BMC Genomics

RESEARCH

Open Access

Single-cell RNA transcriptome uncovers distinct developmental trajectories in the embryonic skeletal muscle of Daheng broiler and Tibetan chicken

Jie Li[1†], Dongmei Yang[1,2,3†], Chuwen Chen[1], Jiayan Wang[1,2,3], Zi Wang[1,2,3], Chaowu Yang[4], Chunlin Yu[4] and Zhixiong Li[1,2,3*]

Abstract

Different chicken breeds exhibit distinct muscle phenotypes resulting from selective breeding, but little is known about the molecular mechanisms responsible for this phenotypic difference. Skeletal muscle is composed of a large number of heterogeneous cell populations. Differences in differentiation and interaction of cell populations play a key role in the difference of skeletal muscle phenotype. In the current study, we performed a single-cell RNA sequencing (scRNA-seq) on the leg muscle of Daheng broiler (DH, cultivated breed) and Tibetan chicken (TC, native breed) at embryonic (E) 10, E14 and E18. A comprehensive cell atlas of embryonic chicken skeletal muscle, consisting of 29,579 high-quality cells representing 6 distinct cell types was built. The differentiation trajectory of Myoblasts and fibro-adipogenic progenitors (FAPs) was constructed through pseudotemporal trajectory analysis. Our results revealed the different developmental trajectories and dynamic gene expression profiles in 3 subtypes of myoblasts and 5 FAPs subtypes of the two chicken breeds. Tibetan chicken showed earlier embryonic myogenesis and less myoblasts compared with Daheng broiler. By comparing the switch status and switch time of genes in the two breeds, *SNRPG,SNRPE,EIF4EBP1* and *HSP90AB1* were considered as potentially critical genes for embryonic myogenesis, and the genes *MYOG,MYBPH,APOA1*, and *MGP* played dominant roles in the embryonic adipogenesis. Intercellular interaction networks showed that strong and complex intercellular communication was contained during embryonic skeletal muscle growth and development. These findings revealed the differences of molecular mechanisms in the skeletal muscle development between TC and DH chickens. Our data provide a better understanding of skeletal muscle developmental differences between cultivated and native breeds and valuable information for genetic breeding of chicken.

Keywords ScRNA-seq, Transcriptional dynamics, Myoblasts, Fibro-adipogenic progenitor, Chicken breeds

†Jie Li and Dongmei Yang contributed equally to this work.

*Correspondence:
Zhixiong Li
lizhixiong@swun.edu.cn
[1]Key Laboratory of Qinghai-Tibetan Plateau Animal Genetic Resource Reservation and Utilization (Southwest Minzu University), Ministry of Education, Chengdu 610041, China
[2]Key Laboratory of Animal Science of National Ethnic Affairs Commission of China, Southwest Minzu University, Chengdu 610041, China
[3]Institute of Qinghai-Tibetan Plateau, Southwest Minzu University, Chengdu 610041, China
[4]Sichuan Animal Science Academy, Chengdu 610066, China

Introduction

Skeletal muscle is a highly complex and heterogeneous tissue and is the most widely distributed type of muscle (approximately 40% of body mass) serving a multitude of functions in the organism [1–3]. Myogenesis is an extremely complex process. For vertebrates, except for the skeletal muscle of the head derived from the occipital and cephalic lateral somite, the rest of all skeletal muscle originates from dermomyotome [4]. In the dermomyotome, there are many types of muscle progenitors, and both Myoblasts and Muscle satellite cells originate from the dermomyotome [5]. Under the control of a large number of genes, transcription factors, and signaling pathways, biological activities such as proliferation, differentiation, fusion, migration, and intercellular communication are carried out in these muscle cells to promote the growth and development of skeletal muscle [6–9]. So far, great breakthroughs have been made in the research of myogenesis process. However, due to the complexity of skeletal muscle growth and development process and technological limitations, many macro phenomena still cannot be explained, and the mechanisms of some biological processes are still unclear.

The process of muscle development is very conservative among vertebrates. The avian embryo is an ideal model system for researching skeletal muscle myogenesis cause of the advantage of easy observation and tracking [10, 11]. Our detailed knowledge of muscle cell lineages and myogenesis in the embryonic stage is mostly based on the manipulation of avian models [10, 12]. During the embryonic period, myoblasts form primary and secondary muscle fibers through proliferation, differentiation, migration and fusion and the total number of muscle fibers is already fixed at hatching [13, 14]. The time points at embryonic (E) 10, E14 and E18 represent embryonic chicken skeletal muscle in the stage of formation of primary and secondary muscle fibers (E10) [15], during the formation and differentiation of independent muscle tubes (E14) [16], and in the period of hypertrophy of muscle fibers (E18) [11]. After hatching, the number of muscle fibers hardly changes.

Chickens are significant agricultural animals that have become the primary meat choice globally over the last 50 years [17]. The growth and development of skeletal muscle determines, at least in part, muscle production and mass [18]. Therefore, genetic selection for the rapid growth of skeletal muscles in poultry has led to the emergence of cultivated breeds like white and yellow feather broilers, which improves the quality and speed of muscle development, resulting in faster muscle growth than native breeds. For example, as a cultivated breed, the skeletal muscle growth rate of Daheng broiler was much faster than that of Tibetan chicken, which is native breed. Although previous research has advanced our understanding of the skeletal muscle, knowledge of the transcriptional dynamics of skeletal muscle heterogeneity in different breeds is currently limited, such as cell types in the skeletal muscle, cell status at different developmental stages of skeletal muscle, and gene expression profiles at single-cell resolution. Therefore, it is necessary to conduct a more comprehensive analysis of skeletal muscle cells and prospectively characterize the molecular features of the skeletal muscle in different breeds at a single-cell level.

Single-cell RNA sequencing (scRNA-seq) is a powerful tool for obtaining a global view of developmental processes. This strategy enables us to explore gene expression in single cells, offering an opportunity to fully grasp the heterogeneity of cells and to more accurately define rare cell types [19, 20]. In this study, Daheng broiler (DH) and Tibetan chicken (TC) were selected as representatives of cultivated breeds and native breeds, we performed systematical and sequential analysis on 29,579 (DH) and 30,060 (TC) single-cell transcriptomes from leg muscle at embryonic (E) 10, E14, and E18 which encompasses the main stage of the increase of myoblast number, respectively. Seventeen major cell populations were identified by using t-distributed stochastic neighbor embedding (tSNE) analysis, we identified cell-clusters-specific marker genes and annotated the cell types in combination with classic marker genes collected. According to pseudotime analysis, we successfully decoded the dynamic differentiation process of skeletal muscle cells and the molecular events that occur during the differentiation process in the two chicken breeds. In addition, gene-switch status and intercellular interactions were revealed during skeletal muscle growth and development. These results provide insights into the regulation of embryonic skeletal muscle development in different breeds at a single-cell level.

Materials and methods

Experimental animals

In this experiment, Fertilized chicken eggs of DH were provided by Sichuan Daheng poultry breeding Co. LTD (Chengdu, Sichuan Province, China). Fertilized chicken eggs of TC were provided by Mao Xian Jiuding Original Ecological Livestock and Poultry Breeding Co., LTD. (Aba Tibetan and Qiang Autonomous, Sichuan Province, China). Fertilized eggs of DH and TC (*Gallus gallus domesticus*) were incubated at 37.5℃ and 55% humidity until they reached the desired developmental stages. Eggs were opened and a small piece of muscle was taken for PCR genetic sexing, as described previously [21]. After determining the gender of the chicken embryos, the leg muscle tissues from healthy male embryos at three different developmental stages (E10, E14, and E18). Three biological replicate samples of leg muscle from two chicken

breeds at each stage were mixed and labeled as DH_E10, DH_E14, DH_E18, TC_E10, TC_E14, and TC_E18, respectively. All embryos were quickly isolated and euthanized via decapitation. The experiments were approved by the Institutional Animal Care and Use Committee of Southwest Minzu University, China (Permit number: 2020MDLS44). All experiments were performed in accordance with the relevant guidelines and regulations.

Single-cell RNA-seq using 10 x Genomics

The muscle tissues isolated from E10, E14, and E18 were digested with collagenase I for 30 min and filtered with a 40 μm nylon cell strainer. After the red blood cells lyse, the single-cell suspension was obtained by mechanical dissociation using a pipetting system. The quality of single-cell suspension was detected by AO/PI fluorescence staining. The single-cell suspension was used to construct a single-cell library after the following conditions: cell activity > 85%, fragments < 4%, cell agglomeration rate < 5%, and total number of cells > 10^5.

Single-cell libraries were constructed using the 10× Genomics single-cell RNA sequencing platform (10× Genomics, Pleasanton, CA, USA). Reverse transcription and 10×barcoded cDNA libraries were performed by the 10× Genomics Chromium Single Cell 3' Library & Gel Bead Kit v2 (10×Genomics, Pleasanton, CA, USA) and 10× Genomics Chromium barcoding system following the instructions provided by the manufacturer. The sequencing of prepared libraries was performed using Illumina NovaSeq 6000 (Illumina, San Diego, CA, USA) as per the manufacturer's instructions.

10 x sequencing data preprocessing

The raw sequencing data was analyzed by the Cell Ranger software (v3.1.0) provided by 10× Genomics to ensure the availability and validity of data for subsequent analysis. The STAR software was used to align sequences to the reference genome (GRCg6a). The matrix of gene expression versus cells were then used for downstream visualization analysis. According to gene number, UMI number, and mitochondrial gene expression distribution characteristics, the abnormal data of unbound cells and low-quality cells such as double cell, multicellular and dead cells were eliminated using the R package Seurat (v3.1.1). The delocalized value was found by fitting a distribution model based on the linear relationship between UMI number and gene number in each cell, and then the delocalized cells were filtered. Cells with gene and UMI numbers within the mean value ± 2 standard deviations and the proportion of mitochondrial genes < 30% were retained as high-quality cells for further analysis. After applying these quality control criteria, 10,344 cells in the DH_E10 group, 9,481 cells in the DH_E14 group, 9,754 cells in the DH_E18 group, 10,978 cells in the TC_E10 group, 12,406 cells in the TC_E14 group and 6,676 cells in the TC_E18 group were included in downstream analyses respectively.

Identification and characterization of cell clusters

Top variable genes across single cells were identified using the method described by Macosko et al. [22]. The most variable genes were selected using FindVariableGenes function in Seurat [23]. A principal component analysis (PCA) was performed to reduce the dimensionality of single-cell sequencing data by the RunPCA function implemented in Seurat. The batch effect between chicken leg muscle tissue samples was corrected by the Mutual nearest neighbors (MNN) algorithm in the R package before clustering. Cell clustering was performed based on the gene expression profile of cells using the FindClusters function implemented in Seurat. Cell clusters were visualized using a 2-dimensional tSNE algorithm with the RunTSNE function implemented in Seurat. We used the Seurat implemented FindAllMarkers function was used to detect differences in gene expression between the specified cluster and all other clusters, and the genes with up-regulated expression ($P < 0.05$) were screened as potential specific marker genes in each cluster. The CellMarker database (http://bio-bigdata.hrbmu.edu.cn/CellMarker/) and collected marker genes of known cell types form a reference data set for marker genes of cell types. Cell types were determined based on the Spearman correlation between the expression profiles of the cells to be identified and the reference data set calculated by the singleR software package.

Immunofluorescence staining

For immunofluorescence staining analysis, paraffin blocks encasing DH or TC leg muscle tissue were deparaffinized and rehydrated. The citrate buffer (pH 6.0) was used to fully expose the antigenic epitopes of paraffin Sect. (97.5℃, 20 min), cooled to room temperature (RT), and rinsed three times with 1×PBS for 5 min each time. To block endogenous peroxidase, paraffin sections were incubated with 3% hydrogen peroxide for 25 min (RT, in the dark). After rinsing three times with 1×PBS for 5 min each time, tissues were blocked for 40 min with 3% BSA at RT, followed by overnight incubation (4℃) with primary antibodies, all of which were commercially prepared by Affinity biosciences Co., Ltd, China, including *DCN* (#DF6543), *PAX7* (#AF7584), *PDGFRA* (#AF5489), *PRRX1* (#DF4274) and *MYF5* (#DF4274) and Beijing Biosynthesis Biotechnology Co., Ltd, China, including *MYOG* (#Bs-3550r). All antibodies at a dilution ratio of 1:100. Following three consecutive washes with 1× PBS for 5 min each time, tissues were incubated with a secondary antibody, FITC conjugated Goat Anti-Rabbit IgG (#GB22303) from Wuhan Servicebio Technology Co.,

Ltd, China for 30 min at 37℃. Paraffin sections were subsequentially washed three times with 1× PBS for 5 min each time and stained with DAPI (#G1012) from Wuhan Servicebio Technology Co., Ltd, China for 10 min at RT. Finally, paraffin sections were washed three times with 1×PBS for 5 min each time and the slides were sealed. The fluorescence images were observed and analyzed using an inverted fluorescence microscope (Olympus, Japan).

Pseudotime analysis

Single-cell trajectories were analyzed by using Monocle according to the instruction [24]. We used Monocle to perform machine learning based on the expression patterns of key genes to simulate the dynamic changes of temporal development. Firstly, the significantly affected genes were obtained among the clusters. Next, the spatial dimensionality reduction was performed according to the expression profile of selected genes, and then the minimum spanning tree (MST) was constructed. Finally, the cell differentiation trajectory with similar transcriptional characteristics was visualized through the longest path of MST. We used Monocle to implement the plot_pseudotime_heatmap function to plot pseudotemporal heatmap. To investigate gene functions in each gene module, gene ontology (GO) analysis was performed using clusterProfiler R package. Gene expression was plotted using the "plot_genes_in_ pesudotime" function to track changes along pseudotime.

GeneSwitches analysis

GeneSwitches analysis was performed to discover the ordering of critical gene (named switching gene) expression changes over pseudotime. Firstly, gene expression data in pseudotime trajectory were binarized to screen out potential switching genes with "on" or "off" gene expression states using the GeneSwitches package [25]. Subsequently, logistic regression analysis was performed for these potential switch genes to calculate the switching time point of each switching gene. The correlation analysis between switching genes and pseudotime was performed using McFadden's Pseudo R^2. Among them, the switching genes that were positively correlated with the pseudotime ($R^2 > 0$) were defined as the activated switching genes or up-regulated switching genes, while the silenced switching genes or down-regulated switching genes were negatively correlated with the pseudotime ($R^2 < 0$). The higher the pseudotemporal correlation ($|R^2|$), the closer the relationship between switching genes and the trajectory process. Switching genes were considered significant at $|R^2| > 0.2$ and the gene expression rate > 10%.

Cell-cell communication analysis

We used the R package CellChat (v1.6.0) with the CellChatDB databases including the information of ligand-receptors and cofactors interactions to infer, visualize, and analyze intercellular communication from scRNA-seq data [26]. We selected chicken genes according to their homologous with humans. CellChat identified differentially expressed signaling genes for all cell groups ($P < 0.05$) and calculated intercellular communication probability represented the intercellular interaction strength. In addition, the KEGG database was combined to provide information on the interactions of pathways and signaling molecules. The cell-cell communication network was calculated by summing the number or strength of ligand-receptor pairs of significant interaction in various cell types.

Results

ScRNA-Seq and characterization of Cellular Heterogeneity during skeletal muscle myogenesis

To decipher the transcriptome regulatory network and cellular fate decisions during skeletal muscle myogenesis, leg musculature from male embryos of DH group and TC group were dissected from embryos at E10, E14, and E18, digested into single cells and performed droplet-based scRNA-seq (Fig. 1A). We detected 20,904 (DH) and 20,107 genes (TC) in total from muscle cells at E10, 20,735 (DH) and 19,979 genes (TC) at E14, 19,960 (DH) and 18,589 (TC) genes at E18, respectively (Fig. S1A and S1B). After removing low-quality cells, we obtained 29,579 single-cell transcriptome profiles (10,344, 9,481, and 9,754 single cells at E10, E14, and E18) and 30,060 single-cell transcriptome profiles (10,978, 12,406, and 6,676 single cells at E10, E14, and E18) in DH and TC group from skeletal muscle cells, respectively. Then we performed tSNE analysis for all the single cells to gain an insight into the cellular heterogeneity during skeletal muscle myogenesis in different chicken breeds. Based on tSNE dimensionality reduction, we identified 17 transcriptionally distinct cell clusters (Fig. 1B). The cell number and proportion detected in different development stages of the two breeds were counted. Cluster 2, 5, 7, and 9 were dominant in E10 of DH (accounting for 34.19%, 10.17%, 12.82%, and 12.00%, respectively), cluster 2, 4, 7, and 11 were dominant in E10 of TC (accounting for 28.27%, 13.34%, 13.62%, and 17.05%, respectively); Cluster 3, 5, 6, and 10 were dominant in E14 of DH (accounting for 13.32%, 13.21%, 14.13%, and 16.76%, respectively), cluster 1, 3, 4, and 8 were dominant in E14 of TC (accounting for 21.64%, 15.45%, 13.20%, and 12.49%, respectively); Cluster 3, 6, 12, and 13 were dominant in E18 of DH (accounting for 30.68%, 10.96%, 10.25%, and 12.12%, respectively), cluster 1, 3, 5, and 12 were dominant in E18 of TC (accounting for 54.06%, 5.71%, 11.38%,

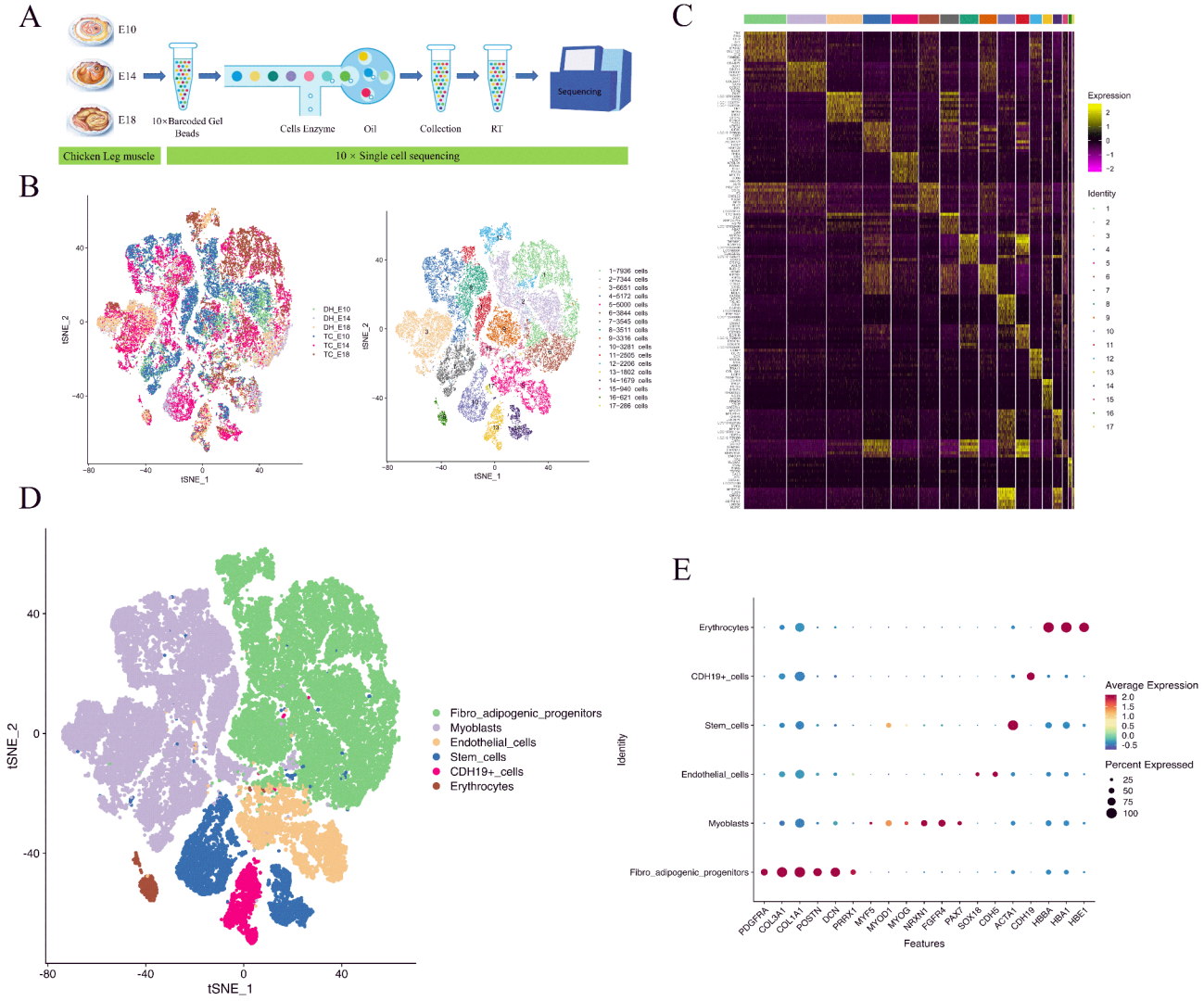

Fig. 1 Single cell transcriptome sequencing reveals multiple cell types in chicken skeletal muscle. (**A**) The scheme of muscle preparation, single-cell isolation, and scRNA-seq in E10, E14, and E18. (**B**) The tSNE visualization of 17 clusters derived from 59,639 high-quality cells filtered from 18 chicken leg muscle samples. (**C**) Heatmap showing the expression of top 10 cluster specific genes in each cluster. The pink-black-yellow color ramp represents an increase in expression. (**D**) The tSNE visualization of 6 cell types. E Dot plot showing the expression of 6 cell types marker genes

and 13.39%, respectively) (Fig. S1C). We found that most of the cell clusters showed significant change from E10 to E18 in two breeds which preliminarily deciphered that skeletal muscle cells were highly heterogenous. We also found that some of the cell clusters showed similar trends in the two breeds while others revealed distinctly different trends. The cluster-specific gene expression was compared across the cell clusters and the top 10 expressed cluster-specific genes showed obvious cluster-specific expression (Fig. 1C).

To further characterize cell cluster identity, we evaluated the expression of a combination of cell marker genes and identified six major cell types: *MYF5,MYOD1,MYOG,NRXN1,FGFR4*, and *PAX7* high expressed in myoblasts [27, 28], *PDGFRA,COL3A1,COL1A1,POSTN,DCN*, and *PRRX1* high expressed in fibro-adipogenic progenitors (FAPs) [29, 30], *ACTA1* high expressed in stem cells [31], *SOX18* and *CDH5* high expressed in endothelial cells [30], *HBBA,HBA1*, and *HBE1* high expressed in erythrocytes [28, 30], and *CDH19* high expressed in *CDH19*$^{+}$ cells (Fig. 1D and E) and (Fig. S2A and S2B). We also performed uniform manifold approximation and projection (UMAP) analysis for all the single cells, which was consistent with the result of tSNE analysis (Fig. S2C). Besides, immunofluorescence staining of cell type marker genes was used to verify myoblasts and FAPs in muscle tissues and the accuracy of cell types (Fig. S3). Taken together, these results identified major cell populations, identified a series of cell identity

specific signature genes, and revealed cellular heterogeneity during myogenesis in the skeletal muscle.

Clustering and pseudotemporal trajectories identified transcriptional dynamics of myoblasts

Unsupervised clustering was performed on the myoblast population, which identified a total of six unique subclusters that we labeled as subclusters 1 to subclusters 6 (Fig. 2A). Specifically, subclusters 3 and 4 expressed high levels of *PAX7* and *MYF5* (Fig. 2B), thus we defined subclusters 3 and 4 as muscle satellite cells. Subclusters 1 and 5 were marked by the expression of myoblast markers, *CDC20* and *PCNA* (Fig. 2B). Most of the cells in subclusters 2 and 6 expressed the myocytes markers, *MYOG* and *MYBPH* (Fig. 2B). Interestingly, we noticed different trends of proportion variation of muscle satellite cells, myoblasts, and myocytes in the two breeds (Fig. 2C).

We then used pseudotime trajectory analysis to analyze the progression of continuous cell states of myoblasts and revealed ordered cells expressing different levels of marker genes in a trajectory. Ordering of cells in trajectory analysis arranged the six subclusters into one major trajectory with four minor bifurcations and nine stages (Fig. 2D) and (Fig. S4). The majority of muscle satellite cells were located toward the origin of the trajectory in both DH and TC. There was a fraction of satellite cells that ended at 3 bifurcations (one at the start and two at the end) in DH while 2 bifurcations (at the start) in TC. Myocytes were located toward or at the end of the major branch and 1 bifurcation at the end of the trajectory in the two breeds. Myoblasts were located toward or at the end of the major branch and 3 bifurcations at the end of the trajectory in DH, whereas myoblasts in TC were located toward or at the origin of the major branch and 3 bifurcations at the start of the trajectory (Fig. S4). The dynamic changes of marker gene expression along the major branch were investigated. High expression of *MYF5* and *PAX7* was observed at the origin of pseudotime, while *MYOG* and *MYBPH* were enriched at the end of pseudotime (Fig. 2E). These results could serve as a validation for the differentiated trajectory of satellite cells. The RNA velocity map showed that myocytes developed from myoblasts, and myoblasts developed from muscle satellite cells, due to the high RNA velocity that began between them (Fig. S5). The RNA velocity of these cell types was consistent with the pseudotime path. Intriguingly, we found that some of the myoblasts turned into muscle satellite cells, which suggested that activated myoblasts either go to differentiation or enter into quiescence. Quiescent satellite cells were mainly responsible for postnatal skeletal muscle development [32].

We next examined pseudotime dynamics of significantly changed genes along the differentiation trajectory. Four modules were arranged according to their pseudotemporal expression patterns (Fig. 3). The pseudotime heatmap showed that module 1 of DH and TC were both enriched in skeletal muscle fiber adaptation (*ACTA1*), skeletal muscle satellite cell commitment (*PAX7*), skeletal muscle cell proliferation (*SIX1* and *WNT4*), and regulation of mitotic cell cycle (*BACH1* and *E4F1*). Modules 2 and 3 of the two breeds revealed clear enrichment in the regulation of myoblast fusion (*MYOG* and *TMEM8C*), muscle cell fate commitment (*MYOD1,MYOG*), regulation of skeletal muscle satellite cell proliferation (*MYOG,GPC1*) and differentiation (*MEF2C,GPC1,DDX17*, and *KLHL41*), skeletal muscle thin filament assembly (*MYBPC3,MYBPH* and *MYOM1*), and skeletal muscle myosin thick filament assembly (*MYBPC3,MYBPH,MYOM1,MYOM3,OBSL1* and *TCAP*). Modules 4 of two breeds were enriched in striated muscle cell development (*SDC1*), and skeletal muscle tissue development (*POGLUT1,ANKRD1,CCNT2,CFL2,COL19A1*, and *RHOA*) (Fig. 3, Table S1 and S2). Collectively, these results provide insights into the developmental trajectory of skeletal muscle cells during cell state transitions.

Interestingly, we found that module 4 of DH was enriched in the regulation of skeletal muscle satellite cell proliferation (*MSTN,MEGF10*, and *GPC1*) and differentiation (*MEGF10*, and *GPC1*) (Fig. 3A). In addition, we found that the expression of *MYF5* and *PAX7* were significantly downregulated at the pseudotime in TC, whereas the levels of *MYF5* and *PAX7* first decreased and then rapidly increased in DH. We also observed that the expression of *MYOG* and *MYBPH* were significantly upregulated at the later pseudotime in TC, while their levels first increased and then rapidly decreased in DH (Fig. S6). These data indicated that satellite cells and myoblasts display distinct proliferation and differentiation in the two breeds.

Dynamic changes of gene expression in myoblasts

To investigate the dynamic changes of gene expression and how the order is maintained in different muscle cell types, we conducted a gene-switch analysis at single-cell resolution based on the genes detected in all clusters. The best-fitting top 35 genes and top 15 transcript factors (TFs) were plotted along the pseudotime trajectory (Fig. 4 and Table S3).

In the myoblasts of DH, 2 genes were significantly inactivated, and 179 genes were significantly activated, including 4 early genes, 82 mid genes, and 93 late genes (Fig. 4A and Table S4). Ribosomal proteins are essential components of the protein biosynthetic. We found 6 small nuclear ribonucleoprotein (SNRP) genes (*SNRPD1,SNRPG,SNRPE,SNRPE,SNRPL*, and *SNRPD2*) were activated in mid and late stages which suggested protein synthesis was active in muscle cells in these two stages.

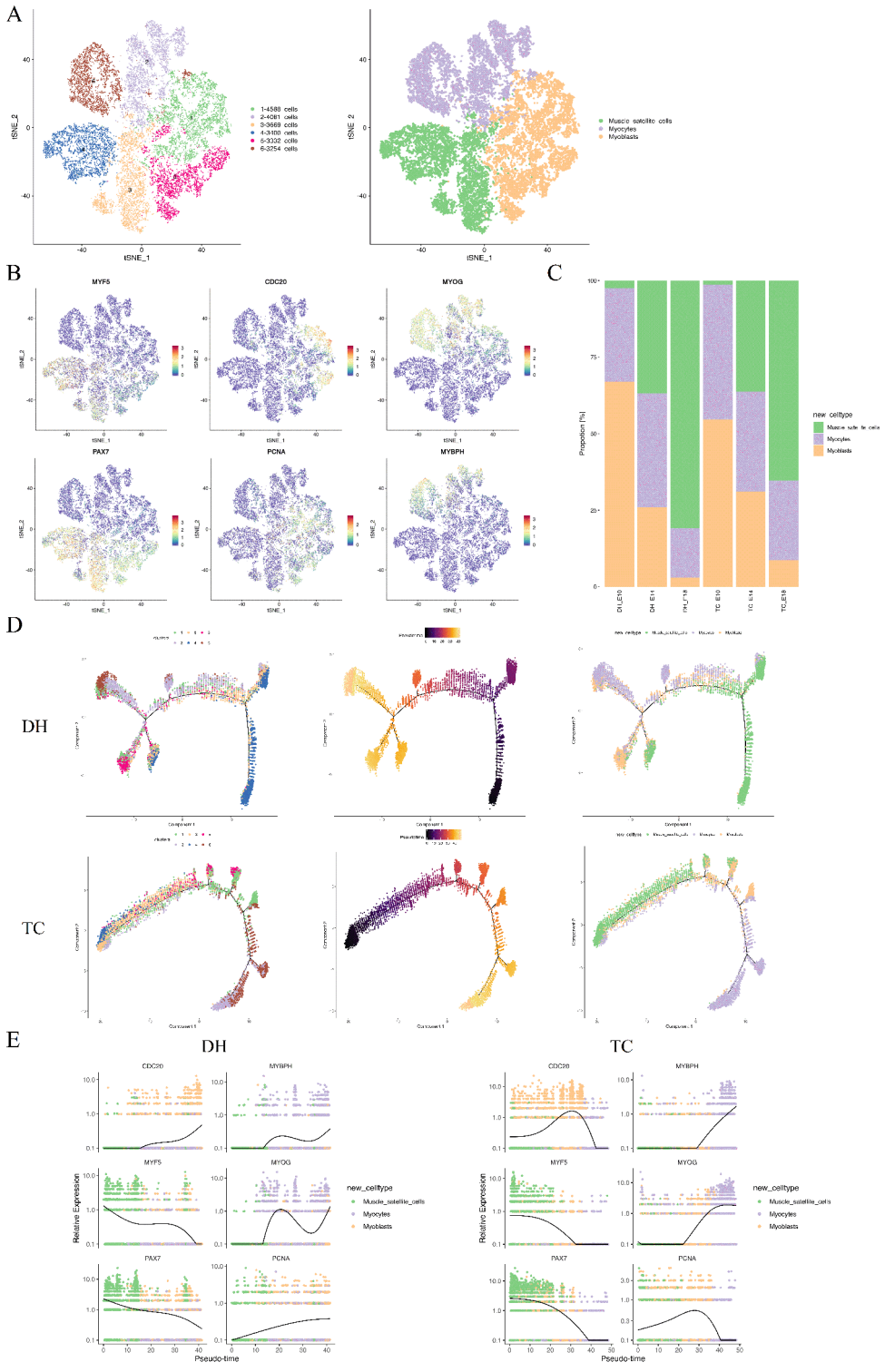

Fig. 2 Identification of myoblasts subtypes and pesudotime trajectory analysis in DH and TC. (**A**) The tSNE visualization of 6 myoblasts subclusters and 3 myoblast subtypes. (**B**) The tSNE plots of marker genes for muscle satellite cells (*MYF5* and *PAX7*), myoblasts (*CDC20* and *PCNA*), myocytes (*MYOG* and *MYBPH*). (**C**) The bar chart showing the proportion of 3 myoblasts subtypes in 6 samples. (**D**) Trajectories of myoblasts in DH and TC along pseudotime. The colors from dark (purple) to light (yellow) represent the forward order of pseudotime. (**E**) Expression patterns of 3 myoblasts subtypes marker genes along pseudotime in DH and TC. Various colored dots represent distinct myoblasts subtypes

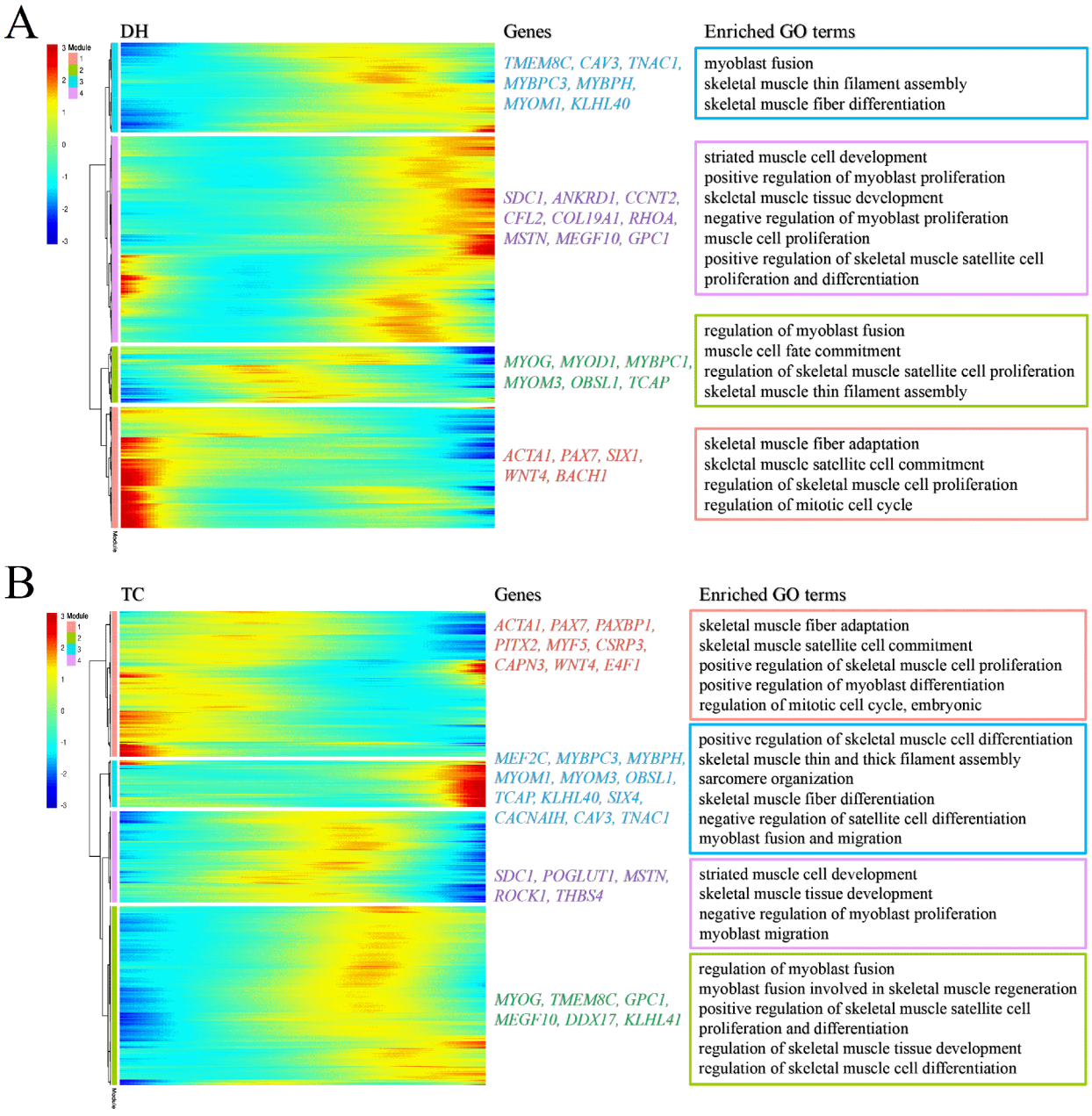

Fig. 3 The pseudotemporal heat map analysis of significantly altered genes on DH (**A**) and TC (**B**) myoblasts differentiation trajectories. Genes (Y-axis) were clustered into four modules, and cells (X-axis) were ordered according to pseudotime. GO terms enriched for a portion of each gene set were labeled in the right panel

Strikingly, *SNRPG*, and *SNRPE* were activated early, indicating that they might be critical genes for protein synthesis and skeletal muscle development. Some mid and late genes that may play crucial roles in muscle cells were identified. For example, *EIF4EBP1* and *HSP90AB1* activated in mid stage regulated mitotic cell cycle. *CHRNA1* may regulate skeletal muscle contraction and muscle cell cellular homeostasis. *PLS3* and *SVIL* were shown to be crucial for the regulation of actin filament binding, actin filament bundle assembly, and actin filament network formation (Fig. 4B and Table S4).

In the muscle cell types of TC, 5 genes were significantly inactivated, and 65 genes were significantly activated, including 10 early genes, 39 mid genes, and 16 late genes (Fig. 4C and Table S4). Some early and mid genes that may play crucial roles in muscle cells were identified. For example, we detected *RBM24* and *MAP1A* activated in early stages which have been indicated to be crucial for the regulation of myoblast differentiation

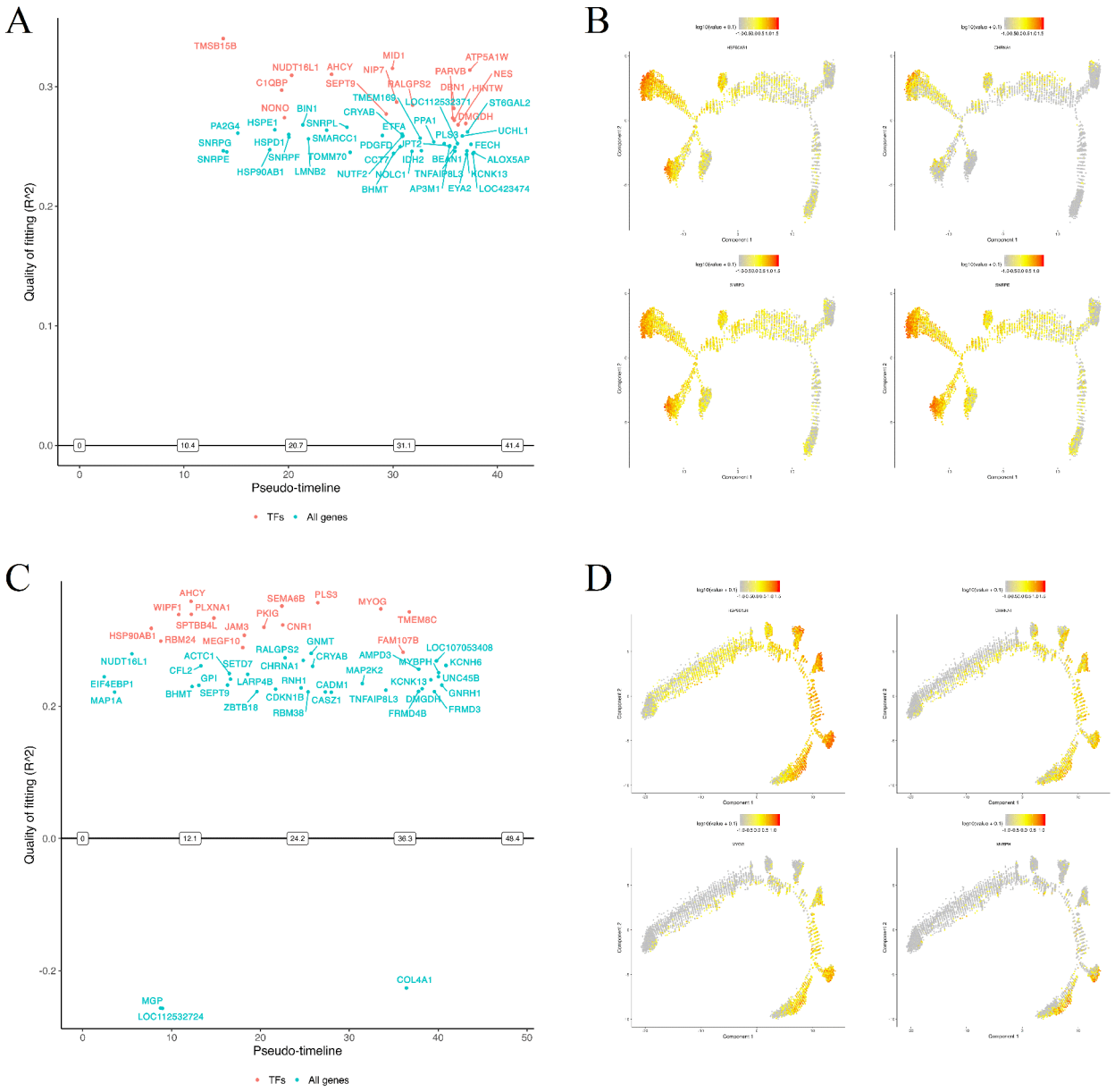

Fig. 4 Schematic diagram of priority activated and inactivated genes in myoblasts. (**A**) and (**C**) Representative activated (top panel) and inactivated (bottom panel) genes in DH and TC during the development of Myoblasts. The X-axis shows the predicted open (upper panel) or closed (lower panel) times. The Y-axis represents the goodness of fit. Early genes, switch-at-time < 10.4 in DH and < 12.1 in TC; late genes, switch-at-time > 31.1 in DH and > 36.3 in TC. (**B**) and (**D**) Trajectory along the pseudotime progression of representative genes that were activated and inactivated in DH and TC

and cytoskeleton organization. Interestingly, we found that *EIF4EBP1* and *HSP90AB1* were activated in the early stage of TC while in the mid stage of DH, and *CHRNA1*,*PLS3* and *SVIL* were activated in the mid stage of TC while in the late stage of DH. As myocyte markers, *MYOG* and *MYBPH* have been indicated to regulate myoblast differentiation and myoblast fusion. Remarkably, *MYOG* and *MYBPH* were activated in the late stage, which suggested they were vital for skeletal muscle fiber development (Fig. 4D and Table S4). The switching state of *MYOG* and *MYBPH* was consistent with the pseudotime path. Collectively, these findings are meaningful for revealing the dynamic changes in gene expression in muscle cells of two breeds, and many insights into gene activation provided here deserve further study.

Clustering and pseudotemporal trajectories identified transcriptional dynamics of fibro-adipogenic progenitors

Of the total clusters, FAPs revealed a gene expression pattern that could be assigned to adipocyte-derived stem cells. Twelve juxtaposed subclusters labeled as subcluster 1 to subcluster 12 were identified by using unsupervised clustering on the FAPs (Fig. 5A). Subclusters 1, 4, 5, and 7 were defined as myofibroblasts, which expressed relatively higher levels of myofibroblast markers (*MYL9,CDH11*, and *PALLD*) (Fig. 5B). Subclusters 2 and 9 expressed high levels of *PDGFRA*, thus we defined subclusters 2 and 9 as FAPs (Fig. 5B). Subclusters 3, 6, 8, and 11 were identified as differentiating adipocytes by the expression of *COL1A1,COL6A3,MGP*, and *APOA1* (Fig. 5B). Subcluster 10 was marked by the expression of classic tenocyte marker (*SCX*) (Fig. 5B). Subcluster 12 was marked by the expression of classic adipocyte marker (*ADIPOQ*) (Fig. 5B). The 6 FAPs subtypes showed different trends of proportion variation in DH and TC (Fig. 5C).

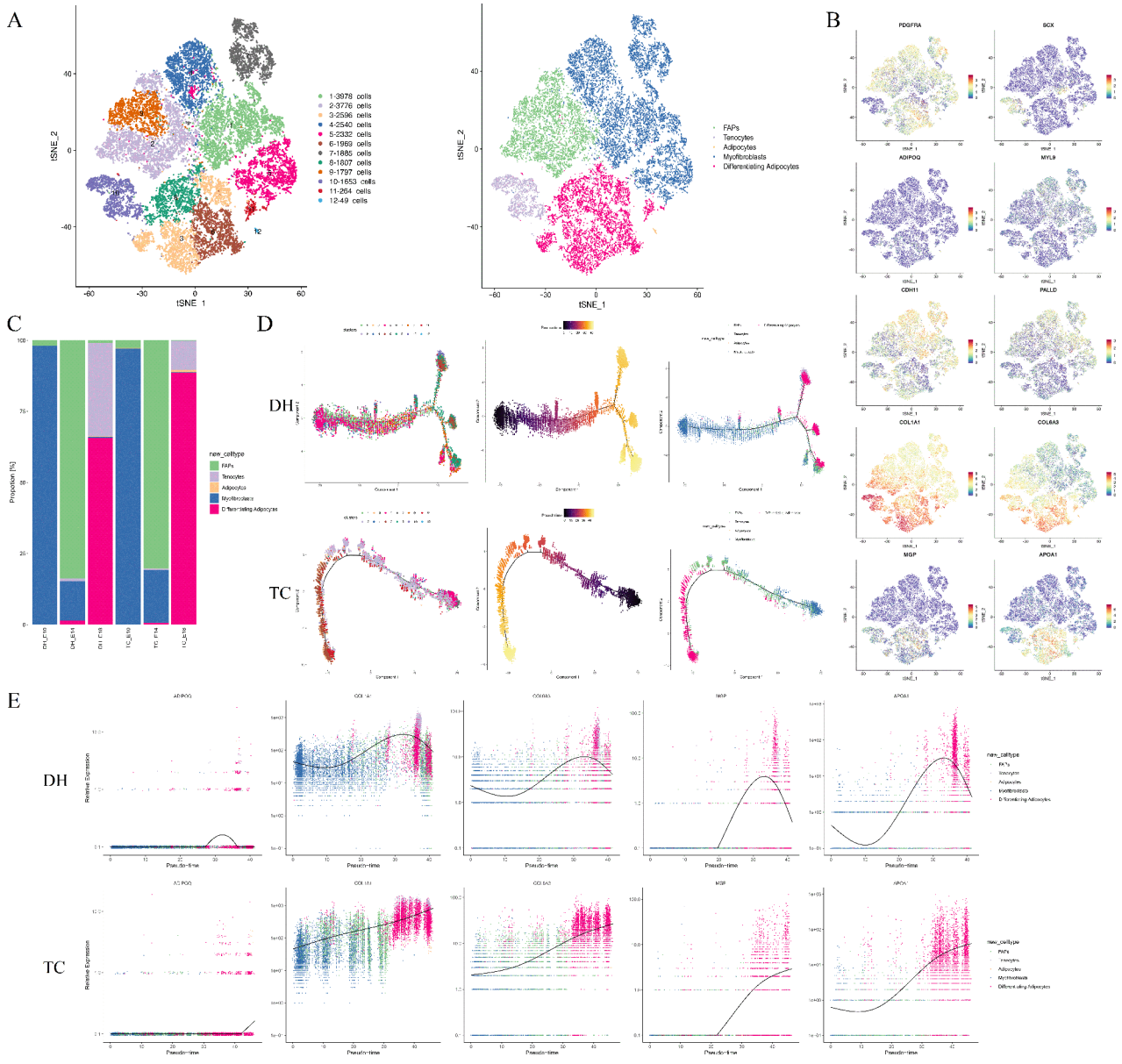

Fig. 5 Identification of FAPs subtypes and pesudotime trajectory analysis in DH and TC. (**A**) The tSNE visualization of 12 FAP subclusters and 5 FAP subtypes. (**B**) The tSNE plots of marker genes for FAPs (*PDGFRA*), tenocytes (*SCX*), adipocytes (*ADIPOQ*), myofibroblasts (*MYL9,CDH11*, and *PALLD*), and differentiating adipocytes (*COL1A1,COL6A3,MGP*, and *APOA1*). (**C**) The bar chart showing the proportion of 5 FAPs subtypes in 6 samples. (**D**) Trajectories of FAPs in DH and TC along pseudotime. The colors from dark (purple) to light (yellow) represent the forward order of pseudotime. (**E**) Expression patterns of the adipocytes and differentiating adipocytes marker genes along pseudotime in DH and TC. Various colored dots represent distinct FAPs subtypes

We then used Monocle to analyze dynamic cell transitions of FAPs during the myogenesis, which arranged the twelve subclusters into one major trajectory with four minor bifurcations and nine stages in DH, whereas only two bifurcations and five stages in TC (Fig. 5D and S7). FAPs were located toward the major branch and bifurcations of the trajectory in two breeds. Myofibroblasts were located toward the major branch and bifurcations of the trajectory in DH, but myofibroblasts in TC were located toward or at the origin of the major branch and two bifurcations of the trajectory. Differentiating adipocytes, tenocytes, and adipocytes were all located toward or at the end of the major branch and bifurcations at the end of the trajectory in the two breeds. Intriguingly, we found that differentiating adipocytes appeared earlier in TC compared to DH, and more adipocytes were detected in TC (Fig. S7). We also investigated the dynamic changes of marker gene expression along the trajectory (Fig. 5E and S8). We found that the expression of differentiating adipocytes marker genes (*COL1A1,COL6A3,MGP*, and *APOA1*) was downregulated at the end of pseudotime in DH, while increased at the end of pseudotime in TC. Similar expression trends of the adipocyte marker gene (*ADIPOQ*) were also observed at the end of pseudotime in two breeds. These results could serve as a validation for the differentiated trajectory of FAPs.

We also examined pseudotime dynamics of significantly changed genes along the differentiation trajectory. Four modules were arranged according to their pseudotemporal expression patterns (Fig. 6). The pseudotime heatmap showed that module 1 of DH and TC were both enriched in fibroblast growth factor production (*SHH*), negative regulation of collagen binding (*GTPBP4*), regulation of fatty acid metabolic process (*ACSL4,EIF2AK3*, and *PRKAG2*), regulation of triglyceride biosynthetic process (*LOC107057318*), regulation of lipid metabolic process (*ACSL1,CHD9,CPT2,DRD3,FADS1,FADS1L1,FDFT1,GOLM1,H2AFY,LACTB,MED1,NCOA2,NCOA6,NCOR1,NCOR2,PLIN2,PPARG,PPP4R3B,SIN3A,SMARCD3,TGS1,TIAM2*, and *TNFRSF21*), and positive regulation of adipose tissue development (*SH3PXD2B*) (Fig. 6, Table S5 and S6). Modules 2 of TC revealed clear enrichment in positive regulation of fatty acid metabolic process (*ADIPOQ*), positive regulation of fat cell proliferation (*PID1*), positive regulation of fat cell differentiation (*CEBPB,CMKLR1,CREBL2,ID2*; *RARRES2,WDFY2,WNT5B*, and *ZNF385A*), regulation of lipid biosynthetic process (*MID1IP1*), very-low-density lipoprotein particle remodeling (*APOA1* and *LPL*), and regulation of cholesterol transport (*APOA1*) (Fig. 6B and Table S6). Modules 4 of DH showed similar enrichment with modules 2 of TC, such as positive regulation of fatty acid metabolic process (*ADIPOQ*), positive regulation of fat cell proliferation (*PID1*), positive regulation of fat cell differentiation (*CCDC71L,CEBPB,CMKLR1,ID2,NOCT,RARRES2,TMEM64,XBP1*, and *ZBTB7C*) (Fig. 6A and Table S5). Altogether, these results improve our understanding of the developmental trajectory of FAPs.

Dynamic changes of gene expression in FAPs

To investigate the dynamic changes of gene expression and how the order is maintained in different FAP cell types, we conducted a gene-switch analysis at single-cell resolution based on the genes detected in all clusters. The best-fitting top 35 genes and top 15 TFs were plotted along the pseudotime trajectory (Fig. 7 and Table S7).

In the FAPs of DH, 1 gene was significantly activated, and 100 genes were significantly inactivated, including 44 early genes, 40 mid genes, and 16 late genes (Fig. 7 and Table S8). In the FAPs of TC, 10 genes were significantly activated, including 1 mid gene, and 9 late genes, and 8 genes were significantly inactivated, including 5 early genes, and 3 mid genes (Fig. 7C and Table S8). We found that *CRABP1* was inactivated in the early stage of TC while in the mid stage of DH. *CRABP1* binds retinoic acid specifically in the cytoplasm, and *Crabp1* knockout mice exhibited an adult-onset amyotrophic lateral sclerosis-like phenotype [33]. We also found that *TMSB15B* and *STMN1* were inactivated in the mid stage of TC while in the late stage of DH. *STMN1* is a microtubule-destabilizing phosphoprotein critically involved in cell cycle progression, cell motility and survival [34]. *TMSB15B*, a small actin-binding protein, plays an important role in the organization of the cytoskeleton. In addition, *FMOD* related to extracellular matrix (ECM) was inactivated in the early stage of DH, but activated in the late stage of TC. Remarkably, as a differentiating adipocyte marker, *APOA1* was only found activated in the mid stage of TC, and *MGP* were only found activated in the late stage of TC (Fig. 7C and D and Table S8). These results collectively provide important insights into the different dynamic changes in gene expression in FAPs of two breeds.

Ligand-receptor interaction prediction during skeletal muscle development

Intercellular communication based on ligand-receptor interactions is crucial for tissue to accomplish complex life activities [35]. Cell-cell contact is common in skeletal muscle growth and development [36]. To further explore cellular interaction during skeletal muscle development, we performed a comparative analysis of cellular communication between 6 identified cell types in skeletal muscle of two chicken breeds, including FAPs (myofibroblasts, FAPs, differentiating adipocytes, tenocytes, and adipocytes), myoblasts (muscle satellite cells, myoblasts, and myocytes), endothelial cells, stem cells, $CDH19^{+}$ cells, and erythrocytes (Fig. 8). We detected a total of 8,175

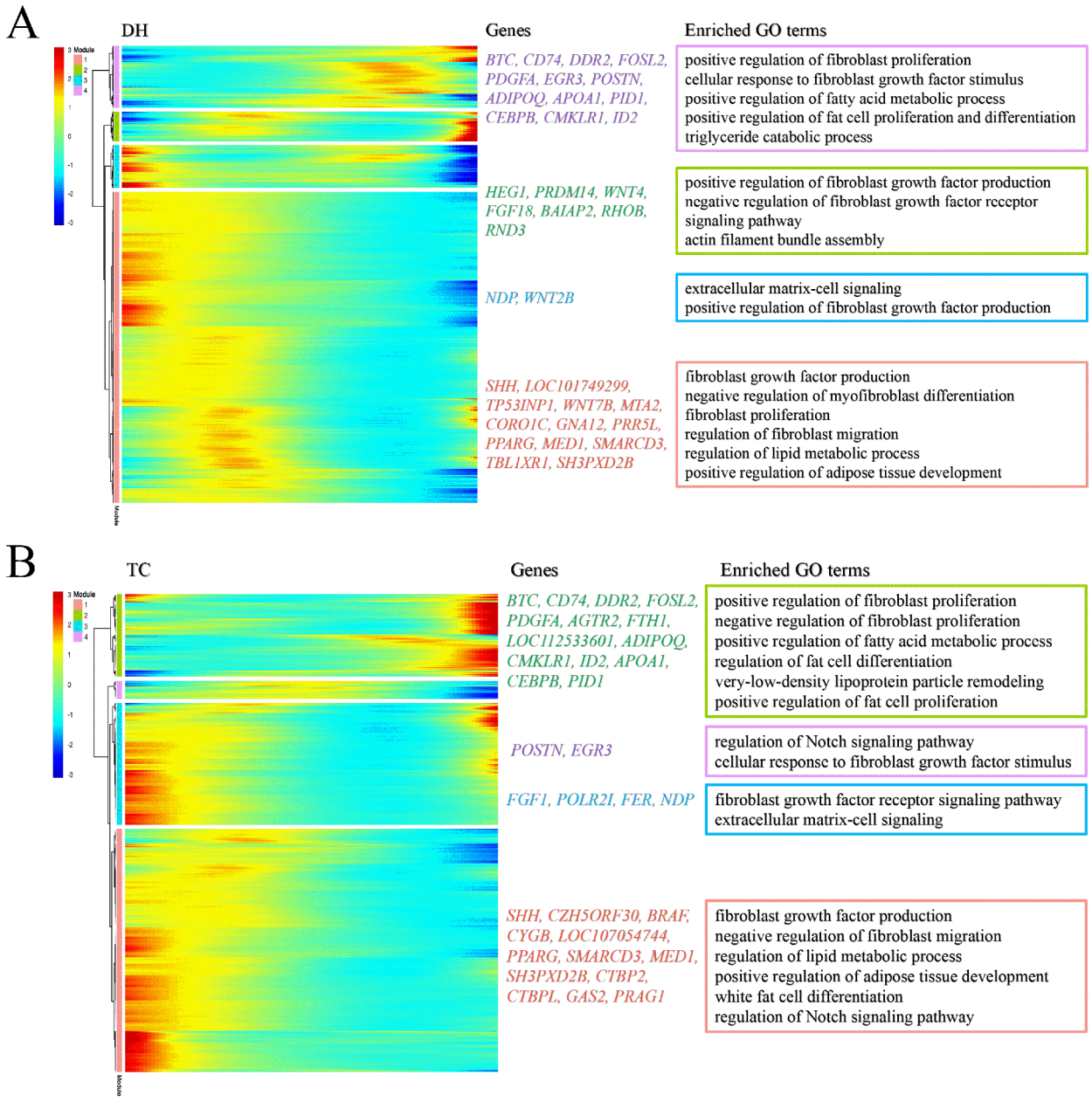

Fig. 6 The pseudotemporal heat map analysis of significantly altered genes on DH (**A**) and TC (**B**) FAPs differentiation trajectories. Genes (Y-axis) were clustered into four modules, and cells (X-axis) were ordered according to pseudotime. GO terms enriched for a portion of each gene set were labeled in the right panel

ligand-receptor pairs during the three developmental stages ($P < 0.05$). At E10, 1,335 and 1,084 ligand-receptor pairs were detected in DH and TC, respectively. Compared with TC, there were more interaction pairs among FAPs, myofibroblasts, myoblasts, and muscle satellite cells in DH.

Interestingly, the interaction number and strength among muscle satellite cells, myocytes, and myoblasts in TC were higher than DH (Fig. 8A, S9A and Table S9). At E14, 1,420 and 1,442 ligand-receptor pairs were detected in DH and TC, respectively. The number of interactions between muscle satellite cells and other cells in DH was greater than that in TC. In addition, we also found that endothelial cells communicated more actively with other cells in DH than in TC (Fig. 8B, S9B and Table S9). At E18, there were 1,298 and 1,596 ligand-receptor pairs in DH and TC, respectively. There were more interaction pairs among myofibroblasts, myoblasts, muscle satellite cells, and tenocytes in DH than in TC and the interaction pairs between differentiating adipocytes and other

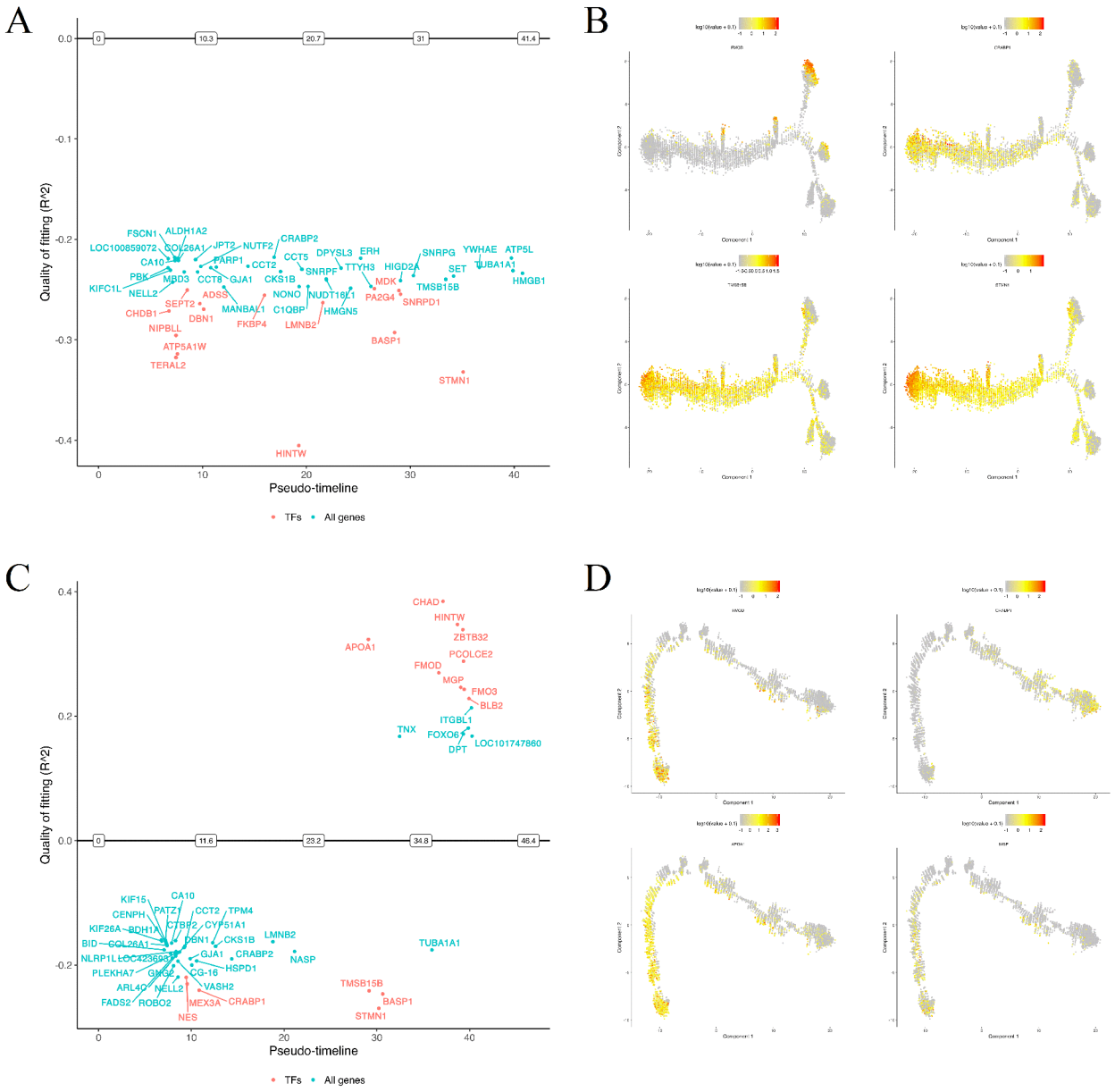

Fig. 7 Schematic diagram of priority activated and inactivated genes in FAPs. (**A**) and (**C**) Representative activated (top panel) and inactivated (bottom panel) genes in DH and TC during the development of FAPs. The X-axis shows the predicted open (upper panel) or closed (lower panel) times. The Y-axis represents the goodness of fit. Early genes, switch-at-time < 10.3 in DH and < 11.6 in TC; late genes, switch-at-time > 31.0 in DH and > 34.8 in TC. (**B**) and (**D**) Trajectory along the pseudotime progression of representative genes that were activated and inactivated in DH and TC

cells was increased. Importantly, no communication information was detected in adipocytes in DH, whereas it was widely detected in TC. In particular, the number of interactions between differentiating adipocytes (source cells) and adipocytes (target cells) was the highest in TC (Fig. 8C, S9C and Table S9). It is worth noting that the interaction of CDH19$^+$ cells with other cells was widespread during all three developmental stages in both two breeds, which suggested that CDH19$^+$ cells may play an important role in skeletal muscle development.

The ligand-receptor pairs were further categorized into 42 signaling pathways, with COLLAGEN and LAMININ pathways showing highly abundant signaling interactions among 6 samples (Fig. 8D). In the COLLAGEN pathway, other cells mainly targeted tenocytes and CDH19$^+$ cells, while interactions between 3 myoblast subtypes and other cells were mainly observed in the

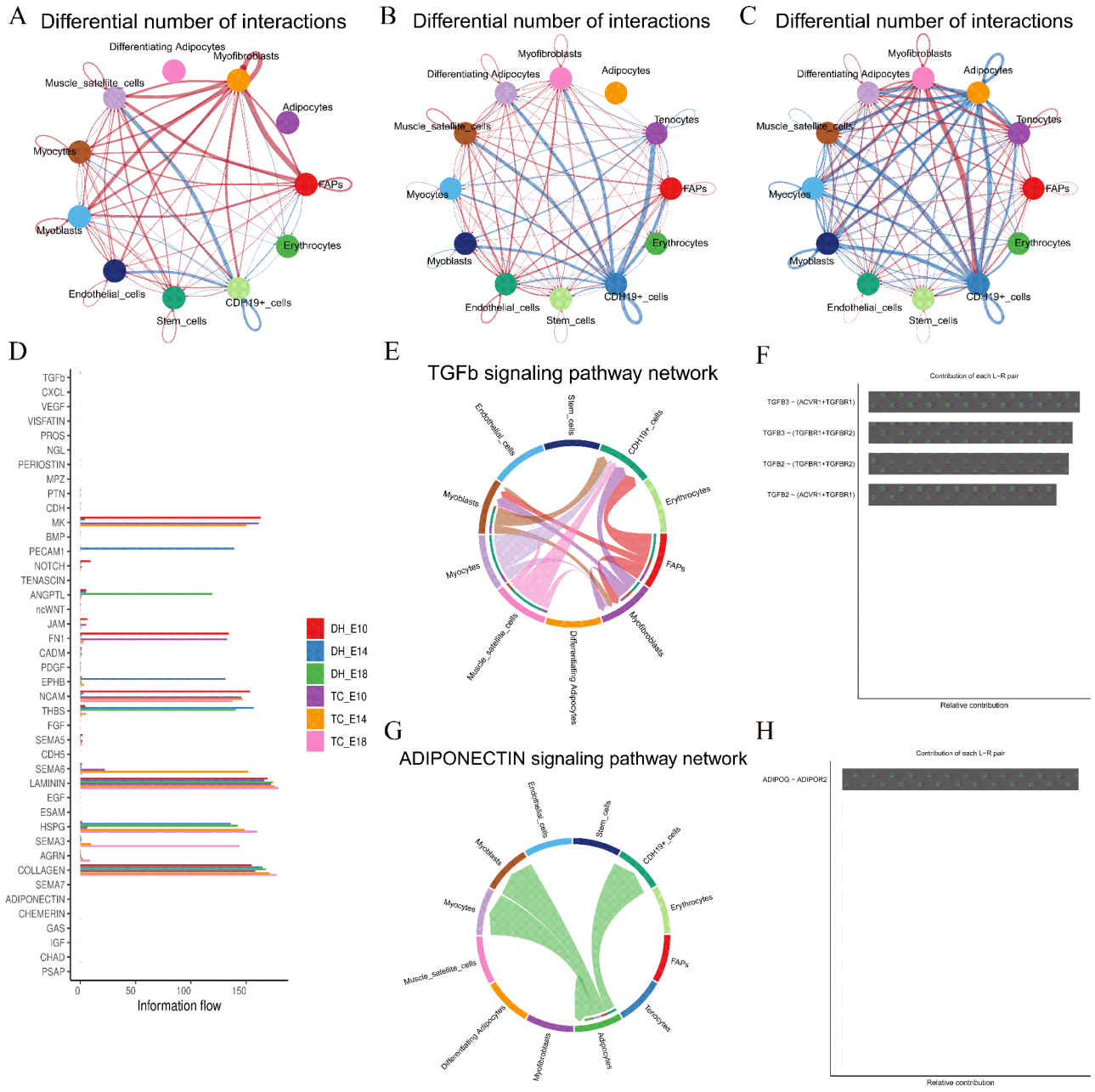

Fig. 8 Cell-cell communications among various cell types in chicken embryonic skeletal muscle. (**A**)-(**C**) Comparison of the number of intercellular communications between DH and TC on E10, E14, and E18. TC was used as control group in all comparison, the red and blue lines indicate an increase and decrease in the intercellular interaction number in DH compared to TC, respectively. (**D**) Comparison of information flow of each signaling pathways in 6 samples. (**E**) and (**G**) Chord plots of intercellular interaction network in TGFb and ADIPONECTIN signaling pathway. Arrows within the circle indicate the information flows of the communication pathways. The outer and inner circles indicate the sender and receiver, respectively. (**F**) and (**H**), Relative contribution of each ligand-receptor pair to the overall TGFb signaling pathway and ADIPONECTIN signaling pathway

LAMININ pathway (Fig. S10A, SS11A and Table S9). We found that ligands *COL1A2* and *COL4A1*, receptors *ITGA1+ITGB1*,*ITGA11+ITGB1*, and *CD44* made higher contributions in the COLLAGEN pathway of all 6 samples (Fig. S10B). In the LAMININ pathway, ligands *LAMC1*,*LAMB1* and *LAMA4*, receptors *ITGA6+ITGB1* and *DAG1* made higher contributions (Fig. S11B). Some signaling pathways appeared only in DH, such as VISFATIN (at E10, E14, and E18), PROS (at E10 and E18), and TGFb (at E10) (Fig. 8E, S12A and S12B). The interactions between muscle satellite muscle, myoblasts, myocytes, and CDH19^{+} cells involved with *NAMPT* as the ligand, and *INSR* as the receptor in the VISFATIN signaling pathway. In the PROS signaling pathway, the main

interaction involved PROS1 as the ligand and TYRO3 as the receptor, with FAPs, myofibroblasts, differentiating adipocytes, myoblasts, and CDH19+. The TGFb signaling pathway only appeared only at E10, FAPs, myofibroblasts, myoblasts, muscle satellite cells, myocytes, and CDH19^{+} cells communicated and interacted via ligands *TGFB2* and *TGFB3*, which were receptors of *ACVR1* + *TGFBR1* and *TGFBR1* + *TGFBR2* (Fig. 8F, S12A and S12B). Additionally, the CHEMERIN *(RARRES2-GPR1*) and ADIPONECTIN (*ADIPOQ-ADIPOR2*) signaling pathways were only found in TC at E18 and were specific to adipocytes communication (Fig. 8G and H and S12c). In the ADIPONECTIN signaling pathway, adipocytes received signals as ligand cells, while adipocytes acted as receptor cells to send signals in the CHEMERIN signaling pathway. Together, the results of cell communication analysis provide insights into the dynamic changes of various cells during skeletal muscle development in the two breeds.

Discussion

Skeletal muscle formation is a highly precise and multi-step process, which requires the proliferation of myoblasts, permanent withdrawal from the cell cycle, differentiation, migration, and final fusion into multinucleated muscle fibers [37]. According to the characteristics of muscle development, the number of muscle fibers is already fixed at the embryonic stage [13]. Therefore, the embryonic stage is the key period for the growth and development of skeletal muscle in livestock and poultry. Poultry is a common model for studying the development of embryonic muscle, but the cell profile corresponding to embryonic skeletal muscle development in chickens is not yet complete. In this study, we selected two chicken breeds and applied scRNA-seq to provide a comprehensive resource describing the molecular signatures of skeletal muscle growth. We constructed a comprehensive single-cell transcriptomic atlas of 29,579 and 30,060 high-quality cells filtered from three DH and TC samples, respectively. Our results revealed the heterogeneity of muscle tissue cells during developmental stages, the developmental trajectories of myoblasts and FAPs, the gene expression status in myoblasts and FAPs, and the spatiotemporal regulation of skeletal muscle development. These findings provide insights into the phenotypic difference in skeletal muscle development of two breeds at single-cell resolution.

In the study, the heterogeneous myoblast population was subdivided into six subclusters and further identified as muscle satellite cells, myoblasts, and myocytes. Ordering of cells in trajectory analysis arranged the three cell subpopulations of two breeds into one major trajectory which is consistent with embryonic myogenesis [3]. Nevertheless, we found that a fraction of satellite cells ended at two bifurcations at the end of the trajectory only in DH. We well known that the growth or hypertrophy of muscle volume and the repair or regeneration of muscle injury after birth are mainly achieved through the cascade of muscle satellite cells [38], which play an important role in the regulation of muscle development. *PAX7* is the most widely recognized marker of satellite cells ubiquitously expressed in all satellite cells, including activated and quiescent ones [39, 40]. *MYF5* is also discovered as a marker of satellite cells, which is specifically expressed by myogenic cells, but not expressed in quiescent satellite cells [41]. The dynamic changes of *PAX7* and *MYF5* were consistent with the pseudotime trajectories in the two breeds. Based on the pseudotemporal trajectories of myoblasts, GO analysis revealed that the regulation of skeletal muscle satellite cell proliferation was activated at the late pseudotime in DH which was also consistent with the RNA velocity. It is acknowledged that *Pax3* and *Pax7* positive cells constitute a reserve of muscle satellite cells. However, *Pax3* was weakly expressed in both two chicken breeds, the reason may be that *Myf5* is genetically downstream of *Pax3* in the myogenic context [42]. *MYOG* is essential for the terminal differentiation of myogenic cells and myogenin-null mice quickly die after birth from severe and global muscle deficiency [43]. *MYOG* was found significantly upregulated at the later pseudotime in TC, which illustrated that satellite cells had thoroughly quit cell proliferation, and began cell differentiation and fusion. However, the upregulation of *PAX7* and *MYF5* and downregulation of *MYOG* in DH indicated that the proliferation of satellite cells was not over yet. TC showed earlier embryonic myogenesis and less myoblasts compared with DH broiler. Consistent with our results, pigs with less meat production show earlier embryonic myogenesis and less myogenic progenitors [44–46]. These differences may be the reasons for the phenotypic difference of skeletal muscle in the two chicken breeds.

We investigated the dynamic changes of gene expression in different muscle cell types by pseudotime and gene-switch analysis for understanding their biological functions in the skeletal muscle. *EIF4EBP1,HSP90AB1,CHRNA1,PLS3* and *SVIL* were found activated earlier in TC than DH. GO enrichment showed that *EIF4EBP1* and *HSP90AB1* were related to positive regulation of mitotic cell cycle, *CHRNA1,PLS3* and *SVIL* were crucial for the regulation of muscle cell cellular homeostasis and actin filament. The results revealed that skeletal muscle development in TC is faster compared to DH. *MYOG* and *MYBPH* were activated in the late stage of TC which also verified skeletal muscle development in TC is faster than DH. The spliceosome is a large ribonucleoprotein complex that is comprised of five SNRPs and numerous proteins that guide pre-mRNA splicing in eukaryotic cells [47]. Six SNRP genes including *SNRPD1,SNRPG,SNRPE,SNRPF,SNRPL*, and *SNRPD2*

were activated in DH, and in which of them, *SNRPG* and *SNRPE* were activated early. Being the core spliceosome associated protein, the downregulation of *SNRPG* induced cell cycle arrest [48]. Together, these findings clearly reveal that protein biosynthesis is more active in DH and skeletal muscle development in TC is faster compared to DH.

The heterogeneous FAPs were subdivided into twelve subclusters and further identified as myofibroblasts, FAPs, differentiating adipocytes, tenocytes, and adipocytes. Ordering of cells in trajectory analysis arranged the five cell subpopulations of two breeds into one major trajectory. Myofibroblasts were located toward the major branch in DH, but myofibroblasts in TC were located toward or at the origin of the major branch. Myofibroblasts are key contributors to fibrosis of skeletal muscle. Some studies showed that myofibroblasts could interact with satellite cells and regulate their proliferation and differentiation [49–51]. The premature depletion of myofibroblasts led to rapid exhaustion and early differentiation of the satellite cells [52]. Consistent with this conclusion, there were fewer myofibroblasts and satellite cells at the later pseudotime in TC compared to DH. We found that differentiating adipocytes and tenocytes appeared earlier in TC compared to DH, and more adipocytes in TC. Skeletal muscle development of embryos involves myogenesis, adipogenesis, and fibrogenesis [53], all of which are derived mainly from mesenchymal stem cells. The commitment of mesenchymal stem cells to myogenic, adipogenic, or fibrogenic lineages can be considered a competitive process [54]. We speculate that the distinguishing competitive process of myogenesis, adipogenesis, and fibrogenesis in two breeds contribute to the different developmental trajectories and ultimately lead to phenotypic differences. We also investigated the dynamic changes of gene expression in different FAPs by pseudotime and gene-switch analysis. *FMOD,TMSB15B* and *STMN1* were found inactivated earlier in TC than DH. GO enrichment showed that *FMOD* was related to the extracellular matrix, and *TMSB15B* was related to the sequestering of actin monomers, actin monomer binding, and actin filament organization. As differentiating adipocyte markers, *APOA1* and *MGP* were only found activated in TC. The dynamic changes in gene expression revealed that myogenesis and adipogenesis in TC are faster compared to DH.

In the process of skeletal muscle development, extremely complex biological activities occur between muscle tissue cells, which require accurate and efficient intercellular communication [55]. To explore the differences in ligand-receptor interaction networks underlying distinct muscle phenotypes, the ligand-receptor interactions were analyzed in three developmental stages of two chicken breeds. Among the 42 obtained signaling pathways, COLLAGEN and LAMININ pathways which accounting for a large proportion were highly abundant in the two breeds. These two pathways belong to the same intercellular interaction type known as ECM-receptor. Ligand-receptor pairs in both COLLAGEN and LAMININ pathways were also extensively detected in bovine skeletal muscle and pig skin [56, 57]. These results suggested that COLLAGEN and LAMININ pathways play important roles in multiple complex tissues of organisms, which indicated the importance of ECM in skeletal muscle development. CHEMERIN, also known as RARRES2 and ADIPONECTIN are two important adipokines that regulate adipogenesis and adipocyte lipid metabolism [58–60]. Interestingly, these two pathways were detected only in TC, suggesting that adipogenesis in TC is faster than in DH. The VISFATIN, PROS, and TGFb pathways were detected only in DH mainly involved in the interactions between FAPs, myofibroblasts, and 3 myoblast subtypes, and the number of interactions between these cells was higher in DH than TC. These findings suggested that communication between muscle cells was more active during myogenesis in DH.

Conclusions

In summary, we build a single-cell atlas of the chicken skeletal muscle, which is composed of 6 distinct cell types and reveal their developmental trajectories, gene expression status, and ligand-receptor interaction during embryonic skeletal muscle development. Our findings indicate that the skeletal muscle cell population showed different dynamic changes in the two chicken breeds. TC showed earlier embryonic myogenesis and less myoblasts compared with DH, which may underpin the distinct muscle phenotypes of the two chicken breeds. This study also detected some functional genes in the molecular mechanisms of skeletal muscle development between TC and DH. Our study provides new insights into the different genetic mechanisms between cultivated and native chicken breeds at single-cell level.

Abbreviations

DH	Daheng broiler
TC	Tibetan chicken
FAPs	Fibro-adipogenic progenitors
scRNA-seq	Single-cell RNA sequencing
tSNE	t-distributed stochastic neighbor embedding
PCA	Principal component analysis
MNN	Mutual nearest neighbors
RT	Room temperature
MST	Minimum spanning tree
GO	Gene ontology
UMAP	Uniform manifold approximation and projection
TF	Transcript factors
SNRP	Small nuclear ribonucleoprotein

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11363-w.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Supplementary Material 9

Supplementary Material 10

Acknowledgements
We thank all members of this work for their advice and technical assistance.

Author contributions
JL and ZXL conceived and designed the research. CWC performed the experiments. JL, DMY, JYW, and ZW collected the samples and analyzed the data. CWY and CLY provided resources. JL and DMY wrote the manuscript. All the authors have read and approved the final manuscript.

Funding
This work was supported by the Science and Technology Support Program of Sichuan Province (2021YFYZ0031), the Natural Science Foundation of Sichuan Province (2023NSFSC1142), and the Project of Qinghai-Tibetan Plateau Research in Southwest Minzu University (2024CXTD05).

Data availability
All data generated or analyzed during this study are included in this published article and its Additional files. The datasets generated and/or analyzed during the current study are available in the [NCBI gene expression omnibus database]. Accession GEO number is GSE251682. BioProject number is PRJNA1053748.

Declarations

Ethics approval and consent to participate
The present study was were approved by the Institutional Animal Care and Use Committee of Southwest Minzu University (2020MDLS034). All methods and experiments were performed in accordance with the relevant guidelines and regulations. The study involved only farm owned fertilized eggs, and informed consent was obtained from the chicken farm for their participation in this study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 18 October 2024 / Accepted: 13 February 2025
Published online: 24 February 2025

References

1. Goodpaster BH, Sparks LM. Metabolic flexibility in Health and Disease. Cell Metab. 2017;25(5):1027–36.
2. Schroder EA, Esser KA. Circadian rhythms, skeletal muscle molecular clocks, and exercise. Exerc Sport Sci Rev. 2013;41(4):224–9.
3. Bentzinger C, Wang Y, Rudnicki M. Building muscle: molecular regulation of myogenesis. CSH PERSPECT BIOL. 2012;4(2):a008342.
4. Biressi S, Molinaro M, Cossu G. Cellular heterogeneity during vertebrate skeletal muscle development. Dev Biol. 2007;308(2):281–93.
5. Gros J, Manceau M, Thomé V, Marcelle C. A common somitic origin for embryonic muscle progenitors and satellite cells. Nature. 2005;435(7044):954–8.
6. Parker MH, Seale P, Rudnicki MA. Looking back to the embryo: defining transcriptional networks in adult myogenesis. Nat Rev Genet. 2003;4(7):497–507.
7. Shinji S, Umezawa K, Nihashi Y, Nakamura S, Shimosato T, Takaya T. Identification of the myogenetic oligodeoxynucleotides (myoDNs) that promote differentiation of skeletal muscle myoblasts by targeting nucleolin. Front Cell Dev Biol. 2020;8:616706.
8. Buckingham M, Rigby PW. Gene regulatory networks and transcriptional mechanisms that control myogenesis. Dev Cell. 2014;28(3):225–38.
9. Delfini MC, Hirsinger E, Pourquié O, Duprez D. Delta 1-activated notch inhibits muscle differentiation without affecting Myf5 and Pax3 expression in chick limb myogenesis. Development. 2000;127(23):5213–24.
10. Scaal M, Marcelle C. Chick muscle development. Int J Dev Biol. 2018;62(1–2–3):127–36.
11. Buckingham M, Bajard L, Chang T, Daubas P, Hadchouel J, Meilhac S, Montarras D, Rocancourt D, Relaix F. The formation of skeletal muscle: from somite to limb. J Anat. 2003;202(1):59–68.
12. Hirst CE, Marcelle C. The avian embryo as a model system for skeletal myogenesis. Results Probl Cell Differ. 2015;56:99–122.
13. Velleman SG. Muscle development in the embryo and hatchling. Poult Sci. 2007;86(5):1050–4.
14. Du M, Tong J, Zhao J, Underwood K, Zhu M, Ford S, Nathanielsz P. Fetal programming of skeletal muscle development in ruminant animals. J Anim Sci. 2010;88:E51–60.
15. Bandman E, Rosser BW. Evolutionary significance of myosin heavy chain heterogeneity in birds. Microsc Res Tech. 2000;50(6):473–91.
16. Stockdale FE, Miller JB. The cellular basis of myosin heavy chain isoform expression during development of avian skeletal muscles. Dev Biol. 1987;123(1):1–9.
17. Petracci M, Cavani C. Muscle growth and poultry meat quality issues. Nutrients. 2012;4(1):1–12.
18. Dransfield E, Sosnicki AA. Relationship between muscle growth and poultry meat quality. Poult Sci. 1999;78(5):743–6.
19. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. Nature. 2019;566(7745):496–502.
20. Sachs S, Bastidas-Ponce A, Tritschler S, Bakhti M, Böttcher A, Sánchez-Garrido MA, Tarquis-Medina M, Kleinert M, Fischer K, Jall S, et al. Targeted pharmacological therapy restores β-cell function for diabetes remission. Nat Metab. 2020;2(2):192–209.
21. Clinton M, Haines L, Belloir B, McBride D. Sexing chick embryos: a rapid and simple protocol. Br Poult Sci. 2001;42(1):134–8.
22. Macosko E, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas A, Kamitaki N, Martersteck E, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015;161(5):1202–14.
23. 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.
24. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol. 2014;32(4):381–6.
25. Cao E, Ouyang J, Rackham O. GeneSwitches: ordering gene expression and functional events in single-cell experiments. Bioinformatics. 2020;36(10):3273–5.
26. Jin S, Guerrero-Juarez C, Zhang L, Chang I, Ramos R, Kuan C, Myung P, Plikus M, Nie Q. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021;12(1):1088.
27. 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.
28. Feregrino C, Sacher F, Parnas O, Tschopp P. A single-cell transcriptomic atlas of the developing chicken limb. BMC Genomics. 2019;20(1):401.
29. Giordani L, He G, Negroni E, Sakai H, Law J, Siu M, Wan R, Corneau A, Tajbakhsh S, Cheung T, et al. High-dimensional single-cell cartography reveals novel skeletal muscle-resident cell populations. Mol Cell. 2019;74(3):609–21.

30. Estermann M, Williams S, Hirst C, Roly Z, Serralbo O, Adhikari D, Powell D, Major A, Smith C. Insights into gonadal sex differentiation provided by single-cell transcriptomics in the chicken embryo. Cell Rep. 2020;31(1):107491.
31. Ho MH, Heydarkhan S, Vernet D, Kovanecz I, Ferrini MG, Bhatia NN, Gonzalez-Cadavid NF. Stimulating vaginal repair in rats through skeletal muscle-derived stem cells seeded on small intestinal submucosal scaffolds. Obstet Gynecol. 2009;114(2 Pt 1):300–9.
32. Seale P, Sabourin LA, Girgis-Gabardo A, Mansouri A, Gruss P, Rudnicki MA. Pax7 is required for the specification of myogenic satellite cells. Cell. 2000;102(6):777–86.
33. Lin Y-L, Nhieu J, Liu P-Y, Le G, Lee DJ, Wei C-W, Lin Y-W, Oh S-H, Lowe D, Wei L-N. CRABP1-CaMKII-Agrn regulates the maintenance of neuromuscular junction in spinal motor neuron. Cell Death Differ. 2022;29(9):1744–56.
34. Machado-Neto JA, Saad STO, Traina F. Stathmin 1 in normal and malignant hematopoiesis. BMB Rep. 2014;47(12):660–5.
35. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan C-H, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021;12(1):1088.
36. Krauss R, Joseph G, Goel A. Keep your friends close: cell-cell contact and skeletal myogenesis. Cold Spring Harb Perspect Biol. 2017;9(2):a029298.
37. Guo R, You X, Meng K, Sha R, Wang Z, Yuan N, Peng Q, Li Z, Xie Z, Chen R, et al. Single-cell RNA sequencing reveals heterogeneity of Myf5-derived cells and altered myogenic fate in the absence of SRSF2. Adv Sci. 2022;9(18):e2105775.
38. Dumont N, Bentzinger C, Sincennes M, Rudnicki M. Satellite cells and skeletal muscle regeneration. Compr Physiol. 2015;5(3):1027–59.
39. Kassar-Duchossoy L, Giacone E, Gayraud-Morel B, Jory A, Gomès D, Tajbakhsh S. Pax3/Pax7 mark a novel population of primitive myogenic cells during development. Gene Dev. 2005;19(12):1426–31.
40. Montarras D, Morgan J, Collins C, Relaix F, Zaffran S, Cumano A, Partridge T, Buckingham M. Direct isolation of satellite cells for skeletal muscle regeneration. Science. 2005;309(5743):2064–7.
41. Kuang S, Kuroda K, Le Grand F, Rudnicki M. Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell. 2007;129(5):999–1010.
42. Buckingham M, Bajard L, Daubas P, Esner M, Lagha M, Relaix F, Rocancourt D. Myogenic progenitor cells in the mouse embryo are marked by the expression of Pax3/7 genes that regulate their survival and myogenic potential. Anat Embryol (Berl). 2006;211(Suppl 1):51–6.
43. Hasty P, Bradley A, Morris JH, Edmondson DG, Venuti JM, Olson EN, Klein WH. Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. Nature. 1993;364(6437):501–6.
44. Cai S, Hu B, Zhu Q, Duo T, Wang X, Tong X, Luo X, Yuan R, Chen Y, Wang J, et al. Notch signaling leads to a slower progression of embryonic myogenic differentiation in Landrace than in Langtang pigs. Acta Biochim Biophys Sin. 2022;54(8):1122–32.
45. Zhao Y, Li J, Liu H, Xi Y, Xue M, Liu W, Zhuang Z, Lei M. Dynamic transcriptome profiles of skeletal muscle tissue across 11 developmental stages for both Tongcheng and Yorkshire pigs. BMC Genomics. 2015;16(1):377.
46. Yang Y, Yan J, Fan X, Chen J, Wang Z, Liu X, Yi G, Liu Y, Niu Y, Zhang L, et al. The genome variation and developmental transcriptome maps reveal genetic differentiation of skeletal muscle in pigs. PLoS Genet. 2021;17(11):e1009910.
47. Will CL, Luhrmann R. Spliceosome structure and function. CSH PERSPECT BIOL. 2010;3(7):a003707–003707.
48. Lan Y, Lou J, Hu J, Yu Z, Lyu W, Zhang B. Downregulation of SNRPG induces cell cycle arrest and sensitizes human glioblastoma cells to temozolomide by targeting Myc through a p53-dependent signaling pathway. Cancer Biol Med. 2020;17(1):112–31.
49. Thooyamani A, Mukhopadhyay A. PDGFRα mediated survival of myofibroblasts inhibit satellite cell proliferation during aberrant regeneration of lacerated skeletal muscle. Sci Rep. 2021;11(1):63.
50. Klingler W, Jurkat-Rott K, Lehmann-Horn F, Schleip R. The role of fibrosis in Duchenne muscular dystrophy. Acta Myol. 2012;31(3):184–95.
51. Zhou L, Lu H. Targeting fibrosis in Duchenne muscular dystrophy. J Neuropathol Exp Neurol. 2010;69(8):771–6.
52. Mathew SJ, Hansen JM, Merrell AJ, Murphy MM, Lawson JA, Hutcheson DA, Hansen MS, Angus-Hill M, Kardon G. Connective tissue fibroblasts and Tcf4 regulate myogenesis. Development. 2011;138(2):371–84.
53. Du M, Yan X, Tong J, Zhao J, Zhu M. Maternal obesity, inflammation, and fetal skeletal muscle development. Biol Reprod. 2010;82(1):4–12.
54. Yan X, Zhu M-J, Dodson MV, Du M. Developmental programming of fetal skeletal muscle and adipose tissue development. J Genomics. 2013;1:29–38.
55. Ellison D, Mugler A, Brennan M, Lee S, Huebner R, Shamir E, Woo L, Kim J, Amar P, Nemenman I, et al. Cell-cell communication enhances the capacity of cell ensembles to sense shallow gradients during morphogenesis. Proc Natl Acad Sci U S A. 2016;113(6):E679–688.
56. Wang L, Gao P, Li C, Liu Q, Yao Z, Li Y, Zhang X, Sun J, Simintiras C, Welborn M, et al. A single-cell atlas of bovine skeletal muscle reveals mechanisms regulating intramuscular adipogenesis and fibrogenesis. J Cachexia Sarcopenia Muscle. 2023;14(5):2152–67.
57. Zou Q, Yuan R, Zhang Y, Wang Y, Zheng T, Shi R, Zhang M, Li Y, Fei K, Feng R, et al. A single-cell transcriptome atlas of pig skin characterizes anatomical positional heterogeneity. eLife. 2023;12:e86504.
58. Li Y, Sun F, Li C, Mo H, Zhou Y, Lv D, Zhai J, Qian H, Ma F. RARRES2 regulates lipid metabolic reprogramming to mediate the development of brain metastasis in triple negative breast cancer. Mil Med Res. 2023;10(1):34.
59. Nguyen T. Adiponectin: role in physiology and pathophysiology. Int J Prev Med. 2020;11(1):136.
60. Goralski K, McCarthy T, Hanniman E, Zabel B, Butcher E, Parlee S, Muruganandan S, Sinal C. Chemerin, a novel adipokine that regulates adipogenesis and adipocyte metabolism. J Biol Chem. 2007;282(38):28175–88.

Publisher's note