Genomics, Proteomics & Bioinformatics, 2024, **22(3)**, qzae005
https://doi.org/10.1093/gpbjnl/qzae005
Advance access publication: 10 January 2024
Database

SMARTdb: An Integrated Database for Exploring Single-cell Multi-omics Data of Reproductive Medicine

Zekai Liu [1,#], Zhen Yuan [1,#], Yunlei Guo [1,#], Ruilin Wang [1], Yusheng Guan [1], Zhanglian Wang [1], Yunan Chen [1], Tianlu Wang [1], Meining Jiang [1], Shuhui Bian [1,2,*]

[1]State Key Laboratory of Reproductive Medicine and Offspring Health, School of Public Health, Nanjing Medical University, Nanjing 211166, China
[2]Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing 211166, China

*Corresponding author: bianshuhui@njmu.edu.cn (Bian S).
#Equal contribution.

Handling Editor: Jin Gu

Abstract

Single-cell multi-omics sequencing has greatly accelerated reproductive research in recent years, and the data are continually growing. However, utilizing these data resources is challenging for wet-lab researchers. A comprehensive platform for exploring single-cell multi-omics data related to reproduction is urgently needed. Here, we introduce the single-cell multi-omics atlas of reproduction (SMARTdb), an integrative and user-friendly platform for exploring molecular dynamics of reproductive development, aging, and disease, which covers multi-omics, multi-species, and multi-stage data. We curated and analyzed single-cell transcriptomic and epigenomic data of over 2.0 million cells from 6 species across the entire lifespan. A series of powerful functionalities are provided, such as "Query gene expression", "DIY expression plot", "DNA methylation plot", and "Epigenome browser". With SMARTdb, we found that the male germ cell-specific expression pattern of *RPL39L* and *RPL10L* is conserved between human and other model animals. Moreover, DNA hypomethylation and open chromatin may collectively regulate the specific expression pattern of *RPL39L* in both male and female germ cells. In summary, SMARTdb is a powerful platform for convenient data mining and gaining novel insights into reproductive development, aging, and disease. SMARTdb is publicly available at https://smart-db.cn.

Key words: Database; Single-cell multi-omics; Embryo; Germ cell; Gonad.

Introduction

Single-cell multi-omics technique has greatly promoted our molecular understanding of early embryos [1–6], fetal germ cells and gonadal somatic cells [7–11], developing testis and ovary [12–16], adult testis and ovary [17–26], aging testis and ovary [27–29], endometrium [30,31], maternal–fetal interface [32–34], and reproductive diseases [35–39] in the last decade. Multiple molecular layers (including gene expression patterns, DNA methylation dynamics, and open chromatin features) have been profiled with high-throughput sequencing technology, and the data have been growing continually. Although previous single-cell multi-omics studies typically provide comprehensive atlas rather than dig into individual genes or genomic sites, they offer valuable resources for more wet-lab researchers to perform further investigations into specific questions, genes, and genomic sites of interest. For example, it is important for researchers to use model animals to investigate if the expression pattern of a specific gene is conserved in humans, and whether knockout of this gene may potentially result in significant phenotype changes.

However, analyzing the massive and complex single-cell multi-omics data requires experienced skills and substantial computing resources, which is challenging and time-consuming for wet-lab researchers especially. For instance, as each of those studies typically focuses on a narrow time span, researchers who wish to query the expression levels of a particular gene in child, young adult, aging adult, and infertile man need to download and process large amount of data from several separate studies, and they might find this process difficult and laborious due to a lack of analytical skills and computing resources. This inconvenience hinders the optimal utilization of the published data, thus limiting potential novel discoveries. Hence, integrated and user-friendly platforms are needed to bridge the gap between rich data resources and wet-lab researchers.

Efforts such as Human Cell Atlas (HCA, https://www.humancellatlas.org/) and Human Protein Atlas (HPA, https://www.proteinatlas.org/) have served as excellent platforms for exploring massive single-cell data conveniently. However, these databases do not contain much reproduction-related data so far. Currently, there are a few small-scale reproduction-related websites available, which typically only include data of their own study, are limited in data amount and time span, only provide simple function of querying gene expression of a specific gene, and lack long-term maintenance. Consequently, there is still an urgent need for an integrated and user-friendly platform designed to promote the efficient utilization of public single-cell multi-omics data related to reproduction across full life cycle.

Here, we introduce the single-cell multi-omics atlas of reproduction (SMARTdb), which is a manually curated, integrative, and interactive platform aiming to bridge the gap between mass data resources and wet-lab researchers and promote the exploration of the molecular dynamics of reproductive development, aging, and disease.

Received: 4 September 2023; **Revised:** 29 October 2023; **Accepted:** 3 November 2023.

Data collection and processing

Data collection

We curated single-cell transcriptomic and epigenomic (DNA methylation and chromatin accessibility) data related to reproduction from published studies (Table S1). The gene expression matrix, DNA methylation, and chromatin accessibility data were downloaded from Gene Expression Omnibus (GEO) in National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/geo), European Bioinformatics Institute of European Molecular Biology Laboratory (EMBL-EBI) (https://www.ebi.ac.uk/), and Zenodo (https://zenodo.org/). In addition to data of developmental stages and normal physiological states, data of aging and human infertility (such as non-obstructive azoospermia) were also incorporated. In summary, the single-cell data of over 2.0 million individual cells across six species (human, monkey, mouse, pig, buffalo, and goat) were curated, involving early embryos, fetal germ cells and gonadal somatic cells, developing testis and ovary, adult testis and ovary, aging testis and ovary, endometrium, maternal–fetal interface, and human infertility. These data cover the entire life cycle, and span > 120 specific time points of the six species (Table S1).

Processing of single-cell transcriptomic data and cell clustering

To reproduce the cell clustering results of original studies as much as possible, for the datasets with metadata provided by original studies, only the cells with cell type identities were used in downstream analyses. For the datasets without metadata provided, the cells were filtered according to the quality control standards mentioned in the original studies. Seurat (v4.3.0) [40] was used to perform the following processing steps. Briefly, the unique molecular identifier (UMI) count of each gene for each cell was first divided by the total count for that cell, multiplied by 10,000 (or author-designated parameters), and then transformed by natural log. The top highly variable genes among cells were selected using the "FindVariableFeatures" function with default or author-designated parameters. The normalized data were centered and scaled, resulting in the mean expression of each gene across all cells being zero and the standard deviation being one. Principal component analysis (PCA) was performed on the scaled data with the previously determined highly variable genes as input. Then, the cells were clustered with the Louvain algorithm. The Uniform Manifold Approximation and Projection (UMAP) was used to visualize the cell clustering in low-dimensional space. For the datasets with obvious batch effects, the anchors of different samples were identified with the "FindIntegrationAnchors" function, and then different samples were integrated to correct batch effects with the "IntegrateData" function. For the integration of datasets from different studies, strict quality control standards were applied to eliminate the potential effects of poor-quality cells. Only cells with a gene number more than 1000 and a mitochondrial gene ratio less than 20% were used in downstream analyses. For the datasets without obvious batch effects, no additional correction was performed.

Cell type annotation and differential expression analysis of single-cell transcriptomic data

A series of analyses were performed to ensure the accuracy of cell identities for each dataset and integrated datasets. For the datasets with metadata provided by original studies, the cells were annotated correspondingly, and classical marker genes for each cell type were checked to ensure accuracy. For the datasets without metadata, the top differentially expressed genes (DEGs) of each cell cluster were identified using the Wilcoxon rank sum test with the "FindAllMarkers" function in Seurat (v4.3.0) [40] with default parameters (logfc.threshold = 0.25, min.pct = 0.1, return.thresh = 0.01), and their consistency with classical marker genes for each cell type was carefully checked. For the integrated datasets, the cell identities after integration were further checked with the cell type information identified in each dataset before integration. After these steps, the cell type identities can be annotated accurately.

Visualization of gene expression levels

To visualize gene expression levels, Seurat objects were first converted into "anndata" objects with the R package sceasy (v0.0.7). Then, dot plots, heatmaps, and violin plots were generated with SCANPY (v1.9.3) [41]. For dot plots and heatmaps, both the relative expression levels with and without standard scaling were provided. The standard scale normalization strategy in SCANPY was performed with the parameter of standard_scale = 'var', which standardizes the expression level of a given gene between 0 and 1 by subtracting the minimum and dividing each by its maximum.

Processing of single-cell epigenomic data

For the epigenetic data generated by single-cell bisulfite sequencing (scBS-seq) [42] and single-cell chromatin overall omic-scale landscape sequencing (scCOOL-seq) [43], the processed single-cell DNA methylation and chromatin accessibility data provided by original studies were downloaded from GEO database, and were further processed to standard BigWig format with BEDTools (v2.30.0) and UCSC bedGraphToBigWig (v2.9) for downstream application. For the data generated by scBS-seq, only CpG sites were retained. For the data generated by scCOOL-seq, WCG (W = A/T) sites were used for DNA methylation analysis and GCH (H = A/C/T) sites were used for chromatin accessibility analysis. For chromatin accessibility data generated by single-cell assay for transposase-accessible chromatin by sequencing (scATAC-seq), the peak regions of each cell type were mainly collected from their original studies.

Considering the sparse coverage of single-cell epigenome data, we merged single-cell epigenomic data according to their cell type identities, developmental stages (*e.g.*, 2-cell embryos and 8-week female mitotic germ cells), and sequencing techniques (*e.g.*, scBS-seq and scCOOL-seq) using custom scripts. Briefly, for the genomic sites only covered by one cell, they were directly retained in the merged files. For the genomic sites covered by multiple cells within a group, the mean values of covered cells were calculated to represent the final states of the sites. For single-cell DNA methylation data, only the genomic sites with DNA methylation levels < 0.1 or > 0.9 in each individual cell were used in downstream merging. The merged single-cell data of cell groups are more convenient and efficient for downstream exploration.

The DNA methylation "Tanghulu" plot was implemented by a custom Python script to visualize the DNA methylation state of each CpG site in a narrow genomic region of interest. For single-cell DNA methylation data, only the genomic sites with DNA methylation levels < 0.1 or > 0.9 in each individual cell will be plotted. The DNA methylation level is defined

as the percentage of methylated sites in all sites. The WashU Epigenome Browser [44] was incorporated into SMARTdb for convenient visualization of single-cell DNA methylation levels and chromatin accessibility states.

Implementation of the database website

SMARTdb was developed by utilizing Python 3 in conjunction with the Django framework (https://djangoproject.com). The data housed within SMARTdb are systematically organized and accessible through PostgreSQL (https://www.postgresql.org). The user-friendly front-end interface was created using NG-ZORRO and Angular. The platform is hosted using the Nginx web server (http://nginx.org) and is freely accessible online at https://smart-db.cn. For optimal performance, SMARTdb is best used on major web browsers such as Google Chrome, Microsoft Edge, Safari, and Mozilla Firefox, with JavaScript enabled.

Database content and usage

Overall design

SMARTdb is an integrative and interactive platform for exploring single-cell multi-omics data related to reproduction. It was designed to have three main features (**Figure 1**, Figure S1). First, single-cell multi-omics data of over 2.0 million individual cells were manually curated and analyzed (Table S1). Users can access gene expression levels, DNA methylation levels, and chromatin accessibility levels at single-cell resolution conveniently. Second, the main stages of the entire life cycle were covered, including zygote, early embryo, fetus, infant, child, puberty, adult, and aging, with more than 120 specific time points included (Table S1). Additionally, data of human infertility were also incorporated. Users can not only explore the molecular dynamics during the entire life cycle, but also compare normal physiological state with disease. Third, considering that cross-species conservation analyses are very important, data across six species were included, including human, non-human primate (monkey), mouse, pig, buffalo, and goat. For example, researchers using mice as model animals can investigate if the expression pattern of the gene of their interest is conserved in humans, which will help select more meaningful research targets.

As a user-friendly platform, SMARTdb offers multiple featured function modules to users for exploring the rich single-cell multi-omics data, such as "Query gene expression", "DIY expression plot", "DNA methylation plot", and "Epigenome browser". The featured function modules are described below.

"Overview" and "Query gene expression" modules

The "Overview" module allows users to obtain a quick overview of cell clustering, sample information, cell counts, and DEGs of each cell type for each single-cell RNA sequencing (scRNA-seq) dataset (**Figure 2A**). The "Query gene expression" module enables users to conveniently visualize the expression levels of a selected gene across various cell types within the chosen dataset through UMAP and violin plots (Figure 2B). Additionally, this module offers a concise introduction to each human gene.

"DIY expression plot" module

This module provides a very practical function for users to customize expression plots according to their personalized requirements (Figure 2C). Users can select multiple genes and multiple cell types of interest to plot simultaneously, and choose appropriate plot forms for visualization. Three main visualization plots are offered in this module, including dot plots, heatmaps, and violin plots. Moreover, the orders of genes and cell types displayed in the plots can be designated by users freely. Users can also choose whether to standardly scale the expression levels within each gene when using dot plots and heatmaps. The generated figures are of publication quality.

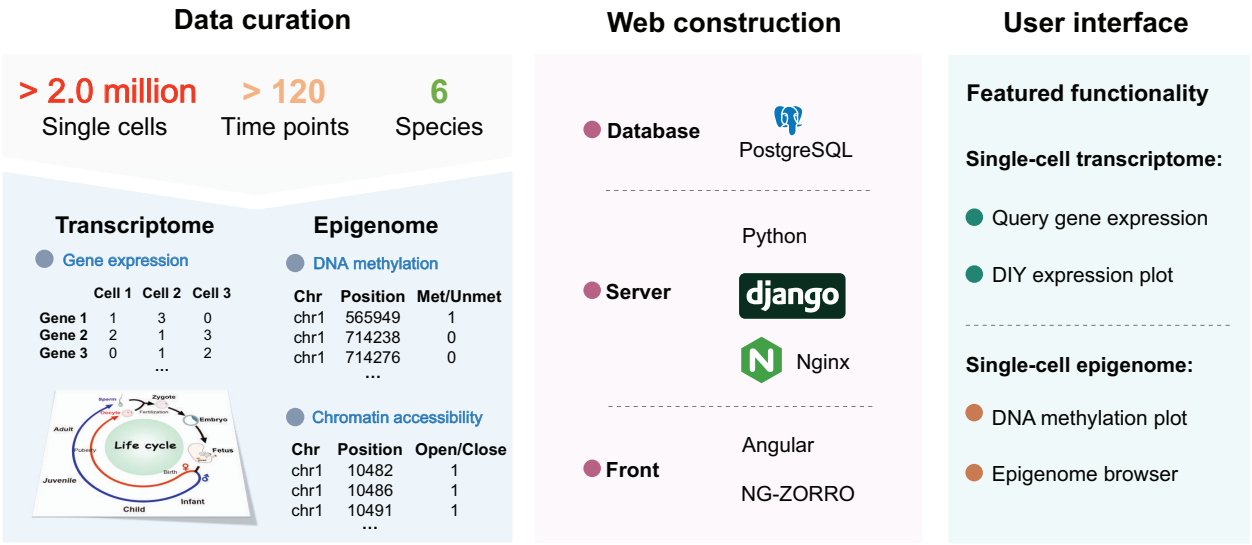

Figure 1 A schematic overview of data curation, web construction, and featured functionalities
The single-cell data of over 2.0 million individual cells across six species (human, monkey, mouse, pig, buffalo, and goat) were curated, covering the entire life cycle and spanning > 120 specific time points of six species. Multiple featured function modules were provided, including "Query gene expression", "DIY expression plot", "DNA methylation plot", and "Epigenome browser".

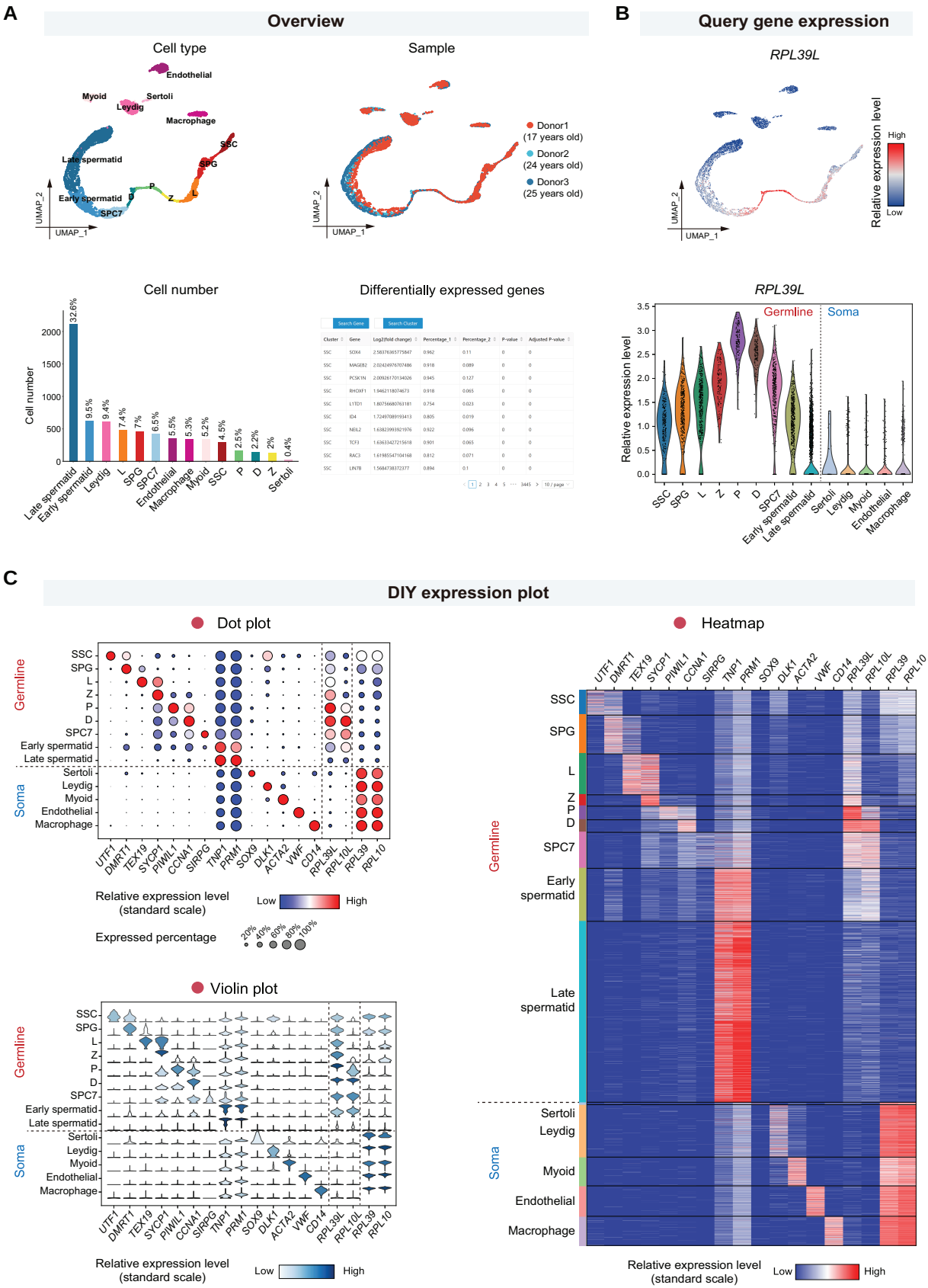

Figure 2 Featured functionalities for exploring single-cell transcriptomic datasets

A. The "Overview" module provides UMAP plots showing cell types and sample information, bar plots showing cell number and proportion of each cell type, and tables showing differentially expressed genes for each single-cell dataset. The scRNA-seq data of human adult testis (GEO: GSE112013) were shown as an example. **B**. The "Query gene expression" module enables visualization of the expression levels for each selected gene with the UMAP plot (upper) and the violin plot (lower). The expression level of a germ cell-specific ribosomal gene *RPL39L* in the human adult testis was shown. **C**. The

"DNA methylation plot" module

This module provides "Tanghulu" plots to visualize the DNA methylation status of each cytosine site in CpG or WCG context (**Figure 3A**). Users can conveniently generate a customized "Tanghulu" plot by selecting datasets and cell types, as well as designating a specific genomic region range of interest (such as promoters and gene bodies). In "Tanghulu" plot, each row represents an individual cell, and each circle represents a cytosine site in CpG or WCG context. Black circles stand for methylated sites, while the white circles refer to unmethylated sites. The DNA methylation levels of selected regions in each single cell are labeled on "Tanghulu" plots.

"Epigenome browser" module

This module incorporates the WashU Epigenome Browser tool [44] and offers a state-of-the-art solution for browsing epigenetic datasets (Figure 3B). Users can select single-cell DNA methylation and chromatin accessibility datasets of various cell types based on their interest, such as oocytes, sperms, early embryos, fetal germ cells, and gonadal somatic cells. The DNA methylation and chromatin accessibility levels of selected samples can be directly and conveniently visualized in the browser. Especially, since single-cell data are typically sparse, the single-cell data of each cell type or sample type are merged for optimal visualization, thus enabling more efficient exploration. The genomic region locator allows users to navigate to a specific gene or specific genomic regions (such as promoters and gene bodies), zoom in and out, and move to new regions by dragging the mouse left and right. Additionally, the browser interface is allowed to be customized to meet personalized requirements, such as display modes (bar plot or heatmap), colors, axis scales, and text labels.

Examples for using SMARTdb

Here, we illustrate how to make good use of SMARTdb to facilitate the exploration of single-cell multi-omics data and promote new findings. A recent study has revealed that RPL39L is a male germ cell-specific ribosome protein and plays critical roles in spermatogenesis and male fertility in mice, whereas the corresponding core ribosome protein RPL39 is mainly expressed in somatic cells [45]. Inspired by the study, we further ask the following questions. (1) Whether the germ cell-specific expression pattern is conserved in the testes of human, non-human primates, and other model animals? (2) When does the germ cell-specific expression pattern of RPL39L form? (3) Does epigenetic regulation function in regulating the male germ cell-specific expression pattern? We will demonstrate how to explore the three questions using SMARTdb below.

Cross-species conservation

We utilized the "DIY expression plot" module to plot the expression levels of the ribosomal large subunit protein genes *RPL39L* and *RPL10L*, and the corresponding core ribosome protein genes *RPL39* and *RPL10* in human [23], as well as their homologous genes in mouse [46], monkey [19], and pig [15] (Figure 2C, **Figure 4A**). In human adult testis [23], we found that the expression levels of *RPL39L* and *RPL10L* peaked at spermatocyte, and then reduced in later stages of spermatogenesis at the messenger RNA (mRNA) level (Figure 2B and C). In contrast, *RPL39* and *RPL10* decreased their expression levels during spermatogenesis, and were mainly expressed in somatic cells (Figure 2C). Moreover, the germ cell-specific expression patterns of *RPL39L* and *RPL10L* were conserved in human and other species (Figure 4A).

Developing dynamics

To address the second question, we plotted the expression levels of *RPL39L* in the fetal stage [10] and the peri-pubertal stage [13] with the "DIY expression plot" module (Figure 4B). We found that the germ cells in fetal stages started to highly express *RPL39L* as early as in migrating primordial germ cells (PGCs). Moreover, we noticed that *RPL39L* transcripts were highly expressed in both male and female germ cells in human fetal stage, but diverge between male and female in adulthood. In human adulthood, *RPL39L* was relatively highly expressed in male germ cells and female growing oocytes, but lowly expressed in fully-grown, metaphase I (MI), and metaphase II (MII) oocytes [21] (**Figure 5A**). The divergence in the *RPL39L* expression among oocytes of various stages in human adult ovary was detected at the mRNA level, indicating potential transcriptional regulation of this gene.

Regulation mechanism

Then we further explored the transcriptional regulatory mechanism using the "Epigenome browser" module. Utilizing epigenetic data of the human fetal stage [7], we discovered that germ cells in fetal stage had hypomethylated and more accessible regions around the promoter of *RPL39L* compared with somatic cells (Figure 3B, Figure S2), which is consistent with the relatively high expression of *RPL39L* observed in fetal germ cells (Figure 4B). In human adult females [21], growing oocytes displayed similar DNA methylation and chromatin accessibility patterns with fetal germ cells (Figure 5B). In contrast, fully-grown, MI, and MII oocytes showed increased DNA methylation levels and decreased chromatin accessibility levels around the *RPL39L* promoter (Figure 5B), which may explain the decreased expression levels of *RPL39L* in fully-grown, MI, and MII oocytes. Hence, epigenetic regulation potentially plays important roles in shaping the expression pattern of *RPL39L* in both male and female human germ cells during both fetal and adult stages.

Figure 2 Continued

"DIY expression plot" module allows users to create customized plots displaying the expression levels of selected genes across designated cell types. Three types of plots (dot plot, violin plot, and heatmap) are provided. The expression levels of representative marker genes and four ribosomal genes in major cell types of the human adult testis were shown. UMAP, Uniform Manifold Approximation and Projection; scRNA-seq, single-cell RNA sequencing; GEO, Gene Expression Omnibus; SSC, spermatogonia stem cell; SPG, spermatogonia; L, leptotene spermatocyte; Z, zygotene spermatocyte; P, pachytene spermatocyte; D, diplotene spermatocyte; SPC7, a mixture of diakinesis, metaphase, anaphase, telophase, and secondary spermatocytes.

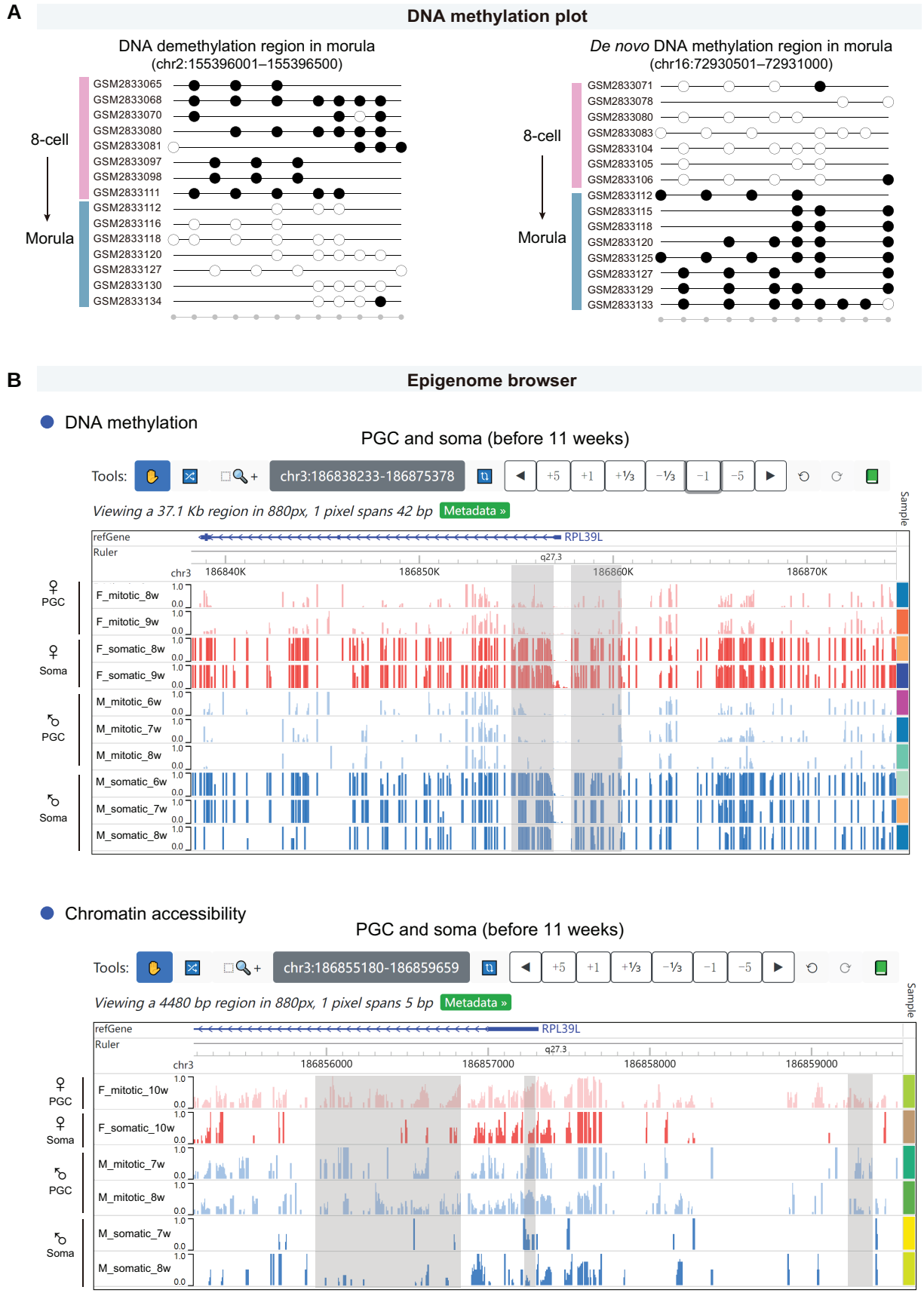

Figure 3 Featured functionalities for exploring single-cell epigenomic datasets

A. The representative DNA demethylation and *de novo* methylation regions between human 8-cell embryo and morula were shown. Each row in the "Tanghulu" plot corresponds to an individual cell, and each circle represents a CpG site. Methylated CpG sites are represented by black dots, while unmethylated CpG sites are indicated by white dots. The gray dots at the bottom indicate covered CpG sites in selected cells. The CpG sites which were

Discussion

The mystery of reproduction-related development, aging, and disease has long been explored. Single-cell multi-omics sequencing techniques have revolutionized research in reproduction in recent years, and have led to accelerated accumulation of massive datasets. The public data provide valuable resources for researchers to perform further data mining. However, difficulties in processing big data for broad wet-lab researchers prevent them from being effectively utilized. To deal with the situation of growing data and promote a more efficient usage, we present SMARTdb — a single-cell multi-omics platform for reproductive research. To the best of our knowledge, SMARTdb is the first integrated platform with multiple powerful functionalities to explore single-cell multi-omics data in the field of reproductive medicine. Moreover, it will be continuously updated by adding high-quality single-cell multi-omics data.

Our database will promote reproductive medicine research in several aspects. For example, researchers using model animals can speculate whether knockout of a specific gene could potentially lead to important phenotypic changes. Users are also able to know whether the expression pattern of a specific gene is conserved in human, which will help select more meaningful research targets. For researchers working with human samples, cross-species conservation analyses provide important clues for selecting suitable model animals, while cross-stage dynamic analysis will help select the right time window for further investigations. Moreover, the epigenetics data resources will provide novel insights on potential regulatory mechanisms.

We showed examples for using SMARTdb and demonstrated its power by exploring germ cell-specific genes. We not only validated the germ cell-specific expression of the ribosome gene *RPL39L* but also gained a deeper understanding of its cross-species conservation, cross-stage expression dynamics, and potential epigenetic regulation. Moreover, taking the advantages of single-cell transcriptome sequencing, the gene expression patterns in multiple cell types can be observed more accurately and in detail. For example, we found that *RPL39L* was not only specifically expressed in male germ cells in human adult testis but also highly expressed in the growing oocytes (one subtype of oocytes) in human adult ovary. Furthermore, DNA methylation and chromatin accessibility may cooperate in regulating the divergence of *RPL39L* expression in germ cells of human adult at mRNA level. These results demonstrate the power of SMARTdb, providing valuable clues for further investigations.

In summary, we firmly believe that with the assistance of SMARTdb, researchers in the field of reproduction can easily access and explore massive single-cell multi-omics data, and hopefully generate novel ideas and new findings.

Data availability

SMARTdb is publicly available at https://smart-db.cn.

CRediT author statement

Zekai Liu: Investigation, Data curation, Formal analysis, Methodology, Software, Visualization, Validation. **Zhen Yuan:** Investigation, Data curation, Formal analysis, Methodology, Software, Visualization, Validation, Writing – original draft. **Yunlei Guo:** Investigation, Formal analysis, Methodology, Software, Visualization, Validation. **Ruilin Wang:** Investigation, Formal analysis, Validation. **Yusheng Guan:** Investigation, Data curation, Methodology, Validation. **Zhanglian Wang:** Investigation, Data curation. **Yunan Chen:** Investigation, Data curation. **Tianlu Wang:** Investigation, Data curation. **Meining Jiang:** Investigation, Data curation. **Shuhui Bian:** Conceptualization, supervision, Writing – original draft, Writing – review & editing, Visualization, Investigation, Data curation, Methodology, Formal analysis, Funding acquisition, Project administration. All authors have read and approved the final manuscript.

Supplementary material

Supplementary material is available at *Genomics, Proteomics & Bioinformatics* online (https://doi.org/10.1093/gpbjnl/qzae005).

Competing interests

The authors have declared no competing interests.

Acknowledgments

We thank Profs. Fuchou Tang (Peking University), Jingchu Luo (Peking University), Zhibin Hu (Nanjing Medical University), Jiahao Sha (Nanjing Medical University), Fan Guo (Institute of Zoology, Chinese Academy of Sciences), Ran Huo (Nanjing Medical University), and Boqiang Hu (Zhejiang University) for discussions and suggestions. This work was supported by the Young Elite Scientists Sponsorship Program by China Association for Science and Technology (Grant No. 2020QNRC001) and the research start-up funding from Nanjing Medical University, China (Grant No. KY116RC20200007).

ORCID

0009-0003-2261-1216 (Zekai Liu)
0000-0003-4980-6681 (Zhen Yuan)
0009-0001-2455-5508 (Yunlei Guo)
0000-0002-8420-9033 (Ruilin Wang)

Figure 3 Continued

not covered by any cell were omitted. The single-cell DNA methylation data of human early embryo (GEO: GSE100272) were used. **B**. The screenshots of Epigenome browser displaying the DNA methylation levels and chromatin accessibility levels around the promoter region of *RPL39L* in PGCs. The gray rectangles highlight regions showing significant differences between somatic niche cells and germline cells. PGC, primordial germ cell.

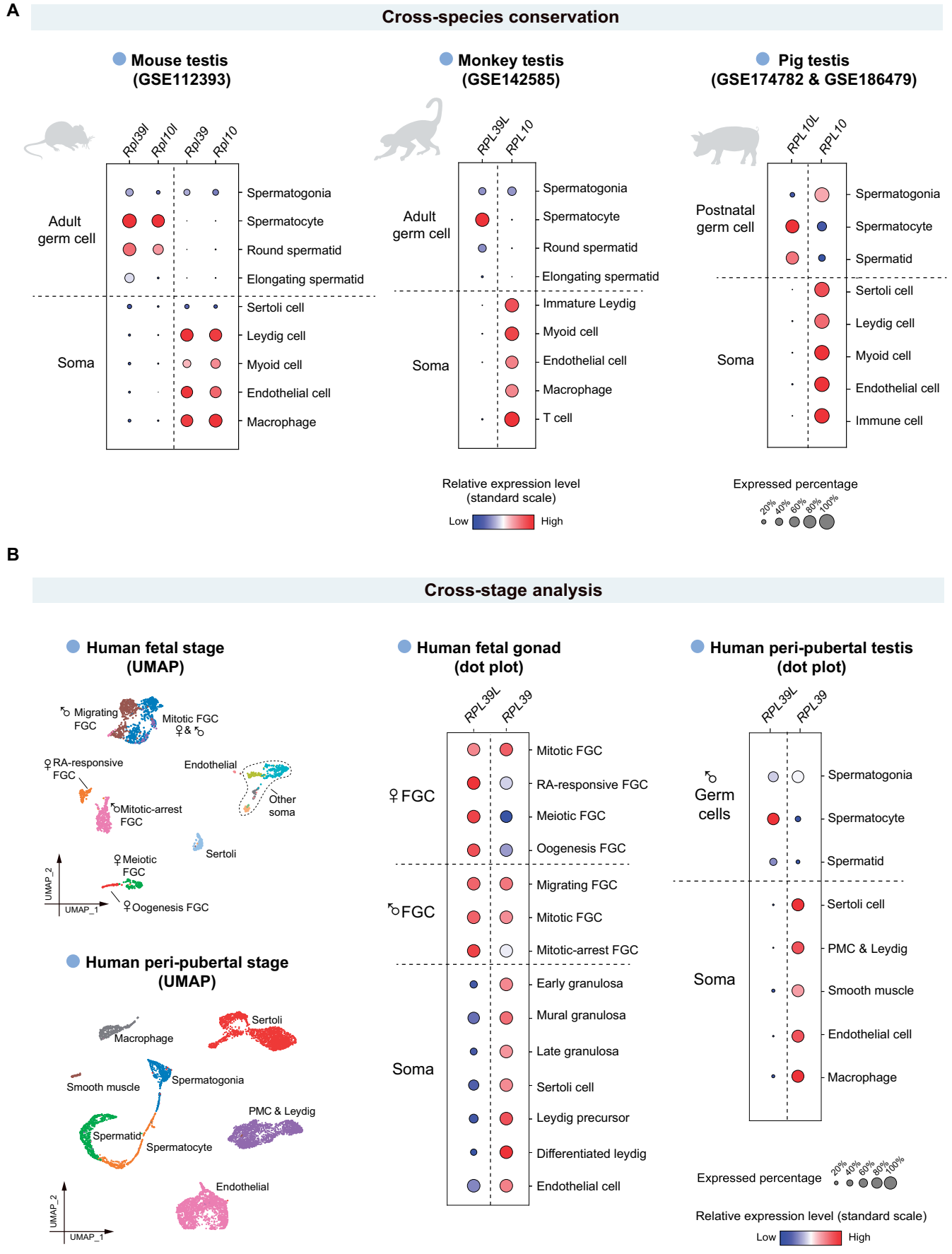

Figure 4 Cross-species and cross-stage analyses of ribosomal genes using SMARTdb

A. The dot plots show the expression levels of four ribosomal genes in mouse, cynomolgus monkey, and pig. **B**. The UMAP plots show the cell type identities of each cell cluster. The expression levels of two ribosomal genes (*RPL39L* and *RPL39*) in human germline and somatic niche cells of fetal stage and peri-pubertal stage were shown in dot plots. FGC, fetal germ cell; PMC, peritubular myoid cell.

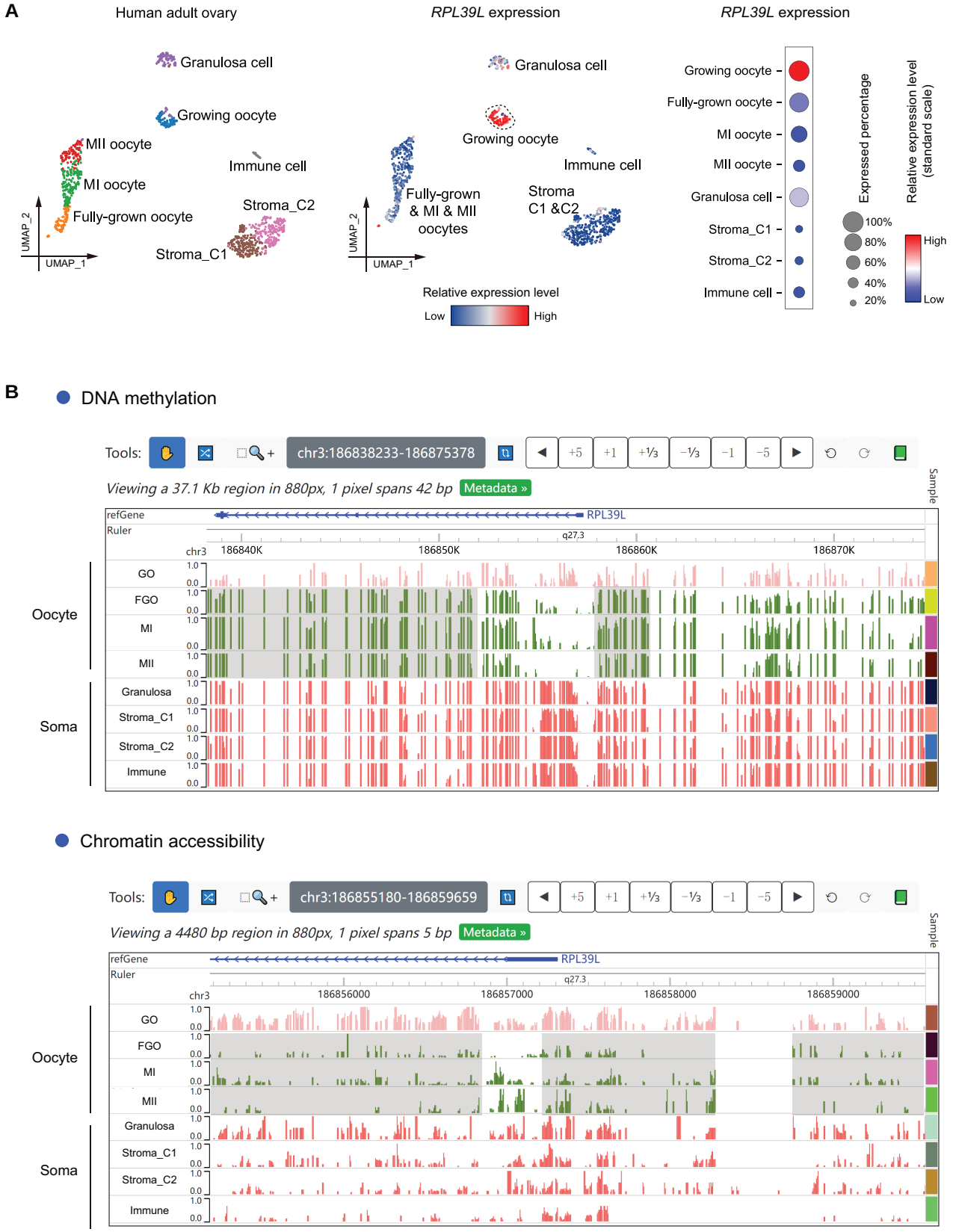

Figure 5 Gene expression, DNA methylation, and chromatin accessibility patterns of *RPL39L* in human ovary
A. The leftmost UMAP plot shows the cell type identities of human adult ovary. The middle UMAP plot and the rightmost dot plot illustrate the expression levels of *RPL39L* in various cell types of human adult ovary. **B**. The screenshots of Epigenome browser displaying the DNA methylation and chromatin accessibility levels around the promoter region of *RPL39L* in human adult ovary. The gray rectangles highlight regions showing changes in fully-grown, MI, and MII oocytes compared with growing oocyte. MI, metaphase I; MII, metaphase II.

0000-0002-1319-7784 (Yusheng Guan)
0009-0001-5068-3853 (Zhanglian Wang)
0009-0006-6175-0042 (Yunan Chen)
0009-0000-1145-8306 (Tianlu Wang)
0009-0001-9423-0151 (Meining Jiang)
0000-0002-9662-113X (Shuhui Bian)

References

[1] Zhu P, Guo H, Ren Y, Hou Y, Dong J, Li R, et al. Single-cell DNA methylome sequencing of human preimplantation embryos. Nat Genet 2018;50:12–9.
[2] Zhou F, Wang R, Yuan P, Ren Y, Mao Y, Li R, et al. Reconstituting the transcriptome and DNA methylome landscapes of human implantation. Nature 2019;572:660–4.
[3] Wang Y, Yuan P, Yan Z, Yang M, Huo Y, Nie Y, et al. Single-cell multiomics sequencing reveals the functional regulatory landscape of early embryos. Nat Commun 2021;12:1247.
[4] Argelaguet R, Clark SJ, Mohammed H, Stapel LC, Krueger C, Kapourani CA, et al. Multi-omics profiling of mouse gastrulation at single-cell resolution. Nature 2019;576:487–91.
[5] Pijuan-Sala B, Griffiths JA, Guibentif C, Hiscock TW, Jawaid W, Calero-Nieto FJ, et al. A single-cell molecular map of mouse gastrulation and early organogenesis. Nature 2019;566:490–5.
[6] Li L, Guo F, Gao Y, Ren Y, Yuan P, Yan L, et al. Single-cell multi-omics sequencing of human early embryos. Nat Cell Biol 2018;20:847–58.
[7] Li L, Li L, Li Q, Liu X, Ma X, Yong J, et al. Dissecting the epigenomic dynamics of human fetal germ cell development at single-cell resolution. Cell Res 2021;31:463–77.
[8] Chitiashvili T, Dror I, Kim R, Hsu FM, Chaudhari R, Pandolfi E, et al. Female human primordial germ cells display X-chromosome dosage compensation despite the absence of X-inactivation. Nat Cell Biol 2020;22:1436–46.
[9] Garcia-Alonso L, Lorenzi V, Mazzeo CI, Alves-Lopes JP, Roberts K, Sancho-Serra C, et al. Single-cell roadmap of human gonadal development. Nature 2022;607:540–7.
[10] Li L, Dong J, Yan L, Yong J, Liu X, Hu Y, et al. Single-cell RNA-seq analysis maps development of human germline cells and gonadal niche interactions. Cell Stem Cell 2017;20:858–73.
[11] Chen M, Long X, Chen M, Hao F, Kang J, Wang N, et al. Integration of single-cell transcriptome and chromatin accessibility of early gonads development among goats, pigs, macaques, and humans. Cell Rep 2022;41:111587.
[12] Guo J, Sosa E, Chitiashvili T, Nie X, Rojas EJ, Oliver E, et al. Single-cell analysis of the developing human testis reveals somatic niche cell specification and fetal germline stem cell establishment. Cell Stem Cell 2021;28:764–78.
[13] Guo J, Nie X, Giebler M, Mlcochova H, Wang Y, Grow EJ, et al. The dynamic transcriptional cell atlas of testis development during human puberty. Cell Stem Cell 2020;26:262–76.
[14] Zhao J, Lu P, Wan C, Huang Y, Cui M, Yang X, et al. Cell-fate transition and determination analysis of mouse male germ cells throughout development. Nat Commun 2021;12:6839.
[15] Zhang L, Guo M, Liu Z, Liu R, Zheng Y, Yu T, et al. Single-cell RNA-seq analysis of testicular somatic cell development in pigs. J Genet Genomics 2022;49:1016–28.
[16] Gu C, Liu S, Wu Q, Zhang L, Guo F. Integrative single-cell analysis of transcriptome, DNA methylome and chromatin accessibility in mouse oocytes. Cell Res 2019;29:110–23.
[17] Chen Y, Zheng Y, Gao Y, Lin Z, Yang S, Wang T, et al. Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. Cell Res 2018;28:879–96.
[18] Sohni A, Tan K, Song HW, Burow D, de Rooij DG, Laurent L, et al. The neonatal and adult human testis defined at the single-cell level. Cell Rep 2019;26:1501–17.
[19] Shami AN, Zheng X, Munyoki SK, Ma Q, Manske GL, Green CD, et al. Single-cell RNA sequencing of human, macaque, and mouse testes uncovers conserved and divergent features of mammalian spermatogenesis. Dev Cell 2020;54:529–47.
[20] Fan X, Bialecka M, Moustakas I, Lam E, Torrens-Juaneda V, Borggreven NV, et al. Single-cell reconstruction of follicular remodeling in the human adult ovary. Nat Commun 2019;10:3164.
[21] Yan R, Gu C, You D, Huang Z, Qian J, Yang Q, et al. Decoding dynamic epigenetic landscapes in human oocytes using single-cell multi-omics sequencing. Cell Stem Cell 2021;28:1641–56.
[22] Zhang Y, Yan Z, Qin Q, Nisenblat V, Chang HM, Yu Y, et al. Transcriptome landscape of human folliculogenesis reveals oocyte and granulosa cell interactions. Mol Cell 2018;72:1021–34.
[23] Guo J, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie X, et al. The adult human testis transcriptional cell atlas. Cell Res 2018;28:1141–57.
[24] Hermann BP, Cheng K, Singh A, Roa-De La Cruz L, Mutoji KN, Chen IC, et al. The mammalian spermatogenesis single-cell transcriptome, from spermatogonial stem cells to spermatids. Cell Rep 2018;25:1650–67.
[25] Huang Y, Li L, An G, Yang X, Cui M, Song X, et al. Single-cell multi-omics sequencing of human spermatogenesis reveals a DNA demethylation event associated with male meiotic recombination. Nat Cell Biol 2023;25:1520–34.
[26] Wang M, Liu X, Chang G, Chen Y, An G, Yan L, et al. Single-cell RNA sequencing analysis reveals sequential cell fate transition during human spermatogenesis. Cell Stem Cell 2018;23:599–614.
[27] Nie X, Munyoki SK, Sukhwani M, Schmid N, Missel A, Emery BR, et al. Single-cell analysis of human testis aging and correlation with elevated body mass index. Dev Cell 2022;57:1160–76.
[28] Huang D, Zuo Y, Zhang C, Sun G, Jing Y, Lei J, et al. A single-nucleus transcriptomic atlas of primate testicular aging reveals exhaustion of the spermatogonial stem cell reservoir and loss of Sertoli cell homeostasis. Protein Cell 2022;14:888–907.
[29] Wang S, Zheng Y, Li J, Yu Y, Zhang W, Song M, et al. Single-cell transcriptomic atlas of primate ovarian aging. Cell 2020;180:585–600.
[30] Garcia-Alonso L, Handfield LF, Roberts K, Nikolakopoulou K, Fernando RC, Gardner L, et al. Mapping the temporal and spatial dynamics of the human endometrium *in vivo* and *in vitro*. Nat Genet 2021;53:1698–711.
[31] Wang W, Vilella F, Alama P, Moreno I, Mignardi M, Isakova A, et al. Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. Nat Med 2020;26:1644–53.
[32] Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, et al. Single-cell reconstruction of the early maternal–fetal interface in humans. Nature 2018;563:347–53.
[33] Arutyunyan A, Roberts K, Troulé K, Wong FCK, Sheridan MA, Kats I, et al. Spatial multiomics map of trophoblast development in early pregnancy. Nature 2023;616:143–51.
[34] Greenbaum S, Averbukh I, Soon E, Rizzuto G, Baranski A, Greenwald NF, et al. A spatially resolved timeline of the human maternal–fetal interface. Nature 2023;619:595–605.
[35] Zhao L, Yao C, Xing X, Jing T, Li P, Zhu Z, et al. Single-cell analysis of developing and azoospermia human testicles reveals central role of Sertoli cells. Nat Commun 2020;11:5683.
[36] Di Persio S, Tekath T, Siebert-Kuss LM, Cremers JF, Wistuba J, Li X, et al. Single-cell RNA-seq unravels alterations of the human spermatogonial stem cell compartment in patients with impaired spermatogenesis. Cell Rep Med 2021;2:100395.
[37] Alfano M, Tascini AS, Pederzoli F, Locatelli I, Nebuloni M, Giannese F, et al. Aging, inflammation and DNA damage in the somatic testicular niche with idiopathic germ cell aplasia. Nat Commun 2021;12:5205.
[38] Chen Y, Liu X, Zhang L, Zhu F, Yan L, Tang W, et al. Deciphering the molecular characteristics of human idiopathic nonobstructive azoospermia from the perspective of germ cells. Adv Sci 2023;10:2206852.
[39] Ferrero H, Corachán A, Aguilar A, Quiñonero A, Carbajo-García MC, Alamá P, et al. Single-cell RNA sequencing of oocytes from

ovarian endometriosis patients reveals a differential transcriptomic profile associated with lower quality. Hum Reprod 2019; 34:1302–12.
[40] Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell 2021;184:3573–87.
[41] Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 2018;19:15.
[42] Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods 2014;11:817–20.
[43] Guo F, Li L, Li J, Wu X, Hu B, Zhu P, et al. Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. Cell Res 2017;27:967–88.
[44] Li D, Purushotham D, Harrison JK, Hsu S, Zhuo X, Fan C, et al. WashU Epigenome Browser update 2022. Nucleic Acids Res 2022;50:W774–81.
[45] Li H, Huo Y, He X, Yao L, Zhang H, Cui Y, et al. A male germ-cell-specific ribosome controls male fertility. Nature 2022; 612:725–31.
[46] Green CD, Ma Q, Manske GL, Shami AN, Zheng X, Marini S, et al. A comprehensive roadmap of murine spermatogenesis defined by single-cell RNA-seq. Dev Cell 2018;46:651–67.

Genomics, Proteomics & Bioinformatics, 2024, 22, 1–11
https://doi.org/10.1093/gpbjnl/qzae005
Database