Single-cell analysis of *Schistosoma mansoni* identifies a conserved genetic program controlling germline stem cell fate

Pengyang Li\(^1\), Dania Nanes Sarfati\(2,6\), Yuan Xue\(^1,6\), Xi Yu\(^1\), Alexander J. Tarashansky\(^1\), Stephen R. Quake\(^{1,3,4}\) & Bo Wang\(^{1,5}\)

Schistosomes are parasitic flatworms causing one of the most prevalent infectious diseases from which millions of people are currently suffering. These parasites have high fecundity and their eggs are both the transmissible agents and the cause of the infection-associated pathology. Given its biomedical significance, the schistosome germline has been a research focus for more than a century. Nonetheless, molecular mechanisms that regulate its development are only now being understood. In particular, it is unknown what balances the fate of germline stem cells (GSCs) in producing daughter stem cells through mitotic divisions versus gametes through meiosis. Here, we perform single-cell RNA sequencing on juvenile schistosomes and capture GSCs during de novo gonadal development. We identify a genetic program that controls the proliferation and differentiation of GSCs. This program centers around *onecut*, a homeobox transcription factor, and *boule*, an mRNA binding protein. Their expressions are mutually dependent in the schistosome male germline, and knocking down either of them causes over-proliferation of GSCs and blocks germ cell differentiation. We further show that this germline-specific regulatory program is conserved in the planarian, schistosome’s free-living evolutionary cousin, but the function of *onecut* has changed during evolution to support GSC maintenance.
Schistosomes are parasitic flatworms that cause one of the most prevalent but neglected infectious diseases, schistosomiasis\(^1\). With over 250 million people infected worldwide and a further 800 million at risk of infection, schistosomiasis imposes a global socioeconomic burden comparable to that of tuberculosis, HIV/AIDS, and malaria\(^2-3\). The disease transmission requires the passage of parasites through two hosts during its life cycle, a molluscan intermediate host and a mammalian definitive host (e.g., human). This complex life cycle requires the deployment of multiple specialized body plans in order to infect and reproduce within each host. These transitions are enabled by a population of stem cells that undergo several waves of proliferation and differentiation\(^4-6\).

The schistosome life cycle begins with the parasite egg being excreted from the mammalian host into freshwater. It hatches into a free-swimming larva serving as a vehicle for approximately a dozen stem cells and transforms into a sporocyst once inside a snail host\(^5,7\). Stem cells in the sporocyst can undergo either self-renewal or enter embryogenesis, producing more sporocysts or massive numbers of infectious progeny (called cercariae)\(^8-10\). Cercariae then emerge from the snail back into the water, burrow through the skin of a mammalian host, migrate to species-specific niches in the host vasculature, and transform into juvenile parasites. This transition initiates the sexual portion of the parasite’s life cycle. Stem cells packed in cercariae proliferate to initiate the growth of juveniles and eventually build sexual reproductive organs de novo, through a developmental program that appears to be modulated by host hormonal and immune cues\(^11,12\). Sexually mature male and female worms mate, and each pair can produce hundreds of fertilized eggs daily, which are excreted to initiate the next life cycle. It is worth noting that the eggs, not the worms, cause pathology as they become excreted to initiate the next life cycle. It is worth noting that the eggs, not the worms, cause pathology as they become excreted to initiate the next life cycle. It is worth noting that the eggs, not the worms, cause pathology as they become excreted to initiate the next life cycle.

We recently conducted a single-cell transcriptomic study comparing stem cells from *Schistosoma mansoni* asexual (intramolluscan) and sexual (intramammalian) stages\(^5\). Among the stem cells that are carried by cercariae and transmitted from snail to mammal, only a subset of them commits to the germline fate. These primordial germ cells initially express a schistosome-specific factor, *eledh* (*Sm-eled*, for brevity the prefix “Sm” will be omitted from gene names in the rest of the paper wherever species origin is unambiguous), and during germline specification activate one of the schistosome homologs of *nanos* (*nanos-1*), an RNA binding protein conserved across metazoans\(^4,13-16\). Unlike most animals that segregate their germline from soma during embryonic development\(^17-19\), schistosomes specify their germ cells from a somatic cell lineage at the onset of juvenile development.

While our previous work has identified the origin of the schistosome germline, it remains unclear which molecular regulatory program(s) underlies the separation of somatic and germ cell lineages\(^5\). In contrast to their distinct cellular fates, somatic and germline stem cells share strikingly similar gene expression signatures, including genes involved in regulating the cell cycle, a set of conserved RNA binding proteins that are often associated with stem cell multipotency\(^20-22\), and transcription factors such as nuclear factor Y subunits\(^18\). Disrupting these key regulators leads to defects in stem cell maintenance and proliferation in both somatic and germline lineages\(^4,18\). The molecular distinction between somatic and germline stem cells is thought to be similarly blurry in other flatworms, such as schistosome’s free-living evolutionary cousin, the planarian\(^20,21\). Planarian germline stem cells resemble their somatic stem cells both in terms of morphology and molecular signatures, and can even contribute to somatic tissues during tissue regeneration after injury\(^20-24\).

In this study, we perform single-cell RNA sequencing (scRNAseq) on juvenile *S. mansoni* to characterize the germline stem cells (GSCs). We focus on this life-cycle stage because juvenile schistosomes contain abundant stem cells, including cells in the transition between somatic and germline fates\(^6\). We succeed in capturing the GSCs and identify their transcriptional signatures using the self-assembling manifolds (SAM) algorithm\(^25\), a method that excels in identifying subtle transcriptional differences between otherwise similar groups of cells. Through RNA interference (RNAi) mediated gene knockdown, we evaluate the function of GSC-enriched transcripts and find that a schistosome homolog of *onecut* homeobox transcription factor, *onecut-1*, is the key regulator that balances the proliferation and differentiation of GSCs (in particular, male GSCs, or spermatogonial stem cells). *onecut-1* functions through complex epistatic interactions with several other genes, including *eled*, *nanos*, and *boule*. As *nanos* and *boule* are broadly conserved germline developmental regulators\(^6,18-21,23,24\), and *onecut* expression has been detected previously in mammalian testes\(^22-25\), we posit that this gene set may represent a conserved germline specific regulatory program. Consistent with this idea, we find that *onecut* is indeed essential in the male germline of planarian *Schmidtea mediterranea*, though its specific function has changed. While *onecut* protein family has been extensively studied across animals (e.g., fly and mouse)\(^26-30\), its role in regulating germline development has been previously unknown.

Results

**scRNAseq identifies GSC-specific transcriptional signatures.** We have previously performed scRNAseq on flow-sorted G2/M phase cells to enrich for stem cells that are actively cycling\(^31,32\). However, this approach failed to capture *nanos-1* expressing GSCs. To overcome this limitation, we here sequenced cells from all tissues in juvenile schistosomes using Smart-seq\(^33,34\). We collected 10945 single-cell transcriptomes, of which 7657 passed the initial quality control (see “Methods” section). Figure 1a shows that SAM analysis divided these cells into several major clusters, but only one population expresses *ago2-1*, an argonaute homolog that is ubiquitously and specifically expressed in all schistosome stem cells\(^4,6,25\). Performing SAM analysis on *ago2-1*+ cells (1464 cells) identified nine clusters (Fig. 1b), four of which correspond to stem cell subtypes we reported earlier: *μ, δ*, *εα*, and *εβ*\(^25,32\). The five newly found populations include one that contains *nanos-1* expressing cells. The other four contain cells expressing genes often associated with differentiated tissues (Supplementary Data 1), for example, *tripomin* (a common muscle gene\(^41\)), *complexin* (*cpx*, a neural gene), *cathepsin B endopeptidase* (*cb2*, a gene expressed in intestinal and parenchymal cells\(^43,44\)), and *tetraspanin-2* (*tp2-2*) co-expressed with *Sm25* (both are known markers for schistosome epidermal lineage\(^45\)). Since these populations still express common stem cell markers (e.g., *ago2-1*, *cyclin B*, *pcna*, *polo-like kinase*, histone *h2a*, and *h2b*) and are actively dividing (Supplementary Fig. 1), they are likely tissue type-specific progenitors, though further functional validation would be required to prove their respective identities and fates.

Hierarchical clustering of the cell populations (see “Methods” section) divided them into two major groups, one grouping *μ*, *δ*, *εα*, and *εβ*-cells together with *nanos-1*+ cells, and the other containing the four putative progenitor populations (Fig. 1c). Unlike progenitor populations that express a large number of population-enriched genes, *μ*, *δ*, *εα*, *εβ*-cells and *nanos-1*+ cluster only express a small number of population-enriched genes beyond the common stem cell markers (all population-enriched genes are provided in Supplementary Data 1). Instead, this group
is separated from the progenitor group primarily based on the lack of marker gene expression (yellow boxes in Fig. 1c), consistent with the idea that multipotent stem cells repress genes that are involved in cell differentiation.

Focusing on the nanos-1+ cells, we examined the expression pattern of their enriched genes (green box in Fig. 1c). This list includes transcription factor onecut-1 (oc-1, Smp_196950, Supplementary Fig. 2), which is expressed throughout this population (Fig. 1d). We also identified two clearly distinguishable subclusters marked by boule (Smp_144860) and a nuclear receptor (nr, Smp_248100), respectively. As nr is known to be expressed in the nanos-1+ stem cells (also known as S1 cells) in primordial vitellaria, a yolk cell-producing somatic reproductive organ in female schistosomes, these observations suggest that the nanos-1+ cluster contains two transcriptionally similar but distinct cell types: S1 cells (nr+) and GSCs (boule+).
is expressed in in situ hybridization (FISH) experiments further showed that is expressed in male and female gonadal primordia. Fluorescence testes and vitellaria, their expression (Fig. 2a).

Quantification of EdU nuclei abundance, GSC number density, and the fraction of differentiated germ cells are provided in Fig. 3c-e along with other RNAi experiments for direct comparison. Quanti- cations of EdU nuclei abundance, GSC number density, and the fraction of differentiated germ cells are provided in Fig. 3c-e along with other RNAi experiments for direct comparison. d, e FISH images showing broader expression of nanos-1 (d) and eled (e) in testes after oc-1 RNAi. Arrowheads: differentiated germ cells that do not express nanos-1 or eled. n: number of samples exhibiting the reported phenotype out of the total number of samples analyzed. RNAi experiments were repeated on at least three biological replicates.

oc-1 homeobox transcription factor suppresses proliferation and promotes differentiation of male GSCs. To understand the function of the genes specifically expressed in nanos-1+ cells, we first studied the only three TFs that were identified in this population: oc-1 in both GSCs and S1 cells, nr only in S1 cells, and a homolog of irququis homeobox TF (irx, Smv_063520) in GSCs. Whole-mount in situ hybridization (WISH) confirmed their expression (Fig. 2a). oc-1 is expressed in both primordial testes and vitellaria, nr is specific to primordial vitellaria, and irx is expressed in male and female gonadal primordia. Fluorescence in situ hybridization (FISH) experiments further showed that oc-1 is expressed in nanos-1+ GSCs in testes (Supplementary Fig. 3a).

We then performed RNAi to knock down them in juvenile parasites. For these experiments, parasites were retrieved from mice 3.5 weeks post-infection and soaked in double-stranded RNA (dsRNA) continuously in vitro for 2 weeks. We focused on the male germine because female development is retarded under in vitro culture and therefore associated phenotypes are difficult to evaluate5.

While neither nr nor irx RNAi showed discernible phenotypes (Supplementary Fig. 4), oc-1 RNAi led to a significant increase in the number of GSCs and a marked reduction in the abundance of differentiated germ cells, including spermatocytes, spermatids, and sperm (Fig. 2b, c). Consistently, GSC markers including nanos-1 and eled were expressed more broadly in testes after oc-1 RNAi (Fig. 2d, e). The expansion of the GSC population caused these cells to be packed at higher density compared to control samples, though the size of the testis lobules remained mostly unchanged. The majority of GSC nuclei were also labeled by 5-Ethynyl-2’-deoxyuridine (EdU) after oc-1 RNAi, suggesting that they were active in DNA synthesis. To test whether the loss of differentiated germ cells was due to the lack of differentiation or apoptosis, we performed TUNEL staining and noticed no increase in apoptotic activity after oc-1 RNAi beyond the baseline level in control animals (Supplementary Fig. 5). Together, these results suggest that oc-1 is required for preventing GSCs from over-proliferating and for promoting the initiation of the differentiation program.
An RNAi screen identifies boule as a germline regulator that phenocopies oc-1 after knockdown. To identify the pathway through which oc-1 functions, we performed an RNAi screen to knockdown genes enriched in GSCs identified by scRNAseq, focusing on predicted RNA binding proteins, enzymes, and receptors. The goal of these experiments was to identify genes that phenocopy oc-1 upon RNAi. Out of the 39 genes knocked down (Supplementary Data 2), we observed male germline developmental defects in RNAi experiments of 6 genes (Fig. 3a). Of these, only the knockdown of boule gave rise to a similar but stronger phenotype compared to that of oc-1 RNAi (Fig. 3b–g): testis lobules contained EdU+ nanos-1+/eled+ GSCs at a higher density compared to controls and drastically reduced number of differentiated germ cells. Consistent with its phenotype, we...
noticed that boule is co-expressed with oc-1 in the male GSCs (Supplementary Fig. 3). boule is a member of the Deleted in Azoosperma (DAZ) family of RNA-binding proteins, which are known germline regulators with important roles in both GSC maintenance and gametogenesis. To investigate whether other genes associated with DAZ function in other animals is also involved in schistosome germline development, we searched for potential DAZ-interacting partners, DAZAP and DZIP2,49,50 in the S. mansoni genome based on BLAST similarity. We only found one DZIP homolog, RNAi of which did not generate a germline phenotype (Supplementary Fig. 6), indicating that boule function in the schistosome is either independent of DZIP or involves other unidentified redundant interacting proteins.

The other five genes identified in this screen affected germline differentiation but not proliferation. RNAi of a putative ankyrin repeat-containing protein kinase (akkn, Smp_131630) and an LSM14 domain-containing protein (lsml4, Smp_129960) blocked the differentiation of germ cells. RNAi of a zinc finger CCHC domain-containing protein (zc3h3l, Smp_056280) and an E3 ubiquitin-protein ligase (ccnb1ip1, Smp_344490) caused defects in spermatogenesis. RNAi of a homolog of urochrome adherence factor A (uaf, Smp_105780) resulted in premature accumulation of sperm and abnormal GSC nuclear morphology (Fig. 3b, c).

Finally, to examine whether these genes play similar functional roles in homeostasis beyond development, we performed RNAi experiments on sexually mature parasites, retrieved from mice 5.5 weeks post-infection (Fig. 4a). Figure 4b, c shows that most phenotypes, including those induced by oc-1 and boule knockdowns, were identical between juveniles and adults, except for akkn, which did not yield a loss-of-function phenotype in adult parasites (Supplementary Fig. 7), suggesting that its function is restricted to regulating development.

**oc-1 and boule expressions are mutually dependent and exhibit similar genetic interactions with nanos and eled.** The observation that RNAi of oc-1 and boule led to similar phenotypes motivated us to investigate the nature of their genetic relation. We followed up RNAi of either gene with WISH to evaluate the changes in the expression of both. We found that boule expression was undetectable after oc-1 RNAi, and oc-1 expression was also eliminated upon boule RNAi (Fig. 5a). This result suggests that oc-1 and boule regulate each other (directly or indirectly) at the transcriptional level.

If oc-1 and boule overlap through the same functional pathway, we expect them to exhibit identical epistatic interactions with other GSC regulators. Previously, we found that nanos-1 is required for GSC proliferation and differentiation, knockdown of which causes degeneration of testes and loss of differentiated germ cells. In contrast, eled is thought to inhibit GSC differentiation, RNAi of which ectopically activates spermatogenesis and leads to premature accumulation of sperm. As their RNAi phenotypes are distinct from those resulting from oc-1 and boule RNAi, it is feasible to test genetic interactions among this set of genes through double RNAi knockdowns. In order to rule out the possibility that treating worms with multiple dsRNA sequences might reduce the knockdown efficacy due to their competition for the RNAi machinery, we also validated the efficient gene knockdown in double RNAi experiments (Supplementary Fig. 8).

Using this approach, we found that both nanos-1;oc-1 and nanos-1;boule double knockdowns exhibited the phenotype of nanos-1 single knockdown (Fig. 5b), suggesting that nanos-1 likely functions upstream of oc-1 and boule as a genetic suppressor. By contrast, eled;oc-1 or eled;boule double knockdowns alleviated the defects observed in oc-1 or boule single knockdowns. In particular, the number of EdU+ germ nuclei was significantly reduced after knocking down eled along with oc-1 or boule (Fig. 5b, c). These observations indicate that oc-1, boule and eled collectively maintain the balance between differentiation and proliferation of GSCs, with oc-1 and boule suppressing proliferation and promoting differentiation versus eled suppressing differentiation (Fig. 5d). When the functions of both sides were lost after double knockdowns, the balance was reestablished.

**Planarian homolog of onecut is also required for male germline development.** To determine whether the function of onecut in germline development is evolutionarily conserved, we studied the planarian flatworm S. mediterranea, a free-living cousin of schistosomes. Using BLAST sequence similarity search, we identified three planarian onecut homologs (Supplementary Fig. 2), but only one (Smed-oc-1, dS_msm1_v1_39638_1_1) was detected in the male germline (Supplementary Fig. 9a). Specifically, it is expressed in testis lobules, which are distributed beneath the dorsal epidermis of sexually mature planarians; in sexually immature planarians, Smed-oc-1 expression is specific to testis primordia, which are essentially GSC clusters (Fig. 6a, b). High-resolution images of individual testis lobules showed that Smed-oc-1 expression is restricted to GSCs (spermatogonial stem cells) and mitotic spermatogenesis, but not detected in differentiated germ cells (Fig. 6c). Consistently, in asexual planarians, Smed-oc-1 is also present in presumptive GSCs (Supplementary Fig. 9b).
Knockdown of Smed-oc-1 led to the complete loss of testes (Fig. 6d). Even in early stages of RNAi, the erosion of the spermatogonial layer had already begun, while differentiated germ cells were largely unaffected (Fig. 6e), suggesting that Smed-oc-1 is required for the maintenance of GSCs and spermatogonia. This function is distinct from that of Smed-oc-1, which is dispensable for maintaining the identity of GSCs but necessary for regulating GSC proliferation and differentiation.

We next evaluated the genetic relations between Smed-oc-1 and other known planarian germline regulators such as nanos and boule. The planarian nanos homolog, Smed-nanos, is a GSC marker and required for GSC specification and maintenance. The planarian boule homologs, Smed-boule-1 and Smed-boule-2, are expressed in GSCs and spermatogonia. Smed-boule-1 is involved in spermatogenesis and loss of its function leads to the elimination of differentiated germ cells, whereas Smed-boule-2 is necessary for the maintenance of GSCs and spermatogonia.

Fig. 5 Epistatic interactions of oc-1, boule, nanos-1, and eled. a WISH images showing that the expression of oc-1 and boule are mutually dependent. Knockdown of either one causes the elimination of both. Asterisks: testes. b Confocal images showing representative individual testis lobules stained by DAPI and EdU in juvenile parasites after RNAi treatments. Dashed circles: testis lobule boundary. n: number of samples exhibiting the reported phenotype out of the total number of samples analyzed. c Fraction of EdU+ nuclei in testes in control and RNAi parasites. Each data point represents the average across all testis lobules in a single parasite. N = 10 (control RNAi); N = 10 (eled RNAi); N = 10 (oc-1 RNAi); N = 11 (eled;oc-1 RNAi); N = 10 (boule RNAi); N = 10 (eled;boule RNAi). nanos-1 RNAi experiments were not quantified because only a few GSC nuclei are present per testis lobule, which is insufficient for reliable statistics. Data are plotted as average ± standard deviation and p-values are calculated using two-sided Welch’s t-test. RNAi experiments were repeated on at least three biological replicates. d Proposed model of genetic interactions between oc-1, boule, nanos-1, and eled in regulating male GSC identity, proliferation, and differentiation.
Fig. 6 Smed-oc-1 is required for the maintenance of spermatogonial cells. 

a Schematic of planarian testes lobules containing spermatogonial cells (GSCs and mitotic spermatogonia) at the periphery and differentiated germ cells (spermatocytes, round spermatids, elongating spermatids, and sperm) distributed in the internal layers. 

b WISH images showing the expression of Smed-oc-1 in sexually immature (left) and mature (right) planarians. The punctuated signal in the immature planarian corresponds to presumptive testis primordia. In the mature planarian, the signal is specific to testes. 

c FISH images of Smed-oc-1 showing its expression in testis primordia within sexually immature planarians (top) and in spermatogonial cells (including GSCs and mitotic spermatogonia) in sexually mature planarians (bottom). N: number of samples showing similar results over three independent experiments. 

d Dorsal view of sexually mature planarians after control and Smed-oc-1 RNAi. Note that testes, which are distributed beneath the dorsal epithelium and visualized by strong DAPI signal, are completely eliminated after Smed-oc-1 RNAi. Animals were fixed after eight feedings of dsRNA. 

e Tests of control (left) and Smed-oc-1 (right) RNAi planarians after three feedings of dsRNA. In control animals, Smed-germinal histone H4 (gH4) labels the spermatogonial cells at the periphery of testis lobules, which are mostly absent after Smed-oc-1 RNAi, while differentiated germ cells remain at the center of the testis lobules. gH4 expression in individually isolated somatic stem cells maintains after Smed-oc-1 RNAi. 

f WISH images showing that the expression of Smed-oc-1, Smed-nanos, and Smed-boule-2 are disrupted after three feedings of Smed-oc-1 RNAi. The imaged areas correspond to the dashed box in the schematic on the left. n: number of samples exhibiting the reported phenotype out of the total number of samples analyzed. RNAi experiments were repeated on at least three biological replicates, each containing a group of planarians fed with dsRNA separately.
(Supplementary Fig. 9c)\textsuperscript{28}. We found that, Smed-oc-1 RNAi quickly abolished Smed-nanos and Smed-boule-2 expression (Fig. 6f). These results suggest that Smed-oc-1 function overlaps with that of Smed-nanos and Smed-boule-2, but not Smed-boule-1.

Based on these observations, we concluded that the expression of onecut, boule, and nanos in the male germline appears to be conserved between planarian and schistosome, but the functional relationships within this gene set have been rewired significantly through evolution.

Discussion

Schistosome germline is essential for both disease transmission and pathology, as it produces eggs that serve as the transmissible foundation by identifying the key members of this previously important question for future research. Our study lays the basis for combating this major parasitic disease.

We expect the conservation of this germline-specific regulatory program to extend beyond flatworms. boule homologs, along with other members of the DAZ protein family, are known germline regulators in diverse animals from basal invertebrates to humans\textsuperscript{29–31}, though their premeiotic functions in GSCs as reported here are considered to be rare among invertebrates\textsuperscript{29}. By contrast, onecut homologs have been studied in the context of neural, liver, and pancreatic development using different animal models\textsuperscript{33–38}, but its function in germline development and its genetic interaction with boule were previously unknown. Of interest to our study is the observation that onecut homologs are indeed expressed in rodent and human testes\textsuperscript{32–35}, suggesting that its germline function may be more universal than previously acknowledged.

An important question raised by our results is what factors maintain the identity of GSCs. Our previous work suggests that GSCs are specified from eled\textsuperscript{+} stem cells during the early juvenile development by activating nanos-1 expression\textsuperscript{3}. In the current scRNAseq data set, we noticed that GSCs are distinguished from other stem cell populations in the presumptive multipotent group only by a small number of genes, while sharing a large set of common transcriptional signatures. The GSC-specific genes only contain two TFs (oc-1 and irx), but neither of them is required for maintaining GSC identity. Knockdown of other genes enriched in GSCs mostly caused defects in differentiation and spermatogenesis but did not affect the identity of GSCs either. These observations favor the hypothesis that GSC identity may be maintained by extrinsic cues, though regulation by other intrinsic determinants such as post-transcriptional regulatory factors cannot be excluded.

In mammalian testes, these extrinsic cues are provided by discrete somatic niche cells (e.g., Sertoli cells) to support male GSC\textsuperscript{36,57}; in Drosophila and C. elegans, stromal cap and distal tip cells form the niche, respectively, instructing GSC fates through several signaling pathways\textsuperscript{58,59}. In the planarian, somatic gonadal cells required for male GSC specification and maintenance have also been defined\textsuperscript{60,61}. All these “niche” cells are in close spatial proximity with their GSC counterparts. However, in the schistosome testes, most GSCs are not in contact with any somatic tissues. This is particularly clear in oc-1 RNAi or boule RNAi animals where many testis lobules contain exclusively a large number of densely packed GSCs. These observations suggest that the extrinsic cues for maintaining GSC identity in schistosomes are likely systemic, potentially through neural or hormonal regulations. This hypothesis is particularly compelling because schistosome maturation is known to depend on host immune cues\textsuperscript{12,13}. The host signals need to be sensed by parasite cells at the host-parasite interface, then translated and relayed internally to instruct GSC activities\textsuperscript{62}. In the planarian, germline development is regulated by neuroendocrine signals, such as neuropeptides\textsuperscript{63}, implicating that there may be similar crosstalk between the nervous system and the gonads in schistosomes. Here our characterization of the intrinsic regulatory program of the schistosome GSCs opens the door for exploring these important questions. Given the significance of parasite germline in disease transmission and pathology, a better understanding of the basic biology of this developmental process will help to define vulnerable points for combating this major parasitic disease.

Methods

Animals. S. mansoni juveniles and adults were retrieved from infected female Swiss Webster mice (NR-21963) by hepatic portal vein perfusion with 37 °C DMEM supplemented with 5% heat-inactivated FBS at 3.5 weeks and 5.5 weeks post-infection, respectively. Worms were cultured at 37 °C in Basch Media 169 supplemented with 1x Antibiotic-Antimycotic\textsuperscript{11}. In adherence to the Animal Welfare Act and the Public Health Service Policy on Humane Care and Use of Laboratory Animals, all experiments with and care of mice were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC) of Stanford University (protocol approval number 30366).
resulting data were added with a pseudocount of 1 and Log2-transformed. This cut-off was chosen to bisect the bimodal distribution of the number of genes detected per cell. For each cluster, marker genes were identified by computing partial sums of the PC centroids to generate the dendrogram shown in Fig. 1c. For each cluster, marker genes were identified by computing partial sums of the k-nearest-neighbor-averaged data:

\[ S_i = \sum_{j=1}^{k} \left( E_{ij} - \mu_i \right)^2 / \mu_i \]

where \( E_{ij} \) is the nearest neighbor averaged expression value for cell \( j \) and gene \( i \), \( \mu_i \) is the mean expression of gene \( i \), \( n_i \) is the set of cells in cluster \( x \), and \( S_i \) is the normalized marker weight for gene \( i \) in cluster \( x \). To account for differences across clusters, we normalized marker score by the maximum value, \( Z = S_i / \max(S_i) \). The ranked genes based on their marker scores, and then kept those with >0.5 SAM weight, >0.1 normalized marker score, and >2 fold enrichment with respect to cells outside the cluster. The complete list of marker genes is provided in Supplementary Data 1. For ease of visualization, only the top 50 markers for each stem cell cluster are shown in Fig. 1c.

**RNAi**

To synthesize dsRNA, gene fragments were amplified from cDNA using oligonucleotide primers listed in Supplementary Data 2. 3 and cloned into the vector pFC3.2 (Addgene Plasmid ID: 26356)\(^{66}\). 10–15 juvenile or 6–8 adult parasites were soaked in Bacitracin media 16970 supplemented with 20 µg/mL dsRNA against each target for 2 weeks. The media with dsRNA were refreshed daily. Each RNAi was performed with at least two biological replicates each containing at least two technical replicates to evaluate phenotypic changes. A small population of parasites exhibiting visible damage or reduced mobility via microscopic examination was excluded from downstream analyses. Potential off-target effects were ruled out by using dsRNA sequences to target non-overlapping regions of the same gene of interest (Supplementary Fig. 10). For gene knockdown in planarians, dsRNA was mixed with liver paste at a concentration of 70 ng/µl dsRNA and fed to planarians every 4 days for 3–8 feedings. In all RNAi experiments, dsRNA matching the ecdB and camK-containing insert of pFC3.2 was used as the negative control. The two-sided Welch’s t-test and p-value calculation were performed via t-test in scipy library with default parameters.

**In situ hybridization.** RNA WISH and FISH experiments were performed following the established protocol\(^{52,55}\). Briefly, schistosomes were killed in 6 M MgCl\(_2\) for 30 s, fixed in 4% formaldehyde supplemented with 0.2% Triton X-100 and 1% NP-40 for 4 h, and then dehydrated in methanol. Adult parasites were relaxed by 0.25% ethyl-3-aminobenzothiazole methanesulfonate for ~1 min to separate male and female parasites before fixation. Dehydrated juveniles were bleached in 3% H\(_2\)O\(_2\); in methanol for 30 min, rehydrated, permeabilized by 10 µg/mL proteinase K for 20 min, and then post fixed with 4% formaldehyde. Dehydrated adults were rehydrated, bleached in the bleaching solution (5% formamide, 0.5x SSC, 1.2% H\(_2\)O\(_2\)) for 45 min under bright light, permeabilized by 10 µg/mL proteinase K for 30 min, and then post fixed with 4% formaldehyde. The hybridization was performed at 52 °C with riboprobes, which were then detected through either alkaline phosphate-catalyzed NBT/BCIP reaction (for WISH) or peroxidase-based tyramide signal amplification (for FISH). Planarians were relaxed on ice and killed in 5% N-Acetyl Cysteine (NAC) solution for 10 min. They were fixed for 2 h in 4% formaldehyde supplemented with 1% NP-40, dehydrated in methanol, and stored at ~20 °C. Samples were then rehydrated for 2 h in the bleaching solution under bright light, permeabilized with 10 µg/mL proteinase K for 10 min, and post fixed with 4% formaldehyde. The rest of the steps were performed similarly to those described for schistosome in situ hybridization experiments, except that the hybridization was performed at 56°C. To compare gene expression between control and RNAi conditions using WISH, the development of WISH signal was progressed in parallel and simultaneously stopped.

**TUNEL assay.** Juvenile parasites were fixed and permeabilized as in situ hybridization experiments. EDU incorporation was detected by click reaction with 25 µM of Cy5-azide conjugates or Carboxyfluorescein 110 Azide conjugates (Click Chemistry Tools)\(^{52}\). To combine EDU detection with FISH, click reaction was performed after tyramide signal amplification.

**EDU labeling.** EDU (TCI Chemicals) was dissolved in DMSO at 10 mM and added into the medium at a final concentration of 10 µM. Both juvenile and adult schistosomes were pulsed with EDU overnight. Parasites were fixed and permeabilized as in situ hybridization experiments. EDU incorporation was detected by click reaction with 25 µM of Cy5-azide conjugates or Carboxyfluorescein 110 Azide conjugates (Click Chemistry Tools)\(^{52}\). To combine EDU detection with FISH, click reaction was performed after tyramide signal amplification.

**Code availability.** The SAM analysis source code, tutorials, and documentation are available through Github [https://github.com/atarashansky/self-assembling-organoid].
References

1. Hoffmann, K. F., Brindley, P. J. & Berriman, M. Halting harmful helminths. Science 346, 168–169 (2014).

2. King, C. G. & Dangerfield-Cha, M. The unacknowledged impact of chronic schistosomiasis. Chronic Illn. 4, 65–79 (2008).

3. Hotez, P. J. & Fenwick, A. Schistosomiasis in Africa: an emerging tragedy in our new global health decade. PLoS Negl. Trop. Dis. 3, e485 (2009).

4. Wang, B., Collins, J. J. III & Newmark, P. A. Functional genomic characterization of neoblast-like stem cells in larval Schistosoma mansoni. eLife 2, e00768 (2013).

5. Wang, B. et al. Stem cell heterogeneity drives the parasitic life cycle of Schistosoma mansoni. eLife 7, e35449 (2018).

6. Collins, J. J. III et al. Adult somatic stem cells in the human parasite Schistosoma mansoni. Nature 494, 476–479 (2013).

7. Pan, S. C. The fine structure of the exmidacum of Schistosoma mansoni. J. Invertebr. Pathol. 36, 307–372 (1980).

8. Cort, W. W., Ameel, D. J. & Van der Woude, A. Germinal development in the flatworm *Macrostomum lignano*. Dev. Biol. 357, 117–132 (2011).

9. Olivier, L. & Mao, C. P. The early larval stages of *Schistosoma mansoni* Sambon, 1907 in the snail host, *Astrorabdilus glabratus* (Say, 1818). J. Parasitol. 35, 267–275 (1949).

10. Whitfield, P. J. & Evans, N. A. Parthenogenesis and asexual multiplication in planarian regeneration. I. A karyological investigation. J. Parasitol. 67, 179–185 (1981).

11. Davies, S. J. et al. Modulation of blood fluke development in the liver by hepatic CD4+ lymphocytes. Science 294, 1358–1361 (2001).

12. Olivier, L. & Mao, C. P. The early larval stages of *Schistosoma mansoni* Sambon, 1907 in the snail host, *Astrorabdilus glabratus* (Say, 1818). J. Parasitol. 35, 267–275 (1949).

13. Whitfield, P. J. & Evans, N. A. Parthenogenesis and asexual multiplication in planarian regeneration. I. A karyological investigation. J. Parasitol. 67, 179–185 (1981).

14. Davies, S. J. et al. Modulation of blood fluke development in the liver by hepatic CD4+ lymphocytes. Science 294, 1358–1361 (2001).

15. Juliano, C. E., Swarts, S. Z. & Wessel, G. M. A conserved germline multipotency program. Development 137, 4113–4126 (2010).

16. Wang, Z. & Lin, H. Nanos maintains germline stem cell self-renewal by preventing differentiation. Science 303, 2016–2019 (2004).

17. Wang, Y., Zayas, R. M., Guo, T. & Newmark, P. A. Nanos function is essential for development and regeneration of planarian germ cells. Proc. Natl Acad. Sci. USA 104, 5901–5906 (2007).

18. Iyer, H., Collins, J. J. III & Newmark, P. A. NF-YB regulates spermatogonial stem cell self-renewal and proliferation in the planarian *Schmidtea mediterranea*. PLoS Genet. 12, e1006109 (2016).

19. Whittle, C. A. & Extavour, C. G. Causes and evolutionary consequences of meiotic gene conserved in *Drosophila*. Cell 66, 637–647 (1991).

20. Wisten, J. M. & Lehmann, R. Nanos is the localized posterior determinant in *Drosophila*. Cell 57, 585–588 (1989).

21. D’Silva, C., Murti, V. & Lang, P. A. The unacknowledged impact of chronic schistosomiasis. Nat. Rev. Immunol. 2, 499–511 (2002).

22. Witschi, H. & Lehmann, R. N. Route for transcriptional regulation of new markers for the vitelline lineage. Int. J. Parasitol. 46, 405–410 (2016).

23. Bauen, A. & Zanetti, G. K. Cytosine methylation regulates oviposition in the planarian *Schmidtea mediterranea*. Development 140, 2731–2741 (2013).

24. Sieg, G. A. & Balakrishnan, K. E. Neoblasts are adult somatic stem cells and germ cells. Development 244, 615–628 (2006).

25. Tarashansky, A. J., Xue, Y., Li, P., Quake, S. R. & Wang, B. Self-assembling manifolds in single-cell RNA sequencing data. eLife 8, e48994 (2019).

26. Xu, E. Y., Moore, F. L. & Reijo Pera, R. A. A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in metazoans. Proc. Natl Acad. Sci. USA 98, 7414–7419 (2001).

27. Steiner, J. K., Tasaki, J. & Roubaha, L. Germline defects caused by *Smed-boule* RNA-interference reveal that egg capsule deposition occurs independently of fertilization, ovulation, mating, or the presence of gametes in planarian flatworms. PLoS Genet. 12, e1006030 (2016).

28. Iyer, H., Issigonis, M., Sharma, P. P., Extavour, C. G. & Newmark, P. A. A primate-specific function for *boule* in the planarian *Schmidtea mediterranea*. Proc. Natl Acad. Sci. USA 113, E3509–E3518 (2016).
58. Xie, T. & Spradling, A. C. A niche maintaining germ line stem cells in the Drosophila ovary. *Science* **290**, 328–330 (2000).
59. Lehmann, R. Germ line stem cell origin and destiny. *Cell Stem Cell* **10**, 729–739 (2012).
60. Chong, T., Collins, J. J. III, Brubacher, J. L., Zarkower, D. & Newmark, P. A. A sex-specific transcription factor controls male identity in a simultaneous hermaphrodite. *Nat. Commun.* **4**, 1814 (2013).
61. Saberi, A., Jamal, A., Beets, I., Schoofs, L. & Newmark, P. A. GPCRs direct germline development and somatic gonad function in planarians. *PLoS Biol.* **14**, e1002457 (2016).
62. Beckmann, S. et al. Schistosoma mansoni: signal transduction processes during the development of the reproductive organs. *Parasitology* **137**, 497–520 (2010).
63. Collins, J. J. III et al. Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLoS Biol.* **8**, e1000509 (2010).
64. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
65. King, R. S. & Newmark, P. A. In situ hybridization protocol for enhanced detection of gene expression in the planarian Schmidtea mediterranea. *BMC Dev. Biol.* **13**, 8 (2013).
66. Hama, H. et al. Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. *Nat. Neurosci.* **14**, 1481–1488 (2011).

**Acknowledgements**

*S. mansoni* (strain: NMRI) was provided by the NIAID Schistosomiasis Resource Center for distribution through BEI Resources, NIH-NIAID Contract HHSN272201000005I. We thank M. Khariton, Y. Fan, S. Korullu, and R. Jones for experimental help, and J. Lee and J. Gao for the critical comments on the manuscript. P.L. is partially supported by an Agilent Bioengineering Fellowship. P.L., D.N.S., and X.Y. performed the functional experiments, P.L. and B.W. wrote the paper with input from all other authors, S.R.Q. and B.W. supervised the project and provided conceptual advice.

**Competing interests**
The authors declare no competing interests.

**Additional information**

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-20794-w.

**Correspondence**

and requests for materials should be addressed to B.W.

**Peer review information**

*Nature Communications* thanks Mireya Plass, Hong You and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

**Publisher’s note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[Open Access] This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021