



HAL
open science

A fast Myh super enhancer dictates adult muscle fiber phenotype through competitive interactions with the fast Myh genes

Matthieu Dos Santos, Stéphanie Backer, Frédéric Auradé, Matthew Wong, Maud Wurmser, Rémi Pierre, Francina Langa, Marcio Do Cruzeiro, Alain Schmitt, Jean-Paul Concordet, et al.

► To cite this version:

Matthieu Dos Santos, Stéphanie Backer, Frédéric Auradé, Matthew Wong, Maud Wurmser, et al.. A fast Myh super enhancer dictates adult muscle fiber phenotype through competitive interactions with the fast Myh genes. 2021. pasteur-03709535v1

HAL Id: pasteur-03709535

<https://hal.science/pasteur-03709535v1>

Preprint submitted on 9 Nov 2021 (v1), last revised 29 Jun 2022 (v2)

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

A fast *Myh* super enhancer dictates adult muscle fiber phenotype through competitive interactions with the fast *Myh* genes

Matthieu Dos Santos¹, Stéphanie Backer¹, Frédéric Auradé^{2,3}, Matthew Wong⁴, Maud Wurmser¹, Rémi Pierre¹, Francina Langa⁵, Marcio Do Cruzeiro¹, Alain Schmitt¹, Jean-Paul Concordet⁶, Athanassia Sotiropoulos¹, Jeffrey Dilworth⁴, Daan Noordermeer⁷, Frédéric Relaix², Iori Sakakibara^{1,8*}, Pascal Maire^{1*}

1. Université de Paris, Institut Cochin, INSERM, CNRS. 75014 Paris, France.
2. Univ Paris Est Creteil, INSERM, EnvA, EFS, AP-HP, IMRB, 94010 Creteil, France
3. Sorbonne Université, INSERM U974, Center for Research in Myology, 75013 Paris, France
4. Regenerative Medicine Program. Ottawa Hospital Research Institute, Canada
5. Institut Pasteur. Paris. France
6. MNHN. Paris. France
7. Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France
8. Institute of Medical Nutrition, Tokushima University Graduate School, Tokushima 770-8503, Japan

* Corresponding authors:

Pascal Maire

24 rue du Faubourg Saint-Jacques, F-75014 Paris, France

Telephone number: +33 1 44 41 24 13

Fax number: +33 1 44 41 24 21

Email address: pascal.maire@inserm.fr

Iori Sakakibara: iori.f.sakakibara@gmail.com

Summary

The contractile properties of adult myofibers are shaped by their Myosin heavy chain (MYH) isoform content. We identify by snATAC-seq a 42kb super-enhancer (SE) at the locus regrouping the fast *Myh* (*fMyh*) genes. By 4C-seq we show that active *fMyh* promoters interact with the SE by DNA looping, leading to the activation of a single promoter per nucleus. A rainbow mouse transgenic model of the locus including the SE recapitulates the endogenous spatio-temporal expression of adult *fMyh* genes. *In situ* deletion of the SE by CRISPR/Cas9 editing demonstrates its major role in the control of associated *fMyh* genes, and deletion of two *fMyh* genes at the locus reveals an active competition of the promoters for the shared SE. Last, by disrupting the organization of *fMyh*, we uncover positional heterogeneity within limb skeletal muscles that may underlie selective muscle susceptibility to damage in certain myopathies.

Introduction

Skeletal muscles constitute the most abundant organ in an adult human, about 40% of its total body mass. Most skeletal muscles are composed of a mixture of myofibers with distinct contractile, metabolic, resistance to fatigue properties, as well as differential vulnerability in pathophysiological situations¹. These different myofibers can be classified as slow or fast subtypes that selectively express genes responsible for their specific properties²⁻⁴. The most widely used classification of myofibers types is based on their Myosin heavy chain (MYH) expression profile⁴⁻⁷. MYH, one of the most abundant proteins present in adult myofibers, is a major determinant of myofiber speed of contraction. Each of the mammalian MYH isoform is coded by a specific gene and adult slow-type myofibers express *Myh7* (also known as *MyHCI*, β or *slow*), adult fast-type myofibers express *Myh2* (*MyHCIIA*), *Myh1* (*MyHCIIX*), *Myh4* (*MyHCIIB*) or *Myh13* (*MyHCEO*). During embryonic development *Myh7* and two specific fast *Myh* (*fMyh*) genes, *Myh3* (*MyHCemb*) and *Myh8* (*MyHCperi*) are expressed⁸.

The *fMyh* (*Myh3*, *Myh2*, *Myh1*, *Myh4*, *Myh8*, and *Myh13*) are organized as a cluster within a 350kb region on mouse chromosome 11⁹. The adult fast *Myh2*, *Myh1* and *Myh4* genes are expressed at a low level during embryogenesis and start to be expressed at a much higher level after birth^{8,10-12}. The mechanisms controlling the robust coordinated expression of *fMyh* genes in the hundreds nuclei of a myofiber are not understood. Special regulatory elements called super enhancers (SE) have been shown to control high expression levels for cell lineage identity genes. These SE are composed of multiple enhancer elements spanning 10 to 50 kb of DNA and allowing efficient expression of associated genes¹³⁻¹⁷. As identity genes expressed at high levels in specific fast myofiber subtypes, *fMyh* genes are good candidates to be controlled by a SE in the skeletal muscle lineage. The clustered organization and strict temporal regulation of the *fMyh* locus shows similarities with that of the human β -globin locus¹⁸. At the β -globin locus a common regulatory sequence called locus control region (LCR) interacts dynamically with the different promoters within the locus to activate a single *Globin* isoform in erythroid cells¹⁹⁻²¹. We hypothesized that a LCR/SE at the *fMyh* locus may coordinate the expression of selective *fMyh* genes in adult myofibers to finely control their identity.

To characterize the cis-regulatory elements required for the complex regulation of the specific *fMyh* genes we performed snATAC-seq and 4C-seq experiments with adult skeletal muscles and identified a 42kb opened chromatin region interacting in an exclusive manner with the activated *fMyh* promoter at the locus through 3D chromatin looping as revealed by 4C-seq experiments. A mouse rainbow transgenic line including this SE recapitulates the spatio-temporal expression of endogenous *Myh2*, *Myh1* and *Myh4* genes. We further show by CRISPR/Cas9 editing that *in situ*

deletion of this 42kb SE region prevents expression of fetal *Myh8* and adult *fMyh* genes at the locus leading to fetal myofibers devoid of sarcomeres, unable to contract and precluding breathing at birth. We also tested the hypothesis of promoter competition for the shared SE, and show that absence of *Myh1* and *Myh4* leads to increased expression of *Myh2*, *Myh8* or *Myh13* in specific subregions of limb muscles. Altogether our studies suggest that the *fMyh* SE is responsible for the non-stochastic robust coordinated *fMyh* gene expression in the hundreds of body myonuclei present in adult myofibers. Analysis of the phenotype of all forelimbs and hindlimbs muscles in genetic perturbations within the *fMyh* locus reveals different categories of muscle susceptibility reminiscent of the selective muscle vulnerability observed in different neuromuscular diseases.

Results

Identification of a super enhancer acting as a locus control region in the *fMyh* locus.

The majority of adult myofibers express a single *Myh* gene among the subfamily of fast *Myh4*, *Myh1*, *Myh2*, or slow *Myh7* genes. Fast muscles like the quadriceps are composed of myofibers expressing predominantly *Myh4* or *Myh1* genes while slow muscles like the soleus are composed of myofibers expressing predominantly *Myh7* or *Myh2* genes (Figure 1A-B). To identify the regulatory elements controlling the expression of *fMyh* genes, we performed snATAC-seq experiments with nuclei isolated from adult fast quadriceps and slow soleus¹⁰. Myonuclei were classified based on the chromatin accessibility in the promoter and gene body of *Myh* genes (Figure 1C-D). In *fMyh* myonuclei (*Myh2*, *Myh1*, and *Myh4*), we observed 7 chromatin accessibility peaks in an intergenic region between *Myh3* and *Myh2*. This chromatin region is “closed” in nuclei that do not express *fMyh* genes like slow *Myh7* myonuclei and Fibro Adipogenic Progenitors (FAPs) nuclei where no snATAC-seq peak is detected (Figures 1D, S1A). These chromatin accessibility peaks cover the *Linc-Myh* gene²², and end 25kb upstream of *Myh2* promoter (Figure S1B). Because of its large size of 42kb, this element could correspond to a conserved super enhancer (SE) controlling the *fMyh* genes of the locus in mammals (Figure S1B).

SE were first identified by Chip-seq by their higher level of transcription coactivators and active histone marks accumulation (H3K4me2 and H3K27ac) than conventional enhancers, and by their larger size compared with classic enhancers²³. They regulate the expression of highly transcribed genes specifying cell identity. To test if the element that we identified corresponds to a SE, we compared snATAC-seq results with available ChIP-seq against H3K4me2 and H3K27ac histone marks performed in skeletal muscle²⁴. We observed specific enrichment of these two active histone marks in the same snATAC-seq peaks of chromatin accessibility in the intergenic region between *Myh3* and *Myh2* in quadriceps and soleus (Figures 1E, S1A,²⁴). To determine if this sequence is a SE, we classified the slow and fast specific muscle active enhancers according to the enrichment in H3K27ac histone marks. The 42kb regulatory region of the *fMyh* locus shows a strong enrichment in H3K27ac marks compared to the other enhancers (Figure 1F). This showed that this 42kb intergenic region between *Myh3* and *Myh2* possesses the characteristics of a SE¹⁵ that could control the expression of adjacent *fMyh* genes in the fast quadriceps but also in the slow soleus where around 50% of myofibers express *Myh2* or *Myh1* (Figures 1G, S1A). Based on these criteria, we called this regulatory element *fMyh*-SE.

One of the first SE discovered is the locus control region (LCR) of the β -globin locus²⁵. Like the *fMyh* locus, the human β -globin locus contains a cluster of *globin* gene isoforms expressed sequentially during embryonic, fetal, and adult erythropoiesis²⁶. The LCR of the β -globin locus forms dynamical and specific chromatin loops with the promoter of the gene transcribed at the locus. The similarities between clustered organization and temporal expression at the β -globin and *fMyh* loci suggested that the *fMyh*-SE could act by chromatin looping. To verify this, we performed 4C-seq by purifying nuclei from fast quadriceps and slow soleus. We designed specific primers to quantify the DNA regions interacting with *Myh4* and *Myh2* promoters when these genes are expressed or not. We observed stronger interactions between the *Myh4* promoter and the *fMyh*-SE in the quadriceps where this gene is more transcribed than in the soleus (Figure 1H). On the contrary, we observed more interactions between the *Myh2* promoter and the *fMyh*-SE in the soleus where this gene is more transcribed than in the quadriceps. We confirmed these results by quantifying the interactions between the *fMyh*-SE and other DNA regions in fast muscles. We observed strong and specific interactions between the *fMyh*-SE and the *Myh4* promoter in muscles expressing predominantly *Myh4* gene (Figure S2B). These results showed that the *fMyh*-SE forms specific and dynamic chromatin loops with the promoters of the genes transcribed at the locus, with 3D spatial proximity directly coinciding with activity of the promoter in the fiber type.

The *fMyh* locus is organized in two topological associated domains (TADs).

In mammals, interactions between enhancers and promoters take place preferentially within TADs that are delimited by CTCF insulator binding sites. These CTCF sites can prevent enhancers from activating a gene present in another TAD²⁷⁻³⁰. TAD organization and CTCF insulator sites are conserved between cells and mammalian genomes³¹. We collected data of TAD organization in *fMyh* genes from available Hi-C experiments in embryonic stem cells³². As shown in Figure S 2, the *fMyh* genes are clustered in two distinct TADs separated by CTCF binding on boundary elements observed in ChIP-seq experiments³³. One TAD includes the embryonically expressed *Myh3*, and another adjacent TAD includes all the other *fMyh* genes. To confirm this 3D organization of the fast *Myh* locus, we performed 4C-seq experiments with different viewpoints all along the locus. These experiments confirmed that the *Myh3* promoter interacted mostly with DNA sequences present in its TAD (TAD-1), while the other *fMyh* promoters interacted almost exclusively with sequences present in the *fMyh* TAD. We also observed that the *fMyh*-SE interacted mostly with sequences present in the *fMyh* TAD (Figure

S2B). This suggesting that the adult *fMyh* genes, the fetal *Myh8* gene and the extraocular muscle-specific *Myh13* gene, that are all located in the same TAD, could be controlled by the same *fMyh*-SE. On the contrary, either the regulatory element(s) that control the spatio-temporal expression of *Myh3* should be distinct from the ones controlling the other *fMyh* genes or the TAD boundary should be dynamically reorganized in cells where this gene is active.

A transgenic mouse model of the *fMyh* locus fully recapitulates *Myh1*, *Myh2*, and *Myh4* expression

To create a transgenic mouse model for *fMyh* expression, we inserted into a 222kb bacterial artificial chromosome (BAC) that partially covered the *fMyh* locus (end of *Myh3* to the middle of *Myh8*), the cDNAs coding for YFP at the ATG of *Myh2*, Tomato at the ATG of *Myh1*, and CFP at the ATG of *Myh4* to test their correct expression. A stop codon and a polyA tail were also inserted at the end of each transgene, preventing the expression of fusion proteins between cDNAs and the associated *fMyh*. The recombined BAC was injected in mouse oocytes and 2 separate transgenic animals were obtained and analyzed. We determined by qPCR on genomic DNA that one transgenic line called Enh⁺ integrated 2 complete copies of the entire length of the BAC including the SE. The second independent mouse line called Enh⁻, possesses an incomplete copy of the BAC devoid of the *fMyh* SE (Figure 2A). We observed efficient YFP, Tomato, and CFP expression in all skeletal muscles of Enh⁺ animals (Figures 2B-D, S3A). Expression of the transgenes was not detected in the lung, liver, heart, or kidney (Figure S3A). Next, we compared the expression of the three transgenes with the accumulation of endogenous MYH proteins and mRNAs. As seen in Figure 2E, YFP myofibers were detected in the slow soleus, in agreement with endogenous MYH2 expression, Tomato myofibers were detected in bracoradial muscles, and CFP myofibers in the quadriceps. By immunohistochemistry we observed a strong correlation between the expression of endogenous MYH2 proteins and YFP⁺ myofibers, and between MYH1 proteins and Tomato myofibers (Figure 2F-G). We did not observe the expression of the transgenes in slow MYH7 myofibers of the soleus (Figure S3B). This correlation between transgene and endogenous gene expression was confirmed by RT-qPCR. Efficient *YFP* and *Myh2* mRNA accumulation was found in the soleus of Enh⁺ mice. *Tomato* mRNAs accumulated in both quadriceps and soleus like *Myh1* mRNA. *CFP* mRNA accumulated more in quadriceps than in soleus like *Myh4* mRNAs (Figure 2F-H). The three transgenes were detected in all skeletal muscles of the body including extraocular muscles and Esophagus (Figure S4A,B). Tomato expression was first detected at P0 in the diaphragm when corresponding *fMyh* genes expression become detected (Figure S4C) ¹⁰.

Interestingly, most adult Enh⁺ myofibers expressed only one transgene although hybrid fibers³⁴ were also observed (Figure S4D-F).

While efficient expression of CFP, Tomato and YFP was detected in the skeletal muscles of Enh⁺ animals, very low expression of the three transgenes was observed in Enh⁻ animals (Figure 2I). This decreased transgene expression was also observed by immunostaining on adult muscle sections: much fewer YFP fibers in soleus and much fewer CFP fibers in gastrocnemius were detected in Enh⁻ mice as compared to Enh⁺ mice (Figures 2J, S5A). Transgenes mRNA level was at least 100-fold lower in Enh⁻ than in Enh⁺ animals, as estimated by RT-qPCR (Figure 2K). Altogether, our results show that all regulatory sequences to fully recapitulate the spatiotemporal expression patterns of the *fMyh* genes are present in the modified 222kb BAC in Enh⁺ mice, which roughly overlaps the *fMyh* TAD, and that the *fMyh*-SE and/or other sequences absent in Enh⁻ transgenic animals are required for efficient *Myh2-YFP*, *Myh1-Tomato* and *Myh4-CFP* transgenes expression. Lastly the Enh⁺ rainbow mouse line allows visualizing the fiber-type switches occurring during denervation, aging, in muscle specific *Six1* conditional knock out mouse models, and in other conditions at an individual scale (Figure S6A-F) and is thus a powerful tool to study fiber-type changes in pathophysiological conditions^{4,5}.

The *fMyh*-SE is required for adult *fMyh* and neonatal *Myh8* expression

To assess the requirement of the SE for efficient *fMyh* genes expression *in vivo*, we generated by CRISPR/Cas9 genome editing a knock-out mouse line deleted of this 42kb element (Figures 3A, S7A-B). Heterozygote mutant mice were viable and fertile and presented no obvious deleterious phenotype. In contrast, homozygote mutants died at birth, potentially due to impairment of respiratory skeletal muscle contractions as suggested by the absence of air in their lungs (Figure 3B). E18.5 mutant fetuses showed no major visible skeletal muscle hypoplasia (Figures 3C, G, S7C). In muscles of E18.5 mutant fetuses, the deletion of the *fMyh*-SE induced a strong decrease of the expression of adult *fMyh* (*Myh2*, *Myh1*, and *Myh4*) and of neonatal *Myh8* genes detected by RNAscope on isolated fibers from the diaphragm and quantified by RT-qPCR on leg skeletal muscles (Figure 3D-F). We detected at this embryonic stage regionalized low expression of adult *fMyh* along a few mutant fibers (Figure S 7D), indicating that the *Myh4* gene can be expressed in rare myonuclei in absence of the *fMyh*-SE. Thus, the *fMyh*-SE allows sustained expression of *fMyh* in the syncytium, although not all myonuclei at E18.5 have yet activated the expression of these adult forms (Dos Santos, 2020). *Myh4* or *Myh1* simple KO^{35,36} and *Myh1/Myh4* double KO (see below) mouse are viable and fertile. *Myh2* mutant mice have not been analyzed in detail but seem to have no major phenotype³⁷. The absence of breathing and survival observed in P0 42kb *fMyh*-

SE mutants could be due to the loss of *Myh8* expression, or to the loss of the expression of a combination of several *fMyh* genes (Figure 3D, G, H). At the limb level expression of *Myh3* and *Myh7* was not affected as shown by RT-qPCR experiments and by immunocytochemistry against MYH3 and MYH7 (Figures 3G, H, S7C). In the 42kb *fMyh*-SE E18.5 mutants many limb myofibers presented absence of sarcomeres associated with Actin aggregates around their myonuclei with only a few fibers that did not present these defects (Figures 3J, S7E). Electronic microscopy experiments showed an accumulation of fibrillar materials in mutant diaphragm myofibers that may correspond to Actin accumulation in absence of MYH proteins (Figure 3K), and the absence of sarcomere in many myofibers. These defects of sarcomere formation in mutant myofibers did not impair their innervation but seemed to affect neuromuscular junctions distribution in the diaphragm (Figure 3I). We suspect that unaffected fibers could be primary fibers expressing *Myh7* and or *Myh3*, whose expression appeared normal, while affected myofibers could be secondary myofibers that normally activate the expression of *Myh8* (Figure 3E). The absence of MYH8 could thus lead to sarcomere formation defects leading to Actin aggregates. Altogether these results showed that the *fMyh* SE controls the expression of adult *fMyh* and neonatal *Myh8* and that these isoforms are required for correct sarcomere formation in secondary myofibers and important for efficient muscle contraction at birth.

The *fMyh*-SE is composed of distinct cis-regulatory modules (CRM)

SEs are composed of multiple enhancers and each with a specific role in promoter activation^{38,39}. To characterize the role of two individual CRM identified by snATAC-seq experiments in the SE, we generated their deletion by CRISPR/Cas9 genome editing and obtained two distinct mouse mutant lines. The first CRM enhancer A (*EnhA*) corresponds to a 5Kb region located at the most 3' snATAC-seq peaks of the *fMyh*-SE (Figure 4A). The second CRM enhancer B (*EnhB*) corresponds to two snATAC-seq peaks located in the middle of the *fMyh*-SE. We previously showed that this CRM can activate the expression of *Myh1*, *Myh2*, and *Myh4* promoters in transient adult muscle transfection assays²². In contrast to homozygote mice deleted for the *fMyh*-SE that died at birth, we obtained viable and fertile adult *EnhA* and *EnhB* homozygote mutant mice. We determined the expression of MYH7, MYH2, and MYH4 in the distal hindlimb by immunohistochemistry (Figure 4B) of these mutants. *EnhA*^{-/-} mice showed a strong decrease of MYH4 expression in certain specific muscles (Figure 4B-C). MYH4 was no more detected in the TP and the FHL limb muscles of *EnhA*^{-/-}, while the number of MYH1 fibers increased in these mutant muscles (Figures 4B-D, S8D). This MYH4 fiber-type switch associated with the absence of the *EnhA* was also observed in other muscles (TA, EDL, PB,

PL, FDL, and Plant) while other muscles (Gas and Sol) were spared. This result was confirmed by RT-qPCR experiments showing downregulation of *Myh4* expression in the TA of *EnhA* mutants (Figures 4E, S8E). These results showed that enhancer A dominates regulation of *Myh4* in specific muscles, probably through the recruitment of key *Myh4* identity factors, while dispensable in others and showed also that MYH4 myofibers are not all equivalent. A low expression of *Myh8* and *Myh13* was also detected in adult WT TA which was strongly decreased in *EnhA*^{-/-} TA, demonstrating that the expression of these two genes is also controlled by the enhancer A present in the SE (Figure 4E).

In contrast to the *EnhA* mutant mice, we observed no major modification of slow MYH7 and fast MYH2 and MYH4 expression in muscles of *EnhB* mutant animals by immunostaining (Figure 4B). As for *EnhA* mutant muscles, we observed a decrease of *Myh8* and *Myh13* expression in adult *EnhB* mutant TA compared to the WT (Figure 4E). *Linc-Myh* expression was no more detected in *EnhB*^{-/-} TA. We further generated a transgenic mouse line carrying an *nls-LacZ* transgene under the control of *EnhB* DNA sequences (Figure S8F-G). *Nls-LacZ* transgene expression was detected only in fast and not in slow muscles. These results showed that even if the deletion of *EnhB* do not induce major alterations of adult *fMyh* expression, this DNA element has an enhancer activity in fast adult fibers, as already suggested²². Altogether analysis of these mutant mouse lines revealed that the SE is composed of distinct enhancer elements possessing distinct functions, two of which activate *Myh1*, *Myh2*, *Myh4*, *Myh8* or *Myh13* genes in specific muscles.

The *fMyh* gene promoters compete for the SE

To further elucidate the mechanisms controlling the specific and exclusive activation of *fMyh* promoters, we tested whether these promoters competed for the SE. A mouse model harboring a 72kb deletion of the *Myh1* and *Myh4* genes (*Myh(1-4)*^{Del}) was generated by CRISPR/Cas9 genome editing (Figures 5A, S9A-B). In the deleted allele, *Myh8* and *Myh13* genes are brought closer to the *fMyh*-SE, while the *Myh2* promoter remains at the same distance from the *fMyh*-SE than in the wt allele. *Myh(1-4)*^{Del/+} and *Myh(1-4)*^{Del/Del} animals were viable. No expression of *Myh1* and *Myh4* was detected in *Myh(1-4)*^{Del/Del} animals. These mutants presented a strong hypotrophy in specific areas of individual skeletal muscles, while other areas of the same muscle seemed preserved: the deeper regions of the TA and Gas were more spared than the superficial regions where small myofibers accumulated (Figure 5B). This selective partitioning seemed to less affect deep muscles (Plantaris, PB) compared to the superficial areas of peripheral muscles like the TA or the Gas (Figure 5B). In *Myh(1-4)*^{Del/+} and *Myh(1-4)*^{Del/Del}

mouse, we observed increased *Myh2* expression showing that the deleted allele for *Myh1* and *Myh4* does ectopically activate *Myh2* in the deep regions of muscle masses (Figures 5B, E, S9E-F). We also detected an increased expression of *Myh8* and *Myh13* in both *Myh(1-4)^{Del/+}* and *Myh(1-4)^{Del/Del}* mutant muscles (Figure 5C, D, E). Interestingly we observed in this *Myh(1-4)^{Del/Del}* mutant a deep to peripheral gradient of MYH8 and MYH13 positive myofibers, with increased MYH13 fibers at the peripheral regions of muscle masses (Figure 5C, D). Thus, in absence of *Myh1* and *Myh4* genes, the f*Myh*-SE can activate the expression of either *Myh2*, *Myh8* or *Myh13*, with a degree of plasticity of the myofibers depending on their position inside each individual muscle. These results show that each f*Myh* promoter competes for interaction with the SE and that this competition is influenced by specific muscle subvolumes in agreement with a selective partitioning⁴⁰, and by the deep or superficial position of the muscle itself.

With the sgRNA used to delete *Myh1* and *Myh4*, we obtained two additional mouse lines with a complete inversion of *Myh1* and *Myh4* genes (*Myh(1-4)^{Inv}* and *Myh(1-4)^{Inv3'}*), allowing to test the hypothesis that the order of the *Myh1* and *Myh4* genes in the locus is important for their correct expression. In both these lines, the order of the f*Myh* genes in the locus was modified (*Myh2*, *Myh4*, *Myh1* then *Myh8*). The homozygote *Myh(1-4)^{Inv/Inv}* mutant mice were viable at the homozygous state and showed a strong decrease of *Myh1* expression, a weaker decrease of *Myh4* expression and no difference of *Myh2* expression compared to WT mice (Figure 5B-F). This indicates that a closer proximity of the *Myh4* promoter to the SE did not increase its activity at the adult stage. The strong decrease of *Myh1* expression could be due to the increased distance between its promoter and the f*Myh*-SE, to the disorder of the genes at the locus, or more probably to missing elements in the *Myh1* promoter, since only 575bp upstream of the transcription start site is associated with *Myh1* promoter in the inverted allele. We also observed an upregulation of *Myh8* in this mutant line (Figure 5F).

In the other *Myh(1-4)^{Inv3'}* line, a deletion at the 3' end of *Myh4* was observed, precluding MYH4 synthesis. The homozygote *Myh(1-4)^{Inv3'/Inv3'}* mutant mice were viable, but presented a severe skeletal muscle atrophy. In this mutant mouse line, we observed a strong decrease of *Myh1* and *Myh4* expression (Figure 5B, G). Quantification of *Myh1* and *Myh4* pre-mRNA levels showed that the transcription at the *Myh4* gene in TA was modestly decreased in *Myh(1-4)^{Inv/Inv}* and in *Myh(1-4)^{Inv3'/Inv3'}* mutant as compared with WT, while *Myh1* transcription level was severely downregulated (Figure S9I-J). This showing that *Myh4* promoter can act as a decoy for the SE in *Myh(1-4)^{Inv3'/Inv3'}* since no MYH4 protein is produced. Similarly to the *Myh(1-4)^{Del/Del}* mouse line, we observed an upregulation of *Myh8* and *Myh13* expression in *Myh(1-4)^{Inv3'/Inv3'}* muscles (Figure 5G). Interestingly in *Myh(1-4)^{Inv3'/Inv3'}* animals we observed many MYH2/MYH8

hybrid fibers and many pure MYH13 fibers preferentially in superficial areas of peripheral muscles like the TA or the Gas. MYH13 positive fibers were atrophic (Figures 5D, 6C). We failed to detect MYH3 on *Myh(1-4)^{Inv3'/Inv3'}* adult hindlimb sections (not shown). These results showed that in the inverted allele, the SE could activate misoriented *Myh4* gene, but less efficiently, and activated the expression of *Myh8* and *Myh13* in the myofibers. Altogether these results suggested that competition between the different *Myh* promoters for a shared SE controls their activation and that the order of the genes at the locus does not dictate their correct spatial expression.

Limb skeletal muscles can be classified into 3 major categories with specific genetic programs

At least 640 different skeletal muscles can be identified in the human body, each with a specific form, architecture, position, and function. In several myopathies, skeletal muscles can be specifically affected depending on their anatomic position¹. Distinct genetic programs controlling the identity of each skeletal muscle in its specific environment may determine this selective vulnerability. The different mutants that were generated in this study presented distinct muscle phenotype depending on their location in the body. By comparing the fiber-type composition and fiber size in WT, *EnhA^{-/-}* and *Myh(1-4)^{Inv3'/Inv3'}* mutant mice (Figure 6A-C), we identified 3 different categories of skeletal muscles. The first category corresponded to muscles like the soleus, principally composed of small MYH7 and of MYH2 fibers (Figure 6A). The soleus muscle was not affected in *EnhA^{-/-}* and *Myh(1-4)^{Inv3'/Inv3'}* mouse. The second category of muscles included muscles similar to the Tibialis posterior principally composed of MYH2, MYH1 and MYH4 fibers (Figure 6B). These muscles were affected in *EnhA^{-/-}* and *Myh(1-4)^{Inv3'/Inv3'}* mutant mice and did not express MYH4 anymore. The last category of muscle regrouped muscles similar to the gastrocnemius expressing mainly MYH4 (Figure 6C). The fibers of these groups of muscles presented a drastic decrease of fiber cross section area in the *Myh(1-4)^{Inv3'/Inv3'}* mutants. In contrast, these muscles were not affected in *EnhA^{-/-}* mice. We next extended this study in proximal and distal muscles of the fore- and hindlimbs (Figure 6D-G). As observed at the distal hindlimb level, muscles in forelimbs and proximal hindlimb showed distinct phenotype depending on their deep or superficial position⁴¹. We could detect specific localization of these 3 groups of muscles in the different parts of the hindlimb and forelimbs but with spatial patterns that seemed similar. The category of muscles with similar properties to the Soleus (shown in red) was the most internal in the limb. In contrast, the category of muscles with similar properties to the Gastrocnemius (shown in blue) was the most external.

The category of muscles similar to the Tibialis posterior (shown in green) was located between these two groups (Figure 6D-G). In the proximal part of the hindlimb, the group of muscles shown in blue was the most important and were severely affected in *Myh(1-4)^{Inv3/Inv3'}* mutants, whereas the same group of muscles was almost not affected in *EnhA^{-/-}* mutants (Figure 6E). In the distal part of the forelimb, the group of muscles shown in green prevailed over the other (Figure 6F) whereas in the proximal part, the distribution of these muscles groups was more heterogeneous (Figure 6G). Altogether these results revealed that limb skeletal muscles could be classified in 3 major categories with distinct properties and possessing different codes of transcription factors controlling their plasticity.

Discussion

In adult muscles the contraction and general metabolic properties of the specialized myofibers are dictated by the expression of specific slow MYH7 and fMYH subtypes (MYH2, MYH1, MYH4, MYH13)^{4,5}. Transient transfection experiments of GFP reporters previously suggested that the proximal (800-1000bp) promoters of the *Myh2*, *Myh1* and *Myh4* genes were sufficient to drive their spatial expression in adult muscles⁴². By combining single-nucleus ATAC-seq, ChIP-seq and 4C-seq data from adult fast and slow skeletal muscles, we show here that *fMyh* genes, with the exception of *Myh3*, are regulated by a shared super enhancer. In fast-type myonuclei this SE interacts dynamically with the activated promoters of the locus by 3D chromatin looping. By using rainbow transgenic mouse models of the locus and knock-out mouse models of the SE, we show that this SE controls the level and the spatio-temporal specificity of *fMyh* genes expression in myonuclei and myofibers through exclusive interactions with their promoters. By disrupting the organization of the *fMyh* locus, we uncover positional heterogeneity within limb skeletal muscles that may underlie selective muscle vulnerability observed in certain human neuromuscular diseases.

Properties of the *fMyh*-SE: competition of *fMyh* promoters for the shared SE

SEs, composed of multiple enhancers, allow a more efficient recruitment of coactivators than conventional enhancers. During this process, multimolecular assemblies form by phase separation, allowing aggregation of the transcriptional machinery^{16,23}. Known SEs have been described to achieve a relatively constant high transcriptional activity, contrasting with the transcriptional bursts provided by typical enhancers that lead to episodic gene expression^{15,43}. We show here that the fine spatio-temporal expression of the *fMyh* genes is governed by a SE, which interacts with *fMyh* promoters by 3D chromatin looping, and is engaged in exclusive interactions with a single promoter at the locus. RNAscope experiments with *fMyh* premRNA probes demonstrated previously the coordinated firing of both alleles of specific *fMyh* genes in adult myonuclei¹⁰, suggesting that specific SE-promoter loops may form simultaneously on both alleles of a given *fMyh* gene in the majority of body nuclei of each adult myofiber. Altogether, these results show that the *fMyh*-SE activates a single gene at the *fMyh* locus, suggesting the *fMyh*-SE cannot simultaneously activate two promoters and arguing against flip-flop mechanisms between the *fMyh*-SE and the different promoters of the locus^{17,44,45}. Whether this apparently non-stochastic gene expression in adult myofibers is true for all muscle genes governed by the SE remains to be established. At the *fMyh* locus, we suspect that these exclusive interactions between the SE and specific promoters are responsible for the high level

of *fMyh* expression and to prevent the expression of two different MYH in adult myofibers. To test if these exclusive interactions result from a competition between the SE and the associated *fMyh* promoters, we analyzed the consequences of *Myh1* and *Myh4* deletion. Muscles of adult *Myh(1-4)^{Del/Del}* mutant were composed of myofibers expressing MYH2, MYH8 or MYH13. Interestingly, *Myh2*, which is closest to the SE, was upregulated only in the deep regions of skeletal muscles, while in more peripheral myofibers where *Myh4* is normally predominantly expressed *Myh8* or *Myh13* were activated. These results suggested that *Myh2* cannot be activated in these peripheral myofibers, even in absence of *Myh1* and *Myh4*, and that competition between the promoters varies depending on the muscle position inside the limb, probably due to the differential enrichment of specific transcription factors in deep and peripheral muscles. These experiments demonstrated that some myofibers have the ability to switch from one specific promoter to another non-random promoter, suggesting that the transcription factors bound to *Myh8* and *Myh13* promoters in adult WT limb myofibers are able to interact with the SE, but compete less efficiently than those bound to *Myh4* probably due to a lower frequency of interactions. In adult wt limb muscles these preferential interactions concur to favor *Myh4* at the expense of *Myh8* and *Myh13* expression. Interestingly, even in *Myh(1-4)^{Del/Del}* mutant, very few hybrid fibers were detected^{10,34}, suggesting that most nuclei within each fiber activated a single gene, and that the SE was still contacting a single promoter in an exclusive manner. Such exclusive interactions were also detected at the *β-Globin* locus where the LCR/SE interacts with a single promoter and where the order and the distance between the LCR and the *Globin* genes dictates their temporal expression^{19,46} but were not detected at *α-Globin* locus, where all promoters interact with the SE in a common nuclear compartment⁴⁷.

The SE is composed of individual enhancer elements with redundant activities

To precise the role of the potential enhancer elements composing the SE, we focused on two elements, enhancer A and enhancer B. Deletion of enhancer A led to an upregulation of *Myh1* in myofibers of peroneal muscles, without up-regulation of the nearest *Myh2* promoter, again suggesting that the SE deleted of enhancer A still contacts a single promoter at a time with physically exclusive interactions. Down-regulation of *Myh4* in enhancer A mutant was observed only in certain muscles. This implies that the SE is composed of individual enhancer elements that may have redundant activities under the control of muscle identity genes, as in *Drosophila*^{48,49}. This redundancy may contribute to the expression of a single gene at the locus. Altogether, we did not identify within the SE a specific enhancer responsible for driving the

expression of either *Myh1*, *Myh2* or *Myh4* in all skeletal muscles. We cannot exclude that analysis of deletion of other snATAC-seq peaks may reveal a “master” enhancer element inside the SE. Alternatively the f*Myh*-SE may be composed of redundant elements, none of which absolutely required for driving efficient and specific gene expression, as found at the *α-globin* locus and for enhancers controlling limb and digits morphogenesis where partially redundant enhancers are suspected to provide both flexibility and robustness of gene expression^{38,50-53}.

Importance of the f*Myh*-SE for human pathologies

Our experiments reveal the importance of the f*Myh*-SE for muscle integrity and function. Deletion of the SE induced impaired ability to breathe leading to death at birth. This deletion impaired *Myh1*, *Myh2*, *Myh4* and *Myh8* gene expression in skeletal muscles of E18.5 fetuses (one day before birth in C57BL/6N mouse strain), demonstrating the involvement of the SE to control their expression. Absence of *Myh8* expression may be involved in the death of the mutant animals. The requirement of *Myh8* expression during fetal development for efficient muscle contraction at birth is supported by the phenotype of *Myod*^{-/-};*Nfatc2*^{-/-} mice, where *Myh8* is no more expressed in intercostal muscles. These mutant mice do not survive after birth due to their inability to breathe⁵⁴. In agreement we identified strong sarcomerisation defects associated with Actin aggregates in f*Myh*-SE^{-/-} E18.5 mutant myofibers at the limb and diaphragm level, suggesting a complete absence of MYH and their inability to contract. The SE is present as well in the human f*MYH* locus, and could be involved in the control of the f*MYH* genes as in mice. *MYH8* and *MYH2* are expressed during human fetal development, *MYH1* is detected after birth^{8,55}, while *MYH4* is only expressed in extra ocular myofibers due to mutations in its promoter region^{36,56}. Absence of *MYH2* is associated with early onset myopathy characterized by mild generalized muscle weakness with predominant involvement of muscles of the lower limbs, and by ophthalmoplegia⁵⁷. In contrast, *MYH1* mutations have not yet been reported and *MYH8* mutations do not seem to be associated with trismus-pseudocamptodactyly⁵⁸, contrarily to what was previously suspected. Mutations or deletions of the f*MYH*-SE have not yet been identified in human pathologies. Congenital myopathies can be associated with Actin aggregates, fiber type disproportion or arthrogryposis⁵⁹⁻⁶¹, but not all of these myopathies have been characterized at the genetic level.

The human body is composed of more than 600 different skeletal muscles, each with specific functions and properties. This heterogeneity is reflected in the spectrum of clinical manifestations of neuromuscular diseases where some specific muscles are affected while other are spared, depending on the pathology¹. Whole-body magnetic resonance imaging and muscle

ultrasound in patients affected by Collagen VI deficiency, Dystrophin deficiency or in ALS showed that specific muscles or specific group of myofibers inside a muscle mass can be specifically affected, while others are spared⁶²⁻⁶⁴. Little is known about the mechanisms driving this variability in susceptibility and understanding the underpinning mechanisms is a major challenge to develop adapted targeted therapies. By disrupting the organization of *fMyh* at the locus, we uncovered positional heterogeneity within limb skeletal muscles and defined three major categories of limb muscles. These three categories of stereotyped muscles are differentially positioned in distal and proximal forelimbs and hindlimbs. We suspect that this heterogeneity may be the consequence of different genetic programs that lead to the activation of groups of genes associated with *Myh4*, or *Myh1* or *Myh2* expression. We and others^{10,65,66} revealed recently an unsuspected genetic diversity of *Myh4*⁺ and other myofiber types with, in mouse hindlimbs, at least three subclasses of *Myh4* + myonuclei and several subclasses of *Myh1*⁺ and *Myh2*⁺ myonuclei. Whether this diversity is at the origin of the deep/superficial gradient of muscle susceptibility observed in the present study and in certain neuromuscular diseases remains to be tested.

Methods

Animals

Animal experimentations were carried out in strict accordance with the European STE 123 and the French national charter on the Ethics of Animal Experimentation. Protocols were approved by the Ethical Committee of Animal Experiments of the Institut Cochin, CNRS UMR 8104, INSERM U1016. We used 6-8 weeks old C57BL/6N mouse female for most of our experiments. Mice were anesthetized with intraperitoneal injections of ketamine and xylazine and with subcutaneous buprenex injections before denervation was performed by sectioning of the sciatic nerve in one leg. All efforts were made to minimize animal suffering, and to reduce the number of animals required for the experiments.

BAC targeting constructs and *Myh* locus modifications

For the construction of the targeting vector pGEM-T-EasyMyh2YFP, C57BL/6N mouse DNA was first used as a template to clone 5' arm and 3'arm of *Myh2* with forward 5'- GAA TGA TTT CAT TGC TAC TTC -3' and reverse HindIII 5'- GCT CAT GAC TGC TGA ACT CAC -3', and forward HindIII 5'- AGT CCG AAA AGG AGC GAA TC -3' and reverse 5'- GGT GAC TTC TAG TGA CTG AG -3', respectively. The 5' arm and 3'arm fragments were cloned into a pGEM-T-Easy vector with HindIII in-between to make pGEM-T-EasyMyh2. The Yellow Fluorescent Protein (YFP) coding sequence was PCR amplified (PHUSION, Thermofisher) and cloned in pBluescriptSK+ using EagI-XbaI sites provided by the primers. Fragments containing three polyA sequences (rabbit β -globin, HSV-TK, and BGH) and LoxP-kanamycin-LoxP were then extracted from preexisting constructs and introduced downstream of YFP. The whole YFP-3pA-LoxP-kana-LoxP fragment was amplified (PHUSION, Thermofisher) with forward 5'- CAG CAG TCA TGA GCA TGG TGA GCA AGG GCG AGG AG-3' and reverse 5'- CTC CTT TTC GGA CTA CGA CTC ACT ATA GGG CGA ATT G-3' primers. The resulting amplicon features 15bp homology in 5' and 3' extremities with the targeting arms allowing Sequence and Ligation Independent Cloning (SLIC) into the HindIII digested pGEM-T-EasyMyh2 plasmid (GeneArt Seamless Cloning and Assembling kit, Thermofisher).

Similarly, for the construction of the targeting vector pGEM-T-EasyMyh1Tomato, targeting arms were PCR generated from C57BL/6N mouse DNA and assembled together with HindIII in-between (pGEM-T-Easy-Myh1 : 5' arm forward 5'- CAT CCA GCA TGT GTT CTC AGA GGT -3', reverse HindIII 5'- ACT CAT GGC TGC GGG CTA TT -3' ; 3'arm forward HindIII 5'- GTC TGA AAA GGA GCG AAT CGA G -3', reverse 5'- AGT AGG TCT GCA TCA AGA GAG GG -3'). The PCR amplified tandem-dimer-Tomato (TdTomato) coding sequence was

cloned in Bsp120I-XbaI of pBluescriptSK+. The three polyA signals and Lox2272-kanamycin-Lox2272 cassettes were subsequently added downstream of TdTomato. For SLIC, 5'- CCG CAG CCA TGA GTA TGG TGA GCA AGG GCG AGG AG -3' and 5'- GCT CCT TTT CAG ACA CGA CTC ACT ATA GGG CGA ATT G -3' primers were used and pGEM-T-Easy-Myh1 linearized with HindIII. The targeting vector pGEM-T-EasyMyh4CFP was generated by SLIC of a CFP-3pA-LoxN-KanamycinLoxN PCR fragment into HindIII linearised pGEM-T-EasyMyh4 (C57BL/6N mouse DNA targeting arms : 5'arm forward 5'- CCC AAG TGC TGG AAT TGA AAG TGT -3', reverse HindIII 5'- ACT CAT GGC TGC GGG CTA TT -3' ; 3'arm forward HindIII 5'- GTC TGA AAA GGA GCG AAT CG -3', reverse 5'- GCT AAC TAT CAG CAC GTG CA -3') using forward 5'- CCG CAG CCA TGA GTA TGG TGA GCA AGG GCG AGG AG -3' and reverse 5'- GCT CCT TTT CAG ACA CGA CTC ACT ATA GGG CGA ATT G -3' primers.

A 222kb Bacterial Artificial Chromosome (BAC) from a C57BL/6J mice genomic library ⁶⁷ containing the whole *Mhy2* to *Mhy4* locus surrounded by 80kb of genomic DNA upstream and 46kb downstream is chosen (RP23-61C14 ; CHORI BACPAC resources) to carry out genetic alterations using λ -red recombination ⁶⁸. To remove Lox motifs preexisting on the pBACe3.6 backbone which will later interfere with our strategy of recombination, BAC DNA amplified in DH10b is extracted (Nucleobond MIDI XTRA, Macherey-Nagel), checked by Acc65I-NotI complex restriction profile, and transformed by electroporation into SW105 competent cells. BAC DNA from several transformants is extracted and checked using the same complex restriction profile against the parental one. Removal of LoxP is carried out on one bacterial clone made competent then induced for recombinase expression by 15 minutes incubation at 42°C by electroporation of a 1.85kb BamHI-NotI DNA fragment purified from pTamp-BACe3.6 (gift of Dr V. Besson) conferring ampicillin resistance. BAC DNA from recombinant ampicillin resistant clones is extracted and checked against parental DNA using Acc65-NotI or MfeI-NotI complex restriction profiling. Similarly, removal of Lox511 is performed on one ampicillin resistant clone using a 2.2kb KpnI-BamHI fragment purified from pSKTHygroBACe3.6Lox511 (gift of Dr J. Hadchouel) which confers hygromycin resistance to recombinant clones. DNA from one clone is then transformed into SW106 cells harboring Cre-inducible expression under arabinose treatment ⁶⁹ for further targeting step.

Sequential *Myh2*, *Myh4* and *Myh1* locus modifications are performed by 3 rounds of competent bacterial clone electroporation using a 3.75kb NotI transgene purified from each respective pGEMTe-based targeting vector described above followed by kanamycin selection of recombinant clones, BAC DNA extraction, complex restriction profiling against parental DNA,

then from a proper recombinant clone floxing-out kanamycin resistance by 0.1% arabinose treatment, BAC DNA extraction and again complex restriction profiling against parental DNA. Enzymes combinations are as follows: KpnI+NotI and MfeI+NotI for *Myh2*-YFP, *Myh4*-CFP-kana and *Myh1*-TdT; MfeI+NotI and BamHI+NotI for *Myh4*-CFP. The final transgenic BAC DNA is then transferred back to DH10b cells for better extraction yield (Nucleobond BAC100, Macherey-Nagel). DNA is resuspended in 10mM Tris-HCL pH 7.0, 1mM EDTA, 100mM NaCl. The final transgenic BAC DNA is then transferred back to DH10b cells for better extraction yield (Nucleobond BAC100, Macherey-Nagel). DNA is resuspended in injection buffer (10mM Tris-HCL pH 7.0, 1mM EDTA, 100mM NaCl), and 200ng filtrated through drop dialysis against the filtration buffer for 1h using Millipore cellulose ester disc membranes VMWP 0.05 μ m (Ref# VMWP02500).

Immunohistochemistry

Immunostaining against YFP and MYH2 were performed on soleus and immunostaining against Tomato and Myh1 were performed on quadriceps. Muscles were fixed 30 minutes in PFA 2% with 0,2% Triton at 4°C. After overnight 10% sucrose treatment, muscles were embedded with TissuTEK OCT (Sakura) and frozen in cold isopentane cooled in liquid nitrogen. For immunostaining against MYH4, MYH2, MYH7, and Laminin, adult legs without fixation and without skin were embedded with TissuTEK OCT and directly frozen in cold isopentane cooled in liquid nitrogen. Muscles were conserved at -80°C and cut with Leica cryostat 3050s with a thickness of 10 μ m. Cryostat sections were washed 3 times 5 minutes with PBS and then incubated with blocking solution (PBS and 10% goat serum) 30 minutes at room temperature. Sections were incubated overnight with primary antibody solution at +4°C, then washed 3 times for 5 minutes with PBS and incubated with secondary antibody solution 1 hour at room temperature. Sections were further washed 3 times 5 minutes and mounted with mowiol solution and a glass coverslip. Images were collected with an Olympus BX63F microscope and a Hamamatsu ORCA-Flash 4.0 camera. Images were analyzed with ImageJ program. The references of the antibodies used are listed in Table S1.

RNA extraction and quantification

RNA extractions from adult skeletal muscles were performed using TRIzol reagent (ThermoFischer) following the manufacturer's protocol. Muscles were lysed with Tissue lyser (Quiagen) in TRIzol solution. RNA was precipitated with isopropanol. cDNA synthesis was performed with Superscript III kit (Invitrogen) using 1 μ g of RNA. RT-qPCR were performed

using Light Cycler 480 (Roche) with the Light Cycler 480 SYBR Green I Master Kit (Roche) following the manufacturer's protocol with 40 cycles at 95°C for 15 seconds, 60°C for 15 seconds, and 72°C for 15 seconds. We used 36B4 housekeeping gene to normalize the expression level between different samples. The sequences of the oligonucleotides used are listed in Table S2.

Single nucleus ATAC-seq from skeletal muscle

We use the 10X genomic nuclei Isolation for Single Cell ATAC Sequencing protocol (CG000169 | Rev B) with some changes. 12 quadriceps and 12 soleus were dissected and pulled in cold PBS. PBS was removed and muscles were minced 2 minutes in 1 ml of cold ATAC-lysis buffer (10mM Tris-HCl pH7.4, 10mM NaCl, 3mM MgCl₂, 1% BSA and 0.1% Tween-20 in Nuclease-Free Water). 6ml of cold ATAC-lysis buffer were added and muscles were lysed on ice. After 3 minutes the lysate was dounced with 10 strokes of loose pestle avoiding too much pressure and air bubbles. After douncing, 8 ml of wash buffer were added and the homogenate was filtered with 70µm, 40µm and 20µm cell strainers. Nuclei were pelleted by centrifugation for 5min at 500g at +4°C. Next, we used the Chromium Single Cell ATAC kit according to the manufacturer's protocol. Nuclei were resuspended in nuclei buffer from the kit, transposed 1 hour at 37°C. We loaded around 6000 nuclei into the 10X Chromium Chip. GEM incubation and amplification were performed in a thermal cycler: 72°C for 5 min, 98°C for 30 sec and 12 repeated cycles of 98°C for 10 sec, 59°C for 30 sec and 72°C for 1 min. Post GEM Cleanup using DynaBeads MyOne Silane Beads was followed by library construction (98°C for 45 sec, cycled 12 x 98°C for 20 sec, 67°C for 30 sec, 72°C for 1 min). The libraries were constructed by adding sample index PCR, and SPRIselect size selection. The fragment size estimation of the resulting libraries was assessed with High Sensitivity™ HS DNA kit run on 2100 Bioanalyzer (Agilent) and quantified using the Qubit™ dsDNA High Sensitivity HS assay (ThermoFisher Scientific). Libraries were then sequenced by pair with a HighOutput flowcell using an Illumina Nextseq 500.

Single-nucleus ATAC-seq analysis

A minimum of 10 000 reads per nucleus were sequenced and analyzed with Cell Ranger Single Cell Software Suite 3.0.2 by 10X Genomics. Raw base call files from the Nextseq 500 were demultiplexed with the cellranger-atac mkfastq pipeline into library specific FASTQ files. The FASTQ files for each library were then processed independently with the cellranger count pipeline. This pipeline used STAR21 to align reads to the *Mus musculus* genome. Once aligned,

barcodes associated with these reads –cell identifiers and Unique Molecular Identifiers (UMIs), underwent filtering and correction. The subsequent visualizations, clustering and differential expression tests were performed in R (v 3.4.3) using Seurat36 (v3.0.2) ⁷⁰, Signac (v0.2.4) (<https://github.com/timoast/signac>) and Chromvar (v1.1.1) ⁷¹. Quality control on aligned and counted reads was performed keeping cells with peak_region_fragments > 3000 reads and < 100000, pct reads in peaks > 15, blacklist ratio < 0.025, nucleosome_signal < 10 and TSS.enrichment > 2. We get 6037 nuclei in total and we detected 132 966 peaks (Figure S1). The motif activity score was analyzed by running chromVAR (Figure 3C-E).

ChIP-seq analysis

Fastq files of quadriceps femoris and soleus H3K27ac ChIP-seq ²⁴ were download from the GEO database (accession number GSE123879). The reads were aligned to the mouse mm10 genome using bowtie2 ⁷² and peaks were called by MACS2 ⁷³ using q value cutoff = 0.05. ROSE algorithm ¹⁴ was applied to identify and rank the enhancers based on H3K27ac ChIP-seq signal, with a stitching distance of 12.5 kb.

Nuclei purification from adult skeletal muscle for 4C-seq

Nuclei purification from adult skeletal muscle has been performed as previously described ⁷⁴ with some modifications. After dissection, 16 soleus or 8 quadriceps were resuspended in 1mL of hypotonic buffer (25mM Hepes-KOH pH 7.8, 10mM KCL, 1.5mM MgCl₂, 0.1% NP40, PIC 1X (complete protease inhibitor Roche), PMSF 1mM) in a 2ml tube for 5 min at +4°C. Muscles were sliced with a scissor for 30 sec. The small pieces of muscles were transferred into a round tube of 15mL at +4°C and 4ml of cold hypotonic buffer was added. After 5 min the solution was homogenized for 15 seconds with an Ultra-Turrax (IKA) at a speed of 17,500 rpm. The solution was transferred in a 15 ml Falcon tube and crosslinked with 2% formaldehyde (in a volume of 10ml of hypotonic buffer) at room temperature during 10 min. 1,43 ml of cold glycine (1M) was added to quench the formaldehyde for 5 min at +4°C while shaking. The crosslinked nuclei were dounced 10 times with a loose pestle and then centrifuged at 1000 Rcf for 10 min at +4°C. The nuclear pellet was resuspended in 5ml of hypotonic buffer and filtered with 70µm and 40µm cells strainers. The nuclei were pelleted with centrifugation at 1000 rcf for 10 min, snap frozen into liquid nitrogen and stored at -80°C.

4C-seq

4C-seq experiments have been performed as previously described⁷⁵ with some modifications. Purified crosslinked nuclei from 160 soleus and 80 quadriceps were pooled together to have 10⁷ nuclei per conditions. PCR primers were designed for each viewpoint according to the protocol. The first digestion was done with DpnII (New England Biolabs) and the second with NlaIII (New England Biolabs). For each viewpoint 800ng of 4C template was amplified by PCR. The samples were sequenced on the Illumina NextSeq 500 platform, using 75 bp single end reads. The analysis of the data has been done using the HTSstation 4C-seq pipeline⁷⁶. Briefly, the sequences were demultiplexed, then aligned to the reference genome (mm10) and normalized. Hi-C data in mouse ES cells were obtained from the 3D Genome Browser website (<http://promoter.bx.psu.edu/hi-c/view.php>). ChIP-seq data against CTCF in mouse ES cells and DNase I hypersensitive site in adult fast skeletal muscle were obtained from the ENCODE database. The sequences of the oligonucleotides used for 4C-seq are listed in Table S3.

Mouse generation by CRISPR/Cas9

SgRNA and Cas9 purified protein were produced by the TACGENE platform. The SgRNA were designed with the Crispor program (<http://crispor.tefor.net/>)⁷⁷. SgRNA are produced by T7 Hiscribe transcription kit (New England Biolabs) and purified by EZNA microelute RNA clean up kit (Omega biotek). The DNA used for transcription was produced by overlapping PCR. For each cut sites, 3 different sgRNA were designed and tested *in vitro* by transfection in MEF cells. The deletions were performed by injecting into oocytes between 1 and 5 pg of sgRNA (60ng/μl) cutting at both sides of the deletion and of the Cas9 protein (30 μM). Oocytes were reimplanted into a pseudopregnant females. Mutant mice were screened by PCR and confirmed by sequencing. The list of the sgRNA and PCR primers used for screening are listed in Table S4.

FISH with amplification (RNAscope) on isolated fibers

RNAscope® Multiplex Fluorescent Assay V2 was used to visualize fast *Myh* pre-mRNAs and mRNAs. Twenty different pairs of probes against the first intron of each fast *Myh* transcript were designed by ACDbio. Muscles were dissected and immediately fixed in 4% PFA at +4°C during 30 minutes. After fixation muscles were washed 3 times in PBS for 5 min. Myofibers were dissociated mechanically with small tweezers and fixed onto Superfrost plus slides (Thermo Fischer) coated with Cell-Tak (Corning) by dehydration at +55°C during 5 min. Slides were then proceeded according to the manufacturer's protocol: ethanol dehydration, 10 min of H2O2 treatment and 30 min of protease IV treatment. After hybridization and revelation, the

fibers were mounted under a glass coverslip with Prolong Gold Antifade Mountant (Thermofischer). Myofibers were imaged with a Leica DMI6000 confocal microscope composed by an Okogawa CSU-X1M1 spinning disk and a CoolSnap HQ2 Photometrics camera. Images were analyzed with Fiji Cell counter program.

Statistical analysis

The graphs represent mean values \pm SEM. Significant differences between mean values were evaluated using two-way ANOVA for Fig 2H, one way ANOVA with multiple comparisons for Fig 3D, Fig 4E, Fig 5C, Fig 5D, Fig 5E, Sup 8E, Sup 9F, Sup 9G, Sup 9H and student t test for Fig. 1B, Fig 2K with Graphpad 8.4.3 software.

GEO data accession number

4C-seq data have been deposited in the NCBI Gene Expression Omnibus database (<https://www.ncbi.nlm.nih.gov/geo/>): accession number GSE168074

The reviewers can have (anonymous) access to the data using this link <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE168074>
And using the password odklgemslhuvjgx

Acknowledgments

We thank D.Duboule, S.Schiaffino, H.Amthor, M.Schuelke and F.Spitz for helpful discussions. We thank U.Schibler, S.Gautron and F.Britto for helpful critical reading of the manuscript. We thank T. Guilbert and F. Letourneur at the Cochin IMAG'IC and GENOM'IC platform for helpful advice. We acknowledge the High-throughput sequencing facility of the I2BC (Gif-sur-Yvette, France) for its sequencing and bioinformatics expertise. M.D.S was supported successively by a PhD fellowship from the Association Française contre les Myopathies (AFM), by a post doc fellowship from EUR-Gene and from the Agence Nationale pour la Recherche (ANR Myolinc R17062KK). Financial support to this work was provided by the AFM (n°16427, 21711 and 23012), the ANR (Myocodes R09108KK and Myolinc R17062KK), the Institut National de la Santé et de la Recherche Médicale (INSERM) and the Centre National de la Recherche Scientifique (CNRS) to P.M, by AFM (n° 21711) to J.D, and by AFM (n°19507 and 22946 Translamuscle) to F.R.

Designed experiments, M.D.S, D.N, I.S, F.A, P.M. Performed experiments, M.D.S, I.S, S.B, M.Wu, F.A, R.P, F.L, M.D.C, A.S, J-P.C. Interpreted data, M.D.S, I.S, D.N, A.S, P.M. Bioinformatic analysis, M.Wo, M.D.S. Wrote the manuscript, M.D.S, P.M. with input from D.N, A.S, and J.D. Funding Acquisition, P.M, J.D, F.R.

Competing interests

The authors declare no competing interests.

REFERENCES

- 1 Emery, A. E. Muscular dystrophy into the new millennium. *Neuromuscular disorders* **12**, 343-349 (2002).
- 2 Mantilla, C. B. & Sieck, G. C. Invited review: Mechanisms underlying motor unit plasticity in the respiratory system. *J Appl Physiol (1985)* **94**, 1230-1241, doi:10.1152/jappphysiol.01120.2002 (2003).

- 3 Kanning, K. C., Kaplan, A. & Henderson, C. E. Motor neuron diversity in development and disease. *Annu Rev Neurosci* **33**, 409-440, doi:10.1146/annurev.neuro.051508.135722 (2010).
- 4 Schiaffino, S. & Reggiani, C. Fiber types in mammalian skeletal muscles. *Physiological reviews* **91**, 1447-1531 (2011).
- 5 Gundersen, K. Excitation-transcription coupling in skeletal muscle: the molecular pathways of exercise. *Biol Rev Camb Philos Soc* **86**, 564-600, doi:10.1111/j.1469-185X.2010.00161.x (2011).
- 6 Egan, B. & Zierath, J. R. Exercise metabolism and the molecular regulation of skeletal muscle adaptation. *Cell Metab* **17**, 162-184, doi:10.1016/j.cmet.2012.12.012 (2013).
- 7 Murgia, M. *et al.* Single Muscle Fiber Proteomics Reveals Fiber-Type-Specific Features of Human Muscle Aging. *Cell Rep* **19**, 2396-2409, doi:10.1016/j.celrep.2017.05.054 (2017).
- 8 Schiaffino, S., Rossi, A. C., Smerdu, V., Leinwand, L. A. & Reggiani, C. Developmental myosins: expression patterns and functional significance. *Skelet Muscle* **5**, 22, doi:10.1186/s13395-015-0046-6 (2015).
- 9 Shrager, J. B. *et al.* Human skeletal myosin heavy chain genes are tightly linked in the order embryonic-IIa-IId/x-IIb-perinatal-extraocular. *Journal of muscle research and cell motility* **21**, 345-355 (2000).
- 10 Dos Santos, M. *et al.* Single-nucleus RNA-seq and FISH identify coordinated transcriptional activity in mammalian myofibers. *Nat Commun* **11**, 5102, doi:10.1038/s41467-020-18789-8 (2020).
- 11 Richard, A. F. *et al.* Genesis of muscle fiber-type diversity during mouse embryogenesis relies on Six1 and Six4 gene expression. *Developmental biology* **359**, 303-320 (2011).
- 12 Lu, B. D., Allen, D. L., Leinwand, L. A. & Lyons, G. E. Spatial and temporal changes in myosin heavy chain gene expression in skeletal muscle development. *Dev Biol* **216**, 312-326 (1999).
- 13 Pott, S. & Lieb, J. D. What are super-enhancers? *Nat Genet* **47**, 8-12, doi:10.1038/ng.3167 (2015).
- 14 Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-947, doi:10.1016/j.cell.2013.09.053 (2013).
- 15 Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A Phase Separation Model for Transcriptional Control. *Cell* **169**, 13-23, doi:10.1016/j.cell.2017.02.007 (2017).
- 16 Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, doi:10.1126/science.aar3958 (2018).
- 17 Misteli, T. The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell* **183**, 28-45, doi:10.1016/j.cell.2020.09.014 (2020).
- 18 Palstra, R. J., de Laat, W. & Grosveld, F. Beta-globin regulation and long-range interactions. *Advances in genetics* **61**, 107-142 (2008).
- 19 Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* **10**, 1453-1465, doi:10.1016/s1097-2765(02)00781-5 (2002).
- 20 Palstra, R. J. *et al.* The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet* **35**, 190-194, doi:10.1038/ng1244 (2003).
- 21 Li, Q., Peterson, K. R., Fang, X. & Stamatoyannopoulos, G. Locus control regions. *Blood* **100**, 3077-3086 (2002).

- 22 Sakakibara, I., Santolini, M., Ferry, A., Hakim, V. & Maire, P. Six homeoproteins and a linc-RNA at the fast MYH locus lock fast myofiber terminal phenotype. *PLoS genetics* **10**, e1004386 (2014).
- 23 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319, doi:10.1016/j.cell.2013.03.035 (2013).
- 24 Ramachandran, K. *et al.* Dynamic enhancers control skeletal muscle identity and reprogramming. *PLoS Biol* **17**, e3000467, doi:10.1371/journal.pbio.3000467 (2019).
- 25 Grosveld, F., van Assendelft, G. B., Greaves, D. R. & Kollias, G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51**, 975-985, doi:10.1016/0092-8674(87)90584-8 (1987).
- 26 Noordermeer, D. & de Laat, W. Joining the loops: beta-globin gene regulation. *IUBMB Life* **60**, 824-833, doi:10.1002/iub.129 (2008).
- 27 Downen, J. M. *et al.* Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* **159**, 374-387, doi:10.1016/j.cell.2014.09.030 (2014).
- 28 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).
- 29 Lupianez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012-1025, doi:10.1016/j.cell.2015.04.004 (2015).
- 30 Chang, L. H., Ghosh, S. & Noordermeer, D. TADs and Their Borders: Free Movement or Building a Wall? *J Mol Biol* **432**, 643-652, doi:10.1016/j.jmb.2019.11.025 (2020).
- 31 Merkenschlager, M. & Odom, D. T. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* **152**, 1285-1297, doi:10.1016/j.cell.2013.02.029 (2013).
- 32 Bonev, B. *et al.* Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**, 557-572 e524, doi:10.1016/j.cell.2017.09.043 (2017).
- 33 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).
- 34 Medler, S. Mixing it up: the biological significance of hybrid skeletal muscle fibers. *J Exp Biol* **222**, doi:10.1242/jeb.200832 (2019).
- 35 Allen, D. L., Harrison, B. C. & Leinwand, L. A. Inactivation of myosin heavy chain genes in the mouse: diverse and unexpected phenotypes. *Microsc Res Tech* **50**, 492-499, doi:10.1002/1097-0029(20000915)50:6<492::AID-JEMT6>3.0.CO;2-J (2000).
- 36 Harrison, B. C., Allen, D. L. & Leinwand, L. A. IIb or not IIb? Regulation of myosin heavy chain gene expression in mice and men. *Skelet Muscle* **1**, 5, doi:10.1186/2044-5040-1-5 (2011).
- 37 Geurts, A. M. *et al.* Gene mutations and genomic rearrangements in the mouse as a result of transposon mobilization from chromosomal concatemers. *PLoS Genet* **2**, e156, doi:10.1371/journal.pgen.0020156 (2006).
- 38 Hay, D. *et al.* Genetic dissection of the alpha-globin super-enhancer in vivo. *Nat Genet* **48**, 895-903, doi:10.1038/ng.3605 (2016).
- 39 Hornblad, A., Bastide, S., Langenfeld, K., Langa, F. & Spitz, F. Dissection of the Fgf8 regulatory landscape by in vivo CRISPR-editing reveals extensive intra- and inter-enhancer redundancy. *Nat Commun* **12**, 439, doi:10.1038/s41467-020-20714-y (2021).
- 40 English, A. W., Wolf, S. L. & Segal, R. L. Compartmentalization of muscles and their motor nuclei: the partitioning hypothesis. *Phys Ther* **73**, 857-867, doi:10.1093/ptj/73.12.857 (1993).

- 41 Charles, J. P., Cappellari, O., Spence, A. J., Hutchinson, J. R. & Wells, D. J. Musculoskeletal Geometry, Muscle Architecture and Functional Specialisations of the Mouse Hindlimb. *PLoS One* **11**, e0147669, doi:10.1371/journal.pone.0147669 (2016).
- 42 Allen, D. L., Weber, J. N., Sycuro, L. K. & Leinwand, L. A. Myocyte enhancer factor-2 and serum response factor binding elements regulate fast Myosin heavy chain transcription in vivo. *J Biol Chem* **280**, 17126-17134, doi:10.1074/jbc.M501207200 (2005).
- 43 Fukaya, T., Lim, B. & Levine, M. Enhancer Control of Transcriptional Bursting. *Cell* **166**, 358-368, doi:10.1016/j.cell.2016.05.025 (2016).
- 44 Wijgerde, M., Grosveld, F. & Fraser, P. Transcription complex stability and chromatin dynamics in vivo. *Nature* **377**, 209-213, doi:10.1038/377209a0 (1995).
- 45 Bartman, C. R. *et al.* Transcriptional Burst Initiation and Polymerase Pause Release Are Key Control Points of Transcriptional Regulation. *Mol Cell* **73**, 519-532 e514, doi:10.1016/j.molcel.2018.11.004 (2019).
- 46 Dillon, N., Trimborn, T., Strouboulis, J., Fraser, P. & Grosveld, F. The effect of distance on long-range chromatin interactions. *Mol Cell* **1**, 131-139, doi:10.1016/s1097-2765(00)80014-3 (1997).
- 47 Oudelaar, A. M. *et al.* A revised model for promoter competition based on multi-way chromatin interactions at the alpha-globin locus. *Nat Commun* **10**, 5412, doi:10.1038/s41467-019-13404-x (2019).
- 48 Bataille, L., Boukhatmi, H., Frenodo, J. L. & Vincent, A. Dynamics of transcriptional (re)-programming of syncytial nuclei in developing muscles. *BMC Biol* **15**, 48, doi:10.1186/s12915-017-0386-2 (2017).
- 49 de Jossineau, C., Bataille, L., Jagla, T. & Jagla, K. Diversification of muscle types in Drosophila: upstream and downstream of identity genes. *Curr Top Dev Biol* **98**, 277-301, doi:10.1016/B978-0-12-386499-4.00011-2 (2012).
- 50 Montavon, T. *et al.* A regulatory archipelago controls Hox genes transcription in digits. *Cell* **147**, 1132-1145, doi:10.1016/j.cell.2011.10.023 (2011).
- 51 Osterwalder, M. *et al.* Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239-243, doi:10.1038/nature25461 (2018).
- 52 Dukler, N., Gulko, B., Huang, Y. F. & Siepel, A. Is a super-enhancer greater than the sum of its parts? *Nat Genet* **49**, 2-3, doi:10.1038/ng.3759 (2016).
- 53 Bolt, C. C. & Duboule, D. The regulatory landscapes of developmental genes. *Development* **147**, doi:10.1242/dev.171736 (2020).
- 54 Daou, N. *et al.* A new role for the calcineurin/NFAT pathway in neonatal myosin heavy chain expression via the NFATc2/MyoD complex during mouse myogenesis. *Development* **140**, 4914-4925, doi:10.1242/dev.097428 (2013).
- 55 Butler-Browne, G. S., Barbet, J. P. & Thornell, L. E. Myosin heavy and light chain expression during human skeletal muscle development and precocious muscle maturation induced by thyroid hormone. *Anat Embryol (Berl)* **181**, 513-522, doi:10.1007/BF00174624 (1990).
- 56 Stirn Kranjc, B., Smerdu, V. & Erzen, I. Histochemical and immunohistochemical profile of human and rat ocular medial rectus muscles. *Graefes Arch Clin Exp Ophthalmol* **247**, 1505-1515, doi:10.1007/s00417-009-1128-0 (2009).
- 57 Tajsharghi, H. & Oldfors, A. Myosinopathies: pathology and mechanisms. *Acta Neuropathol* **125**, 3-18, doi:10.1007/s00401-012-1024-2 (2013).

- 58 Dai, Z. *et al.* Caution in interpretation of disease causality for heterozygous loss-of-function variants in the MYH8 gene associated with autosomal dominant disorder. *Eur J Med Genet* **60**, 312-316, doi:10.1016/j.ejmg.2017.03.012 (2017).
- 59 Gonorazky, H. D., Bonnemann, C. G. & Dowling, J. J. The genetics of congenital myopathies. *Handb Clin Neurol* **148**, 549-564, doi:10.1016/B978-0-444-64076-5.00036-3 (2018).
- 60 Sewry, C. A., Laitila, J. M. & Wallgren-Pettersson, C. Nemaline myopathies: a current view. *J Muscle Res Cell Motil* **40**, 111-126, doi:10.1007/s10974-019-09519-9 (2019).
- 61 Pelin, K. & Wallgren-Pettersson, C. Update on the Genetics of Congenital Myopathies. *Semin Pediatr Neurol* **29**, 12-22, doi:10.1016/j.spen.2019.01.005 (2019).
- 62 Mercuri, E. *et al.* Muscle magnetic resonance imaging involvement in muscular dystrophies with rigidity of the spine. *Ann Neurol* **67**, 201-208, doi:10.1002/ana.21846 (2010).
- 63 Tomas, X. *et al.* Whole-body MRI and pathological findings in adult patients with myopathies. *Skeletal Radiol* **48**, 653-676, doi:10.1007/s00256-018-3107-1 (2019).
- 64 Nakamura, R. *et al.* Spinal Muscular Atrophy Type 3 Showing a Specific Pattern of Selective Vulnerability on Muscle Ultrasound. *Intern Med*, doi:10.2169/internalmedicine.6396-20 (2021).
- 65 Kim, M. *et al.* Single-nucleus transcriptomics reveals functional compartmentalization in syncytial skeletal muscle cells. *Nat Commun* **11**, 6375, doi:10.1038/s41467-020-20064-9 (2020).
- 66 Petrany, M. J. *et al.* Single-nucleus RNA-seq identifies transcriptional heterogeneity in multinucleated skeletal myofibers. *Nat Commun* **11**, 6374, doi:10.1038/s41467-020-20063-w (2020).
- 67 Osoegawa, K. *et al.* Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. *Genome Res* **10**, 116-128 (2000).
- 68 Copeland, N. G., Jenkins, N. A. & Court, D. L. Recombineering: a powerful new tool for mouse functional genomics. *Nat Rev Genet* **2**, 769-779, doi:10.1038/35093556 (2001).
- 69 Lee, E. C. *et al.* A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**, 56-65, doi:10.1006/geno.2000.6451 (2001).
- 70 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821, doi:10.1016/j.cell.2019.05.031 (2019).
- 71 Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* **14**, 975-978, doi:10.1038/nmeth.4401 (2017).
- 72 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
- 73 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
- 74 Joshi, S., Ueberschlag-Pitiot, V., Metzger, D. & Davidson, I. Improved Protocol for Chromatin Immunoprecipitation from Mouse Skeletal Muscle. *J Vis Exp*, doi:10.3791/56504 (2017).
- 75 Matelot, M. & Noordermeer, D. Determination of High-Resolution 3D Chromatin Organization Using Circular Chromosome Conformation Capture (4C-seq). *Methods Mol Biol* **1480**, 223-241, doi:10.1007/978-1-4939-6380-5_20 (2016).

- 76 David, F. P. *et al.* HTSstation: a web application and open-access libraries for high-throughput sequencing data analysis. *PLoS One* **9**, e85879, doi:10.1371/journal.pone.0085879 (2014).
- 77 Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* **17**, 148, doi:10.1186/s13059-016-1012-2 (2016).

Legend figures

Figure 1. Identification of a super enhancer in the intergenic region of *Myh3* and *Myh2*.

(A) Adult myofibers expressed different MYH isoforms. Immunostaining against fast (MYH4, MYH2) and slow (MYH7) MYH of adult fast quadriceps (Quad) and slow soleus (Sol) muscle sections, MYH1+ myofibers by default appear black. (B) Quantification by RT-qPCR of *Myh* mRNA expression in adult Quad and Sol. (C) Graphical scheme of the experiments used for snATAC-seq experiments performed with slow soleus and fast quadriceps adult skeletal muscle. (D) Chromatin accessibility of the different types of myonuclei in the fast *Myh* locus. In fast myonuclei, we identified a 42kb region with multiple chromatin accessibility peaks in the intergenic region of *Myh3* and *Myh2* genes. In slow *Myh7* myonuclei this region of chromatin is not accessible. (E) H3K27Ac and H3K4me2 ChIP-seq signals²⁴ were highly enriched in the 42kb region of snATAC-seq peaks in the intergenic region of *Myh3* and *Myh2* genes. (F) Plot showing ranked H3K27ac ChIP-seq signals for quadriceps enhancers. (G) Same as (F) in soleus. (H) 4C-seq experiments showing the interaction between the *Myh4* (Up) and *Myh2* (down) promoters in quadriceps and soleus: the promoter of the gene activated at the locus interacts with the 42kb cis-regulatory DNA module. Numerical data are presented as mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

Figure 2. Transgenic models to study f*Myh* genes expression.

(A) Schematic representation of mouse f*Myh* locus and the recombined 222kb Bacterial Artificial Chromosome (BAC) of the same locus. YFP, Tomato and CFP cDNAs were inserted in the first exon of *Myh1*, *Myh2* and *Myh4* genes respectively in the BAC. Two transgenic mouse lines were obtained, one called Enh+ that integrated 2 complete copies of the BAC and the other called Enh- devoid of the SE region and the 3' region of the locus. The transgenes YFP, Tomato and CFP are not to scale. (B) Picture of a five days old Enh+ transgenic mouse, lateral view. Red; Tomato, green; YFP and blue; CFP. All skeletal muscles expressed the transgenes. (C) Same as (B), zoom in intercostal muscles. (D) Same as (B) in intercostal and abdominal muscles of a 2-month-old Enh+ transgenic mouse. (E) Transgene expression in adult soleus (Sol), bracoradial (Braco) and quadriceps (Quad) showing predominant expression of YFP in green, Tomato in red and CFP in blue for each muscle. (F) Expression of the transgenes correlates with endogenous MYH protein expression in Enh+ line. Up: immunofluorescence against endogenous MYH2 (red) and of YFP (green) in adult soleus transverse section of Enh+

mice. Down: immunofluorescence against endogenous MYH1 (green) and of Tomato (red) in adult quadriceps transverse section of Enh⁺ mice. (G) Quantification of the percentage of MYH2 or MYH1 fibers expressing YFP or Tomato respectively, (n=3). All MYH2 fibers are YFP⁺ and almost all MYH1 fibers are Tomato⁺. (H) Relative expression level of mRNA in adult Sol and Quad of endogenous *Myh* genes and of transgenes, in wild type (wt) and in Enh⁺ mice (n=3). (I) Pictures of the adult leg of Enh⁺ (left) and Enh⁻ (right) mouse. The expression of the three transgenes is much higher in the Enh⁺ line compared to Enh⁻ mouse. (J) Immunostaining with GFP antibodies revealing YFP fibers on a section of adult Sol in Enh⁺ and Enh⁻ mice. In Enh⁺ mouse, all MYH2 fibers expressed YFP whereas in Enh⁻ only 10% of MYH2 fibers expressed YFP. (K) Quantification by RT-qPCR of transgenes expression in Enh⁺ and Enh⁻ mouse line. Numerical data are presented as mean ± S.E.M. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. Scale bars: 100µm for F, and 50µm for J.

Figure 3. The f*Myh*-SE is required for adult f*Myh* and neonatal *Myh8* genes expression.

(A) A mouse line deleted for the f*Myh*-SE element was generated by injecting specific sgRNAs and Cas9 protein into mouse oocytes. (B) f*Myh*-SE^{-/-} mice died at birth (P0) without breathing and air in their lungs. (C) f*Myh*-SE^{-/-} E18.5 fetuses did not present severe visible malformations. (D) Quantification of *Myh* mRNAs by RT-qPCR in control and f*Myh*-SE^{-/-} E18.5 forelimb skeletal muscles. Mutant muscles showed decreased *Myh2*, *Myh1*, *Myh4* and *Myh8* mRNAs levels. (E) RNAscope experiments against *Myh3* and *Myh8* mRNAs on isolated E18.5 forelimb fibers of control and mutant mice. (F) Same as (E) showing a decreased accumulation of *Myh2* and *Myh4* mRNAs in mutant mice compared to their littermate controls. (G) Immunostaining at the distal hindlimb level of E18.5 control and mutant fetuses revealing MYH3 and MYH8 positive myofibers. (H) Same as (G), zoom in the EDL of control and mutant. (I) *In toto* immunostaining of diaphragms from E18.5 mutant and control fetuses showing in red Actin filaments (phalloidine), in green AchR (alpha-bungarotoxin), and in pink neurofilaments showing altered repartition of NMJ and punctated Actin aggregates in mutant diaphragms. (J) Myofibers from mutant diaphragm showed defects in sarcomeres organization as shown by phalloidine staining. (K) Electronic microscopy pictures of the sarcomeres defects present in mutant E18.5 fetuses compared to their littermate controls. For D (n=3). For E and F, scale bar: 50 µm. For G, scale bar: 500 µm. For H, scale bar: 100 µm. Numerical data are presented as mean ± s.e.m. **P* < 0.05, ***P* < 0.01.

Figure 4. Role of the different enhancers composing the SE.

(A) Schematic representation of the snATAC-seq peaks along the 42kb SE and the enhancers A and B deleted by CRISPR/Cas9 editing. (B) Immunostaining against fast MYH2, MYH4 and slow MYH7 on adult leg sections of 2-3-month-old mouse female deleted for enhancer A or B. (C) Same as (B), zoom in Tibialis posterior and FHL muscle of WT and *EnhA*^{-/-} mutant. (D) Immunostaining against fast MYH1 in Tibialis posterior and FHL muscle of WT and *EnhA*^{-/-} mutant. The absence of *EnhA* induced an increased number of MYH1 positive fibers. (E) Quantification of *fMyh* mRNA and of *Linc-Myh* in adult TA of control and *EnhA* and *EnhB* mutant by RT-qPCR experiments. Schema of the interactions identified. For E, n=3. Numerical data are presented as mean ± S.E.M. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. Scale bars: 100µm for B.

Figure 5. The promoters of *fMyh* genes compete for the shared SE

(A) Schema of the distinct *fMyh* alleles generated by CRISPR/Cas9 editing. (B) Immunostaining against MYH4 (blue), MYH2 (green) and slow MYH7 (red) of adult distal hindlimb sections of wt, *Myh(1-4)*^{Del/Del}, *Myh(1-4)*^{Inv/Inv} and of *Myh(1-4)*^{Inv3'/Inv3'} mutants. (C) Immunostaining against neonatal MYH8 of adult leg sections in wt, *Myh(1-4)*^{Del/Del}, *Myh(1-4)*^{Inv/Inv} and of *Myh(1-4)*^{Inv3'/Inv3'}. (D) Same as (C) against extraocular MYH13. (E) Quantification of *Myh2*, *Myh1*, *Myh4*, *Myh8* and *Myh13* mRNAs of adult wt, *Myh(1-4)*^{Del/+} and *Myh(1-4)*^{Del/Del} TA by RT-qPCR experiments. (F) Quantification of *Myh2*, *Myh1*, *Myh4*, *Myh8* and *Myh13* mRNAs of adult wt, *Myh(1-4)*^{Inv/+} and *Myh(1-4)*^{Inv/Inv} TA by RT-qPCR experiments. (G) Quantification of *Myh2*, *Myh1*, *Myh4*, *Myh8* and *Myh13* mRNAs of adult wt, *Myh(1-4)*^{Inv3'/+} and *Myh(1-4)*^{Inv3'/Inv3'} TA by RT-qPCR. For E, F, G (n=3). Numerical data are presented as mean ± S.E.M. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. Scale bars: 50µm for G.

Figure 6. MYH expression established different groups of limb skeletal muscles in mutant animals.

(A) Immunostaining against MYH7 (red), MYH2 (green) and MYH4 (blue) in soleus of WT, *EnhA*^{-/-}, and *Myh(1-4)*^{Inv3'/Inv3'} adult mice. The soleus is not affected in these mutant mice. (B) Immunostaining against MYH7 (red), MYH2 (green) and MYH4 (blue) in Tibialis posterior of WT, *EnhA*^{-/-}, and *Myh(1-4)*^{inv3'/inv3'} adult mice. In this muscle, the expression of MYH4 is lost in *EnhA*^{-/-}, and in *Myh(1-4)*^{inv3'/inv3'}. In this muscle an upregulation of MYH1 is observed in *EnhA*^{-/-} mice and an upregulation of MYH2 in *Myh(1-4)*^{Inv3'/Inv3'} mice. (C) Immunostaining against MYH7 (red), MYH2 (green) and MYH4 (blue) in peripheral Gastrocnemius of WT, *EnhA*^{-/-}, and *Myh(1-4)*^{Inv3'/Inv3'} adult mice. The peripheral fibers of *Myh(1-4)*^{Inv3'/Inv3'} Gas present

a severe atrophy. In contrast these fibers are not affected in *EnhA*^{-/-} mouse. (D) Comparison of the phenotypes of adult muscles in *EnhA*^{-/-} and in *Myh(1-4)*^{Inv3'/Inv3'} allowed classification of distal hindlimb muscles in three major categories. The first group shown in red corresponded to the soleus that is not affected. The second group shown in green corresponded to muscles affected in *EnhA*^{-/-} and *Myh(1-4)*^{Inv3'/Inv3'} mutants. The third group shown in blue corresponded to muscles strongly affected in *Myh(1-4)*^{Inv3'/Inv3'} but not in *EnhA*^{-/-} mutants. (E) Same as (D) in the proximal part of the hindlimb. (F) Same as (D) in the distal part of the forelimb. (G) Same as (D) in the proximal part of the forelimb. For D, scale bar: 100µm. D-G: drawings of hindlimbs and forelimbs are from Charles et al ⁴¹.

Figure S1. Active histone marks in slow and fast muscles. (A) Chromatin accessibility as identified in snATAC-seq experiments in FAPs and *Myh4* myonuclei, in ChIP-seq experiments ²⁴ for H3K27Ac and H3K4me2 in quadriceps and soleus in the *fMyh* locus. (B) Same as (A) zoom in the *fMyh* SE, and its seven snATAC-seq peaks, placental base base conservation by PhiloP and multi alignment of vertebrate DNA sequences of the 42kb SE.

Figure S2. Distinct TADs partitioned the *fMyh* locus as determined by HiC-seq experiments.

(A) Alignment of Hi-C data from mES cells ³². (B) 4C-seq data from adult fast skeletal muscles and ChIP-seq data against CTCF in C2C12 myoblast ³³ in the *fMyh* locus and adjacent TADs. HiC-seq data revealed that the *fMyh* locus is organized in one TAD delimited by CTCF borders at the 3' end of *Myh3* gene and at the 3' end of *Myh13*, showing that *Myh3* does not belong to the same TAD that the other *fMyh* genes. The size of the TAD including *Myh1*, *Myh2* and *Myh4* is estimated at 350kb. For 4C-seq experiments, we used 7 distinct viewpoints in the *fMyh* locus to determine its organization in adult leg muscles.

Figure S3. *fMyh* regulation in *Enh*⁺ transgenic animals. (A) Transgenes expression in several tissues in *Enh*⁺ line. The expression of YFP, Tomato and CFP is only detected in skeletal muscle (TA and soleus in the picture) and not in other organs like lung, heart, kidney, or liver. (B) Immunostaining against endogenous MYH7 (red) and YFP (with an anti GFP antibody in green) on adult soleus of *Enh*⁺ mice showing that MYH7 fibers does not express YFP.

Figure S4. Expression of the transgenes in skeletal muscles of *Enh*⁺ mice.

(A) Schema of the three inserted transgenes in the *fMyh* BAC, and expression of the transgenes in extraocular muscles. Red; Tomato, green; YFP and blue; CFP. (B) same as in (A) showing the esophagus muscle of an adult *Enh*⁺ mouse. (C) Same as (A) showing the diaphragm in P0 *Enh*⁺ mouse. (D) Transgenes expression in isolated fibers from the EDL. We could detect pure YFP, Tomato and CFP fibers and the expression of the transgene is similar all along the fibers. At the top, a pink Tomato-CFP hybrid fiber. (E) The majority of soleus fibers express only one transgene; Tomato or YFP. (F) A minority of fibers co-expressed Tomato and YFP in adult soleus muscle.

Figure S5. Transgene expression in *Enh*⁻ mouse. (A) Immunostaining with GFP antibodies revealing YFP and CFP proteins on sections of adult soleus, plantaris and gastrocnemius in *Enh*⁺ (left) and *Enh*⁻ (right) mice.

Figure S6. Transgenic *Enh*⁺ mice allowed identification of fiber-type modifications under various pathophysiological conditions.

(A) Control and 4 weeks denervated hindlimb legs of *Enh*⁺ mice. Distal muscles are more affected by the absence of innervation than proximal muscles. A strong atrophy of the gastrocnemius is observed. (B) Same as (A), zoom in Gastrocnemius, Tibialis anterior, EDL and soleus. Denervation induced a fast to slow transition in these 3 muscles visible by the expression of the transgenes. (C) 2-month-old *Enh*⁺ female and male mice. Skeletal muscles of females have more Tomato/MYH1 fibers and less CFP/MYH4 fibers. (D) Same as (C) in 30-month-old *Enh*⁺ female and male. (E) hindlimb of 2- and 30-month-old *Enh*⁺ mice. (F) Control and *Six1*^{flox/flox}; *HSA-CRE* hindlimb muscles of *Enh*⁺ mice. The absence of *Six1* in skeletal muscles induced a fast to slow fiber type switch. Gas: Gastrocnemius, TA: tibialis anterior, Sol: soleus.

Figure S7. Phenotype of *fMyh-SE* deletion in E18.5 mutant fetuses. (A) Diagram showing the strategy for *in vivo* deletion of the *fMyh-SE* by CRISPR/Cas9. (B) Sequence of the deleted allele. (C) Immunostaining of distal hindlimb sections of control and *fMyh-SE*^{-/-} showing MYH8 (green), MYH7 (purple) and Laminin (white) expression. (D) RNAscope experiments against *Myh2* and *Myh4* mRNAs on isolated E18.5 fibers of control and mutant fetuses. (E)

Myofibers from mutant diaphragm showed defects in sarcomeres organization as revealed by phalloidine staining. For E, scale bar: 500 μ m.

Figure S8. Muscle phenotype of enhancers A and B deletion. (A) Diagram showing the strategies for *in vivo* deletions of enhancer A or B by CRISPR/Cas9. (B) PCR showing the DNA fragment amplified after *EnhA* deletion, and sequence of the deleted allele. (C) Sequence of the deleted *EnhB* allele. (D) Immunostaining against MYH1 on adult distal hindlimb sections in control and *EnhA*^{-/-}. (E) Quantification of *Myh3*, *Myh2*, *Myh1*, *Myh4*, *Myh8* and *Myh13* mRNAs and of *Linc-Myh* by RT-qPCR in Soleus of control, *EnhA* and *EnhB* mutants. (F) Schema of the transgene used to generate a transgenic mouse line expressing nuclear beta-galactosidase under the control of the 1.8kb *EnhB* DNA element. (G) Beta-galactosidase positive nuclei are detected in the fast tibialis posterior but not in the slow soleus. For E, n=3. Numerical data are presented as mean \pm S.E.M. **P* < 0.05, ***P* < 0.01, ****P* < 0.001

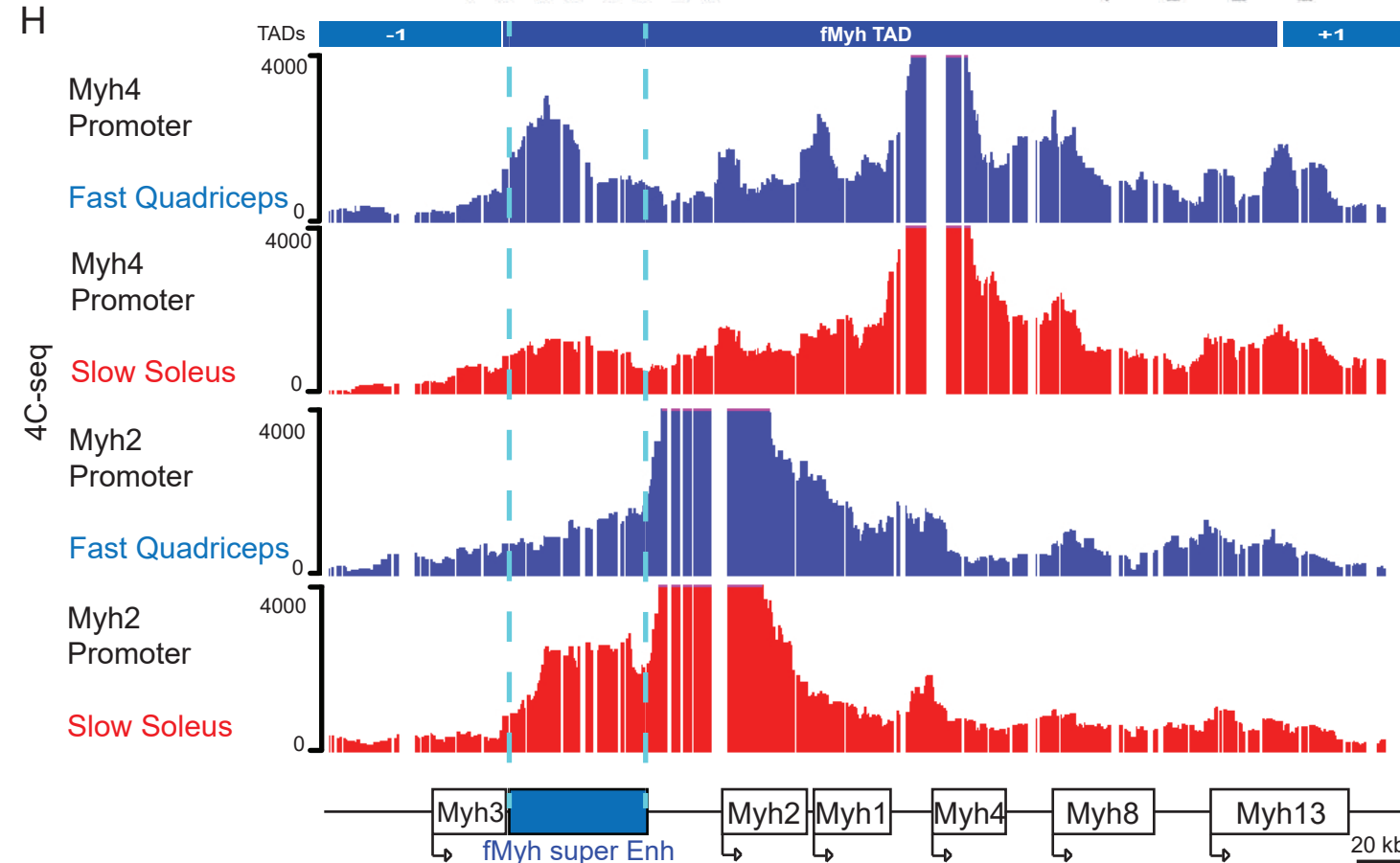
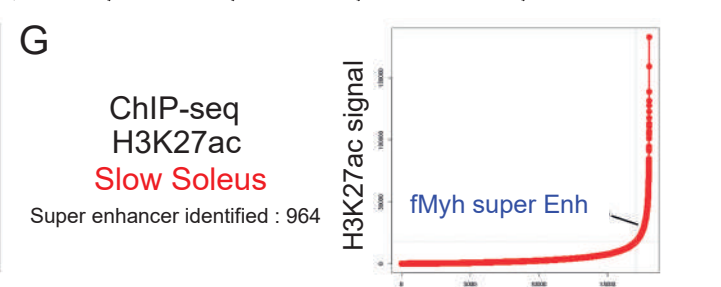
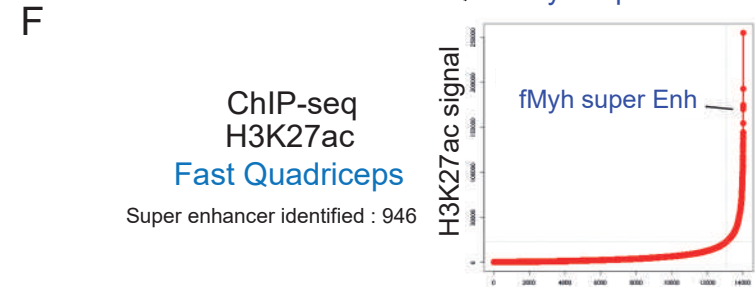
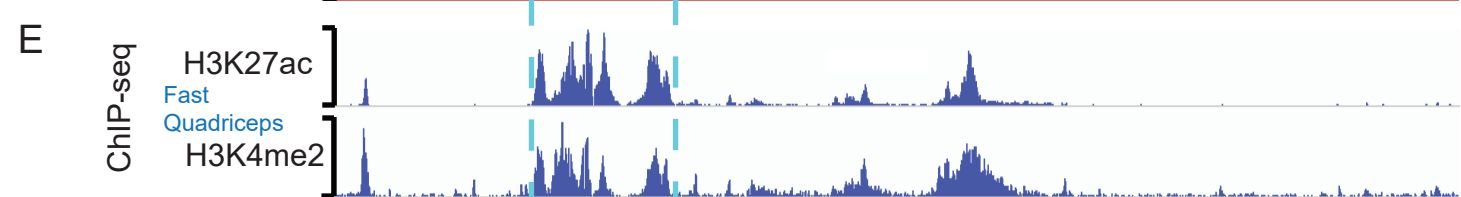
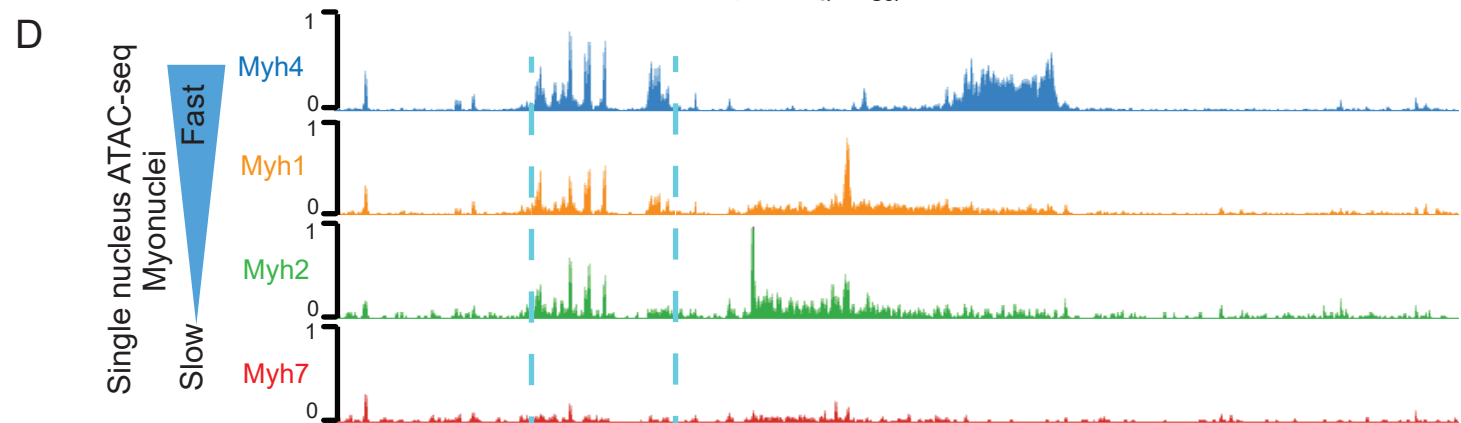
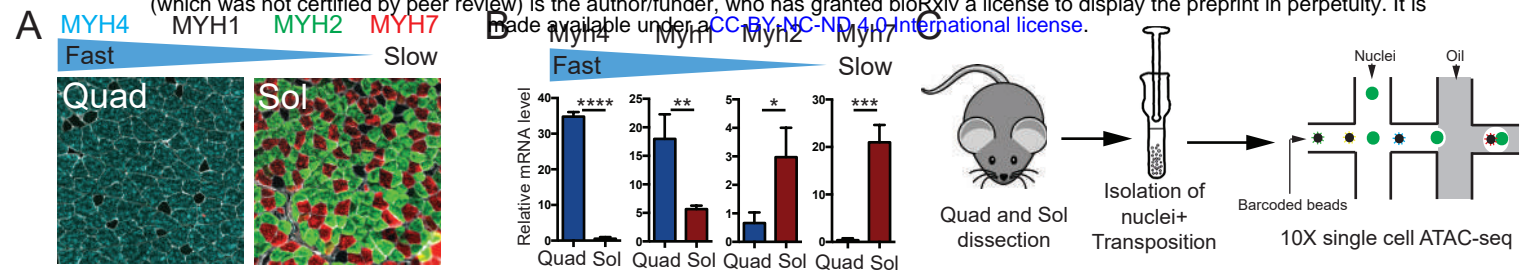
Figure S9. Competition of the fMyh gene promoters for the shared SE. (A) Diagram showing the strategy for *in vivo* deletion of *Myh1* and *Myh4* by CRISPR/Cas9. (B) PCR used for the screening of mutant animals and sequence of the deleted allele. (C) Diagram showing the strategy for *in vivo* inversion of *Myh1* and *Myh4* by CRISPR/Cas9. (D) PCR used for the screening of mutant animals and sequence of the inverted allele. (E) Immunostaining against MYH2, MYH4 and MYH7 proteins of adult distal hindlimb sections of 2-3-month-old WT, *Myh(1-4)*^{Del/+}, *Myh(1-4)*^{Inv/+}, and of *Myh(1-4)*^{Inv3/+} animals. (F) Quantification of *Myh2*, *Myh1*, *Myh4*, *Myh8* and *Myh13* mRNAs by RT-qPCR from adult Soleus of WT, *Myh(1-4)*^{Del/+}, and *Myh(1-4)*^{Del/Del} animals. (G) Quantification of *Myh2*, *Myh1*, *Myh4*, *Myh8* and *Myh13* mRNAs by RT-qPCR from adult Soleus of WT, *Myh(1-4)*^{Inv/+} and *Myh(1-4)*^{Inv/Inv} animals. (H) Quantification of *Myh2*, *Myh1*, *Myh4*, *Myh8* and *Myh13* mRNAs by RT-qPCR from adult Soleus of WT, *Myh(1-4)*^{Inv3/+} and of *Myh(1-4)*^{Inv3/Inv3'} animals. (I) Quantification of *Myh4* premRNA from adult TA of WT, *Myh(1-4)*^{Inv/Inv} and of *Myh(1-4)*^{Inv3/Inv3'} animals. (J) Quantification of *Myh1* premRNA from adult TA of WT, *Myh(1-4)*^{Inv/Inv} and of *Myh(1-4)*^{Inv3/Inv3'} animals. Numerical data are presented as mean \pm S.E.M. **P* < 0.05, ***P* < 0.01, ****P* < 0.001

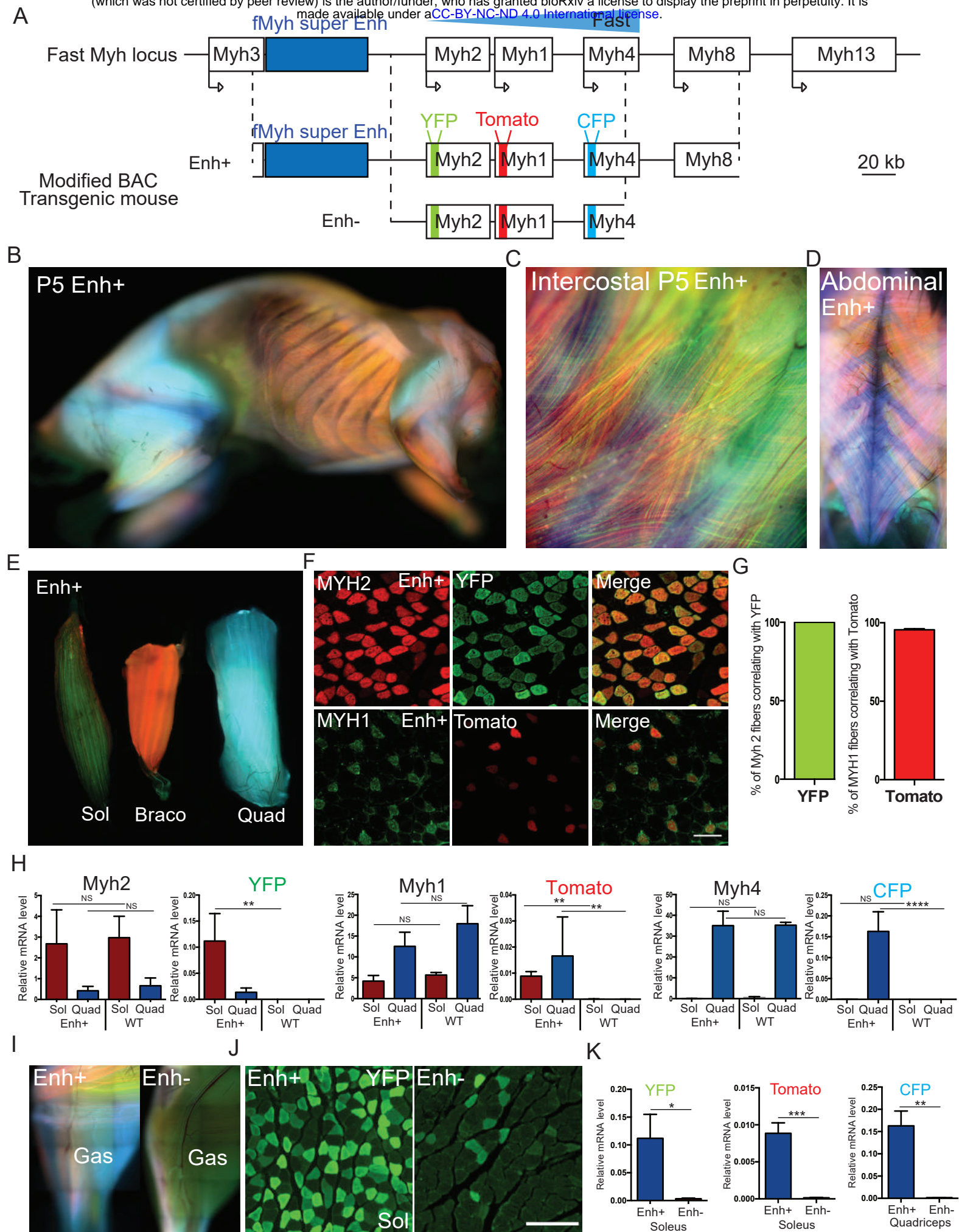
Table S1. Antibodies used in the study.

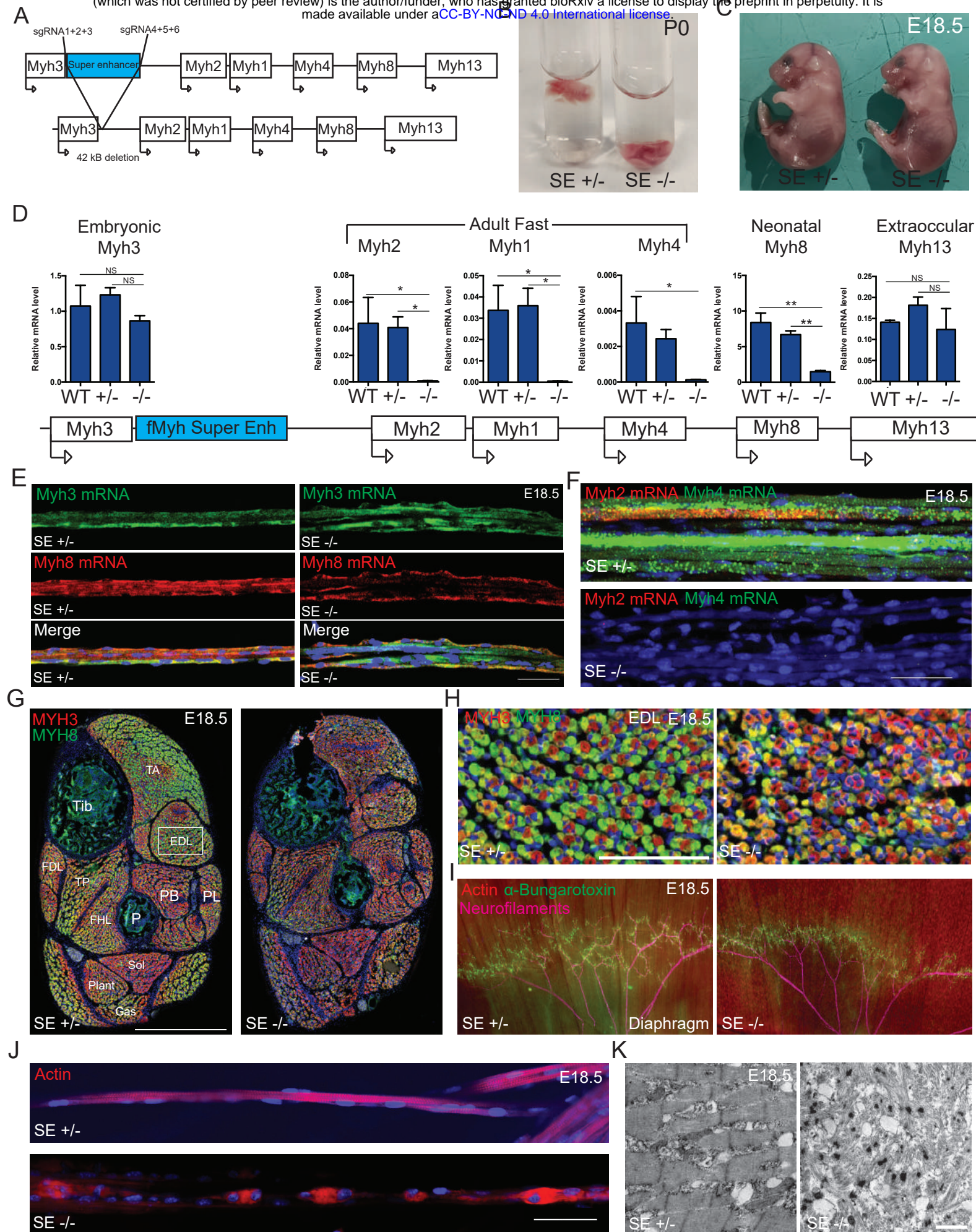
Table S2. Oligonucleotides used for RT-qPCR experiments.

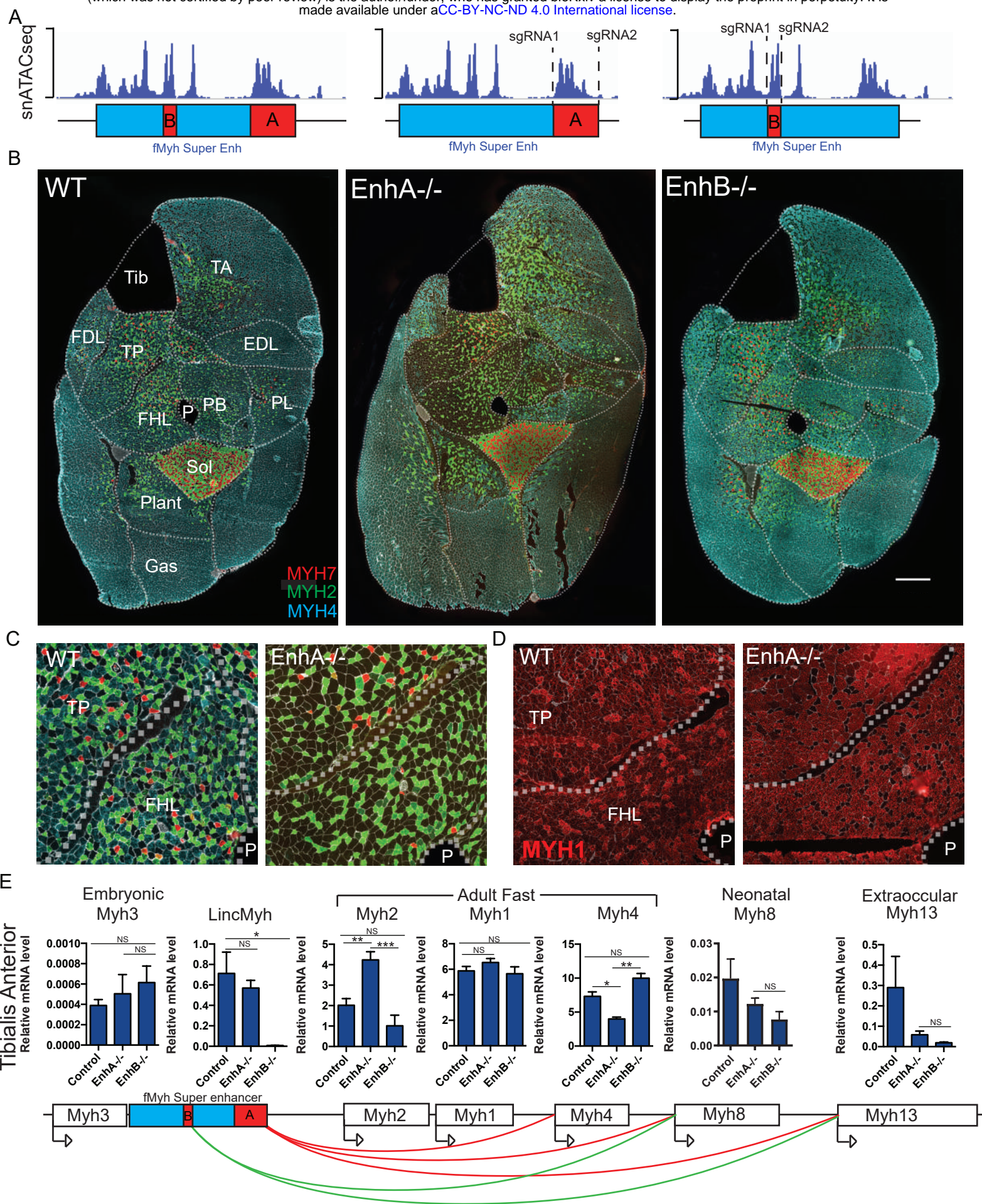
Table S3. Oligonucleotides used for 4C-seq experiments.

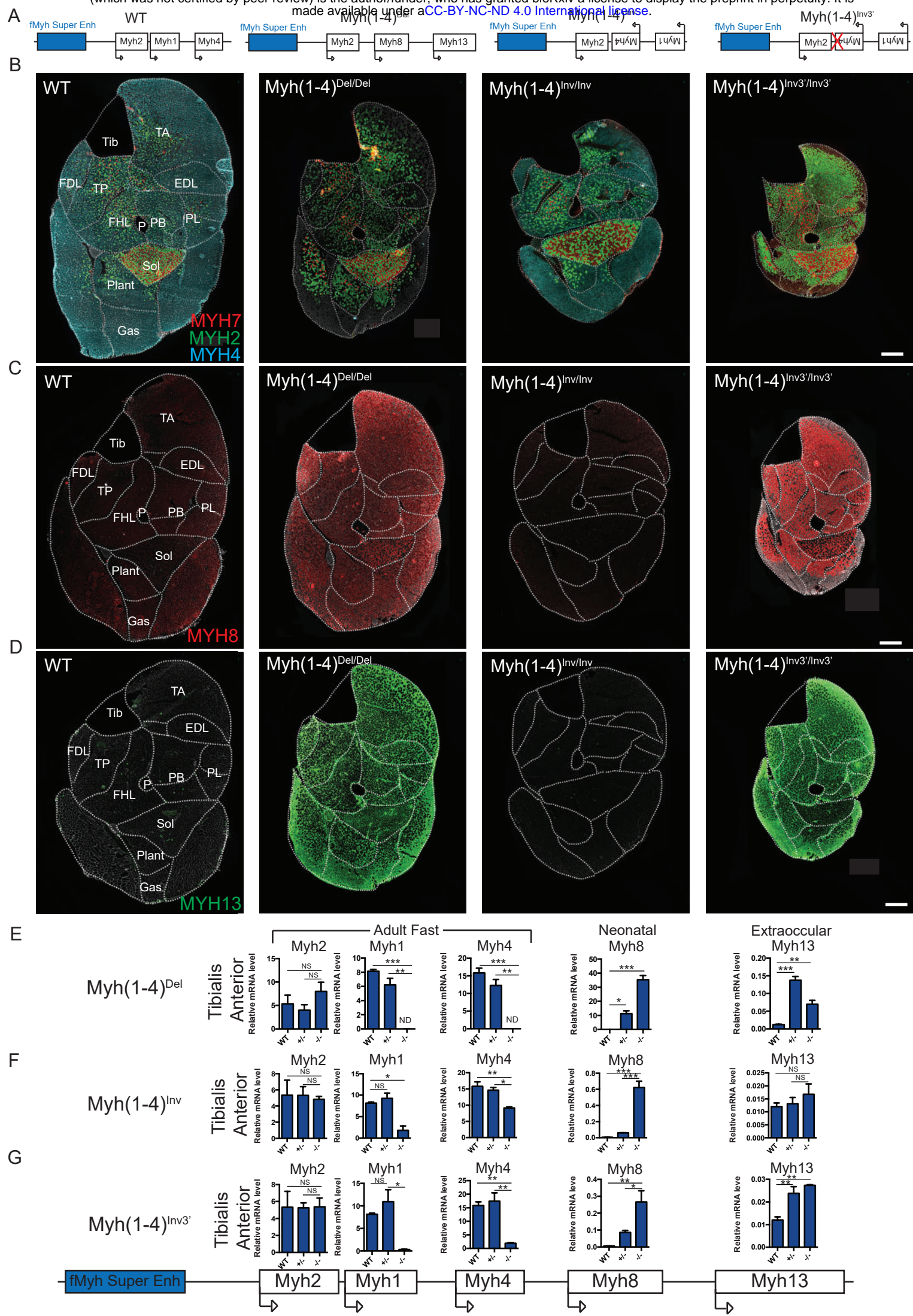
Table S4. sgRNA and PCR primers used for screening of CRISPR/Cas9 edited genome.

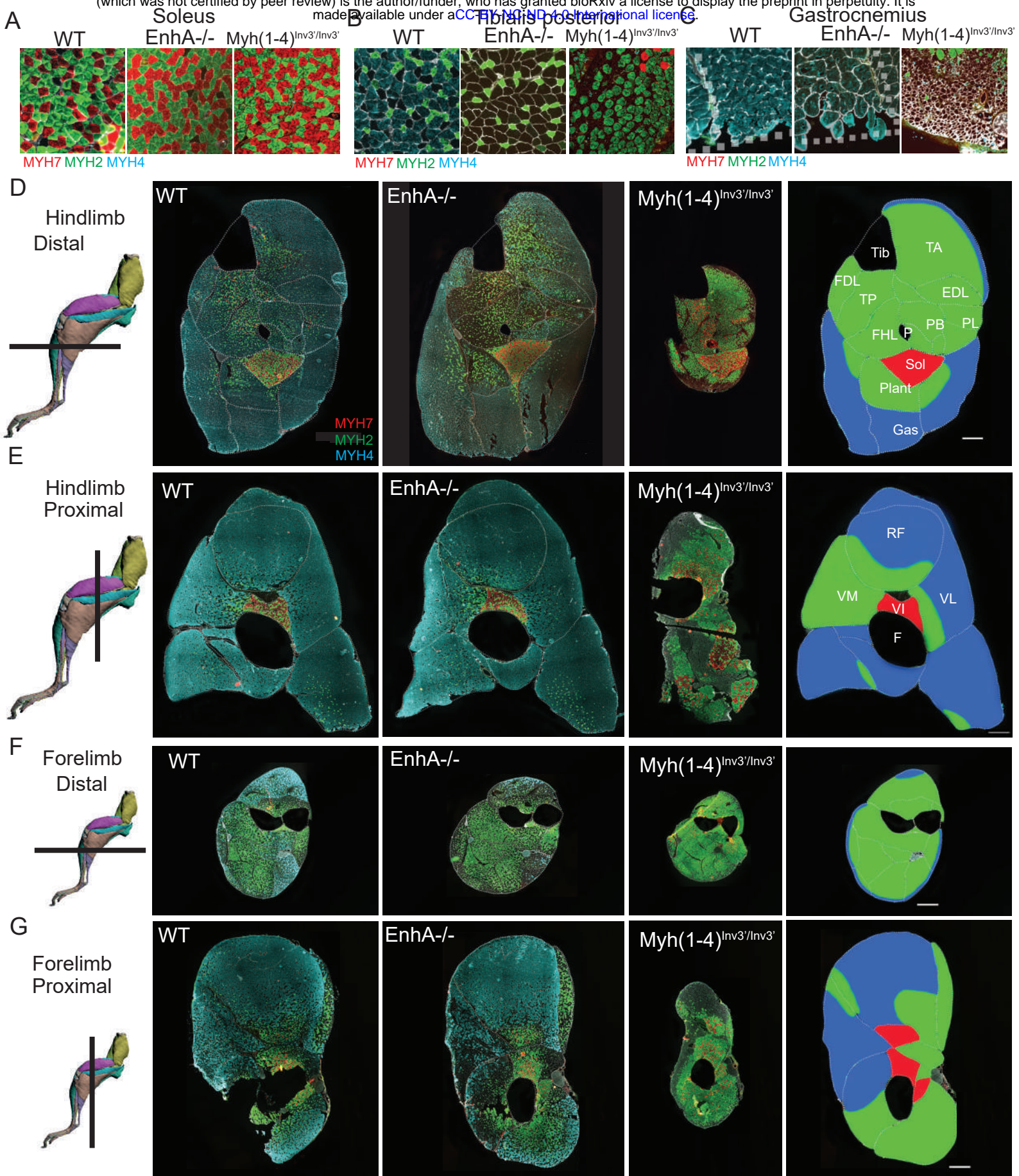




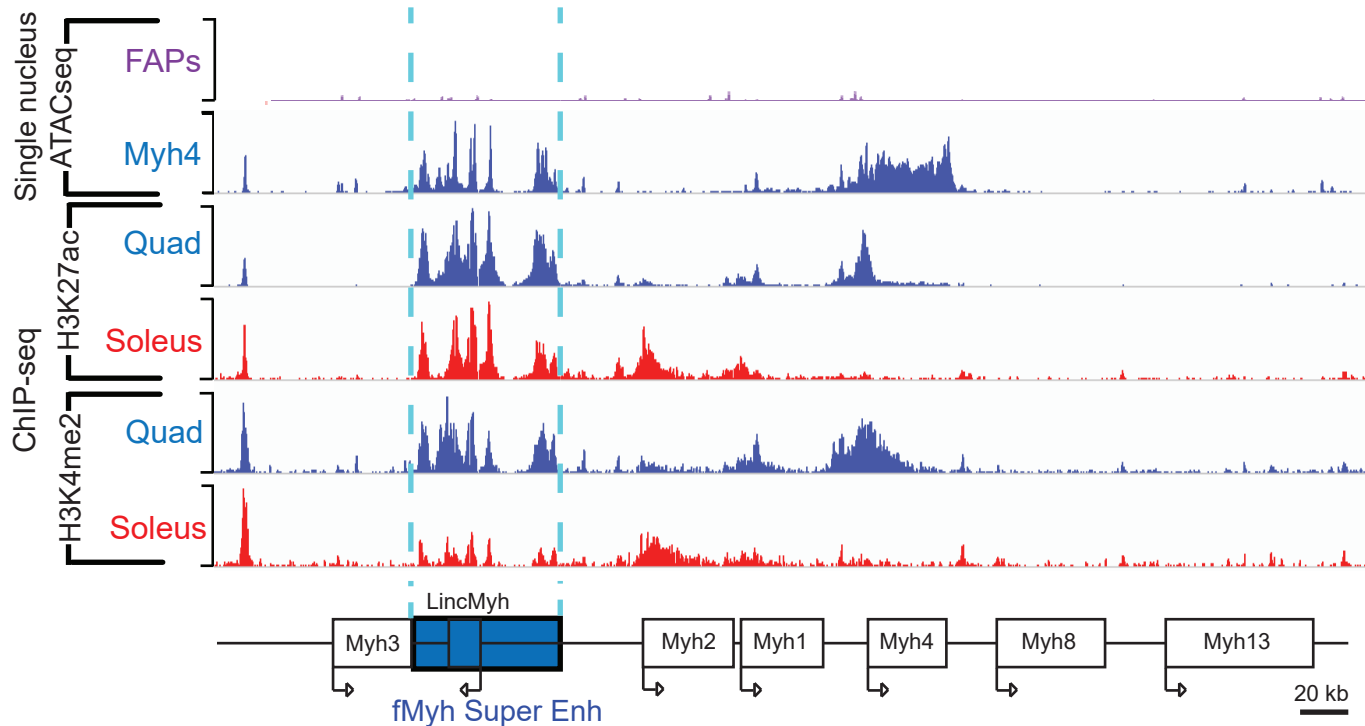




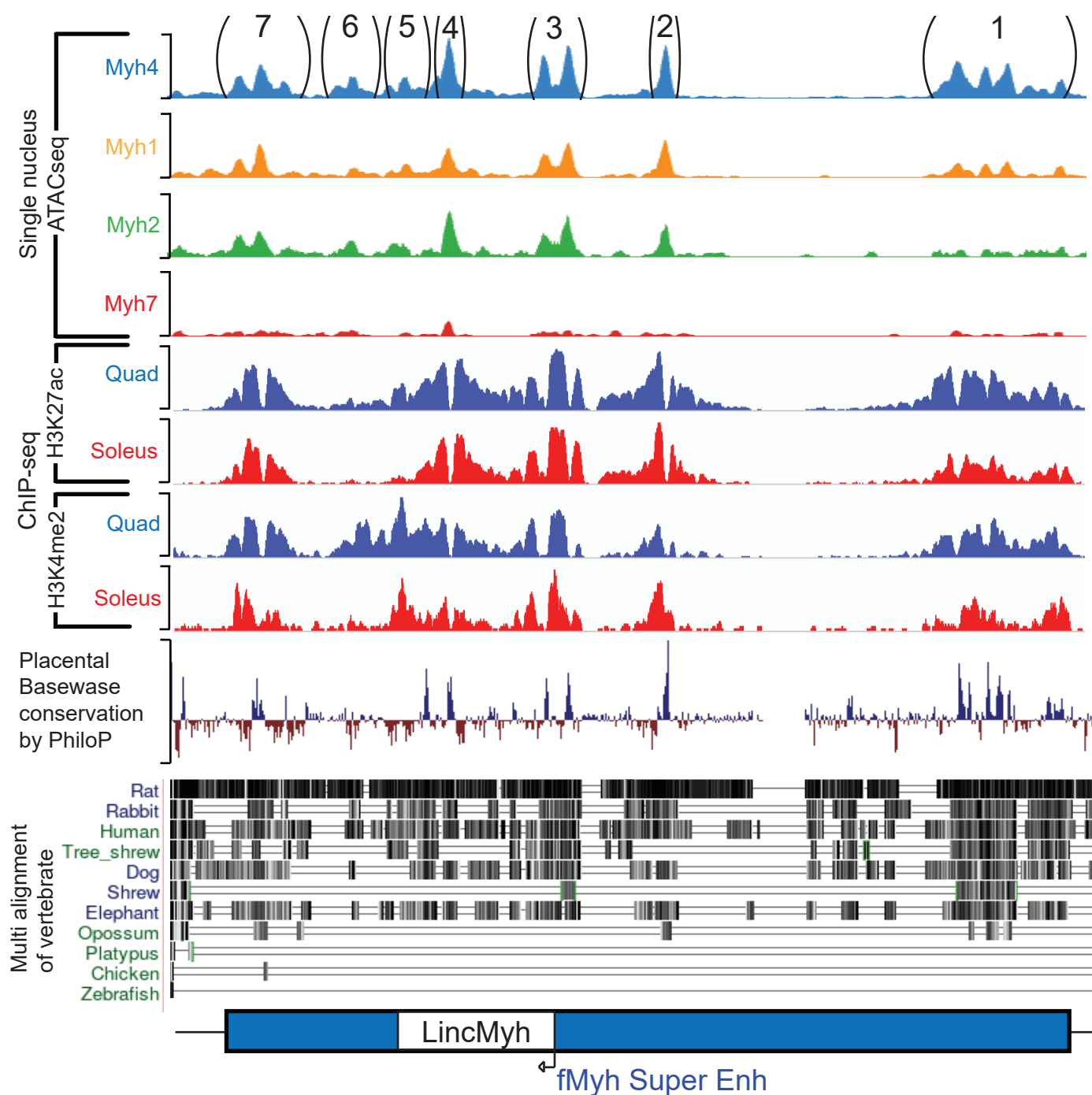




A

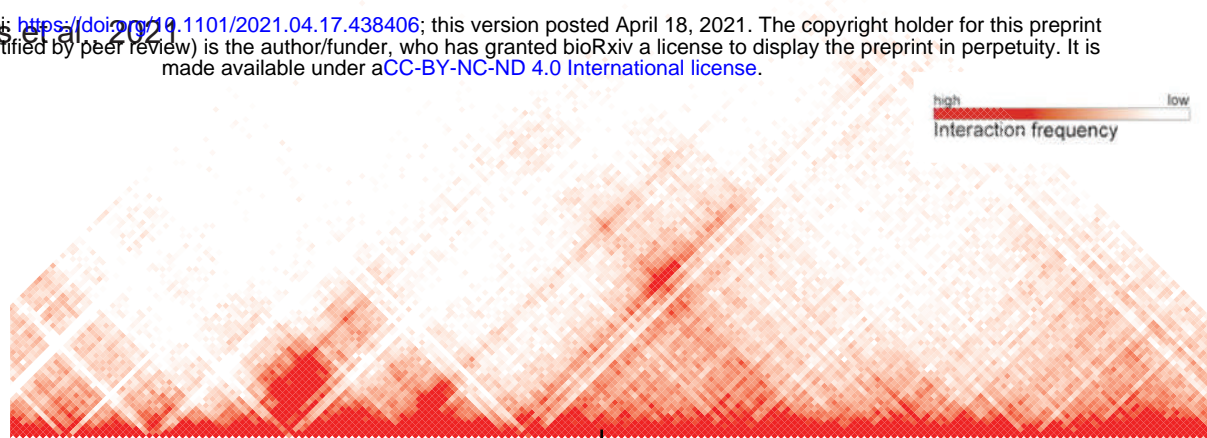


B



A

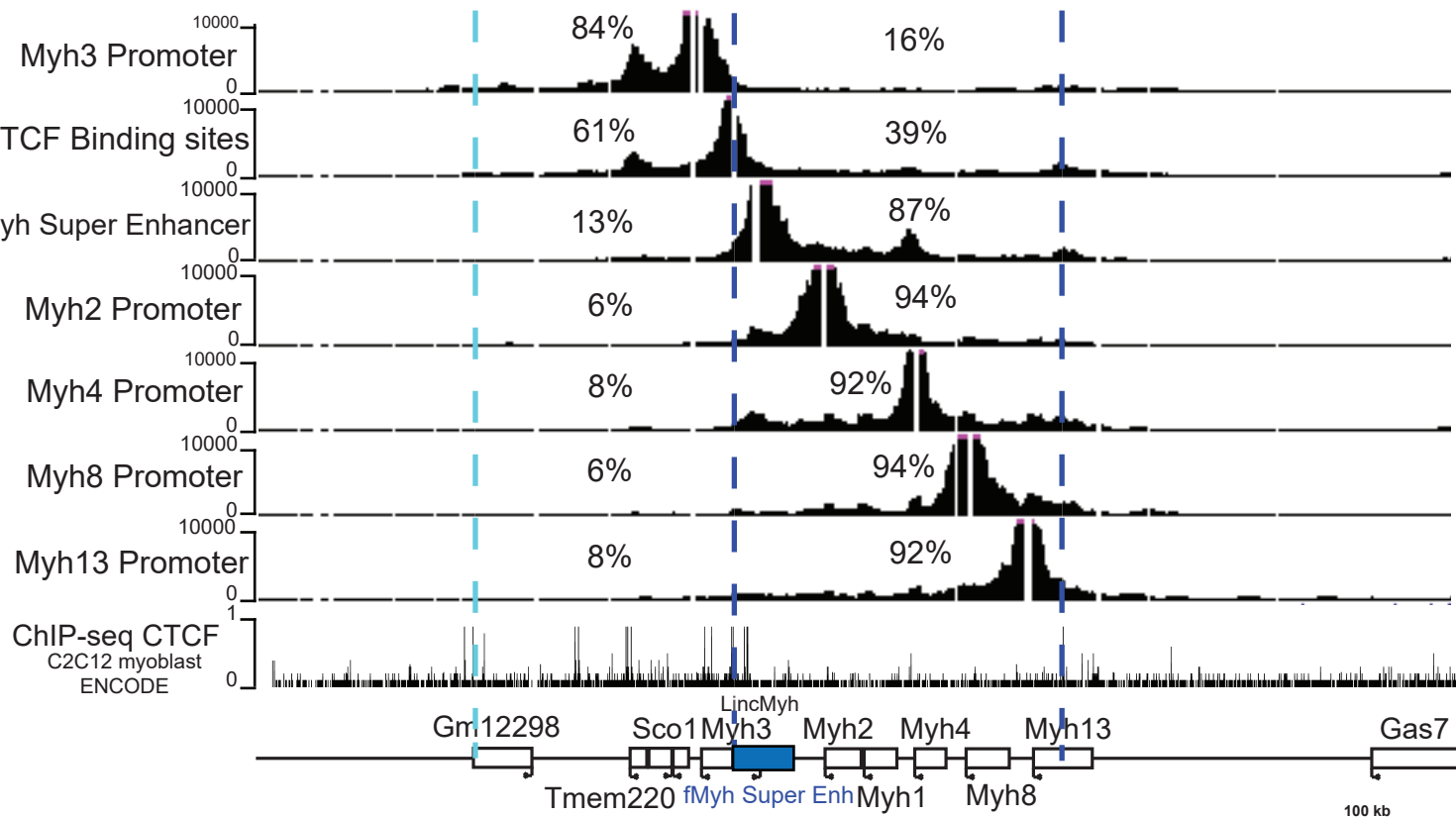
Hi-C mES cells
 Bonev et al., 2017



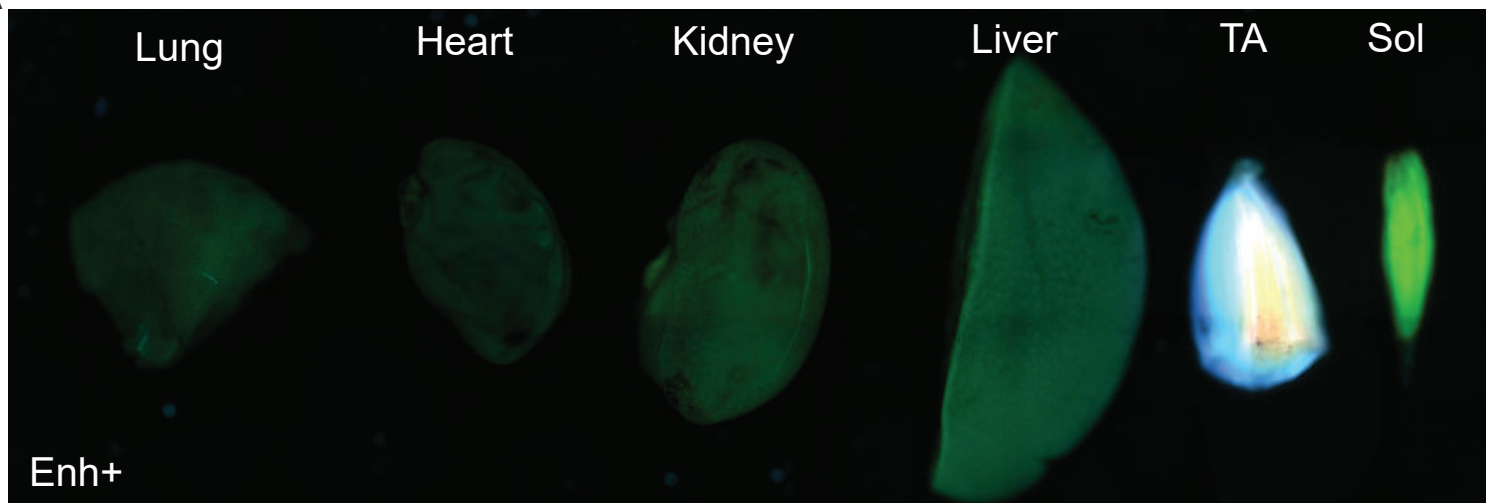
TADs -2 -1 *Fast Myh TAD* +1

B

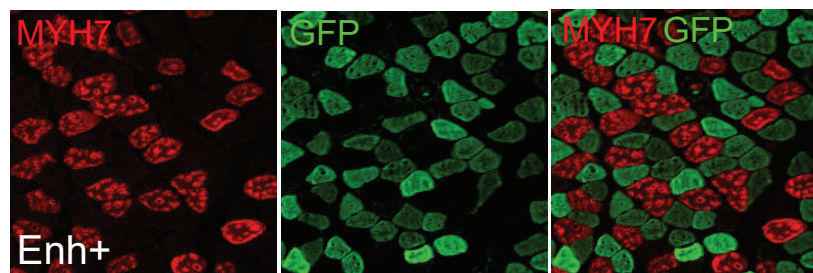
4C-seq Fast skeletal muscle



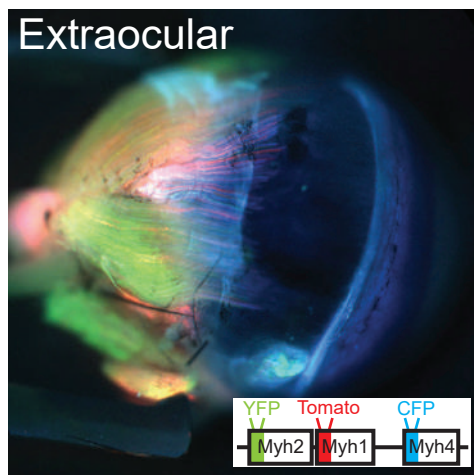
A



B



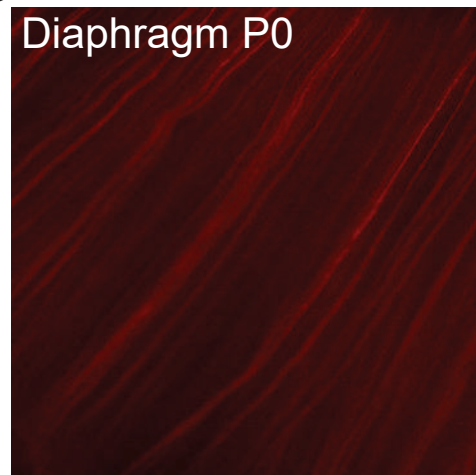
A



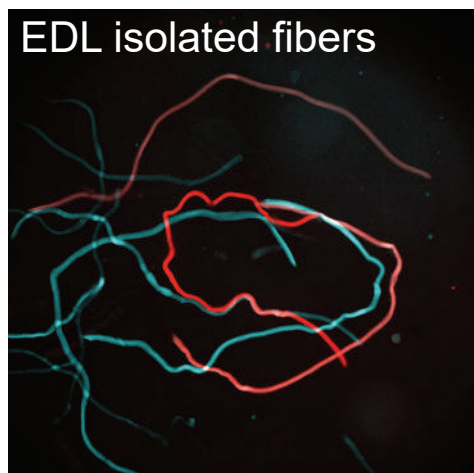
B



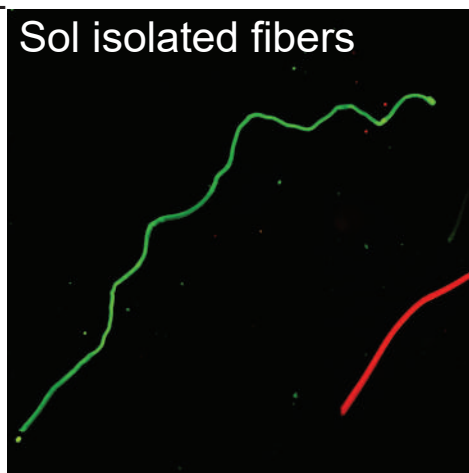
C



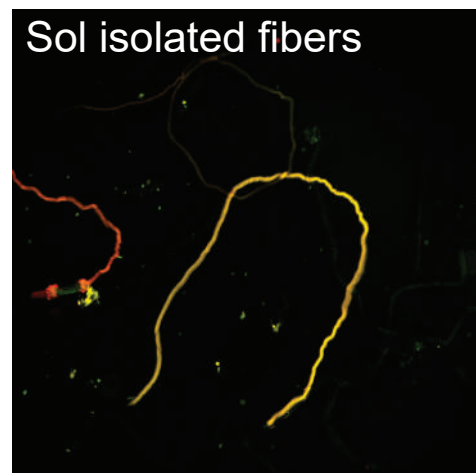
D



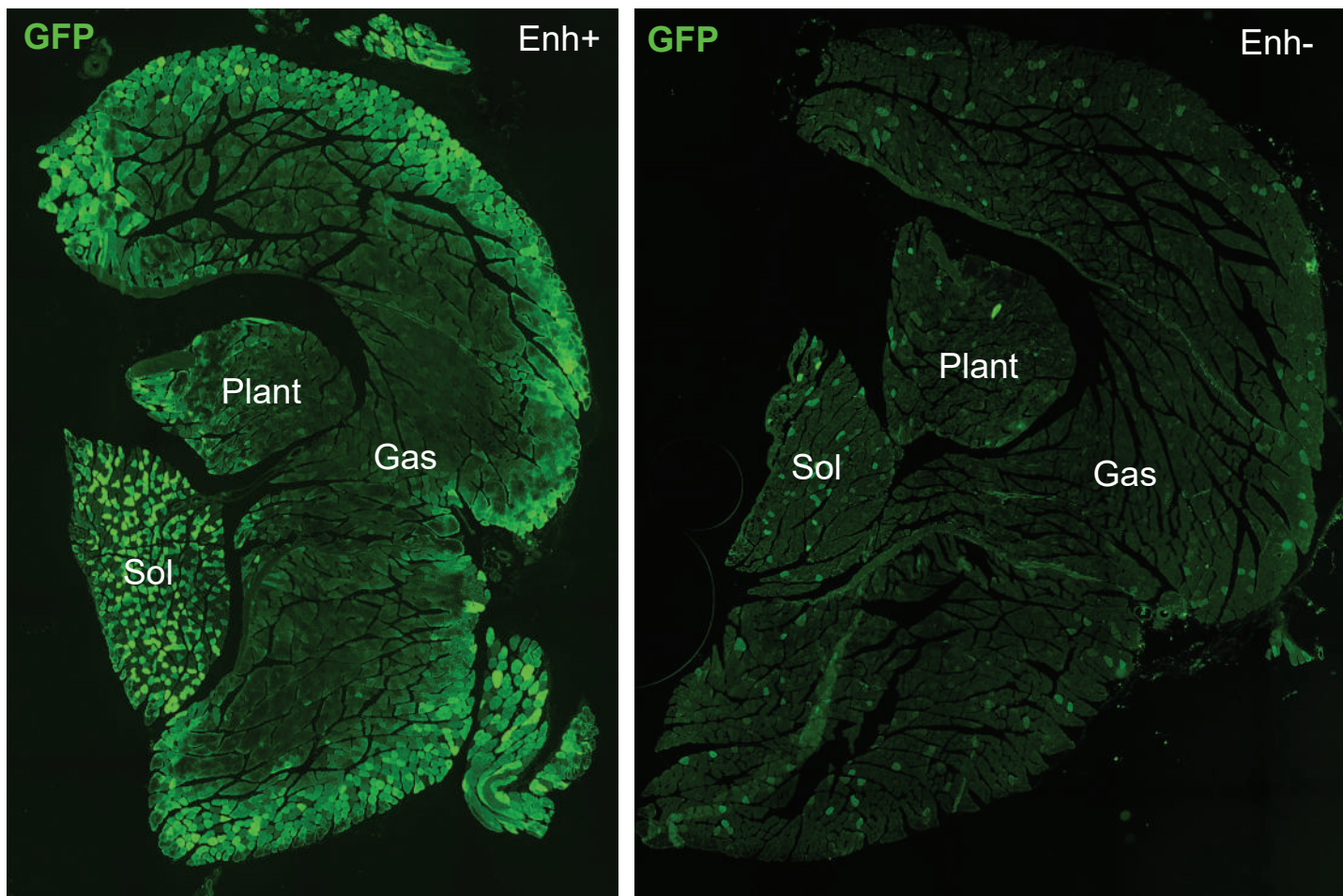
E

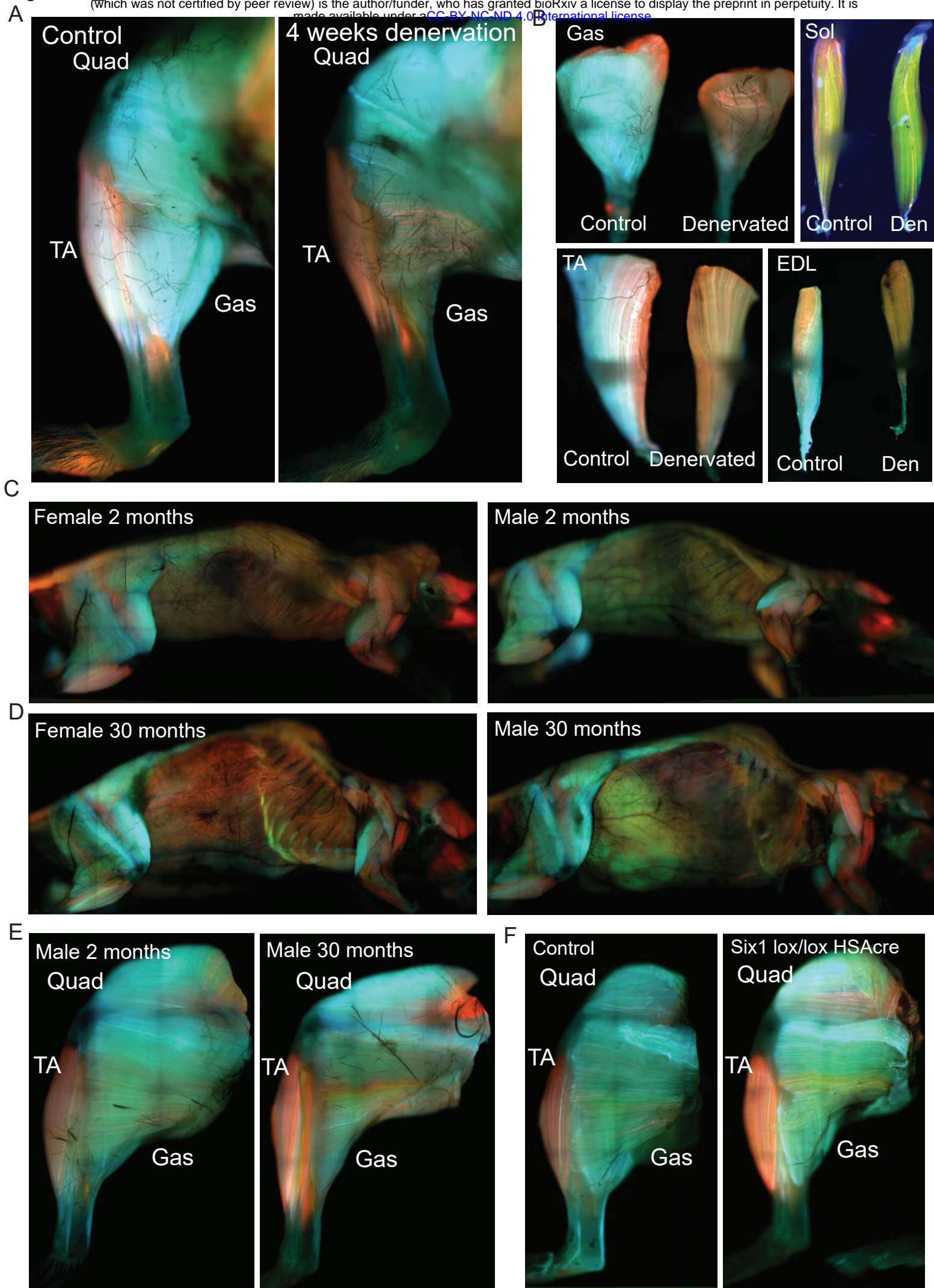


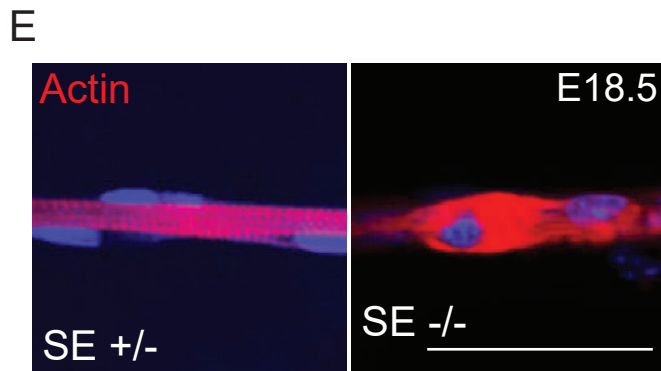
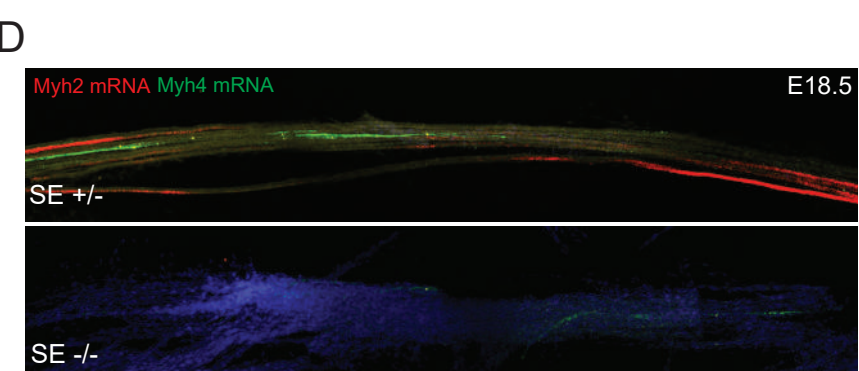
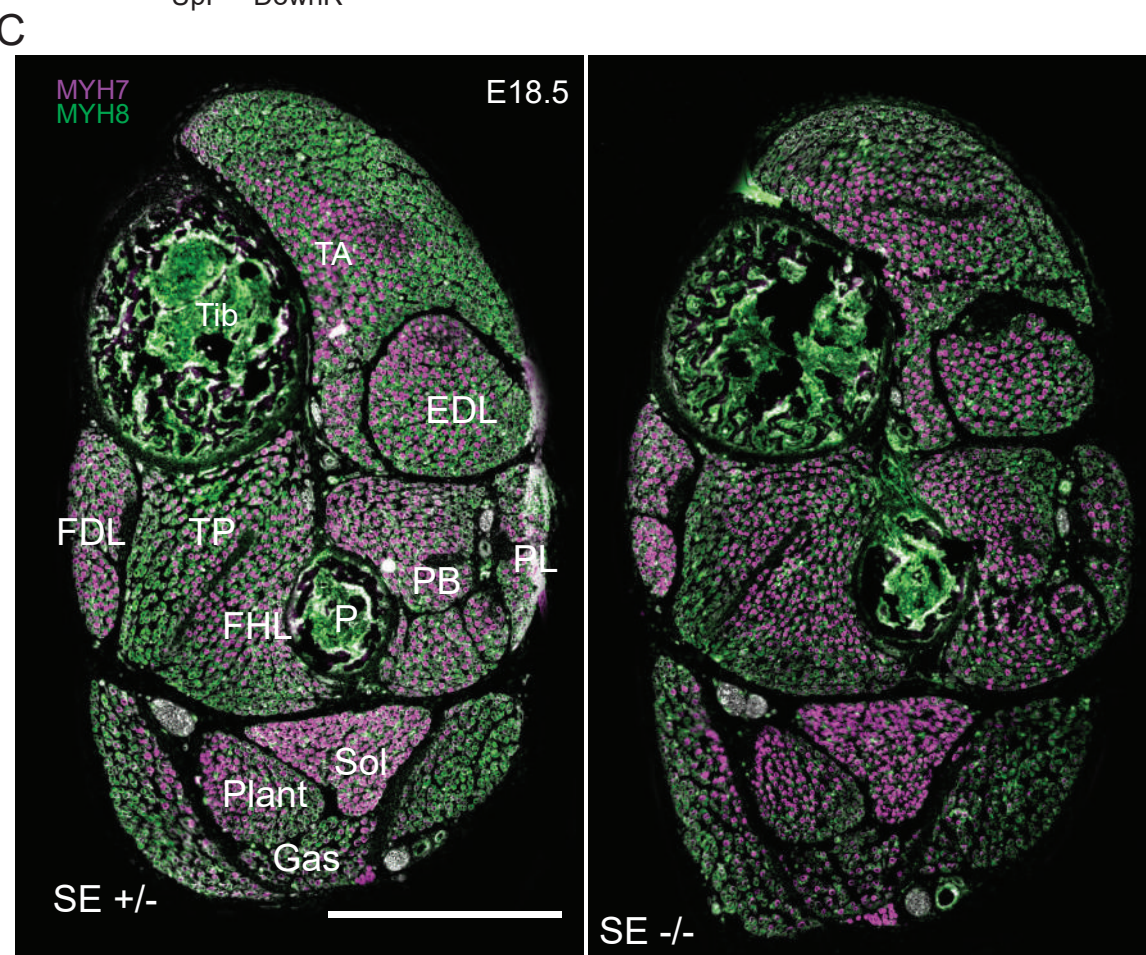
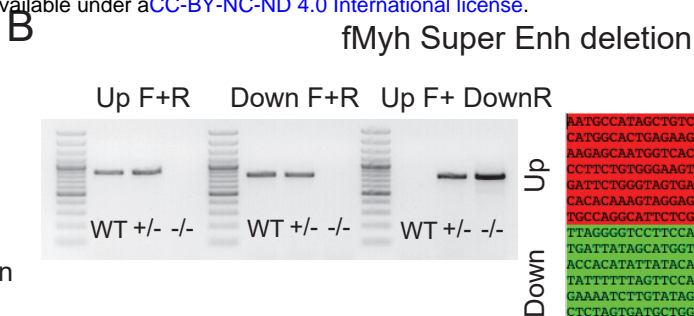
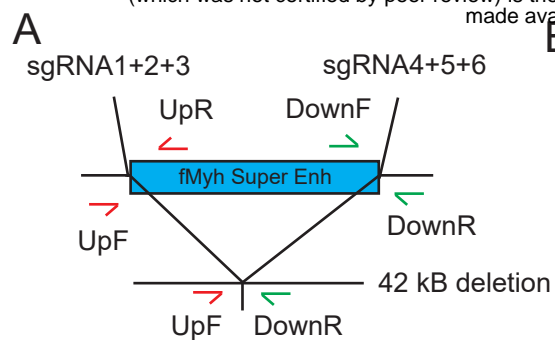
F

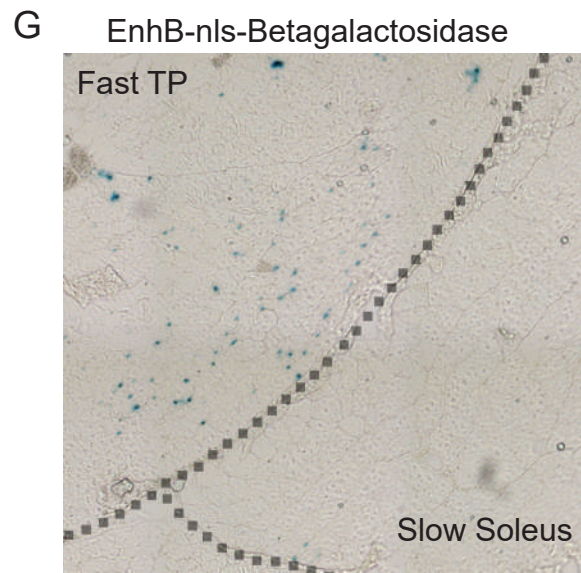
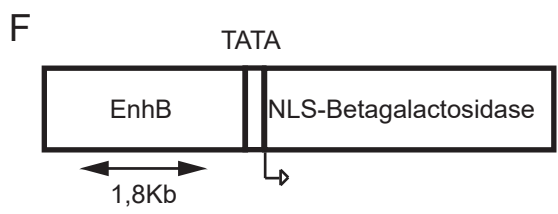
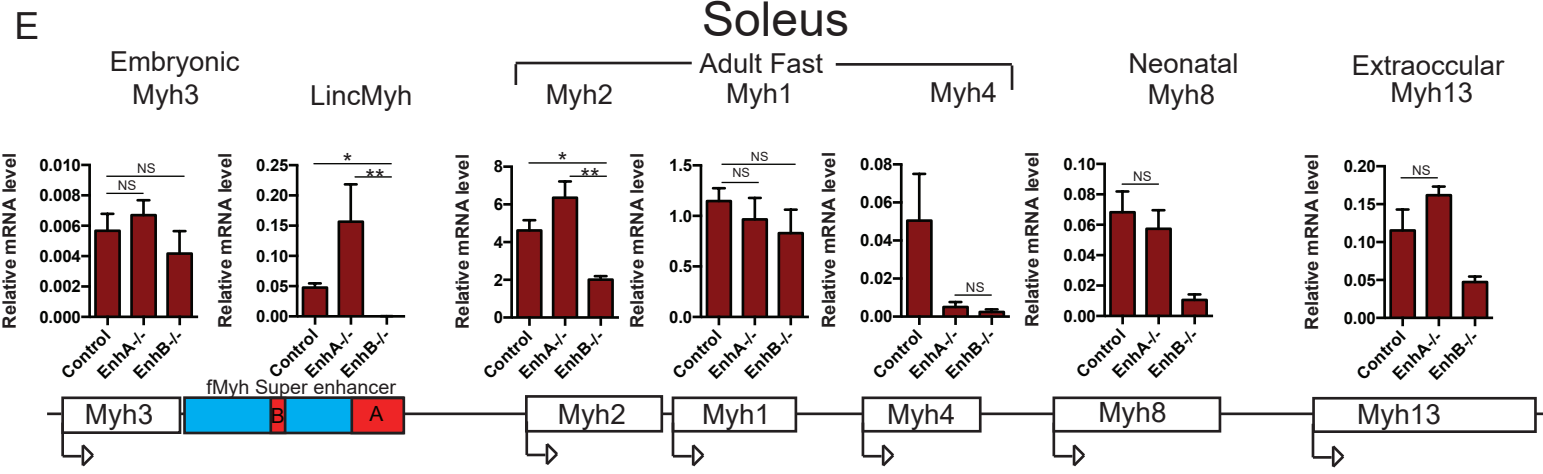
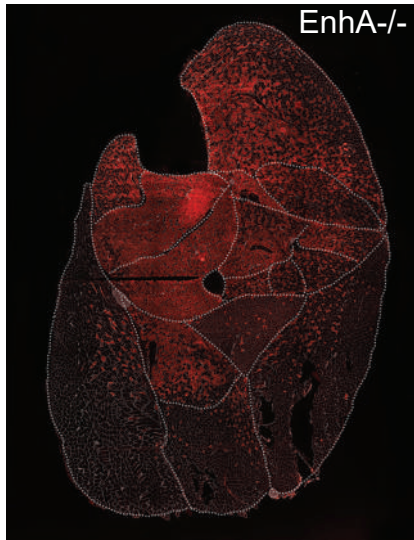
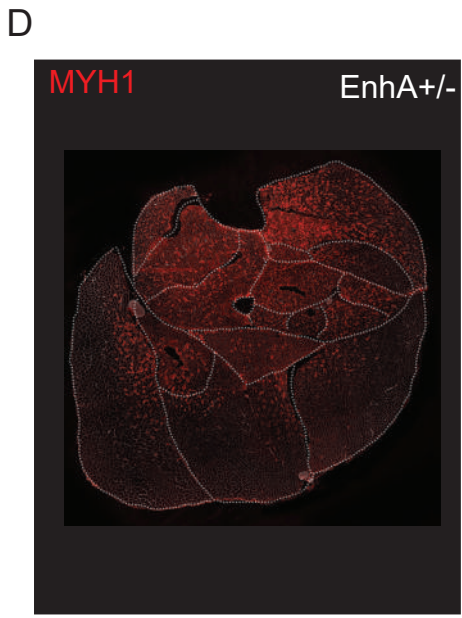
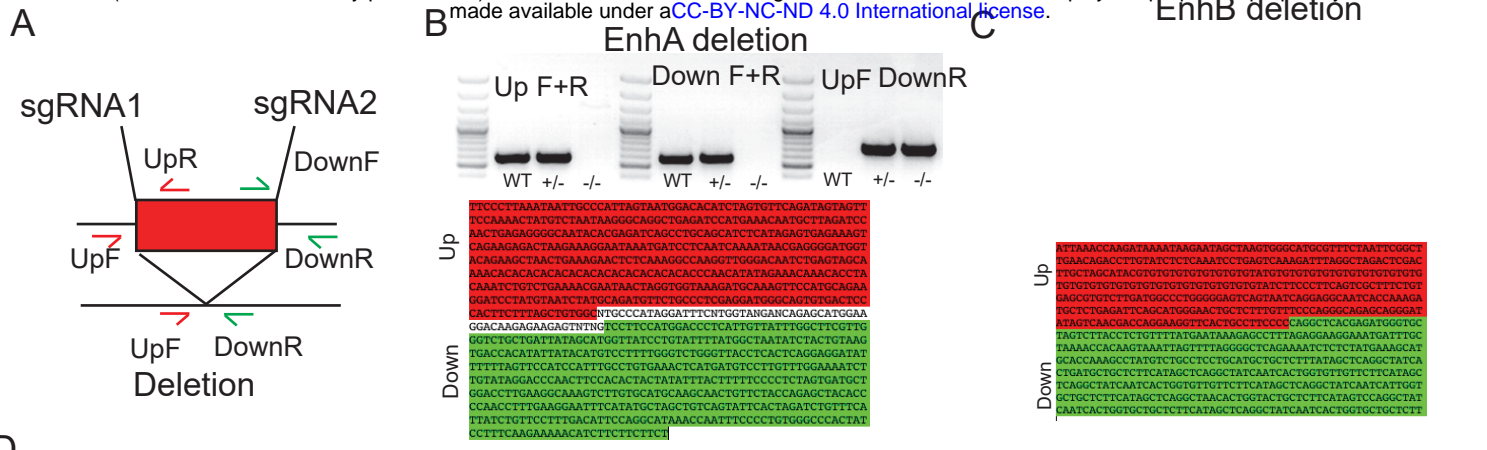


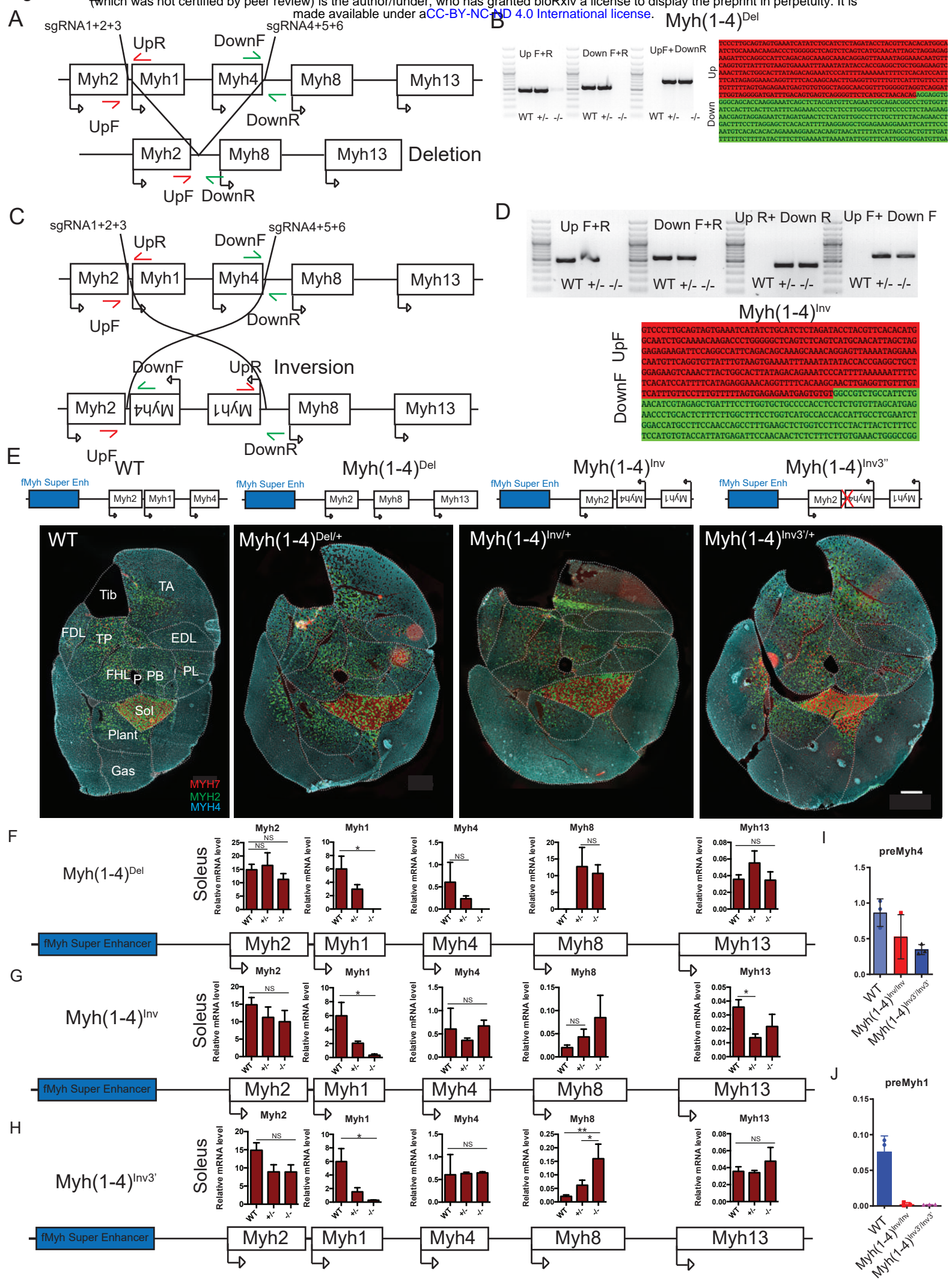
A











Dos Santos et al. Supplementary table 1.

Target	Antibody reference	Supplier	Species	dilution
MYH7	BA-F8	DHSB	mouse IgG2B	1/40
MYH2	SC-71	DHSB	mouse IgG1	1/200
MYH1	6H1	DHSB	mouse IgM	1/40
MYH4	BF-F3	DHSB	mouse IgM	1/200
MYH8	N3.36	DHSB	mouse IgM	1/200
MYH13	4A6	DHSB	mouse IgM	1/200
MYH3	Bf-45	DHSB	mouse IgG1	1/200
Laminin	L9393	Sigma	Rat	1/500
GFP	ab290	Abcam	Chicken	1/200
Neurofilament	841001	BioLegend	Rabbit	1/1000
alpha bungarotoxin	B13422	Thermofischer		1/500

Dos Santos et al. Supplementary table 2.

	Forward	reverse
Myh7	5'CTCAAGCTGCTCAGCAATCTATTT	5'GGAGCGCAAGTTTGTGATAAGT
Myh2	5'CGG GAG TCT TGG TTT CAT TG	5'CCAAGA AAG GTG CCAAGA AG
Myh1	5'CAG GAG TCT TGG TTT CAT T	5'CGG TGG TGG AAA GAA AGG
Myh4	5'GCT TGA AAA CGA GGT GGA AA	5'CCT CCT CAG CCT GTC TCT TG
Myh8	5'CAGGAGCAGGAATGATGCTCTGAG	5'AGTTCCTCAAACCTTCAGCAGCCAA
Myh13	5'GAAGCTCCTGAACTCCATCG	5'GGTCACCAGCTTCTCGTCTC
LincMyh	5'GTG CAG CCA GAA CAA GAC AG	5'CAA GAT GGG AGG CTC TCA AA
YFP	5'TTT-CCC-AGC-TGC-ACC-TTC-TC	5'GCT-GAA-CTT-GTG-GCC-GTT-TA
Tomato	5'AGT-TCA-TGC-GCT-TCA-AGG-TG	5'TGG-AGC-CGT-ACA-TGA-ACT-GG
CFP	5'TTG-TCT-GAC-TCA-AGC-CTG-CC	5'CCG-GTG-GTG-CAG-ATG-AAC-TT

Dos Santos et al. Supplementary table 3.

4C_My3_s2_iF1+adap	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCAGGGGCTCCTAAAGCAAAGGA
4C_My3_s2_iR1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACAGGGAGCTATGCCACCAG
4C_CTCF_s1_iF1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCGAGAATCTCTGACTGGGAAATA
4C_CTCF_s1_iR1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTCTCTCTAGATGGTGGTCC
4C_Mf6_s2_iF1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCCAATCGCAGCCCCCTTGCT
4C_Mf6_s2_iR1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACCCAAGAGGTGAGTGACAGT
4C_DHS_s1_iF1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCCAGAACTGAACACCACC
4C_DHS_s1_iR1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGCAGAACTGAACACCAC
4C_MYH2_s1_iF1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTCCTATGGATAATTTGAGAAGAT
4C_MYH2_s1_iR1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGAGGCGTATGGAGAAT
4C_MYH4_s2_iF1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCAACTATGAAGGCTGGGCT
4C_MYH4_s2_iR1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGTGAAGAGGAGCAGTC
4C_MYH1_s1_2_iF1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCAGTGTCAACATCCAGAGT
4C_MYH1_s1_2_iR1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGAGACTTAAGCATTGATAGGT
4C_MYH13_s3_iF1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCGAGGAAGTGAGCTGTTACG
4C_MYH13_s3_iR1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTTCTACAATGCACCCC

Dos Santos et al. Supplementary table 4.

Super Enhancer deletion		
sgRNA Up1	GGGGACTCCTATGTAAGCAA	
sgRNA Up2	TCTAGGAGGCCGAGAATGCC	
sgRNA Up3	TTGGGAGCAGTCCCATCTGG	
sgRNA Down1	CCATGGAAGGACCCCTAATT	
sgRNA Down2	CAGACAAGGCCTTCAATTAG	
sgRNA Down3	TCTCTGACATCAAGCCATGC	
PCR primer	forward	reverse
super_enh_up	TTGCGGCTATGGAAGGAAGG	GAGTTCAGTCCCCAAGGCAA
super_enh_down	GGGTTCCCTGCCTCTGTACC	TCCAGCATCACTAGAGGGGA
EnhA deletion		
sgRNA Up1	GAGAAGAGTCTCGTCATTGTGG	
sgRNA Up2	GTCACACTGCCCATCCTCGAGGG	
sgRNA Up3	GATGTTCTGCCCTCGAGGATGGG	
sgRNA Down1	GATGGAAGGACCCCTAATTAGG	
sgRNA Down2	GCAGACAAGGCCTTCAATTAGGGG	
sgRNA Down3	GTCTCTGACATCAAGCCATGCAGG	
PCR primer	forward	reverse
EnhA_up	CACACACACACACACCCCAAC	TTCCCATCACCTGAGTCTCC
EnhA_down	CCCCATTTACCTTCAGAGA	CCTTCAAGGTCCAGCATCAC
EnhB deletion		
sgRNA Up1	GTGACTATCCAGAAGTCTGG	
sgRNA Down1	CCTTTCAGGCTCACGAGATGG	
PCR primer	forward	reverse
EnhB_up	GCTAAGTGGGCATGCGTTTC	CAAAGCAGGACCCAAACAGC
EnhB_down	TCCCGTTCCCTTGATCCCTCT	GGCTTTGGTGCATGCTTTCA
Myh1-4 deletion and inversion		
sgRNA Up1	ATGAGTGTGTGGCTAGGCAACGG	
sgRNA Up2	GTGGCTAGGCAACGGTTTGGGGG	
sgRNA Up3	ATGATTTGACAGTGAGTCAGAGG	
sgRNA Down1	GGGTTCTCATGCTAACACAGAGG	
sgRNA Down2	AAGTGAAGTGGATAACCACAGGG	
sgRNA Down3	CAGAATGGCAGACGGCCCTGTGG	
PCR primer	forward	reverse
Myh1_up	GGCCATTCAGACAGCAAAGC	GCTCCTCAGTGTAGCTGCAA
Myh4_down	GGACCAGAGCTTCAAAGGCT	CTCAAAGTGGCCTTCCCGAT