Unexpected contribution of fibroblasts to muscle lineage as a mechanism for limb muscle patterning

Joana Esteves de Lima, Cédrine Blavet, Marie-Ange Bonnin, Estelle Hirsinger, Glenda Comai, Laurent Yvernogeau, Marie-Claire Delfini, Léa Bellenger, Sébastien Mella, Sonya Nassari, Catherine Robin, Ronen Schweitzer, Claire Fournier-Thibault, Thierry Jaffredo, Shahragim Tajbakhsh, Frédéric Relaix, Delphine Duprez

Positional information driving limb muscle patterning is contained in connective tissue fibroblasts but not in myogenic cells. Limb muscles originate from somites, while connective tissues originate from lateral plate mesoderm. With cell and genetic lineage tracing we challenge this model and identify an unexpected contribution of lateral plate-derived fibroblasts to the myogenic lineage, preferentially at the myotendinous junction. Analysis of single-cell RNA-sequencing data from whole limbs at successive developmental stages identifies a population displaying a dual muscle and connective tissue signature. BMP signalling is active in this dual population and at the tendon/muscle interface. In vivo and in vitro gain- and loss-of-function experiments show that BMP signalling regulates a fibroblast-to-myoblast conversion. These results suggest a scenario in which BMP signalling converts a subset of lateral plate mesoderm-derived cells to a myogenic fate in order to create a boundary of fibroblast-derived myonuclei at the myotendinous junction that controls limb muscle patterning.
Skeletal muscle patterning is a developmental process that controls the formation of each muscle at the right position and time, leading to the highly complex organisation of over 600 individual skeletal muscles in human, each of them with a specific shape, size, innervation and attachment sites to the skeletal system. The cellular and molecular mechanisms that regulate limb muscle patterning remain poorly understood, although we know from embryological experiments that the positional information is not contained in myogenic cells themselves but in connective tissues (CTs)1–7. CTs, primarily composed of fibroblasts and extracellular matrix, support and connect organs together2. Muscle attachments are composed of two types of CTs, the tendons that link muscle to bone and the muscle CT that is a continuum of tendon surrounding individual muscles and also present in between muscle fibres2. The long-standing consensus emerging from experimental embryology in avians is that limb myogenic cells are derived from somitic mesoderm, while CT fibroblasts originate from the lateral plate mesoderm1,2,8–10. Consistently, genetic lineage tracing experiments in mice during development have shown that Pax7+ muscle progenitors originate from the Pax3 lineage in the limb11–13 and that limb muscle attachment originates from CT fibroblasts lineages14,15. Previous embryological surgical experiments in chicken embryos demonstrated that the positional information for limb muscles and associated innervation (motoneuron axons) is contained in the lateral plate mesoderm and not in somite-derived tissues1,5,6,16–18. This positional information for muscle patterning is maintained over time in lateral plate-derived tissues, such as muscle CT5 and tendons4. The molecular signals produced by lateral plate-derived tissues that drive the correct positioning of limb muscles during development are not fully identified. To date, the TBX4/5, TCF4 and OSR1 transcription factors expressed in the TA muscle [lineage] describes1,8,9, this type of graft contains periosteum + QCPN (Supplementary Fig. 4d, e), unlike tissue graft experiments (Supplementary Fig. 3). Despite this mosaic labelling, we observed a fraction of Pax7+ muscle progenitors, MYOG+ myoblasts and QCPN+ myonuclei that were not somite-derived (Fig. 1h). As the QCPN/PAX7+ double immunolabelling could not be performed due to antibody incompatibility, we performed presomitic mesoderm grafts using the available cytoplasmic GFP-chicken reporter line26. This analysis identified a population of Pax7+ cells at muscle tips in close proximity to tendons (Supplementary Fig. 2). From these presomitic mesoderm grafting experiments, we conclude that a small fraction of Pax7+ muscle progenitors, MYOG+ myoblasts and QCPN+ myonuclei are not derived from somites.

To determine the embryonic origin of these non-somite-derived muscle cells, we performed the converse experiment and replaced the limb lateral plate mesoderm of chicken embryos by the equivalent region of quail embryos. Consistent with classical lineage descriptions1,8,9, this type of graft confirmed the lateral plate mesoderm origin of cartilage, tendons and muscle connective tissue (Supplementary Fig. 3). However, we also found lateral plate-derived myonuclei (QCPN+) at the extremities of myosin+ myotubes in close vicinity to tendons visualised with SCX expression at E9 (Fig. 1i–l). To assess lateral plate contribution to the muscle lineage, we electroporated the lateral plate mesoderm of chicken embryos at the forelimb level with a plasmid co-expressing membrane-Tomato and nuclear-GFP reporters under the control of a ubiquitous promoter27. Successfully electroporated forelimbs were screened for Tomato expression (Supplementary Fig. 4a–c) and nuclear-GFP labelling was used to accurately assess the contribution of lateral plate cells to sncytial myofibers and interstitial myogenic cells. As expected, GFP+ nuclei encompassed all lateral plate-derived tissues but in a mosaic manner (Supplementary Fig. 4d, e), unlike tissue graft experiments (Supplementary Fig. 3). Despite this mosaic labelling, we observed sporadic GFP+ myonuclei, GFP+MYOG+ myoblasts and GFP+PAX7+ muscle progenitors located at muscle tips in E10 chicken limbs (Supplementary Fig. 4f–i).

Altogether, these grafting and electroporation experiments performed in avian embryos show an unexpected contribution of lateral plate-derived cells to the myogenic lineage in limb muscles with a preferential location at the muscle tips close to tendons.

Scx and Osr1 lineages contribute to myogenic cells. It is well recognised that the Pax3 lineage corresponds to the somite-derivatives in mouse limbs28. To assess if the lateral plate contribution to limb muscles observed in chicken was conserved in mice, we analysed Pax7 expression with respect to the Pax3 lineage. Using a Tomato reporter mouse line to trace the Pax3 lineage28,29, we observed the expected contribution of Pax3-derived cells to limb foetal muscles at E14.5 and E15.5 when the muscle pattern is set (Fig. 2a, b and Supplementary Fig. 5a). Notably, we observed a fraction of Pax7+ cells that were not Pax3-derived (Tomato-negative cells), with a preferential location close to tendons at foetal stages (Fig. 2a–d and Supplementary Fig. 5a–d). Quantifications showed that 4.6% (±1.35 SD)
of PAX7+ cells were Tomato-negative at E15.5 (Fig. 2d), while 7.5% (±2.3 SD) of PAX7+ cells were Tomato-negative at E14.5 (Supplementary Fig. 5d). In addition, using a nuclear H2B-GFP reporter mouse line to trace the Pax3 lineage28,30, we further identified 3.1% (±1.1 SD) of PAX7+ muscle progenitors and 2.6% (±1.2 SD) of MYOD+ myoblasts that were not Pax3-derived (nuclear-GFP-negative cells), at muscle tips close to tendons at E15.5 (Fig. 2e–h). We conclude that a small fraction of PAX7+ and MYOD+ myogenic cells are not derived from the Pax3 lineage in mouse limb muscles at the foetal stage. These results are in contrast to the established model supporting that Pax3+ somitic cells give rise to the entirety of PAX7+ cells in mouse limbs11–13.

To determine the developmental origin of these non-Pax3-derived myogenic cells, we performed genetic lineage tracing analyses with the CT fibroblast markers, Scx and Osr1.

The bHLH transcription factor SCX is a recognised tendon cell marker during development and is involved in tendon
Fig. 1 A subpopulation of muscle cells does not originate from somites but from the lateral plate in the chicken limb. a–c Limb muscle organisation after quail-into-chicken presomitic mesoderm grafts. a Schematic of limb presomitic mesoderm grafts. b, c Transverse and adjacent forelimb sections of E9 presomitic mesoderm grafted embryos immunostained with QCPN (quail nuclei, red) and MF20 (myosins, green) (b), and with QCPN (quail nuclei, red) and ColXII (tendons, green) antibodies (c). d High magnification of FCU, a ventral and posterior muscle immunostained with QCPN (quail nuclei, red), MF20 (myosins, green) and ColXII (tendons, green) antibodies combined with DAPI (nuclei, blue). e High magnification of muscle tips (close to tendons), squared in d, showing non-quail nuclei within myosin+ cells (arrows). f High magnification of FCU immunostained with QCPN (quail nuclei, red) and MYOG (green) antibodies. g Focus on muscle tip regions showing MYOG+ nuclei (green) that are not quail+ (red) (white arrows) close to myosin+ cells (grey). Yellow arrowheads point to MYOG+ (green) that are quail+ (red). h Percentage of MYOG+/QCPN− cells versus total MYOG+ cells. Data are presented as mean values ± SD from 3 grafted limbs. i–l Limb muscle organisation after quail-into-chicken lateral plate mesoderm grafts. i Schematic of limb lateral plate mesoderm grafts. j, k Transverse and adjacent limb sections were hybridised with SCX probe (blue) to label tendons (j) and immunostained with QCPN (quail nuclei, red), MF20 (myosins, green) antibodies combined with DAPI (blue) (k), focused on the FCU muscle. l High magnification of muscle tip regions (close to tendons) showing the high density of quail nuclei of lateral plate mesoderm origin. Arrows point to quail myonuclei (red) in myosin+ fibres (green).

Fig. 2 A subpopulation of muscle cells is not derived from the Pax3 lineage in mouse limbs. a Transverse forelimb sections of E15.5 Pax3Cre:R26stop/Tom mice immunostained with Tomato (Pax3 lineage, red) antibody and DAPI staining (blue). b Focus on a muscle immunostained with Tomato (Pax3 lineage, red) and PAX7 (green) antibodies. c High magnifications showing PAX7+ cells (green) that are tomato-negative, i.e. not of Pax3 lineage (arrows). d Percentage of PAX7+/Tomato− cells versus PAX7+ cells in muscles. Data are presented as mean values ± SD from 3 limbs. e Longitudinal sections along a limb muscle of E15.5 Pax3Cre:R26stop/Tom mice immunostained with GFP (Pax3 lineage, green) and PAX7 (red) antibodies. Arrows point to PAX7+ cells that are GFP-negative. f Percentage of PAX7+/GFP− cells versus PAX7+ cells in muscles. Data are presented as mean values ± SD from 3 limbs. g Longitudinal sections along a limb muscle of E15.5 Pax3Cre:R26stop/Tom mice immunostained with GFP (Pax3 lineage, green) and MYOD (red) antibodies. Arrows point to MYOD+ cells that are GFP-negative. h Percentage of MYOD+/GFP− cells versus MYOD+ cells in E15.5 muscles. Data are presented as mean values ± SD from 3 limbs.

formation31,32. We used the ScxCre;R26stop/Tom,Scx-GFP mice14,29 to analyse contemporary Scx expression together with the Scx lineage contribution in developing limbs. As described previously14, the Scx lineage (Tomato+ cells) was observed in the limb mesenchyme with a pattern overlapping that of the Scx-GFP transgene in tendons (Fig. 3a). At E16.5, an overlap between myosins and Tomato labelling was observed in Tomato+ junctional regions between muscles (Supplementary Fig. 5e, f). In addition to tendon cells, we identified PAX7+ and MYOD/ MYOG+ myogenic cells within the Scx lineage that were
preferentially concentrated close to tendons at E14.5 (Fig. 3b–e). Quantifications showed that 7.4% (±1.7 SD) of PAX7+ cells and 4.8% (±2.2 SD) of MYOD/MYOG+ cells were derived from Scx lineage (Tomato+ cells) (Fig. 3f). At E12.5, when the muscles are not yet individualised, we found 4.12% (±1.3 SD) of PAX7+ and 4.13% (±1.15 SD) of MYOD/MYOG+ myogenic cells surrounding the muscle masses to be Scx-derived (Tomato+) (Supplementary Fig. 5g–k). These values are in accordance with those obtained from the Pax3 lineage analysis.

The zinc-finger transcription factor OSR1 is a key regulator of muscle CT differentiation. As already reported, lineage tracing using Osr1CreERT2 as a driver with a ubiquitous Tomato reporter led to Tomato expression in limb muscle CT around individual muscles and in between myotubes (Fig. 3g–i).
To address the Osr1 lineage contribution to the muscle lineage specifically, we performed lineage tracing with Osr1CreERT2 as a driver and the Pax7 reporter (Pax7GFP, Pax7GPL), a reporter gene that is not expressed in CT cells34. As such, only cells that expressed CreERT2 from the Osr1 locus and belong to the myogenic lineage (Pax7+ and descendants) will express nuclear β-galactosidase (betaGal) upon Tamoxifen induction34. Using this genetic combination, we found 4.5% (±2.5 SD) of Pax7+ muscle progenitors to be of Osr1 lineage (betaGal+ cells) at E15.0 upon tamoxifen induction at E15.0 (Fig. 3k–j). Similarly, 3.6% (±2.4 SD) of MYOD/MYOG+ myogenic cells were derived from Osr1 lineage (betaGal+ cells) (Fig. 3k–m). A slightly higher contribution of the Osr1 lineage to Pax7+ cells (6.2% ± 2.8 SD) and MYOD/MYOG+ cells (6.9% ± 3 SD) was noted at an earlier time point (E12.5) in limb muscle masses (Supplementary Fig. 6c–i). To define the time window of the contribution of Osr1 lineage to myogenic cells, we performed the tamoxifen injection at later time points. We observed only sparse (induction at E12.5) or no (induction at E13.5) betaGal injection at later time points. We observed only sparse (induction at E12.5) or no (induction at E13.5) betaGal injection at later time points. We observed only sparse (induction at E12.5) or no (induction at E13.5) betaGal injection at later time points.

**scRNAseq analysis identifies dual CT/myogenic cells in limbs.** The unexpected contribution of CT fibroblasts to the muscle lineage as early as with the Pax7+ muscle progenitor stage suggested a scenario in which CT fibroblasts acquire progressively a myogenic signature to be incorporated at the tips of existing myotubes during development. In this scenario, CT fibroblasts would convert to a myogenic fate before fusion. To search for cells exhibiting a transient CT/muscle signature, we turned to the global transcriptomic analysis of limb cells with single-cell resolution by single-cell RNA-sequencing (scRNAseq) of chicken whole-limb cells at successive developmental stages: E4, a progenitor stage; E6, when major spatial re-arrangements occur for muscle and CTs and E10, when the final muscle pattern is set. At each stage, we identified CT and muscle clusters, along with their respective CT and muscle markers (Fig. 4a and Supplementary Table 1). In parallel, we identified a subpopulation of cells co-expressing, at the cellular level, at least one CT-marker (among PRRX138, PDGFRA39, TWIST240, OSR115 and SCX31) and one muscle marker (among Pax7, MYF5, MYOD and MYOG) (Supplementary Fig. 8a). The size of this dual CT/muscle (CT/M) population increased from 0.5% of limb cells at E4 to reach 4% of limb cells at E6 and was maintained at 3.3% at E10 (Fig. 4b). At E10, 11% of the MYOG+ myoblasts had a dual phenotype signature, which is consistent with the 9.5% of MYOG+ myoblasts that were not somite-derived in our quail-chick grafting experiments (Fig. 1h). We found that quantifications from the scRNAseq analysis were lower compared to those of the mouse genetic lineage tracing experiments. For example, at E10, 1.6% of Pax7+ cells were Pax7+/OSR1+ dual cells compared to 4.5% in the Osr1CreERT2:Pax7GFP mouse embryos at E15.5 (Fig. 3m). This was expected as in the latter, the analysis is based on lineage history whereas contemporary expression was examined in the transcriptome analysis. CT and M scores were calculated as the corrected average expression levels of the CT and muscle markers, respectively (Supplementary Table 1) and plotted for each cell grouped by its CT, CT/M or M identity (Fig. 4c). Except at E4 when the CT score was the same for the CT and CT/M populations, the CT/M cells exhibited intermediate CT and M scores, compared to those of cells with a strict CT or muscle identity (Fig. 4c), pointing to a hybrid CT/muscle identity. The analysis of genes differentially expressed in the CT/M population versus the combined CT and muscle populations showed that these CT/M cells did not express specific genes other than CT markers or muscle markers (Supplementary Table 2, gene examples in Fig. 4d). As a consequence, the CT/M population does not cluster but rather is included in both the CT and muscle clusters (Fig. 4e). CT/M cells with a high CT score are found within CT clusters, while those with a high M score are found within the muscle clusters (Supplementary Fig. 8b). Another interesting feature is that CT markers were systematically
downregulated (except at E4) and muscle markers were systematically upregulated in the CT/M cells (Supplementary Table 2). This bi-phasic expression profile suggests that the CT/M cells were in transit from a CT identity to a muscle identity. At E4, CT markers are not yet downregulated and M markers are already upregulated in CT/M cells (Supplementary Table 2), suggesting that the transition process starts before E6 by the upregulation of muscle markers followed by a subsequent downregulation of CT markers in the CT/M cells.

Among CT markers, PRRX1 was expressed in the majority of CT/M cells at all stages (85% of CT/M cells at E4, 72% at E6 and 66% at E10). The fraction of PRRX1+ cells among CT/M cells was maintained at 66%, while the fraction of PRRX1+ cells among all limb cells dropped with time from 88% at E4 to 25% at E10 (Supplementary Fig. 8c), suggesting that PRRX1 expression is actively maintained in CT/M cells. This is consistent with the contribution of Prrx1 lineage to mouse postnatal muscles (Yaseen Badarneh et al., back-to-back submission). TWIST2
was also expressed in CT/M cells (15% at E4, 25% at E6 and 29.5% at E10), indicating that a TWIST2 cell population contributes to developmental muscles in addition to adult skeletal muscles as previously described. The myogenic markers, PAX7 and MYOD1 showed similar distributions in CT/M cells, as a majority of CT/M cells expressed PAX7 (42–57%) and MYOD1 (36–57%), while much fewer cells expressed MYOG (10–15%). These observations are suggestive of an immature state of these CT/M cells (Supplementary Fig. 8c).

From this scRNAseq analysis, we conclude that limb cells exhibit a small population with a CT fibroblast/myogenic signature that is in transit between a CT phenotype and a muscle fate.

**BMP signalling promotes a fibroblast-to-myoblast conversion.** Cell communication signals, acting through secreted ligands, govern a multitude of cell-fate choices during development and postnatal life. To explore the mechanisms underlying the fibroblast-to-myoblast switch we examined which signalling pathways were enriched in the CT/M cells. ID genes are recognised transcriptional readouts of BMP activity. ID1, ID2 and ID3 genes were expressed at a similar or greater level in CT/M cells compared to CT and M cells (Fig. 5a). In addition, the fraction of ID+ cells was significantly higher in CT/M cells versus CT and M cells at E6 and E10 (Fig. 5b). Consistently, ID2 expression was enriched at the CT/muscle interface, labelled with SCX expression in E6 chicken limbs (Fig. 5c, d). The nuclear location of pSMAD1/5,9, another readout of active BMP, was increased at the periphery of dorsal and ventral muscle masses close to sites of BMP4 expression (Fig. 5e–h) and at muscle tips (Fig. 5i) at E6. By the time the muscle pattern was fully set, i.e. between E7 and E9, ID2 transcripts and pSMAD1/5,9 were observed at the tips of myotubes (myosins+) close to tendons on longitudinal or transverse sections (Fig. 5j–n). Interestingly, in our quail-chick grafting experiments, we detected 2 populations of pSMAD1/5,9+ myonuclei at muscle tips, one somite-derived (yellow arrows in Fig. 6) and one non-somite-derived (white arrows in Fig. 6) at both E6 and E9. This is reminiscent of the two distinct myonucleus populations recently identified in adult mouse muscles at the myotendinous junction, with one of them being enriched in CT-associated collagen genes. Thus, our findings support the notion that myonuclei with a fibroblast signature could be of CT origin. Altogether, these results show the existence of a BMP-responsive population with a dual CT/muscle identity that could correspond to the fibroblast-derived myogenic cells found at the CT/muscle interface.

Based on BMP activity in this junctional cell population in chicken limbs (Fig. 5), we hypothesised that BMP signalling would be involved in the recruitment of a subpopulation of CT fibroblasts towards a muscle fate. To test the ability of BMP signalling to control a fibroblast-to-myoblast conversion, we modified BMP activity in chicken limbs and assessed the consequences for PAX7+ myogenic progenitors and TCF4+ CT fibroblasts (Fig. 7a–i). Exposure of chicken limbs to retroviral-BMP4 led to an increase in the number of PAX7+ muscle progenitors (Fig. 7e, f) as previously observed. However, consistent with the cell-fate conversion hypothesis, the increase of PAX7+ progenitors occurred at the expense of TCF4+ fibroblasts (Fig. 7i) with no change in the total cell number (Fig. 7g, h). Moreover, no change in the proliferation rate of PAX7+ cells was observed upon BMP4 exposure in chicken limbs (Supplementary Fig. 9a–g). This abnormal increase in PAX7+ cells led to muscle patterning defects, such as muscle splitting defects visualised by the fusion of individual muscles (Fig. 7c–f), which was probably due to the reduction of TCF4+ fibroblasts known to control muscle patterning. BMP overexpression was also applied to myoblast primary cultures in proliferation culture conditions. Consistent with our conversion hypothesis, BMP4 treatment did not alter the number of PAX7+ cells in myoblast cultures (Supplementary Fig. 9h–j). To assess the fibroblast-to-myoblast conversion, we applied retroviral-expressing BMP4 or BMPRI1Aca (a constitutive active form of the BMPRI1A) to CEs (chicken embryonic fibroblasts) that were isolated from whole embryos. BMP over-activation induced the appearance of PAX7+ cells compared to control cultures, albeit with a low efficacy (Fig. 7j–m). This shows that a subset of this mixed fibroblast population can convert into PAX7+ cells when exposed to BMP in vitro.

We then performed BMP loss-of-function experiments in chicken limbs by overexpression of the BMP antagonist NOGGIN. This approach induced the converse phenotype, i.e. an increase in the number of TCF4+ fibroblasts at the expense of PAX7+ muscle progenitors with no global change in cell number (Fig. 8a–i). The behaviour of TCF4+ fibroblasts upon BMP misexpression is consistent with the respective decrease and increase in the expression of the OSR1 CT-marker upon BMP gain- and loss-of-function experiments in chicken limbs. Similar to the in vivo experiments in chicken embryos, the inhibition of BMP signalling in myoblast cultures led to a fibroblast phenotype (Fig. 8j–n and Supplementary Fig. 9k, l). BMP inhibition was achieved with the overexpression of SMAD6 (Fig. 8j–n), an inhibitor of the pSMAD1/5,9 pathway or by using the BMP antagonist NOGGIN (Supplementary Fig. 9k, l). BMP inhibition led to a reduction in the number of PAX7+, MYOD+ and MYOG+ myogenic cells and myotubes in myoblast cultures in differentiation culture conditions (Fig. 8j–l). This was associated with a decrease in the mRNA expression levels of muscle lineage markers from progenitor to differentiation state (PAX7, MYF5, MYOD, OSR1, PAX7, MYOD, ID2).
**MYOG, MYHC** (Fig. 8m). The SMAD6-treated muscle cells adopted a fibroblast-like phenotype compared to myotubes observed in control myoblast cultures (Fig. 8j–l) and displayed an increased expression of CT markers, such as COL3A1, COL6A1 and OSR1 (Fig. 8n). These results show that BMP inhibition converts myoblasts towards a fibroblast phenotype.

Altogether, these in vivo and in vitro BMP functional experiments unravel an unexpected role for BMP signalling in driving a fibroblast/myoblast switch during limb muscle patterning.

**Discussion**

Here, we identified a subpopulation of limb myogenic cells that did not originate from the somitic Pax3-lineage in chicken and mouse embryos. This subpopulation of myogenic cells originated from lateral plate-derived cells (chicken embryos) and CT-marker
Fig. 5 BMP signalling is active in cells displaying a CT/M transcriptional identity and in cells at the muscle/tendon interface in chicken limbs. a Violin plots showing Log-normalised expression levels of ID1, ID2 and ID3 genes in cells grouped by their CT (red), CT/M (green) or M (blue) identity at E4, E6 and E10. The statistical test used is a two-sided Wilcoxon rank sum test; Sample size: n = 4758 cells from two E4 embryos, 3268 cells from three E6 embryos and 21256 cells from three E10 embryos. The exact p-values are indicated on the violin plots. b Bar plot showing the percentage of cells expressing ID1 genes in CT (red), CT/M (green) or M (blue) cells at E4, E6 and E10. The statistical test used is the two-sided Fisher’s exact test. Sample size: n = 4758 cells from two E4 embryos, 3268 cells from three E6 embryos and 21256 cells from three E10 embryos. The exact p-values are indicated on the graphs. c, d Focus on ventral limb muscle masses from adjacent and transverse limb sections of E6 chicken embryos hybridised with SCX (e) and ID2 (d) probes. e, f Transverse adjacent limb sections of E6 chicken embryos were hybridised with BMP4 probe (e) and immunostained with pSMAD5/5/9 antibody (red) (f). e, f Arrows point to BMP4 and pSMAD5/5/9 expression spots. g, h High magnification of dorsal (g) and ventral (h) limb muscle masses immunostained with pSMAD5/5/9 (red) and MF20 (myosins, green) antibodies and counterstained with DAPI, arrows point to pSMAD5/5/9 expression spots at muscle periphery. i Longitudinal view of a muscle labelled with pSMAD5/5/9 (red) and MF20 (myosins, green) antibodies, arrows point to pSMAD5/5/9 enriched expression at muscle tips. j, k Adjacent and longitudinal limb sections of E7 chicken embryos hybridised with BMP4 (j) or ID2 (k) probes (blue staining), arrow points to ID2 expression at muscle tips close to tendons (labelled with BMP4 and ID2 transcripts). l Longitudinal muscle sections of E9 chicken hybridised with ID2 probe (blue) and immunostained with MF20 antibody (myosins, brown). The arrow points to ID2 expression at muscle tips. High magnification of one muscle tip. m Transverse limb sections of E9 chicken embryos were hybridised with ID2 probe (blue) and immunostained with MF20 (myosins, brown) antibody. Focus on dorsal limb muscles. n Transverse limb sections of E9 chicken embryos immunostained with pSMAD1/5/9 (green) and MF20 (myosins, grey) antibodies. Focus on dorsal limb muscles.

lineages (mouse embryos). Based on quantification in chicken and mouse developing limb muscles, this myogenic population of CT fibroblast origin represents 4–9% of muscle cells and was preferentially located at muscle tips close to tendons during foetal myogenesis. Interestingly, the recruitment of mesenchymal stromal cells has already been observed in the myotome in somites of mouse embryos, although the regionalisation of the recruitment had not been addressed in that study. The unexpected contribution of lateral plate-derived CT fibroblasts to the muscle lineage opens new perspectives for the mechanisms underlying limb muscle patterning. Previous BrdU/EdU-based experiments have shown preferential incorporation of nuclei in myotubes at muscle tips during foetal and postnatal limb myogenesis. Although these BrdU/EdU-based experiments never addressed the origin of the newly incorporated nuclei, it was always assumed that this preferential incorporation at the muscle tips was from somite-derived cells. We now propose that the incorporation of new nuclei at muscle tips of forming myotubes also includes CT lineage-derived nuclei. As the lateral plate-derived cells are recognised to drive limb muscle patterning, one compelling hypothesis is that these CT lineage-derived muscle cells would regulate the location of myoblast fusion events and thereby direct the shape and orientation of limb muscle growth during foetal development. These CT lineage-derived muscle cells are reminiscent of muscle founder cells that drive muscle identity and position in Drosophila embryos. In addition to playing a role in limb muscle patterning, these regionalised myonuclei of CT origin could participate to myotendinous junction formation. Consistent with this idea, the myonuclei at the myotendinous junction display a fibroblast gene signature. A possible scenario would be that the CT fibroblasts incorporated in the myotubes at muscle tips close to tendons maintain a fibroblast genetic programme in addition to being reprogrammed to muscle. This molecular fibroblast signature enriched in extracellular matrix components could then participate in myotendinous junction formation.

The reprogramming of non-muscle nuclei to muscle nuclei after fusion has been already demonstrated in heterokaryons in vitro. Moreover, reprogramming of mesenchymal stem cell nuclei to muscle nuclei was shown to rely on fusion and to be mediated by IL4/NFAT signalling. Based on our data, we envision an additional step of fibroblast-to-muscle conversion before fibroblast incorporation to myotubes during limb foetal myogenesis. Our data supporting this scenario are (1) the identification of PAX7+ muscle progenitors that were of Scx or Osr1 lineage in mouse limbs, (2) the identification of a sub-population of chicken limb cells with a dual CT/muscle identity by scRNAseq analysis, (3) and the ability of BMP signalling to tilt the balance between muscle and CT fate. Interestingly, this limb cell population with a dual CT/muscle identity is maintained in postnatal mouse limb muscles (Yaseen Badarneh et al., back-to-back submission). These dual CT/muscle identity cells can be related to the bi-fated cells identified at the bone/tendon insertion in mice during development.

We identified BMP as molecular signalling involved in the recruitment of CT fibroblast towards the myogenic lineage in chicken limb cells. BMP signalling is active at the right place at the CT/muscle interface during foetal development. The presence of BMP activity in this junctional cell population combined with the BMP gain- and loss-of-function experiments in chicken limbs and in cell cultures lead us to propose a scenario in which BMP locally drives the conversion of CT fibroblasts to myoblasts, which are incorporated into myotubes at muscle tips close to tendons. Persistent BMP signalling regionalised at muscle tips could provide an in vivo safeguard to maintain the muscle identity of fibroblast-derived muscle cells. We believe that the BMP effect on fibroblast-muscle conversion during foetal myogenesis is unrelated to the BMP function in the equilibrium between PAX7+ progenitor expansion and differentiation during postnatal growth. However, the BMP effect on fibroblast/myoblast conversion during foetal myogenesis could be related to the function of BMP signalling in the specification of PAX7+ cells in the dorsal lateral plate tissue (lateral plate equivalent) and not in the paraxial mesoderm (somite equivalent) in Xenopus embryos, although we are not looking at the same stage or process. Consistently, BMP signalling is also enriched in the dual CT/muscle cells in postnatal mouse limb muscles (Yaseen Badarneh et al., back-to-back submission). Interestingly, BMP signalling has been already shown to be involved in the formation of another junctional population at the level of the tendon/bone insertion, by regulating Scx expression during mouse limb development. Indeed, BMP signalling has been described in numerous processes related to cell-fate control and reprogramming during development, wound healing and pathologies.

In summary, our studies show that the generation of the tendon/muscle interface partly results from a fibroblast-to-myoblast conversion driven by localised BMP signalling during foetal myogenesis. This unexpected scenario provides a novel cellular and molecular mechanism underlying limb muscle patterning and myotendinous junction formation. The existence of dual
identity cells could be part of a universal mechanism to generate tissue junctions and deal with cell heterogeneity to finally shape developing adjacent tissues. However, it is not yet clear whether the existence of dual identity cells would be part of the original mechanism underlying muscle patterning or whether it is a later evolutionary acquisition to enable the remarkable variation in muscle shapes and size that occurred through mammalian evolution. Future studies will be also needed to explore whether this mechanism can be reactivated in adult tissues and where it could be relevant in the context of muscle regeneration.

Methods

Chicken and quail embryos. Fertilised chicken and quail eggs from commercial sources (White Leghorn and JA57 strain, Japanese quail strain, EARL Les Brûyères, Dangers, France) were incubated at 37.5 °C. Before 2 days of incubation, chicken and quail embryos were staged according to somite number. After 2 days of

Fig. 6 pSMAD1/5/9⁺ myonuclei label two myonucleus populations of distinct embryological origins at muscle tips. a Schematic of isotopic presomitic mesoderm grafts from quail to chicken embryos. b Longitudinal view of a limb muscle from E6 chimera labelled with pSMAD1/5/9 (green), QCPN (quail nuclei, red) and MF20 (myosins, grey) antibodies, with a focus on muscle tips. c Adjacent section to b labelled with COLXII (tendon, green), QCPN (quail nuclei, red) and MF20 (myosins, grey) antibodies. d Focus on muscle tips, labelled with pSMAD1/5/9 (green), QCPN (quail nuclei, red) and MF20 (myosins, grey) antibodies, combined with DAPI (blue). White arrows point to pSMAD1/5/9⁺ myonuclei (green) that are not of quail origin (QCPN⁻), while yellow arrows point to pSMAD1/5/9⁺ myonuclei (green) that are of quail origin (QCPN⁺, red). e Transverse view of the FCU muscle from E9 chimera labelled with pSMAD1/5/9 (green), QCPN (quail nuclei, red) and MF20 (myosins, grey) antibodies. f Adjacent section to e labelled with COLXII (tendon, green), QCPN (quail nuclei, red) and MF20 (myosins, grey) antibodies, combined with DAPI. g Focus on muscle tips. White arrows point to pSMAD1/5/9⁺ myonuclei (green) that are not of quail origin (QCPN⁻), while yellow arrows point to pSMAD1/5/9⁺ myonuclei (green) that are of quail origin (QCPN⁺, red).
incubation, embryos were staged according to days of in ovo development. GFP
+ chicken were generated and provided by the Roslin Institute\textsuperscript{26}.

Mouse strains. Animals were handled according to European Community
guidelines, implementing the 3Rs rules. Protocols were validated by the ethic
committee of the French Ministry, under the reference numbers APAFIS\#13695-
2018021408521124xv2, APAFIS\#24357-2020041613396163xv3 and APAFIS\#6354-
20160890208839v4 and Institut Pasteur ethics committee (CETEA, reference
2015–0008). The following strains were described previously: \textit{Pax3Cre} \textsuperscript{26}, \textit{Wnt1Cre} \textsuperscript{37}, \textit{ScxCre} \textsuperscript{56}, \textit{Osr1GCE} allele (\textit{Osr1GFP-CreERT2})\textsuperscript{15}, \textit{ScxGFP} \textsuperscript{62}, \textit{Rosa26Tomato} (\textit{Ai9})\textsuperscript{28,29}, \textit{Rosa26LoxP-stop-LoxHTB} \textsuperscript{30}, and \textit{Fucci2AR}\textsuperscript{35} (Supplementary Table 3). For temporal
fate mapping using the \textit{Osr1GCE} allele, we used the lineage-specific \textit{Pax7} driven

Fig. 7 BMP gain-of-function experiments in limbs and fibroblast cultures. a–i BMP4/RCAS-producing cells were grafted to the right limbs of E5 chicken
embryos (a). Embryos were fixed 5 days later at E10. In situ hybridisation experiments with mBmp4 probe to transverse limb sections indicate the extent of
ectopic BMP4 expression in dorsal limb muscles (b). c–f Limb sections were immunostained with MF20 (myosins, green) to visualise the muscle pattern.
Asterisks indicate fused muscles in BMP4-grafted limbs (d) versus individual muscles in control limbs (c). e, f High magnifications of dorsal limb muscles
immunostained with MF20 (myosins, green), PAX7 (cyan) and TCF4 (red) antibodies in BMP4 (f) and control limbs (e). g, h Representative fields of
nuclei (DAPI+) in muscles of control and BMP4 limbs (g) and quantification of the number of nuclei per unit area (h). i Data are presented as mean of
values ± SD from 5 BMP4 and 9 control limbs. i Percentages of PAX7+ and TCF4+ nuclei per total number of nuclei in BMP4-grafted limbs versus
controls. Data are presented as mean values ± SD from 3 BMP4 and 6 control limbs. The \textit{p}-value was obtained using the two-tailed Mann–Whitney test.
Asterisks indicate the \textit{p}-values, ****\textit{p}<0.0001. j–l Chicken fibroblast cultures transfected with empty/RCAS (control) (j), BMP4/RCAS (k) or BMPR1ca/
RCAS (l). Immunostaining with PAX7 antibody (red) and DAPI in BMP4- and BMPR1ca-transfected fibroblast cultures compared to controls. m
Quantification of PAX7+ cells versus the total cell number. Data are presented as mean values ± SD from 3 independent experiments.
reporter (Pax7GFP-stop-nlacZ) named Pax7GPL, in which locus accessibility to Cre-mediated recombination allows higher sensitivity for tracking Pax7-expressing cells and their progeny. Adult mice were kept in ventilated cages on a 12-h/12-h light/dark cycle, 22 °C ambient temperature and 34% humidity. Male and female adult mice (Mus musculus, 6–12 weeks of age) were kept on mixed genetic backgrounds C57BL/6JRj or DBA/2JRj (B6D2F1, Janvier Labs). To induce recombination, 5 mg of tamoxifen (Sigma #T5648) were administered by gavage to pregnant females at different time points between E10.5 and E13.5. A 25 mg/ml stock solution in 5% ethanol and 95% sunflower seed oil was prepared by thorough resuspension with rocking at 4 °C.

Surgical grafting experiments in avian embryos

Isotopic/Isochronic quail- or GFP + chicken-into-chicken grafts of presomitic mesoderm. Quail, GFP + chicken and chicken embryos were incubated until they reached 15-somite stage. Of note, limb buds are not formed at this stage and brachial somitic cells have not yet migrated into the forelimb presumptive regions. An artificial dark field was obtained by injecting Indian ink, diluted 1:1 in PBS, beneath the chicken host embryos. Microsurgery was performed on the right side of the host embryos. The 15th somite and the non-segmented presomitic mesoderm was removed over a length corresponding to 5–8 somites. Presomitic mesoderms from either quail (N = 6) or GFP + chicken (N = 4) donor embryos were rinsed in DMEM (PAA Laboratories)/10% foetal calf serum (FCS, Eurobio) and transplanted into chicken hosts submitted to the same ablation. Grafts were
Fig. 8 BMP loss-of-function experiments in limbs and myoblast cultures. a-f NOGGIN/RCAS-producing cells were transfected to right limbs of E5 chicken embryos (a). Embryos were fixed 5 days later at E10. b In situ hybridisation experiments with NOGGIN probes to transverse limb sections indicate the extent of ectopic NOGGIN expression in dorsal limb musculature. c-f Limb sections were immunostained with MF20 (myosins, green) to visualise muscles. e, f High magnifications of dorsal limb muscles immunostained with PAX7 (cyan) and TCF4 (red) antibodies in NOGGIN and control limbs. g, h Representative fields of nuclei (DAPI+) in muscles of control and NOGGIN limbs (g), and quantification of the number of nuclei per unit area (h). h Data are presented as mean ± SD from 4 NOGGIN and 4 control limbs. i Percentages of PAX7+ and TCF4+ nuclei per total number of nuclei in NOGGIN-grafted limbs versus controls. Data are presented as mean values ± SD from 3 NOGGIN and 6 control limbs. The p-value was obtained using the two-tailed Mann–Whitney test. Asterisks indicate the p-values. *p < 0.0001. j-n Myoblast cultures transfected with empty/RCAS (control) or SMAD6/RCAS were induced to differentiation. j–l Immunostaining with PAX7 (j), MYOG (k), and MF20 (myosins) (l) antibodies (green) in control and SMAD6 cultures. m–n RT-qPCR analyses of the expression levels of muscle genes, PAX7, MYF5, MYOD, MYOG, MYHC, BMP-responsive genes ID1, ID2, ID3 (m) and connective tissue genes, SCX, COL3A1, COL6A1, OSRI, TCF4, PDGFRα (n) in SMAD6-transfected and control cell cultures. Gene mRNA levels were normalised to GAPDH and RPS17. The relative mRNA levels were calculated using the 2−ΔΔCt method using the control condition as controls. For each gene, the mRNA levels of the control condition were normalised to 1. Graphs show means ± standard deviations of 8 biological samples (m) and 6 biological samples (n). The exact p-values were obtained using the two-tailed Mann–Whitney test and are noted on the graphs.

Isotopic/isochronic quail-into-chick grafts of limb lateral plate. Quail and chicken embryos were incubated until they reached 18-somite stage. At this stage, neural crest cells did not yet leave the neuromeric plate to colonise the limb bud. The neural tube was excised over a length corresponding to 5–8 somites from GFp + chicken donor embryos and grafted in place of the neural tube of chick host embryos (N = 4) (Supplementary Fig. 6).

For all types of grafts, quail-chicken and GFp + chicken–chicken chimeras were allowed to grow for another 4–8 days and treated for either immunohistochemistry or in situ hybridisation to tissues sections.

Lateral plate electroporation. Chicken embryos were incubated until they reached 19/20-somite stage (E2). The DNA solution was composed of both pT2A-LCMV/factin-membrane/Tomato–T2A-H2B/GFP and CMV/factin-Transposase at a molar ratio of 1:3. The transposase allows the stable integration of membrane- and nuclear-GFP into the chicken genome27. The DNA solution was electroporated with 3 pulses of 50 V, 5 ms duration, 50 ms interpulse interval, 25 μs pulse width. Tendons and connective tissues were detected using Collagen type XII polyclonal antibody48 (Clone 5522, 1:100) and TCF4 polyclonal antibody (Cell signalling cat. # 2865, 1:50, respective). Immunohistochemical staining was detected using the Click-iT Plus EdU Cell Proliferation Kit for Imaging (Invitrogen, C10639). Immune detection was performed using a biotinylated secondary antibody and a streptavidin–alkaline phosphatase conjugate. Sections were counterstained with DAPI (Sigma-Aldrich cat. # D9542, 1:1000) and visualised under a fluorescent microscope (Zeiss). For double staining, the anti-beta-gal (MP Biomedicals cat. # MP 55976, 1:5000), anti-GFP (Abcam cat. # ab13970, 1:1000) and anti-Tomato (Takara cat. # 632496, 1:400) polyclonal antibodies were used.

Grafting experiments in chicken limbs. Chicken embryonic fibroblasts (CEFs) obtained from E10 chicken embryos were transfected with BMP4-RCAS23 or NOGGIN-RCAS23 using the Calcium Phosphate Transfection Kit (Invitroin, France). Cell pellets of ~50–100 μm in diameter were grafted into the limbs of E5 embryos. BMP4-RCASs (N = 8) and NOGGIN-RCASs (N = 4) grafted embryos were harvested at E9/E10 and processed for in situ hybridisation and immunohistochemistry.

EdU labelling. Fifty microliters of EdU (Invitrogen, A10044) solution (5 mg/ml) was injected into the circulation in E10 chick embryos for 1.5 h. Embryos were then fixed and immunohistochemistry was performed on limb transverse sections. EdU was detected using the Click-iT Plus EdU Cell Proliferation Kit for Imaging (Invitroin, C10639).

Immunohistochemistry. For antibody staining, control and manipulated chicken forelimbs were fixed in a paraformaldehyde (PFA) 4% solution in PBS overnight, embedded in gelatin/sucrose and then cut in 12-μm cryostat sections. Mouse tissues were fixed for 2 h in 4% PFA 0.1% triton at 4 °C, washed extensively in PBS, equilibrated in 30% sucrose in PBS, and embedded in OCT for cryo-sectioning. All antibodies (sources, conditions of use, references) are reported in Supplementary Table 3. Quail cells were detected using the QCPN monoclonal antibody (DSHB cat. # QCPN, undiluted supernatant). Differentiated muscle cells were detected using the MF20 monoclonal antibody, MF20 (DSHB cat. # MF20, undiluted supernatant), recognising sarcomeric myosin heavy chains. Mouse myoblasts were detected using the MYOD monoclonal antibody (BD Biosciences cat. # 554130, 1:500) or with a mixture of MYOD and MYOG monoclonal antibodies (DSHB cat. # F5D, 1:30). Chicken myoblasts were detected using the MYOG monoclonal antibody49 (undiluted). Muscle progenitors were detected using the PAX7 monoclonal antibody (DSHB cat. # A3-PAX7, 1:200). Active BMP signalling was detected using the pSMAD polyclonal antibody (Cell signalling cat. # 9516, 1:1000) recognising the complex BMP-activated receptor-phosphorylated SMAD1/5/9. Tendons and connective tissues were detected using Collagen type XII polyclonal antibody48 (Clone 5522, 1:100) and TCF4 polyclonal antibody (Cell signalling cat. # 2865, 1:50, respective). Immunohistochemical staining was detected using the Click-iT Plus EdU Cell Proliferation Kit for Imaging (Invitrogen, C10639). Immune detection was performed using a biotinylated secondary antibody and a streptavidin–alkaline phosphatase conjugate. Sections were counterstained with DAPI (Sigma-Aldrich cat. # D9542, 1:1000). Stained sections were examined using ApoTome.2 (Zeiss), SP5 AOBS confocal (Leica) or LSM700 confocal (Zeiss) microscopes.

Single-cell RNA-sequencing analysis from limb cells. For single-cell RNA sequencing analysis of chicken limb cells at different developmental stages, forelimbs from 2 different E4 embryos, 3 different E6 embryos and 3 different E10 embryos were dissected and dissociated by collagenase and mechanical treatments. Cell concentration was adjusted to 500 cells/μl in the buffer. 500 cells per condition were loaded into the 10x Chromium Chip with the Single Cell 3′ Reagent Kit v3 according to the manufacturer’s protocol. Samples were pooled together at each developmental stage (E4 (N = 2), E6 (N = 3) and E10 (N = 3)) before making the libraries. Libraries were then sequenced by pair with a HighOutput flowcell using an Illumina NextSeq500 sequencer (with the following I7 index pair) 28 bp single-end. After processing the reads (Read 2) and 8 bp (i7 Index). A minimum of 50 000 reads per cell were sequenced and analysed with Cell Ranger Single Cell Software Suite 3.0.2 by 10× Genomics. Raw base call files from the Nextseq 500 were demultiplexed with the cellranger barcode file. This pipeline used STAR21 to align cDNA reads to the Gallus gallus genome (Sequence: GRCg6a, Gallus gallus reference). The used sample size in scRNAseq allowed us to robustly identify cell populations as small as 18 cells at E4, 6 cells at E6 and 66 cells at E10.

The Seurat package (v3.0.1)55 under R (v3.6.1)86 was used to perform downstream clustering analysis on scRNAseq data87. Cells went through a classical downstream clustering analysis on scRNAseq data87. Cells went through a classical downstream clustering analysis on scRNAseq data87.
selections were normalised to the total expression and Log-transformed with the NormalizeData function of Seurat using the nCount median as a scale factor. Highly variable genes were detected with the FindVariableFeatures function (default parameters). The cell cycle effect was regressed out using the ScaleData function. Using highly variable genes as input, principal component analysis was performed on the scaled data in order to reduce dimensionality. Statistically significant principal components were determined by using the JackStrawPlot and the ElbowPlot functions. Cell clusters were generated with the FindNeighbors/FindClusters functions (default parameters except for the number of selected PCs). Different clustering results were generated at different resolutions and for different sets of PCs. Non-linear dimensional reduction (UMAP) and clustering trees using Clustree were used to visualise clustering results and select the most robust and relevant result. The CT and muscle clusters represented the majority of limb cells (90–95% of limb cells), in addition to other clusters encompassing the expected cell populations present in developing limb tissues such as vessels/blood, innervation and ectoderm (Fig. 4a). Muscle clusters gather mononucleated muscle cells, since the plurinucleated myotubes are excluded from the single-cell isolation protocol.

Markers for each cluster were found using the FindAllMarkers function of Seurat (using highly variable genes as an input, default parameters otherwise) that ran Wilcoxon rank-sum tests (p-val adjusted < 0.05). CT and M scores were calculated using the AddModuleScore function for the CT and muscle markers, respectively. For the analysis of the CT/M cells, cells were grouped according to 4 identities (CT, CT/M, M and Other). The CT/M identity is defined by the co-expression (i.e. gene Log-normalised count > 0) within a cell of at least one of the CT markers (PRRX1, TWIST2, PDGFRA, OSRI, SCX) with at least one of muscle markers (PA7X, MYF5, MYOD1, MYOG). CT identity is conferred to all cells allocated to a CT cluster, except the CT/M cells. M identity is conferred to all cells allocated to a muscle cluster, except the CT/M cells. Other identity is conferred to clusters that are neither CT nor muscle, except for the CT/M cells. Differentially expressed genes between the CT/M cells and the combined CT and M cells were found using the FindMarkers function of Seurat (using highly variable genes as an input, default parameters otherwise) that ran Wilcoxon rank-sum tests (p-val adjusted < 0.05). The scRNAseq datasets were then analysed with the angle of these 4 identities using Seurat tools such as Feature plots, Heatmaps and Violin plots. 

Cell cultures. Chicken embryonic fibroblast (CEF) cultures were obtained from E10 chicken embryos. Embryos without heads and viscera were mechanically dissociated. After centrifugation, cells were plated. Myoblast primary cultures were obtained from limbs of E10 chicken embryos.67,68. Empty/RCAS, BMP4/RCAS and BMPR1Ac/RCAS constructions were transplanted to secondary fibroblast cultures. Transfected fibroblasts were left for 5–7 days in culture with cell splitting to allow virus spread. Empty/RCAS, BMP4/RCAS, BMPR1Ac/RCAS fibroblasts were then fixed and processed for immunohistochemistry.

Empty/RCAS, BMP4/RCAS, NOGGIN/RCAS and SMAD6/RCAS constructions and Fucci/RCAS were transplanted to myoblast primary cultures. Transfected myoblasts were left for 5–7 days in culture with cell splitting to allow virus spread. Transfected myoblasts were either analysed in proliferation or in differentiation culture conditions. Empty/RCAS, NOGGIN/RCAS, SMAD6/RCAS or Fucci/RCAS myoblasts were fixed and processed for immunohistochemistry or RT-q-PCR assays to analyse gene expression.

Quantitative real-time PCR. qPCR was performed as described in. Total RNAs were extracted from SMAD6 or control myoblasts. 500 ng of RNA was reverse-transcribed. qPCR was performed on the Applied Biosystems StepOnePlus Real-Time PCR System using SYBR Green PCR master mix (Life Technology, cat. # 4385614) with primers listed in Supplementary Table 4. The relative mRNA levels were calculated using the 2-ΔΔCt method.61. The ΔCt were obtained from Ct normalised with GAPDH and RPS17 levels in each sample.

Cell counting. Cell counting was performed with Fiji (v2.1.0) software. Quantifications were performed on at least 3–5 pictures per experiments taken randomly within each sample. Pictures encompassed the central zone and tips of the muscles. Four to six limbs were counted depending on the experiment. Although labelling in the picture was preferentially restricted to muscle tips, we counted all the muscle cells present on the picture to compare with the numbers derived from the scRNAseq analysis and because it was not possible to define the border between the central zone and the tips. For most experimental settings, the percentage of non-somite-derived/CT-derived XX+ myogenic cells were calculated with the total number of XX+ myogenic cells (for example, Scx-derived PA7X+ cells versus total PA7X+ cells).

Statistics and reproducibility. GraphPad Prism 6 software was used for statistical analysis. Non-parametric tests were used to determine statistical significance, which was set at p-values < 0.05. Unless stated in the figure legend, immunohistochemistry experiments have been performed on at least three limbs.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.
progenitors and satellite cells during development. Dev. Cell 18, 643–654 (2010).

24. Estebanezi de Lima, J., Bonnin, M.-A., Birchmeier, C. & Duprez, D. Muscle contraction is required to maintain the pool of muscle progenitors via YAP and NOTCH during fetal myogenesis. Elife 5, e15593 (2016).

25. Kraft-Sheleg, O. et al. Localized LoxL3-dependent fibropectin oxidation regulates myofiber stretch and integrin-mediated adhesion. Dev. Cell 36, 559–572 (2016).

26. McGrew, M. J. et al. Efficient production of germline transgenic chickens using lentiviral vectors. EMBO Rep. 5, 728–733 (2004).

27. Bourgeois, A. et al. Stable and bicistronic expression of two genes in somite- and lateral plate-derived tissues to study chick limb development. BMC Dev. Biol. 15, 39 (2015).

28. Englert, C. G. A. K. et al. Insertion of Cre into the Pax3 locus creates a new allele of Splotch and identifies unexpected Pax3 derivatives. Dev. Biol. 280, 396–406 (2005).

29. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13, 133–140 (2010).

30. Li, Y. et al. Molecular layer perforant path-associated cells contribute to feed-forward inhibition in the adult dentate gyrus. Proc. Natl Acad. Sci. USA 110, 9106–9111 (2013).

31. Schweitzer, R. et al. Analysis of the tendon cell fate using Scleraxis, a specific marker for tendons and ligaments. Development 128, 3855–3866 (2001).

32. Murchison, N. D. et al. Regulation of tendon differentiation by scleraxis distinguishes force-transmitting tendons from muscle-anchoring tendons. Development 134, 2697–2708 (2007).

33. Esteves de Lima, J. et al. Specification of posterior and lateral plate-derived tissues to study chick limb development. Dev. Biol. 324, 88–98 (2008).

34. Sambasivan, R. et al. Embryonic founders of adult muscle stem cells are primed by the determination gene Mrf4. Dev. Biol. 381, 241–253 (2013).

35. Mort, R. L. et al. Fucci2a: a bicistronic cell cycle reporter that allows Cre using lentiviral vectors. Nat. Commun. 5, 4063 (2014).

36. Yaseen-Badarneh, W. et al. Fibroblast fusion to the muscle extracellular matrix by NFAT/IL-4-mediated cell fusion. Dev. Cell 37, 224–234 (2016).

37. Macias, M. J., Martin-Malpartida, P. & Massagué, J. Structural determinants of myostatin promotes the terminal differentiation of embryonic muscle progenitors. Genes Dev. 22, 668–681 (2008).

38. Koch, M. et al. A novel marker of tissue junctions, collagen XXII. J. Biol. Chem. 279, 22514–22521 (2004).

39. Stuart, T. et al. Comprehensive integration of single-cell data. Cell 177, 1888–1902.e21 (2019).

40. Nature Communications | https://doi.org/10.1038/s41467-021-24157-x | www.nature.com/naturecommunications

41. Ney, T. et al. A language and environment for statistical computing. https://www.r-project.org/

42. Acknowledgements

We thank Sophie Gournet for the illustration. We thank the Roslin Institute (Prof Helen Sang and Dr. Adrian Sherman) for providing us with GFP chicken embryos. The production of the GFP–chicken embryos was supported by grants from BBSRC and the Wellcome Trust. Sequencing and genome mapping of the scRNAseq datasets were performed by the GENOMIC platform at the Cochin Institute (Paris). This work was supported by the CNRS, Inserm, SU, AEFM and FRM. ARTBio was supported by the CNRS, SU, the Institut Français de Bioinformatique (IFB) and by a grant from the SIRIC CURAMUS.

Author contributions

J.E.D.L. initiated the project, conducted the BMP experiments, contributed and analysed the mouse work and performed lateral plate electroporations. C.B. conducted lateral plate grafting experiments, lateral plate electroporations, in situ hybridisation and immunohistochimistry to experimental chicken and mouse limbs. M.-A.B. contributed to somite grafting experiments, performed cell preparation for scRNAseq and analysed the mouse work. E.H. performed confocal imaging of chicken embryos and conducted the bioinformatic analysis of the scRNAseq datasets. G.C. conducted the mouse work and immunohistochimistry for myogenic markers. L.Y. performed GFP chicken and qPCR to experimental chicken and mouse limbs. The authors declare no competing interests.
