http://hdl.handle.net/1765/129060

Erasmus MC
Universitair Medisch Centrum Rotterdam

Investigation of shared genetics between bone mineral density and lean mass in elderly

Katerina Trajanoska, Niki L Dimou, David Karasik, Fernando Rivadeneira on behalf of GEFOS consortium

In preparation

Erasmus University Rotterdam

ABSTRACT

Background: The intimate relationship between bone and muscle is governed by various mechanical and biochemical interactions; regulated by a complex network of pleiotropic genes within common biological pathways. We aimed to assess pleiotropy between bone and muscle by performing a bivariate genome-wide association study (GWAS).

Methods: We included 15 cohorts from populations across Europe, the US and Australia (N_{total}= 30,531). Total body BMD (TB-BMD; g/cm^2) and lean mass (TB-LM; kg) were measured in all cohorts using DXA. Genetic variants were imputed to the Haplotype Reference Consortium (HRC) reference panel. Each cohort performed univariate GWAS adjusting for age, gender, height, fat percentage and principal components using 8,027,751 SNPs. After GWAS meta-analysis genome-wide significance (GWS) was set at $P<5\times10^{-8}$.

Results: The univariate TB-BMD GWAS identified 18 signals all mapping to known BMD loci. Only four loci were associated with TB-LM, three of which were novel associations mapping to ZC3H11B (rs1810144-G, beta=0.06SD) *SPEG* (rs1810144-C, beta=0.05SD) and *MYPN* (rs12415105-G, beta=0.04SD). Both *SPEG* and *MYPN* have evidence from animal models for implication in muscle biology. Notably, the bivariate meta-analysis yielded 17 pleiotropic signals from 15 different loci mapping in or near *MEF2C,* XYLB/ACVR2B, *PPP6R3/LRP5, LRP4* and *SLC27A6* among others, linked with both bone and muscle biology.

Conclusion: Our findings provide basis for future functional assessments aimed at unravelling biological mechanisms underlying complex bone-muscle interactions; with the potential of pin-pointing strategies for the joint prevention and intervention of osteoporosis and sarcopenia.

INTRODUCTION

The intimate relationship between bone and muscle is governed by various mechanical and biochemical interactions. To begin with, bone and muscle precursors originate from the mesenchymal stem cells, go through well-orchestrated, multifaceted mechanisms of cell proliferation and differentiation, and eventually give rise to bone or muscle tissue. The bone tissue begins to experience mechanical stimulations during embryogenesis with the start of the first muscle contractions; contributing to fetal bone development and growth. The mechanical stimulation and loading continues after birth and is pivotal for reaching peak bone mass in early adulthood and maintaining bone mass later in life.[1] At the same time, bone and muscle tissues release variety of endocrine, paracrine, and autocrine factors which influence the surrounding tissues or each other.[2] All these mechanisms are regulated by complex network of genes that can affect either one or both tissues. Therefore, bone and muscle may share common biological pathways i.e. pleiotropy; a genetic variant that influences multiple traits.

Multitude genetic factors have been associated with the variation of bone and lean mass in the general population. Several of these loci such as *MC4R, FTO a*nd *MGMT* harbor genetic variants associated with both bone and muscle;[3,4,5] indicating possible pleiotropy. However, simple cross-phenotype association is not an evidence of pleiotropy. Although pleiotropy is a challenging to be determine several approaches have been developed which can be generally classified as multivariate or univariate. [6]The former approach requires all individuals included in the study to have phenotype information of all traits of interest. On the other hand, the univariate approach combines summary statistics of single-trait GWAS. Combining summary statistics several genes have been proposed to have pleotropic effect on both bone and muscle traits such as *METTL21C* and *GLYAT.*[7,8] However, these results are yet to be replicated in large-scale efforts and validated by functional studies. Recently, a large pediatric effort presented a novel gene, *SREBF1,* exerting opposite effects on total body bone and muscle mass.[9] We have now extended this effort by including individuals across different age categories in order to identify novel genes affecting bone and muscle using much more powerful study settings.

SUBJECTS AND METHODS

Study Populations

This effort included 15 cohorts from populations across Europe, the USA and Australia from the GEFOS consortium. The majority of the cohorts were of European ancestry with only two cohorts with admixed background (GENR and RAINE). In addition, two

out of the 15 cohorts were pediatric cohorts (GENR and ALSPAC). Each of the studies was approved by their local Medical Ethical Committee and only participants with written informed consent were included.

Phenotype Assessment

Total body BMD (g/cm^2) and total body lean mass (kg) were assessed in all cohorts using dual-energy X-ray absorptiometry (DXA). The pediatric cohorts used total body less head as recommended by the International society for clinical densitometry.

Genotyping and Imputation

All cohorts used commercially available genome-wide arrays to genotyped their participants. All cohorts did pre-imputation quality control using already established protocols. SNPs were imputed to the Haplotype Reference Consortium (HRC) reference panel[10] (build 37). The imputation was done using free imputation servers (Michigan or Sanger) and resulted in approximately 40,000,000 SNPs.

Univariate Genome-wide Association Analysis

Each study preformed linear regression models for both BMD and lean mass adjusted for age, gender, height, fat percentage and principal components. For both traits residuals were computed separately by sex while in the family studies sex was included as a covariate in the model. Prior meta-analysis quality controls (QC) was performed with easy QC.[11] Genetic variants with low imputation quality (r^2<0.3) or low minor allele frequency (MAF<0.01) were excluded. Each cohort was tested for possible spurious inflation by computing lambda and visual inspection of the QQ plots. Finally, all cohorts were combined using inverse variance weighted fixed meta-analysis in Metal.[12] Genome wide significant threshold was set to $5x10^{-8}$. Independent SNPs were defined using clumping methods in Plink (parameters of --clump-p1 5e-8 --clump-kb 500 --clump-r2 0.2).

Multi-trait Analysis

Next, we preformed multi-trait analysis of genome-wide association summary statistics using MTAG in order to boost statistical power to detect additional genetic loci from the univariate GWAS for each trait analyzed.[13] The advantages of this method are that 1) statistics don't need to come from independent efforts and overlap does not influence the results as it is based on the LD score repression which accounts for sample overlap, 2) it can included more than two traits and 3) it is computationally very quick. Finally, if a SNP is null for one trait but not for the other can lead to false positives however MTAG also account for this credibility.

Bivariate Analysis

Finally we used the sum rank method, a novel approach, to test for possible pleiotropic variants influencing both BMD and lean mass.[14] The sum ranking method is a simple statistical test that is based on the GWAS summary p-values which covert to an individual rank of every variant for a given trait and computes a final pleiotropy p-values and is robust to extreme values.

Characterization and Annotation of the Genomic Loci

First, for each of the lead SNP we evaluated the DNA features and regulatory elements in noncoding regions of the human genome using RegulomeDB.[15] Next we used the CAAD tool to score the deleteriousness of the lead SNPs.[16,17] The higher the score the more deleterious the variant is. We then searched the closest genes in the Mouse Genome Informatics (MGI) database to explore presence of musculoskeletal phenotypes in mice mutants (http://www.informatics.jax.org).[18] Finally, we performed eQTL analysis using the online GTEx database based on 54 non-diseased tissue sites across nearly 1000 individuals (https://gtexportal.org/home/).[19]

RESULTS

Univariate and Multi-trait Analyses

In the univariate GWAS meta-analysis of TB-BMD 2,081 SNPs surpassed the genome-wide significant threshold (**Figure 1, upper half; Supplementary Table 1**) which mapped to 18 different genomic loci. All discovered loci are known BMD loci and the top genetic variants mapped in or near *WNT4* (rs10493013-C, beta=0.09SD), *WLS* (rs2566752-C, beta=0.05SD), *EN1* (rs141795717-A, beta=0.21 SD), *CSRNP3/GALNT3* (rs6719426-G, beta=0.05SD), *DGKD* (rs7584554-G,beta=0.05SD), *CTNNB1* (rs7609599-G, beta=0.05SD), *RSPO3* (rs4580892-T, beta=0.05SD), *C7orf76/SHFM1* (rs7787512-T, beta=0.06SD), *CPED1/WNT16* (rs2908004-A, beta=0.10SD), *TNFRSF11B* (rs7014574-C, beta=0.05SD), *MBL2* (rs7902708-G, beta=0.07SD), *CCDC34* (rs60212556-C, beta=0.08SD), *C11orf49* (rs79232684-A, beta=0.09SD), *PPP6R3* (rs4752957-A, beta=0.07SD), *TMEM135* (rs618926-C, beta=0.05SD), *SP7* (rs10747668-C, beta=0.06SD), *FABP3P2* (rs9533094-A, beta=0.05SD) and *MEOX1/DUSP3* (rs66838809-A, beta=0.09SD). On the other hand, only four loci harbored GWS variants for TB-LM which mapped near or to *ZC3H11B* (rs1810144-G, beta=0.06SD), *SPEG* (rs1810144-C, beta=0.05SD) MYPN (rs12415105-G, beta=0.04SD) and *FTO* (rs56094641-G, beta=0.05SD) (**Figure 2, upper half; Supplementary Table 1**). Both meta-analyses did not show any inflation in the test statistics (**Supplementary Figure 1A and 1B**). In the multi-trait analysis, the genetic correlation between TB-BMD and

TB-LM was 0.40. Although moderate, no additional SNPs reached GWS for TB-BMD (**Figure 1, lower half**) while for TB-LM only one additional variant surpassed the GWAS threshold (**Figure 2, lower half**) which mapped to *PPP6R3*; known TB-BMD loci that have not been reported with TB-LM before.

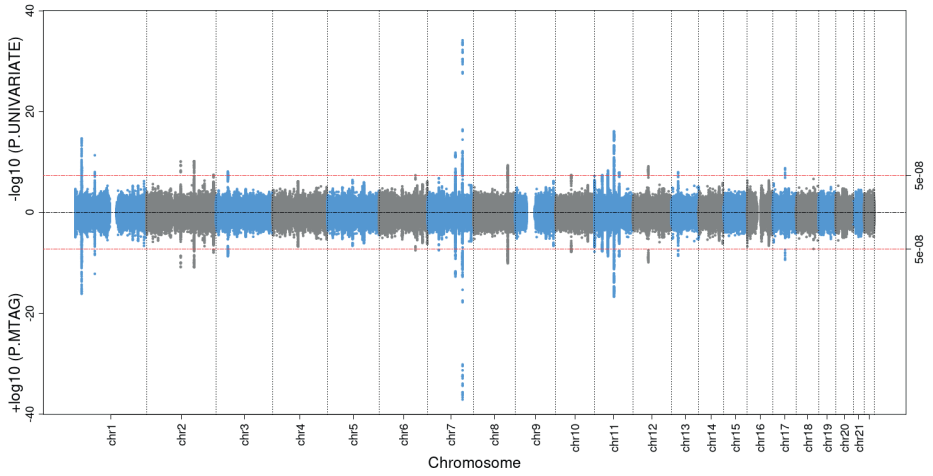


Figure 1 | Miami plot displaying the univariate –log10 (P values) (top half) and multi-trait GWAS (MTAG) –log10 (P values) (bottom half) for the SNP associations with TB-BMD. Dashed red lines mark the GWS threshold ($P < 5 \times 10^{-08}$).

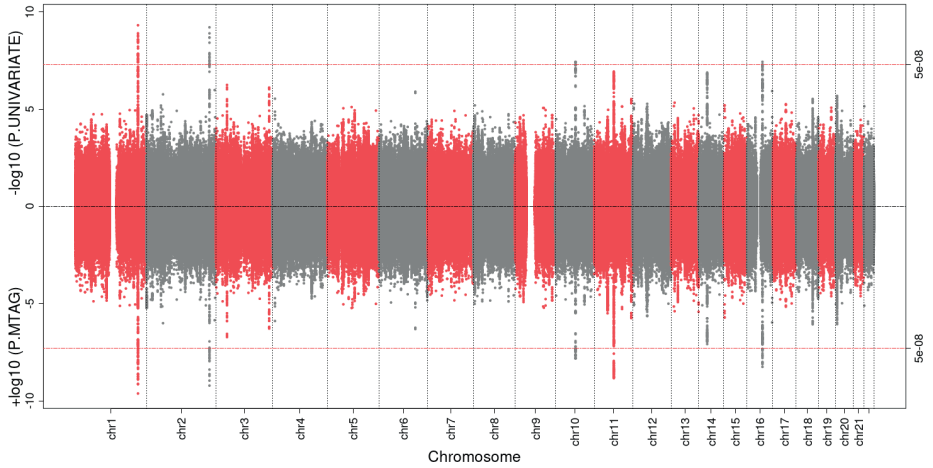


Figure 2 | Miami plot displaying the univariate –log10 (P values) (top half) and multi-trait GWAS (MTAG) –log10 (P values) (bottom half) for the SNP associations with TB-LM. Dashed red lines mark the GWS threshold ($P < 5 \times 10^{-08}$).

Table 1 | Lead SNPs for the 17 bivariate association signals with TB-BMD and TB-LM.

						Total Body BMD			Total Body LM			Bivariate	Closest
Locus	**SNP**	**BP**	**EA**	**NEA**	**EAF**	**Effect**	**se**	**P**	**Effect**	**SE**	**P**	**P**	**Gene**
1p13.3	rs4970731	109995459	T	C	0.30	0.04	0.01	$4.74x10^{-5}$	0.04	0.01	$6.23x10^{-5}$	$6.01x10^{-9}$	*SYPL2/PSMA5*
1q41	rs10779360	219731273	T	C	0.61	0.03	0.01	$2.56x10^{-4}$	0.04	0.01	$2.18x10^{-7}$	$3.27x10^{-8}$	*ZC3H11B*
3p22.2	rs1381798	38437471	A	G	0.29	0.03	0.01	$1.21x10^{-4}$	0.04	0.01	$1.32x10^{-5}$	$8.96x10^{-9}$	*XYLB*
5q14.3	rs254777	88008871	C	G	0.74	0.05	0.01	$9.15x10^{-7}$	0.04	0.01	$1.59x10^{-5}$	$1.41x10^{-10}$	*MEF2C*
5q23.3	rs7706547	128000893	T	C	0.16	-0.05	0.01	$2.97x10^{-6}$	-0.04	0.01	$1.71x10^{-4}$	$1.52x10^{-8}$	*SLC27A6*
6q22.33	rs9491689	127398595	A	C	0.28	0.05	0.01	$1.41x10^{-7}$	0.04	0.01	$5.55x10^{-5}$	$1.55x10^{-9}$	*RSPO3*
6q27	rs2609312	168816483	A	C	0.59	-0.04	0.01	$4.10x10^{-6}$	-0.03	0.01	$8.39x10^{-5}$	$3.87x10^{-9}$	*SMOC2*
11q13.2	rs7113287	68372186	A	T	0.73	0.07	0.01	$1.17x10^{-15}$	0.05	0.01	$1.10x10^{-7}$	$6.00x10^{-15}$	*PPP6R3*
11p11.2	rs61882720	46538988	A	C	0.17	0.06	0.01	$3.92x10^{-7}$	0.05	0.01	$2.48x10^{-5}$	$3.17x10^{-10}$	*AMBRA1*
11p11.2	rs12806687	46924665	C	G	0.35	0.05	0.01	$1.26x10^{-7}$	0.03	0.01	$7.96x10^{-5}$	$3.17x10^{-9}$	*LRP4*
11p11.2	rs1976143	47225534	A	G	0.67	-0.04	0.01	$3.09x10^{-7}$	-0.03	0.01	$1.23x10^{-4}$	$7.64x10^{-9}$	*PACSIN3/DDB2*
12q24.31	rs10846578	124400278	T	C	0.28	-0.04	0.01	$6.55x10^{-5}$	-0.04	0.01	$4.83x10^{-5}$	$6.47x10^{-9}$	*DNAH10*
12q13.11	rs11168357	48412138	A	G	0.23	-0.04	0.01	$1.61x10^{-4}$	-0.04	0.01	$5.50x10^{-5}$	$2.34x10^{-8}$	*COL2A1*
13q34	rs2701321	115058005	A	G	0.09	0.06	0.01	$1.05x10^{-5}$	0.06	0.01	$2.05x10^{-5}$	$4.79x10^{-10}$	*UPF3A*
13q14.11	rs9590675	42749192	T	G	0.40	-0.03	0.01	$2.15x10^{-4}$	-0.03	0.01	$4.40x10^{-5}$	$3.35x10^{-8}$	*DGKH*
16q12.2	rs55872725	53809123	T	C	0.39	0.03	0.01	$2.96x10^{-5}$	0.05	0.01	$5.76x10^{-8}$	$4.41x10^{-10}$	*FTO*
16p13.13	rs56092102	10671035	T	G	0.09	-0.07	0.01	$7.56x10^{-6}$	-0.05	0.01	$2.39x10^{-4}$	$3.03x10^{-8}$	*EMP2*

EA=effect allele; NEA= non-effect allele; Beta coefficients and allele frequencies (EAF) are reported for the EA.

Table 2 | Functional annotation of the leads SNPs

SNP	Closest Gene	GTEx tissue	GTEx Gene	RegulomeDB	CAAD	Mouse Models Tissue	Other phenotypes*
rs4970731	*SYPL2/PSMA5*	Muscle-Skeletal, Testis, Lungs	*SYPL2, MYBPHL*	-	3.30	Muscle, Growth	glomerular filtration rate, BMI, intelligence
rs10779360	*ZC3H11B*	-	-	-	0.17	Aging	BMI., adiposity, smoking behavior
rs1381798	*XYLB*	Thyroid, Cerebellum, Pancreas	*EXOG*	5	8.73	Skeleton, Growth	glomerular filtration rate
rs254777	*MEF2C*	-	-	-	1.82	Skeleton, Muscle, Growth, Aging	platelet count, erythrocyte count, height
rs7706547	*SLC27A6*	Adrenal Gland, Hear, Skin	*FBN2*	6	1.30	-	red blood cell distribution width, atrial fibrillation
rs9491689	*RSPO3*	-	-	-	1.92	Muscle, Aging	BMI, smoking behavior
rs2609312	*SMOC2*	Muscle-Skeletal, Thyroid	*SMOC2*	5	2.25	Skeleton	Height, pulse pressure, Vitiligo
rs7113287	*PPP6R3*	Whole Blood, Tibial Nerve	*GAL*	-	2.04	-	educational attainment, red cell distribution width
rs61882720	*AMBRA1*	Skin. Esophagus -Muscularis	*ATG13*	-	1.03	Aging	schizophrenia
rs12806687	*LRP4*	Muscle-Skeletal, Pancreas, Tibial Nerve	*LRP4, ARHGAP1*	4	4.40	Skeleton, Growth	Thrombosis, systolic blood pressure, risk-taking behavior
rs1976143	*PACSIN3/DDB2*	Tibial Nerve, Pancreas, Skin	*LRP4, DDB2*	5	0.11	Growth, Aging	Educational attainment, HDL cholesterol
rs10846578	*DNAH10*	-	-	5	2.55	Aging	BMI, grip strength, height
rs11168357	*COL2A1*	Thyroid. Skin, Lungs	*RP1-228P16.1*	4	0.38	Skeleton, Growth	Height
rs2701321	*UPF3A*	Muscle-Skeletal	*CDC16*	-	1.15	-	CVD, Pulse pressure, Hodgkin lymphoma
rs9590675	*DGKH*	-	-	4	14.84	-	glomerular filtration rate
rs55872725	*FTO*	Muscle-Skeletal	*FTO*	4	1.57	Skeleton	BMI
rs56092102	*EMP2*	-	-	5	0.30	-	facial asymmetry measurement, FEV/FEC ratio, erythrocyte count

*this column indicates if the gene has been related with other phenotypes.

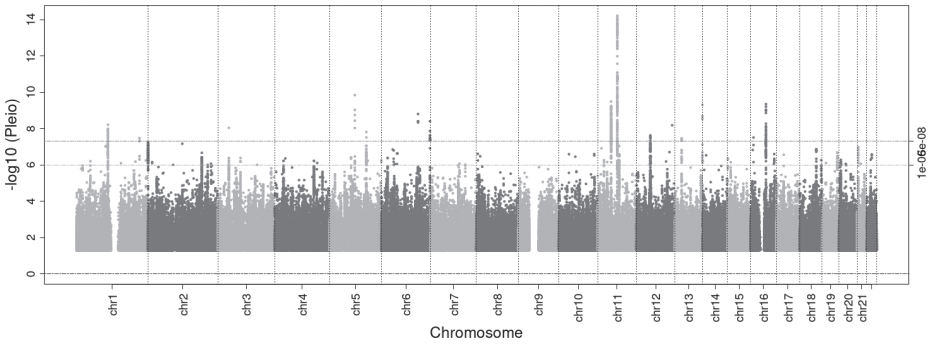


Figure 3 | Manhattan plot displaying the bivariate –log10 (P values) for the SNP associations with TB-BMD and TB-LM. Dashed red and yellow lines mark the GWS threshold ($P < 5\times10^{-08}$) and suggestive threshold ($P < 1\times10^{-6}$), respectively.

Bivariate Analysis

The bivariate meta-analysis yielded 17 GWAS hits mapping to 15 different loci (**Table 1; Supplementary Figure 2a-p**) Eleven variants were intronic and mapped to *ZC3H11B* (rs10779360-T), *XYLB* (rs1381798-A), *SLC27A6* (rs7706547-C), *PPP6R3* (rs7113287-A), *AMBRA1* (rs61882720-A), *LRP4* (rs12806687-C), *DNAH10* (rs10846578-C), *UPF3A* (rs2701321-A), *DGKN* (rs9590675-G), FTO (rs55872725-T) and *EMP2* (rs56092102-G). The other six were intergenic mapping near *SYPL2/PSMA5* (rs4970731-T, kb=13.64), *MEF2C* (rs254777-C, kb=5.10), *RSPO3* (rs9491689-A, kb=41.15), *SMOC2* (rs2609312-C, kb=25.35), PACSIN3 (rs1976143-G, kb=17.52) and *COL2A1* (rs11168357-G, kb=13.85). The top most significant locus (*PPP6R3*) displayed stronger association with TB-BMD and was only suggestive for TB-LM in the univariate GWAS analyses. All variants exhibited the same direction of effect on both lean mass and BMD.

Characterization and Annotation of the Genomic Loci

All of the lead variants had high RegulomeDB score (>=4) indicating that the SNPs are less likely to be functional, thus, do not affect gene biding and expression. Similarly, all lead SNPs had low CAAD score; demonstrating they are not deleterious. In the MGI database only the Mef2c knock-out (KO) mice presented with both muscle and bone abnormalities such as dilated cardiomyopathy, abnormal skeletal muscle fiber type ratio, decreased bone length, abnormal trabecular bone morphology abnormal endochondral bone ossification among others. In addition, only the Sypl2 and Rspo3 KO mice presented only with muscle abnormalities, whereas, Xylb, Smoc2, Lrp4, Col2a1, and Fto KO mice had only skeletal deformities. Finally, five of the lead SNPs (rs4970731, rs2609312, rs12806687, rs2701321, rs55872725) were cis-eQTL variants that can regulate gene expression (*SYPL2, SMOC2, LRP4, CDC16,* and *FTO*

respectively) in skeletal muscle tissue. However, the variants also showed significant eQTLs in other tissues such as lungs, testis, thyroid and skin.

DISCUSSION

We performed a large scale bivariate GWAS between TB-BMD and TB-LM including up to 30,000 individuals and discovered 15 loci harboring variants associated with both TB-BMD and TB-LM. The top variant at the 5q14.3 locus mapped near the Myocyte enhancer factor–2C (*MEF2C*) gene that have been previously linked to various aspects of bone and muscle biology. On top of this we also identified other genes with potential pleiotropic effects such as *XYLB, PPP6R3, COL2A1* and *FTO*. Finally, our univariate TB-LM GWAS yielded three novel lean mass associated variants mapping to 1q41 (*ZCH311B*), 2q35 (*SPEG*) and 10q21.3 (*MYPN*).

The 5q14.3 locus harbored the strongest pleiotropic signal (rs254777-C) which maps near ***MEF2C*** (5.10kb). This gene belongs to the MEF2 transcription factor family which encodes four muscle-specific transcription factors: Mef2a, -b, -c, and -d, that regulate different aspects of skeletal, cardiac and smooth muscle cell growth and differentiation.[20,21] In mice, Mef2c is the first expressed gene during embryogenesis from the mef2[22] and it is crucial for myocardiogenesis and morphogenesis[23]. The mef2 factors are expressed in all developing muscle cell types[22] and are important for muscle fiber differentiation, elongation and fusion;[24] thus, may influence muscle mass variation in the general population. The loss of Mef2c and Mef2d proteins causes a defect in myofibrillogenesis and sarcomere formation.[25] Notably, in matured muscles the mef2 factors have been associated with improved skeletal muscle regeneration after muscle injury;[26] fortifying the important role of the MEF2 family in muscle biology. *MEF2C* is also a known BMD locus[27,28] which was suggestive GWAS signal in our univariate TB-BMD GWAS. In mice, it has been shown to promote chondrocyte hypertrophy and bone development.[29] Furthermore, *MEF2C* also controls the cortical expression of Sfrp2 and Sfrp3 (WNT inhibitors) and it is vital for adequate Sost expression;[30] important regulator of bone formation. Nevertheless, the increased bone mass in Mef2c mutants is not directly related to the reduced Sost expression.[30] Overall, mutations in the *MEFC2* gene lead to both impaired muscle and bone structure. Elucidating the underlying pathways may be crucial for preventing and treating osteoporosis and sarcopenia simultaneously.

Another interesting pleiotropic region is the 11q13.2 locus harboring GWAS variants mapping to the ***PPP6R3*** gene which are also in high LD (r2>0.8) with variants from the *LRP5* gene region. The *PPP6R3/LRP5* region is a known BMD locus and have been associated with both bone and lean mass in children.[28] *LRP5* is an important WNT

co-receptor and it is involved in bone formation and mechanotransduction.[31] Mutations in the LRP5 gene case *autosomal recessive osteoporosis-pseudoglioma syndrome* characterized by infancy-onset vision loss and severe juvenile osteoporosis[32] but also muscular hypotonia and ligamentous laxity.[33] However, the exact effects of LPR5 on muscle mass remain unknown. Furthermore, on chromosome 11 there is additional locus (11p11.2) that harbor four possible pleiotropic genes such as ***LPR4, AMBRA1, PACSIN3*** and ***DDB2***. The ***LRP4*** gene is an important mediator of the SOST-dependent inhibition of bone formation.[34] In addition, *LRP4* is essential for neuromuscular junction (NMJ) formation and maintenance.[35] Mutations in this gene have been associated with i) *Cerani-Lenz syndrome* characterized by syndactyly and oligodactyly of the fingers and toes with malformations of the forearm bones and lower limbs[36] and ii) *Myasthenic syndrome* characterized by failure in neuromuscular transmission resulting in increased fatigue and generalized muscle weakness.[37] Next, ***AMBRA1*** is a known BMD gene and the top SNP is in LD (r^2>0.8) with variants mapping to the surrounding genes also related to BMD and variety of neurological traits. The third lead SNP in this locus was intergenic mapping between the ***PACSIN3*** and ***DDB2***. Both genes have been previously related to total body and heel BMD.[28,38] The exact mechanisms of action of these genes on bone and muscle are yet to be determine.

The 1p13.3 GWAS signals are intergenic and map between the ***SYPL2*** and ***PSMA5*** gene regions that have been previously associated with glomerular filtrationrate,[39] human intelligence[40] and BMI[41]. The lead SNP from this locus is in strong LD (r^2>0.8) with variants mapping to the *SORT1* gene that has been recently linked with variation in appendicular lean mass[42] and variants mapping to the *MYBPHL* gene which have been associated with increased risk for cardiomyopathy and arrhythmia in humans.[43] This gene is mainly expressed in heart ventricles,[43,44] whereas, its paralog *MYBPH* is highly expressed in skeletal muscles only[43,44]. Several GWAS variants also mapped to the ***XYLB*** gene. The lead variant from this locus is in moderate LD with variants mapping to the *ACVr2B* gene. This gene encodes the activin type 2 receptor protein and it is part of the TGF-beta superfamily of signaling proteins. Administration of soluble form of ACVR2B in wild-type mice leads to substantial increase in muscle mass by up to 60%[45] as well as increase in bone and muscle mass in a mouse model of osteogenesis imperfect[46] and in dystrophic mice.[47] In addition, mice lacking ACVR2B and ACVR2A demonstrated sustained increases in trabecular bone volume.[48] Next, the top signal in 5q23.3 is located in the vicinity of the ***SLC27A6*** gene which encodes several members of the fatty acid transport protein family. Importantly, this gene is part of the PPAR signally pathway. Par is an important regulator of bone turnover in both mice and humans.[49] On the other hand, PPARα is expressed in skeletal muscles in both humans and mice and its activation in rare cases (<1%) may lead to muscle weakness, muscle pain and even breakdown of muscle.[50] Furthermore, ***RSPO3*** may

also have pleiotropic effects on both bone and muscle. This gene is a well-known BMD and fracture risk; however, it has not been associated with muscle mass before. What we know so far is that the Rspo3 protein promotes angioblast, vascular and cardiac development.[51,52] Similar, the top variants from the 6q27 locus mapped to the ***SMOC2*** gene which have been also implicated in angiogenesis. Mutations is this gene have been related to Dentin Dysplasia.[53] ***DNAH10*** have been previously associated with appendicular lean mass[42] and BMI.[54] There is no evidence for association with BMD or other bone phenotypes. On the other hand, the ***COL2A1*** gene located on 12q13.11 is a well-established bone region. In addition, the lead SNP from this region is in a moderate LD (r^2>0.6) with SNPs associated with appendicular lean mass.[42] *COL1A1* is relevant for the production of type II collagen which is found primary in the bone cartilage but also in the muscle tissue among others. Mutations in this gene are characterized by variety of skeletal phenotypes such as achondrogenesis type II and hypochondrogenesis.[55] Two additional pleiotropic signals mapped on two distinct loci on chromosome 13. The signals on 13q14.11 maps to the ***DGKH*** gene which has been associated with glomerular filtration rate[56] and height.[57] Now we also show evidence for its possible implication in bone and muscle mass variation. Notably, its paralog *DGKD* have been associated with bone density and calcium levels among other traits.[28,58] The second locus 13q34 contained the ***UPF3A*** gene which is involved in both mRNA nuclear export and mRNA surveillance. Its effects on bone and muscle tissue are still unknown. Last but not least chromosome 16 also harbored two loci with possible pleiotropic effects. On 16q12.2 the GWS variants mapped to ***FTO*** a well-known obesity gene which have been also related with both BMD and muscle mass in previous GWAS efforts;[3,38] indicating that *FTO* may be important link between bone, muscle and fat tissue. Finally, the signals on the second locus (16p13.13) mapped near ***EMP2*** a known heel BMD gene[38] that have not been related to any muscle trait till date. Notably, this gene is a positive relator of the vascular endothelial growth factor A (VEGF). Interestingly, we discovered several genes related to vascular development and angiogenesis. Given bone and muscle tissue are highly vascularized, these genes may play important role in osteo- and or myogenesis.

Our univariate analysis for TB-BMD yielded 18 loci that have been previously reported by large GWAS efforts.[27,28,38] Notably, we identified three novel loci associated with TB-LM which have not been previously reported. The most significant variant on locus 2q35 mapped to ***SPEG*** (Striated muscle preferentially expressed protein kinase) which encodes a protein similar to the members of the myosin light chain kinase family and it is required for myocyte cytoskeletal development.[59] Mutations in this gene have been associated with centronuclear myopathy which is characterized by muscle weakness and wasting[60] and often occur due to interaction between *SPEG* and myotubularin (*MTM1*). In addition, Speg-KO mice were weaker and smaller compared

to their wild-mice controls.[61] Next, we also identified the *MYPN* (Myopalladin) gene as novel muscle-related gene which have been previously associated with cardiomyopathy and nemaline myopathy.[62] *MYPN* is relevant for the interaction between the sarcomere and nucleus in cardiac and skeletal muscles.[63,64] In the MGI database Mypn mice-KOs are characterized by abnormal myocardial fiber morphology, abnormal cardiac muscle relaxation, abnormal intercalated disk morphology, abnormal Z line morphology, and myopathy.

Our study is the largest bivariate GWAS meta-analysis that have been done so far on TB-BMD and TB-LM providing adequate power to detect loci with pleiotropic effects. In addition, we used a novel rank-based approach to detect these pleiotropic variants and our findings should not be biased by strong association with one of the phenotypes; producing false positive results. Finally, both TB-BMD and TB-LM were measured using the same technique (DXA) and each study used the same scanner. We cannot rule out possible false positive findings. Replication and functional studies need to confirm and further characterize our findings. Currently we are working on embarking meta-analysis using individual level data in order to prove more robust evidence of pleiotropy.

In summary, through bivariate genome-wide association study we were able to provide strong evidence for genetic overlap between TB-BMD and TB-LM in the general population. Our findings provide basis for future functional studies aimed to unravel the biological mechanisms behind the complex bone-muscle interaction; bring us one step closer to the joint prevention and intervention of osteoporosis and sarcopenia.

REFERENCES

1. Bergmann, P. *et al.* Loading and Skeletal Development and Maintenance. *J. Osteoporos.* **2011**, 1–15 (2011).
2. Karsenty, G. & Olson, E. N. Leading Edge Review Bone and Muscle Endocrine Functions: Unexpected Paradigms of Inter-organ Communication. (2016). doi:10.1016/j.cell.2016.02.043
3. Zillikens, M. C. *et al.* Large meta-analysis of genome-wide association studies identifies five loci for lean body mass. *Nat. Commun.* **8**, 80 (2017).
4. Willems, S. M. *et al.* Large-scale GWAS identifies multiple loci for hand grip strength providing biological insights into muscular fitness. **8**, 16015 (2017).
5. Kemp, J. P. *et al.* Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nat. Genet.* **49**, 1468–1475 (2017).
6. Hackinger, S. & Zeggini, E. Statistical methods to detect pleiotropy in human complex traits. *Open Biol.* **7**, 170125 (2017).
7. Huang, J. *et al. METTL21C* Is a Potential Pleiotropic Gene for Osteoporosis and Sarcopenia Acting Through the Modulation of the NF-κB Signaling Pathway. *J. Bone Miner. Res.* **29**, 1531–1540 (2014).
8. Guo, Y.-F. *et al.* Suggestion of GLYAT gene underlying variation of bone size and body lean mass as revealed by a bivariate genome-wide association study. *Hum. Genet.* **132**, 189–199 (2013).
9. Medina-Gomez, C. *et al.* Bivariate genome-wide association meta-analysis of pediatric musculoskeletal traits reveals pleiotropic effects at the SREBF1/TOM1L2 locus. *Nat. Commun.* **8**, 121 (2017).
10. Consortium, the H. R. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
11. Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–212 (2014).
12. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
13. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229–237 (2018).
14. RePub, Erasmus University Repository: Imaging genetics : Methodological approaches to overcoming high dimensional barriers. Available at: https://repub.eur.nl/pub/106452. (Accessed: 2nd January 2020)
15. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22**, 1790–1797 (2012).
16. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
17. Rentzsch P, Witten D, Cooper GM, Shendure J, K. M. CADD: predicting the deleteriousness of variants throughout the human genome.
18. Bult, C. J. *et al.* Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* **47**, D801–D806 (2019).
19. Carithers, L. J. *et al.* A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv. Biobank.* **13**, 311–317 (2015).

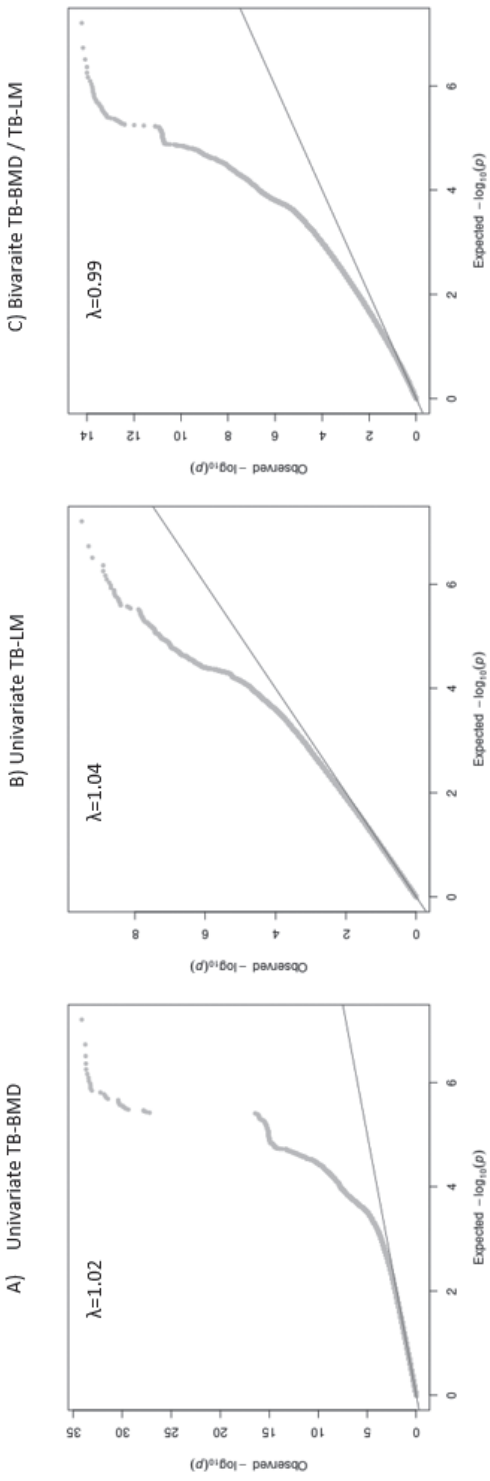
20. Olson, E. N., Perry, M. & Schulz, R. A. Regulation of Muscle Differentiation by the MEF2 Family of MADS Box Transcription Factors. *Dev. Biol.* **172**, 2–14 (1995).
21. Black, B. L. & Olson, E. N. TRANSCRIPTIONAL CONTROL OF MUSCLE DEVELOPMENT BY MYOCYTE ENHANCER FACTOR-2 (MEF2) PROTEINS. *Annu. Rev. Cell Dev. Biol.* **14**, 167–196 (1998).
22. Edmondson, D. G., Lyons, G. E., Martin, J. F. & Olson, E. N. Mef2 gene expression marks the cardiac and skeletal muscle lineages during mouse embryogenesis. *Development* **120**, 1251–63 (1994).
23. Lin, Q., Schwarz, J., Bucana, C. & Olson, E. N. Control of Mouse Cardiac Morphogenesis and Myogenesis by Transcription Factor MEF2C. *Science (80-.).* **276**, 1404–1407 (1997).
24. Hinits, Y. & Hughes, S. M. Mef2s are required for thick filament formation in nascent muscle fibres. *Development* **134**, 2511–2519 (2007).
25. Potthoff, M. J. *et al.* Regulation of skeletal muscle sarcomere integrity and postnatal muscle function by Mef2c. *Mol. Cell. Biol.* **27**, 8143–51 (2007).
26. Liu, N. *et al.* Requirement of MEF2A, C, and D for skeletal muscle regeneration. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4109–14 (2014).
27. Estrada, K. *et al.* Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat. Genet.* **44**, 491–501 (2012).
28. Medina-Gomez, C. *et al.* Life-Course Genome-wide Association Study Meta-analysis of Total Body BMD and Assessment of Age-Specific Effects. *Am. J. Hum. Genet.* **102**, 88–102 (2018).
29. Arnold, M. A. *et al.* MEF2C Transcription Factor Controls Chondrocyte Hypertrophy and Bone Development. *Dev. Cell* **12**, 377–389 (2007).
30. Kramer, I., Baertschi, S., Halleux, C., Keller, H. & Kneissel, M. *Mef2c* deletion in osteocytes results in increased bone mass. *J. Bone Miner. Res.* **27**, 360–373 (2012).
31. Kang, K. S. & Robling, A. G. New Insights into Wntâ€"Lrp5/6â€"Î²-Catenin Signaling in Mechanotransduction. *Front. Endocrinol. (Lausanne).* **5**, (2015).
32. Ai, M., Heeger, S., Bartels, C. F., Schelling, D. K. & Osteoporosis-Pseudoglioma Collaborative Group. Clinical and Molecular Findings in Osteoporosis-Pseudoglioma Syndrome. *Am. J. Hum. Genet.* **77**, 741–753 (2005).
33. Teebi, A. S., Al-Awadi, S. A., Marafie, M. J., Bushnaq, R. A. & Satyanath, S. Osteoporosis-pseudoglioma syndrome with congenital heart disease: A new association. *J. Med. Genet.* **25**, 32–36 (1988).
34. Leupin, O. *et al.* Bone overgrowth-associated mutations in the LRP4 gene impair sclerostin facilitator function. *J. Biol. Chem.* **286**, 19489–19500 (2011).
35. Barik, A. *et al.* LRP4 is critical for neuromuscular junction maintenance. *J. Neurosci.* **34**, 13892–13905 (2014).
36. Kariminejad, A. *et al.* Severe Cenani-Lenz syndrome caused by loss of LRP4 function. *Am. J. Med. Genet. Part A* **161**, 1475–1479 (2013).
37. Ohkawara, B. *et al.* LRP4 third β-propeller domain mutations cause novel congenital myasthenia by compromising agrin-mediated MuSK signaling in a position-specific manner. *Hum. Mol. Genet.* **23**, 1856–1868 (2014).
38. Morris, J. A. *et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nat. Genet.* **51**, 258–266 (2019).
39. Wuttke, M. *et al.* A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat. Genet.* **51**, 957–972 (2019).

40. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
41. Turcot, V. *et al.* Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nat. Genet.* **50**, 26–35 (2018).
42. #1, Y.-F. P. *et al.* Genome-wide association study of appendicular lean mass in UK Biobank cohort 1. doi:10.1101/643536
43. Barefield, D. Y. *et al.* Experimental Modeling Supports a Role for MyBP-HL as a Novel Myofilament Component in Arrhythmia and Dilated Cardiomyopathy. *Circulation* **136**, 1477–1491 (2017).
44. Lahm, H. *et al.* Myosin binding protein H-like (MYBPHL): a promising biomarker to predict atrial damage. *Sci. Rep.* **9**, (2019).
45. Lee, S.-J. *et al.* Regulation of muscle growth by multiple ligands signaling through activin type II receptors. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 18117–22 (2005).
46. DiGirolamo, D. J., Singhal, V., Chang, X., Lee, S.-J. & Germain-Lee, E. L. Administration of soluble activin receptor 2B increases bone and muscle mass in a mouse model of osteogenesis imperfecta. *Bone Res.* **3**, 14042 (2015).
47. Puolakkainen, T. *et al.* Treatment with soluble activin type IIB-receptor improves bone mass and strength in a mouse model of Duchenne muscular dystrophy. *BMC Musculoskelet. Disord.* **18**, 20 (2017).
48. Goh, B. C. *et al.* Activin receptor type 2A (ACVR2A) functions directly in osteoblasts as a negative regulator of bone mass. *J. Biol. Chem.* **292**, 13809–13822 (2017).
49. Scholtysek, C. *et al.* PPARβ/δ governs Wnt signaling and bone turnover. *Nat. Med.* **19**, 608–613 (2013).
50. Burri, L., Thoresen, G. H. & Berge, R. K. The role of PPAR activation in liver and muscle. *PPAR Research* (2010). doi:10.1155/2010/542359
51. Kazanskaya, O. *et al.* The Wnt signaling regulator R-spondin 3 promotes angioblast and vascular development. *Development* **135**, 3655–3664 (2008).
52. Cambier, L., Plate, M., Sucov, H. M. & Pashmforoush, M. Nkx2-5 regulates cardiac growth through modulation of Wnt signaling by R-spondin3. *Dev.* **141**, 2959–2971 (2014).
53. Chen, D. *et al.* Dentin dysplasia type I—A dental disease with genetic heterogeneity. *Oral Diseases* **25**, 439–446 (2019).
54. Lotta, L. A. *et al.* Association of Genetic Variants Related to Gluteofemoral vs Abdominal Fat Distribution with Type 2 Diabetes, Coronary Disease, and Cardiovascular Risk Factors. *JAMA - J. Am. Med. Assoc.* **320**, 2553–2563 (2018).
55. Spranger, J., Winterpacht, A. & Zabel, B. The type II collagenopathies: A spectrum of chondrodysplasias. *European Journal of Pediatrics* **153**, 56–65 (1994).
56. Graham, S. E. *et al.* Sex-specific and pleiotropic effects underlying kidney function identified from GWAS meta-analysis. *Nat. Commun.* **10**, (2019).
57. Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).
58. O'Seaghdha, C. M. *et al.* Meta-Analysis of Genome-Wide Association Studies Identifies Six New Loci for Serum Calcium Concentrations. *PLoS Genet.* **9**, e1003796 (2013).
59. Sussman, M. A. Developing hearts need their SPEG. *Circulation* **119**, 213–4 (2009).
60. Wang, H. *et al.* Insights from genotype–phenotype correlations by novel SPEG mutations causing centronuclear myopathy. *Neuromuscul. Disord.* **27**, 836–842 (2017).

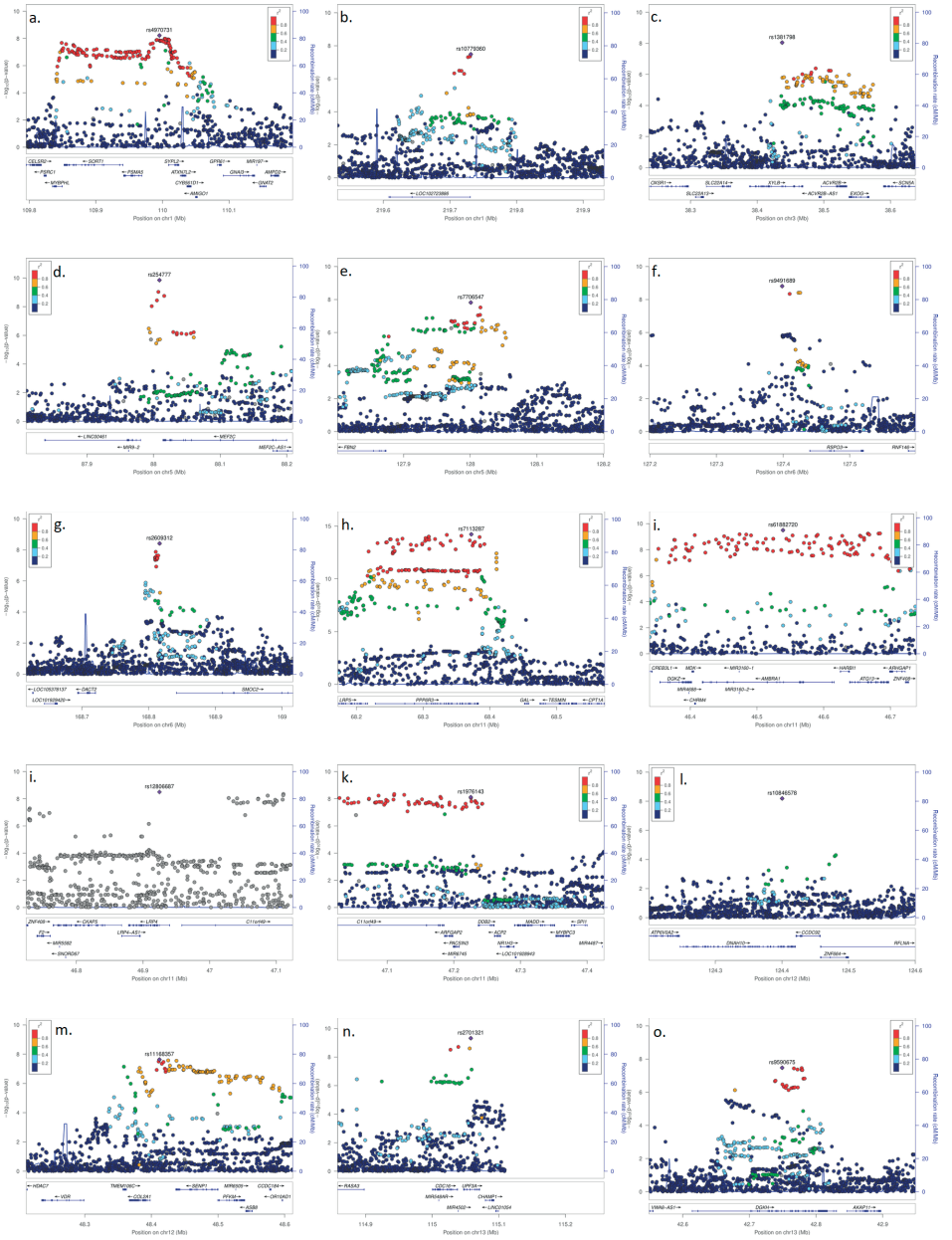
61. Huntoon, V. *et al.* SPEG-deficient skeletal muscles exhibit abnormal triad and defective calcium handling. *Hum. Mol. Genet.* **27**, 1608–1617 (2018).
62. Miyatake, S. *et al.* Biallelic Mutations in MYPN, Encoding Myopalladin, Are Associated with Childhood-Onset, Slowly Progressive Nemaline Myopathy. *Am. J. Hum. Genet.* **100**, 169–178 (2017).
63. Bang, M. L. *et al.* Myopalladin, a novel 145-kilodalton sarcomeric protein with multiple roles in Z-disc and I-band protein assemblies. *J. Cell Biol.* **153**, 413–427 (2001).
64. Ma, K. & Wang, K. Interaction of nebulin SH3 domain with titin PEVK and myopalladin: implications for the signaling and assembly role of titin and nebulin. *FEBS Lett.* **532**, 273–278 (2002).

Supplementary Table 1 | Significant loci from the univariate TB-BMD and TB-LM GWAS.

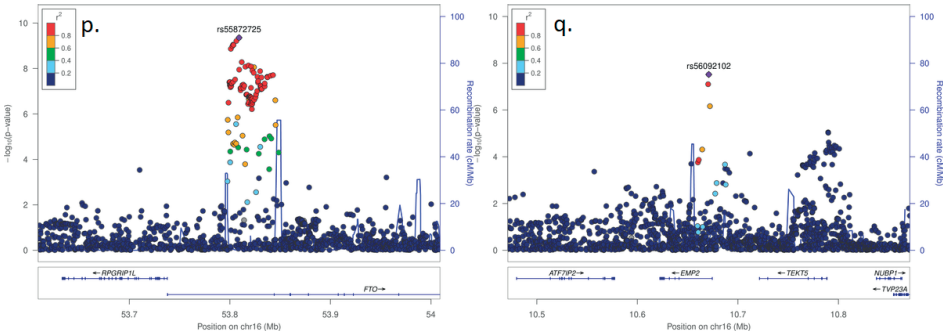
Trait	Locus	SNP	BP	EA	NEA	EAF	Effect	se	P	Closest Gene
TB-BMD	1p36.12	rs10493013	22703035	T	C	0.82	-0.09	0.01	2.20×10^{-15}	*WNT4*
	1p31.3	rs2566752	68656697	T	C	0.62	-0.06	0.01	4.64×10^{-12}	*WLS*
	2q14.2	rs141795717	119539208	A	G	0.02	0.21	0.03	7.73×10^{-11}	*EN1*
	2q24.3	rs6719426	166585566	A	G	0.48	-0.05	0.01	7.08×10^{-11}	*CSRNP3 / GALNT3*
	2q37.1	rs7584554	234308782	T	G	0.44	-0.05	0.01	3.36×10^{-8}	*DGKD*
	3p22.1	rs7609599	41164756	A	G	0.46	-0.05	0.01	7.74×10^{-9}	*CTNNB1*
	6q22.33	rs4580892	127409882	T	C	0.29	0.05	0.01	4.39×10^{-8}	*RSPO3*
	7q21.3	rs7787512	96131103	T	C	0.65	0.06	0.01	1.42×10^{-12}	*C7orf76/ SHFM1*
	7q31.31	rs2908004	120969769	A	G	0.46	0.10	0.01	7.31×10^{-35}	*CPED1/WNT16*
	8q24.12	rs7014574	119977077	T	C	0.56	-0.05	0.01	4.41×10^{-10}	*TNFRSF11B*
	10q21.1	rs7902708	54439728	C	G	0.11	-0.07	0.01	2.96×10^{-8}	*MBL2*
	11p14.1	rs6416055	27281007	T	C	0.63	-0.05	0.01	4.20×10^{-8}	*CCDC34*
	11p11.2	rs79232684	47171714	A	G	0.09	0.09	0.01	4.86×10^{-9}	*C11orf49*
	11q13.2	rs60212556	68305494	T	C	0.27	-0.08	0.01	8.11×10^{-17}	*PPP6R3*
	11q14.2	rs618926	86867252	C	G	0.68	0.05	0.01	1.19×10^{-8}	*TMEM135*
	12q13.13	rs10747668	53737461	T	C	0.71	-0.06	0.01	7.08×10^{-10}	*SP7*
	13q14.11	rs9533094	42965837	A	G	0.54	0.05	0.01	1.08×10^{-8}	*FABP3P2*
	17q21.31	rs66838809	41798621	A	G	0.08	0.09	0.02	1.72×10^{-9}	*MEOX1/DUSP3*
TB-LM	1q41	rs1810144	219649771	T	G	0.30	-0.06	0.01	4.84×10^{-10}	*ZC3H11B*
	2q35	rs1810144	220350452	T	C	0.66	0.05	0.01	6.19×10^{-10}	*SPEG*
	10q21.3	rs12415105	69958631	A	G	0.49	-0.04	0.01	3.69×10^{-8}	*MYPN*
	16q12.2	rs56094641	53806453	A	G	0.60	-0.05	0.01	3.69×10^{-8}	*FTO*



Supplementary Figure1 | QQ plots for the genome-wide meta-analysis for A) TB-BMD; B) TB-LM and C) Bivariate meta-analysis



Supplementary Figure 2 | Regional association plots for each of the genome-wide significant loci 1p13.3 (a), 1q41(b), 3p22.2 (c), 5q14.3 (d), 5q23.3 (e), 6q22.33 (f), 6q27 (g),11q13.2 (h),11p11.2 (i, j, k),12q24.31 (l),12q13.11 (m),13q34 (n),13q14.11 (o),16q12.2 (p), 16p13.13 (q)



Supplementary Figure2 | Regional association plots for each of the genome-wide significant loci 1p13.3 (a), 1q41(b), 3p22.2 (c), 5q14.3 (d), 5q23.3 (e), 6q22.33 (f), 6q27 (g),11q13.2 (h),11p11.2 (i, j, k),12q24.31 (l),12q13.11 (m),13q34 (n),13q14.11 (o),16q12.2 (p), 16p13.13 (q) *(continued)*