

RESEARCH

Open Access



# Single-cell RNA sequencing revealed the liver heterogeneity between egg-laying duck and ceased-laying duck

Xue Du<sup>1,2</sup>, Shujing Lai<sup>3</sup>, Wanqiu Zhao<sup>4</sup>, Xiaoqin Xu<sup>5</sup>, Wenwu Xu<sup>1</sup>, Tao Zeng<sup>1</sup>, Yong Tian<sup>1</sup> and Lizhi Lu<sup>1\*</sup>

## Abstract

**Background:** In the late phase of production, ducks untimely cease laying, leading to a lower feed conversion. Liver plays a vital role in the synthesis and transport of yolk materials during egg formation in birds. However, the molecular mechanism of liver in ceased-laying duck is far from clear, higher resolution and deeper analysis is needed. Sing-cell RNA-sequencing of  $10 \times$  Genomics platform can help to map the liver single cell gene expression atlas of Shaoxing duck and provide new insights into the liver between egg-laying and ceased-laying ducks.

**Results:** About 20,000 single cells were profiled and 22 clusters were identified. All the clusters were identified as 6 cell types. The dominant cell type is hepatocyte, accounted for about 60% of all the cells. Of note, the heterogeneity of cells between egg-laying duck and ceased-laying duck mainly occurred in hepatocytes. Cells of cluster 3 and 12 were the unique hepatocyte states of egg-laying ducks, while cells of cluster 0 and 15 were the unique hepatocyte states of ceased-laying ducks. The expression mode of yolk precursor transporters, lipid metabolizing enzymes and fibrinogens were different in hepatocytes between egg-laying duck and ceased-laying duck. *APOV1*, *VTG2*, *VTG1*, *APOB*, *RBP*, *VTDB* and *SCD* might be activated in egg-laying ducks, while *APOA1*, *APOA4*, *APOC3*, *FGB* and *FGG* might be activated in ceased-laying ducks.

**Conclusions:** Our study further proves that *APOV1* and *APOB* play key roles in egg production, rather than *APOA1* and *APOA4*. It is also the first to detect a correlation between the higher expression of *APOC3*, *FGB*, *FGG* and ceased-laying in duck.

**Keywords:** Ceased-laying duck, Single-cell RNA-seq, Cell heterogeneity, Hepatocytes, Cell cluster

## Background

Eggs are well received as an important source of nutrients all around the world, as they are rich in proteins, lipids, vitamins and minerals. In Asia, duck eggs are traditional ingredients in both fresh style and processed style. However, the per capita share of duck eggs is still low, which

results from farming scale with short peak laying period. In the group, some ducks meet untimely ceased-laying at the egg producing period which leads to a rapid downward trajectory of laying rate after the peak production. In late phase, ducks are known to be characterized by the declined production performance compared with those at peak production, resulting in a restricted economic benefit of egg-laying duck industry.

As a great metabolic factory, liver supports the energy demands of yolk formation in avian species. Liver provides lipids, proteins, vitamins, and choline, and so on, which are essential materials for yolk precursor synthesis [1]. Some studies detected the change of lipid metabolism

\*Correspondence: lulizhibox@163.com

<sup>1</sup> State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Institute of Animal Husbandry and Veterinary Medicine, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, Zhejiang, China  
Full list of author information is available at the end of the article



in liver between pre-laying and peak-laying of laying hens [2, 3]. In consequence, the laying performance of duck might be improved by affecting hepatic lipid metabolism. For instance, dietary betaine improves chicken egg-laying rate through hypomethylation and glucocorticoid receptor-mediated activation of lipogenesis related genes in liver [4]. Previously, we revealed lipid metabolic genes in liver relate to yolk ratio of duck [5]. Based on these studies, we speculated that liver plays heterogeneous functions in the late phase of lower production compared with peak production.

The balance between lipogenesis and export of very-low-density apolipoprotein (VLDL) particles in liver is extremely important for egg production [6]. Peroxisome proliferator activated receptor (PPAR) might play great role in liver metabolism in the switch from egg-laying period to ceased-laying period. The lipid metabolism-related genes such as fatty acid binding proteins (*FABPs*), lipoprotein lipase (*LPL*) and apolipoprotein A1 (*APOA1*) that mapped to PPAR signaling pathway were downregulated in the late phase of production of hens [7]. Hawthorn-leaves flavonoids could improve the reproduction performance of aged breeder hens (60-wk-old) through upregulating the mRNA expression of vitellogenin II (*VTG2*), Apolipoprotein B (*APOB*), and apolipoprotein-VLDL II (*APOV1*) in the liver [8].

*APOB* and *APOV1* are major components of yolk oriented VLDL (VLDL-yolk), which are produced by liver and function to transport lipids to the follicles for yolk deposition. The expression levels of these apolipoproteins dramatically elevate in laying hens compared with non-laying hens [9]. In sexually mature hens, enzymes participate in lipids synthesis and transport are basically controlled by sex hormones from hypothalamus, pituitary and other glands [10]. In hypothalamus, deep-brain photoreceptors (DBPs) extend fibers to the external zone of the median eminence adjacent to the pituitary to translate light information into downstream neuroendocrine responses [11]. Expression of thyroid-stimulating hormone  $\beta$  subunit (*TSHB*), deiodinase 2 (*DIO2*), organic anion transporting polypeptides (*OATPs*) are subsequently enhanced, which modulate gonadotropin-releasing hormone (GnRH) secretion from the hypothalamus [12–14]. GnRH regulates the release of gonadotropins, such as luteinizing hormone (LH), follicle-stimulating hormone (FSH), estrogen and progesterone hormones which subsequently arouses egg laying [15].

The conventional RNA sequencing of bulk sample helps to systematically assess gene expression level and reveal molecular regulatory networks in biological processes. It shows limitation for low-abundance cells, which unable to reveal key messages from complex and highly heterogeneous organisms. High-throughput single-cell

RNA sequencing (scRNA-seq) breaks the conventional bulk sample-based experimental object, can reveal genes expression state of individual cell and reflect the heterogeneity between cells. scRNA-seq has been widely applied in the studies of reproduction, immunization, tumorigenesis and differentiation. In recent years, the cell atlas of mouse and human have been published successively [16, 17], which illustrate the trajectory of cell development and gene expression dynamics in cell differentiation. In poultry science, using scRNA-seq, Li et al. [18] described the heterogeneity of chicken skeletal muscle at two developmental stages, and identified *APOA1* (apolipoprotein A1) and *COL1A1* (the type I collagen encoding gene) as biomarkers for chicken intramuscular fat cells. Shaoxing duck is a sexually mature early, high-yielding duck species. In this study, we constructed a single-cell resolution transcriptomic atlas of liver in laying and ceased laying ducks based on Shaoxing duck in the late phase of production. Our study might reveal the inner view of the mechanisms of untimely ceased-laying in duck liver, thereby providing potential targets for improving the performance of egg-laying ducks in the late phase of production.

## Results

### Sequences statistic and quality control

We obtained over 28.4 M reads in each sample. The sequence data showed a high quality, with valid barcodes over 96%, Q30 bases in barcode over 95%, Q30 bases in RNA read over 93% and Q30 bases in UMI over 92% (Supplementary Table S1). On the whole, over 86.6% of the reads matched to the reference, with an average of 538 and 640 genes detected per cell in L\_C and L\_L, respectively. Median number of genes detected per cell as a function of mean reads per cell was shown in Supplementary Figure S1. Over 29 k of mean reads were detected per cell, and over 16 k of the total genes were detected (Supplementary Table S2).

### Transcriptional clustering of cell populations

After quality filtering, a total of 19,096 single cells were captured, 10,583 were from L\_C and 8,513 were from L\_L (Supplementary Table S3 and Supplementary Figure S2). The 10 main factors were studied by principal components analysis, and the results were visualized in tSNE (Fig. 1A). Cells were clustered into 22 clusters. Normalized expression of the top variable marker genes from 22 clusters are shown in the heatmap (Supplementary Figure S3). Each cluster shows a high specific gene expression pattern. Details were shown in Fig. 1B. *ALB* (albumin), *TAT* (tyrosine aminotransferase isoform X1) and *CEBPA* (CCAAT/enhancer-binding protein alpha) are marker genes of hepatocytes [19, 20], *ALB* and *TAT* were highly expressed

in cluster 0,3,12,15, *CEBPA* were highly expressed in cluster 0,6,8,5,14. These faithfully defined cluster 0, 3, 5, 6, 8, 12, 14 and 15 as hepatocytes. Cluster 1 and 2 were marked by *CD3E* (glycoprotein CD3 epsilon chain) [21], indicating its T cells identity. Meanwhile, cluster 9 and 17 showed B cell characteristics, with *CD80* [22] and *CD24* [23] uniquely expressed. Interestingly, we founded a progenitor subpopulation (cluster 19) with extremely high level of *NCAMI* (neural cell adhesion molecule 1 isoform X8) [24]. In addition, cluster 4 and 20 are defined as endothelial cells by genes like *CD34*. High expression level of *FLT* (vascular endothelial growth factor receptor) [25] indicates cluster 11 and 21 are endothelial precursor cells. According to specific gene expression pattern, cluster 7, 10, 13, 16, 18 were defined as neuroendocrine cell, podocyte, megakaryocyte, fibroblast and natural killer T cell, respectively [26–28]. Hepatocytes of cluster 3 and 12 were unique states in egg-laying ducks, while hepatocytes of cluster 0 and 15 were unique states in ceased-laying duck. This is the main difference between laying and ceased-laying ducks in liver-cell heterogeneity, while in other cell types, they were almost in homogeneity.

Among all the cell types of two experimental samples, hepatocytes were in dominant (Fig. 1C). In egg-laying duck, hepatocytes accounted for 43.00%, T cells accounted for 26.27%, Endothelial cells accounted for 8.52%, B cells accounted for 6.04%. In ceased-laying duck, hepatocytes accounted for 47.00%, T cells accounted for 19.34%, Endothelial cells accounted for 6.86%, B cells accounted for 5.48%. In general, the percentages of different cell types were similar between two samples. The difference between hepatocytes from two samples were shown in Fig. 1D. Among hepatocytes, cluster 0 and cluster 15 occupied 51.43 and 6.83% in L\_C, respectively, while they occupied only 0.98 and 0.12% in L\_L. Cluster 3 and cluster 12 occupied 40.96 and 14.41% of hepatocytes in L\_L, respectively, while they occupied only 1.12 and 1.40% in L\_C. These suggested cluster 0 and cluster 15 were mainly presented in ceased-laying duck, while cluster 3 and cluster 12 were mainly presented in egg-laying duck.

Gene expression heatmap of top 10 marker genes in cluster 0, 3, 12 and 15 was shown in Fig. 1E. For ease of description and understanding, all the genes were divided into 4 parts, I, II III and IV, which were marked in Fig. 1E. From the comparison, cluster 3 and 12 in L\_L had higher expression in marker genes of group I and group III, while cluster

0 and 15 in L\_C had higher expression in marker genes of group II and group IV. Egg-laying duck highly expressed genes code yolk precursors or transporters, like very-low-density apolipoprotein II (*ApoVLDL-II*, *APOV1*), vitellogenin I (*VTG1*), vitellogenin II (*VTG2*), apolipoprotein B (*APOB*), riboflavin-binding protein (*RBP/RTBDN*), vitamin D-binding protein (*VTDB/GC*), enzymes in hepatic lipid disposal and in the metabolic utilization of fatty acids, like stearoyl-acyl-CoA desaturase (*SCD*), liver-type fatty acid-binding protein (*FABP1*) and heart-type fatty acid-binding protein (*FABP3*), and enzymes catalyzing intra oocytic break down of protein components of both vitellogenin and VLDL, like cathepsin E (*CTSE*). Ceased-laying duck highly expressed gallinacin-10 (*GAL10*), fibrinogens (*FGB*, *FGG*), apolipoproteins (*APOA1*, *APOA4*, *APOC3*) and genes related to lipid metabolism, like cytochrome P450 (*CYP3A8*, *CYP2H1*), mid1-interacting protein 1-B (*MIIIB*) and fatty acid synthase (*FAS/FASN*), and amino acid metabolism related genes, like 4-hydroxyphenylpyruvate dioxygenase (*HPD*), glutathione S-transferase (*GSTM*), and guanidinoacetate N-methyltransferase (*GAMT*).

### Functional analysis of differentially expressed genes

#### KEGG analysis

The 20 most enriched pathways of cluster 0, 3, 12 and 15 were shown in Fig. 2A–D and Supplementary File 1. All these hepatocytes participated in basic hepatic metabolic pathways, such as primary bile acid biosynthesis, fatty acid degradation, and glycine, serine and threonine metabolism. Cluster 0 and 3 showed a higher function in oxidative phosphorylation than cluster 15 and 12. Genes of Cluster 12, from L\_L, enriched in primary bile acid biosynthesis and biosynthesis of unsaturated fatty acids.

#### Protein interaction analysis

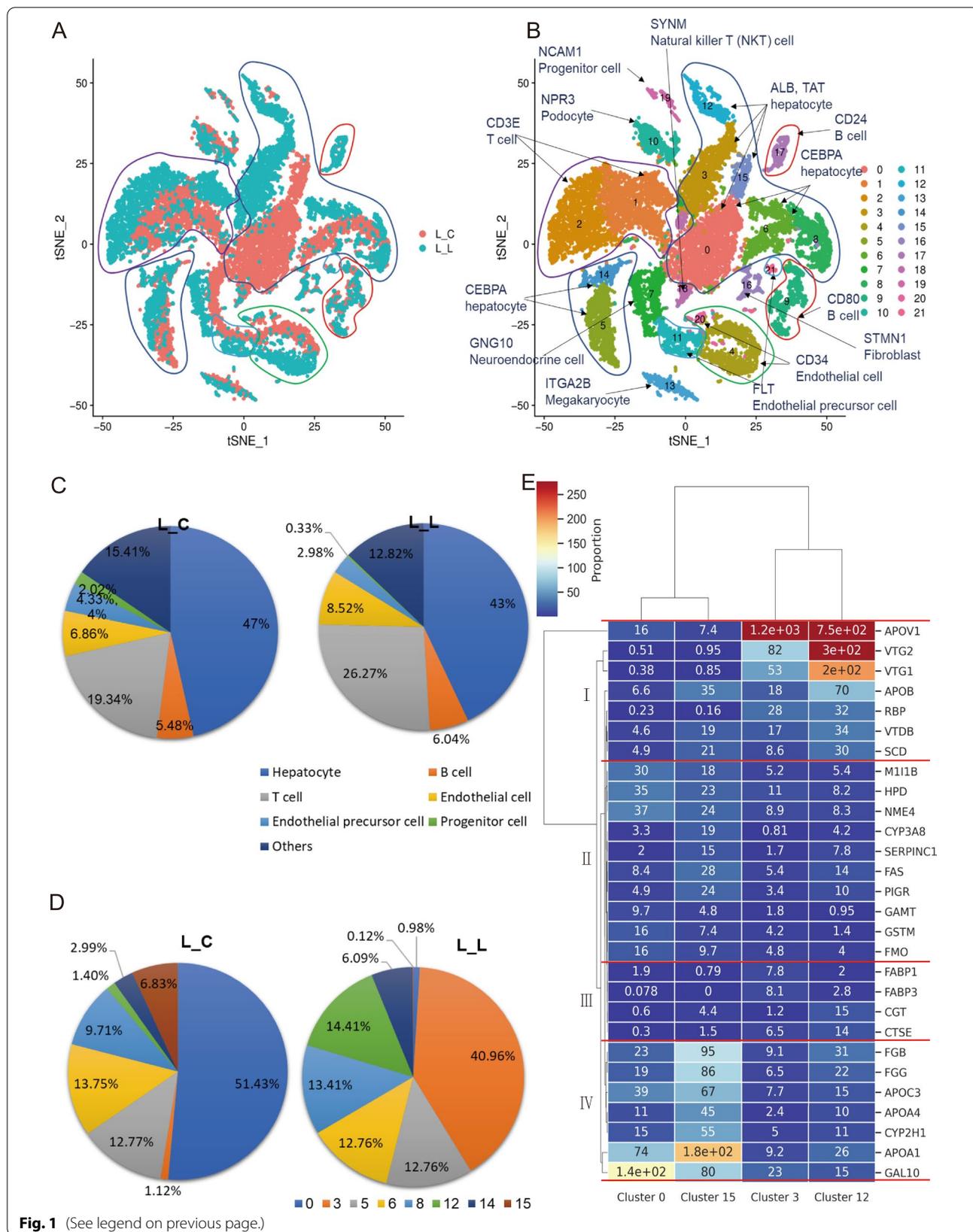
The interaction among proteins coding by marker genes of cluster 0, 3, 12, 15 were predicted and shown in Fig. 2E. Basically, all the genes might be divided into three groups, lipids transport, lipid metabolic process and fibrinogen complex (Fig. 2E).

### Discussion

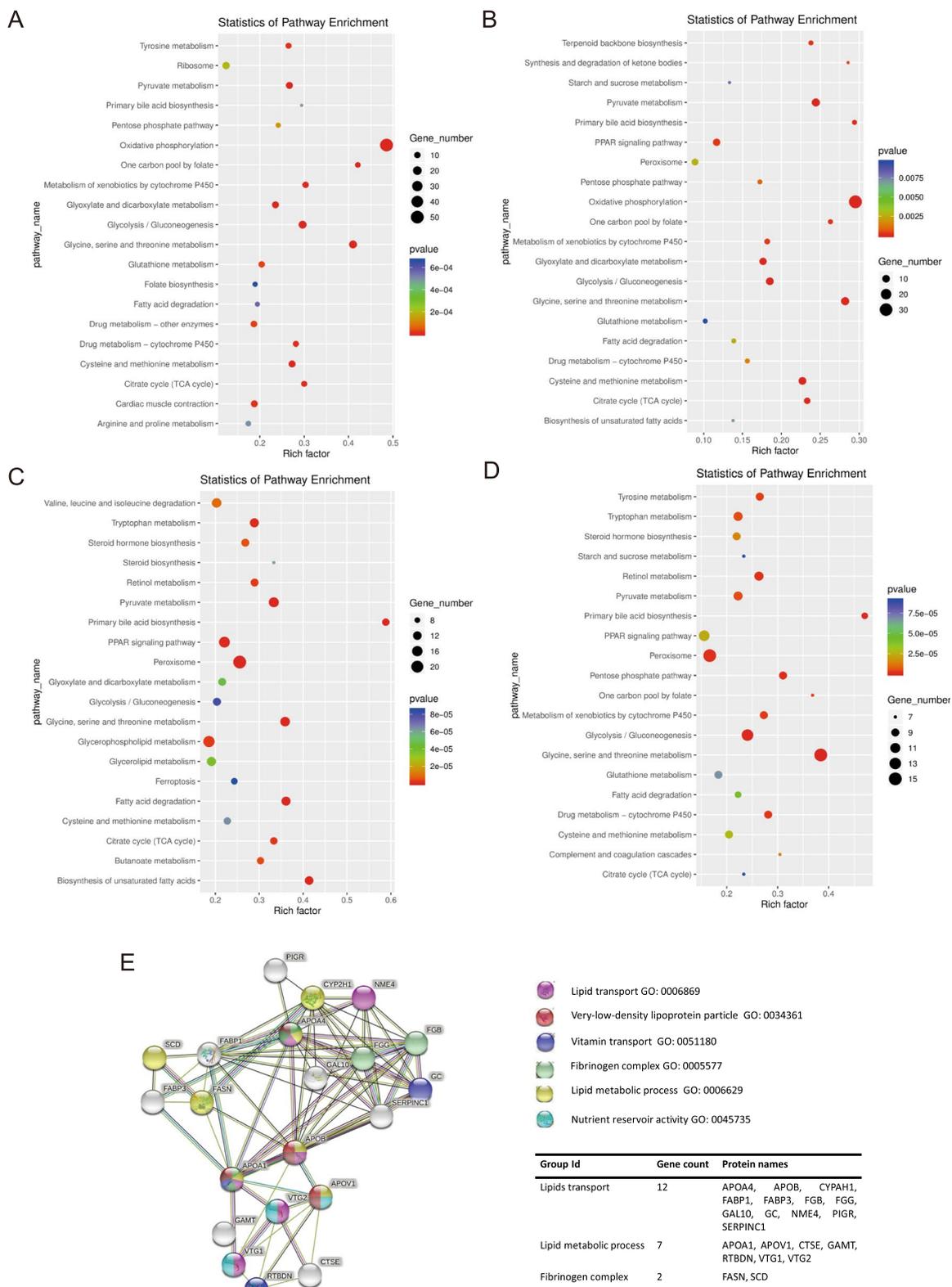
Here, we present a microdroplet-based scRNA-seq technology ( $10 \times$  Genomics) that enables encapsulation of tens of thousands of single cells within minutes. We analyzed the transcriptomes of ~20 k single cells across L\_L

(See figure on next page.)

**Fig. 1** Comparative genomics analysis of liver cells between laying and ceased-laying ducks. **A–B** t-distributed stochastic neighbor embedding (t-SNE) analysis of duck liver cell data. Cells are colored by experimental samples (**A**). Cells are colored by cell type cluster (**B**). Circle of same color represents the cell type, cell types and the marker genes were shown by arrows. **C** The percentage of different cell types in different samples. **D** The percentage of different cell clusters of hepatocytes in two samples. **E** A gene expression heatmap showing the expression of marker genes for cluster 0, 3, 12 and 15 in duck liver data. Rows correspond to the top 10 marker genes in individual clusters; columns are individual cells, ordered by clusters. Red corresponds to high expression level; blue corresponds to low expression level



**Fig. 1** (See legend on previous page.)



**Fig. 2** A-D The KEGG pathway enrichment of the up-regulated genes in cluster 0 (A), cluster 3 (B), cluster 12 (C) and cluster 15 (D). Rich factor represents the ratio of expressed gene number to annotated gene number. Lower number of pvalue equates to higher enrichment. E Protein interaction modules of the top 10 marker genes in cluster 0, cluster 3, cluster 12 and cluster 15

and L\_C. The high-quality transcriptomes, collectively expressing about 17 k genes, provided strong statistical power for unbiased decomposition of these cell populations. Using a conservative statistical threshold, we identified genes that are differentially expressed between cells enabling us to functionally categorize several distinct clusters. To our knowledge, this dataset provides the first cell atlas of duck liver by single-cell transcriptomics profiling. It also provides a fundamental resource to the research of evolution of species.

The relationship between hepatocyte heterogeneity and liver function has been revealed versatily in human studies. For instance, Ho DW et al. [29] identified a  $CD24^+/CD44^+$ -enriched cell subpopulation within the  $EPCAM^+$  cells which had specific signature genes and might indicate a novel stemness-related cell subclone in hepatocellular carcinoma. Zheng H et al. [30] proves that hepatic cancer stem cells (CSCs) at the single-cell level are heterogeneous and different CSC subpopulations are independently associated with hepatocellular carcinoma prognosis. In this study, the effects of hepatocyte heterogeneity on liver function in ceasing laying were focused, which is of scientific significance.

Based on the gene expression, we performed the analysis of cell clustering, cell type analysis and KEGG enrichment analysis. CellMarker is a database that provides a comprehensive and accurate resource of cell markers for various cell types in tissues of human and mouse [31]. CellMarker shows 23 cell types in human and mouse liver. Using the marker genes of liver in CellMarker, we identified 6 cell types, including hepatocytes, T cells, B cells, progenitor cells, endothelial cells and endothelial precursor cells. Other cell types in CellMarker were not identified, such as embryonic cells, stem cells, hematopoietic cells, kupffer cells, dendritic cells, etc. The absence of embryonic cells and stem cells might attribute to the state of the senile ducks. The absence of hematopoietic cells, kupffer cells and dendritic cells might attribute to the loss in cell digestion. In addition, the species differences between birds and mammals might play a role on the bias, e.g., goose and duck have higher liver lipid storage capacity than chicken and mammals [32].

The main function of liver are synthesis, metabolism and decomposition. A large fraction of the corresponding enzymes are mainly expressed in hepatocytes. Our study provided the ocular proof that hepatocytes are dominant in quantitative terms in liver. We also detected a large number of T cells (~19–27%) and B cells (~5.48–6.04%). A promotion effect of lymph node to liver regeneration has been demonstrated by Fontes et al. [33]. The lymphocytes (T cells and B cells) population in our study might provide a high quality environment for liver.

Among all the clusters, we focused on cluster 0, 3, 12 and 15, which showed the main difference between two samples. The cells of these four clusters are hepatocytes. Cluster 3 and 12 specifically belong to egg-laying duck, while cluster 0 and 15 specifically belong to ceased-laying duck. The reference marker gene were used to identify the heterogeneous functions of hepatocytes between ceased-laying duck and egg-laying duck. In egg-laying duck, specific cluster 3 and 12 expressed more yolk materials, like vitellogenins (*VTG1*, *VTG2*) and unsaturated fatty acids (*SCD*), more binding/transport proteins, like *APOV1*, *APOB*, *RBP*, *VTDB*, and more *CTSE*. This is in accordance with the material needs of egg-laying duck, as the yolk formation requires a massive increase in yolk precursor synthesis and package in the liver. As one of the major yolk protein, VTG expressed exclusively in the liver of poultry [34]. *SCD* is vital to monounsaturated fatty acids (MUFAs) synthesis [9], which is important composition of yolk lipids. *APOV1* and *APOB* are major transport proteins of VLDL-yolk [4]. Richards et al. [35] revealed that hepatic expression of *APOV1* and *APOB* genes increased significantly in broiler breeders after photostimulation at first egg. Riboflavin and vitamin D are important yolk nutrition, here we detected the higher gene expression of their binding proteins, *RBP* [36] and *VTDB* [37]. *CTSE* is also a key enzyme for yolk formation as it is capable of catalyzing intra oocytic break down of protein components of both vitellogenin and VLDL during yolk deposition [38]. For the first, these results show a comprehensive atlas of function genes related to egg laying in a higher resolution of single cells.

In ceased-laying duck, specific cluster 0 and cluster 15 expressed more *FAS*, more other apolipoproteins, like *APOA1*, *APOA4* and *APOC3*, and more blood coagulation related genes, like *FGB*, *FGG* and *SERPINC*. *FAS*, *APOA1* and *APOA4* are involved in tissue-specific fat deposition in chickens [39, 40]. *APOA4* promotes the proliferation of lipoprotein particles and reduces the fat burden in the liver, while not increasing the formation of apolipoprotein B lipoprotein particles, which would increase the possibility of arterial sclerosis [41]. *APOC3* attenuates hydrolysis and clearance of triglyceride-rich lipoproteins [42]. *FGB* and *FGG* are polymerize to form an insoluble fibrin matrix, which has a major function in hemostasis, while *SERPINC* is serine protease inhibitor in plasma that regulates the blood coagulation cascade [43]. These results might explain the higher tissue-specific fat deposition and higher vascular burden in ceased-laying avians.

Our study further proofs that *APOV1* and *APOB* play key roles in egg production, rather than *APOA1* and *APOA4*. It is also the first to detect a correlation

between the higher expression of *APOC3*, *FGB*, *FGG* and ceased-laying in duck.

### Conclusions

In conclusion, our study is the first to describe the heterogeneity in the duck liver at different egg-laying period. This single-cell transcriptomic atlas provides a comprehensive genomics resource to study duck in ceasing laying in unprecedented detail. Hepatocytes showed a dominant heterogeneity among all cell types between egg-laying duck and ceased-laying duck. The expression mode of yolk precursor transporters, lipid metabolizing enzymes and fibrinogens were different. *APOV1*, *VTG2*, *VTG1*, *APOB*, *RBP*, *VTDB* and *SCD* might be activated in egg-laying ducks, while *APOA1*, *APOA4*, *APOC3*, *FGB* and *FGG* might be activated in ceased-laying ducks. Our findings provide a solid foundation for future studies on the molecular mechanisms underlying ceasing laying in late phase duck and on strategies for the improvement of egg production.

### Materials and methods

#### Animals

Fifty female Shaoxing ducks (1550 ± 92 g, 450 days of age) were randomly selected from one hereditary line at Guowei Poultry Industry Co., Ltd. China, and grown in individual cages under natural temperature and light conditions in October of Zhejiang province. The study is reported in compliance with the ARRIVE guidelines. Feed and water were provided for ad libitum consumption. The ingredients and nutrients of feed were shown in Supplementary Table S4. Successive twelve days of egg production was recorded from 450 days of age. During the twelve days, ducks produced eggs in each day were marked as the egg-laying ducks (L-L), while ducks didn't produce any egg were marked as the ceased-laying ducks (L-C). One egg-laying duck and one ceased-laying duck were randomly selected. Livers were sampled from egg-laying duck and ceased-laying duck, respectively.

#### Preparation of single-cell suspension

The right liver anterior tips were isolated from ducks. The following processing referred to Han et al. [16]. Firstly, the samples were quickly transferred into cold PBS (Cat# GNM1514, Genomcell.bio, China) and then minced into ~1 mm pieces with scissors. Secondly, they were incubated with 0.25% trypsin–EDTA (Cat# 25,200,056, Gibco, Australia) on ice for 10 min, after washing by cold FBS (Cat# 10,100,147, Gibco, Australia), they were digested by 0.5% collagenase IV (Cat# 17,100,019, Invitrogen, USA) at 37°C for 40 min with gentle agitation for enzymatic dissociation into single-cell suspensions. Thirdly, the cell suspensions were passed through a 40 µm

cell strainer to generate cleaner single-cell suspensions, then, the filtrates were centrifuged for 10 min at 300 × g (4°C), and the supernatants were removed. Fourthly, the cells were incubated for 10 min with 1 × RBC lysis buffer at room temperature and centrifuged for 10 min at 300 × g (4 °C) to remove the contaminating RBCs. Finally, after removal of the supernatants, the cells were washed and suspended by DPBS (Cat# GNM14190, Genomcell.bio, China) with 1% BSA (Cat# PC0001, Solarbio, China) to the concentration of 7 × 10<sup>5</sup> cells/cm<sup>2</sup>. The viabilities were assessed using a 0.4% (w/v) trypan blue dye solution (Cat# 15,250,061, Life, USA), and the total numbers of cells were counted using an automated cell counter.

#### Library construction and Sequencing

The library construction was referred to Feregrino et al. [44]: ① Cell suspensions were loaded on a GemCode Single-Cell Instrument (10 × Genomics, USA) to generate single-cell GEMs (Gel Beads-In-Emulsions). ② Single-cell RNA-Seq libraries were prepared using GemCode Single-Cell 3'Gel Bead and Library Kit (Cat# 1,000,121, 10 × Genomics, USA). ③ GEM-RT was performed in a PCR Amplifier (Cat# 1,851,197, Bio-Rad, USA): 55 °C for 2 h, 85 °C for 5 min; held at 4 °C. ④ cDNA products were cleaned up with the SPRI select Reagent Kit (Cat# B23317, Beckman, USA) and subsequently sheared to 200~300 bp using enzymatic fragmentation system (Cat# 500,295, Covaris, USA). ⑤ Indexed sequencing libraries were constructed using the reagents in the Gem-Code Single-Cell 3' Library Kit, following these steps: (1) end repair and A-tailing; (2) adapter ligation; (3) post ligation cleanup with SPRI select; (4) sample index PCR and cleanup. ⑥ The barcode sequencing libraries were quantified by quantitative PCR (Cat# KK4824, KAPA Biosystems Library Quantification Kit for Illumina platforms, USA).

Sequencing libraries were loaded on an Illumina NextSeq 500 with paired-end kits. As shown in Supplementary Figure S4, each sequence contains Illumina P5 adaptor, P5 primer (TruSeq Read 1), Barcode (16 bp), UMI (12 bp), poly (dT) VN, cDNA, P7 primer (TruSeq Read 2), Index (i7 index read) and Illumina P7 adaptor. Reads 1 were used to distinguish different transcripts of different cells. Reads 2 were used to determine the genetic information.

#### Data processing

The Cell Ranger Single Cell Software Suite (<https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome>) was applied to perform sample demultiplexing, barcode processing, and single-cell gene counting. The cDNA insert was mapped to an appropriate genome reference (*Anas Platyrhynchos*, <ftp://ftp>.

[ncbi.nlm.nih.gov/genomes/all/GCF/003/850/225/GCF\\_003850225.1\\_IASCAAS\\_PekingDuck\\_PBH1.5/GCF\\_003850225.1\\_IASCAAS\\_PekingDuck\\_PBH1.5\\_genom.ic.fna.gz](https://ncbi.nlm.nih.gov/genomes/all/GCF/003/850/225/GCF_003850225.1_IASCAAS_PekingDuck_PBH1.5/GCF_003850225.1_IASCAAS_PekingDuck_PBH1.5_genom.ic.fna.gz)) using STAR (Spliced Transcripts Alignment to a Reference) [45]. The filter thresholds for mapped data were adapted for each sample, depending on the different library complexities. Such cells were filtered out: cells with an UMI count more than 20,000, cells with a gene count less than 200, and cells with a mitochondrial or ribosomal contribution to UMI count more than 20%. Using the R package Seurat [46], the UMI counts were then Log-normalized.

### Dimensionality reduction and visualization

Significant principal components were determined for each sample as those falling outside of a Marchenko-Pastur distribution. A dimensionality reduction step was carried out following a similar approach to Macosko et al. [47]. K-means [48] clustering on the first 10 principal components were visualized in two-dimensional projection of t-distributed stochastic neighbour embedding projection (tSNE) [49]. Clusters were identified using the Seurat FindClusters function with a resolution parameter of 0.6 for the two physiological stages. The up-regulated genes with  $P < 0.01$  and  $|\text{Log}_2\text{FC}| > 0.25$  were defined as the marker genes in each cluster compared to all other clusters. The top 10 marker genes in each cluster were selected to construct the heatmap. Cell types were identified by marker genes in CellMarker (<http://bio-bigdata.hrbmu.edu.cn/CellMarker/index.jsp>) [50] and other published papers.

### KEGG enrichment

KEGG (Kyoto Encyclopedia of Genes and Genomes; <http://www.kegg.jp/>) [51] is utilized for bioinformatics research in different cell clusters. KOBAS [52] was used to test the statistical enrichment of the up-regulated genes in KEGG pathways.

### Protein interaction prediction

We used STRING (version 11.5) to integrate and rank marker genes associations of cluster 0, 3, 12, 15 by benchmarking them against a common reference set, and presents evidence in a consistent and intuitive web interface [53].

### Abbreviations

scRNA-seq: Single-cell RNA sequencing; tSNE: T-distributed stochastic neighbor embedding; UMIs: Unique molecular identifiers; APOV1: Very-low-density apolipoprotein-VLDL II; VTG2: Vitellogenin II; VTG1: Vitellogenin I; APOB: Apolipoprotein B; RBP: Riboflavin-binding protein; VTDB/GC: Vitamin D-binding protein; SCD: Stearoyl-acyl-CoA desaturase; APOA1: Apolipoprotein A1; APOA4: Apolipoprotein A4; APOC3: Apolipoprotein C3; FGB: Fibrinogen B; FGG: Fibrinogen G; VLDL: Very-low-density lipoprotein; PPAR: Peroxisome

proliferator activated receptor; LPL: Lipoprotein lipase; DBP: Deep-brain photoreceptor; TSHB: Thyroid-stimulating hormone  $\beta$ ; DIO2: Deiodinase 2; OATPs: Organic anion transporting polypeptides; GnRH: Gonadotropin-releasing hormone; LH: Luteinizing hormone; FSH: Follicle-stimulating hormone; *COL1A1*: Type I collagen encoding gene; *ALB*: Albumin; *TAT*: Tyrosine aminotransferase isoform X1; *CEBPA*: CCAAT/enhancer-binding protein alpha; *CD3E*: Glycoprotein CD3 epsilon chain; *CD8B*: Glycoprotein CD8 beta chain; *NCAM1*: Neural cell adhesion molecule 1 isoform X8; *FLT*: Vascular endothelial growth factor receptor; *FABP1*: Liver-type fatty acid-binding protein; *FABP3*: Heart-type fatty acid-binding protein; *FAS/FASN*: Fatty acid synthase; *M11B*: Mid1-interacting protein 1-B; *HPD*: 4-Hydroxyphenylpyruvate dioxygenase; *GSTM*: Glutathione S-transferase; *GAMT*: Guanidinoacetate N-methyltransferase.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-022-09089-0>.

**Additional file 1.** The enriched pathways of gene from cluster 0. The enriched pathways of gene from cluster 15. The enriched pathways of gene from cluster 3. The enriched pathways of gene from cluster 12.

**Additional file 2: Table S1.** Statistical results of the sequencing data of liver samples in different laying status.

**Additional file 3: Table S2.** Statistical results of quality by Cell Ranger of liver samples in different laying status.

**Additional file 4: Table S3.** Statistical table of cell filtration.

**Additional file 5: Table S4.** Ingredients and nutrient composition of basal diet.

**Additional file 6: Figure S1.** Median number of genes detected per cell as a function of mean reads per cell. L\_C: liver of ceased-laying duck; L\_L: liver of laying duck.

**Additional file 7: Figure S2.** The distribution of basic information of cells in each sample before and after the filter. **A** The distribution of basic information of cells before the filter; **B** The distribution of basic information of cells after the filter; The left image of each group of pictures shows the distribution of gene number of cells in each sample; The middle image of each group of pictures shows the distribution of UMI number of cells in each sample; The right image of each group of pictures shows the distribution of mitochondrial gene expression level of cells in each sample; L\_C: liver of ceased-laying duck; L\_L: liver of laying duck.

**Additional file 8: Figure S3.** Numbers at the top indicate cluster number, with connecting lines indicating the hierarchical relationship between clusters. Representative markers from each cluster are shown on the left.

**Additional file 9: Figure S4.** Structure of the library. P5: Illumina P5 adaptor; TruSeq Read 1: P5 primer; 10×Barcode: 16 bp, one bead has one kind of barcode; UMI: 12bp, one transcript matches one kind of UMI; TruSeq Read 2: P7 primer; Index: i7 index read; P7: Illumina P7 adaptor.

### Acknowledgements

The authors wish to acknowledge Hao Qiu for help with data analysis.

### Authors' contributions

Conceptualization, XD, TZ and LZL; methodology, XD, SJL, WQZ and XQX; analyses WXX and YT; writing—original draft preparation, XD; writing—review and editing, XD, SJL and LZL. All of the authors read and agreed the published version of manuscript.

### Funding

This work was supported by China Agriculture Research System of MOF and MARA (No. CARS-42–6), and Zhejiang Science and Technology Major Program on Agricultural New Variety Breeding (No. 2021C02068-10).

### Availability of data and materials

The datasets generated and/or analysed during the current study are available in the NCBI Sequence Read Archive repository, the accession numbers are

SRR13402420 and SRR13402419 (<https://dataview.ncbi.nlm.nih.gov/object/PRJNA690971>).

## Declarations

### Ethics approval and consent to participate

All experimental protocols were approved by Laboratory Animal Welfare Ethics Committee, Zhejiang Academy of Agricultural Sciences (Permit No. 2019ZAASLA1901). All methods were carried out in accordance with relevant guidelines and regulations. The study is in accordance with the ARRIVE Guidelines.

### Consent for publication

Not applicable.

### Competing interests

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work.

### Author details

<sup>1</sup>State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Institute of Animal Husbandry and Veterinary Medicine, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, Zhejiang, China. <sup>2</sup>College of Animal Science and Technology, College of Veterinary Medicine, Zhejiang A & F University, Hangzhou, China. <sup>3</sup>Shanghai Institute of Immunology, Shanghai Jiao Tong University School of Medicine, Shanghai, China. <sup>4</sup>Institute of Horticulture, Zhejiang Academy of Agricultural Sciences, Hangzhou 310022, Zhejiang, China. <sup>5</sup>Institute of Ecology, China West Normal University, Nanchong 637002, Sichuan, China.

Received: 19 September 2022 Accepted: 19 December 2022

Published online: 28 December 2022

## References

- Zhu W, Liu L, Yang W, Wei C, Geng Z, Chen X. Comparative analysis of metabolites in the liver of Muscovy ducks at different egg-laying stages using nontargeted ultra-high-performance liquid chromatography-electrospray mass spectrometry-based metabolomics. *J Proteome Res*. 2020;19(9):3846–55.
- Li H, Ma Z, Jia L, Li Y, Xu C, Wang T, Han R, Jiang R, Li Z, Sun G, Kang X, Liu X. Systematic analysis of the regulatory functions of microRNAs in chicken hepatic lipid metabolism. *SCI REP-UK*. 2016;6:31766.
- Li H, Wang T, Xu C, Wang D, Ren J, Li Y, Tian Y, Wang Y, Jiao Y, Kang X, Liu X. Transcriptome profile of liver at different physiological stages reveals potential mode for lipid metabolism in laying hens. *BMC Genomics*. 2015;16:763.
- Omer NA, Hu Y, Idriss AA, Abobaker H, Hou Z, Yang S, Ma W, Zhao R. Dietary betaine improves egg-laying rate in hens through hypomethylation and glucocorticoid receptor-mediated activation of hepatic lipogenesis-related genes. *Poultry Sci*. 2020;99(6):3121–32.
- Du X, Ren JD, Xu XQ, Chen GH, Huang Y, Du JP, Tao ZR, Cai ZX, Lu LZ, Yang H. Comparative transcriptome analysis reveals genes related to the yolk ratio of duck eggs. *Anim Genet*. 2019;50(5):484–92.
- Liu Z, Li Q, Liu R, Zhao G, Zhang Y, Zheng M, Cui H, Li P, Cui X, Liu J, Wen J. Expression and methylation of microsomal triglyceride transfer protein and acetyl-CoA carboxylase are associated with fatty liver syndrome in chicken. *Poultry Sci*. 2016;95(6):1387–95.
- Wang WW, Wang J, Zhang HJ, Wu SG, Qi GH. Transcriptome analysis reveals mechanism underlying the differential intestinal functionality of laying hens in the late phase and peak phase of production. *BMC Genomics*. 2019;20(1):970.
- Dai H, Lv Z, Huang Z, Ye N, Li S, Jiang J, Cheng Y, Shi F. Dietary hawthorn-leaves flavonoids improves ovarian function and liver lipid metabolism in aged breeder hens. *Poultry Sci*. 2021;100(12).
- Li H, Li Z, Liu X. An Overall View of the Regulation of Hepatic Lipid Metabolism in Chicken Revealed by New-Generation Sequencing. London: IntechOpen; 2017.
- Chen J, Okimura K, Yoshimura T. Light and hormones in seasonal regulation of reproduction and mood. *Endocrinology*. 2020;161(9):bqaa130.
- Nakane Y, Ikegami K, Ono H, Yamamoto N, Yoshida S, Hirunagi K, Ebihara S, Kubo Y, Yoshimura T. A mammalian neural tissue opsins (Opsin 5) is a deep brain photoreceptor in birds. *P Natl Acad Sci USA*. 2010;107(34):15264–8.
- Nakao N, Ono H, Yamamura T, Anraku T, Takagi T, Higashi K, Yasuo S, Katou Y, Kageyama S, Uno Y, Kasukawa T, Iigo M, Sharp PJ, Iwasawa A, Suzuki Y, Sugano S, Niimi T, Mizutani M, Namikawa T, Ebihara S, Ueda HR, Yoshimura T. Thyrotrophin in the pars tuberalis triggers photoperiodic response. *Nature*. 2008;452(7185):317–22.
- Yoshimura T, Yasuo S, Watanabe M, Iigo M, Yamamura T, Hirunagi K, Ebihara S. Light-induced hormone conversion of T4 to T3 regulates photoperiodic response of gonads in birds. *Nature*. 2003;426(6963):178–81.
- Nakao N, Takagi T, Iigo M, Tsukamoto T, Yasuo S, Masuda T, Yanagisawa T, Ebihara S, Yoshimura T. Possible involvement of organic anion transporting polypeptide 1c1 in the photoperiodic response of gonads in birds. *Endocrinology*. 2006;147(3):1067–73.
- Hanafy AM, Elnesr SS. Induction of reproductive activity and egg production by gonadotropin-releasing hormone in non-laying hens. *Reprod Domest Anim*. 2021;56(9):1184–91.
- Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, Saadatpour A, Zhou Z, Chen H, Ye F, Huang D, Xu Y, Huang W, Jiang M, Jiang X, Mao J, Chen Y, Lu C, Xie J, Fang Q, Wang Y, Yue R, Li T, Huang H, Orkin SH, Yuan GC, Chen M, Guo G. Mapping the mouse cell atlas by microwell-seq. *cell*. 2018;172(5):1091–107.
- Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y, Chen H, Wang J, Tang H, Ge W, Zhou Y, Ye F, Jiang M, Wu J, Xiao Y, Jia X, Zhang T, Ma X, Zhang Q, Bai X, Lai S, Yu C, Zhu L, Lin R, Gao Y, Wang M, Wu Y, Zhang J, Zhan R, Zhu S, Hu H, Wang C, Chen M, Huang H, Liang T, Chen J, Wang W, Zhang D, Guo G. Construction of a human cell landscape at single-cell level. *Nature*. 2020;581(7808):303–9.
- Li J, Xing S, Zhao G, Zheng M, Yang X, Sun J, Wen J, Liu R. Identification of diverse cell populations in skeletal muscles and biomarkers for intramuscular fat of chicken by single-cell RNA sequencing. *BMC Genomics*. 2020;21(1):752.
- Yuan J, Li W, Huang J, Guo X, Li X, Lu X, Huang X, Zhang H. Transplantation of human adipose stem cell-derived hepatocyte-like cells with restricted localization to liver using acellular amniotic membrane. *Stem Cell Res Ther*. 2015;6:217.
- Cipriano M, Correia JC, Camoes SP, Oliveira NG, Cruz P, Cruz H, Castro M, Ruas JL, Santos JM, Miranda JP. The role of epigenetic modifiers in extended cultures of functional hepatocyte-like cells derived from human neonatal mesenchymal stem cells. *Arch Toxicol*. 2017;91(6):2469–89.
- Lauer FT, Denson JL, Burchiel SW. Isolation, cryopreservation, and immunophenotyping of human peripheral blood mononuclear cells. *Curr Protoc Toxicol*. 2017;74:18–20.
- Qian L, Zhang M, Wu S, Zhong Y, Van Tol E, Cai W. Alkylglycerols modulate the proliferation and differentiation of non-specific agonist and specific antigen-stimulated splenic lymphocytes. *PLoS ONE*. 2014;9(4): e96207.
- Bofill M, Janossy G, Janossa M, Burford GD, Seymour GJ, Wernet P, Kelemen E. Human B cell development. II. Subpopulations in the human fetus. *J Immunol*. 1985;134(3):1531–8.
- Oishi N, Kumar MR, Roessler S, Ji J, Forgues M, Budhu A, Zhao X, Andersen JB, Ye QH, Jia HL, Qin LX, Yamashita T, Woo HG, Kim YJ, Kaneko S, Tang ZY, Thorgeirsson SS, Wang XW. Transcriptomic profiling reveals hepatic stem-like gene signatures and interplay of miR-200c and epithelial-mesenchymal transition in intrahepatic cholangiocarcinoma. *Hepatology*. 2012;56(5):1792–803.
- Nava S, Westgren M, Jaksch M, Tibell A, Broome U, Ericzon BG, Sumitran-Holgersson S. Characterization of cells in the developing human liver. *Differentiation*. 2005;73(5):249–60.
- Plasschaert LW, Zilionis R, Choo-Wing R, Savova V, Knehr J, Roma G, Klein AM, Jaffe AB. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature*. 2018;560(7718):377–81.
- Park J, Shrestha R, Qiu C, Kondo A, Huang S, Werth M, Li M, Barasch J, Susztak K. Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science*. 2018;360(6390):758–63.
- Young MD, Mitchell TJ, Vieira BF, Tran M, Stewart BJ, Ferdinand JR, Collard G, Botting RA, Popescu DM, Loudon KW, Vento-Tormo R, Stephenson E,

- Cagan A, Farndon SJ, Del CVM, Guzzo C, Richoz N, Mamanova L, Aho T, Armitage JN, Riddick A, Mushtaq I, Farrell S, Rampling D, Nicholson J, Filby A, Burge J, Lisgo S, Maxwell PH, Lindsay S, Warren AY, Stewart GD, Sebire N, Coleman N, Haniffa M, Teichmann SA, Clatworthy M, Behjati S. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science*. 2018;361(6402):594–9.
29. Ho DW, Tsui YM, Sze KM, Chan LK, Cheung TT, Lee E, Sham PC, Tsui SK, Lee TK, Ng IO. Single-cell transcriptomics reveals the landscape of intratumoral heterogeneity and stemness-related subpopulations in liver cancer. *CANCER LETT*. 2019;459:176–85.
  30. Zheng H, Pomyen Y, Hernandez MO, Li C, Livak F, Tang W, Dang H, Greten TF, Davis JL, Zhao Y, Mehta M, Levin Y, Shetty J, Tran B, Budhu A, Wang XW. Single-cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma. *Hepatology*. 2018;68(1):127–40.
  31. Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, Luo T, Xu L, Liao G, Yan M, Ping Y, Li F, Shi A, Bai J, Zhao T, Li X, Xiao Y. Cell Marker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res*. 2019;47(D1):D721–8.
  32. Lu L, Chen Y, Wang Z, Li X, Chen W, Tao Z, Shen J, Tian Y, Wang D, Li G, Chen L, Chen F, Fang D, Yu L, Sun Y, Ma Y, Li J, Wang J. The goose genome sequence leads to insights into the evolution of waterfowl and susceptibility to fatty liver. *Genome Biol*. 2015;16:89.
  33. Fontes P, Komori J, Lopez R, Marsh W, Lagasse E. Development of ectopic livers by hepatocyte transplantation into swine lymph nodes. *Liver Transplant*. 2020;26(12):1629–43.
  34. Omer NA, Hu Y, Hu Y, Idriss AA, Abobaker H, Hou Z, Dong H, Zhao R. Dietary betaine activates hepatic VTGII expression in laying hens associated with hypomethylation of GR gene promoter and enhanced GR expression. *J Anim Sci Biotechnol*. 2018;9:2.
  35. Richards MP, Poch SM, Coon CN, Rosebrough RW, Ashwell CM, McMurtry JP. Feed restriction significantly alters lipogenic gene expression in broiler breeder chickens. *J Nutr*. 2003;133(3):707–15.
  36. Maehashi K, Matano M, Kondo A, Yamamoto Y, Udaka S. Riboflavin-binding protein exhibits selective sweet suppression toward protein sweeteners. *Chem Senses*. 2007;32(2):183–90.
  37. Nagasawa H, Uto Y, Sasaki H, Okamura N, Murakami A, Kubo S, Kirk KL, Hori H. Gc protein (vitamin D-binding protein): Gc genotyping and GcMAF precursor activity. *Anticancer Res*. 2005;25(6A):3689–95.
  38. Retzek H, Steyrer E, Sanders EJ, Nimpf J, Schneider WJ. Molecular cloning and functional characterization of chicken cathepsin D, a key enzyme for yolk formation. *DNA Cell Biol*. 1992;11(9):661–72.
  39. Zhang M, Li F, Ma XF, Li WT, Jiang RR, Han RL, Li GX, Wang YB, Li ZY, Tian YD, Kang XT, Sun GR. Identification of differentially expressed genes and pathways between intramuscular and abdominal fat-derived preadipocyte differentiation of chickens in vitro. *BMC Genomics*. 2019;20(1):743.
  40. Moreira G, Boschiero C, Cesar A, Reecy JM, Godoy TF, Trevisoli PA, Cantao ME, Ledur MC, Ibelli A, Peixoto JO, Moura A, Garrick D, Coutinho LL. A genome-wide association study reveals novel genomic regions and positional candidate genes for fat deposition in broiler chickens. *BMC Genomics*. 2018;19(1):374.
  41. Wang Q, Liu M, Xu L, Wu Y, Huang Y. Transcriptome analysis reveals the molecular mechanism of hepatic fat metabolism disorder caused by Muscovy duck reovirus infection. *Avian Pathol*. 2018;47(2):127–39.
  42. Chan DC, Chen MM, Ooi EM, Watts GF. An ABC of apolipoprotein C-III: a clinically useful new cardiovascular risk factor? *Int J Clin Pract*. 2008;62(5):799–809.
  43. Mourey L, Samama JP, Delarue M, Choay J, Lormeau JC, Petitou M, Moras D. Antithrombin III: structural and functional aspects. *Biochimie*. 1990;72(8):599–608.
  44. Feregrino C, Sacher F, Parnas O, Tschopp P. A single-cell transcriptomic atlas of the developing chicken limb. *BMC Genomics*. 2019;20(1):401.
  45. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
  46. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36(5):411–20.
  47. Macosko EZ, Basu A, Satija R, Nemeshe J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martnersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*. 2015;161(5):1202–14.
  48. Sherlock G. Analysis of large-scale gene expression data. *Curr Opin Immunol*. 2000;12(2):201–5.
  49. Liu S, Maljovec D, Wang B, Bremer PT, Pascucci V. Visualizing high-dimensional data: advances in the past decade. *IEEE Trans Vis Comput Graph*. 2017;23(3):1249–68.
  50. Aizarani N, Saviano A, Sagar, Mailly L, Durand S, Herman JS, Pessaux P, Baumert TF, Grun D. A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature*. 2019;572(7768):199–204.
  51. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
  52. Mao X, Cai T, Olyarchuk JG, Wei L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics*. 2005;21(19):3787–93.
  53. von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA, Bork P. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res*. 2005;33(Database issue):D433–7.

## Publisher's Note

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

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

