Germline stem cell isolation, lineage tracing, and aging in a protochordate

Tom Levy[1,2*], Chiara Anselmi[1,2], Katherine J. Ishizuka[1,2], Tal Gordon[1,2], Yotam Voskoboynik[3], Erin McGeever[4], Angela M. Detweiler[4], Liron Levin[5], Karla J. Palmeri[1,2], Daniel Dan Liu[1], Rahul Sinha[1], Benjamin F. Ohene-Gambill[1], Tal Raveh[1], Maurizio Morri[4], Virginia Vanni[6], Lucia Manni[6], Debashis Sahoo[3,7], Norma F. Neff[4], Benyamin Rosental[8], Irving L. Weissman[1,2,9,10*†], Ayelet Voskoboynik[1,2,*†]

[1] Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA
[2] Department of Biology, Hopkins Marine Station, Stanford University, Pacific Grove, CA, USA
[3] Department of Bioinformatics and System Biology, Jacobs School of Engineering, University of California San Diego, San Diego, CA, USA
[4] Chan Zuckerberg Biohub, San Francisco, CA, USA
[5] Bioinformatics Core Facility, Ilse Katz Institute for Nanoscale Science & Technology, Ben-Gurion University of the Negev, Beer Sheva, Israel
[6] Dipartimento di Biologia, Università degli Studi di Padova, 35131 Padova, Italy
[7] Department of Pediatrics, University of California San Diego, CA, USA
[8] The Shraga Segal Department of Microbiology, Immunology, and Genetics, Faculty of Health Sciences. Center for Regenerative Medicine and Stem Cells. Ben Gurion University of the Negev, Beer Sheva, Israel
[9] Ludwig Center for Cancer Stem Cell Research and Medicine, Stanford University School of Medicine, Stanford, CA, USA
[10] Department of Pathology, Stanford University, Stanford, CA, USA

[†] These authors contributed equally and share last authorship
[*] Corresponding Authors E-mail: Tom Levy: levyt@stanford.edu
Irving L. Weissman: irv@stanford.edu
Ayelet Voskoboynik: ayeletv@stanford.edu

Abstract
Germline stem cells (GSCs), the source of gametes, are the only stem cells capable of passing genes to future generations and are therefore considered units of natural selection. Yet, the factors that influence GSC fitness, and thus govern GSC competition, which exist in both protochordates and mammals, remain poorly understood. We studied how aging affects GSC fitness in the protochordate *Botryllus schlosseri*, an evolutionary crosspoint between invertebrates and vertebrates. GSCs were isolated and distinguished from developing and mature gametes using flow cytometry and scRNA-Seq, facilitated by a new PacBio genome assembly. Moreover, their function was validated through a novel lineage tracing approach that combines membrane-labeled GSC transplantation with scRNA-Seq. Leveraging our method to isolate them, single-cell transcriptomics showed significant age-related changes between young and old GSCs. Spermatids and sperm, however, showed minimal changes, suggesting that reproductive aging is governed by GSCs rather than by gametes. Reduced expressions of markers like DDX4 and PIWIL1 in aged GSCs mirrored trends in mammalian datasets, pointing to a conserved GSC-driven aging mechanism across chordate evolution. This study provides new techniques that lay the foundation to investigate further drivers of GSC fitness and highlights fertility-related genes as promising targets for therapies to preserve reproductive health.

Main
Age-related decline in stem cell function and number within a given tissue can adversely affect tissue homeostasis[1]. This aging process may result from cell-intrinsic changes[2], deterioration of niche support[1,3] or both. While stem cells are responsible for the development and regeneration of tissues and organs, only GSCs, which form gonads and are the source of gametes, transmit their genomes to future generations via sexual reproduction. For this genetic transmission to occur, an organism must remain fertile. Therefore, it is highly important to investigate factors such as GSC aging, which may impair their fitness and potentially lead to reduced fertility[4-6].

Research on the protochordate colonial tunicate *B. schlosseri* has significantly advanced our understanding of fundamental principles in stem cell biology, including stem cell competition at the germline level[7-11] and stem cell aging[12]. GSC competition occurs when adjacent *Botryllus* colonies fuse blood vessels and create chimeras. In these chimeras, GSCs from one colony outcompete those from the other, resulting in gonads derived only from the 'winner' colony, which determines the genetic contribution to the next generation through sexual reproduction[7-11]. The germline winner outcome is heritable among independent individuals tested against loser colonies. Since winner phenotypes are inherited across sexually reproductive generations, it is conceivable that winner phenotypes derive from genetic alleles. The phenomenon of GSC competition has been also demonstrated in mouse testes, but the controlling mechanisms, as in *Botryllus*, remain a mystery[13]. In addition to sexual reproduction, *Botryllus* colonies expand through weekly asexual budding[14] (Fig. 1a). This unique process results in the coexistence of three generations within the same colony: adult zooids (first generation), primary buds (second generation), and secondary buds (third generation). This developmental cycle is mediated by

stem cells that migrate from resorbed zooids to the developing buds and create all the tissues and organs in subsequent generations[11,14-16]. Consequently, a new gonad forms weekly from GSCs in the secondary bud and matures in the adult zooid two weeks later, enabling continuous sexual reproduction. Furthermore, since stem cells are the only cells that persist throughout the colony's lifespan, and age as they move across generations, colony aging directly reflects the aging of its stem cells[12]. This makes *B. schlosseri* a unique model for studying gonadal development at the cellular level and GSC aging.

Studying specific types of stem cells in an organism requires their identification and isolation or enrichment. In *B. schlosseri*, stem cell enrichment was previously achieved using fluorescence-activated cell sorting (FACS) coupled with non-species-specific stem cell features and markers like increased alkaline phosphatase (AP) or aldehyde dehydrogenase (ALDH) activities[11,17]. Laird et al. confirmed stem cell function by transplanting FACS-sorted cells expressing high ALDH activity from donor colonies and showing that these cells successfully contributed to the genetics of recipient colonies[11]. However, a key limitation of this approach is that a high proportion of *Botryllus* cells naturally express ALDH activity, and the cells were enriched from suspensions of entire dissociated colonies and not from a specific cell population or tissue. Therefore, the fact that those cells contributed to either somatic or germline fates suggested that GSCs are separate from mixture of somatic tissue stem cells[11]. Also, because GSC migrations span at least 3 asexually derived generations, they could be 'contaminants' of tissue disaggregation of adult zooids. This underscores the critical need for strategies to identify and isolate single stem cells in a tissue-specific manner. In fact, Rosental et al. identified hematopoietic stem cells (HSCs) from *B. schlosseri* and demonstrated multilineage differentiation into several blood cell types[17]. However, to date, no protocol has been established for the specific enrichment of GSCs in *Botryllus*, and their biology, aging and fitness-controlling factors remain largely unexplored.

In this set of studies, we found ways to investigate how aging affects reproductive fitness at the cellular level. We developed a method to isolate GSCs and other cell types involved in spermatogenesis in *B. schlosseri*. We characterized their cellular morphology, validated their function in vivo, and tracked their differentiation pathways in vitro and in situ during the organism's weekly developmental cycle. To identify age-related changes, we generated high-resolution single-cell transcriptomic profiles and compared them between young and old animals, aided by a newly assembled *B. schlosseri* genome. Finally, to shed light on the evolutionary conserved elements of GSC aging from invertebrate chordates to mammals, we analyzed our *Botryllus* single-cell RNA sequencing (scRNA-Seq) dataset in comparison to published mouse and human single-cell germline datasets[18,19].

Isolation and characterization of germline stem cells from the testis

B. schlosseri is a hermaphrodite species containing both male and female gonads. However, since the female gonad activity in *Botryllus* is limited to the reproductive season, and contain none to only a few oocytes per zooid[20,21], we focused on the male gonads, which are more

abundant and develop weekly. We searched for spermatogonial stem cells (spermatogonia), the type of GSCs that, through spermatogenesis in the testis, give rise to mature sperm (Fig. 1b). Using FACS, we sorted 3 different cell populations based on size and granularity of cells expressing increased AP activity (Fig. 1c-h). AP activity has been used to select for GSCs in mammals[22] and we hypothesized that it may be a characteristic of stem cells in the gonads of protochordates. Increased ALDH activity was also used previously to select for stem cells in mammals[23,24] and marine invertebrates[25], including *B. schlosseri*[11], and FACS analysis of cells expressing high ALDH activity in the testis of *Botryllus* yielded the same cell populations as AP (Extended Data Fig. 1).

Using this sorting strategy, we identified the sperm and spermatid cell populations (Fig. 1e, f) but could not define the cells in the other population (Fig. 1g). However, since the sperm and spermatid populations were clearly distinguishable, and since the AP activity level in the morphologically undefined cell population was much higher (Fig. 1h), we hypothesized that the latter contains the spermatogonial stem cells and named it as candidate germline stem cell (cGSC) population. To identify the GSCs and further characterize the different cell types in *Botryllus* testis, we index sorted 3,264 cells, using FACS, from the sperm, spermatid and cGSC populations into 96-well plates for full-length scRNA-Seq using Smart-seq3[26]. Index sorting was used because it records the fluorescence intensities for each sorted cell, allowing for direct correlation between its transcriptome and its exact location on the FACS plot. To include all possible cell types involved in spermatogenesis, the cells for the scRNA-Seq were isolated from testes at different time points along the *Botryllus* developmental cycle including adult zooids in stage A and primary buds and adult zooids in stage C. The developmental stages are illustrated in Fig. 1a and were defined according to Lauzon et al.[27] (as explained in Supplementary Table 1). Alignment of the scRNA-Seq and FACS data showed that the transcriptome-based cell clusters (Fig. 1i) clearly distinguished the different sorted cell populations (cGSC, spermatid and sperm) (Fig. 1j). We annotated each cluster based on gene profiles of known cell-type-specific markers (Fig. 1k, Extended Data Fig. 2, Supplementary Table 2). To aid with the analysis we generated, using PacBio platform, a new *B. schlosseri* genome assembly and annotation whose contiguity as measured by the N50 value (1,097 kb) was much higher than our previous published genome[28] (7 kb) (Details in Supplementary Table 3). Based on the scRNA-Seq analyses, *DDX4,* (*VASA*)[18,29,30] and *PIWIL1*[30-32] were used as representative markers for spermatogonia; *SYCP3*[18,30,33] and *H1F0*[34,35] were markers for spermatocytes, *PSME3*[36-38], *UBE2R2*[36,37,39] and *SAXO4*[36,37,40] (formerly *PPP1R32*) for spermatids and *PRM1*[41], *SPA17*[42] and *TSSK1B*[43] for sperm (Fig. 1l, Supplementary Table 4). Cluster 16 did not express specific markers for cells involved in spermatogenesis but expressed high levels of *ACTIN* transcripts. We hypothesize that this cluster may contain *ACTIN* RNAs from myocytes and other cells expressing cytoskeletal actins for their own functions. Most of the highly expressed genes of Cluster 6 were not annotated and those that were annotated could be correlated with either spermatocytes, spermatids or sperm. Therefore, we hypothesize that this cluster contains actively differentiating cells from spermatocytes to spermatids and sperm. Importantly, Cluster 10, which contains cells from the

cGSC FACS gate, exhibited strong expression of spermatogonia-associated genes such as *DDX4*[18,29,30] and *PIWIL1*[30-32], germline markers like *TDRD7*[44], along with genes linked to proliferating cells such as spermatogonia, like *PCNA*[45]. This suggests that we successfully identified GSC population in *B. schlosseri* testis and developed a reproducible FACS-based method to isolate them, enabling us to further investigate their capacity to engraft and differentiate in vivo, as well as to explore their aging dynamics.

Identifying primordial germ cells in the cell islands niche

The *B. schlosseri* cell islands (Extended Data Fig. 3a) form a putative niche for both germ and somatic stem cells[15]. To evaluate the presence of GSCs in that niche, we index-sorted cells with high AP activity from dissected cell islands isolated from adult Stage A zooids (Extended Data Fig. 3b). A total of 384 index-sorted cells were analyzed by scRNA-Seq (Extended Data Fig. 3c). Cluster 10 expressed prominent markers of primordial germ cells (PGCs) such as *NANOS3*[46] and *AGO1*[47] (Extended Data Fig. 3d, e, Supplementary Table 4). This indicates that the cell islands may present a niche for PGCs which later differentiate into spermatogonia upon migration to the gonadal niche in which they establish the testis. This supports the previously proposed role of the niche as a source of GSCs[15]. Based on gene expression (Extended Data Fig. 4, Supplementary Table 5), the other clusters were found to include gonadial cells, muscle cells, epithelial cells and hemocytes (Extended Data Fig. 3f). Due to the small size, embedded location and proximity to surrounding tissues, some of these cells could have been inadvertently picked up from tissues adjacent to the cell islands during dissection.

Germline stem cell engraftment and differentiation in vivo

A fundamental trait of stem cells is their ability to give rise to different cell types while simultaneously maintaining a self-renewing stem cell pool. In vivo transplantation and tracing of cells in non-model organisms is challenging due to the lack of ubiquitous reporter lines. Therefore, to investigate the in vivo capacity of the cGSC population to engraft and differentiate into sperm in *B. schlosseri*, we developed a novel method combining transplantation and scRNA-Seq. This approach involved transplantation of dye-labeled donor cGSCs into compatible syngeneic recipients (to avoid histocompatibility-related rejection[48-50]) followed by index sorting and scRNA-Seq. Early developing GSCs begin to accumulate in the secondary buds[14] (Extended Data Fig. 5). However, due to their small size, we could not dissect their gonadal niche. Therefore, to enrich for GSCs from both the buds and zooids, for the transplantation assay, we applied the FACS gates that we developed from pure testes to a cell-suspension produced from whole animals at stage B. From this suspension, we sorted cells with high ALDH activity. ALDH was selected to incorporate another marker that was previously used for stem cell identification in *B. schlosseri*[11], and it showed similar staining patterns compared to AP as part of our FACS gating strategy (Extended Data Fig. 1). ALDH+ cells were sorted based on the cGSC gate (Fig. 2a), labeled with a fluorescent membrane dye (DiD), and transplanted into recipient animals at stage B through injection into their vasculature. Seven days post-

transplantation (one developmental cycle), we used FACS to analyze cells from the gated "Testes" population (combining cGSCs, spermatids, and sperm) from the whole animals, that reached to stage B again and captured 1,726 cells from the DiD+ gate for scRNA-Seq (Fig. 2b). Control animals were subclones of the same genotype and developmental stage but not transplanted. Cells from these control animals, which exhibit natural autofluorescence in the DiD+ gate, were sorted for sequencing and labeled as "control". Cells from the transplanted animals which potentially contained gametes from engrafted GSCs, in addition to autofluorescent cells in the DiD+ gate, were sorted for sequencing and labeled as "transplanted" (Fig. 2b). The proportion of DiD+ cells in the control animals (n=2), was lower compared to the transplanted animals (n=3) (Fig. 2c). Clustering analysis of the DiD+ cells (Fig. 2d), emphasized Cluster 10, which highly expressed the spermatid and sperm gene markers previously investigated in this study (*PSME3*, *UBE2R2*, *SAXO4*, *PRM1*, *SPA17*, *TSSK1B*) (Fig. 2e). Analyzing the UMAP expression plots of these genes (Fig. 2f), combined with the observation that their expression levels were higher in the DiD+ cells from transplanted animals compared to controls (Fig. 2g), allowed us to identify the sperm and spermatid progeny of the engrafted GSCs (Fig. 2h). These results indicate that we not only successfully isolated a GSC-enriched cell population but also revealed that these cells were functionally viable to home to the gonadal niche of the recipients, engraft, and undergo differentiation into spermatids and sperm in host animals.

Spatiotemporal dynamics of gonadal development at the cellular level

B. schlosseri colonies undergo a complete whole-body regeneration cycle every week[14,27]. As a crucial part of this process, GSCs must migrate from resorbed zooids to the developing buds to establish the gonads of the next generation[14,20]. To follow testis development at the cellular level, we dissected testes from primary buds at stages C and D, and from adult zooids at stages A, B and C. Testis dissection from primary buds at earlier stages and from secondary buds was not feasible due to their small size. Using FACS, we analyzed the fraction of the cGSC, spermatid and sperm cell populations at each developmental stage (Fig. 3a-e). While the cGSC population was consistently present in testes across all tested parts and stages, spermatids first appeared in the primary bud at stage D, and sperm first observed in the adult zooid at stage A. The proportions of each population along the developmental cycle show that GSCs gradually differentiate into spermatids and sperm, with the highest fraction of sperm cells observed in the testes of adult zooids at stage C (Fig. 3f). These results reveal that at stage C, while the testis of the primary bud contains mostly GSCs, in the adult zooid within the same animal, most GSCs have already differentiated into spermatids and sperm. This is a remarkable coexistence of two generations within the same individual having testis at different developmental stages. Since we couldn't dissect testis from secondary buds, we used histology and 3D reconstructions to identify the earliest developmental stage at which gonads begin to form (Extended Data Fig. 5). A few PGCs were found in the gonadal niche of the secondary bud at stage B (Extended Data Fig. 5a,

b), and more organized PGCs along with early developing oocytes were observed in the secondary bud at stage C (Extended Data Fig. 5c-f).
To further investigate gonad development and validate our scRNA-Seq results, we examined the expression pattern of 3 marker genes from our scRNA-Seq analysis using in situ hybridization chain reaction (HCR) on whole-mount animals along the weekly developmental cycle. The markers chosen were *DDX4* (*VASA*; for GSCs[29]), *H1F0* (for spermatocytes[34,35]) and *PRM1* (for spermatids/sperm[41]) (Fig. 4a-d). *DDX4* expression began in PGCs in the secondary bud at stage B (Fig. 4g) and in oocytes in the secondary bud at stage C (Fig. 4h). Its expression was clearly visible in the developing testis in primary buds from stages B to D (Fig. 4g-i), but disappeared once the primary buds transitioned to stage A zooids (Fig. 4f). *H1F0* expression was testis-specific, beginning in the primary bud at stage C (Fig. 4m). It remained strong at stage D (Fig. 4n) but began to fade in the adult zooids at stages A and B. This fading coincided with the differentiation of spermatogonia and spermatocytes into spermatids and sperm, as demonstrated by the increasing expression of *PRM1* in the center of the testis lobes (Fig. 4k, l). Strikingly, aligned with our FACS data of testes from primary buds and adult zooids (Fig. 3a, e), the in situ gene expression profiles of *H1F0* and *PRM1* at stage C also demonstrated the coexistence of testes from two different generations within the same animal (Fig. 4m).

Germline stem cells drive reproductive aging across species

Reproductive capacity declines with age in many animals and stem cell-based aging mechanisms are conserved across species[5,18,19,51,52]. Using our newly developed strategy to isolate GSCs in a protochordate, we investigated their role in gonadal aging at the single-cell-transcriptomic level and compared these findings to mice and humans. Using FACS, we index-sorted and analyzed 1,968 single cells with high ALDH activity from the testes of young (6-months old) and old (163-months old) *Botryllus* colonies in stage A. The cells were sorted from both the GSC population which is enriched with spermatogonia and from the STD+SZ population containing spermatids and sperm (Fig. 5a). Clustering analysis demonstrated clear separation between cells derived from young and old colonies (Fig. 5b, c). Remarkably, based on gene expression of the top differentially expressed genes, we observed a higher variation between young and old GSCs compared to young and old spermatids/sperm (Fig. 5d, e, Supplementary Table 6). These results suggest that reproductive aging may be driven mainly by differential gene expression at the GSC level rather than at the level of mature gametes. To further investigate the germline single-cell-transcriptomic patterns in young and old *Botryllus* colonies, we analyzed the expression level of the main gene markers studied throughout this project. The expression of markers associated with spermatogonia and spermatocytes (*DDX4*, *PIWIL1*, *SYCP3*, *H1F0*) was higher in young compared to old colonies, while markers associated with spermatids and sperm (*PSME3*, *UBE2R2*, *SAXO4*, *PRM1*, *SPA17*, *TSSK1B*) showed the opposite pattern (Fig, 6a, b, Extended Data Fig. 6, Extended Data Fig. 7a). A comparative meta-analysis of scRNA-Seq dataset from mouse testes[19] showed a similar expression pattern for the same set of genes among cells from young (2-months old) and old (24-months old) animals (Fig. 6c, d, Extended Data Fig. 7b).

However, the expression pattern of those genes was slightly different in humans. The meta-analysis of scRNA-Seq data of human testis cells[18] showed higher expression in younger men for *DDX4*, *PIWIL1*, *SYCP3*, *SAXO4*, *PRM1*, *SPA17* and *TSSK1B* but lower expression for *H1F0*, *PSME3* and *UBE2R2* (Fig. 6f, Extended Data Fig. 7c). Still, it is outstanding that even with the high variation between human samples of the same age (Fig. 6e), which may be attributed to factors like elevated body weight, smoking[18] or simply diverse genetics, the higher expression levels of the GSC markers (*DDX4*, *PIWIL1*) in young compared to old samples was consistent with the pattern observed in mice and *Botryllus*. This strongly suggests the existence of a conserved reproductive aging mechanism, primarily governed by GSCs rather than by gametes, that has been maintained across hundreds of millions of years of evolution since the divergence of protochordates and mammals.

Discussion

Reproductive aging, a universal phenomenon impacting fertility across the animal kingdom, remains a significant biological puzzle. Previous single-cell studies reported varying degrees of age-related transcriptomic differences in GSCs, ranging from distinct[19,53] to modest[18]. However, these studies used droplet-based scRNA-Seq platforms on unsorted testis cell suspensions, which likely included transcriptional noise from unrelated cell types present in the heterogeneous cell samples. To overcome this limitation, we employed a plate-based method combined with index sorting to analyze purified testis cell populations, enabling high-resolution analysis of age-related changes within defined cell types.

With the aim to characterize the different cell types involved in *B. schlosseri* spermatogenesis, we combined two previous approaches that used FACS and AP/ALDH activity indicators as non-species-specific stem cell markers[11,17,25]. We employed it on cell suspension from dissected germline niches (testis and cell islands) followed by scRNA-Seq analysis, in vivo transplantation and in situ tracing for validation of stem cell differentiation. This is a generalized method for characterizing tissue-specific stem cells in non-classical model organisms, which often lack transgenic tools and gene-specific monoclonal antibodies. Using this approach, and based on known prominent gene markers, we report here the first high-resolution identification and isolation of the different cell types involved in *Botryllus* spermatogenesis including spermatogonia, spermatocytes, spermatids and mature sperm. While the cell islands in *B. schlosseri* were previously suggested to include GSCs[15], the presence of diverse cell types within this niche made their definitive identification elusive. Our results here provide a cellular-level analysis that reveals the transcriptomic signature of PGCs, thus confirming that the cell islands are indeed a GSC niche. Furthermore, by tracking germline markers, we precisely mapped the weekly developmental cycle of the *B. schlosseri* testis at the cellular level. Our results suggest that this process begins with GSC accumulation in the secondary bud at Stage B, progresses through differentiation, and ends two weeks later with sperm release from the testis of the adult zooid at Stage C. The expression pattern of the sperm/spermatid-related marker, *PRM1*, aligned with the sperm release temporal pattern that was previously reported for this organism[54].

Accordingly, we suggest that PGCs either initially develop in the cell islands and then migrate to the secondary bud, where they initiate testis formation and differentiate into spermatogonia, or that they originate in the secondary buds and then transit through the cell islands to the developing testes.

The strategy we have developed to isolate GSCs and other germline cell types allowed us to investigate their aging dynamics at the single-cell level. A core strength of *Botryllus* as a model lies in its unique life history, in which the aging of a colony is dictated by the aging of its stem cells, which self-renew and migrate each week across generations. The remarkable 163-months old colony that was used in this study represents approximately 700 generations of persistent stem cell activity. Given that aged stem cells may be prone to higher mutation rates[5], we hypothesize that this constant migration of GSCs can lead to an accumulation of genetic alterations which affect gene expression and impact fertility.

Our study reveals a strikingly distinct aging patterns among different cell types within the *Botryllus* male germline. GSCs exhibit substantial age-related transcriptional changes, while their differentiated gamete progeny maintains relatively stable transcriptomic profiles. This central finding strongly suggests that GSCs, rather than differentiated germ cells, are the primary drivers of reproductive aging. Specifically, we found a significant age-related decline in the expression of the prominent GSC markers *DDX4* and *PIWIL1*. These genes are essential for GSC function and fertility in many organisms[29,31,55,56], including protochordates[57-59] and thus may constitute a conserved mechanism for maintaining reproductive fitness across species.

Our comparative analysis revealed similar expression patterns of spermatogenesis-related gene markers in both *Botryllus* and mice between young and old testes cells but a slightly different pattern when comparing cells from young and old human testes. The substantial variation between human samples, potentially due to lifestyle factors like smoking and increased body weight[18,60], might have contributed to differences in the gene expression pattern compared to *Botryllus* and mice. Interestingly, despite this high variability observed between human samples which can be attributed to genetic or health-related factors, the consistent age-related decline in *DDX4* and *PIWIL1* expression in human germline cells mirrored the pattern observed in *B. schlosseri* and mouse. These results support our central hypothesis that reproductive fitness and its underlying molecular and physiological factors are driven by GSCs and are likely conserved across evolution, from simpler to more complex organisms.

This is not the first demonstration of a GSC-related mechanism showing a conserved pattern among protochordates and mammals. The GSC competition phenomenon observed in *Botryllus*[9,10] prompted us to search for a similar mechanism in mammals. We have previously shown that the germline in mice is derived from only 4 cells at the epiblast stage. Transplantation of color-coded embryonic stem cells into uncolored blastocysts produced tetrachimeric mice with cells of all colors in most of their tissues, including the fetal and neonatal testes. However, during the reproductive phase, most GSCs underwent apoptosis, and only the surviving cells contributed to the mature gonad[13]. These remarkable findings suggest the existence of GSC competition

across protochordates and mammals, but the underlying mechanisms that determine GSC fitness remain largely unclear.

Investigating the driving factors of 'winner' GSCs is highly important since GSCs are the only stem cells capable of passing genes to future generations, making them fundamental units of natural selection[8,11]. Therefore, the present study, which reports the first isolation of GSCs in this unique protochordate model is a foundational step towards identifying the molecular and cellular factors that govern GSC competition. Moreover, the evolutionary position of protochordates, as the likely closest living invertebrate relatives to vertebrates[61], makes the present study of their GSCs highly valuable from a multidisciplinary perspective, as insights from these organisms may be further applied to more complex organisms, including mammals. From an ecological standpoint, understanding how GSCs function can pave the way for innovative strategies to conserve endangered species by addressing the impacts of environmental stress on their GSC fitness and on their reproductive health. On a broader scale, in the fields of regenerative medicine and reproductive aging, the fact that basic biological mechanisms, such as the aging of GSCs explored in this study, and GSC competition, seem to be conserved across protochordates and mammals, suggest that our findings and the techniques established to isolate GSCs in a non-classical model organism like *B. schlosseri* can potentially be harnessed in the future to develop therapies to improve human reproductive health.

Methods

Animal husbandry

B. schlosseri mariculture procedures have been previously described[62]. Animals were collected from the marina in Monterey, California and transferred to the mariculture facility at Hopkins Marine Station. The animals were tied to 3 × 5-cm glass slides while another empty glass slide (to collect hatches) was placed in a slide rack in front of the slide with the tied animal. The slide rack was placed in an aquarium with filtered seawater, and within a few days the sexually reproduced larvae hatched, swam to the empty slide, settled and metamorphosed into the adult body plan. Single hatched animals were then transferred to individual slides, allowing them to grow and form colonies of genetically identical zooids and buds that were used in the experiments.

Tissue dissociation

To avoid food contaminants coming from the digestive system in our experiments, the animals were isolated without food for 24 hours prior to dissection. Insulin syringes were used to dissect testes and cell islands from laboratory grown *B. schlosseri* colonies. The dissection procedure was performed under a light stereomicroscope and dissected tissues were collected into 96-well clear flat bottom plates filled with a serum-free "*Botryllus* medium" (Medium 199 Hanks' Balanced Salts with phenol red and 3.3x calcium/magnesium-free HBSS) and placed on ice until used for further analysis. The testis is a cloudy-looking white tissue located on one or both sides

of the developing buds and adult zooids. The cell islands are cell aggregates aligned laterally on both sides of the endostyle niche[15-17] (Extended Data Fig. 3a).

FACS analysis

Cell suspensions from dissected testes (n=15) and cell islands (n=30) were generated by mechanical dissociation using insulin syringe plunger and filtered through a 70-µm mesh followed by a 35-µm mesh into a 5 ml tube containing serum-free *Botryllus* medium. No enzymatic dissociation was carried out. Cells were centrifuged at a speed of 500 G for 5 min, supernatant was discarded, and cells were resuspended in 1 ml of fresh *Botryllus* medium.
To enrich for stem cells, 1 µl of Alkaline Phosphatase Live Stain (Thermo Fisher) or ALDH (ALDEFLUOR Kit, Stemcell Technologies) was added. Cells were incubated for 30 min on ice under dark conditions, and 1 µl of Propidium Iodide (PI) (1µg/ml) was added before flow analysis in order to separate live from dead cells. After gating for live PI negative cells, we analyzed the cells using forward (FSC) and back-scatter (BSC), and fluorescence from the AP/ALDH staining using a Sony MA900 FACS instrument (Fig. 1b-d, Extended Data Fig. 3a, b, Extended Data Fig. 1). Flow cytometry data were analyzed using Flowjo v10.10 Software (BD Life Sciences). For morphological observations, the sorted cells were collected in 5 ml tubes containing 1 ml *Botryllus* medium and centrifuged at a speed of 500 G for 5 min. The supernatant was discarded and the remaining medium (~30 µl) was spined down, gently mixed and transferred to µ-slide 18 well glass bottom chamber (Ibidi). The chamber was placed on ice for 20 min to allow cells to settle down followed by imaging using Zeiss LSM700 confocal microscope.

In vivo transplantation of GSCs

For the transplantation experiment we used reproductive (visible testes) syngeneic animals (compatible subclones of the same genotype). Cell suspension from a donor whole animal (stage B) was generated using mechanical dissociation as described above. Cells with high ALDH activity from the cGSC FACS population were sorted, centrifuged, resuspended in 500 µl of *Botryllus* medium and labeled with a 5 µl of Vybrant DiD Cell-Labeling Solution (Invitrogen). The cells were incubated for 1 h in room temperature under dark conditions, washed twice with *Botryllus* medium (first with 2% FCS and then with serum-free medium) and resuspended in 30 µl of fresh serum-free medium before injection. For injection, glass needles were prepared using a P-1000 micropipette puller (Sutter Instruments). Needles were loaded with 5 µl of cell suspension (equivalent to ~30,000 cells) and injected using manual micromanipulator (World Precision Instruments) under a stereoscope into the proximal ampulla of recipient animals, also in stage B (n=3). Non-transplanted, stage B, subcloned animals of the same genotype (n=2) served as a control. Seven days after transplantation, following one regeneration cycle and takeover, we used our FACS gating strategy to sort GSCs, spermatids and sperm to analyze the DiD+ cells and captured them for scRNA-Seq to test whether the cells we injected the week before differentiated into spermatids and sperm. The DiD+ cells were captured from whole-

animal cell suspension produced from the reproductive recipient animals that reached again to stage B. *B. schlosseri* cells have natural autofluorescence in the far-red channel (used to detect DiD+ cells) and the DiD+ cells that were captured for scRNA-Seq from non-transplanted control animals were attributed to this autofluorescence.

Single cell RNA sampling and sequencing

To identify the GSCs, which represent a small fraction of the total cells in the mature testis and in the cell islands niche, we employed index sorting followed by the Smart-seq3 method[26]. This plate-based single cell approach provided paired-end reads and full-length transcripts, offering a higher sensitivity and coverage per cell that allowed us to identify a rare cell population (GSCs), which might not have been possible if we used droplet-based platforms for scRNA-Seq. Single cells were index-sorted using Sony MA900 FACS and processed using a modified Smart-seq3 protocol[26,63]. Cells were sorted into 96-well plates containing 2 µl of lysis buffer and plates were immediately centrifuged at a speed of 3000 G for 30 sec, snap frozen on dry ice, and stored at −80 °C. Reagents for all procedures including lysis plates preparation, reverse transcription, PCR amplification, plates clean-up and library preparation were used as previously described[63]. Before reverse transcription we thawed the cell-containing lysis plates on ice and incubated them at 72 °C for 3 min followed by immediate snap chill on ice to anneal the oligo $dT_{30}VN$ primer. Reverse transcription reagents (3 µl per well) were added using a Mantis liquid handler (Formulatrix), and plates were incubated in a C1000 Touch Thermal Cycler (BioRad) at 42° C for 90 min, and then 70° C for 15 min to terminate the reaction. For cDNA amplification, 7.5 ul of PCR mix was added to each well and cycled using the following program: 98 °C for 3 min, followed by 25 cycles of 98 °C for 20 sec, 67 °C for 15 sec, and 72 °C for 6 min, and then by a final elongation step of 72 °C for 5 min. Clean-up of amplified cDNA was performed using 0.65–0.75x volume of calibrated AMPure XP beads (Beckman Coulter) to remove residual reaction components and oligo fragments smaller than 400 base pairs, and eluted in 12.5 µl of ultra-pure water. The clean-up step was carried out using the Bravo automated liquid handling platform (Agilent). For quality control, 1 µl of purified cDNA was taken from each well. cDNA concentrations and size distributions were determined using Fragment Analyzer (Advanced Analytical). The wells within the 96-well plates were combined into new 384-well plates using the Mosquito X1 liquid handler (SPT Labtech). In the destination plate, the cDNA concentration was normalized, by diluting with ultra-pure water, to ensure it did not exceed 1 ng/µl.

Normalized cDNA was used to prepare Illumina sequencing libraries. Tagmentation was performed by combining 0.8 µl cDNA with 1.2 µl homebrew Tn5 mix as previously described[63]. The reaction was terminated by adding 0.4 µl neutralization buffer (0.1% SDS). Indexing PCR reactions were performed by adding 5 µM i5 and i7 index primers, 0.4 µl each (Integrated DNA Technologies, custom made 7680-plex unique dual index-primer set), and 1.2 µl KAPA HiFi HotStart ReadyMix (Kapa Biosystems). PCR amplification was performed using the following program: 72 °C for 3 min, 95 °C for 30 sec followed by 11 cycles of 98 °C for 10 sec, 67 °C for 30 sec, and 72 °C for 1 min. We pooled 1 µl of each well within the 384-well plates followed by

a purification step using 0.8x volume of AMPure XP beads (Beckman Coulter). The 384-cell library pool from each plate was analyzed for concentration and size distribution using Fragment Analyzer (Advanced Analytical). All 384-cell library pools were normalized for concentration and pooled into a single tube that contained all the cell libraries. Following another purification and concentration step using 0.8x AMPure beads, the final library pool was sequenced on the NovaSeq 6000 S4 flow cell (Illumina) to obtain ~1.5 million 2×150 base-pair paired-end reads per cell.

Single cell RNA analysis

For read mapping, sequences were demultiplexed followed by quality trimming with TrimGalore v0.5.0 and aligned to our newly generated *B. schlosseri* genome with STAR program[64] (parameters: --outFilterMultimapNmax 20 --outFilterMismatchNoverLmax 0.04 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --sjdbScore 1 --alignIntronMax 1,000,000 --alignMatesGapMax 1,000,000). The bam files from the STAR alignment were used to calculate expression levels of genes and transcripts and incorporate them into gene count tables. Gene counting was executed using HTSeq v2.0.1_py39[65]. Gene count tables were combined with metadata using the Scanpy python package v.1.10.1[66]. We filtered out cells that expressed less than 5 genes. The data were normalized using size factor normalization so that every cell has 10,000 read counts, log transformed and scaled to a maximum value of 10. Highly-variable genes were computed using default parameters, and principle component analysis (PCA) was used to compute the neighborhood graph, followed by data clustering using the Leiden method[67]. Step-by-step instructions and detailed code to reproduce our data preprocessing and analysis are available in the supplementary information. To assist with cluster annotation, we used GeneAnalytics[68] to identify the highest cell type score based on the top 200 most highly expressed genes in each cluster.

Differential expression analysis

To further investigate differential expressions of the main genes discussed in this study (*DDX4*, *PIWIL1*, *SYCP3*, *H1F0*, *PSME3*, *UBE2R2*, *SAXO4*, *PRM1*, *SPA17*, *TSSK1B*) between young and old *B. schlosseri* cells involved in spermatogenesis, we used the gene count tables generated for the scRNA-Seq analysis. The data was normalized (Log2 CPM) and gene expression was compared across age groups and cell types for the selected genes and plotted using seaborn's boxplot function. Statistical difference was tested using scipy.stats independent student t-test (ttest_ind) with no assumption of equal variance across the tested groups. Significance levels were set to 0.05*, 0.01** and 0.001***. Detailed code to reproduce our differential expression analysis is available in the supplementary information.

Genomic DNA isolation

Genomic DNA (gDNA) was extracted from a whole system of a 3.5-year-old *B. schlosseri* colony (strain: 356a) that was collected in Monterey Bay, CA. This strain was also used for the

assembly of our previously generated The *B. schlosseri* genome[28] to simplify comparison between the gene models of the two different genome assemblies. The gDNA was prepared using a modified silica based DNA extraction technique[69]. The *B. schlosseri* sample was ground under liquid nitrogen to a fine powder, then transferred into a 1.5 ml tube containing a solution of 10 M Guanidine Isothiocyanate lysis buffer and GeneClean Glass Milk (MP Biomedicals). Lysed samples were incubated at 56 °C, washed and pelleted by centrifugation, then eluted in 0.1M TE. The final sample yielded 2.02 µg of high molecular weight gDNA with an average fragment size of 16 kb, which was quantified using Qubit (dsDNA broad range kit) and visualized using an Agilent Fragment Analyzer (genomic DNA 50 kb kit).

Genome sequencing and assembly

A sequencing library was prepared according to PacBio's "Preparing Whole Genome and Metagenome Libraries Using SMRTbell Prep Kit 3.0" protocol, excluding shearing and size selection to preserve gDNA input. The final library size was verified using an Agilent 4150 TapeStation and quantified with Qubit. The Sequel II Binding Kit 2.0 was used for primer annealing and polymerase binding of the library, followed by loading (140 pM) on two 8M SMRT cells and sequenced on PacBio's Sequel IIe system. The sequencing "HiFi" application, circular consensus sequencing (CCS), and genome assembly were set up and performed in SMRT Link (v 11.0). The productivity rate (P1) for the two SMRT cells was 74%, on average, yielding a combined 4.9M HiFi reads with an average read length of 7.7 kb and Q38 read quality. SMRT Link's genome assembly analysis (default settings) yielded 2.1K contigs with an N50 contig length of 1.1 Mb. This represents a substantial improvement over our original *B. schlosseri* genome assembly[28], demonstrating the enhanced resolution achieved through long-read sequencing and assembly programs that were specifically designed for it.

Genome annotation

We used raw sequence reads (NCBI accessions: PRJNA732987, PRJNA579844, PRJNA205369 and SRP022042) of previously published data[14,28,48,70] to annotate the newly generated *B. schlosseri* genome. The raw sequence reads were analyzed using the NeatSeq-Flow platform[71]. The sequences were quality trimmed and filtered using Trim Galore (v 0.4.5) and cutadapt (v 1.15 parameters: --length 50 –quality 20 --max_n 1 --trim-n). Quality trimmed reads were mapped to our newly generated *B. schlosseri* genome. Mapping was done using STAR program[64] (v 2.7.11a parameters: --outFilterMultimapNmax 20 --outFilterMismatchNoverLmax 0.04 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --sjdbScore 1 --alignIntronMax 20,000 --alignMatesGapMax 20,000). The resulting alignment was used to create Genome guided transcriptome assembly using the Trinity program[72] (V2.8.4 parameters: --min_kmer_cov 2 --genome_guided_max_intron 20,000). The Genome guided transcriptome assembly was performed for each strain independently (944axByd196 and Sc109e), and then the created transcriptomes were merged into one fasta file to be used for downstream gene model prediction and annotation. The new *B. schlosseri* genome was masked using the funannotate mask

program[73] (v 1.8.15 parameter: -s "Tunicata" -m "repeatmasker"). The Masked genome was then used alongside the merged transcriptome and the raw reads to train the genes model with the funannotate train program (parameters: --max_intronlen 20,000 --aligners gmap). Next, we used funannotate predict program (using the following parameters: --busco_db metazoan --max_intronlen 20,000 --repeats2evm --organism other) that takes the PASA[74] gene models from the previous step to train Augustus[75]. Then we added UTR data to the predictions and fixed gene models that disagreed with the RNA-seq data using the funannotate update program (parameter: --max_intronlen 20,000). Finally, we added functional annotation to the genes model by running funannotate iprscan and funannotate annotate (parameter: --busco_db metazoan). In the funannotate annotate, we have changed the code of the step (annotate.py) to not restrict annotation by pident (percent identity) and to not exclude annotations that do not match the "CuratedNames" database (funnotate ncbi_cleaned_gene_products.txt file).

Histology and 3D reconstructions

For histology, samples were fixed in 1.5% glutaraldehyde in 0.2 M sodium cacodylate and 1.6% NaCl buffer for 2 h. Samples were washed 3 times in 0.2 M sodium cacodylate and 1.6% NaCl buffer followed by post-fixation in 1% OsO4 in 0.2 M cacodylate buffer for 1.5 h at 4 °C. The samples were dehydrated, soaked in Epon and propylene solution at 37 °C, 45 °C, and 60 °C and then embedded in resin. Sectioning (1 µm thick) was carried out using a Leica ultramicrotome followed by staining with toluidine blue. For 3D reconstructions, chains of histological sections stained with toluidine blue were photographed using Leica DMR optical microscope. The Images were then aligned using Adobe Photoshop and the 3D anatomical model of each sample, based on image stacks, was created with Amira software (Thermo Fisher).

In situ hybridization

Short in situ hybridization antisense DNA probes were ordered as lyophilized oligo pools from Integrated DNA Technologies and were resuspended in nuclease-free water to a concentration of 0.5 µM. For each gene, the probe was designed based on the split-probe design of HCR v.3.0[76] using HCR 3.0 Probe Maker v0.3.2[77] with adjacent B1 or B3 amplification sequences (Supplementary Table 7).

For in situ hybridization, *B. schlosseri* samples were incubated in fixation buffer (4% PFA in 1x PBS) overnight at 4 °C, followed by 100% methanol dehydration and storage at −20 °C for at least 24 h. The samples were gradually rehydrated in 75%, 50% and 25% methanol diluted with PBS (5 min for each step) followed by two washes in 1x PBS for 5 and 10 min, respectively. The samples were permeabilized for 30 min at room temperature in detergent solution (1% SDS, 0.5% Tween-20, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8), 150 mM NaCl). The samples were then washed 2 × 5 min with PTw (0.1% Tween-20 in 1x PBS) followed by post-fixation in 4% PFA for 25 min and washed again 3 × 5 min with PTw. The samples were prehybridized in hybridization buffer (Molecular Instruments) for 30 min at 37 °C. The probes were then added to the hybridization buffer at a final concentration of 0.02 µM and the samples were allowed to

hybridize at 37 °C for overnight but no more than 16 h. Following hybridization, the samples were washed 2 × 30 min in probe wash buffer (Molecular instruments) at 37 °C and then for 5 min in 5x SSCT (5x sodium chloride sodium citrate (SSC), 0.1% Tween-20) at room temperature. They were then pre-amplified in amplification buffer (Molecular Instruments) for 30 min. Meanwhile, H1 and H2 components of the HCR hairpins B1 or B3 coupled to either 546 or 647 amplifiers (Molecular Instruments) were incubated separately at 95 °C for 90 s, cooled down to room temperature in the dark for 30 min and then pooled together before being added to the amplification buffer at a final concentration of 30 nM. The amplification was then performed at room temperature for overnight but no more than 16 h. The samples were subsequently washed 2 × 5 min and additional 2 × 30 min in 5x SSCT followed by incubation for 2 h in PBS containing 0.8:500 DAPI. Finally, the samples were washed 3 × 10 min in PBS, transferred to 50% glycerol for at least 1 h and then mounted on a glass slide for imaging.

Imaging was done using Zeiss LSM700 confocal microscope. For each sample, a series of optical sections were taken with a *z*-step interval of 3–5 µm. Multichannel acquisitions were obtained by sequential imaging and the confocal optical sections were processed using ImageJ v.1.54j[78].

Data availability

The sequencing data generated during this study is available in NCBI. The scRNA-Seq data is under accessions: PRJNA1250573, PRJNA1248000, PRJNA1247989 and PRJNA1247985. The new genome is under accession PRJNA1246227. Step-by-step instructions and detailed code to reproduce our data preprocessing and analysis are available in the supplementary information. The new *B. schlosseri* genome browser can be accessed via the following link (user: colony, password: join):
https://share.nautilus.stanford.edu/new_genome/?session=local-2i7_lmeM1tYlc7-rFKDcO

References

1. Boyle, M., Wong, C., Rocha, M. & Jones, D. L. Decline in self-renewal factors contributes to aging of the stem cell niche in the *Drosophila* testis. *Cell Stem Cell* **1**, 470-478 (2007).
2. Cheng, J. et al. Centrosome misorientation reduces stem cell division during ageing. *Nature* **456**, 599-604 (2008).
3. Conboy, I. M. et al. Rejuvenation of aged progenitor cells by exposure to a young systemic environment. *Nature* **433**, 760-764 (2005).
4. Jones, D. L. Aging and the germ line: where mortality and immortality meet. *Stem Cell Rev.* **3**, 192-200 (2007).
5. Schultz, M. B. & Sinclair, D. A. When stem cells grow old: phenotypes and mechanisms of stem cell aging. *Development* **143**, 3-14 (2016).
6. Zhang, X., Ebata, K. T., Robaire, B. & Nagano, M. C. Aging of male germ line stem cells in mice. *Biol. Reprod.* **74**, 119-124 (2006).
7. Weissman, I. L. Stem cells: units of development, units of regeneration, and units in evolution. *Cell* **100**, 157-168 (2000).

8. Weissman, I. L. Stem cells are units of natural selection for tissue formation, for germline development, and in cancer development. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8922-8928 (2015).
9. Stoner, D. S., Rinkevich, B. & Weissman, I. L. Heritable germ and somatic cell lineage competitions in chimeric colonial protochordates. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9148-9153 (1999).
10. Stoner, D. S. & Weissman, I. L. Somatic and germ cell parasitism in a colonial ascidian: Possible role for a highly polymorphic allorecognition system. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15254-15259 (1996).
11. Laird, D. J., De Tomaso, A. W. & Weissman, I. L. Stem cells are units of natural selection in a colonial ascidian. *Cell* **123**, 1351-1360 (2005).
12. Voskoboynik, A. & Weissman, I. L. *Botryllus schlosseri*, an emerging model for the study of aging, stem cells, and mechanisms of regeneration. *Invertebr. Reprod. Dev.* **59**, 33-38 (2015).
13. Ueno, H., Turnbull, B. B. & Weissman, I. L. Two-step oligoclonal development of male germ cells. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 175-180 (2009).
14. Kowarsky, M. et al. Sexual and asexual development: two distinct programs producing the same tunicate. *Cell Rep.* **34**, 108681 (2021).
15. Rinkevich, Y. et al. Repeated, long-term cycling of putative stem cells between niches in a basal chordate. *Dev. Cell* **24**, 76-88 (2013).
16. Voskoboynik, A. et al. Identification of the endostyle as a stem cell niche in a colonial chordate. *Cell Stem Cell* **3**, 456-464 (2008).
17. Rosental, B. et al. Complex mammalian-like haematopoietic system found in a colonial chordate. *Nature* **564**, 425-+ (2018).
18. Nie, X. et al. Single-cell analysis of human testis aging and correlation with elevated body mass index. *Dev. Cell* **57**, 1160-1176 (2022).
19. Zhang, W. et al. A single-cell transcriptomic landscape of mouse testicular aging. *J. Adv. Res.* **53**, 219-234 (2023).
20. Sabbadin, A. & Zaniolo, G. Sexual differentiation and germ-cell transfer in the colonial ascidian *botryllus schlosseri*. *J. Exp. Zool.* **207**, 289-304 (1979).
21. Rodriguez, D., Kassmer, S. H. & De Tomaso, A. W. Gonad development and hermaphroditism in the ascidian *Botryllus schlosseri*. *Molecular Reproduction and Development* **84**, 158-170 (2017).
22. Chiquoine, A. D. The identification, origin, and migration of the primordial germ cells in the mouse embryo. *Anat. Rec.* **118**, 135-146 (1954).
23. Storms, R. W. et al. Isolation of primitive human hematopoietic progenitors on the basis of aldehyde dehydrogenase activity. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9118-9123 (1999).
24. Jones, R. J. et al. Assessment of aldehyde dehydrogenase in viable cells. *Blood* **85**, 2742-2746 (1995).
25. Talice, S. et al. Candidate stem cell isolation and transplantation in Hexacorallia. *Cell Rep.* **43**, 114944 (2024).
26. Hagemann-Jensen, M. et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat. Biotechnol.* **38**, 708-714 (2020).
27. Lauzon, R. J., Ishizuka, K. J. & Weissman, I. L. Cyclical generation and degeneration of organs in a colonial urochordate involves crosstalk between old and new: a model for development and regeneration. *Dev. Biol.* **249**, 333-348 (2002).

28. Voskoboynik, A. et al. The genome sequence of the colonial chordate, *Botryllus schlosseri*. *Elife* **2** (2013).
29. Castrillon, D. H., Quade, B. J., Wang, T. Y., Quigley, C. & Crum, C. P. The human *VASA* gene is specifically expressed in the germ cell lineage. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 9585-9590 (2000).
30. Moses, E. et al. The killifish germline regulates longevity and somatic repair in a sex-specific manner. *Nat. Aging* **4**, 791-813 (2024).
31. Deng, W. & Lin, H. *miwi*, a murine homolog of *piwi*, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell* **2**, 819-830 (2002).
32. Cox, D. N., Chao, A. & Lin, H. *piwi* encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development* **127**, 503-514 (2000).
33. Ozaki, Y., Saito, K., Shinya, M., Kawasaki, T. & Sakai, N. Evaluation of *Sycp3*, *Plzf* and *Cyclin B3* expression and suitability as spermatogonia and spermatocyte markers in zebrafish. *Gene Expr. Patterns* **11**, 309-315 (2011).
34. Cavalcanti, M. C. et al. Expression of histone 1 (H1) and testis-specific histone 1 (H1t) genes during stallion spermatogenesis. *Anim. Reprod. Sci.* **111**, 220-234 (2009).
35. Bucci, L. R., Brock, W. A. & Meistrich, M. L. Distribution and synthesis of histone 1 subfractions during spermatogenesis in the rat. *Exp. Cell Res.* **140**, 111-118 (1982).
36. Karlsson, M. et al. A single-cell type transcriptomics map of human tissues. *Sci. Adv.* **7** (2021).
37. Uhlen, M. et al. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
38. Huang, L., Haratake, K., Miyahara, H. & Chiba, T. Proteasome activators, PA28gamma and PA200, play indispensable roles in male fertility. *Sci. Rep.* **6**, 23171 (2016).
39. Lv, P., Liu, J. & Liu, X. The role of ubiquitin-conjugating enzyme in the process of spermatogenesis. *Reprod. Biol. Endocrinol.* **22**, 110 (2024).
40. Leung, M. R. et al. Structural specializations of the sperm tail. *Cell* **186**, 2880-2896 e2817 (2023).
41. Merges, G. E. et al. Loss of *Prm1* leads to defective chromatin protamination, impaired PRM2 processing, reduced sperm motility and subfertility in male mice. *Development* **149**, dev200330 (2022).
42. Lea, I. A., Widgren, E. E. & O'Rand, M. G. Association of sperm protein 17 with A-kinase anchoring protein 3 in flagella. *Reprod. Biol. Endocrinol.* **2**, 57 (2004).
43. Salicioni, A. M. et al. Testis-specific serine kinase protein family in male fertility and as targets for non-hormonal male contraception. *Biol. Reprod.* **103**, 264-274 (2020).
44. D'Orazio, F. M. et al. Germ cell differentiation requires Tdrd7-dependent chromatin and transcriptome reprogramming marked by germ plasm relocalization. *Dev. Cell* **56**, 641-656 e645 (2021).
45. Zhao, W. P. et al. Positive PCNA and Ki-67 expression in the testis correlates with spermatogenesis dysfunction in fluoride-treated rats. *Biol. Trace Elem. Res.* **186**, 489-497 (2018).
46. Tsuda, M. et al. Conserved role of nanos proteins in germ cell development. *Science* **301**, 1239-1241 (2003).
47. Yang, L. et al. Argonaute 1 regulates the fate of germline stem cells in *Drosophila*. *Development* **134**, 4265-4272 (2007).

48. Voskoboynik, A. et al. Identification of a colonial chordate histocompatibility gene. *Science* **341**, 384-387 (2013).
49. Sabbadin, A. & Astorri, C. Chimeras and histocompatibility in the colonial ascidian *Botryllus schlosseri*. *Dev. Immunol.* **12**, 737-747 (1988).
50. Scofield, V. L., Schlumpberger, J. M., West, L. A. & Weissman, I. L. Protochordate allorecognition is controlled by a MHC-like gene system. *Nature* **295**, 499-502 (1982).
51. Tolkin, T. & Hubbard, E. J. A. Germline stem and progenitor cell aging in *C. elegans*. *Front. Cell Dev. Biol.* **9**, 699671 (2021).
52. Sharpless, N. E. & DePinho, R. A. How stem cells age and why this makes us grow old. *Nat. Rev. Mol. Cell Biol.* **8**, 703-713 (2007).
53. Xia, K. et al. Single-cell RNA sequencing reveals transcriptomic landscape and potential targets for human testicular ageing. *Hum. Reprod.* **39**, 2189-2209 (2024).
54. Stewart-Savage, J. & Yund, P. O. Temporal pattern of sperm release from the colonial ascidian, *Botryllus schlosseri*. *J. Exp. Zool.* **279**, 620-625 (1997).
55. Spichal, M. et al. Germ granule dysfunction is a hallmark and mirror of Piwi mutant sterility. *Nat. Commun.* **12**, 1420 (2021).
56. Houwing, S. et al. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **129**, 69-82 (2007).
57. Brown, F. D. & Swalla, B. J. Vasa expression in a colonial ascidian, Botrylloides violaceus. *Evol. Dev.* **9**, 165-177 (2007).
58. Rosner, A., Moiseeva, E., Rinkevich, Y., Lapidot, Z. & Rinkevich, B. Vasa and the germ line lineage in a colonial urochordate. *Dev. Biol.* **331**, 113-128 (2009).
59. Sunanaga, T., Inubushi, H. & Kawamura, K. Piwi-expressing hemoblasts serve as germline stem cells during postembryonic germ cell specification in colonial ascidian, *Botryllus primigenus*. *Dev. Growth Differ.* **52**, 603-614 (2010).
60. Cui, L. et al. Single-cell transcriptomic atlas of the human testis across the reproductive lifespan. *Nat. Aging* **5**, 658-674 (2025).
61. Delsuc, F., Brinkmann, H., Chourrout, D. & Philippe, H. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **439**, 965-968 (2006).
62. Boyd, H. C., Brown, S. K., Harp, J. A. & Weissman, I. L. Growth and sexual maturation of laboratory-cultured Monterey *Botryllus schlosseri*. *Biol. Bull.* **170**, 91-109 (1986).
63. Liu, D. D. et al. Purification and characterization of human neural stem and progenitor cells. *Cell* **186**, 1179-1194 e1115 (2023).
64. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
65. Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E. & Zanini, F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* **38**, 2943-2945 (2022).
66. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
67. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
68. Ben-Ari Fuchs, S. et al. GeneAnalytics: an integrative gene set analysis tool for next generation sequencing, RNAseq and microarray data. *OMICS J. Integr. Biol.* **20**, 139-151 (2016).
69. Hoss, M. & Paabo, S. DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Res.* **21**, 3913-3914 (1993).

70. Anselmi, C. et al. Two distinct evolutionary conserved neural degeneration pathways characterized in a colonial chordate. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2203032119 (2022).
71. Sklarz, M. Y., Levin, L., Gordon, M. & Chalifa-Caspi, V. Neatseq-flow: A lightweight high throughput sequencing workflow platform for non-programmers and programmers alike. *bioRxiv*, 173005 (2018).
72. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644-652 (2011).
73. Palmer, J. M. & Stajich, J. Funannotate v1.8.1: Eukaryotic genome annotation (v1.8.1). *Zenodo*, https://doi.org/10.5281/zenodo.1134477 (2020).
74. Haas, B. J. et al. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654-5666 (2003).
75. Stanke, M. et al. AUGUSTUS: *ab initio* prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435-439 (2006).
76. Choi, H. M. T. et al. Third-generation *in situ* hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145** (2018).
77. Kuehn, E. et al. Segment number threshold determines juvenile onset of germline cluster expansion in *Platynereis dumerilii*. *J. Exp. Zoolog. B Mol. Dev. Evol.* **338**, 225-240 (2022).
78. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675 (2012).

Acknowledgements
We thank Federico Caicci, Imaging Facility at Dept. Biol. University of Padova, for technical support in preparing 3D reconstructions; Noah Gordon for assembling the new *B. schlosseri* genome browser and server; Wan-Jin Lu and Stuart Thompson for the helpful and valuable discussions; Christopher Lowe for valuable discussions and for providing resources and support allowing the execution of this project; Linda Quinn and Tejaswitha Naik for laboratory management and support; the Chan Zuckerberg BioHub (CZ Biohub) for long and short-read sequencing processing and expertise; the Stanford Protein and Nucleic Acid (PAN) Facility for technical support. T.L. was supported by the Gruss Lipper Postdoctoral Research Fellowship. This study was supported by NIH 5R01AG076908 grant to A.V. and I.L.W.; Big Ideas for Oceans grant from the Stanford Oceans Department and Stanford Woods Institute for the Environment to A.V.; Stanford Bio-X Seed Grant to A.V.; Wu Tsai Human Performance Alliance (WTHPA) to Y.V. and D.S.

Author information
These authors contributed equally and share last authorship: Irving L. Weissman, Ayelet Voskoboynik

Authors and Affiliations
Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA

Tom Levy, Chiara Anselmi, Katherine J. Ishizuka, Tal Gordon, Karla J. Palmeri, Daniel Dan Liu, Rahul Sinha, Benjamin F. Ohene-Gambill, Tal Raveh, Irving L. Weissman & Ayelet Voskoboynik

Department of Biology, Hopkins Marine Station, Stanford University, Pacific Grove, CA, USA
Tom Levy, Chiara Anselmi, Katherine J. Ishizuka, Tal Gordon, Karla J. Palmeri, Irving L. Weissman & Ayelet Voskoboynik

Department of Bioinformatics and System Biology, Jacobs School of Engineering, University of California San Diego, San Diego, CA, USA
Yotam Voskoboynik & Debashis Sahoo

Chan Zuckerberg Biohub, San Francisco, CA, USA
Erin McGeever, Angela M. Detweiler, Maurizio Morri & Norma F. Neff

Bioinformatics Core Facility, Ilse Katz Institute for Nanoscale Science & Technology, Ben-Gurion University of the Negev, Beer Sheva, Israel
Liron Levin

Dipartimento di Biologia, Università degli Studi di Padova, 35131 Padova, Italy
Virginia Vanni & Lucia Manni

Department of Pediatrics, University of California San Diego, CA, USA
Debashis Sahoo

The Shraga Segal Department of Microbiology, Immunology, and Genetics, Faculty of Health Sciences. Center for Regenerative Medicine and Stem Cells. Ben Gurion University of the Negev, Beer Sheva, Israel
Benyamin Rosental

Ludwig Center for Cancer Stem Cell Research and Medicine, Stanford University School of Medicine, Stanford, CA, USA
Irving L. Weissman

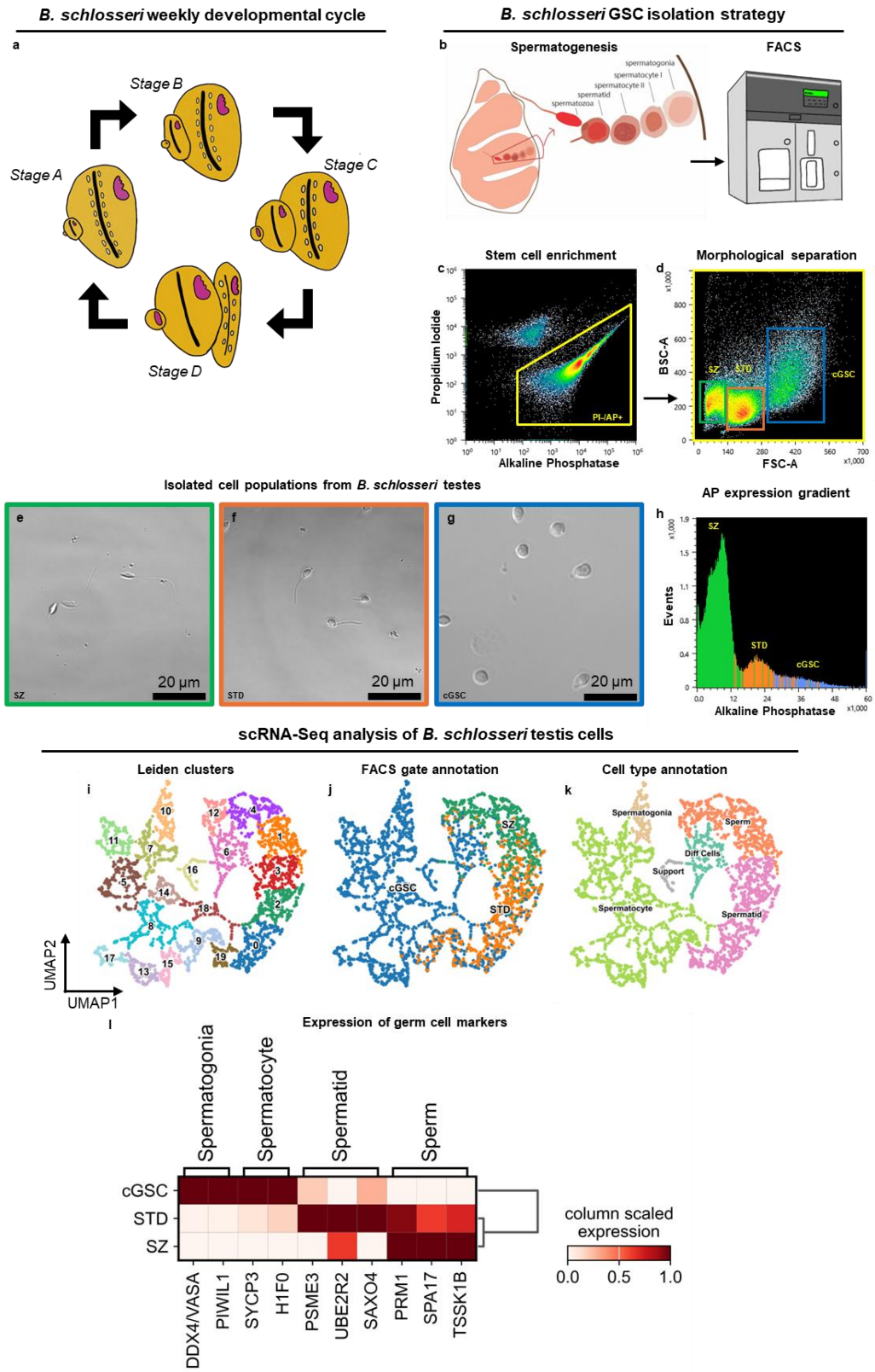
Department of Pathology, Stanford University, Stanford, CA, USA
Irving L. Weissman

Contributions

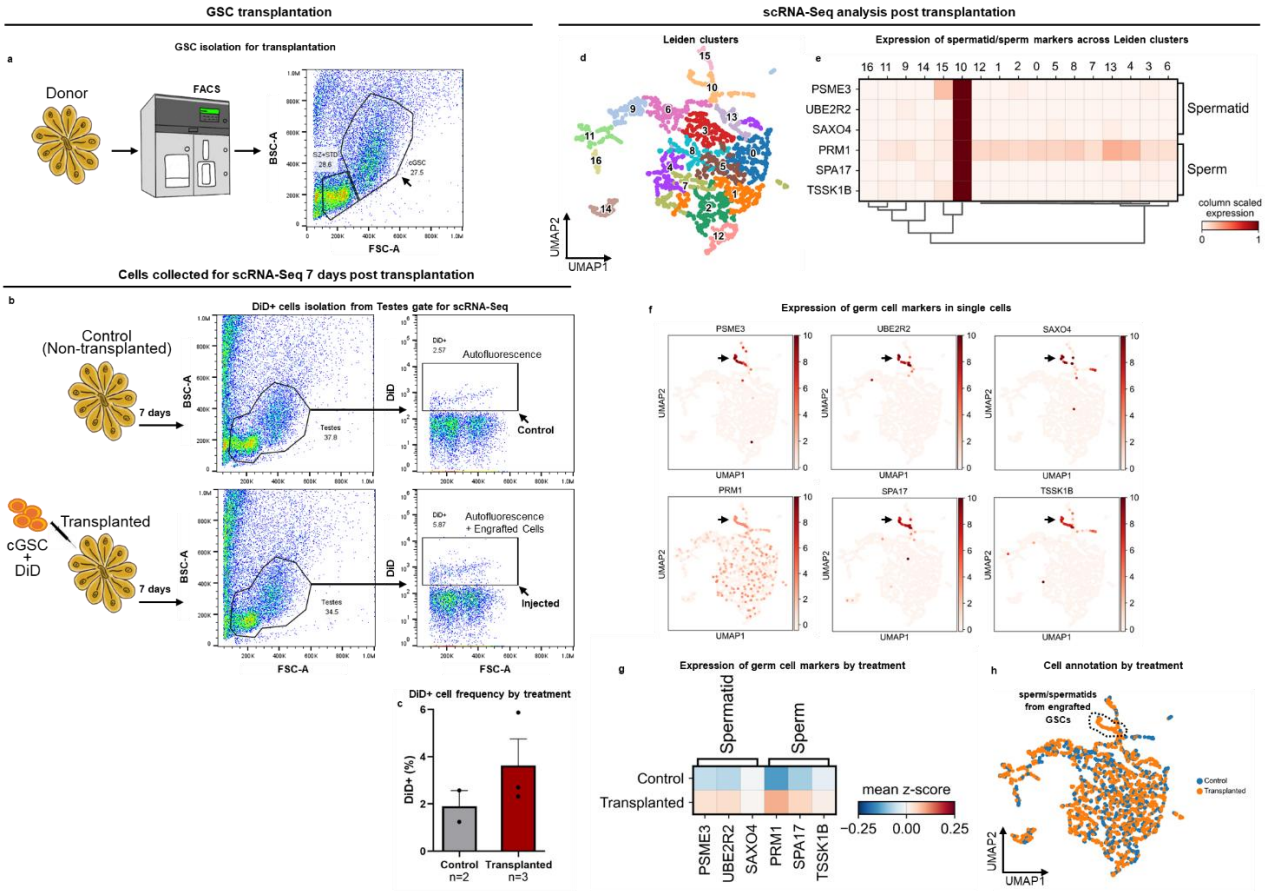
Conception and design: T.L., A.V. and I.L.W.; mariculture: K.J.I. and K.J.P.; flow cytometry and sorting: T.L.; confocal microscopy: T.L.; scRNA isolation and library preparation: T.L., C.A., T.G., D.D.L., R.S., B.F.O.G.; gDNA isolation: K.J.P.; gDNA and scRNA sequencing: A.M.D, M.M., N.F.N.; genome assembly: A.M.D.; transplantation assays: T.L. and K.J.I.; in situ hybridization: T.L., T.G. and K.J.I.; histology and 3D reconstructions: V.V. and L.M.; sequencing analysis and bioinformatics: T.L., Y.V., E.M., L.L., D.D.L. and D.S.; writing of manuscript: T.L., T.R., B.R., A.V. and I.L.W.; all of the authors reviewed, edited and approved the manuscript. A.V. and I.L.W. contributed equally to this work and share senior authorship.

Corresponding authors
Correspondence to Tom Levy, Irving L. Weissman and Ayelet Voskoboynik

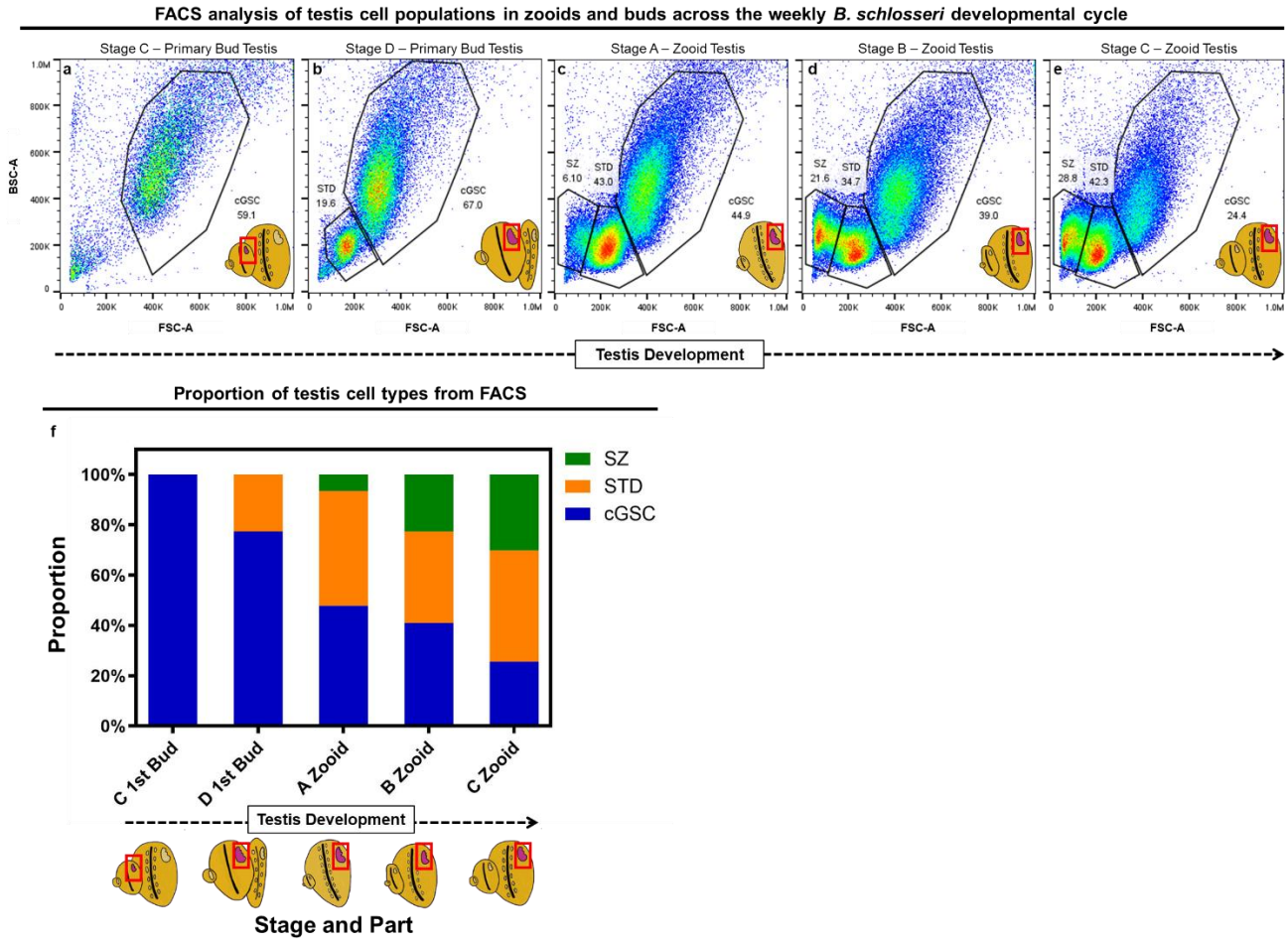
Ethics declarations
The authors declare no competing interests.



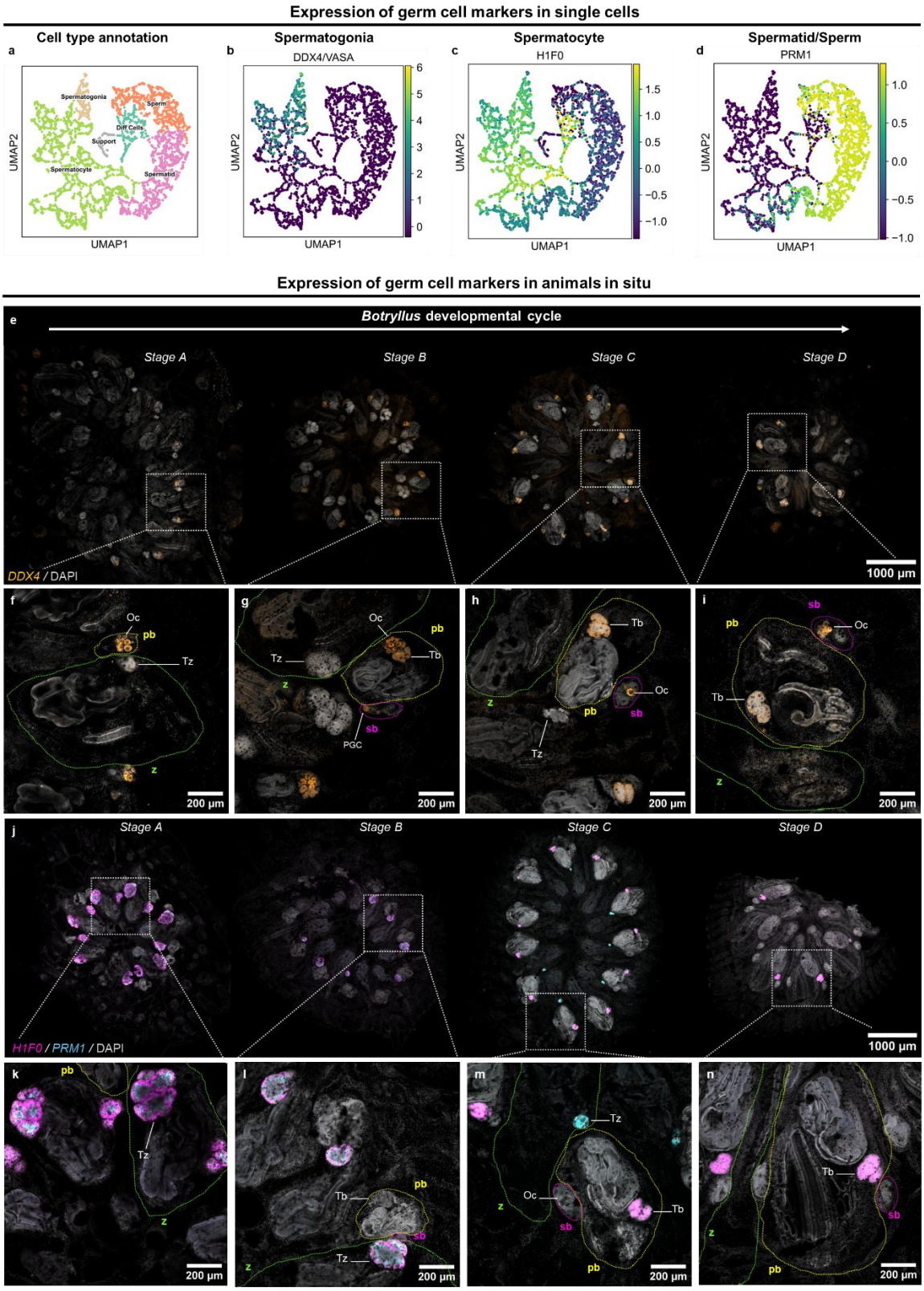
Levy et al., Fig. 1 – High-resolution identification of spermatogenic germline cell types using FACS and scRNA-Seq.



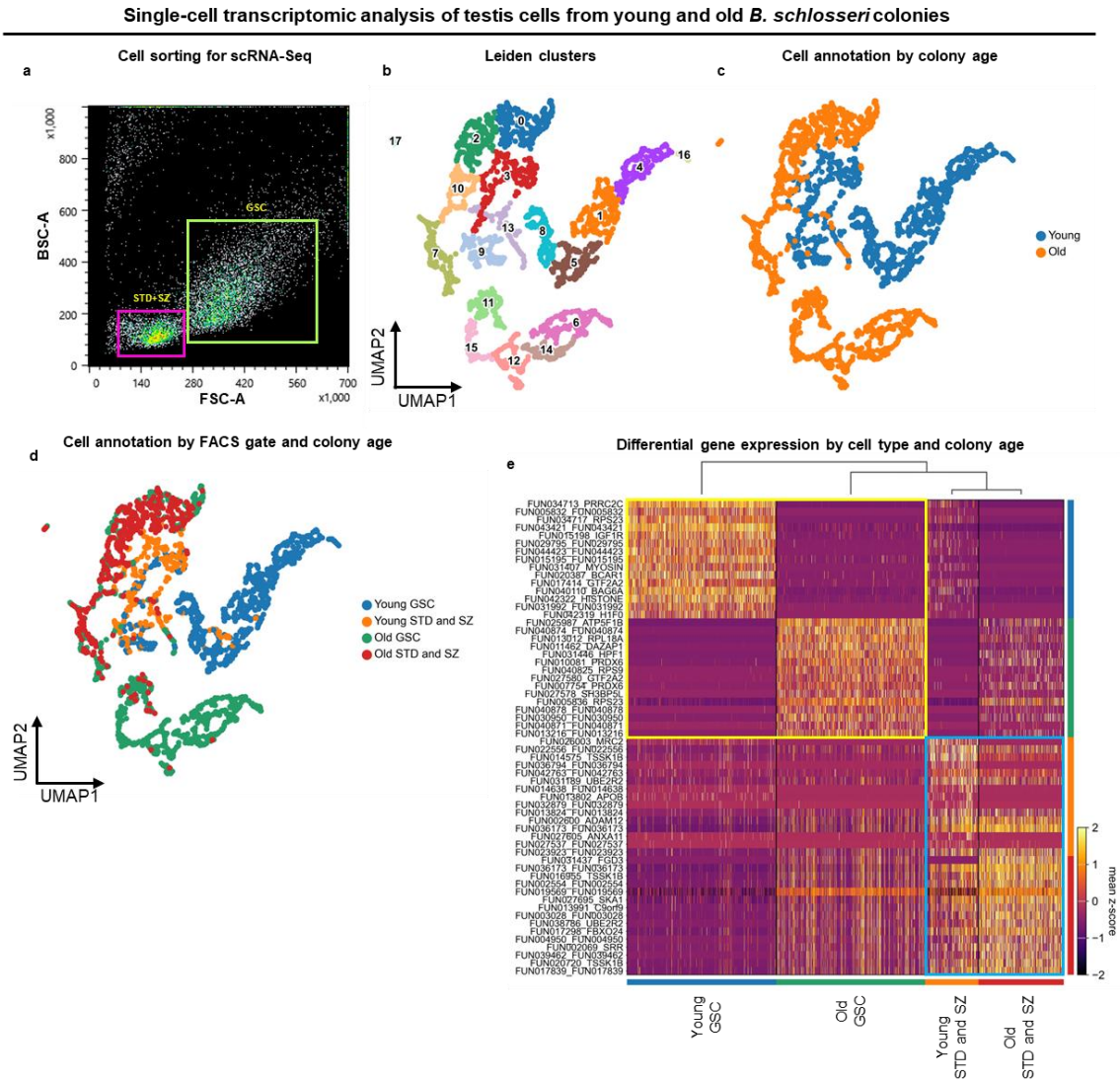
Levy et al., Fig. 2 - Homing and differentiation of transplanted candidate GSCs revealed by in vivo tracing and scRNA-Seq.



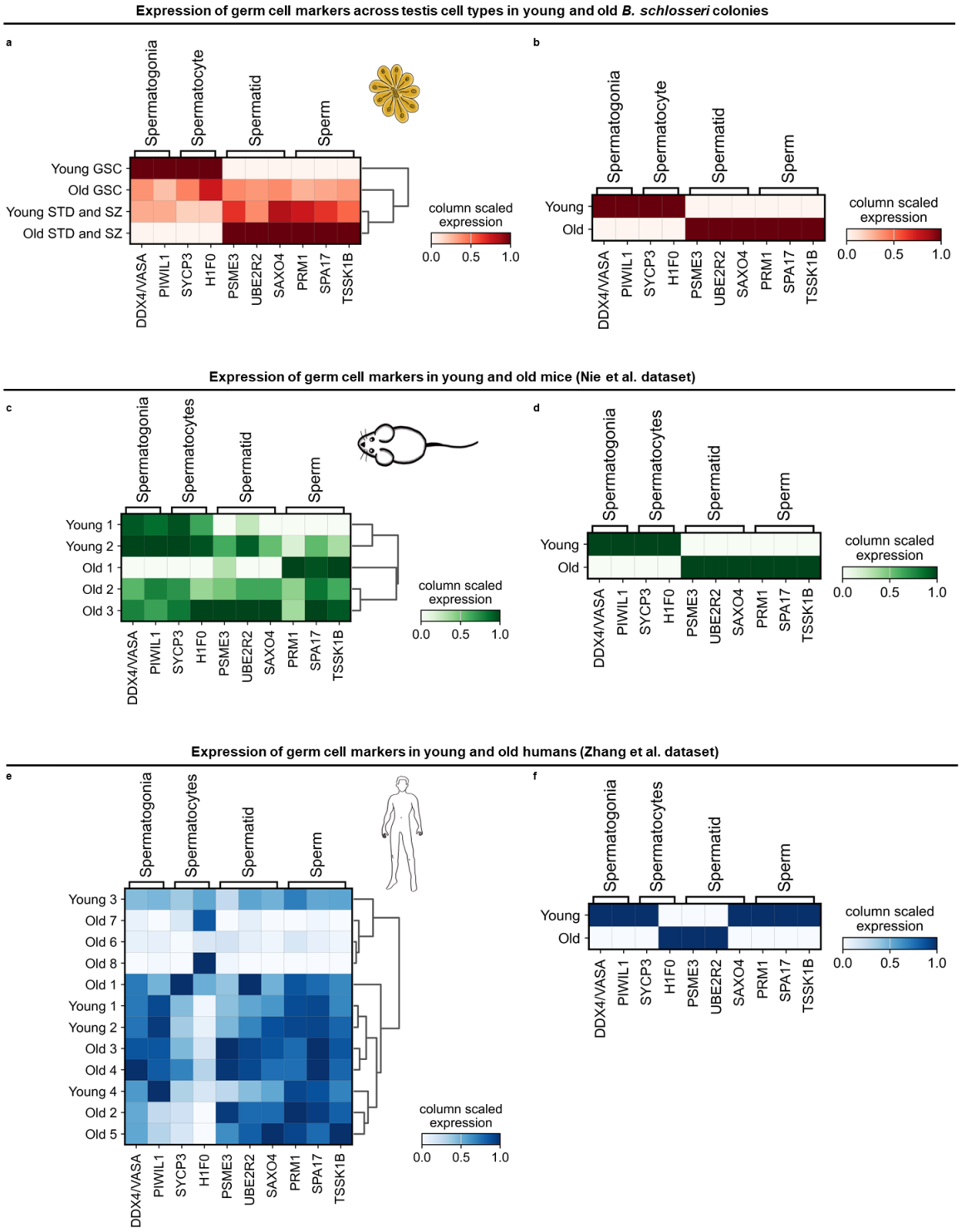
Levy et al., Fig. 3 - Testis development in *B. schlosseri* resolved at the single-cell level.



Levy et al., Fig. 4 - Spatiotemporal expression of germ cell markers during gonadal development.

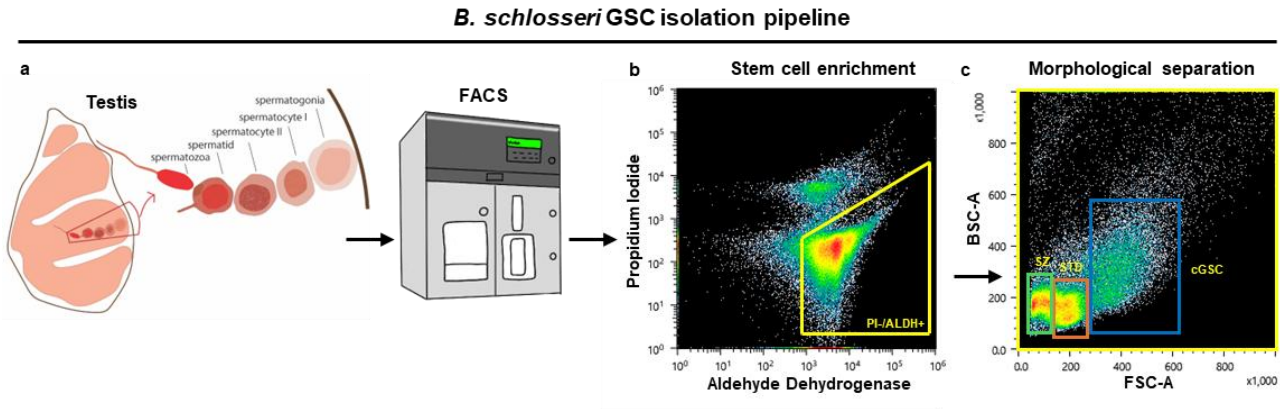


Levy et al., Fig. 5 - Germline stem cell transcriptomic alterations govern reproductive aging.

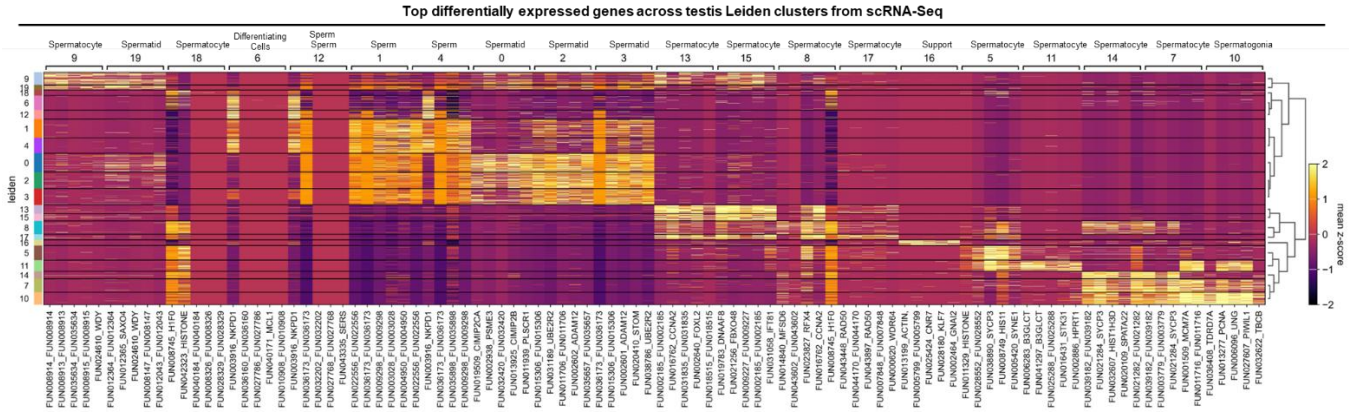


Levy et al., Fig. 6 - Conserved germline stem cell marker profiles underlie reproductive aging across chordate evolution.

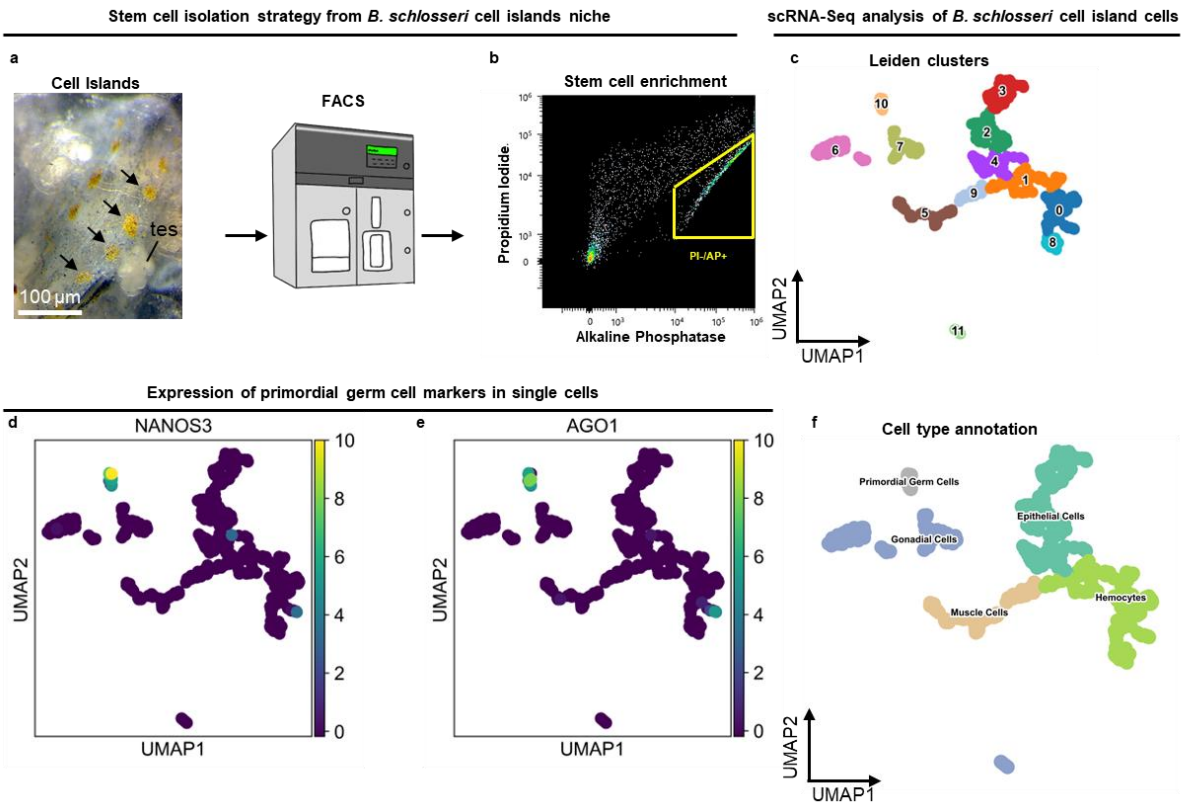
Extended Data Figures



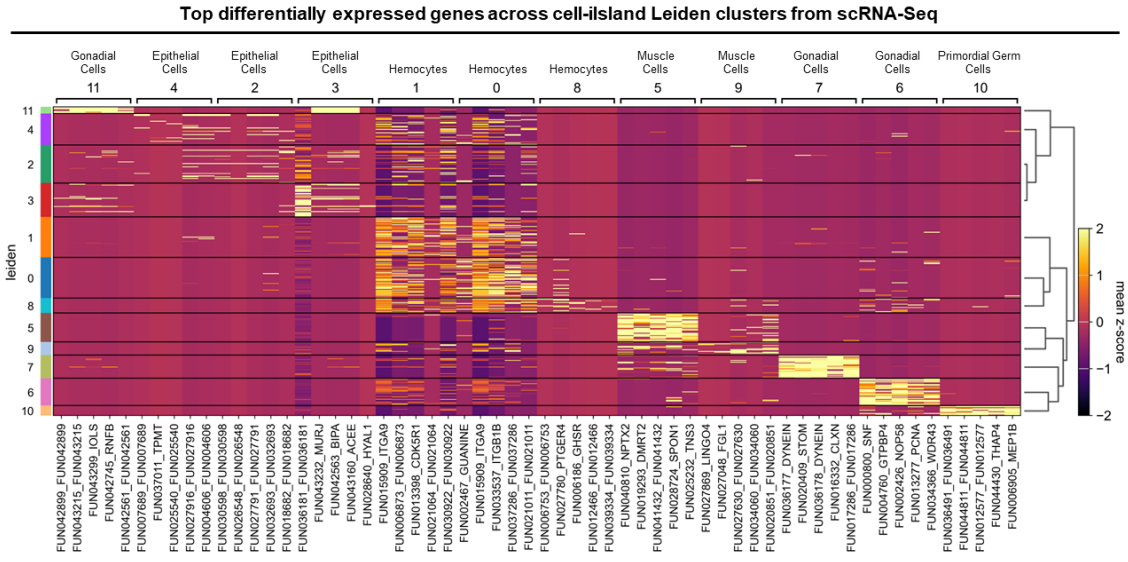
Levy et al., Extended Data Fig. 1 - ALDH stains the same cell populations as AP in the testis of *B. schlosseri.*



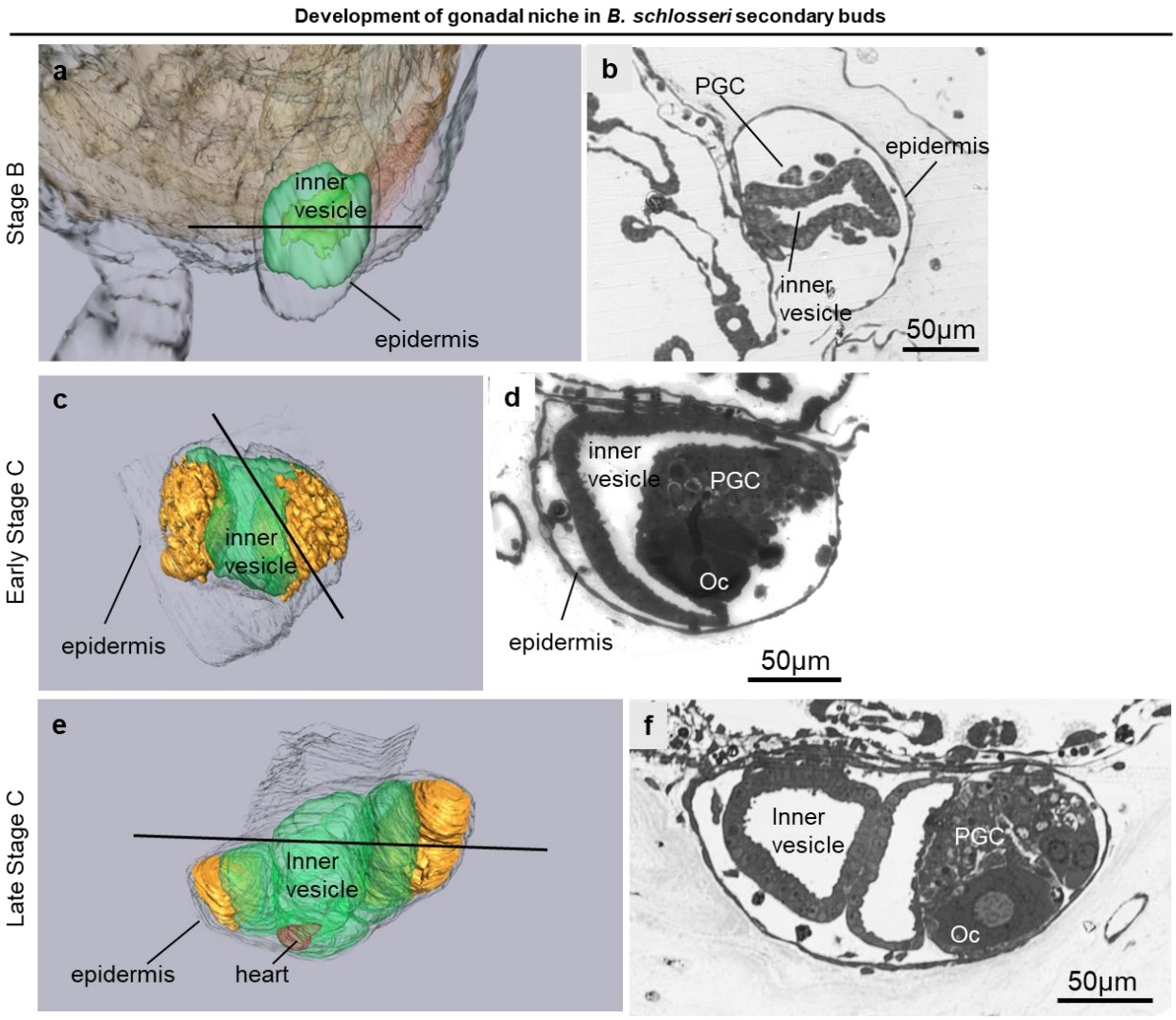
Levy et al., Extended Data Fig. 2 – Top differentially expressed genes across Leiden clusters in *B. schlosseri* testis.



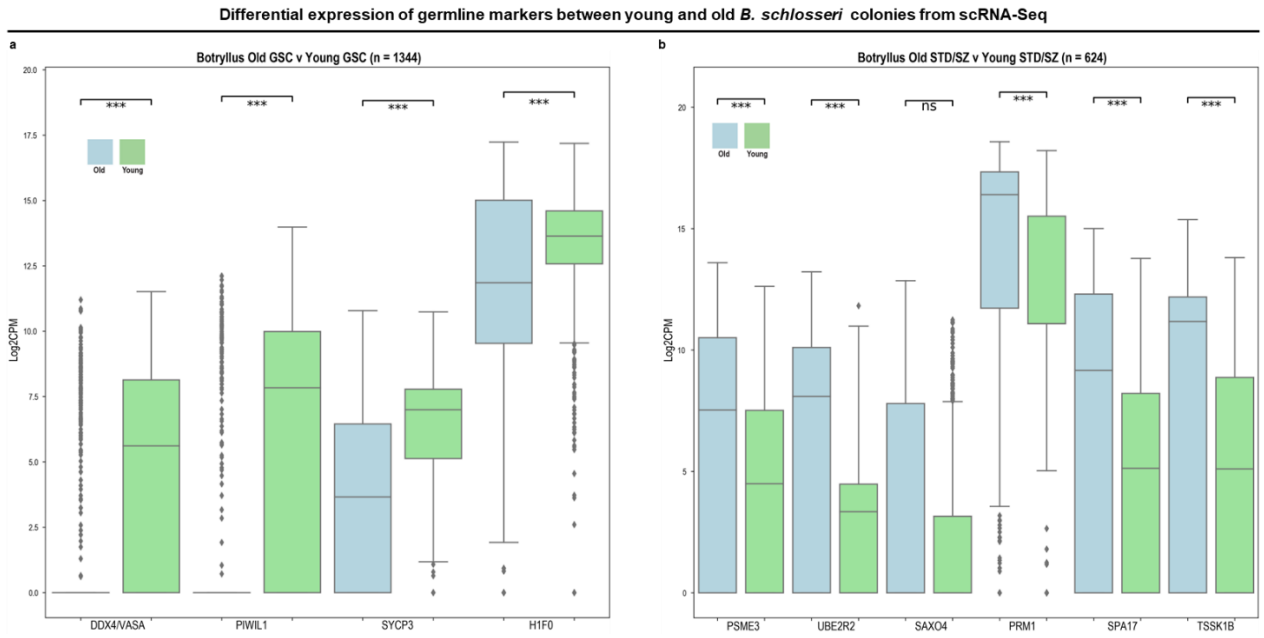
Levy et al., Extended Data Fig. 3 – scRNA-Seq reveals primordial germ cells in the cell islands niche.



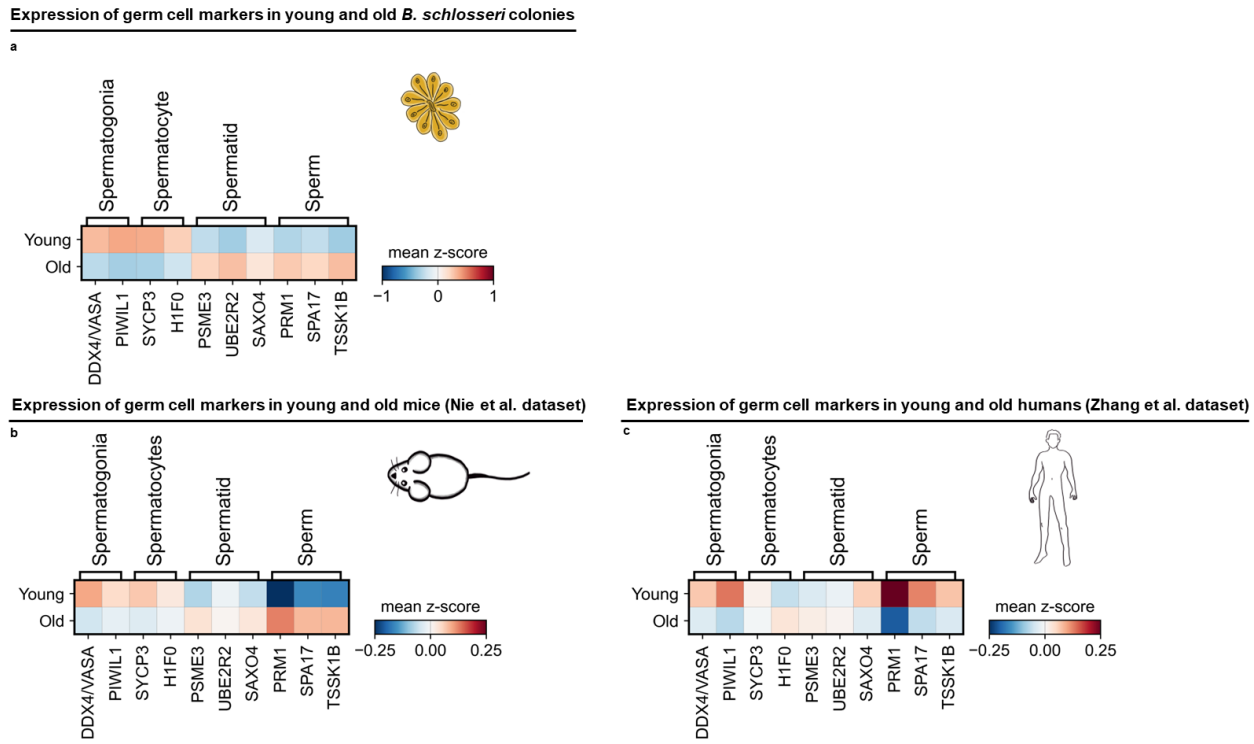
Levy et al., Extended Data Fig. 4 – Top differentially expressed genes across Leiden clusters in *B. schlosseri* cell islands niche.



Levy et al., Extended Data Fig. 5 – Migration of primordial germ cells for gonadal development starts at the secondary bud.



Levy et al., Extended Data Fig. 6 – Age-related differential expression of gene markers between different cell types in *B. schlosseri* testis.



Levy et al., Extended Data Fig. 7 – Gene expression (mean z-scores) in young and old samples.

Figure Legends

Fig. 1: High-resolution identification of spermatogenic germline cell types using FACS and scRNA-Seq.

a, Schematic illustration of the *B. schlosseri* weekly developmental cycle. **b,** Cell types identified during spermatogenesis. Spermatogonia are located at the lobe's periphery, followed by primary and secondary spermatocytes and spermatids (STD). Fully mature sperm (spermatozoa; SZ) are found in the lobe's center. **c,** AP+/PI- cells from *B. schlosseri* testes were analyzed using FACS based on **d,** size (x-axis) and granularity (y-axis). **e-h,** Images of cell populations isolated from *B. schlosseri* testes according to the FACS gates set in **d**. Sperm cells from the "SZ" gate **(e, green)**, spermatids from the "STD" gate **(f, orange)**, morphologically undefined cells from the candidate GSC "cGSC" gate **(g, blue)**, and a histogram showing the AP activity gradient in cells from those gates **(h)**. **i-k,** scRNA-Seq of index-sorted AP+/PI- testes cells using Smart-seq3. **i,** UMAP plot showing the numbered leiden clusters. **j,** UMAP plot with annotated clusters based on the FACS gate by which the cells were sorted. **k,** UMAP plot with annotated cell type identities based on expressed transcripts of known genes (Supplementary Table 2, Extended Data Fig. 2). **l,** Matrix plot showing marker gene expression for transcriptomically defined clusters from scRNA-Seq and the FACS gate they were sorted from.

Fig. 2: Homing and differentiation of transplanted candidate GSCs revealed by in vivo tracing and scRNA-Seq.

a, ALDH+ cGSC cells (black arrow) were sorted, based on our gating strategy, from a whole animal using FACS and transplanted into recipient colonies. The cells from the cGSC gate were labeled with DiD prior to transplantation to recipients (n=3) while control animals were not transplanted (n=2). **b,** Seven days post transplantation (after one regeneration cycle and takeover) DiD+ cells (black arrows) from the Testes gates (cGSC+STD+SZ) applied on cells coming from a whole-animal cell suspension were index-sorted for scRNA-Seq, using Smart-seq3, from transplanted (n=3) and non-transplanted control (n=2), which were at the same stage and had visible testes. In the control animals, the cells from the DiD+ gate displayed natural autofluorescence, while in the transplanted animals, this gate contained both naturally autofluorescent cells and engrafted cells. **c,** The proportion of DiD+ cells was higher in transplanted compared to non-transplanted control animals. **d,** UMAP plot showing the Leiden clusters with cluster numbers from scRNA-Seq of DiD+ index-sorted cells from both transplanted and control animals. **e,** Matrix plot of spermatid and sperm marker gene expression (*PSME3*, *UBE2R2*, *SAXO4*, *PRM1*, *SPA17*, *TSSK1B*) for each Leiden cluster showing the highest expression in Cluster 10. **f,** UMAP expression plots of the above gene markers showed high expression in the cells constituting the left branch of Leiden Cluster 10 (black arrows). **g,** Matrix plot showing higher expression of the above markers in transplanted compared to control animals. **h,** UMAP plot with cells annotated by treatment (control / transplanted). Sperm and spermatids from engrafted GSCs, based on **d-g**, are labeled with dotted line.

Fig. 3: Testis development in *B. schlosseri* resolved at the single-cell level.
a-f, FACS analysis of testis cells throughout development. AP+/PI- cells from *B. schlosseri* testes from primary buds and adult zooids in different developmental stages were analyzed based on size (x-axis) and granularity (y-axis) using FACS. Testis (see red frame) cells from the primary bud of a stage C colony **(a)** and stage D colony **(b)**. Testis cells from the zooid of a stage A colony **(c)**, stage B colony **(d)** and stage C colony **(e)**. Cell fractions in each sorting gate are indicated. **f,** Proportions of the different sorted cell types at each stage from FACS analysis.

Fig. 4: Spatiotemporal expression of germ cell markers during gonadal development.
a-d, scRNA-Seq expression analysis of germ cell markers. UMAP expression plots of gene markers for different cell types that were used for in situ expression analysis (HCR). **b,** *DDX4* – spermatogonia. **c,** *H1F0* - spermatocytes. **d,** *PRM1* – spermatids and sperm. **e-n,** Germ cell marker expression during the *B. schlosseri* developmental cycle. **e,** HCR images showing *DDX4* (orange) expression which begins in the developing gonad of the secondary bud in stage B. The gray frames in **e** mark the close-ups shown in **f-i**. Close-up images of a colony in stage A **(f)**, stage B **(g)**, stage C **(h)** and stage D **(i)**. **j,** HCR images showing *H1F0* (magenta) and *PRM1* (cyan) expression which begins in the developing gonad of the primary bud in stage C and the zooid in stage A, respectively. The gray frames in **j** mark the close-ups shown in **k-n**. Close-up images of a colony in stage A **(k)**, stage B **(l)**, stage C **(m)** and stage D **(n)**. Specimens are counterstained with DAPI (gray). The zooid (z), primary bud (pb) and secondary bud (sb) are outlined with green, yellow and magenta frames, respectively. The primordial germ cells (PGC), developing oocytes (Oc), zooid's testis (Tz) and primary bud's testis (Tb) are marked.

Fig. 5: Germline stem cell transcriptomic alterations govern reproductive aging.
a-e, scRNA-Seq of index-sorted testes cells from young and old *B. schlosseri* colonies using Smart-seq3. **a,** ALDH+/PI- cells from the GSC and STD+SZ gates were sorted using FACS based on size (x-axis) and granularity (y-axis). **b,** UMAP plot showing the Leiden clusters with cluster numbers. **c,** UMAP plot with annotated clusters based on the age of the animal from which the cells were sorted. **d,** UMAP plot with annotated clusters based on the FACS gate and the age of the animal from which the cells were sorted reveals distinct cell clusters of young (blue) and old (green) GSCs, as well as young (orange) and old (red) sperm and spermatids. **e,** A heatmap plot displaying the top 15 differentially expressed genes (based on Supplementary Table 6) in young and old GSCs (yellow frame) and young and old sperm and spermatids (cyan frame) further illustrates the higher aging-related transcriptomic variation within the GSC population compared to spermatids and sperm.

Fig. 6: Conserved germline stem cell marker profiles underlie reproductive aging across chordate evolution.
Matrix plot showing marker gene expression for defined cell types (*DDX4*, *PIWIL1*, *SYCP3*, *H1F0*, *PSME3*, *UBE2R2*, *SAXO4*, *PRM1*, *SPA17*, *TSSK1B* – details in Supplementary Table 4) from scRNA-Seq in cells according to the animal's age and FACS gate they were sorted from **(a)**

and according to the animal's age only **(b)** in *B. schlosseri*. Matrix plot showing gene expression of the same markers from scRNA-Seq in cells according to the animal's age and the individual sample they were isolated from and according to the animal's age only **(c,d)** in *M. musculus* (10X Genomics data from Zhang et al., 2023) and **(e,f)** in *H. sapiens* (10X Genomics data from Nie et al., 2022). Matrix plots with z-scores are given in Extended Data Fig. 7.

Extended Data Figures

Extended Data Fig. 1: ALDH stains the same cell populations as AP in the testis of *B. schlosseri*.

a, Cell types present during spermatogenesis. Spermatogonia are located at the lobe's periphery, followed by primary and secondary spermatocytes and spermatids (STD). Fully mature sperm (spermatozoa; SZ) are found in the lobe's center. **b,** Similar to AP+/PI- (Fig. 1), ALDH+/PI- cells from *B. schlosseri* testes were analyzed using FACS based on **c,** size (x-axis) and granularity (y-axis) and showed similar results.

Extended Data Fig. 2: Top differentially expressed genes across Leiden clusters in *B. schlosseri* testis.

Heatmap showing the top 5 differentially expressed genes across Leiden cell clusters according to scRNA-Seq of *B. schlosseri* testis. The rank is according to Wilcoxon rank-sum test. Leiden numbers and annotations are indicated. Data is correlated with Supplementary Table 2.

Extended Data Fig. 3: scRNA-Seq reveals primordial germ cells in the cell islands niche.

a-c, scRNA-Seq of index-sorted AP+/PI- cell-islands cells using Smart-seq3. **a,** Representative photo of an adult zooid. Cell-islands are labeled with black arrows and the testis (tes) is also visible. **b,** The AP+/PI- cell population that was sorted for sequencing. **c,** UMAP plot showing the Leiden clusters with cluster numbers. **d-f,** Cell type annotation. UMAP expression plots of primordial germ cell markers: *NANOS3* **(d)** and *AGO1* **(e). f,** UMAP plot with annotated cell type identities based on expressed transcripts of known genes (Supplementary Table 4, Supplementary Table 5, Extended Data Fig. 4).

Extended Data Fig. 4: Top differentially expressed genes across Leiden clusters in *B. schlosseri* cell islands niche.

Heatmap showing the top 5 differentially expressed genes across Leiden cell clusters according to scRNA-Seq of *B. schlosseri* cell islands niche. The rank is according to Wilcoxon rank-sum test. Leiden numbers and annotations are indicated. Data is correlated with Supplementary Table 5.

Extended Data Fig. 5: Migration of primordial germ cells for gonadal development starts at the secondary bud.

a-e, 3D reconstructions and cross sections of secondary buds with its inner vesicle (green) and developing gonad (orange) in a stage B colony **(a,b)**, an early-stage C colony **(c,d)** and a late-stage C colony **(e,f)**. The black line indicates the level of the section shown to the right of each

3D reconstruction. The bud's primordial germ cells (PGC), developing oocytes (Oc), inner vesicle, epidermis and heart are marked.

Extended Data Fig. 6: Age-related differential expression of gene markers between different cell types in *B. schlosseri* testis.
Box plots showing differential marker gene expression for defined cell types from scRNA-Seq in cells according to the animal's age (t-test, ***p-value ≤ 0.01). **a,** *DDX4*, *PIWIL1*, *SYCP3*, *H1F0* are highly expressed in young (n=672) compared to old (n=672) GSCs. **b,** *PSME3*, *UBE2R2*, *PRM1*, *SPA17*, *TSSK1B* are highly expressed in old (n=384) compared to young (n=240) spermatids and sperm while *SAXO4* is not statistically significant. Data is correlated with Fig. 6a.

Extended Data Fig. 7: Gene expression (mean z-scores) in young and old samples.
Marker gene expression (mean z-score) for defined cell types (*DDX4*, *PIWIL1*, *SYCP3*, *H1F0*, *PSME3*, *UBE2R2*, *SAXO4*, *PRM1*, *SPA17*, *TSSK1B* – details in Supplementary Table 4) from scRNA-Seq in cells according to the animal's age in **(a)** *B. schlosseri*, **(b)** *M. musculus* (10X Genomics data from Zhang et al., 2023) and **(c)** in *H. sapiens* (10X Genomics data from Nie et al., 2022). Data is correlated with Fig. 6b, d, f.

Supplementary Information

Supplementary Table 1: Definition of stages along the developmental cycle of *B. schlosseri*.
Stage definition is according to Lauzon et al., 2002.

Supplementary Table 2: Rank of gene expression for each Leiden cluster and annotated cell types from scRNA-Seq of *B. schlosseri* testis.
Based on Wilcoxon Rank Sum Test. Data is correlated with Extended Data Fig. 2.

Supplementary Table 3: *B. schlosseri* new genome assembly statistics.
The genome was sequenced using the PacBio Sequel IIe platform. Sample setup using the "HiFi" application, circular consensus sequencing (CCS), and genome assembly were performed in SMRTLink (v. 11.0).

Supplementary Table 4: Key germline marker genes used in this study
Listed are gene names, aliases, full gene names, the main cell type in which each gene is expressed, proposed functions, and literature references supporting their use as germline markers in this study.

Supplementary Table 5: Rank of gene expression for each Leiden cluster and annotated cell types from scRNA-Seq of *B. schlosseri* cell islands.
Based on Wilcoxon Rank Sum Test. Data is correlated with Extended Data Fig. 4.

Supplementary Table 6: Rank of differentially expressed genes from scRNA-Seq of young and old GSCs and sperm/spermatids from *B. schlosseri* testis.
Based on Wilcoxon Rank Sum Test. Data is correlated with Fig. 5, Fig. 6, Extended Data Fig. 6 and Extended Data Fig. 7.

Supplementary Table 7: Short antisense DNA probes used for in situ hybridization in this study

Probes were designed with adjacent B1 or B3 amplification sequences.