from a recently published SCS dataset [28]. *agat-1* is expressed in all post-mitotic epidermal progenitors (cells below arrow). Arrow direction corresponds to stages of epidermal maturation ("Early Stage", "Late Stage", and "Later Stage"). Middle panel: *agat-1* is highly expressed in all post-mitotic epidermal progenitors, but not in other cell types. Shown is a violin plot, where each group (x-axis) corresponds to a cell type and the y-axis corresponds to the $\log_2$ expression of *agat-1* in the cells of that group. Expression of *agat-1* is highest in Late Stage progenitors, but it is also highly expressed in Early Stage and Later Stage progenitors as well. Therefore, SCS data suggests that *agat-1* expression is appropriate for finding epidermal lineage progenitors from all stages of post-mitotic epidermal maturation. Labels are shown only for post-mitotic epidermal progenitor clusters. Right panel: tSNE plot of *prog-1* ranked expression across different cell types. *prog-1* is highly expressed in Early Stage progenitors, as previously reported [22]. However, *prog-1* is also lowly expressed in all other cell types profiled in this dataset [28], which makes it not ideal for using as a stage-specific FISH marker. Color bar of ranked expression: blue to red, low to high, respectively. (B) FISH demonstrates substantial co-expression between *prog-1* and *agat-1* (up to 100% depending on the site examined). (C) tSNE plots of previously described epidermal lineage gene expression markers [20], in the epidermal lineage. See also Fig S1D. Shown is ranked expression. Blue to red, low to high, respectively. As previously reported, the highest expression of *zpuf-6* was observed in Later Stage

progenitors (black arrowhead). (D) Mean cluster expression [log_2(CPM+1)] of previously described gene markers for different stages of planarian epidermal lineage development. Stages, in columns, correspond to the color-coded tSNE plot (top) of the SCS data used in these analyses. (E) PCA plot of the 303 single cells (dots) gene expression used in this study. Principal components 1 and 2 distinguish the different stages of epidermal maturation and suggest of a developmental trajectory (gray dotted line), in agreement with previous reports [20]. (F) WISH analysis of two TFs that were identified as having DV epidermal expression (scale bar = 200 μm). Black arrow points to DV expression domain of BARHL1.

Figure S3

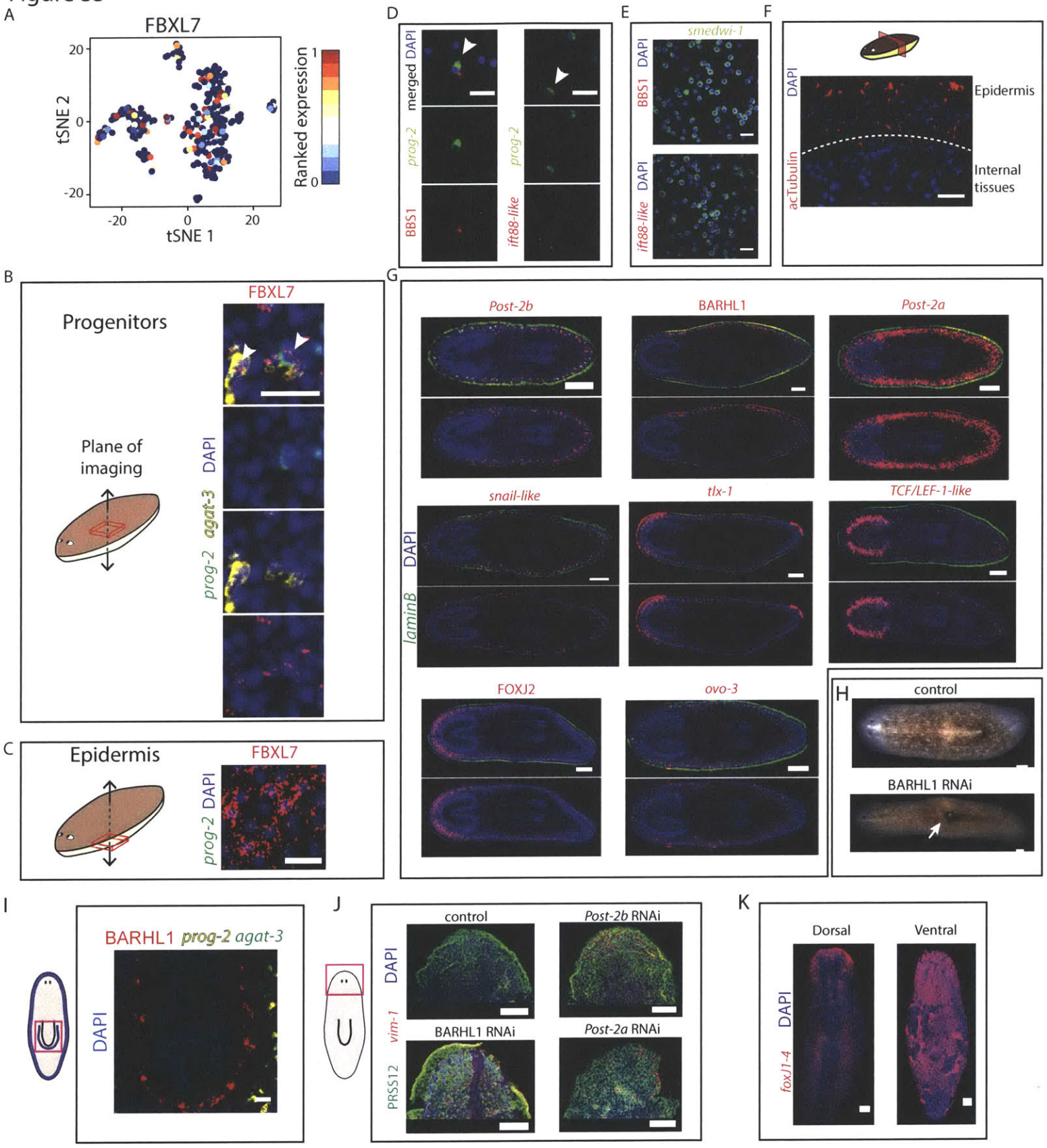
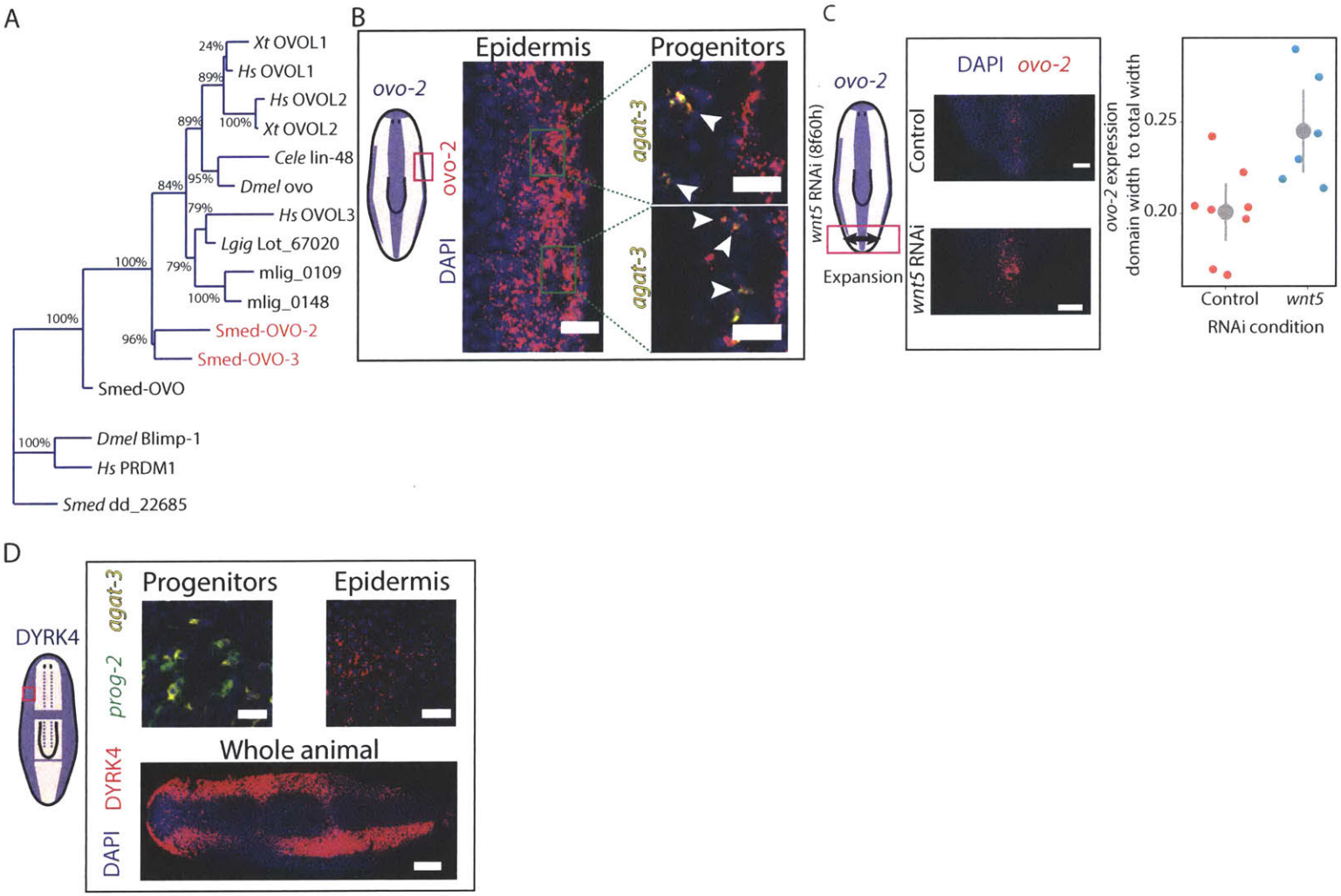

A
FBXL7
tSNE 2
tSNE 1
Ranked expression
B
Progenitors
Plane of imaging
FBXL7
DAPI
agat-3
prog-2
C
Epidermis
FBXL7
DAPI
prog-2
D
merged DAPI
prog-2
BBS1
ift88-like
E
smedwi-1
DAPI
BBS1
ift88-like
F
DAPI
Epidermis
acTubulin
Internal tissues
G
Post-2b
BARHL1
Post-2a
snail-like
tlx-1
TCF/LEF-1-like
DAPI
laminB
FOXJ2
ovo-3
H
control
BARHL1 RNAi
I
BARHL1 prog-2 agat-3
DAPI
J
control
Post-2b RNAi
DAPI
vim-1
BARHL1 RNAi
Post-2a RNAi
PRSS12
K
Dorsal
Ventral
DAPI
foxJ1-4

Figure S3 related to figure 2. Mature epidermal genes are expressed in epidermal progenitors. (A) tSNE plot of FBXL7, which was expressed in the mature epidermis, and was also expressed in epidermal progenitors. Shown is ranked expression in single cells (blue to red, low to high expression, respectively). (B) FISH validated the expression of FBXL7 in progenitors (white arrowheads). (C) The expression of FBXL7 was stronger and more abundant in the mature epidermis. Scale bar = 20 μm. (D-E) FISH was performed on cells isolated by FACS (cell FISH; Supplementary Methods) from the progenitor enriched gate (X2) or the dividing neoblasts gate (X1) [42]. (D) Early Stage epidermal progenitors (*prog-2+*) express genes that encode cilia components (BBS1 and *ift88-like*, left and right panel, respectively) Double positive cells are marked with a white arrowhead. (E) The expression of the same two genes (BBS1 and *ift88-like*) was not detectable in FACS-isolated dividing neoblasts (n > 500 cells), suggesting that they are expressed only in post-mitotic cells. Expression of other cilia-encoding genes, however, was detectable by SCS data (Table S4) in agreement of recent report [33]. (F) Transverse section of an animal showing the protein distribution of acetylated Tubulin (acTubulin), a cilia component. acTubulin expression is highly enriched in the mature epidermis (above dashed line), and is largely restricted from deeper tissues, demonstrating that functional cilia are found in mature epidermal cells and not in epidermal progenitors. A maximal intensity projection is shown. Scale bar = 20 μm. (G) Shown are the expression patterns of the DV-boundary epidermis enriched TFs. Animals were also labeled with the DV-boundary epidermis marker *laminB*. The co-expression of *TCF-LEF-1-like* and *laminB* was not detectable by FISH. (H) Inhibition of BARHL1 led to a lesion above the pharynx (white arrow) in all animals analyzed. Scale

bar = 100 µm. (I) BARHL1, a DV boundary epidermal TF, is also expressed in the pharyngeal cavity. The BARHL1+ cells in the pharyngeal cavity are not co-expressing Early Stage (*prog*-2) or Late Stage (*agat-3*) markers. Scale bar = 20 µm. (J) The expression of mature and Later Stage epidermal markers (PRSS12 and *vim-1*, respectively) were not affected by inhibition of the DV boundary epidermal TFs. (K) Shown is the expression pattern of *foxJ1-4*, a TF that is essential to ciliogenesis. (J-K) Scale bar = 100 µm.

Figure S4


A
Xt OVOL1
Hs OVOL1
Hs OVOL2
Xt OVOL2
Cele lin-48
Dmel ovo
Hs OVOL3
Lgig Lot_67020
mlig_0109
mlig_0148
Smed-OVO-2
Smed-OVO-3
Smed-OVO
Dmel Blimp-1
Hs PRDM1
Smed dd_22685
24%
89%
89%
100%
84%
95%
79%
100%
79%
100%
100%
96%
100%
B
ovo-2
DAPI
Epidermis
Progenitors
agat-3
agat-3
C
ovo-2
wnt5 RNAi (8f60h)
Expansion
DAPI ovo-2
Control
wnt5 RNAi
ovo-2 expression domain width to total width
0.25
0.20
Control
wnt5
RNAi condition

D
DYRK4
agat-3
prog-2
DYRK4
DAPI
Progenitors
Epidermis
Whole animal

Figure S4 related to figure 3. Spatially-restricted epidermal gene expression is pre-patterned in progenitors. (A) Phylogenetic tree for *Smed-ovo-2* and *Smed-ovo-3 in Schmidtea mediterranea* and model organisms are shown (*Dmel*: *D. melanogaster*; *Cele*: *C. elegans*; *Spur*: *S. purpuratus*; *Hs*: *H. sapiens*; *Xt*: *X tropicalis*). Maximum-likelihood phylogenetic inference was performed with RaXML (Supplementary Methods); bootstrap support for phylogeny appears as percent. (B) *ovo-2* expression was detectable in *agat-3*+ cells near the lateral mature epidermal domain of *ovo-2* (see cartoon). Left panel: maximal intensity projection of *ovo-2* expression in epidermal cells. Right panel: single confocal planes showing *ovo-2*+ Late Stage progenitors (white arrowheads) located under *ovo-2*+ mature epidermis (green rectangles). Scale bar = 20 μm. (C) Left panel: *ovo-2* domain of expression is affected by *wnt5* inhibition, which leads to the expansion of the planarian midline [34,35]. Scale bar = 100 μm. Right panel: *ovo-2* expression domain width compared to the entire animal width. Colored dots (red and blue, control and *wnt5* RNAi, respectively). Gray dot is the mean, and gray line is the standard mean error. (D) Top panel: The expression of DYRK4 was not observed in epidermal progenitors, but only in the mature epidermis. Scale bar = 20 μm. Bottom panel: expression pattern of DYRK4 in the mature dorsal epidermis (shown is a maximal intensity projection). Scale bar = 100 μm.

Figure S5

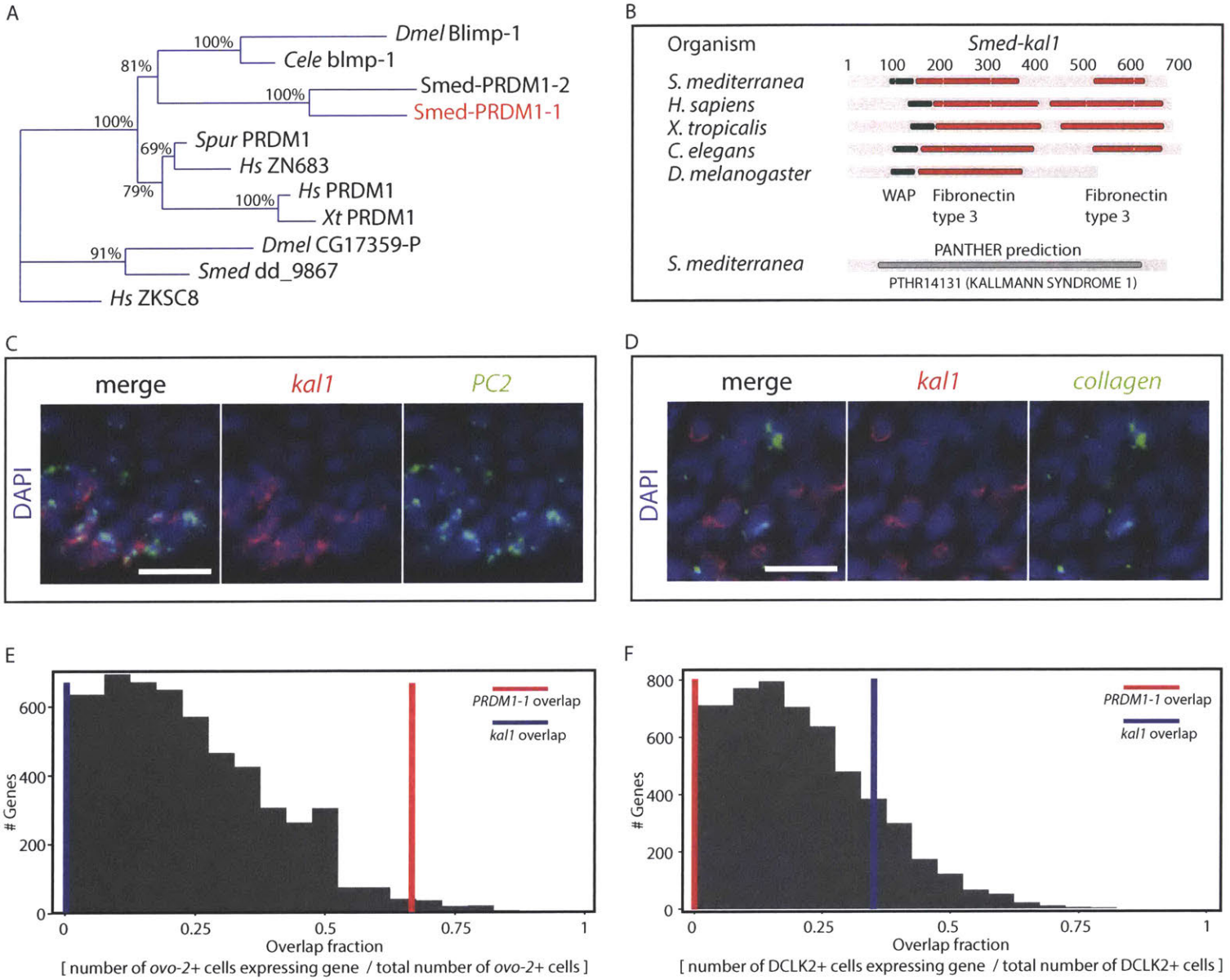

A
100%
81%
100%
100%
69%
79%
100%
91%
Dmel Blimp-1
Cele blmp-1
Smed-PRDM1-2
Smed-PRDM1-1
Spur PRDM1
Hs ZN683
Hs PRDM1
Xt PRDM1
Dmel CG17359-P
Smed dd_9867
Hs ZKSC8
B
Organism
Smed-kal1
1 100 200 300 400 500 600 700
S. mediterranea
H. sapiens
X. tropicalis
C. elegans
D. melanogaster
WAP
Fibronectin type 3
Fibronectin type 3
PANTHER prediction
S. mediterranea
PTHR14131 (KALLMANN SYNDROME 1)
C
merge
kal1
PC2
DAPI
D
merge
kal1
collagen
DAPI
E
PRDM1-1 overlap
kal1 overlap
Genes
0
200
400
600
0
0.25
0.5
0.75
1
Overlap fraction
[number of ovo-2+ cells expressing gene / total number of ovo-2+ cells]
F
PRDM1-1 overlap
kal1 overlap
Genes
0
200
400
600
800
0
0.25
0.5
0.75
1
Overlap fraction
[number of DCLK2+ cells expressing gene / total number of DCLK2+ cells]

Figure S5 related to Figure 4. Dorsal and ventral-specific epidermal genes. (A) Phylogenetic tree for *PRDM1-1* in *Schmidtea mediterranea* and model organisms is shown (*Dmel*: *D. melanogaster*; *Cele*: *C. elegans*; *Spur*: *S. purpuratus*; *Hs*: *H. sapiens*; *Xt*: *X tropicalis*). Maximum-likelihood phylogenetic inference was performed with RaXML (Supplementary Methods); bootstrap support for phylogeny appears as percent. (B) Protein domain analysis of a putative planarian *kal1*, which was identified by reciprocal BLAST analysis (Supplementary Methods), was performed. In green shown a WAP domain and in red fibronectin type 3 domains. Numbers above the domain structure indicate the amino acid number in the protein. Below shown is the PANTHER [43] prediction for the putative planarian *kal1*. (C-D) Expression of *kal1* is not restricted to the epidermal lineage, and was found in other cell types on the ventral side. Shown is *PC2*, a marker for neural tissues (C) and *collagen*, a marker for planarian muscle (D). Scale bar = 20 μm. (E-F) The overlap fraction was calculated for each gene (Supplementary Methods) in the SCS dataset. A histogram was generated from the calculated overlap fractions. (E) The overlap fraction, for *ovo-2*, of the dorsal marker *PRDM1-1* is highlighted with a red line (67%; percentile rank of overlap = 0.98), and of the ventral marker, *kal1* (0%), with a blue line. (F) The overlap fraction, for DCLK2, of the dorsal marker *PRDM1-1* is highlighted with a red line (0%), and of the ventral marker, *kal1* (35%; percentile rank of overlap = 0.85), with a blue line.

Figure S6

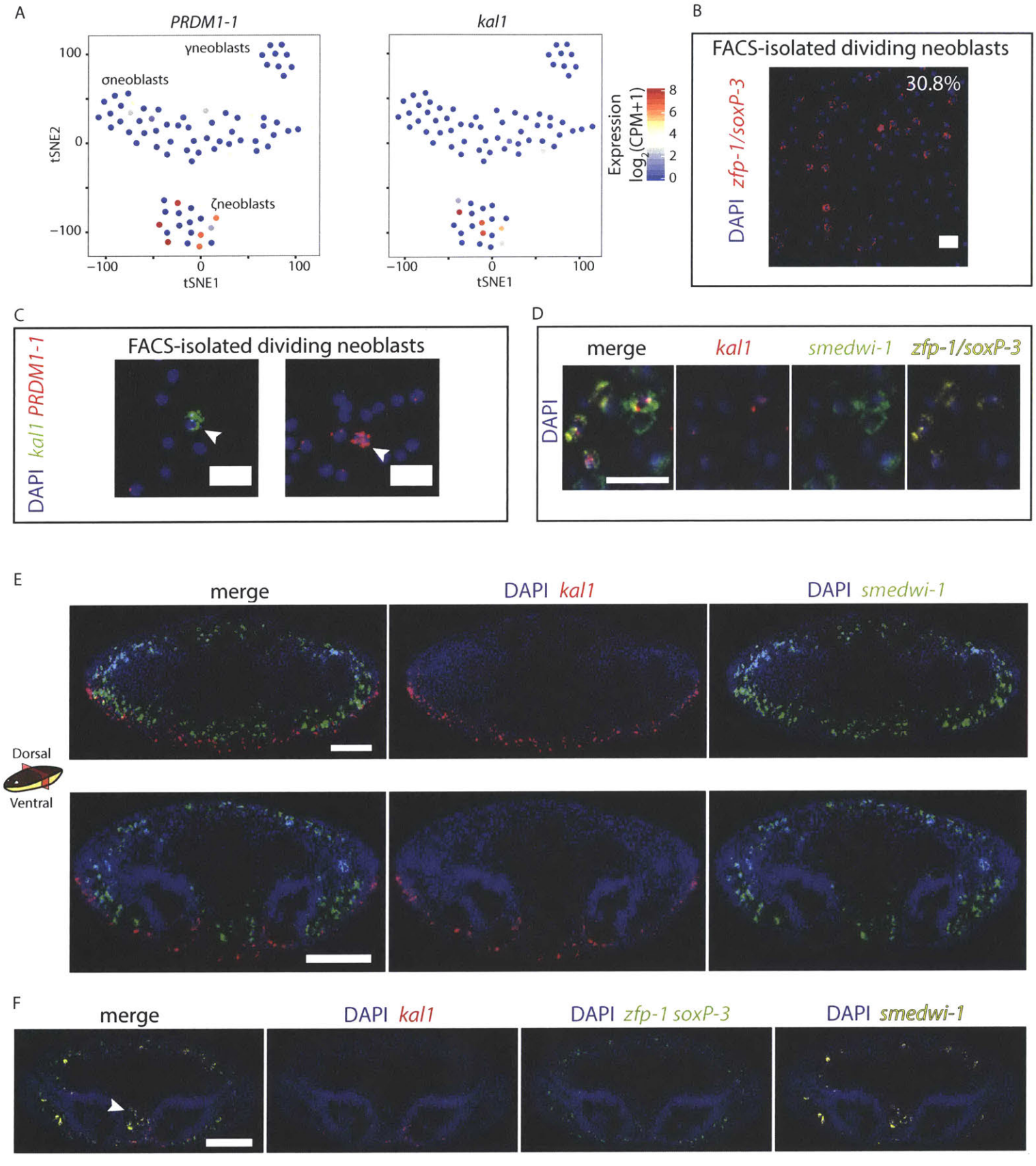

A
PRDM1-1
kal1
γneoblasts
σneoblasts
ζneoblasts
tSNE2
tSNE1
Expression log2(CPM+1)
B
FACS-isolated dividing neoblasts
30.8%
DAPI zfp-1/soxP-3
C
FACS-isolated dividing neoblasts
DAPI kal1 PRDM1-1
D
merge
kal1
smedwi-1
zfp-1/soxP-3
DAPI
E
merge
DAPI kal1
DAPI smedwi-1
Dorsal
Ventral
F
merge
DAPI kal1
DAPI zfp-1 soxP-3
DAPI smedwi-1

Figure S6 related to Figure 5. Expression of *kal1* and *PRDM1-1* in neoblasts. (A) Shown is a tSNE plot of SCS of dividing neoblasts (dots) with the expression of *PRDM1-1* (left) and *kal1* (right) overlaid (blue to red, low to high expression, respectively). The expression of *PRDM1-1* and *kal1* is highly specific to ζneoblasts. Data shown was generated using the online resource we recently published [28]. (B) Shown is cell FISH of FACS-isolated dividing neoblasts, which were labeled with ζneoblast-specific markers (*zfp-1* and *soxP-3*) [19]. The fraction of ζneoblasts in the neoblast pool (30.8%) was estimated by counting cells expressing these markers (n = 1293). (C) FACS-isolated dividing neoblasts were used for FISH for *kal1* and *PRDM1-1* expression. In the 1061 cells we analyzed, we did not observe co-expression of these genes. (D) *kal1* expression is observed, by whole-mount FISH, in ζneoblasts. ζneoblasts were labeled by *smedwi-1* and a combination of *zfp-1* and *soxP-3*. (B-D) Scale bar = 20 μm. (E) Representative transverse sections of FISH on animals labeled with *kal1* and *smedwi-1* reveals that the expression of ventral *smedwi-1* extends more internally than *kal1*. (F) Transverse section showing that *kal1*- ζneoblasts extends more internally than *kal1* expression (white arrowhead). This is a single-channel image of Fig 5F. (E-F) Scale bar = 100 μm.

Figure S7

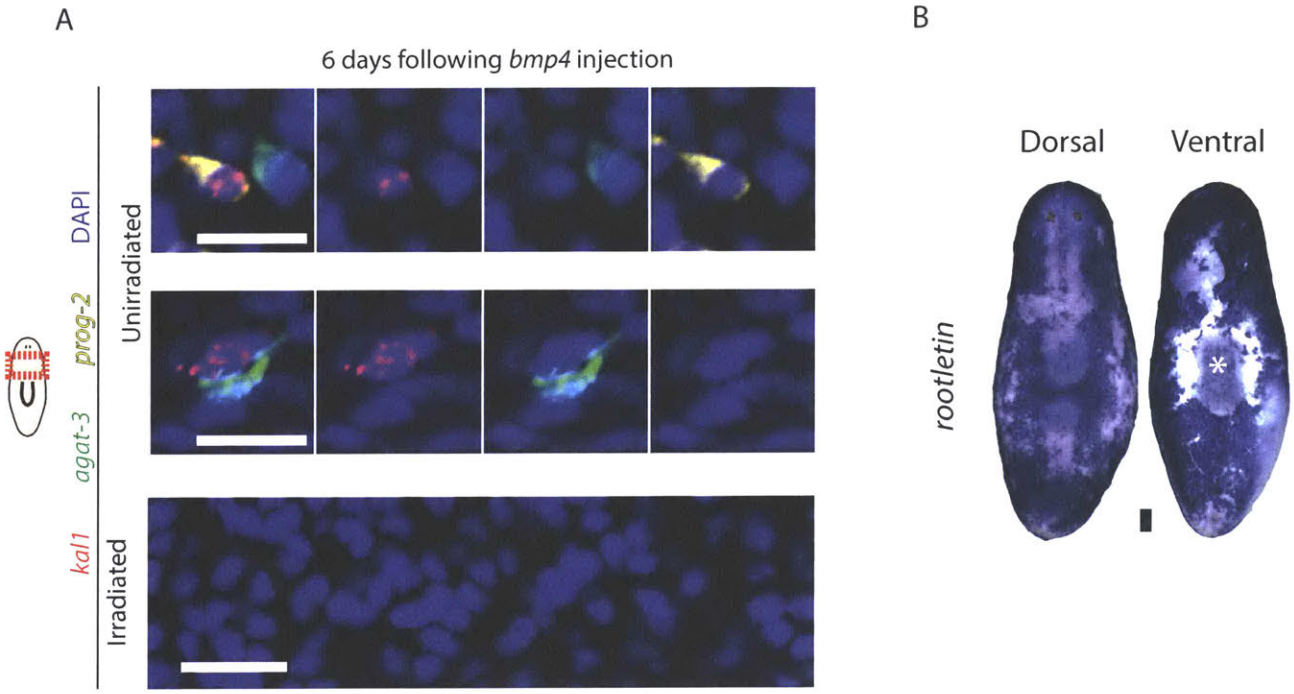

A
6 days following *bmp4* injection
DAPI
prog-2
agat-3
kal1
Unirradiated
Irradiated
B
Dorsal
Ventral
rootletin

Figure S7 related to figure 6. Ectopic expression of *kal1* following *bmp4* injection is neoblast-dependent. (A) Unirradiated and irradiated animals were injected with *bmp4* dsRNA (see Fig 6D-F), and were fixed and analyzed using FISH at day six following injections. Expression of *kal1* in epidermal progenitors on the dorsal region (*prog-2*+ or *agat-3*+ cells, top and middle panels, respectively) was detectable. Conversely, dorsal *kal1*+ cells were not detected in irradiated animals. By day six following irradiation, epidermal progenitors are completely ablated. Scale bar = 20 μm. (B) WISH of *rootletin*. Rips in the ventral epidermis (marked with an asterisk) show the lack of detection of sub-epidermal *rootletin* expression by WISH. Scale bar = 100 μm.

Tables

Table S1 related to figure 1. Overexpressed genes in planarian epidermal samples compared to whole worm controls.

Contig	Description	E-value	*in situ* pattern	Dorsal Enrichment log(fold-change)	Dorsal logCPM	Ventral Enrichment log(fold-change)	Ventral logCPM
dd_Smed_v4_11828_0_1	KIAA1875 (KIAA1875)	1.00E-	Ventral	3.73	7.51	4.87	7.51
dd_Smed_v4_6557_0_1	dynein, axonemal, heavy	0	Ventral	3.72	8.92	4.9	8.92
dd_Smed_v4_8295_0_1	NA	NA	Pan	4.59	6.96	4.7	6.96
dd_Smed_v4_5350_0_1	dynein, axonemal, heavy	0	Ventral	3.48	9.6	4.81	9.6
dd_Smed_v4_8371_0_1	protein kinase, cAMP-	2.00E-	Pan,	3.19	6.02	4.89	6.02
dd_Smed_v4_2802_0_1	NA	NA	Ventral	3.15	10.4	5.07	10.4
dd_Smed_v4_6790_0_1	dynein, axonemal, heavy	0	Ventral	3.6	9.47	4.9	9.47
dd_Smed_v4_10292_0_1	NA	NA	Ventral	2.77	6.82	4.44	6.82
dd_Smed_v4_10248_0_1	NA	4.00E-	Ventral	2.91	7.48	4.53	7.48
dd_Smed_v4_3993_0_1	otoferlin (OTOF)	8.00E-	Pan	4.54	8.39	3.9	8.39
dd_Smed_v4_4940_0_1	NA	NA	Pan	4.03	8.29	4.04	8.29
dd_Smed_v4_4586_0_1	notch-1 (notch-1)	0	Dorsal	4.14	8.4	4.09	8.4
dd_Smed_v4_3340_0_1	NA	0	Pan	4.58	9.6	3.31	9.6
dd_Smed_v4_1404_0_1	ciliary rootlet coiled-coil,	0	Ventral	3.31	11.99	4.85	11.99
dd_Smed_v4_1795_0_1	tyrosyl-tRNA synthetase	7.00E-	Pan	4.25	10.18	4.03	10.18
dd_Smed_v4_2502_0_1	NA	NA	Pan	4.09	8.15	3.68	8.15
dd_Smed_v4_4448_0_1	NA	NA	Pan	3.66	7.19	3.95	7.19
dd_Smed_v4_16204_0_1	NA	NA	Ventral	2.2	5.46	4.94	5.46
dd_Smed_v4_6979_0_1	NA	NA	Ventral	1.54	5.88	4.32	5.88
dd_Smed_v4_9419_0_1	sperm flagellar 2 (SPEF2)	1.00E-	Ventral	3.25	8.37	4.43	8.37
dd_Smed_v4_6330_0_1	NA	NA	Dorsal	3.7	7.23	1.58	7.23
dd_Smed_v4_950_0_1	Labeled: *vim-3*; vimentin	7.00E-		4.32	11.67	3.25	11.67
dd_Smed_v4_11894_0_1	tetratricopeptide repeat	3.00E-	Ventral	3.34	6.46	4.51	6.46
dd_Smed_v4_9955_0_1	natriuretic peptide receptor	0	Dorsal	3.95	6.94	0.97	6.94
dd_Smed_v4_6941_0_1	dynein, axonemal, heavy	0	Ventral	3.19	9.69	4.39	9.69
dd_Smed_v4_351_0_1	protease, serine, 12	2.00E-	Pan	4.07	13.42	4.15	13.42
dd_Smed_v4_9617_0_1	adenylate kinase 9 (AK9)	7.00E-	Ventral	3.17	8.2	4.66	8.2
dd_Smed_v4_9329_1_1	NA	NA	Ventral	3.46	7.21	3.54	7.21
dd_Smed_v4_9462_0_1	sperm-tail PG-rich repeat	4.00E-	Ventral	3.21	6.91	4.58	6.91
dd_Smed_v4_10281_0_1	TNF receptor-associated	3.00E-	Pan	3.74	5.37	4.27	5.37
dd_Smed_v4_6538_0_1	dynein, axonemal, heavy	0	Ventral	2.86	8.41	3.95	8.41
dd_Smed_v4_1688_0_1	ATPase, Na+/K+ transporting,	4.00E-	Dorsal	4.08	8.94	3.68	8.94
dd_Smed_v4_5341_0_1	NA	6.00E-	Ventral	2.06	7.65	4.02	7.65
dd_Smed_v4_10058_0_1	RecName:	1.00E-	Ventral	3.75	7.49	3.92	7.49
dd_Smed_v4_9904_0_1	solute carrier family 24	8.00E-	Dorsal	4.01	6.86	3.05	6.86
dd_Smed_v4_11909_0_1	EF-hand domain family,	6.00E-	Ventral	1.05	5.59	4.08	5.59
dd_Smed_v4_11398_0_1	NA	NA	Pan	4.16	5.17	4.27	5.17
dd_Smed_v4_5255_0_1	NIMA-related kinase 3 (NEK3)	9.00E-	Ventral	4.57	9	4.41	9
dd_Smed_v4_9780_0_1	NA	1.00E-	Ventral	2.5	6.4	3.76	6.4
dd_Smed_v4_6321_0_1	spermatogenesis associated	3.00E-	Ventral	1.28	7.65	3.98	7.65
dd_Smed_v4_12161_0_1	WD repeat domain 93	8.00E-	ND	2.63	5.89	3.7	5.89
dd_Smed_v4_13549_0_1	armadillo repeat containing 4	2.00E-	Ventral	2.49	5.47	4.24	5.47
dd_Smed_v4_12591_0_1	dual-specificity tyrosine-(Y)-	6.00E-	Dorsal	4.62	6.05	0.92	6.05
dd_Smed_v4_10647_0_1	F-box and leucine-rich repeat	2.00E-	Dorsal	3.5	6.28	3.46	6.28
dd_Smed_v4_1507_0_1	NA	NA	Pan	3.93	9.13	3.25	9.13
dd_Smed_v4_9671_0_1	CEP78	0	Ventral	3.57	6.62	3.68	6.62
dd_Smed_v4_7765_0_1	transmembrane protein 67	1.00E-	Ventral	3.44	6.82	3.74	6.82
dd_Smed_v4_11049_0_1	NA	NA	Pan	2.18	5.17	3.87	5.17
dd_Smed_v4_7161_0_1	NA	NA	Pan	3.33	6.98	3.38	6.98
dd_Smed_v4_5570_0_1	NA	NA	Pan, no	3.37	6.55	3.82	6.55
dd_Smed_v4_6927_0_1	ADP-ribosylation factor 3	2.00E-	Dorsal	3.94	5.87	2.54	5.87
dd_Smed_v4_8104_0_1	SRY (sex determining region	5.00E-	Dorsal	3.72	6.56	2.57	6.56
dd_Smed_v4_7087_0_1	sorting nexin 33 (SNX33)	2.00E-	Pan	3.52	7.41	3.62	7.41
dd_Smed_v4_13531_0_1	NA	NA	Pan	3.34	5.57	3.61	5.57
dd_Smed_v4_11693_0_1	T-box 2 (TBX2)	4.00E-	Dorsal	3.81	6.01	0.01	6.01
dd_Smed_v4_4986_0_1	caspase 3, apoptosis-related	6.00E-	Dorsal	3.64	6.75	2.16	6.75
dd_Smed_v4_2937_0_1	NA	NA	Pan	3.22	8.43	3.74	8.43
dd_Smed_v4_1881_0_1	NA	NA	Pan /	3.72	8.14	2.27	8.14
dd_Smed_v4_10121_0_1	NA	1.00E-	Pan	3.37	6.69	3.6	6.69
dd_Smed_v4_12038_0_1	NA	NA	Dorsal	3.97	4.99	0.45	4.99
dd_Smed_v4_7496_0_1	fibronectin type III and	2.00E-	ND	1.88	5.01	3.65	5.01
dd_Smed_v4_8275_0_1	mitogen-activated protein	2.00E-	Sparse	3.46	6.42	3.61	6.42

dd_Smed_v4_10089_0_1	NA	NA	Dorsal	4.4	5.51	1.51	5.51
dd_Smed_v4_8536_0_1	NA	NA	Dorsal	4.05	5.77	2.19	5.77
dd_Smed_v4_7995_0_1	epithelial splicing regulatory	3.00E-	Dorsal	2.74	6.35	0.09	6.35
dd_Smed_v4_10032_0_1	WD repeat domain 60	4.00E-	ND	3.63	6.64	4.1	6.64
dd_Smed_v4_10486_0_1	NA	NA	Dorsal	3.49	5.6	0.59	5.6
dd_Smed_v4_9968_0_1	NA	NA	Pan	3.25	5.48	3.51	5.48
dd_Smed_v4_16524_0_1	msh2 protein	0	Dorsal,	4.28	2.99	1.11	2.99
dd_Smed_v4_17564_0_1	NA	NA	Ventral	0.41	4.14	3.81	4.14
dd_Smed_v4_14341_0_1	NA	NA	Dorsal	3.97	4.34	0.13	4.34
dd_Smed_v4_20693_0_1	NA	NA	Pan	4.76	2.07	1.97	2.07
dd_Smed_v4_11151_0_1	NA	NA	ND	3.74	3.27	0.82	3.27
dd_Smed_v4_6559_0_1	ATPase, Na+/K+ transporting,	7.00E-	Dorsal	3.73	6.77	1.67	6.77
dd_Smed_v4_2231_0_1	NA	NA	Pan	3.69	7.85	4.06	7.85
dd_Smed_v4_15280_0_1	NA	NA	Dorsal	4.08	3.74	0.98	3.74
dd_Smed_v4_10673_0_1	*ovo-2*	Describ	Midline	3.41	3.74	0.77	3.74
dd_Smed_v4_10380_0_1	unc-5 homolog B (C. elegans)	4.00E-	Dorsal,	2.22	5.27	-0.67	5.27
dd_Smed_v4_10593_0_1	phosphodiesterase 8A	8.00E-	Ventral	0.79	6.12	3.61	6.12
dd_Smed_v4_16438_0_1	NA	NA	Dorsal	3.64	6	0.58	6
dd_Smed_v4_5936_0_1	chromosome 2 open reading	1.00E-	Ventral	2.2	6.92	2.56	6.92
dd_Smed_v4_14635_0_1	forkhead box J1-like protein 1	0	Ventral	1.35	3.84	3.75	3.84
dd_Smed_v4_12869_0_1	NA	NA	Dorsal	2.8	4	-0.28	4
dd_Smed_v4_2828_0_1	ATPase, Na+/K+ transporting,	0	Pan	2.13	8.38	1.79	8.38
dd_Smed_v4_13733_0_1	NA	NA	Dorsal	3.43	3.58	-0.08	3.58
dd_Smed_v4_15516_0_1	otxb-like	0	Dorsal	1.25	4.34	4.91	4.34
dd_Smed_v4_7735_0_1	phosphodiesterase 4C,	4.00E-	ND	0.43	7.19	3.33	7.19
dd_Smed_v4_15542_0_1	NA	6.00E-	ND	0.25	4.64	3.13	4.64
dd_Smed_v4_15224_0_1	NA	NA	Ventral	0.8	4.29	4.3	4.29
dd_Smed_v4_16647_0_1	HECT and RLD domain	1.00E-	Ventral	1.52	4.16	2.88	4.16
dd_Smed_v4_8559_0_1	NA	NA	ND	3.81	3.97	0.55	3.97
dd_Smed_v4_24171_0_1	NA	NA	ND	0.95	3.44	4.27	3.44
dd_Smed_v4_11840_0_1	*PRDM1-1*	Describ	Dorsal	3.46	4.21	0.02	4.21
dd_Smed_v4_12485_0_1	deoxyhypusine synthase	2.00E-	Ventral	0.71	4.37	3.64	4.37
dd_Smed_v4_13875_0_1	NA	NA	Broad	3.83	3.54	0.89	3.54
dd_Smed_v4_7346_0_1	NA	NA	Ventral	0.33	5.05	3.85	5.05
dd_Smed_v4_12058_0_1	NA	7.00E-	Ventral	1.36	6.32	4.95	6.32
dd_Smed_v4_5871_0_1	NA	NA	Dorsal	2.04	6.27	1.18	6.27
dd_Smed_v4_3391_0_1	nephrosis 1, congenital,	7.00E-	ND	2.12	7.42	-0.09	7.42
dd_Smed_v4_14685_0_1	calcium/calmodulin-	3.00E-	ND	1.13	3.02	3.19	3.02
dd_Smed_v4_11062_0_1	bambi	0	Dorsal	2.83	4.19	-0.87	4.19
dd_Smed_v4_13154_0_1	NA	NA	Ventral	0.77	3.99	4.26	3.99
dd_Smed_v4_13623_0_1	NA	NA	Ventral	0.69	5.99	4.07	5.99
dd_Smed_v4_11767_0_1	natriuretic peptide receptor	0	Ventral	0.73	6.4	4.11	6.4
dd_Smed_v4_13304_0_1	phosphodiesterase 9A	6.00E-	Dorsal	2.49	4.14	0.1	4.14
dd_Smed_v4_17565_0_1	cyclic nucleotide binding	1.00E-	Ventral	0.81	4.76	4.26	4.76
dd_Smed_v4_1756_0_1	TBC1 domain family, member	2.00E-	Pan	1.59	7.25	1.17	7.25
dd_Smed_v4_7222_0_1	calmodulin-like 3 (CALML3)	7.00E-	Ventral	0.67	5.27	4.33	5.27
dd_Smed_v4_4890_0_1	family with sequence	5.00E-	Pan	1.64	5.99	0.8	5.99
dd_Smed_v4_12689_0_1	doublecortin-like kinase 3	4.00E-	Ventral	0.86	5.96	4.31	5.96
dd_Smed_v4_6024_0_1	chromosome 2 open reading	2.00E-	Ventral	-0.61	5.23	2.82	5.23
dd_Smed_v4_17655_0_1	NA	NA	Ventral	1.11	4.05	4.6	4.05
dd_Smed_v4_11406_0_1	regulatory factor X, 4	2.00E-	Ventral	1.13	5.91	4.55	5.91
dd_Smed_v4_10585_0_1	unc-5 homolog D (C. elegans)	1.00E-	Dorsal,	1.56	5.15	-0.09	5.15
dd_Smed_v4_14679_0_1	NA	NA	Ventral	0.61	4.59	4.19	4.59
dd_Smed_v4_14268_0_1	NA	NA	ND	0.77	4.18	4.37	4.18
dd_Smed_v4_7249_0_1	NA	NA	Ventral	0.08	5.82	3.47	5.82
dd_Smed_v4_6261_0_1	NA	NA	Ventral	0.12	6.29	2.61	6.29
dd_Smed_v4_14906_0_1	NA	NA	Ventral	0.77	4.08	4.38	4.08
dd_Smed_v4_13850_0_1	otoferlin (OTOF)	2.00E-	Ventral	0.2	2.65	3.24	2.65
dd_Smed_v4_14395_0_1	calmodulin 1 (phosphorylase	1.00E-	Ventral	-0.18	3.8	3.44	3.8
dd_Smed_v4_6567_0_1	RNA binding protein, fox-1	5.00E-	Ventral	-1.05	4.29	1.53	4.29
dd_Smed_v4_12435_0_1	spermidine/spermine N1-	3.00E-	Dorsal	2.31	3.14	-0.33	3.14
dd_Smed_v4_15092_0_1	NA	NA	Dorsal	2.99	2.95	-0.15	2.95
dd_Smed_v4_11357_0_1	ras-related C3 botulinum	3.00E-	ND	1.46	4.83	3.94	4.83

Table S1 related to figure 1. Overexpressed genes in planarian epidermal samples compared to whole worm controls.

Differentially expressed genes were called using EdgeR v3.6.8 [41], requiring minimal fold-change of 2 and maximal FDR of 1E-4. Tabulated are the 125 genes enriched in epidermal samples as compared to intact worm controls that were selected for *in situ* validation. Indicated in the table is the contig ID, best BLAST hit, e-value, *in situ* pattern, dorsal enrichment log(fold-change), dorsal logCPM, ventral enrichment log(fold-change) and ventral logCPM.

Table S2 related to figure 1. Cluster assignment of single cell RNA sequencing samples.

Cell ID	# expressed genes	Sample or source	Maturation stage	# mapped mRNA reads	Chromosome content (by Hoechst stain)	Data source
D14.100252	7747	Wurtzel et al, Developmental	ζNeoblast	1395980	4C	Wurtzel,
D14.100313	7735	Wurtzel et al, Developmental	ζNeoblast	660675	4C	Wurtzel,
D14.100312	7725	Wurtzel et al, Developmental	ζNeoblast	1230210	4C	Wurtzel,
D14.100259	7720	Wurtzel et al, Developmental	ζNeoblast	1298222	4C	Wurtzel,
D14.100265	7522	Wurtzel et al, Developmental	ζNeoblast	1015778	4C	Wurtzel,
D14.100268	7509	Wurtzel et al, Developmental	ζNeoblast	1209895	4C	Wurtzel,
D14.100244	7203	Wurtzel et al, Developmental	ζNeoblast	727033	4C	Wurtzel,
D14.100240	7074	Wurtzel et al, Developmental	ζNeoblast	695009	4C	Wurtzel,
D15.101157	6803	Dorsal	ζNeoblast	401625	2C	This
D14.100247	6412	Wurtzel et al, Developmental	ζNeoblast	784661	4C	Wurtzel,
D15.101239	7680	Dorsal	Early Stage	1254504	2C	This
D15.101165	7153	Dorsal	Early Stage	1979214	2C	This
D15.101432	7114	Ventral	Early Stage	636201	2C	This
D15.101203	7101	Dorsal	Early Stage	1005202	2C	This
D15.101437	7090	Ventral	Early Stage	1044110	2C	This
D15.101237	7014	Dorsal	Early Stage	1519026	2C	This
D15.101385	6676	Lateral	Early Stage	349587	2C	This
D14.30300	6657	Wurtzel et al, Developmental	Early Stage	2521975	2C	Wurtzel,
D15.101387	6637	Lateral	Early Stage	206125	2C	This
D15.101236	6571	Dorsal	Early Stage	1106266	2C	This
D15.101434	6521	Ventral	Early Stage	598265	2C	This
D14.400260	6521	Wurtzel et al, Developmental	Early Stage	1109069	2C	Wurtzel,
D15.101196	6512	Dorsal	Early Stage	628001	2C	This
D15.101184	6508	Dorsal	Early Stage	615026	2C	This
D15.101167	6494	Dorsal	Early Stage	1460474	2C	This
D15.101409	6469	Lateral	Early Stage	900781	2C	This
D15.101267	6460	Ventral	Early Stage	938152	2C	This
D15.101205	6361	Dorsal	Early Stage	778484	2C	This
D15.101269	6351	Ventral	Early Stage	763784	2C	This
D15.101261	6340	Ventral	Early Stage	1183491	2C	This
D15.101348	6301	Dorsal	Early Stage	698222	2C	This
D15.101413	6299	Ventral	Early Stage	555027	2C	This
D15.101195	6276	Dorsal	Early Stage	543667	2C	This
D15.101414	6266	Ventral	Early Stage	563443	2C	This
D15.101405	6250	Lateral	Early Stage	1046519	2C	This
D15.101407	6238	Lateral	Early Stage	984149	2C	This
D15.101175	6235	Dorsal	Early Stage	340459	2C	This
D15.101346	6223	Dorsal	Early Stage	630976	2C	This
D15.101396	6202	Lateral	Early Stage	590811	2C	This
D14.30308	6198	Wurtzel et al, Developmental	Early Stage	2344715	2C	Wurtzel,
D15.101166	6151	Dorsal	Early Stage	460823	2C	This
D15.101153	6148	Dorsal	Early Stage	922212	2C	This
D15.101389	6147	Lateral	Early Stage	582432	2C	This
D15.101240	6146	Dorsal	Early Stage	811870	2C	This
D15.101209	6127	Dorsal	Early Stage	744492	2C	This
D15.101422	6105	Ventral	Early Stage	427872	2C	This
D14.400193	6090	Wurtzel et al, Developmental	Early Stage	698352	2C	Wurtzel,
D15.101191	6085	Dorsal	Early Stage	568410	2C	This
D15.101189	6081	Dorsal	Early Stage	1634890	2C	This
D15.101202	6078	Dorsal	Early Stage	1427273	2C	This
D15.101174	6073	Dorsal	Early Stage	439256	2C	This
D15.101350	6063	Dorsal	Early Stage	693903	2C	This
D15.101170	6058	Dorsal	Early Stage	417224	2C	This
D15.101345	6022	Dorsal	Early Stage	505571	2C	This
D15.101392	6005	Lateral	Early Stage	449606	2C	This
D15.101369	5992	Lateral	Early Stage	1067109	2C	This
D15.101154	5990	Dorsal	Early Stage	973357	2C	This
D15.101217	5985	Dorsal	Early Stage	669645	2C	This
D15.101402	5979	Lateral	Early Stage	701023	2C	This
D15.101259	5962	Ventral	Early Stage	1388348	2C	This
D15.101248	5956	Dorsal	Early Stage	541771	2C	This
D14.400218	5913	Wurtzel et al, Developmental	Early Stage	394812	2C	Wurtzel,
D15.101281	5884	Ventral	Early Stage	466262	2C	This
D15.101273	5880	Ventral	Early Stage	539858	2C	This
D15.101431	5867	Ventral	Early Stage	612610	2C	This
D14.400294	5863	Wurtzel et al, Developmental	Early Stage	833039	2C	Wurtzel,
D15.101155	5862	Dorsal	Early Stage	794469	2C	This

D15.101183	5831	Dorsal	Early Stage	387185	2C	This
D15.101426	5816	Ventral	Early Stage	572167	2C	This
D15.101247	5794	Dorsal	Early Stage	360428	2C	This
D14.400205	5772	Wurtzel et al, Developmental	Early Stage	491414	2C	Wurtzel,
D15.101263	5770	Ventral	Early Stage	482065	2C	This
D15.101271	5762	Ventral	Early Stage	504633	2C	This
D15.101193	5758	Dorsal	Early Stage	405535	2C	This
D15.101433	5757	Ventral	Early Stage	478566	2C	This
D15.101192	5746	Dorsal	Early Stage	330406	2C	This
D15.101395	5737	Lateral	Early Stage	458618	2C	This
D15.101275	5717	Ventral	Early Stage	650612	2C	This
D15.101219	5700	Dorsal	Early Stage	341396	2C	This
D15.101265	5687	Ventral	Early Stage	318933	2C	This
D15.101270	5681	Ventral	Early Stage	851063	2C	This
D15.101186	5675	Dorsal	Early Stage	407439	2C	This
D15.101238	5643	Dorsal	Early Stage	464225	2C	This
D15.101182	5635	Dorsal	Early Stage	563999	2C	This
D15.101406	5605	Lateral	Early Stage	227270	2C	This
D15.101201	5593	Dorsal	Early Stage	921536	2C	This
D15.101358	5549	Lateral	Early Stage	472297	2C	This
D15.101197	5536	Dorsal	Early Stage	326514	2C	This
D15.101282	5507	Ventral	Early Stage	343410	2C	This
D15.101163	5489	Dorsal	Early Stage	584652	2C	This
D15.101349	5468	Dorsal	Early Stage	413173	2C	This
D15.101425	5454	Ventral	Early Stage	341059	2C	This
D15.101171	5423	Dorsal	Early Stage	1009162	2C	This
D15.101199	5414	Dorsal	Early Stage	186114	2C	This
D15.101227	5377	Dorsal	Early Stage	548481	2C	This
D15.101353	5364	Dorsal	Early Stage	758276	2C	This
D14.30348	5356	Wurtzel et al, Developmental	Early Stage	654239	2C	Wurtzel,
D15.101268	5345	Ventral	Early Stage	292721	2C	This
D14.400360	5340	Wurtzel et al, Developmental	Early Stage	288537	2C	Wurtzel,
D15.101160	5316	Dorsal	Early Stage	783502	2C	This
D15.101286	5294	Ventral	Early Stage	817946	2C	This
D15.101284	5274	Ventral	Early Stage	247205	2C	This
D15.101285	5267	Ventral	Early Stage	566047	2C	This
D15.101215	5262	Dorsal	Early Stage	465638	2C	This
D15.101427	5247	Ventral	Early Stage	374906	2C	This
D15.101272	5224	Ventral	Early Stage	370937	2C	This
D15.101245	5219	Dorsal	Early Stage	361713	2C	This
D15.101262	5211	Ventral	Early Stage	438330	2C	This
D15.101168	5207	Dorsal	Early Stage	965135	2C	This
D15.101347	5190	Dorsal	Early Stage	418963	2C	This
D15.101383	5189	Lateral	Early Stage	351331	2C	This
D15.101439	5185	Ventral	Early Stage	267592	2C	This
D15.101244	5178	Dorsal	Early Stage	365590	2C	This
D15.101214	5169	Dorsal	Early Stage	922469	2C	This
D15.101228	5159	Dorsal	Early Stage	318751	2C	This
D15.101194	5156	Dorsal	Early Stage	179899	2C	This
D15.101416	5153	Ventral	Early Stage	325873	2C	This
D15.101412	5135	Ventral	Early Stage	480101	2C	This
D15.101242	5131	Dorsal	Early Stage	269182	2C	This
D15.101351	5125	Dorsal	Early Stage	363838	2C	This
D15.101355	5110	Dorsal	Early Stage	450111	2C	This
D14.30310	5104	Wurtzel et al, Developmental	Early Stage	489490	2C	Wurtzel,
D15.101234	5075	Dorsal	Early Stage	255969	2C	This
D15.101382	5074	Lateral	Early Stage	290968	2C	This
D15.101161	5051	Dorsal	Early Stage	282748	2C	This
D15.101159	5041	Dorsal	Early Stage	625059	2C	This
D15.101260	5028	Ventral	Early Stage	701923	2C	This
D15.101423	5019	Ventral	Early Stage	340908	2C	This
D15.101276	5016	Ventral	Early Stage	1291823	2C	This
D15.101403	5007	Lateral	Early Stage	204431	2C	This
D15.101210	5005	Dorsal	Early Stage	480129	2C	This
D15.101266	4972	Ventral	Early Stage	390767	2C	This
D15.101283	4964	Ventral	Early Stage	268162	2C	This
D15.101172	4962	Dorsal	Early Stage	257915	2C	This
D15.101213	4924	Dorsal	Early Stage	936907	2C	This
D15.101216	4906	Dorsal	Early Stage	399552	2C	This
D15.101231	4883	Dorsal	Early Stage	319885	2C	This
D14.400269	4839	Wurtzel et al, Developmental	Early Stage	69093	2C	Wurtzel,
D15.101381	4837	Lateral	Early Stage	422569	2C	This
D15.101158	4797	Dorsal	Early Stage	316059	2C	This
D15.101361	4788	Lateral	Early Stage	272634	2C	This

D15.101280	4771	Ventral	Early Stage	475649	2C	This
D15.101223	4767	Dorsal	Early Stage	231544	2C	This
D15.101371	4759	Lateral	Early Stage	341743	2C	This
D15.101169	4728	Dorsal	Early Stage	215904	2C	This
D15.101206	4699	Dorsal	Early Stage	428977	2C	This
D15.101211	4687	Dorsal	Early Stage	315340	2C	This
D15.101221	4666	Dorsal	Early Stage	275016	2C	This
D15.101212	4637	Dorsal	Early Stage	181019	2C	This
D15.101430	4635	Ventral	Early Stage	429205	2C	This
D15.101399	4632	Lateral	Early Stage	475016	2C	This
D15.101428	4625	Ventral	Early Stage	243156	2C	This
D15.101185	4616	Dorsal	Early Stage	325489	2C	This
D15.101187	4611	Dorsal	Early Stage	293166	2C	This
D15.101229	4604	Dorsal	Early Stage	300741	2C	This
D15.101372	4592	Lateral	Early Stage	453987	2C	This
D15.101222	4541	Dorsal	Early Stage	196793	2C	This
D15.101380	4502	Lateral	Early Stage	226808	2C	This
D15.101218	4462	Dorsal	Early Stage	248648	2C	This
D14.30319	4460	Wurtzel et al, Developmental	Early Stage	452291	2C	Wurtzel,
D14.30302	4439	Wurtzel et al, Developmental	Early Stage	375541	2C	Wurtzel,
D15.101156	4387	Dorsal	Early Stage	345669	2C	This
D15.101366	4380	Lateral	Early Stage	299912	2C	This
D15.101249	4293	Ventral	Early Stage	402241	2C	This
D15.101220	4284	Dorsal	Early Stage	299972	2C	This
D15.101367	4258	Lateral	Early Stage	231930	2C	This
D15.101226	4256	Dorsal	Early Stage	386530	2C	This
D15.101279	4253	Ventral	Early Stage	267928	2C	This
D15.101295	4196	Ventral	Early Stage	465220	2C	This
D15.101379	4145	Lateral	Early Stage	174472	2C	This
D15.101356	4126	Lateral	Early Stage	59195	2C	This
D15.101190	4125	Dorsal	Early Stage	1226824	2C	This
D15.101250	4083	Ventral	Early Stage	267178	2C	This
D15.101264	4082	Ventral	Early Stage	120588	2C	This
D14.30447	4068	Wurtzel et al, Developmental	Early Stage	606990	2C	Wurtzel,
D15.101278	4060	Ventral	Early Stage	222554	2C	This
D15.101384	4045	Lateral	Early Stage	198841	2C	This
D15.101177	4031	Dorsal	Early Stage	625477	2C	This
D15.101200	3996	Dorsal	Early Stage	81650	2C	This
D15.101404	3954	Lateral	Early Stage	263821	2C	This
D15.101251	3896	Ventral	Early Stage	295945	2C	This
D14.30353	3872	Wurtzel et al, Developmental	Early Stage	385200	2C	Wurtzel,
D15.101424	3837	Ventral	Early Stage	214872	2C	This
D14.30418	3785	Wurtzel et al, Developmental	Early Stage	196875	2C	Wurtzel,
D14.30320	3750	Wurtzel et al, Developmental	Early Stage	287863	2C	Wurtzel,
D15.101232	3686	Dorsal	Early Stage	279718	2C	This
D14.400267	3441	Wurtzel et al, Developmental	Early Stage	42290	2C	Wurtzel,
D14.30424	3429	Wurtzel et al, Developmental	Early Stage	770397	2C	Wurtzel,
D15.101288	3423	Ventral	Early Stage	526756	2C	This
D15.101287	3405	Ventral	Early Stage	552012	2C	This
D15.101376	3359	Lateral	Early Stage	109415	2C	This
D15.101291	3074	Ventral	Early Stage	423296	2C	This
D15.101255	3049	Ventral	Early Stage	220389	2C	This
D15.101292	2959	Ventral	Early Stage	190240	2C	This
D15.101178	2755	Dorsal	Early Stage	290093	2C	This
D15.101290	2685	Ventral	Early Stage	354260	2C	This
D15.101294	2585	Ventral	Early Stage	337305	2C	This
D15.101357	2429	Lateral	Early Stage	287956	2C	This
D15.101293	2415	Ventral	Early Stage	476039	2C	This
D15.101180	2414	Dorsal	Early Stage	345937	2C	This
D15.101235	2395	Dorsal	Early Stage	208946	2C	This
D15.101274	2329	Ventral	Early Stage	799073	2C	This
D15.101252	2262	Ventral	Early Stage	228589	2C	This
D15.101198	2191	Dorsal	Early Stage	288736	2C	This
D15.101352	2168	Dorsal	Early Stage	214020	2C	This
D15.101254	1899	Ventral	Early Stage	211985	2C	This
D15.101289	1568	Ventral	Early Stage	199410	2C	This
D15.101296	1275	Ventral	Early Stage	85733	2C	This
D15.101179	1138	Dorsal	Early Stage	257309	2C	This
D14.400247	7439	Wurtzel et al, Developmental	Late Stage	1202187	2C	Wurtzel,
D14.400259	7286	Wurtzel et al, Developmental	Late Stage	1526129	2C	Wurtzel,
D14.400223	6960	Wurtzel et al, Developmental	Late Stage	1090917	2C	Wurtzel,
D15.101429	6806	Ventral	Late Stage	929403	2C	This
D15.101438	6760	Ventral	Late Stage	630594	2C	This
D15.101415	6725	Ventral	Late Stage	576435	2C	This

D15.101411	6652	Ventral	Late Stage	969937	2C	This
D15.101258	6643	Ventral	Late Stage	540088	2C	This
D14.30332	6579	Wurtzel et al, Developmental	Late Stage	1435796	2C	Wurtzel,
D14.400195	6476	Wurtzel et al, Developmental	Late Stage	386109	2C	Wurtzel,
D14.400236	6441	Wurtzel et al, Developmental	Late Stage	884238	2C	Wurtzel,
D15.101360	6390	Lateral	Late Stage	1204745	2C	This
D14.400250	6350	Wurtzel et al, Developmental	Late Stage	452503	2C	Wurtzel,
D15.101391	6339	Lateral	Late Stage	267253	2C	This
D15.101421	6327	Ventral	Late Stage	611724	2C	This
D15.101418	6322	Ventral	Late Stage	678212	2C	This
D14.400217	6294	Wurtzel et al, Developmental	Late Stage	312490	2C	Wurtzel,
D14.30340	6279	Wurtzel et al, Developmental	Late Stage	1093304	2C	Wurtzel,
D15.101436	6205	Ventral	Late Stage	398611	2C	This
D14.400199	6186	Wurtzel et al, Developmental	Late Stage	1565348	2C	Wurtzel,
D15.101388	6151	Lateral	Late Stage	558894	2C	This
D14.30343	6120	Wurtzel et al, Developmental	Late Stage	1056784	2C	Wurtzel,
D14.400302	6048	Wurtzel et al, Developmental	Late Stage	696479	2C	Wurtzel,
D14.400226	5967	Wurtzel et al, Developmental	Late Stage	179849	2C	Wurtzel,
D15.101363	5926	Lateral	Late Stage	649225	2C	This
D14.400256	5910	Wurtzel et al, Developmental	Late Stage	227182	2C	Wurtzel,
D14.400208	5799	Wurtzel et al, Developmental	Late Stage	480107	2C	Wurtzel,
D14.400248	5760	Wurtzel et al, Developmental	Late Stage	678521	2C	Wurtzel,
D14.400181	5707	Wurtzel et al, Developmental	Late Stage	266318	2C	Wurtzel,
D14.400207	5696	Wurtzel et al, Developmental	Late Stage	360706	2C	Wurtzel,
D15.101398	5664	Lateral	Late Stage	559523	2C	This
D15.101420	5629	Ventral	Late Stage	476976	2C	This
D14.400286	5598	Wurtzel et al, Developmental	Late Stage	573611	2C	Wurtzel,
D14.30336	5598	Wurtzel et al, Developmental	Late Stage	417730	2C	Wurtzel,
D14.20001	5564	Wurtzel et al, Developmental	Late Stage	545943	2C	Wurtzel,
D14.400319	5466	Wurtzel et al, Developmental	Late Stage	2116883	2C	Wurtzel,
D14.400241	5429	Wurtzel et al, Developmental	Late Stage	196682	2C	Wurtzel,
D14.30341	5350	Wurtzel et al, Developmental	Late Stage	2564557	2C	Wurtzel,
D14.400291	5344	Wurtzel et al, Developmental	Late Stage	720909	2C	Wurtzel,
D15.101230	5289	Dorsal	Late Stage	292573	2C	This
D14.30338	5265	Wurtzel et al, Developmental	Late Stage	565476	2C	Wurtzel,
D14.20075	5130	Wurtzel et al, Developmental	Late Stage	295227	2C	Wurtzel,
D14.20007	5110	Wurtzel et al, Developmental	Late Stage	773499	2C	Wurtzel,
D14.20053	5105	Wurtzel et al, Developmental	Late Stage	266302	2C	Wurtzel,
D14.400219	5080	Wurtzel et al, Developmental	Late Stage	104471	2C	Wurtzel,
D14.20064	5048	Wurtzel et al, Developmental	Late Stage	366859	2C	Wurtzel,
D14.20015	5042	Wurtzel et al, Developmental	Late Stage	315521	2C	Wurtzel,
D15.101378	5002	Lateral	Late Stage	272829	2C	This
D14.20073	4973	Wurtzel et al, Developmental	Late Stage	343391	2C	Wurtzel,
D14.30337	4930	Wurtzel et al, Developmental	Late Stage	306230	2C	Wurtzel,
D14.400215	4920	Wurtzel et al, Developmental	Late Stage	303263	2C	Wurtzel,
D14.400263	4903	Wurtzel et al, Developmental	Late Stage	226723	2C	Wurtzel,
D14.20051	4852	Wurtzel et al, Developmental	Late Stage	292914	2C	Wurtzel,
D14.20055	4756	Wurtzel et al, Developmental	Late Stage	233830	2C	Wurtzel,
D14.30481	4481	Wurtzel et al, Developmental	Late Stage	370773	2C	Wurtzel,
D14.400365	4437	Wurtzel et al, Developmental	Late Stage	173084	2C	Wurtzel,
D15.101377	4390	Lateral	Late Stage	335280	2C	This
D15.101368	4351	Lateral	Late Stage	225595	2C	This
D14.30479	4347	Wurtzel et al, Developmental	Late Stage	284221	2C	Wurtzel,
D14.30494	4286	Wurtzel et al, Developmental	Late Stage	324759	2C	Wurtzel,
D14.20067	4237	Wurtzel et al, Developmental	Late Stage	179869	2C	Wurtzel,
D14.30490	4142	Wurtzel et al, Developmental	Late Stage	358468	2C	Wurtzel,
D14.400350	4019	Wurtzel et al, Developmental	Late Stage	428672	2C	Wurtzel,
D14.20024	3611	Wurtzel et al, Developmental	Late Stage	902731	2C	Wurtzel,
D14.20042	3098	Wurtzel et al, Developmental	Late Stage	45051	2C	Wurtzel,
D14.400179	2708	Wurtzel et al, Developmental	Late Stage	30804	2C	Wurtzel,
D14.20048	2642	Wurtzel et al, Developmental	Late Stage	909206	2C	Wurtzel,
D14.400258	7357	Wurtzel et al, Developmental	Later Stage	1136265	2C	Wurtzel,
D14.400214	6388	Wurtzel et al, Developmental	Later Stage	373310	2C	Wurtzel,
D14.400253	6268	Wurtzel et al, Developmental	Later Stage	364510	2C	Wurtzel,
D14.400270	6219	Wurtzel et al, Developmental	Later Stage	447686	2C	Wurtzel,
D14.400224	5939	Wurtzel et al, Developmental	Later Stage	1201226	2C	Wurtzel,
D14.400209	5895	Wurtzel et al, Developmental	Later Stage	423936	2C	Wurtzel,
D14.400311	5551	Wurtzel et al, Developmental	Later Stage	537911	2C	Wurtzel,
D14.400255	5526	Wurtzel et al, Developmental	Later Stage	167351	2C	Wurtzel,
D14.400300	5480	Wurtzel et al, Developmental	Later Stage	825229	2C	Wurtzel,
D14.400287	5215	Wurtzel et al, Developmental	Later Stage	635441	2C	Wurtzel,
D14.400238	5122	Wurtzel et al, Developmental	Later Stage	167645	2C	Wurtzel,
D14.30362	4970	Wurtzel et al, Developmental	Later Stage	713829	2C	Wurtzel,
D14.400178	4441	Wurtzel et al, Developmental	Later Stage	113353	2C	Wurtzel,

D14.400180	3854	Wurtzel et al, Developmental	Later Stage	37975	2C	Wurtzel,
D14.20028	5484	Wurtzel et al, Developmental	Non-ciliated	392979	2C	Wurtzel,
D14.20046	4844	Wurtzel et al, Developmental	Non-ciliated	442056	2C	Wurtzel,
D14.20021	3044	Wurtzel et al, Developmental	Non-ciliated	584536	2C	Wurtzel,
D14.20027	4257	Wurtzel et al, Developmental	Ciliated	558827	2C	Wurtzel,
D14.20066	3951	Wurtzel et al, Developmental	Ciliated	889986	2C	Wurtzel,
D14.20003	2785	Wurtzel et al, Developmental	Ciliated	419477	2C	Wurtzel,
D14.20030	2567	Wurtzel et al, Developmental	Ciliated	346795	2C	Wurtzel,
D15.101397	6940	Lateral	Dorsal ventral boundary	774182	2C	This
D15.101401	6763	Lateral	Dorsal ventral boundary	632917	2C	This
D15.101400	6026	Lateral	Dorsal ventral boundary	636282	2C	This
D15.101393	5981	Lateral	Dorsal ventral boundary	600761	2C	This
D15.101370	5612	Lateral	Dorsal ventral boundary	453682	2C	This
D15.101408	4727	Lateral	Dorsal ventral boundary	1012775	2C	This

Table S2 related to figure 1. Cluster assignment of single cell RNA sequencing samples.

Tabulated are all of the cells used for single-cell sequencing (SCS) analysis. SCS gene expression data was visualized by t-distributed stochastic neighbor embedding (tSNE) [30], and the data was clustered using the Seurat package [31] (Fig 1E; Table S2; Methods). The clustering analysis recapitulated known epidermal lineage cell states [20], here defined as: "ζNeoblast", "Early Stage", "Late Stage", "Later Stage", "Non-ciliated", "Ciliated", and "Dorsal ventral boundary." The table displays information about each cell, including the experimental source of the cell and its cluster assignment.

Table S3 related to figure 1. Genes enriched in epidermal clusters identified by single cell RNA sequencing data.

Contig	Cluster	Fold-enrichment	P-value	FDR	Best-blast hit
dd_Smed_v4_14806_0	ζNeoblast	8.001827528	0	0	NA
dd_Smed_v4_3517_0	ζNeoblast	7.880165197	0	0	NA
dd_Smed_v4_7837_0	ζNeoblast	7.394674818	0	0	Smed09349_V2 cyclin B-2
dd_Smed_v4_11583_0	ζNeoblast	7.202441136	0	0	excision repair cross-complementing rodent repair
dd_Smed_v4_21084_0	ζNeoblast	6.61959667	0	0	NA
dd_Smed_v4_11558_0	ζNeoblast	6.562924256	0	0	primase, DNA, polypeptide 1 (49kDa) (PRIM1)
dd_Smed_v4_16468_0	ζNeoblast	6.541433008	0	0	NA
dd_Smed_v4_7519_0	ζNeoblast	6.251794491	0	0	TNF receptor-associated factor 6, E3 ubiquitin protein
dd_Smed_v4_659_0_1	ζNeoblast	6.07685891	0	0	PIWI-like protein 1 (wi-1)
dd_Smed_v4_11487_0	ζNeoblast	5.965544948	0	0	Smed04901_V2 cyclin B-1
dd_Smed_v4_648_0_1	ζNeoblast	5.673951095	0	0	NA
dd_Smed_v4_4804_0	ζNeoblast	5.614949621	0	0	NA
dd_Smed_v4_5956_0	ζNeoblast	5.462241676	0	0	minichromosome maintenance complex component 4
dd_Smed_v4_3777_0	ζNeoblast	5.397305339	0	0	NA
dd_Smed_v4_6783_0	ζNeoblast	5.3591808	0	0	NA
dd_Smed_v4_3549_0	Early Stage	7.666250394	0	0	NA
dd_Smed_v4_332_0_1	Early Stage	7.249963769	0	0	clone NB.21.11E prog-1
dd_Smed_v4_13596_0	Early Stage	6.454915604	0	0	ATP-binding cassette, sub-family C (CFTR/MRP),
dd_Smed_v4_69_0_1	Early Stage	6.352017648	0	0	prog-2
dd_Smed_v4_17406_0	Early Stage	6.107465496	0	0	synaptotagmin II (SYT2)
dd_Smed_v4_363_0_1	Early Stage	5.568190187	0	0	NA
dd_Smed_v4_17734_0	Early Stage	5.432615253	0	0	NA
dd_Smed_v4_61_0_1	Early Stage	5.292693331	0	0	NA
dd_Smed_v4_478_0_1	Early Stage	5.181075253	0	0	NA
dd_Smed_v4_3616_0	Early Stage	5.109529268	0	0	mucolipin 2 (MCOLN2)
dd_Smed_v4_20486_0	Early Stage	4.771754646	0	0	NA
dd_Smed_v4_12111_0	Early Stage	4.622668126	0	0	slit1
dd_Smed_v4_5942_0	Early Stage	3.851155588	0	0	SOXP-3
dd_Smed_v4_15985_0	Early Stage	3.694627771	0	0	delta-like 1 (Drosophila) (DLL1)
dd_Smed_v4_8717_0	Early Stage	3.68194757	0	0	NA
dd_Smed_v4_379_0_1	Late Stage	5.651815655	0	0	NA
dd_Smed_v4_20920_0	Late Stage	5.125156002	0	0	NA
dd_Smed_v4_1321_0	Late Stage	5.087037796	0	0	NA
dd_Smed_v4_2664_0	Late Stage	5.038058333	0	0	NA
dd_Smed_v4_21183_0	Late Stage	5.024720536	0	0	NA
dd_Smed_v4_480_0_1	Late Stage	4.919773794	0	0	NA
dd_Smed_v4_2873_0	Late Stage	4.881085745	0	0	NA
dd_Smed_v4_26705_0	Late Stage	4.857887028	0	0	gdf-like protein
dd_Smed_v4_3226_0	Late Stage	4.675650474	0	0	NA
dd_Smed_v4_468_0_1	Late Stage	4.574445683	0	0	NA
dd_Smed_v4_2147_0	Late Stage	4.501555226	0	0	NA
dd_Smed_v4_13786_0	Late Stage	4.433950243	0	0	NA
dd_Smed_v4_135_0_1	Late Stage	4.433447566	0	0	NA
dd_Smed_v4_1413_0	Late Stage	4.418527443	0	0	NA
dd_Smed_v4_12720_0	Late Stage	4.408747852	0	0	SV2 related protein homolog (rat) (SVOP)
dd_Smed_v4_18_0_1	Later Stage	5.846838249	0	0	NA
dd_Smed_v4_306_0_1	Later Stage	5.778877923	0	0	NA
dd_Smed_v4_5234_0	Later Stage	4.467242475	0	0	hemipterous
dd_Smed_v4_364_0_1	Later Stage	4.431292962	0	0	vimentin (VIM)
dd_Smed_v4_4648_0	Later Stage	3.876027432	0	0	dysferlin, limb girdle muscular dystrophy 2B (autosomal
dd_Smed_v4_8497_0	Later Stage	3.722732927	0	0	ferric-chelate reductase 1 (FRRS1)
dd_Smed_v4_2132_0	Later Stage	3.181777703	0	0	NA
dd_Smed_v4_2326_0	Later Stage	3.073466324	0	0	NA
dd_Smed_v4_327_0_1	Later Stage	2.736664647	0	0	NA
dd_Smed_v4_4660_0	Later Stage	2.343085226	0	0	NA
dd_Smed_v4_1988_0	Later Stage	4.332324891	1.11E-	6.09E-	GLI pathogenesis-related 2 (GLIPR2)
dd_Smed_v4_9329_1	Later Stage	3.668902528	1.11E-	6.09E-	NA
dd_Smed_v4_2449_0	Later Stage	3.297918389	1.11E-	6.09E-	solute carrier family 25, member 45 (SLC25A45)
dd_Smed_v4_1189_0	Later Stage	3.120484488	1.11E-	6.09E-	3-oxoacid CoA transferase 1 (OXCT1)
dd_Smed_v4_1925_0	Later Stage	2.833449623	1.11E-	6.09E-	NA
dd_Smed_v4_19793_0	Non ciliated	6.68710989	1.54E-	8.16E-	NA
dd_Smed_v4_2037_0	Non ciliated	6.263151146	2.01E-	9.99E-	dihydropyrimidine dehydrogenase (DPYD)
dd_Smed_v4_351_0_1	Non ciliated	2.815264719	8.40E-	4.12E-	protease, serine, 12 (neurotrypsin, motopsin) (PRSS12)
dd_Smed_v4_11905_0	Non ciliated	5.046662609	5.45E-	2.61E-	par-3 partitioning defective 3 homolog (C. elegans)
dd_Smed_v4_12355_0	Non ciliated	4.054878574	8.40E-	3.99E-	Smed10693_V2 angiotensin converting enzyme-1
dd_Smed_v4_13346_0	Non ciliated	5.436352938	1.72E-	8.04E-	NA

dd_Smed_v4_28182_0	Non ciliated	5.056827359	5.52E-	0.000	NA
dd_Smed_v4_4619_0	Non ciliated	2.808987287	1.13E-	0.000	dual specificity phosphatase 10 (DUSP10)
dd_Smed_v4_3952_0	Non ciliated	7.138331686	1.76E-	0.000	TNF receptor-associated factor 2 (TRAF2)
dd_Smed_v4_7881_0	Non ciliated	2.522073279	2.99E-	0.001	mll5-2
dd_Smed_v4_26148_0	Non ciliated	5.468836448	3.55E-	0.001	PCM1
dd_Smed_v4_10407_0	Non ciliated	2.068306742	4.30E-	0.001	TNF receptor-associated factor 6, E3 ubiquitin protein
dd_Smed_v4_364_0_1	Non ciliated	2.860060237	5.94E-	0.002	vimentin (VIM)
dd_Smed_v4_1882_0	Non ciliated	4.741448021	7.82E-	0.003	NA
dd_Smed_v4_12567_0	Non ciliated	5.942267369	1.07E-	0.004	NA
dd_Smed_v4_351_0_1	Ciliated	5.884683588	0	0	protease, serine, 12 (neurotrypsin, motopsin) (PRSS12)
dd_Smed_v4_6790_0	Ciliated	7.48059832	4.44E-	2.42E-	dynein, axonemal, heavy chain 1 (DNAH1)
dd_Smed_v4_17692_0	Ciliated	6.771344517	3.33E-	1.80E-	chromosome X open reading frame 22 (CXorf22)
dd_Smed_v4_3833_0	Ciliated	5.817826015	1.01E-	5.45E-	chromosome 1 open reading frame 194 (C1orf194)
dd_Smed_v4_13612_0	Ciliated	7.042598874	1.52E-	8.18E-	NA
dd_Smed_v4_298_0_1	Ciliated	7.701862723	1.99E-	1.07E-	NA
dd_Smed_v4_181_0_1	Ciliated	8.705665184	1.72E-	9.11E-	aminopeptidase-like 1 (NPEPL1)
dd_Smed_v4_3340_0	Ciliated	6.019217525	3.56E-	1.88E-	NA
dd_Smed_v4_6321_0	Ciliated	5.627955058	4.40E-	2.33E-	spermatogenesis associated 18 (SPATA18)
dd_Smed_v4_7087_0	Ciliated	5.401596531	7.95E-	4.18E-	sorting nexin 33 (SNX33)
dd_Smed_v4_4191_0	Ciliated	7.356886647	1.32E-	6.92E-	coiled-coil domain containing 63 (CCDC63)
dd_Smed_v4_5727_0	Ciliated	6.852120545	6.91E-	3.57E-	coiled-coil domain containing 147 (CCDC147)
dd_Smed_v4_5255_0	Ciliated	8.548627064	8.32E-	4.29E-	NIMA-related kinase 3 (NEK3)
dd_Smed_v4_1404_0	Ciliated	6.63591766	1.29E-	6.60E-	ciliary rootlet coiled-coil, rootletin (CROCC)
dd_Smed_v4_2935_0	Ciliated	6.124307522	2.39E-	1.22E-	EF-hand domain (C-terminal) containing 1 (EFHC1)A
dd_Smed_v4_5204_0	Dorsal ventral	10.40236717	0	0	NA
dd_Smed_v4_10286_0	Dorsal ventral	10.24695847	0	0	NA
dd_Smed_v4_3871_0	Dorsal ventral	9.250083077	0	0	clone D21.6.Contig unknown
dd_Smed_v4_399_0_1	Dorsal ventral	8.890262402	0	0	NA
dd_Smed_v4_3065_0	Dorsal ventral	8.869687838	0	0	intermediate filament b (ifb) / laminB
dd_Smed_v4_10873_0	Dorsal ventral	8.342029465	1.11E-	6.09E-	NA
dd_Smed_v4_16021_0	Dorsal ventral	6.526147115	3.66E-	1.98E-	homeobox B9 (HOXB9) (planarian gene: *Post-2a*)
dd_Smed_v4_7515_0	Dorsal ventral	8.723975418	5.00E-	2.70E-	NA
dd_Smed_v4_13493_0	Dorsal ventral	8.147659667	2.65E-	1.42E-	NA
dd_Smed_v4_21347_0	Dorsal ventral	5.589463257	3.80E-	2.03E-	nuclear receptor TLX-1 (tlx-1)
dd_Smed_v4_3234_0	Dorsal ventral	6.413678121	7.61E-	4.06E-	1.G9.2
dd_Smed_v4_13412_0	Dorsal ventral	7.333859553	5.10E-	2.69E-	protocadherin 1 (PCDH1)
dd_Smed_v4_1230_0	Dorsal ventral	3.765307383	8.15E-	4.20E-	secretagogin, EF-hand calcium binding protein (SCGN)
dd_Smed_v4_14748_0	Dorsal ventral	4.3472585	1.42E-	7.25E-	polypyrimidine tract binding protein 2 (PTBP2)
dd_Smed_v4_5372_0	Dorsal ventral	2.921807366	3.36E-	1.71E-	NA

Table S3 related to figure 1. Genes enriched in epidermal clusters identified by single cell RNA sequencing data.

Gene expression comparison between single-cell clusters identified 1,452 genes enriched in particular clusters (Table S3; FDR < 0.001, fold-change > 4; Methods). Tabulated here are the 15 most highly enriched genes from each cluster. For each gene, the table indicates the contig ID, cluster assignment, fold enrichment, p-value, FDR and best BLAST hit.

Table S4 related to Figure 2. Core centrosome components in the epidermal lineage.				
***S. mediterranea* centrosome components from Azimzadeh et al., Science, 2012**	**Contig (In the assembly used by Azimzadeh et al., Science, 2012)**	**Contig (Dresden transcriptome assembly v4)**	**(1 - Ciliated cells, 2 - Ciliated cells or neoblasts, 3 - All cells)**	**(1 - Early Stage, 2 - Late Stage, 3 - Later Stage, 4 - Mature epidermis)**
SMED-SAS-6	mk4.026484.00.01	dd_Smed_v4_8420_0_1	3	1
SMED-SAS-4	mk4.002572.02.01	dd_Smed_v4_13670_0_1	1	1
SMED-MKS1	mk4.004309.01.01	dd_Smed_v4_14053_0_1	2	1
SMED-HYLS1	de_novo.28401.1	dd_Smed_v4_13289_0_1	1	1
SMED-FAM161	mk4.000530.00.01	dd_Smed_v4_15201_0_1	1	1
SMED-CEP135	mk4.012776.01.01	dd_Smed_v4_5218_0_1	1	2
SMED-POC1	mk4.013333.00.01	dd_Smed_v4_10806_0_1	3	1
SMED-ANA1	de_novo.6545.1	dd_Smed_v4_11068_0_1	1	2
SMED-ANA3	mk4.006165.00.01	dd_Smed_v4_83405_0_1	Not detected	Not detected
SMED-CEN2	mk4.004035.00.02	dd_Smed_v4_3644_0_1	3	1
SMED-CEP120	mk4.001538.00.01	dd_Smed_v4_10444_0_1	2	1
SMED-CEP164	mk4.000030.11.01	dd_Smed_v4_12228_0_1	3	1
SMED-CEP131	ASA.00124.01	dd_Smed_v4_13976_0_1	2	2
SMED-CEP290	mk4.001346.02.01	dd_Smed_v4_7145_0_1	2	2
SMED-WDR67	JQ036196	dd_Smed_v4_9022_0_1	3	1
SMED-CEP78	JQ036187	dd_Smed_v4_9671_0_1	1	3
SMED-Plk4	mk4.000854.08	dd_Smed_v4_12637_0_1	1	3
SMED-CEP152	de_novo.8133.1	dd_Smed_v4_12630_0_1	2	1
SMED-ODF2	mk4.002548.02.01	dd_Smed_v4_3609_0_1	3	1
SMED-ROOTLETIN	de_novo.21524.1	dd_Smed_v4_6573_0_1	1	4
SMED-PCNT	de_novo.1766.1	dd_Smed_v4_8417_0_1	3	1
SMED-SAS-5	de_novo.14261.1	dd_Smed_v4_13944_0_1	3	2
SMED-PCM1	JQ036191	dd_Smed_v4_26148_0_1	1	4
SMED-CP110	mk4.005271.00.01	dd_Smed_v4_13022_0_1	3	2
SMED-CEP97	JQ036188	dd_Smed_v4_11066_0_1	1	3
SMED-ALMS1	mk4.001634.00.01	dd_Smed_v4_5064_0_1	3	1
SMED-CCDC123	de_novo.16907.1	dd_Smed_v4_13073_0_1	2	2
SMED-LRRC45	mk4.004069.00.01	dd_Smed_v4_9495_0_1	2	1
SMED-CEP70	mk4.006145.02.01	dd_Smed_v4_14545_0_1	3	2
SMED-CCDC52	de_novo.25840.1	dd_Smed_v4_10100_0_1	2	1
SMED-NA14	de_novo.10830.1	dd_Smed_v4_3576_0_1	2	2
SMED-OFD1	JQ036190	dd_Smed_v4_13339_0_1	2	2
SMED-POC5	JQ036192	dd_Smed_v4_17398_0_1	1	3
SMED-CEN3	mk4.006556.00.01	dd_Smed_v4_11582_0_1	1	1
SMED-SFI1	JQ036193	dd_Smed_v4_12484_0_1	3	1
SMED-TUBD1	JQ036194	dd_Smed_v4_15819_0_1	3	1
SMED-TUBE1	JQ036195	dd_Smed_v4_11274_0_1	3	2
SMED-CEP76	JQ036186	dd_Smed_v4_9101_0_1	1	1
SMED-LRRCC1	JQ036189	dd_Smed_v4_10086_0_1	3	1
SMED-CCDC61	de_novo.17607.1	dd_Smed_v4_14065_0_1	3	1
SMED-CCDC77	mk4.026003.00.01	dd_Smed_v4_4812_0_1	1	2
SMED-CEP41	JQ036184	dd_Smed_v4_9621_0_1	1	3
SMED-CAP350	de_novo.2749.1	dd_Smed_v4_12389_0_1	2	1
SMED-FOP	de_novo.26140.1	dd_Smed_v4_5371_0_1	3	1
SMED-CEP72	JQ036185	dd_Smed_v4_12835_0_1	1	2
SMED-CEP763	de_novo.26061.1	dd_Smed_v4_13335_0_1	3	1

SMED-CENTRIOLIN	JQ036183	dd_Smed_v4_10369_0_1	3	1

Table S4 related to Figure 2. Core centrosome components in the epidermal lineage.

Cilia components selected from Azimzadeh et al., Science, 2012 used for analysis. Single-cell gene expression was used to characterize the degree to which these components were expressed in neoblasts, or in other stages of epidermal lineage. Tabulated for each gene is the gene name, contig ID used in Azimzadeh et al., Science, 2012, Dresden contig ID, assignment of whether the gene is expressed in ciliated cells or neoblasts, and the assignment of when the gene is expressed in the epidermal lineage.

Table S5 related to figure 4. Genes enriched in dorsal or ventral Early epidermal progenitors.							
Contig	**Gene title in figure**	**Class**	**AUC**	**Fold change**	**Power**	**P-value**	**FDR**
dd_Smed_v4_10730_0_1	UNC5C	Dorsal	0.88	5.644570194	0.758	0	0
dd_Smed_v4_11840_0_1	*PRDM1-1*	Dorsal	0.85	2.774267535	0.69	3.33E-16	2.85E-13
dd_Smed_v4_8690_0_1	LRRC4	Dorsal	0.84	2.047608287	0.688	2.02E-12	1.70E-09
dd_Smed_v4_11062_0_1	*bambi*	Dorsal	0.82	5.223708163	0.638	1.48E-14	1.26E-11
dd_Smed_v4_17350_0_1	GCNT1; N-acetyl	Dorsal	0.82	4.050489376	0.634	2.55E-15	2.18E-12
dd_Smed_v4_10380_0_1	UNC5B	Dorsal	0.81	2.13913138	0.622	8.82E-14	7.48E-11
dd_Smed_v4_11693_0_1	TBX2	Dorsal	0.75	3.840177621	0.502	3.74E-11	3.13E-08
dd_Smed_v4_6561_0_1	ACSS1	Dorsal	0.74	2.091715113	0.486	8.35E-06	0.00678777
dd_Smed_v4_12381_0_1	GCNT1; N-acetyl	Dorsal	0.74	4.893681653	0.482	1.11E-11	9.29E-09
dd_Smed_v4_8902_0_1	MSI2	Ventral	0.7	-2.146935801	0.408	9.19E-07	0.00075686
dd_Smed_v4_3027_0_1	LTA4H	Ventral	0.71	-4.791937131	0.414	1.95E-09	1.63E-06
dd_Smed_v4_15516_0_1	*otxb-like*	Ventral	0.74	-5.664615728	0.482	2.24E-13	1.89E-10
dd_Smed_v4_10647_0_1	FBXL7	Ventral	0.75	-5.722537996	0.494	1.69E-14	1.44E-11
dd_Smed_v4_6816_0_1	MYOF	Ventral	0.76	-5.095899243	0.522	7.87E-14	6.68E-11
dd_Smed_v4_7249_0_1	dd_7249	Ventral	0.78	-4.164683201	0.564	2.29E-14	1.95E-11
dd_Smed_v4_12320_0_1	dd_12320	Ventral	0.78	-5.000742402	0.566	0	0
dd_Smed_v4_1422_0_1	dd_1422	Ventral	0.81	-2.123027429	0.624	6.85E-12	5.77E-09
dd_Smed_v4_9783_0_1	dd_9783	Ventral	0.82	-6.070519048	0.63	0	0
dd_Smed_v4_6345_0_1	LRRC4B	Ventral	0.83	-6.355948681	0.652	0	0
dd_Smed_v4_7911_0_1	ROR1	Ventral	0.83	-6.73430508	0.668	0	0
dd_Smed_v4_6261_0_1	dd_6261	Ventral	0.84	-4.858036953	0.672	0	0
dd_Smed_v4_12450_0_1	dd_12450	Ventral	0.87	-5.763017674	0.734	0	0
dd_Smed_v4_6746_0_1	*kal1*	Ventral	0.9	-5.394559145	0.808	0	0

Table S5 related to figure 4. Genes enriched in dorsal or ventral Early epidermal progenitors.

Gene expression of Early Stage progenitors from dorsal and ventral regions were compared. Twenty-three genes were identified as significantly enriched (Fold-change > 4; FDR < 0.1; Power > 0.4) in dorsal or ventral progenitors. Tabulated for each gene is the contig ID, gene title, class assignment, AUC, fold change, power, p-value and FDR.

Materials and Methods

Accession numbers

The data reported in this publication is available at Sequence Read Archive (SRA) under accession PRJNA353867.

Gene labeling and nomenclature

Previously published planarian genes, and genes for which we performed phylogenetic (*PRDM1-1*, *ovo-2*, *ovo-3*) or domain-structure analysis (*kal1*) appear in *italics* throughout the text and the figures. Uppercase labels are the human best-blast hits (blastx; E-value < 10^{-5}), or genes not named with detailed analysis (such as phylogenetic analysis). Numeric labels with "dd_" prefix are shown for genes without human best-blast hit; instead the contig id in the dd_v4 transcriptome assembly was used [44]. The matching dd contig id for each gene label is found in the Supplementary Methods.

Planarian epidermis isolation and RNAseq library preparation

Epidermis was isolated by incubating planarians in 3.8% Ammonium Thiocyanate as previously described [27] in phosphate-buffered saline (PBS) for 20 minutes. Next, epidermis from either the ventral or dorsal surface was scraped off, separately, with a needle pulled from a borosilicate glass capillary into a collection tube on ice. Cells were spun down at 300G for 5 minutes at 4°C, resuspended in PBS, and spun down again. Finally, cells were resuspended in 0.25 mL of TRIzol (ThermoFisher Scientific), and stored at -80C. In parallel, whole animal controls were put in TRIzol. RNA extraction for samples in TRIzol was performed according to manufacturer's protocol. For each sample, 1 µg of RNA was used for RNA sequencing library preparation using the TruSeq RNA library prep kit V2 (Illumina, Inc.) following the manufacturer's protocol. Libraries were sequenced on Illumina HiSeq.

Differential expression analysis of epidermis-enriched RNAseq libraries

Sequenced RNA-seq libraries were mapped as recently described [28] using Novoalign v2.08.02 with parameters [-o SAM -r Random] to the *S. mediterranea* dd_Smed_v4 transcriptome assembly [44]. Read count per contig was calculated with bedtools v2.20.1 [45] for each library. Short contigs (<350) were removed from further analysis, and reads mapped to different contig isoforms summed to represent a single contig. Differentially expressed genes were called using EdgeR v3.6.8 [41], requiring minimal fold-change of 2 and maximal FDR of 1E-4.

Single-cell collection by FACS and qPCR

Animals were subjected to two transverse amputations to generate pre-pharyngeal fragments. The fragments were further dissected into dorsal, ventral, and lateral fragments (Fig 2A). Samples were macerated, stained with Hoechst (1:25) and propidium iodide (1:500) and 2C cells were sorted to plates containing lysis buffer (TCL buffer; QIAgen, inc.), as recently described [28]. Following reverse transcription and cDNA amplification [28,29], qPCR was performed on plates (total 1,096 cells) to identify and isolate *agat-1* expressing cells (forward and reverse sequences, CCTAAAAGGCGAAGGTGTGACT and TGCAACATCCAAACCGACAGA, respectively), with the following program [95°C (30s), 40 cycles (95°C 3s, 60°C 30s)]. Similarly, cells from the lateral region were also screened for the expression of the DV epidermal boundary marker *laminB* (forward and reverse sequences, TGTGGGTAGCCTTTTCTTCTCCC and CGCAAGGTTCAGGTGATCCG, respectively). Cells displaying Ct value of 25 or less were considered as expressing the assayed gene and were used for SCS library preparation.

Single-cell RNA sequencing library preparation

Libraries were prepared following a published protocol [29] with previously described modifications [29].

Analysis of single cell RNA sequencing data

Gene expression data was mapped to the dd_Smed_v4 assembly [44] as previously described [28]. Recently published SCS data prepared using the same protocol and

equipment, was combined if the cells were from the epidermal lineage. Gene expression data for all cells were processed together as follows: Following mapping using Novoalign v2.08.02, a raw expression matrix was generated using custom Perl script, and expression matrix was normalized using edgeR v3.6.8 [41]. Contigs smaller than 450 nt were removed. Cells expressing less than 1000 or more than 7800 genes were removed from further analysis to ensure low-quality cells, or potentially contaminated samples are not included in the analysis [28]. Initial determination of significant principal components was performed by calling function mean.var.plot with parameters [y.cutoff = 2.5, x.low.cutoff = 2, fxn.x = expMean,fxn.y = logVarDivMean], which found 267 meeting these criteria. The jackstraw function [num.replicate = 200] determined principal components 1-12 as significantly contributing to cell-to-cell gene expression variance. The function pca.sig.genes was called with parameters [pcs.use = 1:12, pval.cut = 1e-5, max.per.pc = 300] and the resultant list was used for a second pca analysis. Then, four cells were removed from the dataset ["D15.101224_DMX", "D15.101225_DMX", "D15.101162_DLX", "D15.101359_L0X"], as they expressed multiple neural genes (including *PC2*, *synapsin*, and *synaptotagmin*), and were unlikely to be part of the epidermal lineage. tSNE was performed using the function run_tsne [dims.use = c(1:18), perplexity = 20, do.fast = T]. Next, data was clustered by DBclust_dimension [reduction.use = "tsne", G.use = 2, set.ident = TRUE, MinPts = 3] and clusters were sorted by the buildClusterTree function [do.reorder = TRUE, reorder.numeric = TRUE, pcs.use = 1:18]. Following gene expression analysis cluster 3 was determined to be muscle cells based on expression of multiple muscle markers and was removed. Clusters 8, 11 and 12 included small number of low complexity cells and were removed.

qPCR of dividing neoblasts isolated from dorsal or ventral regions

Animals were first cut transversely and then dorsal and ventral fragments were isolated. Fragments from different animals were pooled (>20) and macerated to a single cell suspension, as recently described [28]. Groups of 500 dividing neoblasts, from the dorsal and ventral samples, were sorted, separately (dorsal replicates n = 8, ventral replicates = 6) into plates containing 30 µl of lysis buffer (TCL buffer; QIAgen, inc.).

RNA from samples was converted to cDNA and amplified using the smart-seq V2 [29] protocol with small modifications [28]. The amplified cDNA library concentrations were measured using Qubit fluorometer and sample concentrations were normalized to 5 ng/μl. Expression of the target genes was measured in each sample using qPCR. qPCR was performed on Applied Biosystems 7500-fast RT-PCR machine for *smedwi-1* (forward and reverse primer sequences: GTCTCAGAAAACAACTAAAGGTACAGCA and TGCTGCAATACACTCGGAGACA, respectively), *PRDM1-1* (forward and reverse primer sequences: CGGTGAACGACCTTTCAAGT and TCAAACAAACCGAACACTCG, respectively), and *kal1* (forward and reverse primer sequences: TCTGTGTGCCCTCTTGTACG and CAGATTTTCCGGCTGAGAAG, respectively) with the following program [95°C (30s), 40 cycles (95°C 3s, 60°C 30s)], with at least two technical replicates per sample. The measured Ct values of either *kal1* or *PRDM1-1* was normalized by the sample expression of *smedwi-1*, a pan-expressed neoblast gene [46]. Student t-test was used for differential expression analysis of between dorsal- and ventral-isolated samples, and p-values were corrected for multiple hypothesis testing using a Bonferroni correction.

Gene cloning and transformation

Genes were cloned as previously described [28]. Briefly, gene-specific primers were used to amplify gene sequences of planarian cDNA. Sequences were cloned into pGEM vectors following the manufacturer's protocol (Promega), and transformed to E. coli DH10B by heat-shock. Bacteria were plated on agarose plates containing 1:500 carbenicillin, 1:200 Isopropylthio-b-D-galactoside (IPTG), and 1:625 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) for overnight growth. Colonies were screened by colony PCR and gel electrophoresis. Plasmids were extracted from positive colonies and subsequently validated by Sanger-sequencing (Genewiz, Inc.).

In situ Hybridizations

Nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) colorimetric whole-mount in situ hybridizations (ISH) were performed as previously described[40].

Fluorescence in situ hybridizations (FISH) were performed as previously described [47] with minor modifications. Briefly, animals were killed in 5% NAC and treated with proteinase K (2ug/ml). Following overnight hybridizations, samples were washed twice in each of pre-hybridization buffer, 1:1 pre- hybridization-2X SSC, 2X SSC, 0.2X SSC, PBST. Subsequently, blocking was performed in 0.5% Western Blocking Reagent (Roche 11921673001) and 5% inactivated horse serum PBST solution when anti-DIG or anti-DNP antibodies were used, and in in 1.0% Western Blocking Reagent PBST solution when an anti-FITC antibody was used. Post-antibody binding washes and tyramide development were performed as described [47]. Peroxidase inactivation with 1% sodium azide was done for 90 minutes at RT.

Immunostainings

Immunostainings for acetylated-tubulin were performed as previously described [25]. Immunofluorescence (IF) and fluorescent in situ hybridization (FISH) signals were developed using tyramide conjugated fluorophores generated from AMCA, fluorescein, rhodamine (Pierce), and Cy5 (GE Healthcare) N-hydroxysuccinimide (NHS) esters as previously reported [48].

Double-stranded RNA synthesis and RNAi experiments

Gene inhibition was done by feeding the animals with dsRNA corresponding to the target gene. dsRNA was synthesized as previously described [28] following a published protocol [49]. Animals were starved for 7 days prior to experiments and were kept in the dark for at least 2 hours before each feeding. Animals were fed 6 times, unless stated otherwise, every 3 days by mixing the dsRNA with homogenized beef liver (1:3).

Following the RNAi feedings animals were cut and allowed to regenerate for eight days, while being monitored for defects. Following the eight days recovery period, animals were fixed and analyzed by FISH. dsRNA injections, using Drummond Nanoject II, were performed in the prepharyngeal region. *bmp4* inhibition by injection was performed using a single injection.

Irradiation

Animals were irradiated using Gammacell 40 dual ^{137}cesium sources (6000 rads) and were used for experiments 24 hours following irradiation, when all neoblasts are ablated [22].

Contigs used for computing correlation of epidermal gene expression in progenitors

The transcriptome assembly was searched for contigs with best-blast hit description including the keywords IFT or "intraflagellar transport". The expression of all contigs was visualized using the SCS resource [28] and genes that are not exclusively expressed in ciliated cells (e.g. epidermal cells, protonephridia, ciliated neurons) were discarded. For each gene pair, Pearson correlation (r) was calculated, per cluster, and then z-transformed using Fisher transformation. Average was calculated on z-transformed values, per cluster, and then transformed to r.

Contig	Best Blast hit description	ID	E-value	Organism
dd_Smed_v4_10638_0_1	IFT172-like protein	GQ337484.1	0	Smed
dd_Smed_v4_11300_0_1	intraflagellar transport 140 homolog (Chlamydomonas) (IFT140)	uc002cmb.3	0	Human
dd_Smed_v4_5043_0_1	IFT52-like protein	GQ337481.1	0	Smed

dd_Smed_v4_5484_0_1	IFT88-like protein	GQ337482.1	0	Smed
dd_Smed_v4_7099_0_1	intraflagellar transport 57 homolog (Chlamydomonas) (IFT57)	uc003dwx.4	3.00E-91	Human
dd_Smed_v4_7533_0_1	intraflagellar transport 172 homolog (Chlamydomonas) (IFT172)	uc002rku.3	1.00E-53	Human
dd_Smed_v4_8803_0_1	intraflagellar transport 81 homolog (Chlamydomonas) (IFT81)	uc001tqh.3	0	Human
dd_Smed_v4_9110_0_1	intraflagellar transport 80 homolog (Chlamydomonas) (IFT80)	uc021xgr.1	0	Human

Comparison of cilia-encoding gene expression in dorsal and ventral early progenitors

Cilia-encoding genes were selected by searching for ciliated-neuron-enriched genes [28] that were also expressed in mature epidermis and protonephridia cells, which represent the three ciliated planarian cell types. Genes conforming to this criterion, and not expressed in any other cell types were selected. Only genes expressed (log_2(CPM+1) > 2) in at least 15 early progenitor cells were tested.

Contig	FDR	Best blast hit	ID	E-value	Organism
dd_Smed_v4_7145_0_1	0.000367	centrosomal protein 290kDa (CEP290)	uc001tar.3	0	Human
dd_Smed_v4_4669_0_1	0.464633	calmodulin 2 (phosphorylase kinase, delta) (CALM2)	uc002rvt.2	4.00E-25	Human
dd_Smed_v4_7298_0_1	2.58E-05	Bardet-Biedl syndrome 1 (BBS1)	uc001oij.1	9.00E-172	Human
dd_Smed_v4_11366_0_1	0.009266	NIMA-related kinase 1 (NEK1)	uc003isd.2	2.00E-100	Human
dd_Smed_v4_8117_0_1	0.519185	NA	NA	NA	NA
dd_Smed_v4_12717_0_1	0.632699	tubulin tyrosine ligase-like family, member 3 (TTLL3)	uc010hco.1	4.00E-109	Human
dd_Smed_v4_12105_0_1	0.278922	leucine-rich repeat kinase 2 (LRRK2)	uc001rmg.4	1.00E-33	Human
dd_Smed_v4_10737_0_1	0.060868	ATPase, class V, type 10B (ATP10B)	uc003lym.1	7.00E-116	Human
dd_Smed_v4_6941_0_1	0.068332	dynein, axonemal, heavy chain 7 (DNAH7)	uc002utj.4	0	Human
dd_Smed_v4_7302_1_1	0.246504	NA	uc002ezr.3	0	Human
dd_Smed_v4_6686_0_1	0.001253	NA	uc002qji.1	2.00E-72	Human
dd_Smed_v4_10152_0_1	0.088794	forkhead box J1-like protein 4	JX842846.1	0	Smed

dd_Smed_v4_9816_0_1	0.464633	NA	NA	NA		NA
dd_Smed_v4_9136_0_1	0.000367	dynein, cytoplasmic 2, heavy chain 1 (DYNC2H1)	uc001pho.2		0	Human
dd_Smed_v4_3323_0_1	0.057932	WD repeat domain 65 (WDR65)5	uc021omk.1		0	Human

Identification of DV-boundary epidermis transcription factors

TFs that were overexpressed in the epidermal lineage were collected from three datasets: (1) TFs enriched in one or more epidermal progenitor or mature epidermis cluster when compared to previously published SCS gene expression of other cell types [28], using the Seurat package [31] find.markers function with thresholds: [FDR < 1E-3, FC > 2]; (2) TFs that were overexpressed in bulk epidermal RNA-seq compared to whole-worm controls [FDR < 0.01, FC > 2]; or (3) that were downregulated in animals with depleted epidermal lineage (*zfp-1* RNAi) using recently published RNAseq data [19] that was mapped to the dd transcriptome [44] and analyzed using the same methods described in the Method section for the bulk epidermal RNAseq libraries data analysis with the following thresholds for significance (FDR < 0.01, FC > 2). TFs enriched in the DV boundary epidermis SCS data (Table S3) or expressed in at least half of the DV-boundary epidermis cells but in none of the other mature cell types (Fig 2E) were further analyzed by FISH and RNAi.

Microscopy and image analysis

FISH and IF images were collected using a confocal microscope (Zeiss LSM 700). Cell counting was performed using ImageJ (1.5) by cropping images to a set size and counting cells using the "Cell Counter" component. The same region was used to count cells for both the controls and the experiment. Transverse sections were generated by

using a scalpel and generating two transverse cuts. Measurement of spatial distribution of *kal1*+ neoblasts (Fig 5D) was done by imaging transverse sections. Then, dividing each section to 10 segments of half concentric ellipses, and counting the number of *kal1*+ neoblasts in each section. Sections were evaluated according to their position compared to major landmarks (*bmp4* expressing muscle cells, DV boundary epidermal cells, and the intestine). The number of *kal1*+ cells in irradiated and unirradiated *bmp4* RNAi animals and controls was quantified by taking prepharyngeal images (400 μm^2, 10 confocal slices, Fig 6H-I) for all animals in four groups (irradiated control, unirradiated control, irradiated *bmp4* RNAi, unirradiated *bmp4* RNAi). Then the file names of the images were randomized and the images were analyzed by a different researcher. Positive cells were counted using the "Cell Counter" component in ImageJ, and documented for each file. Following measurements, the file names were derandomized, and data of different groups was analyzed.

Calculation of the overlap fraction in epidermal progenitors

Overlap fraction was calculated for *ovo-2* as follows: [Number of *ovo-2*+ cells expressing gene X / Number of *ovo-2*+ cells], where gene X represents any gene in the transcriptome for which gene expression was greater than $\log_2(CPM+1) = 5$. Overlap fraction for DCLK2 was calculated similarly. Genes expressed in a small fraction of the entire cell population (<25%) were removed as their overlap fraction was minor. Genes expressed throughout the cell population (> 66%) were also removed, to avoid confounding the results with genes expressed abundantly, such as housekeeping genes.

Phylogenetic analysis of epidermal genes

Phylogenetic analysis of *PRDM1-1*, *ovo-2,* and *ovo-3* was done by multiple alignment of the protein sequences with known sequences from representative model organisms using mafft v7.017b [50] with flag [--auto]. Then gap sequences were trimmed, and trees resolved using RAxML [51] with flags [Substitution model: PROTCAT DAYHOFF; Matrix name: DAYHOFF; Bootstrap: 100]. *kal-1* analysis was done using multiple methods: (1) Reciprocal BLASTP search [e-value < 1E-30) against the human proteome resulted in best similarity with KAL1; (2) Conserved protein domain analysis using CDART [52] and InterProScan v5.18-57.0 [53] of domain structure in model organisms and *S. mediterranea*; and (3) Classification based on the PANTHER database prediction using the protein sequence of the predicted planarian *kal1* [43].

Acknowledgements

We thank Kellie Kravarik and the Reddien lab members for manuscript comments. OW was supported by an EMBO long-term fellowship, and is the Howard Hughes Medical Institute Fellow of The Helen Hay Whitney Foundation. We acknowledge NIH (R01GM080639) support. PWR is a Howard Hughes Medical Institute Investigator and an associate member of the Broad Institute of Harvard and MIT.

References

1. Sánchez Alvarado, A., and Yamanaka, S. (2014). Rethinking differentiation: stem cells, regeneration, and plasticity. Cell. *157*, 110–9.

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

3. Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S., and Amit, I. (2014). Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell. *159*, 1312–26.

4. Baxendale, S., Davison, C., Muxworthy, C., Wolff, C., Ingham, P.W., and Roy, S. (2004). The B-cell maturation factor Blimp-1 specifies vertebrate slow-twitch muscle fiber identity in response to Hedgehog signaling. Nat. Genet. *36*, 88–93.

5. Gautier, E.L., Shay, T., Miller, J., Greter, M., Jakubzick, C., Ivanov, S., Helft, J., Chow, A., Elpek, K.G., Gordonov, S., et al. (2012). Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. Nat. Immunol. *13*, 1118–28.

6. Reddien, P.W., and Sánchez Alvarado, A. (2004). Fundamentals of Planarian Regeneration. Annu. Rev. Cell Dev. Biol. *20*, 725–57.

7. Wagner, D.E., Wang, I.E., and Reddien, P.W. (2011). Clonogenic Neoblasts Are Pluripotent Adult Stem Cells That Underlie Planarian Regeneration. Science. *332*, 811–6.

8. Scimone, M.L., Srivastava, M., Bell, G.W., and Reddien, P.W. (2011). A

regulatory program for excretory system regeneration in planarians. Development. *138*, 4387–98.

9. Forsthoefel, D.J., James, N.P., Escobar, D.J., Stary, J.M., Vieira, A.P., Waters, F. a, and Newmark, P. a. (2012). An RNAi screen reveals intestinal regulators of branching morphogenesis, differentiation, and stem cell proliferation in planarians. Dev. Cell. *23*, 691–704.

10. Adler, C., Seidel, C., McKinney, S., and Alvarado, A. (2014). Selective amputation of the pharynx identifies a FoxA-dependent regeneration program in planaria. eLife, 1–22.

11. Scimone, M.L., Kravarik, K.M., Lapan, S.W., and Reddien, P.W. (2014). Neoblast specialization in regeneration of the planarian *Schmidtea mediterranea*. Stem Cell Reports. *3*, 339–52.

12. Cowles, M.W., Brown, D.D.R., Nisperos, S. V, Stanley, B.N., Pearson, B.J., and Zayas, R.M. (2013). Genome-wide analysis of the bHLH gene family in planarians identifies factors required for adult neurogenesis and neuronal regeneration. Development, 1–12.

13. Lapan, S.W., and Reddien, P.W. (2011). *dlx* and *sp6-9* Control Optic Cup Regeneration in a Prototypic Eye. PLoS Genet. *7*, e1002226.

14. Lapan, S.W., and Reddien, P.W. (2012). Transcriptome Analysis of the Planarian Eye Identifies *ovo* as a Specific Regulator of Eye Regeneration. Cell Rep. *2*, 1–14.

15. Scimone, M.L., Lapan, S.W., and Reddien, P.W. (2014). A *forkhead* Transcription Factor Is Wound-Induced at the Planarian Midline and Required for Anterior Pole Regeneration. PLoS Genet. *10*, e1003999.

16. Reddien, P.W. (2013). Specialized progenitors and regeneration. Development. *140*, 951–7.

17. Glazer, A.M., Wilkinson, A.W., Backer, C.B., Lapan, S.W., Gutzman, J.H., Cheeseman, I.M., and Reddien, P.W. (2010). The Zn finger protein Iguana impacts Hedgehog signaling by promoting ciliogenesis. Dev. Biol. *337*, 148–56.

18. Tazaki, A., Kato, K., Orii, H., Agata, K., and Watanabe, K. (2002). The body margin of the planarian *Dugesia japonica*: characterization by the expression of an intermediate filament gene. Dev. Genes Evol. *212*, 365–73.

19. van Wolfswinkel, J.C., Wagner, D.E., and Reddien, P.W. (2014). Single-Cell Analysis Reveals Functionally Distinct Classes within the Planarian Stem Cell Compartment. Cell Stem Cell. *15*, 1–14.

20. Tu, K.C., Cheng, L.-C., TK Vu, H., Lange, J.J., McKinney, S.A., Seidel, C.W., and Sánchez Alvarado, A. (2015). *Egr-5* is a post-mitotic regulator of planarian epidermal differentiation. eLife. *4*, 1–27.

21. Vij, S., Rink, J.C., Ho, H.K., Babu, D., Eitel, M., Narasimhan, V., Tiku, V., Westbrook, J., Schierwater, B., and Roy, S. (2012). Evolutionarily ancient association of the FoxJ1 transcription factor with the motile ciliogenic program. PLoS Genet. *8*, e1003019.

22. Eisenhoffer, G.T., Kang, H., and Sánchez Alvarado, A. (2008). Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea*. Cell Stem Cell. *3*, 327–39.

23. Molina, M.D., 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. Dev. Biol. *311*, 79–94.

24. Orii, H., and Watanabe, K. (2007). Bone morphogenetic protein is required for dorso-ventral patterning in the planarian *Dugesia japonica*. Dev. Growth Differ., 345–9.

25. Reddien, P.W., Bermange, A.L., Kicza, A.M., and Sánchez Alvarado, A. (2007). BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration. Development. *134*, 4043–51.

26. Witchley, J.N., Mayer, M., Wagner, D.E., Owen, J.H., and Reddien, P.W. (2013). Muscle Cells Provide Instructions for Planarian Regeneration. Cell Rep. *4*, 1–9.

27. Trost, A., Bauer, J.W., Lanschützer, C., Laimer, M., Emberger, M., Hintner, H., and Onder, K. (2007). Rapid, high-quality and epidermal-specific isolation of RNA from human skin. Exp. Dermatol. *16*, 185–90.

28. Wurtzel, O., Cote, L.E., Poirier, A., Satija, R., Regev, A., and Reddien, P.W. (2015). A Generic and Cell-Type-Specific Wound Response Precedes Regeneration in Planarians. Dev. Cell. *35*, 632–45.

29. Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., and

Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods. *10*, 1096–8.

30. Van Der Maaten, L.J.P., and Hinton, G.E. (2008). Visualizing high-dimensional data using t-sne. J. Mach. Learn. Res. *9*, 2579–605.

31. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

32. Azimzadeh, J., Wong, M.L.M., Downhour, D.M.D., Sánchez Alvarado, A., and Marshall, W.F. (2012). Centrosome loss in the evolution of planarians. Science. *335*, 461–3.

33. Duncan, E.M., Chitsazan, A.D., Seidel, C.W., and Alvarado, A.S. (2015). Set1 and MLL1/2 target distinct sets of functionally different genomic loci in vivo. Cell Rep. *13*, 2741–55.

34. Adell, T., Salò, E., Boutros, M., and Bartscherer, K. (2009). Smed-Evi/Wntless is required for β-catenin-dependent and -independent processes during planarian regeneration. Development. *136*, 905–10.

35. Gurley, K.A., Elliott, S.A., Simakov, O., Schmidt, H.A., Holstein, T.W., and Sánchez Alvarado, A. (2010). Expression of secreted Wnt pathway components reveals unexpected complexity of the planarian amputation response. Dev. Biol. *347*, 24–39.

36. Wenemoser, D., and Reddien, P.W. (2010). Planarian regeneration involves distinct stem cell responses to wounds and tissue absence. Dev. Biol. *344*, 979–

91.

37. Zhu, S.J., Hallows, S.E., Currie, K.W., Xu, C., and Pearson, B.J. (2015). A *mex3* homolog is required for differentiation during planarian stem cell lineage development. eLife. *4*, 1–23.

38. Petersen, C.P., and Reddien, P.W. (2009). A wound-induced Wnt expression program controls planarian regeneration polarity. Proc. Natl. Acad. Sci. U. S. A. *106*, 17061–6.

39. Adler, C.E., and Sánchez Alvarado, A. (2015). Types or States? Cellular Dynamics and Regenerative Potential. Trends Cell Biol. *xx*, 1–10.

40. Pearson, B.J., Eisenhoffer, G.T., Gurley, K.A., Rink, J.C., Miller, D.E., and Sánchez Alvarado, A. (2009). A Formaldehyde-based Whole-Mount In Situ Hybridization Method for Planarians. Dev. Dyn. *238*, 443–50.

41. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. *26*, 139–40.

42. Hayashi, T., Asami, M., Higuchi, S., Shibata, N., and Agata, K. (2006). Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. Society, 371–80.

43. Mi, H., Poudel, S., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2016). PANTHER version 10: Expanded protein families and functions, and analysis tools. Nucleic Acids Res. *44*, D336–42.

44. Liu, S.-Y., Selck, C., Friedrich, B., Lutz, R., Vila-Farré, M., Dahl, A., Brandl, H., Lakshmanaperumal, N., Henry, I., and Rink, J.C. (2013). Reactivating head regrowth in a regeneration-deficient planarian species. Nature. *500*, 81–4.

45. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics. *26*, 841–2.

46. Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sánchez Alvarado, A. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. Science. *310*, 1327–30.

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

48. Hopman, A.H., Ramaekers, F.C., and Speel, E.J. (1998). Rapid synthesis of biotin-, digoxigenin-, trinitrophenyl-, and fluorochrome-labeled tyramides and their application for In situ hybridization using CARD amplification. J. Histochem. Cytochem. *46*, 771–7.

49. Rouhana, L., Weiss, J.A., Forsthoefel, D.J., Lee, H., King, R.S., Inoue, T., Shibata, N., Agata, K., and Newmark, P.A. (2013). RNA interference by feeding in vitro synthesized double-stranded RNA to planarians: methodology and dynamics. Dev. Dyn. *242*, 1–43.

50. Katoh, K., Asimenos, G., and Toh, H. (2009). Multiple alignment of DNA sequences with MAFFT. Methods Mol. Biol. *537*, 39–64.

51. Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. *22*, 2688–90.

52. Geer, L.Y., Domrachev, M., Lipman, D.J., and Bryant, S.H. (2002). CDART : Protein Homology by Domain Architecture CDART : Protein Homology by Domain Architecture. Genome Res. *12*, 1619–23.

53. Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: Genome-scale protein function classification. Bioinformatics. *30*, 1236–40.

Chapter 4

Conclusions

I. The Anterior Pole as an Adult Tissue Organizer

Is the planarian anterior pole an organizer? The planarian anterior pole has some molecular and functional similarities to other developmental structures, such as the Spemann-Mangold organizer. Both structures express Wnt and TGF-β inhibitors, and are required for the proper formation of axial structures, albeit during different developmental timeframes [1]. Additionally, Forkhead and Zic family transcription factors are expressed in important signaling centers in other organisms [2–4]. These similarities have led to speculation in the literature that the anterior pole may indeed be an organizer [5,6], although the classical transplantation test had not been done.

Here, we performed transplantation of the head tip into the posterior of host animals, which resulted in the formation of patterned, head-like outgrowths. In planarians, transplantations resulting in positional mismatch lead to outgrowths [7], raising the question of whether any region of the animal could be considered an organizer. However, an organizer would not only induce an outgrowth but also ectopic axial structures. Therefore, as a control for our experiment, we also transplanted regions from the head tip offset from the midline. These pieces also generated outgrowths, but for the most part they were spikes of tissue lacking head structures. The comparison between the head and flank transplantation suggests that the head tip may indeed have organizing activity. The pole outgrowths were much more likely to have anterior PCGs in the correct order (*sFRP-1* and *ndl-2*) as well as a midline (*slit*), and two eyes (*opsin*). We speculate that the few flank transplantations that gave rise to patterned outgrowths likely did so because a few anterior pole cells were present in the donor piece.

One concern with these kinds of experiments is distinguishing between the donor cells inducing changes in host tissue (organization) and the donor cells giving rise themselves to the outgrowth (self-repair). It is suggested in the literature that the lack of lineage tracing techniques in planarians make distinguishing between these cases difficult [5]. However, we circumvented this with our decision to irradiate the donors prior to transplantation. Because neoblasts are the only dividing cells in the animal [8] and irradiation can completely ablate neoblasts [9], our findings demonstrate that most of the cells in the outgrowth must be composed of host cells, which changed their behavior in response to the donor cells. Thus, the outgrowth is a result of organizing activity.

The transplantation experiments as performed here do have some caveats. The techniques available for the specific isolation of a given cell type are limited to a few cell types: the 4C neoblasts using FACS [10], and the epidermis, using the techniques a described in Chapter 3. Thus, the head tip that was selected as the donor for transplantation contains not only the anterior pole, but also other tissues in that region: epidermis, body-wall muscle, and neurons. It is possible that inductive signals from these tissues were instructive for outgrowth, and not the anterior pole. There are two experiments that are technically feasible to perform in planarians that would address these possibilities. RNAi of *foxD* [11] or *zic-1* [12] in uninjured animals leads to the genetic ablation of the pole. Importantly, anterior gene expression domains outside of the anterior pole are still present in these animals [12]. Performing transplantation using these animals as the donors should result in the formation of fewer patterned outgrowths as compared to control RNAi animals. Another potential experiment would

be to inject a cocktail of recombinant anterior pole proteins (*sFFP-1*, *notum*, *follistatin*) and see if they induce patterned, head-like outgrowths.

How are regenerating tails patterned? It turns out there is an analogous signaling structure that forms at posterior-facing wounds known as the posterior pole. The posterior pole is defined by the expression of *wnt1*, and coalesces early in regeneration at about 72 hpa [13]. It requires a different set of transcription factors for its formation including the homeobox genes *islet* [14] and *pitx* [15,16]. Genetic ablation of the posterior pole leads to defects in the regeneration of posterior gene expression domains, as well as the absence of midline markers [15–18]. In addition to *wnt1*, the posterior pole also expresses *wnt11-1*, *wnt11-2*, and *wntP2*, which are broader posterior markers [19]. Further characterization of the posterior pole will be required to assess its potential role as an organizer.

There are still many unanswered questions about the patterning role of secreted proteins from the anterior pole. *notum* is a secreted Wnt antagonist, so it likely acts to preserve the anterior blastema as a low-Wnt environment during regeneration, yet which Wnts are targets of *notum* activity is unclear. *follistatin* expression in muscle is required for the appropriate initiation of the missing tissue response, but its function in anterior pole cells remains to be determined. Are there additional signals secreted by the anterior pole, or can all the functions of the pole be attributed to *notum* and *follistatin*? Finally, uninjured animals that have had their anterior poles genetically ablated appear morphologically normal, suggesting that the anterior pole may not be required for maintenance of tissue pattern once established. More work is needed to fully describe the functions of the pole during regeneration, and potentially homeostasis.

II. Specification of Anterior Pole Progenitors at the Midline

There is a reciprocal relationship between the anterior pole and the midline. RNAi conditions that ablate the anterior pole inevitably lead to defects in the regeneration of the midline at anterior-facing wounds [11,12,17,18,20]. An ectopic pole, achieved through transplantation, results in the formation of a new midline in the outgrowth. These findings suggest that the anterior pole may be both necessary and sufficient for midline induction. Conversely, anterior pole formation is biased for midline [11,12,20], even in fragments that are highly asymmetric. Furthermore, it appears that neoblast specialization is also biased for the midline, suggesting that regulation of the anterior pole's ML position may be determined by this step.

How is the ML axis controlled in planarians? The axon guidance molecule *slit* marks the planarian midline, and *slit* RNAi results in the collapse of structures that are normally separated by the midline [21]. These include neuronal structures, such as the brain and the eyes, and other tissue types, such as the intestine. A collapse in the midline also leads to a narrower anterior pole. *wnt5* is expressed in a lateral ventral domain, running up to the DV boundary. *wnt5* RNAi leads to an expansion of the midline, and the widening or thickening of structures normally present at the midline, such as the brain [19,22]. In the most extreme cases, it can lead to lateral duplications of medial structures such as the pharynx [19], or as reported here, the eyes. As expected, *wnt5* RNAi leads to an expanded or widened anterior pole, and a widening of the specialization domain for anterior pole progenitors.

How is the process of neoblast specialization into anterior pole progenitors at the midline controlled? Neoblasts could be responding to *wnt5* or *slit* directly, or there could be an intermediate signal. Of interest to this question is the wound-induced expression of *foxD*. Very early in regeneration (6 hpa), *foxD* is expressed in the ventral midline abutting the wound. It is possible that the transient expression of *foxD* at anterior-facing wounds may promote the formation of the anterior pole later in regeneration. The molecular signals produced by these cells are unknown.

Why is the anterior pole placed at the prior midline? There are numerous planarian cell types that have symmetry about the midline: brain, gut, eyes, pharynx, secretory cells, and epidermis. An important feature of epimorphic regeneration is that new tissues, once regenerated, must be successfully integrated into pre-existing tissues. By placing the anterior pole at the old midline, the plane of symmetry for regenerating structures will be aligned the plane of symmetry for pre-existing tissues. For example, during head regeneration the bi-lobed brain will need to be regenerated, with one lobe on each side of the midline. The brain will have to send out axonal projections to wire in the pre-existing nerve cords, a process aided by the appropriate alignment of the new brain relative to the old ventral nerve cords. In general, alignment of the new and old midlines (mediated by the anterior pole) will facilitate the integration of new and old tissues during regeneration.

Given the significance of the prior midline for anterior pole placement, what happens to anterior pole formation in a tissue fragment lacking a midline? To understand this, we surgically generated fragments virtually lacking the pre-existing midline. We found that expression of *slit* was ramped up at the wound face, followed by

formation of the anterior pole. Subsequently, the midline was regenerated extending into the body fragment, in some cases almost perpendicular to the old midline. This finding was also documented by Morgan at the turn of the century [23]. Our results suggest that the animals generate a midline *de novo* to ensure anterior pole formation occurs, and then utilize the anterior pole to refine the midline during subsequent steps of regeneration.

III. Anterior Pole Coalescence at the DV Boundary

Primary morphogenetic fields are groups of cells that respond to some positional cue to differentiate in a concentration-dependent manner [24]. It has been suggested that the boundaries between the domains created by a primary morphogenetic field may acquire the ability to act as signaling centers to pattern secondary morphogenetic fields [25]. One of the well-studied examples of this is the *Drosophila* wing-imaginal disc. In the imaginal disc, transcription factors mark the dorsal and ventral domains of the disc. The boundary between these two domains is actively maintained, and is the site of outgrowth and distalization during wing development [26].

Planarians have three distinct orthogonal body axes, and importantly, regeneration of these body axes respects this orthogonality. How is this achieved? Perturbation of *β-catenin-1* by RNAi results in the formation of heads at posterior-facing wounds. It also has a remarkable phenotype in uninjured animals, which develop ectopic heads. It was noted that these ectopic heads form almost exclusively at the DV boundary in intact animals [27–29]. These findings have led to the proposition that the

DV boundary in planarians may act as a signaling center, and that only regions close to the DV boundary are competent for pole formation [30]. Our results suggest instead that the specification of pole progenitors may not rely on the DV boundary, but that the DV boundary may act to guide the positioning of anterior pole progenitor during coalescence.

Why does the anterior pole localize to the DV boundary? The anterior pole mediates AP axial patterning. It may also regulate the vector of growth for head regeneration. Genetic ablation of the anterior pole results in smaller blastemas [11,12,17,18,20], and pole transplantation results in a new vector of growth out of the animal. Placement of the pole at the DV boundary, would in most cases, set the vector of growth for the new AP axis orthogonal to the old DV axis, thus ensuring perpendicularity during regeneration. Perturbation of Bmp signaling results in both dorsally misplaced poles, and aberrant head shape with an indentation at the midline [31–33], suggesting that in these animals the vectors for the AP and DV axes may no longer be orthogonal.

What is the molecular nature of this guidance cue? A number of signaling molecules (*admp* [34,35], *ephR1* [11], *nog1* [36], *serrate*, *wnt5* [19,22]) are expressed at the DV boundary in intact animals. An elegant mechanism would be if a factor expressed at the boundary could serve as an attractive cue for anterior pole progenitors. Important for this model would be the identification of many of the signaling genes expressed at the boundary, as well as the cell types that constitute the DV boundary. There are at least four different cell types present (epidermis [37], muscle [34], neurons [38], and secretory cells [39]) and any of these could potentially be

responsible for secreting the attractive cue. Also significant would be a characterization of DV boundary formation and regeneration in planarians. Bmp signaling is important [31,34], but the exact nature of boundary formation remains mysterious. If the DV boundary is responsible for anterior pole positioning, candidate gene expression domains would need to be regenerated prior to anterior pole formation.

Could the placement of the anterior pole along the DV axis be important for DV patterning? Studies thus far have indicated that ablation of the anterior pole does not appear to affect the DV polarization of regenerating heads [18]. A close examination of the pole transplant outgrowths revealed that they possess few *laminB*+ cells, which may be some of the donor cells, implying that they do not possess a DV boundary. Furthermore, dorsal pole transplant outgrowths have expression of *nlg8*+ throughout the muscle layer of the outgrowth, suggesting that they have entirely dorsal identity. These results, taken together with published findings, suggest that the anterior pole does not have a role in DV patterning.

IV. The Epidermis as a Model to Study Cell Fate Specification

The epidermis is an important organ that exists in essentially all animals. It serves to act as a barrier to the external environment and has acquired many specialized functions in different species. In planarians, the epidermis is a single-cell thick epithelium that covers the entire animal [40]. The ventral epidermal cells are ciliated, and synchronous beating of these cilia are the primary method of locomotion [15]. Dorsal cells are ciliated as well, although the function of these cilia is unclear. The epidermis also contains

rhabdites, which are rod-shaped glycoprotein organelles secreted as a part of mucous production, wound healing, and defense against pathogens [41].

The epidermis is relatively abundant, and its superficial location makes it easy to visualize. It is essential for viability, and thus RNAi conditions that causes defects in epidermal specification or mature epidermal function can be assessed readily. The epidermis turns over rapidly; it is estimated that half of the epidermis is replaced every two weeks [42]. As a large organ, it has regions that are specialized, three of which were known prior to this work: the ciliated epidermis, the non-ciliated epidermis, and the DV boundary epidermis. Markers for each step of the epidermal lineage have been identified and there are established assays for probing epidermal integrity [42]. Finally, as we demonstrated, the mature epidermis can be specifically isolated. For these reasons, we selected the epidermis as a system to study cell fate specification.

Through the course of our work, we generated two datasets valuable for those using the epidermis. Our bulk RNAseq dataset contains genes highly enriched in the mature epidermis. We used this dataset to identify genes expressed in restricted spatial domains, but others could use it to identify genes required for epidermal migration following wounding, ciliogenesis, and mucous secretion. Work prior to ours demonstrated a role for *foxJ1-4* as a master regulator of ciliogenesis in planarians that was highly conserved [43]. The epidermis could be used to study the conserved role of genes in other biological processes, as well as conserved regulatory modules present in epidermis.

Our single-cell sequencing (SCS) dataset is significant in that we have generated transcriptomes for each state of epidermal differentiation in planarians. Our interest was

in understanding how and when spatial patterns emerge during epidermal maturation. Also of interest are the genes required for driving the expression of known markers during differentiation, as well as the genes required for controlling progression from one lineage step to the next. The planarian epidermal lineage could also serve as a model for studying mesenchymal-to-epithelial transition (MET). All lineages start out as mesenchymal neoblasts, but for the epidermis, in addition to a few other cell types, terminal differentiation is accompanied by a MET in order to integrate into the mature tissue. Our transcriptomes likely capture cells before, during, and after this transition. Thus, the dataset presented here opens up many research avenues using the planarian epidermis as a model.

V. Stem Cells Respond to Positional Cues

One of the central questions in regeneration is how animals can appropriately replace their missing tissues. In planarians, a two component system has been proposed to explain how this problem is solved. Neoblasts, which are the only dividing cells in the animal [8], are pluripotent stem cells that are capable of differentiating into any and all of the missing cell types [44]. Body-wall muscle secretes patterning molecules that are the instructions for regeneration [45]. Implicit in this model is the assumption that neoblasts are able to use the information secreted from muscle to guide their decision-making during regeneration and homeostasis. Although not required by the model, it would be particularly elegant if neoblasts could directly read out signals from muscle.

Our work established that mature epidermal cells have DV identities that are not necessarily related to their functional identities (i.e. ciliogenesis). By sorting dorsal and

ventral cells separately, we also showed that these identities were not restricted to mature cells but could be found in progenitors, and identified the DV markers *PRDM1-1* and *kal1*. Using *PRDM1-1* and *kal1*, we were able to show the epidermal neoblasts also possessed DV identity. Perturbation of Bmp signaling resulted in the presence of ventral neoblasts dorsally. The process of ventralization of the lineage was found to require neoblasts. This represents the best evidence that neoblasts can directly respond to positional cues from muscle to decide what to make, a significant step in understanding the nature of planarian regeneration.

VI. Conclusion

Regeneration faces many of the same challenges as development: cell fate specification, coordination of growth, and the patterning of new tissues. Yet it must accomplish these tasks in the context of unpredictable injuries, and integrate new tissues with old to produce functional organs. How these tasks are successfully achieved in the absence of developmental cues is a central question in regeneration. We used the planarian *Schmidtea mediterranea* to ask and answer questions related to patterning during head regeneration and cell fate specification during epidermal differentiation, as situations that exemplify the challenges facing all regenerative organisms.

Animals may use adult tissue organizers during regeneration to guide patterning events [5]. We demonstrated that the anterior head tip in planarians has organizing activity, which can likely be attributed to the anterior pole. The anterior pole integrates

information from pre-existing tissue (the midline) with information about wound geometry (DV boundary) in order to localize correctly during regeneration. Once properly placed, it can secrete signaling molecules to pattern the AP and ML axes of the regenerating head.

Adult animals capable of regeneration, whether they use resident stem cells or dedifferentiation, must properly specify cells to produce appropriate missing tissues. We characterized the planarian epidermal lineage with single-cell resolution, and identified spatial patterns in the mature epidermis and progenitors. Additionally, our work has presented the most convincing evidence to date that neoblasts are capable of responding to gradients established by body-wall muscle.

These findings highlight the importance of using model systems such as planaria to dissect the principles of regeneration. They also underscore how animals are able to regenerate tissues in the absence of embryonic patterning information. Ultimately, understanding which mechanisms are reused from development and which are novel to regeneration will be crucial for advances in regenerative medicine.

References

1. De Robertis, E.M., Larraín, J., Oelgeschläger, M., and Wessely, O. (2000). The establishment of Spemann's organizer and patterning of the vertebrate embryo. Nat. Rev. Genet. *1*, 171–81.

2. Pohl, B.S., and Knöchel, W. (2001). Overexpression of the transcriptional repressor FoxD3 prevents neural crest formation in *Xenopus* embryos. Mech. Dev. *103*, 93–106.

3. Steiner, A.B., Engleka, M.J., Lu, Q., Piwarzyk, E.C., Yaklichkin, S., Lefebvre, J.L., Walters, J.W., Pineda-Salgado, L., Labosky, P.A., and Kessler, D.S. (2006). FoxD3 regulation of Nodal in the Spemann organizer is essential for *Xenopus* dorsal mesoderm development. Development. *133*, 4827–38.

4. Fujimi, T.J., Hatayama, M., and Aruga, J. (2012). *Xenopus* Zic3 controls notochord and organizer development through suppression of the Wnt/β-catenin signaling pathway. Dev. Biol. *361*, 220–31.

5. Vogg, M.C., Wenger, Y., and Galliot, B. How Somatic Adult Tissues Develop Organizer Activity. Curr. Top. Dev. Biol., vol. 116. 1st ed., Elsevier Inc.; 2016, p. 391–414.

6. Owlarn, S., and Bartscherer, K. (2016). Go ahead, grow a head! A planarian's guide to anterior regeneration. Regeneration. *3*, 139–55.

7. Sugino, H. (1937). Transplantation experiments in *Planaria gonocephala*. Jour. Zool. *7*, 373–439.

8. Baguna, J., Saló, E., and Auladell, C. (1989). Regeneration and pattern formation in planarians. III . Evidence that neoblasts are totipotent stem cells and the source of blastema cells. Development. *86*, 77–86.

9. Dubois, F. Contribution a l'etude de la migration des cellules de regeneration chez les Planaires dulcicoles. Bulletin Biologique de la France et de le Belgique, 1949.

10. Hayashi, T., Asami, M., Higuchi, S., Shibata, N., and Agata, K. (2006). Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. Society, 371–80.

11. Scimone, M.L., Lapan, S.W., and Reddien, P.W. (2014). A *forkhead* Transcription Factor Is Wound-Induced at the Planarian Midline and Required for Anterior Pole Regeneration. PLoS Genet. *10*, e1003999.

12. Vásquez-Doorman, C., and Petersen, C.P. (2014). *zic-1* Expression in Planarian Neoblasts after Injury Controls Anterior Pole Regeneration. PLoS Genet. *10*, e1004452.

13. Petersen, C.P., and Reddien, P.W. (2009). A wound-induced Wnt expression program controls planarian regeneration polarity. Proc. Natl. Acad. Sci. U. S. A. *106*, 17061–6.

14. Hayashi, T., Motoishi, M., Yazawa, S., Itomi, K., Tanegashima, C., Nishimura, O., Agata, K., and Tarui, H. (2011). A LIM-homeobox gene is required for differentiation of Wnt-expressing cells at the posterior end of the planarian body. Development. *138*, 3679–88.

15. Currie, K.W., and Pearson, B.J. (2013). Transcription factors *lhx1/5-1* and *pitx* are required for the maintenance and regeneration of serotonergic neurons in planarians. Development. *3588*, 3577–88.

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

17. Blassberg, R.A., Felix, D.A., Tejada-Romero, B., and Aboobaker, A.A. (2013). *PBX/extradenticle* is required to re-establish axial structures and polarity during planarian regeneration. Development. *739*, 730–9.

18. Chen, C.-C.G., Wang, I.E., and Reddien, P.W. (2013). *pbx* is required for pole and eye regeneration in planarians. Development. *729*, 719–29.

19. Gurley, K.A., Elliott, S.A., Simakov, O., Schmidt, H.A., Holstein, T.W., and Sánchez Alvarado, A. (2010). Expression of secreted Wnt pathway components reveals unexpected complexity of the planarian amputation response. Dev. Biol. *347*, 24–39.

20. Vogg, M.C., Owlarn, S., Pérez Rico, Y.A., Xie, J., Suzuki, Y., Gentile, L., Wu, W., and Bartscherer, K. (2014). Stem cell-dependent formation of a functional anterior regeneration pole in planarians requires Zic and Forkhead transcription factors. Dev. Biol. *390*, 136–48.

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

22. Adell, T., Salò, E., Boutros, M., and Bartscherer, K. (2009). Smed-Evi/Wntless is required for β-catenin-dependent and -independent processes during planarian regeneration. Development. *136*, 905–10.

23. Morgan, T.H. (1900). Regeneration in Planarians. Arch. Fur Entwickelungsmechanik Der Org. *10*, 58–119.

24. Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. J. Theor. Biol. *25*, 1–47.

25. Meinhardt, H. (1983). Cell determination boundaries as organizing regions for secondary embryonic fields. Dev. Biol. *96*, 375–85.

26. Dahmann, C., Oates, A.C., and Brand, M. (2011). Boundary formation and maintenance in tissue development. Nat. Rev. Genet. *12*, 43–55.

27. Petersen, C.P., and Reddien, P.W. (2008). *Smed-βcatenin-1* is required for anteroposterior blastema polarity in planarian regeneration. Science. *319*, 327–30.

28. Gurley, K.A., Rink, J.C., and Sánchez Alvarado, A. (2008). β-Catenin Defines Head Versus Tail Identity During Planarian Regeneration and Homeostasis. Science. *319*, 323–7.

29. Iglesias, M., Gomez-Skarmeta, J.L.J., Saló, E., and Adell, T. (2008). Silencing of *Smed-βcatenin1* generates radial-like hypercephalized planarians. Development. *135*, 1215–21.

30. Meinhardt, H. (2009). Beta-catenin and axis formation in planarians. Bioessays.

31, 5–9.

31. Reddien, P.W., Bermange, A.L., Kicza, A.M., and Sánchez Alvarado, A. (2007). BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration. Development. *134*, 4043–51.

32. Molina, M.D., 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. Dev. Biol. *311*, 79–94.

33. Orii, H., and Watanabe, K. (2007). Bone morphogenetic protein is required for dorso-ventral patterning in the planarian *Dugesia japonica*. Dev. Growth Differ., 345–9.

34. Gaviño, M.A., and Reddien, P.W. (2011). A Bmp/Admp Regulatory Circuit Controls Maintenance and Regeneration of Dorsal-Ventral Polarity in Planarians. Curr. Biol. *21*, 294–9.

35. Molina, M.D., Neto, A., Maeso, I., Gómez-Skarmeta, J.L., Saló, E., Cebrià, F., Go, L., and Salo, E. (2011). Noggin and noggin-like genes control dorsoventral axis regeneration in planarians. Curr. Biol. *21*, 300–5.

36. Molina, M.D., Saló, E., and Cebrià, F. (2009). Expression pattern of the expanded noggin gene family in the planarian *Schmidtea mediterranea*. Gene Expr. Patterns. *9*, 246–53.

37. Tazaki, A., Kato, K., Orii, H., Agata, K., and Watanabe, K. (2002). The body margin of the planarian *Dugesia japonica*: characterization by the expression of

an intermediate filament gene. Dev. Genes Evol. *212*, 365–73.

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

39. Zayas, R.M., Cebrià, F., Guo, T., Feng, J., and Newmark, P.A. (2010). The use of lectins as markers for differentiated secretory cells in planarians. Dev. Dyn. *239*, 2888–97.

40. Isao, H., Hori, I., and Isao, H. (1978). Possible role of rhabdite-forming cells in cellular succession of the planarian epidermis. J. Electron Microsc. (Tokyo). *27*, 89–102.

41. Martin, G.G. (1978). A new function of rhabdites: Mucus production for ciliary gliding. Zoomorphologie. *91*, 235–48.

42. van Wolfswinkel, J.C., Wagner, D.E., and Reddien, P.W. (2014). Single-Cell Analysis Reveals Functionally Distinct Classes within the Planarian Stem Cell Compartment. Cell Stem Cell. *15*, 1–14.

43. Vij, S., Rink, J.C., Ho, H.K., Babu, D., Eitel, M., Narasimhan, V., Tiku, V., Westbrook, J., Schierwater, B., and Roy, S. (2012). Evolutionarily ancient association of the FoxJ1 transcription factor with the motile ciliogenic program. PLoS Genet. *8*, e1003019.

44. Wagner, D.E., Wang, I.E., and Reddien, P.W. (2011). Clonogenic Neoblasts Are

Pluripotent Adult Stem Cells That Underlie Planarian Regeneration. Science. *332*, 811–6.

45. Witchley, J.N., Mayer, M., Wagner, D.E., Owen, J.H., and Reddien, P.W. (2013). Muscle Cells Provide Instructions for Planarian Regeneration. Cell Rep. *4*, 1–9.