

Sahinyan et al. 2021 manuscript

Application of ATAC-Seq for genome-wide analysis of the chromatin state at single myofiber resolution

Korin Sahinyan^{1,2*}, Darren M. Blackburn^{1,2*}, Marie-Michelle Simon^{1,3*}, Felicia Lazure^{1,2}, Tony Kwan^{1,3}, Guillaume Bourque^{1,3,4}, Vahab D. Soleimani^{1,2,5}

¹ Department of Human Genetics, McGill University, 3640 rue University, Montréal, QC, H3A 0C7, Canada.

² Lady Davis Institute for Medical Research, Jewish General Hospital, 3755 Chemin de la Côte-Sainte-Catherine, Montréal QC, H3T 1E2, Canada.

³ McGill University and Genome Québec Innovation Centre, Montréal, QC, Canada.

⁴ Canadian Centre for Computational Genomics, Montréal, QC, Canada; Institute for the Advanced Study of Human Biology, Kyoto University, Kyoto, Japan

⁵ Corresponding author: Tel: +1 514 340 8222 ext. 26136 Fax: +1 514 340 7502
vahab.soleimani@mcgill.ca

* These authors have contributed equally.

ABSTRACT

Skeletal myofibers are the main components of skeletal muscle which is the largest tissue in the body. Myofibers are highly adaptive in nature and they can vary in different biological and disease conditions. Therefore, transcriptional and epigenetic studies on myofibers are crucial to discover how chromatin alterations occur in the skeletal muscle under different conditions. However, due to the heterogenous nature of skeletal muscle, studying myofibers in isolation proves to be a challenging task. Single cell sequencing has permitted for the study of the epigenome of isolated myonuclei. While this provides sequencing with high dimensionality, the sequencing depth is lacking, which makes comparisons between different biological conditions difficult. Here we report the first implementation of single myofiber ATAC-Seq, which permits for the sequencing of an individual myofiber at a depth sufficient for peak calling and for comparative analysis of chromatin accessibility under various physiological, physical and disease conditions. Application of this technique revealed significant differences in chromatin accessibility between resting and regenerating myofibers. This technique can lead to wide application in identifying chromatin regulatory elements and epigenetic mechanisms in muscle fibers during development and in muscle-wasting diseases.

INTRODUCTION

Skeletal muscle evolved for contraction and production of force. The main component of skeletal muscle are myofibers which are formed from the fusion of myogenic precursor cells (1) and myofibers are large postmitotic syncytia that are composed of repeating contractile units, called sarcomeres (2). Myofibers exhibit wide variations in their metabolic activity and contractile properties (3). In addition, they have a highly adaptive nature where their size, and the overall skeletal muscle mass, are regulated by complex processes involving rates of protein turnover (4,5) transcriptional (6) and posttranscriptional (7) control of gene expression. Due to their adaptive nature, myofibers can change in response to exercise (3,8,9), aging (10) and diseases, such as sarcopenia (11,12) and cachexia (13). Therefore, the study of the myofiber transcriptome/epigenome can provide key insights into how skeletal muscle adapts and changes under different conditions, and it can potentially lead to discovery of novel therapeutic venues for muscle related diseases.

Myofibers also act as a key signaling component of muscle stem cells (MuSCs) (14), which are in turn required for the regeneration of nascent muscle fiber after injury. Skeletal muscle is a very heterogenous tissue composed not only of myofibers and their associated MuSCs, but also numerous different non-myogenic cell types (15). Previous studies using whole muscle sequencing captures not only the myofibers but also the other resident cell types in the muscle, making it challenging to attribute changes

in transcriptome and epigenome specifically to myofibers as they could be due to changes in other cell types. Recent advances in Next Generation Sequencing (NGS) now allows for high dimensional analysis at a single cell level. Recent studies using these technologies to study muscle tissue, such as single nucleus RNA-Seq and ATAC-Seq have analyzed the transcriptome and epigenome of the myonuclei within the muscle fiber (16-18). However, they present certain limitations where they sequence all myonuclei present in the muscle and cannot distinguish between different myofibers as well as having low sequencing depth with limited capacity of analysis.

Chromatin states play key roles in transcriptional regulation and determination of cellular identity (19). Although the accessible regions make up only 3% of the total genome, it represents over 90% of known transcription factor binding sites (20). Differential chromatin accessibility is a determinant of transcriptional gene regulation of different cell types and changes in chromatin accessibility have been shown in different biological and disease conditions such as during development (21,22), cancers (23,24) and neurological disorders (25,26) among many others. Thus, in recent years, the study of epigenetics and chromatin accessibility has become a promising field for the development of novel therapeutics. Today, ATAC-Seq is the widely used method that allows for the successful mapping of the accessible chromatin regions in the genome. ATAC-Seq relies on the hyperactive Tn5 transposase that fragments the accessible regions in the genome while simultaneously ligating sequencing compatible adaptors (27,28). Over the years, ATAC-Seq has been applied to many different cell types and tissues (29-31). However, to our knowledge, it has not been performed on a single myofiber, possibly due to the rigidity of their membrane, high levels of mitochondria (32-34) and the low number of myonuclei that is present in a single myofiber (35,36).

Here, we have adapted OMNI-ATAC-seq to determine genome-wide chromatin accessibility of myonuclei contained within a single Extensor Digitorum Longus (EDL) muscle fiber in mouse. The Single myofiber ATAC-Seq (smfATAC-Seq) method that we applied in this study allows for the investigation of the accessible chromatin state of a single myofiber, without the presence of other confounding cell types, with high sequencing depth. Through this method we provide comparative analysis of chromatin accessibility between resting and regenerating myofibers as well as their MuSC progenitors. This method can be used in the future to profile the epigenetic state of myofibers in diseases and under various physiological and physical stimuli and to identify active cis-regulatory elements in muscle fibers.

RESULTS

Generation of ATAC-Seq libraries from a single myofiber

A single EDL myofiber of a mouse contains an average of 200-300 myonuclei (35,36), making genome-wide analyses of chromatin state difficult. With the advancements in next generation sequencing (NGS) and the development of the OMNI ATAC-Seq protocol (27), analysis of chromatin accessibility of samples with an input of as low as 500 cells is now possible. However, myofibers present additional challenges with their rigid membrane and high levels of mitochondria (32-34). Here we report a robust protocol for the successful application of ATAC-Seq on a single myofiber from the EDL muscle. Our method relies on the lysis and permeabilization of a single myofiber followed by transposition with a hyperactive Tn5 transposase (27) (Figure 1). DNA fragment sizes obtained from the smfATAC-Seq were of a similar range in size as those obtained from conventional OMNI ATAC-Seq that we have performed on MuSCs that were freshly sorted by Fluorescence Activated Cell Sorting (FACS) (Supplemental Figure 1). Furthermore, post sequencing analysis showed that only 0.9- 2.09 % of reads were derived from the mitochondria in smfATAC-Seq (Table 1), suggesting that this method is highly efficient for the removal of mitochondria from mitochondria-rich myofibers.

smfATAC-Seq can be used to study chromatin accessibility of myofibers under different physiological conditions

In addition to adapting the ATAC-Seq method to study the chromatin accessibility of a single myofiber, we also demonstrate the application of this technique for comparative analysis of chromatin accessibility between myofibers under different conditions. For that purpose, we performed ATAC-Seq on myofibers that were in a resting (uninjured) or regenerating (injured) state. Uninjured and injured (7 days post cardiotoxin (CTX) induced injury) myofibers were isolated from wild type (WT) C57Bl/6 mice and smfATAC-Seq was performed to compare the changes in chromatin accessibility during regeneration. In addition, we compared the chromatin accessibility between myonuclei within a single myofiber and freshly isolated MuSCs. This analysis not only identified accessible regions of chromatin in myofibers and MuSCs but it also revealed a repertoire of active cis-regulatory elements in each sample.

Apart from the myofibers and their associated MuSCs, skeletal muscle also contains many non-myogenic cells such as endothelial cells, adipocytes, hematopoietic cells, fibroblasts, fibro/adipogenic progenitors (FAPs) and macrophages (15,37,38). Our smfATAC-Seq method allows for the analysis of chromatin accessibility of a single myofiber without the confounding effect of these contaminating cell types. To illustrate the absence of the aforementioned cell types in the ATAC-Seq samples, peaks at the promoter

regions of marker genes of muscle resident cells were searched for. Specifically, Platelet and Endothelial Cell Adhesion Molecule 1 (*Pecam1*) was used to determine whether endothelial cells were present (39). Similarly, Resistin (*Retn*) and *Cd45* were used as markers for adipocytes and hematopoietic cells, respectively (40,41). The cell Surface Antigen *Thy1* was the marker selected for fibroblasts (42). In addition, Lymphocyte antigen 6a (*Ly6a*) and Adhesion G protein-coupled receptor E1 (*ADGRE1*) were selected for fibro/adipogenic progenitors (FAPs) and macrophages, respectively (43,44). None of these marker genes had ATAC-Seq peaks at their promoters, indicating that only a single myofiber is processed without any other contaminating cell types. (Supplemental Figure 2).

smfATAC-Seq may be applied to myofibers under different conditions. For instance, we applied smf-ATAC-Seq to analyze the chromatin accessibility of myofibers under resting and CTX-mediated injury conditions. In a disease condition or in injury, not all myofibers undergo damage or regenerate simultaneously. Therefore, individual myofibers within a muscle can be in different physiological and disease condition asynchronously (45). Damaged or regenerating myofibers can be visualized by their characteristic feature of centrally located nuclei (46). Injured myofibers in this method were visually selected by the presence of the centrally located myonuclei by Hoechst staining and the selected myofiber was used for downstream processing with smfATAC-Seq (Figure 2A, 2B). The selection of a specific myofiber that our smfATAC-Seq allows for, as well as the application of trypsin to remove any confounding resident cell types present, results in the sequencing of DNA fragments corresponding purely to the myonuclei within a specific myofiber.

smfATAC-Seq can identify the accessible chromatin regions of a single myofiber

To validate the quality of the ATAC-Seq data generated from a single EDL myofiber, we first investigated the profiles of the ATAC-seq samples from both injured and uninjured myofibers as well as freshly sorted MuSCs. We investigated the similarity between biological replicates for each condition (i.e., uninjured and injured myofibers and MuSCs) by visualization of ATAC-Seq peaks for the muscle-specific gene muscle creatine kinase (*Ckm*) (47), the housekeeping gene *Gapdh* and the MuSC specific gene myogenic factor 5 (*Myf 5*) (48). This analysis not only indicates that smfATAC-Seq can reliably detect chromatin accessibility in a single myofiber but also the presence of comparable peaks in each biological replicate within the specific condition shows the similarity between the samples (Supplemental Figure 3A-C). In addition, Pearson correlation analysis between the biological replicates showed high correlations of ATAC-seq reads between the replicates, indicating consistency within the samples (Supplemental Figure 3D-J). Furthermore, we mapped ATAC-Seq peaks with DNase-Seq from skeletal muscle (Sequence Read Archive, accession # SRX191047) and the similarity between our ATAC-Seq and

previous DNase-Seq was confirmed through the similar peaks present for representative genes, as visualized on the IGV (Supplemental Figure 3A-C). These analyses suggest that smfATAC-Seq can robustly measure chromatin accessibility and identify active cis-regulatory elements in a single EDL myofiber.

Following the initial quality control and the correlation analysis, the biological replicates from the same condition were pooled for further analysis. Peak annotation analysis for MuSCs revealed that more than half of the peaks were in the intron/distal intergenic regions (i.e. enhancer regions) and about 25% of the peaks were in the promoter region (Figure 2C, Supplemental Table 2). Peak annotations for the uninjured and injured single myofibers also showed a great proportion of peaks in the enhancer and promoter regions (Figure 2D,2E, Supplemental Table 2). In addition, an enrichment of ATAC-seq reads around Transcription Start Sites (TSS) from all datasets was observed, which is a typical result expected from ATAC-Seq (30) (Figure 2F-H).

To further assess the quality of the ATAC-Seq data, we analyzed select genes that are expressed by either MuSCs or myofibers. For instance, in the myofiber samples, we confirmed the presence of ATAC-seq peaks in the promoter regions of *Ckm*, Actin alpha 1 (*Acta1*), Myogenic factor 6 (*Myf6*) and Myosin heavy chain 4 (*Myh4*), all of which are expressed by myofibers but not MuSCs (14,47,49,50) (Figure 2I-L). On the other hand, in MuSCs we observed peaks in the promoter regions of Paired box 7 (*Pax7*), and *Myf5*, genes that are known to be expressed in MuSCs (48,51) (Figure 2M-N). *Gapdh* was used a housekeeping gene for all samples (Figure 2O) and *Pou5f1* was used as a negative control (Figure 2P). These observed peaks for known expressed genes demonstrate that our method, smf-ATAC-Seq can reliably analyse chromatin accessibility in a single myofiber.

Uninjured and injured myofibers and MuSCs display distinct chromatin states

To show the global differences in chromatin accessibility between MuSCs, uninjured and injured myofibers, we first performed heatmap clustering of Pearson correlation coefficients on all the replicates/samples, which shows that the biological replicates within conditions are more similar to one another than when compared to those from the other conditions (Figure 3A). This can also be observed through Principal Component Analysis (PCA) where each condition clusters separately, with the injured and uninjured myofibers being more similar to one another than to MuSCs (Figure 3B).

To ensure that the differences seen between myofibers were not due to differences in fiber types, we investigated the chromatin accessibility of known marker genes for slow and fast fiber types. Troponin I2

(*Tnni2*) and Troponin T3 (*Tnnt3*), markers of fast fiber types, (52,53) had a high level of chromatin accessibility while Troponin T1 (*Tnnt1*) and Myosin heavy chain 7 (*Myh7*), expressed in slow fiber types showed no chromatin accessibility (53,54) (Supplemental Figure 4). This data indicates that only fast fiber types were analyzed in this study and that the differences in the chromatin state between the injured and uninjured myofibers were not due to the differences in the fiber types.

For further comparison of accessible regions between the conditions, we performed differential analysis of ATAC-Seq peaks based on the regions defined by the consensus peak sets derived from the uninjured and injured myofibers and MuSCs conditions. The clustering analysis based on the consensus peak set shows the overall chromatin state differences between the MuSCs and injured and uninjured myofibers where the replicates of a condition are more similar to one another than to other conditions (Figure 3C).

The unique chromatin state in each condition can be observed through the pile up analysis that we have performed for the more accessible peaks ($LFC > 2$) and less accessible peaks ($LFC < 0.5$) between uninjured myofibers and MuSCs as well as between uninjured and injured myofibers (Figure 3D). In addition, the proportion of the differential peaks corresponding to various genomic regions such as promoters and enhancers, differs depending on whether we compare MuSCs to myofibers or compare the myofibers during regeneration and homeostasis. For instance, the differential peaks between uninjured myofibers and MuSCs were mostly found close to the promoter region (≤ 1 kb) whereas in the uninjured and injured myofibers comparison, a greater proportion of differential peaks were found in the intron/distal intergenic regions (i.e. enhancer regions) (Figure 3E). This implies that MuSCs and myofibers mostly differ in their promoter accessibility, whereas myofibers during homeostasis and regeneration differ mostly at the level of distal regulatory elements.

Furthermore, we performed occupancy analysis (using DiffBind) in order to determine the unique and common peaks for each comparison between the conditions. Since the occupancy analysis relies on the peak score, the distribution pattern of the peak scores for all the conditions was assessed and found similar despite the differences in the total number of peaks between the conditions (Supplemental Figure 5A). Occupancy analysis between uninjured myofibers and MuSCs revealed that MuSCs have 45533 unique peaks while myofibers contain only 535 unique peaks which are not present in MuSCs (Figure 3F). There are also many common accessible regions as seen from 5930 peaks that are common to both uninjured myofibers and MuSCs. This analysis suggests that myonuclei share large number of open chromatin with their parental stem cells. On the other hand, the occupancy analysis between uninjured and injured myofibers revealed that there are 6352 overlapping peaks between the regenerating and resting myofibers. However, this analysis also revealed that there are 15352 unique peaks in the injured

myofibers and only 83 peaks that are unique to the resting myofiber (Figure 3F). Furthermore, when comparing the read count between conditions around the center of unique peaks, it can be observed that each condition displays a unique open chromatin signature (Supplemental Figure 5B, 5C).

Comparative analysis of the chromatin state between MuSCs and myofibers

To get a better understanding of the functional differences in chromatin accessibility between MuSCs and myofibers, we first performed Gene Ontology (GO Biological Process) analysis on the genes associated with the nearest peaks from the uninjured myofiber and on the genes associated with the nearest unique peaks in the myofiber compared to MuSCs. As expected, this revealed myofiber specific biological processes such as myofiber structure and organization (Figure 4A and Supplemental Figure 6A). On the other hand, GO term analysis on the genes associated with the nearest peaks from MuSCs as well as the genes associated with the nearest unique peaks to MuSCs revealed biological processes such as adherens junction organization, membrane permeability and regulation of notch signaling which play key roles in MuSCs quiescence and function (55,56) (Figure 4B and Supplemental Figure 6B). The analysis above also revealed that genomic regions that remain in an open chromatin state when muscle stem cells fully differentiate into myofibers correspond to genes that are involved in processes such as mitochondrial transport, regulation of transcription and regulation of metabolites and energy (Supplemental Figure 6C). The changes in the chromatin state between MuSCs and myofibers can also be observed from the volcano plots showing differential peaks between conditions labeled by their nearest gene (Figure 4D). For instance, genomic regions associated with genes such as muscle-specific titin-capping protein (*Tcap*), a component of the skeletal muscle z-disc, as well as genes that are involved in regulatory and structural functions in skeletal muscle such as titin gene (*Ttn*) are associated with more accessible chromatin regions in the myofiber compared to MuSCs (57,58) (Figure 4D).

Additionally, we performed GO term analysis between uninjured and injured myofibers, which revealed that the accessible regions in the resting and regenerating myofibers corresponded to genes involved in similar processes. GO term analysis on the genes associated with the nearest peaks from uninjured myofibers and from the injured myofibers as well as the genes associated with the nearest common peaks between uninjured and injured myofibers revealed biological processes involved in striated muscle cell development, actomyosin structure and sarcomere organization, which are important for myofiber structural formation and for the proper function of myofibers (Figure 4A, 4C, Supplemental Figure 6D). On the other hand, genes associated with the nearest unique peaks from the injured myofibers mostly belong to processes involved in structural components of the myofiber while the genes associated with the

nearest unique peaks from uninjured myofibers correspond to genes involved in ion transport and metabolism (Supplemental Figure 6E-F).

Identification of cell type specific pathways by global analysis of chromatin accessibility

To further understand the functional differences in chromatin accessibility between different cell types, we investigated the cell-type-specific pathways. To accomplish this, Gene Set Enrichment Analysis (GSEA) was performed on genes associated with differentially accessible peaks between the conditions. Importantly, the GSEA between uninjured and injured myofibers revealed that inflammatory response and Il2-Stat5 signaling and injury related pathways are still operational even after seven days of CTX-mediated injury to muscle (59) (Figure 5A).

The GSEA between uninjured myofibers and MuSCs revealed that one of the top significantly enriched pathways is myogenesis where we can observe that genes associated with differentiation and myofiber function such as *Myf6*, *Ckm* and Tropomyosin 2 (*Tpm2*) (14,47,60,61) have higher accessibility in the myofiber compared to MuSCs (Figure 5B- 5D). On the other hand, genes associated with quiescence and MuSCs such as alpha 7 integrin (*Itga7*) and *Gpx3* are more accessible in MuSCs compared to the myofiber, as expected (62,63) (Figure 5C, 5E).

Moreover, differential chromatin accessibility between the MuSCs and their myofiber derivatives show differences in pathways that are known to be important for muscle, such as Notch and TGF β signalling (55,56,64,65) (Supplemental Figure 7). For example, increased accessibility of *Notch1* is seen in MuSCs while increased accessibility of Jagged-2 (*Jag2*) is observed in the myofibers regardless of whether they are regenerating or homeostatic as seen by the height of the peaks at their promoters (Supplemental Figure 7A-C). On the other hand, for TGF β signalling, Noggin (*Nog*) shows more accessibility in the myofibers while bone morphogenetic protein-4 (*Bmp4*) has increased accessibility in MuSCs (Supplemental Figure 7D-F). Taken together, this data shows that smfATAC-Seq is an effective method to analyze chromatin accessibility and to identify active cis-regulatory elements in a single muscle fiber as well as to compare muscle fibers under different physiological conditions.

DISCUSSION

Analysis of the myofiber-specific chromatin state and gene expression profile is very limited due to the heterogenous nature of muscle with the presence of numerous non-myogenic cells in the tissue. Whole muscle or muscle biopsies represents a pooled result of numerous cell types which are present in the

muscle tissue. To overcome this limitation, single nucleus RNA-Seq (snRNA-Seq) and ATAC-Seq (snATAC-Seq) have been developed and performed on myonuclei to allow for the computational removal of other cell types (16-18). However, these methods still sequence all myonuclei present in the muscle and cannot distinguish between different myofibers within a muscle. Although snATAC-Seq provides high dimensionality, it is limited in sequencing depth due to the generation of sparse reads. Although computational pseudo bulking of snATAC-Seq can increase read numbers for comparative analysis between samples and conditions, the pooled reads represent average of all myonuclei within the sample.

In this study we have introduced a highly effective protocol based on adaption of OMNI ATAC-Seq (27) to quantify chromatin accessibility of a single EDL myofiber with high resolution and sequencing depth. This method allows comparative analysis of chromatin accessibility within and between muscle types with a potential wide-spread use in future studies to investigate fiber-specific epigenetic alterations in skeletal muscle.

The smfATAC-Seq protocol that we introduce in this study investigates the open chromatin state of myofibers at a single myofiber resolution, and with a high sequencing depth that allows for peak calling and differential peak analysis. Using this method, we have demonstrated that accessible chromatin regions of myonuclei contained within a single EDL myofiber can be tagged and that high-quality sequencing ready libraries can be generated from these fragments. Sequencing of these libraries allow for sufficient depth and peak calling that can be used for genome-wide analysis of chromatin accessibility between myofibers. In this study, we have also demonstrated that smfATAC-Seq strictly investigates a single myofiber without the confounding presence of muscle resident, non-myogenic cell types. Additionally, application of trypsin to isolated myofibers effectively removes MuSCs that are associated with myofibers (66) and was confirmed by the absence of peaks at the promoters of known genes associated with muscle stem and niche cells. Although all of the smfATAC-Seq samples sequenced in this study were fast type myofibers, this protocol can be used to distinguish between different fiber types which could be applied to study myofiber heterogeneity under various physiological conditions. smfATAC-Seq peaks associated with genes involved in muscle structure and function such as *Acta1*, *Ckm* and the myosin heavy chain cluster, indicates that our smfATAC-Seq is a robust technique to investigate genome-wide chromatin accessibility of a single myofiber.

A key implication of this technique is its applicability to the study of changes in chromatin accessibility between myofibers in different contexts. As a demonstration of this, we have performed smfATAC-Seq on uninjured myofibers as well as injured myofibers isolated seven days post-injury to investigate the

changes in chromatin accessibility that occurs during regeneration. Through Pearson correlations and PCA analysis, we showed that each cell type including resting and injured myofibers cluster separately, indicating the power of smfATAC-Seq to determine chromatin signature from minute materials such as a single myofiber. In addition, through occupancy analysis, we showed that there is a large difference in the number of unique peaks present in the injured myofibers compared to uninjured myofibers. This indicates that there are major modifications to chromatin accessibility in the context of regeneration. However, GO term analysis of genes associated with the accessible chromatin regions in both injured and uninjured fibers display significant similarities in the biological processes and pathways such as striated muscle cell development, actomyosin structure and sarcomere organization, which are key factors for the proper structure and function of muscle. Despite these similarities, there are certain trends in which uninjured myofibers have increased accessibility in genes involved in energy metabolism, while injured myofibers have greater accessibility in genes involved in myogenesis and inflammatory response which is what would be expected in the case of an injury and regeneration (67). Despite the increase in chromatin accessibility during injury, the accessible chromatin regions in both injured and uninjured fibers are associated with genes involved in similar biological processes. This similarity at the gene network despite differences in chromatin profile may suggest activation of multiple enhancers on core muscle structural genes in the case of injury. Another possible reason could be the length of the recovery time where at seven days post injury, a significant portion of early activated genes may have gone back to steady state. Previously, a study investigating the changes in the transcriptional profile of MuSCs and various muscle resident cells throughout different time points of muscle injury using single cell RNA-Seq, revealed that after seven days of regeneration most cell types returned to a state that was similar to homeostasis (37). Therefore, it is possible that harvesting the injured EDL myofibers seven days post injury allowed these myofibers to return to a state reminiscent of homeostatic myofibers. smfATAC-Seq can be further implemented to study the effects of regeneration on chromatin accessibility at different time points. Further, our analyses of the myofibers in this study indicates that this technique can effectively compare samples between conditions and could see future use in the study of chromatin accessibility of fibers under different biologically relevant conditions.

Lastly, we have also used smfATAC-Seq to compare changes in chromatin accessibility between MuSCs and myofibers. Our data shows that the regions of open chromatin in the myofibers correspond to genes involved in structural components of the muscle, such as the z-disc, which are important for the proper functioning of the muscle. On the other hand, open regions of chromatin in the MuSCs mostly correspond to genes involved in membrane permeability, adherens junction organization and signalling pathways implicated in the regulation of MuSC function (68). The analysis also revealed that chromatin regions that

are accessible in both MuSCs and myofibers correspond to genes that are crucial for general function of cells such as those involved in mitochondrial transport, regulation of transcription and regulation of metabolites and energy.

Overall, smfATAC-Seq is a robust molecular tool that can be used to analyze genome-wide chromatin accessibility of a single myofiber. The sequencing depth from this approach, allows for in depth analysis, peak calling, quantitative analysis of chromatin accessibility and to identify active enhancers and promoters in a single muscle fiber. smfATAC-Seq can be used to study epigenetic alterations that occur in muscle fibers during development, and in response to exercise.

MATERIALS AND METHODS

ATAC-Seq on a Single Myofiber

Isolation of Extensor Digitorum Longus (EDL) from cardiotoxin-induced injured muscle

The Extensor Digitorum Longus (EDL) muscle was injured by intramuscular injection of 50 μ L of 5 μ M cardiotoxin (CTX) (Sigma, 11061-96-4). Mice were treated with carprofen 20 minutes prior to CTX injection and were injected with CTX under anesthesia by isoflurane. Mice were sacrificed seven days post injury and the EDL was collected from the hind limb of each mouse with the contra lateral EDL being used for the isolation of uninjured myofibers.

Dissection of EDL muscle

The EDL muscle was dissected as previously described (66). Briefly, the skin of the hindlimb was removed and the tibialis anterior (TA) muscle was excised with a pair of dissection scissors. The tendons of the EDL were exposed and cut from tendon to tendon with scissors.

Isolation of a single EDL myofiber

Individual myofibers were isolated from the EDL muscle as previously described (66). Briefly, the intact EDL muscle was placed in a 1.5 mL eppendorf tube with 800 μ L of myofiber digestion buffer containing 1000 U/mL of collagenase from *Clostridium histolyticum* (Sigma, C0130) in un-supplemented DMEM (Gibco, 11995-065) for 1 hour. Trypsin was added to the myofiber digestion buffer at a final concentration of 0.25% to remove the myofiber associated muscle stem cells. The EDL myofibers were then transferred into 2 mL of 1X PBS (Wisent, 311-425-CL) in a 6 well-plate that had previously been

coated with DMEM supplemented with 10% horse serum (HS) (Wisent, 065250). The EDL was then gently pipetted up and down with a large-bore glass pipette to disassociate the myofibers.

Selection of injured and uninjured myofibers

Live myofibers in the 6-well plate were stained with 2 μ L of 5mg/mL of Hoechst (Molecular Probes, H1399) in 2 mL of 1X PBS for 5 minutes in a 37 °C with 5% CO₂ incubator. The myofibers were then visualized under a microscope and selected based on the myonuclei location, where myofibers with a pattern of centrally located nuclei were determined to be regenerating and picked for the injury condition. Individual myofibers were then transferred to 0.2 mL microtubes using a small-bore glass pipette coated with HS.

Lysis and permeabilization of the myofiber

Residual media was removed with pipette under a microscope. Individual myofibers in 0.2 mL microtubes were put in 10 μ L of ddH₂O for 5 minutes on ice. The ddH₂O was removed with a pipette under a microscope, ensuring that the myofiber remained in the tube. The myofiber was then permeabilized with 20 μ L of 0.5% Triton X-100 (Sigma, T9284) in PBS for 15 minutes at room temperature (RT). The permeabilization buffer was removed with a pipette under a microscope and the myofiber was washed twice with 200 μ L of 1X PBS.

Tagmentation of the myofiber by Tn5 transposase

Transposition and ATAC-seq library preparation for a single myofiber was adapted from previously described OMNI ATAC-Seq protocol (27). The permeabilized myonuclei were tagmented with tagmentation mixture optimized for use on a myofiber (20 μ L Tagment DNA Buffer (TD Buffer) (Illumina, 20034197), 13.3 μ L PBS, 0.2% Tween-20 (Sigma, P1379-1L), 0.02% Digitonin (Promega, G9441), 1.39 μ L Tn5 (Illumina, 20034197) and 4.61 μ L water). Each single myofiber was incubated with 6 μ L of the tagmentation mixture at 37 °C for 56 minutes with periodic shaking of the tubes every 5-7 minutes. Following the transposition with Tn5, DNA was purified using a QIAquick PCR Purification Kit (Qiagen, 28104) according to the manufacturer's guidelines.

Library preparation

The purified DNA was PCR amplified for 15 cycles using Q5 High Fidelity DNA polymerase (New England Biolabs, M0491S) with the incorporation of Illumina Nextera XT adaptors (Illumina, FC-131-1001). The libraries were then size selected with AmpureXP Beads (Beckman, Cat# A63880) at a 1: 0.85 ratio (v/v). The size selected libraries were verified for quality control by bioanalyzer as well as

verification of the library size via visualization on an agarose gel stained with GelGreen dye (Biotium, 41005). Libraries were then sequenced on NovaSeq6000 Sprime Paired End (PE) 150 bp.

ATAC-Seq on MuSCs

Isolation of MuSCs by fluorescence-activated cell sorting (FACS) for ATAC-Seq

MuSCs were isolated by Fluorescence Activated Cell Sorting (FACS) as previously described (69). Briefly, hindlimb muscles from Pax7/GFP⁺ mice were dissected and chopped. The minced muscles were then transferred into a 15 mL Falcon tube and digested in un-supplemented F10 media (Gibco, 11550043) with 2.4 U/mL Collagenase D (Roche, 11088882001), 12 U/mL Dispase II (Roche, 39307800) and 0.5mM CaCl₂. Digestion was performed on a shaker in an incubator at 37 °C with 5% CO₂ for 30 minutes. Following the first digestion, digested muscles were centrifuged at 600 g for 20 seconds and the supernatant was transferred to a 50 mL Falcon tube with 9 mL FBS (Wisent, 080450) and was kept on ice. The remaining pellet was triturated and was digested for another 15 minutes with additional digestion buffer added. After the final digestion, the digested muscle mixture was transferred to the 50 ml Falcon tube containing the previously digested mixture. The digested muscle mixture was then filtered through a 40 µm cell strainer (Falcon, C352340) and was centrifuged at 600 g for 18 minutes at 4°C. The pelleted cells were then resuspended in 800 µL FACS buffer that is composed of 2% FBS/ PBS (v:v), 0.5 mM EDTA (Invitrogen, AM9261) and with 0.5 µL DAPI (5mg/mL) (Invitrogen, D3671). Resuspended cells were then filtered through 40 µm cell strainer and were transferred into polypropylene round-bottom FACS compatible tubes (Falcon, 352063). MuSCs were sorted with a FACS Aria Fusion cytometer (BD Biosciences) based on negative selection for DAPI and positive selection for GFP.

Lysis and transposition of MuSCs

ATAC-Seq on MuSCs was performed based on the previously established OMNI-ATAC-Seq protocol (27). Briefly, five thousand MuSCs were sorted by FACS into 30 µL of the ATAC lysis buffer containing 10 mM Tris-HCl (pH 7.5), 10 mM NaCl (Bioshop, 7647-14-5), 3 mM MgCl₂ (Sigma, 7786-30-3), 0.1% Tween-20 (Sigma, P1379-1L), 0.1% NP-40 (Sigma, 74385), and 0.01% Digitonin (Promega, G9441) in a 0.2 mL microtube. Cells were incubated in the lysis buffer for 5 minutes on ice and then 3 minutes at room temperature (RT). Cells were then washed with 100 µL of wash buffer composed of 10mM Tris-HCl (pH 7.5), 10mM NaCl, 3mM MgCl₂ and 0.1% Tween-20, and were centrifuged at 800 g for 10 minutes. The pellet was resuspended in 10 µL of transposition mixture (5 µL TD buffer, 3.2 µL PBS, 0.89 µL Tn5 (Illumina, 20034197), 0.1% Tween-20, 0.01% Digitonin and 0.75 µL nuclease free water). Transposition was performed for 20 minutes at 37 °C while shaking the tubes every 5-7 minutes. The

DNA was then purified using a QIAquick PCR Purification Kit according to the manufacturer's guidelines.

Library preparation for MuSCs ATAC-Seq

The eluted tagmented DNA was PCR amplified for 12 cycles with the incorporation of Illumina Nextera XT adapters using Q5 High Fidelity DNA polymerase. The libraries were then size selected with AmpureXP Beads at a 1: 0.85 ratio (v/v). The libraries were then verified by bioanalyzer and agarose gel visualization. Finally, the samples were sequenced on NovaSeq6000 Sprime Paired End (PE) 150 bp.

ATAC-Seq data processing

The sequencing data was processed using the GenPipes pipeline v.3.1.5 (70). The raw reads were trimmed using Trimmomatic v.0.36 (71) and aligned to the mm10 genome assembly using the Burrows-Wheeler Aligner v.0.7.12 (72). Reads were filtered to keep only high quality alignments (MAPQ score > 20) and duplicates were removed using SAMtools v.1.3.1 (73). Peak calling was performed with MACS2 v.2.1.1 (74) using piling up of paired-end fragment mode (*--format BAMPE*). The peak files (bed) were filtered by removing the ENCODE black listed regions (<https://www.encodeproject.org/files/ENCFF547MET>) using BEDTools v2.29.1 (75). Mitochondrial reads were also removed before the analysis.

Correlation analysis between the biological replicates and clustering

In order to perform a quantitative comparison of the read counts within accessible regions, the overlapping peaks of all replicates were merged using BEDTools v2.29.1 (75). This set of merged peaks and the BAM alignment files were used as input for the *featureCounts* function of Rsubread v.2.2.6 (76) to generate a raw-count matrix. The raw counts were normalized by rlog transformation using DESeq2 (77) with respect to library size. Pearson correlation coefficients were calculated based on the normalized counts for each pairwise comparison. Principal component analysis (PCA) and hierarchical clustering were also performed to evaluate the similarity between the replicates.

Peak annotation analysis

For each condition, the BAM alignment files of the replicates were merged and peak calling was performed with MACS2 v.2.1.1 (74). Peak sets for each condition were annotated using the ChIPseeker v.1.24.0 (78) *annotatePeak* function, and the UCSC Genome Browser knownGene (mm10) table.

Obtaining coverage tracks

The BAM alignment files were converted to bigWig format and normalized by scaling factor (--scaleFactor) with the deepTools v.2.5.0.1 (79) *bamCoverage* function.

Enrichment of genomic signal around TSS

The bigWig files and the TSS coordinates obtained from the UCSC Genome Browser knownGene (mm10) table were used as input for the *computeMatrix* function of deepTools v.2.5.0.1 (79). This matrix was used for *plotHeatmap* function to generate the heatmap.

Identification of overlapping/unique accessible regions

For each comparison between the conditions, overlapping and unique accessible regions were identified with DiffBind v.2.16.2(80) based on the measure of confidence in the peak call by MACS2 v.2.1.1 (74).

Analysis of differentially accessible regions

The identification of differentially accessible regions (DARs) between the conditions was done using DiffBind v.2.16.2 (80) and edgeR v.3.30.1 (81). Log fold changes were calculated, and their associated p-values were corrected for multiple hypothesis testing via the Benjamini–Hochberg procedure to obtain adjusted p-values. The DARs were annotated by their nearest gene using the *annotatePeaks.pl* function of Homer v.4.11(82).

Gene set enrichment analysis

Genes nearby the DARs were ranked based on the log-fold change calculated with edgeR v.3.30.1 (81). This ranked list of genes was used as input to perform gene set enrichment analysis with the *fgseaMultilevel* function of the R package fgsea v.1.14.0 (83). The *FGSEA-multilevel* method is based on an adaptive multi-level split Monte Carlo scheme, which allows the estimation of very low p-values. The Hallmark gene sets collection from the Molecular Signatures Database (MSigDB) (83) was used as a reference to identify the biological processes that were significantly enriched.

Animal care

All procedures that were performed on animals were approved by the McGill University Animal Care Committee (UACC).

Data availability

The data discussed in this study have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession numbers GSE173676 and GSE171534.

Author contributions

Conceptualization: VDS; Methodology: KS, DMB, VDS; Investigation: VDS, MMS, KS, DMB, TK; Visualization: KS, VDS, MMS, DMB; Funding acquisition: VDS; Project administration: VDS; Supervision: VDS, GB; Software: MMS; Writing – original draft: KS, DMB; Writing – review & editing: KS, DMB, VDS, MMS, FL.

Acknowledgments

We thank Christian Young at the Lady Davis Institute for Medical Research—Jewish General Hospital-core facility for his help with fluorescence-activated cell sorting (FACS) of muscle stem cells. We thank Dr. Michael Witcher at McGill University Department of Oncology for his critical comments and careful review of an early draft of this manuscript.

Funding

This work was supported and by discovery grant from Natural Sciences and Engineering Research Council (NSERC) to VDS.

Declaration of Interests

The authors declare no competing interests.

References

1. Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S., Montarras, D., Rocancourt, D., and Relaix, F. (2003) The formation of skeletal muscle: from somite to limb. *J Anat* **202**, 59-68
2. Huxley, H., and Hanson, J. (1954) Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. *Nature* **173**, 973-976
3. Zierath, J. R., and Hawley, J. A. (2004) Skeletal muscle fiber type: influence on contractile and metabolic properties. *PLoS Biol* **2**, e348
4. Bohe, J., Low, J. F., Wolfe, R. R., and Rennie, M. J. (2001) Latency and duration of stimulation of human muscle protein synthesis during continuous infusion of amino acids. *J Physiol* **532**, 575-579

- 539 5. Mittendorfer, B., Andersen, J. L., Plomgaard, P., Saltin, B., Babraj, J. A., Smith, K., and
540 Rennie, M. J. (2005) Protein synthesis rates in human muscles: neither anatomical
541 location nor fibre-type composition are major determinants. *J Physiol* **563**, 203-211
- 542 6. Quiat, D., Voelker, K. A., Pei, J., Grishin, N. V., Grange, R. W., Bassel-Duby, R., and Olson,
543 E. N. (2011) Concerted regulation of myofiber-specific gene expression and muscle
544 performance by the transcriptional repressor Sox6. *Proc Natl Acad Sci U S A* **108**, 10196-
545 10201
- 546 7. Weskamp, K., Olwin, B. B., and Parker, R. (2021) Post-Transcriptional Regulation in
547 Skeletal Muscle Development, Repair, and Disease. *Trends Mol Med* **27**, 469-481
- 548 8. Wilson, J. M., Loenneke, J. P., Jo, E., Wilson, G. J., Zourdos, M. C., and Kim, J. S. (2012)
549 The effects of endurance, strength, and power training on muscle fiber type shifting. *J*
550 *Strength Cond Res* **26**, 1724-1729
- 551 9. Dons, B., Bollerup, K., Bonde-Petersen, F., and Hancke, S. (1979) The effect of weight-
552 lifting exercise related to muscle fiber composition and muscle cross-sectional area in
553 humans. *Eur J Appl Physiol Occup Physiol* **40**, 95-106
- 554 10. Deschenes, M. R. (2004) Effects of aging on muscle fibre type and size. *Sports Med* **34**,
555 809-824
- 556 11. Thompson, L. V. (2002) Skeletal muscle adaptations with age, inactivity, and therapeutic
557 exercise. *J Orthop Sports Phys Ther* **32**, 44-57
- 558 12. Nilwik, R., Snijders, T., Leenders, M., Groen, B. B., van Kranenburg, J., Verdijk, L. B., and
559 van Loon, L. J. (2013) The decline in skeletal muscle mass with aging is mainly attributed
560 to a reduction in type II muscle fiber size. *Exp Gerontol* **48**, 492-498
- 561 13. Roberts, B. M., Frye, G. S., Ahn, B., Ferreira, L. F., and Judge, A. R. (2013) Cancer cachexia
562 decreases specific force and accelerates fatigue in limb muscle. *Biochem Biophys Res*
563 *Commun* **435**, 488-492
- 564 14. Lazure, F., Blackburn, D. M., Corchado, A. H., Sahinyan, K., Karam, N., Sharanek, A.,
565 Nguyen, D., Lepper, C., Najafabadi, H. S., Perkins, T. J., Jahani-Asl, A., and Soleimani, V.
566 D. (2020) Myf6/MRF4 is a myogenic niche regulator required for the maintenance of the
567 muscle stem cell pool. *EMBO Rep* **21**, e49499
- 568 15. Giordani, L., He, G. J., Negroni, E., Sakai, H., Law, J. Y. C., Siu, M. M., Wan, R., Corneau,
569 A., Tajbakhsh, S., Cheung, T. H., and Le Grand, F. (2019) High-Dimensional Single-Cell
570 Cartography Reveals Novel Skeletal Muscle-Resident Cell Populations. *Mol Cell* **74**, 609-
571 621 e606
- 572 16. Petrany, M. J., Swoboda, C. O., Sun, C., Chetal, K., Chen, X., Weirauch, M. T., Salomonis,
573 N., and Millay, D. P. (2020) Single-nucleus RNA-seq identifies transcriptional
574 heterogeneity in multinucleated skeletal myofibers. *Nat Commun* **11**, 6374
- 575 17. Dos Santos, M., Backer, S., Saintpierre, B., Izac, B., Andrieu, M., Letourneur, F., Relaix, F.,
576 Sotiropoulos, A., and Maire, P. (2020) Single-nucleus RNA-seq and FISH identify
577 coordinated transcriptional activity in mammalian myofibers. *Nat Commun* **11**, 5102
- 578 18. Kim, M., Franke, V., Brandt, B., Lowenstein, E. D., Schowel, V., Spuler, S., Akalin, A., and
579 Birchmeier, C. (2020) Single-nucleus transcriptomics reveals functional
580 compartmentalization in syncytial skeletal muscle cells. *Nat Commun* **11**, 6375

19. Zhu, F., Farnung, L., Kaasinen, E., Sahu, B., Yin, Y., Wei, B., Dodonova, S. O., Nitta, K. R., Morgunova, E., Taipale, M., Cramer, P., and Taipale, J. (2018) The interaction landscape between transcription factors and the nucleosome. *Nature* **562**, 76-81
20. Thurman, R. E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M. T., Haugen, E., Sheffield, N. C., Stergachis, A. B., Wang, H., Vernot, B., Garg, K., John, S., Sandstrom, R., Bates, D., Boatman, L., Canfield, T. K., Diegel, M., Dunn, D., Ebersol, A. K., Frum, T., Giste, E., Johnson, A. K., Johnson, E. M., Kutayavin, T., Lajoie, B., Lee, B. K., Lee, K., London, D., Lotakis, D., Neph, S., Neri, F., Nguyen, E. D., Qu, H., Reynolds, A. P., Roach, V., Safi, A., Sanchez, M. E., Sanyal, A., Shafer, A., Simon, J. M., Song, L., Vong, S., Weaver, M., Yan, Y., Zhang, Z., Zhang, Z., Lenhard, B., Tewari, M., Dorschner, M. O., Hansen, R. S., Navas, P. A., Stamatoyannopoulos, G., Iyer, V. R., Lieb, J. D., Sunyaev, S. R., Akey, J. M., Sabo, P. J., Kaul, R., Furey, T. S., Dekker, J., Crawford, G. E., and Stamatoyannopoulos, J. A. (2012) The accessible chromatin landscape of the human genome. *Nature* **489**, 75-82
21. Liu, L., Leng, L., Liu, C., Lu, C., Yuan, Y., Wu, L., Gong, F., Zhang, S., Wei, X., Wang, M., Zhao, L., Hu, L., Wang, J., Yang, H., Zhu, S., Chen, F., Lu, G., Shang, Z., and Lin, G. (2019) An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. *Nat Commun* **10**, 364
22. Trevino, A. E., Sinnott-Armstrong, N., Andersen, J., Yoon, S. J., Huber, N., Pritchard, J. K., Chang, H. Y., Greenleaf, W. J., and Pasca, S. P. (2020) Chromatin accessibility dynamics in a model of human forebrain development. *Science* **367**
23. Corces, M. R., Granja, J. M., Shams, S., Louie, B. H., Seoane, J. A., Zhou, W., Silva, T. C., Groeneveld, C., Wong, C. K., Cho, S. W., Satpathy, A. T., Mumbach, M. R., Hoadley, K. A., Robertson, A. G., Sheffield, N. C., Felau, I., Castro, M. A. A., Berman, B. P., Staudt, L. M., Zenklusen, J. C., Laird, P. W., Curtis, C., Cancer Genome Atlas Analysis, N., Greenleaf, W. J., and Chang, H. Y. (2018) The chromatin accessibility landscape of primary human cancers. *Science* **362**
24. Rendeiro, A. F., Schmidl, C., Strefford, J. C., Walewska, R., Davis, Z., Farlik, M., Oscier, D., and Bock, C. (2016) Chromatin accessibility maps of chronic lymphocytic leukaemia identify subtype-specific epigenome signatures and transcription regulatory networks. *Nat Commun* **7**, 11938
25. Wang, Y., Zhang, X., Song, Q., Hou, Y., Liu, J., Sun, Y., and Wang, P. (2020) Characterization of the chromatin accessibility in an Alzheimer's disease (AD) mouse model. *Alzheimers Res Ther* **12**, 29
26. Bastle, R. M., and Maze, I. (2019) Chromatin Regulation in Complex Brain Disorders. *Curr Opin Behav Sci* **25**, 57-65
27. Corces, M. R., Trevino, A. E., Hamilton, E. G., Greenside, P. G., Sinnott-Armstrong, N. A., Vesuna, S., Satpathy, A. T., Rubin, A. J., Montine, K. S., Wu, B., Kathiria, A., Cho, S. W., Mumbach, M. R., Carter, A. C., Kasowski, M., Orloff, L. A., Risca, V. I., Kundaje, A., Khavari, P. A., Montine, T. J., Greenleaf, W. J., and Chang, H. Y. (2017) An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**, 959-962
28. Buenrostro, J. D., Wu, B., Chang, H. Y., and Greenleaf, W. J. (2015) ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21 29 21-21 29 29

29. Liu, C., Wang, M., Wei, X., Wu, L., Xu, J., Dai, X., Xia, J., Cheng, M., Yuan, Y., Zhang, P., Li, J., Feng, T., Chen, A., Zhang, W., Chen, F., Shang, Z., Zhang, X., Peters, B. A., and Liu, L. (2019) An ATAC-seq atlas of chromatin accessibility in mouse tissues. *Sci Data* **6**, 65
30. Yan, F., Powell, D. R., Curtis, D. J., and Wong, N. C. (2020) From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. *Genome Biol* **21**, 22
31. Rocks, D., Jaric, I., Tesfa, L., Grealley, J. M., Suzuki, M., and Kundakovic, M. (2021) Cell type-specific chromatin accessibility analysis in the mouse and human brain. *Epigenetics*, 1-18
32. Janssen, I., Heymsfield, S. B., Wang, Z. M., and Ross, R. (2000) Skeletal muscle mass and distribution in 468 men and women aged 18-88 yr. *J Appl Physiol* (1985) **89**, 81-88
33. Ortenblad, N., Nielsen, J., Boushel, R., Soderlund, K., Saltin, B., and Holmberg, H. C. (2018) The Muscle Fiber Profiles, Mitochondrial Content, and Enzyme Activities of the Exceptionally Well-Trained Arm and Leg Muscles of Elite Cross-Country Skiers. *Front Physiol* **9**, 1031
34. Mishra, P., Varuzhanyan, G., Pham, A. H., and Chan, D. C. (2015) Mitochondrial Dynamics is a Distinguishing Feature of Skeletal Muscle Fiber Types and Regulates Organellar Compartmentalization. *Cell Metab* **22**, 1033-1044
35. Neal, A., Boldrin, L., and Morgan, J. E. (2012) The satellite cell in male and female, developing and adult mouse muscle: distinct stem cells for growth and regeneration. *PLoS One* **7**, e37950
36. Cramer, A. A. W., Prasad, V., Eftestol, E., Song, T., Hansson, K. A., Dugdale, H. F., Sadayappan, S., Ochala, J., Gundersen, K., and Millay, D. P. (2020) Nuclear numbers in syncytial muscle fibers promote size but limit the development of larger myonuclear domains. *Nat Commun* **11**, 6287
37. De Micheli, A. J., Laurillard, E. J., Heinke, C. L., Ravichandran, H., Fraczek, P., Soueid-Baumgarten, S., De Vlaminck, I., Elemento, O., and Cosgrove, B. D. (2020) Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. *Cell Rep* **30**, 3583-3595 e3585
38. De Micheli, A. J., Spector, J. A., Elemento, O., and Cosgrove, B. D. (2020) A reference single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated muscle stem cell populations. *Skelet Muscle* **10**, 19
39. Khan, S. S., Solomon, M. A., and McCoy, J. P., Jr. (2005) Detection of circulating endothelial cells and endothelial progenitor cells by flow cytometry. *Cytometry B Clin Cytom* **64**, 1-8
40. Steppan, C. M., Bailey, S. T., Bhat, S., Brown, E. J., Banerjee, R. R., Wright, C. M., Patel, H. R., Ahima, R. S., and Lazar, M. A. (2001) The hormone resistin links obesity to diabetes. *Nature* **409**, 307-312
41. McKinney-Freeman, S. L., Naveiras, O., Yates, F., Loewer, S., Philitas, M., Curran, M., Park, P. J., and Daley, G. Q. (2009) Surface antigen phenotypes of hematopoietic stem cells from embryos and murine embryonic stem cells. *Blood* **114**, 268-278
42. Agorku, D. J., Langhammer, A., Heider, U., Wild, S., Bosio, A., and Hardt, O. (2019) CD49b, CD87, and CD95 Are Markers for Activated Cancer-Associated Fibroblasts Whereas CD39 Marks Quiescent Normal Fibroblasts in Murine Tumor Models. *Front Oncol* **9**, 716

43. Waddell, L. A., Lefevre, L., Bush, S. J., Raper, A., Young, R., Lisowski, Z. M., McCulloch, M. E. B., Muriuki, C., Sauter, K. A., Clark, E. L., Irvine, K. M., Pridans, C., Hope, J. C., and Hume, D. A. (2018) ADGRE1 (EMR1, F4/80) Is a Rapidly-Evolving Gene Expressed in Mammalian Monocyte-Macrophages. *Front Immunol* **9**, 2246
44. Joe, A. W., Yi, L., Natarajan, A., Le Grand, F., So, L., Wang, J., Rudnicki, M. A., and Rossi, F. M. (2010) Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. *Nat Cell Biol* **12**, 153-163
45. Folker, E. S., and Baylies, M. K. (2013) Nuclear positioning in muscle development and disease. *Front Physiol* **4**, 363
46. Roman, W., and Gomes, E. R. (2018) Nuclear positioning in skeletal muscle. *Semin Cell Dev Biol* **82**, 51-56
47. Tai, P. W., Fisher-Aylor, K. I., Himeda, C. L., Smith, C. L., Mackenzie, A. P., Helterline, D. L., Angello, J. C., Welikson, R. E., Wold, B. J., and Hauschka, S. D. (2011) Differentiation and fiber type-specific activity of a muscle creatine kinase intronic enhancer. *Skelet Muscle* **1**, 25
48. Rudnicki, M. A., Schnegelsberg, P. N., Stead, R. H., Braun, T., Arnold, H. H., and Jaenisch, R. (1993) MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75**, 1351-1359
49. Nowak, K. J., Wattanasirichaigoon, D., Goebel, H. H., Wilce, M., Pelin, K., Donner, K., Jacob, R. L., Hubner, C., Oexle, K., Anderson, J. R., Verity, C. M., North, K. N., Iannaccone, S. T., Muller, C. R., Nurnberg, P., Muntoni, F., Sewry, C., Hughes, I., Sutphen, R., Lacson, A. G., Swoboda, K. J., Vigneron, J., Wallgren-Pettersson, C., Beggs, A. H., and Laing, N. G. (1999) Mutations in the skeletal muscle alpha-actin gene in patients with actin myopathy and nemaline myopathy. *Nat Genet* **23**, 208-212
50. Stuart, C. A., Stone, W. L., Howell, M. E., Brannon, M. F., Hall, H. K., Gibson, A. L., and Stone, M. H. (2016) Myosin content of individual human muscle fibers isolated by laser capture microdissection. *Am J Physiol Cell Physiol* **310**, C381-389
51. Seale, P., Sabourin, L. A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M. A. (2000) Pax7 is required for the specification of myogenic satellite cells. *Cell* **102**, 777-786
52. Mullen, A. J., and Barton, P. J. (2000) Structural characterization of the human fast skeletal muscle troponin I gene (TNNT2). *Gene* **242**, 313-320
53. Wei, B., and Jin, J. P. (2016) TNNT1, TNNT2, and TNNT3: Isoform genes, regulation, and structure-function relationships. *Gene* **582**, 1-13
54. Meredith, C., Herrmann, R., Parry, C., Liyanage, K., Dye, D. E., Durling, H. J., Duff, R. M., Beckman, K., de Visser, M., van der Graaff, M. M., Hedera, P., Fink, J. K., Petty, E. M., Lamont, P., Fabian, V., Bridges, L., Voit, T., Mastaglia, F. L., and Laing, N. G. (2004) Mutations in the slow skeletal muscle fiber myosin heavy chain gene (MYH7) cause laing early-onset distal myopathy (MPD1). *Am J Hum Genet* **75**, 703-708
55. Bjornson, C. R., Cheung, T. H., Liu, L., Tripathi, P. V., Steeper, K. M., and Rando, T. A. (2012) Notch signaling is necessary to maintain quiescence in adult muscle stem cells. *Stem Cells* **30**, 232-242
56. Mourikis, P., and Tajbakhsh, S. (2014) Distinct contextual roles for Notch signalling in skeletal muscle stem cells. *BMC Dev Biol* **14**, 2

57. Markert, C. D., Meaney, M. P., Voelker, K. A., Grange, R. W., Dalley, H. W., Cann, J. K., Ahmed, M., Bishwokarma, B., Walker, S. J., Yu, S. X., Brown, M., Lawlor, M. W., Beggs, A. H., and Childers, M. K. (2010) Functional muscle analysis of the Tcap knockout mouse. *Hum Mol Genet* **19**, 2268-2283
58. Hackman, P., Vihola, A., Haravuori, H., Marchand, S., Sarparanta, J., De Seze, J., Labeit, S., Witt, C., Peltonen, L., Richard, I., and Udd, B. (2002) Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. *Am J Hum Genet* **71**, 492-500
59. Laurence, A., Tato, C. M., Davidson, T. S., Kanno, Y., Chen, Z., Yao, Z., Blank, R. B., Meylan, F., Siegel, R., Hennighausen, L., Shevach, E. M., and O'Shea J, J. (2007) Interleukin-2 signaling via STAT5 constrains T helper 17 cell generation. *Immunity* **26**, 371-381
60. Buckingham, M. (1994) Muscle differentiation. Which myogenic factors make muscle? *Curr Biol* **4**, 61-63
61. Jin, Y., Peng, Y., Lin, Z., Chen, Y. C., Wei, L., Hacker, T. A., Larsson, L., and Ge, Y. (2016) Comprehensive analysis of tropomyosin isoforms in skeletal muscles by top-down proteomics. *J Muscle Res Cell Motil* **37**, 41-52
62. El Haddad, M., Jean, E., Turki, A., Hugon, G., Vernus, B., Bonniieu, A., Passerieux, E., Hamade, A., Mercier, J., Laoudj-Chenivesse, D., and Carnac, G. (2012) Glutathione peroxidase 3, a new retinoid target gene, is crucial for human skeletal muscle precursor cell survival. *J Cell Sci* **125**, 6147-6156
63. Pasut, A., Oleynik, P., and Rudnicki, M. A. (2012) Isolation of muscle stem cells by fluorescence activated cell sorting cytometry. *Methods Mol Biol* **798**, 53-64
64. Carlson, M. E., Conboy, M. J., Hsu, M., Barchas, L., Jeong, J., Agrawal, A., Mikels, A. J., Agrawal, S., Schaffer, D. V., and Conboy, I. M. (2009) Relative roles of TGF-beta1 and Wnt in the systemic regulation and aging of satellite cell responses. *Aging Cell* **8**, 676-689
65. Girardi, F., Taleb, A., Ebrahimi, M., Datye, A., Gamage, D. G., Peccate, C., Giordani, L., Millay, D. P., Gilbert, P. M., Cadot, B., and Le Grand, F. (2021) TGFbeta signaling curbs cell fusion and muscle regeneration. *Nat Commun* **12**, 750
66. Blackburn, D. M., Lazure, F., Corchado, A. H., Perkins, T. J., Najafabadi, H. S., and Soleimani, V. D. (2019) High-resolution genome-wide expression analysis of single myofibers using SMART-Seq. *J Biol Chem* **294**, 20097-20108
67. Tidball, J. G. (2011) Mechanisms of muscle injury, repair, and regeneration. *Compr Physiol* **1**, 2029-2062
68. Relaix, F., Bencze, M., Borok, M. J., Der Vartanian, A., Gattazzo, F., Mademtzoglou, D., Perez-Diaz, S., Prola, A., Reyes-Fernandez, P. C., Rotini, A., and Taglietti, t. (2021) Perspectives on skeletal muscle stem cells. *Nat Commun* **12**, 692
69. Tichy, E. D., Sidibe, D. K., Greer, C. D., Oyster, N. M., Rompolas, P., Rosenthal, N. A., Blau, H. M., and Mourkioti, F. (2018) A robust Pax7EGFP mouse that enables the visualization of dynamic behaviors of muscle stem cells. *Skelet Muscle* **8**, 27
70. Bourgey, M., Dali, R., Eveleigh, R., Chen, K. C., Letourneau, L., Fillon, J., Michaud, M., Caron, M., Sandoval, J., Lefebvre, F., Leveque, G., Mercier, E., Bujold, D., Marquis, P., Van, P. T., Anderson de Lima Morais, D., Tremblay, J., Shao, X., Henrion, E., Gonzalez, E.,

- Quirion, P. O., Caron, B., and Bourque, G. (2019) GenPipes: an open-source framework for distributed and scalable genomic analyses. *Gigascience* **8**
71. Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120
72. Li, H., and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760
73. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079
74. Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., and Liu, X. S. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137
75. Quinlan, A. R. (2014) BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr Protoc Bioinformatics* **47**, 11.12.11-34
76. Liao, Y., Smyth, G. K., and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930
77. Love, M. I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550
78. Yu, G., Wang, L. G., and He, Q. Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383
79. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A., and Manke, T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187-191
80. Stark, R. B. G. (2011) DiffBind: differential binding analysis of ChIP-Seq peak data. in *Bioconductor*
81. Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140
82. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589
83. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M. N., and Sergushichev, A. (2021) Fast gene set enrichment analysis. *bioRxiv*, 060012
84. McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., and Bejerano, G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501

Figure 1: Schematic of ATAC-seq performed on a single myofiber

Schematic showing the steps and reactions involved in the preparation of sequencing ready libraries of single myofiber DNA for ATAC-Seq. Briefly, myofibers were isolated from the EDL muscle and an individual myofiber was transferred to a 0.2mL microtube. The myofiber was then lysed with ddH₂O and the myonuclei were permeabilized with 0.5% Triton X-100. Then, open chromatin regions were tagmented with hyperactive Tn5 transposase and the DNA fragments were purified through column purification. The tagmented DNA was then amplified by PCR and Nextera adapters were incorporated. Finally, size selection and purification was performed using 0.85X AMPure beads, resulting in sequencing ready libraries.

Figure 1

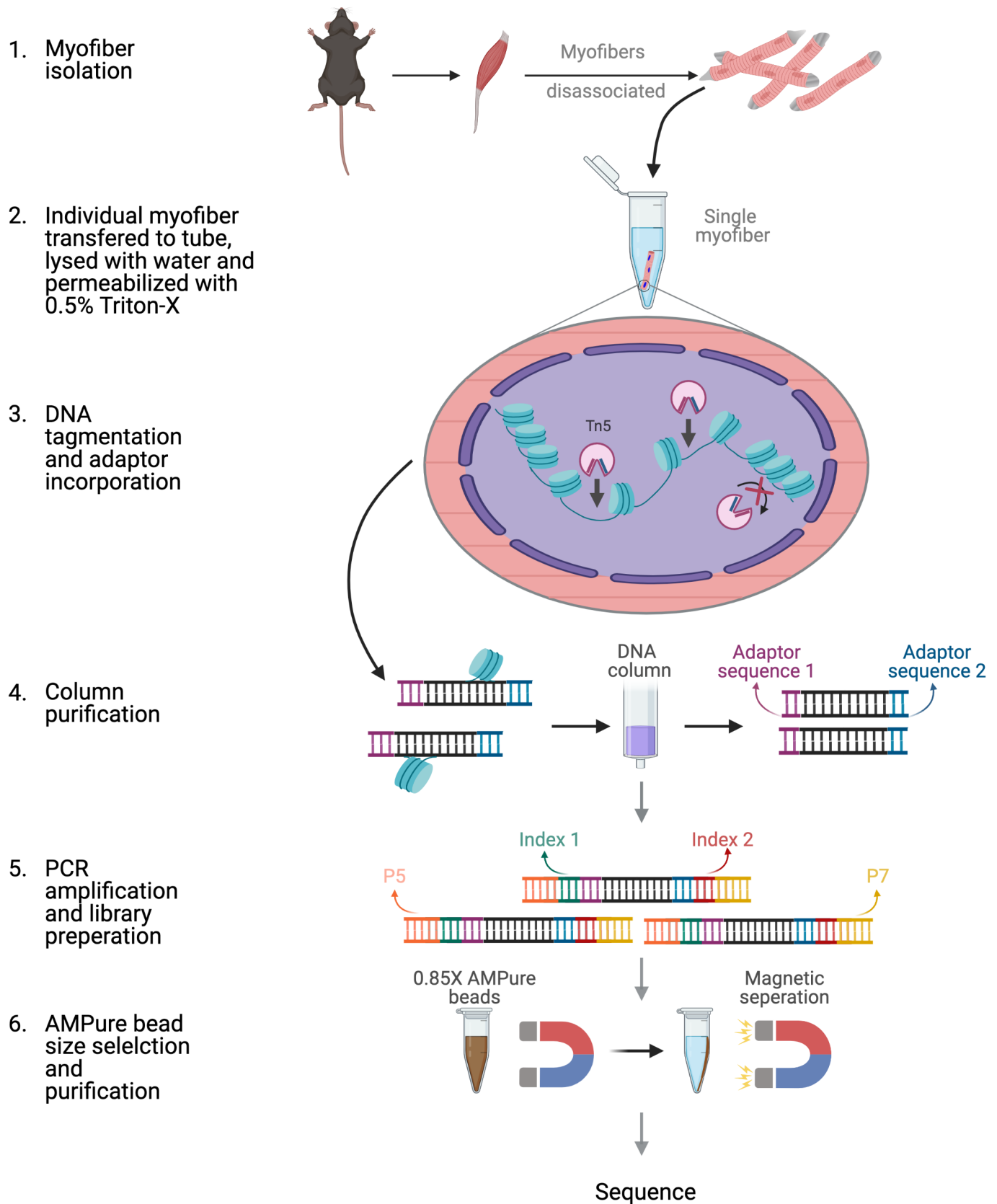


Figure 2: smfATAC-Seq can effectively identify the accessible regions on a single myofiber

(A) Representative picture of an isolated WT C57BL/6J uninjured myofiber stained for DAPI showing presence and location of myonuclei. Scale bar = 50 μ m. (B) Representative picture of an isolated WT C57BL/6J injured myofiber (7 days after cardiotoxin induced injury) stained for DAPI showing centrally located myonuclei as a marker of a regenerating fiber. Scale bar = 50 μ m. (C-E) Peak annotation pie charts for ATAC-Seq peaks of MuSCs, injured myofibers and uninjured myofibers, respectively. (F-H) Heatmaps showing enrichment at transcription start site (TSS) for the ATAC-Seq libraries of MuSCs, injured myofibers and uninjured myofibers, respectively. (I-P) IGV snapshots of known genes expressed in muscle fiber and/or MuSCs displaying accessibility on their respective TSS. (I) The muscle creatine kinase (*Ckm*). (J) Actin alpha 1 (*Acta1*). (K) Part of the myosin heavy chain (*Myh*) gene cluster. (L) Myogenic factor 6 (*Myf6*). (M) Paired box 7 (*Pax7*). (N) Myogenic factor 5 (*Myf5*). (O) Housekeeping gene *Gapdh*. (P) POU Class 5 homeobox 1 (*Pou5f1*) as a negative control.

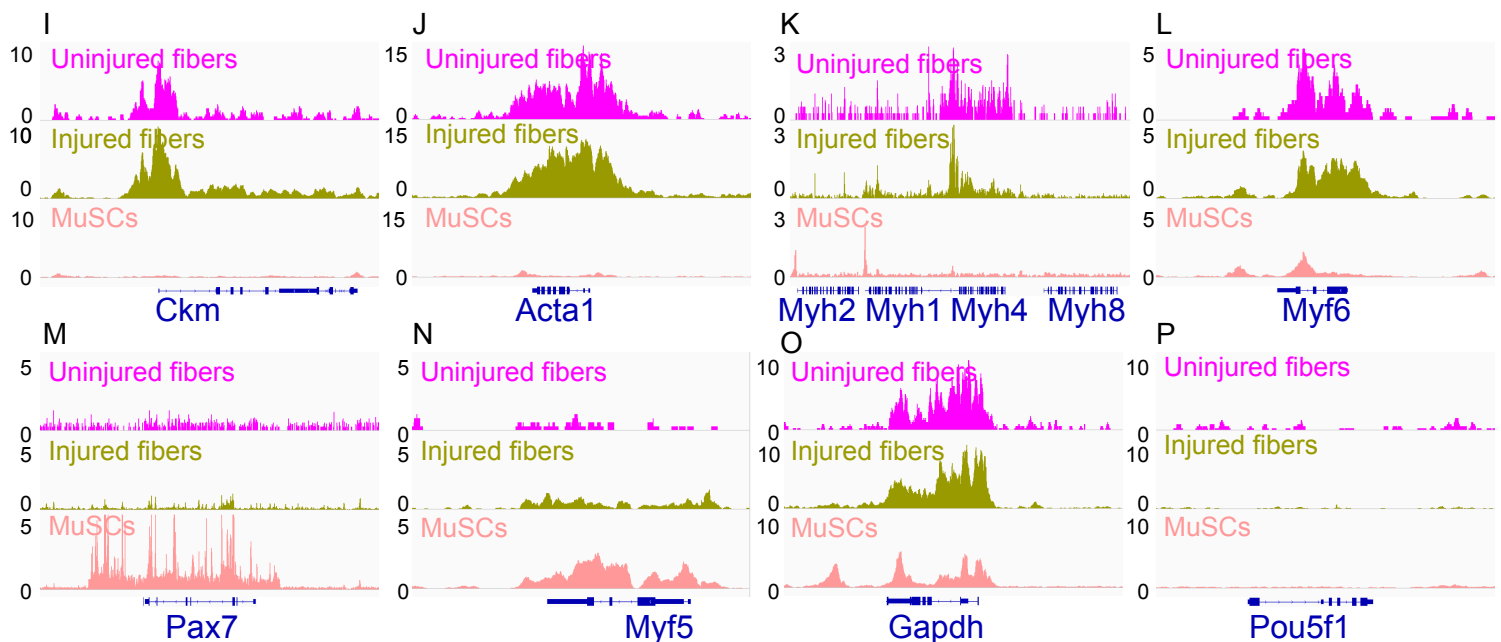
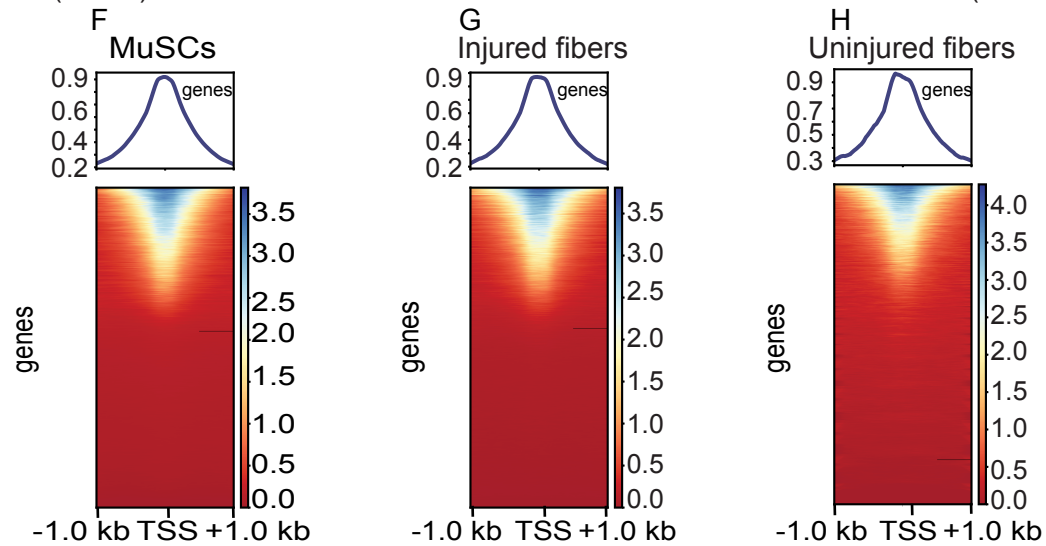
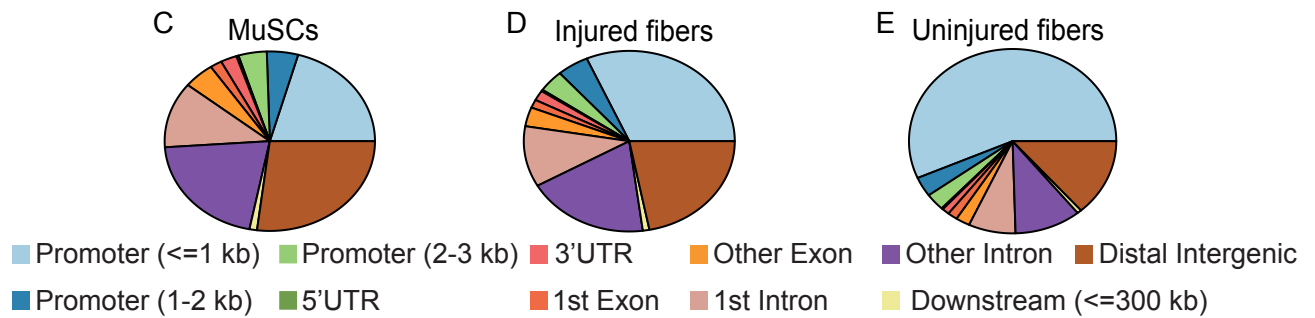
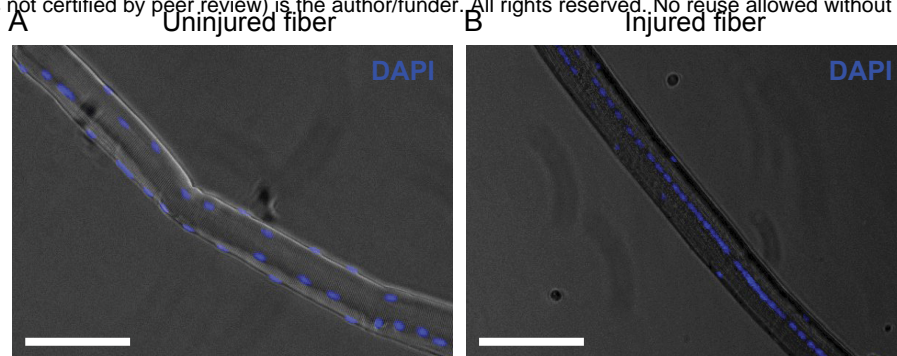


Figure 3: Uninjured and injured myofibers and MuSCs display distinct chromatin states

(A) Heatmap clustering of Pearson correlation coefficients showing the correlation between the replicates of the conditions in the regions defined by the union peakset (merged peaks of all replicates/samples). (B) Projection of samples along first two principal components found by PCA showing the separate clustering of different samples and the clustering of each replicate of the same condition together. (C) Pile up analysis of differentially accessible peaks between uninjured myofibers and MuSCs and between injured myofibers and uninjured myofibers. Less accessible regions: $FDR < 0.05$ and $LFC < 0.5$. More accessible peaks: $FDR < 0.05$ and $LFC > 2$. (D) Heatmap clustering of Pearson correlation coefficients indicating the correlation between the replicates in the regions defined by the consensus peakset derived from the uninjured myofibers, injured myofibers and MuSCs. (E) Peak annotation pie charts for the differentially accessible peaks between uninjured myofibers vs MuSCs and uninjured myofibers vs injured myofibers. (F) Venn diagram of the number of ATAC-Seq peaks that are unique or overlapping between uninjured myofibers vs MuSCs and uninjured myofibers vs injured myofibers.

Figure 3

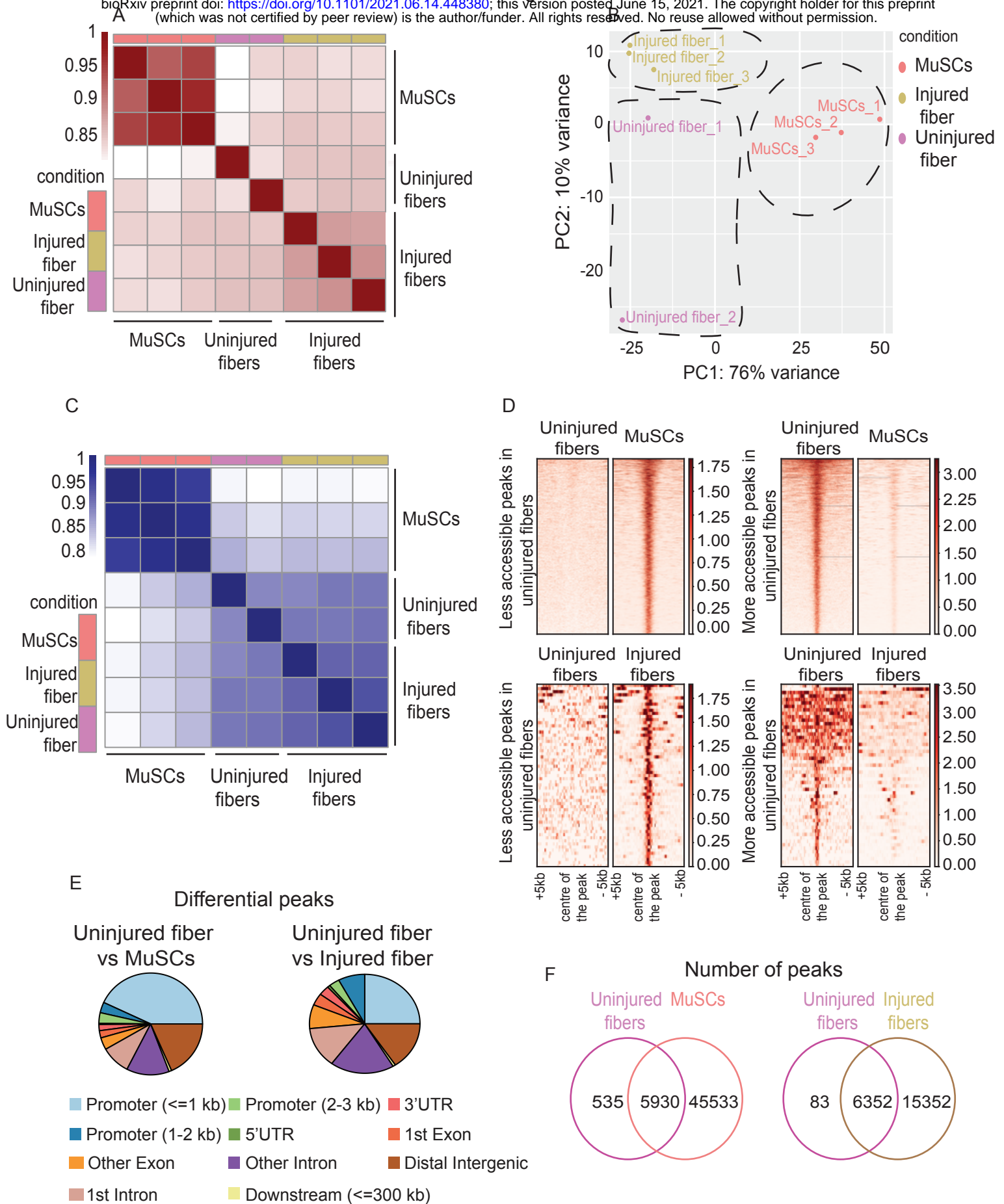
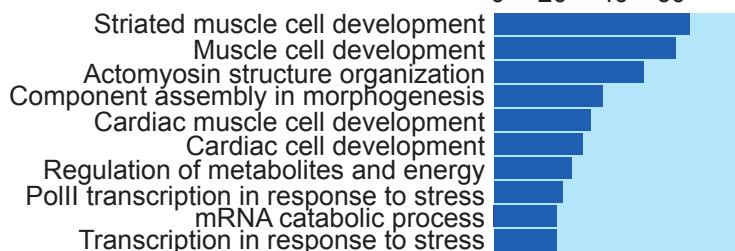


Figure 4: Comparative analysis of chromatin state between uninjured myofibers and MuSCs and between uninjured myofibers and injured myofibers

(A-C) Gene Ontology (GO Biological Process) analysis of genes associated with ATAC-Seq peaks based on association by proximity using Genomic Regions Enrichment of Annotations Tool (GREAT) (84) for all peaks present in the uninjured myofibers, MuSCs and injured myofibers respectively. (D) Volcano plot of differentially accessible regions/peaks identified by $FDR < 0.05$ and $LFC \geq 1$ between uninjured myofibers and MuSCs. Each dot represents a differentially accessible region/peak and the distance to the nearest gene is annotated. (E) Volcano plot of differentially accessible regions/peaks identified by $FDR < 0.05$ and $LFC \geq 1$ between uninjured myofibers and injured myofibers. Each coloured dot represents a differentially accessible region/peak and the distance to the nearest gene is annotated.

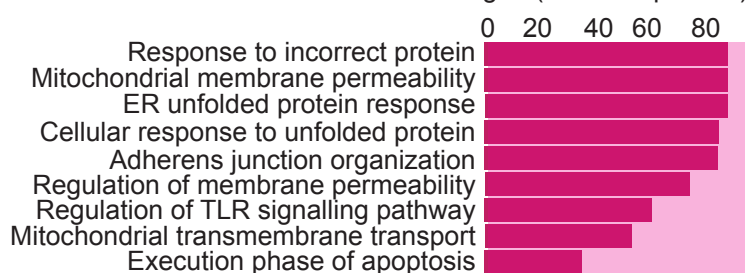
A

Uninjured fibers
GO Biological Process
-log₁₀ (Binomial p value)



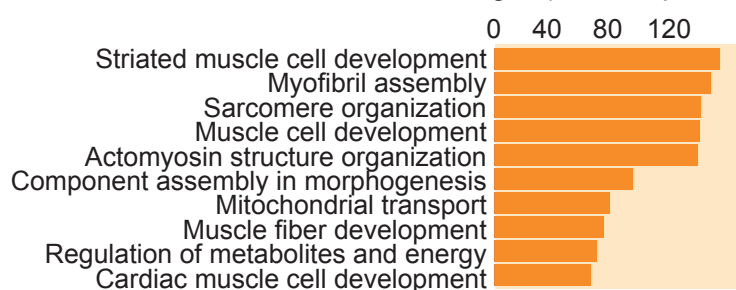
B

MuSCs
GO Biological Process
-log₁₀ (Binomial p value)



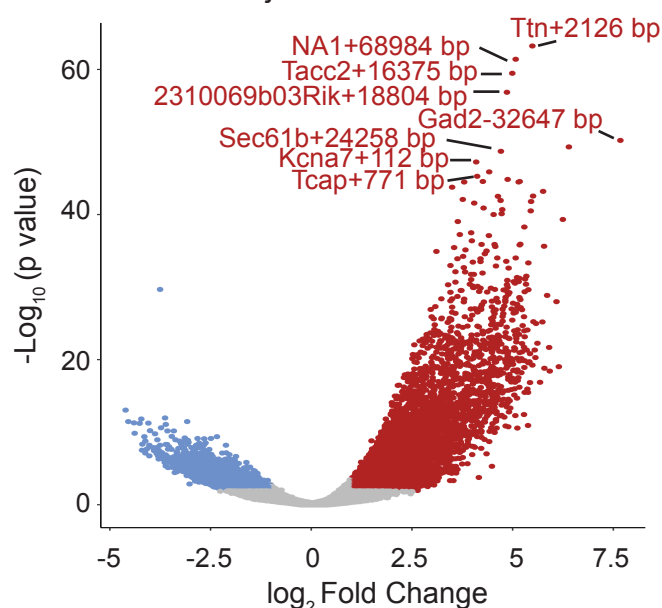
C

Injured fibers
GO Biological Process
-log₁₀ (Binomial p value)



D

Uninjured fibers vs MuSCs



E

Injured fibers vs Uninjured fibers

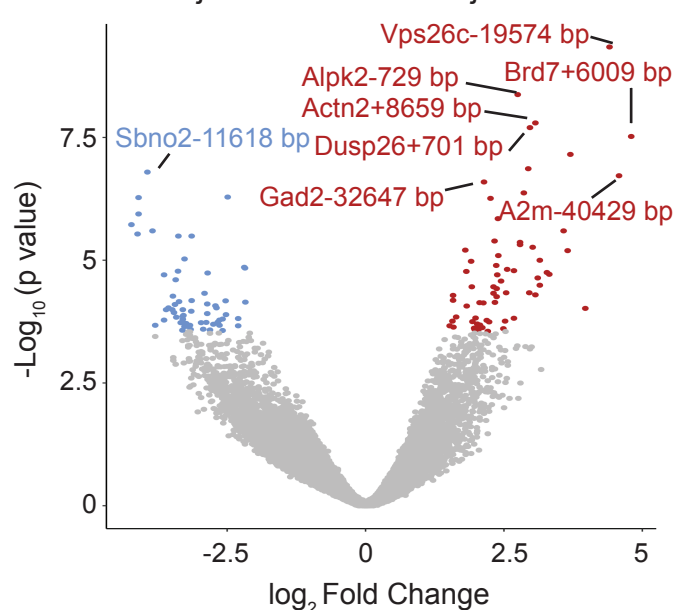


Figure 5: Identification of cell type specific pathways by global analysis of chromatin accessibility

(A) Gene Set Enrichment Analysis performed on nearby genes of differentially accessible regions/peaks for uninjured myofibers compared to injured myofibers. Top 10 enriched pathways are shown although do not reach significance. (B) Gene Set Enrichment Analysis performed on nearby genes of differentially accessible regions/peaks for uninjured fibers compared to MuSCs. Top 10 significantly enriched pathways are shown (FDR < 0.01). (C) Heatmap for myogenesis based on read counts of MuSCs and uninjured fibers +/- 1 kb of the TSS of each gene in the myogenic pathway. (D) IGV snapshot of Tropomyosin 2 (*Tpm2*). (E) IGV snapshot of Glutathione Peroxidase 3 (*Gpx3*).



Figure S1: **Quality control of ATAC-Seq libraries**

(A) Bioanalyzer profile of an ATAC-Seq library prepared from 5000 MuSCs. (B) Example bioanalyzer profile of ATAC-Seq library prepared from a single myofiber. (C) Representative picture of a ready to sequence MuSC ATAC-seq library after size selection, visualized on an agarose gel. (D) Fold enrichment for the TSS of MyoD compared to negative control region of Chromosome 17 qE5 between MuSC ATAC-Seq libraries and untagmented DNA as seen by qPCR (n=5, 3 biological replicates and 2 technical duplicates), two-tailed t-test, error bars = \pm SD. (E) Representative picture of a ready to sequence ATAC-seq on a single myofiber library after size selection, visualized on an agarose gel. (F) qPCR for the TSS of MyoD compared with a negative control region of Chromosome 17 qE5 for the ATAC-Seq libraries prepared from single myofibers (n=3, biological triplicates), two-tailed t-test, error bars = \pm SD.

Supplemental Figure 1

bioRxiv preprint doi: <https://doi.org/10.1101/2021.06.14.448380>; this version posted June 15, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

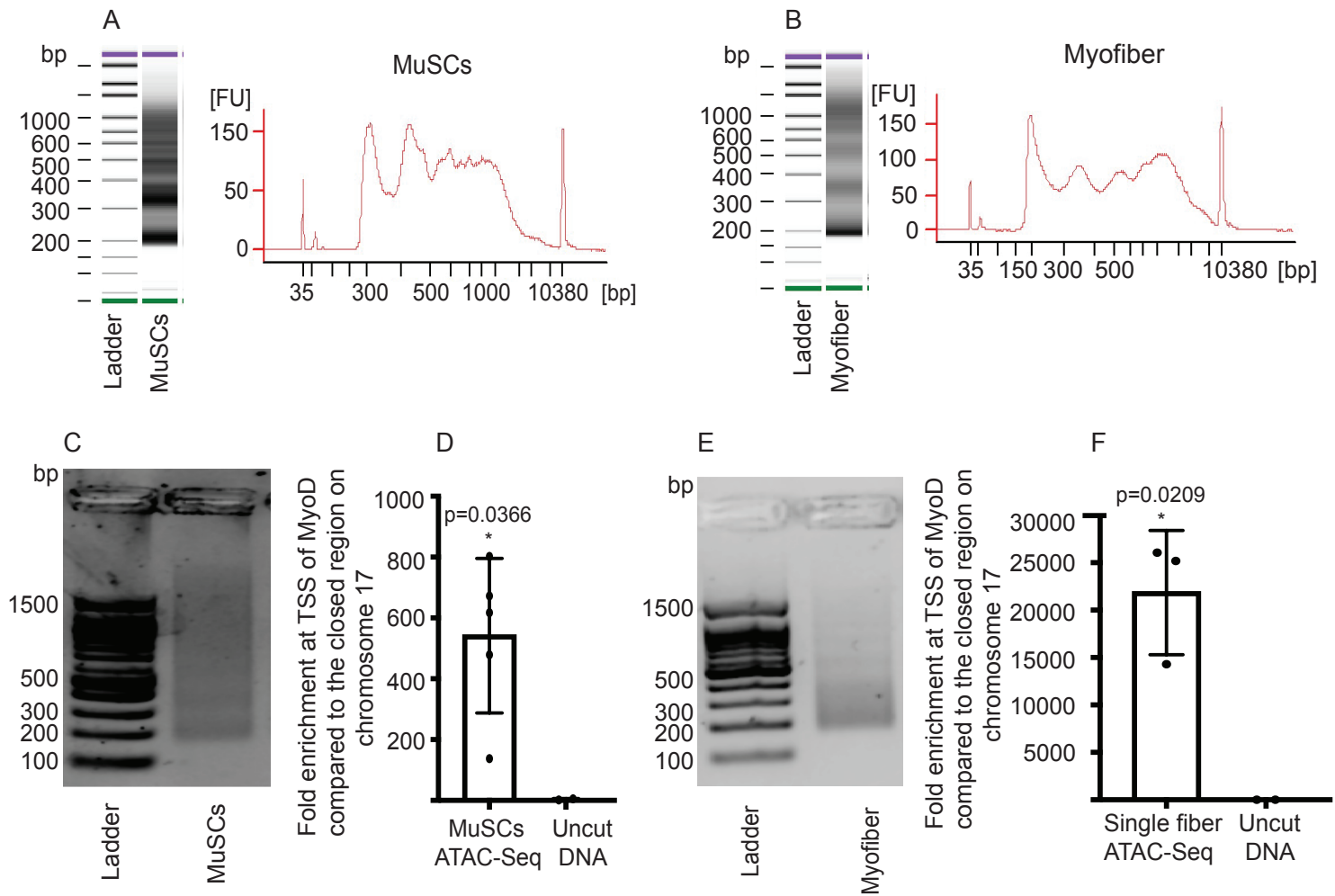


Figure S2. IGV snapshots of non-myogenic genes

(A) Platelet and Endothelial Cell Adhesion Molecule 1 (*Pecam1*) expressed in endothelial cells. (B) Resistin (*Retn*) as a marker of adipocytes. (C) *CD45* expressed in hematopoietic cells. (D) CD90 (*Thy1*) expressed in fibroblasts. (E) Lymphocyte antigen 6a (*Ly6a*) expressed in fibro/adipogenic progenitors (FAPs). (F) Adhesion G protein-coupled receptor E1 (*ADGRE1*) gene expressed in macrophages. (G) The housekeeping gene *RPS2* used as a positive control. (H) Housekeeping gene TATA-Box Binding protein (*Tbp*) was also used as a positive control.

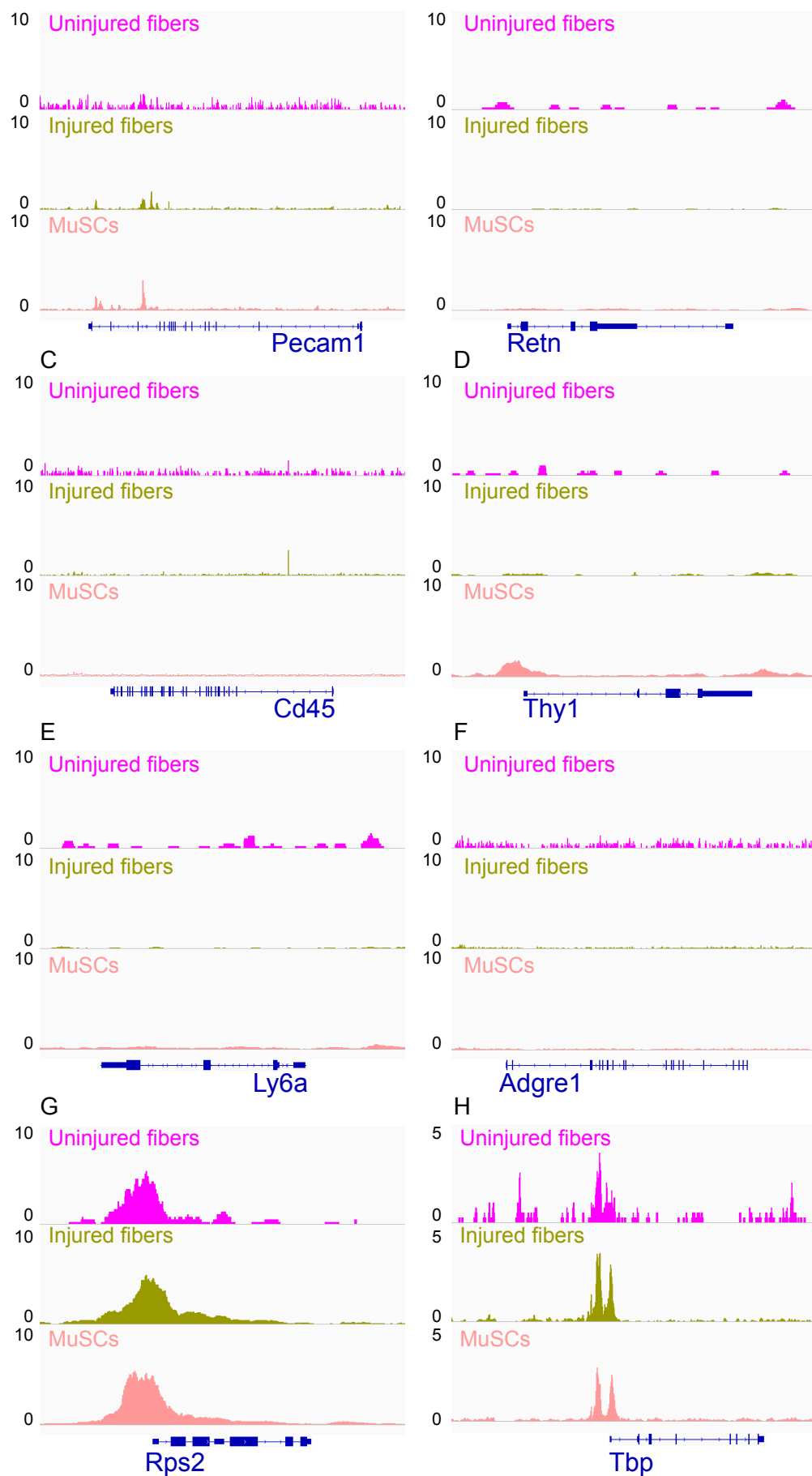


Figure S3: Correlation analysis between biological replicates of each condition

(A-C) IGV snapshots of genes expressed in myofibers and MuSCs for all the replicates of each condition that were pooled together for further analysis. DNase-Seq track added to demonstrate correlation of the myofiber ATAC-Seq with previously performed DNase-Seq on skeletal muscle. (A) Housekeeping gene *Gapdh*. (B) The muscle creatine kinase (*Ckm*). (C) Myogenic Factor 5 (*Myf5*). (D-J) Scatter plot showing the Pearson correlation between the replicates.

Supplemental Figure 3

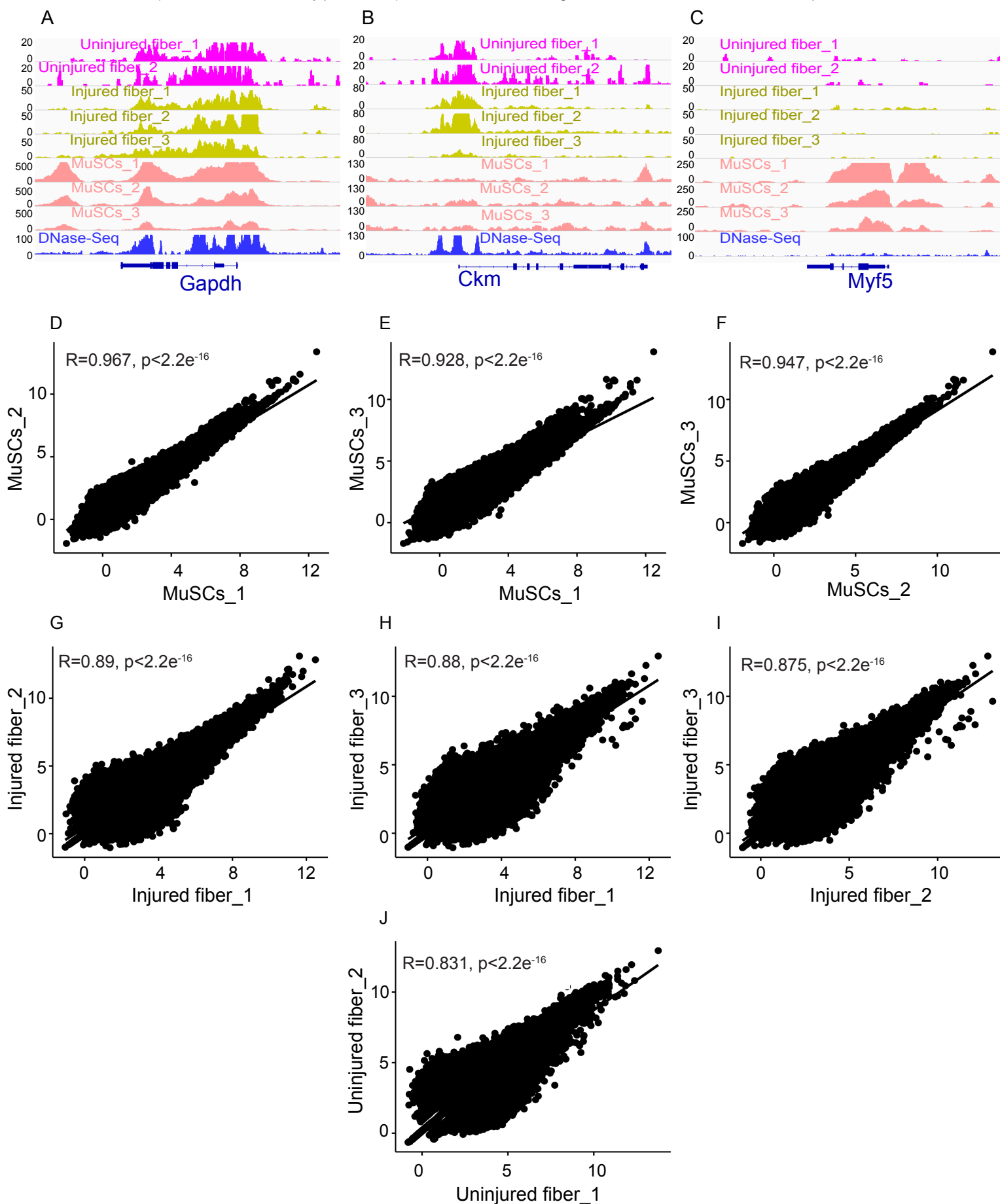


Figure S4. IGV snapshots of genes expressed in fast and slow muscle fiber types

(A) Troponin I2 (*Tnni2*) expressed in fast skeletal muscle fiber. (B) Troponin T3 (*Tnnt3*) expressed in fast skeletal muscle fiber. (C) Troponin T1 (*Tnnt1*) expressed in slow skeletal muscle fibers. (D) Myosin heavy chain 7 (*Myh7*) expressed in slow skeletal muscle fibers

Supplemental Figure 4

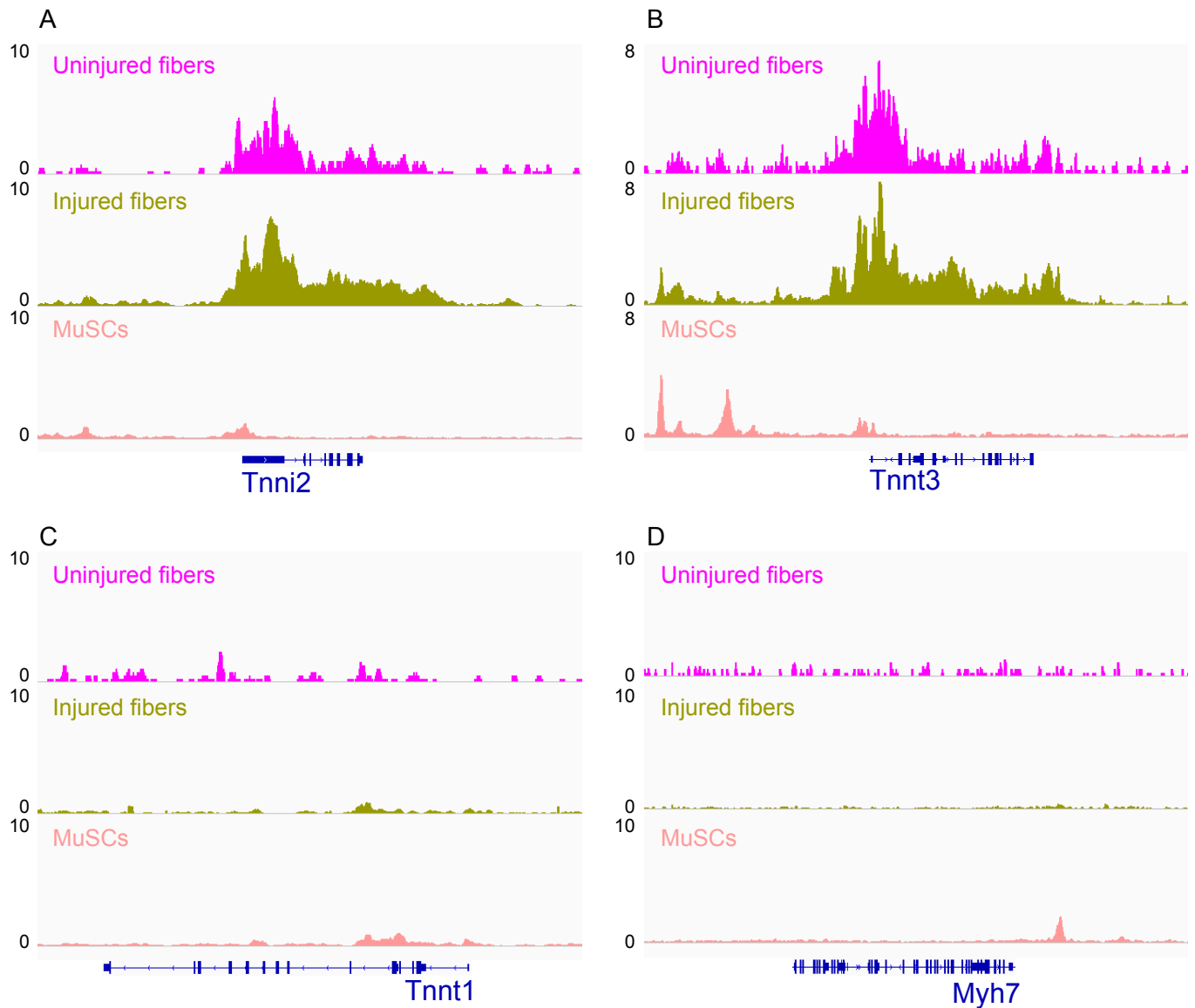


Figure S5: Unique peaks between different conditions indicate a distinct chromatin state for each cell type

(A) Heatmap showing the read count ± 500 bp of the center of unique peaks to MuSCs compared to uninjured myofibers. (B) Heatmap showing the read count ± 500 bp of the center of unique peaks to injured myofibers compared to uninjured myofibers. (C) Peak score distribution (calculated by MACS2 peak calling algorithm) for each of the different conditions. Peak score = $-\log_{10}(\text{FDR})$.

Supplemental Figure 5

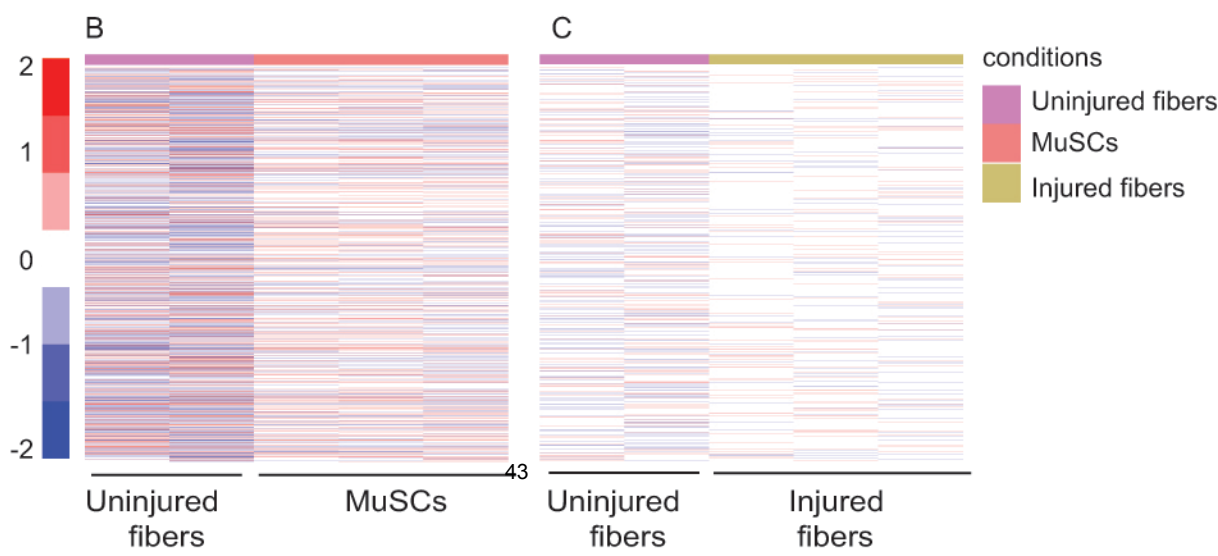
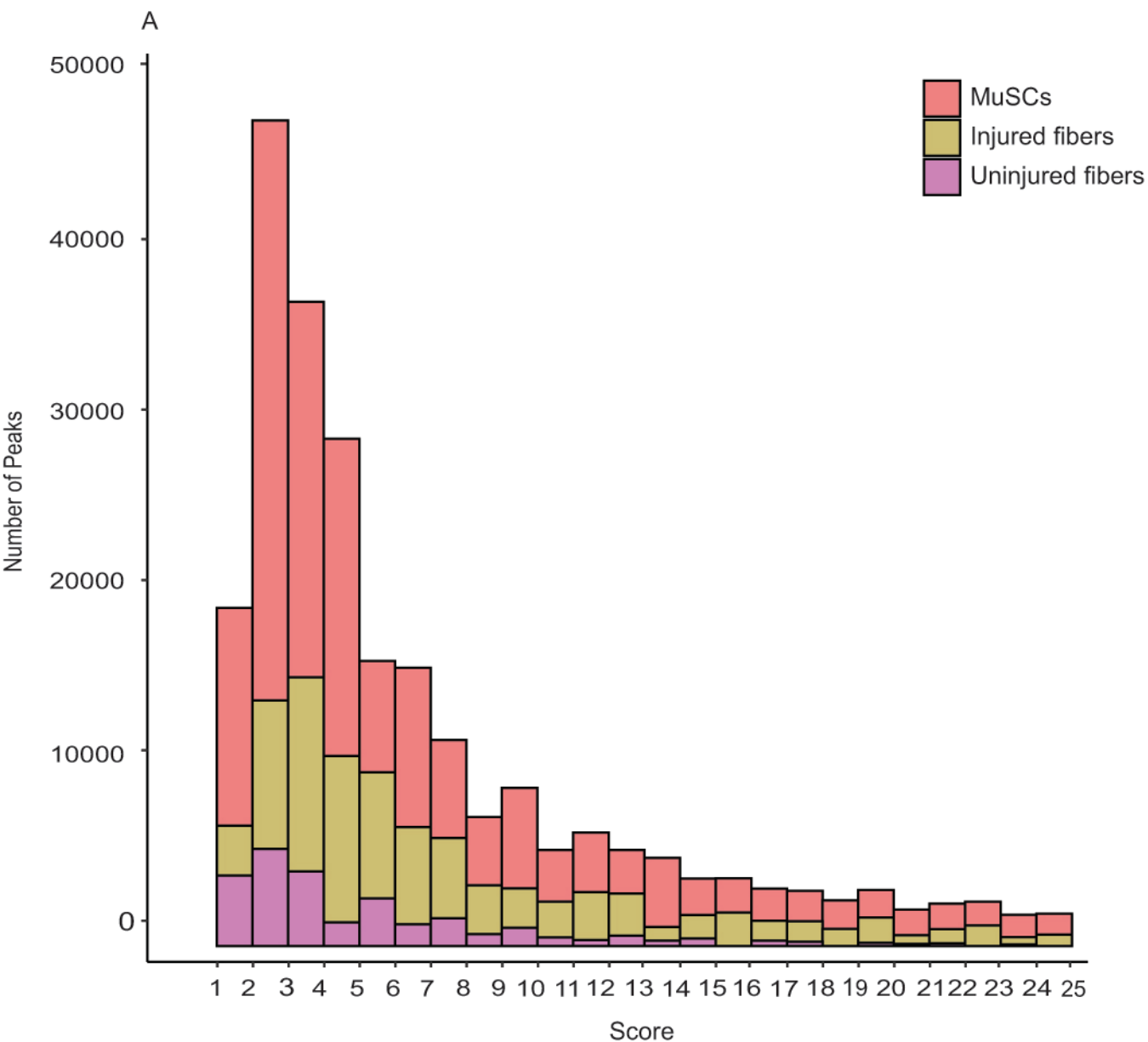
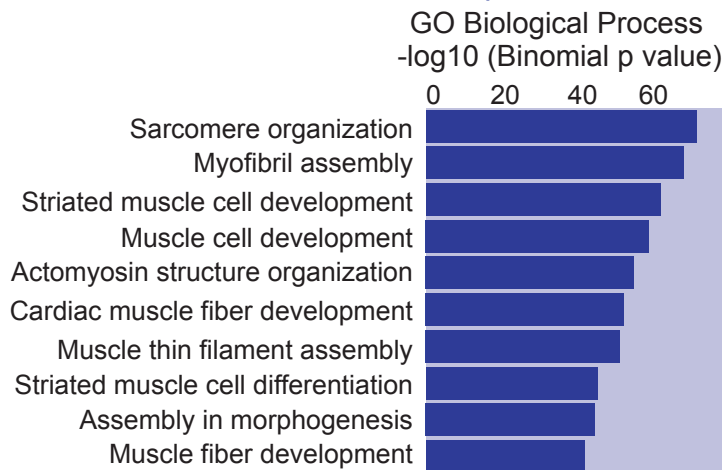


Figure S6: Gene Ontology analysis of unique and common peaks between conditions

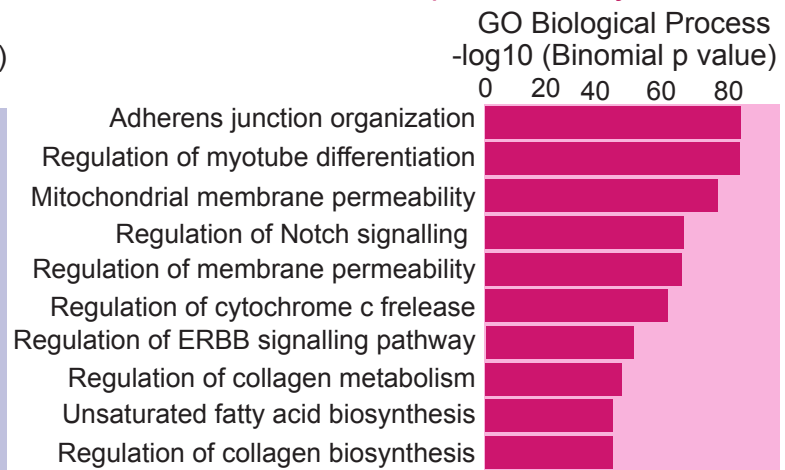
(A) Gene Ontology (GO Biological Process) analysis of genes associated with unique peaks present in the uninjured myofiber compared to MuSCs, based on the proximity of the peaks to the genes. (B) GO term analysis of genes associated with unique peaks in MuSCs compared to uninjured myofibers. (C) GO term analysis of genes associated with peaks that are common between MuSCs and uninjured myofibers. (D) GO term analysis of genes associated with peaks that are common between injured and uninjured myofibers. (E) GO term analysis of genes associated with unique peaks in injured myofibers compared to uninjured myofibers. (F) GO term analysis of genes associated with unique peaks in uninjured myofibers compared to injured myofibers.

Supplemental Figure 2

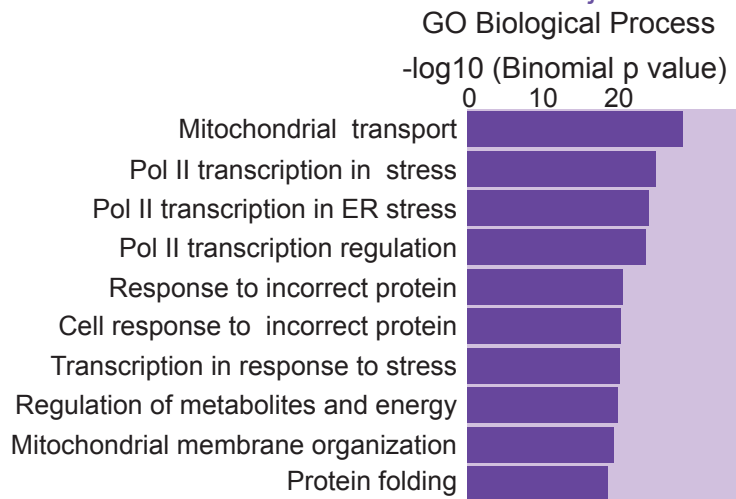
A Peaks unique to uninjured fibers compared to MuSCs



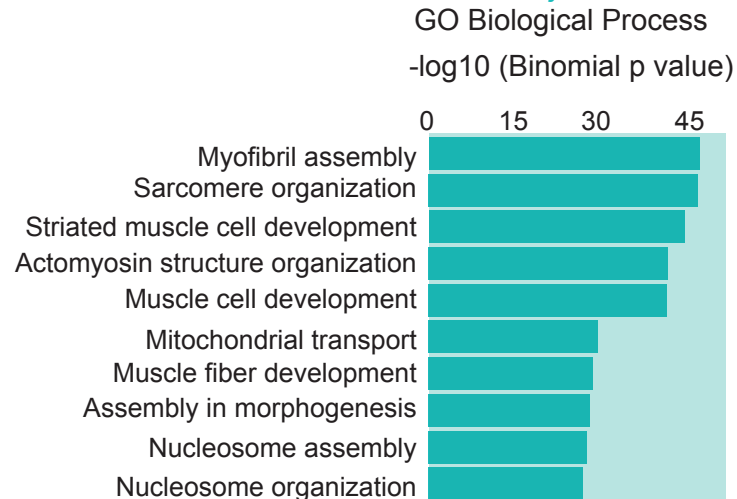
B Peaks unique to MuSCs compared to uninjured fibers



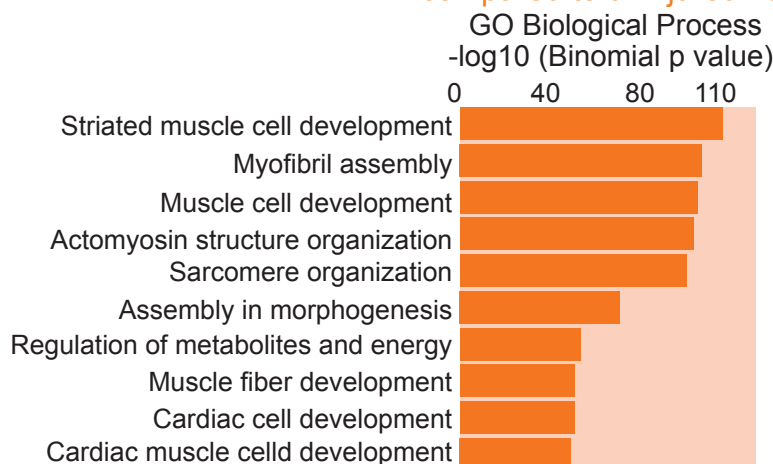
C Peaks common to MuSCs and uninjured fibers



D Peaks common to injured and uninjured fibers



E Peaks unique to injured fibers compared to uninjured fibers



F Peaks unique to uninjured fibers compared to injured fibers

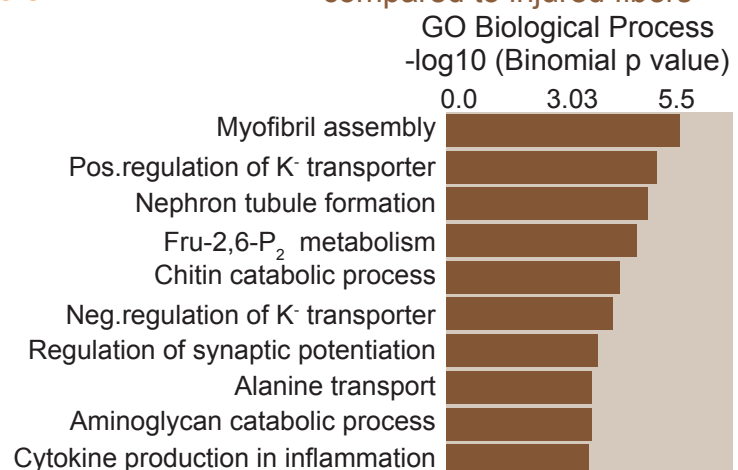


Figure S7: Analysis of Notch and TGF β signalling pathways reveal differential accessibility between MuSCs and uninjured myofibers and injured and uninjured myofibers

(A) Heatmap showing genes involved in the Notch signalling pathway based on read counts of MuSCs, uninjured fibers and injured fibers, +/- 1 kb of the TSS of each gene in the pathway. (B) IGV snapshot of Notch homolog 1 (*Notch1*). (C) IGV snapshot of Protein jagged 2 (*Jag2*). (D) Heatmap showing genes involved in the TGF β signalling pathway based on read counts of MuSCs, uninjured fibers and injured fibers, +/- 1 kb of the TSS of each gene in the pathway. (E) IGV snapshot of Noggin (*Nog*). (F) IGV snapshot of Bone morphogenetic protein (*Bmp4*).

Supplemental Figure 7

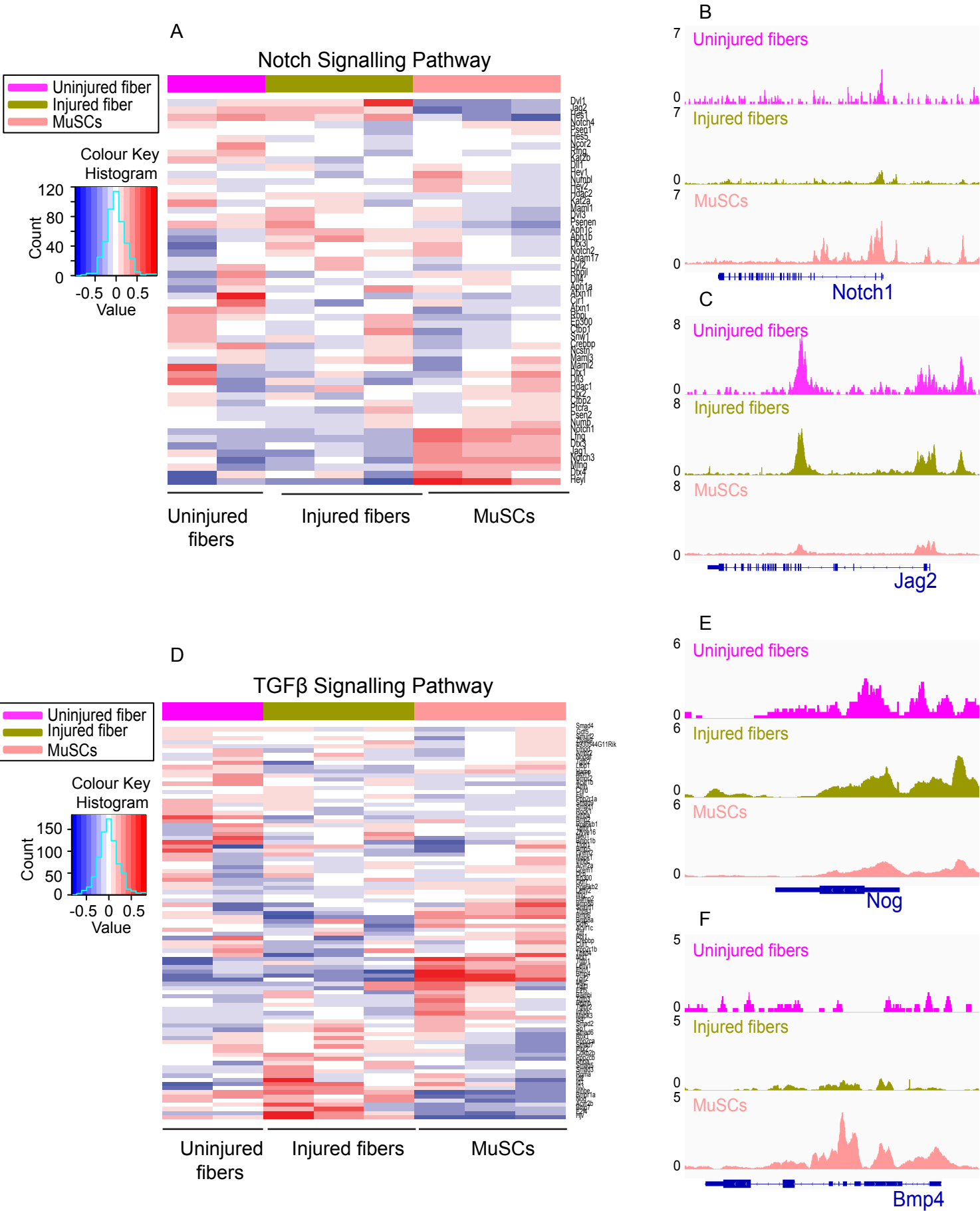


Table 1: Sequencing read information for smfATAC-Seq and MuSCs ATAC-Seq libraries

Library	Number of Raw Reads	Number of Surviving Reads	Aligned Filtered Reads (mm10 reference)	Duplicate Reads	Mitochondrial Reads	Percentage of Mitochondrial Reads (%)	Final Reads Aligned	Number of Peaks	Fraction in peaks (FrIP)
Muscle Stem Cells_1	1.76E+08	1.14E+08	103130186	47623836	529967	0.51	5E+07	65568	0.364
Muscle Stem Cells_2	1.75E+08	1.17E+08	103570009	43672484	374176	0.36	6E+07	68658	0.197
Muscle Stem Cells_3	1.32E+08	91261584	79944121	31299456	223540	0.28	5E+07	69573	0.13
Injured_1	2.3E+08	1.17E+08	90040002	81024926	830215	0.92	8E+06	32853	0.289
Injured_2	1.95E+08	1.3E+08	98752157	88549329	1300615	1.32	9E+06	28351	0.286
Injured_3	1.42E+08	62888552	52132455	42271079	868808	1.67	9E+06	25002	0.233
Uninjured_1	1.45E+08	75781456	61034569	52588315	1274332	2.09	7E+06	12276	0.218
Uninjured_2	1.51E+08	64192706	50120282	45914841	965037	1.93	3E+06	14742	0.321

Table 2: Percentage of total peaks found in each genomic feature

	Muscle Stem Cells (%)	Injured Myofiber (%)	Uninjured Myofiber (%)
Promoter (<=1 kb)	20.66	31.61	56.54
Promoter (1-2 kb)	4.81	4.84	3.45
Promoter (2-3 kb)	4.37	3.92	3.01
5'UTR	0.34	0.27	0.23
3'UTR	2.50	1.82	1.15
1 st Exon	1.83	1.47	1.53
Other Exon	4.75	3.42	2.19
1 st Intron	11.85	10.87	7.35
Other Intron	20.80	18.84	10.35
Downstream (<=300 kb)	1.16	1.01	0.69
Distal Intergenic	26.95	21.93	13.51

Table 3: Percentage of differential peaks found in each genomic feature

	Uninjured Myofiber vs. MuSCs	Uninjured vs Injured Myofiber
Promoter (≤ 1 kb)	43.07	25
Promoter (1-2 kb)	3.36	7.81
Promoter (2-3 kb)	3.39	3.12
5'UTR	0.37	0.78
3'UTR	1.95	3.12
1 st Exon	2.29	3.91
Other Exon	3.85	7.81
1 st Intron	9.16	13.28
Other Intron	13.34	18.75
Downstream (≤ 300 kb)	0.83	0.78
Distal Intergenic	18.39	15.62