



HHS Public Access

Author manuscript

Matrix Biol. Author manuscript; available in PMC 2021 July 01.

Published in final edited form as:

Matrix Biol. 2020 July ; 89: 1–10. doi:10.1016/j.matbio.2019.12.004.

Single cell RNA-sequencing reveals cellular heterogeneity and trajectories of lineage specification during murine embryonic limb development

Natalie H. Kelly^{a,b,c}, Nguyen P.T. Huynh^{a,b,c,d}, Farshid Guilak^{a,b,c}

^aDepartment of Orthopaedic Surgery, Washington University, 4515 McKinley Ave, St. Louis MO 63110, USA

^bShriners Hospital for Children – St. Louis, 4400 Clayton Ave, St. Louis, MO 63110, USA

^cCenter of Regenerative Medicine, Washington University, St. Louis, MO 63110, USA

^dUniversity of Rochester, Rochester, NY 14627, USA

Abstract

The coordinated spatial and temporal regulation of gene expression in the murine hindlimb determines the identity of mesenchymal progenitors and the development of diversity of musculoskeletal tissues they form. Hindlimb development has historically been studied with lineage tracing of individual genes selected *a priori*, or at the bulk tissue level, which does not allow for the determination of single cell transcriptional programs yielding mature cell types and tissues. To identify the cellular trajectories of lineage specification during limb bud development, we used single cell mRNA sequencing (scRNA-seq) to profile the developing murine hindlimb between embryonic days (E)11.5-E18.5. We found cell type heterogeneity at all time points, and the expected cell types that form the mouse hindlimb. In addition, we used RNA fluorescence in situ hybridization (FISH) to examine the spatial locations of cell types and cell trajectories to understand the ancestral continuum of cell maturation. This data provides a resource for the transcriptional program of hindlimb development that will support future studies of musculoskeletal development and generate hypotheses for tissue regeneration.

Corresponding author: Farshid Guilak, Ph.D., Departments of Orthopaedic Surgery, Washington University in St. Louis, 4515 McKinley Ave, Box 8233, Saint Louis, MO 63110, Tel: (314) 362-7239, guilak@wustl.edu.

Authors' roles: Study design: NHK, NPTH, FG. Study conduct: NHK, NPTH. Data collection: NHK, NPTH. Data analysis: NHK. Data interpretation: NHK, NPTH, FG. Drafting manuscript: NHK. Revising manuscript content: NHK, NPTH, FG. Approving final version of manuscript: all authors. NHK takes responsibility for the integrity of the data analysis.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

GEO ACCESSION NUMBER

The data in this manuscript can be retrieved using Geo accession number GSE142425 (in progress).

DISCLOSURES

All authors state that they have no competing interests conflicts of interest with this work.

Keywords

Limb bud; Hindlimb development; Cartilage; Muscle; Tendon; Bone

INTRODUCTION

The development of the mammalian limb, and subsequently the synovial joints, involves a complex and coordinated sequence of events whereby a relatively homogeneous limb bud differentiates into multiple cell types forming the tissues in the limb. This process involves the time-dependent synthesis and degradation of various matrix proteins and signaling molecules in a highly-coordinated manner [1–3]. While specific aspects of limb development have been deciphered, a more thorough and unbiased understanding of the developmental trajectories of the various cell populations of the limb may provide important insights into new methods for enhancing regeneration of individual musculoskeletal tissues, as well as complex processes such as joint formation.

Mapping of the precise sequence of events involved in limb development may provide new targets to stimulate regeneration of musculoskeletal tissues. Although limb development has been studied via the expression of pre-determined markers and fate-mapping, the trajectories of gene expression that drive individual cell differentiation toward distinct lineages are not fully understood due to limitations in *a priori* selection of limited lineage markers and the effects of combining all cells for conventional “bulk” transcriptomic analyses.

In this regard, advances in single cell gene expression analysis have made it possible to decipher cell transcription at increasingly high resolution. Massively parallel droplet based assays can determine the transcriptional profiles of thousands of single cells simultaneously [4–6]. The availability of these new tools has led to an increasing number of initiatives to create comprehensive cell atlases in the human [7, 8], and mouse [9, 10]. These cell atlases focus on adult tissues, while recent work in the mouse has examined gastrulation and organogenesis of whole organisms at single-cell resolution [11, 12]. Other studies have focused on particular organ or cell-type development, revealing previously unappreciated heterogeneity in neuronal, myocardial, and pulmonary development [13–17]. Recently, a single-cell transcriptomic atlas of limb development in the chick highlighted the transcriptional complexity of the 23 distinct cell populations in the autopod [18].

These studies have enhanced the understanding of development of these tissues and provided resources and data to advance regenerative medicine. However, such single-cell data has not been reported for the development of the mammalian limb. The goal of this study was to quantify the transcriptional landscape during mouse limb development via single cell RNA-sequencing and to determine the trajectories of cellular lineage specification into joint tissues during this process.

RESULTS

To examine the transcriptional landscape of murine hindlimb development, the limb bud was isolated at four developmental time points: embryonic day 11.5 (E11.5) to examine early

limb bud formation, E13.5 when cartilaginous condensations form, E15.5 for joint cavitation, and E18.5 when limb morphogenesis has progressed. Single cells were dissociated from hindlimb tissues and droplet-based high-throughput single-cell RNA-sequencing was performed (10x Genomics) (Fig. 1A) [6]. In addition, a species mixing experiment of mouse E18.5 hindlimb cells combined 1:1 with human 293T cells from culture was analyzed to determine single cell capture efficiency. Single cell capture was highly reliable as demonstrated by a low (1.4%) multiplet rate (Supplemental Fig. 1). mRNA from an average of 2,500 cells per experimental time point was captured and sequenced with an average of 200K sequencing reads per cell. We detected the expression of an average of 2,700 genes per cell in each sample.

Cluster analysis within each time point yielded between 5–7 distinct clusters which separated across time (Fig. 1B–E) [19]. At E11.5, we identified 5 cell types representing cartilage precursors (*Hoxd13* expression), bone/tendon precursors (*Tbx13*), skin (*Krt14*), vasculature (*Cdh5*), and blood (*Lyz2*) (Fig. 1B, F and Supplemental Fig. 2A). At E13.5, cartilage and bone/tendon clusters expressed more mature marker genes (*Col2a1* and *Col1a1*, respectively). In addition to skin, vasculature, and blood clusters there were additionally clusters representing the growth plate (*Col10a1*) and muscle (*Myog*) at E13.5 (Fig. 1C, G and Supplemental Fig. 2B). At E15.5, the number of clusters and cell type percentages did not differ from E13.5 (Fig. 1D, H and Supplemental Fig. 2C). At both E13.5 and E15.5 the majority of cells were cartilaginous. Finally, at E18.5 there was no longer a cluster representing the growth plate, and cell type percentages were more evenly distributed (Fig. 1E, I and Supplemental Fig. 2D).

To determine the similarity of cell types across time points, canonical correlation analysis was performed [21]. Across all time points there were 9,466 cells in 7 clusters, with each cluster containing cells from all time points (Fig. 2). Even following cell cycle regression, a cluster emerged with gene expression enriched for mitotic nuclear division (GO:0140014). Cells in this cluster were primarily from the earliest time point (E11.5, 49%), while at E18.5 only 8% of cells were assigned to the cell cycle cluster. The six other cell type clusters recapitulated the cell types found in the individual time points: cartilage, bone/tendon, skin, muscle, blood, and vasculature. The cartilage cluster was mostly comprised of cells from the middle time points (E13.5, 41% and E15.5, 44%), reflecting the prevalence of cartilaginous skeletal elements. As seen in the individual time point clustering, there were very few cells at E11.5 (4%) that were clustered as muscle. In fact, the muscle cluster was comprised mostly of cells from E18.5 (64%). Likewise, the blood cluster was mostly E18.5 (54%).

To examine developmental trajectories, we used an unsupervised algorithm, p-Create, that produces multi-branching graphs from single-cell data [22]. E11.5 and E18.5 cells were found at the ends of the cell trajectory, with E13.5/E15.5 cells along the middle of the trajectory branches (Fig. 3A). A large branch consisting of E11.5 cells was comprised of musculoskeletal precursors which gave rise to cartilage (Fig. 3A, purple arrow; Fig. 3B *Col2a1* expression overlay), and bone/tendon (Fig. 3A, blue arrow; Fig. 3B *Col1a1* expression overlay). While muscle cells arose from later cells (E13.5) along this same branch (Fig. 3A, green arrow; Fig. 3B *Myog* expression overlay). A separate branch defined the developmental trajectory of skin (Fig. 3A, brown arrow; Fig. 3B *Krt14* expression

overlay). Lastly, another branch expressed both vasculature (Fig. 3B *Cdh5* expression overlay) and blood (Fig. 3A, red arrow; Fig. 3B *Lyz2* expression overlay) markers.

To verify the spatial distribution of specific marker genes and time point expression profiles, we performed RNA FISH for two sets of probes. One set examined the expression of a cartilage marker gene (*Col2a1*), a skin marker gene (*Krt14*), and a muscle marker gene (*Myog*). We found that cartilage and skin were evident at all time points, while muscle was not visible until the latest time point (E18.5). *Col2a1* was expressed across a diffuse area of the hindlimb bud at E11.5, which coalesced into cartilaginous rays indicating the future sites of bones in the hind paw at E13.5, while at E15.5 and E18.5 the cartilaginous rays were now interrupted, indicating joint cavitation had occurred. Furthermore, at E18.5, *Col2a1* expression was further restricted to the joint surfaces (Fig. 4).

The second set of marker probes were selected to better understand the cell clusters determined via tSNE that were marked by *Colla1* expression. Because we did not identify a specific cell cluster at any time point marked by tendon gene expression (e.g. Scleraxis, *Scx*), we wanted to spatially verify the expression of *Colla1*, *Scx* (to identify tendon), and *Runx2* (to identify bone). At E11.5 there was not clear expression of the three markers. At E13.5 there was faint staining of *Colla1*, along the rays of the cartilaginous condensations. By E15.5 *Runx2* was found along the nascent bone, both overlapping and, separate from, *Colla1*. Minimal *Scx* was seen at E15.5, which did not overlap with *Colla1* or *Runx2*. Finally, at E18.5 bone and skin were marked by *Runx2* expression, while *Colla1* was in locations on bone that overlapped with *Runx2*, and along the sides of the bone, possibly marking tendon. However, *Scx* expression was low or sporadic (Fig. 5).

DISCUSSION

We successfully generated single cell RNA-sequencing data profiling the transcriptional landscape of murine limb development. We observed significant cell type heterogeneity, even at the earliest developmental time point that a limb bud could be manually isolated (E11.5). The identified cell clusters recapitulated the known cell types present in hindlimb development, i.e., cartilage, bone, muscle, skin, vasculature and blood. Trajectories of musculoskeletal cells identified five main developmental branches that help to explain the ancestral continuum of cell maturation. These findings provide the basis for further understanding of the process of early limb and joint development, which will hopefully serve to inform new approaches for deconstructing and eventually recapitulating the regenerative process.

Cell type heterogeneity at the early limb bud stage (E11.5) was observed, with musculoskeletal precursors clustering into two populations which seem primed even at this early stage to develop into cartilage or bone/tendon. Similarly, five cell type clusters were found at a roughly equivalent developmental time point in the chick autopod (HH25, approximately mouse E12 [23]), although muscle cells were found in chick, but not mouse [18]. Genes indicating the presence of muscle cells were apparent at E13.5 by scRNA-seq, but were not observed with RNA FISH until E18.5. The lack of signal in RNA FISH until

the latest time point could be explained by the single, 2D section. Other cell type clusters included blood, vasculature, bone, cartilage, and skin.

Hypertrophic chondrocytes are typically marked by *Col10a1* expression and are prevalent in the growth plate. We found clusters defined by *Col10a1* expression at the middle developmental time points, E13.5 and E15.5, but not at E18.5. This, again, may be due to cell type drop-out due to limitations in cell capture or it may be that hypertrophic chondrocytes are more similar to bone or cartilage cells at the most mature time point.

As spatial expression information is lost with scRNA-seq due to the dissociation of single cells, we used RNA FISH to examine the spatial expression of markers for cartilage, skin, and muscle in intact embryonic tissues. In addition, we used a second set of markers to determine the spatial location of cells with high expression of *Col1a1*, and to see if these cells were at locations of bone, tendon/ligament, or both. It appears that *Col1a1* may be expressed both at sites of bone (overlap with bone marker, *Runx2*), and possibly at sites of tendon/ligament as well (due to spatial location at edges of cartilage/bone surface).

Although some low and sporadic expression of *Scx* was observed, we were unable to detect a cell cluster that was defined by *Scx* expression. To determine if the absence of *Scx* in the sequencing data was due to a lack of expression *in situ* or to scRNA-seq artefact (i.e., cell type drop out) we performed FISH analysis, which confirmed low levels of expression of *Scx*, indicating that this gene it was not strongly expressed at these time points. These findings are generally consistent with a recent study that performed scRNAseq on 6 week old mouse Achilles tendons, and also did not demonstrate widespread *Scx* expression [24]. These authors attributed this finding to the possibility that the widely used ScxGFP mice contain only portions of the full length regulatory elements of *Scx*, raising the possibility that these mice overestimate *Scx* expression [24].

Cells were aligned along developmental trajectories that defined cartilage, bone, skin, muscle, and blood. Future work will further explore the transitional gene expression that determines whether an early mesenchymal progenitor becomes cartilage or bone, for example. This information may be used in *in vitro* systems to more robustly convert embryonic, mesenchymal, or induced pluripotent cells to musculoskeletal tissues for tissue engineering applications. It is important to note that these markers of gene expression will be modified by local gradients in extracellular matrix properties and signaling molecules to drive embryonic development, and a further understanding of these cell-matrix interactions will provide new insights into enhancing limb regeneration [25–27].

A well characterized marker of joint formation, *Gdf5*, was not highly expressed in the cells in our study. *Gdf5* is a transient marker of the interzone, and is present in a restricted region that is only a few cell layers thick and is thus only expressed in a small number of cells [28, 29]. This could be due to limitations in cell capture or the need for larger cell numbers to robustly quantify rare cell populations. Furthermore, it was difficult to classify the clusters high in collagens type I and III as bone, tendon or ligament because there was not clear expression of tendon markers scleraxis (*Scx*) or tenomodulin (*Tnmd*) or bone markers *Runx2* or osteocalcin (*Bglap*). Deeper sequencing may be necessary to more accurately

define this cell population and gain insight into transcripts that may have relatively low expression levels.

In conclusion, this work provides an in-depth resource describing the development of the mouse hindlimb. In combination with other genetic methods such as lineage tracing, this approach could provide new insights into the interaction among different cell types. A thorough understanding of the limb development process will hopefully allow recapitulation of these events to enhance joint repair or, eventually, joint or limb regeneration.

METHODS

Animals

Murine hindlimbs were manually dissected at 4 time points during embryonic limb development, including points marking formation of the limb bud (E11.5), formation of cartilaginous condensations (E13.5), joint cavitation (E15.5), and joint morphogenesis (E18.5). Embryonic pups (7–9/litter) were removed by caesarean section from a timed-pregnant female C57BL/6 mouse per time point (The Jackson Laboratory, Bar Harbor, ME) and euthanized. Hindlimbs from 6–7 pups/litter were digested with collagenase (Type II, Worthington-biochem, Lakewood, NJ) and pronase (EMD Millipore, Billerica, MA) in 15-minute increments with agitation at 37°C for up to one hour. Every 15 minutes, mechanical dissociation was applied by pipetting up and down. Following digestion, cells from each litter (time point) were pooled, strained through a 100µm cell strainer, pelleted, resuspended at 1×10^6 cells per 500µl in freeze media (80% FBS, 10% DMEM, 10% DMSO), and stored in liquid nitrogen until single cell capture. All experiments were approved by the Washington University IACUC.

Single Cell Capture, and RNA-sequencing

Cells in freeze media were rapidly thawed in a 37°C water bath, and resuspended in 1X PBS and 0.04% bovine serum albumin. Dead cells were removed using magnetic bead-bound antibodies against apoptotic and necrotic cells (Dead Cell Removal Kit, Miltenyi Biotec) and cells were washed twice. Single cell capture was performed with the 10x Chromium Controller (10x Genomics, Pleasanton, CA) [6]. It was necessary to include a freeze/thaw cycle in the preparation of our samples from different developmental time points. This may preferentially select for some cell types over others. However, it has been shown that cryopreserved cells and tissues have similar gene expression profiles and proportions of cell type subpopulations as fresh samples [6, 30].

Briefly, single cell suspensions were brought to a concentration of 1,000 cells/µl and approximately 6,000 cells per time point (4 samples total) were loaded into the 8-channel microfluidic chip that encapsulates thousands of cells in gel beads in emulsion (GEMs). Cell capture, cDNA generation, and library preparation were performed using Chromium Single Cell 3' v2 Reagent kit, following the manufacturer's instructions. To test the reliability of single-cell capture, a 1:1 mixture of E18.5 mouse hindlimb cells and human 293T cells was sequenced and the multiplet rate was determined by the fraction of mouse reads in human barcodes and vice versa. Single cell RNA-seq libraries were pooled and run on the Illumina

NovaSeq S1 flow cell with 26×98bp reads to generate ~500 million reads per library (time point).

Data Analysis

Cell Ranger software (v2, 10x Genomics) was used to demultiplex samples, process barcodes, align to the mouse (GRCm38/mm10) or human (hg19; for mixed species) genome assembly, and count single cell genes [6]. The gene/barcode matrix was imported to Seurat (<https://github.com/satijalab/seurat>, v2.3.4) and filtering was performed to exclude cells with less than 200 or more than 6,000 genes. Cells with more than 10% unique molecular identifier (UMI) counts associated with mitochondrial genes were also excluded. The data was log-normalized and library size and mitochondrial UMI counts were regressed.

The cell cycle stage of each cell was assigned as S, G1, or G2/M based on a pre-determined list of cell cycle markers [31] and the difference between the G2/M and S scores was calculated and variance-corrected.

Cluster analysis was performed via t-distributed stochastic neighbor embedding (tSNE) [32] at each time point to examine cellular heterogeneity and cell types present in hindlimb tissues [21, 33]. In addition, the four time points were combined using Seurat's Multi-canonical correlation analysis (CCA) command to identify projection vectors that maximize the overall correlation across all data sets [21, 33]. Cluster specific genes were determined by calculating the expression difference of each gene between that cluster and the average of the rest of the clusters. The top 100 cluster-specific genes were input into Enrichr to determine major cellular subtypes based on gene ontology (enriched biological processes) and cell type (enriched ARCHS4 Tissues) [34, 35].

Cell trajectories were determined using p-Creode (<https://github.com/KenLauLab/pCreode>) [22]. First, neighborhood variance ratio (NVR) gene selection was used to select genes with local and monotonic variation such that the selected genes possess specific expression patterns over the entire data space amenable to trajectory analysis (<https://github.com/KenLauLab/NVR>) [36, 37]. Then, cells in two-dimensional expression space were downsampled and a density-based k-nearest neighbor network was constructed. Downsampling was performed with a radius of 30, noise of 8, and target density of 25, which downsampled the original 9,924 input cells to 5,615 (57%). End states were identified by K-means clustering and silhouette scoring of cells with low closeness values (<mean), and the number of end states was doubled to account for rare cell types. A topology was created using a hierarchical placement strategy of cells on path nodes between end states which allowed for the placement of data points along an ancestral continuum. Finally, a representative topology was extracted using p-Creode scoring from an ensemble of n=100 topologies. Graph ID number 3 was chosen based on biological knowledge.

RNA FISH

Hindlimbs from one pup per litter (time point) were placed in OCT and frozen. Blocks were cryo-sectioned at 10µm thickness and fixed in 4% paraformaldehyde for 15 minutes. Sections were hybridized with probes for mouse collagen, type II, alpha 1 (*Col2a1*), keratin 14 (*Krt14*) and myogenin (*Myog*). A second set was hybridized with probes for collagen,

type I, alpha 1 (*Col1a1*), scleraxis (*Scx*), and Runx family transcription factor 2 (*Runx2*) following the manufacturer's instructions (Advanced Cell Diagnostics, Inc., RNAscope® Fluorescent Multiplex Assay) [38]. Sections were imaged with confocal microscopy (Zeiss LSM 880).

Histology

Hindlimbs from one pup per litter (time point) were fixed in 4% paraformaldehyde for 24 hours and processed for paraffin embedding. Paraffin embedded samples were sectioned at 5µm thickness and stained with Hematoxylin, Safranin-O and Fast Green.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

Supported in part by grants from the NIH (AR50245, AR46927, AG15768, P30 AR74992, P30 AR73752), Arthritis Foundation, and Nancy Taylor Foundation for Chronic Diseases. We thank the Genome Technology Access Center in the Department of Genetics at Washington University School of Medicine. The authors thank Sara Oswald for her editing and imaging expertise and Dr. Samantha Morris for discussions in the early stages of this work.

Financial support: Supported in part by grants from the NIH (AR50245, AR46927, AG15768, P30 AR74992, P30 AR73752), Arthritis Foundation, and Nancy Taylor Foundation for Chronic Diseases.

REFERENCES

- [1]. Gauci SJ, Golub SB, Tatarczuch L, Lee E, Chan D, Walsh NC, Little CB, Stanton H, Lokmic Z, Sims NA, Mackie EJ, Fosang AJ, Disrupted type II collagenolysis impairs angiogenesis, delays endochondral ossification and initiates aberrant ossification in mouse limbs, *Matrix Biol* 83 (2019) 77–96. [PubMed: 31381970]
- [2]. Hubmacher D, Taye N, Balic Z, Thacker S, Adams SM, Birk DE, Schweitzer R, Apte SS, Limb- and tendon-specific *Adamsl2* deletion identifies a role for *ADAMTSL2* in tendon growth in a mouse model for geleophysic dysplasia, *Matrix Biol* 82 (2019) 38–53. [PubMed: 30738849]
- [3]. Garantziotis S, Savani RC, Hyaluronan biology: A complex balancing act of structure, function, location and context, *Matrix Biol* 78-79 (2019) 1–10. [PubMed: 30802498]
- [4]. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW, Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells, *Cell* 161(5) (2015) 1187–201. [PubMed: 26000487]
- [5]. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA, Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets, *Cell* 161(5) (2015) 1202–14. [PubMed: 26000488]
- [6]. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, Gregory MT, Shuga J, Montesclaros L, Underwood JG, Masquelier DA, Nishimura SY, Schnall-Levin M, Wyatt PW, Hindson CM, Bharadwaj R, Wong A, Ness KD, Beppu LW, Deeg HJ, McFarland C, Loeb KR, Valente WJ, Ericson NG, Stevens EA, Radich JP, Mikkelsen TS, Hindson BJ, Bielas JH, Massively parallel digital transcriptional profiling of single cells, *Nat Commun* 8 (2017) 14049. [PubMed: 28091601]
- [7]. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, Bodenmiller B, Campbell P, Carninci P, Clatworthy M, Clevers H, Deplancke B, Dunham I, Eberwine J, Eils R, Enard W, Farmer A, Fugger L, Gottgens B, Hacohen N, Haniffa M, Hemberg M, Kim S, Klenerman P, Kriegstein A, Lein E, Linnarsson S, Lundberg E, Lundberg J, Majumder P, Marioni JC, Merad M, Mhlanga M, Nawijn M, Netea M, Nolan G, Pe'er D, Phillipakis A, Ponting CP, Quake S,

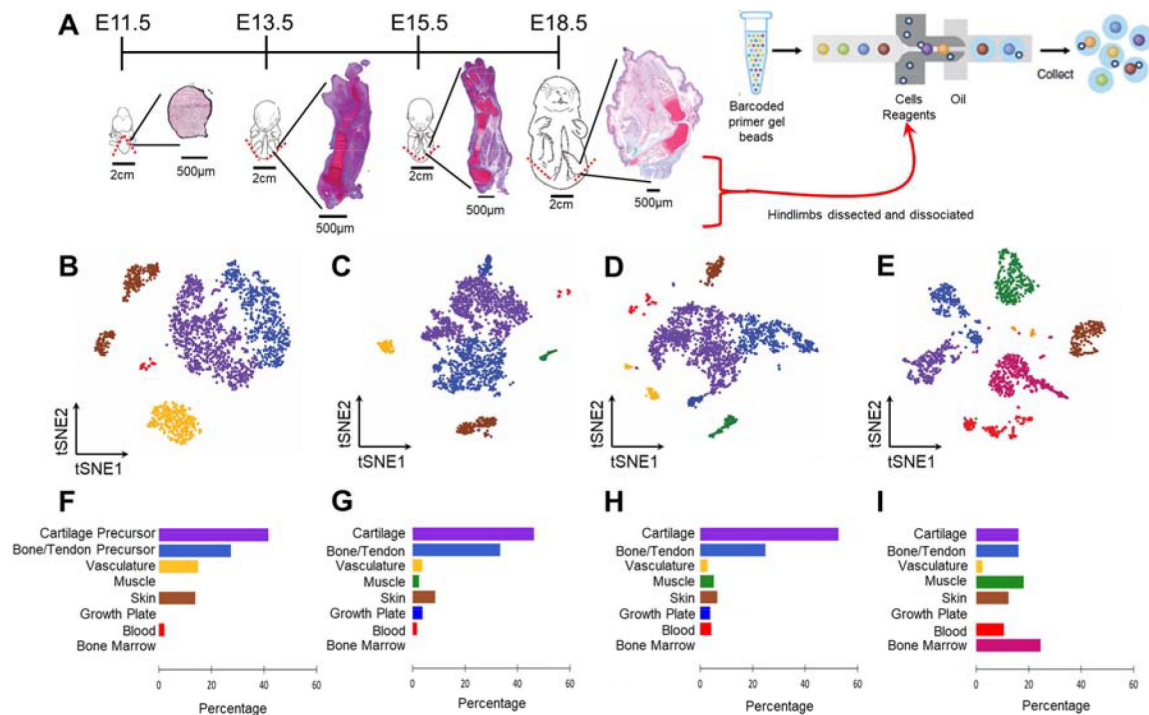
Reik W, Rozenblatt-Rosen O, Sanes J, Satija R, Schumacher TN, Shalek A, Shapiro E, Sharma P, Shin JW, Stegle O, Stratton M, Stubbington MJT, Theis FJ, Uhlen M, van Oudenaarden A, Wagner A, Watt F, Weissman J, Wold B, Xavier R, Yosef N, Human P Cell Atlas Meeting, The Human Cell Atlas, *Elife* 6 (2017).

- [8]. Rozenblatt-Rosen O, Stubbington MJT, Regev A, Teichmann SA, The Human Cell Atlas: from vision to reality, *Nature* 550(7677) (2017) 451–453. [PubMed: 29072289]
- [9]. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, Saadatpour A, Zhou Z, Chen H, Ye F, Huang D, Xu Y, Huang W, Jiang M, Jiang X, Mao J, Chen Y, Lu C, Xie J, Fang Q, Wang Y, Yue R, Li T, Huang H, Orkin SH, Yuan GC, Chen M, Guo G, Mapping the Mouse Cell Atlas by Microwell-Seq, *Cell* 173(5) (2018) 1307. [PubMed: 29775597]
- [10]. Tabula Muris C, c. Overall, c. Logistical, c. Organ, processing, p. Library, sequencing, a. Computational data, a. Cell type, g. Writing, g. Supplemental text writing, i. Principal, Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris, *Nature* 562(7727) (2018) 367–372. [PubMed: 30283141]
- [11]. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C, Shendure J, The single-cell transcriptional landscape of mammalian organogenesis, *Nature* 566(7745) (2019) 496–502. [PubMed: 30787437]
- [12]. Pijuan-Sala B, Griffiths JA, Guibentif C, Hiscock TW, Jawaid W, Calero-Nieto FJ, Mulas C, Ibarra-Soria X, Tyser RCV, Ho DLL, Reik W, Srinivas S, Simons BD, Nichols J, Marioni JC, Gottgens B, A single-cell molecular map of mouse gastrulation and early organogenesis, *Nature* 566(7745) (2019) 490–495. [PubMed: 30787436]
- [13]. Delile J, Rayon T, Melchionda M, Edwards A, Briscoe J, Sagner A, Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord, *Development* 146(12) (2019).
- [14]. La Manno G, Gyllborg D, Codeluppi S, Nishimura K, Salto C, Zeisel A, Borm LE, Stott SR, Toledo EM, Villaescusa JC, Lonnerberg P, Ryge J, Barker RA, Arenas E, Linnarsson S, Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells, *Cell* 167(2) (2016) 566–580 e19. [PubMed: 27716510]
- [15]. Lescroart F, Wang X, Lin X, Swedlund B, Gargouri S, Sanchez-Danes A, Moignard V, Dubois C, Paulissen C, Kinston S, Gottgens B, Blanpain C, Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq, *Science* 359(6380) (2018) 1177–1181. [PubMed: 29371425]
- [16]. Mayer C, Hafemeister C, Bandler RC, Machold R, Batista Brito R, Jaglin X, Allaway K, Butler A, Fishell G, Satija R, Developmental diversification of cortical inhibitory interneurons, *Nature* 555(7697) (2018) 457–462. [PubMed: 29513653]
- [17]. Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH, Desai TJ, Krasnow MA, Quake SR, Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq, *Nature* 509(7500) (2014) 371–5. [PubMed: 24739965]
- [18]. Feregrino C, Sacher F, Parnas O, Tschopp P, A Single-cell Transcriptomic Atlas of the Developing Chicken Limb, *bioRxiv*(2019).
- [19]. Satija R, Farrell JA, Gennert D, Schier AF, Regev A, Spatial reconstruction of single-cell gene expression data, *Nat Biotechnol* 33(5) (2015) 495–502. [PubMed: 25867923]
- [20]. Richardson L, Venkataraman S, Stevenson P, Yang Y, Moss J, Graham L, Burton N, Hill B, Rao J, Baldock RA, Armit C, EMAGE mouse embryo spatial gene expression database: 2014 update, *Nucleic Acids Res* 42(Database issue) (2014) D835–44. [PubMed: 24265223]
- [21]. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R, Integrating single-cell transcriptomic data across different conditions, technologies, and species, *Nat Biotechnol* 36(5) (2018) 411–420. [PubMed: 29608179]
- [22]. Herring CA, Banerjee A, McKinley ET, Simmons AJ, Ping J, Roland JT, Franklin JL, Liu Q, Gerdes MJ, Coffey RJ, Lau KS, Unsupervised Trajectory Analysis of Single-Cell RNA-Seq and Imaging Data Reveals Alternative Tuft Cell Origins in the Gut, *Cell Syst* 6(1) (2018) 37–51 e9. [PubMed: 29153838]
- [23]. Wanek N, Muneoka K, Holler-Dinsmore G, Burton R, Bryant SV, A staging system for mouse limb development, *J Exp Zool* 249(1) (1989) 41–9. [PubMed: 2926360]

- [24]. Swanson JB, De Micheli AJ, Disser NP, Martinez LM, Walker NR, Cosgrove BD, Mendias CL, A single-cell transcriptional atlas identifies extensive heterogeneity in the cellular composition of tendons, *bioRxiv* 801266 (2019).
- [25]. Guilak F, Nims RJ, Dicks A, Wu CL, Meulenbelt I, Osteoarthritis as a disease of the cartilage pericellular matrix, *Matrix Biol* 71-72 (2018) 40–50. [PubMed: 29800616]
- [26]. Krishnan Y, Grodzinsky AJ, Cartilage diseases, *Matrix Biol* 71-72 (2018) 51–69. [PubMed: 29803938]
- [27]. Urbanczyk M, Layland SL, Schenke-Layland K, The role of extracellular matrix in biomechanics and its impact on bioengineering of cells and 3D tissues, *Matrix Biol* (2019).
- [28]. Decker RS, Koyama E, Pacifici M, Genesis and morphogenesis of limb synovial joints and articular cartilage, *Matrix Biol* 39 (2014) 5–10. [PubMed: 25172830]
- [29]. Koyama E, Shibukawa Y, Nagayama M, Sugito H, Young B, Yuasa T, Okabe T, Ochiai T, Kamiya N, Rountree RB, Kingsley DM, Iwamoto M, Enomoto-Iwamoto M, Pacifici M, A distinct cohort of progenitor cells participates in synovial joint and articular cartilage formation during mouse limb skeletogenesis, *Dev Biol* 316(1) (2008) 62–73. [PubMed: 18295755]
- [30]. Guillaumet-Adkins A, Rodriguez-Esteban G, Mereu E, Mendez-Lago M, Jaitin DA, Villanueva A, Vidal A, Martinez-Marti A, Felipe E, Vivancos A, Keren-Shaul H, Heath S, Gut M, Amit I, Gut I, Heyn H, Single-cell transcriptome conservation in cryopreserved cells and tissues, *Genome Biol* 18(1) (2017) 45. [PubMed: 28249587]
- [31]. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, Fallahi-Sichani M, Dutton-Regester K, Lin JR, Cohen O, Shah P, Lu D, Genshaft AS, Hughes TK, Ziegler CG, Kazer SW, Gaillard A, Kolb KE, Villani AC, Johannessen CM, Andreev AY, Van Allen EM, Bertagnolli M, Sorger PK, Sullivan RJ, Flaherty KT, Frederick DT, Jane-Valbuena J, Yoon CH, Rozenblatt-Rosen O, Shalek AK, Regev A, Garraway LA, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq, *Science* 352(6282) (2016) 189–96. [PubMed: 27124452]
- [32]. van der Maaten L, Hinton GE, Visualizing Data using t-SNE, *J Mach Learn Res* 9 (2008) 2579–2605.
- [33]. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM., Stoeckius M, Smibert P, Satija R, Comprehensive integration of single cell data *bioRxiv* (2018).
- [34]. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A, Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool, *BMC Bioinformatics* 14 (2013) 128. [PubMed: 23586463]
- [35]. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, *Nucleic Acids Res* 44(W1) (2016) W90–7. [PubMed: 27141961]
- [36]. Welch JD, Hartemink AJ, Prins JF, SLICER: inferring branched, nonlinear cellular trajectories from single cell RNA-seq data, *Genome Biol* 17(1) (2016) 106. [PubMed: 27215581]
- [37]. Chen B, Herring CA, Lau KS, pyNVR: Investigating factors affecting feature selection from scRNA-seq data for lineage reconstruction, *Bioinformatics* (2018).
- [38]. Wang F, Flanagan J, Su N, Wang LC, Bui S, Nielson A, Wu X, Vo HT, Ma XJ, Luo Y, RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues, *J Mol Diagn* 14(1) (2012) 22–9. [PubMed: 22166544]

Highlights

- The coordinated spatial and temporal regulation of gene expression during embryonic development determines the diversity of tissue formation
- We used single cell mRNA sequencing to profile the developing murine hindlimb between embryonic days E11.5-E18.5.
- Cell heterogeneity was present throughout development and specific genes were validated by RNA-FISH.
- This data provides a resource for studying the transcriptional program of hindlimb development to inform strategies for tissue regeneration

**Fig. 1.**

Distinct populations can be determined in mouse embryonic hindlimb cells. (A) Schematic of sampling area at the four stages of hindlimb development (whole embryo schematics adapted from EMAP eMouse Atlas Project <http://www.emouseatlas.org> [20]). Both hindlimbs were dissected (indicated by red dashed lines) from $n=6-7$ pups per litter. Insets demonstrate cartilaginous morphology at each time point via Safranin-O/Fast Green/Hematoxylin staining. Hindlimbs were dissociated to single cells and captured using 10X single cell RNA-seq technology. (B-E) tSNE projection of the four developmental time points, containing 2,627 (E11.5), 2,815 (E13.5), 2,596 (E15.5) and 1,901 (E18.5) cells, where each cell is grouped into clusters (distinguished by their colors). Comparable cell populations identified in multiple samples are visualized using the same color. (F-I) Percentage of cells represented by each cluster. Colors correspond to the clusters identified in (B-E).

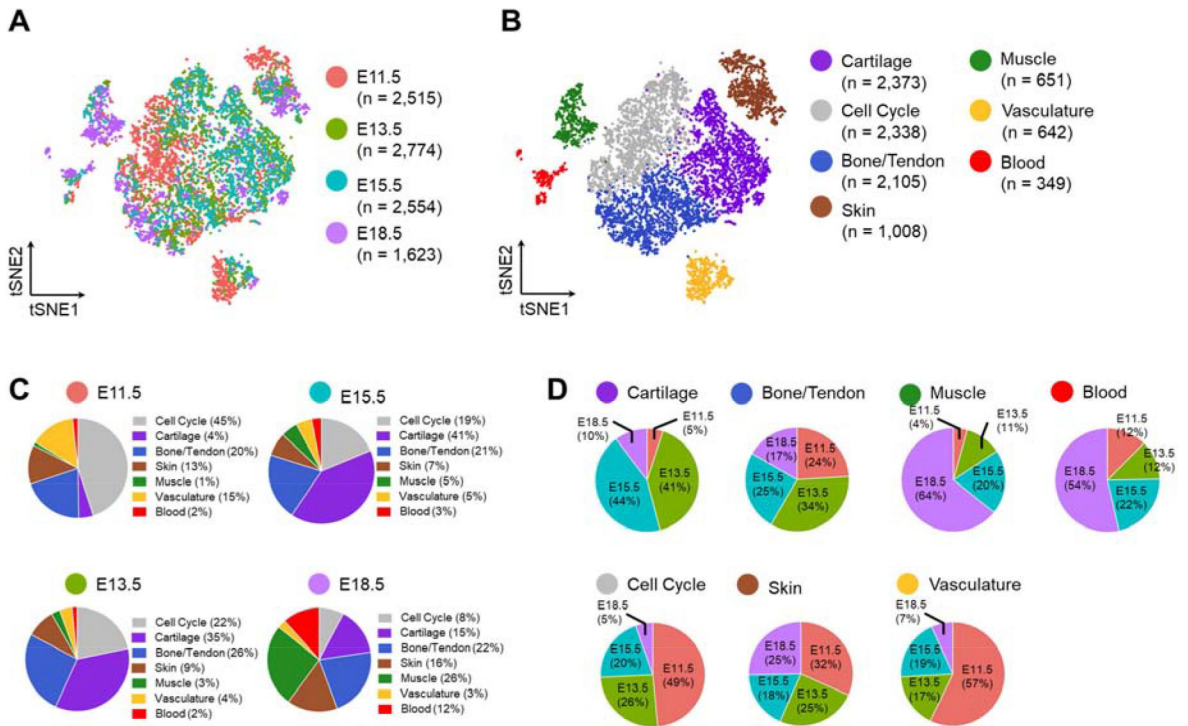


Fig. 2. tSNE projection of the merged dataset with cells colored by (A) time point and (B) cell type determined by unsupervised clustering. (C) Percentage of cell type at each time point and (D) percentage of cells in each type by time point.

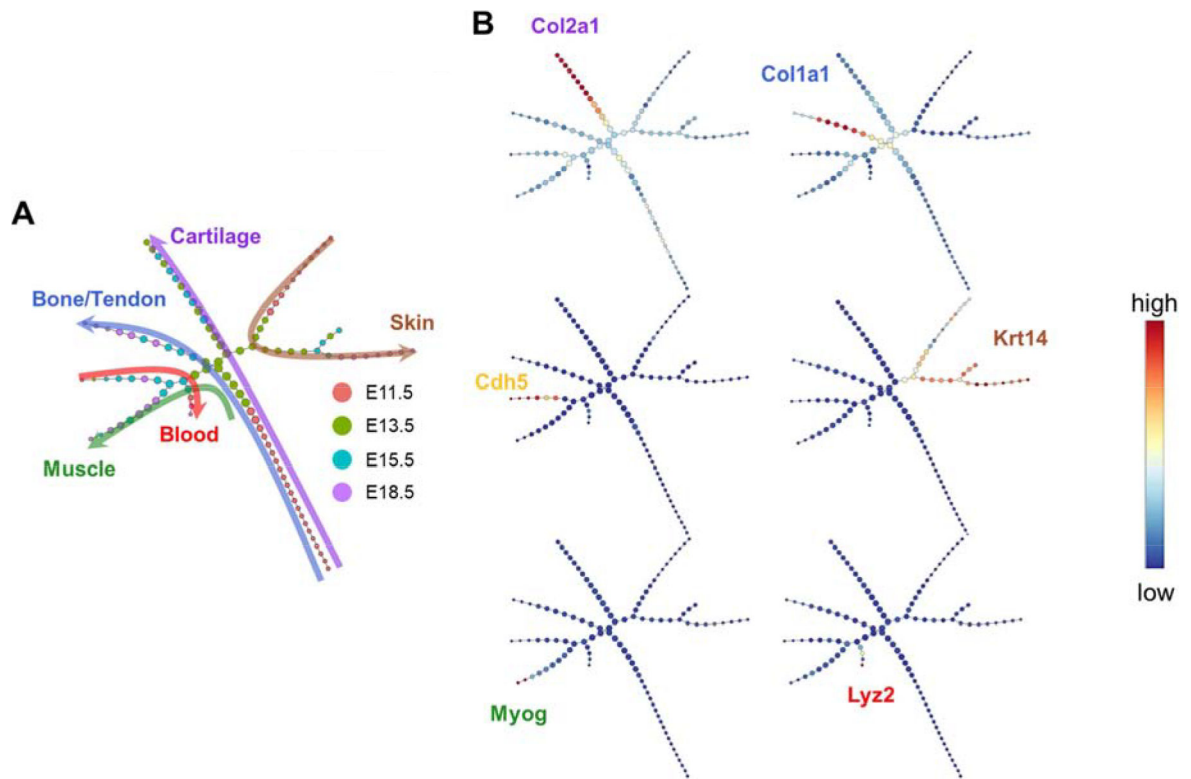


Fig. 3.
 (A) pCreode trajectory analysis with node colors indicating time point and colored arrows indicating tissue type across developmental trajectory. (B) Overlay of selected transcripts depicting cartilage (Col2a1), bone/tendon (Col1a1), vasculature (Cdh5), skin (Krt14), muscle (Myog), and blood (Lyz2) development in the hindlimb on the pCreode topology generated in (A).

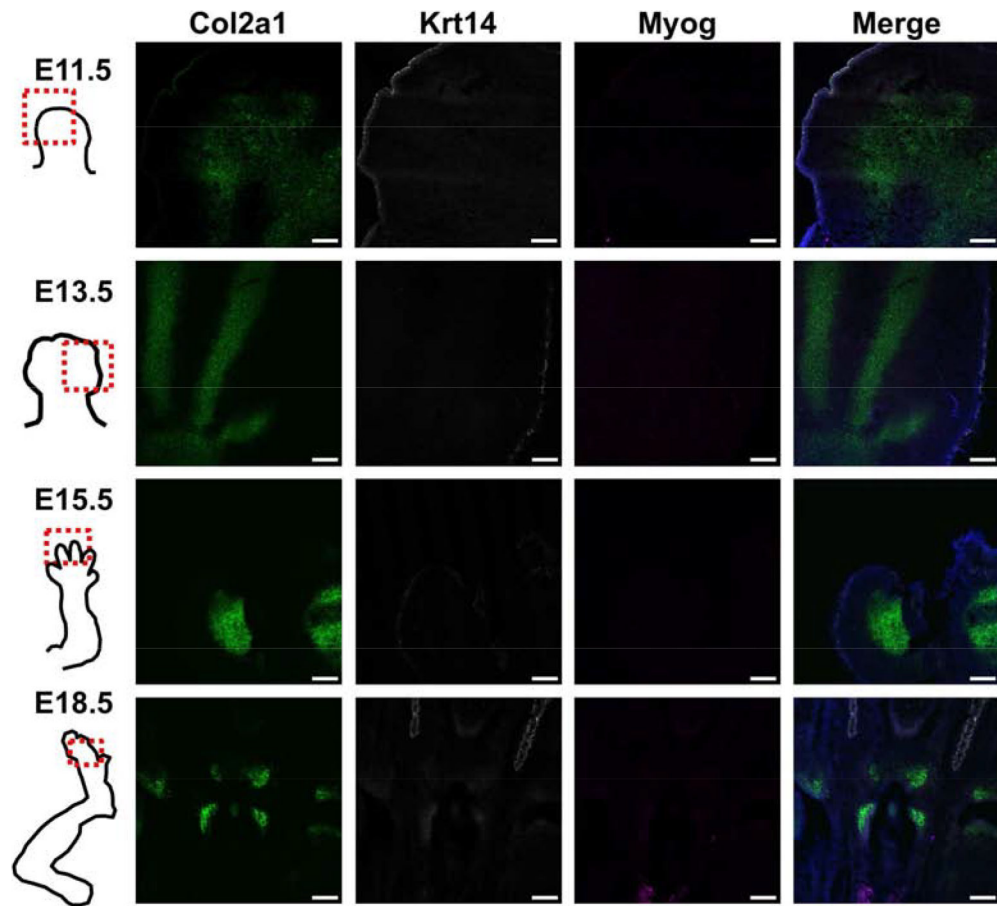


Fig. 4. Representative RNA FISH images of hindlimbs at four developmental time points to verify spatial locations of cartilage (*Col2a1*), skin (*Krt14*), and muscle (*Myog*). Scale bar = 100 μ m. A grayscale version of the individual channels is provided in Supplemental Fig. 4.

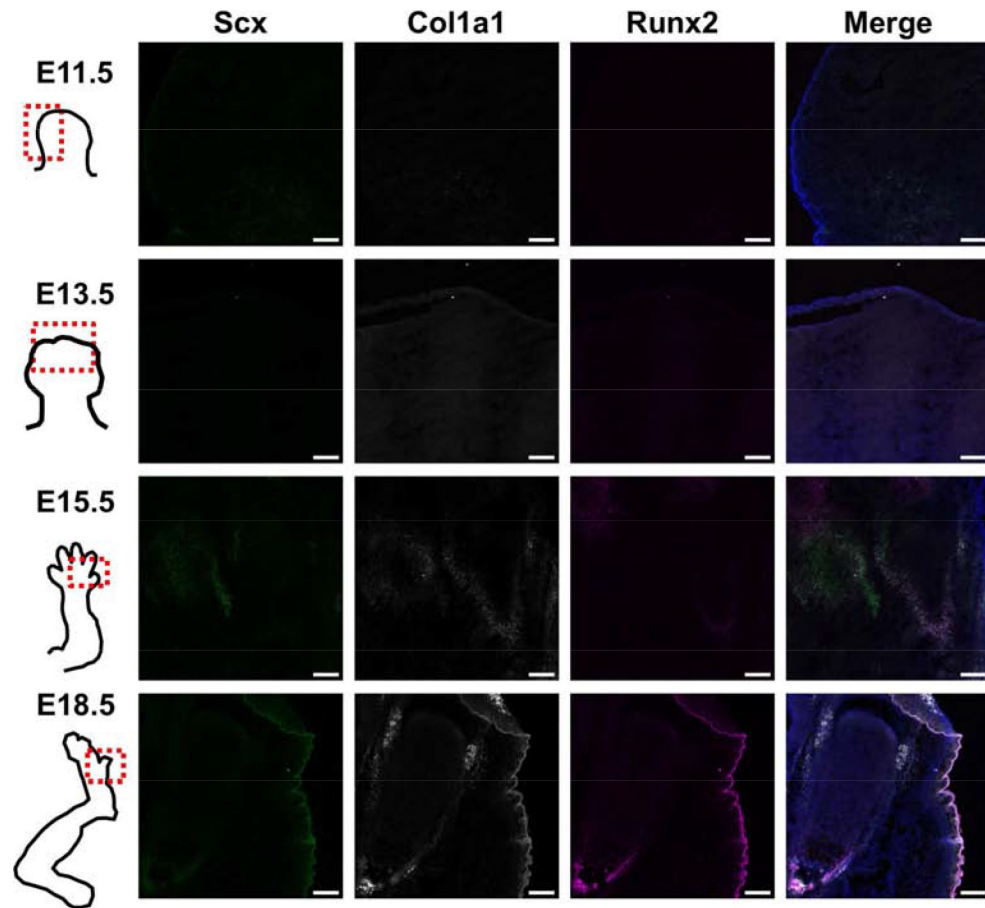


Fig. 5. Representative RNA FISH images of hindlimbs at four developmental time points of tendon (*Scx*) and bone (*Runx2*) markers, as well as *Col1a1*, to verify what cell types are defined by clusters high in *Col1a1*. Scale bar = 100 μ m. A grayscale version of the individual channels is provided in Supplemental Fig. 5.