Molecular Signatures of the Three Stem Cell Lineages in *Hydra* and the Emergence of Stem Cell Function at the Base of Multicellularity

Georg Hemmrich,[†,1] Konstantin Khalturin,[†,1] Anna-Marei Boehm,[†,1] Malte Puchert,[1,2] Friederike Anton-Erxleben,[1] Jörg Wittlieb,[1] Ulrich C. Klostermeier,[3] Philip Rosenstiel,[3] Hans-Heinrich Oberg,[2] Tomislav Domazet-Lošo,[4,5] Toshimi Sugimoto,[6] Hitoshi Niwa,[6] and Thomas C.G. Bosch[*,1]

[1]Zoological Institute, Christian-Albrechts University, Kiel, Germany

[2]Institute of Immunology, University Hospital Schleswig-Holstein, Kiel, Germany

[3]IKMB, University Hospital Schleswig-Holstein, Kiel, Germany

[4]Laboratory of Evolutionary Genetics, Max-Planck-Institute for Evolutionary Biology, Plön, Germany

[5]Laboratory of Evolutionary Genetics, Division of Molecular Biology, Ruđer Bošković Institute, Zagreb, Croatia

[6]Laboratory for Pluripotent Stem Cell Studies, RIKEN Center for Developmental Biology, Kobe, Japan

[†]These authors contributed equally to this work.

***Corresponding author:** E-mail: tbosch@zoologie.uni-kiel.de.

Associate editor: Billie Swalla

Abstract

How distinct stem cell populations originate and whether there is a clear stem cell "genetic signature" remain poorly understood. Understanding the evolution of stem cells requires molecular profiling of stem cells in an animal at a basal phylogenetic position. In this study, using transgenic *Hydra* polyps, we reveal for each of the three stem cell populations a specific signature set of transcriptions factors and of genes playing key roles in cell type-specific function and interlineage communication. Our data show that principal functions of stem cell genes, such as maintenance of stemness and control of stem cell self-renewal and differentiation, arose very early in metazoan evolution. They are corroborating the view that stem cell types shared common, multifunctional ancestors, which achieved complexity through a stepwise segregation of function in daughter cells.

Key words: aging, adult stem cell, self-renewal, differentiation, evolution.

Introduction

The evolution of multicellular animals most likely started from a number of multifunctional cell types that existed in the metazoan ancestor (Arendt 2008; Arendt et al. 2009). To maintain tissue homeostasis and repair function, these ancestral cell types, which simultaneously carry out a number of different functions, must have had properties known from adult multipotent stem cells. The evolutionary origin of stem cells is unknown. Stem cells might have originated from flagellated cells at the surface of the body of the ur-metazoan (King 2004). Proliferative cells in sponges also have features that justify referring them as unipotent stem cells (Funayama 2010). Yet, the origins and mechanisms for establishing stem cell populations, however, remain obscure and the search for a clear stem cell "genetic signature" continues. Cnidarians are not only among the earliest-known phyletic lineages known to contain stem cells (fig. 1*A*) (Hemmrich et al. 2007) but also possess most of the gene families found in bilaterians (Putnam et al. 2007; Dunn et al. 2008; Philippe et al. 2009; Schierwater et al. 2009). They have retained many ancestral genes that have been lost in *Drosophila* and *Caenorhabditis elegans* (Kortschak et al. 2003; Kusserow et al. 2005; Miller et al. 2005; Technau et al. 2005; Chapman et al. 2010). The genome of *Hydra magnipapillata* has been sequenced and analyzed and is a major resource for understanding the molecular "toolbox" of the earliest common ancestors of metazoans (Chapman et al. 2010). Several cnidarians are used as model organisms in developmental biology, including *Nematostella* (Anthozoa), *Clytia* (Hydrozoa), *Hydractinia* (Hydrozoa), *Acropora* (Anthozoa), and *Hydra* (Hydrozoa). All of them have their own benefits and are informative for an understanding of bilaterian evolution and development (Technau and Steele 2011). *Hydra* is the only cnidarian where the cell lineages and the differentiation pathways have been completely understood and a transgene technology is developed to a level that fluorescence-activated cell sorting (FACS) of individually labeled cell lineages is possible. There is some evidence that cnidarians differ in their stem cell differentiation pathways because a interstitial cell lineage seems to be present only in hydrozoans (Technau and Steele 2011). Tissue function, behavioral traits, and sexual reproduction in *Hydra* are

based on three tissue-specific stem cells: ectodermal and endodermal epithelio-muscular cells and interstitial stem cells (Bosch 2009) (fig. 1*B* and *C*). Two lineages of epithelio-muscular stem cells shape the diploblastic body of the polyps and are responsible for all morphogenetic processes (Fujisawa and Sugiyama 1978). Multipotent interstitial stem cells located in the ectoderm of the gastric region have a developmental potency, which is much wider than that of epithelial cells by being capable not only to give rise to a number of somatic cell types but also to gametes (Bosch and David 1986; Bosch 2009; Bosch et al. 2010). Over the past years, we and others have characterized some of the features of *Hydra*'s stem cell system (Wittlieb et al. 2006; Kasbauer et al. 2007; Khalturin et al. 2007; Hemmrich and Bosch 2008; Siebert et al. 2008; Watanabe et al. 2009; Hartl et al. 2010). In *Hydra*, both epithelial and interstitial stem cells seem to rely on signaling pathways involving Notch (Kasbauer et al. 2007) and glycogen synthase kinase-3 β (GSK-3b) (Khalturin et al. 2007; Bosch 2009; Hartl et al. 2010). Moreover, in silico screening of the genome and the Expressed Sequence Tag (EST) data banks of *H. magnipapillata* (Chapman et al. 2010) have demonstrated the existence of common stem cell signatures across adult tissues of various organisms (Hemmrich and Bosch 2008), suggesting that stem cells may be identifiable through expression of certain gene subsets. However, the regulatory events and signaling pathways that control ancestral stem cells, the nature of the transcription factors, and the interlineage communication of stem cell populations remain poorly defined in any of the early branching metazoans (Hemmrich and Bosch 2008).

To uncover stem cell signatures in the bilaterian ancestor and to understand the mechanisms of communication between different stem cell lineages, we generated transgenic polyps expressing Enhanced Green Fluorescent Protein (*eGFP*) specifically in each of the three stem cell lineages. Up to now transgene technology in the genus *Hydra* is limited to *H. vulgaris* strain AEP. No transgenic lines are available yet in *H. magnipapillata*, the species with the recently sequenced genome (Chapman et al. 2010). As shown previously (Hemmrich et al. 2007), *H. vulgaris* strain AEP is genetically

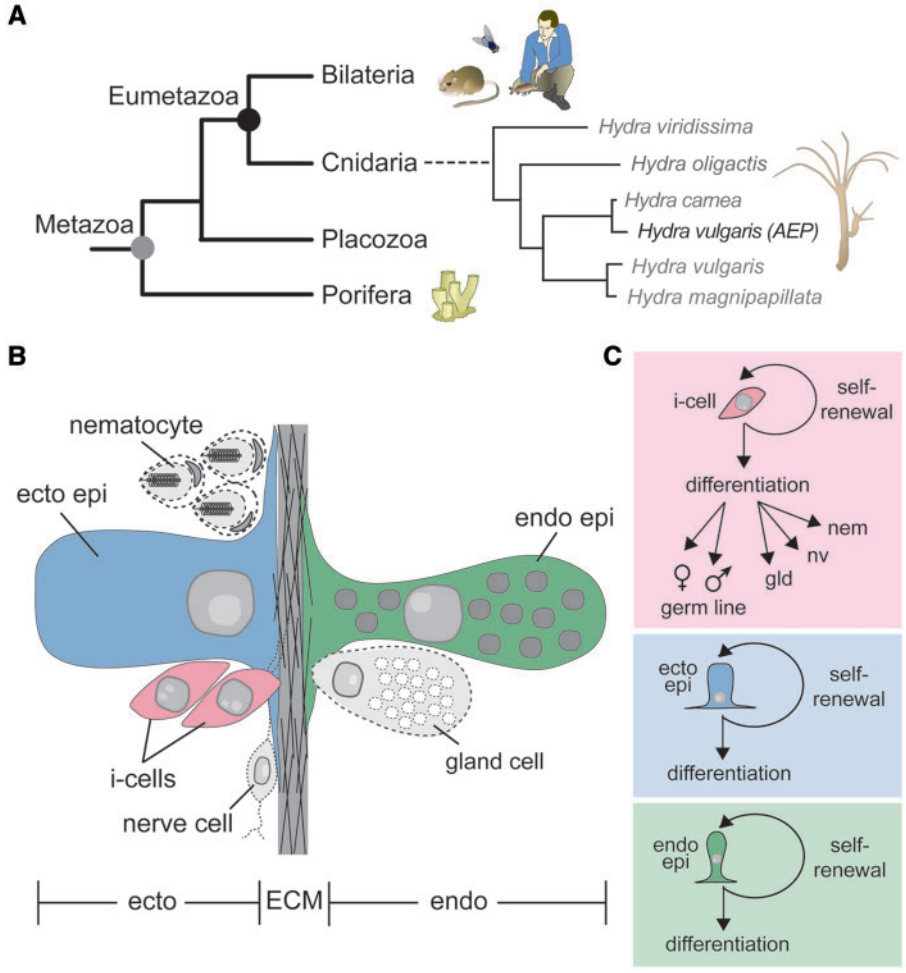

Fig. 1. The three stem cell lineages in *Hydra*. (*A*) A schematic phylogenetic tree showing the main branches in metazoan evolution. (*B*) The major cell types in *Hydra*. Stem cell lineages are colored; derivatives of the interstitial cell lineage are shown in gray. (*C*) Three stem cell systems in *Hydra*. Both epithelial cell lineages represent unipotent stem cells, whereas interstitial cells exhibit multipotent features as they are able to differentiate into various derivatives. Ecto, ectoderm; endo, endoderm; ECM, extracellular matrix; ecto epi, ectodermal epithelial cell; endo epi, endodermal epithelial cell; nv, nerve cell; gld, gland cell; nem, nematocyte. The illustrations used in (*A*) are courtesy of the Integration and Application Network (ian.umces.edu/symbols), University of Maryland Center for Environmental Science (MD).

not identical to *H. vulgaris* strain Basel but groups together with *H. carnea* (fig. 1*A*). The labeled stem cells of *H. vulgaris* strain AEP were separated by FACS, and the commonalities and differences among the three stem cell lineages were addressed by ribonucleic acid (RNA) sequencing and gene expression studies. Finally, the evolutionary age of *H. vulgaris* strain AEP stem cell genes was estimated by a phylostratigraphic analysis. Our results provide unprecedented insights of stem cells in the bilaterian ancestor and indicate that the common key signaling pathways seem to orchestrate stem cell behavior throughout the animal kingdom.

Materials and Methods

Animals and Culture Conditions

All experiments were carried out using *H. vulgaris* strain AEP. All animals were cultured according to standard procedures at 18°C (Lenhoff and Brown 1970).

Generation of Transgenic *H. vulgaris* Strain AEP Expressing *eGFP* in Each of the Three Stem Cell Lineages

Transgenic lines expressing *eGFP* in the ectodermal and endodermal epithelial cell lineages have been reported previously (Wittlieb et al. 2006; Khalturin et al. 2007, 2008). Transgenic founder polyps expressing *eGFP* under the control of the 1 kb 5′ flanking region of the *H. vulgaris* strain AEP *cnnos1* (*nanos1*) gene (GenBank XM_002161814.1) were produced at the University of Kiel Transgenic *Hydra* Facility (http://www.transgenic-hydra.org/); 921 bp of *nanos1* promoter was amplified from *H. vulgaris* strain AEP genomic deoxyribonucleic acid (gDNA) using Platinum High Fidelity Taq polymerase (Invitrogen) and cloned into the modification of HoTG expression vector in front of the reporter gene *eGFP*. The resulting transfection construct was sequenced, and plasmid deoxyribonucleic acid (DNA) was purified using Qiagen MidiPrep Kit and injected into *H. vulgaris* strain AEP embryos as described previously (Wittlieb et al. 2006). Three of the 49 injected embryos showed a positive eGFP signal in the interstitial cells after hatching. All initial founder transgenic animals were further expanded into a mass culture by clonal propagation by budding and used in further experiments.

Fluorescence-Activated Cell Sorting

For each transgenic line, two samples (500 polyps each) were treated for 1.5 h in 5 ml dissociation medium containing 250 U of Pronase E (Greber et al. 1992). The resulting cell solutions were carefully filtered through a 100-μm mesh to get rid of remaining tissue clumps. TO-PRO-3 iodide (Invitrogen, Carlsbad, NM, USA) DNA dye was used (1:1,000) to stain cell nuclei. Single cell suspensions were sorted according to eGFP and TO-PRO-3 fluorescence using the FacsAria™ cell sorting system (BD Biosciences, San Jose, CA, USA) utilizing the 100-μm nozzle. Reanalysis was performed using 100 μl of sorted cell fractions. Immediately after FACS analysis, sorted cells were centrifuged at 1,000 U/min for 5 min, and the resulting pellet was dissolved directly in messenger RNA (mRNA)-extraction buffer.

Complementary Deoxyribonucleic Acid (cDNA) Library Preparation and Transcriptome Sequencing

Polyadenylated RNA was isolated from sorted cell fractions and from whole animals using illustra™ QuickPrep Micro mRNA purification kit (GE Healthcare, Buckinghamshire, UK). Double-stranded cDNA libraries were constructed using SMART™ PCR cDNA Synthesis Kit (Clontech, Mountain View, CA, USA). After all libraries were prepared according to the manufacturer's protocol, individual samples were pyrosequenced on a Roche 454 FLX sequencer (Roche, Penzberg, Germany). Raw sequencing read data was submitted to the Sequence Read Archive at the National Center for Biotechnology Information (NCBI) (GenBank SRX019485 and SRX019488).

Bioinformatic Methods

After removal of adaptor sequences, raw reads from all sequenced libraries were assembled de novo into contigs using the Celera v5.04 assembly pipeline (Miller et al. 2008) (unitigger = bog; utgErrorRate = 0.03) followed by a merging step using Minimus2 (minid = 94%; overlap = 40; maxtrim = 20; wiggle = 16; conserr = 0.06) from the AMOS v2.08 software package (Sommer et al. 2007). The complete assembly is available for downloading and blast searches at our local bio-computational platform, http://www.compagen.org. Sequence homology of contigs from cell fractions was analyzed using a semiautomatic annotation procedure, i.e., Basic Local Alignment Search Tool (BLAST) analyses (Altschul et al. 1990) and HMMer searches (Eddy 2009) against Pfam domain database (Finn et al. 2008).

Access to Gene Sequences

Raw sequence data were submitted to the Sequence Read Archive at NCBI (GenBank SRX019485 and SRX019488). Sequences of selected candidates used for biological validation were submitted to NCBI (see later). In addition, to facilitate access and analysis of the transcriptomic data presented in this study, all sequences from the transcriptome assembly (contigs) are accessible at our local bio-computational platform, http://www.compagen.org.

GenBank accession numbers of selected candidates are as follows: *cux1* (GenBank JQ994215; contig ID 48136), *ets1* (GenBank JQ994222; contig ID 3190), *foxA2* (GenBank JQ994211; contig ID 10689), *foxK1* (GenBank JQ994216; contig ID 46845), *HMG-B3b* (GenBank JQ994217; contig ID 6666), *KLF3* (GenBank JQ994225; contig ID 11133), *KLF8* (GenBank JQ994209; contig ID 9495), *KLF11* (GenBank JQ994226; contig ID 5966), *KLF13* (GenBank JQ994218; contig ID 9445), *PRDM4* (GenBank JQ994219; contig ID 12135), *SOD* (GenBank JQ994220; contig ID 43298), *sp4* (GenBank JQ994212; contig ID 6303), *zic2* (GenBank JQ994213; contig ID 9190), *zic3* (GenBank JQ994214; contig ID 9493), *ZNF436* (GenBank JQ994210; contig IDs 6346 + 42837), *ZNF845* (GenBank JQ994221; contig ID

43371), *ZNF69* (GenBank JQ994223; contig ID 8761), and *DLL* (GenBank JQ994224; contig ID 45378).

Biological Validation for Selected Candidate Transcripts

For validation of differential gene expression and to have an experimental replicate available, an independent FACS approach was conducted. The mRNA from the sorted cell fractions was isolated as described earlier and reverse transcribed into single-stranded cDNA using the Fermentas Pure Extreme™ First-strand cDNA synthesis kit (Fermentas, St Leon-Rot, Germany). For selected candidate genes, quantitative real-time polymerase chain reaction (PCR) using the QuantiTect Probe RT-PCR Kit (QIAGEN, Hilden, Germany) and the 7300 real-time PCR system (ABI, Foster City, CA, USA) was performed according to the manufacturer's protocols (for primers, see supplementary table S6, Supplementary Material online). Whole-mount in situ hybridizations as described previously (Khalturin et al. 2007) were used to localize the spatial expression patterns. Sequences of selected candidates used for biological validation can be found at the NCBI server (see Access to Gene Sequences).

Complementation Assay

For heterologous expression of the *Hydra* homolog of *sox2*, we used *hysox1/2/3* (GenBank XM_002161342.1). This gene was selected because according to the phylogenetic analysis (supplementary figs. S1 and S2, Supplementary Material online) it is positioned at the base of *soxB1* and *soxB2* groups. The coding region of *hysox1/2/3* was cloned into pPyCAG_BstXI_IB expression vector. The *hysox1/2/3* expression construct (GenBank JQ994232) was used for transfection of inducible *msox2*-deficient mouse ES cells (mESCs), which were treated with tetracycline to induce deletion of the endogenous *msox2* before transfection (Masui et al. 2007). If *hysox1/2/3* has the potential to compliment *msox2* function, mESCs are expected to maintain self-renewal capacity and to form colonies of embryonic bodies.

Microscopic Analysis

Fluorescent images were taken on a Zeiss Axioscope fluorescence microscope with Axiocam (Zeiss, Jena, Germany) digital camera. Confocal laser microscopy was done using a TCS SP1 CLS (LEICA, Solms, Germany) microscope. Images of in situ preparations were taken on a Zeiss Axioscope microscope with Axiocam (Zeiss, Jena, Germany) digital camera.

Phylostratigraphic Analysis

We performed phylostratigraphic analysis according to the procedures described in previous studies (Domazet-Loso et al. 2007; Domazet-Loso and Tautz 2008, 2010). BLAST searches were done against the curated NR NCBI protein database (Domazet-Loso and Tautz 2010), which we additionally enriched with a set of 102,381 proteins that were predicted from assembled ESTs of several hydrozoan (*H. vulgaris* strain AEP, *H. magnipapillata*, *H. oligactis*, and *H. viridissima*) and one scyphozoan species (*Aurelia aurita*). Obtained BLAST results were used to map in total 23,691 predicted *H. vulgaris* strain AEP proteins according to the evolutionary origin of their founder genes on the currently best supported phylogeny. The final choice of internodes on the phylogeny was a trade-off between the intention to cover the most important evolutionary transitions, availability of sequenced genomes, and reliability of phylogenetic relationships (Hemmrich et al. 2007).

We calculated transcriptome age index (TAI) according to our previous work (Domazet-Loso and Tautz 2010). Only 6,313 contigs that were present in all three stem cell types were considered in the analysis. To estimate relative expression levels of these contigs in the three stem cell types, we used counts of 454-sequencing reads normalized by the total number of reads of the respective libraries. Phylogentic ranks (phylostrata) are assigned to the contigs using the phylostratigraphic map of *H. vulgaris* strain AEP proteins and contig vs. protein correspondence table. Significance of differences between TAI values is tested by one-way analysis of variance (ANOVA). To see how discrete evolutionary levels (phylostrata) contribute to the cumulative differences in the transcriptome age, we calculated, per each contig, the ratio between the expression level in a particular cell type and the median across all tree stem cell types. Significance of average differences between these ratios across the phylostrata and cell types is tested by two-way mixed model ANOVA, where the repeated measures factor was cell type and the second factor was phylostratum.

Results and Discussion

Labeling, Isolation, and Molecular Definition of the Three Stem Cell Lineages in *Hydra*

To achieve an integrated understanding of the three stem cell populations at the molecular level, we introduced *eGFP* in each of the stem cell lineages in *H. vulgaris* strain AEP (fig. 2*A* and *B*). Transgenic polyps expressing *eGFP* specifically in their endodermal and ectodermal epithelial stem cells have been described previously (Wittlieb et al. 2006; Anton-Erxleben et al. 2009). To characterize the stem cells of the multipotent interstitial stem cell lineage, transgenic polyps were generated, which express *eGFP* under the control of the interstitial cell-specific *cnnos1* (*nanos1*) promoter (fig. 2*A* and *B*). *Nanos1* was shown previously (Mochizuki et al. 2000) to be a specific marker of *Hydra*'s multipotent interstitial stem cells. Approximately 1 kb of the *nanos1* promoter is sufficient to activate strong *eGFP* expression in pairs of interstitial cells (fig. 2*C*). As soon as the eGFP-labeled interstitial stem cells start to differentiate, eGFP fluorescence is fading drastically (supplementary fig. S3, Supplementary Material online) (Mochizuki et al. 2000). For FACS analysis, transgenic polyps were dissociated into single cell solutions, and the dissociated cells were fractionated into populations based on eGFP fluorescence (fig. 2*D*, *E*, *G*, *H*, *K*, and *L*, supplementary fig. S4 and supplementary table S1, Supplementary Material online). To avoid a contamination with interstitial cell derivatives, which still contain residual amounts of eGFP (supplementary fig. S3, Supplementary Material online), we sorted only cells

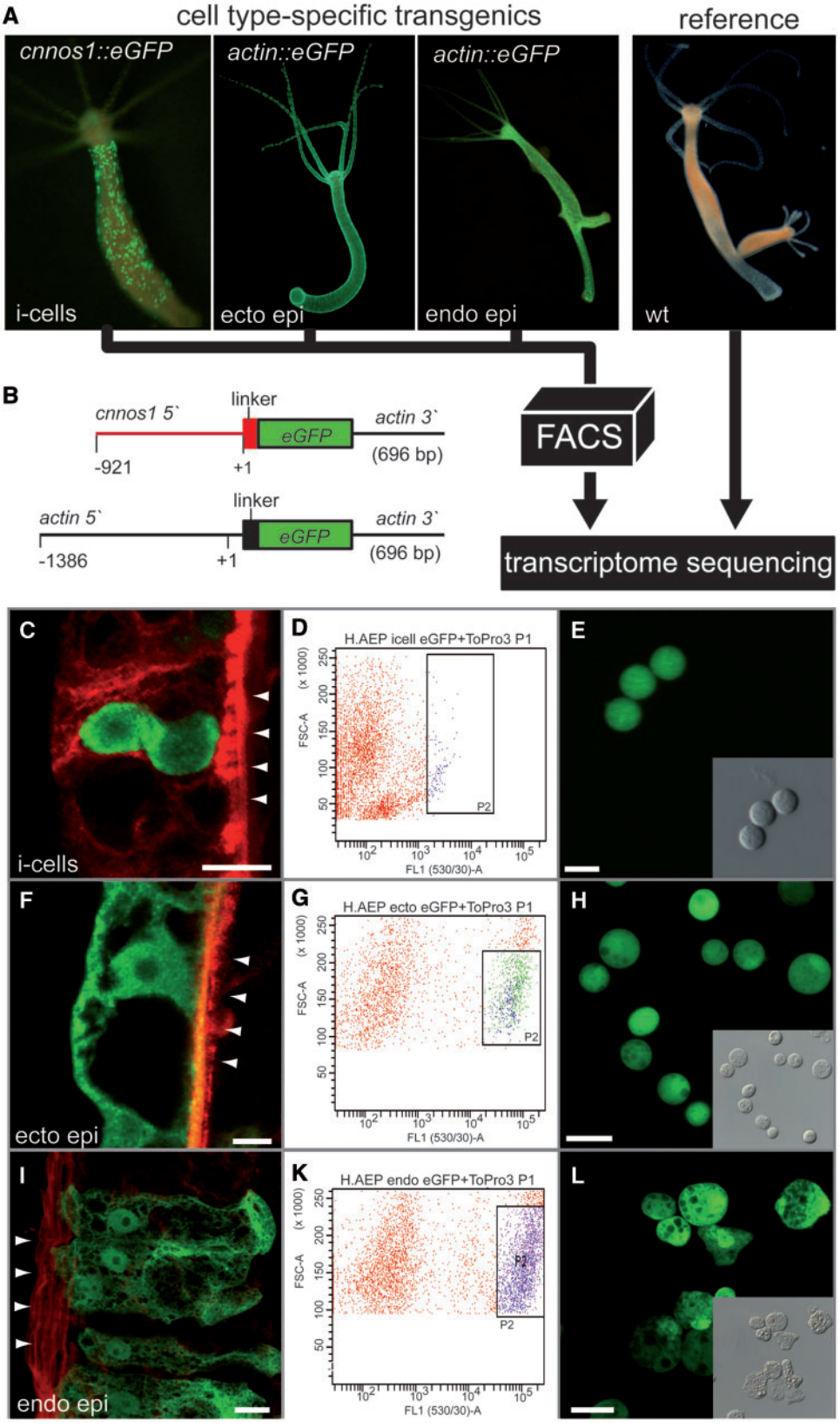

Fig. 2. Labeling and identifying the three stem cell lineages. (*A*) Transgenic polyps carrying *eGFP* in each of the three stem cell lineages; i-cell, interstitial stem cell (*cnnos1::eGFP*); ecto epi, ectodermal epithelial cell (*actin::eGFP*); endo epi, endodermal epithelial cell (*actin::eGFP*). *H. vulgaris* AEP was used as reference. (*B*) Schematic overview of the *eGFP* reporter constructs used to tag stem cells in *Hydra*. (*C, F,* and *I*) Confocal images of eGFP expressing stem cells within *Hydra* tissue. (*D, G,* and *K*) Selected cell fractions in FACS experiment. FL1, GFP fluorescence; FSC-A, forward scatter. (*E, H,* and *L*) Homogenous cell fractions after FACS sorting. Scale bars: (*C, E, F*) = 10 μm (*I*) = 20 μm (*H* and *L*) = 50 μm.

with the highest fluorescence (fig. 2*D*). To uncover stem cell-specific genes, 454 GS-FLX-based sequencing technology was used. cDNA samples were prepared from the three sorted eGFP-positive cell fractions (interstitial cells, ectodermal epithelial cells, and endodermal epithelial cells) and, for reference, from intact nontransgenic *H. vulgaris* strain AEP polyps (fig. 2*A*). An overview of the sequencing results is given in supplementary table S2, Supplementary Material online. A total of eight different libraries were sequenced resulting in 931 85 Mb (4.1 million reads) of raw sequence data, of which 819 47 Mb (3.6 million reads) remained after quality control and adaptor trimming (fig. 3*A* and supplementary table S2, Supplementary Material online). The resulting combined reference assembly of the high-quality sequencing reads produced 49,070 contigs, with 417,908 reads remaining as singletons (fig. 3*A* and supplementary table S3, Supplementary Material online). Contigs ranged from 64 to 25,631 bp in size, with an average of 566 bp and an N50 of 651 bp (i.e., 50% of the assembled bases were incorporated into contigs 651 bp or longer).

To further analyze only contigs that contain sequence information from FACS-sorted cell types exhibiting a significant differential expression when compared between the three lineages, two threshold steps were introduced. First, a minimal threshold of 10 reads per contig was applied, so that only such contigs remained for analysis that contain 10 or more sequencing reads from the above-described cell-type-sorting experiment (fig. 3*A*). Using the above thresholds, a total of 9,188 contigs could be identified (fig. 3*A*). Second, for all resulting contigs, the relative expression levels were determined to identify possible differential expression between the three stem cell lineages. This allowed subdividing cell type-specific transcripts ($\geq$2.5-fold difference in gene expression), transcripts predominantly expressed in two of the three cell types (2- to 2.5-fold difference in gene expression), and common transcripts for all three cell types ($\leq$2-fold difference in gene expression). These contigs were further used in two analytical steps: 1) annotation of sequence and predicted function to obtain a general overview of the functional properties of the three cell populations (fig. 3*B*) and 2) identification of distinct stem cell-specific signatures (fig. 3*C*).

Functional Features of the Three Stem Cell Types

To identify putative homologs of known genes and to uncover conserved protein domains, we subjected all 9,188 contigs to a BLAST and HMMer-based homology analysis. BlastX searches against the nonredundant protein database were manually curated and supplemented by screening the Pfam collection of protein domains (supplementary table S4, Supplementary Material online; www.compagen.org). As shown in figure 3*B*, this annotation procedure generally divided our sequences into the following three categories: 1) contigs with known homologs from other animals and annotated functions; 2) contigs with known homologs in other metazoan animals but with unknown function; and 3) contigs with no identifiable homology. Although 7–9% of all genes showed no blast hit, more than 50% of the sequences shared homology with proteins from other organisms. Based on the described function of the homologous proteins found, the contigs were manually classified into distinct functional classes. Obvious differences are mainly found comparing the interstitial cell lineage with the two epithelial cell lineages, whereas both ectodermal and endodermal epithelial stem cells exhibit similar functional profiles (fig. 3*B*). Interstitial cells, which are cycling about three times faster than epithelial cells (David and Campbell 1972; Campbell and David 1974), strongly express genes categorized into "cell-cycle, general transcription, translation and DNA replication" and "epigenetic regulation." Epithelial cells contain more sequences belonging to the categories "signal transduction" and "extracellular matrix (ECM), cell adhesion and cytoskeleton" (fig. 3*B*). In sum, all three stem cell types seem to require distinct sets of genes for their function. Given that each of the three stem cell populations in addition to the annotated sequences contains many novel genes, molecules regulating stem cell maintenance and differentiation might exist, which are not uncovered yet.

Distinct Stem Cell Type-Specific Signatures

To dissect the differences and commonalities between the three stem cell lineages and to deduce distinct cell type-specific signatures, we next examined the repertoire of genes that displayed 2.5-fold higher expression in one cell type compared with the others (fig. 3*C*). In interstitial cells, 2,609 genes are specifically upregulated compared with 847 in endodermal epithelial cells and 1,497 in ectodermal epithelial cells. Exclusively expressed in interstitial cells are 39 genes, 59 genes are exclusively expressed in endodermal epithelial cells, and 224 genes are exclusively expressed in ectodermal epithelial cells (fig. 3*C*). We predicted that, given the common ontogenetic origin of interstitial cells and endodermal epithelial cells from embryonic "inner cells" (Martin et al. 1997), there would be a significant commonality between these two stem cell populations. Surprisingly, interstitial stem cells share with endodermal epithelial cells only 652 sequences, whereas they share 1,974 sequences with ectodermal epithelial cells (fig. 3*C*). Ectodermal epithelial cells and endodermal epithelial cells share 664 genes. All three stem cell lineages share 914 contigs. Thus, in terms of molecular fingerprints, the interstitial stem cell transcriptome is more similar to the transcriptome of ectodermal epithelial cells. In addition to general differences caused by variable mitotic activity (see earlier), three functional categories show significant differences in expression levels: 1) extracellular matrix (ECM) components, 2) genes related to signal transduction, and 3) transcription factors. In addition, interstitial cells show an expansion of epigenetic regulators, germ line factors, and genes involved in DNA repair and stress response (fig. 3*C*).

Important mediators of tissue morphogenesis are ECM molecules (Rozario and DeSimone 2010; Tsang et al. 2010). Our data show that genes coding for molecules mediating adhesion and building the ECM are produced mainly by the epithelial cells (fig. 3*C*). Interestingly, ectodermal and endodermal cells complement each other in producing the ECM

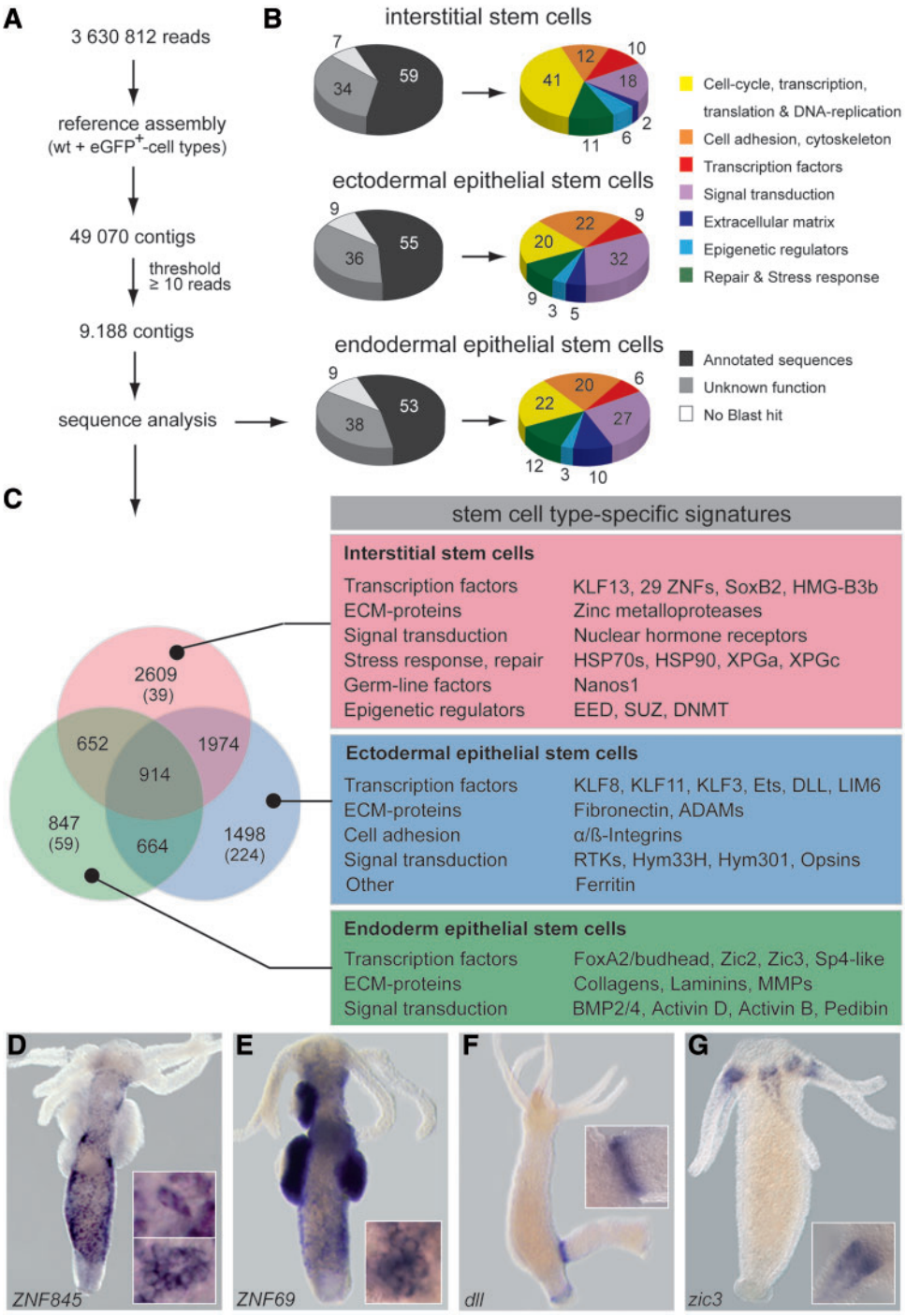

Fig. 3. Lineage-specific signatures. (*A*) Processing pipeline of the sequence data. (*B*) Functional annotation of the cell type-specific transcriptomes; pie charts illustrate assigned functional categories of transcripts found in the three transcriptomes. The annotation procedure used generally divided the sequences into the following three categories: 1) contigs with known homologs from other animals and annotated functions; 2) contigs with known homologs in other metazoan animals but with unknown function; and 3) contigs with no identifiable homology. Functionally relevant differences were mainly found comparing the interstitial cell lineage with the two epithelial cell lineages, whereas both ectodermal and endodermal epithelial stem cells exhibit similar functional profiles. (*C*) Venn diagram of stem cell genes. Genes are subdivided into cell type-specific transcripts (≥2.5-fold difference in gene expression), transcripts strongly expressed in two of the three cell types (2- to 2.5-fold difference in gene expression), and common transcripts for all three cell types (≤2-fold difference in gene expression). Numbers in brackets indicate genes exclusively expressed in one of the three cell lineages. Table on the right shows cell-type-specific signatures (genes that are >2.5-fold higher expressed in a particular cell type). (*D*–*G*) In situ expression analysis of selected signature genes. (*D*) *ZNF845* is expressed in interstitial cells along the whole body column but absent in head and foot tissue and in gametes. (*E*) *ZNF69* appears to define a subpopulation of interstitial cells committed to sperm differentiation. (*F*) A distal-less orthologous gene (*dll*) is expressed in a small ring of ectodermal cells at the base of the bud shortly before foot formation. (*G*) Endodermal epithelial cells express *zic3* at the base of the tentacles.

(mesoglea). Ectodermal cells contribute shedding proteases from the ADAM group of proteins (those that contain a disintegrin and a metalloprotease domain) and ECM glycoprotein fibronectin. Ectodermal epithelial cells are further characterized by the expression of a specific subset of α/β integrins. The endodermal cells contribute all different collagens (11 genes) and laminins and matrix metallo proteinases (MMP)-related proteases. This complementary interaction between ectodermal and endodermal epithelial cells might be the basis for epithelial homeostasis in *Hydra* and the strict maintenance of the ratio between ectodermal and endodermal epithelial cells. In contrast to epithelial cells, the interstitial stem cell lineage seems to provide no structural components of the ECM, but expresses genes needed for ECM modification such as zinc metalloproteases.

Major differences in gene expression levels were observed in transcripts related to signal transduction (fig. 3*C*). Although the general transduction machinery is present in all three cell types (supplementary table S4, Supplementary Material online), many molecules previously described to be involved in patterning the *Hydra* body plan such as Bone Morphogenetic Proteins (BMPs) (Bode et al. 2008) or Pedibin (Grens et al. 1999) are restricted to the endodermal cell lineage. This provides direct molecular support for the view (Fujisawa and Sugiyama 1978) that patterning in *Hydra* is driven by endodermal epithelial cells. The absence of these pattering molecules in interstitial cells underlines the fact that these cells are not involved in patterning processes (Gee et al. 2010). In ectodermal epithelial stem cells, many novel receptor tyrosine kinases with unknown function are found (supplementary table S4, Supplementary Material online). Consistent with previous observations (Takahashi et al. 1997, 2005; Khalturin et al. 2008), epithelio-peptides such as Hym33H or Hym301 are only found in the ectoderm. Also, ectodermal cells express genes related to *opsin* receptors. Although *Hydra* has no light sensing organ, there is photosensitivity. Uncovering *opsin* expression in *Hydra*'s ectodermal epithelial cells points to a potential role in photo transduction and underlines the multifunctionality of this cell type. Interstitial cells show an expansion in epigenetic regulators and germ line factor Nanos1 (fig. 3*C*). *Hydra* homologs of Piwi, another well-characterized germ line regulator (Bosch 2004; Seipel, Yanze and Schmid 2004), can be found in all three stem cell lineages suggesting a critical role in somatic stem-cell maintenance (Boehm AM, Hemmrich G, Khalturin K, Puchert M, Anton-Erxleben F, Wittlieb J, Klostermeier UC, Rosenstiel P, Oberg HH, Bosch TCG, unpublished data). Furthermore, a specific set of stress–response genes (members of the hsp70 and hsp90 family) and DNA damage repair genes are predominantly expressed in interstitial stem cells pointing to differences in DNA damage responses between somatic tissues and the germ line. The tissue-specific expression of selected "signature genes" was confirmed by using in situ hybridization. *ZNF845*, one of 29 zinc finger transcription factors over-represented in the interstitial stem-cell trancriptome, is expressed in interstitial cells along the whole body column (fig. 3*D*). *ZNF845*-expressing cells are not detected in head and foot tissue. In contrast, transcription factor ZNF69 seems to define a subpopulation of interstitial cells committed to sperm differentiation (fig. 3*E*). According to in situ hybridization, the distal-less orthologous gene (*dll*), one of the transcription factors exclusively discovered in the ectodermal epithelial cell fraction, is expressed in a small ring of ectodermal cells at the base of the bud shortly before foot formation (fig. 3*F*). Endodermal epithelial cells express *zic3* (fig. 3*G*) at the base of the tentacles suggesting that *zic3* plays a role in tentacle formation.

Transcriptome-Wide Expression Analysis Points to Functional Cross Talk Between Cells of the Three Stem-Cell Lineages

To elucidate the network of signaling pathways, which allows the three stem-cell lineages to coordinate growth rates and to maintain tissue homeostasis, we next used the three transcriptomes for discovering signaling molecules specifically expressed in one of the three stem-cell lineages. Analysis of conserved receptor–ligand combinations from signal transduction pathways showed a clear separation of receptors being expressed in all three stem-cell populations, whereas the corresponding ligands can be found mainly in one of the epithelial cell lineages (fig. 4*A*). None of the ligands with the exception of chordin is notably expressed in interstitial cells. For example, among the many signaling events in *Hydra*, the interactions between members of the secreted Wnt protein family and their receptors (Frizzled) are prominent (Holstein 2008). Although the ligands seem to be expressed exclusively in ectodermal epithelial cells, both ectodermal epithelial cells and interstitial cells express the frizzled receptor (fig. 4*A*). Similarly, in the fibroblast growth factor (FGF) and vascular endothelial growth factor (VEGF) signaling systems, ligands are expressed in ectodermal epithelial cells and corresponding receptors (*FGFR* and *VEGFR*) in all three cell lineages. These results not only identify the *Hydra* epithelium as a signaling center, which provides information for interstitial stem cells, but also show that the evolution of a tightly integrated biological system requires the segregation of function.

Transcription Factors Controlling Stem-Cell Behavior in *Hydra*

Lineage-specific transcription factors play critical roles in defining cell type-specific gene expression patterns; therefore, we explored in more detail those transcription factors that play key roles in stem-cell decision making in bilaterians (supplementary table S5, Supplementary Material online). Figure 4*B* shows that each cell lineage expresses a distinct set of transcription factor genes. The interstitial stem-cell population contained the largest number of transcription factors of the three cell lines, characterized by a largely expanded group of 29 zinc finger (*ZNF*) genes. Uncovering HMG-B3b, a member of the high-mobility group (HMG) super family of proteins, as overrepresented in the interstitial stem-cell transcriptome was interesting in light of the fact that HMG-B3b in mouse hematopoietic stem cells is thought to regulate the balance between self-renewal and differentiation (Nemeth et al. 2003). KLF13 (Krüppel-like factor 13), a

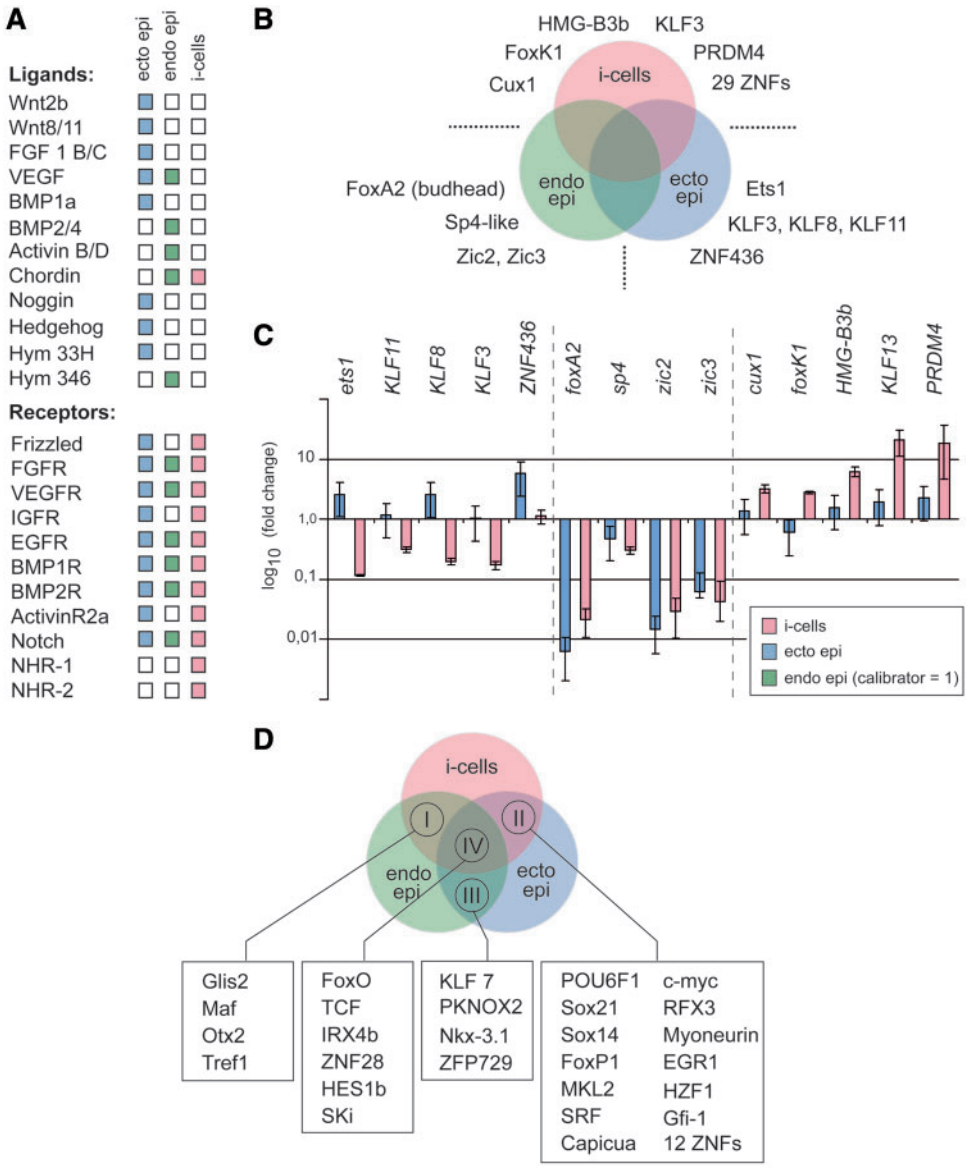

Fig. 4. Lineage-specific signatures and commonly expressed genes. (*A*) Overview of the distribution of conserved receptor–ligand pairs in the three cell types, showing a clear separation of receptors being expressed in all three stem cell populations, whereas the corresponding ligands can be found mainly in the epithelial cell lineages. (*B*) Pie charts of cell type-specific signature genes based on orthologs to transcription factors that exhibit stem cell functions in vertebrates. (*C*) Quantitative real-time PCR confirms the differential gene expression of transcription factors shown in (*B*). (*D*) Venn diagram of conserved transcription factors being expressed in two or all three investigated cell types.

member of a family of transcription factors shown in vertebrates to be involved in hematopoietic development, is also exclusively expressed in interstitial stem cells and a *cux1* (cut homeobox 1) ortholog. Cux1 transcription factors in bilaterians are involved in the control of many cellular processes, including determination of cell identity, cell cycle progression, cell–cell communication, and cell motility (Sansregret and Nepveu 2008). Ectodermal epithelial cells seem to be characterized by distinct members of the conserved KLF (KLF3, KLF8, and KLF11) and two closely related transcription factor genes having the erythroblast transformation (Ets)-specific DNA-binding domain (fig. 4*B*). The molecular signature of endodermal epithelial cells includes budhead (FoxA2), a forkhead transcription factor involved in axis formation (Martinez et al. 1997; Siebert et al. 2005). Because in the mouse FoxA2 regulates a molecular program that induces an endodermal epithelial cellular phenotype (Burtscher and Lickert 2009), identifying a *foxA2* gene as part of the molecular signature of endodermal epithelial cells in *Hydra* further establishes that endoderm specification is one of the ancestral roles of this gene. Other endodermal markers include transcription factor homologs of Zic2 and Zic3 (C2H2 zinc-finger) (Aruga 2004; Aruga et al. 2006). Zic3 plays an important role in the maintenance of pluripotency by preventing endodermal lineage specification (Lim et al. 2007).

All cell lineage-specific expression patterns in *Hydra* could be confirmed by quantitative real-time PCR (fig. 4*C*). These observations extend the current knowledge of cell layer-specific regulators beyond vertebrates and indicate that in

Hydra regulatory pathways in interstitial cells differ from those in epithelial cells.

To identify genes that may play a role in the regulation of differentiation in all three stem-cell lineages, we have searched for transcription factors strongly expressed and shared by all three stem-cell lineages (fig. 4*D*) suggesting that these may play roles in the regulation of differentiation in all three stem-cell lineages. In both endodermal epithelial cells and interstitial cells, transcripts encoding Krüppel-like zinc finger protein *glis2*, basic leucine zipper transcription factor *maf* (previously identified in the hydrozoan *Podocoryne carnea* (Seipel et al. 2004), homeoprotein *otx2* (Sugiyama et al. 2009), and transcription factor *tref1* (TReP-132) (Gizard et al. 2005) are expressed (I in fig. 4*D*). The transcriptomes of ectodermal epithelial cells and interstitial cells share a number of neuron-specific transcription factors with neurogenesis promoting activity (II in fig. 4*D*). Strong expression of these transcription factors in ectodermal epithelial cells and interstitial cells is interesting as neurogenesis in *Hydra* occurs in the ectoderm and involves differentiation of interstitial stem cells into two classes of neurons, ganglion neurons and sensory neurons. Thus, our data seem to provide compelling evidence for an evolutionary conserved function of these transcription factors in the development and/or function of neurons. Are there transcription factors defining the epithelial cell type? Endodermal epithelial cells and ectodermal epithelial cells contain a high number of transcripts for KLF7, a member of the family of C2H2 zinc finger KLFs, leucine zipper protein FosB, and the *Hydra* ortholog of C2H2 Zink Finger Protein 729 (III in fig. 4*D*). Finally, to identify genes that may play a role in the regulation of all three cell lineages, we have searched for transcription factors strongly expressed and shared by all three stem cell lineages (IV in fig. 4*D*). One transcription factor with abundant expression in all three stem cell lineages is the T-cell factor (TCF). This not only underscores the central significance of β-catenin/TCF/Lef signaling but also suggests that β-catenin stabilization might be considered an evolutionary highly conserved mechanism to alter gene expression in stem cells. Another transcription factor strongly expressed and shared by all three stem cell lineages is *foxO*. Functional analysis points to a key role of FoxO in controlling the stem cell populations in *Hydra* (Boehm AM, Hemmrich G, Khalturin K, Puchert M, Anton-Erxleben F, Wittlieb J, Klostermeier UC, Rosenstiel P, Oberg HH, Bosch TCG, unpublished data).

As revealing as the genes that were present in the three stem cell signatures were those that were absent. In higher-level organisms, the Oct3/4-Sox2-Nanog circuitry cooperates in activating transcription of key self-renewal regulators (Chen et al. 2008; Tam and Lim 2008). Consistent with a previous analysis (Hemmrich and Bosch 2008), two transcription factors known in vertebrates to be important in controlling the pluripotent state of stem cells, Nanog and Oct3/4 (Boiani and Scholer 2005; Noggle et al. 2005; Pan and Thomson 2007; de Vries et al. 2008), seem to be conspicuously absent in the *H. vulgaris* strain AEP transcriptomes and in the genomes of *H. magnipapillata* (Chapman et al. 2010), *Nematostella vectensis* (Putnam et al. 2007) and *Acropora digitifera* (Shinzato et al. 2011). Given the absence of Oct3/4 and Nanog, the presence of a *sox2*-related gene (*hysox1/2/3*) stimulated us to investigate the expected stem-cell-specific function in a complementation assay. The gene was selected because it groups at the base of *soxB1* and *soxB2* according to the phylogenetic analysis shown in supplementary figures S1 and S2, Supplementary Material online. *Hysox1/2/3*, therefore, seems to represent the ancestral form of the *soxB1* and *soxB2* gene family. Transfection of *hysox1/2/3* into mESCs where *msox2* activity is under the control of tetracycline (Tc) (see Materials and Methods section for details) revealed that the number of self-renewing primary stem cell colonies obtained in the colony assay in the presence of Tc (i.e., in the absence of *msox2*) is comparable with that of mock transfectants and significantly fewer than that of *msox2* or *msox3* transfectants (fig. 5). These observations suggest that *hysox1/2/3* is unable to substitute the function of *msox2* in mESCs. Taken together, these data suggest that the vertebrate-specific core transcription factor network (Oct3/4, Nanog, and Sox2) seems to be a

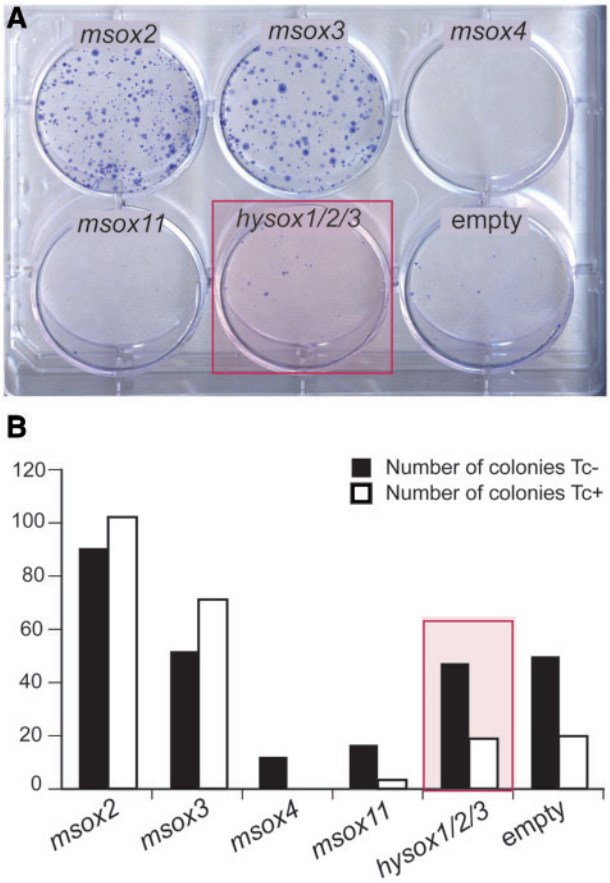

Fig. 5. Complementation assay with *hysox1/2/3*. (*A and B*) *Hysox1/2/3*-transfected mouse ES cells (mESCs) where the activity of *msox2* was under the control of tetracycline (Tc) (highlighted in red). *Msox2*-tansfected mESCs served as positive control, and mESCs transfected with various mouse *sox* genes and an empty vector were used as negative controls. For *hysox1/2/3*, the number of primary colonies obtained in the colony assay in presence of Tc (i.e., in the absence of *msox2*) is comparable with that of mock transfectants and significantly fewer than that of *msox2* or *msox3* transfectants indicating that *hysox1/2/3* is unable to substitute the function of *msox2* in mESCs.

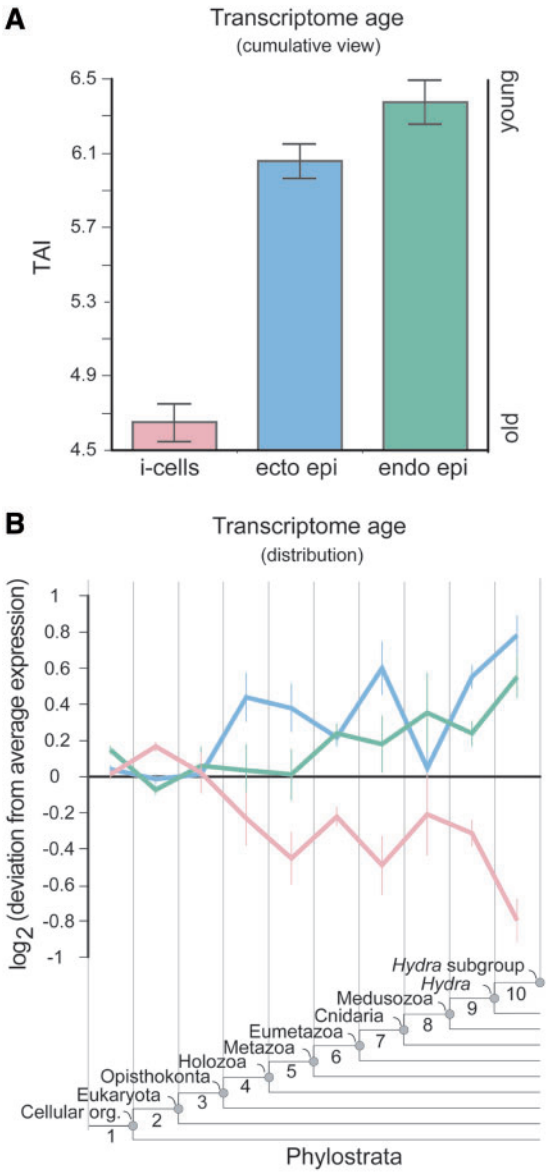

Fig. 6. Phylostratigraphic analysis of *Hydra vulgaris* strain AEP stem cells transcriptome. Using framework of 10 phylogenetic levels (phylostrata), we determined phylogenetic age of 6,313 contigs that were present in all three stem cell types (interstitial, red; ectodermal, blue; and endodermal, green). (*A*) TAI values. The ectodermal and endodermal cells express phylogenetically younger transcriptome compared with the interstitial cells (one-way ANOVA $P = 0$). (*B*) Differences in the expression levels between the three stem cell types across the phylostrata. The ratio between the expression level in a particular cell type and the median across all tree stem cell types is calculated for every contig. The three lines show average values of these ratios across the phylostrata. Error bars represent 1 standard error of mean. Obtained trend suggests that interstitial cells have increasingly lower expression of phylogenetically younger genes compared with ectodermal and endodermal cells. Two-way mixed model ANOVA, where first factor was cell type (repeated measures factor, $P = 1.5 \times 10^{-20}$) and second was phylostratum ($P = 1.2 \times 10^{-4}$, interaction $P = 1.7 \times 10^{-44}$), shows that this trend is highly significant.

later invention in bilaterian evolution and that stem cells in *Hydra* share some but not all the components of the molecular signature of stem cells within vertebrates.

How Old Is the Transcriptome of the Stem Cells? Reconstructing the Past by a Phylostratigraphic Approach

To trace the evolutionary origin and to reconstruct the age of the transcriptome of the three stem-cell lineages, we used the phylostratigraphic approach and the TAI (Domazet-Loso et al. 2007; Domazet-Loso and Tautz 2008, 2010). A framework of 10 phylogenetic levels (phylostrata) was used to assess the phylogenetic age of 6,313 proteins predicted from contigs, which were present in all three stem cell types. Figure 6*A* shows that the phylogenetically oldest transcriptome is expressed in the interstitial cells (i.e., they show the lowest TAI). In contrast, endodermal cells are expressing the youngest transcriptome, whereas the ectodermal cells are in between these extremes but closer to the endodermal cells (fig. 6*A*, one-way ANOVA $P = 0$). To further explore how discrete evolutionary levels (phylostrata) contribute to these cumulative differences in the transcriptome age, we plotted average deviation of the expression levels against 10 phylostrata (fig. 6*B*). The obtained distribution shows that expression levels in interstitial cells is decaying quasilinearly toward evolutionary younger phylostrata. This is completely opposite to the epithelial stem cells that show a reversed pattern where expression levels are increasing in the direction of the younger phylostrata (two-way mixed model ANOVA: cell-type $P = 1.5 \times 10^{-20}$; phylostratum $P = 1.2 \times 10^{-4}$; and interaction $P = 1.7 \times 10^{-44}$). Although these results could not be used to precisely pinpoint the phylogenetic origin of the three stem-cell types, they nevertheless suggest the relative order of their evolutionary emergence with interstitial stem cells preceding the origin of epithelial stem cells. This idea agrees well with our finding that the interstitial stem-cell population contains the largest number of key patterning genes that are in majority of cases conserved and widely distributed across metazoans.

Conclusion

Our observations provided new and comprehensive insight into the complex network that orchestrates patterning and tissue homeostasis in an evolutionary old animal that branched off almost 600 million years ago. The significance of our study is 3-fold. First, our work represents the first detailed characterization of the stem-cell transcriptomes in an animal at the base of evolution revealing that stem cells in the metazoan ancestors were multifunctional. Second, ancestral stem-cell populations bear a defined molecular signature composed of distinct sets of transcription factors, signal transducers, and effector genes. Third, homeostasis between the different stem cell populations in *Hydra* is maintained by cellular interactions in the form of secreted molecules produced mostly by epithelial cells and membrane bound receptors present in all three stem cell lineages.

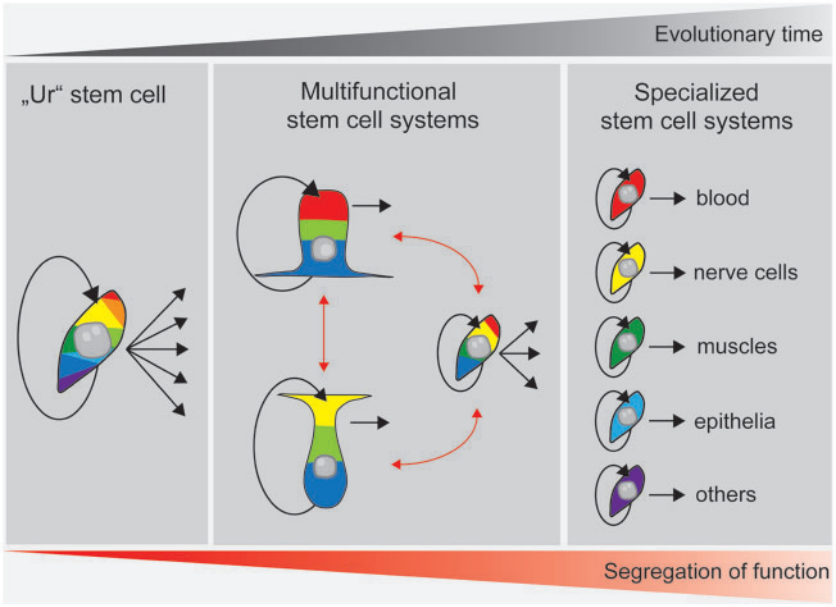

Fig. 7. Scenario for the evolution of metazoan complexity based on the diversification of stem cell systems. “Stemness” evolved in a single-celled ancestor. In unicellular organisms, the cells must have had both the ability to self-renew and to carry out differentiated functions. It, therefore, should be considered a stem cell. In simple multicellular animals such as *Hydra* with relatively few cell types, simple epithelial cells can carry out several steady-state physiological functions and serving as stem cells (Bode 1996). In complex organisms, there is appearance of numerous populations of multipotent, lifetime self-renewing stem cells with restricted developmental potential, each used for a specific task necessary for the organism’s survival.

On the basis of our detailed comparisons of transcriptional profiles from pure populations of the three stem-cell lineages in *Hydra*, we propose a model (fig. 7) in which complexity was achieved through a stepwise segregation of function. Rather than acquiring new functions, ancestral multifunctional cells, which simultaneously carried out a number of different functions, became specialized into distinct and diverse cell types, each with a limited number or even with only one specific function. Taken together, the existing comparative molecular data are most consistent with the view (Arendt et al. 2009) that the evolution of metazoan complexity relied more on cell-type functional segregation and less on the acquisition of entirely new cellular functions. The identification of molecular signatures of stem cells in *Hydra* allows for the first time the opportunity to reconstruct certain aspects of ancestral stem cells and to define molecular stem cell signatures in the bilaterian ancestor. Molecular signatures remain as clues that require detailed evaluation by functional experiments. Future research relies on uncovering the principles that determine the function of ancestral stem cells and the extent to which such functions might apply to more complex cellular systems. The excitement of future research relies on uncovering the common rules and principles that determine the function of ancestral stem cells and the extent to which such laws might apply to more complex cellular systems.

Supplementary Material

Supplementary tables S1–S6 and figures S1–S4 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank the members of the Bosch laboratory for discussion and René Augustin and David Miller for valuable comments on the manuscript. They are grateful to Santiago Insua for the development and maintenance of the transcriptome database at the Compagen server. They also thank two anonymous referees for their positive comments and largely constructive criticism of the manuscript. This work was supported, in part, by grants from the Deutsche Forschungsgemeinschaft (DFG SFB 617-A1) and DFG Cluster of Excellence programs “The Future Ocean” and “Inflammation at Interfaces” (to T.C.G.B.).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403–410.

Anton-Erxleben F, Thomas A, Wittlieb J, Fraune S, Bosch TC. 2009. Plasticity of epithelial cell shape in response to upstream signals: a whole-organism study using transgenic *Hydra*. *Zoology* 112: 185–194.

Arendt D. 2008. The evolution of cell types in animals: emerging principles from molecular studies. *Nat Rev Genet.* 9:868–882.

Arendt D, Hausen H, Purschke G. 2009. The ‘division of labour’ model of eye evolution. *Philos Trans R Soc Lond B Biol Sci.* 364: 2809–2817.

Aruga J. 2004. The role of Zic genes in neural development. *Mol Cell Neurosci.* 26:205–221.

Aruga J, Kamiya A, Takahashi H, et al. (15 co-authors). 2006. A wide-range phylogenetic analysis of Zic proteins: implications for correlations between protein structure conservation and body plan complexity. *Genomics* 87:783–792.

Bode H, Lengfeld T, Hobmayer B, Holstein TW. 2008. Detection of expression patterns in *Hydra* pattern formation. *Methods Mol Biol.* 469:69–84.

Bode HR. 1996. The interstitial cell lineage of *Hydra*: a stem cell system that arose early in evolution. *J Cell Sci.* 109(Pt 6):1155–1164.

Boiani M, Scholer HR. 2005. Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol.* 6:872–884.

Bosch TC. 2004. Control of asymmetric cell divisions: will cnidarians provide an answer? *Bioessays* 26:929–931.

Bosch TC. 2009. *Hydra* and the evolution of stem cells. *Bioessays* 31: 478–486.

Bosch TC, Anton-Erxleben F, Hemmrich G, Khalturin K. 2010. The *Hydra* polyp: nothing but an active stem cell community. *Dev Growth Differ.* 52:15–25.

Bosch TC, David CN. 1986. Male and female stem cells and sex reversal in *Hydra* polyps. *Proc Natl Acad Sci U S A.* 83:9478–9482.

Burtscher I, Lickert H. 2009. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* 136:1029–1038.

Campbell RD, David CN. 1974. Cell cycle kinetics and development of *Hydra* attenuata. II. Interstitial cells. *J Cell Sci.* 16:349–358.

Chapman JA, Kirkness EF, Simakov O, et al. (74 co-authors). 2010. The dynamic genome of *Hydra*. *Nature* 464:592–596.

Chen X, Xu H, Yuan P, et al. (23 co-authors). 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133:1106–1117.

David CN, Campbell RD. 1972. Cell cycle kinetics and development of *Hydra* attenuata. I. Epithelial cells. *J Cell Sci.* 11:557–568.

de Vries WN, Evsikov AV, Brogan LJ, Anderson CP, Graber JH, Knowles BB, Solter D. 2008. Reprogramming and differentiation in mammals: motifs and mechanisms. *Cold Spring Harb Symp Quant Biol.* 73: 33–38.

Domazet-Loso T, Brajkovic J, Tautz D. 2007. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends Genet.* 23:533–539.

Domazet-Loso T, Tautz D. 2008. An ancient evolutionary origin of genes associated with human genetic diseases. *Mol Biol Evol.* 25:2699–2707.

Domazet-Loso T, Tautz D. 2010. A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature* 468: 815–818.

Dunn CW, Hejnol A, Matus DQ, et al. (18 co-authors). 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452:745–749.

Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference. *Genome Inform.* 23:205–211.

Finn RD, Tate J, Mistry J, et al. (11 co-authors). 2008. The Pfam protein families database. *Nucleic Acids Res.* 36:D281–D288.

Fujisawa T, Sugiyama T. 1978. Genetic analysis of developmental mechanisms in *Hydra*. IV. Characterization of a nematocyst-deficient strain. *J Cell Sci.* 30:175–185.

Funayama N. 2010. The stem cell system in demosponges: insights into the origin of somatic stem cells. *Dev Growth Differ.* 52:1–14.

Gee L, Hartig J, Law L, Wittlieb J, Khalturin K, Bosch TC, Bode HR. 2010. Beta-catenin plays a central role in setting up the head organizer in *Hydra*. *Dev Biol.* 340:116–124.

Gizard F, Robillard R, Barbier O, Quatannens B, Faucompre A, Revillion F, Peyrat JP, Staels B, Hum DW. 2005. TReP-132 controls cell proliferation by regulating the expression of the cyclin-dependent kinase inhibitors p21WAF1/Cip1 and p27Kip1. *Mol Cell Biol.* 25:4335–4348.

Greber MJ, David CN, Holstein TW. 1992. A quantitative method for separation of living *Hydra* cells. *Roux's Arch Develop Biol.* 201: 296–300.

Grens A, Shimizu H, Hoffmeister SA, Bode HR, Fujisawa T. 1999. The novel signal peptides, pedibin and Hym-346, lower positional value thereby enhancing foot formation in *Hydra*. *Development* 126: 517–524.

Hartl M, Mitterstiller AM, Valovka T, Breuker K, Hobmayer B, Bister K. 2010. Stem cell-specific activation of an ancestral myc protooncogene with conserved basic functions in the early metazoan *Hydra*. *Proc Natl Acad Sci U S A.* 107:4051–4056.

Hemmrich G, Anokhin B, Zacharias H, Bosch TC. 2007. Molecular phylogenetics in *Hydra*, a classical model in evolutionary developmental biology. *Mol Phylogenet Evol.* 44:281–290.

Hemmrich G, Bosch TC. 2008. Compagen, a comparative genomics platform for early branching metazoan animals, reveals early origins of genes regulating stem-cell differentiation. *Bioessays* 30: 1010–1018.

Holstein TW. 2008. Wnt signaling in cnidarians. *Methods Mol Biol.* 469: 47–54.

Kasbauer T, Towb P, Alexandrova O, David CN, Dall'armi E, Staudigl A, Stiening B, Bottger A. 2007. The Notch signaling pathway in the cnidarian *Hydra*. *Dev Biol.* 303:376–390.

Khalturin K, Anton-Erxleben F, Milde S, Plotz C, Wittlieb J, Hemmrich G, Bosch TC. 2007. Transgenic stem cells in *Hydra* reveal an early evolutionary origin for key elements controlling self-renewal and differentiation. *Dev Biol.* 309:32–44.

Khalturin K, Anton-Erxleben F, Sassmann S, Wittlieb J, Hemmrich G, Bosch TC. 2008. A novel gene family controls species-specific morphological traits in *Hydra*. *PLoS Biol.* 6:e278.

King N. 2004. The unicellular ancestry of animal development. *Dev Cell.* 7:313–325.

Kortschak RD, Samuel G, Saint R, Miller DJ. 2003. EST analysis of the cnidarian *Acropora millepora* reveals extensive gene loss and rapid sequence divergence in the model invertebrates. *Curr Biol.* 13: 2190–2195.

Kusserow A, Pang K, Sturm C, et al. (11 co-authors). 2005. Unexpected complexity of the Wnt gene family in a sea anemone. *Nature* 433: 156–160.

Lenhoff HM, Brown RD. 1970. Mass culture of *Hydra*: an improved method and its application to other aquatic invertebrates. *Lab Anim.* 4:139–154.

Lim LS, Loh YH, Zhang W, Li Y, Chen X, Wang Y, Bakre M, Ng HH, Stanton LW. 2007. Zic3 is required for maintenance of pluripotency in embryonic stem cells. *Mol Biol Cell.* 18:1348–1358.

Martin VJ, Littlefield CL, Archer WE, Bode HR. 1997. Embryogenesis in *Hydra*. *Biol Bull.* 192:345–363.

Martinez DE, Dirksen ML, Bode PM, Jamrich M, Steele RE, Bode HR. 1997. Budhead, a fork head/HNF-3 homologue, is expressed during axis formation and head specification in *Hydra*. *Dev Biol.* 192: 523–536.

Masui S, Nakatake Y, Toyooka Y, et al. (12 co-authors). 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol.* 9:625–635.

Miller DJ, Ball EE, Technau U. 2005. Cnidarians and ancestral genetic complexity in the animal kingdom. *Trends Genet.* 21:536–539.

Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G. 2008. Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics* 24:2818–2824.

Mochizuki K, Sano H, Kobayashi S, Nishimiya-Fujisawa C, Fujisawa T. 2000. Expression and evolutionary conservation of nanos-related genes in *Hydra*. *Dev Genes Evol.* 210:591–602.

Nemeth MJ, Curtis DJ, Kirby MR, Garrett-Beal LJ, Seidel NE, Cline AP, Bodine DM. 2003. Hmgb3: an HMG-box family member expressed in primitive hematopoietic cells that inhibits myeloid and B-cell differentiation. *Blood* 102:1298–1306.

Noggle SA, James D, Brivanlou AH. 2005. A molecular basis for human embryonic stem cell pluripotency. *Stem Cell Rev.* 1:111–118.

Pan G, Thomson JA. 2007. Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Res.* 17:42–49.

Philippe H, Derelle R, Lopez P, et al. (20 co-authors). 2009. Phylogenomics revives traditional views on deep animal relationships. *Curr Biol.* 19:706–712.

Putnam NH, Srivastava M, Hellsten U, et al. (19 co-authors). 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317:86–94.

Rozario T, DeSimone DW. 2010. The extracellular matrix in development and morphogenesis: a dynamic view. *Dev Biol.* 341: 126–140.

Sansregret L, Nepveu A. 2008. The multiple roles of CUX1: insights from mouse models and cell-based assays. *Gene* 412:84–94.

Schierwater B, Eitel M, Jakob W, Osigus HJ, Hadrys H, Dellaporta SL, Kolokotronis SO, Desalle R. 2009. Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoon" hypothesis. *PLoS Biol.* 7:e20.

Seipel K, Yanze N, Muller P, Streitwolf R, Schmid V. 2004. Basic leucine zipper transcription factors C/EBP and MafL in the hydrozoan jellyfish Podocoryne carnea. *Dev Dyn.* 230:392–402.

Seipel K, Yanze N, Schmid V. 2004. The germ line and somatic stem cell gene Cniwi in the jellyfish Podocoryne carnea. *Int J Dev Biol.* 48:1–7.

Shinzato C, Shoguchi E, Kawashima T, et al. (13 co-authors). 2011. Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 476:320–323.

Siebert S, Anton-Erxleben F, Bosch TC. 2008. Cell type complexity in the basal metazoan *Hydra* is maintained by both stem cell based mechanisms and transdifferentiation. *Dev Biol.* 313:13–24.

Siebert S, Thomsen S, Reimer MM, Bosch TC. 2005. Control of foot differentiation in *Hydra*: phylogenetic footprinting indicates interaction of head, bud and foot patterning systems. *Mech Dev.* 122: 998–1007.

Sommer DD, Delcher AL, Salzberg SL, Pop M. 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics* 8:64.

Sugiyama S, Prochiantz A, Hensch TK. 2009. From brain formation to plasticity: insights on Otx2 homeoprotein. *Dev Growth Differ.* 51: 369–377.

Takahashi T, Hatta M, Yum S, Gee L, Ohtani M, Fujisawa T, Bode HR. 2005. Hym-301, a novel peptide, regulates the number of tentacles formed in *Hydra*. *Development* 132:2225–2234.

Takahashi T, Muneoka Y, Lohmann J, et al. (13 co-authors). 1997. Systematic isolation of peptide signal molecules regulating development in *Hydra*: LWamide and PW families. *Proc Natl Acad Sci U S A.* 94:1241–1246.

Tam WL, Lim B. 2008. Genome-wide transcription factor localization and function in stem cells. In *StemBook* [Internet]. Cambridge (MA): Harvard Stem Cell Institute. Available from: http://www.stembook.org/.

Technau U, Rudd S, Maxwell P, et al. (12 co-authors). 2005. Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. *Trends Genet.* 21:633–639.

Technau U, Steele RE. 2011. Evolutionary crossroads in developmental biology: Cnidaria. *Development* 138:1447–1458.

Tsang KY, Cheung MC, Chan D, Cheah KS. 2010. The developmental roles of the extracellular matrix: beyond structure to regulation. *Cell Tissue Res.* 339:93–110.

Watanabe H, Hoang VT, Mattner R, Holstein TW. 2009. Immortality and the base of multicellular life: lessons from cnidarian stem cells. *Semin Cell Dev Biol.* 20:1114–1125.

Wittlieb J, Khalturin K, Lohmann JU, Anton-Erxleben F, Bosch TC. 2006. Transgenic *Hydra* allow in vivo tracking of individual stem cells during morphogenesis. *Proc Natl Acad Sci U S A.* 103: 6208–6211.