DNA Research, 2021, 28(2), 1–11
doi: 10.1093/dnares/dsab003
Advance Access Publication Date: 1 May 2021
Research Article

OXFORD

Research Article

Reorganization of chromatin architecture during prenatal development of porcine skeletal muscle

Renqiang Yuan[1,2†], Jiaman Zhang[2†], Yujie Wang[2], Xingxing Zhu[2], Silu Hu[2], Jianhua Zeng[3], Feng Liang[1], Qianzi Tang[2], Yaosheng Chen[1], Luxi Chen[1,4], Wei Zhu[2], Mingzhou Li[2]*, and Delin Mo[1]*

[1]State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou 510006, China, [2]Institute of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu 611130, China, [3]Guangdong YIHAO Food Co., Ltd, Guangzhou 510620, China, and [4]Guangdong Key Laboratory of Pharmaceutical Bioactive Substances, Guangdong Pharmaceutical University, Guangzhou 510006, China

*To whom correspondence should be addressed. Tel. +86 20 39332940. Fax. +86 20 39332940. Email: mingzhou.li@sicau.edu.cn (M.L.); modelin@mail.sysu.edu.cn (D.M.)

†These authors contributed equally to this work.

Received 31 January 2021; Editorial decision 25 April 2021; Accepted 26 April 2021

Abstract

Myofibres (primary and secondary myofibre) are the basic structure of muscle and the determinant of muscle mass. To explore the skeletal muscle developmental processes from primary myofibres to secondary myofibres in pigs, we conducted an integrative three-dimensional structure of genome and transcriptomic characterization of longissimus dorsi muscle of pig from primary myofibre formation stage [embryonic Day 35 (E35)] to secondary myofibre formation stage (E80). In the hierarchical genomic structure, we found that 11.43% of genome switched compartment A/B status, 14.53% of topologically associating domains are changed intradomain interactions (*D*-scores) and 2,730 genes with differential promoter–enhancer interactions and (or) enhancer activity from E35 to E80. The alterations of genome architecture were found to correlate with expression of genes that play significant roles in neuromuscular junction, embryonic morphogenesis, skeletal muscle development or metabolism, typically, *NEFL*, *MuSK*, *SLN*, *Mef2D* and *GCK*. Significantly, *Sox6* and *MATN2* play important roles in the process of primary to secondary myofibres formation and increase the regulatory potential score and genes expression in it. In brief, we reveal the genomic reorganization from E35 to E80 and construct genome-wide high-resolution interaction maps that provide a resource for studying long-range control of gene expression from E35 to E80.

Key words: pig, skeletal muscle development, Hi-C, TAD, PEI

1. Introduction

Skeletal muscle is one of the most dynamic tissues that comprises ~40% of total body weight and is able to break down energy materials (glucose, amino acid and fatty acid) to produce chemical energy for maintaining posture and movement.[1] Skeletal muscles of agricultural animals (typically, pig, beef cattle and broiler) is a significant dietary protein sources for human consumption. For muscle development, including embryonic muscle development, postnatal

growth and regeneration, its disruption is implicated in various muscle dysfunctions, including muscular atrophy and muscular dystrophy.[2] Given the importance of skeletal muscle in many ways, a comprehensive understanding skeletal muscle development is crucial for human biomedical research related to muscle and improving meat production of livestock.

Skeletal muscle is a heterogeneous and highly complex tissue that composes of fibres with different morphological, contractile and metabolic characteristics including oxidative (Types I and IIa) and glycolytic (Types IIb and IIx) fibres.[3] The number and size of muscle fibres determined the muscle mass. Nevertheless, the total number of muscle fibre remains stable after birth,[4] therefore, the embryonic stage is critical for muscle development. Muscle fibres development from progenitor cells in the paraxial mesoderm that is segmented into somites that lie on either side of the neural tube.[5] Myogenesis in embryonic muscle development involved two waves of fibre formation, including primary fibre formation and secondary fibres formation. In pigs, primary fibres generate from ~35 to 55 days of gestation, followed by the secondary myofibres forming based on the template of the primary myofibre around 50–90 days of gestation.[6] Myogenic regulatory factors (Myf5, MyoD1, myogenin and MRF4), the basic helix–loop–helix family of transcription factors, are well-known that control the determination and differentiation of skeletal muscle cells during embryogenesis and postnatal myogenesis.[7] For example, MyoD1 is a crucial master regulator in controlling muscle-specific gene transcription. Previous report had found three regulatory regions to regulate MyoD1 expression: proximal regulatory region, distal regulatory region and core enhancer region.[8] Although the transcriptional regulation mechanism of MyoD1 is well established, nevertheless, it committed years and manpower for this work.[9] Furthermore, there are a great many of other genes that need to be resolved their expression regulatory mechanisms in myogenesis.

A growing body of evidences have shown that the eukaryotic genome forms a highly ordered, hierarchical structure in the nucleus that closely correlates and may even be causally linked with transcriptional machinery to allow appropriate gene expression.[10,11] Deciphering the reorganization of chromatin architecture during skeletal muscle development is benefit for explain the complex molecular regulatory network that functions in myogenic differentiation.

Here, we conducted an integrative three-dimensional (3D) structure of genome and transcriptomic characterization of porcine longissimus dorsi muscle (LDM) from E35 to E80 for the first time. We aimed to construct the 3D atlas of pig skeletal muscle genome, to explore dynamic changes of genomic spatial structure during development and to analyse the regulatory relationship between genomic structure and gene expression during the stage of secondary myofibres formation.

2. Materials and methods

2.1 Ethics statement

All the animals used in this study were collected according to the guidelines for the care and use of experimental animals that established by the China Council on Animal Care and the Ministry of Agriculture of China. All experiments approved by the Animal Care and Use Committee of Guangdong Province, China. Approval ID and permit numbers are SCXK (Guangdong) 2011-0029 and SYXK (Guangdong) 2011-0112.

2.2 Animals and sample collection

LDM were collected from embryonic Day 35 (E35) and 80 (E80) of purebred Luchuan pigs (a Chinese indigenous breed). E35 and E80 embryos were from three and one sows, respectively. Muscle tissue of E35 for RNA-seq are from three female embryos that from three sows while for Hi-C and chromatin immunoprecipitation (ChIP)-seq are pooled all female embryos. Muscle tissue of E80 for RNA-seq, Hi-C and ChIP-seq are from three female embryos.

As the sexual characteristics of pig foetuses were not visible before Day 49 of gestation period, therefore, we using a PCR-based method to determine the gender of the embryos of E35 as previous.[12] The primer SRYB-5: 5′-TGAACGCTTTCATTGTGTGGTC-3′ and primer SRYB-3: 5′-GCCAGTAGTCTCTGTGCCTCCT-3′ used to amplify a 163-bp region of Sry that located on the Y chromosome. The primer STS-Bo1-5: 5′-GAGCAGACTCCAACTACTCTCAC TCCAC-3′ and primer STS-Bo1-3: 5′-TCAGAAGGATGATTTA GAGTGTCTGTTCAG-3′ used to amplify a 326-bp fragment of STS-Bo1 serving as an internal control.

2.3 Hi-C library preparation

In situ high-throughput chromatin conformation capture (Hi-C) was generated according previous described with some modifications.[11] Briefly, muscle tissue was homogenized and fixed with freshly made 1% formaldehyde solution at room temperature for 30 min. Chromatin were digested with 200 U *Dpn*II enzyme (R0543S, NEB, USA) at 37°C for 90 min, 65°C for 20 min and 25°C for 5 min. Nucleotide fill-in with 0.4 mM Biotin-14-dATP (19524-016, Invitrogen), 10 mM dCTP, 10 mM dGTP, 10 mM dTTP and 5 U/μL Klenow Fragment (M0210L, NEB) at 37°C for 45 min. Ligation was performed by a T4 DNA ligase (L6030-HC-L, Enzymatics, USA) at 20°C for 30 min. DNA was sheared to the length of 300–400 bp and washed by M280 beads at 20°C for 20 min. Hi-C libraries were amplified 10 PCR amplification cycles using mixed universal PCR primer and index primer. The resulting libraries were sequenced on BGISEQ-500 as 100 bp paired-end.

2.4 Hi-C data analysis

Hi-C reads were mapped pig genome (Sscrofa11.1, sex chromosomes are excluded) by BWA software (v.0.7.15) with default parameters. Unaligned read pairs, PCR duplicate paired reads and low-quality alignments defined as one or both reads failing to meet the threshold [Mapping Quality (MAPQ) $\geq$ 30] were removed by using Juicer (Supplementary Fig. S1A). The contact matrices used for further analysis were normalized by Knight–Ruiz algorithm and quantile algorithm. Normalized interaction matrices were separately generated at resolutions of 100, 20 and 5 kb.

2.5 Identification of compartments A and B

Compartments A/B were identified as previously described[13] with minor modifications. Primarily, defined the compartments A and B at 100 kb low resolution as previous.[11] Then constructed the observed/expected matrix at resolution of 20 kb. We defined O/E scores of bins that located in compartment A or B at low resolution as A score or B score, respectively. An A–B index was then created by subtracting the mean A and B scores. A–B indexes are positive values (more association with A) that were called as A compartments, while negative values (more association with B) were called as B compartments. For the compartment switching analysis, we considered only

regions where more than two out of three biological replicates showed changes in A–B index from positive to negative or vice versa.

2.6 Compartment strength

To evaluate the compartmentalization of each chromosome of the genome, we adopted the method as previous described[14] with minor modification. For saddle plot, we rearrange the A–B index (20 kb bin) from the lowest to the highest value for each chromosome. We then reshuffle the O/E map of the chromosome and divided the resulting map into a 50×50 matrix. The compartment strength was defined as (AA + BB)/2AB, where AA and BB are the mean level of interaction in regions belonging to the same compartment, while AB is mean level of interaction in regions belonging to different compartments. We chose the top 20% of strong A compartment and the top 20% of strong B compartment for measure.

2.7 Identification of topologically associating domains and specific topologically associating domain boundaries

Topological domains are called by combining directionality index (DI) score and InsulationScore (IS) at 20 kb resolution. First, we identified topological domains based on the DI score and a Hidden Markov Model as previously described,[15] then using IS[16] to identify and divide the large topologically associating domains (TADs) into more small TADs. The TAD boundaries were <400 kb, and unorganized chromatin regions were longer than 400 kb.

For specific boundary, we merged reads of three replicates of each stage for TAD calling (E35: 4,050 TADs; E80: 4,028 TADs). Boundaries were identified as 'specific' if the boundary regions were called in only one stage and lacked a significant correlation in DI between the two stages compared. The correlation coefficient of DIs between stages was performed as previously described.[15] We made a boundary as a centre and calculated spearman correlation of the DIs of ±10 bins for each centre between E35 and E80. Similarly, 20 bins randomly selected from the two stages and calculated the spearman correlation between the two vector as random correlation. The randomization war repeated 10,000 times to achieve the random distribution.

2.8 Domain score

To compute the change in interaction frequency between stages, we determined a consensus TAD as previous described,[17] which are conserved in at least 50% samples (3,654 TADs, Supplementary Fig. S4A). We calculated an intradomain contact score.[10,18] The *D*-score is the ratio of the number of contacts that occur between regions within the same TAD (intra-TAD contacts) over the total number of intrachromosomal contacts for a TAD (intra-TAD + inter-TAD). To identify TADs that have differential *D*-scores, we used two-tailed *t*-test to evaluate *D*-scores. TADs with an false discovery rate (FDR) <0.05 were selected for subsequent analysis.

2.9 Identification of promoter–enhancer interactions

To identify the promoter–enhancer interactions (PEIs), we merged replicates and called by the PSYCHIC[19] at a resolution of 5 kb. Which use the hierarchical TAD-specific bi-linear model as background model and identify over-represented DNA–DNA interactions according to a piecewise power law regression for each TAD. Then, we filtered promoter–promoter interactions and other interactions, retaining PEIs. We reserved the high-confidence PEIs (FDR ≤ 0.0001, length of PEI ≥ 20 kb) for later analysis. The promoter distal interactions would be defined as poised enhancers (PEs), regular enhancers (REs) and super-enhancers (SEs) if they overlapped without H3K27ac peak, at least 1 bp H3K27ac regular peaks or at least 2.5 kb (50%) H3K27ac SE peaks, respectively.

2.10 Regulatory potential score

To comprehensively consider the influence of multiple enhancers on the expression of a gene, so as to accurately elucidate dynamic PEIs architecture contribute to the functional transcriptomic divergence between E35 and E80, we introduce a regulatory potential score (RPS) for each gene that an enhancer's quantitative effect on a gene should depend on its spatial proximity. The RPS is calculated as $\sum n (\log_{10} I_n)$, of which, I_n indicate the normalized interaction intensity (observed value − expected value). If a promoter without enhancer interaction, the RPS = 0. We divided the differential PEIs in two categories: (i) different RPS: the RPS of gene were subtract between E35 and E80, then made a *Z*-score for delta values to gain normally distributed data, the *Z*-values >|mean ± 2 S.D.| were defined as significant difference and (ii) RPS that was not significantly different but with ≥4 differential active enhancers (RE and SE).

2.11 ChIP-seq and data analysis

ChIP assays were performed by Wuhan IGENEBOOK Biotechnology Co., Ltd. Briefly, muscle tissue samples were homogenized and cross-linked with 1% formaldehyde for 10 min at room temperature and then quenched by 125 mM glycine. After samples lysed on ice, chromatins were obtained and sheared to length of 200–500 bp by sonication. Five micrograms H3K27AC antibodies (ab4729, Abcam) were used in the immunoprecipitation reactions at 4°C overnight. The next day, 30 μL of protein beads was added and further incubated for 3 h. Then the beads were washed with 20 mM Tris/HCL (pH 8.1), 50 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS for once; 10 mM Tris/HCL (pH 8.1), 250 mM LiCl, 1 mM EDTA, 1% NP-40, 1% deoxycholic acid for twice; and TE buffer 1× (10 mM Tris-Cl at pH 7.5. 1 mM EDTA) for twice. The DNA fragments were eluted with 100 mM NaHCO3, 1% SDS. Purified DNA was used to construct libraries by performing the I NEXTFLEX® ChIP-Seq Library Prep Kit according to the manufacturer's instructions (NOVA-5143-02, Bioo Scientific) and sequenced on Illumina Hiseq X ten as 150 bp paired-end.

Raw reads were filtered by quality score and aligned to the pig genome (Sscrofa11.1). Unique aligned and deduplicated reads were used for peak calling. MACS2 was used to determine H3K27ac peak with FDR cut-off 0.05 (Supplementary Fig. S1C). To identify the SE, we used the ROSE (Ranked Order of Super Enhancers) software with default parameters. This algorithm stitches constituent enhancers together if they are within a certain distance (within 12.5 kb) and ranks the enhancers by all H3K27ac ChIP signal. It then separates SEs from REs by identifying an inflection point of H3K27ac signal versus enhancer rank.

2.12 RNA-seq and expression analysis

Muscle tissue was homogenized with the Trizol reagent, and total RNA was extracted according to the manufacture's protocols. Purified RNA was eluted with RNase-free water and quantified by Nanodrop. Strand-specific RNA libraries were constructed with Ribo-Zero kit (RZH1046, Epicentre) and sequenced on Illumina HiSeq X Ten as 150 bp paired-end. Raw read pairs were aligned to

the pig reference genome (Sscrofa11.1) using Hisat2 (v.2.0.4). Stringtie (v.1.3.3) was applied to quantify mRNA expression. Normalized read counts are reported as transcripts per million (TPM) (Supplementary Fig. S1B).

2.13 Functional enrichment analysis

For functional enrichment analysis, we mapped pig genes to their human (*Homo sapiens*) orthologs and choose human as the target analysis specie. The selected genes (TPM > 0.5 at least in a sample) were analysed using Metascape (http://metascape.org May 2, 2021, date last accessed) with default parameters.[20] Enrichment analysis were performed against all genes in the genome as background set with gene ontology-biological processes as ontology sources. The most statistically significant 10 terms in each cluster were depicted using $-\log_{10}$ (*P*-value) bar plots.

2.14 Statistical analysis

All statistical analyses were performed by two-tailed Student's *t*-test or Wilcoxon rank-sum test using GraphPad Prism8 and RStudio. All data were expressed as means ± s.d.

3. Result

3.1 Dynamic chromatin compartment in skeletal muscle development

The spatial organization of chromosomes is intimately linked to cell physiology. To study chromatin organization in mammalian muscle cells, we constructed *in situ* Hi-C libraries for porcine LDM of E35 and E80, separately (Fig. 1A). A total of six libraries generated ~5,394.27 million (M) validly aligned contacts with a depth of ~889.04 M per library that reach a maximum resolution of 5 kb (~90.30% bins have at least 1,000 reads) (Supplementary Fig. S1A and D). After normalization, the contact maps of both stages showed highly reproducible (QuASAR-Rep score > 0.92) (Supplementary Fig. S1E), which consistent with RNA-seq data (Pearson's $r > 0.97$) (Supplementary Fig. S1F and G). The contact possibility is declined with the genome distance increases (Supplementary Fig. S1H) and most of contacts are occur in 20 Mb (~81.41%) (Supplementary Fig. S1I).

Each chromosome be decomposed into compartment regions at a resolution of 20 kb, including accessible compartment A (1,099.21 Mb, ~48.51% of the genome, which are gene and GC-rich, actively transcribed) and densely packed compartment B (1,166.75 Mb, ~51.49% of the genome, which are gene and GC-poor, inactively transcribed) (Supplementary Fig. S2A and B) that presented as plaid patterns in chromatin interaction (Supplementary Fig. S2C) or the derived correlation heatmaps (Fig. 1B). Compared to E80, E35 showed blurred plaid patterns (Fig. 1B and Supplementary Fig. S2C) and a less well-segregated compartment (Supplementary Fig. S2D), which as well supported by the intensive compartment strength in E80 than that in E35 (Fig. 1C and Supplementary Fig. S2E).

As expected, during development, compartment statues of 4.39% genomes were switched from A to B, concomitant with decreased gene expression ($n = 329$); while 7.04% of compartments were switched from B to A, concomitant with increased gene expression ($n = 1,284$) (Fig. 1D and E). Functional enrichment showed that genes with compartments change from A to B were mainly associated with development and growth, such as the pathways of 'embryonic limb morphogenesis', 'regulation of neuron differentiation' and 'negative regulation of cell proliferation' (Fig. 1F). *NEFL* (Fig. 1H) encodes the core subunits of neurofilament, whose mutations are associated with Charcot–Marie–Tooth neuropathy[21,22] and depletion is sufficient to cause mild sensorimotor dysfunctions and spatial deficits.[23] Genes with compartments switched from B to A were involved in metabolism relative pathways, including 'carboxylic acid biosynthetic process' and 'proteolysis involved in cellular protein catabolic process', suggest skeletal muscles have the function of energy metabolism at their foetal development (Fig. 1G). Typically, *SLN* (Fig. 1I), which encodes a protein sarcolipin that is necessary for muscle-based non-shivering thermogenesis by regulating the sarco/endoplasmic reticulum Ca^{2+}-ATPase pump in skeletal muscle.[24] In addition, SLN promotes oxidative metabolism in skeletal muscle in response to increased metabolic demand.[25] Brown adipose tissue is the significantly important heat production tissue that also dependent on non-shivering thermogenesis via the activation of uncoupling protein 1 (*UCP1*) in human and mouse.[26] Whereas brown adipose tissue and *UCP1* gene both are lacked in pig.[27,28] The compartment switching from B to A and increased expression of *SLN* in skeletal muscle may indicate that skeletal muscle-dependent non-shivering thermogenesis probably is the main way for piglets to adapt to low temperature environment. In addition, the pathways related to development were enriched, including 'muscle structure development' and 'osteoblast differentiation' (Fig. 1G). In which, *MYL1* (Fig. 1J) encodes the skeletal muscle fast-twitch specific myosin, and its mutations could result in conspicuously reduced in skeletal muscle; knockdown of *MYL1* in zebrafish disrupted muscle structure and impaired touch-evoked escape responses.[29]

3.2 Dynamic intra-TAD interactions are associate with transcription and compartmentalization

TAD is the structural unit of chromosome that involved in transcription, which displays high frequencies of physical interaction within a given domain but lower frequencies outside of these domains.[15] We called TADs (E35: ~4,074 TADs with size of 460 kb, E80: ~3,965 TADs with size of 440 kb) in resolution of 20 kb (Supplementary Fig. S3A). The accuracy of TAD boundaries was evidenced by the aggregate analysis of DI and IS (Supplementary Fig. S3B and C), and CTCF binding regions are high enrichment in TAD boundaries (Supplementary Fig. S3D). We found that TAD boundaries are largely invariable [Pearson's *r* (DI) > 0.96, Pearson's *r* (IS) > 0.94, Jaccard index > 0.98] between E35 and E80 (Supplementary Fig. S3 E–G). We only identified ~2% differential TAD boundaries between E35 and E80 (Supplementary Fig. S3H), indicating that TAD is a conserved chromatin structure, which is consistent with previous reports.[10,15]

Although TAD boundaries are stable, the contacts within TADs are dynamic. *D*-score was used to explain interactions within a given TAD, defined by the ratio of intra-TAD contacts over all intrachromosomal contacts for a TAD.[18,30] It is positively correlative (Pearson's $r > 0.99$, $P < 2.2 \times 10^{-16}$) to A–B index and with increase of *D*-score the gene expression showed the tendency of increase (Supplementary Fig. S4B and C). There are 531 significantly differential *D*-score TADs ($P \leq 0.001$, two-tailed *t*-test) between E35 and E80 (Fig. 2A). Of which, 483 TADs (1,417 genes) with higher *D*-score showed more active compartment status (higher A–B index and gene expression) in E35 than that in E80 (Fig. 2B and C). Compared to E35, E80 has 48 TADs (133 genes) with higher *D*-score indicating more active compartment status (Fig. 2B and C).

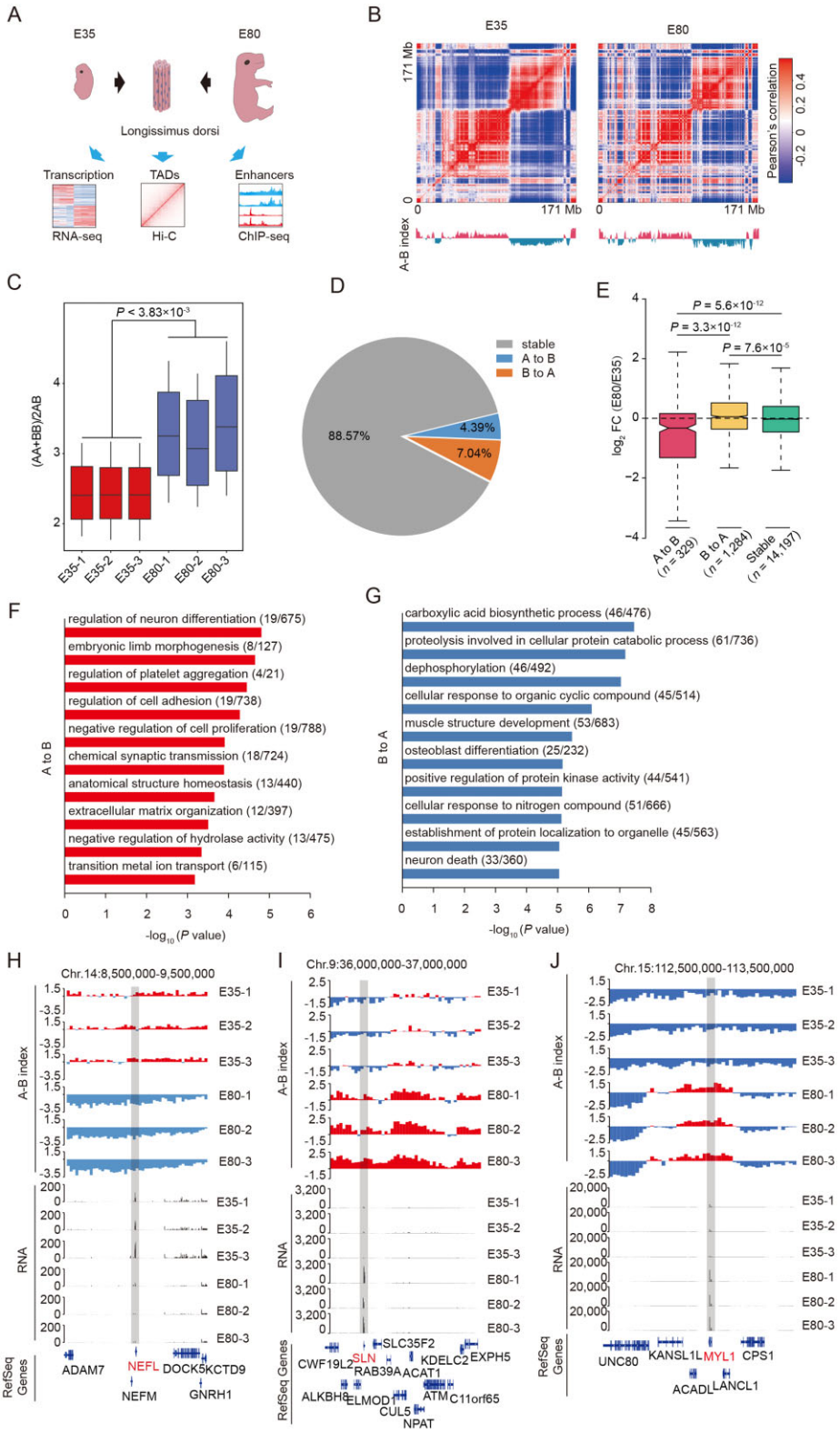

Figure 1. Changes of chromatinic compartments A/B. (A) Schematic overview of the study. (B) Correlation heatmaps show Pearson correlations for their intra-chromosomal interaction frequency patterns (100 kb bin; chromosome 6). (C) Boxplots show the compartment strength for average interaction frequency between the same classes of compartments (AA and BB) compare to those between different classes of compartments (AB) for each euchromosome (see Materials and methods). *P*-values were calculated by Student's *t*-test (two-tailed). (D) The proportion of the genome that changed compartments A/B status between E35 and E80. (E) The expression changes for genes that compartments were switched. *P*-values were calculated by Wilcoxon rank-sum test. (F and G) Function enrichment for genes (TPM $>$ 0.5) with compartment status were switched. (H–J) Genome browser shots show the compartments status and gene expressions at locales of Chr.14: 8.5–9.5 Mb (*NEFL*, H), Chr.9: 36–37 Mb (*SLN*, I) and Chr.15: 112.5–113.5 Mb (*MYL1*, J).

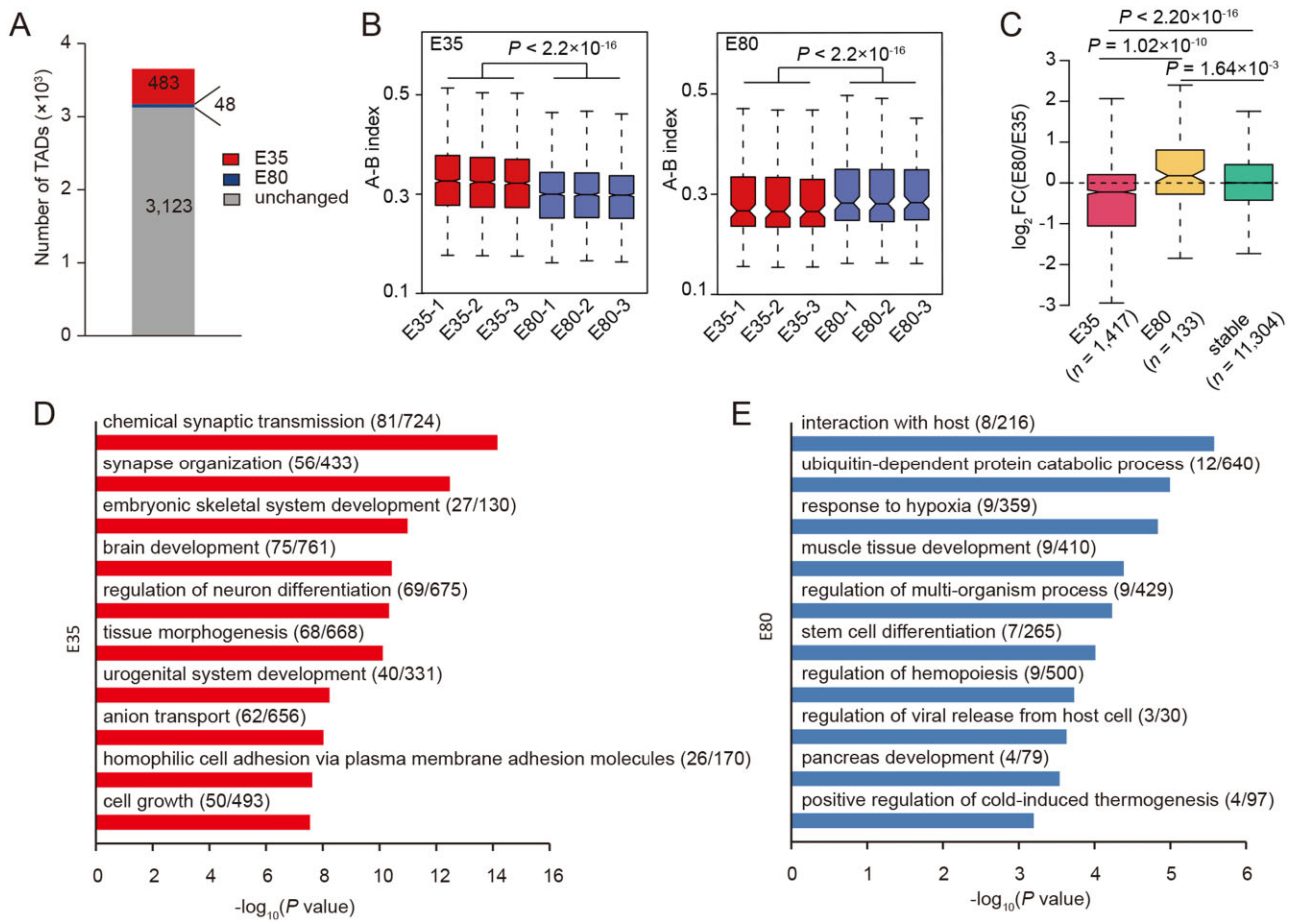

Figure 2. *D*-score dynamics within topological domain during development. (A) Number of significantly differential TADs (two-tailed *t*-test) between E35 and E80. (B) A–B indexes of significantly differential TADs between E35 and E80. *P*-values calculated by Wilcoxon rank-sum test. (C) The average expression changes for genes within TADs with significantly differential *D*-score. The number of genes in TAD are showed. *P*-values calculated by Wilcoxon rank-sum test. (D and E) Function enrichment for genes of significantly differential *D*-score.

Genes that located in TADs with higher *D*-score in E35 were enriched in morphogenesis and synapse-associated pathways (e.g. 'chemical synaptic transmission', 'synapse organization' and 'embryonic skeletal system development') (Fig. 2D). While located in TADs with higher *D*-score in E80, these genes were enriched in skeletal muscle development and metabolism (e.g. 'muscle tissue development') (Fig. 2E).

3.3 Rewiring of PEI in skeletal muscle development

Promoter interactions are considered closely associated with gene expression regulation.[31,32] To uncover an extensive genome-wide catalogue of PEIs in skeletal muscle and define how they response to muscle development, we found 48,909 and 54,350 high-confidence PEIs (FDR $\leq$ 0.0001) overlapping 13,548 and 12,993 genes by PSYCHIC[19] in E35 and E80 at resolution of 5 kb (three libraries were merged and reached 99.54% of bins have at least 1,000 reads), respectively (Supplementary Fig. S5).

Gene expression is positively correlated with the number of promoter interactions.[32,33] We observed that highly expressed genes interacted with more enhancers (Fig. 3A), suggesting additive enhancer effects on target-gene transcription level. To explore the regulatory effects of multi-enhances on a gene, we introduced RPS, which is a measure for the sum of interactions of a gene regulated by multi-enhancers depending on spatial proximity. Indeed, the increase of gene expression is concomitant with enhanced RPS (Fig. 3B). Furthermore, the activity of enhancer is another crucially regulative factor for gene expression.[34,35] H3K27ac are used to annotate the activated enhancers dividing into RE and SE,[36] while enhancer without H3K27ac is defined to PE. As expected, RPS is positively correlated with H3K27ac enrichment ($r > 0.8$, Supplementary Fig. S6A and B). The promoter of gene that interact with SEs showed larger RPS and higher expression than promoters interacting with RE followed by promoters interacting with PE (Fig. 3C and Supplementary Fig. S6C and D).

There are 951 genes (RPS > |mean ± 2 S.D|) were identified that are rewired PEIs for responding to muscle development. Of which, 307 genes lessened interactions accompanied by gene expression decrease, whereas 644 genes gain more interactions correlate with increasing gene expression during muscle development (Fig. 4A). Because RPS does not include the effect of enhancer activity on gene expression, we examined H3K27ac enrichment on the distal enhancer that interacted with gene promoter that RPS are not differential, and identified 1,533 genes that were H3K27ac high enrichment at distal enhancers in E35, while 246 genes were H3K27ac more enrichment at distal enhancers in E80 [|active enhancer$_{(E80)}$ − active enhancer$_{(E35)}$ | $\geq$ 4] (Fig. 4B). The activation of enhancers that

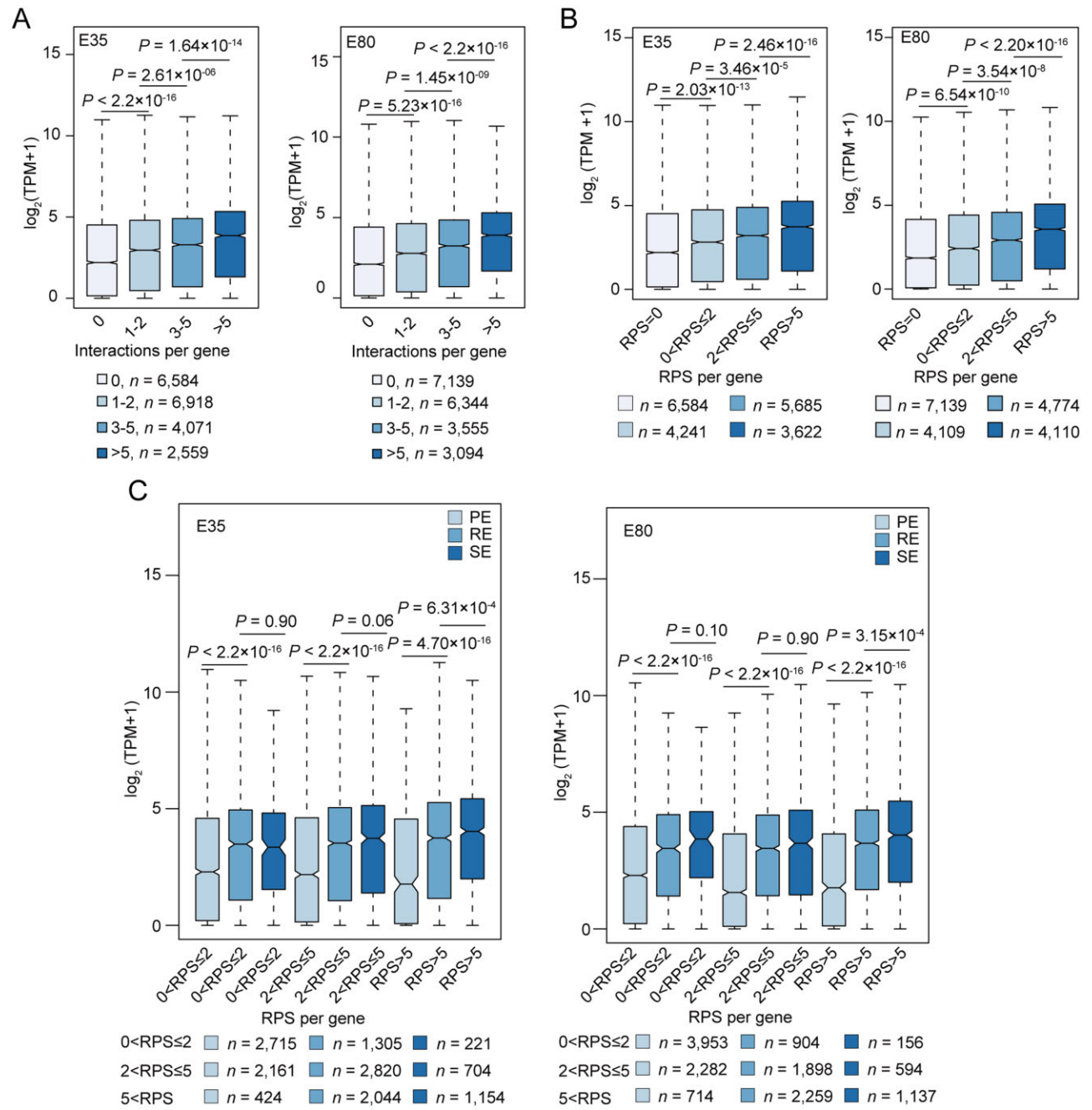

Figure 3. PEIs regulate gene expression. (A) Expression of genes with different contacted enhancer numbers. Enhancer numbers were divided into four groups: number = 0, 1–2, 3–5 and >5. (B) Expression of genes with different RPS, genes with different RPS were divided into three groups: 0< RPS ≤2, 2< RPS ≤5 and RPS >5. (C) Expression of genes with different RPS with PE, RE or SE.

interacted with promoter is significantly positively correlated with gene expression (Fig. 4B).

Next, we merged that RPS and enhancer activity are increased genes and RPS and activity are reduced genes from E35 to E80, respectively. Function enrichment analysis showed that categories are consistent with that of enrichment in genes with changed compartment A/B and *D*-score. Genes with high RPS and enhancer activity in E35 were mainly enriched in pathways related to morphology and neuron differentiation, including 'embryonic morphogenesis', 'tissue morphogenesis' and 'regulation of neuron differentiation' (Fig. 4C). MuSK is a key gene for neuromuscular junction (NMJ) development; its knockout disrupts the formation of NMJ.[37,38] It is highly expressed in primary myotubes while is dramatically down-regulated in skeletal muscle.[39] Compared with E35, E80 presents decreased expression level for *MuSK*, which accompanied by reducing enhancer activation (Fig. 4D). As well as *MuSK*, *MAP1B*,[40,41] *NNAT*,[42,43] *NCALD*,[44] *ZIC1*[45,46] and *ZIC4*[47,48] are associated with neuron development. Their promoters are interacting with more spatially closer enhancers in E35 than E80 (Supplementary Fig. S7A–D).

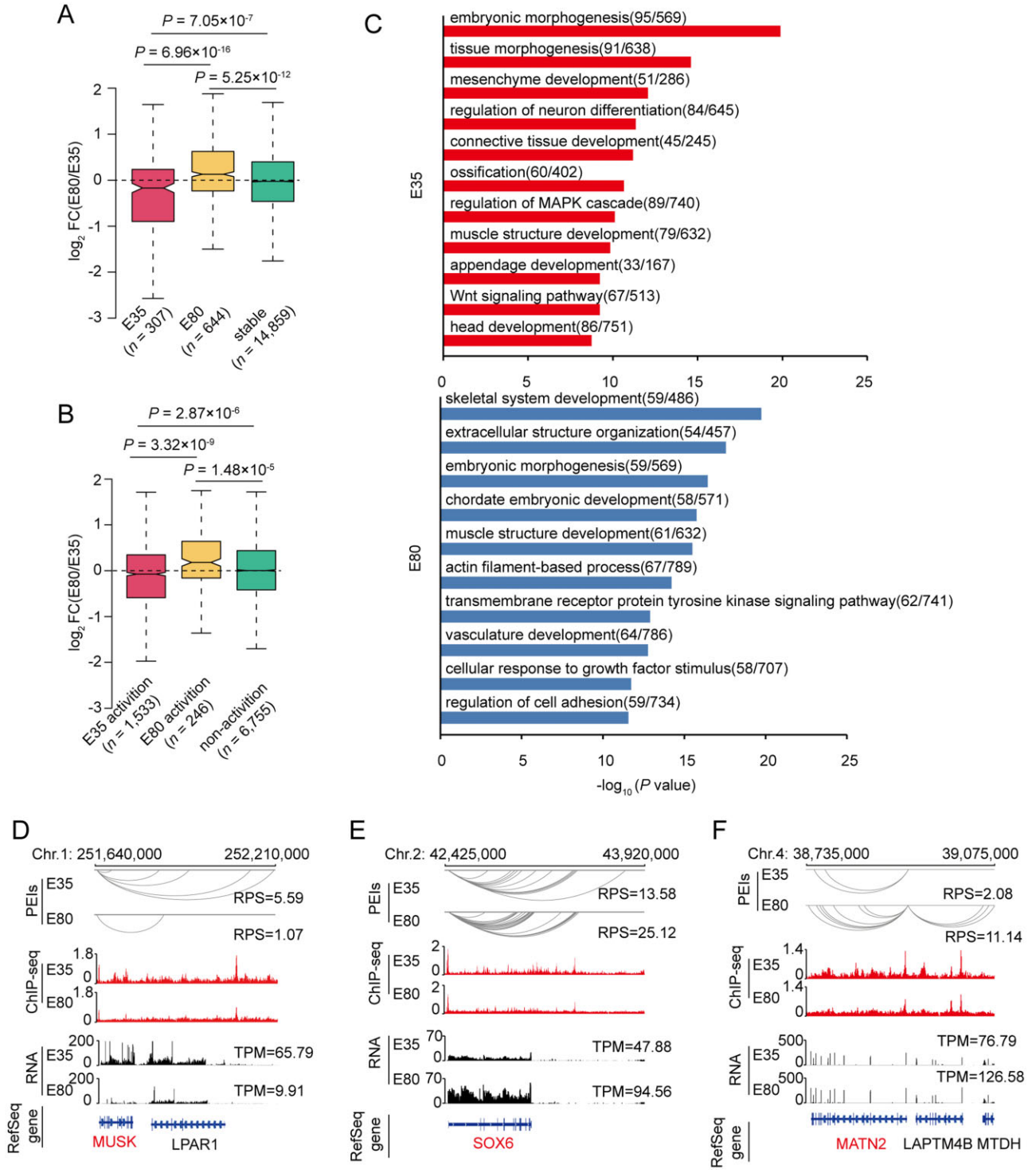

Figure 4. PEIs rewiring between E35 and E80. (A) The expression changes for genes with changed RSP. (B) The expression changes for genes with stable RPS, but changed distal enhancer activation [|active enhancer$_{(E80)}$ − active enhancer$_{(E35)}$| ≥ 4]. (C) Gene ontology enrichment for genes that changed RPS and (or) distal enhancer activation. (D–F) Genome browser view of contact frequency, PEIs, H3K27ac and transcript abundance at the *MuSK* (D), *Sox6* (E) and *MANT2* (F) locus. The RPS and TPM of represent genes are showed.

Genes with high RPS and enhancer activity in E80 were enriched in 'skeletal system development', 'extracellular structure organization', 'muscle structure development' and 'actin filament-based process' (Fig. 4C), such as *DAG1*,[49] *TOMD1*,[50] *MYH8*,[51] *SYNM*,[52] *LAMA2*[53,54] that encode cellular structural proteins in muscle cells (Supplementary Fig. S8A–E); *CAPN3*,[55,56] *PLN*,[57,58] *STIM1*[59,60] that regulate muscle function by Ca^{2+} transport mechanism (Supplementary Fig. S8F–H); *GCK*,[61] *ACO2*[62,63] that encode metabolic enzymes in tricarboxylic acid cycle (Supplementary Fig. S8I and J); *IGF2*[64–66] and *Mef2D*[67,68] that are associated with growth and

development of skeletal muscle (Supplementary Fig. S8K and L). These genes are interacting with more enhancers in E80 than E35. Interestingly, of which, we observed *Sox6*[69,70] and *MATN2*[71] (Fig. 4E and F). They play a crucial role in the timely induction of the global switching in myogenesis, which indicated that more interacting enhancers, enhancers activity and higher gene expression on Sox6 and MATN2 in E80 than E35 correlate with in myoblasts differentiated to secondary to myofibres.

4. Discussion

Previous report had shown that the total number of myofibres is constant after birth,[4] suggesting hyperplasia of myofibre is of prime importance for the muscle mass. Myofibres formation could be divide into two stages in pigs, including primary and secondary fibres formations.[6] We first perform Hi-C to investigate the change of genome architecture from primary to secondary fibre formations in this study.

The researches of early mammalian development had shown that chromatin compartments are gradually established.[14,72] In contrast with previous reports, we found the chromatin compartment almost established before E35, although we observed the blurred plaid in E35 than E80 (Fig. 1B and Supplementary Fig. S2C), which may consistent with the more active progenitors in E35 than in E80. It was found that ~11.43% of compartments were switched between E35 and E80 (Fig. 1D), which is comparable with C2C12 induced differential differentiation of 3 days (8%)[73] and 8 days.[74] Compartment changes are related to gene transcription, but not all genes whose compartment are changed their transcription levels, or vice versa. This is due to compartment is the large genomic structure that hard to fine regulation of gene expression.

As well as compartment A/B, TADs are gradually established in the mammalian embryogenesis.[14,72] However, after the establishment and consolidation of TAD, its stable and conserved across different cell types and species.[15] Of course, the number and size of TAD are significantly different if performed various tools or resolutions.[75] We identified that the number and size of TADs are comparable between E35 and E80 (Supplementary Fig. S3A). Furthermore, over 98% of TADs boundaries were stable (Supplementary Fig. S3H), consistent with the stability of TADs.[15] Although TADs are highly stable, however, the interaction frequencies of intra-TADs are dynamic.[10,30] Increase or decrease of intra-TADs interaction frequencies is associated with up-regulation or down-regulation of gene expressions,[10] respectively, indicating there exist a precise architecture (such as PEI) for gene regulation.

PEIs are the direct architecture to finely regulate gene expression. Nevertheless, it requires a large amount of data to reach an extremely high resolution. We merged our data of replicates and identified ~50,000 PEIs with an extremely rigid cut-off (FDR $\leq$ 0.0001) under resolution of 5 kb (Supplementary Fig. S5A and B). In spite of the rigid criterion, over half of genes are interacting with enhancer, indicating enhancer is a universal and important element in gene regulation. Usually, the enhancer number is positively correlate with expression level,[32] showing the enhancer additivity,[76,77] such as *Sox6* interacted with more enhancers in E80 compared with that in E35 (Fig. 4E). *SVEP1* and *LMNA* have comparable PEIs (or RPS) and their mRNA expression are not differential (Supplementary Fig. S9). However, some gene interacted with different amounts of enhancer while their expression were not changed, showing the enhancer redundancy.[78] For example, the promoter of *MyoD1* separately interacted with 10 (RPS: 12.24) and 17 (RPS: 19.06) enhancers in E35 and E80, whereas the gene expressions (mean TPM, E35: 70.17, E80: 89.83) were not significantly differential. In addition, enhancer activation is another significant factor that affect gene expression. As enhancer number, spatial proximity and enhancer activation are always co-regulate gene expression, however, there lack an algorithm that can integrate all these factors. Therefore, we ignored other factors when analyse one of them. We identified mounts of genes that may play significant role in skeletal development, that were regulated by differentially active enhancer, such as *MuSK* (Fig. 4D).

Primary myoblasts express high levels of MyHC-I and MyHC-emb, while secondary myoblasts express low levels of MyHC-I and high levels of MyHC-neo.[79–81] Nfix is highly expressed in foetal myoblasts and acts as a transcriptional switch from embryonic to foetal myogenesis.[82] We observed that the promoter of Nfix interacting with SEs in E80 but not in E35. However, the Nfix's mRNA expression are comparable between E35 and E80, which may result from the heterogeneity of muscle tissue (data not shown). Sox6 is another gene that required for foetal fibre type differentiation.[69,70] It cooperates with Nfix directly inhibit the expression of MyHC-I in foetal muscle[83] and activate more genes that expressed specifical in secondary myofibres (such as CKM, β-enolase).[82] In addition, MATN2 is an extracellular matrix protein in controlling the early stages of myogenic differentiation.[71] Nfix directly regulates MATN2 expression and MATN2 positive-feedback regulates Nfix expression.[71] In this study, we observed that chromatin interactions and mRNA expression of *Sox6* and *MATN2* both are increased from E35 to E80 (Fig. 4E and F). These results indicated that MATN2 as extracellular regulatory signals and Sox6-Nfix as intracellular transcription factor complex drive the PEIs rewiring of genes in myoblasts and determine myoblasts differentiated into secondary myofibres.

In summary, we revealed the changes of genome architecture from compartment A/B, TAD and PEI three scales in early muscle development and confirmed dynamic spatial reorganization is consistent with gene expression modulation. The most important is that we constructed the high-resolution interaction maps (5 kb resolution), which could provide genome-wide atlases of promoter interaction for studies on the regulation mechanism of skeletal muscle development.

Supplementary data

Supplementary data are available at *DNARES* online.

Funding

This work was supported by grants from the National Key R & D Program of China (2020YFA0509500), the National Natural Science Foundation of China (32072697, 31902134, 31772576 and U19A2036), the Science and Technology Projects of Zhanjiang (2019A01004), and the Sichuan Science and Technology Program (2021YFYZ0009 and 2021YFYZ0030). We appreciate High-performance Computing Platform of Sichuan Agricultural University for providing the powerful support for data analysis.

Data availability

All the sequencing raw data of Hi-C, RNA-seq and ChIP-seq have been deposited to Sequence Read Archive (SRA) under accession number: SRP305313 and the processed data files have been deposited to GEO under accession number: GSE166346.

Conflict of interest

All authors declare that they have no conflict of interest.

References

1. Frontera, W.R. and Ochala, J. 2015, Skeletal muscle: a brief review of structure and function, *Calcif. Tissue Int.*, **96**, 183–95.
2. Chang, N.C., Chevalier, F.P. and Rudnicki, M.A. 2016, Satellite cells in muscular dystrophy—lost in polarity, *Trends Mol. Med.*, **22**, 479–96.
3. Bottinelli, R. 2001, Functional heterogeneity of mammalian single muscle fibres: do myosin isoforms tell the whole story? *Pflügers Arch.*, **443**, 6–17.
4. Ashmore, C.R., Addis, P.B. and Doerr, L. 1973, Development of muscle fibers in the fetal pig, *J. Anim. Sci.*, **36**, 1088–93.
5. Buckingham, M. and Rigby, P.W. 2014, Gene regulatory networks and transcriptional mechanisms that control myogenesis, *Dev. Cell*, **28**, 225–38.
6. Picard, B., Berri, C., Lefaucheur, L., Molette, C., Sayd, T. and Terlouw, C. 2010, Skeletal muscle proteomics in livestock production, *Brief Funct. Genomics*, **9**, 259–78.
7. Hernandez-Hernandez, J.M., Garcia-Gonzalez, E.G., Brun, C.E. and Rudnicki, M.A. 2017, The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration, *Semin. Cell Dev. Biol.*, **72**, 10–8.
8. Chen, J.C. and Goldhamer, D.J. 2004, The core enhancer is essential for proper timing of MyoD activation in limb buds and branchial arches, *Dev. Biol.*, **265**, 502–12.
9. Wardle, F.C. 2019, Master control: transcriptional regulation of mammalian Myod, *J. Muscle Res. Cell Motil.*, **40**, 211–26.
10. Dixon, J.R., Jung, I., Selvaraj, S., et al. 2015, Chromatin architecture reorganization during stem cell differentiation, *Nature*, **518**, 331–6.
11. Rao, S.S., Huntley, M.H., Durand, N.C., et al. 2014, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping, *Cell*, **159**, 1665–80.
12. Li, M., Xia, Y., Gu, Y., et al. 2010, MicroRNAome of porcine pre- and postnatal development, *PLoS One*, **5**, e11541.
13. Rowley, M.J., Nichols, M.H., Lyu, X., et al. 2017, Evolutionarily conserved principles predict 3D chromatin organization, *Mol. Cell*, **67**, 837–52.e7.
14. Ke, Y., Xu, Y., Chen, X., et al. 2017, 3D Chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis, *Cell*, **170**, 367–81.e20.
15. Dixon, J.R., Selvaraj, S., Yue, F., et al. 2012, Topological domains in mammalian genomes identified by analysis of chromatin interactions, *Nature*, **485**, 376–80.
16. Crane, E., Bian, Q., McCord, R.P., et al. 2015, Condensin-driven remodelling of X chromosome topology during dosage compensation, *Nature*, **523**, 240–4.
17. Paulsen, J., Liyakat Ali, T.M., Nekrasov, M., et al. 2019, Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation, *Nat. Genet.*, **51**, 835–43.
18. Krijger, P.H., Di Stefano, B., de Wit, E., et al. 2016, Cell-of-origin-specific 3D genome structure acquired during somatic cell reprogramming, *Cell Stem Cell*, **18**, 597–610.
19. Ron, G., Globerson, Y., Moran, D. and Kaplan, T. 2017, Promoter-enhancer interactions identified from Hi-C data using probabilistic models and hierarchical topological domains, *Nat. Commun.*, **8**, 2237.
20. Zhou, Y., Zhou, B., Pache, L., et al. 2019, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets, *Nat. Commun.*, **10**, 1523.
21. Zuchner, S., Vorgerd, M., Sindern, E. and Schroder, J.M. 2004, The novel neurofilament light (NEFL) mutation Glu397Lys is associated with a clinically and morphologically heterogeneous type of Charcot-Marie-Tooth neuropathy, *Neuromuscul. Disord.*, **14**, 147–57.
22. Sainio, M.T., Ylikallio, E., Maenpaa, L., et al. 2018, Absence of NEFL in patient-specific neurons in early-onset Charcot-Marie-Tooth neuropathy, *Neurol. Genet.*, **4**, e244.
23. Dubois, M., Strazielle, C., Julien, J.P. and Lalonde, R. 2005, Mice with the deleted neurofilament of low molecular weight (Nefl) gene: 2. Effects on motor functions and spatial orientation, *J. Neurosci. Res.*, **80**, 751–8.
24. Bal, N.C., Maurya, S.K., Sopariwala, D.H., et al. 2012, Sarcolipin is a newly identified regulator of muscle-based thermogenesis in mammals, *Nat. Med.*, **18**, 1575–9.
25. Pant, M., Bal, N.C. and Periasamy, M. 2016, Sarcolipin: a key thermogenic and metabolic regulator in skeletal muscle, *Trends Endocrinol. Metab.*, **27**, 881–92.
26. Lidell, M.E., Betz, M.J., Dahlqvist Leinhard, O., et al. 2013, Evidence for two types of brown adipose tissue in humans, *Nat. Med.*, **19**, 631–4.
27. Hou, L.J., Shi, J., Cao, L.B., Xu, G.L., Hu, C.Y. and Wang, C. 2017, Pig has no uncoupling protein 1, *Biochem. Biophys. Res. Commun.*, **487**, 795–800.
28. Trayhurn, P., Temple, N.J. and Van Aerde, J. 1989, Evidence from immunoblotting studies on uncoupling protein that brown adipose tissue is not present in the domestic pig, *Can. J. Physiol. Pharmacol.*, **67**, 1480–5.
29. Ravenscroft, G., Zaharieva, I.T., Bortolotti, C.A., et al. 2018, Bi-allelic mutations in MYL1 cause a severe congenital myopathy, *Hum. Mol. Genet.*, **27**, 4263–72.
30. Stadhouders, R., Vidal, E., Serra, F., et al. 2018, Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming, *Nat. Genet.*, **50**, 238–49.
31. Mifsud, B., Tavares-Cadete, F., Young, A.N., et al. 2015, Mapping long--range promoter contacts in human cells with high-resolution capture Hi-C, *Nat. Genet.*, **47**, 598–606.
32. Schoenfelder, S., Furlan-Magaril, M., Mifsud, B., et al. 2015, The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements, *Genome Res.*, **25**, 582–97.
33. Qin, Y., Grimm, S.A., Roberts, J.D., Chrysovergis, K. and Wade, P.A. 2020, Alterations in promoter interaction landscape and transcriptional network underlying metabolic adaptation to diet, *Nat. Commun.*, **11**, 962.
34. Rubin, A.J., Barajas, B.C., Furlan-Magaril, M., et al. 2017, Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation, *Nat. Genet.*, **49**, 1522–8.
35. Phanstiel, D.H., Van Bortle, K., Spacek, D., et al. 2017, Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development, *Mol. Cell*, **67**, 1037–48.e6.
36. Perez-Rico, Y.A., Boeva, V., Mallory, A.C., et al. 2017, Comparative analyses of super-enhancers reveal conserved elements in vertebrate genomes, *Genome Res.*, **27**, 259–68.
37. Kim, N. and Burden, S.J. 2008, MuSK controls where motor axons grow and form synapses, *Nat. Neurosci.*, **11**, 19–27.
38. DeChiara, T.M., Bowen, D.C., Valenzuela, D.M., et al. 1996, The receptor tyrosine kinase MuSK is required for neuromuscular junction formation in vivo, *Cell*, **85**, 501–12.
39. Valenzuela, D.M., Stitt, T.N., DiStefano, P.S., et al. 1995, Receptor tyrosine kinase specific for the skeletal muscle lineage: expression in embryonic muscle, at the neuromuscular junction, and after injury, *Neuron*, **15**, 573–84.
40. Gordon-Weeks, P.R. and Fischer, I. 2000, MAP1B expression and microtubule stability in growing and regenerating axons, *Microsc. Res. Tech.*, **48**, 63–74.
41. Sole, L., Wagnon, J.L., Akin, E.J., Meisler, M.H. and Tamkun, M.M. 2019, The MAP1B binding domain of Na(v)1.6 is required for stable expression at the axon initial segment, *J. Neurosci.*, **39**, 4238–51.
42. Lin, H.H., Bell, E., Uwanogho, D., et al. 2010, Neuronatin promotes neural lineage in ESCs Via Ca2+ signaling, *Stem Cells*, **28**, 1950–60.
43. Wijnholds, J., Chowdhury, K., Wehr, R. and Gruss, P. 1995, Segment-specific expression of the neuronatin gene during early hindbrain development, *Dev. Biol.*, **171**, 73–84.

44. Upadhyay, A., Hosseinibarkooie, S., Schneider, S., et al. 2019, Neurocalcin delta knockout impairs adult neurogenesis whereas half reduction is not pathological, *Front. Mol. Neurosci.*, **12**, 19.
45. Aruga, J., Tohmonda, T., Homma, S. and Mikoshiba, K. 2002, Zic1 promotes the expansion of dorsal neural progenitors in spinal cord by inhibiting neuronal differentiation, *Dev. Biol.*, **244**, 329–41.
46. Inoue, T., Ota, M., Ogawa, M., Mikoshiba, K. and Aruga, J. 2007, Zic1 and Zic3 regulate medial forebrain development through expansion of neuronal progenitors, *J. Neurosci.*, **27**, 5461–73.
47. Elsen, G.E., Choi, L.Y., Millen, K.J., Grinblat, Y. and Prince, V.E. 2008, Zic1 and Zic4 regulate zebrafish roof plate specification and hindbrain ventricle morphogenesis, *Dev. Biol.*, **314**, 376–92.
48. Gaston-Massuet, C., Henderson, D.J., Greene, N.D. and Copp, A.J. 2005, Zic4, a zinc-finger transcription factor, is expressed in the developing mouse nervous system, *Dev. Dyn.*, **233**, 1110–5.
49. Cohn, R.D., Henry, M.D., Michele, D.E., et al. 2002, Disruption of DAG1 in differentiated skeletal muscle reveals a role for dystroglycan in muscle regeneration, *Cell*, **110**, 639–48.
50. Ono, Y., Schwach, C., Antin, P.B. and Gregorio, C.C. 2005, Disruption in the tropomodulin1 (Tmod1) gene compromises cardiomyocyte development in murine embryonic stem cells by arresting myofibril maturation, *Dev. Biol.*, **282**, 336–48.
51. Schiaffino, S., Rossi, A.C., Smerdu, V., Leinwand, L.A. and Reggiani, C. 2015, Developmental myosins: expression patterns and functional significance, *Skelet. Muscle*, **5**, 22.
52. Garcia-Pelagio, K.P., Muriel, J., O'Neill, A., et al. 2015, Myopathic changes in murine skeletal muscle lacking synemin, *Am. J. Physiol. Cell Physiol.*, **308**, C448–62.
53. Helbling-Leclerc, A., Zhang, X., Topaloglu, H., et al. 1995, Mutations in the laminin alpha 2-chain gene (LAMA2) cause merosin-deficient congenital muscular dystrophy, *Nat. Genet.*, **11**, 216–8.
54. Xu, X.X., Rock, C.O., Qiu, Z.H., Leslie, C.C. and Jackowski, S. 1994, Regulation of cytosolic phospholipase A2 phosphorylation and eicosanoid production by colony-stimulating factor 1, *J. Biol. Chem.*, **269**, 31693–700.
55. Hata, S., Doi, N., Shinkai-Ouchi, F. and Ono, Y. 2020, A muscle-specific calpain, CAPN3, forms a homotrimer, *Biochim. Biophys. Acta Proteins Proteom.*, **1868**, 140411.
56. Vissing, J., Barresi, R., Witting, N., et al. 2016, A heterozygous 21-bp deletion in CAPN3 causes dominantly inherited limb girdle muscular dystrophy, *Brain*, **139**, 2154–63.
57. Song, Q., Young, K.B., Chu, G., et al. 2004, Overexpression of phospholamban in slow-twitch skeletal muscle is associated with depressed contractile function and muscle remodeling, *FASEB J.*, **18**, 974–6.
58. Komatsu, M., Nakada, T., Kawagishi, H., Kato, H. and Yamada, M. 2018, Increase in phospholamban content in mouse skeletal muscle after denervation, *J. Muscle Res. Cell Motil.*, **39**, 163–73.
59. Stiber, J., Hawkins, A., Zhang, Z.S., et al. 2008, STIM1 signalling controls store-operated calcium entry required for development and contractile function in skeletal muscle, *Nat. Cell Biol.*, **10**, 688–97.
60. Kiviluoto, S., Decuypere, J.P., De Smedt, H., Missiaen, L., Parys, J.B. and Bultynck, G. 2011, STIM1 as a key regulator for Ca2+ homeostasis in skeletal-muscle development and function, *Skelet. Muscle*, **1**, 16.
61. Sternisha, S.M. and Miller, B.G. 2019, Molecular and cellular regulation of human glucokinase, *Arch. Biochem. Biophys.*, **663**, 199–213.
62. Beinert, H. and Kennedy, M.C. 1993, Aconitase, a two-faced protein: enzyme and iron regulatory factor, *FASEB J.*, **7**, 1442–9.
63. Tong, W.H. and Rouault, T.A. 2007, Metabolic regulation of citrate and iron by aconitases: role of iron-sulfur cluster biogenesis, *Biometals*, **20**, 549–64.
64. Nezer, C., Moreau, L., Brouwers, B., et al. 1999, An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs, *Nat. Genet.*, **21**, 155–6.
65. Jeon, J.T., Carlborg, O., Tornsten, A., et al. 1999, A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus, *Nat. Genet.*, **21**, 157–8.
66. Van Laere, A.S., Nguyen, M., Braunschweig, M., et al. 2003, A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig, *Nature*, **425**, 832–6.
67. Runfola, V., Sebastian, S., Dilworth, F.J. and Gabellini, D. 2015, Rbfox proteins regulate tissue-specific alternative splicing of Mef2D required for muscle differentiation, *J. Cell Sci.*, **128**, 631–7.
68. Ohkawa, Y., Marfella, C.G. and Imbalzano, A.N. 2006, Skeletal muscle specification by myogenin and Mef2D via the SWI/SNF ATPase Brg1, *EMBO J.*, **25**, 490–501.
69. Hagiwara, N., Ma, B. and Ly, A. 2005, Slow and fast fiber isoform gene expression is systematically altered in skeletal muscle of the Sox6 mutant, p100H, *Dev. Dyn.*, **234**, 301–11.
70. Hagiwara, N., Yeh, M. and Liu, A. 2007, Sox6 is required for normal fiber type differentiation of fetal skeletal muscle in mice, *Dev. Dyn.*, **236**, 2062–76.
71. Deak, F., Mates, L., Korpos, E., et al. 2014, Extracellular deposition of matrilin-2 controls the timing of the myogenic program during muscle regeneration, *J. Cell Sci.*, **127**, 3240–56.
72. Du, Z., Zheng, H., Huang, B., et al. 2017, Allelic reprogramming of 3D chromatin architecture during early mammalian development, *Nature*, **547**, 232–5.
73. Doynova, M.D., Markworth, J.F., Cameron-Smith, D., Vickers, M.H. and O'Sullivan, J.M. 2017, Linkages between changes in the 3D organization of the genome and transcription during myotube differentiation in vitro, *Skelet. Muscle*, **7**, 5.
74. He, M., Li, Y., Tang, Q., et al. 2018, Genome-wide chromatin structure changes during adipogenesis and myogenesis, *Int. J. Biol. Sci.*, **14**, 1571–85.
75. Forcato, M., Nicoletti, C., Pal, K., Livi, C.M., Ferrari, F. and Bicciato, S. 2017, Comparison of computational methods for Hi-C data analysis, *Nat. Methods*, **14**, 679–85.
76. Levine, M., Cattoglio, C. and Tjian, R. 2014, Looping back to leap forward: transcription enters a new era, *Cell*, **157**, 13–25.
77. Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M. and Epstein, D.J. 2006, A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers, *Development*, **133**, 761–72.
78. Osterwalder, M., Barozzi, I., Tissieres, V., et al. 2018, Enhancer redundancy provides phenotypic robustness in mammalian development, *Nature*, **554**, 239–43.
79. Eusebi, V., Rilke, F., Ceccarelli, C., Fedeli, F., Schiaffino, S. and Bussolati, G. 1986, Fetal heavy chain skeletal myosin. An oncofetal antigen expressed by rhabdomyosarcoma, *Am. J. Surg. Pathol.*, **10**, 680–6.
80. Lyons, G.E., Ontell, M., Cox, R., Sassoon, D. and Buckingham, M. 1990, The expression of myosin genes in developing skeletal muscle in the mouse embryo, *J. Cell Biol.*, **111**, 1465–76.
81. Schiaffino, S., Gorza, L., Sartore, S., Saggin, L. and Carli, M. 1986, Embryonic myosin heavy chain as a differentiation marker of developing human skeletal muscle and rhabdomyosarcoma. A monoclonal antibody study, *Exp. Cell Res.*, **163**, 211–20.
82. Messina, G., Biressi, S., Monteverde, S., et al. 2010, Nfix regulates fetal-specific transcription in developing skeletal muscle, *Cell*, **140**, 554–66.
83. Taglietti, V., Maroli, G., Cermenati, S., et al. 2016, Nfix induces a switch in Sox6 transcriptional activity to regulate MyHC-I expression in fetal muscle, *Cell Rep.*, **17**, 2354–66.