Published in final edited form as:

Ann Neurol. 2018 August ; 84(2): 234–245. doi:10.1002/ana.25283.

Long-range genomic regulators of *THBS1* and *LTBP4* modify disease severity in Duchenne muscular dystrophy

Robert B. Weiss, PhD^{1,*}, Veronica J. Vieland, PhD^{2,4,6}, Diane M. Dunn, BS¹, Yuuki Kaminoh, BS³, and Kevin M. Flanigan, MD^{3,4,5} **United Dystrophinopathy Project**

¹Department of Human Genetics, University of Utah, Salt Lake City, Utah

²Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital, Columbus, Ohio

³Center for Gene Therapy, The Research Institute at Nationwide Children's Hospital, Columbus, Ohio

⁴Department of Pediatrics, The Ohio State University, Columbus, Ohio

⁵Department of Neurology, The Ohio State University, Columbus, Ohio

⁶Department of Statistics, The Ohio State University, Columbus, Ohio

Abstract

Objective—Duchenne muscular dystrophy (DMD) is a severe X-linked recessive disease caused by loss-of-function dystrophin (*DMD*) mutations in boys, who typically suffer loss of ambulation by age 12. Previously we reported that coding variants in latent transforming growth factor beta-binding protein 4 (*LTBP4*) were associated with reduced TGF β signaling and prolonged ambulation (p-value = 1.0×10^{-3}) in DMD patients; this result was subsequently replicated by other groups. In this study, we evaluated whether additional DMD modifier genes are observed using whole genome association in the original cohort.

Methods—We performed a genome-wide association study (GWAS) for SNPs influencing loss of ambulation (LOA) in the same cohort of 253 DMD patients used to detect the candidate

*Corresponding author: Robert B. Weiss, PhD, Eccles Institute of Human Genetics, Rm 2260, 15 North 2030 East, Salt Lake City, UT 84112, Phone: (801) 585-3435, bob.weiss@genetics.utah.edu.

AUTHOR CONTRIBUTIONS

R.B.W., V.J.V. and K.F.M conceived and designed the study. R.B.W., V.J.V. D.M.D., Y.K., K.F.M and the UDP Investigators acquired and analyzed the data. R.B.W., V.J.V. and K.F.M drafted the manuscript.

CONSORTIA

UDP Investigators: Kathryn J. Swoboda, MD; Eduard Gappmaier, MPT, PhD; Michael T. Howard, PhD; Jacinda B. Sampson, MD, PhD; Mark B. Bromberg, MD, PhD; Russell Butterfield, MD, PhD; Lynne Kerr, MD, PhD; Alan Pestronk, MD; Julaine M. Florence, MPT; Anne Connolly, MD; Glenn Lopate, MD; Paul Golumbek MD, PhD; Jeanine Schierbecker MHS, PT; Betsy Malkus MHS, PT; Renee Renna, RN; Catherine Siener, MHS, MPT; Richard S. Finkel, MD; Carsten G. Bonnemann, MD, PhD; Livija Medne, MS; Allan M. Glanzman, MPT, DPT, PCS, ATP; Jean Flickinger, RPT; Megan Waldrop, MD; Roxane Alles, PhD; Jerry R. Mendell, MD; Wendy M. King, MPT; Linda Lowes, PhD; Lindsay Alfano, DPT; Laurence Viollet, PhD; Katherine D. Mathews, MD; Carrie Stephan, RN; Karla Laubenthal, MPT, MS, PCS; Kris Baldwin, LPT; Brenda Wong, MD; Paula Morehart, RN; Amy Meyer, MPT; John W. Day, MD, PhD; Cameron E. Naughton; and Marcia Margolis, MPT, ATP; Craig M. McDonald, MD; R. Ted Abresch, MS; Michelle Cregan, CRA; Jay J. Han, MD; Eric Henricson, MPH; Linda Johnson, MPT.

CONFLICTS OF INTEREST

Nothing to report.

association with *LTBP4* coding variants. Gene expression and chromatin interaction databases were used to fine map association signals above the threshold for genome-wide significance.

Results—Despite the small sample size, two loci associated with prolonged ambulation met genome-wide significance and were tagged by rs2725797 (chr15, p-value = 6.6×10^{-9}) and rs710160 (chr19, p-value = 4.7×10^{-8}). Gene expression and chromatin interaction data indicated that the latter SNP tags regulatory variants of *LTBP4*, while the former SNP tags regulatory variants of thrombospondin-1 (*THBS1*): an activator of TGF β signaling via direct binding to LTBP4 and an inhibitor of pro-angiogenic nitric oxide signaling.

Interpretation—Together with prior evidence implicating *LTBP4*, the *THBS1* modifier locus emphasizes the role that common regulatory variants in gene interaction networks can play in mitigating disease progression in muscular dystrophy.

Introduction

Duchenne muscular dystrophy (DMD [MIM: 310200]) is a severe, progressive muscular dystrophy that affects 1 per ~5000 live male births^{1,2}. The extreme size of the *DMD* gene (2.2 Mb) leads to a range of primary mutations, including mutations that produce reduced amounts of functional dystrophin and result in a milder form of the disease, Becker muscular dystrophy (BMD [MIM: 300376]). Complete loss-of-function *DMD* mutations disrupt the dystrophin-glycoprotein complex leading to muscle membrane fragility, with eventual muscle tissue loss and replacement by fat and fibrotic tissue. However, even without dystrophin protein, there is variability in clinical progression of disease symptoms suggesting that genetic background may play a role in modulating the disease process. Consistent with observations of genetic modifiers of muscle disease severity in animal models^{3,4}, several human candidate gene associations with age at loss of ambulation (LOA) have identified variants in *SPP1*⁵, *LTBP4*⁶, and *CD40*⁷ as genetic modifiers of DMD.

In mouse models of muscular dystrophy, a coding variant in latent TGF β binding protein 4 (*Ltbp4*) is associated with disease severity, and modifies both muscle membrane permeability and fibrosis³. Concordant with mouse *Ltbp4*, human *LTBP4* coding polymorphisms are associated with disease severity in multiple DMD cohorts^{6,8,9}. Four nonsynonymous *LTBP4* SNPs (rs2303729, p.Val127Ile; rs1131620, p.Thr720Ala; rs1051303, p.Thr753Ala; and rs10880 p.Thr1074Met; NP_001036010.1) are in moderate linkage disequilibrium (LD), resulting in two major LTBP4 protein coding haplotypes, designated as VTTT (risk) and IAAM (protective) protein isoforms. Both human and mouse protective alleles show reduced TGF β signaling in fibroblast assays, and the human IAAM isoform displays higher avidity for TGF β ¹⁰, consistent with LTBP4 controlling the bioavailability of latent TGF β , thus leading to reduced dystrophic muscle damage.

An earlier candidate gene association with osteopontin (*SPP1*), an extracellular matrix protein and proinflammatory cytokine, implicated a *SPP1* promoter SNP as a modifier of muscle weakness progression in DMD⁵. *SPP1* gene expression is highly induced in dystrophic muscle^{11,12} and knockout of *Spp1* in the *mdx* mouse (a stop codon mutation in exon 23 of the *Dmd* gene) results in a milder muscle phenotype¹³, and also shifts infiltrating macrophages to a regenerative phenotype¹⁴. An analysis that combined both *Spp1* ablation

and protective *Ltbp4* variants in the mouse suggests they work synergistically to reduce the harmful effects of TGF β signaling in dystrophic muscle¹⁵.

To detect genome-wide modifiers of disease severity in DMD, we make use of patients enrolled in the United Dystrophinopathy Project (UDP), a multi-site consortium established to study genotype-phenotype correlations in DMD, BMD, and related dystrophinopathy patients^{16–18}. Using a candidate gene approach, we previously reported associations in the UDP cohort between coding variants in *LTBP4* and age at Loss of Ambulation (LOA) under a recessive model⁶. Here we report a genome-wide association analysis in this same cohort, using SNP data that passed quality control for 253 DMD patients for association with LOA, with steroid use and *DMD* gene mutation type as covariates.

Methods

Subjects and Clinical Classification.

All subjects have a known *DMD* mutation confirmed by either MLPA or DNA sequencing, and have been previously described¹⁶. Informed consent was obtained under institutional review board-approved protocols. Genomic DNA was extracted from venous blood samples and phenotypic data was extracted from clinical records; most subjects underwent prospective clinical examinations as well. Ambulatory loss was defined by full-time wheelchair use (requiring wheelchair within the home) and recorded to the nearest month when available, and otherwise to the nearest half-year of age. All patients in this analysis had ambulatory loss before 20 years of age (LOA mean = 10.6, s.d. = 2.3). Glucocorticoid treatment was defined as use of any steroid regimen for more than six months that began at least six months prior to loss of ambulation (137 subjects, mean age at ambulatory loss was 10.9 ($\pm$ 2.3) years). Untreated was defined as never exposed, exposed to any steroid regimen for less than six months, or had onset of treatment less than six months prior to loss of ambulation (101 subjects, mean age at ambulatory loss was 10.0 ($\pm$ 2.1) years). The mutation class assignment was either ‘truncating’ *DMD* mutations (211 subjects, mean age at ambulatory loss was 10.4 ($\pm$ 2.2) years) or ‘other’ *DMD* mutations (42 subjects, mean age at ambulatory loss was 11.3 ($\pm$ 2.4) years).

Illumina Array Genotyping and Quality Control.

Whole-genome genotyping was performed on archived genomic DNA that had been previously collected for *DMD* mutation analysis. Human Infinium Omni2.5Exome-8 v1.2 BeadChips were processed with standard Illumina protocols, the fluorescence signal collected on an Illumina iScan System scanner, and genotypes called with Illumina GenomeStudio software. The median genotype call rate was 99.87%, with only 4 arrays called below 99.0% (98.90%, 98.79%, 98.41% and 98.14%). Cleaning of the initial 2,608,369 SNPs removed: 1,157 SNPs in the non-pseudoautosomal region of X exhibiting excess heterozygosity (10%; remaining heterozygous genotypes were set to missing); 1,752 SNPs with excess missingness (20%); 1,522 SNPs showing Hardy-Weinberg (HWE) violations at a threshold of $P = 0.0001$, and 1,814 SNPs with ambiguous chromosome/position assignments. For common variant tests, an additional 866,175 markers were dropped because they had ≤ 3 occurrences of the minor allele in this data set (the

default cutoff for MENDEL, see below); of these, 559,912 were monomorphic. This left a total of 1,735,619 SNPs in the analysis. For gene-based rare variant testing, exonic SNPs were filtered by genotyping rate $\geq 98\%$, HWE ≥ 0.0001 , minor allele frequency $\geq 5\%$ and annotation to include non-synonymous, stop gained, stop lost and splice site variants. An initial list 42,514 bi-allelic sites were assigned to 12,671 genes, yielding a Bonferroni-corrected threshold of 3.9×10^{-6} for gene-wise significance. Additionally, variants were filtered by pathogenicity using a standardized scoring metric (C-score ≥ 20) from the Combined Annotation Dependent Depletion method¹⁹, resulting in 19,427 bi-allelic sites assigned to 9,083 genes, with a Bonferroni-corrected threshold of 5.5×10^{-6} for gene-wise significance.

Genotype Imputation and Haplotype Phasing.

SNPs on chromosome 15 and 19 were imputed using EAGLE2 (v2.0.5) pre-phasing with the Haplotype Reference Consortium panels (release 1.1) and imputed with PBWT using the Sanger Imputation Service pipeline²⁰. Phased haplotypes were extracted from VCF output files with filtering for allele frequency and minimum IMPUTE2 info score of 0.8. Linkage disequilibrium bins were assigned using ldSelect²¹ with an r^2 threshold of 0.64.

Statistical Methods.

Genome-wide association analyses were conducted in Mendel (v.16.0)²², using the Ped-GWAS option, which carries out a classical variance-components QTL analysis allowing for the inclusion of related individuals in the dataset. Relationships among individuals were determined using relativeFinder (Goncalo Abecasis, <http://genome.sph.umich.edu/wiki/relativeFinder>) and confirmed against clinical records; the data set included 8 full-sibling pairs and 1 first-cousin pair, which were coded as such in the MENDEL input files (KINSHIP_SOURCE = pedigree_structure). The SmartPCA routine from EIGENSOFT 6.0.1²³ was used to determine ancestral outliers, using the default selection criteria (10 principal components, 6 standard deviations for flagging outliers along those components); this suggested 10 individuals likely to be outliers (including one of the full-sibling pairs). These individuals were left in the primary analyses, but at SNPs of interest (see Results) follow up analyses were conducted with these individuals removed. To further investigate population structure, we combined the UDP samples with Omni 2.5 genotypes from 1000 Genomes reference populations of European ancestry, including CEU (183 samples), GBR (104 samples), FIN (100 samples), IBS (150 samples) and TSI (112 samples). Genotype data in vcf format (ALL.chip.omni_broad_sanger_combined.20140818.snps.genotypes.vcf.gz) was downloaded from ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/supporting/hd_genotype_chip/. The merged genotype set was pruned for linked SNPs using PLINK indep-pairwise functionality, keeping 291,366 variant sites for ancestry analysis. We ran the ADMIXTURE software²⁴ using default parameters with $K=3$ on the combined genotypes and for testing the effect of population stratification, filtered the UDP samples by the minimum ancestry fraction value ($=0.967$) observed for the CEU population.

The phenotype was Age at Loss of Ambulation (LOA), and primary analyses included the covariates Steroid Use and *DMD* Mutation Class. Data were initially analyzed under a

recessive model, which is the model that had previously been established for *LTBP4*; the genome scan was subsequently repeated under additive and dominant models. Global and haplotype-specific tests for association used the haplo.score routine from the haplo.stats²⁵ package in R. Gene-based rare-variant testing used the optimal Sequence Kernel Association Test²⁶ (SKAT-O) as implemented in the Efficient and Parallelizable Association Container Toolbox EPACTS, using the skat-o small sample adjustment version of the mmskat test. Rare variant associations included Steroid Use and *DMD* Mutation Class as covariates and used a linear mixed model (EMMAX) adjustment to account for population structure and relatedness. *LTBP4* haplotypes were estimated from unphased genotype data and evaluated for association with age of ambulatory loss as a quantitative trait or RNA-seq TPM values for *LTBP4* expression levels in macrophages. The regional linkage disequilibrium plots were generated with LocusZoom²⁷.

Genotype-Tissue Expression and Chromatin Interaction Data.

The GTEx portal (GTEx Analysis Release V6p, dbGaP Accession phs000424.v6.p1, <https://www.gtexportal.org>) was used to access pre-calculated eQTL statistics for chromosome 15 SNPs and thrombospondin 1 mRNA expression (*THBS1*, ENSG00000137801.9) in skeletal muscle (gastrocnemius) tissue, and sample-level box plots of rank normalized gene expression. *LTBP4* expression data (ENSG00000090006) in macrophages was obtained from the NCBI GEO data (GEO accession number GSE81046, processed data file with gene-level TPM values downloaded from GSE81046_RAW.tar). Sample-level genotype data for GSE81046 was provided by Dr. Luis B. Barreiro, CHU Sainte-Justine/University of Montreal²⁸. UCSC genome browser BED Files of enhancers from human skeletal muscle myotubes were retrieved from Data S2 zipped file from Hinsz et al.²⁹.

Additional chromatin analysis tracks were retrieved from the UCSC Genome Browser (Human Feb. 2009, GRCh37/hg19 Assembly) included: Human Skeletal Muscle Myoblasts (HSM) Chromatin State Segmentation by HMM from ENCODE/Broad, Transcription Factor ChIP-seq Uniform Peaks from ENCODE/Broad/Analysis for CTCF in Human Skeletal Muscle Myotubes, Transcription Factor ChIP-seq (CTCF) from ENCODE with Factorbook Motifs, and DNaseI Hypersensitivity Clusters in 125 cell types from ENCODE (V3). The Promoter Capture Hi-C (PCHi-C) peak matrix file from 17 human primary hematopoietic cell types, consisting of CHiCAGO scores for all interactions that pass a cutoff of ≥ 5 in at least one cell type were downloaded from the Open Science Framework repository (accession u8tzp)³⁰.

Results

Identification of genome-wide significant variants

Prior candidate gene studies in DMD cohorts using age at loss of ambulation as the phenotype reveal non-additive intra-locus effects of the modifier SNPs. *LTBP4* SNPs/haplotypes show greater protective effects with a recessive model^{6,8,9}, while *SPP1* and *CD40* SNPs show greater risk effects with a dominant model^{5,7}, therefore we performed an initial GWAS for all three (recessive, additive and dominant) genetic models. Three SNPs tagging two loci exceeded a conventional genome-wide significance threshold of p-value

5×10^{-8} ($-\log_{10}[\text{p-value}] = 7.3$): rs2725797 and rs2624259 at 15q14 ($r^2 = 0.95$), and rs710160 at 19q13.2 (Fig 1, Table 1) using the previously supported recessive model for age at loss of ambulation^{6,8,9}. Under additive and dominant models, no single locus exceeded a genome-wide significance threshold, although rs2624259 and rs2725797 at 15q14 remained among the top 15 associated loci ($\text{p-value} = 7.5 \times 10^{-6}$) under an additive model (Table S1). Quantile-quantile plots of the p-value distributions indicated little systematic bias and estimation of ancestry using principal component analysis or model-based estimation showed that the 253 UDP subjects have predominantly European genetic ancestry (Figure S1). After removing 10 subjects classified as ancestral outliers by principal component analysis, the top three SNPs remained within the top ten genome-wide signals ($N=243$, Table 1). A more stringent filter for population stratification required European ancestry equivalent to the CEU reference population (Utah residents with ancestry from northern and western European ancestry), and after removing an additional 54 subjects, the association of rs2725797, rs2624259 and rs710160 with LOA remained robust ($N=189$, Table 1). All SNP results with a nominal $-\log_{10}[\text{p-value}] > 2.0$ in both the $N=253$ and $N=243$ analyses for all three genetic models are shown in Tables S2–S4. We also performed a gene-level association test for binned low frequency ($\text{MAF} \leq 0.05$) exonic variants, using SKAT-O adjusted for steroid use, *DMD* mutation class, and sample structure. The gene-based rare variant test did not identify any significant associations exceeding the multiple testing threshold, either in the $N = 253$ and $N = 243$ sample groups, or after filtering variants by relative pathogenicity. Since only the recessive model yielded SNPs that passed genome-wide significance threshold, we focused on analyzing functional data that would support a biological role for these variants in modifying disease progression.

rs710160 is associated with LTBP4 mRNA expression and the IAAM haplotype

Since rs710160 is located ~12 kb upstream of the *LTBP4* promoter region, we investigated the relationship between rs710160 and LOA associations previously detected at four closely linked nonsynonymous *LTBP4* SNPs: the VTTT risk and IAAM protective *LTBP4* protein isoforms. rs710160 is in weak LD ($r^2 \sim 0.20$ – 0.28 , $D' \sim 0.55$ – 0.79) with VTTT/IAAM SNPs (Fig 1C), but the D' statistics suggested rs710160 alleles were unevenly distributed between VTTT/IAAM haplotypes. Phased haplotypes were estimated from the 253 *DMD* subjects and showed that the rs710160_C allele sub-divided the IAAM haplotype (Fig 2A). These five SNPs (rs710160 [T/C], rs2303729 [V/I]; rs1131620 [T/A], rs1051303 [T/A], and rs10880 [T/M]) were found on three common haplotypes, at frequencies: T-VTTT (49%), C-IAAM (17%) and T-IAAM (13%). When tested for haplotype-specific association with LOA under a recessive model (Table 2), the C-IAAM haplotype was the most protective ($p = 2 \times 10^{-5}$).

Next, we investigated the relationship of regional SNPs with *LTBP4* transcript levels by querying expression quantitative trait locus (eQTL) databases. In monocyte-derived macrophages collected from 91 European-Americans (EA) and 77 African-Americans (AA), 66 SNPs in the *LTBP4* region, including rs710160 and the four coding SNPs (rs2303729, rs1131620, rs1051303 and rs10880), were nominally associated with *LTBP4* transcript levels, with 32 *LTBP4* eQTL associations having an estimated genome-wide false discovery rate $< 1\%$ in macrophages²⁸. We re-analyzed these *LTBP4* eQTLs in the 91 EA subjects and

observed that minor alleles of rs710160 and the individual VTTT/IAAM SNPs were associated with lower *LTBP4* transcript levels in macrophages (Fig 2B). When haplotype-specific tests for expression levels were examined (Table 2), the C-IAAM haplotype associated with lower *LTBP4* transcript levels ($p = 0.004$), T-IAAM was not significantly associated ($p = 0.38$) and T-VTTT associated with higher transcript levels ($p = 0.0003$). Additionally, we searched for rs710160 eQTLs in 43 tissues and cell lines collected by the Genotype-Tissue Expression (GTEx) Project and observed that the top rs710160 association with *LTBP4* expression was in transformed fibroblasts (272 subjects: $p = 0.00039$, $\beta = -0.13$, s.e. = 0.037), where the C allele was also associated with lower *LTBP4* transcript levels (Fig 2C).

Long range chromatin interactions link rs2725797 to the thrombospondin-1 promoter

We next imputed SNPs in the chromosome 15q14 region tagged by the genome-wide significant SNPs (rs2725797 and rs2624259) and observed a strong LD block ($r^2 > 0.8$) comprising 48 SNPs located between chr15: 39,055,584–39,110,903 (Fig 1B). This LD block is in a large gene desert, however the thrombospondin-1 gene (*THBS1*) is located ~750 kb telomeric from rs2624259. Thrombospondin-1 (TSP-1, encoded by the *THBS1* gene) is a major activator of TGF β signaling and directly interacts with LTBP4 in the extracellular matrix (ECM)^{31,32}. Since *THBS1* expression is elevated in *mdx* mice³³, we examined the region between the rs2725797-tagged LD block and the *THBS1* gene for features that regulate gene expression.

We searched for evidence of long range chromatin interactions between the rs2725797-tagged LD block and the *THBS1* promoter using chromosome conformation capture data, specifically with promoter capture Hi-C (PCHi-C) data from 17 human primary blood cell types³⁰. The 3.7 kb *HindIII* fragment containing the *THBS1* promoter (chr15:39872009–39875707) pulls down 142 flanking *HindIII* fragments with high-confidence (Capture Hi-C Analysis of Genomic Organization ‘CHiCAGO’ interaction score ≥ 5 in at least one cell type) and the genomic locations of these promoter-interacting fragments are shown as arcs connecting the captured *HindIII* fragments to the *THBS1* promoter (Fig 3). We sorted the 142 captured *HindIII* fragments by their median *THBS1* promoter interaction score across all 17 cell types and observed that the two highest scoring fragments contained the two top SNPs associated with LOA (rs2725797 in 1.1 kb fragment E chr15:39108568–39109689, median score = 17.6; rs2624259 in 11.0 kb fragment F chr15:39109690–39120732, median score = 10.6).

rs2725797 and rs2624259 flank a strong enhancer feature and CTCF binding site

RNA-Seq data showed that *THBS1* is highly expressed in pediatric skeletal muscle, and features associated with transcriptionally active regulatory elements are found throughout the region between the rs2725797-tagged LD block and *THBS1* in human skeletal muscle myotubes (HSM), including stretch enhancers and chromatin marks for histone H3 acetylated at lysine 27 (H3K27ac) and CCCTC-binding factor (CTCF) sites from ChIP-Seq data (Fig 3). A prominent CCCTC-binding factor (CTCF) binding site, which mediates long-range interactions between distant enhancers and their promoters, is observed between rs2725797 and rs2624259, in HSMs and in multiple cell types. A DNase I hypersensitive

site, indicative of open chromatin and a cis-regulatory element, is seen in this same location in 122 out of 125 cell types, suggesting this region binds regulatory/modulatory proteins in a broad range of cell types.

rs2725797 is associated with reduced THBS1 mRNA expression in skeletal muscle

Given the density of regulatory features in the vicinity of rs2725797, we searched for *THBS1* eQTLs in 43 tissues and cell lines collected by the GTEx Project and observed that the top rs2725797: *THBS1* eQTL association was seen in skeletal muscle tissue (361 subjects: $p = 0.0049$, $\beta = -0.18$, s.e. = 0.063). SNPs in strong LD with rs2725797, including rs2624259, were also associated with *THBS1* expression in skeletal muscle (Fig 4A) and the minor allele (protective for LOA) was also associated with lower expression. The pattern of SNP association signals for *THBS1* skeletal muscle expression and LOA were comparable when SNPs in the entire rs2725797 LD block were examined (Fig 4B). Finally, when 1,569 SNPs from 923 kb upstream to 20 kb downstream of the *THBS1* promoter were ranked by their *THBS1* skeletal muscle eQTL statistic, 64 SNPs in strong LD with rs2725797 ($r^2 > 0.6$) ranked in the top 72 (including 58 of the top 59) regional *THBS1* cis-eQTLs.

Joint LTBP4 and THBS1 variant effects on LOA

It is noteworthy that *LTBP4* and *THBS1* encode directly interacting proteins that control TGF β bioavailability and that the longest ambulating DMD patient in our cohort (LOA at age 18) was homozygous for protective alleles at both loci. We therefore partitioned the 253 DMD patients by their number of *LTBP4* and *THBS1* alleles, and the joint effect of protective *THBS1* alleles and a homozygous rs710160 protective genotype resulted in an average delay in LOA of 1.2, 3.5 and 6.8 years for 0, 1 or 2 protective *THBS1* rs2725797 alleles (Fig 5A), respectively.

DISCUSSION

We have found evidence that regulatory variants associated with reduced expression of *LTBP4* and *THBS1* prolong ambulation in this cohort of DMD patients. Functional enhancer SNPs have been major candidates for causal variants underlying GWAS loci, and the availability of functional high throughput data from the ENCODE³⁴ and GTEx³⁵ projects were critical to interpreting the functional role of the *THBS1* enhancer SNPs. These DMD modifier loci are an example of common regulatory variants typically found by GWAS having a role in the phenotypic variability of a rare Mendelian disease. The genome-wide association strengthened and refined the known *LTBP4* IAAM association with prolonged ambulation. Fine mapping of regulatory variants in linkage disequilibrium with the IAAM coding SNPs suggests that the most protective human *LTBP4* haplotype combines reduced mRNA expression with tighter latent TGF β binding, as has been demonstrated for the IAAM protein isoform¹⁰. When transgenic human LTBP4 VTTT protein isoform is overexpressed in the *mdx* model, muscle disease is aggravated³⁶ and sarcolemmal damage of dystrophic muscle fibers is increased³⁷. The combination of reduced expression and the LTBP4 IAAM protein isoform known to have less TGF β release is consistent with mouse models where reduced latent TGF β bioavailability lessens disease progression in the injured myofiber.

TSP-1 is a multifunctional, extracellular matrix glycoprotein that is both an activator of TGF β signaling^{32,38} and an anti-angiogenic factor^{39,40}. The protective effect of the *THBS1* distal regulatory SNPs may be exerted through multiple distinct TSP-1 functions (Fig 5B). LTBP4 sequesters latent TGF β complex in the ECM through a disulfide linked latency-associated peptide (LAP) and TSP-1 activates ECM-bound latent TGF β ³⁸ via direct interaction with LAP³². TSP1-null mice have similar histological abnormalities as seen in TGF- β 1 null mice, and TSP1-null fibroblasts produce only 15% of the amount of active TGF- β 1 as wild-type fibroblasts³⁸. Upregulation of TSP-1 expression via angiotensin II mediated signaling leads to TGF β activation in both cardiac fibroblasts⁴¹ and human mesangial cells⁴². In the diaphragm of dystrophin-deficient *mdx* mice, TSP-1 expression is elevated and administration of the angiotensin II type 1 receptor blocker losartan decreases TSP-1 expression and reduces excessive TGF β signaling³³. In the dysferlin-deficient mouse model of limb-girdle muscular dystrophy type 2B, TSP-1 accumulates at high levels in skeletal muscle and is correlated with both inflammatory activity and disease progression^{43,44}. Differentiation in primary cell cultures of myotubes from dysferlin-deficient myoblasts results in upregulation of *THBS1* mRNA⁴⁵, indicating active regulation of *THBS1* transcription in muscle cells. The location of SNPs associated with LOA within the strongest enhancer-*THBS1* promoter chromatin contact across multiple cell types suggests these enhancer variants may exert their protective effect by altering the regulation of *THBS1* transcription in the injured myofiber, thus reducing TGF β activation.

TSP-1 also has direct effects on inhibiting angiogenesis through CD36- and CD47-mediated suppression of nitric oxide (NO) signaling and by modulating the activity and bioavailability of vascular endothelial cell growth factor (VEGF)⁴⁰. Activation of the CD47 receptor by TSP-1 binding leads to downstream inhibition of pro-angiogenic NO/cGMP signaling. Reduced angiogenesis via a VEGFA-dependent pathway leads to impaired regeneration of skeletal muscle fibers in *mdx* mice⁴⁶ and an increased vascular density in skeletal muscle is associated with reduced dystrophic pathology, as observed in *mdx* mice crossed to a heterozygous knockout of the VEGF receptor-1 (*mdx:Flt-1^{+/-}*)⁴⁷. TSP-1 null mice also have increased capillary density in skeletal and cardiac muscle⁴⁸. Thus, a second mechanism for the protective effects of *THBS1* regulatory SNPs may be to decrease TSP-1 levels, which promotes a pro-angiogenic response in dystrophic muscle by modulating the VEGF pathway.

The *LTBP4* and *THBS1* regulatory SNPs reported here join additional examples of expression levels ameliorating the DMD phenotype, including *Jagged1* overexpression in the golden retriever muscular dystrophy model⁴, and in human, rs28357094 in the *SPP1* promoter⁵ and rs1883832 in the *CD40* Kozak consensus sequence⁷. As in previous candidate modifier gene studies with LOA in human DMD cohorts, these new associations show non-additive (recessive) protective effects of the modifier variants on the expressivity of DMD mutations. Whether this is a general observation for DMD modifiers will require additional examples, however we suggest that recessivity in this case may arise from protein-protein and pathway interactions involving LTBP4 and TSP-1 that jointly perturb the physiological robustness of the TGF β pathway. A limitation of this study is the small sample size, although it remains one of the largest LOA genetic cohorts reported. These results warrant replication in additional cohorts, but we note that this sample was adequate to

identify a protective effect of the *LTBP4* gene, a finding that has subsequently been confirmed in additional cohorts^{8,9}. Previous candidate associations have supported *SPP1*⁵, *LTBP4*^{8,9} and *CD40*⁷ as modifiers of Duchenne muscular dystrophy, but to our knowledge, this is the first study to identify genome-wide significant modifiers. That genome-wide significant noncoding variants in two interacting genes were identified in just 253 patients suggests that larger cohorts will uncover additional modifiers. Identification of such protective modifiers is important, as they provide additional pathways for therapeutic intervention in DMD which will be independent of mutation class and complementary to dystrophin restorative therapies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (NINDS NS043264 and NS085238) to KMF, RBW, and VJV. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from: the GTEx Portal on 07/16/17 using the GTEx Analysis Release V6p (dbGaP Accession phs000424.v6.p1). We thank Dr. Luis B. Barreiro, CHU Sainte-Justine/University of Montreal for sharing sample-level genotype and macrophage *LTBP4* expression data, and John Burian, the Battelle Center for Mathematical Medicine, for help with genotype data cleaning.

REFERENCES

1. Mendell JR, Shilling C, Leslie ND, et al. Evidence-based path to newborn screening for duchenne muscular dystrophy. *Ann. Neurol.* 2012;71(3):304–313. [PubMed: 22451200]
2. Moat SJ, Bradley DM, Salmon R, et al. Newborn bloodspot screening for Duchenne Muscular Dystrophy: 21 years experience in Wales (UK). *Eur. J. Hum. Genet* 2013;21(10):1049–1053. [PubMed: 23340516]
3. Heydemann A, Ceco E, Lim JE, et al. Latent TGF-beta-binding protein 4 modifies muscular dystrophy in mice. *J. Clin. Invest* 2009;119(12):3703–12. [PubMed: 19884661]
4. Vieira NM, Elvers I, Alexander MS, et al. Jagged 1 Rescues the Duchenne Muscular Dystrophy Phenotype. *Cell* 2015;163(5):1204–1213. [PubMed: 26582133]
5. Pegoraro E, Hoffman EP, Piva L, et al. SPP1 genotype is a determinant of disease severity in Duchenne muscular dystrophy. *Neurology* 2011;76(3):219–226. [PubMed: 21178099]
6. Flanigan KM, Ceco E, Lamar KM, et al. LTBP4 genotype predicts age of ambulatory loss in duchenne muscular dystrophy. *Ann. Neurol.* 2013;73(4):481–488. [PubMed: 23440719]
7. Bello L, Punetha J, Gordish-Dressman H, et al. Association Study of Exon Variants in the NF- κ B and TGF β Pathways Identifies CD40 as a Modifier of Duchenne Muscular Dystrophy. *Am. J. Hum. Genet.* 2016;99(5):1163–1171. [PubMed: 27745838]
8. van den Bergen JC, Hiller M, Bohringer S, et al. Validation of genetic modifiers for Duchenne muscular dystrophy: a multicentre study assessing SPP1 and LTBP4 variants. *J. Neurol. Neurosurg. Psychiatry* 2014;1–6. [PubMed: 24337225]
9. Bello L, Kesari A, Gordish-Dressman H, et al. Genetic modifiers of ambulation in the cooperative international Neuromuscular research group Duchenne natural history study. *Ann. Neurol.* 2015;77(4):684–696. [PubMed: 25641372]
10. Lamar K-M, Miller T, Dellefave-Castillo L, McNally EM. Genotype-Specific Interaction of Latent TGFbeta Binding Protein 4 with TGFbeta. *PLoS One* 2016;11(2):e0150358. [PubMed: 26918958]
11. Haslett JN, Sanoudou D, Kho AT, et al. Gene expression comparison of biopsies from Duchenne muscular dystrophy (DMD) and normal skeletal muscle. *Proc. Natl. Acad. Sci. U. S. A.* 2002;99(23):15000–5. [PubMed: 12415109]

12. Porter JD. A chronic inflammatory response dominates the skeletal muscle molecular signature in dystrophin-deficient mdx mice. *Hum. Mol. Genet.* 2002;11(3):263–272. [PubMed: 11823445]
13. Vetrone SA, Montecino-Rodriguez E, Kudryashova E, et al. Osteopontin promotes fibrosis in dystrophic mouse muscle by modulating immune cell subsets and intramuscular TGF- β . *J. Clin. Invest.* 2009;119(6):1583–1594. [PubMed: 19451692]
14. Capote J, Kramerova I, Martinez L, et al. Osteopontin ablation ameliorates muscular dystrophy by shifting macrophages to a proregenerative phenotype. *J. Cell Biol.* 2016;213(2):275–288. [PubMed: 27091452]
15. Quattrocelli M, Capote J, Ohiri JC, et al. Genetic modifiers of muscular dystrophy act on sarcolemmal resealing and recovery from injury. *PLoS Genet.* 2017;13(10):1–22.
16. Flanigan KM, Dunn DM, Von Niederhausern A, et al. Mutational spectrum of DMD mutations in dystrophinopathy patients: Application of modern diagnostic techniques to a large cohort. *Hum. Mutat.* 2009;30(12):1657–1666. [PubMed: 19937601]
17. Flanigan KM, Dunn DM, Von Niederhausern A, et al. Nonsense mutation-associated Becker muscular dystrophy: Interplay between exon definition and splicing regulatory elements within the DMD gene. *Hum. Mutat.* 2011;32(3):299–308. [PubMed: 21972111]
18. Soltanzadeh P, Friez MJ, Dunn D, et al. Clinical and genetic characterization of manifesting carriers of DMD mutations. *Neuromuscul. Disord.* 2010;20(8)
19. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 2014;46(3):310–315. [PubMed: 24487276]
20. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* 2016;48(10):1279–1283. [PubMed: 27548312]
21. Carlson CS, Eberle MA, Rieder MJ, et al. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am. J. Hum. Genet.* 2004;74(1):106–20. [PubMed: 14681826]
22. Lange K, Papp JC, Sinsheimer JS, et al. Mendel: The Swiss army knife of genetic analysis programs. *Bioinformatics* 2013;29(12):1568–1570. [PubMed: 23610370]
23. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet.* 2006;2(12):2074–2093.
24. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 2009;19(9):1655–1664. [PubMed: 19648217]
25. Schaid DJ, Rowland CM, Tines DE, et al. Score Tests for Association between Traits and Haplotypes when Linkage Phase Is Ambiguous. *Am. J. Hum. Genet* 2002;70:425–434. [PubMed: 11791212]
26. Lee S, Emond MJ, Bamshad MJ, et al. Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am. J. Hum. Genet.* 2012;91(2):224–237. [PubMed: 22863193]
27. Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* 2010; 26(18):2336–2337. [PubMed: 20634204]
28. Nédélec Y, Sanz J, Baharian G, et al. Genetic Ancestry and Natural Selection Drive Population Differences in Immune Responses to Pathogens. *Cell* 2016;167(3):657–669.e21. [PubMed: 27768889]
29. Hnisz D, Abraham BJ, Lee TI, et al. Super-enhancers in the control of cell identity and disease. *Cell* 2013;155(4):934–47. [PubMed: 24119843]
30. Javierre BM, Sewitz S, Cairns J, et al. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell* 2016;167(5):1369–1384.e19. [PubMed: 27863249]
31. Schultz-Cherry S, Murphy-Ullrich JE. Thrombospondin causes activation of latent transforming growth factor-beta secreted by endothelial cells by a novel mechanism. *J Cell Biol* 1993;122:923–932. [PubMed: 8349738]
32. Murphy-Ullrich JE, Poczatek M. Activation of latent TGF-beta by thrombospondin-1: mechanisms and physiology. *Cytokine Growth Factor Rev.* 2000;11(1–2):59–69. [PubMed: 10708953]

33. Cohn RD, van Erp C, Habashi JP, et al. Angiotensin II type 1 receptor blockade attenuates TGF-beta-induced failure of muscle regeneration in multiple myopathic states. *Nat Med* 2007;13(2): 204–210. [PubMed: 17237794]
34. Bernstein BE, Birney E, Dunham I, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489(7414):57–74. [PubMed: 22955616]
35. Aguet F, Ardlie KG, Cummings BB, et al. Genetic effects on gene expression across human tissues. *Nature* 2017;550(7675):204–213. [PubMed: 29022597]
36. Ceco E, Bogdanovich S, Gardner B, et al. Targeting latent TGFβ release in muscular dystrophy. *Sci. Transl. Med.* 2014;6(259):259ra144.
37. Quattrocelli M, Spencer MJ, McNally EM. Outside in: The matrix as a modifier of muscular dystrophy. *Biochim. Biophys. Acta - Mol. Cell Res* 2017;1864(3):572–579. [PubMed: 28011285]
38. Crawford SE, Stellmach V, Murphy-Ullrich JE, et al. Thrombospondin-1 is a major activator of TGF-β1 in vivo. *Cell* 1998;93(7):1159–1170. [PubMed: 9657149]
39. Good DJ, Polverini PJ, Rastinejad F, et al. A tumor suppressor-dependent inhibitor of angiogenesis is immunologically and functionally indistinguishable from a fragment of thrombospondin. *Proc. Natl. Acad. Sci. U. S. A* 1990;87(17):6624–8. [PubMed: 1697685]
40. Lawler PR, Lawler J. Molecular basis for the regulation of angiogenesis by thrombospondin-1 and -2. *Cold Spring Harb. Perspect. Med* 2012;2(5):a006627–a006627. [PubMed: 22553494]
41. Zhou Y, Poczatek MH, Berecek KH, Murphy-Ullrich JE. Thrombospondin 1 mediates angiotensin II induction of TGF-beta activation by cardiac and renal cells under both high and low glucose conditions. *Biochem Biophys Res Commun* 2006;339(2):633–641. [PubMed: 16310163]
42. Naito T, Masaki T, Nikolic-Paterson DJ, et al. Angiotensin II induces thrombospondin-1 production in human mesangial cells via p38 MAPK and JNK: A mechanism for activation of latent TGF-β 1. *Am. J. Physiol. - Ren. Physiol.* 2004;286(2):F278–F287.
43. Urao N, Mirza RE, Heydemann A, et al. Thrombospondin-1 levels correlate with macrophage activity and disease progression in dysferlin deficient mice. *Neuromuscul. Disord.* 2016;26(3): 240–251. [PubMed: 26927626]
44. Urao N, Mirza RE, Corbiere TF, et al. Thrombospondin-1 and disease progression in dysferlinopathy [Internet]. *Hum. Mol. Genet* 2017;26(24):4951–4960. [PubMed: 29206970]
45. de Luna N, Gallardo E, Sonnet C, et al. Role of thrombospondin 1 in macrophage inflammation in dysferlin myopathy. *J. Neuropathol. Exp. Neurol* 2010;69(6):643–653. [PubMed: 20467328]
46. Miyazaki D, Nakamura A, Fukushima K, et al. Matrix metalloproteinase-2 ablation in dystrophin-deficient mdx muscles reduces angiogenesis resulting in impaired growth of regenerated muscle fibers. *Hum. Mol. Genet.* 2011;20(9):1787–1799. [PubMed: 21320869]
47. Verma M, Asakura Y, Hirai H, et al. Flt-1 haploinsufficiency ameliorates muscular dystrophy phenotype by developmentally increased vasculature in mdx mice. *Hum. Mol. Genet.* 2010;19(21): 4145–4159. [PubMed: 20705734]
48. Malek MH, Olfert IM. Global deletion of thrombospondin-1 increases cardiac and skeletal muscle capillarity and exercise capacity in mice. *Exp. Physiol.* 2009;94:749–760. [PubMed: 19297388]

Figure 1. Manhattan plots of age at loss of ambulation from the recessive model GWAS. (A) The x axis represents chromosomal location, and the y axis represents $-\log_{10}(P\text{ value})$. The dashed horizontal line denotes a genome-wide significance threshold of $P = 5.0 \times 10^{-8}$ and the dotted line denotes a suggestive threshold of $P = 1.0 \times 10^{-5}$. New (*THBS1*) and known (*LTBP4*) loci exceeding the GWAS threshold are labeled in red and blue, respectively. (B) LocusZoom plot spanning 892 kb of the *THBS1* genomic region from chr15:39000001–39892000 (hg19). The linkage disequilibrium (LD) value (r^2) with rs2725797 (purple diamond) is indicated by the plotted SNP color (minor allele frequency > 0.15). (C) LocusZoom plot spanning 96 kb of the *LTBP4* region (chr19:41056790–41152850, hg19, MAF > 0.15) with rs710160 marked by a purple diamond. The arrows show the location of the four nonsynonymous *LTBP4* SNPs previously associated with LOA.

Figure 2. *LTBP4* haplotypes found in the DMD cohort and genotype-specific expression in macrophages.

(A) The locations of 82 SNPs from the *LTBP4* genomic region (66.5 kb) are plotted from rs1165836 (chr19:41074938) to rs7254268 (chr19:41141434). 506 phased haplotypes from 253 DMD patients are shown in register with the plotted SNPs, with the minor alleles at each SNP position colored by LD bin (16 bins, $r^2 > 0.64$). The relative positions and allele frequencies of rs716160 and the four nonsynonymous *LTBP4* SNPs are shown below the haplotypes. (B) Boxplots of *LTBP4* macrophage RNA-Seq expression levels, grouped by SNP genotype and colored by LD bin. Expression data in macrophages was obtained from the NCI GEO accession number GSE81046²⁸. (C) GTEx boxplots of the rs710160:*LTBP4* eQTL in transformed fibroblasts; circles indicate subject level data.

THBS1 promoter capture HiC interactions in 17 human primary blood cell types

Figure 3. SNPs, chromatin interactions and enhancer features in the 862 kb region flanking the *THBS1* gene.

THBS1 promoter chromatin interactions from promoter capture Hi-C (PChI-C) data are displayed as colored arcs connecting 142 promoter-captured *HindIII* restriction fragments, where the color scale represents the median promoter interaction (CHiCAGO) score in 17 human primary blood cell types (grey arcs have a median score < 5). UCSC genome browser tracks (chr15:39000001–39892000) are shown below the chromatin interaction arcs, and in order are: protein coding (RefSeq) transcripts, RNA-Seq read depth from unaffected pediatric skeletal muscle, stretch enhancer regions from human skeletal muscle myotubes²⁹, and human skeletal muscle myotube ChIP-Seq tracks for H3K27ac histone modification and CTCF binding from ENCODE data. The highlighted region (yellow) contains SNPs associated ($P < 1 \times 10^{-6}$) with LOA. The lower panel expand this 25 kb region

(chr15:39097537–39122787) to show SNP locations, median CHiCAGO scores and chromatin features on 8 consecutive *HindIII* fragments that include rs2725797 and rs2624259. The eight *HindIII* fragments (fragments A to H) are colored by their median promoter interaction score in 17 human primary blood cell types. Chromatin analysis tracks include: chromatin state segmentation from human skeletal muscle myoblasts, transcription factor (CTCF) ChIP-seq, and DNaseI hypersensitivity clusters.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 4. rs2725797 and rs2624259 are *THBS1* expression QTLs in skeletal muscle. (A) Sample-level GTEx boxplots and eQTL statistics for *THBS1* rank normalized gene expression in skeletal muscle (gastrocnemius) tissue. (B) LocusZoom plots of the 122 kb genomic region containing the *THBS1* distal enhancer from rs7163834 (chr15:39020093) to rs8036595 (chr15:39141983) showing SNP associations for skeletal muscle *THBS1* eQTLs (upper panel) and LOA association, recessive model (lower panel).

Figure 5. Two-locus protective genotypic effects of reduced levels/activity variants of *THBS1* and *LTBP4*.

(A) Genotypes are encoded as 0,1, or 2 by protective allele count. Bar height and numbers show the difference between the global average of LOA (10.6 years) and the two-locus genotype cell averages of LOA. Bar width shows the proportion of subjects within each genotype cell. Lighter cell color indicates average LOA below the group, darker cell color indicates average LOA above the group mean, and the white cell has no genotype counts.

(B) Model for joint variant effects on reduced TGF β signaling and increased NO signaling. Homotrimeric TSP-1 activates latent TGF β via direct binding to the LTBP4-TGF β complex, allowing TGF β to bind its receptor and initiate downstream inflammatory and fibrotic pathways. TSP-1 binding also antagonizes pro-angiogenic nitric oxide signaling through binding cell surface receptors CD36 and CD47.

Table 1.

Top SNPs at the genome-wide significant loci associated with age at ambulatory loss

Chr	Position	SNP	Gene (distance, kb)	Alleles	PAF	<i>P</i> (n=253)	<i>P</i> (n=243)	<i>P</i> (n=189)
15	39109233	rs2725797	<i>THBS1</i> (761)	<u>T</u> /C	0.146	6.6E-09	7.5E-07	3.8E-06
15	39110903	rs2624259	<i>THBS1</i> (760)	<u>T</u> /C	0.144	6.6E-09	8.3E-07	3.8E-06
15	39105587	rs1004866	<i>THBS1</i> (765)	<u>G</u> /A	0.148	5.7E-07	2.6E-04	1.5E-04
19	41087491	rs710160	<i>LTBP4</i> (12)	<u>C</u> /T	0.257	4.7E-08	5.7E-07	2.2E-06
19	41084693	rs814518	<i>LTBP4</i> (15)	<u>A</u> /G	0.296	2.9E-06	2.2E-05	2.0E-05

Associations and genes at $P < 5 \times 10^{-8}$ are highlighted in bold. Intergenic distance is relative to nearest candidate gene. Protective alleles are underlined and allele frequency (PAF) is from the cohort of 253 DMD patients. The distance (kb) of the SNP to the promoter of the candidate gene is shown in parentheses.

Table 2.Haplotype association with age at ambulatory loss and *LTBP4* expression

Haplotype SNPs (LTBP4 amino acid change)								
Haplotype	rs710160	rs2303729	rs1131620	rs1051303	rs10880	Hap Freq.	Score	Haplotype <i>p</i> value
Phenotype ^a : Age at ambulatory loss								
T-VTTT	T	G (V)	A (T)	A (T)	C (T)	0.49	−1.09	0.27
C-IAAM	C	A (I)	G (A)	G (A)	T (M)	0.17	4.26	2.0E-05
T-IAAM	T	A (I)	G (A)	G (A)	T (M)	0.13	0.28	0.78
Phenotype ^b : <i>LTBP4</i> mRNA level in macrophages								
T-VTTT	T	G (V)	A (T)	A (T)	C (T)	0.56	3.64	0.0003
C-IAAM	C	A (I)	G (A)	G (A)	T (M)	0.20	−2.88	0.0040
T-IAAM	T	A (I)	G (A)	G (A)	T (M)	0.06	−0.88	0.38

^a n = 253, DMD cohort (this study); global $P = 0.00032$, for association between haplotypes and age at ambulatory loss as a quantitative trait using a recessive model, with steroid use as a covariate.

^b n = 89, European-American healthy blood donors from Nédélec *et al.*,²⁵; global $P = 0.0047$, for association between haplotypes and macrophage *LTBP4* expression level (TPM from RNA-seq) as a quantitative trait using an additive model.