Developmental Cell

Supplemental Information

A Generic and Cell-Type-Specific Wound Response Precedes Regeneration in Planarians

Omri Wurtzel, Lauren E. Cote, Amber Poirier, Rahul Satija, Aviv Regev, and Peter W. Reddien

Supplementary figures

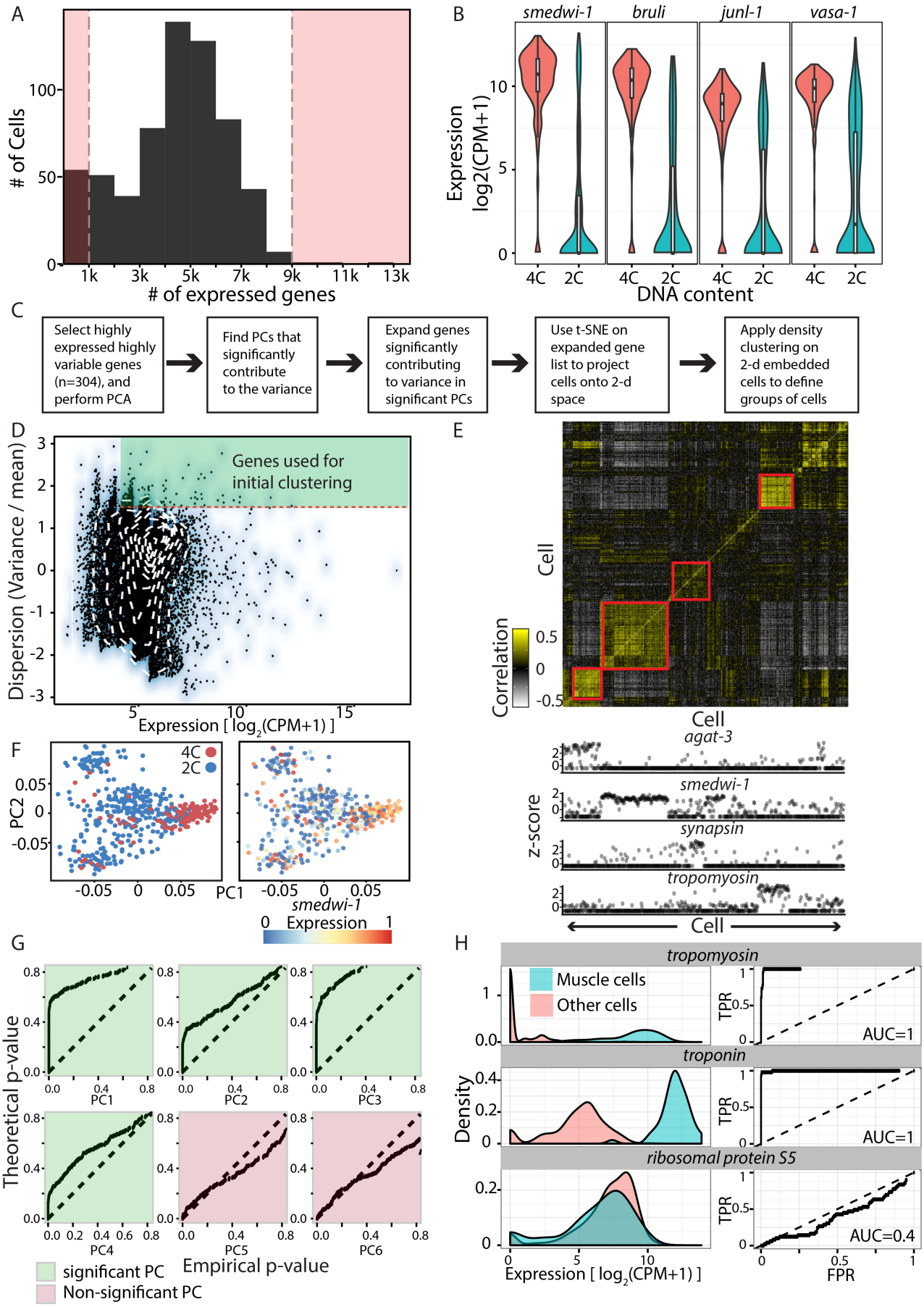

Figure S1, related to Figure 1. Assessment of single-cell sequencing. (A) Histogram of the number of genes expressed (CPM > 0) in the single cells used in this analysis. Shaded red regions represent thresholds used for eliminating cells with extremely low or high number of expressed genes. (B) Violin plots of canonical neoblast markers (Wagner et al., 2012) between cells that were collected using FACS from the 4C gate (Methods) compared to the 2C gate. White rectangle represents the interquartile range, black bar is the median. (C) Schematic describing the different steps of analysis performed for assigning cells to clusters. (D) Genes selected for initial clustering exhibited high dispersion and expression across the sequenced-cells (2C: 0h, 4-6, and 12-14 hpi, 4C: 4-6, 12-14 hpi). Black dots with blue hue represent the mean expression of a gene, white dashed contours represent the density of the dots as obtained by 2d kernel density estimation with 25 bins (kde2d). Green shaded rectangle outlines the selection of genes for the initial clustering. (E) Upper panel: Correlation matrix generated based on the initial set of genes. Cell-order was determined by hierarchal clustering of the cells based on the initial set of genes used for clustering. Red rectangles represent cells co-expressing canonical markers for several cell types: *smedwi-1* for neoblasts, *tropomyosin* for muscle, *synapsin* for neural cells, and *agat-3* for *late epidermal lineage* cells. (F) Left panel: PCA projection of individual sequenced cells (dots), based on the initial set of genes (n=304) used for clustering. PC1 separates the dividing cell fraction from cells that are not dividing (red and blue; 4C and 2C DNA content, respectively; DNA content determine by Hoechst dye analysis during cell isolation with FACS; Methods). Right panel: The same PCA projection is shown with the cells colored based on their rank of expression of the canonical neoblast marker *smedwi-1* (blue, yellow, and red correspond to low, medium and high). Most of the *smedwi-1* expressing cells are separated by PC1. (G) Testing the significance of different principal components through Jackstraw analysis (Chung and Storey, 2015). Each subplot is a quantile-quantile plot (qqplot) of gene p-values in the principal component, as determined by a jackstraw analysis compared to theoretical p-values based on sampling from uniform distribution (Extended experimental procedures). Empirical values near the dashed lines fit a uniform distribution and hence are not considered for further testing (in this case genes were selected from principal components 1 through 4). Green and red backgrounds represent PCs found to be significant and non-significant, respectively, through this analysis. (H) Example of classification of genes to clusters. Shown is the cluster we subsequently determined to be muscle. Left panel: For every cluster, a list of genes that are highly expressed compared to all other clusters was assembled. Shown is the expression of the canonical muscle markers *tropomyosin* and *troponin*, and a negative control *ribosomal protein s5* (top, middle, and

bottom, respectively; blue and red area, muscle cluster, and all other clusters, respectively). Right panel: The ability of individual genes to partition the cells to the tested cluster is plotted by the true positive rate (TPR; sensitivity) and false positive rate (FPR; 1 - specificity) of the assignments, and the area under the curve. The diagonal (dashed black line; AUC=0.5) represents random assignment to the cluster, such as observed for the negative control.

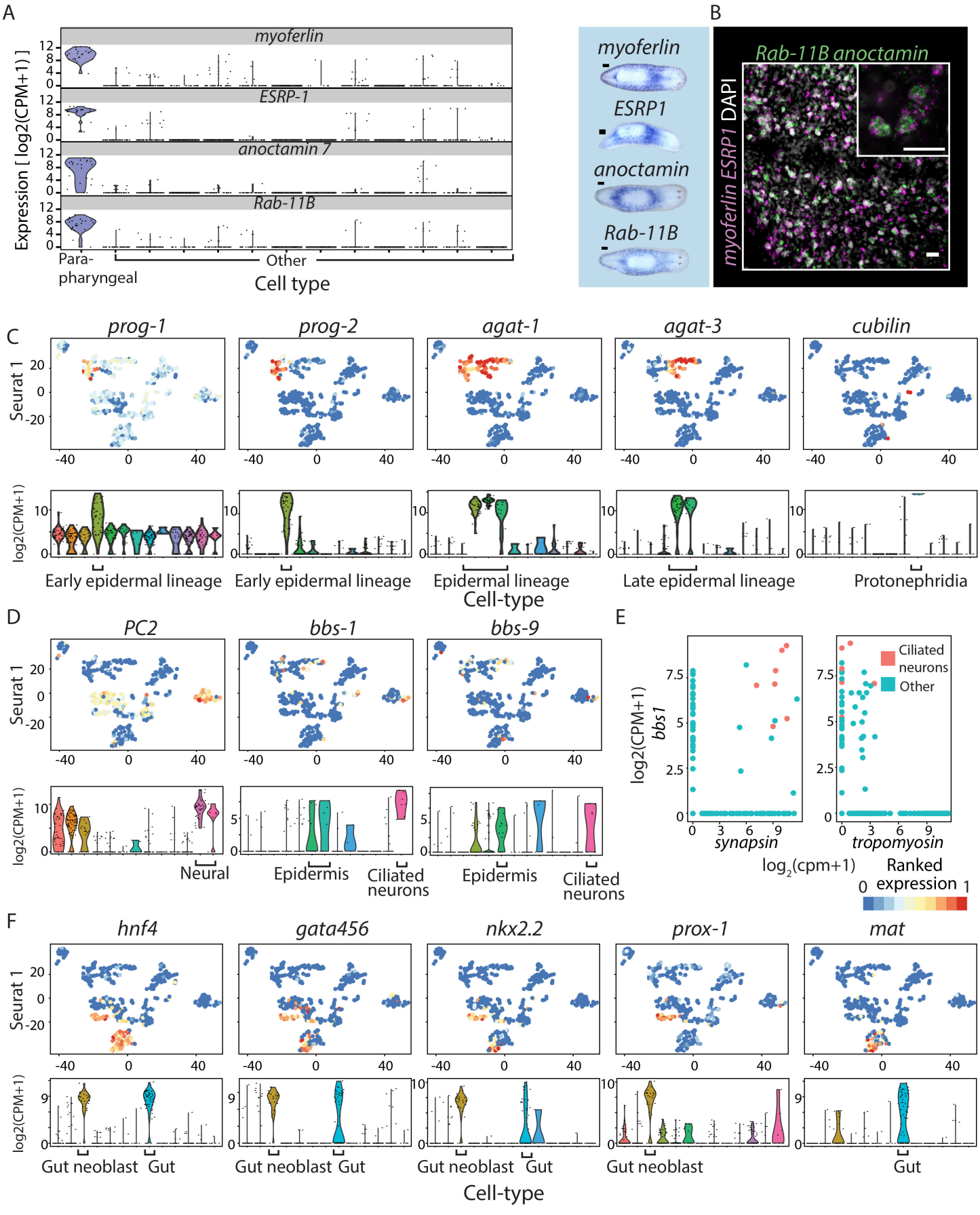
A
myoferlin
ESRP-1
anoctamin 7
Rab-11B
Expression [log2(CPM+1)]
Para-pharyngeal
Other
Cell type
B
myoferlin
ESRP1
anoctamin
Rab-11B
Rab-11B anoctamin
myoferlin ESRP1 DAPI
C
prog-1
prog-2
agat-1
agat-3
cubilin
Seurat 1
log2(CPM+1)
Early epidermal lineage
Early epidermal lineage
Epidermal lineage
Late epidermal lineage
Protonephridia
Cell-type
D
PC2
bbs-1
bbs-9
Seurat 1
log2(CPM+1)
Neural
Epidermis
Ciliated neurons
Epidermis
Ciliated neurons
E
log2(CPM+1) bbs1
synapsin
tropomyosin
Ciliated neurons
Other
log2(cpm+1)
Ranked expression
F
hnf4
gata456
nkx2.2
prox-1
mat
Seurat 1
log2(CPM+1)
Gut neoblast
Gut
Gut neoblast
Gut
Gut neoblast
Gut
Gut neoblast
Gut
Cell-type

Figure S2, related to Figure 1. Single cell gene expression planarian resource. (A) Left panel: violin plots show high specificity to a single cluster (violet; black dots represent single-cell expression). Right panel: WISH analyses of the genes reveal, in all cases, a parapharyngeal localization (scale = 100 μm). (B) dFISH of genes enriched in the parapharyngeal cluster. Pooled probes for *myoferlin* and *ESRP-1* (magenta) and *Rab-11B* and *anoctamin* (green) were used for coexpression analysis (scales = 20 μM; DAPI in gray). (C) Upper panel: Expression of cell type specific markers plotted on *Seurat* maps showing the specificity of genes to cell types (Cells represented by dots; color is the ranked expression of the gene in cells. blue to red, low and high ranked expression, respectively). Lower panels: Violin plots of gene expression across cell-types. X-axis annotation highlights cell types enriched for the plotted gene. (D) Upper panels: *Seurat* maps of the canonical neural marker PC2, and two canonical cilia components (*bbs-1* and *bbs-9*). The components are expressed almost exclusively in differentiated epidermal cells and in a subset of the neural cells (*PC2+/synapsin+)*. Lower panels: violin plots of neural (*PC2*) and cilia (*bbs1* and *bbs9*) related genes. (E) Left panel: Co-expression plots of *bbs-1*, a cilia component, and *synapsin*, a canonical neural marker, shows that a subset of the cells expressing high levels of *synapsin* also express *bbs1*. Right panel: Co-expression of *bbs-1* with *tropomyosin*, a canonical muscle marker, shows that there are no cells highly expressing both genes (cells represented by dots, red and blue colors are cells determined to be ciliated neurons, and other cells, respectively). (F) Upper and lower panels: *Seurat* maps and violin plots of putative gut neoblasts markers, including transcription factors and gut markers. The expression of *hnf4*, *gata4/5/6*, *nkx-2.2* appears in both the differentiated gut cluster, and the gut (γ) neoblasts. The transcription factor *prox-1* is expressed the gut neoblasts cluster, but not in the differentiated gut cells. *mat*, a planarian gut marker, is expressed exclusively in the differentiated gut marker.

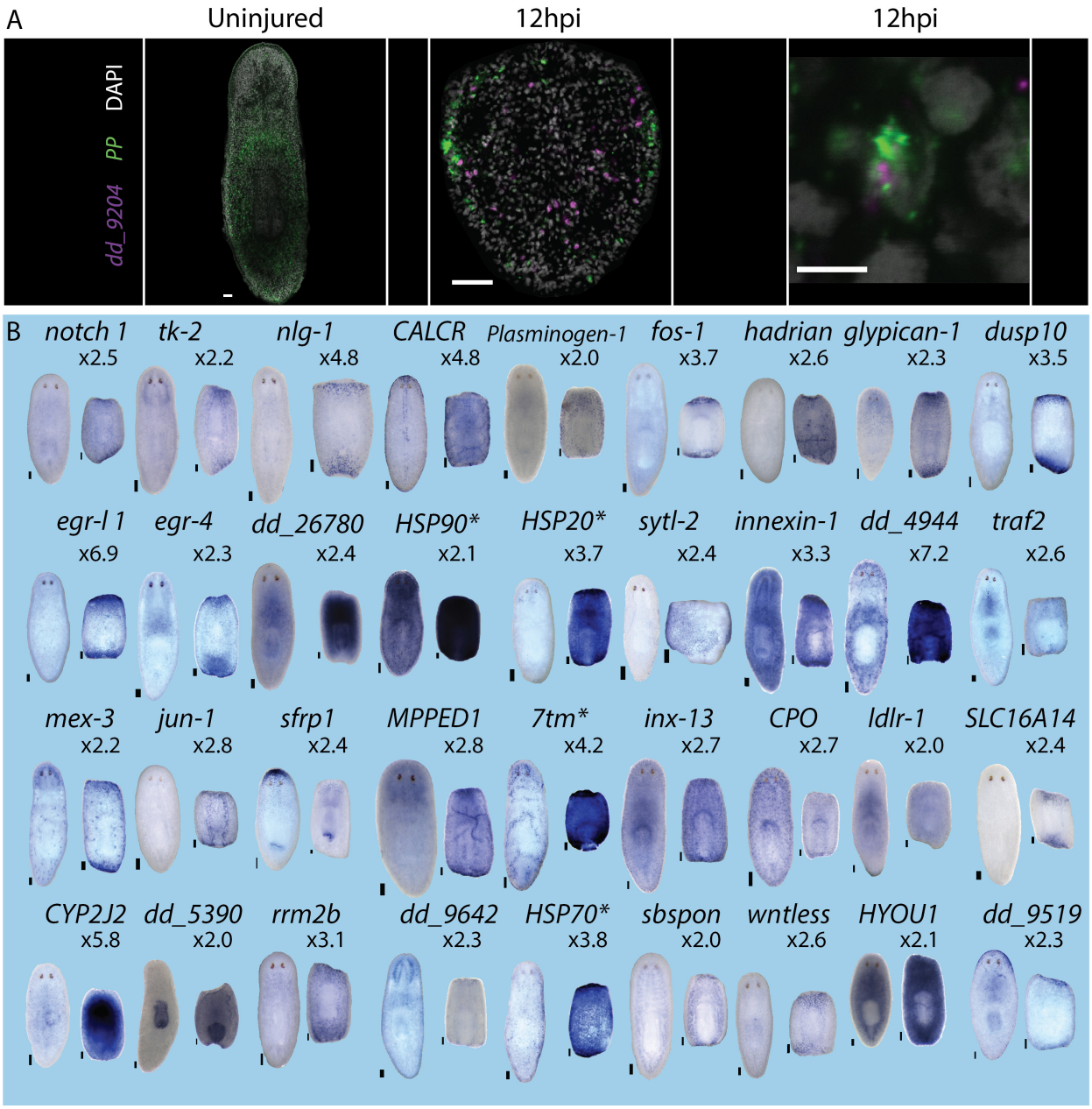

Figure S3, related to Figure 4. Wound induced gene expression. (A) dFISH validation of parapharyngeal-specific gene expression of *dd_9204* (magenta) with a parapharyngeal probe pool (*myoferlin* and *ESRP1*; green) and DAPI (gray) in intact and injured (12 hpi) animals. Scale = 50 µm; right panel scale = 5 µm. (B) WISH analysis of 36 additional genes tested for asymmetry in expression of wound-induced genes. Shown are intact animals and trunks. * denotes annotation based on protein family domains (PFAM; Methods). Scale=100 µm.

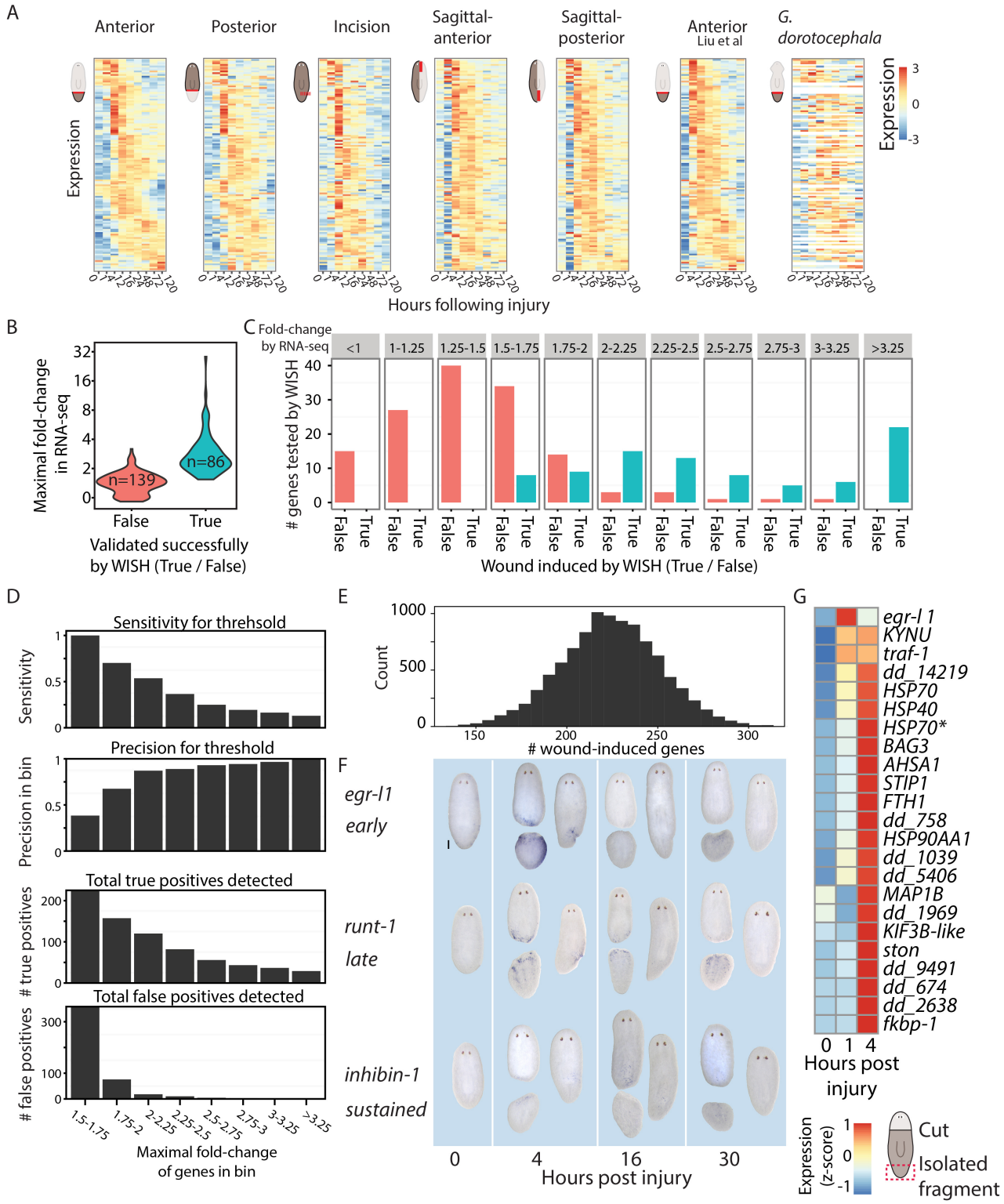

Figure S4, related to Figure 5. Extended time course analyses of distinct injuries. (A) Shown is the expression of 128 wound-induced genes in extended time-courses from multiple injuries. Each row represents a gene, and columns represent the time of isolation (hours post injury; 0-120 h). The colors are z-transformed value (minimal and maximal range was set to -3 and 3, respectively; blue, yellow and red colors correspond to low, medium, and high

expression, respectively). The ordering of the genes is identical in all heatmaps to facilitate comparisons, and furthermore, it is identical to the gene order in figure 5A. The rightmost heatmap presents regeneration timecourse from *G. dorotocephala*. Gene ordering of the orthologs found in *G. dorotocephala* for the 128 wound-induced genes, was retained. In case no ortholog was found, a blank line was plotted. (B) Violin plot summarizing the WISH analyses performed to estimate the sensitivity and precision of RNA-seq for detecting wound-induced genes (n represents the number of WISH analyses in the group it is plotted in). (C) Shown are bar-plots summarizing the number of true positive and false positive found through WISH analyses grouped according to their maximal change in expression (top title, linear scale) up-to 12 hours from amputation (Sampling performed at 0, 3, 6, 12 hpi in anterior and posterior amputations; Methods). (D) Summary of key parameters of false-discovery, sensitivity, and precision, obtained through comparisons of RNA-seq and WISH analyses. Shown are bar plots comparing different groups of genes tested by WISH that were binned by their maximal expression induction following wounding. Shown are estimates for the total number of true and false positives in each bin, through multiplying the sensitivity and precision by the total number of significantly overexpressed genes (FDR < 0.05). (E) Estimation of total number of wound-induced genes in the planarian genome by resampling analysis (n=10,000; extended experimental methods). (F) Shown are the full WISH images corresponding to the fragments displayed in Fig 5E. (G) Heatmap of wound-induced genes that were found by analyzing tissues far from the wound site (row z-score; Table S4).

Extended experimental procedures

Gene cloning and transformation

Genes were amplified from planarian cDNAs using gene-specific primers and cloned into pGEM vector according the manufacturer's protocol (Promega). Vectors were transformed into *E. coli* DH10B by the heat-shock method as follows: 20-100 µl of bacteria were mixed with 10 µl of pGEM vector cloned products and incubated on ice for 30 minutes, and then put at 42°C for 1 minute. The mixtures were then supplemented with 100 ul of SOC medium and following 1 h incubation at 37°C, were plated on agarose plates containing 1:500 carbacyclin, 1:200 Isopropylthio-b-D-galactoside (IPTG), and 1:625 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). Colonies were grown overnight at 37°C and white colonies were screened by colony PCR using primer sequences M13F (GTAAAACGACGGCCAGT) and M13R (CAGGAAACAGCTATGAC) with the following PCR program: a. 5 minutes at 95°C; b. 34 cycles of 45 sec at 95°C, 60 sec at 55°C, and 2:30 minutes at 72°C; c. 10 minutes at 72°C; d. hold at 4°C. Reactions were analyzed by gel-electrophoresis and for each gene a colony showing the correctly sized gene product was transferred to Luria Broth media (LB) supplemented with 1:500 carbacyclin for overnight incubation at 37°C. Plasmids were purified from overnight cultures with the QIAprep Spin Miniprep Kit (CAT #27106; Qiagen). Cloned genes sequences were validated by Sanger-sequencing (Genewiz, Inc.).

Double-stranded RNA synthesis for RNAi experiments

Double stranded RNA (dsRNA) was synthesized as previously described (Petersen and Reddien, 2008). 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 (4 ul) were mixed, in separate reactions, with 16 ul of 10 mM rNTPs (Promega); 1 ul of 100 uM dithiothreitol (DTT; Promega); 1 ul of thermostable Inorganic Pyrophosphatase (TIPP; New-England Biolabs); 0.5 ul of RNasin (Promega); 1.5 ul

of T7 polymerase; and 6 ul of 5x Transcription optimized buffer (Promega). Reactions were incubated for 4-12 h at 37°C and then supplemented with RNase-free DNase for 45 minutes. RNA was purified by phenol extraction followed by ethanol precipitation, and finally resuspended in 30 ul of MiliQ H2O. RNA was analyzed on 1% agarose gel, and quantified by Nanodrop (Thremoscientific) to have at least 5 ug/ul. 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.

Planarian dsRNA feedings

Animals were starved for at least 10 days prior to the first feeding. dsRNA was mixed 1:3 with 100% homogenized beef liver, and supplemented with 1 ul of red food coloring. Animals were kept in dark for at least 2 h before feeding, and were then taken out of the dark and fed the dsRNA-liver mix for at least 2 h. Animals uptake of the food was evaluated by the red coloring of the gut branches. Following a feeding, the culture plates and water were replaced and worms were kept in the dark; water in plates was replaced the day following a feeding as well, and every 3 days, unless another feeding was done.

RNAi feeding protocol

Worms were fed with liver containing dsRNA every three or four days. Three days following the fourth feeding animals were cut to three fragments, and the trunks were immediately soaked in 100 ul of planarian water supplemented with dsRNA against the target gene for 6 hours in the dark. Then, animals were washed and trunks were kept in the dark for 9 days before being fed with liver containing dsRNA against the target gene (booster). Then, 3 days following the booster feeding, trunks were cut to 3 fragments and soaked in planarian water containing dsRNA for 6 hours. Regenerating fragments were screened for defects every other day.

Illumina library preparation for anterior and posterior time-courses

Control prepharyngeal fragments (0 hpi) were isolated in biological triplicates and placed in TRIzol Reagent (Life Technologies) within 5 minutes from tissue isolation. Anterior-facing or posterior-facing wounds were amputated as pre-pharyngeal fragments as follows: A first cut was done either in the anterior or posterior end of the pre-pharyngeal region, and, at a given time-point (3, 6, or 12 hpi), a second cut was done to the opposite end of the pre-pharyngeal region. Prepharyngeal fragments were placed in biological triplicates in TRIzol (Life Technologies). Total RNA was purified following manufacturer's instructions (Life Technologies), followed by a second chloroform extraction to remove residual phenol contamination. Libraries (total 21) were prepared using the TruSeq RNA Sample Preparation Kit v2 (Illumina) and were sequenced on Illumina HiSeq 2000 sequencer (Illumina).

Illumina library preparation for far libraries

Animals were amputated prepharyngeally and were placed in planarian water. Following a recovery period (0, 1, or 4 h) tail fragments were isolated and put immediately in TRIzol (Life Technologies). Total RNA was purified and sequencing libraries were prepared using the TruSeq RNA Sample Preparation Kit v2 (Illumina).

Illumina library preparation for extended time courses

Tissues were isolated and placed in TRIzol (Life Technologies) for RNA extraction as previously described (Liu et al., 2013). Briefly, animals were cut as: postpharyngeally for (1) anterior- and (2) posterior- regeneration time courses; sagittally for (3) anterior and (4) posterior sagittal time-course; a postpharyngeal incision for the (5) incision time-course; and postpharyngeal amputation on *G. dorotocephala* for (6) anterior regeneration. Then, animals were put in planarian water for recovery. At each of the time points (1, 4, 12, 16, 24, 48, 72, 120 hpi) at least 8 animals were killed in 1% HCl for 1 minute, followed by 2 washes in

phosphate buffered saline (PBS). Animals were then put in RNALater (Life Technologies) and wound-sites were isolated on a cold block and put immediately in TRIzol as previously described (Liu et al., 2013). Uninjured fragments were isolated similarly, with the exception that the animals were killed and put in RNALater before isolation of the desired fragment. RNA was extracted according to manufacturers' instructions following tissue lysis in TRIzol with TissueLyser II (Qiagen; 2 minutes at 20Hz, followed by 2 minutes at 30Hz). RNA concentration was measured with Qubit 2.0 Fluorometer (Life technologies). At least 500 ng of purified RNA was used for strand-specific Illumina RNA-sequencing library construction as previously described (Engreitz et al., 2014; Schwartz et al., 2014). Briefly, for each sample poly-adenylated RNA was purified by two rounds of polyA selection (Dynabeads mRNA Purification Kit; Life Technologies) and eluted in 18 ul of H_2O. RNA was then fragmented with RNA Fragmentation reagent for 2 minutes (AM8740; Ambion) and purified on paramagnetic beads (Dynabeads Silane MyOne; Life Technologies 37002D). Then, RNA was incubated with 2U of Turbo DNase (Life Technologies) for 30 minutes followed by addition of FastAP for 10 minutes (Life Technologies). Then, RNA was ligated with an RNA oligo corresponding to a truncated 3' Illumina adapter (AGAUCGGAAGAGCACACGUC; IDT) using T4 RNA ligase 1 (36 units; NEB), and reverse transcribed with a specific primer (AGACGTGTGCTCTTCCG; IDT) with AffinityScript reverse transcriptase (Agilent). Following cDNA synthesis, primers were removed by adding ExoSAP-IT (Affymetrix) directly to the mix. RNA was degraded by adding NaOH, and cDNA was isolated and eluted in H_2O by paramagnetic beads clean-up. The 3' of the cDNA was ligated with a truncated 5' adapter (AGATCGGAAGAGCGTCGTGTAG; IDT), and the cDNAs were amplified for 12 cycles with barcoded Illumina primers.

Sequencing read mapping

Sequencing reads from each library were mapped to the *S. mediterranea* dd_Smed_v4 transcriptome assembly (Liu et al., 2013) with 5 additional sequences (sequences listed below) using Novoalign v2.08.02 with parameters [-o SAM -r Random]. The resulting

Sequence Alignment/Map files (SAM) were converted to sorted BAM format with the samtools v1.1 (Li et al., 2009) command [samtools view -T dd_Smed_v4 -bS IN| samtools sort – OUT] where dd_Smed_v4 is the fasta file of the assembly (http://planmine.mpi-cbg.de/), IN is the name of the SAM file, and OUT is the name of the sorted BAM file. Mapping statistics for each library were calculated by the samtools flagstat, and were examined manually. Read count per contig was calculated by bedtools v2.20.1 (Quinlan and Hall, 2010) using the coverageBed command [-abam IN -b BED > OUT] where IN is an input sorted BAM file; BED is a bed formatted file with all contigs in the assembly and their lengths; and OUT is the resulting read-counting coverage file.

Detection of differentially expressed wound-induced genes

Coverage files from the high-resolution wound-response time courses (0, 3, 6, and 12 h following anterior or posterior amputation) were consolidated to a read count matrix. The expression matrix was filtered for contigs longer than 450 base-pairs (bp). Following TMM data-transformation with edgeR v3.6.8 (Robinson et al., 2010), low expressing contigs were filtered with a cut-off of CPM of 6 in at least 2 out of 21 libraries. Next, differentially expressed genes were determined by the exactTest function by comparing each time point in the two time courses, separately, to the expression at time 0. Following hypothesis testing the p-values were corrected for multiple testing with false discovery rate (FDR). FDR smaller or equal to 0.05 and a fold-change of 2 or more were set as thresholds for determining wound induction. Genes that were found to be upregulated in at least one time point were included in the wound-induced genes list, except for contigs dd_Smed_v4_9491_0_1, dd_Smed_v4_14725_0_1, and dd_Smed_v4_1071_0_1 that were not validated by WISH (Table S3).

Detection of genes with putative asymmetric expression

Expression levels of wound-induced genes were compared between matching time points in the anterior and posterior time courses using the edgeR exactTest function. Genes with

corrected FDR of 0.1 or less, or exhibiting a fold change of at least 1.5 were selected for WISH validation, as well as 50 additional wound-induced genes exhibiting smaller differences in expression or less significant FDR between injuries. WISH validation of the genes was done on (i) intact animals and (ii) trunks of amputated animals that were cut and subsequently fixed at the time-point exhibiting the largest difference in expression between anterior and posterior wound sites.

Estimating the sensitivity of the wound-induced gene detection

Recent surveys of wound-induced gene expression in planaria yielded very partially overlapping results (Kao et al., 2013; Sandmann et al., 2011; Wenemoser et al., 2012), reflecting different instrumentation, analytical methods, and experimental setup. WISH was performed on 225 different genes on intact and amputated animals that covered a wide range of expression changes and FDR following wounding, including 46 negative controls (fold change; FC; and FDR, 0 to 28.6 and 0 to 1, respectively; Fig S3B; Table S5). 38% (86/225) of all of the tested genes were detectibly wound-induced by WISH. None of the genes with maximal FC less than 1.5 (n=82) could be validated by WISH regardless of their FDR. Furthermore, genes with maximal FC between 1.5 and 2 could be called with a precision of only 26% (17/65; Table S4; Fig S4B-D). By contrast, 88% (69/78) of the genes with FC > 2 could be validated by the WISH analysis (Table S4-5; Fig S4B-D). Therefore, 2-fold overexpression in at least one time point following wounding was used as a threshold for calling wound-induced expression. This threshold balanced sensitivity (57%) with precision (88%) compared to alternative thresholds (Fig S4C-D). Estimation of the total number of wound-induced genes was done by sampling 50% of the differentially-expressed genes according to thresholds [FC> 1.5; FDR < 0.05; minimal CPM 6 in at least 2 libraries] 10,000 times. For each sample the total number of wound-induced genes was estimated by multiplying the number of genes in an expression bin by the fraction of genes that were wound induced, as detected by the WISH validated genes in the sample. Estimations were

multiplied by 2 to correct for the sample size. The total number of wound-induced genes was the average of the individual estimations.

Single-cell isolation and Fluorescence-Activated Cell Sorting

Wound sites were collected from post-pharyngeally amputated animals 4-6h or 12-14h following an amputation; control cells were collected from the same region in intact animals and were processed immediately. Cell suspension was prepared and was subjected to FACS as previously described (Hayashi et al., 2006; Reddien et al., 2005). Briefly, isolated tissues were put in 450 ul of CMFB (calcium magnesium free buffer + 1% BSA) with 50 ul of collagenase and incubated at room temperature while gently pipetting the samples. Samples were then filtered through a 40 um filter into CMFB. Samples were span down and re-suspended in 500 ul of CFMB. To each sample, 20 ul of Hoechst was added and incubated in the dark for 45 minutes, followed by addition of 1 ul of propidium iodide. Negative controls devoid of either Hoechst, PI, or both were prepared in parallel. Single cells were sorted to 96-well microplates containing 5 ul Buffer TCL (Qiagen) + 1% 2-mercaptoethanol. Plates were incubated for 5 minutes at room temperature and were then placed on dry ice.

Single-cell sequencing library construction

RNA-sequencing libraries were prepared from Single sorted-cells as previously described (Picelli et al., 2013; Picelli et al., 2014) with few modifications. Each well in a 96-well microplate was supplemented with x2.2 (11 ul) of Ampure XP beads (Agencourt) and incubated for 10 minutes at room temperature, and then put on a 96-well magnet plate (Dynamag 96-side magnet; Life Technologies) for 5 minutes. Supernatant was removed and beads were washed twice with 100 ul of 80% EtOH. EtOH was removed and beads were air-dried for 10 minutes before elution of the beads in a mixture of 1 ul of reverse transcription primer (5′-AAGCAGTGGTATCAACGCAGAGTACT(30)VN-3′, IDT DNA), 1 ul of dNTP mix (10 mM), 0.1 ul of SUPERase RNase-inhibitor (40 U/ul; Life Technologies #AM2696), and 1.9 ul of H2O. The plate was incubated at 72°C for 3 minutes and placed immediately on ice. Each

well was supplemented with 7 ul of a mixture consisting of 1.65 ul H_2O, 2 ul of 5x Maxima reverse-transcription buffer (Thermo-Fischer), 0.9 ul MgCl2 (100mM Sigma-Aldrich; M1028), 2 ul of Betaine (5M; Sigma-Aldrich; B0300-5VL), 0.25 ul of SUPERase RNase-inhibitor (40 U/ul), 0.1 ul of Maxima RNase H- RT (200 U/µL; Thermo-Fischer, EP0753), and 0.1 template switching-oligo (Exiqon; 100 uM; AAGCAGTGGTATCAACGCAGAGTACrGrG+G; r and "+" denote RNA and LNA bases, respectively). Plate was briefly span-down and incubated as follows: 42°C for 90 minutes, followed by 10 cycles of (50°C for 2 minutes, 42°C for 2 minutes), followed by 70°C for 15 min. Following reverse-transcription a pre-amplification mix of 14 ul was added to each well [1 ul of H2O; 0.5 ul of PCR primer (10 uM; 5'-AAGCAGTGGTATCAACGCAGAGT-3′), and 12.5 ul of KAPA HiFi HotStart ReadyMix (Kapa Biosystems; KK2601)]. The cDNA was amplified using the following program: 98°C for 3 min; 20 cycles of (98°C for 15 sec, 67°C for 20 sec, 72°C for 6 min); 72°C for 5 min; hold at 4°C. Following pre-amplification PCR products were purified using x0.8 Ampure XP beads, and eluted in 20 uL of H2O. Amplified cDNA concentrations were measured using Qubit HS-DNA reagents (Life Technologies). Samples were diluted to 0.2 ng/ul, and sequencing libraries were prepared using the Nextera XT library kit (Illumina). For each sample, 1.25 ul of amplified cDNA was combined with 2.5 ul of tagmentation DNA buffer and 1.25 ul of the amplicon tagmentation mix. Samples were mixed and put at 55°C for 10 minutes. Samples were chilled on ice, and 1.25 ul of neutralize tagment buffer was added for 5 minutes incubation at room temperature. An amplification mix was added as follows: 3.75 ul of Nextera PCR mastermix and 1.25 ul of two barcoded amplification primers. The samples were amplified with the following PCR program: 72°C for 3 minutes; 95°C for 30 seconds; 12 cycles of (95°C for 10 seconds, 55°C for 30 seconds, and 72°C for 1 minute); 72°C for 5 minutes; hold 4°C. Following amplification, 2.5 ul of each sample were pooled in groups of 32-96 samples, purified with x0.9 Ampure XP beads, and eluted from beads in 25-50 ul of H_2O.

Single-cell data mapping and clustering

Following Illumina sequencing, cells were eliminated from further analysis if they were found to be contaminated by more than 25% of non-planarian DNA, as detected by mapping to human, mouse, rat, yeast, and *E. coli* genomes. The sequencing reads from all other cells were mapped to the dd_Smed_v4 assembly (Liu et al., 2013) with Novoalign v2.08.02 and the number of reads for each contig was calculated as described above. Following mapping, reads mapped to contigs dd_Smed_v4_10881_0_1 and dd_Smed_v4_5614_0_1 were excluded, as they represented misalignments of primer amplification sequences to the planarian transcriptome. Samples having reads mapped to less than 1,000 or more than 9000 contigs were eliminated from subsequent analyses. Data was analyzed using the single-cell analysis Seurat method (Satija et al., 2015). Briefly, genes showing CPM expression of more than 2^4 and a dispersion of 1.5, were selected for initial PCA. Next, principal components that significantly separated cells were determined by a jackstraw analysis (Chung and Storey, 2015) by running the function jackstraw [num.pc = 15, num.replicate = 100, prop.freq = 0.03]. Based on the analysis PCs 1 through 4 were selected, and the list of genes used for the Seurat analysis was expanded by using the function pca.sig.genes [pcs.use = c(1:4), pval.cut = 1e-5]. Then, a second round of PCA was performed with the expanded list of genes, and a jackstraw analysis determined the significant genes in each PC by running the jackStraw function [num.pc = 15, num.replicate = 100]. The top 50 genes contributing to the variance in PCs 1 through 15 were examined manually to identify technical biases (Satija et al., 2015). PC 3 was eliminated from gene selection because most of its highly variable genes consisted of ribosomal proteins, which indicated that it represented technical difference between cells. Dimensional reduction was performed by t-SNE using the run_tnse function [pcs.use = c(1:2, 4:12), max_iter=500, perplexity=20]. Cells were clustered together by the Mclust_dimension function with parameters [reduction.use="tsne", G.use=3.2, set.stat=TRUE, MinPts=3]. Clusters 5 and 14 were

eliminated since they grouped cells suffering from low complexity, based on number of expressed genes. Cell-specific markers were found by using the find_all_markers function with parameters [thresh.test=4, test.use='roc', return.thresh=0.7]. Clusters having the same markers were merged [9 and 24; 4, 5, and 20; 7 and 8; 12, 19 and 23; 13, 16, and 10; 19, 21, 22, 25, and 27; 2, 6 and 26. Cluster 6 cells showing high expression (CPM > 1024) *prog1* and *prog2* (Eisenhoffer et al., 2008) were consolidated with cluster 18].

Detection of differentially expressed genes between clusters

Differentially expressed genes in clusters were detected by running the Seurat find_all_markers function [thresh.test=2] and area-under-the-curve for each gene was calculated with find_all_markers [thresh.test=2 , test.use='roc', return.thresh=0.5]. In addition, clusters expressing the same canonical cell-type markers (e.g., *smedwi-1* or *synapsin*) were temporarily merged, as they might reflect functional relationships of different clusters (e.g. subtypes of a major class of cells). Genes enriched in the merged clusters were found by comparing them to all other clusters by running find.markers [stat.1 = MERGED, stat.2 = OTHER] where MERGED stands for the merged clusters and OTHER stands for cells that were not in the merged clusters. P-values were corrected using the R function p.adjust with default parameters.

Detection of cell-type-specific wound-induced genes

Cell-type-specific wound-induced genes were determined by three analyses. First, the gene expression of wound-induced genes from cells derived from intact animals was compared with the gene expression of cells from wounded-animals using the bimodal expression hypothesis testing with parameters [FDR ≤ 1E-7] (Shalek et al., 2014). Second, gene expression of cells from the wounded time-points was contrasted between a cell type and all other cells with the following parameters pairs using the bimodal expression hypothesis (McDavid et al., 2013) [log FC ≥ 2, FDR ≤ 0.001; FDR ≤ 1E-7]. Finally, adjacent clusters on the t-SNE plot were combined, and the hypothesis testing was repeated. Particularly, late

epidermal lineage and epidermal cell-types were combined; neuronal types; and neoblast subpopulations.

Clustering of unwounded 4C isolated cells

Expression matrix from uninjured neoblasts (n=90) was generated. The *Seurat* method was applied with the following parameters [min.cells = 10, min.genes = 4000, calc.noise=FALSE, is.expr=0.01, do.scale = TRUE]. Cells expressing more than 9000 genes were discarded from further analysis. Gene selection was performed as previously described with the following parameters: mean.var.plot [y.cutoff = 1.5, x.low.cutoff = 5]; jackstraw [num.pc = 6, num.replicate = 100, prop.freq = 0.03]; pca.sig.genes [pcs.use = 1, pval.cut = 1e-3]; run_tsne [pcs.use = c(1,2,3), max_iter=500, perplexity=23]; Mclust_dimension [G.use = 25].

Transcriptome assembly of *G. dorotocephala*

Sequencing reads from all samples from *G. dorotocephala* were combined. Adapter sequences were trimmed with trimmomatic (v0.30) (Bolger et al., 2014) with the following parameters [LEADING:3 TRAILING:3 ILLUMINACLIP:TruSeq2-PE.fa:2:40:15 SLIDINGWINDOW:4:15 MINLEN:30]. Reads containing long stretches of polyA or polyT (>20 nt) were removed using a grep command. Assembly was performed with trinity (release r20131110) (Haas et al., 2013) with the following parameters [--seqType fq --JM 100G --output dor_PE --CPU 6 --min_contig_length 200 --SS_lib_type FR]. Redundant sequences were detected and removed by running cd-hit-est (v4.5.4) (Fu et al., 2012) on the output file with default parameters. Orthologs with *S. mediterranea* were identified with proteinortho (Lechner et al., 2011) with parameters [-p=blastp+ -e=1e-7 -cov=0.35 -pairs –singles].

Detection of onset and offset of wound induction

Expression data from each time course were used for fitting by the *impulse* model (Chechik and Koller, 2009; Chechik et al., 2008) using a published Matlab implementation (Sivriver et al., 2011). Fit for each wound-induced gene was produced for function fit_impulse_params_constrained [expression [log2(CPM+1], retries=100, time points = (0,

1, 4, 12, 16, 24, 48, 72, 120) for all time courses except for the head regeneration time course by Liu et al (Liu et al., 2013) that lacked 1 hpi time point and was called with (0,4,12,16,24,48,72,120)]. Internally, the fit function was called with constraint parameters [t1 ≥ 0; t2 ≥ 0; h0 ≥ 0; h1 ≥ 0; h2 ≥ 0; β1 ≥ 0; β2 ≤ 0]. Expression values for every time point within the 0-120 hpi range was extracted using the impulse function [fit parameter output, time point 0-120]. Genes used for onset of anterior regeneration analysis were collected from previous publications (Gurley et al., 2010; Reddien, 2011; Scimone et al., 2014; van Wolfswinkel et al., 2014; Vogg et al., 2014), and filtered for extremely lowly expressed genes [minimal expression 2 and standard deviation of 0.3]. A median group fit was produced by using median z-score values in each class [Genes associated with specialized neoblasts, patterning factors, and differentiated tissue markers] and a using baseline value for the three classes at 0 h time point as -1.5; to allow convergence 4 hpi samples were corrected by subtraction of 0.05-0.1, as for regeneration related genes their expression was almost identical to the 12 hpi gene expression. Fit function was called with constraint parameters [t1 ≥ 0; t2 ≥ 0; h0 ≥ 0; h1 ≥ 0; h2 ≥ 0; β1 ≥ 0; β2 ≤ 0]. Following filtering the following contigs were used for *S. mediterranea*: dd_Smed_v4_11372_0_1, dd_Smed_v4_11521_0_1, dd_Smed_v4_13056_0_1, dd_Smed_v4_13215_0_1, dd_Smed_v4_13898_0_1, dd_Smed_v4_14611_0_1, dd_Smed_v4_14633_0_1, dd_Smed_v4_15104_0_1, dd_Smed_v4_15144_0_1, dd_Smed_v4_15178_0_1, dd_Smed_v4_15253_0_1, dd_Smed_v4_15516_0_1, dd_Smed_v4_16375_0_1, dd_Smed_v4_17385_0_1, dd_Smed_v4_17726_0_1, dd_Smed_v4_17731_0_1, dd_Smed_v4_21717_0_1, dd_Smed_v4_21801_0_1, dd_Smed_v4_856_0_1, dd_Smed_v4_9774_0_1, dd_Smed_v4_9893_0_1, dd_Smed_v4_11285_0_1, dd_Smed_v4_12674_0_1, dd_Smed_v4_13487_0_1, dd_Smed_v4_13985_0_1, dd_Smed_v4_15531_0_1, dd_Smed_v4_19866_0_1, dd_Smed_v4_5102_0_1, dd_Smed_v4_6604_0_1, dd_Smed_v4_8832_0_1, dd_Smed_v4_11968_0_1, dd_Smed_v4_12112_0_1,

dd_Smed_v4_12647_0_1, dd_Smed_v4_12653_0_1, dd_Smed_v4_14207_0_1, dd_Smed_v4_16476_0_1, dd_Smed_v4_16581_0_1, dd_Smed_v4_17854_0_1, dd_Smed_v4_20433_0_1, dd_Smed_v4_29533_0_1, dd_Smed_v4_3135_0_1, dd_Smed_v4_6710_0_1, dd_Smed_v4_8392_0_1. Orthologs were identified for *G. dorotocephala* as described above and the following contigs were used: comp14905_c0_seq1, comp25657_c0_seq1, comp28223_c0_seq1, comp28241_c0_seq1, comp28262_c0_seq1, comp28562_c0_seq1, comp28762_c0_seq1, comp29894_c0_seq1, comp29915_c0_seq1, comp30125_c0_seq1, comp30289_c0_seq1, comp31066_c0_seq1, comp31293_c0_seq1, comp31342_c0_seq1, comp31414_c0_seq1, comp32324_c0_seq1, comp37945_c0_seq1, comp3844_c0_seq1, comp4224_c0_seq1, comp5212_c0_seq2, comp8621_c0_seq1, comp8817_c0_seq1, comp27033_c0_seq1, comp27221_c0_seq1, comp27470_c0_seq1, comp28896_c0_seq1, comp3788_c0_seq1, comp4439_c0_seq2, comp5124_c0_seq1, comp5348_c0_seq1, comp7983_c0_seq1, comp17302_c0_seq2, comp25468_c0_seq1, comp29782_c0_seq1, comp29980_c0_seq1, comp31706_c0_seq1, comp32106_c0_seq1, comp43392_c0_seq1, comp5522_c0_seq1, comp6277_c0_seq1.

Primers used in this study

Contig	Primer A	Primer B	Primer C
dd_Smed_v4_10259_0_1	ACGCAGAGGCTTGCAGTT	TTGGTCTGTGTGCAGCCA	GCCACAAATTGCACCGCA
dd_Smed_v4_10337_0_1	AAAAGACGCGATGAGGCA	TGTCCTTTGCAATTTATTCGCGA	CAGCCAGGTCACAGTGGC
dd_Smed_v4_1039_0_1	TGTTTCGATTTCTAGACGAACCG	TTGGCCGGAATATTCTCATCA	GCATCACCACTTTCCACAGG
dd_Smed_v4_1054_0_1	CCGGAATTCACGGGCCAA	TGTAGAATGACTCGAATCTCGGA	TTGAGTGTCCGCTGCTCG
dd_Smed_v4_10569_0_1	CGCGTTCCCAATGACAGC	TGAAGGCGGTGTTCCTGAC	ACAGATAACCCTGCAAGATCCT
dd_Smed_v4_10584_0_1	CCGCCGTACAGTATCATGGA	ACCAATAGAGACAGTTCAGCCA	ACGAAATTGACAACGCTAGTGA
dd_Smed_v4_10624_0_1	ACGAGCCAATGTCCAGCC	TATGTGTTTACGAGTGCGATTTT	CACCGGGTGACGCATGAA
dd_Smed_v4_1071_0_1	ACGGGTCGACGTCAGTTG	TGCAACACAAATCGTAAACAGA	GTCCTGACGCACGAGGAA
dd_Smed_v4_10716_0_1	CGGTGAGCGGTGTGTGAT	TCGATTTCAGTTGCATTTGTGGA	TCCCGGCATACAAGAGCAC
dd_Smed_v4_10730_0_1	GCAATCAGCCAACTCGGG	TTTATTAAAGAACCCGAAAGCGT	TGGGGTGCCGGATACAGT
dd_Smed_v4_10776_0_1	GACATTTGGCAGTCCTTCCTG	CGAACTTGCTCCCGGACA	GGGGTATCTGATTATGACTGAGC
dd_Smed_v4_10868_0_1	TTGGGCTGCGGGATTTGG	GGAGCATTGATAAGTTGTTCTGT	TCGGCAACAAACTCCTCGA
dd_Smed_v4_1087_0_1	ACCAGAACCGGAAGCTCC	TGTCGCTTTCAATAAAGGCAAA	TGTTTGCTCACGTCCTCTCC
dd_Smed_v4_10927_0_1	ACAACGAATGGCAGAGTGAGT	TTTTGGAGTGTGTATGCATGAGA	CAACGCAGAGTTCTGTCAAAA
dd_Smed_v4_10930_0_1	ACCAAATTCTATGCAAAGTCGTT	ACACAGTGTTTTGGTTTCCACC	TGCGGCATTATATTTGCGGA
dd_Smed_v4_11074_0_1	CCGGCTGGTTCTGTCGAG	TCAATGAACATTATGGTCCCACC	CTCCCCGCATCGAAAGCA
dd_Smed_v4_11115_0_1	TGCCTAGAGACGACTGCTCT	TGCATTGAAATTCTGCCTTTGGT	TGCGGTGCTTGCTCATGA
dd_Smed_v4_11134_0_1	GGCCTTCTTAGCGATGCGA	ACTCTGCTCCACCACACAG	CTGGCGCTGACAATCCGA
dd_Smed_v4_11142_0_1	AGGCTTCACTGTCGGTTCG	TGTCCATGTGTTCACCAGTCA	TGTGACTCTGCGCTGACG
dd_Smed_v4_11216_0_1	CTCGAGCTGACGCGGAAA	TGACTGCGTCCATAGTGTTGA	TCTCCAAGGGGGTGCAGT
dd_Smed_v4_11220_0_1	AGGAACTTGAGGACATTTCCGC	GTTCTTCGGATAATTGTCCACCT	CAAATTTTCAATCCATCCCGACA
dd_Smed_v4_11254_0_1	TTCAATTTCATTCACGCATGTGG	TGACATTTTCGATCGTTGCGT	ATATCCTTGGCTTTGTACACTGA
dd_Smed_v4_11501_0_1	TGTCGCTCAATATGCAGGCT	TCGTGCTAACTTCCAGGGA	AATTCGACTTGCGGTGCC
dd_Smed_v4_11512_0_1	AACTCGTCTGTGCTGCGA	TCCCAGCGACATGATTGGT	TGGTGGGACATTCATAATGGC
dd_Smed_v4_11561_0_1	TGGGCAACTGCATTGGGA	CAACGAAAATCCCTCTAGCTCC	TCAAAGCTGCTTCGGGGG
dd_Smed_v4_11608_0_1	GGCCGATCAGTGCACCTT	ACGGAGAAATGTCCCCAGG	CGACTTGATGGGCCCACA
dd_Smed_v4_11629_0_1	TGCTTCCTCATTGGCGGA	GCTCCACATCCAAATGGGC	CCACATGCCATAAACACCCG
dd_Smed_v4_11635_0_1	GAGTGATCTAGCGATTTGATTGG	TCCTCGATGCCTATGGAAACT	ATTTTGCAATAGGCCCATCAGT
dd_Smed_v4_11693_0_1	CAGTGGATGGTTGCCGGT	AGCTGATCCAGAAATGCCTAA	TAGACGGGCTGTTCGGGT
dd_Smed_v4_11824_0_1	TGCTCTGTGGCACTGACG	TGTGAGAAACGCTACGATCAA	ATGTCGCTTCCCACCGTC
dd_Smed_v4_11858_0_1	TCACAGAAAACCCAGTCCCC	TGCAGTTTCAACAAAAGATTCCT	ACTATTTCGTTCAATGGACGACA
dd_Smed_v4_11943_0_1	ACACCATTCCATACGCCGA	TCCAATAACTCGAGCAATATGGT	TTGATTGAGGCCGCTGCA
dd_Smed_v4_12081_0_1	TGGAAACCAGGGGGCTTT	TGTCATCGTTTACTGTGGCT	ACGGTAAATGTGCGATGAACG
dd_Smed_v4_12210_0_1	TCGGACGCAGATTCAGAAACT	CCAATACACAAGCTTATGACACG	GGAATGGCTGTTCCGGGT
dd_Smed_v4_12467_0_1	GCAGTTTGCGATCTGTATTGC	TTGGAATCGACTGACGGAAG	TGGGTTTGCTGTAATTGGCA
dd_Smed_v4_12472_0_1	CCGTTCGATTATGCGGCC	CTCTGTACGGATATTCCCAATCA	CGTCACGCAATTCGACCG
dd_Smed_v4_12575_0_1	CCCCTCTACGAGAAATTGCTGT	TGGGCTAGCTTAATACTTTGCA	TCGGAGAAGGCGAATTCGG
dd_Smed_v4_12619_0_1	AGCATGTCAGGAGCTCGA	ACAATTACCACATCAATGGGACA	GGCTTTGGTTTAGGCTTTGGT
dd_Smed_v4_12634_0_1	GCAGGTCTTGAGGCAGCT	CCTGTCCATATAACACTGGAACA	TGTATCAGGGCAAACGAGTT
dd_Smed_v4_12695_0_1	CCATCGAGACCGCGTTGA	CAAATCGGTTTCGGAAAGTTTCA	TCGGCTGCTGTTTGCTGT
dd_Smed_v4_13056_0_1	ACAGTGGGCGATTTTCTCCT	TCTATGGATTCCCCGAAGTCC	TGACACCAAGGTTGAGGCA
dd_Smed_v4_13061_0_1	TGCAAAACAATACTAGCCAATGC	TGCGAAAGTTGTATCAATCCGT	TCGATGAAGTCATATTTCCCGT

dd_Smed_v4_13186_0_1	TCCCTGCCATTAGTACGACA	AATAGATCCGGATGAATTGCTTG	AGGAAAAGGGGGAGGGCC
dd_Smed_v4_13188_0_1	ATTGAAATTTCTTCACTGACGCT	TGTACTCGTCTATCGCTTGCA	GACTCTAAAATGGATGCCGAGC
dd_Smed_v4_13216_0_1	AAACTGCCGCGACGAAGA	TGTTTGGTGAAATGTTAGAGCAA	CGGCGGACTATGACCTCG
dd_Smed_v4_13318_0_1	CAAGTGGTGTTACATTTTCAGCA	TCAAAGGCCAAATTCTGCCT	TGACATCAATTAGCCCTGGAAA
dd_Smed_v4_13356_0_1	TCCAACTTGAACCATGTCGGA	GTCCAATTCGATTGTGAACCGA	TGTTGCAGTGGGGCTCAG
dd_Smed_v4_13468_0_1	TCCAAGTGGATTCGGGCA	TGGACGAAAATGACAATTCTCCT	AGGAGCATTGTCGTTGGCA
dd_Smed_v4_13487_0_1	ACGCGTGACTGAGTTGGT	TCGGACTACCCCATTTGCAG	TGATTGTTGAGATTGGCGAGT
dd_Smed_v4_13835_0_1	TGACTGCCAGTGTGTTATCAGA	ACACGAATTGGTTGGATCAAACC	TCCACAGAATTGCGAATCCCA
dd_Smed_v4_13843_0_1	GTAAACGGGACCTCGCCA	AGAAAGTTCAACGCAAGATCAGT	TGTCGAATCTTGCGCCCA
dd_Smed_v4_13860_0_1	CGGTTGATCTGCAATACCGC	CGTTCTCGATTGTGATAGAAAGG	TGTTGGTCAGATACAAGTGCGA
dd_Smed_v4_13985_0_1	TGACCAAGATTTTTCCCCTAAGT	TCATTGGAGATTGGCAAGCA	GGCAGACCGATTGTTGGGT
dd_Smed_v4_14011_0_1	ACTTCTCAACTGTTCAAAATGCA	TTCACTTCGGCATTTGCAACT	AGGTTTAAAACAAAAGCTGCCT
dd_Smed_v4_14068_0_1	TTTGGAACATTTTACGAGAACCG	ACTATAGCGGAAGTTTAATCGGA	TCTTAACAGCTACATGTGCAAGA
dd_Smed_v4_14158_0_1	GCCGAATGTTCATTCAAACCG	TGTCATTTTCAGTAAAAACGGCA	TCGAAAAATTTGCCGACAAGA
dd_Smed_v4_14199_0_1	GCCTTAATCGACGTGTTTGGA	CGGTTCCTCAGATTCCGAGA	TCTTGTTCAAAACGGAGGAACA
dd_Smed_v4_14370_0_1	TGATGCGGCTATTGTTGATTTT	TGCGCTTCCATTTTACCAGC	ACTGTTACGCAACAAAATAAGGT
dd_Smed_v4_14391_0_1	GGCTTCAAAGGCCACGGT	ACCTTTGCTGACAGGAGATGG	CCTCGTCATCAAGTCGTCGA
dd_Smed_v4_14392_0_2	TGTCTCAAACAGAAGTTCGTCAG	TCGTCGATTGAAAGAAATGACCT	GATGGGCGGCCGTATGAA
dd_Smed_v4_14656_0_1	TCGACCCGAAAATGTGTTTGC	TGTTCAGACCCAAGCTACCG	ACCATTTGAAACGTTCAGAAGTT
dd_Smed_v4_14711_0_1	TCAGACTGGATATACCCCATTGC	TGCCGGGAATTCATGAATCG	ATGATTTTGTCTGAAATGTCGCA
dd_Smed_v4_14725_0_1	CCCATTGTCTTTATTGCAAGGCT	CAGAAAATGCAGGAGCTCTGA	GCCAGCCATTTCAGCGAC
dd_Smed_v4_15035_0_1	CGCTGATTCCCAAGCGGA	TGCACTCACTAAAGGTACAGAA	ACAACACGAATTTGTGCAAAACA
dd_Smed_v4_15386_0_1	CGGCCGAAAGAGTCTCCA	CCGATTGACAGTGCGTATTCA	CGCTGTCGGTGTTGTCGA
dd_Smed_v4_15499_0_1	TGGTTTAGATGCGGTTCCAT	GCCCTGTTAGAAATTTATCCCGA	TGCTTCGCAGCCTACGTC
dd_Smed_v4_15531_0_1	GTTGGCCTCTCATCCAGCA	TCCGACAATTATCCGCCTGA	CCCTGTTACCGAGCCTGAC
dd_Smed_v4_15647_0_1	TCACTTATAAAGGCCGCCCA	TTTGCTTCTAGATGAGGTCTGCA	CAAAGCCCACCACTCGAGA
dd_Smed_v4_15715_0_1	TGTGAAACTGTAACCTTGTTCTG	TGATTCTCCATCTCTAGACTCCA	AAACCACTACGTTCCCAAACA
dd_Smed_v4_15787_0_1	GCCATCCCAGATGCCTCC	TGCCAGCATTACCACAGATT	ACGGCTGCTTTGACCTCC
dd_Smed_v4_158_0_1	TGCTGCAACTTCTTCGCA	GCCTCTTCAATAACTTCAGCAGC	CTCCGCTGATCAATCACCGA
dd_Smed_v4_1580_0_1	GTTGGTGAAGGCCATCCAGA	AGTGATGCCATTCTAGATGCACA	TGAGGCACTTGCTGAACGT
dd_Smed_v4_1581_0_1	CTCGGACTTGGGTCTGCC	AGGAAACGATCGTGGATGACT	GGTCACACTCTCTGCACGT
dd_Smed_v4_16092_0_1	TGCCGAAAAACGCAAGCA	TGCAGTAGACTCGAAACCAAA	ACCAAAGCAGGAGAGGAAGG
dd_Smed_v4_16209_0_1	TTTGCAGGCTTCGACCAA	CTGTTTGGATTTCTGTGGCGA	GTCCTCGACCGCAACACA
dd_Smed_v4_16222_0_1	CCAGCGATTAATTGTGTCGAACA	CGGTTCAACGGTTTCAGCA	TGATTTCTTTTACGGGGCTCCT
dd_Smed_v4_16227_0_1	GGTCGGTTTTTCCATCGTGG	AGCTCTCAACCTCAAGATCTACA	CGTCGACGTCTTGTGAGGT
dd_Smed_v4_16605_0_1	TTGGCTTTACGTTGGCATTTCT	CTTTCATGTGTATTGGCTGTGAT	AGTCGAAGTGGTCAACGCA
dd_Smed_v4_16842_0_1	AGCGTCCTTTGCGAGACA	CCTCAACTCCAAATGCTAAAACA	GGACCAGCTCATGACCCG
dd_Smed_v4_17385_0_1	TGGAACGCTATAAGTCGGTGA	TGGCGGTTCACATTTCCA	TCGGACCGATTGAAGCGT
dd_Smed_v4_17402_0_1	CGGATAGCGAATACAATTGATGC	ACTCACACAAATAATTGATGCCA	CCATCGGGAAAGCAATTGTCC
dd_Smed_v4_1771_0_1	TTCCTTTACACCGTCCTTTGT	TTGTCACCACAATGGATATCCCG	TCCATATGTTATGAATGGAGGCA
dd_Smed_v4_17726_0_1	GCAAGAAAACCGGCAGGG	CGAGTGATCCTGGAAACATTGC	ACTCCGGAGCGAGACCAT
dd_Smed_v4_1846_0_1	ATGGAACCGCAGCAAGCT	TCAAATGTGGCATGGATTTTCGT	GTCGACAGGGCCACTTGT
dd_Smed_v4_18818_0_1	GCGCTTGTTAATCTGGTCCC	AAGAGTGAAATCAAAATCGCGT	TGGAAAAACCAGCTACAATTCCA
dd_Smed_v4_1921_0_1	TTATCGGCAGTGTCGCCC	TCCTTTATTTTGGCGAGGCA	ACTATGGAGCAATACGCAGGA

dd_Smed_v4_19428_0_1	CCGAAGACGATTTGCAACGT	TGCCATCGGAATTACAGGCT	ACAGTTAGGCCATACTCAAATGA
dd_Smed_v4_19592_0_1	ACTCGGGTTTAAATGCACCAC	ACCAGTGTGACTATCTTTTGTGC	CGGCGATTGGCTGCTTCT
dd_Smed_v4_19826_0_1	CGACAATCGGCCTGAGGT	TGACATATTCGAAAACCAACCTC	AATGGGAATCACGGCGCA
dd_Smed_v4_1986_0_1	GCCGCTGGATCTTTTTGCA	TCTGCATAGCGGGATCACT	AGATCCGCGGCTTTTTGT
dd_Smed_v4_1999_0_1	TGATCGCCACTCCGAACG	CCTGATCGAAGCAGTTCCAGA	TGTCGTAGGAGGACGCCA
dd_Smed_v4_20048_0_1	TCATCGGAAAATCACCTGCT	ATCAGAAACCTGTCCAATGGT	TGTCAGGCTGAATGGTCGG
dd_Smed_v4_20122_0_1	ATTACTTCCGCCGAGAGAAGT	TCATTGGAAATCGACATGAGACA	AGTCATTTTCAACATGAACGGCG
dd_Smed_v4_20133_0_1	CGGCCGATCTCAGCCAAT	GGATTGAAAGCCGCGAAATCA	GCTTCAACAACGCGTCCA
dd_Smed_v4_20318_0_1	TGAATGCCCAATGGTCGCA	TCGAAGAGAGAGTAGAACGAGC	TGGACGCAAGCACTGTCC
dd_Smed_v4_21069_0_1	TGTGGCAATTGCATGGTGT	TGGCTGAAACAAGTCAAATCCG	CGACAAGTCGCAACATTTGT
dd_Smed_v4_21717_0_1	TGACCACTTCATCTGTTGACA	AGGGCCAAAGAAGAAGCCG	AGTGCACATGGAAATGGACCT
dd_Smed_v4_2172_0_1	AGAAGGAATCGGACTGTTTGGA	TGAGAGACCAAGTGACAAAGAA	TGGAATGGCCAAGGCAGA
dd_Smed_v4_22031_0_1	TCGTTTCTTGGGCAGTCGA	ACTCTCTCAGCAATTTTGAGTGA	TGCGGCTGCTGGGTAAAA
dd_Smed_v4_22061_0_1	AGATTTTGACATATGTTGCCTCG	TCGATGTCTCCTTCATCAGACG	AACTTTGACACAACCACAAGAGA
dd_Smed_v4_22479_0_1	TCACAGCGATGTGGAAGACA	AGCAACAATCCAGAACTCGA	GGAGCGGAAGGGAGGAGA
dd_Smed_v4_22918_0_1	TCAAGTTGCGAGGCCTTGT	TGCCAAATATGTACAGCAACGA	AGCCTAATGAATGAGTCGAAAGT
dd_Smed_v4_2324_0_1	GCGCCACCACTGTATCGA	CAGCTATCAGATGGTCAAAGTCA	GTGTTGCTGGACCCTGCT
dd_Smed_v4_23420_0_1	TCCAACTGTGTAAATGGGGTGA	TTTCTTGAAAGTTGCGTCCCG	TCCTCACAACAAGAAAACGGA
dd_Smed_v4_23666_0_1	TCTCCAACAATCTCCATCCGT	TCGGCTTTGGAAAACCGA	AGGAATCTACCGAAATCCTTCAA
dd_Smed_v4_2394_0_1	TGGAATGCCAACATTTTCTCAA	GAACCCTTCAAGATACCATGACA	GCGAATAAAAGGAAGTACTGAGC
dd_Smed_v4_24180_0_1	TGAATGATCCGCAATCCAGT	AACGTTCGCTGCAATGACG	TGAGATACCCAACGATTTCGCA
dd_Smed_v4_2442_0_1	GCTCACTGAGTTTGCGTATGC	TGACAAGTCTTCCCAGAATTCCT	GGTTTCAATGATCATTGTTGGCG
dd_Smed_v4_246_0_1	TGCACACAACCTCATGAGCA	GGTAGATCGTTCTGCAATGCA	AGTTGAACCTCCAGACAACACA
dd_Smed_v4_2575_0_1	TGGAAATTCGCATTGTTGTTGT	TCTGCAGTTCTCGCCGAT	TGTCAATCATCCAAGTCTGACA
dd_Smed_v4_2582_1_1	TCCAAGGAGGGAATGGTGGA	TGTACACGAACTGGGCGG	ACGACAAGATAACCGCTCACA
dd_Smed_v4_2588_0_1	ATGGCAGCCGGTGATGTT	AGCTATGCGAGGAAACTATTGA	TCAAATCCCAATCCTGATCGT
dd_Smed_v4_26705_0_1	TGCCTTCTTTTTCGGTGGA	TCATGTTTGTCTTTTGTCAACGA	GGTACTTAATGACAGTTGCAACT
dd_Smed_v4_26780_0_1	TCGAGTTTTCCCATGTTGTGAC	TGTGTCGTGGTTCGTTCCC	CAAACGGTAAATTTGCCAAGAGA
dd_Smed_v4_2679_0_1	ATATCGGTCAGGCTGGCG	TGCTGGGAGTTGTACTGTCT	CCTCATCTTCGTTATCGTCTTGA
dd_Smed_v4_2789_0_1	ACTCGAAGCGGAAGAAGTGG	CCAATCATAACTGCGTCATCACA	TCTCTGTACACACGCCGC
dd_Smed_v4_28398_0_1	CGTCAATCATCTCAGAAACACCA	CAATATGCTTTCACCAGACACCA	TGACATTCAACTTTGCAACACCA
dd_Smed_v4_2844_0_1	TCAGCAGCAGCAGCATGT	CCGCTGCTGATGCCACTA	AGCAAACGGCCGATGTAGA
dd_Smed_v4_28487_0_1	TGTTGGTGGTCTGTTTTGGTC	TGCCCATTTTGTGTTGCCT	ACAGTAATCGATTTGGAGTTTGG
dd_Smed_v4_30088_0_1	TCTGTCACGGTGATGTTTGT	ACAGCAGTTGATTATCAAGGCG	TCAGCCAATGGAAAATCAGTTGA
dd_Smed_v4_3012_0_1	CTCGTCTCGCAAGCGTCA	ACAAGCTCCATATGGAAGAGGC	CGGCGATGTCTGCTGTGA
dd_Smed_v4_3040_0_1	TGAAGGACGAATGTGACGGT	TTCTCGGTTTATTGTTGGACAAC	AGCTTGTATGGCGCTACACA
dd_Smed_v4_30891_0_1	TCGGGCAGCTTCCTTGTC	AGGTCCATGTGCAATGTGGT	TGTAAGCACATTGAGTTACAGGA
dd_Smed_v4_31236_0_1	TGCTTGGCCTTGTCGGTC	TGGTTAAGCATTTCTGTGGGTC	GGCGACATGACATCGTCCT
dd_Smed_v4_320_0_1	TGAACCAACAGCTGCTGC	CGCTGAACGCAATGTGTT	CAGCCCCTCCTGGTCCT
dd_Smed_v4_3257_0_1	GCGACGTCATTAAGAAGCTAGTG	GATGCAGTGTAGTGAAAATGTCA	ACTGTTTGCCACGCAGGA
dd_Smed_v4_3259_0_1	ACAACGCTTCCATCAACAACA	TCCTCACCTTCATCATCTTCGA	TGGCAGTGCTCAAAAGTACAC
dd_Smed_v4_3269_0_1	TGCAGTTTCTCAATGTCATGACT	AGAATCGCAAGGAGTTGGTGT	CCCATCAGTCCTAGATCGGC
dd_Smed_v4_32934_0_1	AAAGACGACGAAGGGCGC	TCCATCATGCAGAAAGTCGGT	AGGCTTCCAAATCTTTTTCTGTG
dd_Smed_v4_33456_0_1	TCCGACTCAGTTCATGACCA	CTCTTGAAACATCTTTGCCAGGA	GCTCGGAGCGAATGGAAA

dd_Smed_v4_3362_0_1	AATGTGTGGTCATTGGGGATG	TGCAGTTGGGAAAACATGCA	CAAAACTTTGTGCGTTTCCGT
dd_Smed_v4_345_0_1	CGGCGAGTATAAATCGGGGG	ACAAAATGCAATTCAACATGCAA	TCCGTTCTTTTGGATCATGGT
dd_Smed_v4_3541_0_1	AAATGACGGATTTCGCGCC	GCTCAGCTCACATATTGCAGG	ATTCAATGTGGGAAATTTGCACA
dd_Smed_v4_3603_0_1	GCCGCACTAGAGTTGGCA	AGCGAGCGATGTTTATGAAAAGG	AGTGCCACTTCGTGAGCC
dd_Smed_v4_3606_0_1	ACTCTTAATTGTCGCGTTTGTT	AGAATTGACTGAACTCGGAAAGA	AGGTTCATCATAGCATTGGCCA
dd_Smed_v4_3632_0_1	CGTGCTGCGTTTCTTCGG	GCGAAACTTCTGGTGATTGCA	GCACATTTTTGTTGCACAGCA
dd_Smed_v4_3638_0_1	TCCCAAACACTTTGCCAACA	ACAGAAGAAACTTTTTCCCTGCA	TCAGGAAACCGAGGATAAAACT
dd_Smed_v4_3667_0_1	CGTTCTCCGAGTGGCTGG	GGCGAGACACTGAGCTCG	GACCACACGTCGGCCTTT
dd_Smed_v4_3674_0_1	TCTCACAGCCCTCTTCGGA	ACTCAATTTCATAAGAAACGCGT	CCCTCTCGCTCCCTCCTT
dd_Smed_v4_36829_0_1	ATCGACGAAAACCAAATGTTGA	ACAGCAGTCAAGAATACGATGC	GACAGAGATAAATCAGACGGAGC
dd_Smed_v4_3703_0_1	CCAGCAGGGTGCCAGAAT	ACCATGTCTGCTATCAGCTCA	ATTTGGAAATGGCTCGAAGTG
dd_Smed_v4_39545_0_1	ACTAATTCATCGCCACCAACAC	TCGATACAATGAAAGACGACTGG	TCGGAAGTACTTGGAAATTCCCA
dd_Smed_v4_4012_0_1	CAGTGTCACGAAGAAATTGGTCA	AGGGGGTTTCGGAACAGT	ATGAGATCTGACGTTGTCTGAA
dd_Smed_v4_4154_0_1	GGGCTGCTCATGACGTGT	CTCAAAAGCTGATGCCATCGA	GATCGGTCCGCGGGAATC
dd_Smed_v4_4279_0_1	TGGGTTAATTTTATGTTGCACGT	CGTTTTCTGCTTTAACGTTTGCT	ATTTGACAGACAACTGAGTCCA
dd_Smed_v4_4299_0_1	GCAAAGGACCCCATGGCA	ACCCCAAAATGAAACAGTTGCC	ACCGACAAGACACAAAGGACA
dd_Smed_v4_4381_0_1	TGCGTCGACAATGAAATGGA	TGAAATTGGAAAACGGCATGAA	CCCGCGTAGAAATCGGCA
dd_Smed_v4_4392_0_1	TCCGAAATTCTCAGAGCAGATCA	TGGAATCGACAATTGTCTCTGA	TGTTTGGACTTTGATTGCGAGT
dd_Smed_v4_4435_0_1	TGCAATTTATGGGAAAATCGGTG	TCTCAAATGGAAAATCTGTCGCT	TCAAATCTCGACATTCTGCTGA
dd_Smed_v4_4486_0_1	GCCGCCCTCCGTTATGTG	CCCCTCCCAAACTGAAATCCC	TGCCAATTCCACTGCGGT
dd_Smed_v4_4619_0_1	AATCATCCACTTCGATGCCAAC	ATCTATTATCCGAAGAGCCGTCC	TGAAGTTCCCGTAAACAATGTCG
dd_Smed_v4_4793_0_1	GTGGGACTCTGTGCTGTTCA	TCAATTCAAAGTTGTGCACGGA	TGGGAGGTCAGTTGCACTC
dd_Smed_v4_4808_0_1	ACCATCGAAACTCGTGTGCA	ATGGCTCCTAAAGGTAAAGTAGC	ACAACCATCATTGTGGTCCT
dd_Smed_v4_4902_0_1	GGGATTATTCTGGTTTCCGGGT	TGGGCGTCGGTGGAGTAT	ACATCGTATCCAAAACCGCA
dd_Smed_v4_4944_0_1	AGCAGACAAGTGTTTCGTCA	CTCTAATGTGAAATACGGTCAGC	TCTGAGAACAAGAAATCATGCGA
dd_Smed_v4_5081_0_1	TTTCTGTTGTCGCCCCCG	ACAAAGGTGAACTAGGAGTCTTT	ACGGTTCGGCGTACACAG
dd_Smed_v4_5102_0_1	GGTCAGCCAAAGTCCCCC	ACAGTATTTCTTAACACGGGTCA	TGAACCATACGGAGCGGT
dd_Smed_v4_511_0_1	GGCGATACTCACTTGGGAGG	AGGAAAGGATATCACCGATGACA	GCTTGATCTGAGAAAGTGAAAGT
dd_Smed_v4_5120_0_1	TTACAGATCGGCAGGAAGC	TGCACATCGAATGAAAACAGATC	TGAAGTTCTAGAAAATCCAGCCA
dd_Smed_v4_526_0_1	GCTGAATGGGGAAGGAAGACA	TGCAGAAAATGAAATGCCTGGT	AGCCGCTCTAAATGAACCACA
dd_Smed_v4_5390_0_1	GTTGTGGACGTTCCTCTCGA	GGTTTGGCATTGGCATTCAGT	TGCAATCTTGTCAACCATTTCGA
dd_Smed_v4_5406_0_1	TTTCGCTATTTAGATGAGCCGA	TGGCCAGAAATTACTCATGTTGG	TGTTGGTAGCTTCAATTGGGA
dd_Smed_v4_5469_0_1	ACTCTAATGGATCCGAAACTGGA	GGAACTGAAGGATCTGAACCT	ACAACAAGAAATCTCGGTCAGT
dd_Smed_v4_5525_0_1	CTGGTGCTCATAATTTGGAAGCA	GATGTGTTTATGAAACGTCCTGA	GAAGGCTGAGAAATTCGATCGG
dd_Smed_v4_5531_0_1	CGAATCACCCCAGTTCCAGA	GTCTGTCATGACAACCAAACTCT	ACTTGGGGAGTTATCAATTCCGA
dd_Smed_v4_5630_0_1	GTCCTACCGGCGGAAGTG	TGCCGCAAAACTTGTGACTG	GGCACTGACCCTTGCAGT
dd_Smed_v4_5638_0_1	CCTTCTGAAAGGCCTCCATT	ACAACTCTAGGTGTTATTGTCGT	CGCCAACAGTAACATTAGCTGT
dd_Smed_v4_5700_0_1	ACTCAAAATGTCTGGTCGAGGA	GGCTTAGGTAAAGGTGGAGCT	ACTACAATCAACATTTGTGGCCC
dd_Smed_v4_574_0_1	TCGCTGCAGTGCTGATTCA	TGTGCCAGTGTCAAGGCC	AGTGGATCTAAAAAGGCTGTCCA
dd_Smed_v4_5749_0_1	TGCTGTCTTCTAAGCATAAACCA	AGACTACAATAACACCACAGGTC	TCGAGTCTGCTTCATGAATGACA
dd_Smed_v4_5781_0_1	CAGTTGACGCGATCGGGA	ACGCAATTTGACCAGATTCAACA	GAGTGCTGTCCGCTCCAG
dd_Smed_v4_5786_0_1	TGCAAATTTCAGCCGAAAATTTG	TCGATGTTGCAAGGGACAA	TGCTGAGCAATTAAACTTCATCA
dd_Smed_v4_5818_0_1	AAACCATTTTCCCTTGCCAAA	GAGCACCGACACTAGTGGT	AAACCAACTAAAACCGATTTCGA
dd_Smed_v4_5853_0_1	ACTGTTTCAAACATTTCTCCGCA	CAGTCTTCGAATGCAATTAACGA	AGCGAGGTAGTGAATCCTACA

dd_Smed_v4_5862_0_1	ACGATATTTATGCCGCCTATCA	TCCTCAACAATTCGGTACTGAA	GGCTATTTGAATGGATTCTCGCC
dd_Smed_v4_5924_0_1	GGTTCCGGTGCACAAGGA	TGTAGCTGCACTTGATTCGGT	ACCGCTATGTCAAAATCAACCA
dd_Smed_v4_5999_0_1	TTTTTCTGCTCACGGGAAATCC	TCAAATCTCAGTAGGCTAAGGGA	ATGAAAGAATTGATTGCCAACGG
dd_Smed_v4_6047_0_1	GCCCCGAAAACAACAACACA	ACCTGCAAGATCCTCGAGA	TGCCGGATGTTGGTCCAG
dd_Smed_v4_6053_0_1	TGGTGAGGAAATTATGCCTACTG	CCGATCGAATAAGATTTCCAAGC	TGGGCGACGTAGATGTCT
dd_Smed_v4_6075_0_1	TTTTTCAATCTTTCAGCTCTGGC	CTAGAGCGTGTTTTTCTTTACCG	TCAATGATATTGATGATGCAGCC
dd_Smed_v4_6278_0_1	ACATGCCACCGAAGAAACT	TGACTGCATTGAAAAAGGAATCA	TCCTCCTCCTCGTCGAGA
dd_Smed_v4_6349_0_1	AAACCAGTAATTAAGCGACCCT	TGGCTTTTCTTTTATCAGCTGCC	ACCATTGATAAACGTGATGAACC
dd_Smed_v4_6420_0_1	TTTGGAAATTATTGGCGAAGGAG	AGAAAAGCTATTCGTCGATCCGA	TGTTTGTCTTTTTGGAAGAGGTG
dd_Smed_v4_6444_0_1	AAATCCACAAAGACAACAACAGC	GGGTGACCGCTCCTGTG	AAGGTTCTAACTAGCAAATGGGC
dd_Smed_v4_6463_0_1	GACGTTTAACAATCGGCGCT	TGAGTTTTTGTGGGTTCCTGA	ATAGAGAAAGGCCGCAGC
dd_Smed_v4_6562_0_1	ACCGATGCTTGGGGAATGA	GTCGAAGTGAAAGATGTTCGGT	ACAGTGCAGTCGGAACTTGT
dd_Smed_v4_668_0_1	TGATCTTTGCCAAATCAAGCCA	GCTACTTAGCATGGGAGCTACT	TCCAAGTCTAGTCCAAATCGTCA
dd_Smed_v4_678_0_1	CTTTTCAAGCTGAAATCGCACA	GCGATGCACTAGACAAAATTCGA	TAGCGCAGGAAGTCAGCC
dd_Smed_v4_6794_0_1	CACCTTATTTACCCGGGGCA	TGATGCTTACTTTACGAGATGGT	TGGCCAAATTGTAAAACGAGACT
dd_Smed_v4_681_0_1	ACAATTTGCCACTGTGACGTG	CCCCCGATCAGAAAAAGGCT	ACTGCATAGTCCATCATTGCA
dd_Smed_v4_6813_0_1	TGCCTATTTATCCCTTGTCTTGG	CTTCCACAAAATCTCCAATCTGG	ATCATCTGCTGTTGTGGTTTTGC
dd_Smed_v4_6882_0_1	CCTGTTGAAGGGGTCGATT	TGCGGAGAATGTGAATTACCT	GCAATTAACGCTTTGCATCTCC
dd_Smed_v4_6884_0_1	CCGGAGGTTCTTGGCACAA	TGTTGGATATTTGTCGGTGGACT	CGTTTGTGAGTACTTCTTGATCG
dd_Smed_v4_6929_0_1	CCTTGTCACGGTAGCGCA	TCCGTTGTCAATTGTATCTGTCC	CACAGATCCAGCACTCGGT
dd_Smed_v4_6948_0_1	AGCCGGTGTCATTCCTCA	ATTATCTCTGCGAGAACTGGATC	GTGACCGTTTGCGTTTGCT
dd_Smed_v4_7038_0_1	TTCAGCGTGGTCGG	ACAATGCGACAAATGTGCCA	GTTTCTCACCGCTGTGGA
dd_Smed_v4_7063_0_1	TCCTTGCTCATTGCTGCCA	GCTCGGATTAATGGCAGCG	ACGGACGGCTCTTTTCGA
dd_Smed_v4_7128_0_1	TGTGTTCACGAGTTTTTGATTCA	TGATTTAGCTACATCCGAGGAAA	AACGGTGACCAGGCATCG
dd_Smed_v4_7166_0_1	GAAAGTAACCTTTGCCGACGA	ACCACTTGCATTTCAAAAATGGA	TGCCAATTGTGTCATAAACCACT
dd_Smed_v4_7168_1_1	ACAGATGCATGAGTTTGTGAAAT	ACACATCAACAATAGCTCTGACG	AGTTGCAAGGTCAGCGTGA
dd_Smed_v4_7262_0_1	CAACACGCGCAGACACAC	TCCGTTTCTATTTGATCGCCA	TCAGCAATCTGACGAACCTGA
dd_Smed_v4_7295_0_1	GGACTTCGATAAAACACTTGTCA	CACAATTGACATTGGTGTTTCGA	TGTCAACCAGCAAACCGA
dd_Smed_v4_7326_0_1	TGCATATCTGGACGTGGATTAGT	TTCACAAAATTGGAACACGTCA	ACTTTTCTCTTGCAAGTTTCACA
dd_Smed_v4_7413_0_1	TCCATTGAACCAGAAATTCGGC	AGTCGGATGGCAAATGCTGA	AAATTGGGCGCTGAAGCAC
dd_Smed_v4_7444_0_1	ACCAAGACGCAGAGTTGATGA	ATTGCTCCATTTTGGTTTCCAA	CAGCAGCATTAGCATCAGCA
dd_Smed_v4_758_1_1	CCCTGACAGACAGCACCG	CATATTGTCGATACAGGTGTGGG	TTCCCGCTGCTCTTTGGC
dd_Smed_v4_7607_0_1	CATCATGAAGCGAAACACAATGT	TCAAATTGAGACAACTCCGAACA	TGTTACAATGTAGCAGTTGCCA
dd_Smed_v4_766_0_1	TCGTGGCAAAAGGCGTCA	ACAGTTAAAGCGGGAGGC	TGCAACACAGCATAGCACT
dd_Smed_v4_7701_0_1	CCGCTCCAGTACGAACGG	TCAGTGCGATCAAAGAAAAGCA	AAAGCCGACGCCATGTGA
dd_Smed_v4_7731_0_1	AACTCTACCAGTGAAAATCGACA	ACTCCGTTGCCAGGAATTCA	TGCCTGAGCCTTTCATCAGA
dd_Smed_v4_7788_0_1	AGGTACAGGGTTTAAAGCAGCA	ACACCAAGGCGCCAAAGT	TGATTGTCGTTTGTAAATGCCTT
dd_Smed_v4_7921_0_1	ATGGTGCCATTGTCCCGG	TGGATGACGGAAATCAAGGTCA	AACCGAGAGTTGCCGGTG
dd_Smed_v4_8252_0_1	TGACAGTGCCAATTTGCTACA	GGATCCGTGATCATTCTTGGC	ATTTGTGAAGGGCCCCCA
dd_Smed_v4_8302_0_1	TGAAGCTGACAACGGGCA	AGCTGTATCGGTTGAGGCAC	GGTTGACGGTTGAGGGGT
dd_Smed_v4_8340_0_1	GATGCAGTCTGACCCGCA	TGCAACAGGAAGGAAAGTTACTG	ACCCTGTTGAACCAAGCCA
dd_Smed_v4_8356_0_1	GGAACCGTCTATGAATGCGC	CGTAAAGGAAGAATGCCCCCT	CCCGAATATCCCGCTGGG
dd_Smed_v4_8439_0_1	TTGTTACAAATGCACGGTAGTTT	TGAGTTTCGGTGCTATACGGG	TGCAAATGTCCAATTGCAAGACT
dd_Smed_v4_8569_0_1	AGGCTTTGAAACCCAACAGGA	TGCATTGAAAGATCTTATTCCCG	GCCAGCGACAACTTTCGG

dd_Smed_v4_8580_0_1	TCTCGTGCAGTAATTTCTACCGA	AGGAGAAAATGGGATTGCGGA	TCCCCAGTTGCAGTTCGAG
dd_Smed_v4_8829_0_1	TGGGGCAGAATCTTGTGCT	TGGCTGGTCAAGGATTTGGG	TATGTTGACGCAGCGGCC
dd_Smed_v4_8835_0_1	TTCGCCAACCTCCAGCAA	AGGGTGAAGAAGTGTCTCAAGA	GCGCTTTCACACACAAGCA
dd_Smed_v4_8839_0_1	GGATGACGGATTCTCTACGGT	TCAAAATCTTCTGCAAACGTTGA	TCGCCGAAAAATATTTTCCAACA
dd_Smed_v4_8858_0_1	GCGGTTCTTGTCCAGTGGA	TGAGTTGGCCGATATTAACAGT	CGTTCTCCGGTGTGGGTT
dd_Smed_v4_8901_0_1	TGAAGGTTACACTCGGGGG	CTTTGACTGTCAAGCTGGGC	ACTTGCAAGGACACAATTCGAG
dd_Smed_v4_8918_0_1	TCACAGCCTGGGAAAACTCC	CGATAGCATGAACATCATCACAA	TGCGACTGGTAAGCCGTT
dd_Smed_v4_8994_0_1	TAAATGTCGCGGGGCAGT	TGCCAGTATTGGGTGCACA	CCGGCTCCAGAACTGCTC
dd_Smed_v4_9050_0_1	TCCAGTCCGTTGGAAAGGA	TGGTTATGAGGAGAAACTTCGT	TGAATTGTCTGACAAGGCAGGA
dd_Smed_v4_9165_0_1	CGACAGCAAACAGGTAGCC	TGGGGTCAGTACAAAGAAAGAAG	ACGCACAAACCAAACTGACA
dd_Smed_v4_9202_0_1	AGAAATATAACACGGTGTTTGCA	TGACTTGTGCGAATTGTTGCA	ACCTTAAGTGGCGGATGTTGT
dd_Smed_v4_9204_0_1	ACAACTCGATCATTCCTTCTCGT	CGTTCTCGTTTCACCGTCA	TCGAACGCATTATGAGCGA
dd_Smed_v4_9273_0_1	AGATGGCAGTGAACTGGACA	TGGATTAACGCCTCCGCA	TGAGAACTGAACTTTTGGTAGCA
dd_Smed_v4_9402_0_1	GGAGGCTGGGGATGGGTA	TGGTGCATGTATTAGCAGATGGT	TCCATCCTGCCAAGGGGG
dd_Smed_v4_9410_0_1	CTCCTACTGGGAAATTTGGTACA	GACACACCACAACCTTTAGAAGA	TCAAAATTCAGTTTATTGCGGGT
dd_Smed_v4_9416_0_1	TCGAACAACGAGCAACGG	CGTGCCTTCATCATTTTTGGC	AATCGTCCACCCCTCGGA
dd_Smed_v4_9472_0_1	CAATTGTGCGTATTTTGTGGTGT	CGTAATTGGAGCCGGCCA	GATCAAACTAATCGCACCAGCA
dd_Smed_v4_9490_0_1	AGATGACAACCAAAGCCGGA	TGTATCGACAATTTACCGATCGA	AGGGGCCGGTTCAGACT
dd_Smed_v4_9491_0_1	TGAGCCAAAAAGAAGAAAGTGCA	GCATGGAAGATACTCAGGACGT	CGGATCAGATAAGCTCCATTTCG
dd_Smed_v4_9519_0_1	TGCAAAGCTAACGCAGAAGA	CTCTACGGTATTCGACTTTACCA	TCCCATGGAAGCCACGTTC
dd_Smed_v4_953_0_1	AGGACCACCTGGCAGCTA	CCGCAACGGCTGAAACTG	GCTGATCATCCTGCTCACA
dd_Smed_v4_9530_0_1	ACAGCCAGTCTTCGCCAA	TCCCTCGCAGCATTGTGT	ACGCTTCAACCTTTGATCGG
dd_Smed_v4_9546_0_1	CGTTGTTTTCAATGGGTAGCTGT	TTTGGTGAATATTCGCATTCCAT	TCTATCGCGCATGATAGCAA
dd_Smed_v4_961_0_1	GCTTATGCTATGCTCAATGTGGA	TTGGAGACATGGTTCTTAGCCC	AGGCACATCCATAATAGTCTCGT
dd_Smed_v4_9610_0_1	TCAATCTCATTTCTGGACAGTGT	TCCCTCAAATGTCTACGTAGTGG	TAAAATTGCGCTCATTCTGTTGA
dd_Smed_v4_9642_0_1	TGCCACAGACAATCTTGCT	TGGTTCTGCCAATGAGTTAGAT	CTGTCCAACAGCGGCAAC
dd_Smed_v4_9677_0_1	CCGGGGGCCTCAAATTGT	TTCTGCTGACAAAACCTCTCGA	AGCTCATGACGCCCGAAG
dd_Smed_v4_9905_0_1	ACAACAACCGAAAATTGTGCCA	CGTCCTAATTCTCACAATCGCAC	ATTGTCGGTGGGCAGTGG
dd_Smed_v4_996_0_1	ACGGTGTGAATGGATCTTCAGA	AACATGGGAAATGGGTATTGTGA	CCGTTTTGTTCACCGCGG

Additional sequences used for mapping

>rRNA_5s

TCACCCGATCTCGTTCGATCTCGGAAGTTAAGCAGGTTAAGGCCTCGTTAGTACTTGGATGGGTGACCGCCTGGGA
ATACGAGGTGTTGTGGACTTTATACTGTTTGTCCACGACCATACTAATCTGGGTTCACCCGATCTCGTTCGATCTC
GG

>mtRNA_2

TCTAAGCAGTGGTATCAACGCAGAGTACGGGGGGTATTGCACTGTTTAGTTGTGATATTTTCCTTTTTGTATTACG
GTTTGTAGGTATTTTTATGTTTTTATTCTCTCGAATGGAATATGATATATCTTTTGTGTGTTTTGTTTTATTTTCT
TTGGTTAACATTAGAA

>mtRNA_1

AGTTGGTGTTGTTGTTTTGTGCAGGTAAGTTAATTAAAACTAGCAGATTCATGTTCTGTCTATGAGTCCTTTCTCT
GTATATGTGGTTAAGATAGTTTATTCAGAATGTTAATTTGTGGAGTTAATGGTAAAAGACTTGTTTTTCTTAATAT
TTGTTTTAATAGCTTA

>unidentified

CGGCCGGACGTAGCGGTGTGCGTCTGTAATCCAACTACTAGGAGGCCGGGTATATGAATGGTTTGAGATGAGGAGT
TCTGTGAGCATTGCGCCTATGTAGATCGGATGTCCACAGTAAGCTTGGCGTCAACATGGTTATATTGTCGGAGGAT
AGAATACCCAGGTTGT

>SMED_11901_V2

ATGAATGAAATTTTGGAAAAGGATATGAAAGCGATTGAATCCATTAAAGTAAAAGAAAAAAAGGCTGTTGATGGTT
TTATGGGTACCTCATCGTTTCATGGAGTGATTCAAGCATATCATAAACGAAATAAAATTGATAAAGGGAGCTGGTT
CATCAGTTTAGTTATTTGTATGTTTGGCTTAATTGGGCATCTCTACCTAATAATCAGTAGATATATAAGTTTGCCC
ACAACTATTGACATGGTCTCTTCAGTGAATTTTGATCCTTTTCCTGCTGTCGCAATATGTCCGGTTACCTTTATTA
GCAGGGATAAATTCACCAAGTATTACAATACAACTCAAGTTTCCCTTAATAAAAAGCTAGTTGGGGATATTTTCTA
CGTCGATGTAAGTGCCTTGAATTTCTGGAGGTCCCTAAGTAAACAACAAGGCAAAGACATAAACAGTAGTTCAGTT
CTTGGAAAGTATTGGGATGAAGCTGAAACCACTTTCTATAGATTCCAGAAAATGATGAATGTTTCAATAGGTCATC
GAAATTATGAAATGATTTTCTTTTGTGAAATTAACAATAAACCTTGCTCATGGGAACATTTCCTTGAATTCGATCA
TCCGATTTATAAGCGATGTTTTAAATTCTCCTATCCGGTAACTGATGAAGATGAAATTCCAGATAAATTGATATTG
GGGCTTTATGTTGATGATGACTATCAAAGAGACACTGATGATATTAAAACGATAATAACCTCTCATGGAGGAAAGG
TTACTATAAATGAAGCAAGTATTTACCCTGGAACTGAAAGTTCATTTGAACATTTTCCGTCAGGATTCCAAACGAT
GTTTCGATTGAAACAAGAAGGTAGCAGTCAAATCAATAAACCAAGGTCTCCATGCCAAGTTAATACTGATTCAGTG
ATCAACGTTTTCAACGATTATGAATATGATGGCTCAACAAATATCACAATACCATATAAATACAATGTGATACTTT
GCAGACAATACCATCAACAAATAGAATGCGTTAAAAGATGCAAGTGTTTAAATCCGAACATTCCAGTATTTGTTGA
TGCTATTAAGAATTCTGAAAATAAATCATTCTTTTGCGATGAGATTCAGCTTAATTCTTCCTTTTCAAGCATTATT
AATCAGCTTGATTGTCTTTATAATTTAGATTATGATCAGTATTTTAATGAGAATGTTATATCATTATGTTCGGGAT
TGTGTAATCAGGTAGAATATTCAATGTATTCTTATACTATGCCTTGGTTCGGTAAAACAATGATCAAAGAAATGGA
GTTTGTCNAATGAGAAATTCATGGCCCATTACAACAGTCTAATTAATTCCATCCAATTGTGAAGGATTATGGAACC
ATTGAATAGAGCACGCAATTGCGTAATCAAATCCATGAAAGATAATGATCAAGCCAGCTTGTGTTTCGCAATGATT
AATATTCAATTTGAATCTCCCAGAAAAGAAATTATTCGAGAATATGAGGCATATTTATTGGGGAATTTACTCAGTG
ATTTCGGCGGGATTTTAGGACTGTGGATTGGAATGTCTCTGATAACAATTATTGAAATCATATACTTAGCATGCTC
GTTGAGTAAACACAAAACTGAACGCGCTGCTTCAGTTTTCAAAAAGTCAATCCACAAGAGAAGTCTGAAAAGGAAT
TCCGATAAAAACAAAATTATCAGAATCGGAATAGAAAATGAGGCGTATGAAAATTAG

Contig Ids corresponding to genes shown in figures

Fig.	Panel	Gene annotation in figure	Contig	Best-blast hit description	ID	E-value	Organism
1	C	*smedwi-1*	dd_Smed_v4_659_0_1	smedwi-1	DQ186985.1	0	Smed
1	C	*tropomyosin*	dd_Smed_v4_436_0_1	tropomyosin 1 (alpha) (TPM1)	uc002alp.3	1.00E-54	Human
1	C	*vim-1*	dd_Smed_v4_364_0_1	vimentin (VIM)	uc001iou.2	3.00E-29	Human
1	C	*synapsin*	dd_Smed_v4_3135_0_1	synapsin II (SYN2)IIb	uc003bwl.1	1.00E-121	Human
1	E	*znf91*	dd_Smed_v4_7664_0_1	zinc finger protein 91 (ZNF91)	uc002nre.3	3.00E-09	Human
1	E	*zfp-1*	dd_Smed_v4_8720_0_1	ZFP-1	JQ425154.1	0	Smed
1	E	*hnf4*	dd_Smed_v4_1694_0_1	HNF4 (hnf4)	JF802199.1	0	Smed
2	C	*egr-2*	dd_Smed_v4_9273_0_1	clone SMED_20251_V2 early growth response-2	JX010482.1	0	Smed
2	C	*RPSAP58*	dd_Smed_v4_8634_0_1	ribosomal protein SA pseudogene 58 (RPSAP58)	uc002nrn.3	3.00E-22	Human
2	C	*wntless*	dd_Smed_v4_9546_0_1	Evi/Wls	FJ463748.1	0	Smed
2	C	*svopl*	dd_Smed_v4_12695_0_1	SVOP-like (SVOPL)	uc011kqh.2	1.00E-42	Human
2	C	*dd_9490*	dd_Smed_v4_9490_0_1	Smed06730_V2 hypothetical protein	JX010552.1	0	Smed
3	A	*Tob2*	dd_Smed_v4_7444_0_1	transducer of ERBB2, 2 (TOB2)	uc021wqf.1	8.00E-28	Human
3	A	*dd_9519*	dd_Smed_v4_9519_0_1	NA	NA	NA	NA
3	A	*svopl*	dd_Smed_v4_12695_0_1	SVOP-like (SVOPL)	uc011kqh.2	1.00E-42	Human
3	B	*egr-2*	dd_Smed_v4_9273_0_1	clone SMED_20251_V2 early growth response-2	JX010482.1	0	Smed
4	B	*notum*	dd_Smed_v4_24180_0_1	notum	JF725701.1	0	Smed
4	B	*sulfotransferase*	dd_Smed_v4_15647_0_1	sulfotransferase family, cytosolic, 1C, member 3 (SULT1C3)	uc010ywo.2	6.00E-48	Human
4	B	*klf*	dd_Smed_v4_3638_0_1	Kruppel-like factor 13 (KLF13)	uc001zfo.3	9.00E-34	Human
4	B	*TRAF-1*	dd_Smed_v4_4392_0_1	Smed19658_V2 TNF receptor associated factor-1	JX010627.1	8.00E-136	Smed
4	B	*H2B*	dd_Smed_v4_4808_0_1	Smed15708_V2 histone h2b-2	JX010617.1	1.00E-104	Smed
4	B	*dd_6806*	dd_Smed_v4_6808_0_1	NA	NA	NA	NA
4	B	*rhomboid*	dd_Smed_v4_13835_0_1	5B07 rhomboid-like protein	KJ573355.1	0	Smed
5	B	*mex-3*	dd_Smed_v4_6053_0_1	mex-3 homolog A (C. elegans) (MEX3A)	uc001fnd.4	2.00E-08	Human
5	B	*hsp70*	dd_Smed_v4_320_0_1	heat shock 70kDa protein 8 (HSPA8)	uc001pyo.3	0	Human
5	B	*traf2*	dd_Smed_v4_10569_0_1	Smed07121_V2 TNF receptor associated factor-2	JX010549.1	0	Smed
5	B	*sytl2*	dd_Smed_v4_21069_0_1	synaptotagmin-like 2 (SYTL2)h	uc001paw.3	1.00E-35	Human
5	B	*dd_14011*	dd_Smed_v4_14011_0_1	NA	NA	NA	NA
5	B	*CALCR*	dd_Smed_v4_15499_0_1	calcitonin receptor (CALCR)	uc003umw.2	8.00E-22	Human
5	B	*wntless*	dd_Smed_v4_11629_0_1	Evi/Wls	FJ463748.1	0	Smed

5	B	*dd_8302*	dd_Smed_v4_8302_0_1	NA	NA	NA	NA
5	B	*slc16a14*	dd_Smed_v4_9402_0_1	solute carrier family 16, member 14 (monocarboxylic acid transporter 14) (SLC16A14)	uc002vqf.3	8.00E-39	Human
5	B	*mpped1*	dd_Smed_v4_9610_0_1	metallophosphoesterase domain containing 1 (MPPED1)	uc011apy.2	2.00E-16	Human
5	B	*rrm2b*	dd_Smed_v4_5862_0_1	Smed05893_V2 ribonucleoside diphosphate reductase subunit M2	JX010583.1	0	Smed
5	B	*dd_8901*	dd_Smed_v4_8901_0_1	NA	NA	NA	NA
5	B	*dd_9519*	dd_Smed_v4_9519_0_1	NA	NA	NA	NA
5	B	*notch*	dd_Smed_v4_10716_0_1	notch 1 (NOTCH1)	uc004chz.3	8.00E-57	Human
5	B	*jun-1**	dd_Smed_v4_5749_0_1	Smed03061_V2 1-Jun	JX010576.1	0	Smed
5	B	*nlg-1**	dd_Smed_v4_14068_0_1	noggin-like protein 1	EF633691.1	0	Smed
5	B	*inhibin-1**	dd_Smed_v4_7607_0_1	clone SMED_01282_V2 inhibin-1	JX010479.1	0	Smed
5	B	*glypican-1**	dd_Smed_v4_4154_0_1	clone SMED_05117_V2 glypican-1	JX010469.1	0	Smed
5	B	*dd_20048*	dd_Smed_v4_20048_0_1	NA	NA	NA	NA
5	B	*inx-13*	dd_Smed_v4_11501_0_1	INX-13	JQ425145.1	0	Smed
5	B	*cyp2j2*	dd_Smed_v4_2394_0_1	cytochrome P450, family 2, subfamily J, polypeptide 2 (CYP2J2)	uc001czq.3	3.00E-36	Human
5	B	*sbspon*	dd_Smed_v4_5786_0_1	somatomedin B and thrombospondin, type 1 domain containing (SBSPON)	uc003xzf.3	9.00E-15	Human
5	B	*pif1*	dd_Smed_v4_16842_0_1	PIF1 5'-to-3' DNA helicase (PIF1)	uc010uiq.1	1.00E-119	Human
5	B	*dd_13860*	dd_Smed_v4_13860_0_1	NA	NA	NA	NA
5	B	*pxdn*	dd_Smed_v4_3603_0_1	peroxidasin homolog (Drosophila) (PXDN)	uc002qxa.3	0	Human
5	B	*sfrp1*	dd_Smed_v4_13985_0_1	secreted frizzled-related protein 1 (sfrp1)	EU296635.1	0	Smed
5	B	*med12l*	dd_Smed_v4_11943_0_1	mediator complex subunit 12-like (MED12L)	uc003eyp.3	8.00E-64	Human
5	B	*plasminogen-1*	dd_Smed_v4_23420_0_1	Smed27240_V2 plasminogen-1	JX010625.1	0	Smed
5	D-E	*egr-l1*	dd_Smed_v4_7731_0_1	EGR-like protein 1	JF914965.1	0	Smed
5	D-E	*runt-1*	dd_Smed_v4_16222_0_1	runt-like 1 protein	JF720854.1	0	Smed
5	D-E	*Inhibin-1*	dd_Smed_v4_7607_0_1	clone SMED_01282_V2 inhibin-1	JX010479.1	0	Smed
S2	A	*Rab-11B*	dd_Smed_v4_7604_0_1	EF-hand calcium binding domain 4B (EFCAB4B)	uc010sen.1	3.00E-45	Human
S2	A	*anoctamin 7*	dd_Smed_v4_4761_0_1	anoctamin 7 (ANO7)NGEP-L	uc002wax.2	0	Human
S2	A	*ESRP-1*	dd_Smed_v4_5053_0_1	epithelial splicing regulatory protein 1 (ESRP1)	uc003ygt.4	4.00E-117	Human
S2	A	*myoferlin*	dd_Smed_v4_6816_0_1	myoferlin (MYOF)	uc001kio.3	0	Human
S3	C	*plasminogen-1*	dd_Smed_v4_23420_0_1	Smed27240_V2 plasminogen-1	JX010625.1	0	Smed
S3	C	*fos-1*	dd_Smed_v4_2789_0_1	clone SMED_00055_V2 fos-1	JX010471.1	0	Smed
S3	C	*hadrian*	dd_Smed_v4_3606_0_1	clone SMED_02793_V2 hadrian	JX010472.1	0	Smed
S3	C	*glypican-1*	dd_Smed_v4_4154_0_1	clone SMED_05117_V2 glypican-1	JX010469.1	0	Smed
S3	C	*dusp10*	dd_Smed_v4_4619_0_1	dual specificity phosphatase 10 (DUSP10)	uc001hmy.2	7.00E-39	Human

S3	C	*wntless*	dd_Smed_v4_11629_0_1	Evi/Wls	FJ463748.1	0	Smed
S3	C	*egr-l 1*	dd_Smed_v4_7731_0_1	EGR-like protein 1	JF914965.1	0	Smed
S3	C	*egr-4*	dd_Smed_v4_9410_0_1	clone SMED_09938_V2 early growth response-4	JX010483.1	0	Smed
S3	C	*HSP20**	dd_Smed_v4_5406_0_1	NA	NA	NA	NA
S3	C	*innexin-1*	dd_Smed_v4_11254_0_1	Smed09630_V2 innexin-1	JX010623.1	0	Smed
S3	C	*dd_4944*	dd_Smed_v4_4944_0_1	NA	NA	NA	NA
S3	C	*traf2*	dd_Smed_v4_10569_0_1	Smed07121_V2 TNF receptor associated factor-2	JX010549.1	0	Smed
S3	C	*mex-3*	dd_Smed_v4_6053_0_1	mex-3 homolog A (C. elegans) (MEX3A)	uc001fnd.4	2.00E-08	Human
S3	C	*Jun-1*	dd_Smed_v4_5749_0_1	Smed03061_V2 1-Jun	JX010576.1	0	Smed
S3	C	*sfrp1*	dd_Smed_v4_13985_0_1	sFRP1	EU296635.1	0	Smed
S3	C	*MPPED1*	dd_Smed_v4_9610_0_1	metallophosphoesterase domain containing 1 (MPPED1)	uc011apy.2	2.00E-16	Human
S3	C	*7tm**	dd_Smed_v4_20048_0_1	NA	NA	NA	NA
S3	C	*inx-13*	dd_Smed_v4_11501_0_1	INX-13	JQ425145.1	0	Smed
S3	C	*CPO*	dd_Smed_v4_5999_0_1	carboxypeptidase O (CPO)	uc002vby.2	8.00E-63	Human
S3	C	*ldlr-1*	dd_Smed_v4_1581_0_1	Smed05022_V2 low density lipoprotein receptor-1	JX010530.1	0	Smed
S3	C	*SLC16A14*	dd_Smed_v4_9402_0_1	solute carrier family 16, member 14 (monocarboxylic acid transporter 14) (SLC16A14)	uc002vqf.3	8.00E-39	Human
S3	C	*CYP2J2*	dd_Smed_v4_2394_0_1	cytochrome P450, family 2, subfamily J, polypeptide 2 (CYP2J2)	uc001czq.3	3.00E-36	Human
S3	C	*dd_5390*	dd_Smed_v4_5390_0_1	NA	NA	NA	NA
S3	C	*rrm2b*	dd_Smed_v4_5862_0_1	Smed05893_V2 ribonucleoside diphosphate reductase subunit M2	JX010583.1	0	Smed
S3	C	*dd_9642*	dd_Smed_v4_9642_0_1	NA	NA	NA	NA
S3	C	*sbspon*	dd_Smed_v4_5786_0_1	somatomedin B and thrombospondin, type 1 domain containing (SBSPON)	uc003xzf.3	9.00E-15	Human
S3	C	*HYOU1*	dd_Smed_v4_2324_0_1	hypoxia up-regulated 1 (HYOU1)	uc010ryu.1	0	Human
S3	C	*HSP90**	dd_Smed_v4_758_1_1	NA	uc001tkb.1	0	Human
S3	C	*dd_9519*	dd_Smed_v4_9519_0_1	NA	NA	NA	NA
S3	C	*HSP70**	dd_Smed_v4_1087_0_1	heat shock 70kDa protein 4-like (HSPA4L)	uc003ifm.3	0	Human

Supplementary References

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Chechik, G., and Koller, D. (2009). Timing of gene expression responses to environmental changes. Journal of computational biology : a journal of computational molecular cell biology *16*, 279-290.

Chechik, G., Oh, E., Rando, O., Weissman, J., Regev, A., and Koller, D. (2008). Activity motifs reveal principles of timing in transcriptional control of the yeast metabolic network. Nature biotechnology *26*, 1251-1259.

Chung, N.C., and Storey, J.D. (2015). Statistical significance of variables driving systematic variation in high-dimensional data. Bioinformatics *31*, 545-554.

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

Engreitz, J.M., Sirokman, K., McDonel, P., Shishkin, A.A., Surka, C., Russell, P., Grossman, S.R., Chow, A.Y., Guttman, M., and Lander, E.S. (2014). RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent Pre-mRNAs and chromatin sites. Cell *159*, 188-199.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics *28*, 3150-3152.

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. Developmental biology *347*, 24-39.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M*., et al.* (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature protocols *8*, 1494-1512.

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. Development, growth & differentiation *48*, 371-380.

Kao, D., Felix, D., and Aboobaker, A. (2013). The planarian regeneration transcriptome reveals a shared but temporally shifted regulatory program between opposing head and tail scenarios. BMC genomics *14*, 797.

Lechner, M., Findeiss, S., Steiner, L., Marz, M., Stadler, P.F., and Prohaska, S.J. (2011). Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics *12*, 124.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Liu, S.Y., Selck, C., Friedrich, B., Lutz, R., Vila-Farre, 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*, 81-84.

McDavid, A., Finak, G., Chattopadyay, P.K., Dominguez, M., Lamoreaux, L., Ma, S.S., Roederer, M., and Gottardo, R. (2013). Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. Bioinformatics *29*, 461-467.

Petersen, C.P., and Reddien, P.W. (2008). *Smed-betacatenin-1* is required for anteroposterior blastema polarity in planarian regeneration. Science *319*, 327-330.

Picelli, S., Bjorklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nature methods *10*, 1096-1098.

Picelli, S., Faridani, O.R., Bjorklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nature protocols *9*, 171-181.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.
Reddien, P.W. (2011). Constitutive gene expression and the specification of tissue identity in adult planarian biology. Trends in genetics : TIG *27*, 277-285.
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*, 1327-1330.
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*, 139-140.
Sandmann, T., Vogg, M.C., Owlarn, S., Boutros, M., and Bartscherer, K. (2011). The head-regeneration transcriptome of the planarian Schmidtea mediterranea. Genome biology *12*, R76.
Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nature biotechnology.
Schwartz, S., Bernstein, D.A., Mumbach, M.R., Jovanovic, M., Herbst, R.H., Leon-Ricardo, B.X., Engreitz, J.M., Guttman, M., Satija, R., Lander, E.S.*, et al.* (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell *159*, 148-162.
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*, 339-352.
Shalek, A.K., Satija, R., Shuga, J., Trombetta, J.J., Gennert, D., Lu, D., Chen, P., Gertner, R.S., Gaublomme, J.T., Yosef, N.*, et al.* (2014). Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature *510*, 363-369.
Sivriver, J., Habib, N., and Friedman, N. (2011). An integrative clustering and modeling algorithm for dynamical gene expression data. Bioinformatics *27*, i392-400.
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*, 326-339.
Vogg, M.C., Owlarn, S., Perez 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. Developmental biology *390*, 136-148.
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. Cell stem cell *10*, 299-311.
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 & development *26*, 988-1002.