Dissecting chicken germ cell dynamics by combining a germ cell tracing transgenic chicken model with single-cell RNA sequencing

Deivendran Rengaraja,1, Dong Gon Cha b,1, Hong Jo Lee a,1, Kyung Youn Lee a, Yoon Ha Choi c,b, Kyung Min Jung c, Young Min Kim a, Hee Jung Choi a, Hyeon Jeong Choi a, Eunhui Yoo a, Seung Je Woo a, Jin Se Park a, Kyung Je Park b, Jong Kyoungh Kim c,b,a, Jae Yong Han a,*

a Department of Agricultural Biotechnology, and Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul 08826, South Korea
b Department of New Biology, DGIST, Daegu 42988, South Korea
c Department of Life Sciences, Pohang University of Science and Technology (POSTECH), Pohang 37673, South Korea

1. Introduction

Germ cells produce types of cells that transmit genetic information to the next generation. Compared with germ cells of other animal species, avian germ cells have unique developmental characteristics in terms of specification, migration, and development [1]. Two modes of germ cell specification were reported in the vertebrate species: epigenesis mode in mammals; and preformation mode in several lower vertebrate species [2]. In particular, avian germ cells are known to be specified by preformation mode, in which maternally inherited components called “germ plasm” play critical role in the specification of germ cells and maintaining the integrity of germ cells. The germ plasm consists of energy-rich mitochondria, and a set of RNAs and proteins, including the germ-cell-specific markers DEAD-box helicase (DDX4) and deleted in azoospermia like (DAZL) [3,4]. The precursors of chicken germ cells, called primordial germ cells (PGCs), are detected as early as Eyal-Giladi and Kochav (EGK) [5] stage-III of embryogenesis on the basis of expression of the DAZL [4]. Germ cells scattered in the central region of EGK stage-X embryos are incorporated into the germinal crescent region of Hamburger and Hamilton (HH) [6] stage-4 embryos. These germ cells then circulate through the blood to colonize bilateral-embryonic gonads at around HH stage 24. After colonization, germ cells actively proliferate and differentiate in a sexually dimorphic manner [7].

During embryogenesis in several species, common germ-cell developmental events include migration, embryonic gonad settlement, and sex-specific differentiation. Distinct molecular networks regulate these events in both male and female germ cells [8,9]. Avian germ cells also dynamically proliferate and differentiate:
male germ cells enter mitotic arrest, and female germ cells enter meiotic arrest [7,10–12]. However, studies of germ-cell dynamics during embryogenesis are limited because surface markers to isolate germ cells at each developmental stage are not available.

We combined a novel germ-cell–tracing method with time-resolved single-cell RNA sequencing (scRNA-seq) to explore chicken germ-cell dynamics at single-cell resolution. We used the clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (CRISPR/Cas9) gene-editing system to insert a green fluorescent protein (GFP)-expression cassette into DAZL through nonhomologous end joining (NHEJ) to produce DAZL::GFP transgenic chickens. DAZL, an RNA-binding protein and DAZ-family member, is a master factor critical for germ-cell development and maintenance in diverse vertebrate and invertebrate species. DAZL binds and regulates thousands of mRNAs 3′ untranslated regions (3′ UTRs) during different stages of germ-cell development [13]. In chickens, DAZL mRNA and protein are components of germ plasm, in which the zigote inherits to specify germ cells [4]. DAZL mRNA and protein are also continuously expressed at various stages of chicken germ-cell development (including specification, migration, differentiation, and gametogenesis) [4,14]. Our main objectives of this study are to produce DAZL::GFP transgenic chickens without affecting endogenous DAZL expression, to isolate pure population of male and female germ cells from the DAZL::GFP transgenic chickens at embryonic day 2.5 (E2.5) to 1 week post-hatch, and to analyze the isolated cells by using scRNA-seq for investigating the transcriptional landscape, cellular heterogeneity and developmental trajectories of chicken germ cells.

2. Materials and methods

2.1. Animals

The management and experimental use of chickens were approved by the Institute of Laboratory Animal Resources, Seoul National University (SNU-190401-1-1). The experimental animals were cared according to a standard management protocol at the University Animal Farm, Seoul National University, Korea. The procedures for animal management, reproduction, and embryo manipulation adhered to the standard operating protocols of Animal Genetic Engineering Laboratory, Seoul National University.

2.2. Construction of CRISPR/Cas9 and donor plasmids

All-in-one CRISPR/Cas9 plasmids targeting the last intron of chicken DAZL were constructed by using the multiplex-CRISPR/Cas9 plasmid-construction kit (provided by Takashi Yamamoto) (Addgene Kit #1000000054) [15]. A neomycin-resistance gene (regulated by the thymidine-kinase promoter) was ligated into the NotI (New England Biolabs, Ipswich, MA, USA) sites of the CRISPR/Cas9 plasmids. To insert DAZL-targeting gRNA sequences into the CRISPR/Cas9 plasmids, sense and antisense oligonucleotides were designed (Table S1) and synthesized (Bionics, Seoul, Korea). The oligonucleotides were denatured and reannealed with the following thermocycling conditions: 30 sec at 95 °C, 2 min at 72 °C, 2 min at 37 °C, and 2 min at 25 °C. To tag chicken DAZL, a donor plasmid was synthesized to contain the last intron and exon of DAZL (including the gRNA-recognition sequence) in frame with a T2A peptide and GFP expression cassette (Bioneer, Daejeon, Korea). The neomycin-resistance gene with a thymidine-kinase promoter was inserted into the 3′-downstream region of the synthesized plasmid.

2.3. Cell culture

White Leghorn (WL) PGCs retrieved from male gonads at E6 by magnetic-activated cell-sorting (MACS) method [16] were maintained and sub-passaged on knockout Dulbecco’s minimum essential medium (DMEM; Invitrogen, Carlsbad, CA, USA) supplemented with 20% fetal bovine serum (FBS; Invitrogen); 2% chicken serum (Sigma-Aldrich, St. Louis, MO, USA); 1× nucleosides (Millipore, Temecula, CA, USA); 2 mM L-glutamine; 1× nonessential amino acids; 0.2% β-mercaptoethanol; 10 mM sodium pyruvate; 1× antibiotic–antimycotic (Invitrogen); and 10 ng/mL human basic fibroblast growth factor (Sigma-Aldrich). Chicken PGCs were cultured in an incubator at 37 °C under an atmosphere of 5% CO2 and 60–70% relative humidity. The PGCs were sub-cultured onto mitomycin-inactivated mouse embryonic fibroblasts at five-day or six-day intervals via gentle pipetting.

2.4. Transfection and G418 selection of PGCs for targeted gene insertion

To edit chicken PGC genomes, 1 × 10⁵ PGCs were co-transfected with donor plasmids (2 μg) and CRISPR/Cas9 plasmids (2 μg) using 4 μL Lipofectamine 2000 reagent (Thermo Fisher–Invitrogen, Carlsbad, CA, USA) suspended in 1 mL Opti-MEM (Thermo Fisher–Invitrogen). After 4 h, the transfection mixture was replaced with PGC culture medium. Geneticin Selective Antibiotic (300 μg/mL) (G418; GibCO Invitrogen, Grand Island, NY, USA) was added to the culture medium 1 day after transfection. G418 selection was performed for up to 3 weeks. After G418 selection, genomic DNA was extracted from PGCs. Regions encompassing the CRISPR/Cas9-targeted sites were amplified by using specific primer sets (Table S1). For sequencing analysis, the amplicons were cloned into the pGEM-T-easy vector (Promega, Madison, WI, USA) and sequenced by using an ABI Prism 3730 XL DNA Analyzer (Thermo Fisher–Applied Biosystems, Foster City, CA, USA). Sequences were compared with assembled genome sequences by using BLAST (https://blast.ncbi.nlm.nih.gov).

2.5. Production of DAZL::GFP transgenic chickens

To produce a DAZL::GFP transgenic chicken, a window was cut into the sharp end of a Korean-Ogye (i/i)-recipient egg, and over 3,000 DAZL::GFP WL PGCs (I/i) were transplanted into the dorsal aorta of the HH stage 14–17 recipient embryo. The egg window was sealed with parafilm, and the egg was incubated with the sharp end down until hatching. After sexual maturation, sperm from male recipient chickens (I or i) were evaluated by breed-specific PCR, and the chickens with WL sperm (I) were mated with WT females (I/i). Germline-chimeric chickens were identified by offspring feather color (I/i) and the individual donor PGC-derived chickens (I/i) were further validated by subsequent genomic DNA analysis to identify targeted gene insertion.

2.6. Detection of targeted gene insertion and off-target effects

To identify modified alleles in the DAZL::GFP PGCs and in transgenic chickens, genomic DNA was analyzed by using knock-in PCR. All reactions were performed under the same conditions: a 20-μL volume containing genomic DNA (100 ng); 1× PCR buffer; 10 mM each dNTP; 10 pmol each amplification primers (Table S1); and 0.5 U Taq polymerase (Bionics). The following thermocycling conditions were used: 2 min at 95 °C; 35 cycles of 20 sec at 95 °C, 40 sec at 60 °C, and 30 sec at 72 °C; and 5 min at 72 °C. For sequencing, amplicons were ligated into pGEM-T-easy vector and sequenced as described.
To identify off-target effects, putative off-target sites were predicted for the DAZL::GFP transgenic chickens by using web-based CRISPR off-target prediction software (https://crispr.mit.edu/). Each site was validated manually by using TA cloning. Clones were sequenced as described.

2.7. Reverse-transcription PCR and quantitative reverse-transcription PCR analysis

Total RNA was isolated from DAZL::GFP PGCs and WT cultured PGCs by using TRizol reagent (Thermo Fisher Scientific). For reverse-transcription PCR (RT-PCR) and quantitative reverse-transcription PCR (qRT-PCR), total RNA (1 μg) was used as a template for cDNA synthesis with the SuperScript III First-Strand Synthesis System (Thermo Fisher Scientific). The cDNA was serially diluted (5-fold) and quantitatively equalized for PCR amplification. For each target, the same amplification primers were used for both RT-PCR and qRT-PCR (Table S1). RT-PCR was performed in a 20-μL volume containing 2 μL cDNA; 1 × PCR buffer; 10 mM each dNTPs; 10 pmol each amplification primers; and 0.5 U Taq polymerase (Bionics). The following thermocycling conditions were used: 2 min at 95 °C; 35 cycles of 20 sec at 95 °C, 40 sec at 60 °C, 30 sec at 72 °C; and 5 min at 72 °C. qRT-PCR was performed in a 20-μL volume containing 2 μL cDNA; 1 × PCR buffer; 10 mM each dNTP; 10 pmol each amplification primers; 1 × EvaGreen (Biotium, Hayward, CA); and 0.5 U Taq polymerase. Amplification was monitored by using the CFX96 real-time PCR detection system on a C1000 thermal cycler (Bio-Rad Laboratories, Hercules, CA, USA). Thermocycling conditions for qRT-PCR were as follows: 95 °C for 3 min; and 40 cycles of 95 °C for 30 sec, 59 °C for 30 sec, and 72 °C for 30 sec. Melting-curve profiles were analyzed for all amplimers. Each sample was measured in triplicate, and relative gene expression was calculated after normalization with a housekeeping gene and control/reference sample (2^(-ΔΔCT) method, where control/reference = 1).

2.8. Immunocytochemistry analysis

To validate the expression of DAZL protein, DAZL::GFP PGCs and WT cultured PGCs were harvested, fixed with 4% paraformaldehyde at room temperature (RT) for 10 min, and permeabilized with 0.1% Triton X-100 at RT for 10 min. After washing three times with phosphate-buffered saline (PBS), cells were blocked with a blocking buffer (PBS; 5% goat serum; 1% bovine serum albumin [BSA]) at RT for 1 h. Then, the cells were incubated at 4 °C overnight with rabbit anti-cDAZL polyclonal antibodies [4]. After washing three times with PBS, cells were incubated with secondary antibodies (anti-rabbit IgG) conjugated with Alexa Fluor 594 (Invitrogen) at RT for 1 h. Finally, cells were stained with propidium iodide (PI) to isolate live cells, and GFP+/PI– cells were sorted by using a BD FACS Aria III (BD Biosciences, San Jose, CA, USA) for scRNA-seq. In addition, another batch of DAZL::GFP germ cells were isolated at the representative times, including E6, E12, E16, hatch and 1 week post-hatch, for the examination of gene expression by qRT-PCR. The methods for the isolation of total RNA, cDNA synthesis, and gene amplification are same as mentioned above.

2.10. Preparation of DAZL::GFP germ cells for scRNA-seq

DAZL::GFP germ cells at various stages of embryogenesis were collected from either blood (at E2.5) or from gonads (at E6, E8, E12, E16, hatch, and 1 week post-hatch). To sample germ cells from E2.5 embryos, circulating embryonic blood from male and female embryos was extracted separately using a microneedle and was resuspended in PBS containing 1% BSA for fluorescence-activated cell sorting (FACS). The pooled left-side gonads of the male and female DAZL::GFP chicken at respective time (E6, E8, E12, E16, hatch, and 1 week post-hatch) were pooled separately and treated with Hank's Balanced Salt Solution (Gibco Invitrogen) containing 0.05% trypsin–EDTA (Gibco Invitrogen) and incubated at 37 °C for 10 min. During incubation, cells were gently pipetted every 2 min. After incubation, trypsin–EDTA was inactivated by adding the same volume of DMEM containing 5% FBS. Cells were harvested by centrifugation (1,250 rpm; 5 min), and washed with PBS. Cells were resuspended in PBS containing 1% BSA and filtered through a 40-μm cell strainer (Falcon® 352340, Fisher Scientific, Hampton, NH, USA). The sex of E2.5, E6, and E8 embryos was determined (at E2.5) by sex-discriminating PCR of blood samples [17]. The sex of embryos at E12, E16, hatch, and 1 week post-hatch was determined by gonad morphology. Finally, cells were stained with propidium iodide (PI) to isolate live cells, and GFP+/PI– cells were sorted by using a BD FACS Aria III (BD Biosciences, San Jose, CA, USA) for scRNA-seq. In addition, another batch of DAZL::GFP germ cells were isolated at the representative times, including E6, E12, E16, hatch and 1 week post-hatch, for the examination of gene expression by qRT-PCR. The methods for the isolation of total RNA, cDNA synthesis, and gene amplification are same as mentioned above.

2.11. scRNA-seq

Libraries for scRNA-seq were prepared by using the Chromium Single Cell 3’ GEM, Library & Gel Bead Kit v3 (PN-1000075, 10X Genomics, Pleasanton, CA, USA); Chromium Single Cell B Chip Kit (PN-1000073, 10X Genomics); and Chromium i7 Multiplex Kit (PN-120262, 10X Genomics). Cells were resuspended in PBS containing 0.04% BSA and diluted to ~2 × 10^6~ ~1 × 10^5 cells/mL. Cells were mixed with a reverse-transcription master mix and loaded onto B chip channels to capture of ~800 to ~5,000 single-cell transcriptomes. Gel bead-in-emulsions (GEMs) were generated by using Chromium Controller (10X Genomics). Reverse transcription was conducted by using a C1000 Touch thermal cycler (Bio-Rad). DNA was purified and libraries were constructed according to the manufacturer’s instructions. The qualities of amplified cDNAs and of the constructed libraries were assessed by using Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Libraries were sequenced with a 2 x 100-bp paired-end protocol on a Novaseq-6000 platform (Illumina, San Diego, CA, USA) to generate at least 40,000 read pairs per cell.

2.12. scRNA-seq data preprocessing

Raw fastq files were processed using the CellRanger pipeline, version 3.1.0. The fasta and GTF files for chicken genome (GRCg6a) were modified to include the DAZL-GFP insert sequence. The cDNA
sequences were mapped to the modified-chicken genome by using STAR, version 2.5.1b, aligner [18] with the GRCg6a.99 GTF file. A gene-by-cell count matrix was generated by using default parameters. To remove empty droplets while capturing single cells, the EmptyDrops function of DropUtils, version 1.8.0, R package [19] was used (with FDR < 0.05). Low-quality cells were excluded by using different cutoff thresholds for different samples. The cut-off thresholds were determined by visually inspecting outliers in the principal component analysis (PCA) plot on the quality-control metrics using the calculateQCMetrics function of the scater, version 1.16.1, R package [20].

To remove cell-specific biases, cells were clustered by using the quickCluster function of the scan, version 1.16.0, R package [21]. Default parameters and cell-specific size factors were computed by using the computeSumFactors function of the same package. The gene-by-cell count matrix of E2.5 sample was normalized by dividing the raw unique molecular identifier (UMI) counts by cell-specific size factors. The normalized counts were then log2-transformed by adding a pseudo-count of 1. One thousand highly variable genes (HVGs) in E2.5 PGCs were selected with respect to biological variability by using the decomposeVar and the getTopHVGs function of the scan package.

The k-nearest neighbor (kNN) graph was computed with FindNeighbors function of Seurat, version 3.1.5, R package [22] on the first 15 principal components (PCs) and used to compute clusters by using FindClusters function with resolution = 1.0. The 15 PCs were used to calculate uniform manifold approximation and projection (UMAP) by using RunUMAP function of the same package. Six clusters were identified, and a cluster of erythrocytes (which expressed genes such as \textit{HBA1}, \textit{HBAD}, \textit{HBBA} and \textit{HBBR}) was removed. For the remaining clusters, signature scores of the W-chromosome genes were calculated, and clusters with positive scores were annotated as male PGCs, whereas those with negative scores were annotated as male PGCs.

Count matrices of the samples (E2.5–1 w) for each sex were aggregated. Further normalization, HVG selection, dimensionality reduction, and clustering were performed as described above, with ten PCs of 434 HVGs for count matrix and resolution = 0.6 for males; and 750 HVGs, resolution = 0.6 for females. Clusters of putative doublets and low-quality cells were removed. The remaining cells were reclustered by using the method described above. Cell-cycle-specific clustering biases were regressed out during z-scaling of gene expression for the final 12 clusters of male samples and 16 clusters of female samples.

### 2.13. scRNA-seq data analysis of chicken germ cells

By calculating and clustering signature scores of Kyoto encyclopedia of genes and genomes (KEGG) gene sets, clusters from male germ cells were grouped into four developmental stages and those from females, into five stages. Differentially expressed genes (DEGs) of each stage were identified by using FindAllMarkers function of Seurat R package with default parameters. To identify biological processes enriched at each developmental stage, significantly upregulated or downregulated gene ontology biological process (GOBP) terms (P < 0.05) were selected with significant (P < 0.05) DEGs by using the topGO version 2.40.0, R package with the org.Gg.eg.db version 3.11.4, annotation data package.

Developmental trajectories of the chicken germ cells were estimated using the Palantir version 0.2.8, python package [23]. Diffusion components (DCs) were computed by using the run_diffusion_maps function of the package, with the first five PCs for males; and the first 20 for females. The kNN graph was constructed by using the first five DCs for males; and the first ten for females, and the cells were visualized on the t-SNE plot based on the same DCs. Starting cells for calculating pseudotime were defined by choosing the cell with the lowest expression of \textit{FAP} (a cell-cycle-arrest-related gene) for male germ cells and the cell with the highest expression of \textit{CXCR4} (a migration-related gene) for female germ cells. The pseudotime and branch probabilities of the trajectories were specified by using the run_palantir function of the Palantir python package with num_waypoints = 750. Female germ cells were assigned into one of two developmental fates by binomial sampling based on the calculated branch probabilities.

To calculate the transcription factor (TF) activities along pseudotime, chicken-specific TFs were retrieved from the AnimalTFDB3.0 [24]. Putative target genes for 485 TFs in males and 488 TFs in females were specified by using the algorithm for the reconstruction of accurate cellular networks with adaptive partitioning (ARACon-AP) software [25]. Positively regulated target genes were identified using ssmarina, version 1.0.0, R package [26]. To calculate activity scores for the TF modules, the AUCell_ function and visualized by heatmap.

To specify the bifurcation point of the developmental trajectories of female germ cells, the cells were ordered over pseudotime. Cells were grouped into 13 bins along the pseudotime. Eleven bins were defined to contain 1,002 cells each, and the remaining two bins were then assigned to contain 1,003 cells each. Pseudotime bins responsive to the developmental fates were prioritized on the basis of the cross-validation area under the receiver operating characteristic curve (AUC) of the random forest classifier implemented in the Augur, version 1.0.0, R package [28] with default parameters. DEGs between path 1 and path 2 cells after pseudotime bin 9 were identified by using the FindMarkers function of the Seurat R package.

### 2.14. Interspecies comparison of scRNA-seq data

The transcripts per kilobase million count tables of developing human gonadal cells were gathered from the NCBI GEO database (GSE86146) and log2-transformed. One thousand HVGs for males and 1000 for females were identified by using the methods described above. \textit{DAZL}, \textit{DDX4}, \textit{DND1}, and \textit{NANOG}-positive clusters from the gonad samples were annotated as fetal germ cells (FGCs) and were then extracted. Clustering and dimensionality reduction were done by using the methods described on 20 PCs, with resolution = 0.8. Then, eight identified clusters of male FGCs were grouped into three stages and eight clusters of female FGCs into four stages, by using the marker genes described by Li et al [29].

To compare development of chicken germ cells with that of human FGCs, 12,048 orthologs were identified for males and 12,009 orthologs were identified for females. For orthologs from male chicken germ cells (262 HVGs) and female chicken germ cells (565 HVGs), count matrices of chicken germ cells and human FGCs were extracted. The five chicken germ cells closest to each human FGC were identified by using the knn.index.dist function of the KernelKnn version 1.1.0, R package. The average of UMAP coordinates for the nearest neighbor were used to visualize chicken germ cells and human FGCs on a single scatter plot. Each human FGC was annotated as having the same developmental stage as that shared by most (three of five) of the closest chicken germ cells.

Developmental trajectories of human FGCs were estimated by using the Palantir python package [23]. DCs were computed by using the run_diffusion_maps function of the package, with the first 100 PCs for both males and females. The kNN graph was constructed by using the first 50 DCs for males and the first 25 DCs for females. Cells were visualized on the t-SNE plot based on the same DCs. Starting cells for calculating pseudotime were defined by choosing the cell with the highest expression of \textit{POU5F1} (a pluripo-
tency and early-germ-cell marker) for both male and female germ cells. The pseudotime and branch probabilities of the trajectories were specified by using the run_palantir function of the Palantir python package with num_waypoints = 250. The dynamic expression pattern of intersection of HVGs for chicken germ cells and for human FGCs were smoothened over developmental pseudotime. The Z-scores were hierarchically clustered by using the hclust function and visualized by heatmap. Genes that showed convergent or divergent expression patterns were used to select GOBP terms (P < 0.05) by using the topGO, version 2.40.0, R package with the org.Hs.eg.db version 3.11.4, annotation data package.

3. Results

3.1. Genome editing to produce DAZL::GFP chicken PGCs

To trace PGCs during chicken development, we inserted GFP into the last intron of DAZL by using CRISPR/Cas9-NHEJ-mediated genome editing, as described [30]. First, we constructed a CRISPR/Cas9 plasmid to target the last intron of chicken DAZL, without affecting endogenous DAZL expression. Next, to tag DAZL with a GFP-expression cassette, we constructed a donor plasmid comprising portions of the last intron and exon (but not the stop codon) of DAZL (+12,099 to + 12,292) and a cassette to express GFP fused with the T2A self-cleavage peptide. The donor plasmid also included a neomycin-selection marker. To express GFP from the endogenous promoter of the modified DAZL allele, we transplanted the constructed donor plasmid (including the gRNA-recognition sequences) and the CRISPR/Cas9 plasmid into chicken PGCs (Fig. 1A). Seven days after transfection, GFP expression was detected in several PGCs (Fig. 1B). GFP expression in PGCs was induced for an extended period after G418 drug selection (Fig. 1C). Flow cytometric analysis revealed that 70.7% of PGCs expressed GFP (Fig. 1D). We confirmed that we had tagged DAZL by using performing knock-in-specific PCR (Fig. 1E and Table S1). TA sequencing confirmed that the PGCs genomic DNA was edited by NHEJ, with and without indels (Fig. 1F).

The DAZL::GFP PGCs expressed normal levels of DAZL mRNA (as determined by RT-PCR and qRT-PCR) (Fig. 1G, H and Table S1) and DAZL protein (as determined by immunocytochemistry) (Fig. 1I). Besides, we examined the normal expression of candidate pluripotency marker genes (POUV and NANOG), PGC formation genes (TBXT and PRDM14), germ cell marker genes (DDX4 and PIWIL1), epigenetic marker genes (DNMT3B and HDAC3), somatic marker gene (SUZ12), apoptosis marker gene (FADD), and DNA repair gene (H2AFX) in the DAZL::GFP PGCs by RT-PCR (Fig. 1J and Table S1). Among the examined genes, POUV, NANOG, PRDM14, DDX4, and PIWIL1 are the DAZL interacting genes [31]. Also, the expression of FADD and H2AFX reinforces the previous report that the chicken PGCs express lower levels of apoptosis genes and higher levels of DNA repair genes [32]. These results indicate that the DAZL::GFP PGCs are a valuable resource to produce transgenic chickens by germline transmission.

3.2. Production of DAZL::GFP transgenic chickens by germline transmission

To test germline transmission of DAZL::GFP and to produce transgenic chickens, we transplanted genome-edited PGCs into Korean-Ogye-breed recipient embryos at HH stages 14–17. After the recipient males sexually matured, we confirmed donor-derived sperm and produced genome-edited progeny by test-cross analysis. In total, 24 donor-derived progeny (8.2%) were produced from one germline-chimeric chicken (#1066). Thirteen of these progeny (54.2%) had edited genomes (Table S2). We also identified genome-edited progeny (Fig. 2A) by using knock-inspecific PCR (Fig. 2B). Sequencing the amplicons from transgenic chicken revealed modified alleles either with or without indels at the junctions of exogenous and endogenous sequences (Fig. 2C). The transgenic chicken genotypes were identical to those of the genome-edited PGCs (Fig. 1F). To identify off-target effects, we sequenced eight off-target candidate genes in the transgenic chickens by TA cloning, which demonstrated no mutations in these selected genes (Table S3).

Next, we mated a transgenic male rooster (G1) with WT hens, and examined GFP expression at different embryonic stages in the ensuing progeny by using fluorescence microscopy. As expected, GFP expression was specifically detected in PGCs that migrated to the germinal crescent region at HH stage 4 (18–19 h of development) and thereafter in PGCs that settled in the embryonic gonads at HH stage 27 (E6). GFP was expressed in germ cells of both male and female embryonic gonads at HH stage 35 (E9) (Fig. 2D). Immunostaining of gonad sections of male and female transgenic chicken (at E6, E8, E12, E16, hatch, and 1 week post-hatch) to detect GFP expression in germ cells demonstrated, as expected, GFP expression in the male germ cells, which were pervasive in the early gonads (E6–E12) but restricted to the sex cords in the late gonads (E16–1 w). Similarly, we detected GFP expression in the female germ cells, which migrating to the cortex region of early gonads (E6) and which settled in the cortex region of the late gonads (E8–1 w) (Fig. 2E). These results indicate that the DAZL::GFP transgenic chicken model is valuable for tracing and isolating germ cells at various developmental times.

3.3. scRNA-seq reveals distinct developmental stages of chicken germ cells

To define developmental stages during chicken germ-cell development, we performed scRNA-seq of GFP+ and GFP− DAZL::GFP germ cells isolated from transgenic chicken’s blood circulation (at E2.5) and from gonads (at E6, E8, E12, E16, hatch, and 1 week post-hatch) (Fig. 3A and Table S4). We profiled 4,752 male and 13,028 female germ cells that had fulfilled our quality-control criteria (Table S4, Fig. S1A-C). Unsupervised graph-based clustering identified 12 clusters for male germ cells and 16 clusters for female germ cells, which were visualized by using UMAP plots (Fig. S1D). We classified clusters into four stages for male (mS1-mS4) germ cells and five stages for female (fS1-fS5) germ cells on the basis of cluster-specific signature scores of KEGG pathways (Fig. 3B, Fig. S1-D-F, Table S5). In males, each of these stages was enriched for cells at specific developmental time points. mS1 was enriched for cells at E2.5 and E6; mS2, for cells at E8 and E12; mS3, at E16 and hatch; and mS4, at 1 week post-hatch. Female germ cells at fS1 were enriched for cells at E2.5, E6, and E8; fS2, at E12; fS3, at E16 and hatch; fS4, at hatch and 1 week post-hatch; and fS5, at 1 week post-hatch (Fig. 3B). Based on the basis of the activated signature scores and the associated developmental time points of each stage, we annotated male stages as follows: mS1, migrating PGCs; mS2, mitotic gonocytes; mS3–mS4, mitotic-arrested prospermatogonia [7,33]. We annotated female stages as follows: fS1, migration to differentiating PGCs; fS2, mitotic to retinoic acid (RA)-responsive oogonia; fS3, RA-responsive to meiotic-arrest oocytes; fS4, meiotic-arrest to primordial-follicular oocytes; and fS5, primordial-follicular to growing-follicular oocytes [7,33-35].

In male germ cells, we observed that mS1 cells were distinguished by upregulation of focal-adhesion pathway, adhesion-junction pathway, and Wnt-signaling pathway, which are important for PGC migration (Table S5) [36,37]. mS2 cells and mS3 cells were distinguished by upregulated Hedgehog-signaling pathway (which is involved in the mitotic phase in mS2 cells) [38] and steroid-biosynthesis pathway (which is related to cell-cycle arrest
CRISPR/Cas9-NHEJ-mediated genome editing to produce DAZL::GFP chicken primordial germ cells (PGCs). (A) CRISPR/Cas9 plasmid targeting the last intron of DAZL (DAZL #1) and donor plasmid containing the last intron and exon of DAZL (+12,099 to +12,292) in frame with a T2A peptide and green fluorescent protein (GFP) were cotransfected into cells to drive GFP expression from the endogenous DAZL promoter. Insertion did not disrupt the terminal DAZL exon. Blue bars indicate the gRNA-recognition sequence; red bars, the protospacer adjacent motif. (B–C) GFP expression in chicken PGCs seven days after transfection (b', b'') and three weeks after transfection (c', c'') with drug selection. (D) Flow cytometric analysis of genome-edited PGCs. (E–F) Analysis of targeted sites in chicken PGC genomes by using knock-in-specific PCR, and by sequencing TA-cloned amplicons. (G–H) Validation of endogenous DAZL mRNA expression by RT-PCR and qRT-PCR. In (H), the relative expression of DAZL was calculated by normalizing to both GAPDH expression and the control sample. Differences between samples were non-significant (ns), as determined by t-test. (I) Validation of endogenous DAZL protein expression via immunocytochemistry. (J) Examination of several candidate mRNA expression by RT-PCR. Wild-type (WT) PGCs were used as controls in (D, E, G, H, I, and J). Scale bars in (B, C, and I) = 100 μm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
in mS3 cells) [39]. mS4 cells were distinguished by a strong transcriptional signature of the MAPK-signaling pathway, which is critical at different steps of male germ-cell development (including spermatogenesis, sperm maturation, and sperm activation) [40]. In female germ cells, we also found stage-specific activation of KEGG pathways (Table S5). FS1 cells were distinguished by activated Wnt- and TGFβ-signaling pathways, which are important for PGC proliferation and self-renewal [41,42]. FS2 cells (during oogonia development) were distinguished by activation of several DNA-repair pathways that maintain the genetic stability of germ cells [43,44] in the developing oogonia. FS3 cells were distinguished by activation of the MAPK-signaling pathway, which is in accord with the importance of RA for initiating meiosis in female germ cells [45,46]. FS4 cells had elevated G protein-coupled receptor-related sphingolipid-metabolism pathway, phosphatidylinositol-signaling pathway, and calcium-signaling pathway. These observations are in accord with the importance of both G protein-coupled receptor-related pathway and calcium concentration for the meiotic-arrest phase in oocytes [47–49]. Finally, FS5 cells were enriched for pyruvate metabolism and steroid hormone biosynthesis, which are required for meiotic maturation and folliculogenesis [50,51].

To further characterize the molecular features of each stage, we identified stage-specific upregulated genes (adjusted P < 0.05; Fig. 3C and Table S6) and their enriched gene-ontology biological processes (Fig. 3D and Table S7). We found that heat-shock response gene expression was strongly enriched in mS1 and FS1 germ cells annotated as migrating PGCs, which is in line with the role of heat-shock proteins in the zebrafish PGC migration [52]. Gene sets involved in mitotic cell-cycle processes were upregulated in mS2 and FS2 cells and gene sets involved in meiosis were upregulated in FS4 cells. FS5 germ cells were enriched for expression of apoptosis-related genes (such as Becn1, Bnip3, Tollip, Atg3, and Anxa7) (Fig. S1G) [53–56]. We examined the expression (by qRT-PCR) of selected scRNA-seq-derived stage-specific markers in GFP-expressing germ cells: Lin28a for mS1 and FS1 germ cells; Kif4a, for mS2–mS3 cells; Tura3e, for FS2–FS4 cells; Col6a2, for mS4 cells; and Anxa7, for FS5 cells (Fig. 3E, Table S1). We measured expression levels of several housekeeping genes to identify an appropriate normalization control. Although expression levels of commonly used housekeeping genes (including Pdcd4) fluctuated over developmental stages (Fig. S2), we selected Sec61b (an essential factor for the embryonic development and protein trafficking in the oocyte of Drosophila) [57] for normalization. Sec61b expression was constant among developmental stages (Fig. S2).

In addition, we compared the transcriptomes of the mS1 and FS1 cells to provide insights into germ cell sex determination. Our analysis revealed that the mS1 and FS1 cells overlap at E2.5, E6, and E8 (Fig. S1A), however, the mS1 cells show a greater number of significantly upregulated genes (Fig. S3). Especially, genes related to sex determination (Sox9) [58], de novo DNA methylation (Dnmt3b) [17], and PGCs migration (Cxcr4) [59] were enriched in mS1 compared to FS1 (Table S8). Together, these data provide a temporal single-cell molecular atlas of gene expression during chicken germ-cell development and define the developmental stages of germ cells at a finer resolution.

3.4. Developmental trajectories and regulators mediating stage transition of chicken germ cells

We next investigated the developmental trajectories of male and female chicken germ cells by using Palantir [23]. We found that male germ cells linearly transitioned from mS1 to mS4. This trajectory was characterized by gradual acquisition of stage-specific TF activities (for example, during mS1, Sox11, Rcor3, Arid3a, Foxn2, Hlf, and ENSGALGO00000021665; during mS2, Tfdp2, Msx1, and Znfs21; during mS3, Smad6 and Ngflc2; and during mS4, Tftp2d, Iki2f1, St18, and Twist2) (Fig. 4A–C, Table S9). Of the mS1-specific TFs, Sox111 regulates progenitor- and stem-cell behavior during embryogenesis [60]; Rcor3 epigenetically modulates the pluripotency-related gene Nanog in PGCs [61]; and Arid3a regulates different pluripotency genes (Oct4, Sox2, and Nanog) in embryonic stem cells [62]. These results reflect the pluripotent stem-cell characteristics of mS1 germ cells. The two TFs related to the BMP4-signaling pathway were activated at mS2 (Msx1) and mS3 (Smad6). Mdx1 is critical for PGCs migration and proliferation [63]. Smad6 expression is lower in mouse gonocytes but higher in spermatogonia and spermatocytes [64], supporting our annotation of mS3 germ cells as mitotically arrested in a differentiated state.

By contrast, the trajectory of female germ cells bifurcated at FS3 (specifically at E16) to produce either a mixed population of FS4 and FS5 cells (path 1) or only FS4 cells (path 2) (Fig. 4A and B, Fig. S4). Along these bifurcating trajectories, we observed dynamic changes in stage-specific TF activities (such as in F1, Eomes and Msx1; in FS2, Csd1e1, Ncor2, Zfhx3, Atf4, Ssrp1, Hmgcb3, and Gabpa; in FS3, Dmrt1; in FS4, Cerss, Sub1, Six6, and Abnt2; and in FS5, Gli1) (Fig. 4C, Table S9). Of the FS1-specific TFs, Eomes is a stem-cell-related factor [65], and Msx1 is critical for PGCs migration and proliferation [63]. These results suggest that the stem-cell characteristics of FS1 germ cells resemble those of mS1 cells. Msx1 (which was also activated in mS2 cells) is also involved in meiosis initiation [66]. This suggests that whereas MSX1 is involved in PGCs migration in both sexes, MSX1 may specifically prepare female germ cells for early entry into meiosis. During chicken embryogenesis, female germ cells first enter meiosis, whereas male germ cells enter mitotic arrest [7,73]. Of the FS2-specific TFs, we found stem-cell-related factors (Csd1e1 and Gabpa) [67] and oocyte-priming factor (Hmgcb3) [29], indicating asynchronous and heterogeneous germ-cell development. Furthermore, Hmgcb3 (one of the top candidates for regulating the gene-expression network in the RA-responsive phase of human fetal germ cells [FGCs]) [29] was specifically expressed in FS2 cells. This indicated its involvement in both human and mouse oogenesis. The FS3-specific TF Dmrt1 is predominantly important for Sertoli-cell development and spermatogonia in the testis. However, Dmrt1 is also important for developing oogonia and oocytes (until meiosis entry) in the fetal ovary [68,69]. This indicates that the roles of Dmrt1 in male germ cells differ from those in female germ cells. Of the FS4-specific TFs, Sub1 is one of the eight TFs essential to trigger oocyte growth at the transition from primordial to primary follicle. Sub1 swiftly converts pluripotent stem cells into oocyte-like cells competent for fertilization and subsequent cleavage [70]. The FS5-specific Gli1

Fig. 2. Cisreg/Rca9-NheJ-mediated production and validation of Dazl::Gfp transgenic chickens. (A) Donor-PGC-derived progeny were distinguished from WT chickens by feather color. (B–C) Representative results of knock-in-specific PCR of targeted sites, and sequencing of TA-cloned amplicons. (Blue and red bars are used as in Fig. 1.) The sequences of chicken genomic DNA and of the donor plasmid (including the terminal Dazl exon) are presented. A WT chicken was used as a control in (B). (D) GFP expression was observed (by using a fluorescence microscope) in PGCs migrating to the embryonic germinal crescent at HH stage 4 (18–19 h of development); PGCs settled in the embryonic gonads at HH stage 27 (Embryonic day [E] 6); and the germ cells of male and female embryonic gonads at HH stage 35 (E9). The fluorescence signal adjacent to the germinal crescent and gonadal regions was due to autofluorescence. (E) GFP expression was examined by immunohistochemistry in sections of male and female gonads at E6, E8, E12, E16, hatch, and 1 week post-hatch. Scale bars in (D) and (E) = 100 µm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
is a maternal TF enriched in oocytes and one-cell-stage embryos. 

**GLIS1** effectively and directly promotes reprogramming of somatic cells into induced pluripotent stem cells (iPSC) [71]. Moreover, chicken oocytes store maternal factors (including sTFs) for zygotic genome activation and specification of the next generation of PGCs [72].

In the female germ cells, the bifurcating trajectories (paths 1 and 2) were distinguished by mutually exclusive patterns of gene expression. These cells developing along path 1 expressed genes related to apoptosis (ACTB, MAPK8, and ATF4) [73–75]; and those along path 2, meiosis (RBX1, CALM2, and ANAPC10) [76–78]. This observation supported our annotation of these two stages (Fig. 4D and E, Table S10). Moreover, we examined the expression (by qRT-PCR) of above apoptosis and meiosis-related genes in the DAZL::GFP germ cells isolated from the E16, hatch, and 1 week post-hatch ovaries (Fig. 4E). PGCs, oogonia, and oocytes apoptose at various stages of mammalian meiosis. However, a substantial proportion of oocyte apoptosis occurs during meiotic prophase I (mainly in the zygote/pachytene stage) in the fetal and early postnatal ovaries (before follicle formation) [79]. Oocytes may apoptose at this stage because of limited levels of growth factors necessary for oocyte survival and because of defects in chromosome crossover [79,80]. Moreover, oocyte apoptosis occurs in fetal and early postnatal ovaries because of the breakdown of germ-cell cysts and the formation of primordial follicles [81]. Although there is no evidence of this in chickens, our results are consistent with these observations in mammalian fetal and early postnatal ovaries.

To provide further insight into oocyte meiosis, we first analyzed the expression patterns of candidate meiotic initiators in the chicken female germ cells developing along path 2. Our analysis revealed that the meiosis initiators PLCB1, MSX1, and RDH10 express in a population of fS1 (Fig. 5A), earlier than the previously known pre-meiotic marker SIRAR8, which is triggered by retinoic acid and essential for the initiation of meiosis [34]. Next, we sub-clustered the fS3–fS5 cells and analyzed the expression patterns of candidate meiotic prophase I markers. Our analysis revealed 14 sub-clusters in fS3–fS5 (Fig. 5B). By the expression of meiotic prophase I markers, SPO11 (leptotene marker), RAD21LI (zygotene marker), and PIWIL2 (pachytene marker) were particularly expressed relatively high in sub-clusters 4, 6, 9, 10, and 13 (mostly fS3–fS4 cells) (Fig. 5B–C), indicating the accumulation of meiotic-arrested oocytes at these stages. After entry into meiosis, oocytes are usually arrested in the meiotic prophase I and maintained in the primordial follicles, which forms within few days after hatch in chicken [35]. We have also examined (by qRT-PCR) the expression of MSX1, SPO11, and RAD21LI in the DAZL::GFP germ cells isolated from the E12, E16, and hatch ovaries (Fig. 5D). Finally, we sub-clustered the fS5 cells and checked the expression patterns of genes related to apoptosis (ACTB, MAPK8 and ATF4) and meiosis (RBX1, CALM2 and ANAPC10). We found that the subpopulation structure of fS3 cells is largely explained by the expression of MAPK8 (Fig. 5E–F). Together, these results identify stage-specific TF activities that govern the progression of chicken germ cells through developmental stages and that define the path of oocyte meiosis and apoptosis at a finer resolution.

3.5. Comparison of chicken and human germ-cell stages reveals species-specific gene expression

To determine whether the developmental stages and trajectories of chicken germ cells resemble those of human FGCs [29], we compared scRNA-seq data from these cell types by clustering gene expression of male FGCs (4–25 w after fertilization) and female FGCs (5–26 w after fertilization) and using the clusters to define three developmental stages in male cells and four stages in female cells (Fig. 6A, Fig. S5A and S5B). We mapped the human FGC data onto the chicken germ-cell data, and mapped the chicken germ-cell data onto the human FGC data. For males, we mapped human mS1 to chicken mS1; human mS2 to chicken mS2; and human mS3 to chicken mS2–mS4. For females, we mapped human fS1 to chicken fS1; human fS2 to chicken fS2–fS4; human fS3–fS4 to chicken fS4 (Fig. 6B and C, Fig. S6A–C, Table S11).

We then compared the inferred developmental trajectories of human FGCs with those of chicken germ cells by manually inspecting expression dynamics in each species. In males, of 56 orthologs dynamically expressed along the human or chicken trajectories, we identified 32 convergent (evolutionary conserved) and 24 divergent (species-specific) genes. In females, we identified 83 convergent and 36 divergent genes (Fig. 6D, Table S12). The convergent and divergent genes in both sexes were significantly enriched for various cell-cycle-related functions, reflecting the fact that proliferation is essential for germ-cell development. As expected, female-specific convergent and divergent genes were additionally enriched for meiosis-related functions (Table S13).

To add further insight into the germ cell biology between the chicken and humans, we endeavored to find out the general germ-cell-specific genes and germ-plasm-specific genes among the convergent and divergent genes. Even though most germ-cell-specific genes were not dynamically expressed along the male trajectory, TCF7L2 (a mitotic-arrest related gene) [82] was identified as a divergent gene (Fig. 6D). We also identified female-specific convergent genes associated with pluri potency, germ-cell development, and meiosis (including SAL4, KIT, SPO11, SYCPI, SYP2 and SYCE3) (Fig. 6D) [33,83–86]. DDX4 and DAZL are well-known germ-cell-specific as well as germ-plasm-speciﬁc marker genes in chicken [3,4]. In our analysis of germplasm genes, DDX4 and TDRKH (which is not well-known in chicken) were identiﬁed as divergent genes in the male germ cells between chicken and human, while DDX4 and DAZL were identiﬁed as divergent genes in the female germ cells between the species. Most of the other germplasm genes, including MAEL, were identiﬁed as convergent genes in both male and female germ cells between the species (Fig. 6E). Some of the chicken germ-plasm-speciﬁc genes, including DAZL, are not required for the speciﬁcation and migration of PGCs in mammals since the mammals follow epigenesis mode of germ cells speciﬁcation, whereas chickens follow preformation mode of germ cells speciﬁcation [2,87]. Together, these comparisons of human-gene expression with chicken-gene expression during germ-cell development identiﬁed several evolutionarily conserved cell-cycle-related functions and diverse species-speciﬁc biological processes.
Fig. 4. Developmental trajectories and regulators that mediate chicken germ-cell transitions between stages. (A) t-SNE plot showing the developmental pseudotime and annotated stages of male (left) and female (right) germ cells. Cells are colored by calculated pseudotime and by stages. (B) Schematic of developmental time points, stages, and trajectories of germ cells. (C) Heatmaps displaying the stage-specific transcription factor (TF) activities over the calculated pseudotimes for male (left) and female (right) germ cells. The dotted lines segment TFs active in specific stages. Histogram above the heatmaps showing the cell density of each developmental stage in male and female germ cells, respectively. The alphabet on the left side of the heatmaps indicate the cluster of the genes with similar dynamic expression patterns along pseudotime. (D) Volcano plot illustrating the DEGs between developmental path 1 and path 2 of female germ cells. Blue dots indicate genes upregulated in path 1; red dots, in path 2. (E) Violin plots showing the expression levels of apoptosis pathway genes enriched in path 1 (left) and of meiosis-pathway genes enriched in path 2 (right) in female germ cells calculated from scRNA-seq data; bar plots showing normalized expression of the same genes examined by qRT-PCR. For qRT-PCR analysis, the relative expression of genes was calculated by normalizing to the expression of SEC61B and reference sample, and one-way ANOVA with Tukey’s multiple comparisons test was applied to calculate statistical significance. ** P < 0.01, **** P < 0.0001, ns – no significant difference. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
4. Discussion and conclusion

Transcriptome analyses of chicken germ-cell developmental stages have been limited and fragmentary [31,88], and these analyses have not been applied in vivo. The main reasons for these limitations are the lack of a technology to discriminate germ cells and the lack of sophisticated markers to isolate these cells. Recently, scRNA-seq was used to compare undifferentiated germ cells with somatic cells; as well as to compare germ cells at different stages of sexual maturation [89–93]. In this study, we used scRNA-seq to produce a single-cell-resolution atlas of the dynamic transcriptional landscape during chicken germ-cell...
development (from undifferentiated to differentiated stages) in both sexes.

We developed a strategy to monitor and isolate chicken germ cells by inserting GFP into DAZL (a germ-cell-specific gene). Adopting a published method [94], we inserted a T2A-GFP cassette in frame with the last exon of DAZL by using CRISPR/Cas9. Successful editing was confirmed by monitoring GFP expression in the established PGC lines (DAZL::GFP) and no alteration of endogenous DAZL expression was observed. DAZL::GFP transgenic chickens were produced by transplantaing DAZL::GFP PGC lines, indicating that genome-edited PGCs can be transmitted through the germline. Furthermore, in E2 embryos of transgenic chickens, we detected GFP expression specifically in germ cells of both sexes at HH stages 4, 27, and 35. These results indicate that these transgenic genes are useful for monitoring germ-cell development. By isolating germ cells for scRNA-seq analyses, we could also characterize germ-cell-specific mechanisms during avian development. Thus, the CRISPR/Cas9-NHEJ-mediated gene-insertion strategy is a powerful method to edit chicken genomes and to reveal tissue-specific gene expression patterns in both avian and mouse models [95].

Using our method, we identified several cellular and molecular characteristics of germ-cell development in both sexes. Male germ cells were characterized into four developmental stages (mS1, migration stage; mS2, mitotic stage; mS3–mS4, mitotic-arrest stages), according to activated signature scores and the elapsed time of development [7,33]. Sometimes, female germ cells (particularly those in fS1–fS4) were classified simultaneously into multiple adjacent stages, indicating that germ-cell development in females is asynchronous and heterogeneous [7,29,33,35]. These results indicate that (in chicken) the development of female germ cells is more advanced than that of male germ cells and that many female germ cells are in states of transition. We identified LIN28A as an mS1 and fS1 marker. LIN28A is essential for proper PGC development [96], and LIN28A induces somatic cells to become PGCs [97]. Previous studies have attempted to distinguish the phenotypes of E2.5 PGCs from E6 PGCs [98,99], however, the transcriptomic difference has not been clearly identified. We found that the expression profiles of migrating PGCs (at E2.5) resemble those of colonized PGCs (at E6 in males and at E6–E8 in females). In mice, it is known that the transcriptomes of XX and XY cells overlap at the early developmental stage, whereas sex-specific branches are formed in the differentiated cells at the late stages [100]. However, in this study, chickens showed differences in transcriptome levels between mS1 and fS1 (early-stage germ cells).

We found that the gene-expression profiles of male germ cells differed from those of female germ cells after E8 (which corresponds to mS2 in male germ cells, and to fS3 in females). After E8, the number of mitotic male germ cells rapidly declined, and most cells begin to arrest at G0/G1 of mitosis [7]. Reflecting this phenomenon, the transcriptional signatures of cell division, of cellular response to growth factor stimuli, and of positive regulation of cell population proliferation were elevated in mS2 cells but not in mS3 cells. On the other hand, after E8 the female germ cells rapidly proliferated and begin to arrest at G2/M of meiosis [7]. Consistent with this, we observed that the G2/M transition was regulated and that fS2 cells actively divided during meiosis. These transcriptional signatures persisted into fS3, but did not persist into fS4 and fS5. Our analysis shows that the trajectory of female germ-cell development, particularly after fS3 (E16), bifurcates to produce apoptotic cells (from path 1) and meiotic cells (from path 2). In mammals, a substantial proportion of oocyte apoptosis occurs throughout meiotic prophase I in fetal and early postnatal ovaries. This is mainly because of limited levels of growth factors necessary for oocyte survival and because of elimination of oocytes with chromosome-crossover defects [79,80]. Also, the environment of ovaries before birth differs from that after birth (in relation to the breakdown of germ-cell cysts and the formation of primordial follicles). When the mouse cysts undergo programmed breakdown in the fetal ovary, approximately 33% of the oocytes survive to form primordial follicles [81]. We found that fS4–fS5 germ cells developing along path 1 were apoptotic. In chickens, the primordial follicle pool begins to develop within 4 days of hatch by germ-cell cysts breakdown and by enclosure of oocytes with pregranulosa cells [35].

A comparison of chicken germ cells with human FGCs [29] revealed sexually dimorphic similarities and differences. On the one hand, we matched the first two stages of human male FGC development (mS1 and mS2) with the first two stages of chicken male germ-cell development (mS1 and mS2). In later stages as well, the gene expression profiles of human male FGCs and chicken male germ cells were very similar. On the other hand, in females, the human fS2 stage matched with a broad range of chicken germ-cell stages (fS2–fS4). The similarity between human FGC and chicken germ-cell transcriptomes was lower in females than in males. Based on these results, further in-depth analyses of chicken germ-cell characteristics at each stage and of interactions between germ cells and somatic cells are required. In conclusion, we used a novel method to track germ cells and to produce a time-resolved single-cell-resolution atlas of chicken germ-cell development. Our results highlight the cellular and molecular characteristics of different stages in germ-cell development. These results could be valuable for using chicken as a model system to study germ-cell development.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by grants: J.Y.H. (National Research Foundation of Korea grant [NRF-2015R1A3A2033826] funded by the Ministry of Science, Information and Communication Technology, and Future Planning), J.K.K. (2018R1A5A1025511 and DGIST R&D program [20-CoE-BT-04] funded by the Ministry of Science and ICT of Korea), and H.J.L. (National Research Foundation of Korea grant [NRF-2018R1D1A1B07048893] funded by the Ministry of Education).

Fig. 6. Transcriptome-based comparison of developmental stages in chicken germ cells with those in human fetal germ cells (FGCs). (A) UMAP showing the developmental stages of human male (left) and female (right) FGCs. Cells are colored by developmental stage. (B) Scatter plots mapping human male FGCs onto the UMAP plot of chicken male germ cells (top); and human female FGCs onto chicken female germ cells (bottom). Gray points indicate chicken germ cells; colored points, original developmental stages of human FGCs. (C) Bar plots showing proportions of human FGCs in different stages that were mapped onto chicken germ-cell stages, corresponding to (B). (D) Heatmaps showing dynamic patterns of DEGs from chicken male germ cells and human male FGCs (left); and from chicken female germ cells and human female FGCs (right). (E) Heatmaps showing dynamic patterns of germ-plasm-specific genes from chicken male germ cells and human male FGCs (left); and from chicken female germ cells and human female FGCs (right). In D and E, histogram above the heatmaps showing the cell density of each developmental stages in male and female germ cells, and in each species, respectively. Gray lines between gene clusters connect chicken and human orthologs. Notable convergent (blue) and divergent (red) genes are indicated. Cf, Gallus gallus (chicken); Hs, Homo sapiens (human). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Author Statement

Deivendran Rengaraj, Dong Gon Cha, Han Jo Lee, and Young Min Kim analyzed the scRNA-seq data, investigated the results, and wrote the first draft of manuscript. Deivendran Rengaraj, Hong Jo Lee, Kyung Youn Lee, and Yoon Ha Choi prepared single-cells and libraries for scRNA-seq. Hong Jo Lee, Deivendran Rengaraj, Kyung Youn Lee, Kyung Min Jung, Young Min Kim, and Hee Jung Choi contributed to producing transgenic chickens and performed the experiments. Hyeon Jeong Choi, Eunhui Yoo, Seung Je Woo, Jin Se Park, and Kyung Je Park assisted in animal management, sample preparation, scRNA-seq data analysis and experiments. Jae Yong Hong and Jong Kyoung Kim conceived the project, supplied research materials, investigated the results, and critically checked the content in the draft. All authors have read and agreed to the published version of the manuscript.

Data and code availability

The single-cell RNA sequencing data have been deposited in the SRA database under the accession code PRJNA761874 and are available. The scripts and instances used for the analysis of the single-cell RNA sequencing data are uploaded in the GitHub: https://github.com/CB-postech/Chicken_GermCell. The source data underlying gel electrophoresis images and qRT-PCR are provided as a Source Data file with this paper.

Ethics approval

All experimental procedures and care of chickens were approved by the Institute of Laboratory Animal Resources, Seoul National University, Korea, and all methods were carried out in accordance with ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines and approved by the Institutional Animal Care and Use Committee (IACUC, SNU-190401-1-1) of Seoul National University, Korea.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Appendix A. Supplementary data

Supplementary data (Figures and Tables) to this article can be found online at https://doi.org/10.1016/j.csbj.2022.03.040.

References

[1] Kim YM, Han JY. The early development of germ cells in chicken. Int J Dev Biol 2018;62:145–52.
[2] Jamieson-Lucy A, Mullins MC. The vertebrate Balbiani body, germ plasm, and oocyte polarity. Curr Top Dev Biol 2019;135:1–34.
[3] Tsunekawa N, Naito M, Sakai Y, et al. Isolation of chicken vasa homolog gene during early development. Arch Histol Cytol 2001;64:493–500.
[4] Henry GV, Hameleerts B, Pecquet C, et al. The dynamic development of germ cells during embryogenesis. Poult Sci 1993;72:133–42.
[5] Yang SY, Lee HJ, Lee HC, et al. The dynamic development of germ cells during embryogenesis. Poult Sci 2018;97:650–7.
[6] Li L, Dong J, Yang J, et al. DNA methyltransferases in chicken primordial germ cells. PLoS ONE 2011;6:e19524.
[7] Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
[8] Lun ATL, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Research 2016;5:2122.
[9] Kato E, Takahashi N, Sato S, et al. Multiplex genome engineering through adaptive partitioning inference of mutual information. Bioinformatics 2016;32:2233–3235.
[10] Lefebvre C, Rajhandari P, Alvarez MJ, et al. A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. Mol Syst Biol 2010;6:377.
[11] Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and prediction of animal transcription factors. Nucleic Acids Res 2019;47:D31–8.
[12] Lachmann A, Giorgi FM, Lopez G, et al. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinformatics 2016;32:2233–3235.
[13] Horiuchi K, Segawa K, Kojima Y, et al. Identification of germ cell genes involved in gonadal differentiation by single-cell RNA sequencing. Dev Growth Differ 2017;59:540–9.
[14] Lee HD, Yoon JW, Jung KM, et al. Targeted gene insertion into zygotic genome of mouse germ cells using CRISPR/Cas9 system. Nat Biotechnol 2016;34:12–18.
[15] Li H, Liang Z, Yang J, et al. Dynamic analysis of DAZL and its interacting genes during germ cells specification and zygotenic genome activation in chickens. Int J Mol Sci 2020;21:8170.
[16] Zhang YH, Rengaraj D, Choi JW, et al. Expression pattern of meiosis associated genes in chicken primordial germ cells and germ-line cells. Theriogenology 2010;74:765–76.
[17] Satija R, Querec D, Satija R, et al. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 2015;33:495–502.
[18] Setty M, Kiseliovas V, Levine J, et al. Characterization of cell fate probabilities in single-cell data with PantaRhei. Nat Biotechnol 2019;37:451–60.
[19] Hu H, Miao YR, Jia LH, et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome Biol 2019:20:63.
[20] McCarthy DJ, Campbell KR, Lun AT, et al. scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 2017;33:1179–86.
[21] Li H, Liang Z, Yang J, et al. Dynamic analysis of DAZL and its interacting genes during germ cells specification and zygotenic genome activation in chickens. Int J Mol Sci 2020;21:8170.
[22] Richardson BE, Lehmann R. Mechanisms guiding primordial germ cell migration: strategies from different organisms. Nat Rev Mol Cell Biol 2010;11:37–49.
[23] Wang X, Guo G, Zhang X, et al. Effect of RFRP-3, the mammalian ortholog of GnIH, on the epidermis of male rats. Theriogenology 2018;118:196–202.
[24] Western PS, Miles DC, van den Bergen JA, et al. Dynamic regulation of mitotic arrest in female germ cells. Stem Cells 2008;26:339–47.
[25] Miles DC, van den Bergen JA, Sinclair A, et al. Regulation of the female mouse germ cell cycle during entry into meiosis. Cell Cycle 2010;9:468–18.
[26] Hughes CC. The population of germ cells in the developing female chick. J Embryol Exp Morphol 1963;11:513–36.
[27] Ménendez C, Carrasco E, Pedernera E. Adenohypophysis regulates cell proliferation in the gonads of the developing chick embryo. J Exp Zool A Comp Exp Biol 2005;303:179–85.
[28] Nakamura Y, Yamamoto Y, Usui F, et al. Migration and proliferation of primordial germ cells in the early chicken embryo. Poult Sci 2007;86:2182–93.
[29] Li H, Liang Z, Yang J, et al. DAZL is a master translational regulator of murine spermatogenesis. Natl Sci Rev 2019;6:455–68.
[30] Rengaraj D, Zheng YH, Kang KS, et al. Conserved expression pattern of chicken DAZL in primordial germ cells and germ-line cells. Theriogenology 2010;74:765–76.
[31] Sakuma T, Nishikawa A, Kume S, et al. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci Rep 2014;4:5400.
[32] Han JY, Park TS, Kim JN, et al. Gene expression profiling of chicken primordial germ cell ESTs. BMC Genomics 2006;7:220.
[33] Rengaraj D, Lee BR, Lee SI, et al. Expression patterns and miRNA regulation of DNA methyltransferases in chicken primordial germ cells. PLoS ONE 2011;6:e19524.
[34] Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
[35] Lun ATL, Riesenfeld S, Andrews T, et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome Biol 2019:20:63.
[36] McCarthy DJ, Campbell KR, Lun AT, et al. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 2017;33:1179–86.
[37] Liu AT, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Research 2016;5:2122.
[38] Satija R, Farrell JA, Gennert D, et al. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 2015;33:495–502.
[39] Setty M, Kiseliovas V, Levine J, et al. Characterization of cell fate probabilities in single-cell data with PantaRhei. Nat Biotechnol 2019;37:451–60.
[40] Hu H, Miao YR, Jia LH, et al. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. Nucleic Acids Res 2019;47:D31–8.
[41] Lachmann A, Giorgi FM, Lopez G, et al. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinformatics 2016;32:2233–3235.
