eLife

Reviewed Preprint
v2 • July 17, 2025
Revised by authors

Reviewed Preprint
v1 • February 27, 2025

Developmental Biology

Single-nucleus transcriptional and chromatin accessibility analyses of maturing mouse Achilles tendon uncover the molecular landscape of tendon stem/progenitor cells

Hiroki Tsutsumi, Tomoki Chiba, Yuta Fujii, Takahide Matsushima, Tsuyoshi Kimura, Akinori Kanai, Akio Kishida, Yutaka Suzuki, Hiroshi Asahara

Department of Systems BioMedicine, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, Tokyo, Japan • Department of Molecular and Cellular Biology, The Scripps Research Institute, La Jolla, United States • Laboratory for Biomaterials and Bioengineering, Institute of Integrated Research, Institute of Science Tokyo, Tokyo, Japan • Department of Biomedical Engineering, Faculty of Life Science, Toyo University, Saitama, Japan • Laboratory of Systems Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, the University of Tokyo, Chiba, Japan

eLife Assessment

This study presents a **valuable** finding of novel markers that may potentially identify resident tendon stem/progenitor cells (TSPCs). The study also presents a comprehensive single-cell transcriptional dataset that will be of value to the field. The evidence supporting the identification of novel markers of a TSPC is **incomplete**, requiring clarification of current analyses and additional validation experiments to demonstrate that these markers are indeed specific and these cells are indeed TSPCs. This work will be of interest to biologists and engineers focused on tendons and ligaments.

https://doi.org/10.7554/eLife.104768.2.sa3

Abstract

Tendons and ligaments are crucial connective tissues linking bones and muscles, yet achieving full functional recovery after injury remains challenging. We investigated the characteristics of tendon stem/progenitor cells (TSPCs) by focusing on the declining tendon repair capacity with growth. Using single-cell RNA sequencing on Achilles tendon cells from 2-and 6-week-old mice, we identified *Cd55* and *Cd248* as novel surface antigen markers for TSPCs. Combining single-cell RNA sequencing with single-nucleus RNA and ATAC sequencing analyses revealed that *Cd55* and *Cd248* positive fractions in tendon tissue represent TSPCs, as confirmed by their expression of established TSPC markers, with this population decreasing at 6 weeks. We also identified candidate upstream transcription factors regulating these fractions. Functional analyses of isolated CD55/CD248 positive cells demonstrated high

clonogenic potential and tendon differentiation capacity, forming functional tendon-like tissue *in vitro*. This study establishes CD55 and CD248 as novel TSPC surface antigens, potentially advancing tendon regenerative medicine and contributing to the development of new treatment strategies for tendon and ligament injuries.

Introduction

Tendons are connective tissues involved in the contractile movement of muscle to bone. Tendons are rich in extracellular matrix components, such as type 1 collagen and proteoglycans, which both have elastic and viscous properties to withstand overload (Andarawis-Puri et al., 2015; Asahara et al., 2017; Connizzo et al., 2013).

Tendons have a hierarchical structure. Collagen fibrils aggregate to form fibers, which then bundle into fascicles. Each fascicle is surrounded by a thin connective tissue layer called the endotenon. Fascicles are further enclosed by the epitenon, and in some cases, an additional outer layer called the paratenon. The number of fascicles and the presence of a paratenon can vary among different tendons and species (Jozsa et al., 1991; Walia et al., 2019).

Tendon and ligament injuries potentially account for the majority of musculoskeletal disorders and can lead to arthritis and spondylitis. (Gracey et al., 2020). The most popular treatment approaches are surgery and conservation (Steinmann et al., 2020). However, full functional recovery is often not achieved due to scarring and fibrosis in the injured area (Guilak et al., 2014; Thomopoulos S et al., 2015). Furthermore, the risk of postoperative rupture is high (Leong et al., 2020; Loiacono et al., 2019). In this context, tendon regeneration focused on tendon stem/progenitor cells (TSPCs) is attracting attention (Leong et al., 2016). Tendons have few cellular components, including tendon cells (tenocytes) and TSPCs (Tang et al., 2016; Bi et al., 2007). Tenocytes are responsible for tendon homeostasis, while TSPCs self-renew and differentiate into tenocytes. Thus, investigating TSPCs biology is important for understanding tendon regeneration and homeostasis.

Murine TSPCs were first reported in 2007 by Bi Y et al. (Bi et al., 2007). They demonstrated that TSPCs possess self-renewal capacity, colony-forming ability, and multi-differentiation potential *in vitro*. Furthermore, murine TSPCs were reported to simultaneously express stem cell markers such as *Cd44* and *Stem cells antigen-1* (*Ly6a*), as well as tendon-related genes like *Scleraxis* (*Scx)* and Collagen type I alpha 1 chain (*Col1a1*), and were identified as a subset of tendon cells within the tendon fascicle.

Subsequent research on murine TSPC localization has yielded diverse perspectives. Harvey et al. reported that *Tppp3*/*Pdgfra*-positive cells in the epitenon are induced upon tendon injury and contribute to repair (Harvey et al., 2019). Additionally, Yin et al. and Tempfer et al. demonstrated the presence of TSPCs around blood vessels (Yin et al., 2016; Tempfer et al., 2009).

Markers used to characterize murine TSPCs include CD73, CD105, and CD90, which are known criteria for mesenchymal stem cells (MSCs) (Lui et al., 2015). Additionally, TSPCs have been reported to express *POU class 5 homeobox 1* (*Oct-4*), *Nanog*, *Nucleostemin*, stage-specific embryonic anitigen-4 *(SSEA-4)*, *c-Myc*, *SRY-box transcription factor (Sox*2), *Fucosyltransferase 4* (*Fut-4*), and other genes (Zhang et al., 2016). Expression of CD146 and CD44 has been confirmed as well (Ruzzini et al., 2014). The *Tppp3*/*Pdgfra*-positive TSPCs reported by Harvey et al. highly express *Cd34*, which is generally considered to have low expression in MSCs (Harvey et al., 2019; Tachibana, et al., 2022). *Cd34* is known to be highly expressed in mouse embryonic limb buds at E14.5 compared to E11.5 (Havis et al., 2014), suggesting that *Cd34* positive cells might reflect progenitor cells that constitute the limb bud, including tendons. However, these markers are not

specific to TSPCs, thus, a definitive method to distinguish TSPCs from mature tendons *in vivo* has not been established yet (Li et al., 2019; Wang et al., 2008; Chen et al., 2012; Cho et al., 2018; Fang et al., 2022).

Regarding tendon regenerative capacity, Howell et al. reported an interesting observation. Tendons in juvenile mice can regenerate functional tissue after injury, but this ability is lost in mature mice, resulting in scar tissue formation (Howell et al., 2017). This finding suggests the possibility of abundant TSPCs in juvenile mouse tendons.

Given this background, we hypothesized that evaluating tendon tissue heterogeneity at the single-cell level using juvenile mouse tendons would enable a more detailed characterization of TSPCs. This approach is expected to advance the identification and characterization of TSPCs, which has been challenging with conventional methods.

Therefore, we performed scRNA-seq using cells collected from 2-and 6-week-old mouse Achilles tendons and investigated clusters that co-express known TSPCs markers, such as *Tppp3*, *Pdgfra*, and *Ly6a*. Then, we analyzed the expression of surface antigens in these clusters and identified *Cd55* and *Cd248* as novel candidate surface antigens for murine TSPCs.

Furthermore, snATAC-seq and snRNA-seq were performed simultaneously using cells collected from Achilles tendons to evaluate the validity of *Cd55* and *Cd248* as surface antigens for TSPCs and to identify the landscape of transcription factors (TFs) involved in tendon maturation. We sorted mouse Achilles tendon cells based on CD55 and CD248 and confirmed their phenotypes, demonstrating high clonogenicity and highly efficient induction into tendon cells. These results suggested that CD55 and CD248 are novel surface antigens of murine TSPCs and may be useful for understanding the process of tendon maturation and for applications in cell therapy.

Material and Methods

Single cell isolation

Mice were euthanized by exposure to CO_2 followed by cervical dislocation. To remove Achilles tendons, a longitudinal incision through the skin was made down the midline of the posterior aspect of the lower limb. A sharp transverse incision was made just distal to the myotendinous junction and again just proximal to the enthesis, and the Achilles tendon was carefully removed. The procedure was performed bilaterally. All animals were purchased from Sankyo Lab Service (Tokyo, Japan). Achilles tendons were digested to obtain a single-cell suspension. In total, 40 or 60 (2-week and 6-week, respectively) mouse Achilles tendons were processed together as a single sample. After washing with 1× phosphate-buffered saline (PBS) several times and cutting 1-mm-wide sections using scissors, tendons were digested for 1 h at 37°C with continuous shaking at 1,200 rpm and 37°C in a dissociation solution consisting of 30 mg/mL collagenase (Wako, Osaka, Japan). After digestion, the single-cell suspension was filtered for debris using a 40 μm cell strainer and washed twice with 1× PBS. The samples were centrifuged for 15 min at 1,500 rpm and 4°C. Dead cells were removed using a Dead Cell Removal Kit (Miltenyi Biotec, Bergisch Gladbach, Germany). Cell viability was 76% (2-week) and 73% (6-week), respectively. The experiment involving mice was approved by Animal Experimental Committee of the Institute of Science Tokyo.

scRNA-seq library construction and sequencing

Cells isolated from Achilles tendons of both 2-week-old and 6-week-old mice were processed separately. Cells were resuspended in PBS with 1% fetal bovine serum (FBS) at a concentration of 10,000 cells per μL. Then, 10,000 cells per sample were loaded on a Chromium Controller (10x Genomics, Pleasanton, CA, USA) for single cell capture.

Libraries were prepared using Single Cell 3' Library & Gel Bead Kit v3 (10x Genomics) following the manufacturer's instructions. A single-cell emulsion (Gel Bead-In-EMulsions, GEMs) was created by making barcoded cDNA unique to each individual emulsion. A recovery agent was added to break GEM and cDNA was then amplified. A library was produced via end repair, dA-tailing, adapter ligation, post-ligation cleanup with SPRIselect, and sample index PCR. The quality and concentration of the amplified cDNA were evaluated using the Bioanalyzer (Agilent 2100) on a High Sensitivity DNA chip (Agilent, #5065-4401; Santa Clara, CA, USA). Sequencing was performed using the HiSeq X system (Illumina, San Diego, CA, USA) to generate 28/90 bp paired-end reads.

Cell dissociation, nuclei isolation, and snRNA-seq and snATAC-seq library construction and sequencing

After obtaining cells from the mouse Achilles tendons of 2-week-old and 6-week-old mice (processed separately for each age group), nuclei were isolated following the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression Reagent Bundle (10XGenomics) and all of the buffers were made according to the manufacturer's instructions. Briefly, after centrifugation, the supernatant was discarded, and the cells were re-suspended in 1 mL of 1% FBS in PBS. Approximately 500,000 cells were transferred to a new tube for further lysis. To remove the supernatant, cells were centrifuged again for 5 min at 300 rcf and 4°C and resuspended in 100 mL of chilled lysis buffer. Cells were lysed for 9 min on ice and 1 mL of wash buffer was added to stop the reaction. Then, the suspension was filtered through a 40 µm cell strainer, cells were centrifuged and resuspended in 66.2 mL of chilled Diluted Nuclei Buffer aiming to target 7000 nuclei. The final concentration of nuclei was determined, followed by transposition, GEM generation, barcoding, and library construction according to the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression (10XGenomics). Libraries were sequenced with the parameters recommended by the manufacturer, using the NovaSeq 6000 (Illumina) to generate 28/90 bp paired-end reads for gene expression and 50/49 bp paired-end reads for ATAC sequencing.

scRNA-seq and snRNA-seq analyses

Sequencing reads were processed with the Cellranger_arc (10X Genomics, v2.0.0) using the mouse reference mm10. We performed separate analyses for scRNA-seq data (collected from both 2-week and 6-week samples) and snRNA-seq data (collected from both 2-week and 6-week samples) to enable comprehensive characterization of transcriptional landscapes. From the gene expression matrix, downstream analyses were carried out using R. Quality control, filtering, data clustering and visualization, and a differential expression analysis were carried out using Seurat (Butler et al., 2018). For each dataset, cells with unique feature counts over 2,500 or less than 200 and >5% mitochondrial counts were filtered. Then, heterotypic doublets (assuming 5% of barcodes represent doublets) were removed using DoubletFinder (McGinnis et al., 2019).

Unsupervised shared nearest neighbor (SNN) clustering was performed with varying resolution and the results were visualized using uniform manifold approximation and projection (UMAP) (Becht et al., 2019). Differentially expressed genes (DEGs) among each cell cluster were determined using the FindAllMarkers function in Seurat. The criteria of DEGs as a marker gene for the cluster was logFC > 0.25, adjusted $p < 0.05$, expression in >25% of cells. We then annotate each cluster based on DEGs (supplementary table 1-15). To ensure consistency between our different sequencing approaches, we compared cell clusters identified in the scRNA-seq analysis with those from the snRNA-seq analysis (Supplemental Figure 5).

Pseudotime analysis

Monocle3 (Trapnell et al., 2014) was used to convert the snRNA-seq dataset into a cell dataset object (CDS), preprocess data, correct for batch effects, embed with dimensional reduction, and perform pseudotemporal ordering. Cicero (Pliner et al., 2018) was used to generate pseudo-temporal trajectories for the snATAC-seq dataset.

Cell-cell communication analysis

Intercellular communication networks were quantitatively inferred and analyzed using scRNA-seq data. The R package CellChat (Jin et al., 2021) was used to visualize the interactions among different cell groups. Two hundred twenty-nine signaling pathway families were grouped as a library to analyze cell–cell communication.

GO analysis

The R package ClusterProfiler (Yu et al., 2012) was used to perform a gene enrichment analysis. The p-values were corrected by the Benjamini & Hochberg method.

Motif enrichment analysis (ChiP seeker)

Genomic regions containing snATAC-seq peaks were annotated using ChIPSeeker (Yu et al., 2015) and the UCSC database on mouse (mm10).

snATAC-seq analysis

The Cell Ranger ATAC pipeline (1.2.0) (Satpathy et al., 2019) was used to preprocess the data resulting from sequencing. First, Tn5 sites were mapped to the mouse reference transcriptome mm10, and duplicate reads and background cells were removed. This returned barcoded fragment files, which were loaded into Signac (Stuart et al., 2021) for downstream analyses using the standard Signac/Seurat pipeline. Macs2 (Zhang et al., 2008) was run on the fragment files to call peaks using the Signac 'CallPeaks' function.

Fragments were mapped to the Macs2 called peaks and assigned to cells using the Signac 'FeatureMatrix' function. Nucleosome signal strength and transcription start site (TSS) enrichment for each cell were calculated using the Signac 'NucleosomeSignal' and 'TSSEnrichment' functions, respectively. Outliers in the QC metric categories were removed as per Signac's standard processing guidelines. Latent semantic indexing (LSI), a form of dimensional reduction, was performed using the Signac 'RunTFIDF' and 'RunSVD' functions. The UMAP hyperparameters were varied to produce consistent object shapes (using R). Once hyperparameters were chosen, the Signac/Seurat's 'RunUMAP' function was run on the LSI dimensions chosen earlier for UMAP embedding. The Signac/Seurat 'FindNeighbors' function was run using the same LSI dimensions used for UMAP to compute the nearest neighbor graph. Signac/Seurat 'FindClusters' was then run at varying resolutions. A gene activity matrix was constructed by counting ATAC peaks within the gene body and 2 kb upstream of the transcriptional start site for protein-coding genes annotated in the Ensembl database. The gene activity matrix was log-normalized prior to label transfer with the aggregated snRNA-seq Seurat object using a canonical correlation analysis. The aggregated snATAC-seq object was filtered using a 97% confidence threshold for cell-type assignment following label transfer to remove heterotypic doublets. The filtered snATAC-seq object was reprocessed with TFIDF, SVD, and batch effect correction followed by clustering and annotation based on lineage-specific gene activity. Differential chromatin accessibility between cell types was assessed with the Signac FindMarkers function for peaks detected in at least 20% of cells using a likelihood ratio test and a log-fold-change threshold of 0.25. Bonferroni-adjusted p-values were used to determine significance at an FDR < 0.05.

Integration of snRNA-seq and snATAC-seq data

snRNA-seq and snATAC-seq data were integrated using the cluster-label transfer procedure as implemented in Signac and Seurat. Each snRNA-seq sample. as clustered individually and its cluster labels were projected onto the matching, individually clustered snATAC-seq sample or vice versa. Anchors were identified for condition-matched snRNA-and snATAC-seq samples using the FindTransferAnchors function and a canonical correlation analysis (CCA) was performed using the

snRNA expression values and the snATAC-imputed gene expression values. The anchors were used to transfer cluster-label identifiers between the two data types using the TransferData function. Each cell in the query was assigned the cluster label with the highest prediction score, and label transfer was considered successful for query cells with prediction scores above 0.3.

SCENIC analysis

To identify TFs and characterize cell states, a cis-regulatory analysis was performed using the R package SCENIC (Lake et al., 2018), which infers gene regulatory networks based on co-expression and DNA motif analyses. The network activity was then analyzed for each cell to identify recurrent cellular states. In short, TFs were identified using GENIE3 and compiled into modules (regulons), which were subsequently subjected to cis-regulatory motif analysis using RcisTarget with two gene-motif rankings: 10 kb around the TSS and 500 bp upstream. Regulon activity in every cell was then scored using AUCell.

Immunohistochemistry

Immunohistochemistry was performed as previously described (Tsutsumi et al., 2022). Briefly, tissue samples were fixed in 4% paraformaldehyde overnight at 4°C and embedded in paraffin. Sections of 10 μm in thickness were deparaffinized and activated with citric acid buffer (10 mM sodium citrate, 1 mM EDTA; pH 6.0) at 80 °C for 60 min in a decloaking chamber NxGen (BIOCARE Medical, CA, USA) or with 0.1% trypsin/PBS at 37 °C for 30 min. After they were blocked with Blocking One solution (Nacalai tesque, Kyoto, Japan) for 60 min, they were incubated with rabbit anti-CD55 antibody (1:100; A13918, ABclonal, Woburn, MA, USA) or rat anti-CD248 (1:100; MAB7535, R&D Systems, Minneapolis, MN, USA) overnight at 4 °C. Following this step, they were incubated with Alexa Fluor plus 488 goat anti-rabbit antibody (1:1000; A32731, Thermo Fisher Scientific, Waltham, MA, USA) or Alexa Fluor plus 488 donkey anti-rat antibody (1:1000; A21208, Thermo Fisher Scientific) for 60 min. Hoechst 33342 (100ng/ml, Thermo Fisher Scientific) was added during this incubation. The sections were then mounted with ProLong Glass Antifade Mountant (P36980, Thermo Fisher Scientific).

Fluorescence-activated cell sorting (FACS)

Cells harvested from the Achilles tendon were incubated for 60 min on ice with APC-CD55 (#131812; BioLegend, San Diego, CA, USA) and a primary antibody against CD248 (#LS-B2712, LSBio) using a FACS buffer (PBS containing 1 [v/v%] FBS). After washing, cells were reacted with fluorescent-conjugated rabbit secondary antibody. Cell sorting was performed using MoFlo XDP FACS (Beckman Coulter, Brea, CA, USA). DAPI was used as the live/dead discriminator to the gate. CD55/CD248 dual positive and negative cell cluster were sorted.

Primary cell culture

Primary TSPCs were isolated from the Achilles tendons of 2-week-old mice with collagenase digestion described above. Single-cell suspensions were cultured in the culture medium (MEMα + 20% FBS + 1% penicillin-streptomycin + 1 [v/v%] 100× non-essential amino acid solution [NEAA; Gibco], 1 [v/v%] 100× GlutaMAX [Gibco]). At 80%–90% confluence, cells were trypsinized, centrifuged, resuspended in culture medium as passage 1 cells, and incubated in 5% CO_2 at 37°C, with fresh medium every 2–3 days.

Colony formation assay

For the colony formation assay, single-cell suspensions of TSPCs (1000 cells/well) were seeded and incubated in six-well plates for 12–14 days in the growth medium and fixed with 4% paraformaldehyde (PFA) (Sigma-Aldrich, St. Louis, MO, USA). Then, 0.1% crystal violet solution

(Wako) was used to stain the cells. Colonies of >30–50 cells were defined as a single colony unit, and the number of clusters was counted using the ImageJ package ColonyArea (Guzmán et al., 2014)

RT-PCR

Total RNA was purified using TRIzol reagent (Invitrogen, Grand Island, NY, USA). Reverse transcription of mRNA was carried out using the PrimeScript RT Reagent Kit (Takara, Shiga, Japan). Quantitative PCR (qRT-PCR) was performed on cDNA with the Thunderbird SYBR mix (Toyobo, Osaka, Japan). B2m was used as a reference gene, and relative gene expression levels were calculated through the ΔCT method.

Tenocytes, cartilage, and osteocyte differentiation

For the differentiation experiment, TSPCs were cultured in 6-well plates (50,000 cells/well). Osteogenic, chondrogenic, and tenogenic differentiation were induced using a corresponding differentiation medium. The osteogenic differentiation medium contained a culture medium supplemented with 10 nM dexamethasone (Sigma-Aldrich), 5 mM β-glycerophosphate (APEXBIO), and 0.05 mM L-ascorbic acid 2-phosphate (Sigma-Aldrich). The chondrogenic differentiation medium contained a culture medium supplemented with 100 nM dexamethasone and 10 ng/mL BMP2 (Sigma-Aldrich). The tenogenic differentiation medium contained a culture medium supplemented with 10 ng/mL TGF-β1 (Peprotech, Rocky Hill, NJ, USA), 10 ng/mL GDF-5 (R&D Systems, Minneapolis, MN, USA), and 0.05 mM L-ascorbic acid 2-phosphate. After a 2 week induction period, cells were harvested for RT-PCR.

Bio-cultured tendon (Bio-tendon)

Generation of Bio-tendon has reported previously (Kataoka et al., 2020; Tsutsumi et al., 2022) Briefly, sorted cells were embedded in a 3D-culture cocktail, which is consistent with 2 mg/mL Cellmatrix (Type I-A, Nitta Gelatin Inc., Osaka, Japan), pro-survival cocktail (final concentrations: 100 nM Bcl-Xl BH4 4-23 (Merck Millipore, Burlington, MA, USA), 100 μM carbobenzoxy-valyl-alanyl-aspartyl-[*O*-methyl]-fluoromethylketone (Z-VAD-FMK) (Promega, Madison, WI, USA)) in culture medium. The 3D chamber was coated with 1% gelatin and incubated at 37°C and 5% CO_2 for 30 min. After washing with 1× PBS 3 times, 1.0×10^6 cells were transferred to a 3D-culture cocktail mixture and incubated at 37°C and 5% CO_2 for 60 min for gelation. Following gelation, tenogenic differentiation medium was added to the chamber. Following 24 h of incubation, the chamber was placed within a cell stretching system (Shellpa Pro, Menicon Life Science, Aichi, Japan). Cyclic mechanical stretch was applied for 1 week, with gradual increase in stretch load: 1% on day 1, 2% on day 2, 3% on day 3, 4% on day 4, and 5% from days 5 to 7. The stretching cycle was programed at 0.25 Hz for 18 h/day, followed by resting for 6 h/day at 37°C and 5% CO_2. Daily medium changes were also required.

Decellularization by HHP and chemical treatment

The cultured bio-tendon was placed into a plastic pack filled with saline and sealed to prevent implosion and leakage during the procedure. The pack was then pressurized at 1,000 MPa at 30°C for 10 min using an HHP machine (Dr. CHEF; Kobe Steel, Hyogo, Japan). After pressurization, the bio-tendon was washed thrice with 30% ethanol (EtOH) by continuous shaking for 5 min at each step. Finally, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC)/*N*-hydroxysuccinimide (NHS)-based cross-linking was performed by adding 70 mM EDC (Wako) and 70 mM NHS (Wako) with 30% EtOH for 24 h at 4°C. The processed bio-tendon was incubated in 1× PBS at 4°C until the experiment.

Electron microscopy

Bio-tendon tissues were dissected and fixed overnight in 2.5% glutaraldehyde in 0.1 M phosphate buffer (PB). For transmission electron microscopy (TEM), the specimens (n = 3) were rinsed with 0.1 M PB, post-fixed in 1% osmium buffered with 0.1 M PB for 2 h, and dehydrated in a graded ethanol series. The specimens were then embedded in Epon 812, sectioned into ultrathin sections (70 nm), mounted on copper grids, and double-stained with uranyl acetate and lead citrate. TEM observation was conducted JEM-1400Flash (JEOL, Tokyo, Japan). For scanning electron microscopy (SEM), the specimens (n = 2) were dried in a critical-point dryer (HCP-2; Hitachi, Tokyo, Japan) with liquid CO_2 and coated with platinum. The specimens were subsequently observed under SEM (JSM-7900F/JED-2300; JEOL). Fiber orientation within the microscopic sections was analyze using the OrientationJ image processing tool, a plugin for ImageJ.

Stretch test

The mechanical properties were measured using a creep meter (RE-3305S; Yamaden, Tokyo, Japan). After measuring the initial length (mm), diameter (mm), and thickness (mm) using a micrometer, the samples were fixed with two grips, which were pulled at a constant speed of 0.05 mm/s until failure, and the tensile strength (N) and failure strain (mm) were measured. The cross-sectional area (mm^2) was calculated using the initial diameter and thickness. The stiffness was manually determined from the slope in the linear region of the failure-stress curve.

The tensile strength (MPa), failure strain (%), stiffness, and Young's modulus were calculated using the following formulas:

Tensile strength (MPa) = tensile strength (N)/cross-sectional area (mm^2). Failure strain (%) = failure strain (mm)/initial length (mm).

Stiffness = Stress (N)/Strain (mm)

Young's modulus = stiffness × initial length (mm)/cross-sectional area (mm^2)

Statistical analysis

All values are presented as means ± SEM. Statistically significant differences were assessed using unpaired two-tailed Student's t-tests and one-way analysis of variance (ANOVA) with Tukey's post hoc tests. Statistical significance was set at $p < 0.05$.

Results

Identification of Novel Surface Antigens for Tendon Stem/Progenitor Cells (TSPCs)

In mice, the healing capacity for tendon injuries is high up to 2 weeks of age, but decreases as the musculoskeletal system matures at 6 weeks (Howell et al., 2017). To investigate whether this phenomenon is due to changes in the cellular population of tendon tissue, including fluctuations in progenitor cells, we performed single-cell RNA sequencing on mouse Achilles tendons at these two time points.

Achilles tendons were harvested from mice at 2 and 6 weeks of age. Following collagenase digestion and dead cell removal, we employed massively parallel, droplet-enabled scRNA-seq analysis (10X Genomics Chromium). Data processing was conducted using the Cell Ranger pipeline (10X Genomics). We analyzed 10,314 cells from 2-week-old mice (median 3,167 genes/cell and

41,611 mean reads/cell) and 6,513 cells from 6-week-old mice (median 732 genes/cell and 60,386 mean reads/cell). After doublet removal using Doubletfinder, we merged the datasets and performed unbiased clustering using Seurat, identifying a total of 15 clusters (**Fig. 1A**). The top differentially expressed genes (DEGs) for each cluster, based on log2 fold change and statistical significance, are summarized in **Fig. 1B**.

We identified two clusters (0 and 11) with high expression of *Mkx* and *Scx*, transcription factors involved in tendon development and growth. Cluster 0 showed higher expression of ECM-related genes such as *Col1a1* and *Fmod* compared to cluster 11, suggesting these were tenocytes (TC). Cluster 11, with high *Periostin* (*Postn)* expression (Jacobson, et al., 2020), was classified as myotendinous junction (MTJ). Other clusters were characterized as follows: cartilage (CA, high *SRY-box transcription factor 9* (*Sox9*) expression), lymphocytes (LC1, LC2, high *Lysozyme 2* (*Lyz2*) and *Complement component 1, q subcomponent, alpha polypeptide* (*C1qa*)), endothelial cells (EC, high *Platelet and endothelial cell adhesion molecule 1* (*Pecam1*)), red blood cells (RBC, high *Hemoglobin alpha, adult chain 1* (*Hba-a1*)), smooth muscle cells (SM, high *Myosin light chain 9* (*Myl9*)), proliferating cells (PC, high *Mki67* and *Stathmin 1* (*Stmn1*)), Schwann cells (SW, high *Myelin basic protein* (*Mbp*)), vascular endothelial cells (VEC, high *Multimerin 1* (*Mmrn1*)), and macrophages (MC, high *Protein tyrosine phosphatase receptor type C* (*Ptprc*)). In cluster3, *Col1a1* and *Fmod* were expressed; however, *Gm42418* and *AY036118* were highly expressed. These long non-coding RNAs are related to RN45s and therefore cluster3 represents ribosomal contamination (Isola et al., 2024). It is also speculated that the cartilage cluster (CA) reflects enthesis, with expression of *Scx* as well as *Sox9* (Zhan et al., 2023). DEG s of each cluster was summarized in supplementary table 1-15.

To identify progenitor populations within these clusters, we analyzed expression patterns of previously reported markers *Tppp3* and *Pdgfra* (Harvey et al., 2019; Tachibana, et al., 2022), along with the known stem/progenitor cell marker *Ly6a* (Holmes et al., 2007; Sung et al., 2008; Hittinger et al., 2013; Sidney et al., 2014; Fang et al., 2022) (Supplemental Figure 1B). We identified subclusters within clusters 1 and 4 showing high expression of these genes, which we defined as SP1 and SP2. SP2 exhibited the highest expression of these genes, suggesting it had the strongest progenitor characteristics. It has been reported (Harvey et al., 2019) that the *Tppp3*-positive population is localized to peritenon, and SP clusters might reflect peritenon as well.

SP2 also showed strong expression of genes associated with early tendon development in mouse embryos, as identified by RNA-seq of mouse limb tendon cells at E14.5 (Havis et al., 2014) and the Eurexpress mouse embryo transcriptome atlas database (Diez-Roux et al., 2011) (Supplemental Figure 2A). Gene Ontology analysis revealed upregulation of tendon development-related pathways in SP2, including TGFβ production and collagen-containing extracellular matrix (Supplemental Figure 2B). Evaluation of cell numbers in each cluster showed a decrease in SP2, TC, and MTJ at 6 weeks (**Fig. 1C**, Supplemental Figure 1A).

We then examined DEGs in SP2 compared to SP1 (**Fig. 1D**). Whole list of DEGs comparing SP2 and SP1 was summarized in supplementary table16. Among these, we identified *Cd34*, *Cd55*, and *Cd248* as genes encoding surface antigens. *Cd34* is known to be highly expressed in mouse embryonic limb bud E14.5 compared to E11.5 (Havis et al., 2014), suggesting that *Cd34* positive cluster might reflect a progenitor cell that consist of the limb bud, including tendons. However, *Cd55* and *Cd248* have not been discussed with this context. Expression of *Cd55* and *Cd248* was localized to areas corresponding to SP1 and SP2 (**Fig. 1E**). We evaluated the correlation coefficients between the expression of these genes and previously suggested TSPC markers (*Cd73*, *Cd90*, *Cd105*, *Cd44*, *Cd146*) as well as *Tppp3*, *Pdgfra*, and *Ly6a* at 2 weeks of age. *Cd34*, *Cd55*, and *Cd248* showed high correlation, suggesting these genes as new candidate surface antigens for TSPCs (**Fig. 1F**).

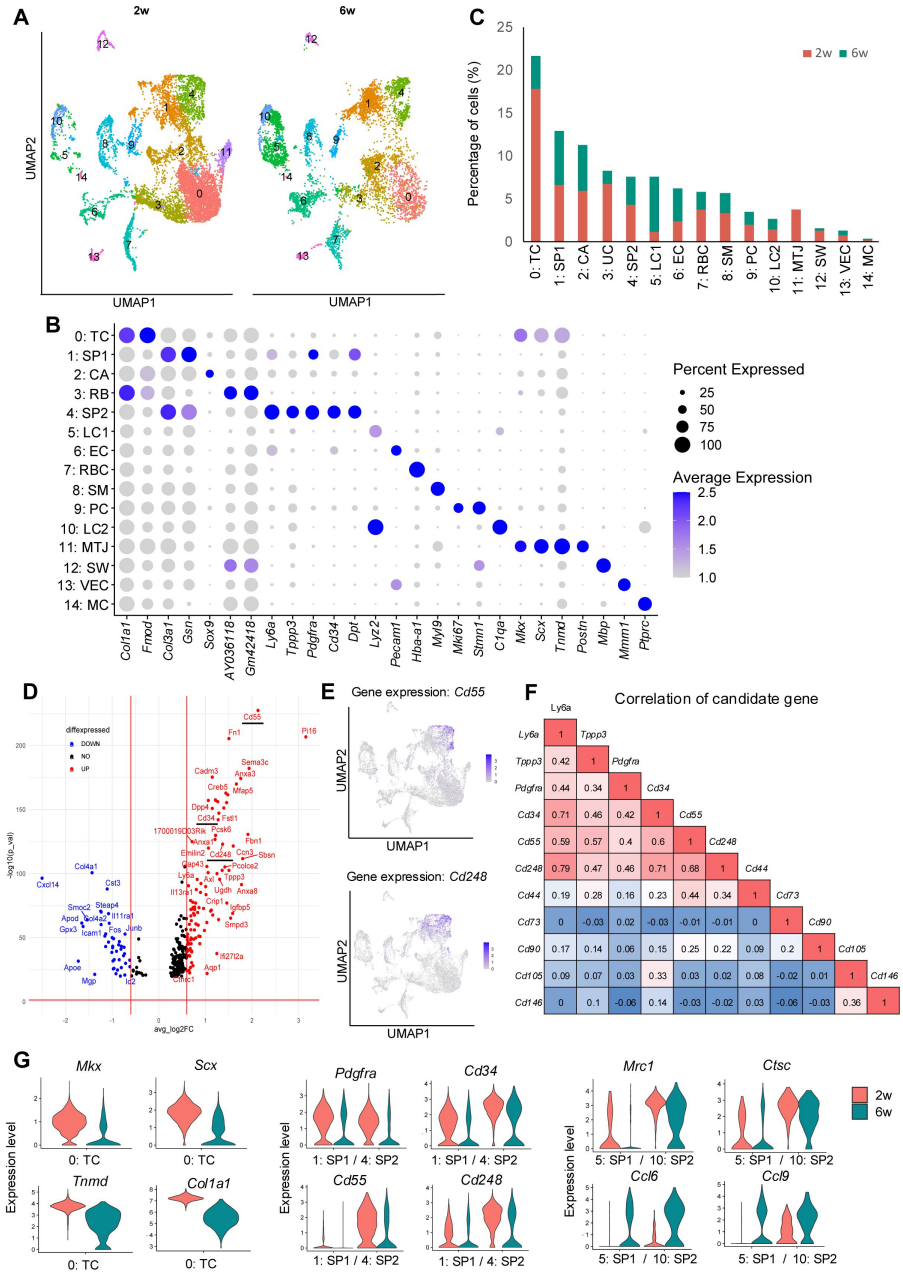


Figure 1.

scRNA-seq of tendon cells from 2-week-old and 6-week-old mice and the identification of surface markers of TSPC.

(A) Integrated uniform manifold approximation and projection (UMAP) scRNA-seq clustering of cells harvested from 2-week-old and 6-week-old mouse Achilles tendons. (B) Dot plot of average gene expression levels of the indicated genes in each scRNA-seq cluster. The size of the dot reflects the percentage of cells in the cluster that express each gene. TC, tenocyte; SP1, tendon stem/progenitor cell_1; CA, cartilage; RB, ribosomal RNA; SP2, tendon stem/progenitor cell_2; LC1, lymphocyte_1; EC, endothelial cell; RBC, red blood cell; SM, smooth muscle cell; PC, proliferating cell; LC2, lymohocytes_2; MTJ, myotendinous junction cell; SW, Schwann cell; VEC, vascular endothelial cell; MC, macrophage. (C) Proportions of cells in clusters identified from scRNA-seq. Clusters are colored according to cluster type. (D) Volcano plot of gene expression in the SP2 cluster and the identification of candidate TSPC marker genes (red under line). (E) Feature plot of *Cd55* and *Cd248* expression. (F) Correlation of gene expression of TSPC candidate genes in 2-week data. (G) Violin plots presenting the gene expression changes for a selection of differentially expressed genes.

Comparing gene expression between 2 and 6 weeks, we observed a decrease in *Mkx* and *Scx* expression, consistent with previous reports (Grinstein et al., 2019). Furthermore, the expression of our candidate tendon progenitor cell markers *Cd34*, *Cd55*, and *Cd248* all decreased at 6 weeks. Inflammatory cells, including macrophages, increased at 6 weeks (**Fig. 1G**). Gene expression analysis revealed a shift from M2 macrophage-related genes (*Mannose receptor C-type 1* (*Mrc1*) and *Cathepsin C* (*Ctsc*)) at 2 weeks to M1 macrophage-related genes (*Chemokine (C-C motif) ligand 6* (*Ccl6*) and *Chemokine (C-C motif) ligand 9* (*Ccl9*)) at 6 weeks. Comparative analysis between 2-week and 6-week tendons was summarized as a heatmap in supplemental Figure 3.

Analysis of cell-cell communication changes in postnatal tendon and surrounding tissues (Lui et al., 2019) using CellChat (Jin et al., 2021) showed a significant decrease in overall interaction at 6 weeks compared to 2 weeks. These results demonstrate that the cellular profile of tendon tissue changes dramatically at 6 weeks, with decreased expression of stem progenitor markers and reduced interaction with surrounding tissues.

Single Nucleus RNA + ATAC Analysis of 2-Week-Old Mouse Achilles Tendon Cells

Gene regulation is mediated by the binding of transcription factors to cis-regulatory elements proximal to the gene. Consequently, epigenetic changes such as chromatin accessibility play a crucial role in gene expression (Miller et al., 2013). Moreover, transcription factors are typically expressed at low levels, potentially leading to false negatives in scRNA-seq due to detection limits. Chromatin accessibility changes often precede gene expression changes, potentially allowing us to predict transcriptional changes (Ranzoni et al., 2020). Therefore, we performed simultaneous ATAC-seq and RNA-seq at the single-cell level to identify gene expression changes and their associated epigenetic alterations, aiming to elucidate the post-natal growth mechanisms of tendon tissue and evaluate the validity of *Cd55* and *Cd248* as markers.

We extracted nuclei from cells isolated from 2-week-old mouse Achilles tendons and conducted droplet-enabled multi-omics analysis (10X Genomics Chromium), performing simultaneous snATAC-seq and snRNA-seq as a reference. We first analyzed the snATAC-seq data. After mapping reads to the genome and calling peaks, we annotated the location of peaks in terms of genomic features. The peaks were associated with promoters, introns, exons, and intergenic regions.

We evaluated 6,571 cells (median high-quality fragments per cell: 17,408) and clustered them using latent semantic indexing and Uniform Manifold Approximation and Projection (UMAP) with the R package Signac (Stuart et al., 2021) (**Figure 2A**). Given the limited knowledge of cell-specific chromatin accessibility, we assessed gene activity by computing counts per cell within the gene body and promoter of protein-coding genes. Using this gene activity data, we performed unsupervised clustering and identified 17 clusters (ground-truth annotation). Clusters expressing tendon-related genes such as *Mkx* and *Scx* were identified as clusters 1 and 6, defined as A1: TC1 and A6: TC2, respectively. Clusters 0, 14, and 15 were identified as expressing stem progenitor markers including *Cd34*, *Ly6a*, *Pdgfra*, *Cd55*, and *Cd248*, and were defined as A2-0: SP1, A2-14: SP2, and A2-15: SP3, respectively. The characteristic gene activities and annotations for each cluster are summarized in **Fig. 2B**.

Next, we analyzed the snRNA-seq data. We evaluated 10,314 cells (median 3,167 genes/cell and 41,611 mean reads/cell). After doublet removal, we performed unbiased clustering using Seurat and identified 17 clusters (Supplemental Figure 4A). Clusters expressing tendon-related genes *Mkx* and *Scx* were clusters 1 and 13, defined as R2-1: TC_1 and R2-13: TC_2, respectively. Clusters 2, 3, and 7 were identified as expressing stem/progenitor-related genes and defined as R2-2: SP_1, R2-3: SP_2, and R2-7: SP_3, respectively. To validate the consistency of our approaches, we conducted a comprehensive comparison between the cell clusters identified in our scRNA-seq analysis (which included both 2-week and 6-week samples) and our snRNA-seq analysis. We confirmed that all

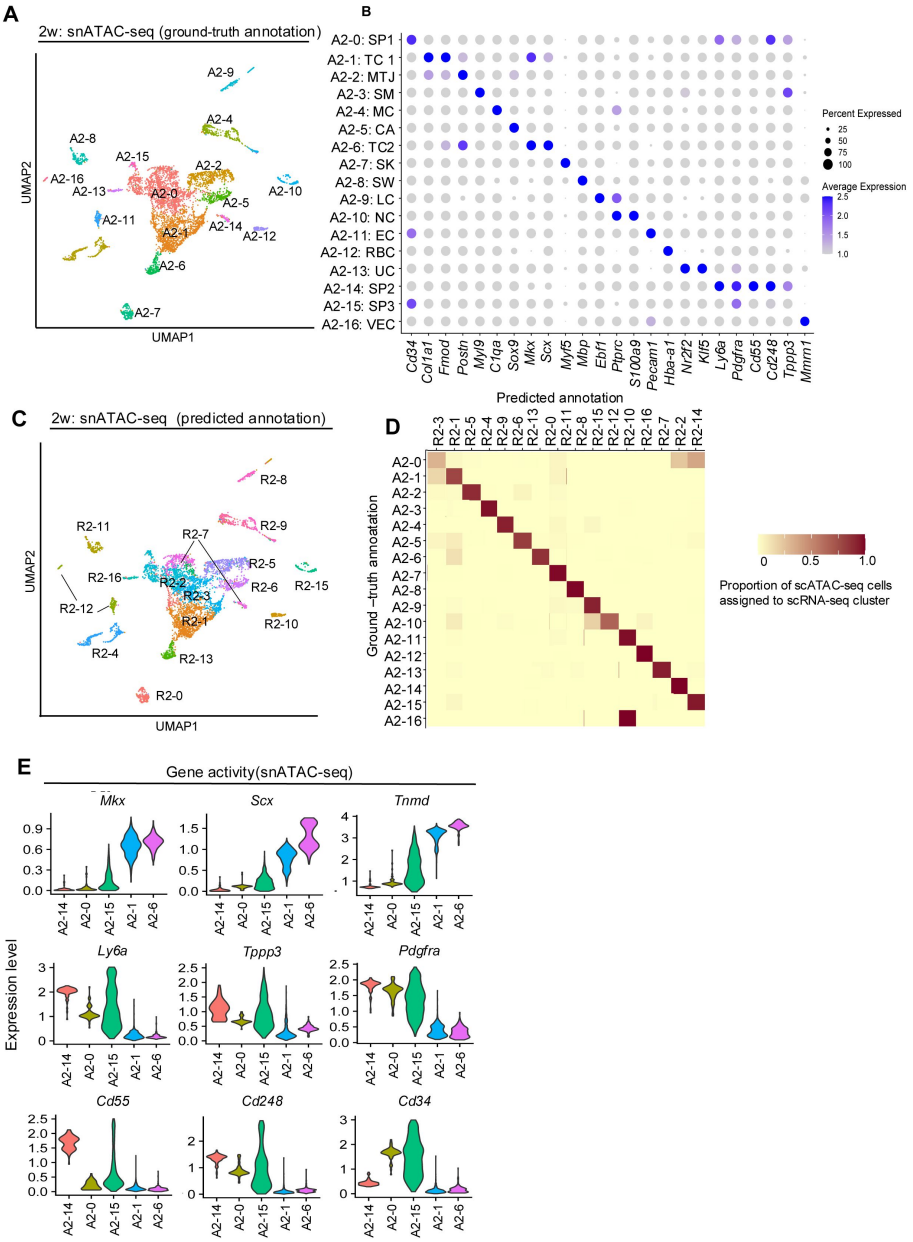


Figure 2.

snATAC-seq of tendon cells from a 2-week-old mouse and the validation of Cd55 and Cd248 as candidate markers of TSPC.

(A) UMAP snATAC-seq clustering of cells derived from the Achilles tendon of a 2-week-old mouse. Annotation was based on each gene activity (ground-truth annotation). SP1, tendon stem/progenitor cell_1; TC1, tenocyte_1; MTJ, myotendinous junction cell; SM, smooth muscle cell; MC, macrophage; CA, cartilage; TC2, tenocyte_2; SK, skeletal muscle cell; SW, Schwann cell; LC, lymphocyte; NC, neutrophil; EC, endothelial cell; RBC, red blood cell; UC, unspecified cell; SP2, tendon stem/progenitor cell_2; SP3, tendon stem/progenitor cell_3; VEC, vascular endothelial cell. (B) Dot plot of average gene activity of the indicated genes in each snATAC-seq cluster. The size of the dot reflects the percentage of cells in the cluster that express each gene. (C) UMAP visualization and predicted annotation of 2-week snATAC-seq after integration and label transfer of 2-week snRNA-seq data. (D) Identification of matching cell clusters between the 2-week snRNA-and 2-week snATAC-seq data from visualized as heatmap. The heatmap shows the proportions of cells from each snATAC-seq cluster across all sample conditions assigned to each snRNA-seq cluster as part of the label-transfer process. (E) Violin plot of tenocytes and TSPC-related gene expression in each cluster.

clusters identified in the scRNA-seq data could be similarly annotated in the snRNA-seq data, demonstrating strong concordance between these complementary approaches (Supplemental Figure 5). This validation supports the reliability of our cell type identification across different single-cell methodologies. Interestingly, R2-5: The MTJ cluster expressed not only *Tnc* but also *Sox9*, consistent with previous reports (Nagakura et al., 2020). Using the Signac package's "FindTransferAnchors" function, we calculated predicted IDs for snATAC-seq clusters based on snRNA-seq annotations (predicted annotation) (**Fig. 2C**). Evaluation of the two annotation methods for snATAC-seq revealed that both methods allowed annotation of major cell types, and correlation between the two was maintained. Thus, the validity of predicted annotations based on snRNA-seq was confirmed (**Figure 2D**).

R2-7: SP3, considered the most undifferentiated stem/progenitor fraction in snRNA-seq, could be divided into A2-0: SP1 and A2-14: SP2 in snATAC-seq.

Comparing gene activity of tendon-related and stem/progenitor genes in each cluster of the ground truth annotation, A2-14: SP2 showed high expression of *Ly6a*, *Pdgfra*, and *Tppp3*, and inverse correlation with *Mkx*, *Scx*, and *Tnmd*. Furthermore, the newly identified surface antigens *Cd55* and *Cd248* also showed high expression in cluster 14. In contrast, *Cd34* did not show a high expression pattern in A2-14: SP2. These results suggest that A2-14: SP2 is the most immature cluster, with *Cd55* and *Cd248* showing characteristic expression (**Fig. 2E**). Subcluster analysis of R7, considered the most immature fraction in snRNA-seq, revealed two clusters based on *Cd55*/*Cd248* and *Cd34* expression, similar to the snATAC-seq analysis. The former (R7-1) showed high expression of *Ly6a*, *Tppp3*, and *Pdgfra* (**Fig. 2F**). Trajectory analysis using R2-7 and A2-14 as roots in snRNA-seq and snATAC-seq, respectively, showed increased expression of tendon-related genes *Mkx* and *Scx* along the trajectory, while expression of *Tppp3*, *Cd55*, and *Cd248* decreased (**Figure 3A**). Evaluation of peaks in each cluster based on gene expression levels in snRNA-seq revealed multiple peaks with heights correlated to gene expression for *Mkx* and *Scx*. Similar evaluation of *Cd55* and *Cd248* showed multiple peaks upstream of the TSS for *Cd248* (**Fig. 3B**). These results indicate correlation between snRNA-seq and snATAC-seq, and consistent with scRNA-seq results, suggest that A2-14 represents the most undifferentiated TSPCs based on *Cd55* and *Cd248* expression.

Comparison of Single Nucleus RNA+ATAC Analysis between 2-week and 6-week Mouse Achilles Tendon Cells

To confirm the presence of the identified *Cd55*/*Cd248* fraction at 6 weeks, we performed snRNA-seq and snATAC-seq on 6-week-old mouse Achilles tendon cells. Using the 2-week snATAC-seq data as a reference, we annotated the 6-week snATAC-seq clusters. Clusters A2-14 and A2-0 corresponded to clusters A6-12 and A6-0, respectively (Supplemental Figure 6). We also annotated the 6-week snATAC-seq clusters using 2-week snRNA-seq with the same results (Supplemental Figure 7). Gene activity evaluation revealed that, similar to the 2-week data, the 6-week A6-12 and A6-0 clusters showed *Cd55*/*Cd248* high and *Cd34* dim, and Cd55/*Cd248* dim and *Cd34* high patterns, respectively. Furthermore, annotation of the 6-week snRNA-seq data using the 2-week snRNA-seq as a reference revealed clusters with expression patterns similar to those observed in snATAC-seq. In both cases, clusters with high *Cd55* and *Cd248* expression showed high gene activity (snATAC-seq) and gene expression (snRNA-seq) of *Tppp3*, *Pdgfra*, and *Ly6a*. These results confirm that the *Cd55* and *Cd248* high clusters identified in the 2-week snRNA-seq+snATAC-seq analysis are similarly detected at 6 weeks.

We also compared the snATAC-seq data between 2 and 6 weeks. No significant differences were observed in genomic annotations between the two time points (Supplemental Figure 8A). After merging the datasets (Supplemental Figure 8B), we compared gene activity of tendon/stem-related

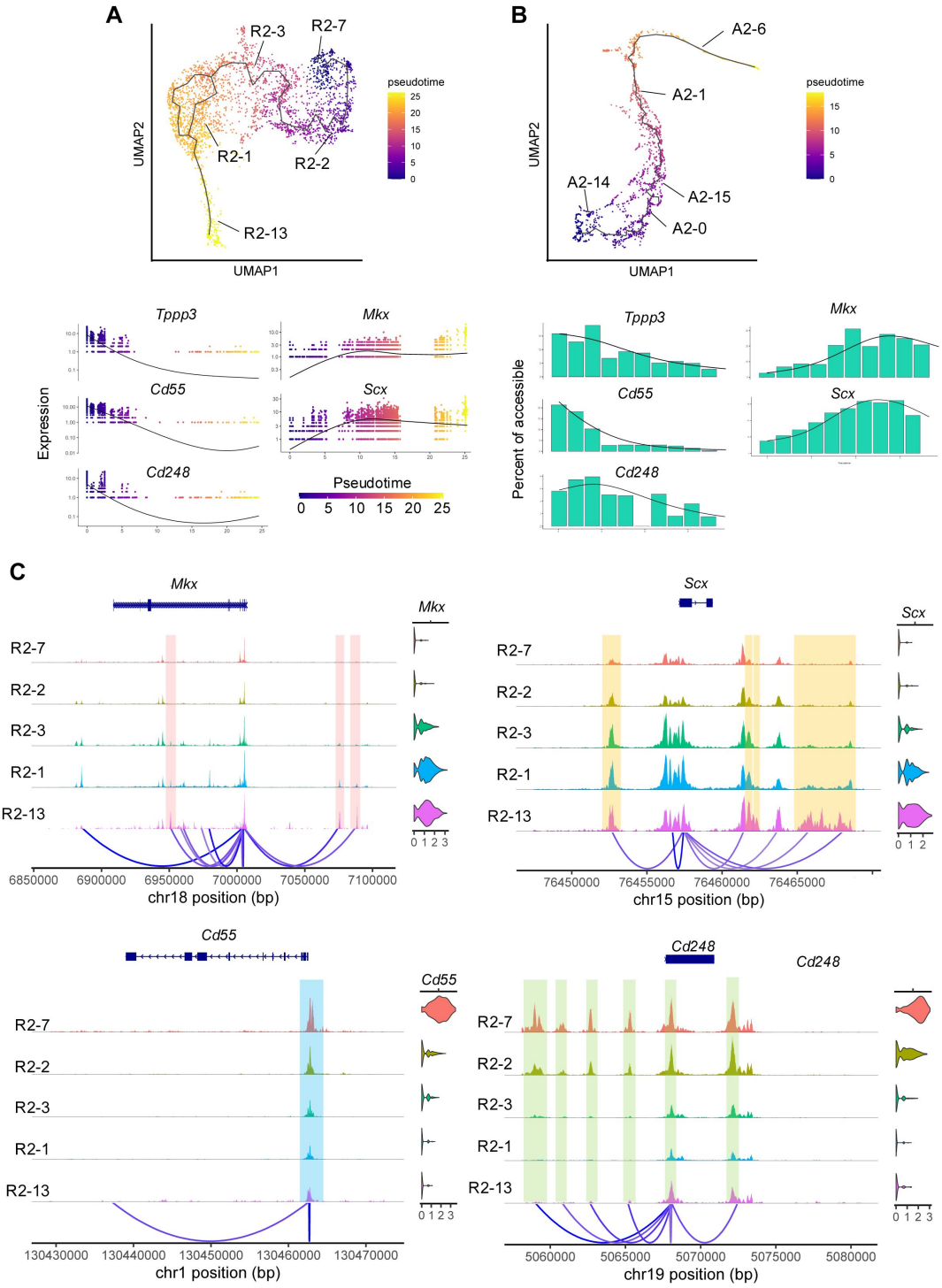


Figure 3.

Trajectory analysis and peak visualization of snRNA-seq and snATAC-seq data for the tendon and tendon stem/progenitor cell-related cluster.

(A) UMAP representation of snRNA-seq differentiation trajectory of tenocytes and TSPC lineage and pseudotime-dependent gene expression changes of *Tppp3*, *Cd55*, *Cd248*, *Mkx,* and *Scx,* as inferred using Monocle3. (B) UMAP representation of snATAC-seq differentiation trajectory of tenocytes and the TSPC lineage and pseudotime-dependent gene expression changes, as inferred using Cicero. (C) Coverage plots of *Mkx*, *Scx*, *Cd55,* and *Cd248*. Selected peaks that differ across each cluster are highlighted.

genes to evaluate changes in gene activity. We found that the activity of *Tppp3*, *Pdgfra*, *Ly6a*, *Cd55*, *Cd248*, and *Cd34* all decreased at 6 weeks (Supplemental Fig 8C). This was consistent with the decreased gene expression observed when comparing 2-week and 6-week data.

Estimation of Transcription Factor Activity in 2-week Mouse Achilles Tendon Cells

To estimate transcription factor activity in each cluster, we used the Single-Cell Regulatory Network Inference and Clustering (SCENIC) package (Lake et al., 2018) to calculate gene regulatory network activity from scRNA-seq gene expression data. SCENIC constructs gene expression networks centered on transcription factors and infers transcription factor activity in each cluster. We summarized the predicted transcription factor activities in tendon and stem/progenitor-related clusters R2-7, R2-2, R2-3, R2-1, and R2-13 identified by snRNA-seq (**Figure 4**).

Next, we evaluated the gene activity of these identified transcription factors in each snATAC-seq cluster, along with motif activity calculated by chromVAR (Schep et al., 2017). We also assessed the expression levels of transcription factors in each snRNA-seq cluster. In the most immature fractions, represented by cluster A2-14 in snATAC-seq and cluster R2-7 in snRNA-seq, *KLF transcription factor 3* (*Klf3*), *KLF transcription factor 4* (*Klf4*), and *cAMP responsive element binding protein 5* (*Creb5*) showed consistent behavior in gene activity, motif activity, and gene expression. Additionally, *Signal transducer and activator of transcription 2* (*Stat2*) and cAMP responsive element binding protein 3 like 1 (*Creb3l1*) showed high values in A2-0/A2-15 and A2-1, respectively. *Stat2* gene expression was not observed, likely due to false positives resulting from low expression levels. The gene activity and motif activity of each transcription factor estimated by snATAC-seq correlated with the transcription factor activity calculated by SCENIC. Therefore, simultaneous analysis of snRNA-seq and snATAC-seq could be used to more accurately evaluate the function of transcription factors in each cluster.

Furthermore, SCENIC can predict candidate transcription factors regulating each gene (**Table 1**). *Cd55* and *Cd248* were predicted to be regulated by *Klf3* and *Klf4*, while *Mkx* and *Scx* were predicted to be under the control of *Creb3l1*. These predictions were consistent with the data calculated from gene activity and motif activity in snATAC-seq.

In vitro Evaluation of CD55 and CD248

Immunostaining evaluation of CD55 and CD248, the identified candidate stem/progenitor markers, revealed expression in the peritenon (**Figure 5C**, 5A). This was consistent with previous reports of localized expression of *Tppp3*/*Pdgfra*-positive cells in the peritenon and our analysis showing co-expression of *Tppp3*/*Pdgfra* and *Cd55*/*Cd248*. The results support the possibility that SP clusters reflect peritenon. Given that *Cd55* and *Cd248* expression appeared to reflect *Tppp3*, *Pdgfra*, and *Ly6a* expression more sensitively than *Cd34*, we extracted tendon cells from 2-week-old mice and sorted them by FACS to determine the biological phenotype of Cd55 and Cd248 positive cells (**Figure 5B**).

Gene expression analysis of sorted cells showed that CD55/CD248 positive cells, compared to negative cells, had lower expression of *Mkx*, *Scx*, *Col1a1*, and *Creb3l1*, but higher expression of *Ly6a*, *Tppp3*, *Pdgfra*, *Creb5*, and *Klf3*, consistent with snRNA-seq analysis (Supplemental Figure 9). To assess stem/progenitor capacity, we performed colony formation assays. CD55/CD248 sorted cells showed significantly increased colony formation compared to CD55/CD248 negative cells (control) (**Figure 5C**).

We then evaluated the differentiation tendencies of these cells towards tendon, cartilage, and bone. Tendon differentiation resulted in clusters of spindle-shaped cells from CD55/CD248 positive cells, suggesting tenogenic differentiation (**Figure 5D**). Gene expression analysis revealed that

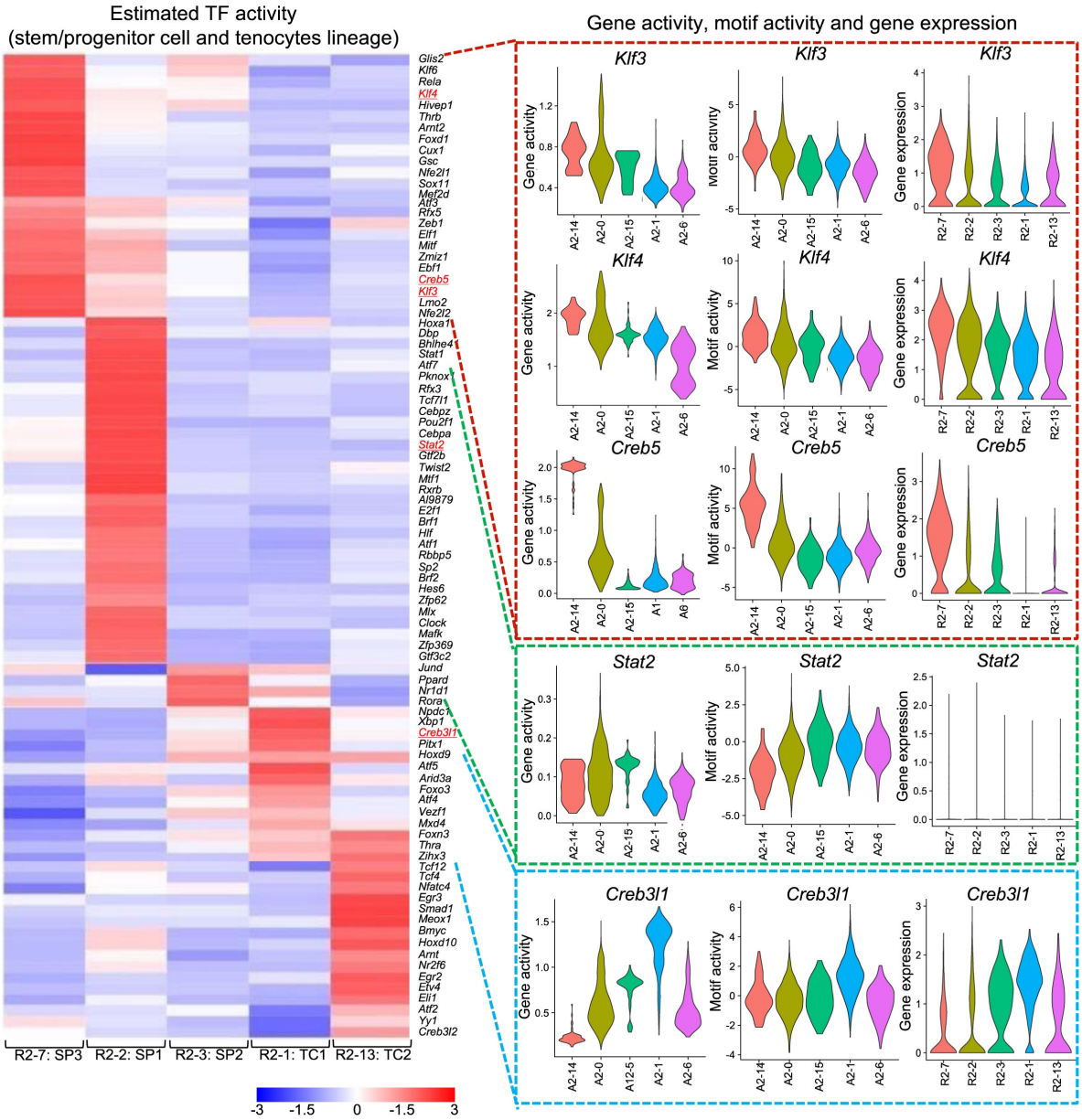


Figure 4.

Transcription factor landscapes of 2-week mouse Achilles tendons.

SCENIC analysis of transcription factor activity based on 2-week snRNA-seq data for tenocytes and the TSPC lineage (Left). Validation was performed based on the gene activity and motif activity of 2-week snATAC-seq and gene expression of 2-week snRNA-seq (right).

Estimated TF activity : Cd55	
TF	**spearman correlation**
Klf4	0.264195541
Klf3	0.262413411
Zeb1	0.219593135
Pbx1	0.204085218
Hic1	0.202481054
Klf2	0.196232716
Nfe2l2	0.196076187

Estimated TF activity : Cd248	
TF	**spearman correlation**
Hic1	0.375565336
Klf3	0.277768834
Zmiz1	0.272991765
Klf4	0.260526766
Zeb1	0.250583629
Sp3	0.2502976
Kdm5b	0.193229132

Estimated TF activity : Mkx	
TF	**spearman correlation**
Creb3l1	0.320638653
Zfhx3	0.169341424
Mxi1	0.149048856
Bhlhe40	0.144301452
Ets2	0.137445497
Max	0.116134957
Elk3	0.098325126

Estimated TF activity : Scx	
TF	**spearman correlation**
Creb3l1	0.316522613
Npdc1	0.163452737
Ets2	0.145463784
Mix1	0.121248221
Bhlhe41	0.115930755
Elk3	0.10803805
Bmyc	0.104968457

Table1

Estimated TF activity in each gene

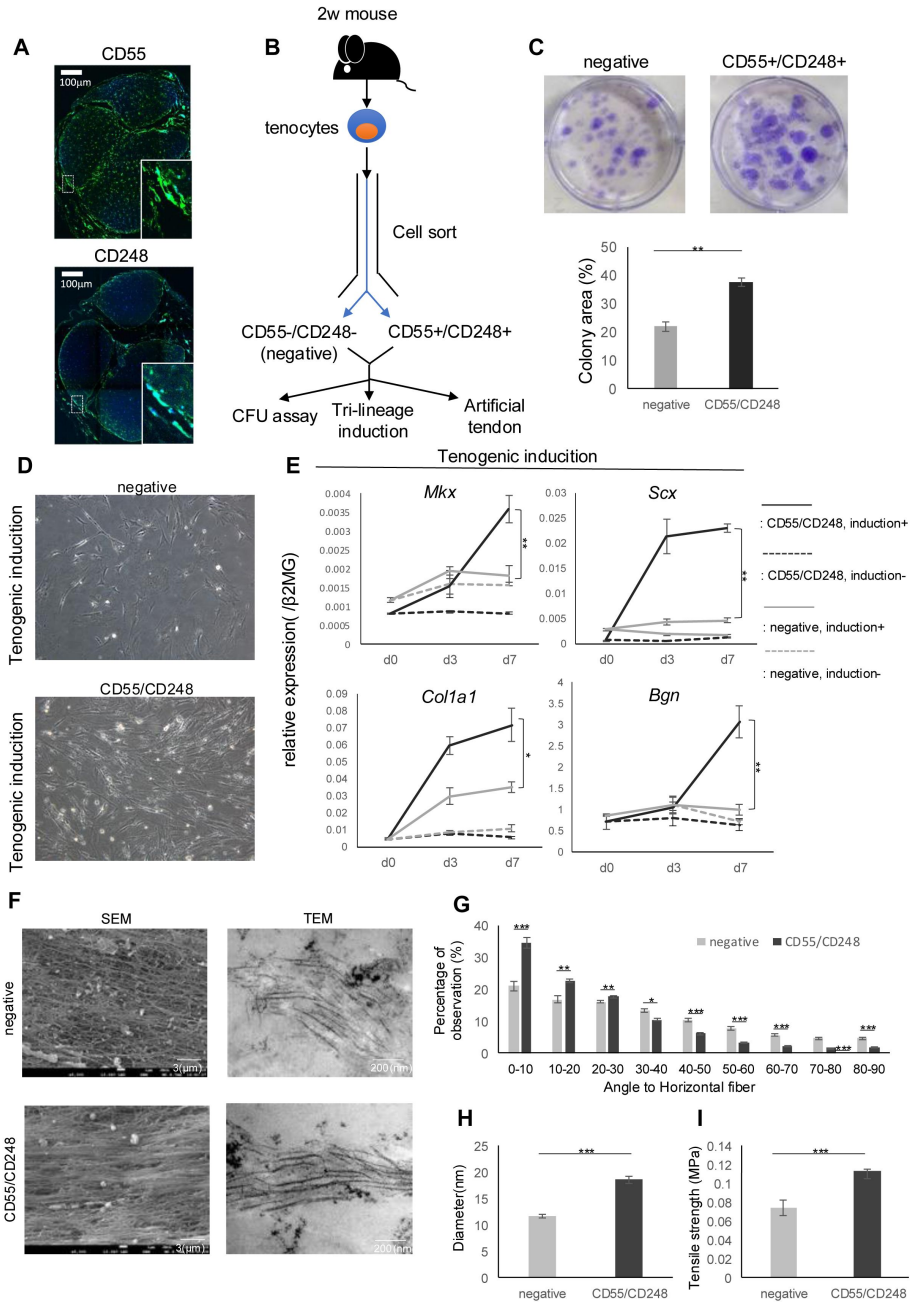


Figure 5.

In vitro analysis of CD55+/CD248+ TSPCs.

(A) Immunohistochemical image of 10-week mouse Achilles tendons. Scale bars show 100μm. CD55 and CD248, green; Hoechst 33342, blue. (B) Schema of the *in vitro* assessment of the capacity of CD55+/CD248+ TSPCs as the differentiation toward tenocytes. (C) Colony-forming efficiency of CD55+/CD248+ and CD55-/CD248-(negative) TSPCs. Colonies were stained with crystal violet (n = 6). CD55+/CD248+ TSPC exhibited higher clonogenic capacity. Data are presented as means ± SEM. $**p < 0.01$. (D) Morphological changes of CD55+/CD248+ and negative TSPCs after tenogenic induction. (E) Quantitative PCR of tendon-related gene expression in CD55+/CD248+ and negative TSPCs after tenogenic induction (n = 3). Data are presented as means ± SD. $**p < 0.01$, $*p < 0.05$. (F) SEM and TEM imaging of artificial tendons derived from CD55+/CD248+ and negative TSPCs. Data are presented as means ± SD. $**p < 0.01$, $*p < 0.05$. (G) Proportions of fiber alignment for each artificial tendon (n = 4). Data are presented as means ± SEM. $***p < 0.005$, $**p < 0.01$, $*p < 0.05$. (H) Diameter of collagen fiber in each artificial tendon based on TEM imaging (n = 4). Data are presented as means ± SD. $***p < 0.005$. (I) Tensile strength (MPa) of each artificial tendon (n = 5). Data are presented as means ± SEM. $***p < 0.005$.

while CD55/CD248 positive cells initially had lower *Mkx* and *Scx* expression than negative cells, this pattern reversed after differentiation (**Figure 5E**). Increased expression of other tendon-related genes such as *Col1a1* and *Bgn* was also observed. In contrast, no significant increases in *Sox9* and *Collagen type II alpha 1 chain* (*Col2a1*) (chondrogenic) or *RUNX family transcription factor 2* (*Runx2*), and *Alkaline phosphatase, biomineralization associated* (*Alpl*) (osteogenic) expression were observed compared to negative cells during cartilage and bone differentiation assays (Supplemental Figure 10). Cells negative for CD55/CD248 could be mixed cell populations, including hematopoietic lineages, cells from tendon mid substance, immune cells, and/or endothelial cells. However, given that CD55/CD248 negative cells have increased expression of genes characteristic of tendons such as *Mkx*, *Scx* and *Col1a1*, we could speculate that this popularity might reflect terminally differentiated tenocytes and CD55/CD248 positive cells possess tenogenic differentiation capacity.

To further evaluate the tenogenic potential of CD55/CD248 positive cells, we created tendon-like tissue (bio-tendon) using our previously reported 3D stretch stimulation culture system. SEM and TEM analysis to assess collagen fiber density and thickness showed that bio-tendons derived from CD55/CD248 positive cells had an increased proportion of collagen fibers parallel to the stretch direction and increased collagen fiber diameter. TEM also revealed characteristic banding patterns and triple helix structures in these bio-tendons, indicating mature collagen organization (Wieczorek et al., 2015) (**Figure 5F-H**).

Stretch tests to evaluate the mechanical capacity of the bio-tendons showed high tensile strength (**Figure 5I**).

These results demonstrate that CD55/CD248 positive cells have a tendency to differentiate into tendon cells.

Discussion

Tendons are known for their low cellular content, making complete functional recovery after injury challenging and increasing the risk of re-rupture. As a result, cell therapy has garnered attention as a novel treatment strategy, distinct from current conservative and surgical approaches. However, this requires a thorough understanding of TSPCs. Research on TSPCs has been limited due to the lack of identified characteristic surface antigens.

To our knowledge, no multi-omics analysis comparing juvenile and mature mouse tendon cells has been conducted. In this study, we performed single-cell RNA-seq along with snRNA-seq+snATAC-seq on nuclei isolated from 2-week and 6-week mouse Achilles tendons. While comparisons using older mice (e.g., 12 weeks or more) were initially considered, they were excluded due to extremely low cell yield and viability, making 2-and 6-week-old mice more suitable for this analysis. This approach allowed us to identify candidate surface antigens for TSPCs and elucidate the dynamism of transcription factors involved in post-developmental tendon growth. As a result, we discovered a novel combination of *Cd55* and *Cd248* as surface antigens showing characteristic expression in the TSPC fraction.

While scRNA-seq analysis also showed *Cd34* expression characteristic of the TSPC fraction, consistent with previous reports, snRNA-seq + snATAC-seq analysis revealed that the *Cd34* high-expression cluster could be further classified into two clusters based on *Cd55* and *Cd248* expression patterns. The cluster showing high expression of both *Cd55* and *Cd248* also exhibited high expression of *Tppp3*, *Pdgfra*, and *Ly6a*, suggesting that isolating *Cd55* and *Cd248* positive fractions may more sensitively capture immature populations compared to *Cd34*.

snRNA-seq+snATAC-seq analysis provided simultaneous information on gene expression levels and open chromatin regions for each cluster. Notably, *Cd248* showed several peaks upstream of the TSS in high-expression clusters, likely reflecting its expression regulation mechanism. We also identified several peaks that increased proportionally with expression for *Mkx* and *Scx*. Given the many unknowns in *Mkx* and *Scx* expression regulation mechanisms (Guerquin et al., 2013; Otabe et al., 2015), we used the SCENIC package to predict upstream transcription factors based on expressed genes in each snRNA-seq cluster. Combined with gene activity and motif activity data from snATAC-seq, Creb3l1 was identified as a transcription factor regulating *Mkx* and *Scx* expression. *Creb3l1* has been reported to increase in expression throughout development (Liu et al., 2015).

Klf3, *Klf4*, and *Creb5* were identified as characteristic transcription factors in the TSPC fraction. Recent reports have highlighted the role of the *Klf* family in limb development (Kult et al., 2021). Given that tendons are components of the developing limb, these transcription factors may have roles in tendon biology. *Creb5* has been reported to show increased expression from E11 to E13 (Liu et al., 2015), suggesting a role in early developmental stages. Furthermore, *Creb5* has been reported to regulate Proteoglycan 4 (*Prg4*) expression (Zhang et al., 2021). Further investigation into the functions of these transcription factors in TSPCs is necessary.

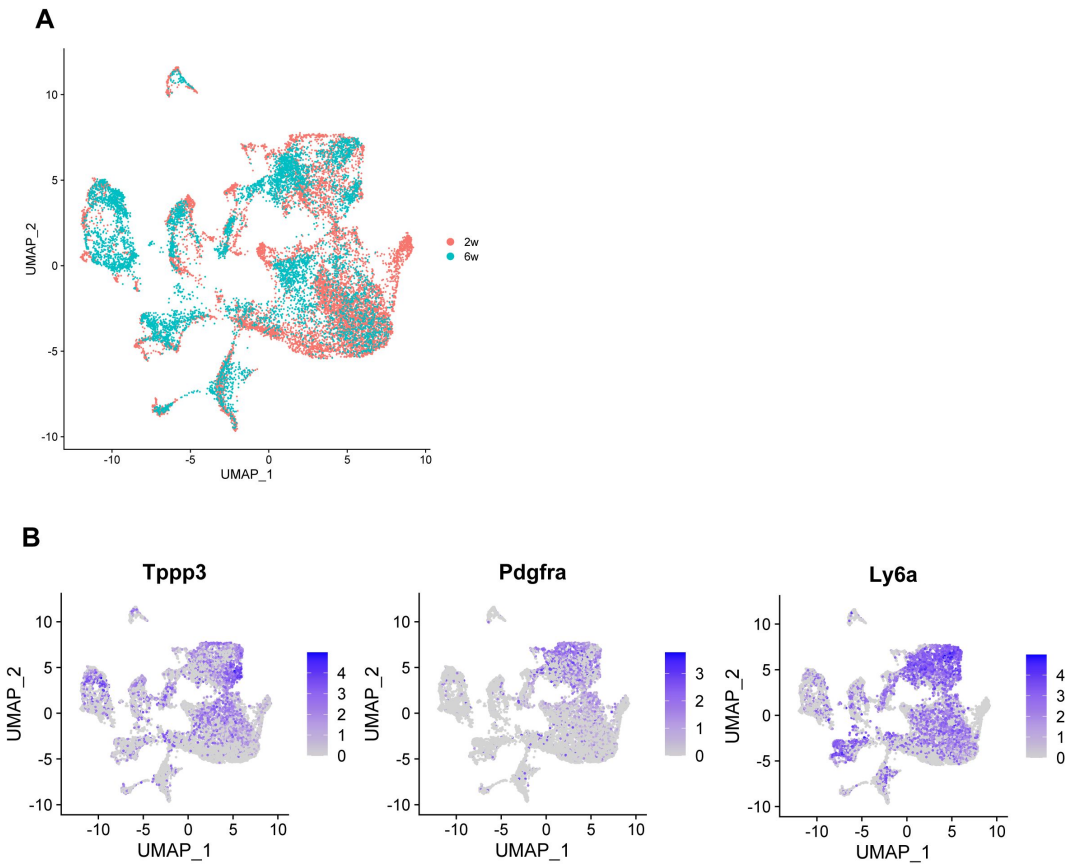
CD55 was identified by Hoffmann et al. in 1969 as a surface antigen functioning as a complement inhibitory factor on erythrocytes. CD55 is known to be expressed from early developmental stages and characterizes the initial differentiation stage of hematopoietic stem cells (Guo et al., 2013). It is also expressed in MSCs and has been reported as an early progenitor marker for mouse mammary epithelial cells (Pal et al., 2017). CD248 was identified in 1992 as an antigen for the FB5 antibody reacting with vascular wall cells (Rettig et al., 1992). CD248 is a transmembrane glycoprotein expressed in pericytes and fibroblasts during developmental stages. Regarding their relationship, CD55 and CD248 have been reported to be expressed in stromal cells during the early stages of arthritis (Choi et al., 2017), but their detailed mechanisms in tendons remain unclear.

Previously suggested TSPC surface antigen candidates such as *Cd73*, *Cd90*, and *Cd105* showed poor correlation with the expression of genes like *Tppp3*. Moreover, the initial report on TSPCs (Bi et al., 2007) described TSPCs as *Scx+Cd34-*. The high expression of CD55 and CD248 in the peritenon, similar to *Tppp3* (Harvey et al., 2019; Staverosky et al., 2019), suggests the possibility of cells with tenogenic differentiation potential exhibiting *Scx+Cd34*-expression patterns within the tendon. Spatial information is important to investigate further. Future studies using lineage tracing experiments with mice labeled for CD55 and CD248 or mouse model of selective ablation of CD55 and CD248 are necessary to analyze the developmental functions of CD55 and CD248 positive cells and their roles in injury healing in more detail.

Clinically, to our knowledge, few studies have sorted TSPCs based on surface antigens and examined their *in vitro* tendon tissue-generating ability. In this study, we found that cells sorted for CD55/CD248 showed higher clonogenicity compared to CD55-/CD248-cells and demonstrated superior tendon tissue-generating ability in an artificial tendon model. In the future, CD55/CD248 double-positive TSPC cells may prove useful in clinical applications such as *in vitro* artificial tendon creation (Tsutsumi et al., 2022) and cell therapy for tendon injuries (Huang et al., 2021).

Data availability

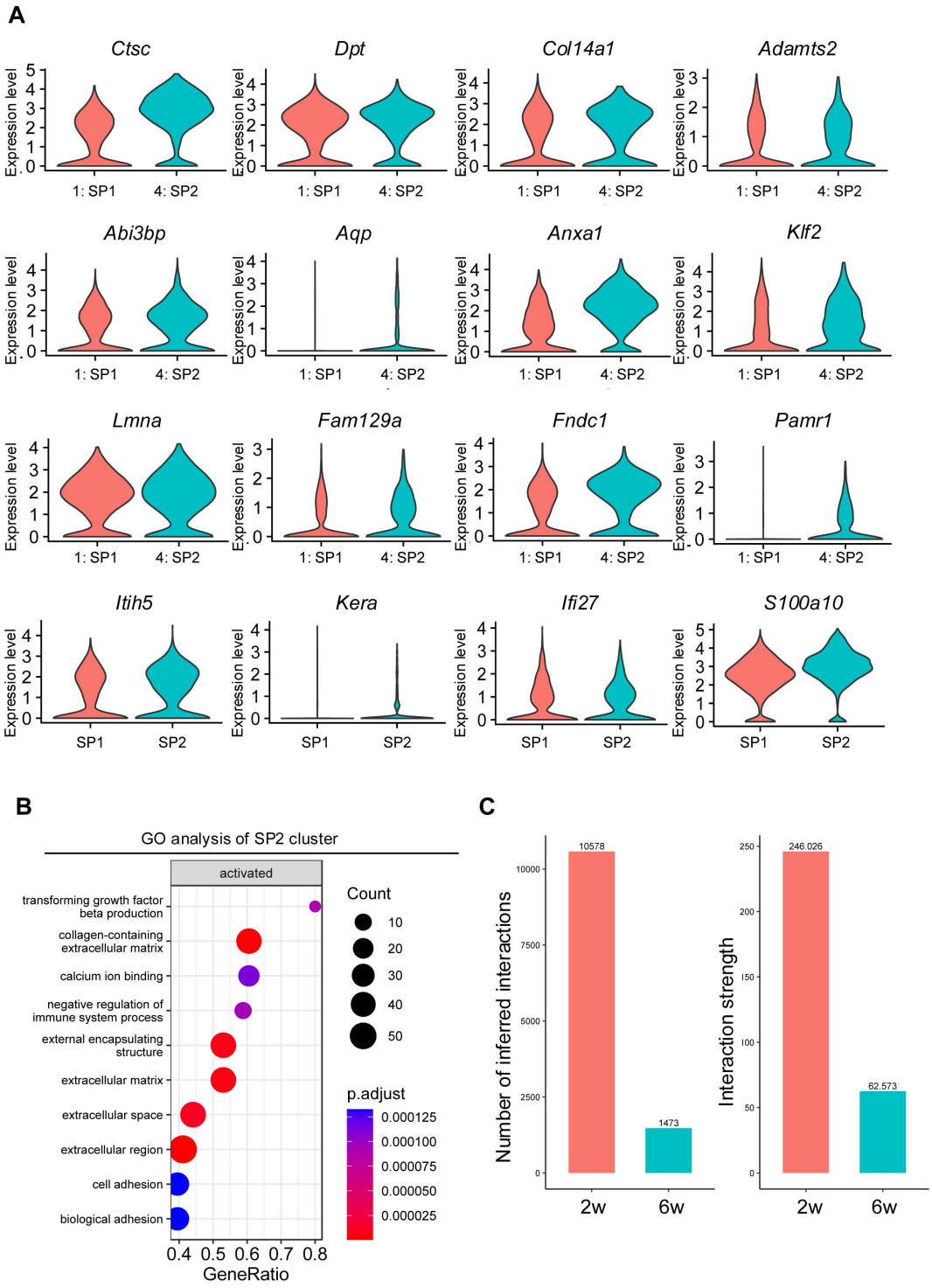
FASTQ data of RNA-Seq and ATAC-seq are deposited in DDBJ under accession number PRJDB18857



Supplemental Figure 1.

2w-and 6w-scRNA seq results.

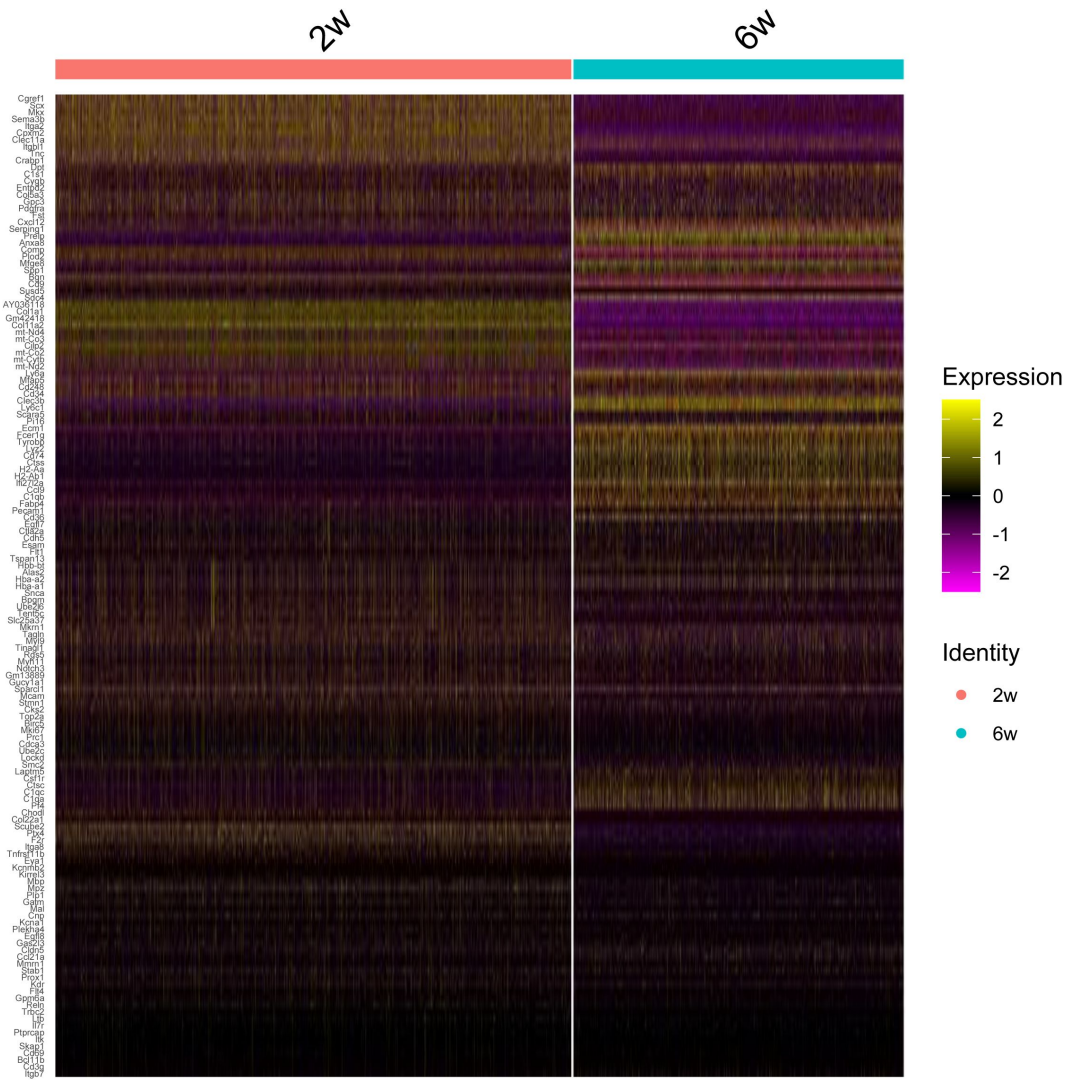
(A) Integrated uniform manifold approximation and projection (UMAP) scRNA-seq clustering of cells harvested from 2-week-old and 6-week-old mouse Achilles tendons. Clusters are colored according to cluster type. (B) Feature plot of *Tppp3*, *Pdgfra* and *Ly6a* expression.



Supplemental Figure 2.

Comparison of 2w-and 6w-scRNA seq results.

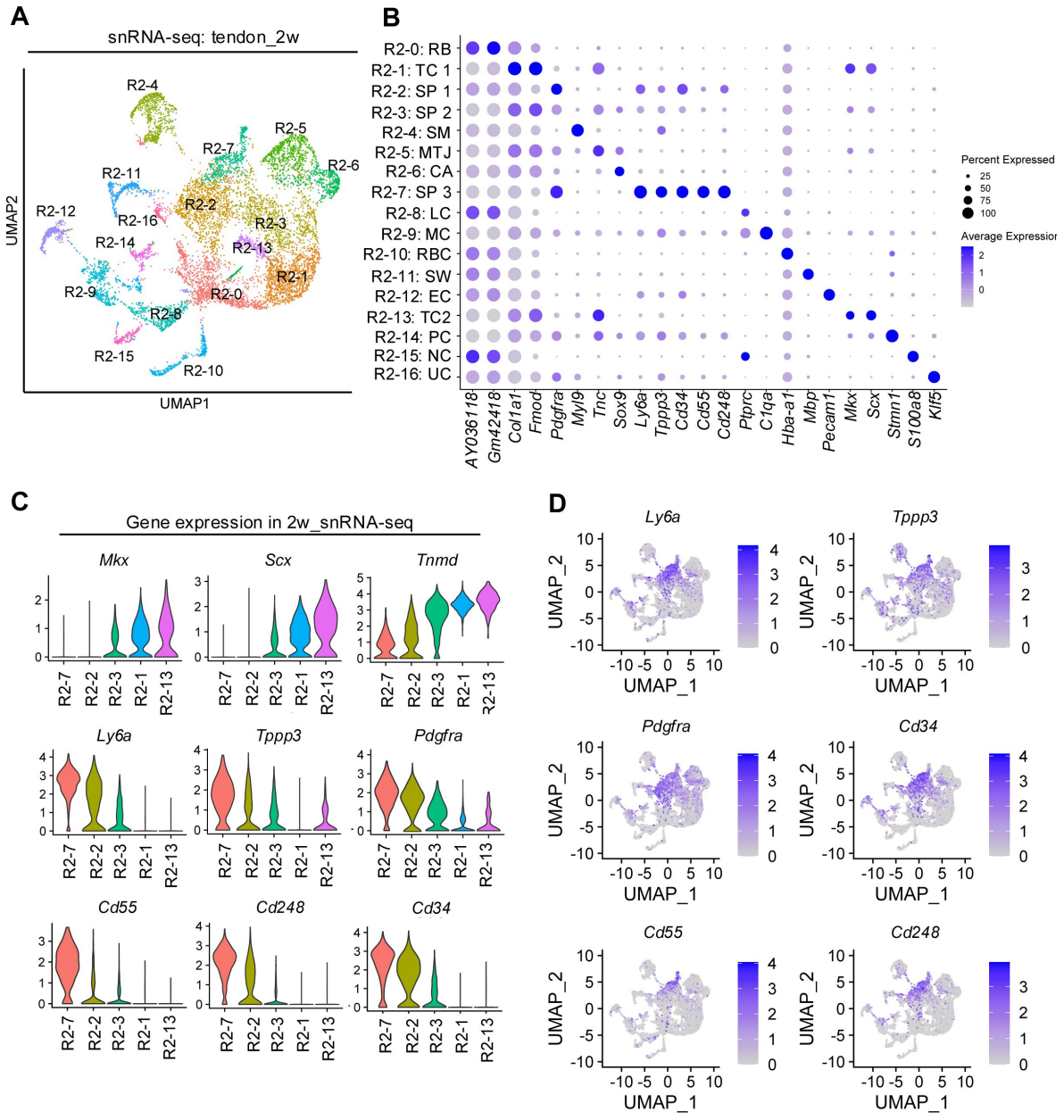
(A) Violin plot of gene expression enriched in the limb bud (Development (2014) 141, 3683-3696) in each cluster. (B) Gene Ontology (GO) terms associated with genes with upregulated expression in the SP2 cluster. (C) Number of inferred interactions and interaction strength of genes expressed within 2w-and 6w-mouse scRNA-seq data.



Supplemental Figure 3.

Comparative analysis of 2w-and 6w-scRNA seq results.

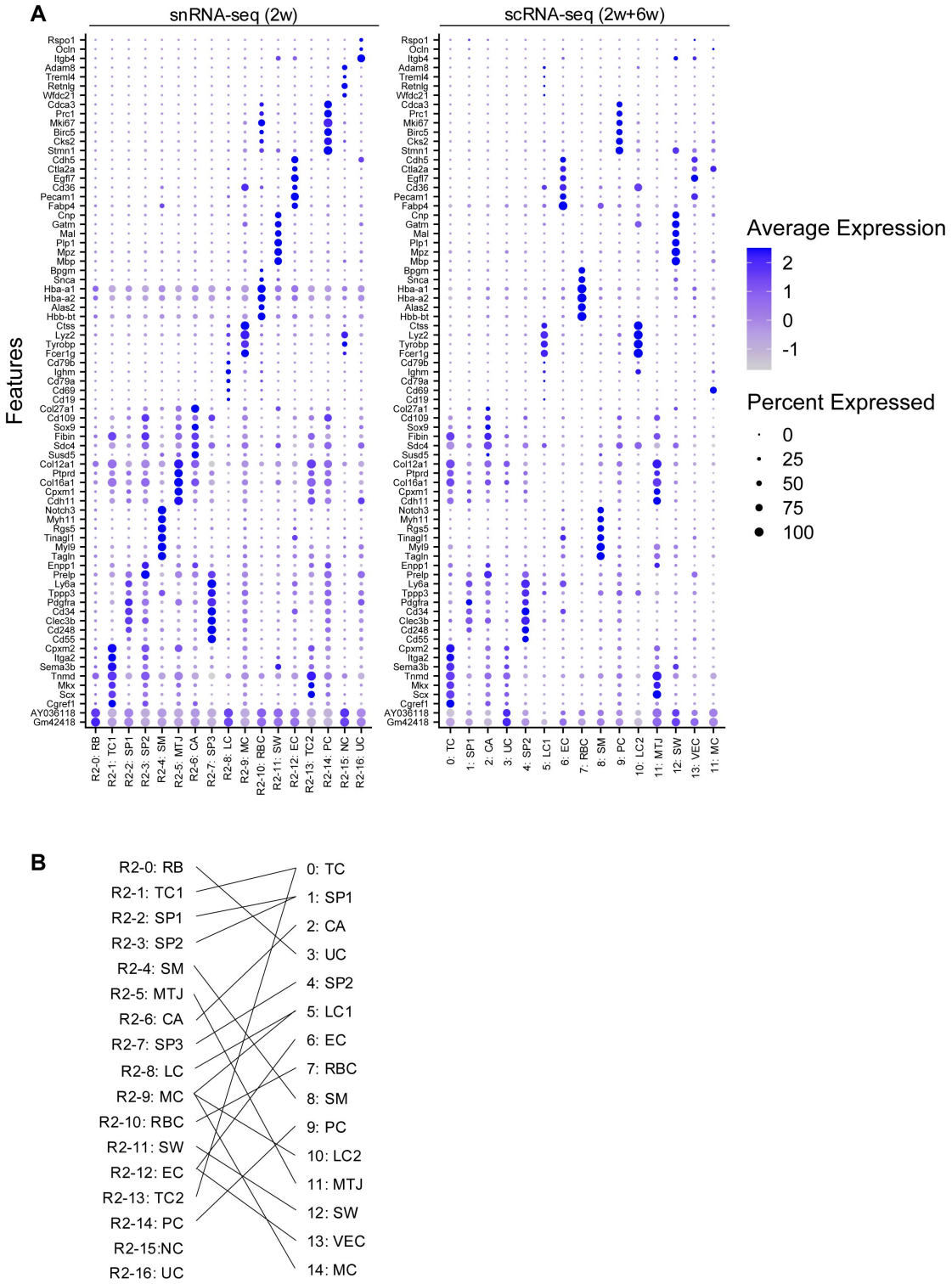
Heatmap showing differential gene expression between red (2w) and green (6w).



Supplemental Figure 4.

Analysis of 2-week snRNA-seq data.

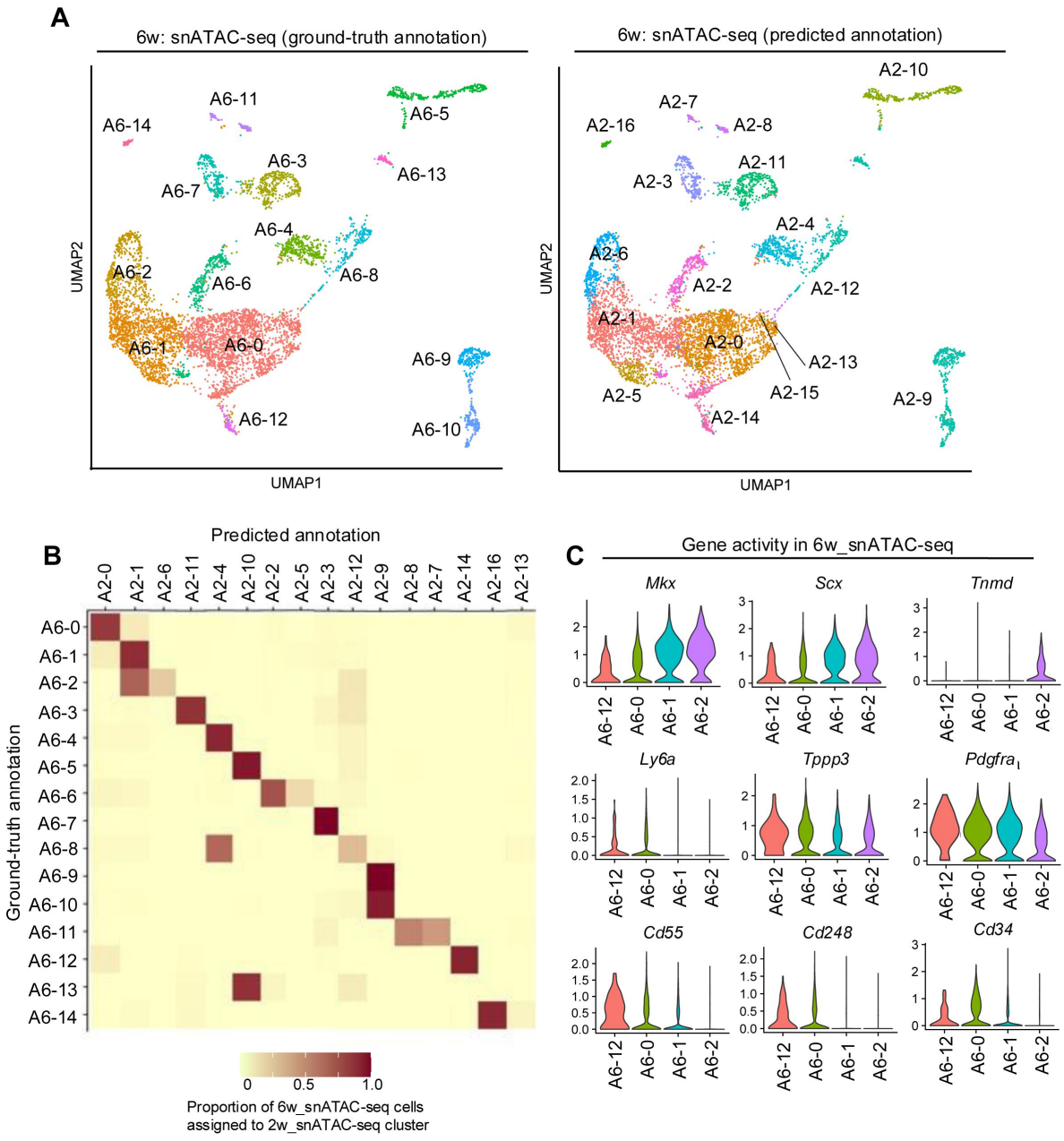
(A) UMAP snRNA-seq clustering of cells harvested from 2-week mouse Achilles tendons. (B) Dot plot of average gene expression levels of the indicated genes in 2-week snRNA-seq clusters. RB, ribosomal RNA; TC1, tenocyte_1; SP1, stem/progenitor cell_1; SP2, stem/progenitor cell_2; SM, smooth muscle cell; MTJ, myotendinous junction cell; CA, cartilage; SP3, stem/progenitor cell_3; LC, lymphocyte; MC, macrophage; RBC, red blood cell; SW, Schwann cell; EC, endothelial cell; TC2, tenocyte_2; PC, proliferating cell; NC, neutrophil; UC, unspecified cell. (C) Violin plot of tenocytes and TSPC-related gene expression in each cluster. (D) Feature plot of TSPC-related gene expression.



Supplemental Figure 5.

Analysis of 6-week snATAC-seq.

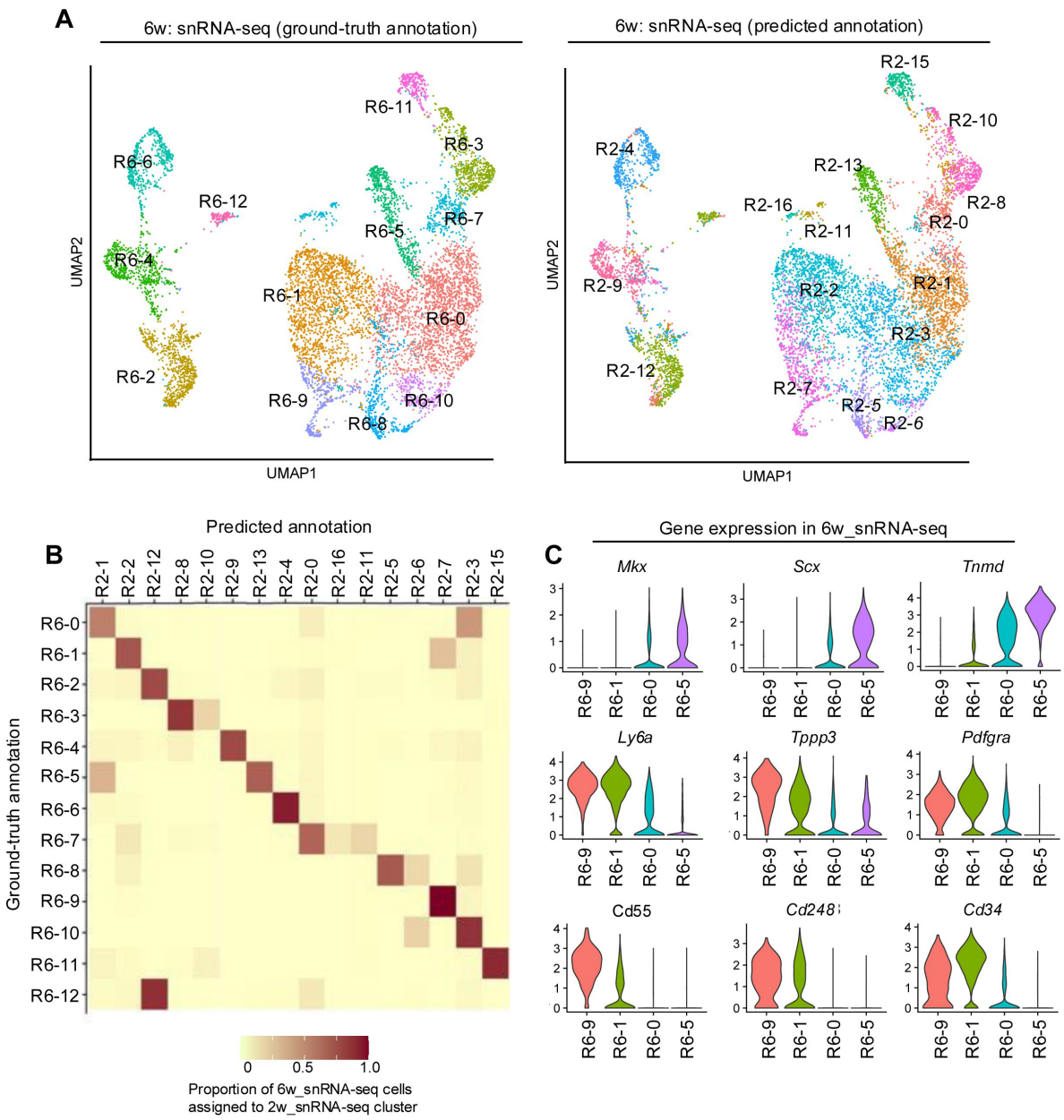
(A) UMAP snATAC-seq clustering of cells harvested from 6-week mouse Achilles tendons. Annotation was based on gene activity (ground-truth annotation, left) and the predicted annotation inferred from 2-week snATAC-seq (right). (B) Identification of matching cell clusters between the 2-week and 6-week snATAC-seq data, visualized as a heatmap. (C) Violin plot of tenocytes and TSPC-related gene activity in each cluster.



Supplemental Figure 6.

Comparison of snRNA-seq and scRNA-seq.

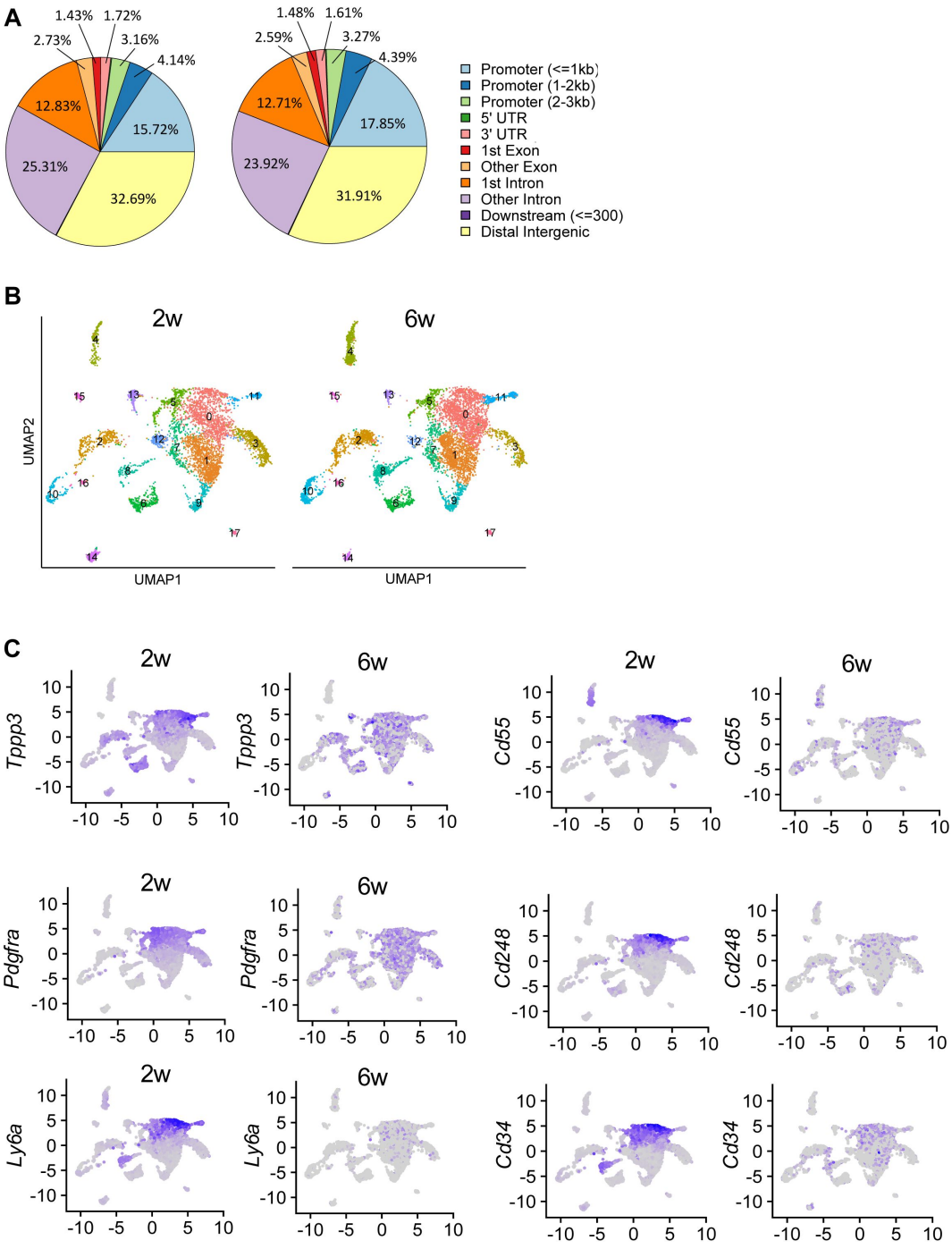
(A) Dot plot of representative DEGs in snRNA-seq (2w) and snRNA-seq (2w+6w). (B) Relevance of annotation in each data.



Supplemental Figure 7.

Analysis of 6-week snRNA-seq.

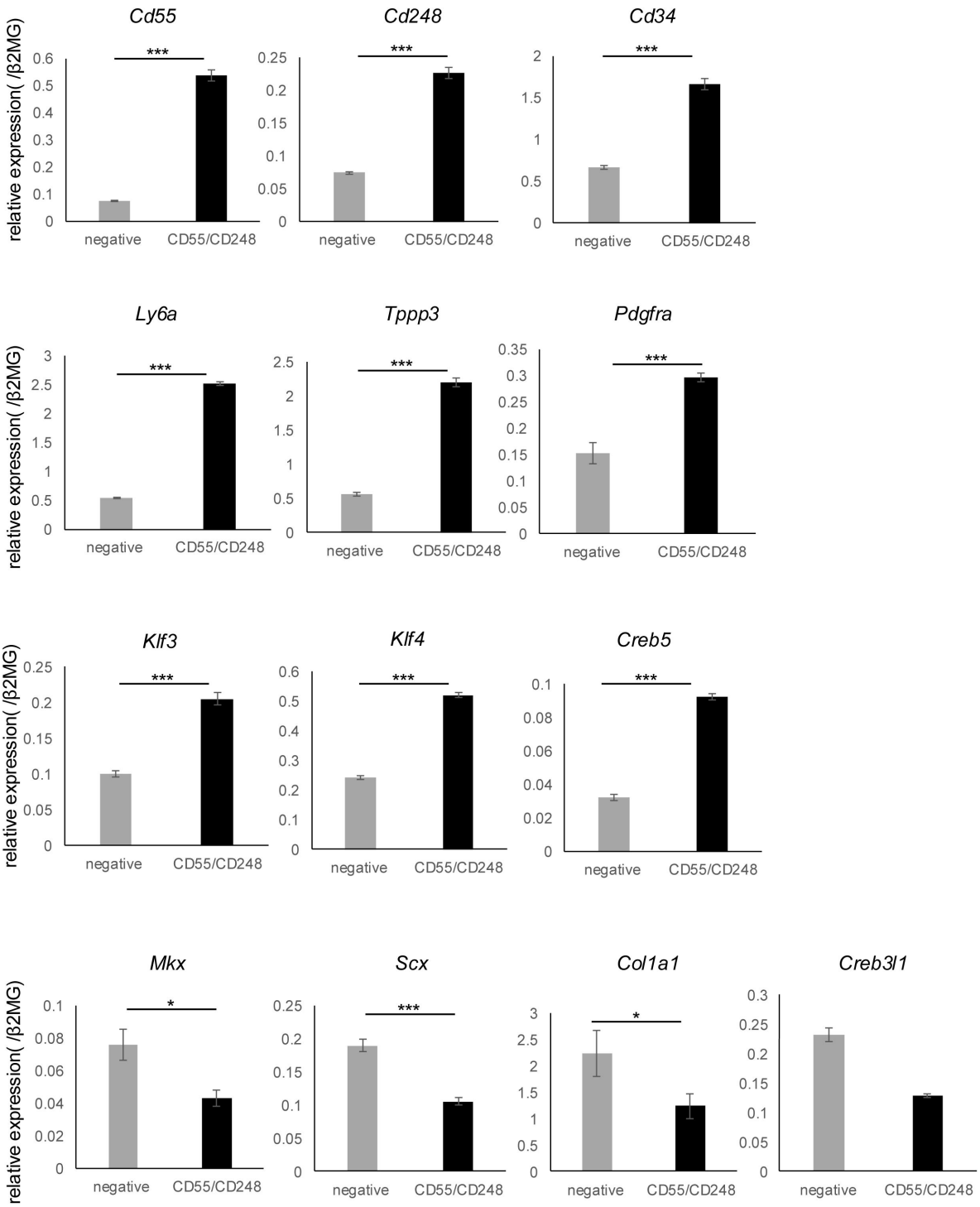
(A) UMAP snRNA-seq clustering of cells harvested from 6-week mouse Achilles tendons. Annotation was based on each gene expression (ground-truth annotation, left) and predicted annotation inferred from 2-week snRNA-seq (right). (B) Identification of matching cell clusters between the 2-week and 6-week snRNA-seq data visualized as a heatmap. (C) Violin plot of tenocytes and TSPC-related gene expression in each cluster.



Supplemental Figure 8.

Comparison of 2-week and 6-week snATAC-seq.

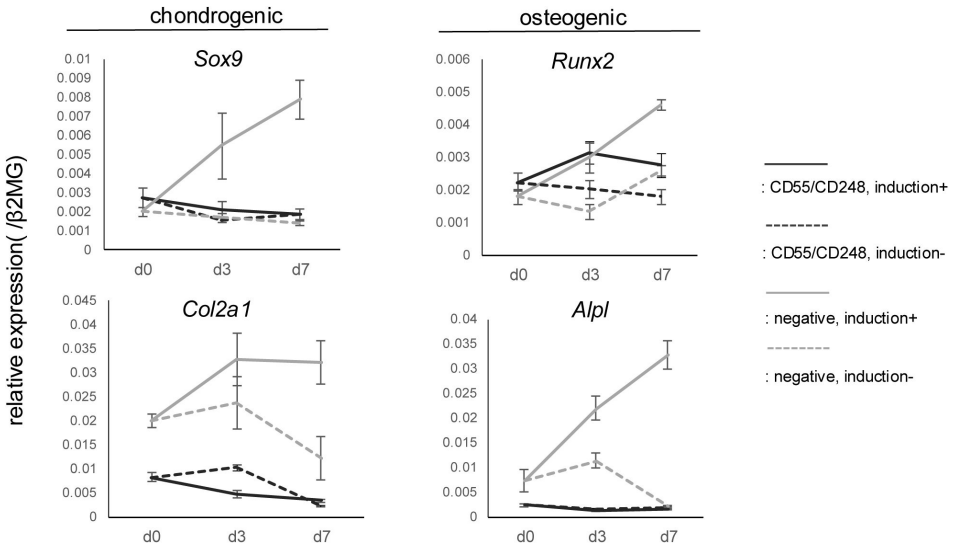
(A) Circle plot of annotated differentially accessible regions for each data. (B) Integrated UMAP snATAC-seq clustering of cells harvested from 2-week and 6-week mouse Achilles tendons. (C) Feature plot of TSPC-related gene activity in each dataset.



Supplemental Figure 9.

Gene expression changes in CD55+/CD248+ and negative TSPCs.

Quantitative PCR of gene expression in CD55+/CD248+ and negative TSPCs (n = 4). Data are presented as means ± SEM. *p < 0.05, ***p < 0.005.



Supplemental Figure 10.

Chondrogenic and osteogenic induction of CD55+/CD248+ and negative TSPCs.

Quantitative PCR of tendon-related genes in CD55+/CD248+ and negative TSPCs after chondrogenic and osteogenic induction (n = 4). Data are presented as means ± SEM. $^{**}p < 0.01$, $^{*}p < 0.05$.

Acknowledgements

We thank all the members of the Department of Systems BioMedicine at Institute of Science Tokyo for their support. We also thank the Research Core at Institute of Science Tokyo for supporting cell sorting.

Additional information

Author contributions

Conceptualization, H.T., T.C., and H.A.; Methodology, H.T., T.C., and H.A.; Formal analysis, H.T.; Investigation, H.T., F. Y., T. M., T. K., A. K., K. A., Y. S.; Writing – Original Draft, H.T.; Writing – Review & Editing, T.C., and H.A.; Funding Acquisition, H.A.; Supervision, H.A.

Funding

This work was supported by JSPS KAKENHI (Grant Numbers JP15H02560, JP20H05696, 16H06279 (PAGS)), AMED (Grant Numbers JP21gm0810008, JP23ym0126805, JP24gm0010009), and NIH (Grant Number R01AR080127) to H.A.

Additional files

Supplementary table 1

Supplementary table 2

Supplementary table 3

Supplementary table 4

Supplementary table 5

Supplementary table 6

Supplementary table 7

Supplementary table 8

Supplementary table 9

Supplementary table 10

Supplementary table 11

Supplementary table 12

Supplementary table 13

Supplementary table 14

Supplementary table 15

References

Andarawis-Puri N., Flatow EL., Soslowsky LJ (2015) **Tendon basic science: Development, repair, regeneration, and healing** *J Orthop Res* **33**:780–4 Google Scholar

Asahara H., Inui M., Lotz MK (2017) **Tendons and Ligaments: Connecting Developmental Biology to Musculoskeletal Disease Pathogenesis** *J Bone Miner Res* **32**:1773–1782 Google Scholar

Becht E., McInnes L., Healy J., Dutertre CA., Kwok IWH., Ng LG., Ginhoux F., Newell EW. (2019) **Dimensionality reduction for visualizing single-cell data using UMAP** *Nat Biotechnol* **37**:38–44 Google Scholar

Bi Y, Ehirchiou D, Kilts TM, Inkson CA, Embree MC., Sonoyama W., Li L., Leet AI., Seo BM., Zhang L., Shi S., Young MF (2007) **Identification of tendon stem/progenitor cells and the role of the extracellular matrix in their niche** *Nat Med* **13**:1219–27 Google Scholar

Butler A., Hoffman P., Smibert P., Papalexi E., Satija R (2018) **Integrating single-cell transcriptomic data across different conditions, technologies, and species** *Nat Biotechnol* **36**:411–420 Google Scholar

Cao Y., Liu Y., Liu W., Shan Q., Buonocore SD., Cui L (2002) **Bridging tendon defects using autologous tenocyte engineered tendon in a hen model** *Plast Reconstr Surg* **110**:1280–9 Google Scholar

Chen B., Wang B., Zhang WJ., Zhou G., Cao Y., Liu W (2012) **In vivo tendon engineering with skeletal muscle derived cells in a mouse model** *Biomaterials* **33**:6086–97 Google Scholar

Chen L., Liu JP., Tang KL., Wang Q., Wang GD., Cai XH., Liu XM (2014) **Tendon derived stem cells promote platelet-rich plasma healing in collagenase-induced rat achilles tendinopathy** *Cell Physiol Biochem* **34**:2153–68 Google Scholar

Cho YT., Hsu WY., Lin LF., Lin YN (2018) **Kinesio taping reduces elbow pain during resisted wrist extension in patients with chronic lateral epicondylitis: a randomized, double-blinded, cross-over study** *BMC Musculoskelet Disord* **19**:193 Google Scholar

Connizzo BK., Yannascoli SM., Soslowsky LJ (2013) **Structure-function relationships of postnatal tendon development: a parallel to healing** *Matrix Biol* **32**:106–16 Google Scholar

Craft A. M., Ahmed N., Rockel J. S., Baht G. S., Alman B. A., Kandel R. A., Grigoriadis A. E., Keller G. M (2013) **Specification of chondrocytes and cartilage tissues from embryonic stem cells** *Development* **140**:2597–2610 Google Scholar

Diez-Roux G., Banfi S., Sultan M., Geffers L., Anand S., Rozado D., Magen A., Canidio E., Pagani M., Peluso I., Lin-Marq N., Koch M., Bilio M., Cantiello I., Verde R., Masi C., Bianchi SA., Cicchini J., Perroud E., Mehmeti S., Dagand E., Schrinner S., Nürnberger A., Schmidt K., Metz K., Zwingmann C., Brieske N., Springer C., Hernandez AM., Herzog S., Grabbe F., Sieverding C., Fischer B., Schrader K., Brockmeyer M., Dettmer S., Helbig C., Alunni V., Battaini MA., Mura C., Henrichsen CN., Garcia-Lopez R., Echevarria D., Puelles E., Garcia-Calero E., Kruse S., Uhr M., Kauck C., Feng G., Milyaev N., Ong CK., Kumar L., Lam M., Semple CA., Gyenesei A., Mundlos S.,

Radelof U., Lehrach H., Sarmientos P., Reymond A., Davidson DR., Dollé P., Antonarakis SE., Yaspo ML., Martinez S., Baldock RA., Eichele G., Ballabio A. (2011) **A High-Resolution Anatomical Atlas of the Transcriptome in the Mouse Embryo** *PLoS Biol* **9**:e1000582 Google Scholar

Gracey E., Burssens A., Cambré I., Schett G., Lories R., McInnes IB., Asahara H., Elewaut D (2020) **Tendon and ligament mechanical loading in the pathogenesis of inflammatory arthritis** *Nat Rev Rheumatol* **16**:193–207 Google Scholar

Fang F., Xiao Y., Zelzer E., Leong KW., Thomopoulos S (2022) **A mineralizing pool of Gli1-expressing progenitors builds the tendon enthesis and demonstrates therapeutic potential** *Cell Stem Cell* **29**:1669–1684 Google Scholar

Franken NA., Rodermond HM., Stap J., Haveman J., van Bree C. (2006) **Clonogenic assay of cells in vitro** *Nat Protoc* **2006**:1 Google Scholar

Grinstein M., Dingwall HL., O'Connor LD., Zou K., Capellini TD., Galloway JL (2019) **A distinct transition from cell growth to physiological homeostasis in the tendon** *eLife* **8**:e48689 https://doi.org/10.7554/eLife.48689 | Google Scholar

Harvey T., Flamenco S., Fan CM (2019) **A Tppp3 + Pdgfra + tendon stem cell population contributes to regeneration and reveals a shared role for PDGF signalling in regeneration and fibrosis** *Nat Cell Biol* **21**:1490–1503 Google Scholar

Guilak F., Butler DL., Goldstein SA., Baaijens FP (2014) **Biomechanics and mechanobiology in functional tissue engineering** *J Biomech* **47**:1933–40 Google Scholar

Gumucio JP., Sugg KB., Mendias CL (2015) **TGF-β superfamily signaling in muscle and tendon adaptation to resistance exercise** *Exerc Sport Sci Rev* **43**:93–9 Google Scholar

Guzmán C., Bagga M., Kaur A., Westermarck J., Abankwa D (2014) **ColonyArea: an ImageJ plugin to automatically quantify colony formation in clonogenic assays** *PLoS One* **9**:e92444 Google Scholar

Havis E., Bonnin MA., Olivera-Martinez I., Nazaret N., Ruggiu M., Weibel J., Durand C., Guerquin MJ., Bonod-Bidaud C., Ruggiero F., Schweitzer R., Duprez D (2014) **Transcriptomic analysis of mouse limb tendon cells during development** *Development* **141**:3683–96 Google Scholar

Hittinger M., Czyz ZT., Huesemann Y., Maneck M., Botteron C., Kaeufl S., Klein CA., Polzer B (2013) **Molecular profiling of single Sca-1+/CD34+,-cells--the putative murine lung stem cells** *PLoS One* **8**:e83917 Google Scholar

Hoffmann EM (1969) **Inhibition of complement by a substance isolated from human erythrocytes. II. Studies on the site and mechanism of action** *Immunochemistry* **6**:405–19 Google Scholar

Holmes C., Stanford WL (2007) **Concise review: stem cell antigen-1: expression, function, and enigma** *Stem Cells* **25**:1339–47 Google Scholar

Howell K., Chien C., Bell R., Laudier D., Tufa SF., Keene DR., Andarawis-Puri N., Huang AH (2017) **Novel Model of Tendon Regeneration Reveals Distinct Cell Mechanisms Underlying Regenerative and Fibrotic Tendon Healing** *Sci Rep* **7**:45238 Google Scholar

Isola JVV., Ocañas SR., Hubbart CR., Ko S., Mondal SA., Hense JD., Carter HNC., Schneider A., Kovats S., Alberola-Ila J., Freeman WM., Stout MB (2024) **A single-cell atlas of the aging mouse ovary** *Nat Aging* **4**:145–162 Google Scholar

Jacobson KR., Lipp S., Acuna A., Leng Y., Bu Y., Calve S (2020) **Comparative Analysis of the Extracellular Matrix Proteome across the Myotendinous Junction** *J Proteome Res.* **19**:3955–3967 Google Scholar

Jozsa L., Kannus P., Balint JB., Reffy A (1991) **Three-dimensional ultrastructure of human tendons** *Acta Anat (Basel)* **142**:306–12 Google Scholar

Kataoka K., Kurimoto R., Tsutsumi H., Chiba T., Kato T., Shishido K., Kato M., Ito Y., Cho Y., Hoshi O., Mimata A., Sakamaki Y., Nakamichi R., Lotz MK., Naruse K., Asahara H. (2020) **In vitro Neo-Genesis of Tendon/Ligament-Like Tissue by Combination of Mohawk and a Three-Dimensional Cyclic Mechanical Stretch Culture System** *Front Cell Dev Biol* **2020**:307 Google Scholar

Kendal AR., Layton T., Al-Mossawi H., Appleton L., Dakin S., Brown R., Loizou C., Rogers M., Sharp R., Carr A (2020) **Multi-omic single cell analysis resolves novel stromal cell populations in healthy and diseased human tendon** *Sci Rep* **10**:13939 Google Scholar

Leong DJ., Sun HB (2016) **Mesenchymal stem cells in tendon repair and regeneration: basic understanding and translational challenges** *Ann. N. Y. Acad. Sci* **1383**:88–96 Google Scholar

Leong NL., Kator JL., Clemens TL., James A., Enamoto-Iwamoto M., Jiang J (2020) **Tendon and Ligament Healing and Current Approaches to Tendon and Ligament Regeneration** *J Orthop Res* **38**:7–12 Google Scholar

Li Y., Wu T., Liu S (2021) **Identification and Distinction of Tenocytes and Tendon-Derived Stem Cells** *Front Cell Dev Biol* **9**:629515 Google Scholar

Liu H., Xu J., Liu CF., Lan Y., Wylie C., Jiang R (2015) **Whole transcriptome expression profiling of mouse limb tendon development by using RNA-seq** *J Orthop Res* **33**:840–848 Google Scholar

Loiacono C., Palermi S., Massa B., Belviso I., Romano V., Gregorio AD., Sirico F., Sacco AM (2019) **Tendinopathy: Pathophysiology, Therapeutic Options, and Role of Nutraceutics. A Narrative Literature Review** *Medicina (Kaunas)* **55**:447 Google Scholar

Lui PP (2015) **Markers for the identification of tendon-derived stem cells in vitro and tendon stem cells in situ - update and future development** *Stem Cell Res Ther* **6**:106 Google Scholar

Lui PPY., Wong CM (2019) **Biology of Tendon Stem Cells and Tendon in Aging** *Front Genet* **10**:1338 Google Scholar

McGinnis CS., Murrow LM., Gartner ZJ (2019) **DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors** *Cell Syst* **8**:329–337 Google Scholar

Mendias CL., Gumucio JP., Bakhurin KI., Lynch EB., Brooks SV (2012) **Physiological loading of tendons induces scleraxis expression in epitenon fibroblasts** *J Orthop Res* **30**:606–12 Google Scholar

Morikawa S., Mabuchi Y., Kubota Y., Nagai Y., Niibe K., Hiratsu E., Suzuki S., Miyauchi-Hara C., Nagoshi N., Sunabori T., Shimmura S., Miyawaki A., Nakagawa T., Suda T., Okano H., Matsuzaki Y (2009) **Prospective identification, isolation, and systemic transplantation of multipotent mesenchymal stem cells in murine bone marrow** *J. Exp. Med* **206**:2483–2496 Google Scholar

Nagakura R., Yamamoto M., Jeong J, Hinata N, Katori Y, Chang WJ, Abe S (2020) **Switching of Sox9 expression during musculoskeletal system development** *Sci Rep* **10**:8425 Google Scholar

Nusspaumer G., Jaiswal S., Barbero A., Reinhardt R., Ishay Ronen D., Haumer A., Lufkin T., Martin I., Zeller R (2017) **Ontogenic identification and analysis of mesenchymal stromal cell populations during mouse limb and long bone development** *Stem Cell Rep* **9**:1124–1138 Google Scholar

Pliner HA., Packer JS., McFaline-Figueroa JL., Cusanovich DA., Daza RM., Aghamirzaie D., Srivatsan S., Qiu X., Jackson D., Minkina A., Adey AC., Steemers FJ., Shendure J., Trapnell C (2018) **Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data** *Mol. Cell* **71**:858–871 Google Scholar

Ruzzini L., Abbruzzese F., Rainer A., Longo UG., Trombetta M., Maffulli N., Denaro V (2014) **Characterization of age-related changes of tendon stem cells from adult human tendons** *Knee Surg Sports Traumatol Arthrosc* **22**:2856–66 Google Scholar

Satpathy AT., Granja JM., Yost KE., Qi Y., Meschi F., McDermott GP., Olsen BN., Mumbach MR., Pierce SE., Corces MR., Shah P., Bell JC., Jhutty D., Nemec CM., Wang J., Wang L., Yin Y., Giresi PG., Chang ALS., Zheng GXY., Greenleaf WJ., Chang HY (2019) **Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion** *Nat Biotechnol* **37**:925–936 Google Scholar

Sidney LE., Branch MJ., Dunphy SE., Dua HS., Hopkinson A (2014) **Concise review: evidence for CD34 as a common marker for diverse progenitors** *Stem Cells* **32**:1380–9 Google Scholar

Somerville JM., Aspden RM., Armour KE., Armour KJ., Reid DM (2004) **Growth of C57BL/6 mice and the material and mechanical properties of cortical bone from the tibia** *Calcif Tissue Int* **74**:469–75 Google Scholar

Staverosky JA., Pryce BA., Watson SS., Schweitzer R (2019) **Tubulin polymerization-promoting protein family member 3, Tppp3, is a specific marker of the differentiating tendon sheath and synovial joints** *Dev Dyn* **238**:685–92 Google Scholar

Steinmann S., Pfeifer CG., Brochhausen C., Docheva D (2020) **Spectrum of Tendon Pathologies: Triggers, Trails and End-State** *Int J Mol Sci* **21**:844 Google Scholar

Still C, Chang WT., Sherman SL., Sochacki KR., Dragoo JL., Qi LS. (2021) **Single-cell transcriptomic profiling reveals distinct mechanical responses between normal and diseased tendon progenitor cells** *Cell Rep Med* **2**:100343 Google Scholar

Stuart T., Hanna J., Gutruf P (2021) **Single-cell chromatin state analysis with Signac** *Nat Methods* **18**:1333–1341 Google Scholar

Sung JH., Yang HM., Park JB., Choi GS., Joh JW., Kwon CH., Chun JM., Lee SK., Kim SJ (2008) **Isolation and characterization of mouse mesenchymal stem cells** *Transplant Proc* **40**:2649–54 Google Scholar

Tachibana N., Chijimatsu R., Okada H., Oichi T., Taniguchi Y., Maenohara Y., Miyahara J., Ishikura H., Iwanaga Y., Arino Y., Nagata K., Nakamoto H., Kato S., Doi T., Matsubayashi Y., Oshima Y., Terashima A., Omata Y., Yano F., Maeda S., Ikegawa S., Seki M., Suzuki Y., Tanaka S., Saito T (2022) **RSPO2 defines a distinct undifferentiated progenitor in the tendon/ligament and suppresses ectopic ossification** *Sci Adv* **8**:eabn2138 Google Scholar

Takakura N., Yoshida H., Ogura Y., Kataoka H., Nishikawa S., Nishikawa S (1997) **PDGFR alpha expression during mouse embryogenesis: immunolocalization analyzed by whole-mount immunohistostaining using the monoclonal anti-mouse PDGFR alpha antibody APA5** *J. Histochem. Cytochem* **45**:883–893 Google Scholar

Tang JB., Wu YF., Cao Y., Chen CH., Zhou YL., Avanessian B., Shimada M., Wang XT., Liu PY (2016) **Basic FGF or VEGF gene therapy corrects insufficiency in the intrinsic healing capacity of tendons** *Sci Rep* **6**:20643 Google Scholar

Tempfer H., Wagner A., Gehwolf R., Lehner C., Tauber M., Resch H., Bauer HC (2009) **Perivascular cells of the supraspinatus tendon express both tendon-and stem cell-related markers** *Histochem Cell Biol* **131**:733–41 Google Scholar

Thomopoulos S., Parks WC., Rifkin DB., Derwin KA (2015) **Mechanisms of tendon injury and repair** *J Orthop Res* **33**:832–9 Google Scholar

Trapnell C., Cacchiarelli D., Grimsby J., Pokharel P., Li S., Morse M., Lennon NJ., Livak KJ., Mikkelsen TS., Rinn JL. (2014) **The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells** *Nat Biotechnol* **32**:381–386 Google Scholar

Tsutsumi H., Kurimoto R., Nakamichi R., Chiba T., Matsushima T., Fujii Y., Sanada R., Kato T., Shishido K., Sakamaki Y., Kimura T., Kishida A., Asahara H (2022) **Generation of a tendon-like tissue from human iPS cells** *J Tissue Eng* **13**:20417314221074018 Google Scholar

Walia B., Huang AH (2019) **Tendon stem progenitor cells: Understanding the biology to inform therapeutic strategies for tendon repair** *J Orthop Res* **37**:1270–1280 Google Scholar

Wang B., Liu W., Zhang Y., Jiang Y., Zhang WJ., Zhou G., Cui L., Cao Y (2008) **Engineering of extensor tendon complex by an ex vivo approach** *Biomaterials* **29**:2954–61 Google Scholar

Yin Z., Hu JJ., Yang L., Zheng ZF., An CR., Wu BB., Zhang C., Shen WL., Liu HH., Chen JL., Heng BC., Guo GJ., Chen X., Ouyang HW (2016) **Single-cell analysis reveals a nestin + tendon stem/progenitor cell population with strong tenogenic potentiality** *Sci Adv* **2**:e1600874 Google Scholar

Yu G., Wang LG., Han Y., He QY (2012) **clusterProfiler: an R package for comparing biological themes among gene clusters** *Omics* **16**:284–7 Google Scholar

Yu G., Wang LG., He QY (2015) **ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization** *Bioinformatics* **31**:2382–3 Google Scholar

Zhang C., Zhu J., Zhou Y., Thampatty BP., Wang JH (2019) **Tendon Stem/Progenitor Cells and Their Interactions with Extracellular Matrix and Mechanical Loading** *Stem Cells Int*

2019:3674647 Google Scholar

Zhang T., Wan L., Xiao H., Wang L., Hu J., Lu H (2023) **Single-cell RNA sequencing reveals cellular and molecular heterogeneity in fibrocartilaginous enthesis formation** *eLife* **12**:e85873 https://doi.org/10.7554/eLife.85873 | Google Scholar

Author information

Hiroki Tsutsumi[#]
Department of Systems BioMedicine, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, Tokyo, Japan

[#]These authors contributed equally to this work.

Tomoki Chiba[#]
Department of Systems BioMedicine, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, Tokyo, Japan
ORCID iD: 0000-0001-5472-9030

[#]These authors contributed equally to this work.

Yuta Fujii
Department of Systems BioMedicine, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, Tokyo, Japan, Department of Molecular and Cellular Biology, The Scripps Research Institute, La Jolla, United States

Takahide Matsushima
Department of Systems BioMedicine, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, Tokyo, Japan

Tsuyoshi Kimura
Laboratory for Biomaterials and Bioengineering, Institute of Integrated Research, Institute of Science Tokyo, Tokyo, Japan, Department of Biomedical Engineering, Faculty of Life Science, Toyo University, Saitama, Japan

Akinori Kanai
Laboratory of Systems Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, the University of Tokyo, Chiba, Japan

Akio Kishida
Laboratory for Biomaterials and Bioengineering, Institute of Integrated Research, Institute of Science Tokyo, Tokyo, Japan

Yutaka Suzuki
Laboratory of Systems Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, the University of Tokyo, Chiba, Japan

Hiroshi Asahara
Department of Systems BioMedicine, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, Tokyo, Japan, Department of Molecular and Cellular Biology, The Scripps Research Institute, La Jolla, United States
ORCID iD: 0000-0002-5215-8745

For correspondence: asahara@scripps.edu

Editors

Reviewing Editor
Nathaniel Dyment
University of Pennsylvania, Philadelphia, United States of America

Senior Editor
Didier Stainier
Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany

Reviewer #1 (Public review):

This study is focused on identifying unique, innovative surface markers for mature Achilles tendons by combining the latest multi-omics approaches and in vitro evaluation, which would address the knowledge gap of controversial identity of TPSCs with unspecific surface markers. The use of multi-omics technologies, in vivo characterization, in vitro standard assays of stem cells, and in vitro tissue formation is a strength of this work and could be applied for other stem cell quantification in the musculoskeletal research. The evaluation and identification of Cd55 and Cd248 in TPSCs have not been conducted in tendon, which is considered as innovative. Additionally, the study provided solid sequencing data to confirm co-expressions of Cd55 and Cd248 with other well-described surface markers such as Ly6a, Tpp3, Pdgfra, and Cd34. Generally, the data shown in the manuscript support the claims that the identified surface antigens mark TPSCs in juvenile tendons.

https://doi.org/10.7554/eLife.104768.2.sa2

Reviewer #2 (Public review):

Summary:

The molecular signature of tendon stem cells is not fully identified. The endogenous location of tendon stem cells within native tendon is also not fully elucidated. Several molecular markers have been identified to isolate tendon stem cells but they lack tendon specificity. Using the declining tendon repair capacity of mature mice, the authors compared the transcriptome landscape and activity of juvenile (2 weeks) and mature (6 weeks) tendon cells of mouse Achilles tendons and identified CD55 and CD248 as novel surface markers for tendon stem cells. CD55+ CD248+ FACS-sorted cells display a preferential tendency to differentiate into tendon cells compared to CD55neg CD248neg cells.

Strengths:

The authors generated a lot of data of juvenile and mature Achilles tendons, using scRNAseq, snRNAseq, ATACseq strategies. This constitutes a resource datasets.

Weaknesses:

The analyses and validation of identified genes are not complete and could be pushed further. The endogenous expression of newly-identified genes in native tendons would be informative. The comparison of scRNAseq and snRNAseq datasets for tendon cell populations would strengthen the identification of tendon cell populations.

https://doi.org/10.7554/eLife.104768.2.sa1

Author Response:

The following is the authors' response to the original reviews

> ***Reviewer #1 (Public review):***
>
> *This study is focused on identifying unique, innovative surface markers for mature Achilles tendons by combining the latest multi-omics approaches and in vitro evaluation, which would address the knowledge gap of the controversial identity of TPSCs with unspecific surface markers. The use of multi-omics technologies, in vivo characterization, in vitro standard assays of stem cells, and in vitro tissue formation is a strength of this work and could be applied for other stem cell quantification in musculoskeletal research. The evaluation and identification of Cd55 and Cd248 in TPSCs have not been conducted in tendons, which is considered innovative. Additionally, the study provided solid sequencing data to confirm co-expressions of Cd55 and Cd248 with other well-described surface markers such as Ly6a, Tpp3, Pdgfra, and Cd34. Generally, the data shown in the manuscript support the claims that the identified surface antigens mark TPSCs in juvenile tendons.*
>
> *However, there are missing links between scientific questions aimed to be addressed in Introduction and Methodology/Results. If the study focuses on unsatisfactory healing responses of mature tendons and understanding of mature TPSCs, at least mature Achilles tendons from more than 12-week-old mice and their comparison with tendons from juvenile/neonatal mice should be conducted. However, either 2-week or 6-weekold mice, used for characterization here, are not skeletally mature, Additionally, there is a lack of complete comparison of TPSCs between 2-week and 6-week-old mice in the transcriptional and epigenetic levels.*
>
> *In order to distinguish TPSCs and characterize their epigenetic activities, the authors used scRNA-seq, snRNA-seq, and snATAC-seq approaches. The integration, analysis, and comparison of sequencing data across assays and/or time points is confusing and incomplete. For example, it should be more comprehensive to integrate both scRNA-seq and snRNA-seq data (if not, why both assays were used for Achilles tendons of both 2-week and 6-week timepoints). snRNA-seq and snATAC-seq data of 6-week-old mice were separately analyzed. No comparison of difference and similarity of TPSCs of 2-week and 6-week-old mice was conducted.*
>
> *Given the goal of this work to identify specific TPSC markers, the specificity of Cd55 and Cd248 for TPSCs is not clear. First, based on the data shown here, Cd55 and Cd248 mark the same cell population which is identified by Ly6a, TPPP3, and Pdgfra. Although, for instance, Cd34 is expressed by other tissues as discussed here, no data/evidence is provided by this work showing that Cd55 and Cd248 are not expressed by other musculoskeletal tissues/cells. Second, the immunostaining of Cd55 and Cd248 doesn't support their specificity. What is the advantage of using Cd55 and Cd248 for TPSCs compared to using other markers?*

Reviewer #2 (Public review):

Summary:

The molecular signature of tendon stem cells is not fully identified. The endogenous location of tendon stem cells within the native tendon is also not fully elucidated. Several molecular markers have been identified to isolate tendon stem cells but they lack tendon specificity. Using the declining tendon repair capacity of mature mice, the authors compared the transcriptome landscape and activity of juvenile (2 weeks) and mature (6 weeks) tendon cells of mouse Achilles tendons and identified CD55 and CD248 as novel surface markers for tendon stem cells. CD55+ CD248+ FACS-sorted cells display a preferential tendency to differentiate into tendon cells compared to CD55neg CD248neg cells.

Strengths:

The authors generated a lot of data on juvenile and mature Achilles tendons, using scRNAseq, snRNAseq, and ATACseq strategies. This constitutes a resource dataset.

Weaknesses:

The analyses and validation of identified genes are not complete and could be pushed further. The endogenous expression of newly identified genes in native tendons would be informative. The comparison of scRNAseq and snRNAseq datasets for tendon cell populations would strengthen the identification of tendon cell populations.

Reviewer #3 (Public review):

Summary:

In their report, Tsutsumi et al., use single nucleus transcriptional and chromatin accessibility analyses of mouse achilles tendon in an attempt to uncover new markers of tendon stem/progenitor cells. They propose CD55 and CD248 as novel markers of tendon stem/progenitor cells.

Strengths:

This is an interesting and important research area. The paper is overall well written.

Weaknesses:

Major problems:

(1) It is not clear what tissue exactly is being analyzed. The authors build a story on tendons, but there is little description of the dissection. The authors claim to detect MTJ and cartilage cells, but not bone or muscle cells. The tendon sheath is known to express CD55, so the population of "progenitors" may not be of tendon origin.

(2) Cluster annotations are seemingly done with a single gene. Names are given to cells without functional or spatial validation. For example, MTJ cells are annotated based on Postn, but it is never shown that Postn is only expressed at the MTJ, and not in other anatomical locations in the tendon.

(3) The authors compare their data to public data based on interrogating single genes in their dataset. It is now standard practice to integrate datasets (eg, using harmony), or at a minimum using gene signatures built into Seurat (eg AddModuleScore).

> *(4) Progenitor populations (SP1, SP2). The authors claim these are progenitors but show very clearly that they express macrophage genes. What are they, macrophages or fibroblasts?*
>
> *(5) All omics analysis is done on single data points (from many mice pooled). The authors make many claims on n=1 per group for readouts dependent on sample number (eg frequency of clusters).*
>
> *(6) The scRNAseq atlas in Figure 1 is made by analyzing 2W and 6W tendons at the same time. The snRNAseq and ATACseq atlas are built first on 2W data, after which the 6W data is compared. Why use the 2W data as a reference?*
>
> *Why not analyze the two-time points together as done with the scRNAseq?*
>
> *(7) Figure 5: The authors should show the gating strategy for FACS. Were non-fibroblasts excluded (eg, immune cells, endothelia...etc). Was a dead cell marker used? If not, it is not surprising that fibroblasts form colonies and express fibroblast genes when compared to CD55-CD248- immune cells, dead cells, or debris. Can control genes such as Ptprc or Pecam1 be tested to rule out contamination with other cell types?*
>
> *Minor problems:*
>
> *(1) Report the important tissue processing details: type of collagenase used. Viability before loading into 10x machine.*
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *(1) Better healing responses in neonatal mice than mature mice have been well appreciated in the field and differences in ECM environment, immune responses, and cell function might account for varied injury results. However, direct evidence/data between better healing and abundant TSPCs needs to be discussed in the Introduction.*

We agree with this insightful comment. We have now enhanced our introduction to include a more direct discussion of the relationship between better healing responses in neonatal mice and the abundance of TSPCs. We specifically highlighted how Howell et al. (2017) demonstrated that tendons in juvenile mice can regenerate functional tissue after injury, while this ability is lost in mature mice. Based on this observation, we articulated our hypothesis that juvenile mouse tendons likely contain abundant TSPCs, which potentially explains their superior healing capacity. Additionally, we have added a statement emphasizing that "investigating TSPCs biology is important for understanding tendon regeneration and homeostasis" (lines 61-62), which clearly articulates the central role that TSPCs play in tendon repair processes and tissue maintenance.

> *(2) 6-week-old mouse Achilles tendons are not mature enough and clinically relevant to understand the deficiency of regenerative capacity of TPSCs for undesired healing. If the goal of this study is to identify TSPCs of mature tendons, evaluation of Achilles tendons from at least 12-week-old mice is more reasonable.*

We agree with this insightful comment. We have now enhanced our introduction to include a more direct discussion of the relationship between better healing responses in neonatal mice and the abundance of TSPCs. We specifically highlighted how Howell et al. (2017) demonstrated that tendons in juvenile mice can regenerate functional tissue after injury, while this ability is lost in mature mice. Based on this observation, we articulated our hypothesis that juvenile mouse tendons likely contain abundant TSPCs, which potentially explains their superior healing capacity. Additionally, we have added a statement emphasizing that "investigating TSPCs biology is important for understanding tendon

regeneration and homeostasis" (lines 61-62), which clearly articulates the central role that TSPCs play in tendon repair processes and tissue maintenance.

> *(3) 40-60 mouse Achilles tendons pooled for one sample seems a lot and there is mixed/missed information about how many total cells were collected for each sample and how they were used for different sequencing assays. This could raise the concern that cell digestion was not complete and possibly abundant resident cells might be missed for sequencing analysis.*

We agree with this insightful comment. We have now enhanced our introduction to include a more direct discussion of the relationship between better healing responses in neonatal mice and the abundance of TSPCs. We specifically highlighted how Howell et al. (2017) demonstrated that tendons in juvenile mice can regenerate functional tissue after injury, while this ability is lost in mature mice. Based on this observation, we articulated our hypothesis that juvenile mouse tendons likely contain abundant TSPCs, which potentially explains their superior healing capacity. Additionally, we have added a statement emphasizing that "investigating TSPCs biology is important for understanding tendon regeneration and homeostasis" (lines 61-62), which clearly articulates the central role that TSPCs play in tendon repair processes and tissue maintenance.

> *(4) The methods section has necessary information missing, which could create confusion for readers. Which time points are used for scRNA-seq and snATAC-seq? Which time points of cells are integrated and analyzed regarding each assay/combined assays? Why is transcriptional expression evaluated by both scRNA-seq and snRNA-seq and is there any technological difference between the two assays?*

We have thoroughly revised the Methods section to clearly specify which time points were used for each assay (line 132-133 and line 148-149). We have also clarified how cells from different time points were integrated and analyzed (lines 167-170, 179-184 and 494-502). Regarding the use of both scRNA-seq and snRNA-seq, we have explained that this complementary approach allowed us to capture both cytoplasmic and nuclear transcripts, providing a more comprehensive view of gene expression profiles while also enabling direct integration with snATAC-seq data. Comparison of similarity between scRNA-seq integration data (2-week and 6-week) and snRNA-seq (2-week) clusters confirmed that the clusters in each data set are almost correlated. We added the dot plot and correlation data in supplemental figure 5. Additionally, we have included comprehensive lists of differentially expressed genes (DEGs) for each identified cluster across all datasets (supplementary tables 1-15), which provide detailed molecular signatures for each cell population and facilitate cross-dataset comparisons.

> *(5) snATAC-sequencing data seems to be used to only confirm the findings by snRNA-seq and snATAC-sequencing data is not well explored. This assay directly measures/predicts transcription factor activities and epigenetic changes, which might be more accurate in inferring transcription factors from RNA sequencing data using the R package SCENIC.*

We appreciate the reviewer's insightful comment regarding the utilization of our snATAC-seq data. We agree that snATAC-seq provides valuable direct measurements of chromatin accessibility and transcription factor binding sites that can complement inference-based approaches like SCENIC. To address this concern, we have revised our manuscript to better emphasize the value of our snATAC-seq data in transcription factor activity evaluation. We have modified our text (lines 570-574). This modification emphasizes that our integrated approach leverages the strengths of both methodologies, with snATAC-seq providing direct measurements of chromatin accessibility and transcription factor binding sites that can validate and enhance the inference-based predictions from SCENIC analysis of RNA-seq data.

> *(6) The image quality of immunostaining of Cd55 and Cd248 is low. The images show that only part of the tendon sheath has positive staining. Co-localization of Cd55 and Cd248 can't be found.*

We agree with the reviewer regarding the limitations of our immunostaining images. To obtain clearer images, we used paraffin sections for our analysis. Additionally, the antibodies for CD55 and CD248 required different antigen retrieval conditions to work effectively, which unfortunately prevented us from performing co-immunostaining to directly demonstrate co-localization. Despite these technical limitations, we have optimized the processing and imaging parameters to improve the quality of the immunostaining images in Figure 5A. These improved images more clearly demonstrate the expression of CD55 and CD248 in the tendon sheath, although in separate sections. The consistent localization patterns observed in these separate stainings, together with our FACS and functional analyses of double-positive cells, strongly support their co-expression in the same cell population. We have also updated the corresponding Methods section (lines 260-272) to include these optimized immunostaining protocols for better reproducibility.

> *(7) Only TEM data of tendon construct formed by sorted cells are shown. Results of mechanical tests will be super helpful to show the capacity of these TPSCs for tendon assembly.*

We appreciate the reviewer's suggestion regarding mechanical testing. We would like to direct the reviewer's attention to Figure 5I in our manuscript, where we have already included tensile strength measurements of the tendon construct. These mechanical test results demonstrate the functional capacity of CD55/CD248+ cells to form tendon-like tissue with appropriate mechanical properties, providing quantitative evidence of their ability for tendon assembly.

> *(8) Cells negative for CD55/CD248 could be mixed cell populations, including hematopoietic lineages, cells from tendon mid substance, immune cells, and/or endothelial cells. Under induction of tri-lineage media, these mixed cell populations could process different, unpredicted phenotypes (shown by no increased gene expression of tenogenic, chondrogenic, and osteogenic markers after induction). Higher tenogenic gene expressions of TPSCs after induction don't mean that TPSCs are induced into tenocytes if compared to unknown cell populations with/without similar induction. Additionally, PCR data in Figure 5 presented as ΔΔCT, with unclear biological meanings, is challenging to interpret.*

We appreciate the reviewer's suggestion regarding mechanical testing. We would like to direct the reviewer's attention to Figure 5I in our manuscript, where we have already included tensile strength measurements of the tendon construct. These mechanical test results demonstrate the functional capacity of CD55/CD248+ cells to form tendon-like tissue with appropriate mechanical properties, providing quantitative evidence of their ability for tendon assembly.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *The aim of this study was to identify novel markers for tendon stem cells. The authors used the fact that tendon cells of juvenile tendons have a greater ability to regenerate versus mature tendons. scRNAseq, snRNAseq, and snATACseq datasets were generated and analyzed in juvenile and mature Achilles tendons (mice).*
>
> *The authors generated a lot of data that could be exploited further to show that these two novel surface tendon markers are more tendon-specific than those previously*

> *identified. Another concern is that there is no robust data indicative of the endogenous location of CD55+ CD248+ cells in the native tendon. Same comments for the transcription factors regulating the transcription of CD55 and CD248 and that of Scx and Mkx. A validation of the ATACseq data with a location in native tendons would be pertinent.*
>
> *The analysis was performed by comparing 2 sub-clusters of the same datasets and not between the two stages. Given the introduction highlighting the differential ability to regenerate between the two stages, the comparison between the two stages was somehow expected. I wonder if there is an explanation for the absence of analysis between the two stages.*
>
> *The authors have all the datasets to (bioinformatically) compare scRNAseq and snRNAseq datasets. This comparative analysis would strengthen the clustering of tendon cell populations at both stages. The labeling/identification of clusters associated with tendon cell populations is not obvious. I am surprised that there is no tendon sheath cluster such as endotenon or peritenon. A discussion on the different tendon cell populations (tendon clusters) is lacking.*
>
> *(1) Choice of the three markers*
>
> *The authors chose three genes known to be markers for tendon stem cells, Tppp3, PdgfRa, and Ly6a, and investigated clusters (or subclusters) that co-express these three genes. Except for Tppp3, the other two genes lack tendonspecificity. Ly6a is a stem cell marker and is recognized to be a marker of epi/perimysium in fetal and perinatal stages in mouse limbs (PMID: 39636726). Pdgfra is a generic marker of all connective tissue fibroblasts. Could it be that the identification of the two novel surface markers was biased with this choice? The identification of CD55 and CD248 has been done by comparing DEGs between cluster 4 (SP2) and cluster 1 (SP1). What about an unbiased comparison of both clusters 4 and 1 (or individual clusters) between mature and juvenile samples? The reader expected such a comparison since it was introduced as the rationale of the paper to compare juvenile and mature tendon cells.*

We selected Tppp3, PdgfRa, and Ly6a based on established literature identifying them as TSPC markers (Harvey et al., 2019; Tachibana et al., 2022). While only Tppp3 has tendon specificity, these genes collectively represent reliable TSPC markers currently available.

Our identification of CD55 and CD248 came from comparing SP2 and SP1 clusters that showed these three markers plus tendon development genes. We did compare juvenile and mature samples as shown in Figure 1G, revealing decreased stem/progenitor marker expression with maturation. Additionally, we performed a comprehensive comparison between 2-week and 6-week samples visualized as a heatmap in Supplemental Figure 3, which clearly demonstrates the transcriptional changes that occur during tendon maturation. We have also provided the complete lists of differentially expressed genes for each identified cluster

(supplementary tables 1-15), allowing for unbiased examination of cluster-specific gene signatures across developmental stages.

Our functional validation confirmed CD55/CD248 positive cells express Tppp3, PdgfRa, and Ly6a while demonstrating high clonogenicity and tenogenic differentiation capacity, confirming their TSPC identity.

> *(2) Concerns with cluster identification*
>
> *The cluster11, named as MTJ cluster, in 2-week scRNAseq datasets was not detected in 6-week scRNAseq datasets (Figure 1A). Does it mean that MTJ disappears at 6 weeks in*

> *Achilles tendons? In the snRNAseq MTJ cluster was defined on the basis of Postn expression. «Cluster 11, with high Periostin (Postn) expression, was classified as a myotendinous junction (MTJ).» Line 379.*
>
> *What is the basis/reference to set a link between Postn and MTJ?*
>
> *Could the CA clusters be enthesis clusters? Is there any cartilage in the Achilles tendon?*
>
> *If there are MTJ clusters, one could expect to see clusters reflecting tendon attachment to cartilage/bone.*
>
> *I am surprised to see no cluster reflecting tendon attachments (endotenon or peritenon).*
>
> *Cluster 9 was identified as a proliferating cluster in scRNAseq datasets. Does the Cell Cycle Regression step have been performed?*

Thank you for highlighting these important questions about our cluster identification. The MTJ cluster (cluster 11) appears reduced but not absent in 6-week samples. We based our MTJ classification on Postn expression, which is enriched at the myotendinous junction, as documented by Jacobson et al. (2020) in their proteome analysis of myotendinous junctions. We have added this reference to the manuscript to provide clear support for our cluster annotation (lines 400-401).

Regarding the CA cluster, these cells express chondrogenic markers but are not enthesis clusters. We have revised our manuscript to acknowledge that these could potentially represent enthesis cells, as you suggested (lines 412-414). While Achilles tendons themselves don't contain cartilage, our digestion process likely captured some adjacent cartilaginous tissues from the calcaneus insertion site.

We acknowledge the absence of clearly defined endotenon/epitenon clusters. We have added more comprehensive explanations about peritenon tissues in our manuscript (lines 431-433 and 584-585), noting that previous studies (Harvey et al., 2019) have reported that *Tppp3*-positive populations are localized to the peritenon, and our SP clusters might also reflect peritenon-derived cells. This additional context helps clarify the potential tissue origins of our identified cell populations.

For the proliferating cluster (cluster 9), we confirmed high expression of cell cycle markers (Mki67, Stmn1) but did not perform cell cycle regression to maintain biological relevance of proliferation status in our analysis. We have clarified this methodological decision in the revised Methods section.

> *(3) What is the meaning of all these tendon clusters in scRNAseq snRNAseq and snATACseq? The authors described 2 or 3 SP clusters (depending on the scRNAseq or snRNAseq datasets), 2 CT clusters, 1 MTJ cluster, and 1CA cluster. Do genes with enriched expression in these different clusters correspond to different anatomical locations in native tendons? Are there endotenon and peritenon clusters? Is there a correlation between clusters (or subclusters) expressing stem cell markers and peritenon as described for Tppp3*

Thank you for this important question about the biological significance of our identified clusters. The multiple tendon-related clusters we identified likely represent distinct cellular states and differentiation stages rather than strictly discrete anatomical locations. The SP clusters (stem/progenitor cells) express markers consistent with tendon progenitors reported in the literature, including Tppp3, which has been described in the peritenon. As we mentioned in our response to the previous question, we have added more comprehensive explanations about peritenon tissues in our manuscript (Lines 432-433 and 584-585), noting that previous studies (Harvey et al., 2019) have reported that Tppp3-positive populations are

localized to the peritenon, and our SP clusters might reflect peritenon-derived cells. Our immunohistochemistry data in Figure 5A further confirms that CD55/CD248 positive cells are localized primarily to the tendon sheath region, similar to the localization pattern of Tppp3 reported by Harvey et al. (2019). The tenocyte clusters (TC) represent mature tendon cells within the fascicles, and their distinct transcriptional profiles suggest heterogeneity even within mature tenocytes. The MTJ cluster specifically expresses genes enriched at the myotendinous junction, while the CA cluster likely represents cells from the enthesis region, as you suggested. In the revised manuscript, we have clarified this interpretation and added additional discussion about the relationship between cluster identity and anatomical localization, particularly regarding the SP clusters and their correlation with peritenon regions.

> *(4) The use of single-cell and single-nuclei RNAseq strategies to analyze tendon cell populations in juvenile and mature tendons is powerful, but the authors do not exploit these double analyses. A comparison between scRNAseq and snRNAseq datasets (2 weeks and 6 weeks) is missing. The similar or different features at the level of the clustering or at the level of gene expression should be explained/shown and discussed. This analysis should strengthen the clustering of tendon cell populations at both stages. In the same line, why are there 3 SP clusters in snRNAseq versus 2 SP clusters in scRNAseq? The MTJ cluster R2-5 expressing Sox9 should be discussed.*

Thank you for highlighting this important gap. We have conducted a comprehensive comparison between scRNA-seq and snRNA-seq datasets, revealing substantial correlation between cell populations identified by both methodologies. We've added a detailed dot plot visualization and correlation heatmap in Supplemental Figure 5 that demonstrates the relationships between clusters across datasets. The additional SP cluster in snRNA-seq likely reflects the greater sensitivity of nuclear RNA sequencing in capturing certain cell states that might be missed during whole-cell isolation. Our analysis shows this SP3 cluster represents a transitional state between stem/progenitor cells and differentiating tenocytes. Regarding the Sox9-expressing MTJ cluster R2-5, we have expanded our discussion in the revised manuscript (lines 500502) to address this finding, incorporating relevant references (Nagakura et al., 2020) that describe Sox9 expression at the myotendinous junction. This expression pattern suggests that cells at this specialized interface may maintain developmental plasticity between tendon and cartilage fates, which is consistent with the transitional nature of this anatomical region.

> *(5) The claim of "high expression of CD55 and CD248 in the tendon sheath" is not supported by the experiments. The images of immunostaining (Figure 5A) are not very convincing. It is not explained if these are sections of 3Dtendon constructs or native tendons. The expression in 3D-tendon constructs is not informative, since tendon sheaths are not present. The endogenous expression of the transcription factors regulating tendon gene expression would be informative to localize tendon stem cells in native tendons.*

Thank you for this important critique. We agree that the original immunostaining images were not sufficiently convincing. To address this, we have used paraffin sections and optimized our staining protocols to improve image quality. It's worth noting that CD55 and CD248 antibodies required different antigen retrieval conditions to work effectively, which unfortunately prevented us from performing coimmunostaining to directly demonstrate co-localization in the same section. Despite these technical limitations, we have significantly improved the quality of the immunostaining images in Figure 5A with enhanced processing and imaging parameters

The improved images more clearly demonstrate the preferential expression of CD55 and CD248 in the tendon sheath/peritenon regions. The consistent localization patterns observed in these separate stainings, together with our FACS and functional analyses of double-positive cells, strongly support their coexpression in the same cell population.

In the revised manuscript, we have also improved the figure legends to clearly indicate the nature of the tissue samples and updated the methods section to provide more detailed protocols for the immunostaining procedures used.

Your suggestion regarding transcription factor visualization is valuable. While beyond the scope of our current study, we agree that examining the endogenous expression of regulatory transcription factors like Klf3 and Klf4 would provide additional insights into tendon stem cell localization in native tendons, and we plan to pursue this in future work

> *Minor concerns:*
>
> *(1) Lines 392-397 « To identify progenitor populations within these clusters, we analyzed expression patterns of previously reported markers Tppp3 and Pdgfra (Harvey et al., 2019; Tachibana, et al., 2022), along with the known stem/progenitor cell marker Ly6a (Holmes et al., 2007; Sung et al., 2008; Hittinger et al., 2013; Sidney et al., 2014; Fang et al., 2022). We identified subclusters within clusters 1 and 4 showing high expression of these genes, which we defined as SP1 and SP2. SP2 exhibited the highest expression of these genes, suggesting it had the strongest progenitor characteristics.» Please cite relevant Figures. Feature and violin plots (scRNAseq) across all cells (not for the only 2 SP1 and SP2 clusters) of Tppp3, Pdgfra and Ly6a are missing.*

Thank you for pointing out this important oversight. We have modified the manuscript to clarify that the text in question describes Figure 1B. Additionally, we have added new feature plots showing the expression of Tppp3, Pdgfra, and Ly6a across all cells in supplymental figure 1B

> *(2) The labeling of clusters with numbers in single-cell, single nuclei RNAseq, and ATACseq is difficult to follow.*

We appreciate your feedback on this issue. We recognize that the numerical labeling system across different datasets (scRNA-seq, snRNA-seq, and snATAC-seq) makes it difficult to track the same cell populations. To address this, we have added Supplemental Figure 5, which clearly shows the correspondence between cell populations in single-cell and single-nucleus RNA-seq datasets.

> *(3) Figure 1C. It is not clear from the text and Figure legend if the DEGs are for the merged 2 and 6 weeks. If yes, an UMAP of the merged datasets of 2 and 6 weeks would be useful.*

We appreciate your feedback on this issue. We recognize that the numerical labeling system across different datasets (scRNA-seq, snRNA-seq, and snATAC-seq) makes it difficult to track the same cell populations. To address this, we have added Supplemental Figure 5, which clearly shows the correspondence between cell populations in single-cell and single-nucleus RNA-seq datasets.

> *(4) Along the Text, there are a few sentences with obscure rationale. Here are a few examples (not exhaustive):*
>
> *Abstract*

> *« Combining single-nucleus ATAC and RNA sequencing analyses revealed that Cd55 and Cd248 positive fractions in tendon tissue are TSPCs, with this population decreasing at 6 weeks. »*
>
> *The rationale of this sentence is not clear. How can single-nucleus ATAC and RNA sequencing analyses identify Cd55 and Cd248 positive fractions as tendon stem cells?*

Thank you for highlighting this unclear statement in our abstract. We agree that the previous wording did not adequately explain how our sequencing analyses identified CD55 and CD248 positive cells as TSPCs. We have revised this sentence to clarify that our multi-modal approach (combining scRNA-seq, snRNA-seq, and snATAC-seq) enabled us to identify *Cd55* and *Cd248* positive populations as TSPCs based on their co-expression with established TSPC markers such as *Tppp3*, *Pdgfra*, and *Ly6a*. This comprehensive analysis across different sequencing modalities provided strong evidence for their identity as tendon stem/progenitor cells, which we further validated through functional assays. The revised abstract now more clearly communicates the logical progression of our analysis and findings

> *Line 80-82*
>
> *« Cd34 is known to be highly expressed in mouse embryonic limb buds at E14.5 compared to E11.5 (Havis et al., 2014), making it a potential marker for TSPCs. »*
>
> *The rationale of this sentence is not clear. How can "the fact to be expressed in E14.5 mouse limbs" be an indicator of being a "potential marker of tendon stem cells"?*

Thank you for highlighting this unclear statement in our abstract. We agree that the previous wording did not adequately explain how our sequencing analyses identified CD55 and CD248 positive cells as TSPCs. We have revised this sentence to clarify that our multi-modal approach (combining scRNA-seq, snRNA-seq, and snATAC-seq) enabled us to identify *Cd55* and *Cd248* positive populations as TSPCs based on their co-expression with established TSPC markers such as *Tppp3*, *Pdgfra*, and *Ly6a*. This comprehensive analysis across different sequencing modalities provided strong evidence for their identity as tendon stem/progenitor cells, which we further validated through functional assays. The revised abstract now more clearly communicates the logical progression of our analysis and findings

> *Line 611*
>
> *« Recent reports have highlighted the role of the Klf family in limb development (Kult et al., 2021), suggesting its potential importance in tendon differentiation »*
>
> *Why does the "role of Klf family in limb development" suggest an "importance in tendon differentiation"?*

Thank you for highlighting this logical gap in our manuscript. You're right that involvement in limb development doesn't necessarily indicate specific importance in tendon differentiation. We've revised this statement to more accurately reflect current knowledge, noting that while Klf factors are involved in limb development, their specific role in tendon differentiation requires further investigation (lines 658-659). This revised text better aligns with our findings of Klf3 and Klf4 expression in tendon progenitor cells without making unsupported claims about their functional significance

> ***Reviewer #3 (Recommendations for the authors):***
>
> *In addition to the points highlighted above some additional points are listed below.*

> *(1) Case in point: the authors claim CD55 and CD248 are found at the tendon sheath (line 541), which is not part of the tendon proper (although the IHC seems to show green in the epi/endotenon).*

Thank you for highlighting this logical gap in our manuscript. You're right that involvement in limb development doesn't necessarily indicate specific importance in tendon differentiation. We've revised this statement to more accurately reflect current knowledge, noting that while Klf factors are involved in limb development, their specific role in tendon differentiation requires further investigation (lines 658-659). This revised text better aligns with our findings of Klf3 and Klf4 expression in tendon progenitor cells without making unsupported claims about their functional significance

> *(2) All cell types seem to express collagen based on Figure 1B, so either there is serious background contamination (eg, ambient RNA), or an error in data analysis.*

Thank you for highlighting this logical gap in our manuscript. You're right that involvement in limb development doesn't necessarily indicate specific importance in tendon differentiation. We've revised this statement to more accurately reflect current knowledge, noting that while Klf factors are involved in limb development, their specific role in tendon differentiation requires further investigation (lines 658-659). This revised text better aligns with our findings of Klf3 and Klf4 expression in tendon progenitor cells without making unsupported claims about their functional significance

> *Minor problems:*
>
> *(1) The figures are confusingly formatted. It is hard to go between cluster numbers and names. Clusters of similar cell types (eg progenitors) are not grouped to facilitate comparison, as ordering is based on cluster number).*

Thank you for highlighting this logical gap in our manuscript. You're right that involvement in limb development doesn't necessarily indicate specific importance in tendon differentiation. We've revised this statement to more accurately reflect current knowledge, noting that while Klf factors are involved in limb development, their specific role in tendon differentiation requires further investigation (lines 658-659). This revised text better aligns with our findings of Klf3 and Klf4 expression in tendon progenitor cells without making unsupported claims about their functional significance

> *(2) The introduction does not distinguish between findings in mice and man. A lot of confusion in the tendon literature probably arises from interspecies differences, which are rarely addressed.*

We appreciate this important point about species distinctions. We have revised our introduction to clearly identify species-specific findings by adding the term "murine" before TSPC references when discussing mouse studies (lines 64, 66, 70, 75, 100, and 108). We agree that interspecies differences are important considerations in tendon biology research, particularly when translating findings between animal models and humans. Our study focuses specifically on mouse models, and we have been careful not to overgeneralize our conclusions to human tendon biology without appropriate evidence. This clarification helps readers better contextualize our findings within the broader tendon literature landscape.

https://doi.org/10.7554/eLife.104768.2.sa0