Histone methyltransferase MLL4 controls myofiber identity and muscle performance through MEF2 interaction

Lin Liu, …, Kai Ge, Zhenji Gan

*J Clin Invest.* 2020. [https://doi.org/10.1172/JCI136155](https://doi.org/10.1172/JCI136155).

**Graphical abstract**

[Diagram showing the relationship between MLL4, MEF2, and myofiber types.]
Histone methyltransferase MLL4 controls myofiber identity and muscle performance through MEF2 interaction

Lin Liu, Chenyun Ding, Tingting Fu, Zhenhua Feng, Ji-Eun Lee, Liwei Xiao, Zhisheng Xu, Yujing Yin, Qiqi Guo, Zongchao Sun, Wanping Sun, Yan Mao, Likun Yang, Zheng Zhou, Danxia Zhou, Leilei Xu, Zezhang Zhu, Yong Qiu, Kai Ge, and Zhenji Gan

1The State Key Laboratory of Pharmaceutical Biotechnology and MOE Key Laboratory of Model Animals for Disease Study, Department of Spine Surgery, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Chemistry and Biomedicine Innovation Center (ChemBIC), Model Animal Research Center of Nanjing University, Nanjing 210061, China.
2Department of Spine Surgery, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Nanjing 210008, China.
3Adipocyte Biology and Gene Regulation Section, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA.

To whom to address correspondence:
Zhenji Gan, Ph.D.
Model Animal Research Center of Nanjing University
12 Xuefu Road, Pukou, Nanjing, China 210061
Tel: 86-25-58641546; Fax: 86-25-58641500
ganzj@nju.edu.cn

Conflicts of Interest: The authors have declared that no conflict of interest exists.
Abstract

Skeletal muscle depends on the precise orchestration of contractile and metabolic gene expression programs to direct fiber type specification and to ensure muscle performance. Exactly how such fiber type-specific patterns of gene expression are established and maintained remains unclear, however. Here, we demonstrate that histone mono-methyltransferase MLL4 (KMT2D), an enhancer regulator enriched in slow myofibers, plays a critical role in controlling muscle fiber identity as well as muscle performance. Skeletal muscle-specific ablation of MLL4 in mice resulted in downregulation of the slow-oxidative myofiber gene program, decreased number of type I myofibers, and diminished mitochondrial respiration, which caused reductions in muscle fat utilization and endurance capacity during exercise. Genome-wide ChIP-seq and mRNA-seq analyses revealed that MLL4 directly binds to enhancers and functions as a coactivator of the myocyte enhancer factor 2 (MEF2) to activate transcription of slow-oxidative myofiber genes. Importantly, we also found that the MLL4 regulatory circuit is associated with muscle fiber type remodeling in humans. Thus, our results uncover a pivotal role for MLL4 in specifying structural and metabolic identities of myofibers that govern muscle performance. These findings provide new therapeutic opportunities for enhancing muscle fitness to combat a variety of metabolic and muscular diseases.
**Introduction**

Muscle fitness is a key determinant of human health and disease. Muscle performance and fatigability are determined by the structural and metabolic properties of the specialized myofibers (1-3). Myofibers differ significantly in their contractile and energy metabolism functions and can be generally classified as slow-twitch (type I) and fast-twitch (type II). Type I myofibers are rich in mitochondria, relying largely on mitochondrial oxidative metabolism and resistant to fatigue (4, 5), whereas type II myofibers generally contain fewer mitochondria, have lower oxidative capacity and are fatigue sensitive, and can be sub-classified as either type IIa, IIx, or IIb in rodents based on the type of myosin heavy chain (MHC) isoform expressed (4, 5). Myofibers maintain remarkable plasticity to undergo metabolic and structural remodeling during development and in response to physiological stimuli and systemic diseases. Exercise enhances muscle performance and endurance by increasing the proportion of slow-oxidative fibers and by promoting capacity of mitochondria to burn chief fuels-fatty acids and glucose (6-10). Conversely, reduced muscle fitness, including a shift away from slow-oxidative fibers and decreased mitochondrial oxidative capacity, is a common consequence of a variety of human illness including metabolic disorders and muscular diseases (3, 7, 11, 12).

Myofibers depend on the precise orchestration of contractile and metabolic gene expression programs to direct fiber type specification and ensure muscle performance. Exactly how such fiber type-specific patterns of gene expression are established and maintained remains unclear, however. Previous studies have identified multiple
transcriptional factors such as nuclear receptor PPARs, ERRs, and MEF2, along with coregulators PGC-1s, NCoR1 and HDACs in the regulation of diverse metabolic and structural gene expression in muscle (13-21). As a critical step for gene activation and the primary site of the gene-environment interaction, enhancer activation is increasingly recognized as a key layer of regulation to adapt gene transcription to environmental cues. Active enhancers are marked by the presence of mono-methylated histone H3 lysine 4 (H3K4me1) and acetylated histone H3 lysine 27 (H3K27ac) (22, 23). Whether and how the epigenetic regulation of enhancer activation orchestrates muscle fiber type-specific patterns of gene expression is unclear.

Mixed-lineage leukemia 4 (MLL4/KMT2D), a major histone H3 lysine 4 (H3K4) mono-methyltransferase, is an essential histone writer for enhancer activation (24-28). MLL4 has been shown to co-localize with cell type-specific transcription factors to establish active enhancers during cell differentiation in multiple cell types (26-29). Loss of MLL4 prevents the establishment of H3K4me1 and H3K27ac signatures on enhancer regions, leading to significant defects in enhancer activation, and consequently, gene transcription (26-28, 30). Studies in mouse models further highlighted the importance of MLL4 in regulating a wide range of biological processes including embryonic development, metabolic homeostasis, and cancer (27, 29, 31-33). Moreover, frequent mutations in MLL4 have been implicated in several human genetic diseases, including Kabuki syndrome, congenital heart disease, and various type of cancer (33-35).
As a critical regulator of cell differentiation and cell fate transition, MLL4 has been documented to direct embryonic muscle and fat tissue development (27, 29). However, it has yet to be explored whether MLL4 has a role in adult skeletal muscle. In this study, we generated three independent skeletal muscle-specific \textit{Mll4}-knockout mouse lines, and found that slow-myofibers enriched MLL4 is required for slow-oxidative type I fiber formation to ensure muscle endurance. Mice with muscle-specific ablation of MLL4 exhibited marked downregulation of the slow-oxidative myofiber gene program, had an decreased number of type I myofibers and mitochondrial activity, resulting in reduced muscle fat utilization and endurance capacity during exercise. Mechanistically, we demonstrated that MLL4 drives the slow-oxidative type I fiber program in cooperation with MEF2. Our studies also strongly suggest that the MLL4-dependent regulation of the type I fiber program is operational in human muscle.
Results

Slow myofibers enriched MLL4 is required for type I fiber formation. As an initial step to explore the potential function of MLL4 in adult muscle, we examined the expression patterns of MLL4 in different muscle types from adult wild-type (WT) mice. Comparison of slow fiber-dominant soleus muscle to fast fiber-enriched white vastus lateralis (WV) by Western blot revealed that MLL4 protein is expressed preferentially (~2 fold) in soleus muscle (Figure 1A). In contrast, the protein levels of EZH2 and SUZ12, key components of the polycomb repressive complex 2, were markedly decreased in soleus muscle compared with those in WV (Figure 1A). These results suggest that MLL4 may activate the expression of slow-fiber genes in slow myofibers.

To directly examine the role of MLL4 in skeletal muscle, we generated two independent skeletal muscle-specific Mll4-knockout mouse lines using Mll4^{floxed} and Mll4^{SET}^{floxed} mice. In the Mll4^{floxed} strain, the exons 16-19 were flanked by two loxP sites (27), while in the Mll4^{SET}^{floxed} strain, the exons 50 and 51 encoding the enzymatic SET domain of MLL4 were floxed (29). Mll4^{floxed} and Mll4^{SET}^{floxed} mice were bred with mice expressing Cre in postnatal skeletal muscle, in which Cre expression is under the control of muscle creatine kinase (Mck) promoter (36) to achieve deletion of Mll4 in both fast- and slow-twitch muscle fiber types (referred to as Mll4 mKO and Mll4^{SET} mKO, respectively). This resulted in efficient postnatal deletion of Mll4 in skeletal muscle. Consistent with previously published data (37, 38), efficient ablation of Mll4 mediated by Mck-Cre did not occur until ~7 days after birth (Supplemental Figure
1A). As expected, the expression of Mill4 mRNA and protein levels was markedly reduced in fast and slow muscles from both muscle-specific Mill4 KO mouse lines (Figure 1B and Supplemental Figure 1, B-E). Notably, deleting the enzymatic SET domain destabilized endogenous MLL4 protein in muscle (Figure 1B and Supplemental Figure 1D), which is consistent with recent published data (29). Both Mill4 mKO and Mill4SET mKO mice were born at normal Mendelian ratios and did not exhibit an overt metabolic phenotype compared to WT littermates on standard chow. This includes similar body weight, food intake, locomotors activity, energy expenditure, and muscle weight (Supplemental Figure 2, 3). This indicates that postnatal ablation of Mill4 neither affects muscle differentiation nor development.

Examination of myofiber cross-sectional area by histochemical staining revealed that fiber size distribution was altered in gastrocnemius (GC) muscle from both muscle-specific Mill4 KO mouse lines, with a shift toward an increased cross-sectional area relative to WT controls (Figure 1C and Supplemental Figure 3M). Muscle phenotyping and transcriptional profiling revealed that fiber type programs were significantly regulated in the muscle-specific Mill4 KO mouse lines. Expression of the gene encoding the major slow-twitch type I myosin isoform MHC1 (Myh7 gene) and slow-twitch troponin genes (Tnnl, Tnnc1 and Tnnt1) was reduced in both Mill4 mKO and Mill4SET mKO GC muscle (Figure 1D). In contrast, expression of fast-twitch troponin gene (Tnnt3) was increased in muscle from the two Mill4 KO mouse lines (Figure 1D). Similar observations were made in soleus and WV muscles from the Mill4SET mKO mice (Supplemental Figure 4). Consistent with the gene expression
results, MHC1 immunofluorescence staining revealed a marked reduction in the number of type I fibers in several muscle groups from both $Mll4$ mKO and $Mll4SET$ mKO mice (Figure 1, E and F). Notably, the $Mll4SET$ mKO mouse line did not exhibit overt muscle phenotype in young mice at stage P10 (Supplemental Figure 5), suggesting that the observed results were not a consequence of changes in muscle development or early fiber-type patterning in $Mll4SET$ mKO mice.

In vitro primary skeletal myocyte culture system further demonstrated that the MLL4 protein levels were induced during differentiation of myoblasts into mature myotubes (Supplemental Figure 6A). Consistent with the established role of MLL4 in maintaining H3K4me1 levels on active enhancers (27, 28, 30), the induced expression of slow myosin paralleled the elevated levels of H3K4me1 on $Myh7$ enhancer during muscle cell differentiation (Supplemental Figure 6, A and B). MLL4 loss-of-function study was also conducted in primary skeletal myocyte. We used adenoviral vectors to express Cre or control viruses in $Mll4SET^{er}$ myoblasts. Cells were then induced to differentiate into myotubes. Adenoviral Cre-mediated KO of $Mll4$ in myoblasts resulted in diminished expression of $Myh7$ and decreased proportion of slow myosin-positive myotubes (Figure 1, G and H). In contrast, the fast myosin gene $Myh4$ expression was not decreased by $Mll4$ deletion (Figure 1, G and H). Consistent with reduction of $Myh7$ mRNA in $Mll4$ KO myotubes, MLL4 deficiency reduced the amounts of H3K4me1 on $Myh7$ enhancer (Supplemental Figure 6C). Thus, in vitro manipulation of $Mll4$ in myocyte provided further evidence that MLL4 exerts control upon slow-fiber gene expression. Taken together, these data demonstrate an essential
role of MLL4 in the regulation of type I muscle fiber type.

**Loss of muscle MLL4 causes reduced running endurance.** Muscle fiber type composition and the capacity to burn chief fuel are important determinants of muscle performance and endurance, and fast-twitch muscle fibers exhibit a glycolytic burst metabolism and are more susceptible to endurance exercise-induced fatigue (5). We next sought to determine the physiological impact of muscle MLL4 deficiency. To rule out possible effects mediated by MLL4 in the heart on exercise phenotypes, *Mll4SET*/*f* mice were bred with human skeletal actin Cre (*HSA-Cre*) mice to establish another skeletal muscle-specific deletion of the *Mll4* alleles. As expected, the protein expression of *Mll4* was dramatically decreased in multiple muscle types, but not in the heart, from *Mll4SET*/*f*/*HSA-Cre* mice (Supplemental Figure 7A). In addition, MHC1-positive fibers were markedly reduced in *Mll4SET*/*f*/*HSA-Cre* GC muscle (Figure 2, A and B and Supplemental Figure 7B). We next assessed acute running endurance performance in the *Mll4SET*/*f*/*HSA-Cre* mouse line using a run-to-exhaustion protocol on a motorized treadmill. Consistent with the observed alterations in muscle fiber type proportion, *Mll4SET*/*f*/*HSA-Cre* mice could run for significantly shorter time and distance (~30%) compared to WT littermates (Figure 2C).

To further evaluate muscle fuel utilization during exercise, WT or *Mll4SET*/*f*/*HSA-Cre* mice were first subjected to a forced maximal exercise capacity test (VO$_{2\text{max}}$ test) consisting of an increasing speed every 2 min at 10° inclination until
exhaustion (Supplemental Figure 8A). Consistent with a shift to muscle glucose utilization, the respiratory exchange ratio (RER) increased with exercise in both $\textit{Mll4SET}^{\text{fl}}\text{HSA-Cre}$ and the WT control group, indicative of a switch to carbohydrates as the chief fuel (Supplemental Figure 8, A-D). Interestingly, in this high intensity exercise challenge, both genotypes demonstrated similar energy substrate utilization (RER) during the course of exercise, and $\textit{Mll4SET}^{\text{fl}}\text{HSA-Cre}$ mice run a similar distance and maximal speed as their control littermates (Supplemental Figure 8, C-F). Mice that performed the VO$_{2}\text{max}$ test were then subjected to a low intensity (endurance-type) protocol, during which animals run at a constant speed of 60% of their maximal running speed (Supplemental Figure 8G). While RER values increase during the beginning of the exercise test, both WT and the $\textit{Mll4SET}^{\text{fl}}\text{HSA-Cre}$ mice displayed a drop in RER after ~14 min of endurance exercise (Supplemental Figure 8, G-I). This decrease indicates a shift in substrate usage from glucose to fat metabolism. Despite no difference in RER during the exercise period, the $\textit{Mll4SET}^{\text{fl}}\text{HSA-Cre}$ mice run significantly shorter compared with the WT control group (Supplemental Figure 8, H-K). These results suggest that muscle MLL4 deficiency resulted in reduced capacity to persistent use of fat as a fuel during endurance exercise. Furthermore, we also measured glucose and fatty acid levels in blood from WT and $\textit{Mll4SET}^{\text{fl}}\text{HSA-Cre}$ mice before and after endurance exercise. $\textit{Mll4SET}^{\text{fl}}\text{HSA-Cre}$ mice showed significantly lower blood glucose levels compared to WT controls at baseline and after 80 min endurance exercise (Figure 2D). Conversely, blood TG and fatty acid levels were significantly higher after exercise in $\textit{Mll4SET}^{\text{fl}}\text{HSA-Cre}$ mice (Figure 2E).
Notably, blood ketone body β-hydroxybutyrate levels mirrored the changes of blood fatty acids, as they were also increased in Mll4SET^dif HSA-Cre mice at basal condition or post-exercise (Figure 2E). Together, these results demonstrate a change in slow type I fibers and muscle metabolism in the absence of muscle MLL4 that compromises running endurance.

**MLL4 coordinates gene programs controlling muscle contraction and energy metabolism.** Type I myofibers depend on the precise orchestration of slow- contractile and oxidative metabolic gene expression to ensure muscle endurance. To more thoroughly analyze the fiber type-specific patterns of gene expression changes that result from loss of muscle MLL4, we performed RNA-seq analysis on mRNA isolated from GC muscle of the Mll4SET mKO mice and littermate controls. We found that MLL4 regulated a total of 1000 genes in skeletal muscles, with 447 up- and 553 down-regulated, respectively (Figure 3A). In addition to the key slow-twitch myosin heavy chain gene Myh7 and troponin complex described above (Figure 1D), gene ontology (GO) analysis of MLL4-regulated genes also revealed significant enrichment in fiber-type specific isoforms of sarcomeric components as well as ion channels involved in excitation-contraction coupling (Figure 3B). The comparative RNA profiling strategy revealed extensive fiber-type switching and novel myofiber-specific gene expression. As shown in Figure 3C, a broad array of contraction-related genes that differ between muscle fiber type were altered in Mll4SET mKO muscle (Figure 3C). Real-time PCR confirmed that the expression of many genes encoding slow
contractile proteins (Myl2, Myl3, Tpm3, and Myom3) was reduced in the GC muscle of Mll4SET mKO mice compared to WT controls (Figure 3D). The expression of slow-fiber calcium handling genes (Atp2a2, Casq2, Smtnl1, and Myoz2) was also reduced in the Mll4SET mKO muscles compared to WT controls (Figure 3D). Moreover, we also found a decreased expression of genes associated with fatty acid and glucose metabolism (Cpt1b, Slc27a1, Fabp3, Dgat2, Fads6, Phyhd1, Ldhb and Ppplrla) in Mll4SET mKO muscle (Figure 3D). The expected LDH isoenzyme activity shifts due to changes in Ldhb expression were confirmed by gel-activity studies (Figure 3E). To further assess the metabolic effects of muscle MLL4, mitochondrial respiration rates were determined in the extensor digital longus (EDL) muscle of Mll4SET mKO mice and WT controls using pyruvate or palmitoylcarnitine as substrates. State 3 (maximal ADP-stimulated) respiration rates were significantly lower in Mll4SET mKO muscle compared to the WT controls (Figure 3F). Moreover, measurement of oxygen consumption rate (OCR) using an extracellular flux analyzer also revealed that Mll4SET KO myotubes had a reduced OCR in the presence of uncoupler FCCP, a sign of reduced maximal mitochondrial respiratory capacity (Figure 3G), whereas no difference in the extracellular acidification rate (ECAR) was observed in Mll4SET KO myotubes compared with Mll4SET^6f^ controls (data not shown). Together, these results demonstrate that MLL4 programs type I muscle metabolism by coordinate regulation of gene programs controlling muscle contraction and energy metabolism.
**MLL4 drives the slow-oxidative muscle fiber program in cooperation with MEF2.**

To define the mechanism involved in the activation of slow-oxidative muscle fiber gene programs by MLL4, we examined the possibility that MLL4 regulates factors known to regulate slow myofiber gene expression. *Esrrg* mRNA levels were modestly increased and *Ppard* mRNA levels were lower in *Mll4SET* mKO muscles (Supplemental Figure 9), whereas *Fnip1, Ppargc1a, Ppargc1b, Ppara, Esrra,* and *Esrrb* gene expressions were not changed in *Mll4SET* mKO muscles (Supplemental Figure 9). These results led us to hypothesis that MLL4 may acts as direct activator of slow-myofiber gene expression.

We next sought to examine the genome-wide MLL4 occupancy on WT mouse muscle chromatin by chromatin immunoprecipitation coupled with high-throughput DNA sequencing (ChIP-seq). As a control for specificity, ChIP-seq experiments were also performed in *Mll4* mKO muscle. 9403 high-confidence MLL4 genomic binding regions were obtained by filtering out non-specific signals observed in MLL4-deficient muscle (Supplemental Figure 10, A-C). MLL4 binding sites in skeletal muscle were predominantly located in the intergenic, intronic and promoter regions (Supplemental Figure 10D), which is consistent with our previous MLL4 ChIP-seq results in myocytes (27). Correlating the MLL4 cistrome in muscle and myocytes with the global mRNA changes upon muscle *Mll4* deletion revealed that ~49% of MLL4-regulated genes were directly bound by MLL4 (based on the nearest genes to the MLL4 binding peaks) (Figure 4A). Furthermore, GO analysis of the 492 directly regulated genes showed enrichment of the similar muscle contraction-related terms as seen in the *Mll4SET* mKO transcriptome (Figure 4B). In addition, we also found that MLL4 could occupy many fatty acid utilization genes (e.g., *Cpt1b* and *Fabp3*) (Supplemental Figure 10E). Together, the strong correlation between the...
cistromic and transcriptomic findings supports a direct role for MLL4 in regulating muscle fiber type gene programs.

Interestingly, De novo motif analysis of the top 5,000 emergent MLL4 binding regions in muscle using Homer revealed a sequence element, 5′-CTAAAAATAG-3′, as the highest-score motif with a P-value of 1e-137 (Figure 4C). This motif closely corresponded to a consensus site that has been previously known as the MEF2-binding site. Indeed, the search for known motifs confirmed that the most enriched motif within MLL4 binding regions was the binding site for MEF2 family members, key regulators of the type I fiber-type program (18, 39) (Figure 4D).

Moreover, we found that MLL4 could occupy both Mef2a and Mef2d genes (Supplemental Figure 10F), indicating that Mef2a and Mef2d are also direct targets of MLL4. These results suggest that MEF2s may be co-regulatory transcription factors that are involved in the MLL4-mediated regulation of the type I fiber-type program.

We next interrogated a published myocyte MEF2 ChIP-seq dataset (40) in conjunction with our previously published myocyte ChIP-seq data for MLL4, and histone modifications (H3K4me1 and H3K27ac) associated with enhancer activities (27). As shown in Figure 4E, heatmap visualization of the MEF2 binding and H3K4me1 and H3K27ac signal at the MLL4 binding events demonstrated a substantial degree of the MLL4 and MEF2 co-occupancy on active enhancers (Figure 4E). Gene set enrichment analysis (GSEA) further demonstrated that Mll4 deletion significantly downregulated expression of MLL4 and MEF2 co-bound target genes in muscle (Figure 4F). These data suggest that MLL4 acts coordinately with MEF2 in
regulating muscle gene expression. Genomic co-localization of MLL4 and MEF2 was confirmed on many slow-oxidative muscle fiber gene loci in myocytes. For example, slow-fiber-related gene enhancers (e.g., Myh7, Tnnc1, Tnni1, Tnnt1, Ldhb, Casq2 and Mybph) showed strong MLL4 and MEF2 binding as well as presence of H3K4me1 and H3K27ac signals (Figure 4, G and H and Supplemental Figure 11). Importantly, deletion of Mll4 markedly decreased levels of H3K4me1 and H3K27ac, which are consistent with pronounced reduction of slow-oxidative muscle fiber gene mRNA in Mll4SET mKO muscle (Figure 4, G and H and Supplemental Figure 11). Together, these results suggest that MLL4 occupies the enhancers in cooperation with MEF2 to activate transcription of the slow-oxidative muscle genes.

A series of co-immunoprecipitation (co-IP) studies were next conducted to determine whether MLL4 interacts with MEF2. HEK293 cells were cotransfected with expression vectors for HA-MEF2A and Flag-MLL4. Anti-HA was found to coimmunoprecipitate MLL4 and MEF2A (Figure 5A). Using MLL4 as the immunoprecipitation target, MEF2A was pulled down (Figure 5B). Furthermore, immunoprecipitation of MLL4 in C2C12 myotubes with an anti-MLL4 antibody confirmed that endogenous MLL4 and MEF2 actually interact in muscle cells (Figure 5C). To explore functional correlates of the MLL4/MEF2 interaction, we focused on the Myh7 and Myh7b loci. These two MLL4-regulated genes encode the major slow myosin and drive expression of miR-208b and miR-499 that further promote muscle fiber remodeling (41, 42). Upon examination of the MLL4 and MEF2 ChIP-seq data, we uncovered a MLL4/MEF2 co-bound region located ~3 kb upstream of the Myh7
TSS. We identified two highly conserved DNA sequences conforming to the consensus binding sites for MEF2 \([(C/T)TA(T/A)_{4}TA(G/A)]\) in the \(Myh7/miR-208b\) cis-proximal enhancer region (Figure 5D). Cell cotransfection studies were next conducted using a rat 3.5 kb \(Myh7\) promoter reporter containing the enhancer element (Figure 5E). The \(Myh7\) luciferase reporter was not activated by MLL4 alone, but when expressed together with MEF2A, synergistic activation was observed (Figure 5E), suggesting a cooperative transcriptional activation by the two factors. To further assess the activation of the \(Myh7/miR-208b\) enhancer by MLL4/MEF2, deletion mapping studies were used to demonstrate that this MLL4/MEF2 synergistic effect was markedly diminished by deleting the upstream enhancer sequences (Figure 5F).

We also detected MLL4 and MEF2 co-localization on \(Myh7b\) promoter region (Figure 5G). Cell cotransfection studies were further conducted using a mouse \(Myh7b\) promoter-reporter containing a highly conserved MEF2 binding site (Figure 5H). As expected, the combination of MLL4 and MEF2 resulted in synergistic activation of the \(Myh7b\) reporter (Figure 5H). This synergistic effect was completely abolished upon mutation of the MEF2 site (Figure 5H), providing further evidence that the MEF2 site cooperates with MLL4 in the activation of \(Myh7b/miR-499\) gene transcription. Taken together, these results demonstrate that MLL4 cooperates with MEF2 to activate slow muscle gene transcription, thereby driving the fiber type remodeling.

*MLL4 levels are regulated in response to physiological and pathophysiological*
**stimuli in human muscle.** To determine the relevance of the MLL4 regulatory circuit in humans, we investigated whether MLL4 function could be altered by physiological or pathophysiological stimuli in human muscle. We first analyzed the gene expression profiles in human muscle (GEO: GSE1718) from a group of sedentary subjects who underwent a 20 weeks of endurance exercise training program (43). This analysis revealed that the MLL4 regulatory circuit was induced in human muscle by exercise training. Increased mRNA levels of *MLL4* associated with elevated expression of its targets (*MYH7B, TNNT1, MYL2* and *ATP2A2*) were observed in the muscle tissue from the “Post-Training” group compared with the “Pre-Training” group (Figure 6A). In contrast, the levels of *TNNI2* and *TNNC2* mRNA showed a decrease in “Post-Training” muscle (Figure 6A).

Skeletal muscle dysfunction, including fiber type switching and reduced oxidative capacity, has been associated with a variety of human diseases. Patients with adolescent idiopathic scoliosis (AIS), the most common type of scoliosis that affects 1-4% adolescents, have generalized muscle dysfunction (44, 45). Although causal mechanisms of AIS remain unclear, paraspinal muscle imbalance due to muscle fiber type switching has long been recognized a possible factor underlying the pathology of AIS (46-48). Therefore, we examined the regulation of the MLL4 signaling in paraspinal muscle samples from 40 patients with AIS. Consistent with previous reports (47, 48), metachromatic ATPase staining revealed that muscle fiber type programs were altered in paraspinal muscle from AIS patients. The biopsy specimens from the concave side (“MV”) of the curvature contained a lower percentage of type I
muscle fibers compared to the convex side (“MX”) (Figure 6B). As expected, muscle biopsies from the “MV” group exhibited lower slow fiber gene (MYH7, TNNI1, TNNC1, TNNT1 and LDHB) expression but higher fast fiber gene (TNNI2 and TNNC2) expression compared to the “MX” group (Figure 6C and Supplemental Figure 12A). Expression levels of MLL4 and MEF2s were also lower in the “MV” group compared to the “MX” group (Figure 6C). In addition, the levels of MLL4 mRNA exhibited a significant positive correlation with slow fiber gene (MYH7, TNNI1, TNNT1 and LDHB) and MEF2s mRNA levels, but not with a marker of the fast-fiber program (TNNI2) (Figure 6D and Supplemental Figure 12B). The relationship between MLL4 expression and levels of the MEF2s with type I fiber percentage was also assessed. As shown in Figure 6E, both MLL4 and MEF2s exhibited a significant positive correlation with type I fiber percentage (Figure 6E). Together, these results strongly suggest that the MLL4 regulatory circuit is operative in the regulation of the type I muscle program in response to physiological and pathophysiological stimuli in humans.
Discussion

The mechanisms underlying the precise orchestration of contractile and metabolic gene expression that specify muscle fiber identity and function remain unclear. Herein, using loss-of-function strategies in mice and primary muscle cells, together with assessment of muscle biopsies from humans, we have uncovered an essential role for the enhancer regulator MLL4 in controlling the structural and metabolic programs that govern myofiber identity and muscle performance (Figure 7). MLL4 is highly expressed in type I muscle fibers, and deletion of Mll4 specifically in skeletal muscle resulted in decreased number of type I myofibers and diminished mitochondrial respiration, leading to significant defects in muscle endurance during exercise. These changes were resulted from marked downregulation of the slow-oxidative muscle gene programs in skeletal muscle lacking MLL4. We found that MLL4 directly binds to enhancers and functions as a coactivator of MEF2 to activate the transcription of slow-oxidative myofiber genes, thereby driving the muscle fiber type remodeling. Moreover, our data also established that the MLL4 regulatory circuit is associated with muscle fiber type switching in humans. Therefore, MLL4 likely represents a previously unrecognized molecular switch that specifies myofiber structural and metabolic identities that govern muscle performance.

Enhancer activation is a critical step for gene activation. While transcriptional regulation of myofiber gene expression through multiple transcription factors such as nuclear receptor PPARs, ERRs, MEF2, SOX6, and Tbx15, along with coregulators have been established (13-21, 49, 50), it remains unclear how fiber type-specific
patterns of gene expression are controlled at enhancer activation layer. In this study, we showed that enhancer regulator MLL4 drives the contractile and metabolic specification of type I muscle fibers in both primary skeletal myotubes and skeletal muscle. We also confirmed a marked reduction in active enhancer hallmarks on slow muscle gene enhancers in Mill4 KO myocytes. Notably, deletion of the enzymatic SET domain destabilized MLL4 protein, thus limiting the study of the role of its enzymatic activity in muscle. It is also worth noting that we used Mck-Cre and HSA-Cre mice for creating mice with Mill4 deletion specifically in skeletal muscle in this study, and we did not observe a change in muscle differentiation or development upon muscle Mill4 deletion. In Mill4SET mKO model using Mck-Cre, which is not fully active until after birth (37, 38), we did not observe significant changes in muscle fiber type in Mill4SET mKO muscle compared with WT controls at stage P10. This could reflect muscle fiber type remodeling during postnatal development. It is possible that other postnatal programs such as neuronal or hormone signals are more dominant during the first weeks after birth (5). Remarkably, we found that MLL4 is required for the maintenance of type I fibers in adult muscle. As such, ablation of muscle MLL4 profoundly affects muscle fatigability and exercise performance. Thus our study has expanded the role of enhancer regulator MLL4 to include pivotal muscle fiber specification and exercise physiological functions other than development and cell differentiation. Conceivably, this muscle specific regulatory action may reflect the multi-functions of MLL4 in utilizing its enhancer remodeling activity to direct context-specific biological processes.
Type I muscle fibers are equipped with slow-twitch contractile machinery and a high capacity fuel burning system poised for endurance exercise. Our data suggest that MLL4 orchestrates the structural and metabolic programs controlling type I muscle metabolism. Genome-wide transcriptional analysis revealed that MLL4 acts by coordinately regulating genes controlling slow-twitch myofibers, calcium handling, and oxidative metabolism. At a functional level, these changes resulted in significant impaired muscle endurance during exercise in muscle specific Mll4 KO mice. Interestingly, we found that muscle MLL4 deficiency did not compromise muscle performance during high intensity (sprint) exercise, perhaps reflecting that fast type II fibers are involved in rapid bursts of contraction. Our data suggest that MLL4 is a critical upstream epigenetic switch that specifies slow-oxidative myofiber identity by co-activating MEF2 transcriptional regulators. Since the expression of slow oxidative muscle genes is also regulated by nuclear receptors such as PPARβ/δ and ERRs (14, 15, 17, 19), it remains to be tested whether MLL4 may also contribute to the activation of PPARβ/δ and ERRs in muscle. Additionally, it will be of interest to determine whether MLL4 plays an opposing role as a direct transcriptional repressor of fast myofiber-specific gene expression. Coordinate control of muscle fiber type specification and fuel burning capacity occurs during exercise training. Interestingly, we found that exercise induces an increase in mRNA abundance of MLL4 and its targets in human muscle, thereby suggesting the MLL4 regulatory circuit unveiled here could be involved in the response to exercise training.
Our results from the global MLL4-directed transcriptional analysis indicate that
the MLL4-driven slow-oxidative muscle gene program acts in cooperation with
MEF2. The observed role of MEF2 is of interest given its known role in the regulation
of muscle development and muscle fiber type remodeling in adulthood (18, 39). We
found that MLL4 interacts and cooperates with MEF2 to activate slow-oxidative
muscle gene transcription. This is consistent with the observation that elevated
expression of MLL4 parallels high MEF2 activity in slow-fiber dominant soleus
muscle. Notably, MLL4 also directly binds to Mef2a and Mef2d gene loci, which
likely adds another regulation layer downstream of MLL4 to enhance the mechanisms
describe here and boosts the highest level of MLL4/MEF2 cooperation. MEF2D has
been shown to recruit the Ash2L methyltransferase complex to MyoD target genes
during myogenesis (51). The interaction between MLL4 and MEF2 and the genomic
co-localization of MLL4 with MEF2 suggest that MEF2 plays critical roles in
recruiting MLL4 to establish active enhancers on slow-oxidative muscle genes. It
remains to be dissected, however, the precise mechanism whereby MLL4
transcriptional activation complex tethered to MEF2 binding sites to establish active
enhancers. The class II HDACs and NCoR1 are known to corepress MEF2 (18, 21),
therefore, it is likely that MEF2-mediated regulation of slow-fiber phenotype is
controlled by the balance between MLL4 and class II HDACs/NCoR1 signaling.
MLL4 and MEF2 cooperation allows enhancer activation and the establishment of
slow fiber-specific gene expression pattern, while the class II HDACs and NCoR1 put
brakes on the activation of MEF2. Interestingly, previous studies also suggest that
muscle-specific switch in exon usage for MEF2 leads to the recruitment of alternate coregulatory complexes (40, 52). It is also possible that MEF2 exon switching could alter its association with MLL4.

Myofiber shift from slow-oxidative toward fast-glycolytic has been associated with a variety of chronic illness including metabolic disorders and muscular diseases (3, 7, 12). In this study, our survey of the human muscle samples demonstrated that the MLL4 regulatory circuit is associated with muscle fiber type switching in AIS patients. Specifically, MLL4 and MEF2s levels were reduced in muscle samples from the “MV” group compared to the “MX” group. We also found that the expression of MLL4 and MEF2s was strongly correlated with expression of slow fiber genes and type I fiber-type proportion. Based on our data, it is possible that downregulated MLL4 signaling leads to reduced type I muscle fibers, contributing to the paraspinal muscle imbalance in AIS patients. Notably, MLL4 has also been identified as a major causative mutation gene in Kabuki syndrome, a human genetic disease that cause multiple malformations including muscle hypotonia (34, 53). It is tempting to speculate that loss-of-function mutations in MLL4 contribute to the muscle dysfunction in Kabuki syndrome.

In summary, we demonstrate an essential role for enhancer regulator MLL4 in specifying myofiber structural and metabolic identities that govern muscle performance. Given that many disease states are associated with reduced muscle endurance, these findings provide new therapeutic opportunities for enhancing muscle fitness to combat a variety of metabolic and muscular diseases.
Methods

Animal studies

Male C57BL/6J wild-type mice were from the Model Animal Research Center (MARC) of Nanjing University. Generation of *Mll4*°° and *Mll4SET*°° mice have been described elsewhere (27, 29). Mice were back-crossed to the C57BL/6J background for more than 6 generations. To generate mice with a muscle-specific disruption of the *Mll4* allele, *Mll4*°° and *Mll4SET*°° mice were crossed with mice expressing Cre recombinase under control of a muscle creatine kinase (*Mck*) promoter (Jackson Laboratory, stock no. 006475) or the human skeletal actin (*HSA*) promoter (Jackson Laboratory, stock no. 006139) to achieve muscle-specific deletion of *Mll4*. Male offspring were genotyped and mice at the age of 1 day to 20 weeks were used. Mice were randomly assigned to various analyses. Littermate controls were used in all cases. Investigators involved in the immunofluorescence imaging, RNA-seq/ChIP-seq, and histological analysis were blinded. Investigators performing animal handling, sampling, and raw data collection were not blinded.

Human studies

Details on subject characteristics are provided in Supplemental Table 1. Adolescent idiopathic scoliosis (AIS) patients who underwent posterior instrumentation and spinal fusion surgery were recruited from Drum Tower Hospital, Nanjing, China. Patients with scoliosis secondary to known etiology were excluded from the present study. Deep paraspinal muscle biopsies were taken at both the concave and convex sides of the curve’s apex during spinal fusion surgery, cleaned, and mounted for fiber typing or flash frozen in liquid nitrogen for RNA isolation.
Exercise stress test

Mice were acclimated (run for 9 minutes at 10 meters (m)/minute followed by 1 minute at 20 m/minute) to the treadmill for 2 consecutive days prior to the experimental protocol. Low intensity (endurance) exercise studies were conducted as described previously (41, 54). In brief, fed mice were run for 10 minutes at 10 m/minute followed by a constant speed of 20 m/minute until exhaustion.

Respiratory exchange ratios (RER) during exercise were determined as described previously (41, 54, 55). Briefly, mice were placed in an enclosed treadmill attached to the Comprehensive Laboratory Animal Monitoring System (CLAMS) (Columbus Instruments) for 10 min at a 10° incline and 0 m/min. To determine maximal exercise capacity, the mice were subjected to a high intensity exercise (wind sprints) test consisting of an increasing speed every 2 min at 10° inclination until exhaustion. The increasing speeds used in the protocol were 10, 14, 18, 22, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46 and 48 m/min. One week later, the same group of mice performed a low intensity exercise (endurance) challenge. After a brief warm-up, the mice were challenged with a constant speed of 60% of their maximal running speed at 10° inclination until exhaustion. Measurements were collected before the exercise challenge and throughout the challenge.

Metabolic measurements in vivo

Mice were housed individually in metabolic cages at a 12-h light and dark cycle with
free access to food and water using the CLAMS (Columbus Instruments). Mice were acclimated in the metabolic cage for 1 day prior to the recording according to the instructions of the manufacturer. Food, energy expenditure, physical activity, VO$_2$ and VCO$_2$ were assessed simultaneously.

**Body composition analyses**

Mice body composition parameters including fat mass and lean tissue mass were determined via dual-energy X-ray absorptiometry (DEXA) using a Lunar PIXImus II densitometer (GE Healthcare) according to the instructions of the manufacturer.

**Blood and tissue chemistry**

Blood glucose levels were determined using a OneTouch ultramini glucose meter (OneTouch). Serum TGs levels were determined using a Triglyceride kit (Wako, 290-63701). Serum fatty acid levels were determined using a NEFAs kit (Wako, 294-63601). Serum β-hydroxybutyrate (Cayman, 700190) levels were measured according to manufacturer’s instruction.

**Histological analyses**

Mice muscle tissues were frozen in isopentane that had been cooled in liquid nitrogen. Immunofluorescence (IF) stains were conducted as previously described (41). For IF stains, the muscle fibers were stained with antibodies directed against MHC1 (BA-D5, #AB 2235587) or MHC2b (BF-F3, #AB 2266724). Tissue sections were stained with hematoxylin and eosin (Sigma-Aldrich) according to the standard protocol. Wheat
germ agglutinin (WGA) staining was performed using FITC-conjugated WGA (Sigma-Aldrich, #L4859). Quantification of cross sectional area of the myofibers was performed with NIH ImageJ software.

**RNA-Seq studies**

Transcriptomics analyses were performed using RNA-sequencing as described previously (56). Total RNA was isolated from the gastrocnemius muscle of 8-week-old male *Mll4SET* mKO and WT control mice using RNAiso Plus (Takara Bio). RNA-seq using Illumina HiSeq 4000 was performed by Beijing Novogene Bioinformatics Technology Co., Ltd. Three independent samples per group were analyzed. Paired-end, 150 nt reads were obtained from the same sequencing lane. Transcriptome sequencing libraries averaged 33 million paired reads per sample, with 81.8% alignment to the mouse genome (UCSC mm10). The sequencing reads were then aligned to the UCSC mm10 genome assembly using TopHat 2.0.14 with the default parameters. Fragments Per Kb of exon per Million mapped reads (FPKM) were calculated using Cufflinks 2.2.1. The criteria for a regulated gene were a fold change greater than 1.5 (either direction) and a significant P-value (< 0.05) versus WT. For pathway analysis, the filtered data sets were uploaded into DAVID Bioinformatics Resources 6.8 to review the bio pathways using the Functional Categories database. The GO analysis was used to interpret data, and the regulated terms were ranked by *P*-value in Figure 3B and Figure 4B. The volcano plot and heat-map analysis of regulated genes were generated by using R software (Version 3.3.2) and
ggplot2/gplots package. The RNA-seq data have been deposited in the NCBI Gene Expression Omnibus and are accessible through GEO Series accession number GSE137368. The following secure token has been created to allow review of record GSE137368 prior to publication: qtabkwmovxqzhap.

**RNA analyses**

Quantitative RT-PCR was performed as described previously (56). Briefly, total RNA was extracted from mouse muscle or primary myotubes using RNAiso Plus (Takara Bio). Isolated total RNA integrity was electrophoretically verified by ethidium bromide staining. 1 µg total RNA samples were then reverse transcribed with the PrimeScript RT Reagent Kit with gDNA Eraser (Takara Bio) using random hexamer primers according to the manufacturer’s instructions. Real-time quantitative PCR was performed using the ABI Prism Step-One system with SYBR® Premix Ex Taq™ (Takara Bio). Specific oligonucleotide primers for target gene sequences are listed in Supplemental Table 2. Arbitrary units of target mRNA were corrected to the expression of 36b4 or GAPDH.

**ChIP and ChIP-seq**

ChIP assays from WT or Mll4 mKO mice or primary skeletal myocyte were conducted as described previously with modifications (56). Briefly, muscle cell chromatin fragmentation was performed by sonication using a Bioruptor (Diagenode). An aliquot of chromatin was pre-cleared with protein G and immunoprecipitated with anti-MLL4, anti-H3K4me1 (Abcam, ab8895) or IgG control (Beyotime, A7016)
antibodies. Following reversal of cross-linking, DNA was isolated using the standard Phenol-chloroform method. Quantitative PCR products were assessed and measured using the ABI Prism Step-One system. Quantitative analysis was performed by the standard curve method. Specific oligonucleotide primers for target regions were listed in Supplemental Table 2. For ChIP-seq, the MLL4 precipitated DNA samples from three independent ChIP experiments were pooled (n = 3 mice each WT or Mll4 mKO group) and then amplified according to ChIP Sequencing Sample Preparation Guide provided by Illumina, DNA library was generated using the NEBNext® Ultra™ II DNA Library Prep Kit (NEB #E7645). Deep sequencing using Illumina HiSeq 2500 was performed by Beijing Novogene Bioinformatics Technology Co., Ltd.

**ChIP-seq data processing**

ChIP-seq data analysis was performed as described previously with modifications (56). Briefly, single end 50 nt reads were mapped to the mouse genome (UCSC mm10) using Bowtie2 (Version 2.2.5). Only the sequences uniquely mapped with no more than 1 mismatch were kept and used as valid reads. The peak caller program MACS2 (Version 2.1.1) was used to identify peaks with the following parameter settings: --keep-dup = 1, -B, --SPMR to generate signal pileup tracks in bedGraph format on a per million reads basis. This allows for direct comparison between the WT and KO ChIP-seq. Genome regions were associated to the gene with the nearest transcription start sites (TSS) from the UCSC genome browser. Region centers that were 1 kb upstream to 200 bp downstream of a RefGene TSS were classified as “Promoter” regions. Remaining region centers were classified based on overlapping RefGene introns, exons coding regions, 5’UTRs and 3’UTRs. All other regions were classified as “Intergenic”. A total of 8334 MLL4 bound genes in skeletal muscle or myocyte and
the global mRNA changes (1000 genes) upon muscle Mill4 deletion were used for the combinatorial analysis to define a set of genes directly regulated by MLL4. GO term enrichment analysis was performed with MLL4 direct targets (492 genes) as input to the Database for Annotation, Visualization and Integrated Discovery (DAVID v6.8; https://david.ncifcrf.gov/). De novo motif discovery of MLL4 ChIP–seq peaks was performed using the findMotifsGenome command in Homer (- size = given) with default options. The known motif enrichment analysis was performed using TRANSFAC database and the findMotifsGenome command in Homer with the following parameter settings: - size = given, -nomotif. P values were calculated on the basis of a binomial test against GC%-matched background. We also interrogated a public myocyte MEF2 ChIP-seq dataset (GSE43223) (40) in conjunction with our previously published myocytes ChIP-seq data for MLL4, and histone modifications (H3K4me1 and H3K27ac) associated with enhancer activities (GSE50466) (27). The co-localization was defined as two regions overlapping for at least 1bp, BEDTools was used for genomic region handling. For heat-maps, tag density matrices were calculated using Homer followed by visualization by Treeview (Version 1.1.6r4). Around 5 kb centered on each peak with 10 bp bin size was used for MLL4, MEF2D, H3K4me1 and H3K27ac ChIP-seq datasets. The heatmap color scale indicates the binding signals per million total reads. Gene set enrichment analysis (GSEA) was carried out by comparing a dataset of MLL4/MEF2D co-bound targets (627 genes) with that from genome-wide mRNA expression profiling upon muscle Mill4 deletion. Enrichment score was obtained with a false positive rate (FDR) q-value by running sum of the ChIP-seq genes and differentially expressed rank-ordered genes in mRNA profiling. Genome browser tracks of ChIP-seq data were visualized in IGV (Version 2.3.70). The ChIP-seq data have been deposited in the NCBI Gene Expression
Omnibus and are accessible through GEO Series accession number GSE138994. The following secure token has been created to allow review of record GSE138994 prior to publication: ghcdwkwkfktvgy.

Antibodies and immunoblotting studies

Antibodies directed against MHC1 (BA-D5, #AB 2235587) and MHC2b (BF-F3, #AB 2266724) were purchased from the Developmental Studies Hybridoma Bank; anti-SUZ12 (#sc-46264, 1:1,000 dilution) was from Santa Cruz, anti-EZH2 (#5246S, 1:1,000 dilution) was from Cell Signaling Technology, anti-α-Tubulin antibody (bs1699, 1:5,000 dilution) was from Bioworld, anti-MEF2C (ab211493, 1:500 dilution) was from Abcam, anti-myosin slow (#M8421, 1:250 dilution), anti-myosin fast (#M4276, 1:250 dilution), anti-Flag (#F1804, 1:1,000 dilution), and anti-HA (#H9658, 1:1,000 dilution) were from Sigma; anti-MLL4 (1:2,000 dilution) antibody was provided by Kai Ge at National Institutes of Health (Bethesda, Maryland, USA) as previously described (27). Western blotting studies were performed as previously described (56).

LDH isoenzyme analysis

LDH isoenzyme patterns were determined as previously described (55). Protein extracted from mouse hearts served as the positive control.

Mitochondrial respiration studies
Mitochondrial respiration rates were measured in saponin-permeabilized extensor digital longus muscle fibers with pyruvate or palmitoylcarnitine as a substrate as described previously (41, 55). In brief, the muscle fibers were separated and transferred to BIOPS buffer (7.23 mM K$_2$EGTA, 2.77 mM CaK$_2$EGTA, 20 mM imidazole, 20 mM taurine, 50 mM potassium 2-[N-morpholino]-ethanesulfonic acid, 0.5 mM dithiothreitol, 6.56 mM MgCl$_2$, 5.7 mM ATP, and 14.3 mM phosphocreatine [PCr], pH 7.1). The muscle fibers bundles were then permeabilized with 50 µg/ml saponin in BIOPS solution. Measurement of oxygen consumption in permeabilized muscle fibers was performed in buffer Z (105 mM potassium 2-[N-morpholino]-ethanesulfonic acid, 30 mM KCl, 10 mM KH$_2$PO$_4$, 5 mM MgCl$_2$, 5 mg/ml BSA, 1 mM EGTA, pH 7.4) at 37°C and in the oxygen concentration range 220-150 nmol O$_2$/ml in the respiration chambers of an Oxygraph 2K (Oroboros Inc., Innsbruck, Austria). Following measurement of basal, pyruvate (10 mM)/malate (5 mM) or palmitoylcarnitine (40 µM)/malate (5 mM) respiration, maximal (ADP-stimulated) respiration was determined by exposing the mitochondria to 4 mM ADP. Uncoupled respiration was evaluated following addition of oligomycin (1 µg/mL). Respiration rates were determined and normalized to tissue wet weight using Datlab 5 software (Oroboros Inc., Innsbruck, Austria), the data were expressed as “pmol O$_2$ s$^{-1}$ mg wet weight$^{-1}$”.

**Oxygen consumption measurements**

Cellular oxygen consumption rates (OCR) were measured using the XF24 analyzer
(Seahorse Bioscience Inc.) per the manufacturer’s protocol as described previously (41). The basal OCR was first measured in XF Assay Media without sodium pyruvate, followed by administration of sodium pyruvate to a final concentration of 10 mM. Uncoupled respiration was evaluated following the addition of oligomycin (2 μM) to inhibit ATP synthase, by addition of the uncoupler FCCP (2 μM), and then followed by the addition of rotenone/antimycin (1 μM). Immediately after measurement, total protein levels were measured with the Micro BCA Protein Assay Kit (Thermo scientific) for data correction.

**Cell transfection and luciferase reporter assays**

pCMV6-flag-myc-MLA4, rat 3.5-kb Myh7, rat 408 Myh7, and mouse 1.0 kb Myh7b promoter reporters have been described previously (15, 57, 58). The MEF2A plasmid was generated by PCR amplification from cDNA of human MEF2A followed by cloning into the pcDNA3.1 vector. Site-directed mutagenesis was performed using the Quik-Change Kit (Stratagene) according to the manufacturer’s protocol with mutated nucleotides shown in bold and with lowercase letters: 5’- GTAcgcgTAG (Myh7b MEF2mut). All constructs were confirmed by DNA sequencing. HEK293T cells were obtained from the American Type Culture Collection, and were cultured at 37°C and 5% CO2 in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal calf serum, 1,000 U/ml penicillin and 100 mg /ml streptomycin. Transient transfections in HEK293T cells were performed using PEI Transfection Reagent (Polysciences) following the manufacturer’s protocol. Luciferase reporter assays were
performed as previously described (56). Briefly, 100 ng of reporter was cotransfected with 50 ng of pcDNA3.1-MEF2A and 100 ng of pCMV6-MLL4 and 25 ng of CMV promoter-driven Renilla luciferase to control for transfection efficiency. Cells were harvested 48h after transfection. The luciferase assay was performed using Dual-Glo (Promega) according to the manufacturer’s recommendations. All transfection data are presented as the mean ± standard error of the mean (SEM) for at least three separate transfection experiments.

Cell culture and adenoviral infection

Primary muscle cells were isolated from skeletal muscles as previously described (14). For differentiation, cells were washed with PBS and re-fed with 2% horse-serum/DMEM differentiation medium and re-fed daily. Primary myoblasts were infected with an adenovirus overexpressing Cre or control virus as previously described (14), 12 hours post-infection, cells were induced to differentiate into myotubes for 60 hours prior to harvest.

Immunoprecipitation

Whole lysate from C2C12 myotubes or lysate from HEK293T cells 48 hours post-transfection were used for co-immunoprecipitation studies. HEK293T and C2C12 Cells were obtained from the American Type Culture Collection and were cultured at 37°C and 5% CO2 in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% FBS, 1,000 U/ml penicillin and 100 g/ml
streptomycin. For myocyte differentiation, C2C12 cells were cultured with 2% horse-serum/DMEM differentiation medium. C2C12 myotubes or HEK293 T cells were collected in lysis buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 2 mM EDTA, 1.5% NP40, 1x Complete (Roche), and 1 mM PMSF). 1 μg of M2 anti-FLAG (Sigma) or anti-HA (Sigma) antibodies were incubated with extract and protein G-conjugated agarose beads and the immunoprecipitated proteins were analyzed by immunoblotting.

**Statistics**

All mouse and cell studies were analyzed by Student’s t test (2-tailed) or ANOVA coupled to a Fisher’s least significant difference (LSD) post-hoc test when more than two groups were compared. No statistical methods were used to predetermine sample sizes, and sample size (range from n = 3 to n = 13) are explicitly stated in the figure legends. All data points were used in statistical analyses. Data represent the mean ± SEM, with a statistically significant difference defined as a value of \( P < 0.05 \). Statistical analyses in human studies were performed using Graphpad Prism 6 software. Gene expression levels in human studies were analyzed using paired Student’s t test, the Pearson correlation or Spearman correlation test. Significant differences were defined as \( P < 0.05 \).

**Study approval**

All animal studies were conducted in strict accordance with the institutional guidelines for the humane treatment of animals and were approved by the IACUC.
committees at the Model Animal Research Center (MARC) of Nanjing University. Human muscle samples were collected from 40 female AIS patients. Clinical ethical approval in comply with the Declaration of Helsinki was obtained from the Ethical Committee of Drum Tower Hospital, Nanjing, China. Details of the research project and biopsy collection were explained to all subjects and/or their parents or guardians before entering into the study. Informed consent was obtained from all subjects.
Author Contributions

L.L. and C.D. contributed equally to this work and performed most of the experiments with assistance from T.F., Z.F., J.L., L.X., Z.X., Y.Y., Q.G., Z.S., W.S., Y.M., L.Y., Z.Z., D.Z., L.X.. Order of co-first authors is based on the length of time spent on the project. L.L., C.D. and Z.G. designed experiments, discussed data, and wrote the manuscript. Z.Z, Y.Q. and K.G. contributed reagents and provided scientific insight and discussion. Z.G. supervised the work. All authors reviewed and contributed to the manuscript.
Acknowledgements

Special thanks to Dr. Daniel P. Kelly (University of Pennsylvania), Dr. Aibin He (Peking University), Dr. Yong Liu (Wuhan University), Drs. Zhongzhou Yang and Min-Sheng Zhu (Nanjing University) for insightful discussions; Dr. Hao Yin (Shanghai Institute for Biological Sciences) for advice on myotube IF staining studies.

This work was supported by grants from the Ministry of Science and Technology of China (National Key R&D Program of China 2018YFA0800700 and 973 Program 2015CB856300), the National Natural Science Foundation of China (No. 31922033, 91857105 and 31771291), Natural Science Foundation of Jiangsu Province (BK20170014 and SWYY-002) to Z.G., Fundamental Research Funds for the Central Universities 090314380036 (to T.F.), 090314380031 and 090314380035 (to Z.G.).

The authors declare that they have no conflict of interest.
References


8. Holloszy JO, and Coyle EF. Adaptations of skeletal muscle to endurance exercise
and their metabolic consequences. *J Appl Physiol Respir Environ Exerc Physiol.*


31. Ang SY, et al. KMT2D regulates specific programs in heart development via


39. Black BL, and Olson EN. Transcriptional control of muscle development by


47. Mannion AF, Meier M, Grob D, and Muntener M. Paraspinal muscle fibre type
alterations associated with scoliosis: an old problem revisited with new evidence.


Liu et al, Figure 1

A

B

C

D

E

F

G

H
Figure 1. Slow myofibers enriched MLL4 is required for type I muscle fiber formation. (A) Representative Western blot analysis of protein extracts prepared from white vastus lateralis (WV) and soleus muscles of wild-type (WT) mice using indicated antibodies. Quantification of the MLL4/Tubulin, EZH2/Tubulin and SUZ12/Tubulin signal ratios were normalized (=1.0) to the WV and presented below the corresponding bands. n = 5-6 mice per group. (B) Representative Western blot analysis of MLL4 expression in gastrocnemius (GC) muscles of indicated mice. n.s., none-specific band, n = 3 mice per group. (C) (Top) Representative wheat germ agglutinin (WGA) staining of GC muscle from 8-week-old male $\text{Mll}4$ mKO and $\text{Mll4SET}$ mKO mice. Scale bars: 50 μm. (Bottom) Cross-sectional areas of GC myofibers were measured by ImageJ, n = 4-5 mice per group. (D) Expression of the slow-twitch myosin gene ($\text{Myh7}$) and representative slow/fast-twitch troponin genes (qRT-PCR) in GC muscle from indicated genotypes. n = 5-8 mice per group. (E) Cross-section of (Top) soleus and (Bottom) GC muscle from 8-week-old male $\text{Mll}4$ mKO and $\text{Mll4SET}$ mKO mice stained for MHC1. MHC1 (green), and MHC2b (red). Scale bar: 250 μm. (F) Quantification of IF data shown in (E). n = 3-5 mice per group. (G, H) Primary myoblasts isolated from $\text{Mll4SET}^{\text{fr}}$ mice were infected with an adenovirus overexpressing Cre or control virus (Ctrl), followed by differentiation into myotubes. (G) Results of qRT-PCR and Western blot analysis in skeletal myotubes. n = 3 independent experiments. (H) (Left) IF staining of skeletal myotubes was performed using antibodies directed against myosin-slow or myosin-fast. Scale bars: 100 μm. (Right) Quantification of the myosin-slow IF data expressed as mean percentage total myotubes. n = 3 independent experiments. Values represent mean ± SEM, *P < 0.05 vs. corresponding controls, P-value was determined using two-tailed unpaired Student’s $t$-test.
**Figure 2. Loss of muscle MLL4 causes reduced running endurance.** (A) MHC fiber typing by IF of GC muscle of indicated genotypes. MHC1 (green), and MHC2b (red). Representative images were shown. Scale bar: 250 μm. (B) Quantification of IF data shown in (A) expressed as type I fibers per section. n = 5-6 mice per group. (C) (Left) Schematic depicts the increments of speed over time. (Right) Bars represent mean running time and distance for 10-week-old male Mll4SET^f/f^HSA-Cre mice and WT on a motorized treadmill. n = 11-13 mice per group. (D, E) Blood glucose, TG, NEFA and β-hydroxybutyrate levels in mice of indicated genotypes at rest or after 80 min exercise. n = 8-9 mice per group. Values represent mean ± SEM, *P < 0.05 vs. corresponding WT controls, P-value was determined using two-tailed unpaired Student’s t-test.
**Liu et al, Figure 3**

**A**

![Bar graph showing fold change and P-values](image)

**B**

<table>
<thead>
<tr>
<th>GO cellular compartments term</th>
<th>Count</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>cytoplasm</td>
<td>390</td>
<td>8.20E-10</td>
</tr>
<tr>
<td>membrane</td>
<td>401</td>
<td>1.30E-08</td>
</tr>
<tr>
<td>myofibril</td>
<td>11</td>
<td>9.30E-06</td>
</tr>
<tr>
<td>sarcoplasmic reticulum</td>
<td>11</td>
<td>2.70E-04</td>
</tr>
<tr>
<td>sarcolemmma</td>
<td>16</td>
<td>3.20E-04</td>
</tr>
<tr>
<td>myosin complex</td>
<td>10</td>
<td>5.30E-04</td>
</tr>
<tr>
<td>I band</td>
<td>7</td>
<td>6.00E-04</td>
</tr>
<tr>
<td>cell surface</td>
<td>47</td>
<td>1.20E-03</td>
</tr>
<tr>
<td>MHC class II protein complex</td>
<td>5</td>
<td>1.60E-03</td>
</tr>
<tr>
<td>neuronal cell body</td>
<td>40</td>
<td>2.80E-03</td>
</tr>
<tr>
<td>voltage-gated potassium channel complex</td>
<td>11</td>
<td>3.20E-03</td>
</tr>
<tr>
<td>sarcoplasmic reticulum membrane</td>
<td>7</td>
<td>3.50E-03</td>
</tr>
<tr>
<td>troponin complex</td>
<td>4</td>
<td>4.50E-03</td>
</tr>
</tbody>
</table>

**C**

Heatmap showing gene expression changes with color gradient indicating Log2(fold change) and Log10(P-value).

**D**

Graphs showing normalized AU for Contraction, Calcium-handling, and Metabolism.

**E**

Images of gels showing LDH isoforms with labels for each isoform (LDH5, LDH4, LDH3, LDH2, LDH1).

**F**

Bar graphs showing O2 flux per mass for Pyruvate and Palmitoylcarnitine.

**G**

Graph showing OCR with time points and different treatments (Ad-Ctrl, Ad-Cre).
Figure 3. MLL4 coordinately regulates gene programs controlling muscle contraction and energy metabolism. (A) Volcano plot showing fold changes versus P-values for the analyzed RNA-seq data generated from the GC muscle of 8-week-old male Mll4SET mKO mice compared to littermate controls (WT). Significantly upregulated genes are represented by red dots, whereas downregulated genes are represented by blue dots. (B) Gene ontology (GO) enrichment analysis of gene transcripts regulated in Mll4SET mKO muscle. (C) Heat-map analysis of contraction-related genes regulated in Mll4SET mKO muscle compared to WT controls. n = 3 independent samples per group, color scheme for fold change is provided. (D) Expression of genes (qRT-PCR) involved in muscle contraction, calcium handling and metabolism in the GC muscle from the indicated genotypes. n = 5-6 mice per group. (E) (Left) LDH isoenzymes were separated by polyacrylamide gel electrophoresis using whole cell extracts from heart (Ht, control) and GC muscle from the indicated mice. A representative gel showing 4 independent mice per group is shown. (Right) Quantification of LDH isoenzyme activity gel electrophoresis shown on the left. (F) Mitochondrial respiration rates were determined from the extensor digital longus muscle of the indicated genotypes using pyruvate or palmitoylcarnitine as a substrate. Pyruvate/malate (Py/M) or palmitoylcarnitine/malate (PC/M)-stimulated, ADP-dependent respiration and oligomycin-induced (oligo) are shown. n = 6-7 mice per group. (G) Oxygen consumption rates (OCR) in skeletal myotubes harvested from Mll4SET f/f mice subjected to adenovirus-based overexpression of Cre compared with control virus (Ctrl). Basal OCR was first measured, followed by administration of 10 mM sodium pyruvate, and 2 μM oligomycin (to inhibit ATP synthase), uncoupler FCCP (2 μM), or rotenone/antimycin (Rot/A; 1 μM) as indicated. n = 6 separate experiments done with 5 biological replicates. Values represent mean ± SEM. *P < 0.05 vs. corresponding controls, P-value was determined using two-tailed unpaired Student’s t-test.
**Skeletal muscle cell differentiation**

**Log**

**10**

(P-value)

**Cardiac muscle contraction**

**Positive regulation of transcription from RNA polymerase II promoter**

**Regulation of cell proliferation**

**Regulation of muscle contraction**

---

**D**

**Motif analysis: MLL4 binding sites**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Motif matrix</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEF2B</td>
<td>CTAxAAAAxTAG</td>
<td>1e-70</td>
</tr>
<tr>
<td>MEF2C</td>
<td>CTAxAAAAxTAG</td>
<td>1e-68</td>
</tr>
<tr>
<td>RLM1</td>
<td>CTAxAAAAxTAG</td>
<td>1e-68</td>
</tr>
<tr>
<td>MEF2D</td>
<td>CTAxAAAAxTAG</td>
<td>1e-63</td>
</tr>
<tr>
<td>MEF2A</td>
<td>GACxAAAAxTAG</td>
<td>1e-34</td>
</tr>
<tr>
<td>NR3C1</td>
<td>GACxAAAAxTAG</td>
<td>1e-11</td>
</tr>
</tbody>
</table>

---

**E**

**MLL4 de novo motif:**

p = 1e-137

---

**G**

**Input**

**MEF2D**

**WT**

**KO**

**MLL4**

**WT**

**KO**

**H3K4me1**

**WT**

**KO**

**H3K27ac**

**WT**

**KO**

**mRNA-seq**

**KO**

---

**Hdhh**

---

**Liu et al, Figure 4**

---

**52**
Figure 4. MLL4 co-localizes with MEF2 on enhancers of slow-oxidative muscle genes. (A) Analysis of the MLL4 ChIP-seq data in muscle and myocytes combined with mRNA-seq dataset upon muscle *Mll4* deletion defines a set of genes directly regulated by MLL4. (B) GO enrichment analysis of MLL4 direct targets, with the top five terms shown. (C) De novo motif analysis of MLL4 binding regions in muscle using Homer. Shown is the top-scoring motif present in the top 5000 emergent MLL4 binding sites. (D) Known TF motifs with the highest relative enrichment in MLL4 binding regions in muscle. (E) MLL4 co-localizes with MEF2 on active enhancers in myocyte. Heat-map shows the ChIP-seq binding signal intensity for MLL4, MEF2D, H3K4me1 and H3K27ac. Binding is ranked from the strongest to the weakest MLL4 binding sites. Active enhancers are defined with enhancer markers H3K4me1 and H3K27ac. (F) Gene set enrichment analysis (GSEA) showing that expression of MLL4/MEF2 co-bound target genes is significantly downregulated in *Mll4SET* mKO muscle compared to WT controls. All genes in mRNA-Seq profiling from *Mll4SET* mKO muscle were ranked by fold difference compared to WT controls and expressed on the x-axis. This data set was compared with the gene list of nearest genes identified in MLL4/MEF2 co-bound regions in muscle. (G, H) MLL4-dependent active enhancers on slow-oxidative gene loci are shown. (Top) ChIP-seq binding profiles for MLL4, MEF2D and histone modifications in WT or *Mll4* KO myocytes. mRNA-seq data from WT and *Mll4SET* mKO muscle is shown at the bottom, indicating a high correlation of the two datasets. Input, genomic DNA from myocytes; Gray box indicates the high confidence MLL4 binding regions corresponding to *Tnncl*, *Tnni1*, and *Ldhb* genes. Data obtained from published data sets GSE50466 and GSE43223, as well as current data set GSE138994 and GSE137368 were analyzed.
Figure 5. MLL4 cooperates with MEF2 to activate slow myofiber genes transcription. (A, B) Co-IP experiments were performed by cotransfecting HA-MEF2A and Flag-MLL4 in HEK293 cells as indicated at the top. Antibodies against the HA or Flag epitope were used for co-IP. The extracts (Input) from the HEK293 cells and the proteins from the IP were analyzed by immunoblotting (IB). Representative results for co-IP are shown. n = 3 independent experiments. (C) Co-IP results with extracts prepared from C2C12 myotubes using anti-MLL4 antibody or control IgG. Representative results are shown. n = 3 independent experiments. (D) MLL4 and MEF2D ChIP-seq tracks from myocyte at the Myh7 locus. Two putative conserved MEF2-binding sites within the cis-proximal enhancer region of the Myh7 gene are shown. (E) MLL4 and MEF2 synergistically activate Myh7 gene promoter. Values represent mean (± SEM) firefly/renilla luciferase activity shown as arbitrary units (AU) normalized (=1.0) to vector control in HEK293 cells after cotransfection with expression vectors indicated and the rat Myh7.Luc.3.5k promoter reporter. n = 4 independent experiments. (F) Results of transient transfection performed with tMyh7.Luc.3.5k and truncation mutant of rMyh7.Luc.408 in HEK293 cells in the presence of expression vectors indicated. n = 4-5 independent experiments. (G) (Left) MLL4 and MEF2D ChIP-seq tracks from myocyte at the Myh7b locus. (Right) The putative conserved MEF2-binding site within the mouse Myh7b promoter regions. (H) (Top) Site-directed mutagenesis was used to abolish the MEF2 response element. (Bottom) The mMyh7b.Luc.1k (WT) or MEF2mut.mMyh7b.Luc.1k promoter reporters was used in cotransfection studies in HEK293 cells in the presence of expression vectors indicated. n = 5 independent experiments. Values represent mean ± SEM, *P < 0.05 vs. corresponding controls, #P < 0.05 compared to MEF2A alone. P-value was determined using one-way ANOVA coupled to a Fisher’s least-significant difference (LSD) post-hoc test.
A Human skeletal muscle microarray (GSE1718) Pre-Training Post-Training

B ATPase (pH 10.2)

C

D

E

Liu et al, Figure 6
Figure 6. MLL4 regulatory circuit is associated with muscle fiber type remodeling in humans. (A) Relative expression of MLL4 and representative fiber type-specific genes in human vastus lateralis muscles before or after an endurance training. Data were extracted from GEO: GSE1718. (B-E) Paraspinal muscle samples from the convex (“MX”) and concave (“MV”) sides of the curvature from 40 patients with adolescent idiopathic scoliosis (AIS) were used for this analysis. (B) (Left) Representative sections of paraspinal muscle from AIS patients stained for myosin ATPase activity (pH=10.2, type II fibers dark, type I fibers light). Scale bar: 100 μm. (Right) Quantification of ATPase staining data shown on the left expressed as mean percentage of total muscle fibers. n = 15 patients. Significant differences were analyzed using paired Student’s t test. (C) mRNA expression levels of MLL4, MEF2s and slow-oxidative muscle fiber genes were determined by qRT-PCR. n = 40 patients. Significant differences were analyzed using paired Student’s t test. (D) Correlation between MLL4 gene expression and that of slow-oxidative myofiber genes and MEF2s. n = 40 patients. Pearson correlation analysis was used to determine the correlation. (E) Correlation between MLL4 and MEF2s expression and the type I fiber percentage. n = 15 patients. Spearman correlation analysis was used to determine the correlation.
Figure 7. Model of MLL4 in the control of muscle fiber type specification and function. The schematic depicts a proposed model for the MLL4 regulatory mechanism that specifies muscle fiber identity and muscle performance.