Published in final edited form as:

J Orthop Res. 2015 June ; 33(6): 840–848. doi:10.1002/jor.22886.

Whole transcriptome expression profiling of mouse limb tendon development by using RNA-seq

Han Liu¹, Jingyue Xu¹, Chia-Feng Liu^{1,3}, Yu Lan^{1,2}, Christopher Wylie¹, and Rulang Jiang^{1,*}

¹Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA

²Division of Plastic Surgery, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA

Abstract

Tendons are fibrous connective tissues that transmit force between muscle and bone. Whereas the molecular and cellular mechanisms of bone and muscle development have been well studied, that of tendon development is poorly understood. Using the *Scx-GFP* transgenic mice, we isolated GFP⁺ cells from the developing mouse limbs at E11.5, E13.5, and E15.5, respectively, and carried out whole transcriptome RNA-seq analysis. Comparing the gene expression profiles of GFP⁺ and GFP⁻ cells in the E13.5 limb isolated over 1500 genes that exhibited enrichment of mRNA expression by at least 1.5-fold in the GFP⁺ cells. Of these, 778 genes showed expression up-regulated by more than 1.5-fold from E11.5 to E13.5 and 516 genes showed expression up-regulated by more than 1.5-fold from E13.5 to E15.5 in the GFP⁺ cell population. Interestingly, over 30 genes encoding transcription factors are among the early-activated genes in the GFP⁺ cells. Whole mount and section in situ hybridization analyses showed that many of these transcription factor genes have distinct patterns of expression during limb development and identified *Foxf2* expression as a specific marker for differentiated dorsal limb tendon cells. Together, these data provide a valuable resource for further investigation of the molecular mechanisms regulating tendon development.

Keywords

tendon; RNA-seq; *Scx*; *Foxf2*; *Six2*

Introduction

Tendons and ligaments provide physical and mechanical connections between muscle and bone, and contribute to the complex architectures of the musculoskeletal system. Mechanical

*Author for correspondence: Rulang Jiang, PhD, Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, MLC 7007, Cincinnati, OH 45229, USA, Phone: 513-636-3212, Rulang.Jiang@cchmc.org.

³Present address: Department of Cellular and Molecular Medicine, Cleveland Clinic Lerner Research Institute, Cleveland, OH 44195, USA

Author Contributions: HL, JX, CFL, YL, CW and RJ designed the experiments. HL, JX and CFL performed sample preparation and RNA-seq experiments. HL and JX performed data analysis and in situ hybridization assay. HL, JX, and RJ prepared initial draft manuscript. HL, JX, CFL, YL, CW and RJ edited and approved manuscript.

forces generated by muscle contractions are transmitted through tendons to drive movement of the skeletal system. During force transmission, tendons and ligaments withstand stress and tension and, thus, are prone to acute injuries. In addition, chronic stress overload and aging related tissue degeneration cause the common tendon disorder known as tendinopathy. Most acute tendon injuries and some chronic tendon diseases require surgical tissue repair, which is clinically challenging since tendon healing is slow and complete recovery is rarely achieved. To improve tendon treatment strategies, recent research has started to focus on understanding the basic mechanisms of tendon development and tendon cell differentiation.

Tendon development initiates with the specification of tendon progenitor cells within the somite, limb mesoderm, and craniofacial mesenchyme, respectively. Tendon progenitors condense and reside between osteoblasts and myoblasts, where they undergo rapid proliferation and differentiate into tenocytes. Tenocytes synthesize various extracellular matrix molecules, including predominantly type I collagen, type III collagen, as well as the small leucine-rich proteoglycans decorin, biglycan, fibromodulin, and lumican, and organize them to form tendon fibrils¹. Highly oriented tendon fibrils are organized into more complex hierarchical structures, tendon fibers, which are surrounded by connective tissue sheath, called epitenon²⁻⁴. Although the anatomy of tendons has been relatively well characterized, mechanistic studies of tendon development have been hampered by the lack of tendon-specific molecular markers at distinct stages of tendon cell differentiation. The discovery of tendon-specific expression of the helix-loop-helix class transcription factor *Scx* in 2001 significantly facilitated molecular and genetic studies of tendon development⁵. Pryce et al. (2007) demonstrated that an 11-kb genomic DNA containing the *Scx* gene is sufficient to drive transgenic reporter expression in all tendon cell lineages, including craniofacial tendon cells derived from the neural crest, axial tendon cells derived from the somites, and limb tendon cells derived from lateral plate mesoderm⁶. Although *Scx* is expressed in the early tendon progenitor cells and its expression persists throughout tendon development, it is not required for tendon cell specification as mice lacking *Scx* exhibit defects in development of specific subpopulations of tendons, i.e., the force transmitting tendons⁷. Two other transcription factors, Mohawk homeobox (*Mkx*) and early growth response 1 (*Egr1*), have since been shown to play critical roles in tendon development as well. *Mkx* is expressed in differentiating tendon cells and regulates tendon fibril growth and organization postnatally^{8,9}. *Egr1* is expressed in differentiated tendon cells and plays an important role in the tendon healing process in damaged tendon tissues¹⁰.

To gain better understanding of the molecular mechanisms of tendon development, we have carried out whole transcriptome RNA-seq analysis of gene expression profiles of limb tendon cells in mice. We report here the RNA-seq datasets and identification of additional transcription factors with distinct patterns of expression during limb development.

Materials and Methods

Mice

The *Scx-GFP* transgenic mice have been described previously⁶ and generously provided by Dr. Ronen Schweitzer (Shriners Hospital for Children, Portland, Oregon). The transgenic mice are maintained by crossing to wildtype CD1 mice. All animal procedures were

approved by the Institutional Animal Care and Use Committee (IACUC) at Cincinnati Children's Hospital Medical Center.

Fluorescence-activating cell sorting (FACS)

The limbs of E11.5, E13.5 and E15.5 *Scx-GFP* transgenic embryos were harvested and the forelimbs and hindlimbs, respectively, were pooled from multiple embryos (nine E11.5 embryos, six E13.5 embryos, six E15.5 embryos). The freshly dissected embryonic limbs were digested with the trypsin-EDTA solution (Invitrogen) at 37°C for 4 minutes. After inactivation of trypsin with DMEM containing 10% FBS, cells were dissociated by pipetting. The dissociated limb cells were suspended in PBS with 2% FBS and 10 mM EDTA, and filtered through a 40 µm nylon cell strainer (BD Falcon, 352340). GFP⁺ cells from E11.5, E13.5, and E15.5 hindlimbs were isolated using BD FACSAria II. In addition, both GFP⁺ and GFP⁻ cells from the pooled E13.5 forelimbs were isolated using BD FACSAria II.

RNA-seq and data analysis

Total RNAs were extracted from pooled FACS sorted cells of three limbs for each stage studied, using the Qiagen RNeasy Micro Kit (Qiagen catalog# 74004). Sequencing libraries were generated by using Illumina Nextera DNA Sample Prep Kit and sequenced using Illumina HiSeq 2000. Sequenced reads were mapped to the reference mouse genome (mm9) using Bowtie version 0.12.7. Single-end reads were aligned using Tophat version 1.4.1 for Illumina¹¹. RNAseq data were analyzed using Avadis NGS software, with the fragments per kilobase exon per million mapped sequences (FPKM) value calculated for each RefSeq gene for relative gene expression. For analyses of differential expression, the fold change cut-off was set at 1.5-fold or higher and P-values <0.01 from the AudicClaverie test were considered to be statistically significant, with Benjamini Hochberg FDR multiple testing correction¹¹. Functional annotations analysis was carried out using online tools at <https://toppgene.cchmc.org>, as previously described^{12, 13}.

In situ hybridization assays

The plasmid containing the *Scx* cDNA for synthesis of cRNA probes for *in situ* hybridization has been described previously⁷ and generously provided by Dr. Ronen Schweitzer (Shriners Hospital for Children, Research Division, Portland, Oregon). Other cDNA templates were amplified by PCR from E11.5 mouse embryonic limb cDNAs using gene specific primers (Supplementary Table 1) and subcloned in the pBluescript II KS+ plasmid vector. All cRNA probes were tested for specificity by whole mount *in situ* hybridization of E12.5 mouse embryos and cDNA templates that did not result in tissue specific expression patterns were either validated with a second, non-overlapping, template or replaced by a more specific template.

For whole mount *in situ* hybridization, embryos were harvested from wildtype CD1 mice at predetermined gestational stages and fixed in 4% paraformaldehyde in PBS overnight at 4°C, dehydrated through methanol series and directly used for experiments. For section *in situ* hybridization, embryos were fixed in 4% paraformaldehyde in PBS overnight at 4 °C, dehydrated through ethanol series and embedded in paraffin. Serial sections of 7-µm

thickness were used for experiments. *In situ* hybridization assays were performed as previously described¹⁴.

Results

Characterization of whole transcriptome expression profiles of limb tendon development by using RNA-seq

The *Scx-GFP* transgenic mice have been shown exhibit strong GFP expression throughout tendon development⁶. However, it has also been reported that the *Scx-GFP* transgene is much more broadly expressed than the endogenous *Scx* mRNA in the early developing limbs, with the difference been most pronounced in E10.5 and E12.5⁶. As limb tendon progenitor cells start to differentiate from E12.5 to E13.5¹⁵, we first compared directly the patterns of GFP expression in *Scx-GFP* transgenic embryos with that of endogenous *Scx* mRNA expression in wildtype embryos at E13.5. We found that the pattern of *Scx-GFP* is highly similar to that of *Scx* mRNA expression in the autopod region while GFP expression domain is broader than the endogenous *Scx* mRNA expression in the proximal limb regions at this stage (Fig. 1A–H). We next performed quantitative real-time RT-PCR analysis of expression of select markers for tendon and other cell types in GFP⁺ and GFP[−] cells isolated from microdissected E13.5 forelimbs by using fluorescence activated cell sorting and found that expression of tendon marker genes, including *Scx*, *tenomodulin* (*Tnmd*), and *Col1a1*, were highly enriched, whereas expression of the chondrocyte marker *Sox9* and of the osteoblast marker *Runx2* were not, in the GFP⁺ cells in comparison with the GFP[−] cells in the E13.5 forelimb (Fig. 1I). These results suggest that comparative gene expression profiling analysis of the GFP⁺ versus GFP[−] cells in the E13.5 limb tissues of *Scx-GFP* transgenic embryos is a valid approach to identify early tendon developmental regulators.

Total RNAs were prepared from FACS-isolated E13.5 forelimb cells and processed for whole-transcriptome RNA-seq. The resulting data were analyzed using the Avadis NGS software, with the fragments per kilobase exon per million mapped sequences (FPKM) value calculated for each RefSeq gene for relative levels of gene expression. With the FPKM cut-off threshold at 5.0, transcripts corresponding to 11035 genes were included for analysis of differential expression in the GFP⁺ versus GFP[−] cells. Of these, transcripts of 1555 genes exhibited more than 1.5 fold higher levels of expression in the GFP⁺ than the GFP[−] cells. Gene ontology (GO) analysis of these 1555 genes showed that the GO terms related to tendon, including “skeletal system development”, “connective tissue development”, “cartilage development”, “collagen binding”, “collagen fibril organization”, “proteoglycan binding”, “abnormal joint morphology”, “abnormal tendon morphology”, are highly represented (Supplementary Table 2).

Although expression of many tendon marker genes is significantly enriched in the GFP⁺ cells in the E13.5 forelimb of *Scx-GFP* transgenic embryos (Supplementary Table 3), many of the GFP⁺ cells are not tendon cells because the *Scx-GFP* transgene is expressed more strongly and in broader domains than the endogenous *Scx* mRNAs⁶. To further facilitate identification of early tendon developmental regulators, we performed RNA-seq of FACS-isolated GFP⁺ cells from E11.5, E13.5, and E15.5 hindlimb tissues, respectively, of *Scx-GFP* transgenic embryos. At E11.5, the limb tendon progenitor cells are beginning to

commit to the tendon cell fate and both the *Scx-GFP* transgene and the endogenous *Scx* mRNAs are expressed in broad domains in the developing limb buds. By E13.5, *Scx*-expressing tendon tissues are morphologically distinguishable from surrounding tissues in the autopod and zeugopod regions in both fore- and hind-limbs¹⁶. By E15.5, most limb tendon progenitor cells have differentiated into tenocytes and expression of the *Scx-GFP* transgenic reporter becomes highly restricted to the tendon tissues^{6, 16}. Thus, we reasoned that important positive regulators of tendon development should be up-regulated from E11.5 to E13.5 and some positive regulators as well as markers of tendon cell differentiation should be up-regulated from E13.5 to E15.5. Indeed, the levels of expression of several known marker genes for tendon differentiation, including *Tnmd*, *lumican (Lum)*, *fibromodulin (Fmod)*, *decorin (Dcn)*, *Col1a1*, *Col1a2*, *Col3a1*, *Col6a1* and *Col6a2*, were significantly increased from E11.5 to E13.5 and E15.5, even though the relative abundance of endogenous *Scx* mRNAs decreased slightly from E13.5 to E15.5, in our GFP⁺ RNA-seq datasets (Fig. 2A), consistent with enrichment of the tendon cell lineage in the GFP⁺ cell population.

Differential expression analysis of the RNA-seq datasets showed that over 1700 genes exhibited increased levels of mRNA expression by more than 1.5-fold in the GFP⁺ populations from E11.5 to E13.5 (Fig. 2B). Of these, 778 were in the GFP⁺ cell enriched dataset from the E13.5 forelimb tissues, suggesting that they were activated in tendon progenitor cells during early tendon development. Expression of 379 of these genes were further up-regulated by more than 1.5-fold from E13.5 to E15.5, indicating that they are good tendon differentiation markers. GO analysis showed that many of these genes encode extracellular matrix components or proteins interacting with extracellular matrix (Table 1), whereas many of the other 399 genes in the 778 gene-set are involved in regulation of cell development or differentiation (Table 2). The top 20 enriched genes in each pair-wise comparison of the RNA-seq datasets from different stages are included in Supplementary Tables 4 – 6.

Characterization of expression of transcription factors in early tendon development

One problem hindering tendon research is that only three transcription factors, *Scx*, *Mkx*, and *Egr1*, have been shown to play important roles in tendon cell differentiation. Thus, we focused on identifying transcription factors that are up-regulated during early tendon development. We found 31 transcription factor genes, including *Mkx*, whose expression levels were enriched in the GFP⁺ forelimb cells at E13.5 and also significantly up-regulated, by more than 1.5-fold, from E11.5 to E13.5 in the FACS-isolated GFP⁺ hindlimb cells (Table 3). Of these, 13 were further up-regulated by more than 1.5-fold from E13.5 to E15.5 in the FACS-isolated GFP⁺ hindlimb cells. To further characterize their expression patterns during tendon development, we performed both whole mount and section in situ hybridization analyses. These analyses revealed that many of these transcription factor genes were broadly expressed during limb development (Supplementary Fig. 1), with only a few showing tendon-specific expression. For example, three members of the early B-cell factor (*Ebf*) family are among the 31 transcription factor genes significantly enriched in the *Scx*-GFP⁺ forelimb cells at E13.5 and significantly up-regulated from E11.5 to E13.5 in the GFP⁺ hindlimb cells. Upon examination by whole mount in situ hybridization analysis, we

found that, although all three *Ebf* genes were expressed in the developing limb mesenchyme, they were not specific for tendon (Fig. 3A–U, and Supplementary Fig. 2). Comparison of their expression with that of *Scx* by using section in situ hybridization on adjacent cross sections of the E13.5 forelimbs showed that expression of *Ebf1* and *Ebf3* was concentrated in the connective tissues surrounding the digit cartilage but absent from the dorsal and ventral limb tendons (Fig. 3J–U). Expression of *Ebf2* was partially overlapping with that of *Scx* mRNAs and was mostly in the connective tissues in between the cartilage and the tendon tissues located dorsally and ventrally (Fig. 3L, P, T, compare to *Scx* expression in Fig. 3J, N, and R, respectively). These results are consistent with a recent report of expression patterns of the *Ebf* family genes during mouse limb development¹⁸. On the other hand, we found that expression of *Six2* and *Foxf2* are highly specific in limb tendon development (Fig. 4, and Supplementary Fig. 3). Although a previous study suggested that both *Six1* and *Six2* are expressed during limb tendon development¹⁹, our differential expression analysis of the RNA-seq datasets did not pick up *Six1*. Direct comparison of the expression patterns of *Six1* and *Six2* with that of *Scx* and *Tnmd* confirmed that *Six1* is primarily expressed in the perichondrium and skeletal muscle in the developing limb whereas *Six2* is highly expressed in the differentiating limb tendon cells by E13.5 (Fig. 4A – Ab, and Supplementary Fig. 3), with much stronger expression in the ventral than the dorsal limb tendons (Fig. 4, K, O, S, W, Aa). Our in situ hybridization results showing lack of *Six1* expression in the developing mouse limb tendon are consistent with another study that showed non-overlapping patterns of expression of *Six1* and *Scx* mRNAs in the developing mouse limbs²⁰. Remarkably, we found that *Foxf2* expression is absent in the early developing limb prior to E12.5 and is only activated in the differentiating dorsal limb tendon cells, but not in ventral limb tendon, at E13.5 (Fig. 4D, H, L, P, T, X, Ab). Consistent with the RNA-seq data showing continued up-regulation of *Foxf2* mRNA expression from E13.5 to E15.5 (Table 3), in situ hybridization analysis detected a graded proximal-distal pattern of *Foxf2* mRNA expression in the E13.5 limbs, with strongest mRNA signals in the digit region (Fig. 4, T, X, Ab). These data identify *Foxf2* expression as a specific marker for differentiated dorsal limb tendon.

Discussion

The discovery of *Scx* as a tendon lineage marker in 2001⁵ and subsequent generation of the *Scx-GFP* transgenic mice⁶ have enabled unprecedented characterization of tendon tissues at the cellular and molecular levels in mice. Havis et al. (2014) reported recently the first transcriptomic analysis of developing mouse limb tendon cells using microarray hybridization analysis of FACS-isolated GFP⁺ cells from E11.5, E12.5, and E14.5 forelimb tissues of *Scx-GFP* transgenic mouse embryos¹⁷. Although microarrays provide a powerful approach for high-throughput analysis at the genomic scale, the hybridization-based approach is known to suffer from cross-hybridization noise and limited dynamic range of detection due to signal saturation. Thus, our RNA-seq datasets generated in this study complements the microarray datasets from Havis et al. (2014) and provide a valuable resource for the tendon research field¹⁷. We have deposited these RNA-seq datasets into the Gene Expression Omnibus database at the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/geo>, accession number GSE65180) so that other researchers

will be able to independently analyze these datasets to incorporate into their own research on tendon biology.

In contrast to bone and muscle development in which the transcriptional regulatory mechanisms have been extensively studied, only three transcription factors have been shown to play important roles in tendon development. Mice lacking *Scx* exhibit the earliest and most severe tendon differentiation defects thus far characterized⁷. However, many tendons and ligaments still form and function in the *Scx*^{-/-} mutant mice⁷. Mice lacking either *Mkx* or *Egr1* exhibit postnatal tendon maturation defects^{8, 9, 21}. To improve the understanding of the molecular regulation of tendon development, we focused on identification of other transcription factors specifically expressed during tendon development. However, although we isolated over 30 transcription factor genes from differential expression analyses of the early limb tendon RNA-seq datasets, in situ hybridization assays showed that most of them did not exhibit restricted expression in the developing tendon cells. On the other hand, GO analysis of the genes whose expression was significantly enriched in the *Scx*-GFP⁺ cells showed a clear association with connective tissue development and many of the top 20 enriched genes in each pair-wise comparison encode known markers of tendon cell differentiation. Thus, the *Scx*-GFP⁺ cell populations were clearly enriched with tendon cells although each of the FACS-isolated population must also contain non-tendon cells, consistent with the initial report of the *Scx*-GFP transgenic mice in which the *Scx*-GFP was shown expressed in much broader domains than that of the endogenous *Scx* mRNAs, particularly during early stages of tendon development⁶. It is possible that the cell dissociation and FACS isolation procedures might alter the RNA profiles of the isolated cells, which, combined with the ectopic *Scx*-GFP expression during early limb tendon development and the fact that genes encoding transcription factors are usually expressed at low abundance compared with genes encoding extracellular matrices or other structural proteins, limited our success in identifying tendon-specific transcription factors. In the future, it will be helpful to combine the *Scx*-GFP reporter with in situ cell isolation procedures, such as laser capture microdissection, for more accurate transcriptome expression profiling of the developing tendon.

The combined RNA-seq and in situ hybridization analyses allowed us to identify *Foxf2* expression as a specific marker of differentiated dorsal limb tendon. *Foxf2* is a member of the large Forkhead family of transcription factors^{22, 23}. It is strongly expressed in the developing facial mesenchyme and mice lacking *Foxf2* die perinatally with cleft palate²⁴. The specific expression of *Foxf2* in the differentiated dorsal limb tendon suggests that distinct molecular pathways are employed in regulating differentiation of different tendons. Consistent with this hypothesis, the levels of *Six2* mRNAs appear much higher in the ventral limb tendons than the dorsal limb tendons. Yamamoto-Shiraishi and Kuroiwa (2013) recently showed that *Six2* expression in the dorsal autopod, but not in the ventral autopod, depends on *Hox13* function²⁵. Further investigation of the relationship between *Foxf2* and *Six2* as well as their function in limb tendon development will provide novel insights into the molecular mechanisms of musculoskeletal system development and organization.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Ronen Schweitzer for providing the *Scx-GFP* transgenic mice. This work was supported by National Institute of Health (NIH) grants R01AR056943 and R01DE013681.

References

1. Birch HL, Thorpe CT, Rumian AP. Specialisation of extracellular matrix for function in tendons and ligaments. *Muscles, ligaments and tendons journal*. 2013; 3:12–22.
2. Benjamin M, McGonagle D. Entheses: tendon and ligament attachment sites. *Scandinavian journal of medicine & science in sports*. 2009; 19:520–527. [PubMed: 19522749]
3. Benjamin M, Ralphs JR. The cell and developmental biology of tendons and ligaments. *International review of cytology*. 2000; 196:85–130. [PubMed: 10730214]
4. Charvet B, Ruggiero F, Le Guellec D. The development of the myotendinous junction. A review. *Muscles, ligaments and tendons journal*. 2012; 2:53–63.
5. Schweitzer R, Chyung JH, Murtaugh LC, et al. Analysis of the tendon cell fate using *Scleraxis*, a specific marker for tendons and ligaments. *Development*. 2001; 128:3855–3866. [PubMed: 11585810]
6. Pryce BA, Brent AE, Murchison ND, et al. Generation of transgenic tendon reporters, *ScxGFP* and *ScxAP*, using regulatory elements of the *scleraxis* gene. *Developmental dynamics: an official publication of the American Association of Anatomists*. 2007; 236:1677–1682. [PubMed: 17497702]
7. Murchison ND, Price BA, Conner DA, et al. Regulation of tendon differentiation by *scleraxis* distinguishes force-transmitting tendons from muscle-anchoring tendons. *Development*. 2007; 134:2697–2708. [PubMed: 17567668]
8. Liu W, Watson SS, Lan Y, et al. The atypical homeodomain transcription factor *Mohawk* controls tendon morphogenesis. *Molecular and cellular biology*. 2010; 30:4797–4807. [PubMed: 20696843]
9. Ito Y, Toriuchi N, Yoshitaka T, et al. The *Mohawk* homeobox gene is a critical regulator of tendon differentiation. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:10538–10542. [PubMed: 20498044]
10. Lejard V, Blais F, Guerquin MJ, et al. *EGR1* and *EGR2* involvement in vertebrate tendon differentiation. *The Journal of biological chemistry*. 2011; 286:5855–5867. [PubMed: 21173153]
11. Brunskill EW, Potter SS. RNA-Seq defines novel genes, RNA processing patterns and enhancer maps for the early stages of nephrogenesis: *Hox* supergenes. *Dev Biol*. 2012; 368:4–17. [PubMed: 22664176]
12. Chen J, Aronow BJ, Jegga AG. Disease candidate gene identification and prioritization using protein interaction networks. *BMC bioinformatics*. 2009; 10:73. [PubMed: 19245720]
13. Chen J, Xu H, Aronow BJ, et al. Improved human disease candidate gene prioritization using mouse phenotype. *BMC bioinformatics*. 2007; 8:392. [PubMed: 17939863]
14. Zhang Y, Zhao X, Hu Y, et al. *Msx1* is required for the induction of *Patched* by *Sonic hedgehog* in the mammalian tooth germ. *Dev Dyn*. 1999; 215:45–53. [PubMed: 10340755]
15. Pryce BA, Watson SS, Murchison ND, et al. Recruitment and maintenance of tendon progenitors by *TGFbeta* signaling are essential for tendon formation. *Development*. 2009; 136:1351–1361. [PubMed: 19304887]
16. Huang AH, Riordan TJ, Wang L, et al. Repositioning forelimb superficialis muscles: tendon attachment and muscle activity enable active relocation of functional myofibers. *Developmental cell*. 2013; 26:544–551. [PubMed: 24044893]
17. Havis E, Bonnin MA, Olivera-Martinez I, et al. Transcriptomic analysis of mouse limb tendon cells during development. *Development*. 2014; 141:3683–3696. [PubMed: 25249460]

18. Mella S, Soula C, Morello D, et al. Expression patterns of the *coo/ebf* transcription factor genes during chicken and mouse limb development. *Gene expression patterns: GEP*. 2004; 4:537–542. [PubMed: 15261831]
19. Oliver G, Wehr R, Jenkins NA, et al. Homeobox genes and connective tissue patterning. *Development*. 1995; 121:693–705. [PubMed: 7720577]
20. Bonnin MA, Laclef C, Blaise R, et al. *Six1* is not involved in limb tendon development, but is expressed in limb connective tissue under *Shh* regulation. *Mechanisms of development*. 2005; 122:573–585. [PubMed: 15804569]
21. Guerquin MJ, Charvet B, Nourissat G, et al. Transcription factor *EGR1* directs tendon differentiation and promotes tendon repair. *The Journal of clinical investigation*. 2013; 123:3564–3576. [PubMed: 23863709]
22. Clark KL, Halay ED, Lai E, et al. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*. 1993; 364:412–420. [PubMed: 8332212]
23. Kaufmann E, Knochel W. Five years on the wings of fork head. *Mechanisms of development*. 1996; 57:3–20. [PubMed: 8817449]
24. Wang T, Tamakoshi T, Uezato T, et al. Forkhead transcription factor *Foxf2* (*LUN*)-deficient mice exhibit abnormal development of secondary palate. *Developmental biology*. 2003; 259:83–94. [PubMed: 12812790]
25. Yamamoto-Shiraishi Y, Kuroiwa A. Wnt and BMP signaling cooperate with *Hox* in the control of *Six2* expression in limb tendon precursor. *Developmental biology*. 2013; 377:363–374. [PubMed: 23499659]

Fig. 1. Analysis of gene expression profiles of developing limb tendons at E13
 (A–D) In situ hybridization detection of *Scx* mRNA expression in the wildtype fore- (A, B) and hind-limbs (C, D). The *Scx* mRNA signals are detected in purple color. (E–H) The patterns of green fluorescent protein (GFP) expression in the *Scx-GFP* transgenic mouse embryos recapitulate the endogenous *Scx* mRNA expression patterns in both fore- (E, F) and hind-limbs (G, H). (I) Data from Quantitative real-time RT-PCR assay of select markers in the FACS-isolated GFP⁺ and GFP⁻ cells from E13 forelimbs.

Fig. 2. Differential gene expression analysis of developing limb tendon by using RNA-seq
 (A) Comparison of levels of expression (FPKM) of *Scx*, *Tnmd*, *Lum*, *Fmod*, *Dcn*, *Colla1*, *Colla2*, *Col3a1*, *Col6a1*, *Col6a2* mRNAs in the RNA-seq datasets from FACS-isolated E11.5, E13.5, and E15.5 hindlimb GFP+ cells. (B) Venn diagrams showing partial overlap of enriched gene sets from differential gene expression analysis of RNA-seq data sets. 1555 genes were found to exhibit at least 1.5-fold increase in expression in GFP+ cells versus GFP- cells from forelimbs of E13.5 *Scx-GFP* transgenic mouse embryos (Red circle). These were compared with the genes whose expression was increased by more than 1.5-fold from E11.5 to E13.5 (blue circle) and from E13.5 to E15.5 (green circle) in the GFP+ hindlimb cell populations.

Fig. 3. Expression of the *Ebf* family genes during early limb tendon development

(A – C) At E11.5, the three *Ebf* genes were expressed in partially overlapping domains in the forelimb mesenchyme. (D – F) At E12.5, *Ebf1* and *Ebf2* mRNAs were mainly expressed in the autopod, around the developing digits (D, E), while *Ebf3* mRNAs were also expressed in the zeugopod (F). (G – U) At E13.5, *Ebf1* and *Ebf3* were expressed in cells surrounding the developing tendon, cartilage and bones, but not in the dorsal or ventral tendons. Expression of *Ebf2* mRNAs partially overlapped with *Scx* mRNAs in the dorsal and ventral tendons in the autopod (L, P, T, compared with J, N, R, respectively). Cross sections (J – U) were from mid-phalangeal (J – M), metacarpal (N – Q) and Wrist (R – U) regions. All sections are shown dorsal side up.

Fig. 4. *Foxf2* and *Six2* exhibit differential expression in dorsal and ventral limb tendons
 Expression patterns of *Six1*, *Six2*, and *Foxf2* mRNAs in E11.5, E12.5, and E13.5 forelimbs were compared with that of *Scx* mRNAs by whole mount *in situ* hybridization (A – P) and with that of *Tnmd* by section *in situ* hybridization (Q – Ab). Cross sections (Q – Ab) are from mid-phalangeal (Q – T), metacarpal (U – X) and Wrist (Y – Ab) regions. All sections are shown dorsal side up. *Six1* is strongly expressed in the proximal limb mesenchyme from E11.5 to E12.5 (B, F, J, N) and its expression in the developing limb becomes restricted to the perichondrium and muscle cells by E13.5 (R, V, Z). *Six2* mRNA expression (C, G, K, O, S, W, Aa) exhibited significant overlap with that of *Scx* (A, E, I, M) and *Tnmd* (Q, U, Y), but is stronger in the ventral limb tendon than the dorsal limb tendon (K, O, S, W). *Foxf2* mRNA expression was not detected in the limbs at E11.5 and E12.5 (D, H). At E13.5, *Foxf2* mRNA expression was specifically detected in the dorsal limb tendon within the autopod (T, X, Ab).

Table 1

GO analysis of 379 genes enriched in Scx-GFP⁺ limb cells at E13.5 and continuously up-regulated from E11.5 to E15.5

GO category	GO ID	GO term	p value	# of Genes
Molecular Function	GO:0005201	Extracellular matrix structural constituent	4.268E-24	24
	GO:0005518	Collagen binding	1.512E-21	22
	GO:0019838	Growth factor binding	1.663E-15	22
	GO:0005539	Glycosaminoglycan binding	1.325E-13	25
Biological Process	GO:0030198	Extracellular matrix organization	2.002E-45	68
	GO:0030199	Collagen fibril organization	6.771E-24	20
	GO:0007155	Cell adhesion	5.741E-19	69
	GO:0001501	Skeletal system development	4.840E-16	41
	GO:0061148	Connective tissue development	2.315E-10	23
	GO:0051216	Cartilage development	3.280E-9	19
	GO:0000904	Cell morphology involved in differentiation	8.627E-9	43
Human Phenotype	HP:0001388	Joint laxity	1.841E-16	21
	HP:0002758	Osteoarthritis	2.113E-11	13
	HP:0001367	Abnormal joint morphology	2.407E-9	46
	HP:0002813	Abnormality of limb bone morphology	5.840E-9	54
	HP:0003549	Abnormality of connective tissue	8.320E-9	39
Mouse Phenotype	MP:0005503	Abnormal tendon morphology	1.487E-13	13
	MP:0002108	Abnormal muscle morphology	1.925E-13	59
	MP:0005508	Abnormal skeleton morphology	5.163E-10	74
	MP:0003098	Decreases tendon stiffness	6.642E-7	4
	MP:0000163	Abnormal cartilage morphology	3.094E-6	24

Table 2GO analysis of 399 genes enriched in Scx-GFP⁺ cells at E13.5 and up-regulated from E11.5 to E13.5 only

GO category	GO ID	GO term	p value	# of Genes
Molecular Function	GO:0005509	Calcium ion binding	6.803E-7	34
	GO:0045499	Chemorepellent activity	4.576E-6	4
	GO:0008307	Structural constituent of muscle	1.265E-4	6
	GO:0043394	Proteoglycan binding	2.120E-4	5
Biological Process	GO:0048468	Cell development	2.444E-13	87
	GO:0009887	Organ morphogenesis	2.122E-11	51
	GO:0001501	Skeletal system development	4.561E-11	34
	GO:0000904	Cell morphogenesis involved in differentiation	1.272E-10	47
	GO:0045595	Regulation of cell differentiation	2.024E-10	64
	GO:0061448	Connective tissue development	4.591E-8	20
	GO:0051216	Cartilage development	7.191E-7	16
Human Phenotype	HP:0009803	Short phalanx of finger	3.354E-6	12
	HP:0009115	Aplasia/hypoplasia involving the skeleton	6.477E-5	27
	HP:0009810	Abnormality of the joints of the upper limbs	1.582E-4	17
Mouse Phenotype	MP:0002113	Abnormal skeleton development	8.617E-7	29
	MP:0000166	Abnormal chondrocyte morphology	1.109E-6	12
	MP:0005076	Abnormal cell differentiation	5.647E-4	29

Table 3

Transcription factors activated during early tendon development

Gene Symbol	Gene Description	FPKM values		
		E11(GFP+)	E13(GFP+)	E15(GFP+)
<i>Mkx^a</i>	Mohawk homeobox	2.92371	29.2162	20.8919
<i>Bnc2^a</i>	basonuclin 2	9.19681	39.9182	31.2196
<i>Creb3l1^b</i>	cAMP responsive element binding protein 3-like 1	0.95854	37.0428	118.088
<i>Creb3l2^a</i>	cAMP responsive element binding protein 3-like 2	56.5512	114.06	123.322
<i>Creb5^a</i>	cAMP responsive element binding protein 5	8.42502	37.2559	29.2659
<i>Dmrt2^a</i>	doublesex and mab-3 related transcription factor 2	3.55406	14.0685	10.5184
<i>Ebf1^a</i>	early B cell factor 1	50.7322	91.7113	73.8705
<i>Ebf2^a</i>	early B cell factor 2	8.74121	42.0117	35.7571
<i>Ebf3^a</i>	early B cell factor 3	30.2354	94.1107	68.8637
<i>Foxf2^b</i>	Forkhead box F2	1.29488	4.47666	7.19177
<i>Foxp1^a</i>	Forkhead box P1	12.2231	46.6546	34.5489
<i>Hic1^a</i>	hypermethylated in cancer 1	5.59986	41.599	40.1295
<i>Hivp2^b</i>	human immunodeficiency virus type 1 enhancer binding protein 2	0.7615	7.2643	19.2601
<i>Klf14^b</i>	Kruppel-like factor 14	1.03155	6.09617	29.6916
<i>Meox1^a</i>	mesenchyme homeobox 1	17.3961	35.7012	53.9161
<i>Myod1^b</i>	myogenic differentiation 1	3.00055	5.56645	15.6337
<i>Myog^b</i>	Myogenin	3.323	8.62464	35.9339
<i>Nfib^b</i>	nuclear factor I/B	7.72681	59.2753	90.4447
<i>Nfic^b</i>	nuclear factor I/C	11.0855	32.8171	79.9389
<i>Nfix^b</i>	nuclear factor I/X	4.60583	55.5602	131.575
<i>Osr1^a</i>	Odd-skipped related 1 (Drosophila)	29.585	56.2089	50.3268
<i>Plagl1^b</i>	pleiomorphic adenoma gene-like 1	18.2221	103.312	457.855
<i>Sim2^b</i>	single-minded homolog 2 (Drosophila)	3.21879	5.816	9.59725
<i>Six2^a</i>	sine oculis-related homeobox 2 homolog (Drosophila)	5.53047	29.5461	15.2109
<i>Tshz2^a</i>	teashirt zinc finger family member 2	16.9844	43.3862	53.111
<i>Zcchc5^b</i>	zinc finger, CCHC domain containing 5	2.1941	25.7706	90.3176
<i>Zfp354c^a</i>	zinc finger protein 354C	19.2049	37.4625	37.6418
<i>Zfp580^a</i>	zinc finger protein 580	12.3642	21.3423	19.844
<i>Zfp9^a</i>	zinc finger protein 9	10.8411	32.3694	30.8232
<i>Zhx3^a</i>	zinc fingers and homeobox-3	8.66091	13.9502	16.6142

Gene Symbol	Gene Description	FPKM values		
		E11(GFP+)	E13(GFP+)	E15(GFP+)
<i>Zim1</i> ^b	zinc finger, imprinted 1	7.57369	16.3261	59.0092

^a 18 genes up-regulated from E11 to E13;

^b 13 genes up-regulated from E11 to E13 and from E13 to E15.