



## *Dcx* expression defines a subpopulation of *Gdf5*<sup>+</sup> cells with chondrogenic potentials in E12.5 mouse embryonic limbs

Ruoxin Lan<sup>a</sup>, Dongxia Ge<sup>b,c</sup>, Yao-Zhong Liu<sup>a,\*\*</sup>, Zongbing You<sup>b,c,d,e,f,g,\*</sup>

<sup>a</sup> Department of Biostatistics and Data Science, School of Public Health and Tropic Medicine, Tulane University, New Orleans, LA, 70112, USA

<sup>b</sup> Department of Structural & Cellular Biology, School of Medicine, Tulane University, New Orleans, LA, 70112, USA

<sup>c</sup> Department of Orthopaedic Surgery, School of Medicine, Tulane University, New Orleans, LA, 70112, USA

<sup>d</sup> Department of Research Service, Southeast Louisiana Veterans Health Care System, New Orleans, LA, 70119, USA

<sup>e</sup> Tulane Cancer Center and Louisiana Cancer Research Consortium, Tulane University, New Orleans, LA, 70112, USA

<sup>f</sup> Tulane Center for Stem Cell Research and Regenerative Medicine, Tulane University, New Orleans, LA, 70112, USA

<sup>g</sup> Tulane Center for Aging, Tulane University, New Orleans, LA, 70112, USA

### ARTICLE INFO

#### Keywords:

Doublecortin

*Gdf5*

Joint interzone

Articular chondrocyte

Joint development

Synovial joint

### ABSTRACT

Growth differentiation factor 5 (*Gdf5*) and doublecortin (*Dcx*) genes are both expressed in joint interzone cells during synovial joint development. In this study, we re-analyzed the single cell RNA-sequencing data (Gene Expression Omnibus GSE151985) generated from *Gdf5*<sup>+</sup> cells of mouse knee joints at embryonic stages of E12.5, E13.5, E14.5, and E15.5, with a new focus on *Dcx*. We found that *Dcx* expression was enriched in clusters of *Gdf5*<sup>+</sup> cells, with high expression levels of pro-chondrogenic genes including sex determining region Y-box transcription factor 5 (*Sox5*), *Sox6*, *Sox9*, *Gdf5*, versican, matrilin 4, collagen type II  $\alpha 1$  chain (*Col2a1*), *Col9a1*, *Col9a2*, and *Col9a3* at E12.5. *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells had differential gene expression profiles. The up-regulated genes in *Dcx*<sup>+</sup> vs. *Dcx*<sup>-</sup> cells at E12.5 and E13.5 were enriched in chondrocyte differentiation and cartilage development, whereas those genes up-regulated at E14.5 and E15.5 were enriched in RNA splicing, protein stability, cell proliferation, and cell growth. Gene expression profiles in *Dcx*<sup>+</sup> cells showed rapid daily changes from E12.5 to E15.5, with limited number of genes shared across the time period. Expression of *Gdf5*, *Sox5*, *Sox6*, melanoma inhibitory activity, noggin, odd-skipped related transcription factor 2, matrilin 4, and versican was positively correlated with *Dcx* expression. Our results demonstrate that *Dcx* expression defines a subpopulation of *Gdf5*<sup>+</sup> cells with chondrogenic potentials in E12.5 mouse embryonic limbs.

### 1. Introduction

It has been well established that embryonic articular joint development starts from appearance of joint interzone – a densely packed region of flattened cells in the cartilaginous anlage [1]. The interzone contains a central intermediate layer and two chondrogenous layers that eventually differentiate into articular chondrocytes. In mouse embryos, forelimb buds appear at embryonic stage of 9.5 days postcoitus (i.e., E9.5) and mesenchymal condensation occurs at E12.5 [2]. Interzones appear at the presumptive sites of articular joints at E13.5 [3]. At E15.5, joint cavitation appears, marking formation of a synovial joint [3]. Growth differentiation factor 5 (*Gdf5*) is first expressed in the forelimb

at E11.5 at the presumptive regions of the shoulder and elbow and its expression decreases at E15.5 [4]. *Gdf5*-Cre reporter mice demonstrate that *Gdf5*<sup>+</sup> interzone cells differentiate into articular chondrocytes, synovial lining, and intrajoint ligaments [5–8]. Bian et al. recently isolated single *Gdf5*<sup>+</sup> cells from mouse knee joints at E12.5, E13.5, E14.5, and E15.5, and performed single cell RNA-sequencing (scRNA-seq) analysis. They found that *Gdf5*-lineage enriched (GLE) cells contain three super-clusters (SC): SC1 comprises a mixture of progenitor cells of mesenchymal character and chondroprogenitors; SC2 contains a mixture of interzone cells and transient chondrocytes; and SC3 largely consists of fibroblast-related cells. Their work further demonstrates the heterogeneity of *Gdf5*<sup>+</sup> cells at the transcriptional level [9]. How the clusters of *Gdf5*<sup>+</sup> cells commit to different lineages remains to be

\* Corresponding author. Department of Structural & Cellular Biology, Tulane University School of Medicine, Southeast Louisiana Veterans Health Care System, 2400 Canal St, 1430 Tulane Avenue, New Orleans, LA, 70112, USA.,

\*\* Corresponding author. Department of Biostatistics and Data Science, Tulane University School of Public Health and Tropic Medicine, 1440 Canal Street, New Orleans, LA, 70112, USA.

E-mail addresses: [rlan@tulane.edu](mailto:rlan@tulane.edu) (R. Lan), [dge@tulane.edu](mailto:dge@tulane.edu) (D. Ge), [yliu8@tulane.edu](mailto:yliu8@tulane.edu) (Y.-Z. Liu), [zyou@tulane.edu](mailto:zyou@tulane.edu), [Zongbing.You@va.gov](mailto:Zongbing.You@va.gov) (Z. You).

<https://doi.org/10.1016/j.bbrep.2022.101200>

Received 25 August 2021; Received in revised form 14 December 2021; Accepted 3 January 2022

2405-5808/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Abbreviations:**

ANOVA	analysis of variance
<i>Col</i>	collagen
<i>Dcx</i>	doublecortin
<i>Gdf5</i>	growth differentiation factor 5
GSEA	Gene Set Enrichment Analysis
<i>Kdr</i>	kinase insert domain receptor
<i>Matn</i>	matrilin
<i>Mia</i>	melanoma inhibitory activity
<i>Nog</i>	noggin
<i>Osr2</i>	odd-skipped related transcription factor 2
<i>Pdgfra</i>	platelet derived growth factor receptor $\alpha$
SC	super-cluster
scRNA-seq	single cell RNA-sequencing
<i>Sox</i>	Sex Determining Region Y-Box transcription factor
tSNE, t	-Distributed Stochastic Neighbor Embedding
<i>Vcan</i>	versican
<i>Wnt</i>	wingless-type MMTV integration site family member

determined.

Doublecortin (*Dcx*) is originally identified in migrating and differentiating neurons [10]. In mouse limbs, *Dcx* is initially expressed in the common mesenchymal precursors at E9.5 through E12.5. *Dcx* expression is sequestered in the joint interzones at E13.5 and then in immature articular chondrocytes, but *Dcx* is not expressed in the synovium, ligaments/tendons, and menisci [11,12]. Thus, *Dcx* appears to be a unique marker of articular chondrocyte lineage [7,13–19]. It is hypothesized that *Dcx*<sup>+</sup> cells are a subpopulation of *Gdf5*<sup>+</sup> cells in the joint interzones. The present study is to test this hypothesis through re-analyzing the scRNA-seq data publicly deposited by Bian et al. [9].

## 2. Materials and methods

**Data source and identification of the target cells.** Raw data were generated using scRNA-seq analysis of *Gdf5*<sup>+</sup> cells isolated from mouse (FVB/NJ crossing with C57BL/6J background) embryonic knee joints at E12.5, E13.5, E14.5, and E15.5, and were deposited to Gene Expression Omnibus under accession number GSE151985 by Dr. Patrick Cahan's lab at Johns Hopkins School of Medicine; *Gdf5*<sup>+</sup> cells expressed enhanced yellow fluorescent protein and were isolated using fluorescence-activated cell sorting (FACS) (see details as described previously [9]). R (version 4.0.5, R Foundation for Statistical Computing, Vienna, Austria)-based Seurat (version 4.0.2) [20] was used to analyze the raw data. We excluded potential cell doublets through removing the cells of the top 5% of total count and excluded potential low-quality libraries that had fewer than 500 genes or >5% of mitochondrially encoded genes. SingleR (version 1.4.1) [21] was used to annotate the scRNA-seq data into 18 main cell types, such as fibroblasts, endothelial cells, macrophages, etc, based on mouse cell type reference database [22]. Fibroblasts were included while the other cell types were excluded from further analysis. Seurat's t-Distributed Stochastic Neighbor Embedding (tSNE) function was used to cluster the fibroblasts. Like in the previous study [9], we excluded the clusters with >20% cells expressing myogenic differentiation 1 (*Myod1*, known to be expressed mainly in myoblasts), premelanosome protein (*Pmel*, known to be expressed mainly in melanocytes), sex determining region Y-box transcription factor 10 (*Sox10*, known to be expressed mainly in neural crest cells), and/or twist basic helix-loop-helix transcription factor 2 (*Twist2*) and iroquois homeobox 1 (*Irx1*) (both genes were known to be expressed mainly in dermis cells). We eventually included a total of 8514 cells in our analysis (Supplementary Table S1), which was 1185 cells more than what were used for downstream analyses by Bian et al. [9]. We

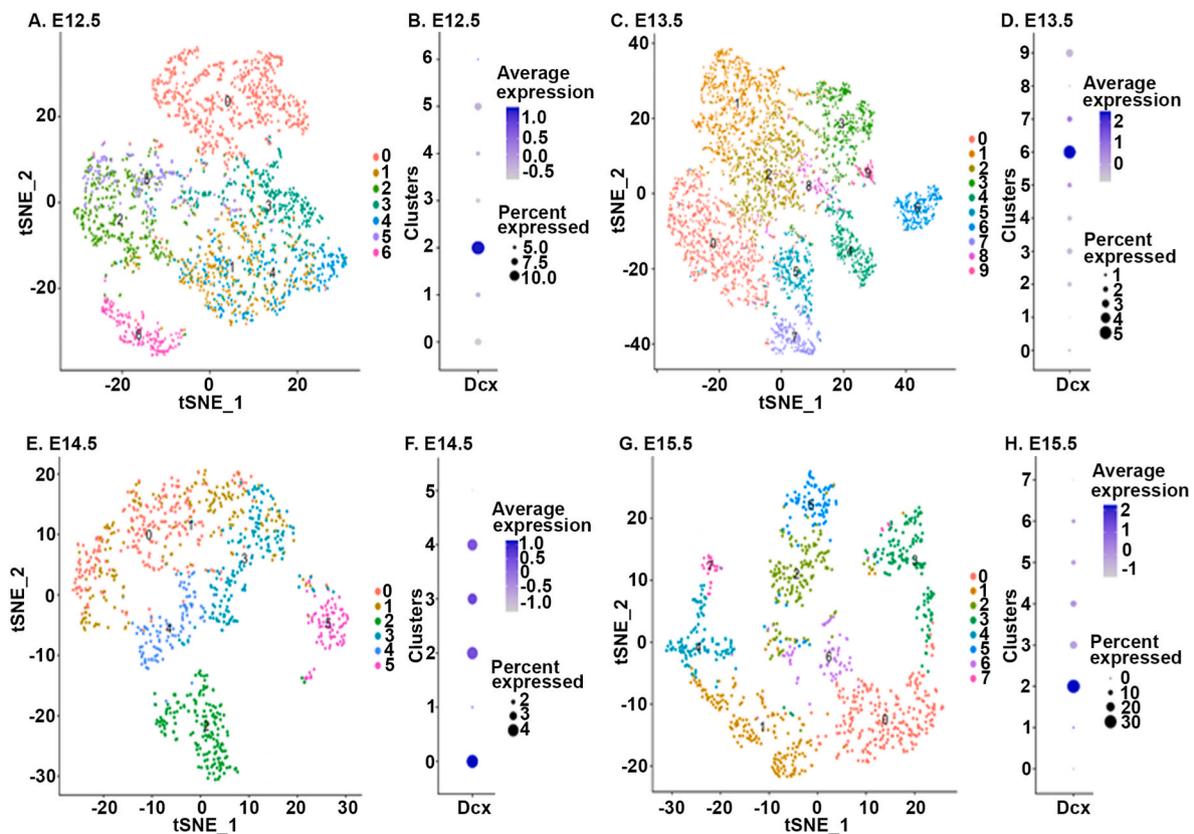
speculated that the difference might be due to the difference in algorithms of cell annotation used by Bian et al. [9] and SingleR [21] used here in our analysis.

**Analysis of the scRNA-seq data.** Software packages used to generate each figure are summarized in Supplementary Fig. S1. Briefly, Seurat was used to generate tSNE plots and visualize cell clusters, and its Dotplot function was used to visualize average expression levels of *Dcx* and percentages of the cells with *Dcx* expression. FindMarkers function of MAST package (version 1.16.0) [23] was used to identify 2000 highly variable genes, and their average expression was ranked from the highest (top) to the lowest (bottom) levels in the clusters with the highest *Dcx* expression. Seurat's DoHeatmap function was used to generate heatmaps in Fig. S2 and FeaturePlot function was used to separate *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells. FindMarkers function was used to identify the genes differentially expressed between *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells. Average log<sub>2</sub> (Fold Change, threshold -2.09 to 1.15) was calculated and Student's *t*-test was used to calculate *p* values (<0.05 was considered significant). The ggplot function of ggpubr (version 0.4.0 created by Alboukadel Kassambara) was used to generate volcano plots based on average log<sub>2</sub>(Fold Change) and -log<sub>10</sub>(*p* value). A set of significantly upregulated genes were analyzed in Gene Set Enrichment Analysis (GSEA) using enrichR package (version 3.0 created by Wajid Jawaid). In order to compare the gene expression levels across the four embryonic stages, we integrated the datasets of four embryonic stages using Seurat's integrateData function. Quality control procedures were performed as described above to exclude the low quality libraries and unwanted cell types, resulting in 9319 target cells. These cells were separated into *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> using Seurat's FeaturePlot function. We selected 139 genes that had been linked to joint development or shown in the previous analysis [9]. The genes were ranked based on their average expression levels in *Dcx*<sup>+</sup> cells at E12.5. The superheat package (version 1.0.0 created by Rebecca Barter and Bin Yu) was used to generate heatmaps in Fig. 3, Fig. 4A, Fig. S4, and Fig. S5, because Seurat's DoHeatmap function does not allow manual selection of individual genes. Eight genes, *i.e.*, *Gdf5*, *Sox5*, *Sox6*, melanoma inhibitory activity (*Mia*), noggin (*Nog*), odd-skipped related transcription factor 2 (*Osr2*), matrilin 4 (*Matn4*), and versican (*Vcan*), showed differential expression patterns similar to *Dcx*. *Matn4* is mainly expressed in articular chondrocytes [24], in contrast to *Matn1* that is mainly expressed in endochondral chondrocytes but not in articular chondrocytes [25]. Kendall's cor.test function in ggpubr package was used to perform correlation between each of the eight genes and *Dcx*.

## 3. Results and discussion

### 3.1. *Dcx* expression is enriched in a subpopulation of *Gdf5*<sup>+</sup> cells

In order to demonstrate the heterogeneity of *Gdf5*<sup>+</sup> cells, Seurat was used to analyze 2435 *Gdf5*<sup>+</sup> cells of the knee joints at E12.5. A tSNE plot showed that the cells were assigned into 7 clusters (Fig. 1A). Cluster #2 had the highest percentage of cells (12.4%) that expressed *Dcx* (Fig. 1B), the average expression of which was significantly higher than the other six clusters (*p* = 0.001 based on analysis of variance -ANOVA). This indicates that *Dcx* expression is enriched in Cluster #2 of *Gdf5*<sup>+</sup> cells at E12.5. At E13.5, tSNE plot showed that 4072 *Gdf5*<sup>+</sup> cells were assigned into 10 clusters (Fig. 1C). Cluster #6 had the highest percentage of cells (5.1%) that expressed *Dcx* (Fig. 1D), the average expression of which was significantly higher than the other nine clusters (*p* = 1.99e-07 based on ANOVA). This indicates that *Dcx* expression is enriched in Cluster #6 of *Gdf5*<sup>+</sup> cells at E13.5. At E14.5, tSNE plot showed that 965 *Gdf5*<sup>+</sup> cells were assigned into 6 clusters (Fig. 1E). Cluster #0 had the highest percentage of cells (4.2%) that expressed *Dcx* (Fig. 1F), but Cluster #2 (4.0%), #3 (3.6%), and #4 (3.8%) also had relatively higher percentages of *Dcx*<sup>+</sup> cells than Cluster #1 (1.6%) or #5 (1.1%). However, the average expression of *Dcx* was not statistically significant among the six clusters (*p* = 0.366 based on ANOVA). This indicates that *Dcx* expression



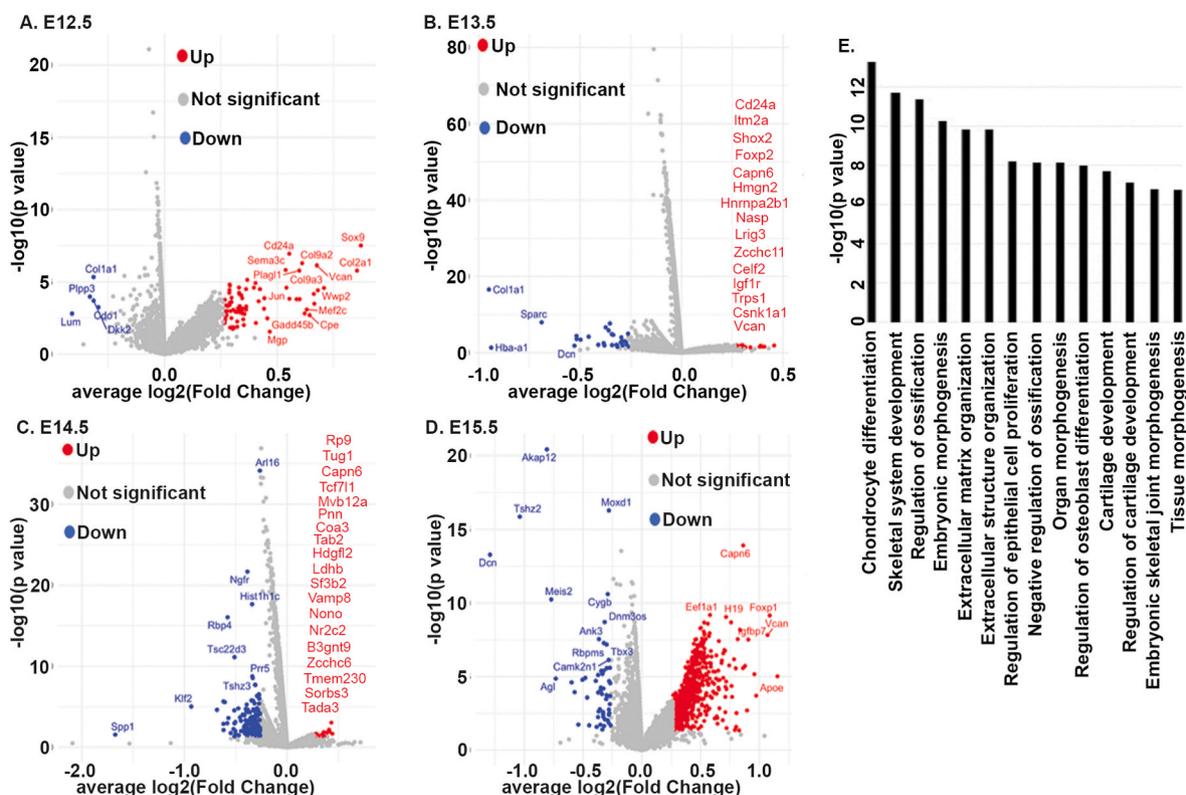
**Fig. 1.** *Dcx* expression is enriched in a subpopulation of *Gdf5*<sup>+</sup> cells. (A, C, E, and G) Cells were assigned to clusters using t-Distributed Stochastic Neighbor Embedding (tSNE) plots. (B, D, F, and H) Dotplots represent average expression levels of *Dcx* and percentages of cells with *Dcx* expression.

is enriched in multiple clusters of *Gdf5*<sup>+</sup> cells at E14.5. At E15.5, tSNE plot showed that 1042 *Gdf5*<sup>+</sup> cells were assigned into 8 clusters (Fig. 1G). Cluster #2 had the highest percentage of cells (30.4%) that expressed *Dcx* (Fig. 1H), the average expression of which was significantly higher than the other seven clusters ( $p = 2e-16$  based on ANOVA). This indicates that *Dcx* expression is enriched in Cluster #2 of *Gdf5*<sup>+</sup> cells at E15.5.

To demonstrate how the clusters with the highest *Dcx* expression were defined at transcriptional levels, we ranked levels of gene expression in the clusters and generated heatmaps to show the relative levels of those genes ranked at top 20 or bottom 20. The top ranked 20 genes in Cluster #2 at E12.5 included *Sox5*, *Sox6*, *Sox9*, *Gdf5*, *Vcan*, *Matn4*, collagen type II  $\alpha$  1 chain (*Col2a1*), *Col9a1*, *Col9a2*, and *Col9a3* that are known to be involved in chondrocyte differentiation (Supplementary Fig. S2A). In contrast, the bottom ranked 20 genes in Cluster #2 at E12.5 included *Col1a2* and *Col3a1* that are related to fibroblasts (Supplementary Fig. S2A). Although the top ranked 20 genes were highly expressed in Cluster #2, they were expressed at low levels in the other six clusters. In contrast, the bottom ranked 20 genes in Cluster #2 were expressed at relatively high levels in the rest clusters (Supplementary Fig. S2A). Since the expression of top 20 genes showed more dramatic difference compared to the rest clusters, we expanded the heatmap to include the top 50 genes in Cluster #2 and found a similar contrast between Cluster #2 and the rest clusters (Supplementary Fig. S2B). These findings suggest that Cluster #2 of *Gdf5*<sup>+</sup> cells at E12.5 had a distinct transcriptional profile that is pro-chondrogenic. We did similar analyses for E13.5 (Supplementary Figs. S2C–D), E14.5 (Supplementary Figs. S2E–F), and E15.5 (Supplementary Figs. S2G–H), and found that the clusters with the highest *Dcx* expression showed distinct transcriptional profiles compared to the other clusters.

### 3.2. *Dcx*<sup>+</sup> cells have distinct gene expression profiles compared to *Dcx*<sup>-</sup> cells

Since the clusters contained both *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells, we separated *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells and compared their transcriptional profiles. A volcano plot showed that many genes including *Sox9*, *Col2a1*, *Col9a2*, *Col9a3*, and *Vcan* were significantly upregulated in *Dcx*<sup>+</sup> cells compared to *Dcx*<sup>-</sup> cells at E12.5 (Fig. 2A,  $p < 0.05$ ). In contrast, some genes, including *Col1a1*, phospholipid phosphatase 3 (*Pllp3*), dickkopf WNT signaling pathway inhibitor 2 (*Dkk2*), Lumican (*Lum*), and cysteine dioxygenase type 1 (*Cdo1*), were significantly downregulated in *Dcx*<sup>+</sup> cells compared to *Dcx*<sup>-</sup> cells (Fig. 2A,  $p < 0.05$ ). Similarly, we found significantly upregulated and downregulated genes in *Dcx*<sup>+</sup> cells compared to *Dcx*<sup>-</sup> cells at E13.5 (Fig. 2B), E14.5 (Fig. 2C), and E15.5 (Fig. 2D). Of note, most of the significantly up- or down-regulated genes had about 1 fold change, which might be due to the fact that the cells were all *Gdf5*<sup>+</sup> cells, thus their differences in gene expression were subtle, rather than dramatically different. These findings suggest that *Dcx*<sup>+</sup> cells had distinct gene expression profiles compared to *Dcx*<sup>-</sup> cells. To explore the functions of the upregulated genes, we performed GSEA of the significantly upregulated genes at each embryonic stage. We found that at E12.5, the upregulated genes were enriched in chondrocyte differentiation, skeletal system development, embryonic morphogenesis, and cartilage development (Fig. 2E). These functions are highly related to chondrogenesis, suggesting that the upregulated genes in *Dcx*<sup>+</sup> cells are involved the development of synovial joints. GSEA for E13.5 showed that the upregulated genes were enriched in mesenchymal cell proliferation, chondrocyte differentiation, and cartilage development (Supplementary Fig. S3A). GSEA for E14.5 showed that the upregulated genes were enriched in RNA splicing, mRNA processing, and protein stability (Supplementary Fig. S3B). GSEA for E15.5 showed that the upregulated genes were enriched in cell growth,



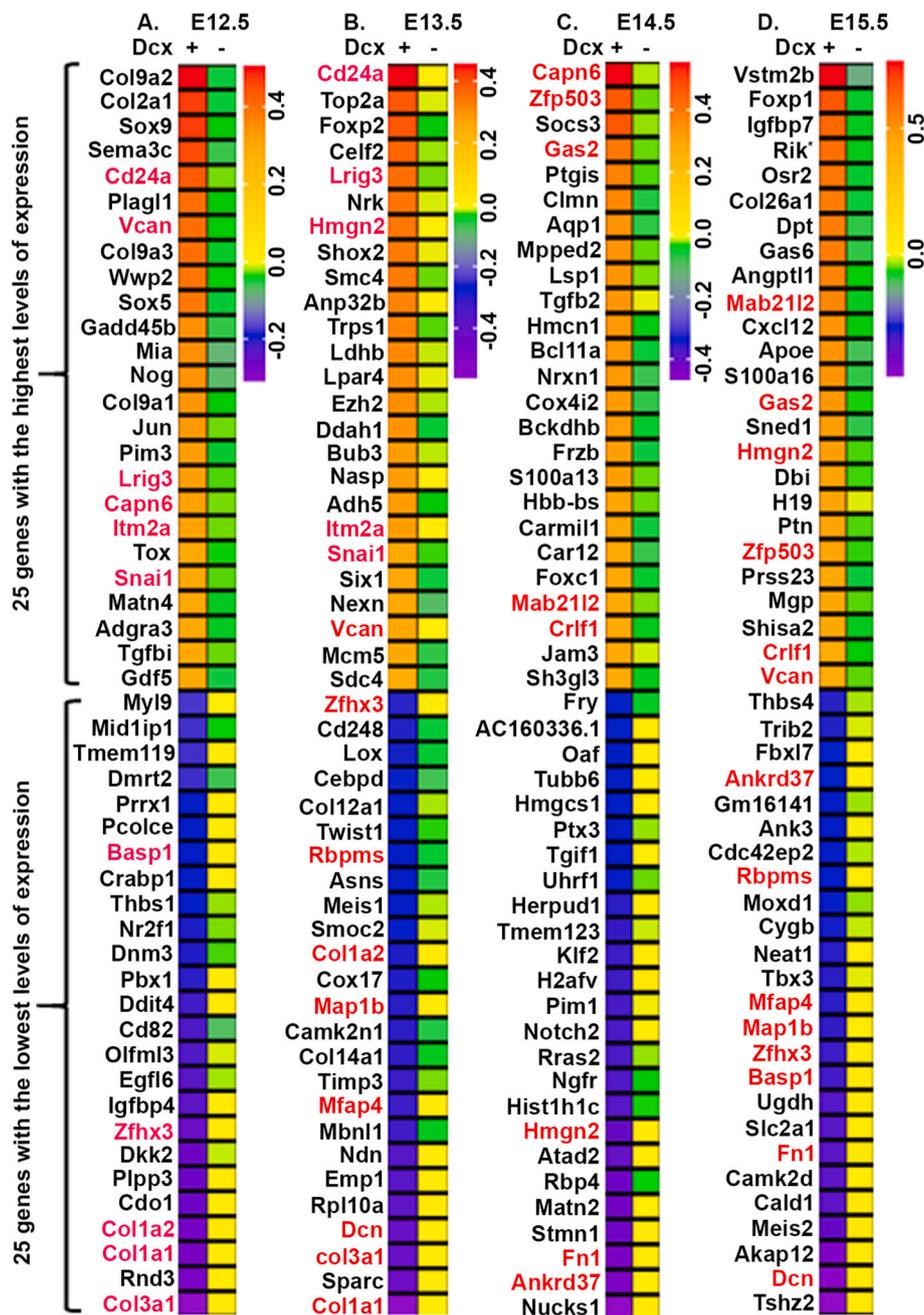
**Fig. 2.** *Dcx*<sup>+</sup> cells have distinct gene expression profiles compared to *Dcx*<sup>-</sup> cells. (A, B, C, and D) Volcano plots represent significantly up-regulated and down-regulated genes in *Dcx*<sup>+</sup> cells compared to *Dcx*<sup>-</sup> cells. Log<sub>2</sub> (Fold Change) > 0.263035 (*i.e.*, fold change >1.2) was considered up-regulated and log<sub>2</sub> (Fold Change) < -0.263035 (*i.e.*, fold change <0.83) was considered down-regulated. It was considered significant if -log<sub>10</sub> (p value) was >1.30103 (*i.e.*, p value < 0.05). (E) Gene Set Enrichment Analysis of significantly up-regulated genes in *Dcx*<sup>+</sup> cells at E12.5. The most significant functions based on gene ontology are shown.

epithelial cell proliferation, cell cycle arrest, and skeletal system development (Supplementary Fig. S3C). We noted that the upregulated genes at E12.5 and E13.5 were enriched mainly in chondrocyte differentiation and cartilage development, whereas the upregulated genes at E14.5 and E15.5 were enriched mainly in RNA splicing, protein stability, cell growth and proliferation. It is well known that E12.5 to E13.5 is the time when joint interzones emerge from the condensed mesenchyme. The joint interzones are destined to give rise to the joint, whereas the non-interzone mesenchymal cells become transient chondrocytes that eventually will be replaced by the bone. More importantly, the interzones maintain *Dcx* expression, whereas the non-interzone mesenchymal cells turn off *Dcx* expression at E13.5 [11]. If we assess our findings under the aforementioned context, we may reasonably speculate that the upregulated genes in *Dcx*<sup>+</sup> cells at E12.5 and E13.5 may play critical roles in determining the choice of cell fates, *i.e.*, to become articular chondrocyte lineage. In contrast, at E14.5 and E15.5, the cell fates have already been determined, thus the upregulated genes in *Dcx*<sup>+</sup> cells serve to maintain cell proliferation and cartilage growth.

### 3.3. Gene expression profiles in *Dcx*<sup>+</sup> cells present rapid changes from E12.5 to E15.5

To further investigate the gene expression profiles in *Dcx*<sup>+</sup> cells in comparison to *Dcx*<sup>-</sup> cells, we ranked the gene expression levels from the highest to the lowest in *Dcx*<sup>+</sup> cells at individual embryonic stages and generated heatmaps to compare *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells. We found that at E12.5, the top 25 genes in *Dcx*<sup>+</sup> cells included *Col9a2*, *Col2a1*, *Sox9*, *Vcan*, *Col9a3*, *Sox5*, *Col9a1*, and *Gdf5* (Fig. 3A), which were similarly highly expressed in Cluster #2 at E12.5 (Supplementary Fig. S2A). The bottom 25 genes in *Dcx*<sup>+</sup> cells included *Col1a2* and *Col3a1* (Fig. 3A), which were similarly weakly expressed in Cluster #2 at E12.5 (Supplementary Fig. S2A). The results suggest that the gene expression

profile in Cluster #2 at E12.5 is likely dominated by *Dcx*<sup>+</sup> cells in the cluster. The top 25 genes in *Dcx*<sup>+</sup> cells were weakly expressed in *Dcx*<sup>-</sup> cells, whereas the bottom 25 genes in *Dcx*<sup>+</sup> cells were expressed at higher levels in *Dcx*<sup>-</sup> cells (Fig. 3A). Of note, the levels of gene expression in *Dcx*<sup>-</sup> cells were not dramatically different (mostly in yellow to green color codes around “0” – the average level). This pattern of differential expression was also observed when comparing the top 50 genes and bottom 50 genes in *Dcx*<sup>+</sup> cells against *Dcx*<sup>-</sup> cells (Supplementary Figs. S4A–B). At E13.5, the top 25 and bottom 25 genes in *Dcx*<sup>+</sup> cells were largely different from E12.5 (Fig. 3B), except for a few genes, including CD24 molecule (*Cd24a*), leucine rich repeats and immunoglobulin like domains 3 (*Lrig3*), high mobility group nucleosomal binding domain 2 (*Hmgn2*), integral membrane protein 2A (*Itm2a*), snail family transcriptional repressor 1 (*Snai1*), *Vcan*, zinc finger homeobox 3 (*Zfhn3*), *Col1a2*, *Col3a1*, and *Col1a1* (Fig. 3B). The top 50 genes in *Dcx*<sup>+</sup> cells were weakly expressed in *Dcx*<sup>-</sup> cells, while the bottom 50 genes in *Dcx*<sup>+</sup> cells were expressed at higher levels in *Dcx*<sup>-</sup> cells (Supplementary Figs. S4C–D). At E14.5, the top 25 and bottom 25 genes in *Dcx*<sup>+</sup> cells were completely different from E13.5 (Fig. 3C). Only one of the top 25 genes, calpain 6 (*Capn6*), was shared with the top 25 genes at E12.5. The top 50 genes in *Dcx*<sup>+</sup> cells were weakly expressed in *Dcx*<sup>-</sup> cells, while the bottom 50 genes in *Dcx*<sup>+</sup> cells were expressed at higher levels in *Dcx*<sup>-</sup> cells (Supplementary Figs. S4E–F). At E15.5, the top 25 and bottom 25 genes were mostly different from E14.5 (Fig. 3D). A few genes including mab-21 like 2 (*Mab21l2*), growth arrest specific 2 (*Gas2*), cytokine receptor like factor 1 (*Crlf1*), ankyrin repeat domain 37 (*Ankrd37*), and fibronectin 1 (*Fn1*) were shared with those of E14.5. *Hmgn2*, *Vcan*, *Zfhn3*, brain abundant membrane attached signal protein 1 (*Basp1*), and decorin (*Dcn*) were shared with those of E12.5 and/or E13.5 (Fig. 3A–D). The top 50 genes in *Dcx*<sup>+</sup> cells were weakly expressed in *Dcx*<sup>-</sup> cells, while the bottom 50 genes in *Dcx*<sup>+</sup> cells were expressed at higher levels in *Dcx*<sup>-</sup> cells (Supplementary Figs. S4G–H). The lists of top and bottom 50 genes



**Fig. 3.** Gene expression profiles in *Dcx*<sup>+</sup> cells present rapid changes from E12.5 to E15.5. (A, B, C, and D) Heatmaps of top 25 and bottom 25 genes expressed in *Dcx*<sup>+</sup> cells. Genes shared across the four embryonic stages are highlighted in red font. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

with annotated functions are shown in [Supplementary Table S2](#). At E13.5, some genes highly expressed in *Dcx*<sup>+</sup> cells (Fig. 3B), e.g., *Itm2a*, *Top2a*, *Lpar4*, *Nexn*, and *Nrk*, were also highly expressed in cluster #6 with dominant *Dcx* expression (Fig. S2D). Similarly, at E15.5, *Foxp1*, *Lgfbp7*, *Col26a1*, *Dpt*, *Gas6*, *Angptl1*, *Cxcl12*, *Apoe*, *Gas2*, *Dbi*, *Ptn*, *Mgp*, and *Vcan* genes in Fig. 3D were shown in Cluster #2 with dominant *Dcx* expression (Fig. S2H). However, at E14.5, such concurrence was not found (comparing Fig. 3C and Fig. S2F), which might be due to the fact that *Dcx* expression at E14.5 was distributed in 4 clusters (Fig. 1F). These findings demonstrate that the gene expression profiles in *Dcx*<sup>+</sup> cells undergo rapid changes during the four-day period from E12.5 to E15.5, implying a quick differentiation of *Dcx*<sup>+</sup> cells.

#### 3.4. Expression of a set of genes is positively correlated with *Dcx* expression

To track the expression of genes from E12.5 to E15.5, we integrated the datasets and normalized the gene expression levels to the same scale across the four embryonic stages and generated a heatmap of 32 genes that have been studied in joint development [19]. We found that *Sox4*, *Sox9*, *Sox11*, *Jun*, paired like homeodomain 1 (*Pitx1*), *Col2a1*, *Vcan*, *Matn4*, and platelet derived growth factor receptor  $\alpha$  (*Pdgfra*) were highly expressed from E12.5 to E15.5 in both *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells, whereas scleraxis basic helix-loop-helix transcription factor (*Scx*), transforming growth factor beta receptor 2 (*Tgfb2*), Indian hedgehog (*Ihh*), leucine rich repeat containing G protein-coupled receptor 5 (*Lgr5*), wingless-type MMTV integration site family member 4 (*Wnt4*), *Wnt9a*,

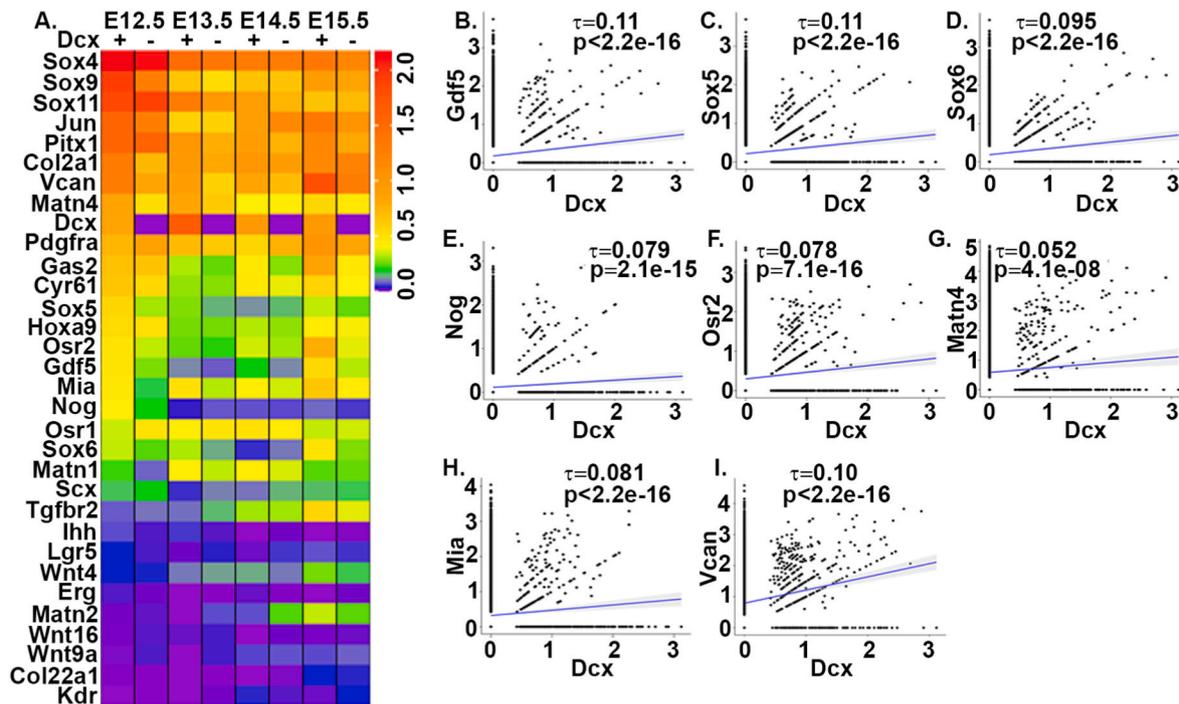


Fig. 4. Expression of a set of genes is positively correlated with *Dcx* expression. (A) Heatmap of 32 genes from the integrated datasets. (B to I) Kendall's correlation analyses.  $\tau$ , correlation coefficient.

*Wnt16*, erythroblast transformation-specific transcription factor *Erg* (*Erg*), *Col22a1*, and kinase insert domain receptor (*Kdr*) were weakly expressed in both *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells (Fig. 4A). *Wnt4* and *Wnt9a* expression was almost absent at E12.5 and slightly increased at E15.5, suggesting that *Wnt* signaling is not involved in interzone specification, but may be involved in joint maintenance. This finding is consistent with a previous study using *Wnt4*<sup>-/-</sup>*Wnt9a*<sup>-/-</sup> double mutant mice, which revealed that the mice developed joint fusion after joint interzone formation [26]. *Sox9*, *Jun*, and *Col2a1* were expressed at higher levels in *Dcx*<sup>+</sup> cells than *Dcx*<sup>-</sup> cells at E12.5, but this pattern was not kept at E13.5 to E15.5. On the other hand, *Tgfr2* and *Matn2* expression was increased in *Dcx*<sup>+</sup> cells only at E15.5 (Fig. 4A). *Matn1* levels were slightly higher in *Dcx*<sup>+</sup> cells than *Dcx*<sup>-</sup> cells at E13.5 to E14.5, but it was weakly expressed at E12.5 and E15.5. A set of eight genes including *Gdf5*, *Sox5*, *Sox6*, *Mia*, *Nog*, *Osr2*, *Matn4*, and *Vcan* were consistently expressed at higher levels in *Dcx*<sup>+</sup> cells than *Dcx*<sup>-</sup> cells from E12.5 to E15.5 (Fig. 4A). *Gdf5* expression varied from E12.5 to E15.5. Kendall's correlation analysis showed that expression of these eight genes was positively correlated with *Dcx* expression (Fig. 4B–I). Heatmaps of other 48 genes (Supplementary Fig. S5A) and 59 genes (Supplementary Fig. S5B) showed some genes that were highly or weakly expressed in both *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells. However, there was not any consistent pattern of differential expression between *Dcx*<sup>+</sup> and *Dcx*<sup>-</sup> cells across the four embryonic stages from E12.5 to E15.5. These 107 genes were analyzed in the previous study by Bian et al. [9].

In summary, we found that *Dcx* expression was enriched in clusters of *Gdf5*<sup>+</sup> cells expressing pro-chondrogenic genes. The up-regulated genes in *Dcx*<sup>+</sup> cells were enriched in chondrocyte differentiation and cartilage development at E12.5 and E13.5, whereas those genes up-regulated at E14.5 and E15.5 were enriched in RNA splicing, protein stability, cell proliferation, and cell growth. Gene expression profiles in *Dcx*<sup>+</sup> cells showed rapid daily changes from E12.5 to E15.5. Expression of *Gdf5*, *Sox5*, *Sox6*, *Mia*, *Nog*, *Osr2*, *Matn4*, and *Vcan* was positively correlated with *Dcx* expression from E12.5 to E15.5. These results demonstrate that *Dcx* expression defines a subpopulation of *Gdf5*<sup>+</sup> cells with chondrogenic potentials in E12.5 mouse embryonic limbs, which provides clues to further investigate *Dcx*'s roles in joint development.

## Declaration of interests

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

## Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbrep.2022.101200>.

## References

- [1] C.W. Archer, G.P. Dowthwaite, P. Francis-West, Development of synovial joints, *Birth Defects Res C Embryo Today* 69 (2003) 144–155, <https://doi.org/10.1002/bdrc.10015>.
- [2] P. Martin, *Tissue patterning in the developing mouse limb*, *Int. J. Dev. Biol.* 34 (1990) 323–336.
- [3] D. Mitrovic, Development of the diarthrodial joints in the rat embryo, *Am. J. Anat.* 151 (1978) 475–485, <https://doi.org/10.1002/aja.1001510403>.
- [4] E.E. Storm, D.M. Kingsley, Joint patterning defects caused by single and double mutations in members of the bone morphogenetic protein (BMP) family, *Development* 122 (1996) 3969–3979.
- [5] R.B. Rountree, M. Schoor, H. Chen, M.E. Marks, V. Harley, Y. Mishina, D. M. Kingsley, BMP receptor signaling is required for postnatal maintenance of articular cartilage, *PLoS Biol.* 2 (2004) e355, <https://doi.org/10.1371/journal.pbio.0020355>.
- [6] E. Koyama, Y. Shibukawa, M. Nagayama, H. Sugito, B. Young, T. Yuasa, T. Okabe, T. Ochiai, N. Kamiya, R.B. Rountree, D.M. Kingsley, M. Iwamoto, M. Enomoto-Iwamoto, M. Pacifici, A distinct cohort of progenitor cells participates in synovial joint and articular cartilage formation during mouse limb skeletogenesis, *Dev. Biol.* 316 (2008) 62–73, <https://doi.org/10.1016/j.ydbio.2008.01.012>.
- [7] R.S. Decker, H.B. Um, N.A. Dymont, N. Cottingham, Y. Usami, M. Enomoto-Iwamoto, M.S. Kronenberg, P. Maye, D.W. Rowe, E. Koyama, M. Pacifici, Cell origin, volume and arrangement are drivers of articular cartilage formation, morphogenesis and response to injury in mouse limbs, *Dev. Biol.* 426 (2017) 56–68, <https://doi.org/10.1016/j.ydbio.2017.04.006>.

- [8] Y. Shwartz, S. Viukov, S. Krief, E. Zelzer, Joint development involves a continuous influx of Gdf5-positive cells, *Cell Rep.* 15 (2016) 2577–2587, <https://doi.org/10.1016/j.celrep.2016.05.055>.
- [9] Q. Bian, Y.H. Cheng, J.P. Wilson, E.Y. Su, D.W. Kim, H. Wang, S. Yoo, S. Blackshaw, P. Cahan, A single cell transcriptional atlas of early synovial joint development, *Development* 147 (2020), <https://doi.org/10.1242/dev.185777>.
- [10] J.C. Corbo, T.A. Deuel, J.M. Long, P. LaPorte, E. Tsai, A. Wynshaw-Boris, C. A. Walsh, Doublecortin is required in mice for lamination of the hippocampus but not the neocortex, [doi.org/22/17/7548](https://doi.org/10.1523/JNEUROSCI.22-02.2002), *J. Neurosci.* 22 (2002) 7548–7557 [pii].
- [11] Q. Zhang, A.D. Cigan, L. Marrero, C. Loprore, S. Liu, D. Ge, F.H. Savoie, Z. You, Expression of doublecortin reveals articular chondrocyte lineage in mouse embryonic limbs, *Genesis* 49 (2011) 75–82, <https://doi.org/10.1002/dvg.20702>.
- [12] Y. Zhang, J.A. Ryan, P.E. Di Cesare, J. Liu, C.A. Walsh, Z. You, Doublecortin is expressed in articular chondrocytes, *Biochem. Biophys. Res. Commun.* 363 (2007) 694–700, <https://doi.org/10.1016/j.bbrc.2007.09.030>.
- [13] L. Yu, L.A. Dawson, M. Yan, K. Zimmel, Y.L. Lin, C.P. Dolan, M. Han, K. Muneoka, BMP9 stimulates joint regeneration at digit amputation wounds in mice, *Nat. Commun.* 10 (2019) 424, <https://doi.org/10.1038/s41467-018-08278-4>.
- [14] A.M. Craft, J.S. Rockel, Y. Nartiss, R.A. Kandel, B.A. Alman, G.M. Keller, Generation of articular chondrocytes from human pluripotent stem cells, *Nat. Biotechnol.* 33 (2015) 638–645, [//doi.org/10.1038/nbt.3210](https://doi.org/10.1038/nbt.3210).
- [15] A.M. Craft, N. Ahmed, J.S. Rockel, G.S. Baht, B.A. Alman, R.A. Kandel, A. E. Grigoriadis, G.M. Keller, Specification of chondrocytes and cartilage tissues from embryonic stem cells, *Development* 140 (2013) 2597–2610, <https://doi.org/10.1242/dev.087890>.
- [16] Z.H. Wen, C.C. Tang, Y.C. Chang, S.Y. Huang, Y.Y. Lin, S.P. Hsieh, H.P. Lee, S. C. Lin, W.F. Chen, Y.H. Jean, Calcitonin attenuates cartilage degeneration and nociception in an experimental rat model of osteoarthritis: role of TGF-beta in chondrocytes, *Sci. Rep.* 6 (2016) 28862, <https://doi.org/10.1038/srep28862>.
- [17] T. Yamagami, A. Molotkov, C.J. Zhou, Canonical Wnt signaling activity during synovial joint development, *J. Mol. Histol.* 40 (2009) 311–316, <https://doi.org/10.1007/s10735-009-9242-1>.
- [18] A.A. Pitsillides, F. Beier, Cartilage biology in osteoarthritis—lessons from developmental biology, *Nat. Rev. Rheumatol.* 7 (2011) 654–663, <https://doi.org/10.1038/nrrheum.2011.129>.
- [19] R.S. Decker, E. Koyama, M. Pacifici, Genesis and morphogenesis of limb synovial joints and articular cartilage, *Matrix Biol.* 39 (2014) 5–10, <https://doi.org/10.1016/j.matbio.2014.08.006>.
- [20] Y. Hao, S. Hao, E. Andersen-Nissen, W.M. Mauck 3rd, S. Zheng, A. Butler, M.J. Lee, A.J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E.P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L.M. Fleming, B. Yeung, A.J. Rogers, J. M. McElrath, C.A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data, *Cell* 184 (2021) 3573–3587, <https://doi.org/10.1016/j.cell.2021.04.048>, e3529.
- [21] D. Aran, A.P. Looney, L. Liu, E. Wu, V. Fong, A. Hsu, S. Chak, R.P. Naikawadi, P. J. Wolters, A.R. Abate, A.J. Butte, M. Bhattacharya, Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage, *Nat. Immunol.* 20 (2019) 163–172, <https://doi.org/10.1038/s41590-018-0276-y>.
- [22] B.A. Benayoun, E.A. Pollina, P.P. Singh, S. Mahmoudi, I. Harel, K.M. Casey, B. W. Dulken, A. Kundaje, A. Brunet, Remodeling of epigenome and transcriptome landscapes with aging in mice reveals widespread induction of inflammatory responses, *Genome Res.* 29 (2019) 697–709, <https://doi.org/10.1101/gr.240093.118>.
- [23] G. Finak, A. McDavid, M. Yajima, J. Deng, V. Gersuk, A.K. Shalek, C.K. Slichter, H. W. Miller, M.J. McElrath, M. Prlic, P.S. Linsley, R. Gottardo, MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data, *Genome Biol.* 16 (2015) 278, <https://doi.org/10.1186/s13059-015-0844-5>.
- [24] A.R. Klatt, D.P. Nitsche, B. Kobbe, M. Macht, M. Paulsson, R. Wagener, Molecular structure, processing, and tissue distribution of matrilin-4, *J. Biol. Chem.* 276 (2001) 17267–17275, <https://doi.org/10.1074/jbc.M100587200>.
- [25] G. Hyde, S. Dover, A. Aszodi, G.A. Wallis, R.P. Boot-Handford, Lineage tracing using matrilin-1 gene expression reveals that articular chondrocytes exist as the joint interzone forms, *Dev. Biol.* 304 (2007) 825–833, <https://doi.org/10.1016/j.ydbio.2007.01.026>.
- [26] D. Spater, T.P. Hill, R.J. O'Sullivan, M. Gruber, D.A. Conner, C. Hartmann, Wnt9a signaling is required for joint integrity and regulation of Ihh during chondrogenesis, *Development* 133 (2006) 3039–3049, <https://doi.org/10.1242/dev.02471>.