



HAL
open science

BMP signalling directs a fibroblast-to-myoblast conversion at the connective tissue/muscle interface to pattern limb muscles

Joana Esteves de Lima, Cédrine Blavet, Marie-Ange Bonnin, Estelle Hirsinger, Glenda Comai, Laurent Yvernogeu, Léa Bellenger, Sébastien Mella, Sonya Nassari, Catherine Robin, et al.

► To cite this version:

Joana Esteves de Lima, Cédrine Blavet, Marie-Ange Bonnin, Estelle Hirsinger, Glenda Comai, et al.. BMP signalling directs a fibroblast-to-myoblast conversion at the connective tissue/muscle interface to pattern limb muscles. 2020. hal-03002436

HAL Id: hal-03002436

<https://hal.science/hal-03002436v1>

Preprint submitted on 12 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 **BMP signalling directs a fibroblast-to-myoblast conversion at the connective**
2 **tissue/muscle interface to pattern limb muscles**

3

4 Joana Esteves de Lima^{1,6*}, Cédrine Blavet^{1*}, Marie-Ange Bonnin^{1*}, Estelle Hirsinger^{1*},
5 Glenda Comai^{2**}, Laurent Yvernogeu^{3**}, Léa Bellenger⁴, Sébastien Mella², Sonya Nassari¹,
6 Catherine Robin³, Ronen Schweitzer⁵, Claire Fournier-Thibault¹, Shahragim Tajbakhsh²,
7 Frédéric Relaix⁶ and Delphine Duprez^{1#}

8

9 * Co-first authors

10 ** Co-second authors

11 # Corresponding author

12

13 ¹ Sorbonne Université, CNRS, Institut Biologie Paris Seine, IBPS-UMR 7622,
14 Developmental Biology Laboratory, F-75005 Paris, France
15 Inserm U1156, F-75005 Paris, France.

16

17 ² Institut Pasteur, Department of Developmental and Stem Cell Biology, CNRS UMR 3738,
18 F-75015 Paris, France.

19

20 ³ Hubrecht Institute-Royal Netherlands Academy of Arts and Sciences (KNAW),
21 Regenerative Medicine Center, University Medical Center Utrecht, Utrecht, The Netherlands.

22

23 ⁴ Sorbonne Université, CNRS, Institut Biologie Paris Seine, IBPS-FR3631, ARTbio
24 bioinformatics platform, Inserm US 037, F-75005 Paris, France.

25

26 ⁵ Research Division, Shriners Hospital for Children, Portland, OR 97239, USA.

27

28 ⁶ Univ Paris Est Creteil, INSERM, EnvA, EFS, AP-HP, IMRB, F-94010 Creteil, France.

29

30

31 **Key words:** Chicken, mouse, limb, patterning, cell fate, skeletal muscle, connective tissues,
32 tendon, PAX7, MYOD, MYOG, OSR1, SCX, BMP4, ID2, PSMAD.

33

34

35

36 **Abstract**

37 Positional information driving limb muscle patterning is contained in lateral plate mesoderm-
38 derived tissues, such as tendon or muscle connective tissue but not in myogenic cells
39 themselves. The long-standing consensus is that myogenic cells originate from the somitic
40 mesoderm, while connective tissue fibroblasts originate from the lateral plate mesoderm. We
41 challenged this model using cell and genetic lineage tracing experiments in birds and mice,
42 respectively, and identified a subpopulation of myogenic cells at the muscle tips close to
43 tendons originating from the lateral plate mesoderm and derived from connective tissue gene
44 lineages. Analysis of single-cell RNA-sequencing data obtained from limb cells at successive
45 developmental stages revealed a subpopulation of cells displaying a dual muscle and
46 connective tissue signature, in addition to independent muscle and connective tissue
47 populations. Active BMP signalling was detected in this junctional cell sub-population and at
48 the tendon/muscle interface in developing limbs. BMP gain- and loss-of-function experiments
49 performed *in vivo* and *in vitro* showed that this signalling pathway regulated a fibroblast-to-
50 myoblast conversion. We propose that localised BMP signalling converts a subset of lateral
51 plate mesoderm-derived fibroblasts to a myogenic fate and establishes a boundary of
52 fibroblast-derived myonuclei at the muscle/tendon interface to control the muscle pattern
53 during limb development.

54

55 **Introduction**

56 Skeletal muscle patterning is a developmental process that controls the formation of each
57 muscle at the right position and time, leading to the highly complex organisation of over 600
58 individual skeletal muscles in human, each with a specific shape, size, innervation and
59 attachment to the skeletal system. The cellular and molecular mechanisms that regulate limb
60 muscle patterning remain poorly understood. The current consensus emerging from
61 experimental embryology in birds is that limb myogenic cells are of somitic mesoderm, while
62 connective tissue (CT) fibroblasts of muscle attachments originate from the lateral plate
63 mesoderm¹⁻⁴. Consistently, genetic lineage tracing experiments in mice have shown that
64 developmental PAX7+ muscle progenitors originate from the *Pax3* lineage in limbs⁵⁻⁷ and
65 that muscle attachments originate from CT fibroblast lineages^{8,9}. Previous embryological
66 surgical experiments in chicken embryos demonstrated that the positional information for
67 limb muscles and associated innervation (motoneuron axons) is contained in the lateral plate
68 mesoderm and not in somitic-derived tissues^{4,10-14}. This positional information for muscle
69 patterning is maintained over time in lateral plate mesoderm-derived tissues, such as muscle
70 CT¹⁵ and tendons¹⁶. The molecular signals produced from lateral plate-derived tissues that
71 drive the correct positioning of limb muscles during development are not fully identified. To
72 date, the TBX4/5, TCF4 and OSR1 transcription factors expressed in limb CT fibroblasts are
73 recognized to act in a non-cell autonomous manner to regulate limb muscle patterning^{9,17,18}.
74 While it is possible that CT fibroblasts and myogenic cells interact through secreted signals at
75 the CT/muscle interface, the precise nature and mechanisms of these interactions are currently
76 unknown. The CT/muscle interface is likely to be the place where muscle patterning occurs
77 and one would expect that these junctional cells show specificities in relation to their interface
78 status. The myotendinous junction is a CT/muscle interface that links muscle to tendon and is
79 fully formed at postnatal stages¹⁹. The mechanisms driving the establishment of the
80 myotendinous junction during development are poorly studied. There is a regionalisation of
81 patterning signals at the tendon/muscle interface, since known signalling molecules display a
82 regionalised expression in a subset of myonuclei at muscle tips close to tendons²⁰⁻²³; however,
83 the functions of these regionalized genes in the establishment of the myotendinous junction
84 are not understood.

85 Here we identify a fibroblast-to-myoblast conversion that takes place at the CT/muscle
86 interface under the control of regionalised BMP signalling to pattern limb muscles.

87
88

89 Results

90 To identify the cell origin operating at the muscle/CT interface in developing limbs, we first
91 revisited the embryological origins of CT and myogenic cell populations using cell lineage
92 experiments in avian embryos and genetic lineage tracing experiments in mice. As previously
93 demonstrated^{1,2}, isotopic/isochronic quail-into-chicken presomitic mesoderm grafts performed
94 at the presumptive limb regions showed that the vast majority of muscle cells derive from the
95 somitic mesoderm (Fig. 1a-d, Extended Data Fig. 1). However, careful examination of these
96 grafts also identified a subpopulation of myonuclei (nuclei of multinucleated myotubes) that
97 did not have a somitic origin. These myonuclei were located at the extremities of myotubes,
98 close to tendons visualized with collagen XII (Fig. 1a-e, Extended Data Fig. 1). Similarly, a
99 subset of MYOG+ myoblasts (Fig. 1f,g) and PAX7+ muscle progenitors (Extended Data Fig.
100 2) located at muscle tips were also not of somitic origin. We conclude that, in contrast to our
101 current understanding, a small fraction of PAX7+ muscle progenitors, MYOG+ myoblasts
102 and myonuclei are not derived from somites. To determine from which embryological regions
103 these non-somitic-derived muscle cells originate, we performed the converse experiments and
104 replaced the limb lateral plate of chicken embryos by the quail equivalent region. Consistent
105 with classical lineage descriptions^{1,2,4}, this type of graft showed the lateral plate mesoderm
106 origin of cartilage, tendon and muscle connective tissue (Extended Data Fig. 3). However, we
107 also found myonuclei of lateral plate mesoderm origin at muscle tips in close vicinity to
108 tendons visualized with *SCX* expression (Fig. 1h-k). These grafting experiments in avian
109 embryos show an unexpected contribution of lateral plate mesoderm-derived cells to the
110 myogenic lineage at the muscle tips close to tendons in limbs.

111 To assess if the unexpected lateral plate mesoderm contribution to limb muscle in
112 chicken is conserved in mice, we analysed PAX7 expression within the *Pax3* lineage. In
113 contrast to the model supporting that *Pax3*+somitic cells give rise to all PAX7+ cells in mice⁵⁻
114 ⁷, we observed sparse PAX7+ cells that were not *Pax3*-derived (Tomato-negative cells), with
115 a preferential location close to tendons at foetal stages (Fig. 2a-c, Extended Data Fig. 4a). To
116 determine the developmental origin of this subpopulation of non-*Pax3*-derived PAX7+cells,
117 we performed genetic lineage tracing analyses with the limb CT fibroblast markers *Scx* and
118 *Osr1*. The bHLH transcription factor, SCX is a recognized tendon cell marker during
119 development, which is involved in tendon formation^{24,25}. Using the *Scx*^{Cre}:*R26*^{stop/Tom} mice^{8,26},
120 we identified PAX7+ and MYOD/MYOG+ myogenic cells within the *Scx* lineage in limbs of
121 E12.5 and E14.5 embryos (Fig. 2d,e, Extended Data Fig. 4b-e). These myogenic cells of CT
122 fibroblast origin were preferentially found surrounding muscle masses at E12.5 and were

123 concentrated close to tendons at E14.5 in mouse limbs (Fig. 2d,e, Extended Data Fig. 4b-e).
124 The zinc finger transcription factor OSR1 is a key regulator of irregular CT differentiation⁹.
125 We performed lineage tracing using *Osr1*^{CreERT2} as a driver^{9,27,28} combined with the *Pax7*
126 reporter mouse (*Pax7*^{GFP-Puro-nlacZ} = *Pax7*^{GPL}), a reporter gene that is not expressed in CT cells,
127 in order to provide a more stringent lineage analysis and circumvent some specificity
128 problems associated with lineage tracing experiments. Using this combination, we identified a
129 subpopulation of PAX7+ muscle progenitors, MYOD/MYOG+ myoblasts and myonuclei that
130 were nlacZ+ (*Osr1* origin) in E15.5 mouse limbs (Fig. 2f-i, Extended Data Fig. 5). Given that
131 PAX7 is expressed in a subset of neural crest cells in chicken limbs, we verified the absence
132 of neural crest cell contribution to the muscle lineage using isotopic/isochronic GFP+chicken-
133 into-chicken grafts of neural tubes and genetic lineage tracing experiments in limbs of
134 *Wnt1*^{Cre}.*R26*^{stop}/*Tom* mice²⁹. Using these approaches, we did not find any contribution of neural
135 crest cells to either myogenic cells or myonuclei, at muscle tips close to tendons in chicken
136 and mouse limbs (Extended Data Fig. 6). The cell lineage tracing experiments in avians and
137 mice led us to conclude that a fraction of myogenic cells at muscle tips close to tendons are
138 derived from limb CT fibroblasts and are not somite/*Pax3*-derived.

139

140 Our results showing a recruitment of CT fibroblasts into the muscle lineage led us to
141 reconsider the mechanisms underlying limb muscle patterning. We propose a fibroblast-to-
142 myoblast conversion model, whereby a population of CT fibroblast progenitors acquire
143 progressively a myogenic signature and are incorporated at the tips of existing myotubes
144 during development. In this model, progenitors would exhibit a dual CT/muscle identity. To
145 search for such cells with dual identity, we performed single-cell RNA-sequencing
146 (scRNAseq) of chicken whole-limb cells at successive developmental stages: E4 a progenitor
147 stage, E6 when major spatial re-arrangements occur for muscle and CTs and E10 when the
148 final muscle pattern is set (Fig. 3). At each stage, we detected muscle and CT clusters that
149 represented the majority of limb cells (90 to 95% of limb cells), in addition to other clusters
150 encompassing the expected cell populations present in developing limb tissues such as
151 vessels, blood, innervation, and ectoderm (Fig. 3a). Skeletal muscle clusters (which only
152 include mononucleated cells) started from 2% at E4 and reached 29% of limb cells at E10,
153 while CT clusters started from 96% at E4 and comprised 63% of limb cells at E10 (Fig. 3a).
154 In addition to the CT and muscle clusters, we identified a subpopulation of cells co-
155 expressing at least one CT marker, *PRRX1*³⁰, *PDGFRA*³¹, *TWIST2*³², *OSR1*⁹, *SCX*²⁴ with one
156 muscle marker, *PAX7*, *MYF5*, *MYOD*, *MYOG* (Fig. 3b, Extended Data Fig. 7a,b). This cell

157 population exhibits a dual CT/muscle (CT/M) identity expressing the Top10 markers
158 identified for CT and for muscle clusters (Fig. 3b). This dual-identity population increased
159 from E4 (0.5% of limb cells) to reach 4% of limb cells at E6, when the major spatial re-
160 arrangements occur for muscle and CTs, and was maintained at 3.3% at E10 when the final
161 muscle and CT pattern is set (Fig. 3b). Among CT markers, *PRRX1* was expressed in the
162 majority of CT/M cells at all stages (85% of CT/M cells at E4, 72% at E6 and 66% at E10).
163 The fraction of *PRRX1*+ cells among CT/M cells was maintained, while the fraction of
164 *PRRX1*+ cells among all limb cells dropped with time from 88% at E4 to 25% at E10
165 (Extended Data Fig. 7c), suggesting that *PRRX1* expression was actively retained in CT/M
166 cells. *TWIST2*, expressed in 15 to 29% of CT/M cells was the second major contributor to
167 CT/M identity (Extended Data Fig. 7c). In contrast, the contribution of *SCX* to the CT/M
168 identity decreased with time from 27% to 6% of CT/M cells (Extended Data Fig. 7c). Twenty
169 to 30% of the CT/M cells expressed more than one CT marker and the major combination at
170 all stages was *PRRX1/TWIST2*. Altogether, these results point to a key role for *PRRX1* and
171 *TWIST2* in the CT/M identity and that a *TWIST2* cell population contributes to
172 developmental muscles in addition to adult skeletal muscles³². The myogenic markers, *PAX7*
173 and *MYOD1* showed similar distributions in CT/M cells, as a majority of CT/M cells
174 expressed *PAX7* (42 to 57% of CT/M cells) and *MYOD1* (36 to 57% of CT/M cells), while
175 much fewer cells expressed *MYOG* (10 to 15% of CT/M cells). These observations are
176 suggestive of an immature state of these bi-potent CT/M cells (Extended Data Fig. 7c).

177 *ID* genes are recognized transcriptional readouts of BMP activity³³. *ID2* and *ID3* genes
178 were expressed in a similar or greater fraction of CT/M cells compared to CT and M cells
179 (Fig. 3c). *ID1* showed an interesting expression kinetic as it peaks at E6 when muscle and CT
180 spatial re-arrangements occur (Fig. 3c). Consistent with *ID* gene expression in this population
181 with a dual identity (Fig. 3c), BMP activity visualized with *ID2* and pSMAD1/5/9 expression
182 was enriched at the CT/muscle interface, in between muscle masses and tendon primordia at
183 E6 (Fig. 3d) and at muscle tips close to tendon at E8/E9 (Fig. 3e,f, Extended Data Fig. 8).
184 Interestingly, we detected 2 populations of pSMAD1/5/9+ myonuclei at muscle tips, one
185 somite-derived (yellow arrows in Extended Data Fig. 9) and one non-somite-derived (white
186 arrows in Extended Data Fig. 9) at both E6 and E9 stages. This is reminiscent of the two
187 distinct myonuclear populations recently identified in adult mouse muscle at the
188 myotendinous junction, with one being enriched with CT-associated collagen genes³⁴; leading
189 to the compelling notion that the myonuclei with a fibroblast signature could be of CT origin.
190 Altogether, these results show the existence of a BMP-responsive population with a dual

191 CT/muscle identity that could correspond to the fibroblast-derived myogenic cells found at
192 muscle tips close to tendon.

193

194 Based on BMP activity at the CT/muscle interface (Fig. 3d-f, Extended Data Fig. 8)
195 and in a cell subpopulation with a dual CT/M identity (Fig. 3c) in chicken limbs, we
196 hypothesised that BMP signalling would be involved in the recruitment of a subpopulation of
197 CT fibroblasts towards a muscle fate. To test the ability of BMP signalling to control the
198 fibroblast-to-myoblast conversion, we overexpressed BMP activity in chicken limbs and
199 assessed the consequences for PAX7+ myogenic progenitors and TCF4+ CT fibroblasts (Fig.
200 4a-g). Exposure of chicken limbs to retroviral BMP4 led to an increase in the number of
201 PAX7+ muscle progenitors (Fig. 4a,c,f) as previously observed²¹. However, consistent with
202 the cell fate conversion hypothesis, the increase of PAX7+ progenitors occurred at the
203 expense of TCF4+ fibroblasts (Fig. 4a-f). The inverse correlation between PAX7+ and
204 TCF4+ cells upon BMP exposure was not accompanied by changes in the total cell number
205 (Fig. 4g). Moreover, no change in the proliferation rate of PAX7+ cells was observed upon
206 BMP4 exposure in chicken limbs (Extended Data Fig. 10a-g). This abnormal increase in
207 PAX7+ cells led to muscle patterning defects as BMP4 overexpression induced muscle
208 splitting defects, visualised by muscle fusion (Fig. 4a-c,e,f), due to the reduction of TCF4+
209 fibroblasts known to control muscle patterning¹⁷. Because there is no patterning process *in*
210 *vitro*, BMP overexpression was then applied to myoblast primary cultures in proliferation
211 culture conditions. Consistent with our conversion hypothesis and patterning defects observed
212 *in vivo*, BMP4 treatment did not alter the number of PAX7+ cells in myoblast cultures
213 (Extended Data Fig. 10h-j). Our interpretation is that, in contrast to the *in vivo* situation where
214 BMP-competent fibroblasts are present to become myoblasts upon BMP exposure, the
215 contaminating fibroblasts inherent to primary myoblast cultures were not present in enough
216 quantity to generate a detectable conversion phenotype. The absence of any BMP4 effect on
217 the number PAX7+ cells in myoblast cultures is in support of the fibroblast-to-myoblast
218 conversion upon BMP4 exposure in chicken limbs. To mimic the *in vivo* conversion, we
219 applied retroviral BMP4 or BMPR1Aca (a constitutive active form of the BMPR1A) to CEFs
220 (chicken embryonic fibroblasts) that were isolated from whole embryos. BMP over-activation
221 induced the appearance of PAX7+ cells compared to control cultures, albeit with a low
222 efficacy (Fig. 4h,i). This indicates that a subset of this mixed fibroblast population can
223 convert into PAX7+ cells when exposed to BMP *in vitro*.

224 BMP loss-of-function experiments in chicken limbs, by overexpression of the
225 antagonist NOGGIN induced the converse phenotype, *i.e.* an increase in the number of
226 TCF4+ fibroblasts at the expense of PAX7+muscle progenitors with no change in cell number
227 (Fig. 4j-n). The behaviour of TCF4+ fibroblasts upon BMP misexpression is consistent with
228 the decrease and increase in the expression of the *OSRI* CT marker upon BMP gain- and loss-
229 of-function experiments in chicken limbs, respectively³⁵. Consistent with the *in vivo*
230 experiments in chicken embryos, the inhibition of BMP signalling in myoblast cultures led to
231 a fibroblast phenotype (Fig. 4o-q, Extended Data Fig. 10k). BMP inhibition was achieved
232 with the overexpression of SMAD6 (Fig. 4o-q), an inhibitor of the pSMAD1/5/8 pathway³³ or
233 the BMP antagonist NOGGIN (Extended Data Fig. 10k). BMP inhibition led to a reduction of
234 PAX7+, MYOD+ and MYOG+ cells and myotubes in myoblast cultures in differentiation
235 culture conditions (Fig. 4o). This was associated with a decrease in the mRNA expression
236 levels of muscle lineage markers from progenitor to differentiation state (*PAX7*, *MYF5*,
237 *MYOD*, *MYOG*, *MYHC*) (Fig. 4p). The SMAD6-treated muscle cells adopted a fibroblast-like
238 phenotype compared with myotubes observed in control myoblast cultures (Fig. 4o) and
239 displayed an increased expression of CT markers, such as *COL3A1*, *COL6A1* and *OSRI* (Fig.
240 4q). These results demonstrate that BMP inhibition converts myoblasts towards a fibroblastic
241 phenotype.

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

245

246 **Conclusion**

247 This work shows that the tendon/muscle interface partly results from a fibroblast-to-myoblast
248 conversion driven by localized BMP signalling. These data lead us to propose a scenario in
249 which BMP locally drives the conversion of CT fibroblasts to myoblasts that are incorporated
250 into myotubes at muscle tips close to tendons. This process leads to myonuclei of fibroblast
251 origin at the tendon/muscle interface. Persistent BMP signalling regionalised at muscle tips
252 then provides an *in vivo* safeguard to maintain the muscle identity of fibroblast-derived
253 muscle cells. Previous BrdU/EdU-based experiments have shown a preferential incorporation
254 of nuclei in myotubes at muscle tips during foetal and postnatal myogenesis³⁶⁻³⁸. Although
255 these BrdU/EdU-based experiments never addressed the origin of the newly incorporated
256 nuclei, it was always assumed that this preferential incorporation at the muscle tips was with
257 somitic-derived nuclei. We now propose that the preferential nucleus incorporation at muscle

258 tips of myotubes occurs with nuclei of fibroblast origin. We see these muscle cells of CT
259 fibroblast origin are a source of positional information for muscle patterning, possibly
260 regulating the localisation of myoblast fusion events and thereby the shaping and orientation
261 of growth of limb skeletal muscles during foetal development. These muscle cells of
262 fibroblast origin are reminiscent of muscle founder cells that drive muscle identity in
263 *Drosophila* embryos³⁹. Future studies will explore whether this mechanism can be reactivated
264 in adult where it could be relevant in the context of muscle regeneration.

265

266

267 **Material and methods**

268

269 **Chicken and quail embryos**

270 Fertilized chicken and quail eggs from commercial sources (White Leghorn and JA57 strain,
271 Japanese quail strain, Morizeau, Dangers, France) were incubated at 37.5°C. Before 2 days of
272 incubation, chicken and quail embryos were staged according to somite number. After 2 days
273 of incubation, embryos were staged according to days of *in ovo* development. GFP+chicken
274 were generated and provided by the Roslin Institute⁴⁰.

275

276 **Mouse strains**

277 Animals were handled according to European Community guidelines and protocols were
278 validated by the ethic committee of the French Ministry. Embryos were dated taking the day
279 of the vaginal plug as E0.5. Mouse lines used in this study were kept on a mixed C57BL/6JRj
280 and DBA/2JRj genetic background (B6D2F1, Janvier862 Labs). The following strains were
281 described previously: *Pax3^{Cre}:R26^{Tomato}* (Ai9)^{26,41}, *Wnt1^{Cre29}*, *Scx^{Cre42}*, *Osr1^{GCE}* allele
282 (*Osr1^{eGFP-CreERT2}*)²⁸, *Pax7^{GPL}* reporter (*Pax7^{nGFP-stop-nlacZ}*)²⁷ and *Scx^{Cre}:R26^{Tomato}* (Extended
283 Ressources Table 1). For temporal fate mapping using the *Osr1^{GCE}* allele, we used the
284 lineage-specific Pax7-driven reporter (*Pax7^{GPL}*), in which locus accessibility to Cre-mediated
285 recombination allows higher sensitivity for tracking Pax7-expressing cells and their
286 progeny²⁷. To induce recombination, 5 mg of tamoxifen (Sigma #T5648) were administered
287 by gavage to pregnant females. A 25 mg/ml stock solution in 5% ethanol and 95% sunflower
288 seed oil was prepared by thorough resuspension with rocking at 4°C.

289

290 **Surgical grafting experiments in avian embryos**

291 *Isotopic/Isochronic quail- or GFP+chicken-into-chicken grafts of presomitic mesoderm.*
292 Quail, GFP+chicken and chicken embryos were incubated until they reached 15-somite stage.
293 Of note, limb buds are not formed at this stage and brachial somitic cells have not yet
294 migrated into the forelimb presumptive regions. An artificial dark field was obtained by
295 injecting Indian ink, diluted 1:1 in PBS, beneath the chicken host embryos. Microsurgery was
296 performed on the right side of the host embryos. The 15th somite and the non-segmented
297 presomitic mesoderm was removed over a length corresponding to 5–8 somites. Presomitic
298 mesoderms from either quail (N=6) or GFP+chicken (N=4) donor embryos were rinsed in
299 DMEM (PAA Laboratories)/10% foetal calf serum (FCS, Eurobio) and transplanted into
300 chicken hosts submitted to the same ablation. Grafts were performed according to the original
301 dorso-ventral and antero–posterior orientations (Fig. 1a-g, Extended Data Fig.1,2,9).

302 *Isotopic/Isochronic quail-into-chicken grafts of limb lateral plate.* Quail and chicken embryos
303 were incubated until they reached 18-somite stage. Of note, limb buds are not formed at this
304 stage, and brachial somitic cells have not yet migrated into the forelimb lateral plate. The
305 forelimb lateral plate mesoderm at the level of the 15th somite and over a length of 3 somites
306 was excised from host chicken embryos and replaced by their quail counterpart (N=3) (Fig.
307 1h-k, Extended Data Fig.3).

308 *Isotopic/isochronic GFP+chicken-into-chicken grafts of neural tube.* GFP+chicken and
309 chicken embryos were incubated until they reached 12- to 14-somite stage. At this stage,
310 neural crest cells did not yet leave the neural tube to colonize the limb bud. The neural tube
311 was excised over a length corresponding to 5–8 somites from GFP+ chicken donor embryos
312 and grafted in place of the neural tube of chicken host embryos (N=4) (Extended Data Fig. 6).

313 For all types of grafts, quail-chicken and GFP+chicken-chicken chimeras were
314 allowed to grow for another 4 to 8 days and treated for either immunohistochemistry or *in situ*
315 hybridization to tissue sections.

316

317 **Grafts of BMP4-RCAS and NOGGIN-RCAS expressing cells in chicken limbs**

318 Chicken embryonic fibroblasts (CEFs) obtained from E10 chicken embryos were transfected
319 with BMP4-RCAS or NOGGIN-RCAS using the Calcium Phosphate Transfection Kit
320 (Invitrogen, France). Cell pellets of approximately 50–100 µm in diameter were grafted into
321 the limbs of E5 embryos as previously described²¹. BMP4-RCAS- (N=8) and NOGGIN-
322 RCAS- (N=4) grafted embryos were harvested at E9/E10 and processed for *in situ*
323 hybridization and immunohistochemistry. Cell counting was performed with Fiji software.

324

325 **Immunohistochemistry**

326 For antibody staining, control and manipulated chicken forelimbs were fixed in a
327 paraformaldehyde 4% solution, embedded in gelatin/sucrose and then cut in 12- μ m cryostat
328 sections. All antibodies (sources, conditions of use, references) are reported in Extended
329 Ressources Table 1. Quail cells were detected using the QCPN antibody. Differentiated
330 muscle cells were detected using the monoclonal antibody, MF20 recognizing sarcomeric
331 myosin heavy chains. Myoblasts were detected using the MYOD and MYOG antibodies.
332 Muscle progenitors were detected using the monoclonal PAX7 antibody. Active BMP
333 signalling was detected using the polyclonal pSMAD antibody recognizing the complex
334 BMP-activated receptor-phosphorylated SMAD1/5/9. Tendons were detected using Collagen
335 type XII antibody. The following secondary antibodies (Molecular probes) were used: goat
336 anti-mouse IgG coupled to Alexa fluor 555, goat anti-mouse IgG1 coupled to Alexa fluor 568,
337 goat anti-mouse IgG2 coupled to Alexa fluor 488, and goat anti-rabbit IgG coupled with
338 Alexa fluor 488 or 555. To label nuclei, sections were incubated for 15 minutes with DAPI.
339 Stained sections were examined using the Apotome.2 microscope (Zeiss) or LSM700
340 confocal microscope (Zeiss).

341

342 ***In situ* hybridization to tissue sections**

343 Normal or manipulated embryos were fixed in a paraformaldehyde 4% solution. Limbs were
344 cut in 12- μ m cryostat transverse sections and processed for *in situ* hybridization. Alternating
345 serial sections from embryos were hybridized with probe 1 and probe 2. The digoxigenin-
346 labelled mRNA probes were used as described: SCX, ID2, BMP4, NOGGIN (chicken
347 probes), and mBmp4 (mouse probe)^{21,35}.

348

349 **Single-cell RNA-sequencing analysis from limb cells**

350 For scRNAseq analysis of chicken limb cells at different developmental stages, forelimbs
351 from 2 different E4 embryos, 3 different E6 embryos and 3 different E10 embryos were
352 dissected and dissociated by collagenase and mechanical treatments. Cell concentration was
353 adjusted to 5000 cells/ μ l in buffer. 5000 cells per conditions were loaded into the 10X
354 Chromium Chip with the Single Cell 3' Reagent Kit v3 according to the manufacturer's
355 protocol. Samples were pooled together at each developmental stage (E4 (N=2), E6 (N=3)
356 and E10 (N=3)) before making the libraries. Libraries were then sequenced by pair with a
357 HighOutput flowcel using an Illumina Nextseq 500 with the following mode (150 HO): 28

358 base-pairs (bp) (Read1), 125 bp (Read 2) and 8 bp (i7 Index). A minimum of 50 000 reads per
359 cell were sequenced and analyzed with Cell Ranger Single Cell Software Suite 3.0.2 by 10X
360 Genomics. Raw base call files from the Nextseq 500 were demultiplexed with the cellranger
361 mkfastq pipeline into library specific FASTQ files. The FASTQ files for each library were
362 then processed independently with the cellranger count pipeline. This pipeline used STAR21
363 to align cDNA reads to the *Gallus gallus* genome (Sequence: GRCg6a, *Gallus gallus*
364 reference). The used sample-size in scRNAseq allowed us to robustly identify cell
365 populations as small as 18 cells at E4, 6 cells at E6 and 66 cells at E10.

366 The Seurat package (v3.0) was used to perform downstream clustering analysis on scRNAseq
367 data⁴³. Cells went through a classical Quality Control using the number of detected genes per
368 cell (nFeatures), the number of mRNA molecules per cell (nCounts) and the percentage of
369 expression of mitochondrial genes (pMito) as cut-offs. Outliers on a nFeature vs nCount plot
370 were manually identified and removed from the dataset. Most importantly for this study,
371 potential doublets were identified by running the Scrublet algorithm⁴⁴ and then removed from
372 the dataset. Gene counts for cells that passed the above selections were normalized to the total
373 expression and Log-transformed with the NormalizeData function of Seurat using the nCount
374 median as scale factor. Highly variable genes were detected with the FindVariableFeatures
375 function (default parameters). Cell cycle effect was regressed out using the ScaleData
376 function. Using highly variable genes as input, principal component analysis was performed
377 on the scaled data in order to reduce dimensionality. Statistically significant principal
378 components were determined by using the JackStrawPlot and the ElbowPlot functions. Cell
379 clusters were generated with the FindNeighbors/FindClusters functions (default parameters
380 except for the number of selected PCs). Different clustering results were generated at different
381 resolutions and for different sets of PCs. Non-linear dimensional reduction (UMAP) and
382 clustering trees using Clustree (Zappia 2018 Gigascience doi:gigascience/diy083) were used
383 to visualize clustering results and select for the most robust and relevant result. Differentially
384 expressed genes were found using the FindAllMarkers function of Seurat (using highly
385 variable genes as an input, default parameters otherwise) that ran Wilcoxon rank sum tests.

386 For the analysis of the CT/M cells, cells were grouped according to 4 identities (CT, CT/M,
387 M and Other). The CT/M identity is defined by the co-expression (i.e gene Log-normalized
388 count > 0) within a cell of at least one of the CT markers (PRRX1, TWIST2, PDGFRA,
389 OSR1, SCX) with at least one of muscle markers (PAX7, MYF5, MYOD1, MYOG). CT
390 identity is conferred to all cells allocated to a CT cluster, except the CT/M cells. M identity is
391 conferred to all cells allocated to a muscle cluster, except the CT/M cells. « Other » identity is

392 conferred to all cells to clusters that are neither CT nor muscle, except for the CT/M cells.
393 The scRNAseq datasets were then analysed with the angle of these 4 identities using Seurat
394 tools such as Feature plots, Heatmaps and Violin plots.

395

396 **Cell cultures**

397 Chicken embryonic fibroblast (CEF) cultures were obtained from E10 chicken embryos.
398 Briefly, 10 embryos without heads and viscera were mechanically dissociated. After
399 centrifugation, cells were plated. Myoblast primary cultures were obtained from limbs of E10
400 chicken embryos as already described⁴⁵.

401 Empty/RCAS, BMP4/RCAS, BMPR1ACa/RCAS²¹ were transfected to primary fibroblast
402 cultures. Transfected fibroblasts were left for 5 to 7 days in culture with cell splitting to allow
403 virus spread. Empty/RCAS, BMP4/RCAS, BMPR1ACa/RCAS fibroblasts were then fixed
404 and processed for immunohistochemistry.

405 Empty/RCAS, BMP4/RCAS, NOGGIN/RCAS²¹ and SMAD6/RCAS were transfected to
406 myoblast primary cultures. Transfected myoblasts were left for 5 to 7 days in culture with cell
407 splitting to allow virus spread. Transfected myoblasts were either analysed in proliferation or
408 in differentiation culture conditions. Empty/RCAS, NOGGIN/RCAS, SMAD6/RCAS
409 myoblasts were fixed and processed for immunohistochemistry or RT-q-PCR assays to
410 analyse gene expression.

411

412 **Quantitative real-time PCR**

413 qPCR was performed as described in²². Total RNAs were extracted from SMAD6 or control
414 myoblasts. 500 ng of RNA was reverse-transcribed. qPCR was performed using SYBR Green
415 PCR master Mix (Applied Biosystems) with primers listed in Extended Resources Table 2.
416 The relative mRNA levels were calculated using the $2^{-\Delta\Delta Ct}$ method⁴⁶. The ΔCt were
417 obtained from Ct normalized with GAPDH and RPS17 levels in each sample.

418

419 **Statistics.** GraphPad Prism 6 software was used for statistical analysis. Non-parametric tests
420 were used to determine statistical significance, which was set at p values <0.05.

421 **GEO data accession number.** Raw sequencing data will be deposited in the NCBI Gene
422 Expression Omnibus database (<https://www.ncbi.nlm.nih.gov/geo/>) upon acceptance.

423 **Code availability.** Data analysis was performed with standard pipelines in Seurat R packages,
424 as described in the Methods. Scripts will be made available upon request.

425

426 **Acknowledgements:**

427 We thank Sophie Gournet for illustration. This work was supported by the CNRS, Inserm,
428 SU, AFM and FRM. We thank the Roslin Intitute (Prof Helen Sang and Dr. Adrian Sherman)
429 for providing us with GFP⁺ chicken eggs. The production of the GFP⁺ chicken embryos was
430 supported by grants from BBSRC and the Wellcome Trust. ARTbio was supported by the
431 CNRS, SU, the Institut Français de Bioinformatique (IFB) and by a grant from the SIRIC
432 CURAMUS.

433

434 **Author contributions:**

435 JEL initiated the project, conducted the BMP experiments and contributed to Pax3 and Scx
436 mouse work. CB conducted lateral plate grafting experiments, in situ hybridization and
437 immunohistochemistry to experimental chicken and mouse limbs. MAB contributed to somite
438 grafting experiments and performed cell preparation for scRNAseq. EH performed confocal
439 imaging of chicken embryos and conducted the bioinformatic analysis of the scRNAseq
440 datasets. GC conducted the mouse work with the Osr1, Pax3, Wnt1 drivers and
441 immunohistochemistry with MYOG/MYOD antibodies. LY, CR performed GFP chicken
442 (somite and neural tube) and quail (somite) grafting experiments. LB, SM contributed to the
443 design of the scRNAseq analysis methods. RS provided the Scx:Cre embryos. SN, CFT
444 performed qPCR on fibroblasts. JEL, EH, GC, LY, CR, CFT, ST, FR discussed about the
445 project and contributed to write the manuscript. DD supervised the study, analysed the results,
446 pictured most of the data and organised the figures, wrote the manuscript and acquired
447 funding.

448

449

450 **References**

- 451 1. Christ, B., Jacob, H. J. & Jacob, M. Experimental analysis of the origin of the wing
452 musculature in avian embryos. *Anat. Embryol.* **150**, 171–186 (1977).
- 453 2. Chevallier, A., Kieny, M. & Mauger, A. Limb-somite relationship: origin of the limb
454 musculature. *J Embryol Exp Morphol* **41**, 245–258 (1977).
- 455 3. Ordahl, C. P. & Le Douarin, N. M. Two myogenic lineages within the developing somite.
456 *Development* **114**, 339–353 (1992).
- 457 4. Duprez, D. Signals regulating muscle formation in the limb during embryonic
458 development. *International Journal of Developmental Biology* **46**, 915–926 (2002).
- 459 5. Relaix, F., Rocancourt, D., Mansouri, A. & Buckingham, M. A Pax3/Pax7-dependent
460 population of skeletal muscle progenitor cells. *Nature* **435**, 948–953 (2005).
- 461 6. Kassar-Duchossoy, L. Pax3/Pax7 mark a novel population of primitive myogenic cells
462 during development. *Genes & Development* **19**, 1426–1431 (2005).
- 463 7. Hutcheson, D. A., Zhao, J., Merrell, A., Haldar, M. & Kardon, G. Embryonic and fetal
464 limb myogenic cells are derived from developmentally distinct progenitors and have
465 different requirements for -catenin. *Genes & Development* **23**, 997–1013 (2009).
- 466 8. Huang, A. H. *et al.* Requirement for scleraxis in the recruitment of mesenchymal
467 progenitors during embryonic tendon elongation. *Development* **146**, (2019).
- 468 9. Vallecillo-García, P. *et al.* Odd skipped-related 1 identifies a population of embryonic
469 fibro-adipogenic progenitors regulating myogenesis during limb development. *Nat*
470 *Commun* **8**, 1218 (2017).
- 471 10. Aoyama, H. & Asamoto, K. Determination of somite cells: independence of cell
472 differentiation and morphogenesis. *Development* **104**, 15–28 (1988).

- 473 11. Lance-Jones, C. Development of neuromuscular connections: guidance of motoneuron
474 axons to muscles in the embryonic chick hindlimb. *Ciba Found. Symp.* **138**, 97–115
475 (1988).
- 476 12. Lance-Jones, C. & Dias, M. The influence of presumptive limb connective tissue on
477 motoneuron axon guidance. *Dev. Biol.* **143**, 93–110 (1991).
- 478 13. Michaud, J. L., Lapointe, F. & Le Douarin, N. M. The dorsoventral polarity of the
479 presumptive limb is determined by signals produced by the somites and by the lateral
480 somatopleure. *Development* **124**, 1453–1463 (1997).
- 481 14. Lance-Jones, C. & Landmesser, L. Motoneurone projection patterns in embryonic chick
482 limbs following partial deletions of the spinal cord. *J. Physiol. (Lond.)* **302**, 559–580
483 (1980).
- 484 15. Kieny, M., Pautou, M. P., Chevallier, A. & Mauger, A. Spatial organization of the
485 developing limb musculature in birds and mammals. *Bibl Anat* 65–90 (1986).
- 486 16. Kardon, G. Muscle and tendon morphogenesis in the avian hind limb. *Development* **125**,
487 4019–4032 (1998).
- 488 17. Kardon, G., Harfe, B. D. & Tabin, C. J. A Tcf4-positive mesodermal population provides
489 a prepattern for vertebrate limb muscle patterning. *Dev. Cell* **5**, 937–944 (2003).
- 490 18. Hasson, P. *et al.* Tbx4 and tbx5 acting in connective tissue are required for limb muscle
491 and tendon patterning. *Dev. Cell* **18**, 148–156 (2010).
- 492 19. Charvet, B., Ruggiero, F. & Le Guellec, D. The development of the myotendinous
493 junction. A review. *Muscles Ligaments Tendons J* **2**, 53–63 (2012).
- 494 20. Edom-Vovard, F., Bonnin, M.-A. & Duprez, D. Misexpression of Fgf-4 in the Chick
495 Limb Inhibits Myogenesis by Down-Regulating Frk Expression. *Developmental Biology*
496 **233**, 56–71 (2001).

- 497 21. Wang, H., Noulet, F., Edom-Vovard, F., Le Grand, F. & Duprez, D. Bmp Signaling at the
498 Tips of Skeletal Muscles Regulates the Number of Fetal Muscle Progenitors and Satellite
499 Cells during Development. *Developmental Cell* **18**, 643–654 (2010).
- 500 22. Esteves de Lima, J., Bonnin, M.-A., Birchmeier, C. & Duprez, D. Muscle contraction is
501 required to maintain the pool of muscle progenitors via YAP and NOTCH during fetal
502 myogenesis. *Elife* **5**, (2016).
- 503 23. Kraft-Sheleg, O. *et al.* Localized LoxL3-Dependent Fibronectin Oxidation Regulates
504 Myofiber Stretch and Integrin-Mediated Adhesion. *Dev. Cell* **36**, 550–561 (2016).
- 505 24. Schweitzer, R. *et al.* Analysis of the tendon cell fate using Scleraxis, a specific marker for
506 tendons and ligaments. *Development* **128**, 3855–3866 (2001).
- 507 25. Murchison, N. D. *et al.* Regulation of tendon differentiation by scleraxis distinguishes
508 force-transmitting tendons from muscle-anchoring tendons. *Development* **134**, 2697–2708
509 (2007).
- 510 26. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization
511 system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
- 512 27. Sambasivan, R. *et al.* Embryonic founders of adult muscle stem cells are primed by the
513 determination gene Mrf4. *Dev. Biol.* **381**, 241–255 (2013).
- 514 28. Mugford, J. W., Sipilä, P., McMahon, J. A. & McMahon, A. P. Osr1 expression
515 demarcates a multi-potent population of intermediate mesoderm that undergoes
516 progressive restriction to an Osr1-dependent nephron progenitor compartment within the
517 mammalian kidney. *Dev. Biol.* **324**, 88–98 (2008).
- 518 29. Danielian, P. S., Muccino, D., Rowitch, D. H., Michael, S. K. & McMahon, A. P.
519 Modification of gene activity in mouse embryos in utero by a tamoxifen-inducible form
520 of Cre recombinase. *Curr. Biol.* **8**, 1323–1326 (1998).

- 521 30. Logan, M. *et al.* Expression of Cre Recombinase in the developing mouse limb bud
522 driven by a Prxl enhancer. *Genesis* **33**, 77–80 (2002).
- 523 31. Olson, L. E. & Soriano, P. Increased PDGFRalpha activation disrupts connective tissue
524 development and drives systemic fibrosis. *Dev. Cell* **16**, 303–313 (2009).
- 525 32. Liu, N. *et al.* A Twist2-dependent progenitor cell contributes to adult skeletal muscle.
526 *Nat. Cell Biol.* **19**, 202–213 (2017).
- 527 33. Macias, M. J., Martin-Malpartida, P. & Massagué, J. Structural determinants of Smad
528 function in TGF- β signaling. *Trends Biochem. Sci.* **40**, 296–308 (2015).
- 529 34. Kim, M. *et al.* Single-nucleus transcriptomics reveals functional compartmentalization in
530 syncytial skeletal muscle cells. *bioRxiv* 2020.04.14.041665 (2020)
531 doi:10.1101/2020.04.14.041665.
- 532 35. Nassari, S. *et al.* The chemokines CXCL12 and CXCL14 differentially regulate
533 connective tissue markers during limb development. *Sci Rep* **7**, 17279 (2017).
- 534 36. Kitiyakara, A. & Angevine, D. M. FURTHER STUDIES ON REGENERATION AND
535 GROWTH IN LENGTH OF STRIATED VOLUNTARY MUSCLE WITH ISOTOPES
536 P32 AND THYMIDINE-H3. *Nihon Byori Gakkai Kaishi* **52**, 180–183 (1963).
- 537 37. Gu, J.-M. *et al.* An NF- κ B--EphrinA5-Dependent Communication between NG2(+)
538 Interstitial Cells and Myoblasts Promotes Muscle Growth in Neonates. *Dev. Cell* **36**, 215–
539 224 (2016).
- 540 38. Zhang, M. & McLennan, I. S. During secondary myotube formation, primary myotubes
541 preferentially absorb new nuclei at their ends. *Dev. Dyn.* **204**, 168–177 (1995).
- 542 39. de Joussineau, C., Bataillé, L., Jagla, T. & Jagla, K. Diversification of muscle types in
543 *Drosophila*: upstream and downstream of identity genes. *Curr. Top. Dev. Biol.* **98**, 277–
544 301 (2012).

- 545 40. McGrew, M. J. *et al.* Efficient production of germline transgenic chickens using lentiviral
546 vectors. *EMBO Rep.* **5**, 728–733 (2004).
- 547 41. Engleka, K. A. *et al.* Insertion of Cre into the Pax3 locus creates a new allele of Splootch
548 and identifies unexpected Pax3 derivatives. *Dev. Biol.* **280**, 396–406 (2005).
- 549 42. Blitz, E. *et al.* Bone ridge patterning during musculoskeletal assembly is mediated
550 through SCX regulation of Bmp4 at the tendon-skeleton junction. *Dev. Cell* **17**, 861–873
551 (2009).
- 552 43. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual
553 Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).
- 554 44. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell
555 Doublets in Single-Cell Transcriptomic Data. *Cell Syst* **8**, 281-291.e9 (2019).
- 556 45. Havis, E. *et al.* Sim2 prevents entry into the myogenic program by repressing MyoD
557 transcription during limb embryonic myogenesis. *Development* **139**, 1910–1920 (2012).
- 558 46. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C T
559 method. *Nat Protoc* **3**, 1101–1108 (2008).
- 560
- 561

562 **Figure legends**

563

564 **Figure 1**

565 **A subpopulation of muscle cells do not originate from somites but from lateral plate in**
566 **chicken limbs**

567 **(a-g)** Limb muscle organisation after quail-into-chicken presomitic mesoderm grafts. **(a)**
568 Schematic of limb presomitic mesoderm grafts. **(b,c)** Transverse and adjacent limb sections of
569 E9 presomitic mesoderm grafted-embryos were immunostained with QCPN (quail nuclei) and
570 MF20 (myosins) (b), and with QCPN (quail nuclei) and ColXII (tendons) antibodies (c). **(d)**
571 High magnification of FCU, a ventral and posterior muscle immunostained with QCPN (quail
572 nuclei), MF20 (myosins) and ColXII (tendons) antibodies combined with DAPI (nuclei). **(e)**
573 High magnification of muscle tips (close to tendons), squared in (d), showing non-quail nuclei
574 within myosin+ cells (arrows). **(f)** High magnification of FCU immunostained with QCPN
575 (quail nuclei) and MYOG antibodies. **(g)** Focus on muscle tip regions showing MYOG+
576 nuclei that are not quail+ (white arrows). Yellow arrowheads point to MYOG+ that are
577 quail+. **(h-k)** Limb muscle organisation after quail-into-chicken lateral-plate mesoderm grafts.
578 **(h)** Schematic of limb lateral plate mesoderm grafts. **(i,j)** Transverse and adjacent limb
579 sections were hybridised with SCX probe (blue) to label tendons (i) and immunostained with
580 QCPN (quail antibody), MF20 (myosins) combined with DAPI, focused on the FCU (j). **(k)**
581 High magnification of muscle tip regions close to tendons showing the high density of quail
582 nuclei of lateral plate mesoderm origin. Arrows point to quail myonuclei in myosin+ cells.

583

584 **Figure 2**

585 **A subpopulation of muscle cells are not derived from the *Pax3* lineage but are derived**
586 **from the fibroblast lineage in mouse limbs**

587 **(a)** Transverse forelimb sections of E15.5 *Pax3^{Cre}:R26^{stop/Tom}* mice immunostained with
588 Tomato antibody and DAPI staining. **(b)** Focus on a muscle immunostained with Tomato and
589 PAX7 antibodies. **(c)** High magnifications showing PAX7+ cells that are tomato-negative i.e.
590 not of *Pax3* lineage (arrows). **(d)** Transverse forelimb sections of E12.5
591 *Scx^{Cre}:R26^{stop/Tom}:GFP(Scx)* mice immunostained with Tomato (red) and MYOD-MYOG
592 (green) antibodies, and stained with DAPI (blue). Combinations of two colours are indicated
593 on panels. Arrows point to MYOD-MYOG+ cells that are Tomato+ (*Scx* lineage). **(e)** High
594 magnifications showing MYOD-MYOG+ cells that are Tomato+ (*Scx* lineage). **(f,h)**
595 Transverse and adjacent forelimb sections of E15.5 *Osr1^{CreERT2}:Pax7^{GPL}* mice immunostained

596 with betaGal (*Osr1*-lineage-derived cells, green) and PAX7 (red) antibodies, combined with
597 DAPI (blue) (f), and with betaGal (*Osr1*-lineage-derived cells, green) and MYOD/MYOG
598 (red) antibodies combined with DAPI (h). (g,i) High magnifications. (g) Arrows point to the
599 betaGal⁺ nuclei (*Osr1*-lineage-derived, green) and PAX7⁺ (red) that are located in between
600 myotubes, while arrowheads point to the betaGal⁺ nuclei (green) that are PAX7⁻ and located
601 inside myotubes. These betaGal⁺ PAX7⁻ nuclei correspond to cells in which PAX7 has been
602 downregulated and betaGal expression kept after cell incorporation into myotubes. (i) Arrows
603 point to the betaGal⁺nuclei (*Osr1*-lineage-derived, green) that are MYOD/MYOG⁺ (red).

604

605 **Figure 3**

606 **BMP signalling is active in cells displaying a CT/M transcriptional identity and in cells** 607 **at the muscle/tendon interface in chicken limbs**

608 (a) UMAP plots showing the whole-limb clustered populations with their respective number
609 of cells at E4 (4820 cells), E6 (3340 cells) and E10 (22912 cells). CT and muscle clusters
610 represent the majority of limb cells at each developmental stage. (b) Heatmaps showing the
611 relative expression of the top 10 markers for each CT and muscle clusters in cells grouped by
612 their identity CT (red), CT/M (green) or M (blue) at E4, E6 and E10. Upregulated genes in
613 yellow, downregulated genes in purple. CT/M identity labels 26 cells (0.5% of limb cells) at
614 E4, 133 cells (4.0% of limb cells) at E6 and 729 cells (3.3% of limb cells) at E10. (c) Violin
615 plots showing Log-normalized expression levels of *ID1*, *ID2* and *ID3* genes, which are
616 transcriptional readouts of the BMP activity, in cells grouped by their identity CT (red), CT/M
617 (green) or M (blue) at E4, E6 and E10. (d) Adjacent and transverse limb sections of E6
618 chicken embryos were hybridized with *ID2* or *SCX* probes and immunostained with
619 pSMAD1/5/9 and MF20 (myosins) antibodies. High magnifications of ventral limb muscle
620 masses. (e) Transverse limb sections of E9 chicken embryos were hybridized with *ID2* probe
621 (blue) and then immunostained with MF20 (myosins) antibody (brown). Focus on dorsal limb
622 muscles. (f) Transverse limb sections of E9 chicken embryos were immunostained with
623 pSMAD1/5/9 (green) and MF20 (myosins, grey) antibodies. Focus on dorsal limb muscles.

624

625

626 **Figure 4**

627 **BMP signalling regulates a fibroblast-myoblast conversion in vivo and in vitro**

628 (a-g) BMP4/RCAS-producing cells were grafted to right limbs of E5 chicken embryos.
629 Embryos were fixed 5 days later at E10. *In situ* hybridisation experiments with mBmp4

630 probes to transverse limb sections indicate the extent of ectopic BMP4 in dorsal limb muscles
631 (a). (b,e) Limb sections were immunostained with MF20 (myosins, green) to visualise the
632 muscle pattern. Stars indicate fused muscles in BMP4-treated limbs (b) versus individual
633 muscles in control limbs (e). (c,f) High magnifications of dorsal limb muscles immunostained
634 with MF20 (myosins, green), PAX7 (cyan) and TCF4 (red) antibodies in BMP4 (c) and
635 respective control limbs (f). (d) Percentages of PAX7+ and TCF4+ nuclei per total number of
636 nuclei in BMP4-treated limbs versus controls. (g) Representative fields of DAPI+ nuclei in
637 muscles of control and BMP4 limb and quantification of the number of nuclei per unit area.
638 (h) Chicken fibroblast cultures treated with BMP4/RCAS or BMPR1ca/RCAS.
639 Immunostaining with PAX7 antibody (red) with DAPI in BMP4- and BMPR1ca-treated
640 fibroblast cultures compared to controls. (i) Quantification of PAX7+ cells versus the total
641 cell number. (j-n) NOGGIN/RCAS-producing cells were grafted to right limbs of E5 chicken
642 embryos. Embryos were fixed 5 days later at E10. (j) *In situ* hybridisation experiments with
643 cNOGGIN probes to transverse limb sections indicate the extent of ectopic NOGGIN in
644 dorsal limb muscles. (k) Limb sections were immunostained with MF20 (myosins, green) to
645 visualise muscles. (m) High magnifications of dorsal limb muscles immunostained with
646 PAX7 (cyan) and TCF4 (red) antibodies in NOGGIN and respective control limbs. (l)
647 Percentages of PAX7+ and TCF4+ nuclei per total number of nuclei in NOGGIN-treated
648 limbs versus controls. (n) Representative fields of DAPI+nuclei in muscles of control and
649 NOGGIN limbs; and quantification of the number of nuclei per unit area. (o-q) Myoblast
650 cultures treated with SMAD6/RCAS were induced to differentiation. (o) Immunostaining
651 with PAX7, MYOG, and MF20 (myosins) antibodies in control and SMAD6 cultures. (p,q)
652 RT-qPCR analyses of the expression levels of muscle genes, *PAX7*, *MYF5*, *MYOD*, *MYOG*,
653 *MYHC*, BMP-responsive genes *ID1*, *ID2*, *ID3* (p) and connective tissue genes, *SCX*, *COL3A1*,
654 *COL6A1*, *OSR1*, *TCF4*, *PDGFRA* (q) in SMAD6-treated and control cell cultures. Gene
655 mRNA levels were normalized to *GAPDH* and *RPS17*. The relative mRNA levels were
656 calculated using the $2^{-\Delta\Delta C_t}$ method using the control condition as controls. For each gene, the
657 mRNA levels of the control condition were normalized to 1. Graph shows means \pm standard
658 deviations of 9 biological samples (p) and 6 biological samples (q).

659

660

661 **EXTENDED DATA**

662

663 **Extended Data Fig. 1**

664 **A subpopulation of myonuclei in skeletal muscles are not derived from presomitic**
665 **mesoderm in chicken limbs**

666 **(a,b)** Forelimbs of presomitic grafts of E9 quail-into-chicken embryos were co-
667 immunostained with QCPN (quail nuclei), MF20 (myosins) and ColXII (tendons) antibodies.
668 Focus on (a) the EMR (anterior muscle) and (b) a longitudinal view of muscle. Arrows point
669 to zones of nuclei (DAPI) that are not of quail origin at the muscle tips close to tendons. **(c,d)**
670 Focus on muscle tips from forelimb sections of presomitic grafts fixed at E8, immunostained
671 with QCPN (quail nuclei), MF20 (myosins) and ColXII (tendons) antibodies. **(e-h)** Presomitic
672 graft (quail-into-chicken embryos) fixed at E9. **(e)** Transverse forelimb sections
673 immunostained with QCPN (quail nuclei). **(f)** Focus on the EMU. **(g,h)** Focus on the FCU. **(h)**
674 Zoom to muscle tip regions with high magnifications showing quail-negative myonuclei (*i.e.*
675 not presomitic-derived).

676

677 **Extended Data Fig. 2**

678 **A subpopulation of PAX7+ cells are not derived from presomitic mesoderm in chicken**
679 **limbs**

680 **(a)** Diagram of presomitic grafts of GFP+chicken into chicken embryos (left panel). GFP
681 fluorescence in forelimb of a presomitic graft E8 chimera (middle panel). Transverse limb
682 section showing cytoplasmic GFP expressed in presomitic-derived cells in muscles (right
683 panel). **(b)** Focus on FCU muscle labelled with GFP (green), MF20 (myosins, red) and
684 ColXII (tendons, red) antibodies. Arrows point to tendons. **(c)** FCU muscle labelled with
685 PAX7 (red) and GFP (green). High magnifications of FCU muscle tips show PAX7+ cells
686 (arrows) that are GFP negative (*i.e.* not presomitic-derived).

687

688 **Extended Data Fig. 3**

689 **Quail lateral plate mesoderm grafts into chicken embryos lead to expected derivatives in**
690 **limbs**

691 **(a)** Limb transverse sections of lateral plate mesoderm grafts quail-into-chicken embryos
692 fixed at E9 immunostained with QCPN (quail nuclei, red) and MF20 (myosins, green)
693 antibodies, combined with DAPI. The quail+ nuclei (red), which are lateral plate-derived, are
694 observed in cartilage **(b)**, tendon **(d,e)**, visualised with *SCX* expression by in situ
695 hybridization on adjacent section **(c)** and in irregular CT in between myosin+ cells inside
696 muscle **(f)**. **(e)** is a zoom of the tendon regions of the FCU muscle **(d)**.

697

698 **Extended Data Fig. 4**

699 **A subpopulation of PAX7+ progenitors and MYOD/MYOG+ myogenic cells are derived**
700 **from Scx-lineage in mouse limbs**

701 (a) Transverse forelimb sections of E14.5 $Pax3^{Cre}:R26^{Tom}$ mice immunostained with Tomato
702 (red) and PAX7 (green) antibodies, combined with DAPI staining. Arrows point to the
703 PAX7+ cells (green) that are Tomato-negative (*i.e.* not Pax3 lineage-derived). (b) Forelimb
704 sections of E12.5 $Scx^{Cre}:R26^{Stop/Tom}:GFP(Scx)$ mice immunostained with Tomato (Scx-
705 lineage-derived cells, red) and PAX7 (green) antibodies. First panel shows GFP reflecting
706 SCX expression in tendon primordia. Arrows point to Tomato+ cells (Scx-derived) that are
707 PAX7+. (c) Adjacent sections of (b) immunostained with Tomato (Scx-lineage-derived cells,
708 red) and MYOD/MYOG (green) antibodies. Arrows point to Tomato+ cells (Scx-derived)
709 that are MYOD/MYOG+. First panel shows MYOD/MYOG+ (red) in between GFP+ cells
710 (green). (d) Forelimb sections of E14.5 $Scx^{Cre}:R26^{Tom}:GFP(Scx)$ mice were immunostained
711 with Tomato (Scx-lineage-derived cells, red) and MYOD/MYOG (green) antibodies. Focus
712 on ventral forelimb regions. First panel shows GFP reflecting SCX expression in tendons.
713 Arrows point to Tomato+ cells (Scx-derived) that are MYOD/MYOG+ closed to tendons. (e)
714 Transverse forelimb sections of E16.5 $Scx^{Cre}:R26^{Tom}$ mice immunostained with Tomato (red)
715 and My32 (myosins, green) antibodies. High magnifications of a Tomato+ junctional region
716 between two dorsal limb muscles (encircled with dashed lines), showing the overlap between
717 myosins and Tomato labelling (Scx-lineage-derived cells).

718

719 **Extended Data Fig. 5**

720 **A fraction of muscle progenitors, myogenic cells and myonuclei are derived from Osr1-**
721 **lineage in mouse limbs**

722 (a) Transverse forelimb sections of E15.5 $Osr1^{CreERT2}:Pax7^{GPL}$ mice immunostained with
723 betaGal (Osr1-lineage-derived cells) and PAX7 antibodies. (b) High magnifications of ventral
724 limb muscles. Arrows point to the betaGal+nuclei (green) that are PAX7+ (red), while
725 arrowheads point to the betaGal+nuclei (green) that are not PAX7+. (c-d) Transverse
726 forelimb sections of E15.5 $Osr1^{CreERT2}:Pax7^{GPL}$ mice immunostained with betaGal (Osr1-
727 lineage-derived cells) and MYOD/MYOG antibodies. (d) High magnifications of ventral limb
728 muscles. Arrows point to the betaGal+nuclei (green) that are MYOD/MYOG+ (red), while
729 arrowheads point to the betaGal+nuclei (green) that are not MYOD/MYOG+. (e-f)
730 Transverse forelimb sections of E15.5 $Osr1^{CreERT2}:Fucci2aR$ mice immunostained with Fucci

731 (Osr1-lineage-derived cells) and My32 (myosins) antibodies. (f) Arrows point to Fucci+nuclei
732 (red) in myosin+ cells (green).

733

734 **Extended Data Fig. 6**

735 **Neural crest cells do not contribute to myogenic lineage at muscle tips of limb skeletal** 736 **muscles**

737 (a) Diagram of isotopic neural tube grafts from GFP chicken to chicken embryos. (b) GFP
738 fluorescence in E8 grafted-forelimbs (6 days after grafting) showing the neural tube cell
739 derivatives. (c) Transverse forelimb sections from neural tube grafts immunostained with GFP
740 (neural tube-derived cells) and MF20 (myosins) antibodies. (d,d') Focus on dorsal limb
741 muscles immunostained with GFP (green) and MF20 (myosins) antibodies, combined with
742 DAPI. White stars indicate tendons. (e,e') Transverse forelimb sections of E14.5
743 *Wnt1^{Cre}:R26^{Tom}* mice immunostained with Tomato (Wnt1-lineage-derived cells) and My32
744 (myosins) antibodies, combined with DAPI. (f,f') Transverse forelimb sections of E14.5
745 *Wnt1^{Cre}:R26^{Tom}* mice immunostained with Tomato (Wnt1-lineage-derived cells) and PAX7
746 antibodies, combined with DAPI. (g,g'), (h,h') are high magnifications of ventral limb
747 regions squared in (e') and (f'), respectively. Arrows point to muscle cells (Myosin+ or
748 PAX7+ cells, green) at muscle tips displaying no Tomato labelling, (Wnt1-lineage-derived
749 cells, red).

750

751 **Extended data Fig. 7**

752 **Characterisation of the cells with a CT/M identity**

753 (a) Feature plots showing the distribution of cells with a CT/M identity (in blue) across
754 whole-limb clustered populations (in grey) at E4, E6 and E10. At E4, 50% of the CT/M cells
755 are found in the CT clusters and 50% in muscle clusters. At E6, 17% and 83% of the CT/M
756 cells are found in CT and muscle clusters, respectively. At E10, 37% and 61% of the CT/M
757 cells are found in CT and muscle clusters, respectively. (b) Heatmaps showing the relative
758 expression of five recognized CT markers (PRRX1, TWIST2, PDFGRA, OSR1, SCX) and
759 four muscle markers (PAX7, MYF5, MYOD, MYOG) in cells grouped by their identity CT
760 (red), CT/M (green) or M (blue) at E4, E6 and E10. Upregulated genes in yellow,
761 downregulated genes in purple. High magnification of the heatmaps showing the relative
762 expression of the same nine CT and muscle markers but only in CT/M cells at E4, E6 and E10.
763 (c) Violin plots showing Log-normalized expression levels of selected CT and muscle
764 markers in cells grouped by their identity CT (red), CT/M (green) or M (blue) at E4, E6 and

765 E10. The expression of PRRX1, and to a lesser extent TWIST2, is progressively found in an
766 increasing number of CT/M cells at the expense of CT cells. Conversely, SCX expression is
767 found in a decreasing number of CT/M cells, like PAX7 and MYOD.

768

769 **Extended Data Fig. 8**

770 **BMP signalling labels a junctional cell population at different stages of development**

771 **(a,b)** Transverse adjacent limb sections of E6 chicken embryos were hybridized with BMP4
772 probe (a) and immunostained with pSMAD1/5/9 antibody (b). Arrows point to BMP4 and
773 pSMAD1/5/9 expression spots. **(c,d)** High magnification of dorsal (c) and ventral (d) muscle
774 masses immunostained with pSMAD1/5/9 and myosin antibodies and counterstained with
775 DAPI. Arrows point to pSMAD1/5/9 expression spots. **(e)** Longitudinal view of a muscle
776 labelled with pSMAD1/5/9 and MF20 (myosins) antibodies. Arrows point to pSMAD1/5/9
777 enriched expression (red) at muscle tips (green). **(f)** Adjacent and longitudinal limb sections
778 of E7 chicken embryos were hybridized with BMP4 and ID2 probes (blue staining). The
779 arrows point to *ID2* expression at muscle tips close to tendons labelled with *BMP4* and *ID2*
780 transcript. **(g)** Longitudinal muscle sections of E9 chicken hybridized with ID2 probe (blue)
781 and then immunostained with MF20 antibody (myosins, brown). The arrow points to *ID2*
782 expression at muscle tips. **(h,i)** Adjacent and transverse limb sections of E8 chicken embryos
783 were co-immunostained with pSMAD1/5/9 and MF20 (myosins) antibodies (h) and COLXII
784 (tendons) and MF20 (myosins) antibodies (i). Arrows point to restricted location of
785 pSMAD1/5/9 in FCU muscle

786

787 **Extended Data Fig. 9**

788 **pSMAD1/5/9+ myonuclei label two myonucleus populations of distinct embryological** 789 **origins at muscle tips**

790 **(a)** Schematic of isotopic pre-somitic mesoderm grafts from quail to chicken embryos. **(b)**
791 Longitudinal view of a limb muscle from E6 chimera labelled with pSMAD1/5/9, QCPN
792 (quail nuclei) and MF20 (myosins) antibodies, with a focus to muscle tips. **(c)** Adjacent
793 section labelled with COLXII (tendon), QCPN (quail nuclei) and MF20 (myosins) antibodies.
794 **(d)** Zoom to muscle tips, labelled with pSMAD1/5/9, QCPN (quail nuclei) and MF20
795 (myosins) antibodies, combined with DAPI. White arrows point to pSMAD1/5/9+ myonuclei
796 (green) that are not of quail origin (QCPN-), while yellow arrows point to pSMAD1/5/9+
797 myonuclei (green) that are of quail origin (QCPN+). **(e)** Transverse view of the FCU muscle
798 from E9 chimera labelled with pSMAD1/5/9, QCPN (quail nuclei) and MF20 (myosins)

799 antibodies. **(f)** Adjacent sections to **(e)** labelled with COLXII (to label tendon), QCPN (quail
800 nuclei) and MF20 (myosins) antibodies, combined with DAPI. **(g)** Focus on muscle tips.
801 White arrows point to pSMAD1/5/9+ myonuclei (green) that are not of quail origin (QCPN-),
802 while yellow arrows point to pSMAD1/5/9+ myonuclei (green) that are of quail origin
803 (QCPN+).

804

805 **Extended Data Fig. 10**

806 **BMP signalling and fibroblast/myoblast conversion**

807 **(a)** Schematics of the grafting procedure of BMP4/RCAS-producing cells. **(b,c)** In situ
808 hybridisation with BMP4 probe to transverse limb sections indicate the extent of ectopic
809 BMP4, in E10 grafted-embryos, 5 days after grafting (BMP4 right **(b)** and control left **(c)**
810 limbs). **(d,e)** Immunohistology with PAX7 (green) and MF20 (myosins, red) antibodies to
811 BMP4 right **(d)** and control left **(e)** limb sections. **(f,g)** Following BMP4 exposure, the
812 proportion of EDU+/PAX7+ cells versus PAX7+ cells was not significantly changed
813 compared to control muscles. **(h-j)** BMP4/RCAS exposure to chicken primary myoblast
814 cultures in proliferation conditions did not increase the number of PAX7+ cells compared to
815 control (empty/RCAS) myoblast cultures **(h,i)** and did not affect cell proliferation assed with
816 the Fucci system **(j)**. **(k)** NOGGIN/RCAS and empty/RCAS (control) exposure to chicken
817 primary myoblast cultures in differentiation conditions. Myoblasts were immunostained with
818 MF20 (myosins), MYOG and PAX7 antibodies (green) combined with DAPI.
819 NOGGIN/RCAS exposure in myoblast cultures led to a drastic disappearance of Myosin+
820 myotubes, MYOG+ cells and PAX7+ cells compared to control cultures, while DAPI nuclei
821 were still observed.

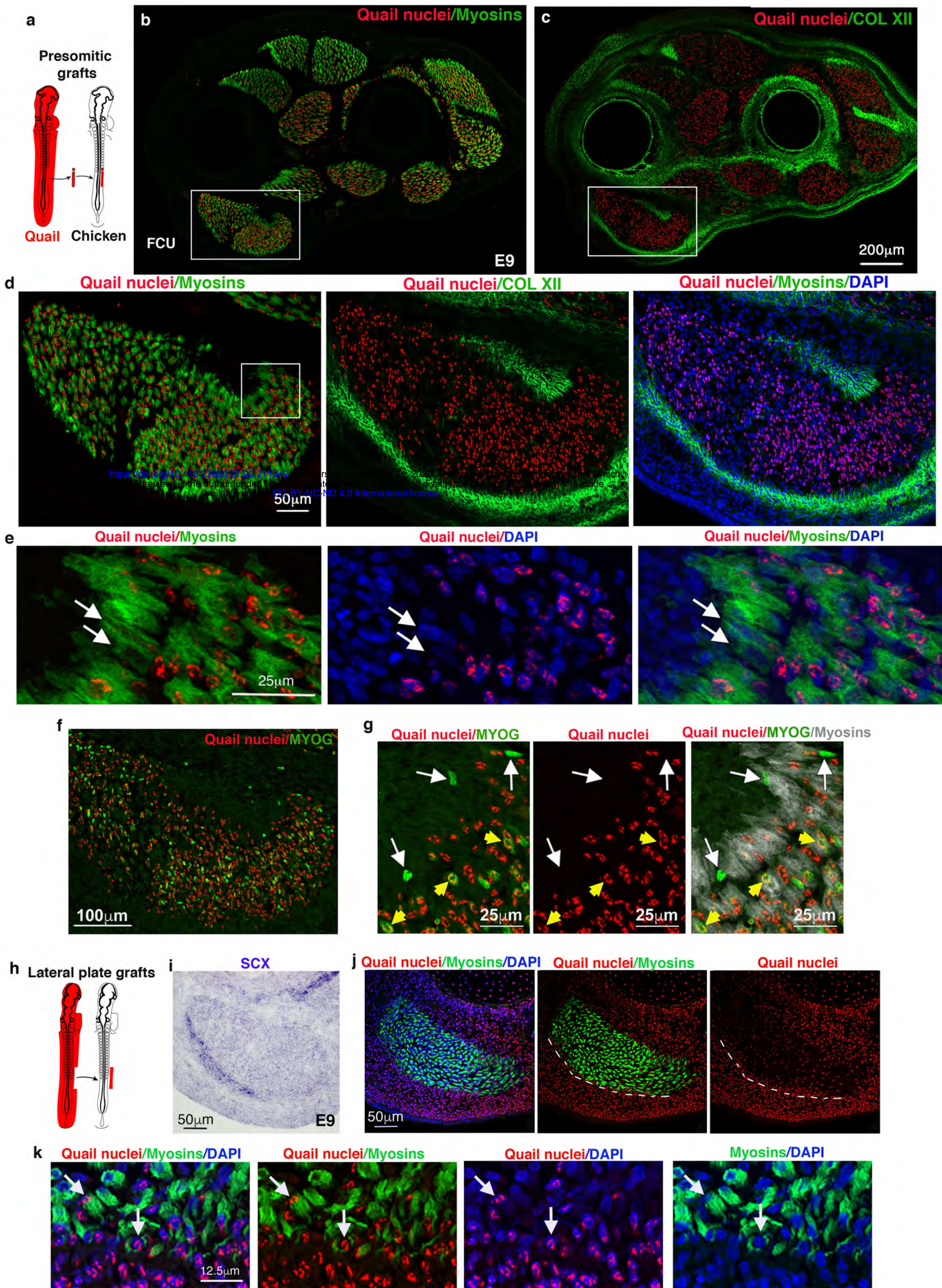


Figure 1

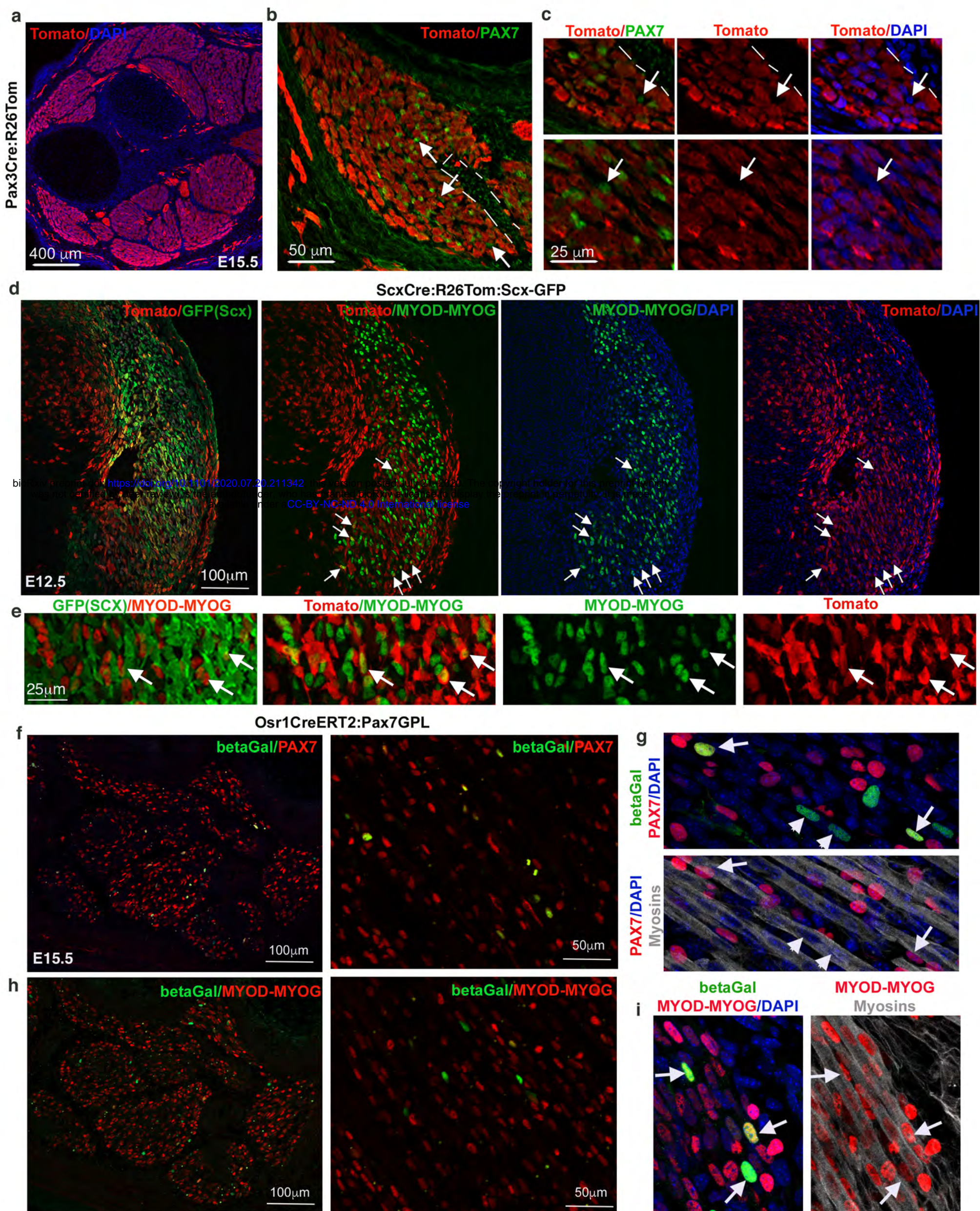


Figure 2

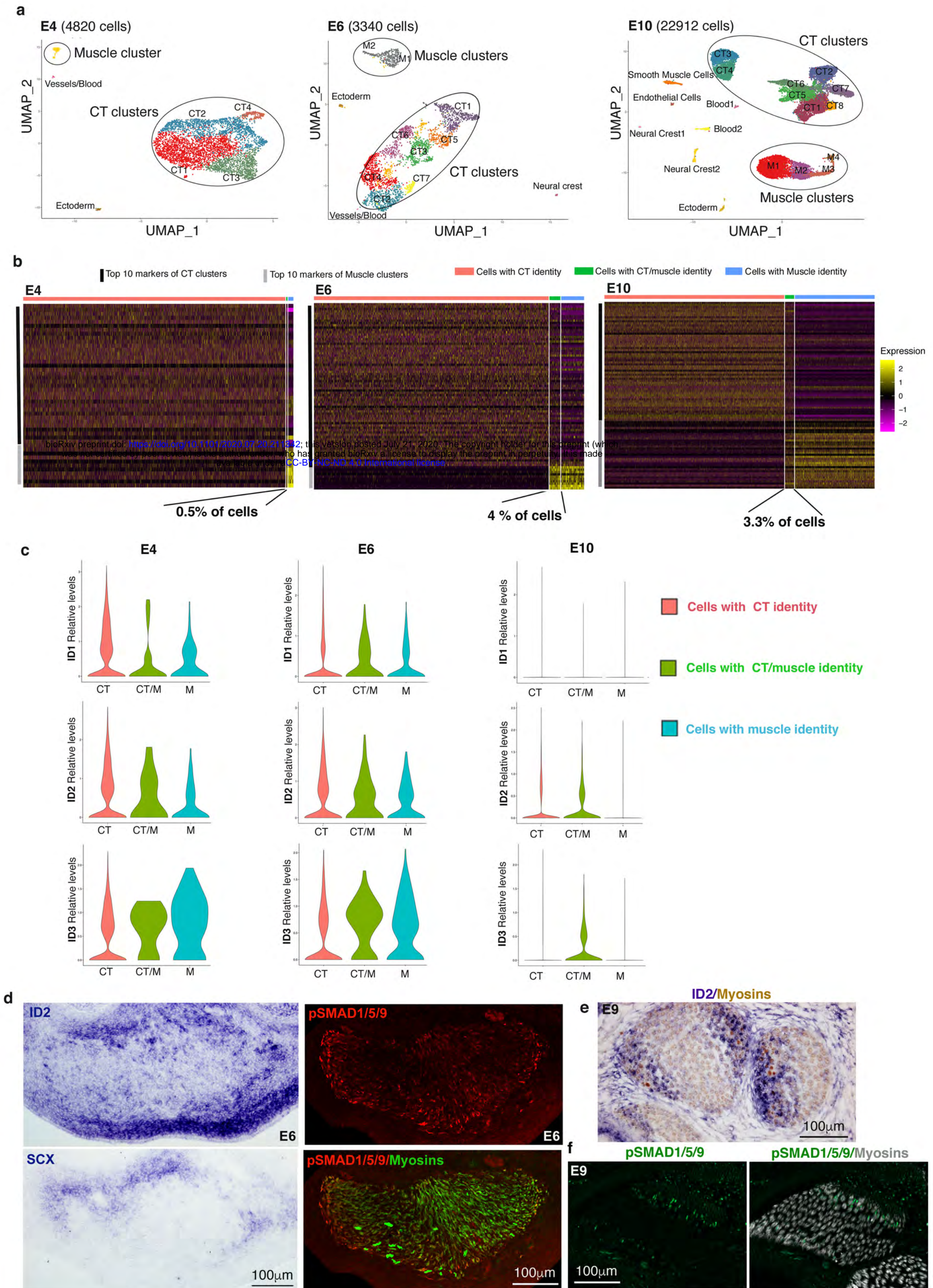


Figure 3

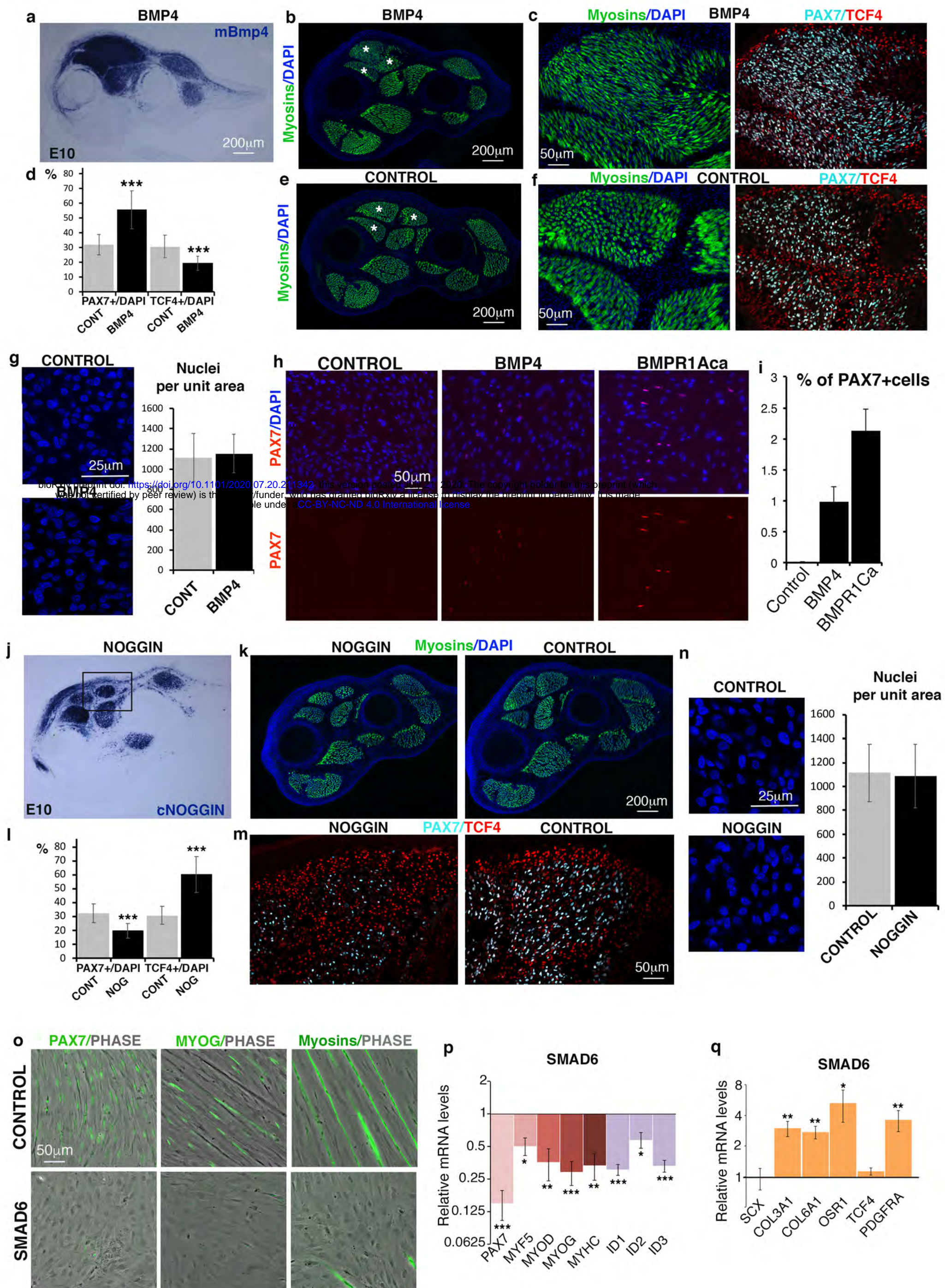
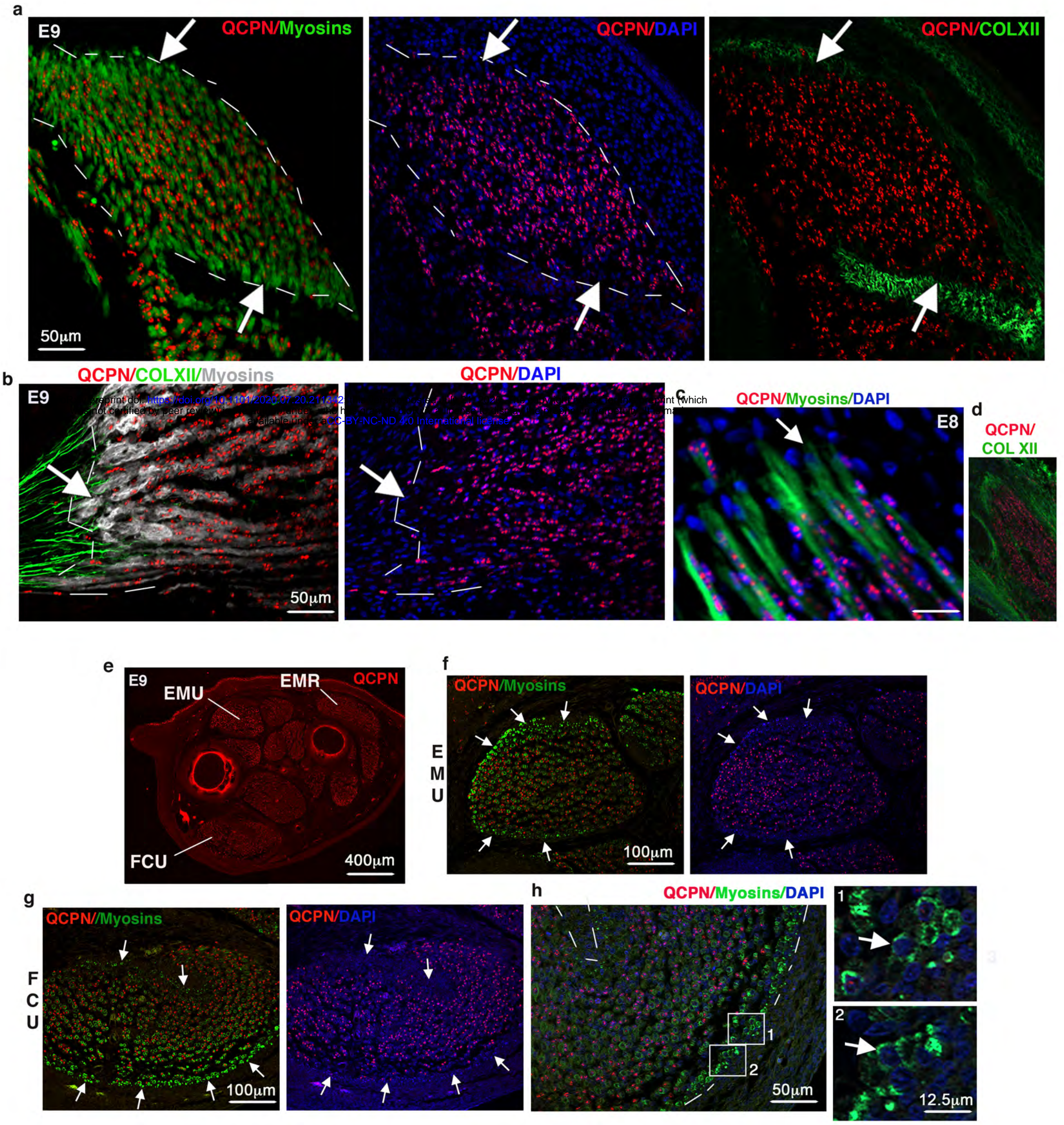


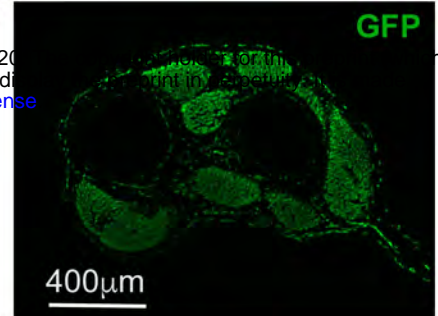
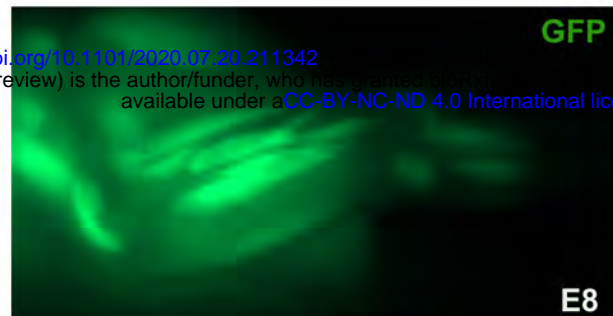
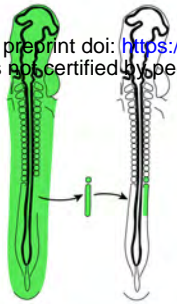
Figure 4



Extended Data Fig. 1

a

bioRxiv preprint doi: <https://doi.org/10.1101/2020.07.20.211342>; this version posted July 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

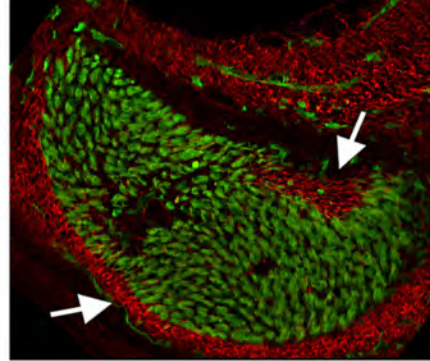
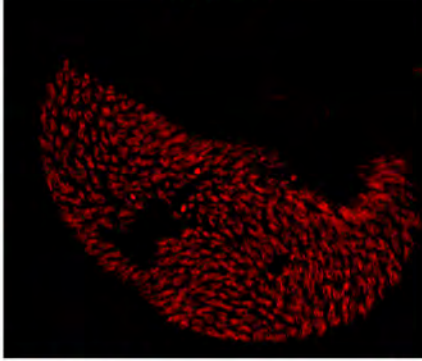
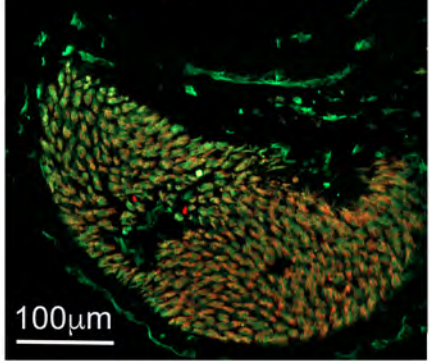


b

Myosins/GFP

Myosins

COLXII/GFP



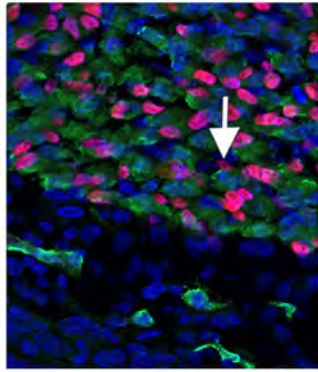
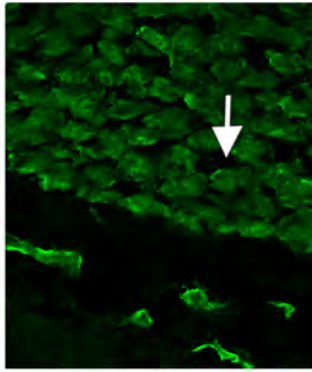
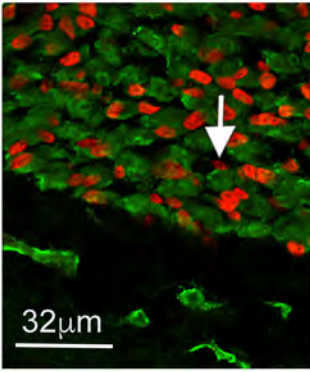
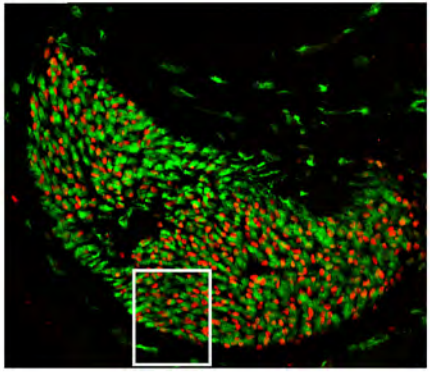
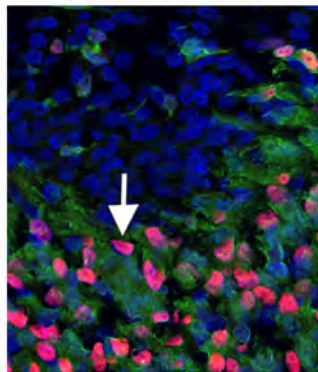
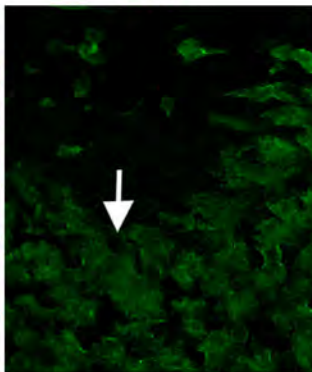
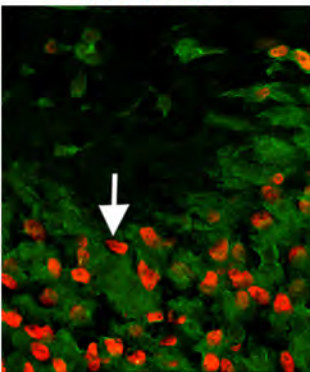
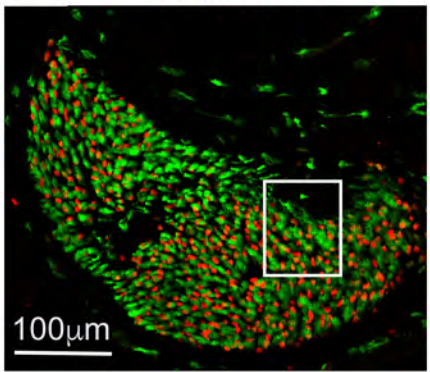
c

PAX7/GFP

PAX7/GFP

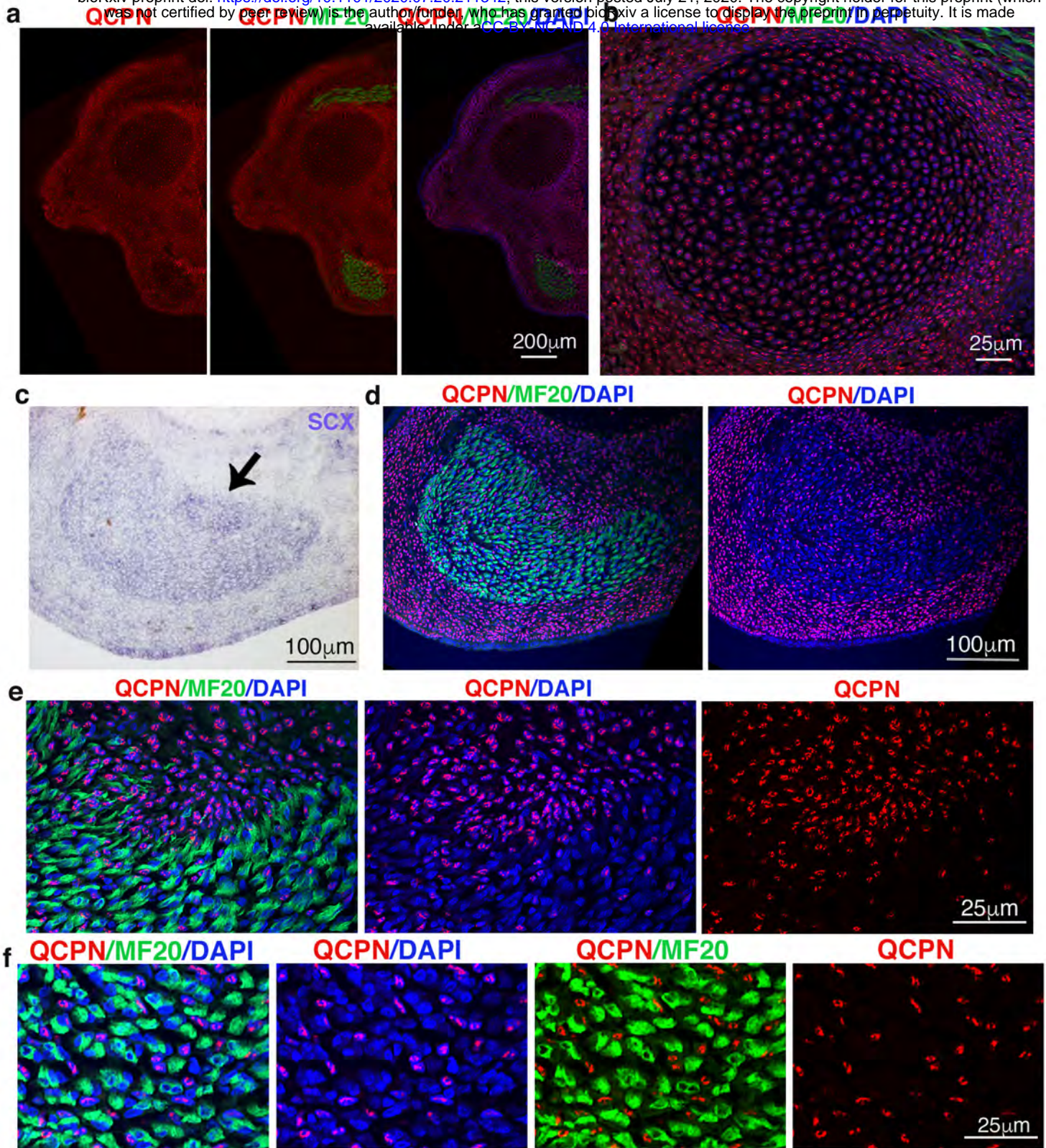
GFP

PAX7/GFP/DAPI

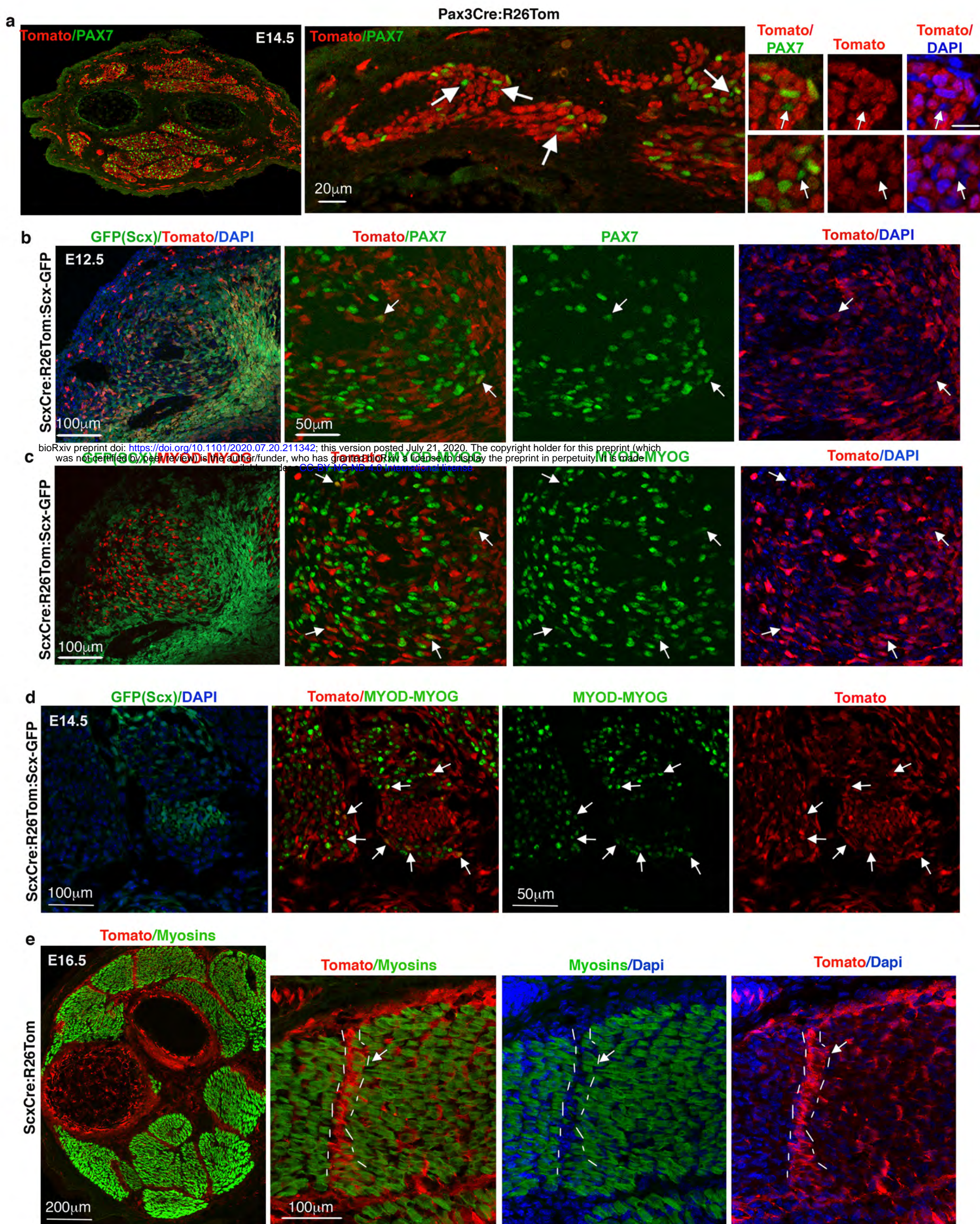


Extended Data Fig.2

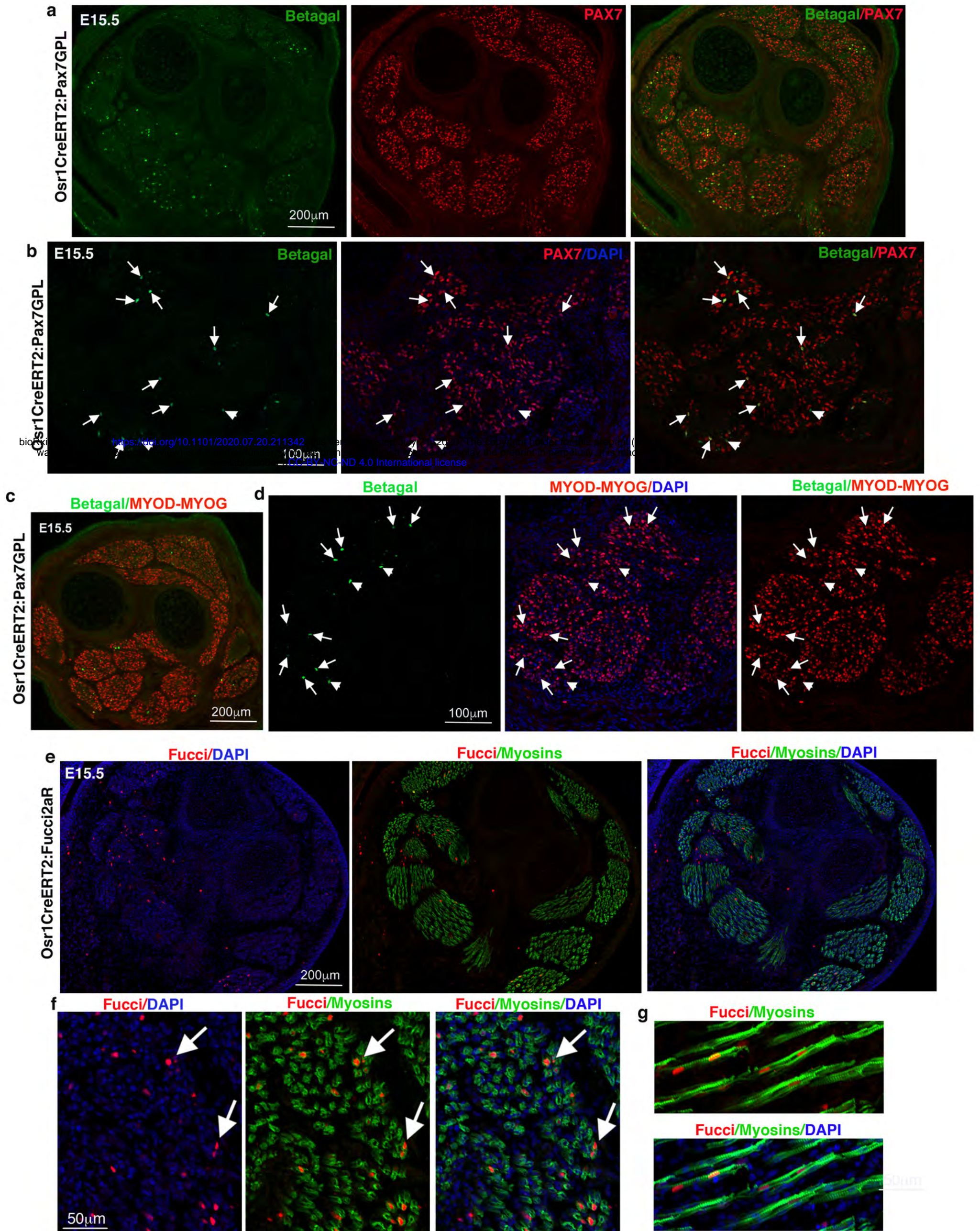




Extended Data Fig. 3

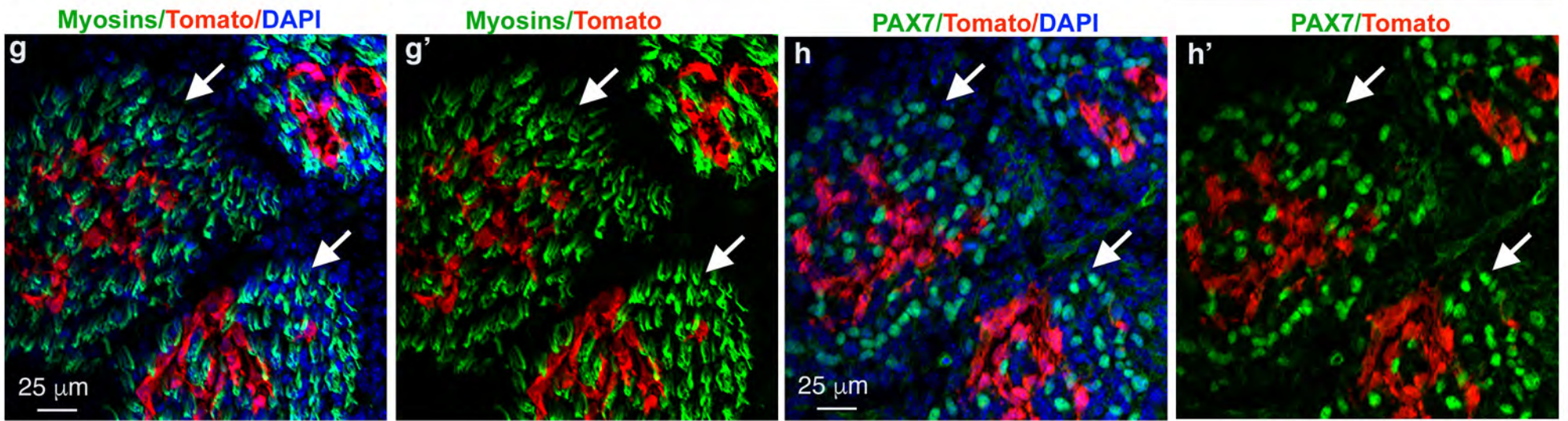
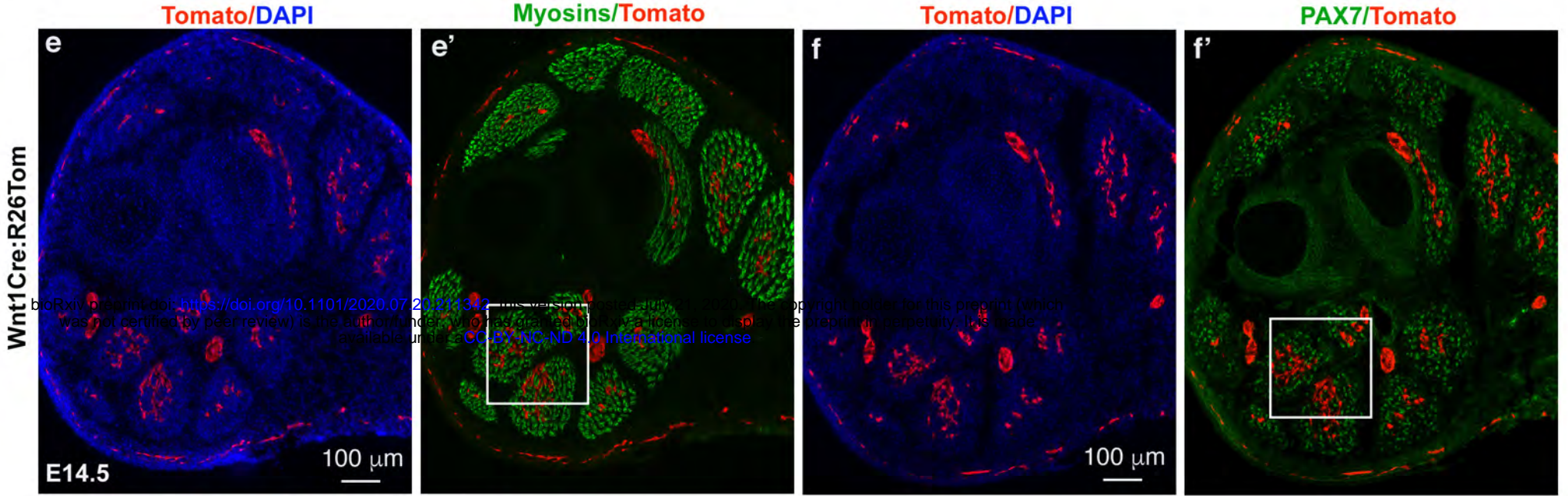
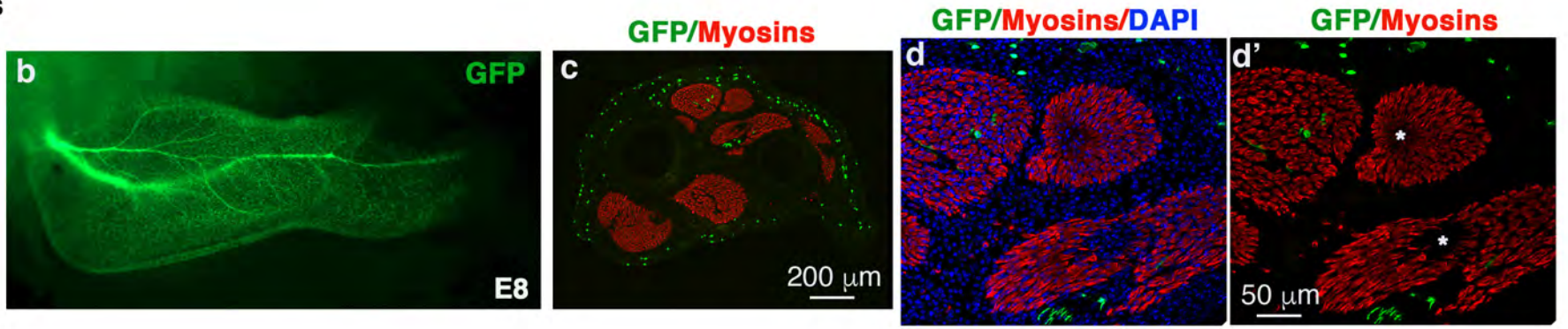
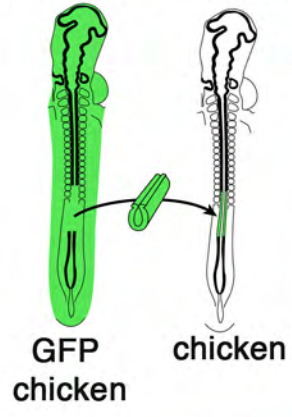


Extended Data Fig. 4

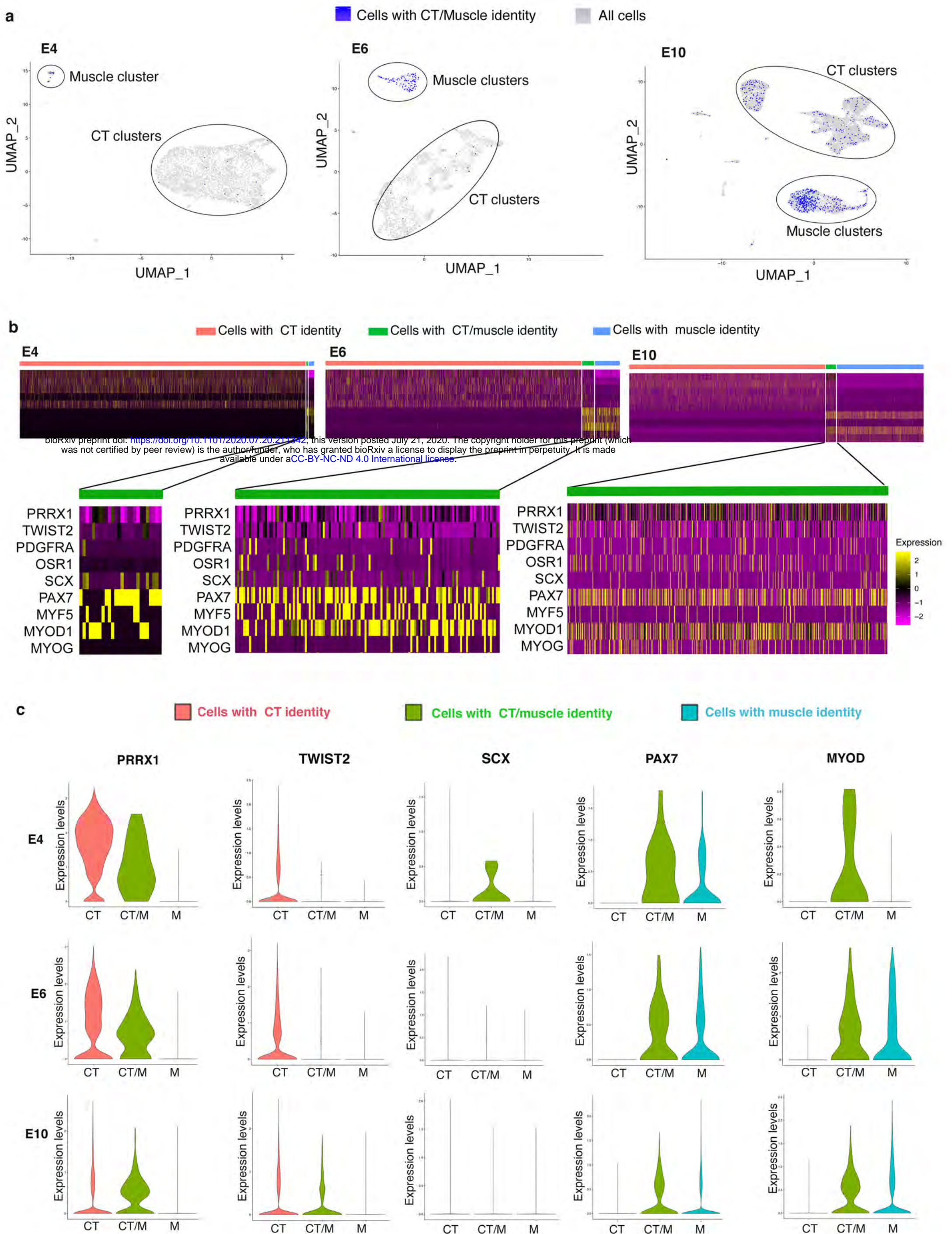


Extended Data Fig. 5

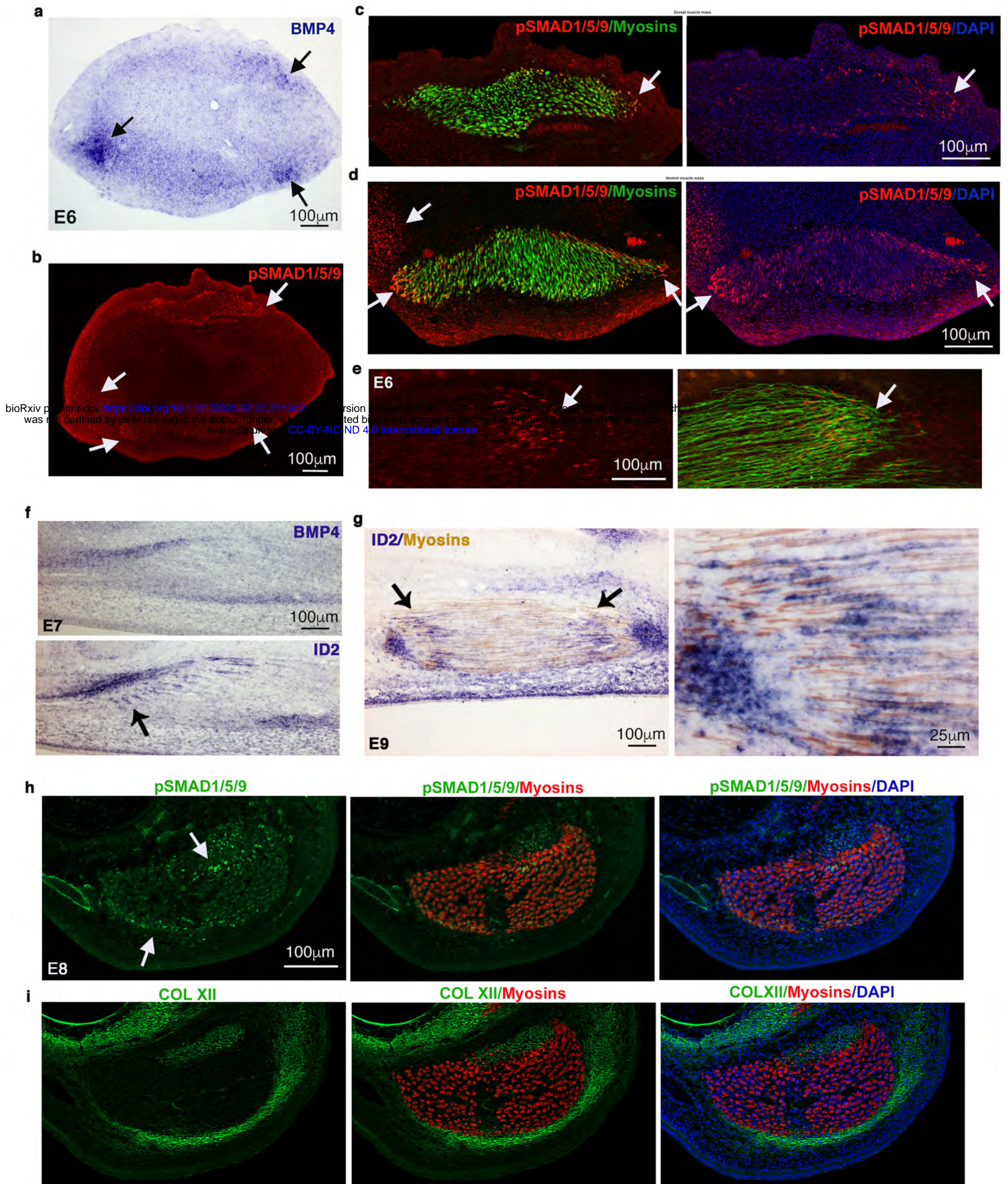
a Neural tube grafts



Extended Data Fig. 6

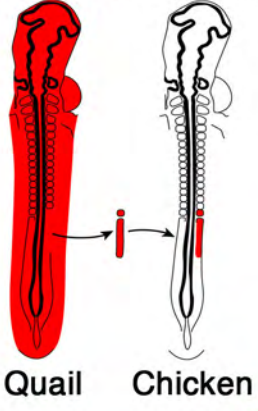


Extended Data Figure 7

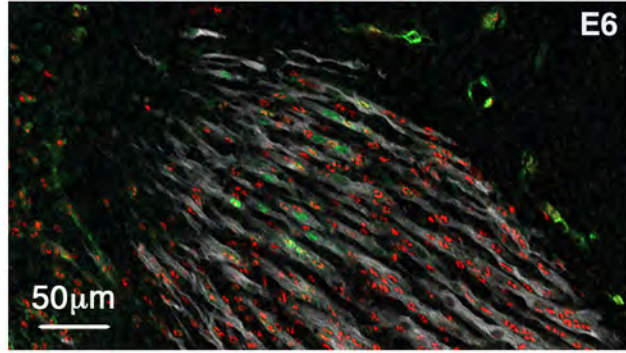


Extended Data Fig. 8

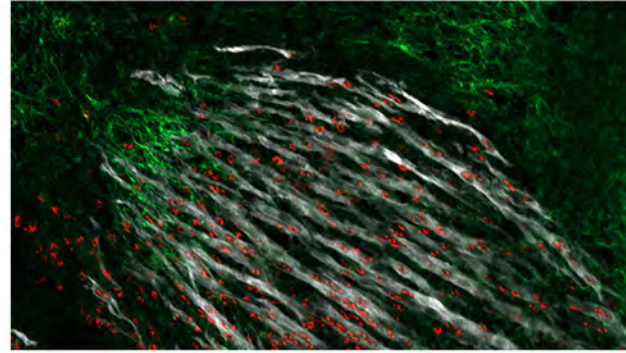
a Presomitic grafts



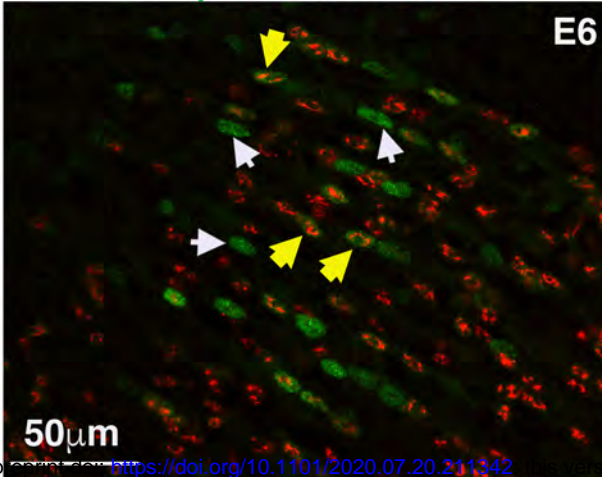
b pSMAD1/5/9/QCPN/Myosins



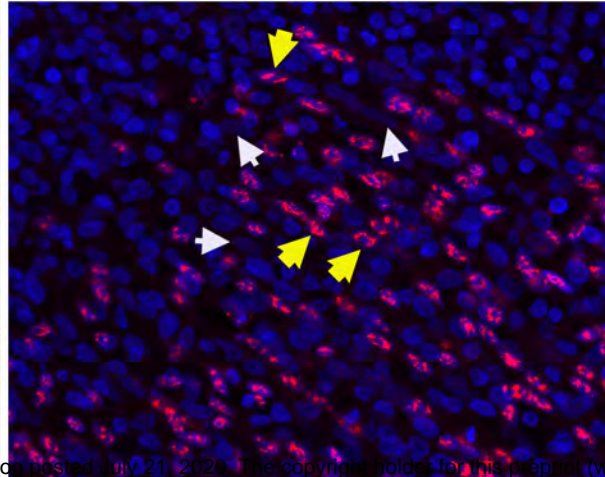
c COLXII/QCPN/Myosins



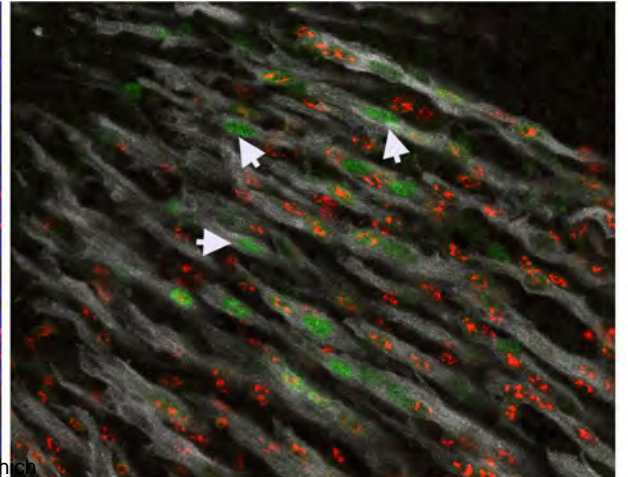
d pSMAD1/5/9/QCPN



QCPN/DAPI

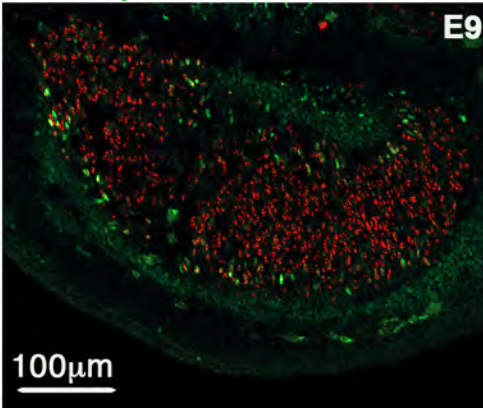


pSMAD1/5/9/QCPN/Myosins

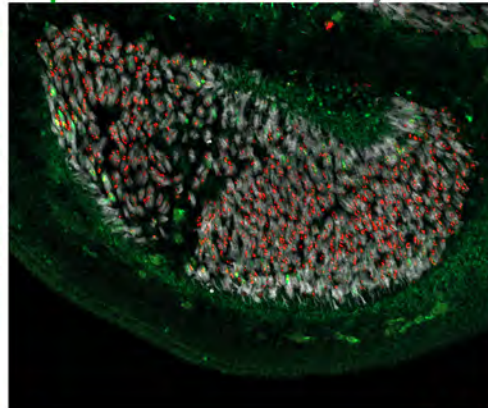


bioRxiv preprint doi: <https://doi.org/10.1101/2020.07.20.211342>; this version posted July 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

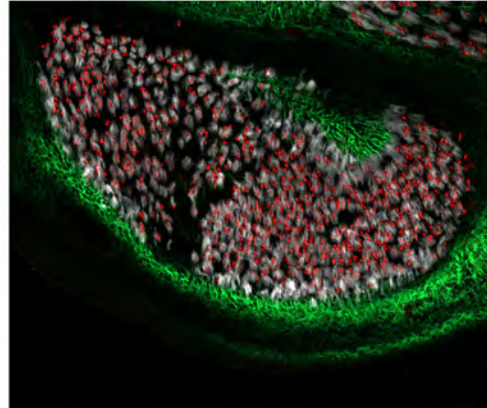
e pSMAD1/5/9/QCPN



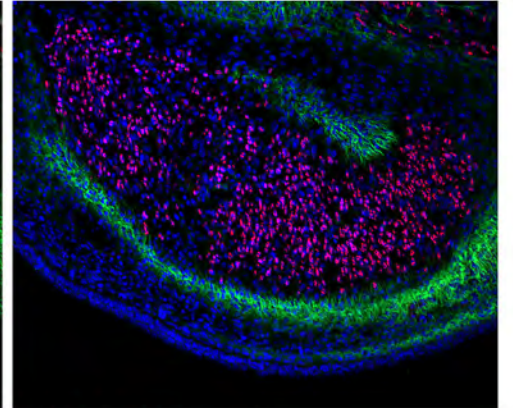
pSMAD1/5/9/QCPN/Myosins



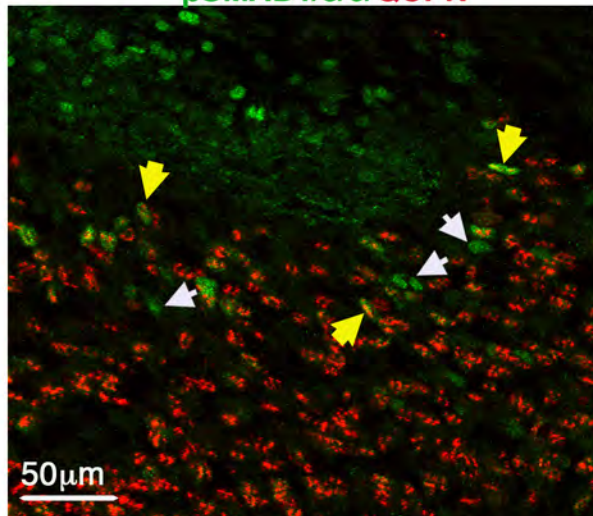
COL XII/QCPN/Myosins



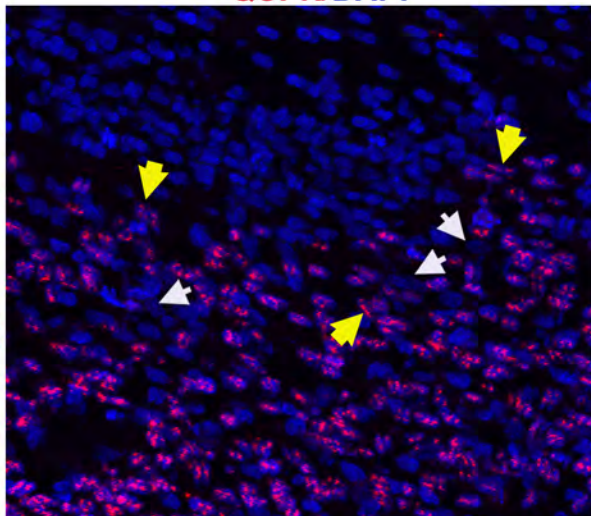
COL XII/QCPN/DAPI



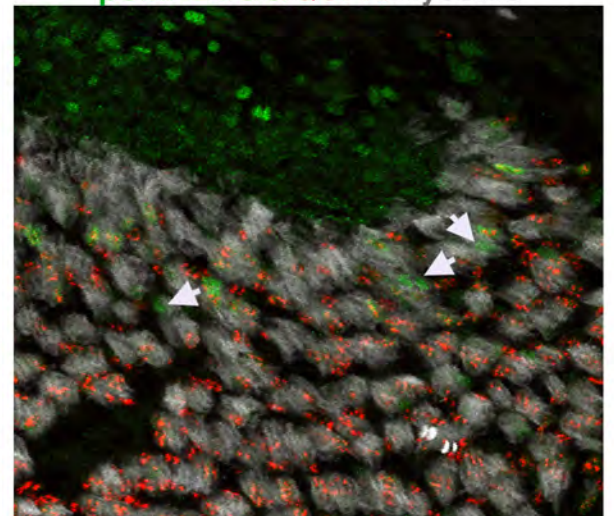
g pSMAD1/5/9/QCPN



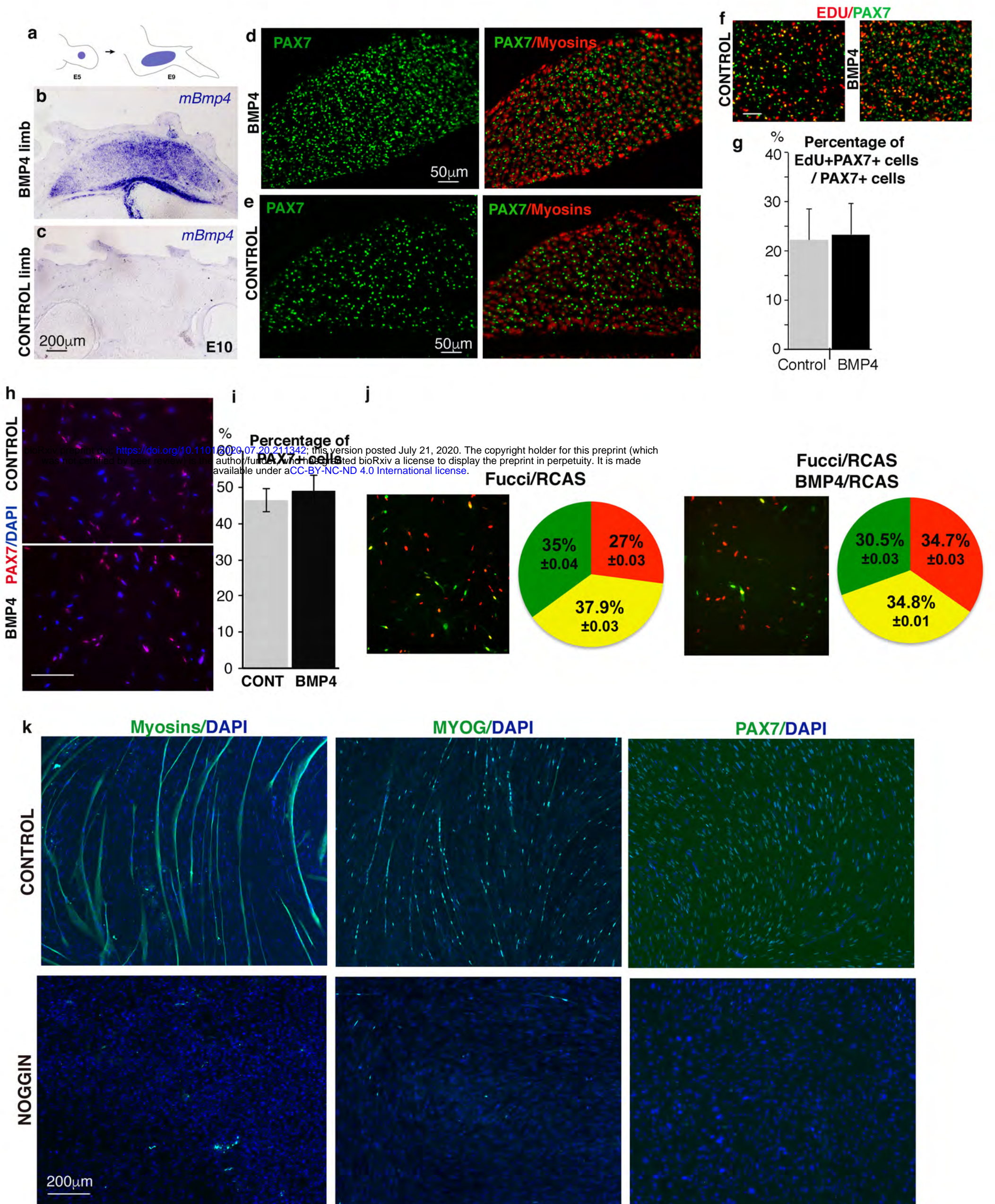
QCPN/DAPI



pSMAD1/5/9/QCPN/Myosins



Extended Data Fig. 9



Extended Data Fig. 10