A Comparative Transcriptomic Analysis Reveals Conserved Features of Stem Cell Pluripotency in Planarians and Mammals

Roselyne M. Labbé[1,3] Manuel Irimia[5], Ko W. Currie[1,2], Alexander Lin[1,2], Shu Jun Zhu[1,3], David D.R. Brown[1,2], Eric J. Ross[4], Veronique Voisin[5], Gary D. Bader[2,3,5], Benjamin J. Blencowe[2,5] and Bret J. Pearson[1,2,3]*

[1] The Hospital for Sick Children, Program in Developmental and Stem Cell Biology; [2] Department of Molecular Genetics, University of Toronto, Toronto, ON, CANADA; [3] Ontario nstitute for Cancer Research, Toronto, ON, CANADA; [4] Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO, 64110, USA.; [5] The Donnelly Centre, University of Toronto, ON, CANADA

ABSTRACT

Many long-lived species of animals require the function of adult stem cells throughout their lives. However, the transcriptomes of stem cells in invertebrates and vertebrates have not been compared, and consequently, ancestral regulatory circuits that control stem cell populations remain poorly defined. In this study, we have used data from high-throughput RNA sequencing (RNA-Seq) to compare the transcriptomes of pluripotent adult stem cells from planarians with the transcriptomes of human and mouse pluripotent embryonic stem cells. From a stringently-defined set of 4,432 orthologs shared between planarians, mice and humans, we identified 123 conserved genes that are ≥ 5-fold differentially expressed in stem cells from all three species. Guided by this gene set, we used RNAi screening in adult planarians to discover novel stem cell regulators, which we found to affect the stem cell-associated functions of tissue homeostasis, regeneration, and stem cell maintenance. Examples of genes that disrupted these processes included the orthologs of *TBL3*, *PSD12*, *TTC27*, and *RACK1*. From these analyses, we concluded that by comparing stem cell transcriptomes from diverse species, it is possible to uncover conserved factors that function in stem cell biology. These results provide insights into which genes composed the ancestral circuitry underlying the control of stem cell self-renewal and pluripotency.

INTRODUCTION

Adult stem cells (ASCs) are capable of substantial cellular feats including organismal longevity, regeneration, and replacement of cells lost during physiological tissue turnover (homeostasis) [1]. In the vertebrate embryo, embryonic stem cells (ESCs) are well-known for their capacity to generate all the tissues of the developing organism (pluripotency) [2]. Both ASCs and ESCs share the ability to self-renew following division. However, unlike ESCs, ASC

daughter cells must adopt a different fate in order to maintain a constant stem cell population in the adult animal. While it is clear that the existence of stem cells is ancestral to multicellular animals [3-5], it is unknown whether stem cells from diverse species share conserved genetic programs for self-renewal, proliferation control, and pluripotency.

Comparative genomic approaches have the potential to shed light on ancestral genetic mechanisms of self-renewal and pluripotency. For example, a comparison of stem cell transcriptomes between different species may yield an overlapping set of conserved genes that are required for fundamental aspects of stem cell biology. In order to discover conserved features of animal stem cells, it is important to compare stem cell expression profiles across large evolutionary distances, such as vertebrates versus invertebrates. Although it is now simple to obtain large quantities of pluripotent ESCs from vertebrates, invertebrate model systems tend to be short-lived and do not produce appreciable numbers of stem cells with which to perform genomic analyses. In addition, pluripotent stem cells have not been described outside of vertebrates, and thus, it is unclear whether genetic programs of pluripotency could be uncovered by comparative methods. A relatively new model system, the freshwater planarian, bypasses these concerns and may be the ideal invertebrate to use for comparative studies of pluripotency.

Planarians are free-living freshwater flatworms of the phylum Platyhelminthes and the super-phylum Lophotrochozoa. Planarians are well known for their extreme regenerative abilities where they can regenerate any amputated body part in about 7 days [6-8]. More recently, studies have shown that planarians have a large population of ASCs that comprise ~10-20% of the cells in the animal [9-12]. Importantly, planarian ASCs have been shown to be pluripotent, which facilitates comparison with other types of pluripotent stem cells [12]. When planarians are irradiated with a lethal dose of

gamma radiation, studies have shown that two populations of cells are sensitive: the stem cell population dies within 24 hours; and the stem cell progeny are lost between 24 hours and 7 days post-irradiation [10,13]. The ability to purify these cell populations in large quantities facilitates comparative genomic analyses [10,14-16]. Finally, gene function can be tested in planarians by RNAi screening in adult animals *in vivo*, and subsequent effects on the stem cell population are readily assayed [9]. The unique biology of planarians can be interrogated with a combination of both RNA deep sequencing (RNA-Seq) and subsequent knockdown of candidate genes [17-20]. These advantages of planarians provide a unique and powerful model organism for the discovery and analysis of conserved, fundamental features of stem cell biology.

In the present study, RNA-Seq was performed on planarian stem cell, stem cell progeny, and differentiated tissues. The sequencing data were pooled to facilitate the *de novo* assembly of a reference planarian transcriptome, which was then used to compare stem cells and stem cell progeny to differentiated tissues. These analyses showed that 2,177 (10.4%) transcripts were specifically expressed $\geq$ 5-fold higher in the planarian stem cell compartment. Interestingly, only 539 of these transcripts were specific to stem cells ($\geq$ 5-fold higher over both stem cell progeny and differentiated tissues), which showed that the stem cell and stem cell progeny share 1,638 transcripts at various levels. Similarly, only 219 transcripts were specific to stem cell progeny ($\geq$ 5-fold higher over both stem cells and differentiated tissues). We then compared the 2,177 planarian stem cell gene signature to both mouse and human ESCs (mESCs and hESCs, respectively). Using multiple criteria of gene homology and expression in planarian and mammalian stem cells, we tested gene function in planarians by RNAi and assayed for defects in stem cell function and maintenance. We provide data for 26 genes that were found to function in planarian stem cells. Most genes were chosen because they

were found to have a single mouse and human ortholog, such as *TBL3*, *PSD12*, *RACK1*, and *TTC27*, which can now be tested for their roles in mammalian stem cells. These data showed that a comparative transcriptomics approach to stem cell biology can identify conserved factors which may reflect conserved mechanisms of pluripotency.

<div align="center">

MATERIALS AND METHODS

</div>

FACS purification, RNA isolation, RNA sequencing

We used flow cytometry (FACS) to obtain a cell population highly purified for planarian stem cells as previously described [13]. To obtain differentiated tissues lacking both stem cells and stem cell progeny, lethally-irradiated worms, 7 days post-irradiation, were collected and sequenced. A lethal dose of gamma irradiation (60 Gy.) renders planarians devoid of stem cells and their lineage-progeny, yet the treated worms are still morphologically normal [10]. The irradiated sample was referred to as "differentiated tissues". Approximately 1.3 million stem cells and 1 million stem cell progeny were isolated on a Becton-Dickinson FACSaria over 7 independent sorts, and then pooled. RNA for each sample was subjected to next generation sequencing for comparison of differential expression enrichment in the stem cell compartment. RNA samples were prepared using Illumina TruSeq RNA-Seq kits. RNA sequencing was performed on an Illumina HiSeq2000 with paired-end 100 base pair reads and approximately 150 million reads were collected for both stem cells and irradiated tissues, and 300 million reads collected for stem cell progeny at the sequencing core facility at the Hospital for Sick Children (www.tcag.ca). RNA-Seq data have been submitted to NCBI Gene Expression Omnibus (GEO), with accession number GSE37910. Mouse and human ESC, NPC and adult tissue sequencing datasets were obtained from GEO projects SRA012498, GSM591658, and published data [21-23].

Transcriptome assembly, sequence alignments, and determining gene orthology

RNA-Seq was pooled and assembled into a transcriptome *de novo* using the Trinity software package with default settings [24]. This produced 48,833 transcripts with an average length of 921.6 nucleotides. We discarded any transcripts encoding ORFs < 100 amino acids, and 21,021 transcripts remained, with an average length of 1004.6 nucleotides. These transcripts were then compared to NCBI NR protein database using BLASTX. In cases where BLAST e-values were > $1 \times e^{-5}$, we determined transcript direction based off of longest predicted ORF.

For gene expression quantification in the three species, all reads were trimmed to 50 nucleotides, and only the forward end was used, when paired-ends existed. Reads were then mapped to the transcriptome using bowtie [25] with −m 1 −v 2 parameters (requiring unique mapping and two or less mismatches across the full alignment). Correction for multiple mapping was done as follows: each position in each transcript using 50-basepair windows was mapped back against the whole transcriptome using bowtie with the same parameters. If this sequence mapped somewhere else in addition to itself it was discarded and discounted from the transcript effective length (length-49). This "effective length" was then used to divide the raw read counts per million mapped reads for each gene to obtain corrected-RPKM values (cRPKM).

For human and mouse, only the canonical transcripts as defined by Ensembl (via Biomart) were used for gene expression quantification. Our cRPKM values with canonical transcripts correlate very well with those obtained by Brawand *et. al.* [21] (r > 0.92 for all correlations between $\log_2$-scaled data), who used a much more sophisticated method, that cannot be used in planarians due to lack of similar annotations (repeat and pseudogene annotation, constitutively spliced exons, etc).

To determine the 1:1:1 orthologs of planarian genes to mammals, we used pairwise BLASTX of all transcripts from each species using –F F (not filtering out low complexity regions). Only best reciprocal hits for the three pairwise comparisons were used for conservation analyses. Cluster analysis was performed using R function hclust with default settings ("complete" method) and Euclidean distance matrix for $\log_2$-scaled cRPKM values. Only 1:1:1 orthologs with ≥ 2-fold in cRPKM ratios for any of the three pairwise comparisons between planarian samples were used (1691 genes in total).

Gene set enrichment analysis (GSEA)
The list of the non-redundant 4,432 planarian-mammal orthologous genes was used to analyze pathways modulated in one tissue (stem cell, stem cell progeny or differentiated tissue compartments) versus another using GSEA pathway analysis software [26]. Data was $\log_2$ transformed and cRPKM values lower than one were discarded. This produced normally distributed data (Figs. S1-2), which were then ranked using the log fold change of each pairwise comparison. The data were analyzed using GSEA with the parameters of 2000 gene-set permutations, gene-sets size between 8 and 500, and weight set to p2. The gene-sets included in the GSEA analyses were obtained from KEGG, MsigDB-c2, NCI, Biocarta, IOB, Netpath, HumanCyc, Reactome and the Gene Ontology (GO) databases, updated January 2011 (http://baderlab.org/GeneSets). An enrichment map (version 1.2 of Enrichment Map software [27,28]) was generated for each comparison using enriched gene-sets with a nominal *p*-value < 0.001, False Discover Rate < 1% and the overlap coefficient set to 0.5.

Animal husbandry, cloning, RNAi, and *in situ* hybridizations
Asexual *Schmidtea mediterranea* strain CIW4 was reared as previously described [29]. Primers were designed and full length genes were cloned by 3' RACE. RNAi experiments were performed using previously described expression constructs and HT115 bacteria [9]. The control RNAi construct was the *C. elegans* gene *unc22*, which does not have homology in the planarian genome [9]. To make RNAi food, bacteria were grown to an O.D.600 of 0.6-0.8 and induced with 1mM IPTG for 2 hours. Bacteria were pelleted and mixed with liver paste at a ratio of 500µl of liver per 100ml of original culture volume. Bacterial pellets were thoroughly mixed into the liver paste and frozen as aliquots. All RNAi food was fed to 7-day starved experimental worms every 3rd day for 3-4 total feedings depending on the gene, although the majority of genes were fed 4 times. From the screen of 100 planarian stem cell specific genes, we have only reported the 26 genes that gave 100% penetrant phenotypes at 4 feedings or less. Primer sequences used to clone these genes are given in Table S6. The time points of images and stains in Figs. 5-7 reflect time from the beginning of the experiment. Images of intact worms were taken as soon as morphological phenotypes were observed. Pre- and post-pharyngeal head amputations were performed 3 days after the final feeding (cutting worms into thirds). All animals used for *in situ* hybridizations (ISH) were 2-3 mm in length and size-matched between experimental and control worms, and individual RNAi experiments were performed on batches of approximately 40 worms. Animals were collected for ISH at 6 days post-RNAi, as well as at a second time point, specific to each gene, when animals began showing more severe phenotypes. ISHs were performed as previously described and at least 10 worms were used per stain, and representative animals shown [30]. All experiments were performed with a minimum of three replicates. All images were taken on either a Leica M80 or a Leica M165 dissecting microscope, and were processed in Adobe Photoshop.

RESULTS

Determining transcripts specific to planarian stem cells, stem cell progeny, and differentiated tissues
To define the planarian stem cell transcriptome, we performed an RNA-Seq analysis on three

distinct planarian cellular compartments: stem cells, stem cell progeny, and differentiated tissues (Fig. 1A; see Methods) [13-15]. To determine differential gene expression (DE) between each compartment, a planarian reference transcriptome was assembled *de novo* using the program Trinity, which produced 21,021 transcripts (Table S1; see Methods) [24]. The reference transcriptome was subsequently used to align reads from each planarian sample, which gave an expression level for each transcript as a cRPKM value. To determine whether a given transcript was DE in one cell compartment over another, cRPKM ratios were determined. It is important to note that because stem cells and stem cell progeny were FACS-purified, transcripts can be ranked by cRPKM values alone, by fold-change over differentiated tissues, or as one over the other.

It has recently been shown that although irradiation renders planarians devoid of stem cells and stem cell progeny, the procedure is not neutral to differentiated tissues, and worms will undergo transcriptional changes as a result [16]. However, it has also been established, by multiple studies, that comparing whole-irradiated to whole-non-irradiated planarians allows for the identification of many genes, specifically expressed in stem cells, despite no FACS purification [10,31]. To determine whether our data and comparisons were accurate, we examined previously published genes and how they ranked in our dataset by both cRPKM alone and fold-DE over differentiated tissues (Table S2). We observed that fold-DE showed a better rank for most previously reported genes, whereby analysis of 44 "control" genes were enriched as expected with an average of 35.8-fold DE in stem cells over differentiated tissues (Table S2) [11,14,32-37]. Furthermore, two genes described to be specific to the stem cell progeny, *prog-1* and *prog-2*, were the top two genes in our stem cell progeny data and were both > 217-fold DE over differentiated tissues. Thus, we concluded our method was valid and our data highly accurate. Using fold-DE as the best method to rank stem cell specific genes, we

found that a total of 2,177 planarian transcripts were ≥ 5-fold DE in stem cells over differentiated tissues (Fig. 1B). Interestingly, only 539 of those transcripts were also ≥ 5-fold DE over stem cell progeny, which suggested that 1,638 transcripts are shared between stem cells and stem cell progeny at various levels below the 5-fold DE cutoff. Finally, we found that only 219 transcripts were specific to the stem cell progeny (≥ 5-fold DE over stem cells and differentiated tissues).

Determining orthology between planarian and mammalian genes establishes a set of 4,432 gene orthologs
Currently, there is no extensive gene ontology (GO) database for planarian genes, and without which, it is difficult to reliably assign genetic pathways expressed in one planarian cell compartment versus another. In addition, to compare the DE of planarian genes in the stem cell compartment with that of mammalian ESCs, it was necessary to assign homology between them. To approach this problem, we mapped gene orthologs between the planarian, human, and mouse, which enabled the use of GO databases generated for mouse and human genes. To minimize false-positive orthology assignments, reciprocal-BLAST searches were performed between translated planarian transcripts and mouse and human protein sequences, resulting in 1:1:1 orthologs shared by all three species (Table S3; see Methods; ref. [38]). Using this method, a total of 4,432 of planarian transcripts were assigned as orthologs to both mouse and human genes. This gene set was used for all subsequent comparative analyses (Figs 3, 4A, 6 and 7).

Gene set enrichment analysis reveals specific genetic pathways associated with planarian stem cells, stem cell progeny, and differentiated tissues
Gene set enrichment analysis (GSEA) is a powerful method to analyze genetic pathways enriched in a particular dataset, and was recently used to show that leukemias expressing a stem cell gene set correlated with poor survival

[26,39]. Using the set of 4,432 gene orthologs, we performed GSEA for each of these pairwise comparisons: 1) stem cells vs. stem cell progeny; 2) stem cells vs. differentiated tissues; and 3) stem cell progeny vs. differentiated tissues (Fig. 2A-C). To minimize false positives, pathways and gene sets were only considered if their significance was < 0.001. Pathway analysis of stem cells vs. stem cell progeny revealed significant enrichment of cell cycle and DNA replication pathways in the stem cell compartment, which is consistent with previous reports that stem cells are the only cycling cells in planarians (Fig. 2A) [10]. In addition, we observed that genes associated with the *Retinoblastoma* (*Rb*) pathway were also specifically enriched in the stem cell fraction, consistent with previous observations (B.J.P. unpublished data, and ref. [35]). In contrast, GSEA revealed that stem cell progeny were enriched in pathways associated with development, differentiation, as well as several other functional categories including cell migration, response to stimulus, and signal transduction pathways (Fig. 2A).

GSEA of stem cells vs. differentiated tissues revealed significant enrichment in the stem cell compartment for pathways that function in cell cycle, DNA repair, and RNA processing/transport, the latter of which could reflect higher overall transcriptional levels in stem cells relative to differentiated tissues (Fig. 2B). Consistent with previous data, we also detected significant enrichment of the *p53*, *Rb* and Fanconi pathways in stem cells [33,35]. Pathways with significant enrichment in differentiated tissues were involved in processes that should not occur in stem cells, as expected. These included gene sets predicted to function in and maintain the nervous system (ion transport, synaptic transmission, synapse development, and neurological system processes), as well as pathways implicated in cell adhesion, muscular maintenance, and osmoregulation (Fig. 2B).

Finally, GSEA was used to compare pathways in stem cell progeny vs. differentiated tissues (Fig.

2C), which also revealed enrichment of the *p53* pathway, consistent with previous results [33]. Additionally, spliceosome and chromatin modification pathways were significantly enriched in stem cell progeny, as were the pathways of DNA replication and cell cycle checkpoints (Fig. 2C). Interestingly, where nervous system pathway enrichment was observed for differentiated tissues relative to stem cells (Fig. 2B), no such enrichment was observed when compared to stem cell progeny. This suggested that both of these cell compartments have comparable enrichment for nervous system pathways. This result was similar to how both stem cells and stem cell progeny are known to require the *p53* pathway, and as such, neither cell type was enriched for it [33]. In conclusion, although only approximately 20% of planarian genes were used in the GSEA, the method successfully detected pathways previously known to be specific to each compartment, with no *a priori* knowledge, using mammalian GO annotations. In addition, GSEA supported the complicated nature of the "stem cell progeny", which were enriched for cell cycle, *p53*, differentiation, and nervous system pathways (see Discussion).

Comparison of planarian and mammalian stem cell signatures

To compare DE in planarian ASCs to mammalian ESCs, DE in mammalian ESCs was first determined. To this end, available RNA-Seq data from hESCs (both H1 and H9 cells) and mESCs (V6.5 cells) were used [22,23]. For the mammalian equivalent of the "differentiated tissues" sample used for planarians, the average cRPKM value from RNA-Seq of six differentiated tissues for each species was determined (brain, cerebellum, heart, kidney, liver, testis) [21]. To validate that this method produced a set of ESC-specific genes, DE in human and mouse ESCs was compared to a previously established list of genes associated with pluripotency in various cell types (but not necessarily specific to them), referred to as the "PluriNet" gene set [40]: 212/278 (76%) and 215/278 (77%) of PluriNet genes were ≥ 2-fold

DE by our method in both hESCs and mESCs, respectively.

Next, in order to discover genes that were DE in both mammalian ESCs and planarian ASCs, only the orthologous set of genes was used for comparisons. Of the 4,432 orthologous genes, 538 in mESCs, 421 in hESCs, and 605 in planarian ASCs were found to be ≥ 5-fold DE in the stem cell compartment over their respective differentiated tissues (Fig. 3D). Remarkably, the percentage of gene overlap was similar when comparing planarians to each mammalian dataset (46% to human and 37% to mouse; Fig. 3D), as it was between ESCs of mouse and human (44-56%; Fig. 3C). Importantly, 123 genes were found to be specifically ≥ 5-fold DE in the stem cells of all three species (Fig. 3D). This overlap was even more striking for PluriNet genes, of which most (average of 79%) were up-regulated in the stem cell populations of all three species (Fig. 3E).

Planarian stem cell progeny are significantly enriched for transcription factors and genes associated with nervous system development
Planarian stem cell progeny are thought to be of restricted cellular potential and GSEA suggested a possible involvement for these cells in nervous system function or maintenance. We thus hypothesized that planarian progeny may share regulatory similarities with neuronal precursor cells in mammals (NPCs), which have restricted potential and perform critical roles in neuronal development and differentiation. To test this, we first used available RNA-Seq data for mouse and human NPCs to identify genes DE over differentiated tissues and ESCs, which were then compared to genes DE in planarian stem cell progeny (Figs. 1B, 4A) [21-23]. The resulting overlap of specific genes among the three samples was much less pronounced than between stem cell types (Fig. 3), and even when a ≥ 1.5-fold DE cutoff over both differentiated tissues and stem cells was applied, we detected only 22 genes shared between all three organisms (Fig. 4A). Despite low gene overlap, the genes shared were well-known regulators of

nervous system development such as: *Eya1, Fyn, Notch2, Pou3f3, Sema6D,* and *Pax6*. This result was consistent with the significant enrichment we observed for the GO categories of "transcription factor activity" and "nervous system development" (Table S4; Fig. 4A; both *p*-values = 0.003).

Even though we found only 219 transcripts to be specific to planarian stem cell progeny (Fig. 1B), it was interesting that this cell type was significantly enriched for the GO category of "transcription factor activity". To further explore this observation, we next assigned all planarian transcripts to major transcription factor families (Fig. 4B-C). This confirmed that despite the small number of stem cell progeny-specific genes compared to stem cells and tissues (219 genes vs. 539 and 3424), stem cell progeny expressed nearly twice as many specific transcription factors than stem cells and differentiated tissues (*p* < 0.0001 by proportion test vs. stem cells). Interestingly, the profile of specific transcription factors was different for each planarian compartment (Table S5). From these analyses, we concluded that high transcription factor diversity in the planarian stem cell progeny reflects real and diverse heterogeneity in this population of cells (see Discussion).

Comparisons between planarian and mammalian stem cells revealed novel regulators of stem cell biology in planarians
RNA-Seq analyses identified over 2,000 genes specifically DE in planarian stem cells. Because these genes cross many arbitrary thresholds of fold-DE, cRPKM values, and homology with mammals, it was logical to divide this gene set into three broad categories prior to functional validation (represented in Fig. 3B, D). Briefly, these categories were: (1) genes with no 1:1:1 ortholog but ≥ 5-fold DE in planarian stem cells (Figs. 3B, 5); (2) genes with a 1:1:1 ortholog and ≥ 5-fold DE in planarian stem cells, but no DE in mammalian ESCs (Figs. 3D, 6), and; (3) genes with a 1:1:1 ortholog and ≥ 5-fold DE in mouse, human and planarian stem cells (Figs. 3D, 7). To

assess whether genes from each category included novel regulators of stem cell biology, RNAi was used to test their functions in planarians. Following knockdown, three different processes involving stem cell function were measured in order to demonstrate gene requirements in planarian stem cells (Figs. 5-7). First, defects in tissue homeostasis were assayed because this process happens at a rapid rate in planarians and depends on the activity of stem cells. Homeostatic defects can be as subtle as epithelial lesions, or as drastic as head loss [32]. Second, we assayed the ability of planarians to regenerate missing head structures, a process that depends on proper stem cell function. Third, as a more direct measure of effects on the stem cell population in whole animals, the expression of the stem cell marker *piwi-1* was monitored following RNAi (as was the stem cell marker *H2A* to ensure that *piwi-1* loss was not simply *piwi-1* down-regulation; Fig. S9) [14]. Using these three assays, 26 genes out of approximately 100 that were screened, gave multiple stem cell defective phenotypes at 100% penetrance. These genes represented all three expression and homology categories described above.

Following RNAi of genes belonging to category (1), a variety of effects on tissue homeostasis were observed. For example, knockdown of a homolog to eukaryotic high mobility group-like nuclear protein 2 (*NHP2-like)*, implicated in RNA processing [41], resulted in head regression, which is typical of a stem cell defect. Knockdown of other genes in category (1) gave dorsal lesions of various patterns, which is also indicative of a stem cell defect (*e.g. SMD1, FBRL, NOP2, RACK1*). Regardless of the severity of the observed homeostatic defects, head regeneration ability, as assayed at 7 days post-amputation, was severely diminished (Fig. 5). Animals were fixed and stained for *piwi-1* at both 3 days post-RNAi and at a gene specific, terminal time point, prior to animal death (Fig. 5). We observed that the stem cell population was decreased in all cases, except for RNAi against the homolog of mammalian guanine

nucleotide-binding protein subunit beta-like protein (*RACK1*). RNAi knockdown of *RACK1* led to minimal disruption of the stem cell compartment, as measured by *piwi-1* staining (Fig. 5). This result was particularly interesting because loss of *RACK1* function in human cells has been shown to enhance tumor growth and migration [42]. Importantly, for all other genes reported, we observed that *piwi-1*-expressing cells could be lost in a spatial-specific manner, depending on the gene. For example, in *NOP2(RNAi)*, animals first lost the lateral population of stem cells (Fig. S3). This was in contrast to *SMD1(RNAi)* animals, which preferentially lost stem cells from the posterior (Fig. S4). From these data, we concluded that genes belonging to category (1) include novel regulators of planarian stem cells.

As in the case of category (1) genes, knockdown of planarian genes with mammalian orthologs in category (2) produced a variety of stem cell defects in tissue homeostasis, such as dorsal lesions (*TBL3, TDRD9*) and head regression (*TDRD9, PHB, RPA2*) (Fig. 6). When head regeneration was assayed, all genes showed loss or dramatic reduction of regenerative ability. Finally, RNAi of each gene resulted in loss of the stem cell population. Similar to category (1) genes, we observed that as the stem cell loss progressed, it often had a spatial bias. For example, we observed regional loss of *piwi-1*-expressing cells in the posterior (*TDRD9, RPA2*), the anterior (*C1orf107, PSMC4*) or a nearly complete loss throughout (*TBL3, PHB*) (Figs. 6 and S5-6). From these data, we concluded that orthologous genes DE in planarian stem cells, but not DE in mammalian ESCs, contain novel stem cell regulators in planarians.

Finally, RNAi of category (3) genes also disrupted stem cell processes in planarians (Fig. 7). Homeostatic defects included dorsal lesions (*PSD12, TTC27, RUVB2, NOP58, NUP93*) and head regression (*CSTF3, CBX3, PLK*). Head regeneration was found to be abolished in all RNAi treatments. Upon assay of the stem cell

population with *piwi-1*, we observed that knockdown of genes belonging to category (3) also had different spatial biases on stem cell loss (Figs. 7 and S7). As we observed for knockdown of genes belonging to the other two categories, RNAi of category (3) genes produced spatial-specific losses of the stem cell population. Patterns of stem cell loss were observed to occur specifically in either the posterior, anterior, midline, and body edge.

Despite substantial differences in homology or expression levels, neither gene category contained a higher percentage of genes that gave phenotypes. In addition, all the genes reported were 100% penetrant and lethal. Finally, even though consistent and specific spatial loss of *piwi-1*-expressing cells for various genes was observed, the expression of the genes themselves did not appear to be spatially restricted (Fig. S8). In total, by using different criteria of homology and expression levels in mammalian ESCs, we have identified novel and highly conserved regulators of stem cell biology in planarians that can now be tested in mammalian stem cell systems.

<div align="center">

DISCUSSION

</div>

In this study, we took a comparative transcriptomic approach between invertebrate and mammalian stem cells. We hypothesized that because planarian ASCs and mammalian ESCs are both pluripotent, pluripotency is an ancestral biological process and therefore pluripotent stem cell types are likely to use conserved genetic programs. Accordingly, we examined the expression of the 4,432 gene orthologs shared between mouse, human, and planarians, we found that 610 (13.8%) displayed $\geq$ 2-fold DE in the stem cells of all three species. This result was particularly striking considering the large evolutionary distances involved, and suggested that gene expression patterns in pluripotent stem cells have been extensively conserved to an unappreciated extent. To support our comparative analyses, we used RNAi screening in adult planarians to functionally

validate the stem cell requirement for a total of 26 novel genes, chosen from 3 categories of gene homology or overlap with DE in mammalian ESCs (Figs. 3, 5-7). Knockdown of the genes in each category displayed a clear phenotype associated with multiple stem cell defects in the processes of tissue homeostasis, regeneration, and stem cell maintenance. Interesting examples of gene knockdowns which displayed pronounced stem cell associated phenotypes include *NOP2*, *RACK1*, and *NHP2L* (category 1), *RPA2*, *TBL3*, and *TDRD9* (category 2), and *NOP58*, *TTC27*, and *PSD12* (category 3). These examples have not been widely implicated in stem cell biology.

Multiple approaches have identified planarian ASC-specific transcripts

There are currently three primary methods to identify transcripts that are specific to planarian ASCs. First, stem cell specific transcripts can be discovered by simply comparing whole, irradiated to whole non-irradiated worms. Using cDNA microarrays, this method was first used by Eisenhoffer, *et. al*., 2008 and more recently with whole-genome microarrays by Wagner, *et. al*., 2012. While it was not possible to determine the false-positive/negative rates in either study, it was clear that this method is valid. However, both Eisenhoffer, *et. al*., and a recent study by Solana, *et. al*., 2012, using RNA-Seq, showed that irradiation is indeed not neutral to tissues, so this caveat must be taken into account when using irradiation. The second method, used by Solana, *et. al*., applied RNAi to ablate stem cells, then used those animals as the equivalent of irradiated animals. This method also led to successful identification of stem cell specific genes, however, it has the caveats of potential off-target effects of the RNAi being used, transcriptional changes invoked by the RNAi process itself, and transcriptional changes caused by the delivery method of RNAi (feeding or injection). While the above study used RNA-Seq to measure transcript levels, no stem cell purification was applied, and thus having all tissues present can also complicate analyses and diminish the dynamic range of the sequencing.

Finally, the third method to discover planarian stem cell specific transcripts, which was used in this study, is to purify the stem cells by FACS, then compare against irradiated tissues. FACS purification has the advantage of containing the least amount of contaminating tissues, so that all transcripts sequenced are expressed in stem cells (within the limits of FACS). We chose to compare purified stem cells with whole, irradiated planarians at 7 days post-irradiation because the worms are still morphologically and behaviorally normal, and many of the transcripts induced by irradiation observed in Eisenhoffer, *et. al.*, have returned to un-irradiated levels. Our method has the same caveats of using irradiated worms discussed above, as well as the accuracy and physical effects of FACS. The physical effect of flow cytometry on the stem cell transcriptional profile is likely to be minor because the stem cells can be recovered following FACS and survive subsequent transplantation into an intact worm [12]. Furthermore, several decades of studies on mammalian hematopoiesis have successfully used FACS to isolate, profile, and transplant living cells. Finally, our analyses identified all previously known planarian stem cell genes, suggesting that the method is accurate and has a large dynamic range not observed using other methods (Table S2). In conclusion, it is clear that each of these three methods is valid, and all have been successfully applied to the identification of stem cell specific genes. In the future, it will be interesting to perform meta-analyses to determine which method is the most accurate for any given application. Perhaps the highest accuracy will be obtained using combinations of approaches such as stem cell FACS-purification compared against whole worms that have been depleted of stem cells by RNAi.

Analysis of the stem cell progeny
Due to the fact that clonal lineage tracing is not yet possible in planarians, there are no conclusive data demonstrating that cells from the stem cell-FACS fraction are heterogeneous in their cellular potential. In contrast, markers from the stem cell progeny-FACS fraction are spatially restricted and are likely to be highly heterogeneous (also supported by our transcription factor analysis) [10]. Although this study did not focus on genes or phenotypes associated with the stem cell progeny, several interesting observations were made during our analyses. Similar to the analysis of previously known stem cell specific genes, the two genes that are known to mark the earliest-born progeny of stem cells, had the two highest DE in our analysis (Table S2) [10]. Another gene associated with slightly later-born progeny, *AGAT1*, had the 19th highest DE in our analysis. However, we were surprised to find that the stem cell marker *piwi-1*, had the 81st highest DE in the stem cell progeny over differentiated tissues. This may reflect one of two possibilities. Either the stem cell progeny still express *piwi-1* RNA as they transit to differentiation, or there are *bona fide* stem cells in the "stem cell progeny" FACS fraction. The former possibility is supported by the facts that: (1) early progeny markers and *piwi-1* have been shown to overlap at the protein and RNA levels; and (2) approximately 10% of sorted stem cell progeny express *piwi-1* [10,11,14,33,36]. The latter possibility is supported by our analysis that stem cells and stem cell progeny share 1,638 transcripts that they express at various fold-DE levels over differentiated tissues, while stem cell progeny only express 219 unique transcripts (Fig. 1B). If significant numbers of stem cells FACS-purify along with the stem cell progeny, this may represent stem cells in G_0/G_1 that would not be expected to sort into the high-DNA gate used to collect stem cells. Finally, because known progeny markers and a disproportionately large number of transcription factors are highly expressed in the stem cell progeny, this suggests that these cells are heterogeneous in their competence and biology.

In conclusion, our results suggest that planarians regulate stem cell-specific programs via the action of both unique and highly conserved genes. Using the combination of FACS purification of stem cells with RNA-Seq, novel stem cell regulatory genes can be identified. We

hypothesize that genes with DE in stem cells in both mammals and planarians represent part of a core, ancestral, stem cell network in metazoans, although further experiments and additional species will be required to test this. As supported by the RNAi experiments presented in this study, genes that were found to have DE in planarian, human, and mouse stem cell populations, were shown to be regulators of planarian stem cells. In future studies, it will be interesting to test whether these genes function similarly in other stem cell systems.

ACKNOWLEDGMENTS

Conflicts

There are no conflicts of interest with this manuscript.

REFERENCES

1. Simons, B. D. & Clevers, H. Strategies for homeostatic stem cell self-renewal in adult tissues. *Cell* 145, 851-862, doi:S0092-8674(11)00594-0 [pii] 10.1016/j.cell.2011.05.033 (2011).
2. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* 414, 105-111 (2001).
3. Noh, K. H. *et al.* Cancer Vaccination Drives Nanog-Dependent Evolution of Tumor Cells toward an Immune-Resistant and Stem-like Phenotype. *Cancer Res* 72, 1717-1727, doi:0008-5472.CAN-11-3758 [pii] 10.1158/0008-5472.CAN-11-3758 (2012).
4. Watanabe, H., Hoang, V. T., Mattner, R. & Holstein, T. W. Immortality and the base of multicellular life: Lessons from cnidarian stem cells. *Semin Cell Dev Biol* 20, 1114-1125, doi:S1084-9521(09)00173-6 [pii] 10.1016/j.semcdb.2009.09.008 (2009).
5. Bosch, T. C. Hydra and the evolution of stem cells. *BioEssays* 31, 478-486, doi:10.1002/bies.200800183 (2009).
6. Morgan, T. H. Experimental studies of the regeneration of Planaria maculata. *Arch. Entw. Mech. Org* 7, 364-397 (1898).
7. Brøndsted, H. V. *Planarian regeneration*. [1st] edn, (Pergamon Press, 1969).
8. Inoue, T. *et al.* Morphological and functional recovery of the planarian photosensing system during head regeneration. *Zoolog Sci* 21, 275-283 (2004).
9. Newmark, P. A., Reddien, P. W., Cebria, F. & Sánchez Alvarado, A. Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians. *Proc Natl Acad Sci U S A* 100 Suppl 1, 11861-11865 (2003).
10. Eisenhoffer, G. T., Kang, H. & Sanchez Alvarado, A. Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian Schmidtea mediterranea. *Cell Stem Cell* 3, 327-339 (2008).
11. Guo, T., Peters, A. H. & Newmark, P. A. A bruno-like Gene Is Required for Stem Cell Maintenance in Planarians. *Dev Cell* 11, 159-169 (2006).
12. Wagner, D. E., Wang, I. E. & Reddien, P. W. Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* 332, 811-816, doi:332/6031/811 [pii] 10.1126/science.1203983 (2011).
13. Hayashi, T., Asami, M., Higuchi, S., Shibata, N. & Agata, K. Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. *Dev Growth Differ* 48, 371-380 (2006).

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

15. Hayashi, T. *et al.* Single-cell gene profiling of planarian stem cells using fluorescent activated cell sorting and its "index sorting" function for stem cell research. *Dev Growth Differ* 52, 131-144, doi:DGD1157 [pii] 10.1111/j.1440-169X.2009.01157.x (2010).

16. Dingli, D. & Pacheco, J. M. Stochastic dynamics and the evolution of mutations in stem cells. *BMC Biol* 9, 41, doi:1741-7007-9-41 [pii] 10.1186/1741-7007-9-41 (2011).

17. Solana, J. *et al.* Defining the molecular profile of planarian pluripotent stem cells using a combinatorial RNA-seq, RNAi and irradiation approach. *Genome Biol* 13, R19, doi:gb-2012-13-3-r19 [pii] 10.1186/gb-2012-13-3-r19 (2012).

18. Blythe, M. J. *et al.* A dual platform approach to transcript discovery for the planarian Schmidtea mediterranea to establish RNAseq for stem cell and regeneration biology. *PLoS One* 5, e15617, doi:10.1371/journal.pone.0015617 (2010).

19. Abril, J. F. *et al.* Smed454 dataset: unravelling the transcriptome of Schmidtea mediterranea. *BMC Genomics* 11, 731, doi:1471-2164-11-731 [pii] 10.1186/1471-2164-11-731 (2010).

20. Sandmann, T., Vogg, M. C., Owlarn, S., Boutros, M. & Bartscherer, K. The head-regeneration transcriptome of the planarian Schmidtea mediterranea. *Genome Biol* 12, R76, doi:gb-2011-12-8-r76 [pii] 10.1186/gb-2011-12-8-r76 (2011).

21. Brawand, D. *et al.* The evolution of gene expression levels in mammalian organs. *Nature* 478, 343-348, doi:10.1038/nature10532 nature10532 [pii] (2011).

22. Guttman, M. *et al.* Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol* 28, 503-510, doi:nbt.1633 [pii] 10.1038/nbt.1633 (2010).

23. Wu, J. Q. *et al.* Dynamic transcriptomes during neural differentiation of human embryonic stem cells revealed by short, long, and paired-end sequencing. *Proc Natl Acad Sci U S A* 107, 5254-5259, doi:0914114107 [pii] 10.1073/pnas.0914114107 (2010).

24. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29, 644-652, doi:10.1038/nbt.1883 nbt.1883 [pii] (2011).

25. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10, R25, doi:gb-2009-10-3-r25 [pii] 10.1186/gb-2009-10-3-r25 (2009).

26. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550, doi:0506580102 [pii] 10.1073/pnas.0506580102 (2005).

27. Merico, D., Isserlin, R. & Bader, G. D. Visualizing gene-set enrichment results using the Cytoscape plug-in enrichment map. *Methods Mol Biol* 781, 257-277, doi:10.1007/978-1-61779-276-2_12 (2011).

28. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5, e13984, doi:10.1371/journal.pone.0013984 (2010).

29. Sánchez Alvarado, A., Newmark, P. A., Robb, S. M. & Juste, R. The Schmidtea mediterranea database as a molecular resource for studying platyhelminthes, stem cells and regeneration. *Development* 129, 5659-5665 (2002).

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

31. Shackleton, M., Quintana, E., Fearon, E. R. & Morrison, S. J. Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* 138, 822-829, doi:S0092-8674(09)01030-7 [pii] 10.1016/j.cell.2009.08.017 (2009).

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

33. Pearson, B. J. & Sanchez Alvarado, A. A planarian p53 homolog regulates proliferation and self-renewal in adult stem cell lineages. *Development* 137, 213-221 (2010).

34. Palakodeti, D., Smielewska, M., Lu, Y. C., Yeo, G. W. & Graveley, B. R. The PIWI proteins SMEDWI-2 and SMEDWI-3 are required for stem cell function and piRNA expression in planarians. *RNA* 14, 1174-1186, doi:rna.1085008 [pii] 10.1261/rna.1085008 (2008).

35. Bonuccelli, L. *et al.* An RbAp48-like gene regulates adult stem cells in planarians. *J Cell Sci* 123, 690-698, doi:jcs.053900 [pii] 10.1242/jcs.053900 (2010).

36. Scimone, M. L., Srivastava, M., Bell, G. W. & Reddien, P. W. A regulatory program for excretory system regeneration in planarians. *Development* 138, 4387-4398, doi:138/20/4387 [pii] 10.1242/dev.068098 (2011).

37. Orii, H., Sakurai, T. & Watanabe, K. Distribution of the stem cells (neoblasts) in the planarian Dugesia japonica. *Dev Genes Evol* 215, 143-157, doi:10.1007/s00427-004-0460-y (2005).

38. Chen, F., Mackey, A. J., Vermunt, J. K. & Roos, D. S. Assessing performance of orthology detection strategies applied to eukaryotic genomes. *PLoS One* 2, e383, doi:10.1371/journal.pone.0000383 (2007).

39. Eppert, K. *et al.* Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat Med*, doi:10.1038/nm.2415 nm.2415 [pii] (2011).

40. Muller, F. J. *et al.* Regulatory networks define phenotypic classes of human stem cell lines. *Nature* 455, 401-405, doi:nature07213 [pii] 10.1038/nature07213 (2008).

41. Henras, A. *et al.* Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *Embo J* 17, 7078-7090, doi:10.1093/emboj/17.23.7078 (1998).

42. Wang, F., Osawa, T., Tsuchida, R., Yuasa, Y. & Shibuya, M. Downregulation of receptor for activated C-kinase 1 (RACK1) suppresses tumor growth by inhibiting tumor cell proliferation and tumor-associated angiogenesis. *Cancer Sci* 102, 2007-2013, doi:10.1111/j.1349-7006.2011.02065.x (2011).

See www.StemCells.com for supporting information available online.

Figure 1. A) On the left are whole-mount *in situ* hybridizations for the stem cell marker *piwi-1* and the stem cell progeny marker *prog-1*. Planarian *piwi-1+* cells (stem cells; blue) sort to the high-DNA fraction. Stem cell progeny sort to the side population (green). Both of these cell fractions are sensitive to irradiation, as shown below, and were purified and deep sequenced. Red delineates the area where differentiated cells sort, however, we used whole-irradiated worms illustrated below for deep sequencing. B) Heat map showing all genes up-regulated in the stem cell or progeny compartments, as well as a subset of the genes that were differentiated tissue specific. C) Flow chart of the experimental design used in this study.
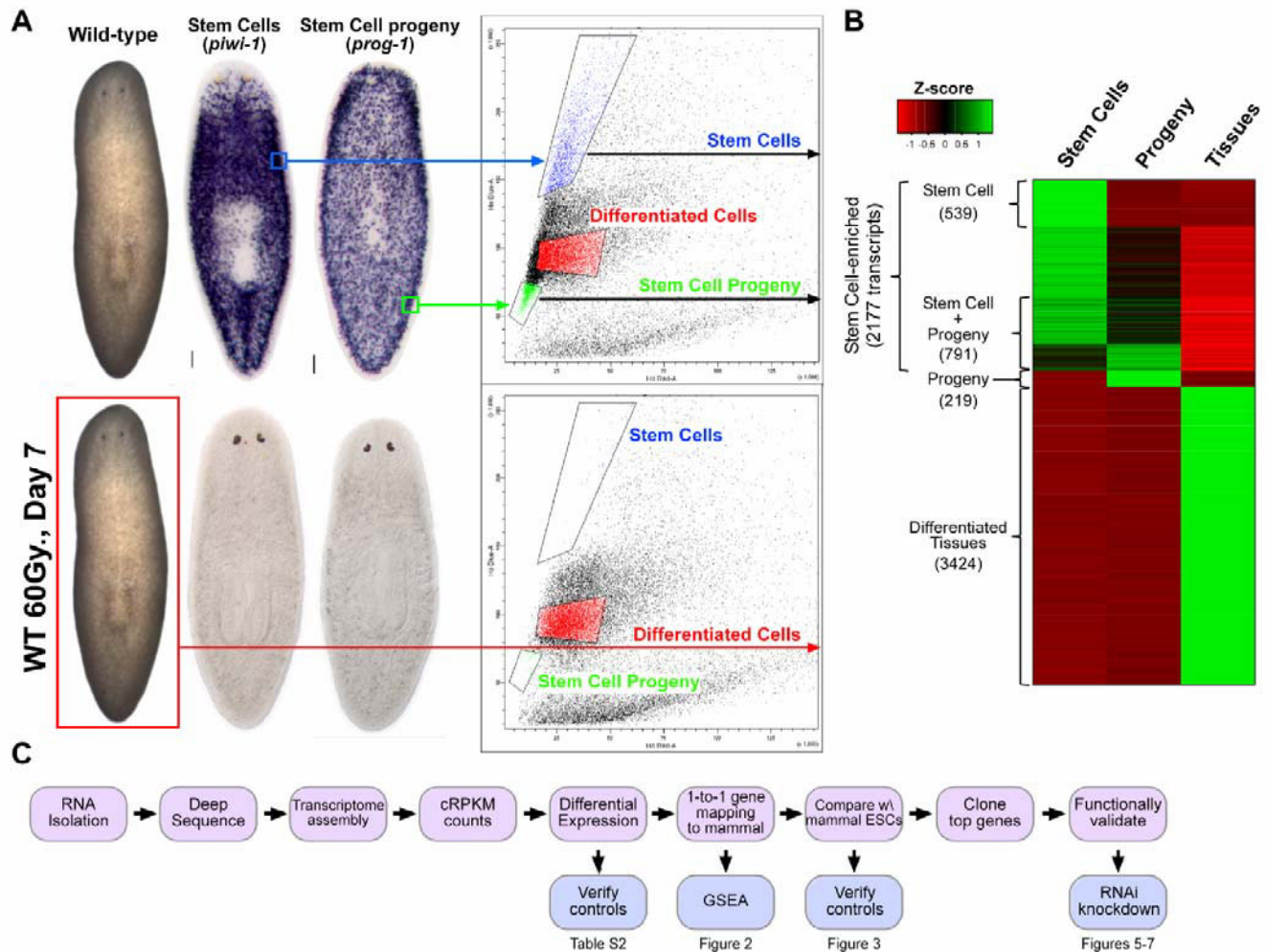
Figure 2. GSEA for planarian cell compartments (stem cell, stem cell progeny and differentiated tissue compartments). Colored circles (nodes) represent GO-associated genetic pathways that were significantly enriched in a given cell type. Nodes are connected when they overlap (*i.e.* they have genes in common), with line width corresponding to the number of shared genes (green lines). The size of the node is proportional to the enrichment score (GSEA nominal enrichment score) for that pathway. To simplify visualization, related pathways are circled and their function labeled. A).Stem cells vs. stem cell progeny showed high enrichment for cell cycle pathways and the *Retinoblastoma* pathway in stem cells, while progeny enrich for processes associated with differentiated cells. B) Stem cells vs. differentiated tissues showed similar enrichment in stem cells for cell cycle and tumor suppressor pathways, while differentiated tissues have enrichment for many pathways in the rest of an organism such as nervous system function, cell signaling, and metabolism. C) Progeny vs. differentiated tissues showed that the *p53* pathway is enriched in progeny as previously described.
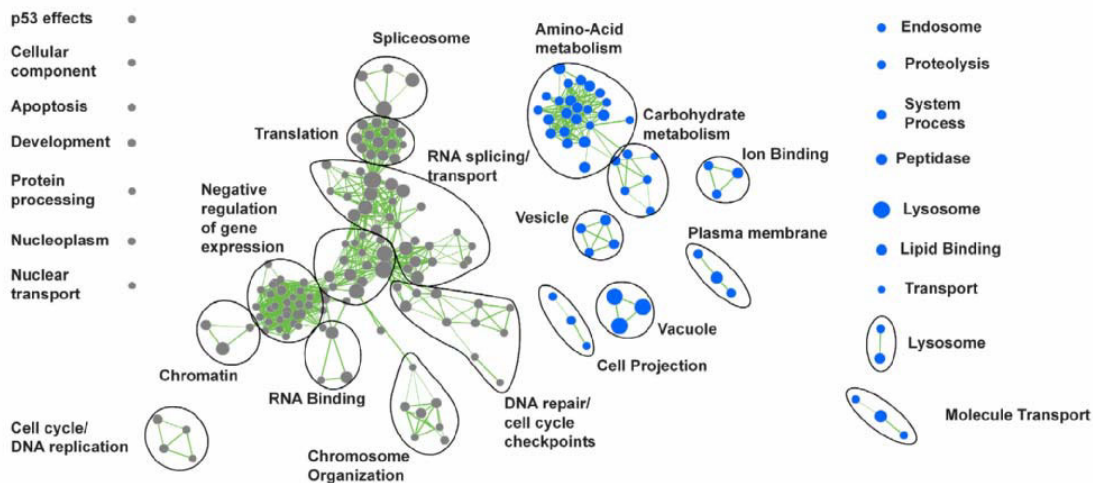
A: Stem cells vs. stem cell progeny
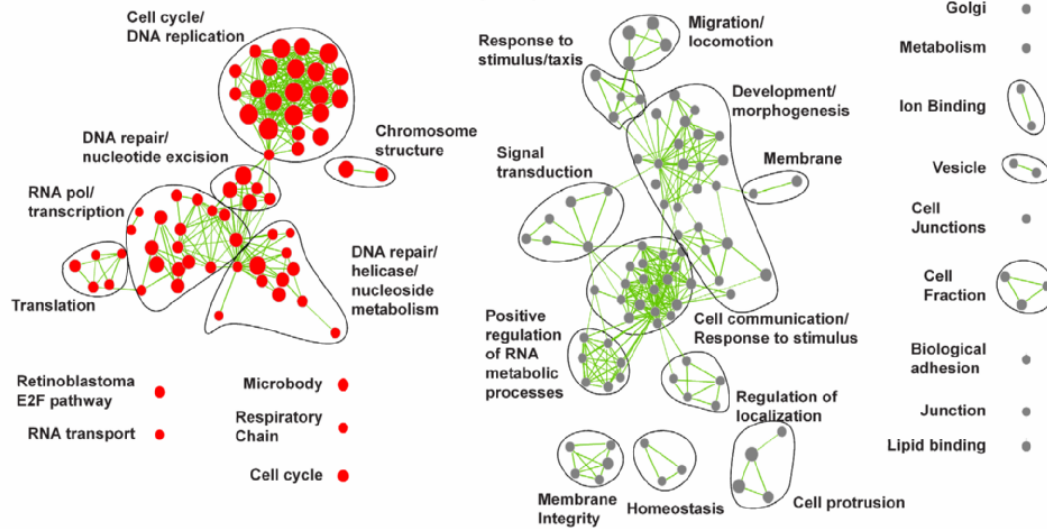
B: Stem cells vs. differentiated tissues
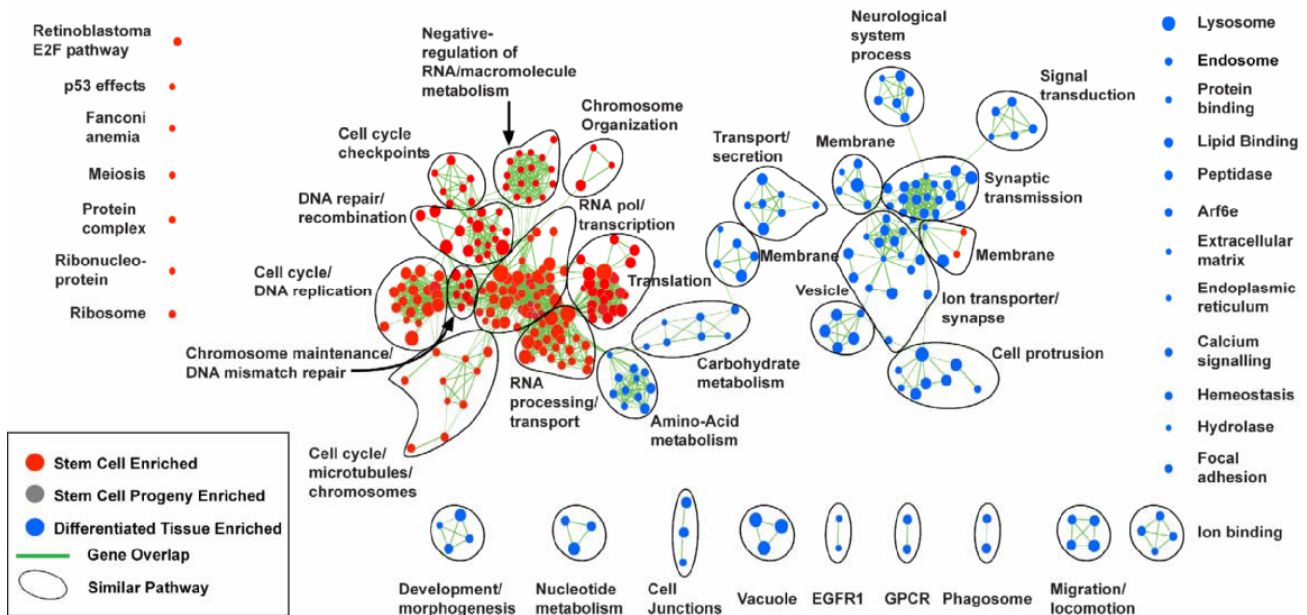
C: Stem cell progeny vs. differentiated tissues

16

Figure 3. A) Hierarchical clustering showed that various stem cell and progenitor types (progeny in planarians and NPCs in mammals) are more closely related than differentiated tissues, which cluster on their own. Smed = planarian; Hsa = Human; Mmu = mouse. B-E) The set of 4,432 gene orthologs was used to determine whether genes were enriched in planarian, mouse and human stem cell types. B) Comparing all 2,177 planarian stem cell specific transcripts from Fig. 1B to the gene-ortholog set showed that 605 genes were both orthologous to mammals and had ≥ 5-fold DE in planarian stem cells C) Comparison of genes with ≥ 5-fold DE in mESCs and hESCs and also included in the ortholog-gene set, showed surprisingly low overlap. D) Combining the 605 planarian genes from (B) and the mammalian genes from (C), showed 123 genes common to all three species' stem cells. E) 127/278 genes on the PluriNet gene list were included in our set of 4,432 orthologs. Of those 79 are shared between all 3 species stem cells at ≥ 5 fold DE. * = genes tested in Fig. 5. ** = genes tested in Fig. 6. *** = genes tested in Fig. 7.
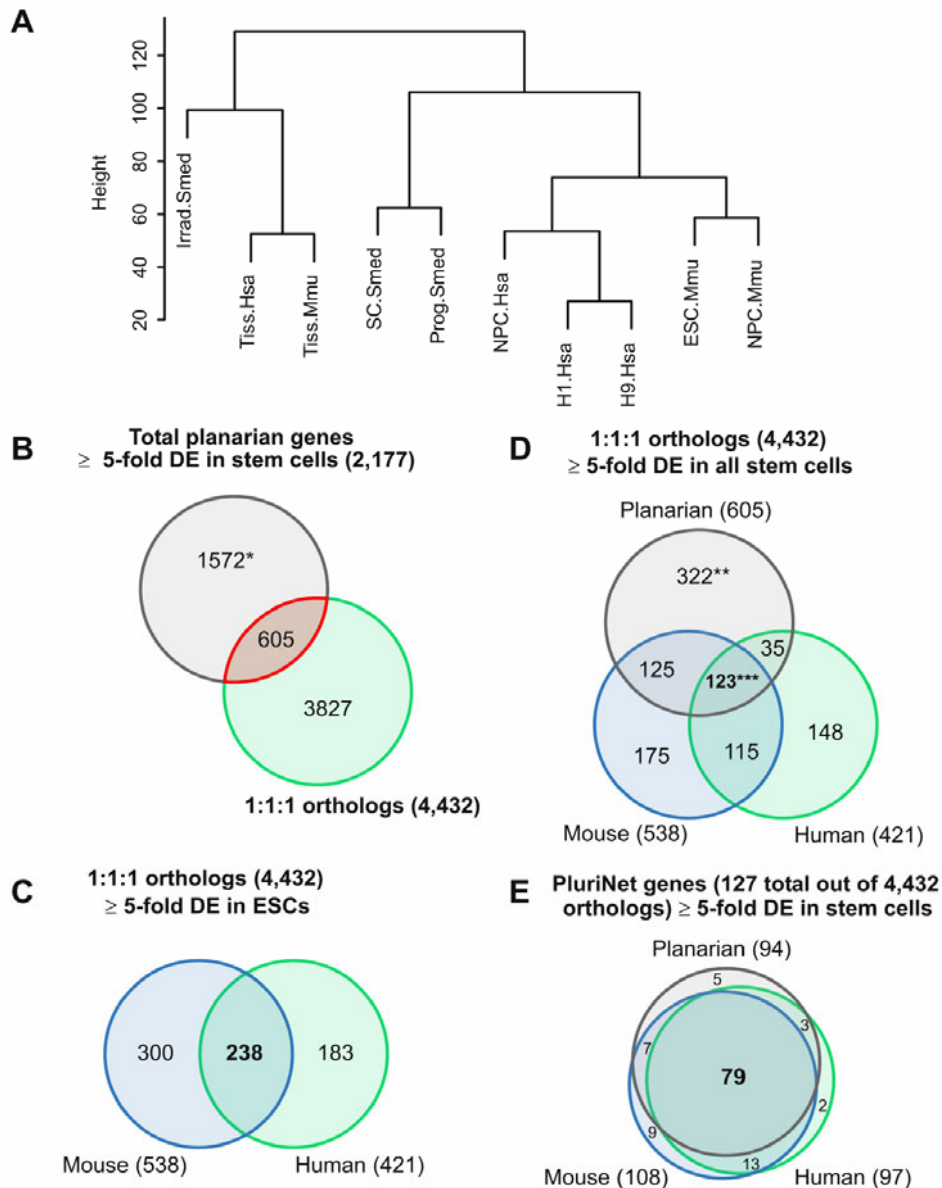
Figure 4. A) Overlapping genes between planarian stem cell progeny and mammalian NPCs (≥ 1.5-fold DE for each pairwise comparison; 4,432 gene orthologs used for comparison). Box shows GO terms significantly enriched in planarian stem cell progeny. B) Fraction of transcription factors from each family expressed in each of the planarian compartments (*i.e.* cRPKM ratios ≥ 1.5 for both pairwise comparisons against the other compartments). The total number of identified members is shown near the name of the family. C) Graphical representation of transcription factor enrichment in specific for each of the compartments. The size of each circle is proportionate to the number of members specific to each compartment. Despite the lowest number of progeny-specific genes in general, the progeny compartment has the highest number of specific TFs. Color codes as in (B).
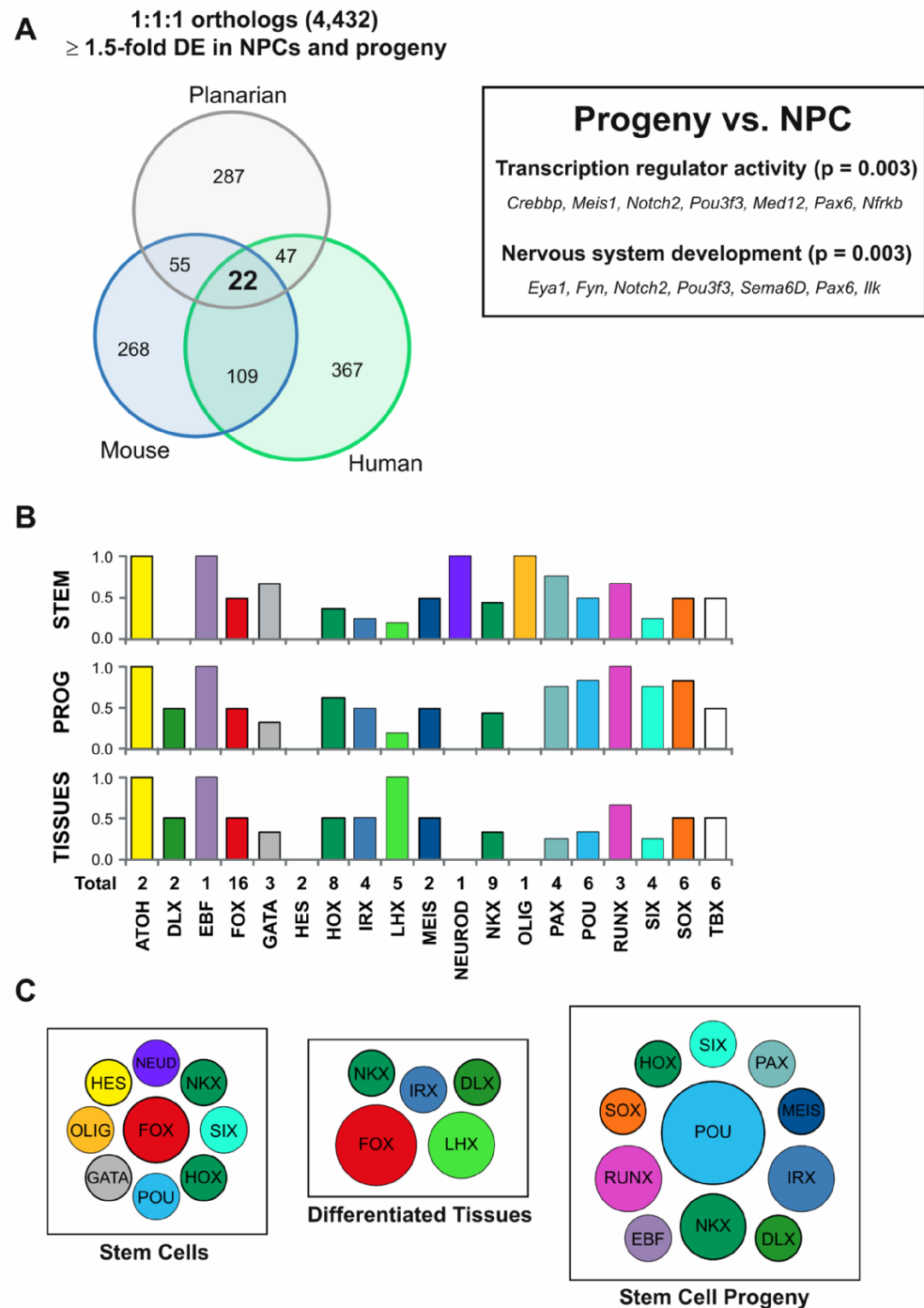
Figure 5. Functional analysis of genes with no 1:1:1 ortholog, but still ≥ 5-fold DE in planarian stem cells. The closest BLAST hit to a mouse gene was chosen for the naming on the left. A) An experimental timeline is shown for feedings, amputations, and stainings. B) The left column shows the various defects observed for tissue homeostasis in uninjured animals. Listed in each panel are the number of days from the beginning of the experiment that the images were taken. The middle column is an ISH for the stem cell marker *piwi-1*. Similar to the various homeostatic phenotypes, we observed varying defects in stem cell maintenance. The right column shows 7 days of head regeneration following amputation. Whereas controls can regenerate a new head in 7 days, regeneration was severely decreased in the RNAi treated worms.
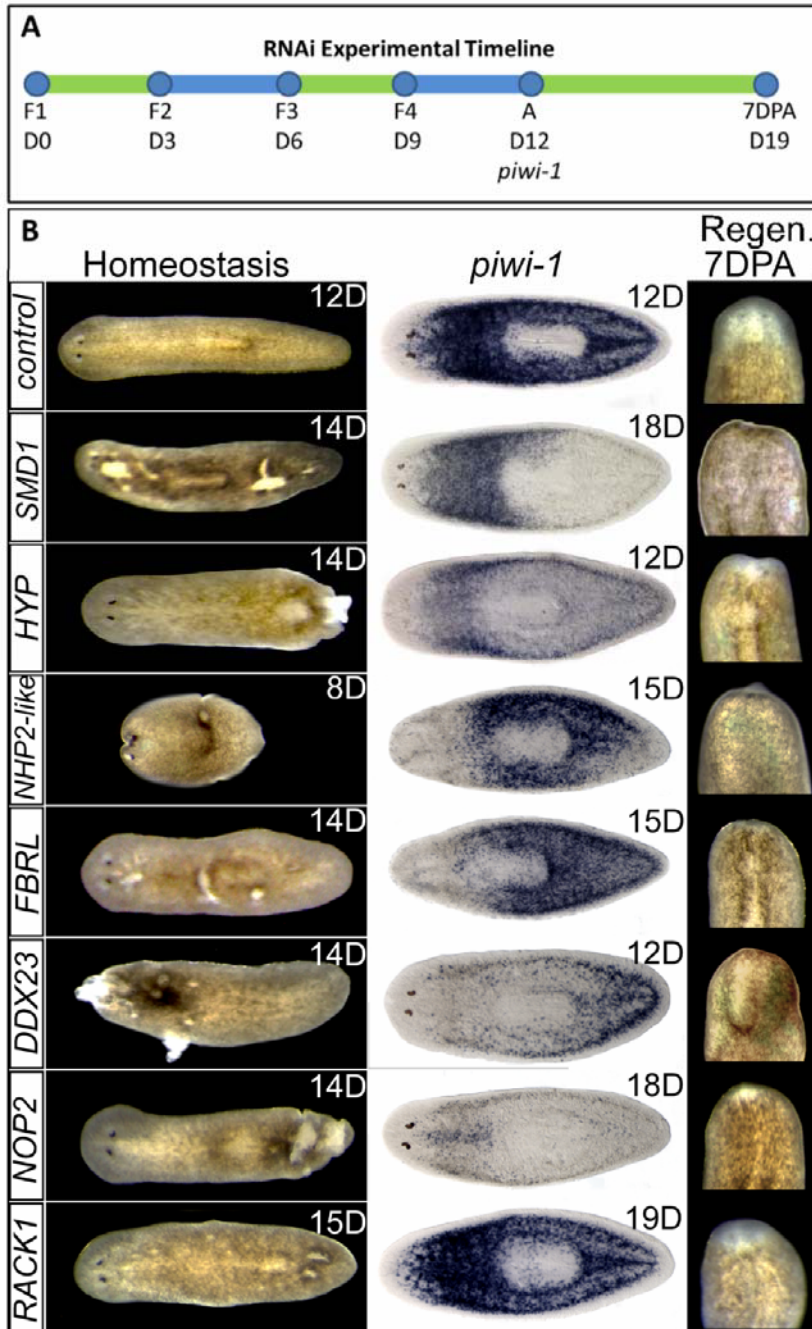
Figure 6. Functional analysis of 1:1:1 gene orthologs that are only DE in planarian stem cells and not DE in mammalian ESCs. The timeline of the experiments is the same as Fig. 5A. The left column shows the various defects observed for tissue homeostasis in uninjured animals. The middle column is an ISH for *piwi-1*. Indicated are the experimental days that the worms were fixed and stained. Similar to the various homeostatic phenotypes, we observed varying effects on the stem cell population. In the right column are amputated planarians after 7 days of head regeneration.
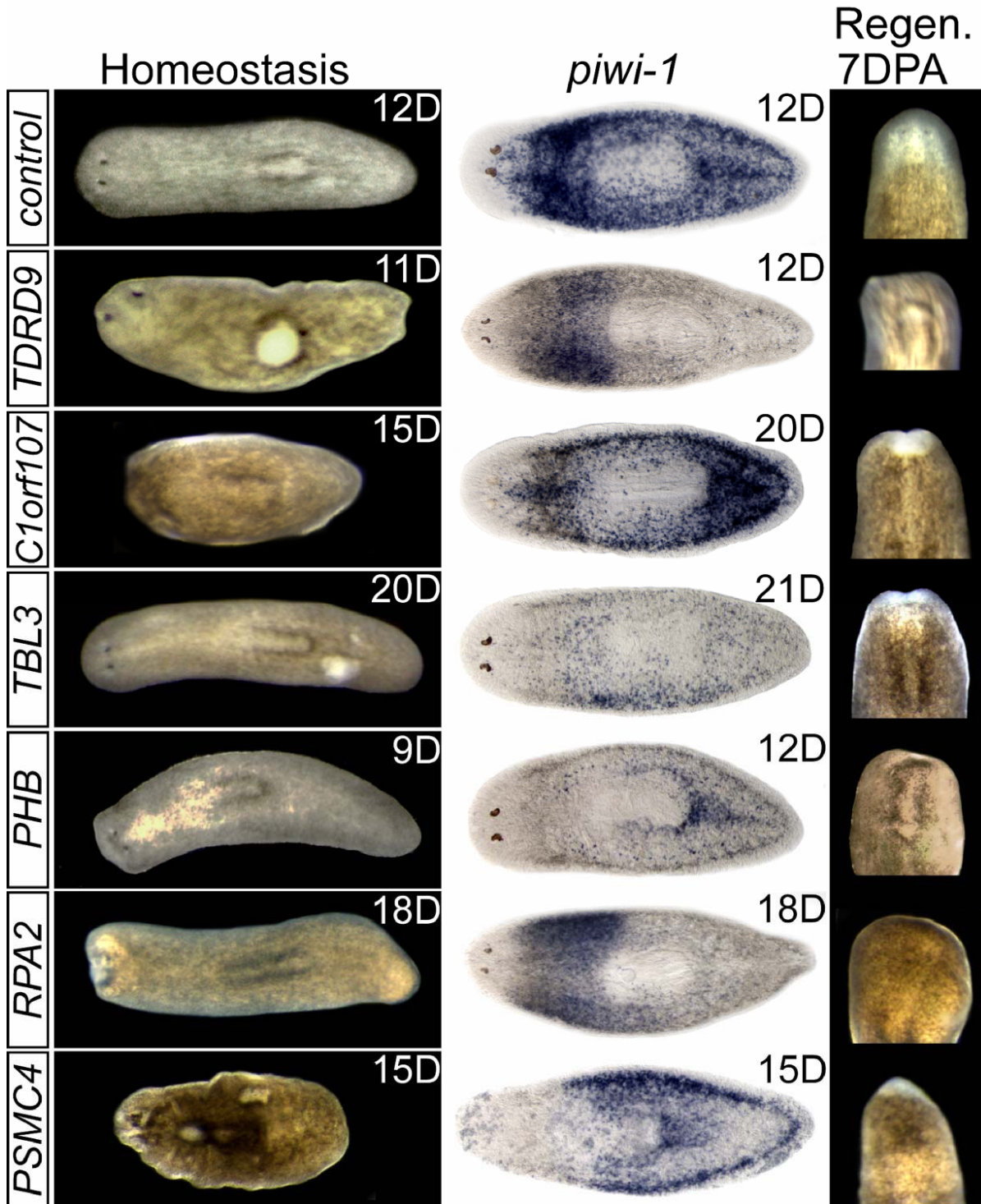
Figure 7. Functional analysis of 1:1:1 gene orthologs that are ≥ 5-fold DE in planarian, mouse and human stem cells. The timeline of the experiments is the same as Fig. 5A. The left column shows the various defects observed for tissue homeostasis in uninjured animals. The middle column is an ISH for *piwi-1*. Indicated are the experimental days that the worms were fixed and stained. Similar to the various homeostatic phenotypes, we observed varying effects on the stem cell population. The right column are amputated planarians after 7 days of head regeneration. Regeneration was severely decreased in the RNAi treated worms with the exception of *NOP58*.