

ARTICLE

DOI: 10.1038/s41467-017-00031-7

OPEN

Large meta-analysis of genome-wide association studies identifies five loci for lean body mass

M. Carola Zillikens *et al.*[#]

Lean body mass, consisting mostly of skeletal muscle, is important for healthy aging. We performed a genome-wide association study for whole body (20 cohorts of European ancestry with $n = 38,292$) and appendicular (arms and legs) lean body mass ($n = 28,330$) measured using dual energy X-ray absorptiometry or bioelectrical impedance analysis, adjusted for sex, age, height, and fat mass. Twenty-one single-nucleotide polymorphisms were significantly associated with lean body mass either genome wide ($p < 5 \times 10^{-8}$) or suggestively genome wide ($p < 2.3 \times 10^{-6}$). Replication in 63,475 (47,227 of European ancestry) individuals from 33 cohorts for whole body lean body mass and in 45,090 (42,360 of European ancestry) subjects from 25 cohorts for appendicular lean body mass was successful for five single-nucleotide polymorphisms in/near *HSD17B11*, *VCAN*, *ADAMTSL3*, *IRS1*, and *FTO* for total lean body mass and for three single-nucleotide polymorphisms in/near *VCAN*, *ADAMTSL3*, and *IRS1* for appendicular lean body mass. Our findings provide new insight into the genetics of lean body mass.

Correspondence and requests for materials should be addressed to D.P.K. (email: kiel@hsl.harvard.edu)

[#]A full list of authors and their affiliations appears at the end of the paper.

Lean body mass consists primarily of skeletal muscle, is an important contributor to physical strength, mobility, stamina, and balance^{1–3}, and has been a very recent focus of an effort to define “sarcopenia” (loss of muscle tissue) for clinical care and drug development⁴. The determinants of adult skeletal muscle mass have not been well characterized. It is known, for example, that exercise produces increases in muscle mass, and there is some evidence that protein intake is directly associated with lean mass⁵. Heavier people have increased muscle mass⁶, which may be due to the loading effect of increased fat mass or may reflect a common genetic background between muscle and adipose tissue⁷. With aging, there is a progressive loss of skeletal muscle mass, and a concurrent increase in fatty infiltration and fibrosis of muscle. This loss of muscle mass may reach a critical point at which time functional impairment and even disability occurs⁸. In fact, the annual healthcare costs of sarcopenia in the United States are estimated to be in excess of 18 billion dollars⁹.

As estimated from family and twin studies, lean mass is a highly heritable phenotype with heritability estimates of 0.52–0.60^{10, 11}. While there have been previous studies related to the genetic background of BMI and fat mass, few studies have searched for genes associated with lean mass^{12–15, 16, 17}. To date, no single-nucleotide polymorphisms (SNPs) have been found to be associated at a genome-wide significance level with lean mass (p -values $< 5 \times 10^{-8}$)^{18, 19}. A copy number variation located in the *GREM1* gene was reported to be associated with lean body mass in a genome-wide association study (GWAS) of 1627 Chinese²⁰. Guo et al.¹⁹ identified in 1627 Chinese and replicated in 2286 European ancestry individuals, a locus near *CNTF* and *GLYAT* genes at 11q12.1 in a bivariate GWAS for bone-size phenotypes and appendicular lean mass. Most recently in a study of Japanese women, the *PRDM16* gene was suggested to be associated with lean mass²¹.

With the advent of relatively simple, inexpensive methods of measuring the fat and lean compartments of the body using dual energy X-ray absorptiometry (DXA) and bioelectrical impedance analysis (BIA), multiple cohort studies have accumulated phenotypic information on body composition that permits large-scale GWAS to be performed. While whole body lean mass incorporates all of the non-fat soft tissue including the internal

organs, appendicular lean mass, estimated by DXA and BIA, may be a better reflection of skeletal muscle mass^{22–24}. To identify genetic loci associated with whole body and appendicular lean mass, we performed a large-scale GWAS meta-analysis in over 100,000 participants from 53 studies yielding sufficient power to identify common variants with small to moderate effect sizes.

Results

GWAS meta-analyses for discovery and replication.

Descriptions and characteristics of the study populations in the discovery stage and the replication stage are shown in Supplementary Tables 4 and 5 and Supplementary Note 2.8. The age of the participants ranged from 18 to 100 years. In the GWAS discovery set, comprising 38,292 participants for whole body lean mass and 28,330 participants for appendicular lean mass, a substantial excess of low p -values compared to the null distribution was observed after genomic control adjustment of the individual studies prior to meta-analysis: $\lambda_{GC} = 1.076$ and $\lambda_{GC} = 1.075$, for whole body and appendicular lean mass, respectively (Supplementary Fig. 1)

Meta-analyses were conducted using the METAL package (www.sph.umich.edu/csg/abecasis/metal/). We used the inverse variance weighting and fixed-effect model approach. Supplementary Table 1 shows the genome-wide significant (GWS) and suggestive (sGWS) results in the discovery set. For whole body lean mass, we observed one GWS result in/near *HSD17B11* and 12 sGWS results (in/near *VCAN*, *ADAMTSL3*, *IRS1*, *FTO* (two SNPs), *MOV10*, *HMCN1*, *RHOC*, *FRK*, *AKR1B1*, *CALCR*, and *KLF12*). For appendicular lean mass, one result was GWS (intronic SNP in *PKIB*) and seven were sGWS (in/near *VCAN*, *ADAMTSL3*, *HSD17B11*, *IRS1*, *FRK*, *TXN*, and *CTNNA3*).

We selected 21 associations (13 for whole body lean mass and 8 for appendicular lean mass; a total of 16 discovery SNPs with 5 SNPs overlapping between the two phenotypes) (Supplementary Table 1) to conduct a replication study in a set of 33 cohorts comprising up to 48,125 participants of European descent for whole body lean mass and 43,258 participants for appendicular lean mass. Both in silico replication and de novo genotyping for replication was conducted. Table 1 shows the results for successfully replicated SNPs in participants of European ancestry,

Table 1 Results for the successfully replicated SNPs in discovery, replication and combined sample

SNP ID	Chrom	Position	Closest gene	Allele ½	EAF	Discovery (n = 38,292)			Replication EU (n = 47,227)			Combined EU (n = 85,519)		
						Beta	SE	p-value	Beta	SE	p-value	Beta	SE	p-value
<i>Whole body lean mass</i>														
rs2943656	2	226830162	<i>IRS1</i>	A/G	0.38	-0.17	0.03	2.5×10^{-7}	-0.13	0.03	8.0×10^{-6}	-0.14	0.02	1.5×10^{-11}
rs9991501	4	88477507	<i>HSD17B11</i>	T/C	0.04	-0.61	0.01	2.9×10^{-8}	-0.26	0.08	1.9×10^{-3}	-0.39	0.07	5.8×10^{-9}
rs2287926	5	82851164	<i>VCAN</i>	A/G	0.12	0.24	0.05	8.6×10^{-7}	0.15	0.04	8.5×10^{-4}	0.19	0.03	7.5×10^{-9}
rs4842924	15	82378611	<i>ADAMTSL3</i>	T/C	0.52	-0.17	0.03	1.4×10^{-7}	-0.08	0.03	3.9×10^{-3}	-0.12	0.02	1.4×10^{-8}
rs9936385	16	52376670	<i>FTO</i>	T/C	0.61	-0.17	0.03	1.1×10^{-6}	-0.11	0.03	1.6×10^{-4}	-0.14	0.02	1.4×10^{-9}
<i>Appendicular lean mass</i>														
rs2943656	2	226830162	<i>IRS1</i>	A/G	0.38	-0.10	0.02	1.1×10^{-6}	-0.06	0.01	2.2×10^{-5}	-0.07	0.01	2.9×10^{-10}
rs2287926	5	82851164	<i>VCAN</i>	A/G	0.13	0.14	0.03	8.1×10^{-7}	0.08	0.02	3.5×10^{-4}	0.10	0.02	4.5×10^{-9}
rs4842924	15	82378611	<i>ADAMTSL3</i>	T/C	0.52	-0.09	0.02	1.2×10^{-6}	-0.05	0.02	1.6×10^{-3}	-0.06	0.01	5.0×10^{-8}
All results reflect analyses in participants of European ancestry														
No significant heterogeneity was observed at $\alpha = 0.00625$ (0.05/8)														
Only mild heterogeneity was indicated in two associations for whole body lean mass when using an uncorrected threshold of $\alpha = 0.05$, <i>FTO</i> /rs9936385 ($p = 0.018$, $I^2 = 34\%$) and <i>HSD17B11</i> /rs9991501 ($p = 0.04$, $I^2 = 31\%$)														
All results were adjusted for the following covariates: sex, age, height and fat mass (kg)														

Table 2 Tissue-specific regulatory-element enrichment analyses of the GWAS loci (GWAS SNPs and SNPs in LD with the GWAS SNPs)

SNP ID	In/near gene	SNP functional role	Coding variant function by Polyphen2	Number of SNPs in LD ^a	p-value ^b of tissue-specific regulatory element enrichment analyses ^c in five tissues ^d					
					Skeletal muscle	Smooth muscle	Fat	Brain	Blood	Gastrointestinal tract
rs2943656	<i>IRS1</i>	Intergenic		86	0.14	0.38	1×10^{-7}	0.04	0.82	1
rs9991501	<i>HSD17B11</i>	Exonic missense	Benign (Arg283Gln)	1	NA ^e	NA	NA	NA	NA	NA
rs2287926	<i>VCAN</i>	Exonic missense	Possibly damaging (Gly428Asp)	5	1×10^{-7}	1×10^{-7}	1×10^{-7}	1×10^{-7}	1	0.98
rs4842924	<i>ADAMTSL3</i>	Intronic		87	1×10^{-7}	1×10^{-7}	1×10^{-7}	1×10^{-7}	1	0.23
rs9936385	<i>FTO</i>	Intronic		91	0.78	0.38	0.49	0.45	1	0.51

^aSNPs in LD: number of SNPs in LD ($r^2 \geq 0.8$ and $MAF \geq 1\%$, based on CEU samples in the 1000 Genome Project) with the lead GWAS SNP in each locus

^bMinimum p-value permutation tests: this analysis included all SNPs in LD with the GWAS lead SNPs. Multiple testing correction was done by the minimum p-value permutation test. Permutation p-values < 0.05 are considered as statistically significant.

^cEnhancers and promoters (regulatory elements) in 25 chromatin states (retrieved from HaploReg4 database): SNPs are located within active regulatory elements, including promoter upstream TSS, promoter downstream TSS 1, promoter downstream TSS 2, transcribed and regulatory (prom/enh), transcribed 5' preferential and enh, transcribed 3' preferential and enh, transcribed and weak enhancer, active enhancer 1, active enhancer 2, active enhancer flank, weak enhancer 1, weak enhancer 2, primary H3K27ac possible enhancer, poised promoter, and bivalent promoter

^dSee Supplementary Note 2.6 for description of human primary cells and tissues that were included in each tissue group

^eWe did not perform enrichment analysis on rs9991501 because rs9991501 has no other SNPs in LD to obtain overlapping regulatory elements

including the discovery phase, replication phase, and the combined results.

For whole body lean mass, joint analysis of the discovery and replication cohorts successfully replicated five SNPs in/near *HSD17B11*, *VCAN*, *ADAMTSL3*, *IRS1*, and *FTO* (p -values between 1.4×10^{-8} and 1.5×10^{-11} and lower than discovery p -values). Three of these five SNPs (located in/near *VCAN*, *ADAMTSL3*, and *IRS1*) were also successfully replicated (p -values between 5×10^{-8} and 2.9×10^{-10}) for appendicular lean mass. None of the eight replicated associations (five for whole body and three appendicular lean mass) had significant heterogeneity at $\alpha = 0.00625$ (0.05/8, Bonferroni-corrected for eight tests). Only mild heterogeneity was indicated in two whole body lean mass SNPs when using an uncorrected threshold of $\alpha = 0.05$: *FTO* ($p = 0.018$, $I^2 = 34\%$) and *HSD17B11* ($p = 0.04$, $I^2 = 31\%$). For appendicular lean mass, p -values for heterogeneity were > 0.05 for all three replicated SNPs.

Supplementary Table 1 shows the results for all participants including those of non-European descent as well. Results were similar showing low heterogeneity using “transethnic meta-analysis” (MANTRA)²⁵ for inclusion of both cohorts of European ancestry and replication cohorts with Asians or African Americans. Heterogeneity probability values were below 0.5 for all replicated SNPs both for whole body and appendicular lean mass. Furthermore, in this combined analysis, except for the *VCAN* locus, \log_{10} Bayes' factors were > 6.0 and p -values were smaller than those found in European-only ancestry analysis.

Additional analyses stratified by sex failed to identify significant sex-specific associations or evidence of an interaction between SNPs and sex (Supplementary Table 2 and Supplementary Note 1). Similarly, we also found no evidence for heterogeneity between measurement techniques (BIA vs DXA; Supplementary Table 3 and Supplementary Note 2). Finally, we failed to replicate previously reported candidate genes for lean mass (Supplementary Note 2.4)

Association with other anthropometric phenotypes. We looked for associations between the lead SNPs in the five replicated loci and other reported anthropometric phenotypes from the GIANT Consortium (Supplementary Table 10)^{26–29}. There were no significant associations ($p < 0.05$) between the SNP in *HSD17B11* and any reported phenotypes. The allele associated with greater lean mass was associated with lower values of various

anthropometric traits for the *IRS1* (hip circumference (HC), waist circumference (WC)), *VCAN* (hip ratio, WC adjusted for BMI, waist to hip ratio adjusted for BMI), and the *ADAMTSL3* locus (height, hip, and WC with the association becoming more significant for hip and waist adjusted for BMI). The replicated *FTO* SNP was very significantly associated with BMI ($p = 2.7 \times 10^{-144}$) and significantly associated with HC ($p = 9.3 \times 10^{-81}$) and WC ($p = 1.3 \times 10^{-96}$) in the same direction as the lean mass association (i.e., the higher lean mass allele was associated with higher values of anthropometric traits).

Annotation and enrichment analysis of regulatory elements.

Among five replicated SNPs, rs9991501 (*HSD17B11* locus) and rs2287926 (*VCAN* locus) are missense SNPs. rs2287926 was predicted as possibly damaging to the protein structure and/or function by PolyPhen-2³⁰. Since the remaining GWAS SNPs are non-coding, to estimate whether these SNPs are located in regulatory elements in specific human tissue/cell types, we performed a tissue-specific regulatory-element enrichment analysis using experimental epigenetic evidence including DNase hypersensitive sites, histone modifications, and transcription factor-binding sites in human cell lines and tissues from the ENCODE Project and the Epigenetic Roadmap Project. As shown in Table 2, 86 SNPs were in high LD ($r^2 \geq 0.8$) with the GWAS lead SNP rs2943656 at the *IRS1* locus. The SNPs in this locus that were in high LD were enriched in enhancers estimated by ChromHMM³¹ (permutation p -values < 0.05 ; after multiple testing corrections), especially enriched in fat and brain tissues, but not in skeletal muscle, smooth muscle, blood and gastrointestinal tract tissues in the ENCODE and Roadmap projects. Although the *IRS1* locus was not specifically enriched with skeletal muscle tissue enhancers, the lead SNP rs2943656 itself at the *IRS1* locus was actually located within a histone mark-identified promoter in an adult skeletal muscle sample and a histone mark-identified enhancer in several muscle samples (such as muscle satellite cultured cells, fetal skeletal muscle, skeletal muscle myoblasts, and skeletal muscle myotubes (Supplementary Figs. 6 and 7)). Based on the position weight matrices (PWMs) score from Chip-seq and other sequencing resources, rs2943656 was found to possibly alter regulatory motifs, including Irf, Foxo, Sox, and Zfp105 in skeletal muscle tissues with a PWM score p -value $< 4^{-7}$ ($\approx 6.1 \times 10^{-5}$)³².

For the *ADAMTSL3* locus, the GWAS lead SNP rs4842924 was located in a histone mark-identified enhancer in smooth muscle

tissues (Supplementary Figs. 11 and 12). The promoter/enhancer enrichment analysis in the *ADAMTSL3* locus showed significant enrichments in skeletal muscle, smooth muscle, fat and brain tissues, but not in blood and gastrointestinal tract tissues. At the *FTO* locus, a few SNPs in high LD with the GWAS lead SNP, rs2287926, were located within a group of enhancers that are not muscle tissue-specific (Supplementary Fig. 13); therefore, no significant enrichment was found in any tissues listed in Table 2.

Expression quantitative trait loci. We queried existing *cis*-expression quantitative trait loci (eQTLs) analyses on the five replicated GWAS SNPs rs2943656, rs9991501, rs2287926, rs4842924, and rs9936385 with transcripts within 2 Mb of the SNP position in skeletal muscle tissues as well as in subcutaneous adipose, omental adipose, liver tissue, lymphocytes, and primary osteoblasts (obtained from bone biopsies). Rs9936385 was associated with *FTO* expression in skeletal muscle tissues in the FUSION samples ($p = 4.4 \times 10^{-11}$) (Supplementary Table 10). However, in sequential conditional analysis, upon addition of the lead *FTO* eQTL SNP (rs11649091, association *p*-value with *FTO* gene expression $= 5.1 \times 10^{-15}$), rs9936385 was no longer significantly associated ($p = 0.16$), whereas rs11649091 remained significantly associated with *FTO* gene expression ($p = 1.9 \times 10^{-5}$). SNP rs11649091 could not be imputed using our HapMap imputation references; thus, we did not have association results between rs11649091 and lean mass in the present study. The T allele of rs9936385, associated with reduced lean mass in the present study, was significantly associated with lower *FTO* gene expression levels in the GTEx project (Supplementary Table 9 and Supplementary Fig. 14). For rs9936385, we also examined *IRX3* and *IRX5* expression in skeletal muscle tissues, as recent reports have implicated GWAS SNPs associated with obesity in intron 2 of the *FTO* gene as being associated with *IRX3* and *IRX5* gene expression in brain³³ and adipose tissue³⁴. SNP rs9936385 was not significantly associated with *IRX3* and *IRX5* gene expression in skeletal muscle tissues. For missense SNP rs9991501 in the *HSD17B11* locus, a significant association with *HSD17B11* gene expression was found in skeletal muscle from the GTEx project ($p = 1.4 \times 10^{-4}$). There were no significant GWAS SNPs in strong LD with this one; thus conditional analyses were not performed.

As shown in Supplementary Table 10, for GWAS lead SNP rs2943656 in the *IRS1* locus, significant eQTLs with the *IRS1* gene expression in omental ($p = 4 \times 10^{-7}$) and subcutaneous fat tissue ($p = 6.44 \times 10^{-6}$) were found.

Finally we found no evidence for differential expression of our five replicated genes in young vs old muscle biopsies (Supplementary Note 2.7 for methods and results).

Discussion

In this first large-scale GWA meta-analysis study for lean mass that included most of the cohorts worldwide with lean mass phenotypes, we identified and successfully replicated five GWS loci (in/near *HSD17B11*, *VCAN*, *ADAMTSL3*, *IRS1*, and *FTO* genes) for whole body lean mass and three of these (in/near *VCAN*, *IRS1*, and *ADAMTSL3* genes) for appendicular lean mass, both important for sarcopenia diagnosis.

This study contributes to a better understanding of the biology underlying inter-individual variation in muscle mass, since lean body mass consists primarily of muscle mass (especially in the extremities). Genetic determinants of lean body mass cannot be studied specifically by using anthropometric measures such as height, waist circumference, hip circumferences, or BMI, as evidenced by our finding of associations between genetic loci and lean mass that were not observed in results from the GIANT

consortium^{26–29}. Four novel GWS loci for lean mass phenotypes harboring *ADAMTSL3*, *VCAN*, *HSD17B11*, and *IRS1* genes have biologic effects supporting their role in skeletal muscle. Although the functional involvement of the *ADAMTSL3* gene (a disintegrin-like and metalloprotease domain with thrombospondin type I motifs-like 3) remains unknown, it has been shown to be consistently associated with adult height in large human samples²⁶, including individuals of African ancestry³⁵. The gene is expressed ubiquitously, including in skeletal muscle but the lead GWAS SNP in this locus was not significantly associated with expression of *ADAMTSL3* in any of the tissues examined. One might hypothesize that our observed association between this gene and lean mass could be a reflection of an allometric relationship between muscle mass/size and body size³⁶. Since our analyses were adjusted for height, it is plausible that variation in *ADAMTSL3* is associated with muscle mass directly. In fact, in a recent study that identified classes of potentially regulatory genomic elements that are enriched in GWAS loci, the height phenotype was almost exclusively enriched in DNase sites in muscle³⁷. In our study, we observed enrichment of regulatory genomic elements in this locus in both skeletal muscle and smooth muscle, which may not fully support a unique role of this gene in only skeletal muscle tissue.

Versican (*VCAN*) plays a role in intercellular signaling and in connecting cells with the extracellular matrix, which also is important for the skeletal muscle. Interestingly, *VCAN* facilitates chondrocyte differentiation and regulates joint morphogenesis³⁸, and therefore has a wider role in the musculoskeletal health. The GWAS SNP in the *HSD17B11* locus was significantly associated with the *HSD17B11* gene expression in skeletal muscle. Hydroxysteroid (17- β) dehydrogenase 11 (*HSD17B11*) functions in steroid biosynthesis, and most relevant to muscle tissue, contributes to androgen catabolic processes. Although not much is known about how variants in this gene associate with sex-hormone related human phenotypes, androgen metabolism is certainly a driver of muscle tissue anabolism and catabolism.

As for the *FTO* locus, variants in the *FTO* gene, which is known to regulate postnatal growth in mice³⁹, have been found to be associated with adiposity/obesity⁴⁰ and related traits, such as BMI⁴¹, metabolic syndrome/type 2 diabetes and even with menarche⁴². Apart from its role in adiposity, in two recent candidate gene studies, SNPs in *FTO* were found to associate both with DXA-derived fat and lean mass⁴³, and in one, the association for lean mass was only slightly attenuated after fat mass adjustment⁴⁴, similar to our results using fat mass-adjusted lean mass. Also, *FTO* knockout mice have been shown to have not only reduced fat mass, but also decrease in lean mass⁴⁵. Recently, obesity associated non-coding sequences in the *FTO* gene were found to be functionally connected to the nearby gene, *IRX3*, by directly interacting with the promoters of this gene³³, suggesting obesity SNPs located inside the *FTO* gene may regulate gene expression other than *FTO*. Our GWAS lead SNP rs9936385 in the *FTO* locus was in LD with the obesity GWAS SNP and located in the same haplotype. A denser fine-mapping study using sequencing of the *FTO* locus with a better resolution will be helpful in narrowing down the *FTO* region to identify potential causal variant(s). The actual function of the variants and the underlying mechanisms of *FTO*'s involvement in skeletal muscle biology still need to be further elucidated by in vitro and animal experiments.

The other body composition-related gene that was successfully replicated is the insulin receptor substrate 1 (*IRS1*), which belongs to the insulin signaling pathway and participates in growth hormone and adipocytokine signaling pathways. Besides being overexpressed in adipocytes, *IRS1* is also highly expressed in skeletal muscle¹³, *IRS1* polymorphisms have been associated

with fasting insulin-related traits⁴⁶, adiposity¹³, and serum triglycerides/HDL cholesterol⁴⁷. Our GWAS lead SNP rs2943656 was associated with *IRS1* gene expression in skeletal muscle obtained from the GTEx project. Interestingly, another SNP, rs2943650, near *IRS1* in high LD with our lead SNP ($R^2 = 0.854$) was previously found to be associated with percent body fat in the opposite direction from the association that we find with lean mass¹³. The body fat percentage decreasing allele in this study was associated with lower *IRS1* expression in omental and subcutaneous fat¹³. Another study reported an association between a SNP in high LD (rs2943641) with *IRS-1* protein expression and insulin-induced phosphatidylinositol 3-OH kinase activity in skeletal muscle¹³. Also rs2943656 was associated with obesity traits in the GIANT Consortium; the allele associated with greater lean mass was inversely associated with obesity traits. Further understanding of the potential functional effects of these variants in the *IRS1* locus are needed to determine whether they have pleiotropic and opposite effects on fat and lean tissue. The same is true for variants in the *FTO* locus for which the allele that was associated with greater lean mass was previously reported to be associated with greater fat mass.

From the current eQTL data analysis, we have no definitive evidence that the non-coding GWAS SNPs (or variants in high LD) functionally influence gene expression of *IRS1*, *HSD17B1*, and *FTO* in skeletal muscle. Use of a larger reference panel for imputation in the GWAS sample, and analysis of larger tissue expression data sets coupled with conditional analysis will help reveal if underlying functional associations exist. It should be emphasized that with the current data we are not able to determine which SNPs may be functional and it is possible that the identified lean mass variants are not driving the associations with expression. Using ENCODE and Epigenetic Roadmap data, we found that the GWAS SNPs (or SNPs in LD with these SNPs) were significantly enriched in the predicted gene regulatory regions (in our case, enhancers) not only in skeletal muscle, but also in smooth muscle, fat and brain tissues. With the complexity of lean mass phenotypes, we cannot rule out the possibility that these genes are involved in regulating lean mass biology via tissues other than skeletal muscle.

Using results from the overall meta-analysis, the percent variance explained by the successfully replicated SNPs was 0.23% and 0.16% for whole body and appendicular lean mass, respectively. Estimates were slightly higher when we used individual level data from the Framingham Study cohorts (percent variance explained of 0.97% and 0.55% for whole body and appendicular lean mass, respectively). This relatively small percentage of explained variance is not dissimilar to other body composition measures such as bone mineral density of the femoral neck, for which 63 SNPs explained 5.8% of the variance⁴⁸. To estimate the proportion of variance for lean mass explained by all genotyped SNPs across the genome in Framingham Study participants, we applied a GREML model implemented in the GCTA package⁴⁹ with the assumption that all 550K genotyped SNPs captured $\geq 80\%$ of the common sequence variance in the Framingham Study participants who are Caucasians of European ancestry. We estimated that the proportion of whole body lean mass and appendicular lean mass variance explained by all genotyped SNPs together was 43.3% (SE: 2.7%) and 44.2% (SE: 2.6%), respectively (after adjustment for age, age², sex, height and fat mass), suggesting most of the heritability of lean mass was not detected in the current study due to the limitations of study design (only common variants in this study). This percent variance is similar to the 45% of the variance explained for height using this method⁴⁹.

Because of the substantial sexual dimorphism in body composition with men having higher muscle mass compared with women²⁴, we performed both sex-combined and sex-stratified

analyses. We examined potential sex differences in the genetic associations. Lean mass is a highly heritable trait and the heritability is of similar magnitude in both genders^{24, 50}. No formal interaction test between SNP and sex was significant for any of these SNPs. Thus, our findings do not support any substantial sex-specific genetic influence for any of the successfully replicated lean mass SNPs. We cannot rule out the possibility that the reported GWS SNPs are false-positive findings, although the chance of such false positives is extremely low due to the robustness of replication in our well-powered study. There are also other limitations to our study. Because lean mass is correlated with fat mass, we adjusted for fat mass to focus our search on genes contributing to lean mass independent of those regulating fat mass. A potential limitation with this strategy of adjusting for fat mass is that the power to identify genetic signals with a similar impact on lean and fat mass will be reduced. Nevertheless, the *FTO* signal was found to be significantly associated with lean mass after fat adjustment and the direction of this association was the same as the association with fat mass. Since androgens have a major impact on muscle mass, it is a limitation of the present study that the X chromosome, harboring the androgen receptor gene, was not included in the present meta-analysis. Another potential weakness of this study is our decision to meta-analyze body composition results using two different techniques (BIA and DXA). Nevertheless, the two methods are highly correlated ($r = 0.83$ for Framingham cohort participants), and by combining them power to detect GWS loci was greatly enhanced.

In conclusion, in this first large-scale meta-analysis of GWAS, five GWS variants in or near the *HSD17B11*, *VCAN*, *ADAMTSL3*, *IRS1*, and *FTO* genes were found to be robustly associated with lean body mass. Three of these loci were found to be significantly enriched in enhancers and promoters in muscle cells, suggesting that our signals have a potential functional role in muscle. Our findings shed light on pathophysiological mechanisms underlying lean mass variation and potential complex interrelations between the genetic architecture of muscle mass, fat mass, body height and metabolic disease.

Methods

Study summary. We focused on two phenotypes: (1) whole body lean mass; (2) appendicular lean mass. For these two phenotypes, we performed a genome-wide meta-analysis of the discovery cohorts (Stage I), then meta-analyzed the genotyped discovery SNPs in replication cohorts (Stage II), followed by a combined analysis with discovery and replication cohorts (Supplementary Fig. 1). The total sample size for the combined analysis was just over 100,000 from 53 studies. Our primary results are based on the ~85,000 individuals of European descent in 47 studies, as was pre-specified. Because whole body and appendicular lean mass are correlated with fat mass and height, analyses were adjusted for these potential confounders in addition to sex, age, age², and other study-specific covariates to focus our search on genes contributing to lean mass independent of those of body height and fat mass.

Study population. The Stage I Discovery sample comprised 38,292 individuals of European ancestry drawn from 20 cohorts with a variety of epidemiological designs and participant characteristics (Supplementary Tables 4–6 and Supplementary Note 2.8). Whole body lean mass was measured using DXA (10 cohorts, $n = 21,074$) and BIA (10 cohorts, $n = 17,218$). Of the 20 cohorts, 15 consisted of male and female subjects, while 2 had male and 3 had female subjects only. In total, the cohorts included 22,705 women and 15,587 men. Appendicular lean mass was estimated in 28,330 subjects from a subset of 15 cohorts (9 using DXA and 6 using BIA).

Subjects from 33 additional studies were used for replication with a total sample size of 63,475 individuals. Of these 63,475, the majority was of European ancestry ($n = 47,227$ in 27 cohorts), and the remaining 16,248 were of African American, South Asian, or Korean ancestry (Supplementary Table 4). All these 33 cohorts had data for whole body lean mass. Among them, 16 studies had DXA measurements ($n = 23,718$) and 17 studies had BIA measurements ($n = 39,757$). Twenty-five cohorts had data for appendicular lean mass in a total of 45,090 individuals (16 cohorts with DXA ($n = 23,718$) and 9 with BIA ($n = 21,372$)). Of these, 42,360 individuals (23 cohorts) were of European ancestry and the remaining 2730 of African American and Korean descent. Our a priori aim was to perform replication

in cohorts with European subjects only and to explore if adding non-European cohorts would increase power or show evidence of heterogeneity due to ethnicity. The Stage II Replication included cohorts with existing GWAS data that were unavailable at the time of the Stage I Discovery, and cohorts who agreed to undergo de novo genotyping. All studies were approved by their institutional ethics review committees and all participants provided written informed consent.

Lean mass measurements. Lean mass was measured in all cohorts using either DXA or BIA. DXA provides body composition as three materials based on specific X-ray attenuation properties; bone mineral, lipid (triglycerides, phospholipid membranes, etc.) and lipid-free soft tissue. For each pixel on the DXA scan, these three materials are quantified. For the cohorts with DXA measures, the phenotype used for these analyses was the lipid free, soft tissue compartment that is referred to as lean mass, and is the sum of body water, protein, glycerol, and soft tissue mineral mass. Two lean mass phenotypes were used: whole body lean mass and appendicular lean mass. The latter was obtained by considering only pixels in the arms and legs collectively, which has been demonstrated to be a valid measure of skeletal muscle mass⁵¹.

Some of the cohorts estimated body composition using BIA, which relies on the geometrical relationship between impedance (Z), length (L), and volume (V) of an electrical conductor. Adapted to the human body, V corresponds to the volume of fat-free mass (FFM) and L to the height of the subject. Z is composed of the pure resistance (R) of the conductor, the FFM, and the reactance (X_c), produced by the capacitance of cellular membranes, tissue interfaces and non-ionic tissues: $Z^2 = R^2 + X_c^2$. A variety of BIA machines were used by the various cohorts (summarized in Supplementary Table 5), and in some cohorts, the specific resistance and reactance measures were not available because the manufacturers provided only summary output on FFM. For BIA cohorts with specific resistance and reactance measures, we used the validated equation from Kyle et al.⁵² with an R^2 of 0.95 between BIA and DXA to calculate the appendicular lean mass.

Stage 1: genome-wide association analyses in discovery cohorts. *Genotyping and imputation.* Genome-wide genotyping was done by each study on a variety of platforms following standard manufacturer protocols. Quality control was performed independently for each study. To facilitate meta-analysis, each group performed genotype imputation with IMPUTE⁵³ or MACH⁵⁴ software using HapMap Phase II release 22 reference panels (CEU or CHB/JPT as appropriate). Overall imputation quality scores for each SNP were obtained from IMPUTE (“proper_info”) or MACH (“rsq_hat”). Details on the genotyping platform used, genotype quality control procedures and software for imputation employed for each study are presented in Supplementary Table 6.

Study-specific genome-wide association analyses with lean mass. In each study, a multiple linear regression model with additive genetic effect was applied to test for phenotype–genotype association using ~2.0 to 2.5 million genotyped and/or imputed autosomal SNPs. Other covariates adjusted in the model included ancestral genetic background, sex, age, age², height, fat mass measured by the body composition device (kg) and study-specific covariates when appropriate such as clinical center for multi-center cohorts. Adjustment for ancestral background was done within cohorts using principal component analyses as necessary. Furthermore, for family-based studies, including the Framingham Study, ERF, UK-Twins, Old Order Amish Study and the Indiana cohort, familial relatedness was taken into account in the statistical analysis within their cohorts by: (1) linear mixed-effects models that specified fixed genotypic and covariate effects and a random polygenic effect to account for familial correlations (the R Kinship package; <http://cran.r-project.org/web/packages/>) in the Framingham Study; (2) the GenABEL⁵⁵ in the ERF and UK-Twins cohorts; (3) the Mixed Model Analysis for Pedigrees (MMAP) program (<http://edn.som.umaryland.edu/mmap/index.php>) in the Amish cohort; and GWAF, an R package for genome-wide association analyses with family data in the Indiana cohort⁵⁶.

Meta-analyses. Meta-analyses were conducted using the METAL package (www.sph.umich.edu/csg/abecasis/metal/). We used the inverse variance weighting and fixed-effect model approach. Prior to meta-analysis, we filtered out SNPs with low minor allele frequency, MAF (<1%) and poor imputation quality (proper_info<0.4 for IMPUTE and rsq_hat<0.3) and applied genomic control correction where the genomic control parameter lambda (λ_{GC}) was >1.0.

We used quantile–quantile (Q–Q) plots of observed vs. expected $-\log_{10}$ (p -value) to examine the genome-wide distribution of p -values for signs of excessive false-positive results. We generated Manhattan plots to report genome-wide p -values, regional plots for genomic regions within 100 Kb of top hits, and forest plots for meta-analyses and study-specific results of the most significant SNP associations. A threshold of $p < 5 \times 10^{-8}$ was pre-specified as being genome-wide significant (GWS), while a threshold of $p < 2.3 \times 10^{-6}$ was used to select SNPs for a replication study (suggestive genome-wide significant, sGWS).

Stage 2: replication. In each GWS or sGWS locus, we selected the lead SNP with the lowest p -value for replication. In addition, GWS or sGWS SNPs that had low-linkage disequilibrium with the lead SNPs ($LD < 0.5$) were also selected for replication. Both in silico replication and de novo genotyping for replication was conducted. In silico replication was done in 24 cohorts with GWAS SNP chip

genotyping that did not have data available at the time of the initial discovery efforts (Supplementary Table 7). De-novo replication genotyping was done using: KBioScience Allele-Specific Polymorphism (KASP) SNP genotyping system (in OPRA, PEAQ25, AGES, CAIFOS, DOPS cohorts), TaqMan (METSIM), Illumina OmniExpress + Illumina Metabochip (PIVUS and ULSAM), or Sequenom’s iPLEX (WHI) (Supplementary Table 8). Samples and SNPs that did not meet the quality control criteria defined by each individual study were excluded. Minimum genotyping quality-control criteria were defined as: SNP call rate > 90% and Hardy–Weinberg equilibrium $p > 1 \times 10^{-4}$.

Meta-analysis of replication and discovery studies. In the replication stage, we meta-analyzed results from: (1) individuals of European descent only (Rep-EUR); and (2) all replication cohorts with multiple ethnicities (Rep-All). Likewise we meta-analyzed results from discovery cohorts and European-descent-only replication cohorts (“Combined EUR”) from discovery cohorts and all replication cohorts (“Combined All”). To investigate and account for potential heterogeneities in allelic effects between studies, we also performed “trans-ethnic meta-analysis” using MANTRA²⁵ in the replication sample that included all ethnic groups (“Rep-All”) and in the combined analysis of the discovery and all ethnic groups in the replication sample (“Combined All”).

A successful replication was considered if: (1) the association p -value in the cumulative-meta-analysis (Combined EUR) was genome-wide significant ($p < 5 \times 10^{-8}$) and less than the discovery meta-analysis p -value; or (2) the association p -value in the meta-analysis of replication-cohorts only (Rep-EUR) was less than $p = 0.0024$ (a Bonferroni-adjusted threshold at $p = 0.05/21$ since there were a total of 21 tests performed for whole body and appendicular lean mass in Rep-EUR cohorts during replication). Using the METAL package we also estimated I^2 to quantify heterogeneity and p -values to assess statistical significance for a total of eight associations that were replicated in the cumulative-meta-analysis (combined EUR, five SNPs for whole body and three for appendicular lean mass).

To estimate the phenotypic variance explained by the genotyped SNPs in the Framingham Heart Study (FHS), we used a restricted maximum likelihood model implemented in the GCTA (Genome-wide Complex Trait Analysis) tool package^{57, 58} and adjusted for the same set of covariates included in our GWAS.

Finally, we examined associations between all imputed SNPs in/near five genes (*THRH*, *GLYAT*, *GREM1*, *CNTF*, and *PRDM16* including 60 kb up and downstream of the gene) and lean mass, as these genes have been implicated to have associations with lean mass in previous association studies^{18–20}.

Annotation and enrichment analysis of regulatory elements. For coding variants, we predicted their function by PolyPhen-2. For all variants, we annotated potential regulatory functions of our replicated GWAS SNPs and loci based on experimental epigenetic evidence including DNase hypersensitive sites, histone modifications, and transcription factor-binding sites in human cell lines and tissues from the ENCODE Project and the Epigenetic Roadmap Project. We first selected SNPs in high LD ($r^2 \geq 0.8$) with GWAS lead SNPs based on the approach of Trynka et al.⁵⁹ We then identified potential enhancers and promoters in the GWAS loci (GWAS SNPs and SNPs in LD with the GWAS SNPs) across 127 healthy human tissues/normal cell lines available in the ENCODE Project and the Epigenetic Roadmap Project from the HaploReg4 web browser⁶⁰ using ChromHMM³¹. To evaluate whether replicated GWAS loci were enriched with regulatory elements in skeletal muscle tissue, we performed a hypergeometric test. Specifically we tested whether estimated tissue-specific promoters and enhancers in a GWAS locus were enriched in eight relevant skeletal muscle tissues/cell lines vs enrichment in non-skeletal muscle tissues (119 tissues/cell lines). The permutation with minimum p -value approach was performed to correct for multiple testing. Permutation p -values <0.05 were considered statistically significant. In addition, we also performed enrichment analyses in smooth muscle tissues/cells, fat tissue, brain, blood cells and gastrointestinal tract tissues. The eight skeletal muscle relevant tissues/cells were excluded when conducting enrichment analyses for other tissue types. The detailed information for tissue types and chromatin state estimation is described in the Supplementary Materials.

cis-eQTL. We conducted cis-eQTL analyses on the five replicated GWS loci, SNPs rs2943656, rs9991501, rs2287926, rs4842924, and rs9936385, with gene expression within 2 Mb of the SNP position. A linear regression model was applied to examine associations between SNP and gene expression. The eQTL analyses were performed in five studies with available human skeletal muscle tissues, including: GTEX⁶¹, STRRIDE^{62, 63}, a study with chest wall muscle biopsies from patients who underwent thoracic surgery for lung and cardiac diseases⁶⁴, the Finland-United States Investigation of NIDDM Genetics (FUSION) Study⁶⁵, and a study of Pima Indians⁶⁶. In addition, eQTL analyses were also conducted in studies with other human tissues, including subcutaneous adipose^{67, 68}, omental adipose^{67, 68}, liver tissue⁶⁷, lymphocytes⁶⁹, and primary osteoblasts⁷⁰ (obtained from bone biopsies). These five GWAS SNPs were either genotyped or imputed in each sample. The detailed methods are described in the Supplementary Materials. Multiple testing was corrected by using false discovery rate (FDR q -value <0.05) to account for all pairs of SNP-gene expression analyses in multiple tissues and studies.

Data availability. All relevant data are available from the authors and summary level results are available on dbGaP.

Received: 17 August 2015 Accepted: 2 May 2017

Published online: 19 July 2017

References

- Janssen, I. Influence of sarcopenia on the development of physical disability: the Cardiovascular Health Study. *J. Am. Geriatr. Soc.* **54**, 56–62 (2006).
- Janssen, I., Heymsfield, S. B. & Ross, R. Low relative skeletal muscle mass (sarcopenia) in older persons is associated with functional impairment and physical disability. *J. Am. Geriatr. Soc.* **50**, 889–896 (2002).
- Visser, M. *et al.* Leg muscle mass and composition in relation to lower extremity performance in men and women aged 70 to 79: the health, aging and body composition study. *J. Am. Geriatr. Soc.* **50**, 897–904 (2002).
- Studenski, S. A. *et al.* The FNIH sarcopenia project: rationale, study description, conference recommendations, and final estimates. *J. Gerontol. A Biol. Sci. Med. Sci.* **69**, 547–558 (2014).
- Houston, D. K. *et al.* Dietary protein intake is associated with lean mass change in older, community-dwelling adults: the Health, Aging, and Body Composition (Health ABC) Study. *Am. J. Clin. Nutr.* **87**, 150–155 (2008).
- Lillioja, S. & Bogardus, C. Obesity and insulin resistance: lessons learned from the Pima Indians. *Diabetes Metab. Rev.* **4**, 517–540 (1988).
- Devaney, J. M. *et al.* Differences in fat and muscle mass associated with a functional human polymorphism in a post-transcriptional BMP2 gene regulatory element. *J. Cell Biochem.* **107**, 1073–1082 (2009).
- Janssen, I., Baumgartner, R. N., Ross, R., Rosenberg, I. H. & Roubenoff, R. Skeletal muscle cutpoints associated with elevated physical disability risk in older men and women. *Am. J. Epidemiol.* **159**, 413–421 (2004).
- Janssen, I., Shepard, D. S., Katzmarzyk, P. T. & Roubenoff, R. The healthcare costs of sarcopenia in the United States. *J. Am. Geriatr. Soc.* **52**, 80–85 (2004).
- Arden, N. K. & Spector, T. D. Genetic influences on muscle strength, lean body mass, and bone mineral density: a twin study. *J. Bone Miner. Res.* **12**, 2076–2081 (1997).
- Hsu, F. C. *et al.* Heritability of body composition measured by DXA in the diabetes heart study. *Obes. Res.* **13**, 312–319 (2005).
- Berndt, S. I. *et al.* Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat. Genet.* **45**, 501–512 (2013).
- Kilpelainen, T. O. *et al.* Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nat. Genet.* **43**, 753–760 (2011).
- Monda, K. L. *et al.* A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. *Nat. Genet.* **45**, 690–696 (2013).
- Randall, J. C. *et al.* Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet.* **9**, e1003500 (2013).
- Chagnon, Y. C. *et al.* Genome-wide search for genes related to the fat-free body mass in the Quebec family study. *Metabolism* **49**, 203–207 (2000).
- Livshits, G., Kato, B. S., Wilson, S. G. & Spector, T. D. Linkage of genes to total lean body mass in normal women. *J. Clin. Endocrinol. Metab.* **92**, 3171–3176 (2007).
- Liu, X. G. *et al.* Genome-wide association and replication studies identified TRHR as an important gene for lean body mass. *Am. J. Hum. Genet.* **84**, 418–423 (2009).
- 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).
- Hai, R. *et al.* Genome-wide association study of copy number variation identified gremlin1 as a candidate gene for lean body mass. *J. Hum. Genet.* **57**, 33–37 (2012).
- Urano, T., Shiraki, M., Sasaki, N., Ouchi, Y. & Inoue, S. Large-scale analysis reveals a functional single-nucleotide polymorphism in the 5'-flanking region of PRDM16 gene associated with lean body mass. *Aging Cell* **13**, 739–743 (2014).
- Lee, R. C., Wang, Z. M. & Heymsfield, S. B. Skeletal muscle mass and aging: regional and whole-body measurement methods. *Can. J. Appl. Physiol.* **26**, 102–122 (2001).
- Heymsfield, S. B., Gallagher, D., Visser, M., Nunez, C. & Wang, Z. M. Measurement of skeletal muscle: laboratory and epidemiological methods. *J. Gerontol. A Biol. Sci. Med. Sci.* **50**, 23–29 (1995).
- Zillikens, M. C. *et al.* Sex-specific genetic effects influence variation in body composition. *Diabetologia* **51**, 2233–2241 (2008).
- Morris, A. P. Transethnic meta-analysis of genomewide association studies. *Genet. Epidemiol.* **35**, 809–822 (2011).
- Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
- Lindgren, C. M. *et al.* Genome-wide association scan meta-analysis identifies three loci influencing adiposity and fat distribution. *PLoS Genet.* **5**, e1000508 (2009).
- Winkler, T. W. *et al.* The influence of age and sex on genetic associations with adult body size and shape: a large-scale genome-wide interaction study. *PLoS Genet.* **11**, e1005378 (2015).
- Wood, A. R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).
- Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
- Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods* **9**, 215–216 (2012).
- Touzet, H. & Varre, J. S. Efficient and accurate P-value computation for position weight matrices. *Algorithms Mol. Biol.* **2**, 15 (2007).
- Smemo, S. *et al.* Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507**, 371–375 (2014).
- Claussnitzer, M. *et al.* FTO obesity variant circuitry and adipocyte browning in humans. *N. Engl. J. Med.* **373**, 895–907 (2015).
- N'Diaye, A. *et al.* Identification, replication, and fine-mapping of loci associated with adult height in individuals of African ancestry. *PLoS Genet.* **7**, e1002298 (2011).
- Handsfield, G. G., Meyer, C. H., Hart, J. M., Abel, M. F. & Blemker, S. S. Relationships of 35 lower limb muscles to height and body mass quantified using MRI. *J. Biomech.* **47**, 631–638 (2014).
- Pickrell, J. K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am. J. Hum. Genet.* **94**, 559–573 (2014).
- Choocheep, K. *et al.* Versican facilitates chondrocyte differentiation and regulates joint morphogenesis. *J. Biol. Chem.* **285**, 21114–21125 (2010).
- Gao, X. *et al.* The fat mass and obesity associated gene FTO functions in the brain to regulate postnatal growth in mice. *PLoS ONE* **5**, e14005 (2010).
- Hinney, A. *et al.* Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (FTO) variants. *PLoS ONE* **2**, e1361 (2007).
- Yang, J. *et al.* FTO genotype is associated with phenotypic variability of body mass index. *Nature* **490**, 267–272 (2012).
- Elks, C. E. *et al.* Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat. Genet.* **42**, 1077–1085 (2010).
- Livshits, G., Malkin, I., Moayyeri, A., Spector, T. D. & Hammond, C. J. Association of FTO gene variants with body composition in UK twins. *Ann. Hum. Genet.* **76**, 333–341 (2012).
- Sonestedt, E. *et al.* Association between fat intake, physical activity and mortality depending on genetic variation in FTO. *Int. J. Obes.* **35**, 1041–1049 (2011).
- Fischer, J. *et al.* Inactivation of the Fto gene protects from obesity. *Nature* **458**, 894–898 (2009).
- Manning, A. K. *et al.* A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat. Genet.* **44**, 659–669 (2012).
- Sharma, R. *et al.* The type 2 diabetes and insulin-resistance locus near IRS1 is a determinant of HDL cholesterol and triglycerides levels among diabetic subjects. *Atherosclerosis* **216**, 157–160 (2011).
- 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).
- Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
- Schouboe, K. *et al.* Twin study of genetic and environmental influences on adult body size, shape, and composition. *Int. J. Obes. Relat. Metab. Disord.* **28**, 39–48 (2004).
- Chen, Z. *et al.* Dual-energy X-ray absorptiometry is a valid tool for assessing skeletal muscle mass in older women. *J. Nutr.* **137**, 2775–2780 (2007).
- Kyle, U. G., Genton, L., Hans, D. & Pichard, C. Validation of a bioelectrical impedance analysis equation to predict appendicular skeletal muscle mass (ASMM). *Clin. Nutr.* **22**, 537–543 (2003).
- Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
- Li, Y. & Abecasis, G. R. Mach 1.0: rapid haplotype reconstruction and missing genotype inference. *Am. J. Hum. Genet.* **579**, 2290 (2006).

55. Aulchenko, Y. S., Ripke, S., Isaacs, A. & van Duijn, C. M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–1296 (2007).
56. Chen, M. H. & Yang, Q. GWAf: an R package for genome-wide association analyses with family data. *Bioinformatics* **26**, 580–581 (2010).
57. Yang, B., Sun, H. & Wang, H. The downstream effects of vitamin D in spermatozoa needs further study. *Hum. Reprod.* **25**, 2152–2153 (2010); author reply 3.
58. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
59. Trynka, G. *et al.* Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat. Genet.* **45**, 124–130 (2013).
60. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **40**, D930–D934 (2012).
61. Consortium, G. T. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
62. Kraus, W. E. *et al.* Studies of a targeted risk reduction intervention through defined exercise (STRRIDE). *Med. Sci. Sports Exerc.* **33**, 1774–1784 (2001).
63. Slentz, C. A. *et al.* Effects of aerobic vs. resistance training on visceral and liver fat stores, liver enzymes, and insulin resistance by HOMA in overweight adults from STRRIDE AT/RT. *Am. J. Physiol. Endocrinol. Metab.* **301**, E1033–E1039 (2011).
64. Schadt E. A data driven approach to diagnosing and treating disease. In *Proc. 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 3 (ACM, 2014).
65. Valle, T. *et al.* Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. *Diabetes Care* **21**, 949–958 (1998).
66. Hanson, R. L. *et al.* A genome-wide association study in American Indians implicates DNER as a susceptibility locus for type 2 diabetes. *Diabetes* **63**, 369–376 (2014).
67. Schadt, E. E. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6**, e107 (2008).
68. Greenawalt, D. M. *et al.* A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. *Genome Res.* **21**, 1008–1016 (2011).
69. Dixon, A. L. *et al.* A genome-wide association study of global gene expression. *Nat. Genet.* **39**, 1202–1207 (2007).
70. Grundberg, E. *et al.* Population genomics in a disease targeted primary cell model. *Genome Res.* **19**, 1942–1952 (2009).

Acknowledgements

We acknowledge the essential role of the Cohorts for Heart and Aging Research in Genome Epidemiology (CHARGE) Consortium in development and support of this manuscript. CHARGE members include the Netherland's Rotterdam Study (RS), Framingham Heart Study (FHS), Cardiovascular Health Study (CHS), the NHLBI's Atherosclerosis Risk in Communities (ARIC) Study, and Iceland's Age, Gene/Environment Susceptibility (AGES) Reykjavik Study. Age, Gene/Environment Susceptibility Reykjavik Study (AGES-Reykjavik): has been funded by NIH contract N01-AG-12100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, (VSN: 00-063) and the Data Protection Authority. The researchers are indebted to the participants for their willingness to participate in the study. Old Order Amish (OOA): this work was supported by NIH research grants U01 HL72515, U01 GM074518, R01 HL088119, R01 AR046838, and U01 HL084756. Partial funding was also provided by the Mid-Atlantic Nutrition and Obesity Research Center of Maryland (P30 DK072488). L.M.Y.-A. was supported by F32AR059469 from NIH/NIAMS. M.F. was supported by American Heart Association grant 10SDG2690004. Cardiovascular Health Study (CHS): This CHS research was supported by NHLBI contracts N01-HC- 85079, N01-HC-85080, N01-HC-85081, N01-HC-85082, N01-HC-85083, N01-HC-85084, N01-HC-85085, N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, N01-HC-85239, and by HHSN268201200036C and NHLBI grants HL080295, HL087652, HL105756, HL103612, HL120393, and HL130114 with additional contribution from NINDS. Additional support was provided through AG-023629, AG-15928, AG-20098, and AG-027058 from the NIA. See also <http://www.chs-nhlbi.org/pi.htm>. DNA handling and genotyping at Cedars-Sinai Medical Center was supported in part by the National Center for Research Resources, grant UL1RR033176, and is now at the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124; in addition to the National Institute of Diabetes and Digestive and Kidney Disease grant DK063491 to the Southern California Diabetes Endocrinology Research Center. CoLaus: The CoLaus study received financial contributions from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 33CS0-122661, 33CS30-139468, and 33CS30-148401). We thank Vincent Mooser and Gérard Waeber, Co-PIs of the CoLaus study. Special thanks to Yolande Barreau, Mathieu

Firmann, Vladimir Mayor, Anne-Lise Bastian, Binasa Ramic, Martine Moranville, Martine Baumer, Marcy Sagette, Jeanne Ecoffey, and Sylvie Mermoud for data collection. Data analysis was supervised by Sven Bergmann and Jacques S. Beckmann. The computations for this paper were performed in part at the Vital-IT Center for high-performance computing of the Swiss Institute of Bioinformatics. deCODE Study: The study was funded by deCODE Genetics, ehf. We thank all the participants of this study, the staff of deCODE Genetics core facilities and recruitment center and the densitometry clinic at the University Hospital for their important contributions to this work. The EPIC Study: The EPIC Obesity study is funded by Cancer Research United Kingdom and the Medical Research Council. I.B. acknowledges support from EU FP6 funding (contract no. LSHM-CT-2003-503041) and by the Wellcome Trust (WT098051). Erasmus Rucphen Family (ERF) Study: The study was supported by grants from The Netherlands Organisation for Scientific Research (NWO), Erasmus MC, the Centre for Medical Systems Biology (CMSB), and the European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE Consortium, grant agreement HEALTH-F4-2007-201413. We are grateful to all general practitioners for their contributions, to Petra Veraart for her help in genealogy, Jeannette Vergeer for the supervision of the laboratory work and Peter Snijders for his help in data collection. Fenland: The Fenland Study is funded by the Wellcome Trust and the Medical Research Council, as well as by the Support for Science Funding programme and CamStrad. We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for help with recruitment. We thank the Fenland Study co-ordination team and the Field Epidemiology team of the MRC Epidemiology Unit for recruitment and clinical testing. Tuomas O. Kilpeläinen was supported by the Danish Council for Independent Research (DFF—1333-00124 and Sapere Aude program grant DFF—1331-00730B). Framingham Osteoporosis Study (FOS)/Framingham Heart Study (FHS): The study was funded by grants from the US National Institute for Arthritis, Musculoskeletal and Skin Diseases and National Institute on Aging (R01 AR 41398 and U24AG051129; D.P.K. and R01AR057118; D.K. D.K. was also supported by FP7-PEOPLE-2012-Marie Curie Career Integration Grants (CIG)). The Framingham Heart Study of the National Heart, Lung, and Blood Institute of the National Institutes of Health and Boston University School of Medicine were supported by the National Heart, Lung, and Blood Institute's Framingham Heart Study (N01-HC-25195) and its contract with Affymetrix, Inc. for genotyping services (N02-HL-6-4278). Analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. A portion of this research was conducted using the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. eQTL Hob Study: The study was supported by Genome Quebec, Genome Canada and the Canadian Institutes of Health Research (CIHR). Gothenburg Osteoporosis and Obesity Determinants Study (GOOD): The study was funded by the Swedish Research Council, the Swedish Foundation for Strategic Research, The ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Emil and Vera Cornell Foundation, the Torsten and Ragnar Söderberg's Foundation, Petrus and Augusta Hedlunds Foundation, the Västra Götaland Foundation, and the Göteborg Medical Society. We would like to thank Dr Tobias A. Knoch, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technology under grants #01 AK 803 A-H and # 01 IG 07015G for access to their grid resources. We also thank Karol Estrada, Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands for advice regarding the grid resources. Health Aging and Body Composition Study (Health ABC): This study was funded by the National Institutes of Aging. This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. Indiana: We thank the individuals who participated in this study, as well as the study coordinators, without whom this work would not have been possible. This work was supported by National Institutes of Health grants R01 AG 041517 and M01 RR-00750. Genotyping services were provided by CIDR. CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Library of Medicine. KORA (KORA F3 and KORA F4): The KORA research platform was initiated and financed by the Helmholtz Center Munich, German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Part of this work was financed by the German National Genome Research Network (NGFN-2 and NGFNPlus: 01GS0823). Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ. The London Life Sciences Population (LOLIPOP): The study was funded by the British Heart Foundation, Wellcome Trust, the Medical Research Council, and Kidney Research UK. The study also receives support from a National Institute for Health Research (NIHR) programme grant. Rotterdam Study (RSI, RSII & RSIII): The generation and management of GWAS genotype data for the Rotterdam Study (RS I, RS II, RS III) was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The

Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research NWO Investments (no. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project no. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera, Marjolein Peters, MSc, and Carolina Medina-Gomez, MSc, for their help in creating the GWAS database, and Karol Estrada, PhD, Yurii Aulchenko, PhD, and Carolina Medina-Gomez, PhD, for the creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. We are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We thank Dr Karol Estrada, Dr Fernando Rivadeneira, Dr Tobias A. Knoch, Anis Abuseiris, and Rob de Graaf (Erasmus MC Rotterdam, The Netherlands) for their help in creating GRIMP, and we thank BigGRID, MediGRID, and Services@MediGRID/D-Grid (funded by the German Bundesministerium fuer Forschung und Technology; grants 01 AK 803 A-H, 01 IG 07015G) for access to their grid computing resources. Rush Memory and Aging Project (MAP): The Memory and Aging Project was supported by National Institute on Aging grants R01AG17917, R01AG15819, and R01AG24480, the Illinois Department of Public Health, the Rush Clinical Translational Science Consortium, and a gift from Ms Marsha Dowd. TwinsUK (TUK): The study was funded by the Wellcome Trust, Arthritis Research UK, and the Chronic Disease Research Foundation. The study also received support from a National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. We thank the staff and volunteers of the TwinsUK study. The study was also supported by Israel Science Foundation, grant number 994/10. Age, Gene/Environment Susceptibility Reykjavik Study (AGES-Reykjavik) has been funded by NIH contract N01-AG-12100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee (VSN: 00-063) and the Data Protection Authority. The researchers are indebted to the participants for their willingness to participate in the study. Berlin Aging Study II (BASE-II) was supported by the German Federal Ministry of Education and Research (BMBF (grants #16SV5536K, #16SV5537, #16SV5538, and #16SV5837; previously #01UW0808)). Additional contributions (e.g., financial, equipment, logistics, personnel) are made from each of the other participating sites, i.e., the Max Planck Institute for Human Development (MPIB), Max Planck Institute for Molecular Genetics (MPIMG), Charite University Medicine, German Institute for Economic Research (DIW), all located in Berlin, Germany, and University of Lübeck in Lübeck, Germany. B-vitamins in the prevention of osteoporotic fractures (B-PROOF): B-PROOF is supported and funded by The Netherlands Organization for Health Research and Development (ZonMw, grant 6130.0031), the Hague; unrestricted grant from NZO (Dutch Dairy Association), Zoetermeer; Orthica, Almere; NCHA (Netherlands Consortium Healthy Ageing) Leiden/Rotterdam; Ministry of Economic Affairs, Agriculture and Innovation (project KB-15-004-003), the Hague; Wageningen University, Wageningen; VU University Medical Center, Amsterdam; Erasmus Medical Center, Rotterdam. All organizations are based in the Netherlands. We thank Dr Tobias A. Knoch, Anis Abuseiris, Karol Estrada, and Rob de Graaf as well as their institutions the Erasmus Grid Office, Erasmus MC Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technology (grants #01 AK 803 A-H and #01 IG 07015G) for access to their grid resources. Further, we gratefully thank all participants. Calcium Intake Fracture Outcome Study (CAIFOS): This study was funded by Healthway Health Promotion Foundation of Western Australia, Australasian Menopause Society and the Australian National Health and Medical Research Council Project Grants (254627, 303169, and 572604). We are grateful to the participants of the CAIFOS Study. The salary of Dr Lewis is supported by a National Health and Medical Research Council of Australia Career Development Fellowship. Danish Osteoporosis Study (DOPS): The study was supported by Karen Elise Jensen foundation. Family Heart Study (FamHS): The study was supported by NIH grants R01-HL-117078, R01-HL-087700, and R01-HL-088215 from NHLBI; and R01-DK-089256 and R01-DK-075681 from NIDDK. GenMets (Health 2000): S.R. was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (213506 and 129680), Academy of Finland (251217), the Finnish foundation for Cardiovascular Research and the Sigrid Juselius Foundation. S.M. was supported by grants #136895 and #141005, V.S. by grants #139635 and 129494, and M.P. by grant #269517 from the Academy of Finland and a grant from the Finnish Foundation for Cardiovascular Research. M.P. was supported by the Yrjö Jahnsson Foundation. Helsinki Birth Cohort Study (HBCS): We thank all study participants as well as everybody involved in the HBCS. HBCS has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Samfundet Folkhälsan, Novo Nordisk Foundation, Liv och Hälsa, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation, University of Helsinki, European Science Foundation (EUROSTRESS), Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation, Juho Vainio Foundation, and Wellcome Trust (grant number WT089062). Johnston County Study: The Johnston County Osteoarthritis Project is supported in part by cooperative

agreements S043, S1734, and S3486 from the Centers for Disease Control and Prevention/Association of Schools of Public Health; the NIAMS Multipurpose Arthritis and Musculoskeletal Disease Center grant 5-P60-AR30701; and the NIAMS Multidisciplinary Clinical Research Center grant 5 P60 AR49465-03. Genotyping services were provided by Algenomics company. Korean Genome Epidemiology Study (KoGES): Korean Genome Epidemiology Study (KoGES): This work was supported by the Research Program funded by the Korea Centers for Disease Control and Prevention (found 2001-347-6111-221, 2002-347-6111-221, 2009-E71007-00, 2010-E71004-00). Kora F3 and Kora F4: The KORA research platform was initiated and financed by the Helmholtz Center Munich, German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Part of this work was financed by the German National Genome Research Network (NGFN-2 and NGFNPlus: 01GS0823). Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ. LOLIP-REP-IA610: The study was supported by the Wellcome Trust. We thank the participants and research teams involved in LOLIPOP. LOLIP-REP-IA_I: The study was supported by the British Heart Foundation Grant SP/04/002. LOLIP-REP-IA_P: The study was supported by the British Heart Foundation Grant SP/04/002. METSIM: The study was supported by the Academy of Finland, the Finnish Diabetes Research Foundation, the Finnish Cardiovascular Research Foundation, the Strategic Research Funding from the University of Eastern Finland, Kuopio, and the EVO grant 5263 from the Kuopio University Hospital. MrOS Sweden: Financial support was received from the Swedish Research Council (2006-3832), the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation, Petrus and Augusta Hedlunds Foundation, the Västra Götaland Foundation, the Göteborg Medical Society, and the Novo Nordisk foundation. Greta and Johan Kock Foundation, A. Pahlsson Foundation, A. Osterlund Foundation, Malmö University Hospital Research Foundation, Research and Development Council of Region Skåne, Sweden, the Swedish Medical Society. MrOS US: The Osteoporotic Fractures in Men (MrOS) Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute on Aging (NIA), the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Center for Advancing Translational Sciences (NCATS), and NIH Roadmap for Medical Research under the following grant numbers: U01 AG027810, U01 AG042124, U01 AG042139, U01 AG042140, U01 AG042143, U01 AG042145, U01 AG042168, U01 AR066160, and U01 TR000128. The National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) provided funding for the MrOS ancillary study "GWAS in MrOS and SOF" under the grant number RC2AR058973. Osteoporosis Prospective Risk Assessment study (OPRA): This work was supported by grants from the Swedish Research Council (K2009-53X-14691-07-3, K2010-77PK-21362-01-2), FAS (grant 2007-2125), Greta and Johan Kock Foundation, A. Pahlsson Foundation, A. Osterlund Foundation, Malmö University Hospital Research Foundation, Research and Development Council of Region Skåne, Sweden, the Swedish Medical Society. We are thankful to all the women who kindly participated in the study and to the staff at the Clinical and Molecular Osteoporosis Research Unit for helping in recruitment of study individuals. Orkney Complex Disease Study (ORCADES): ORCADES was supported by the Chief Scientist Office of the Scottish Government (CZB/4/276, CZB/4/710), the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK (17539) and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney. PEAK 25: This work was supported by grants from the Swedish Research Council (K2009-53X-14691-07-3, K2010-77PK-21362-01-2), FAS (grant 2007-2125), Greta and Johan Kock Foundation, A. Pahlsson Foundation, A. Osterlund Foundation, Malmö University Hospital Research Foundation, Research and Development Council of Region Skåne, Sweden, the Swedish Medical Society. We are thankful to all the women who kindly participated in the study and to the staff at the Clinical and Molecular Osteoporosis Research Unit for helping in recruitment of study individuals. Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS): The study was supported by grants from the Swedish research council (projects 2008-2202 and 2005-8214) and ALF/LUA research grants from Uppsala university hospital, Uppsala, Sweden. Relationship between Insulin Sensitivity and Cardiovascular Disease (RISC): The RISC study is supported by European Union Grant QLGI-CT-2001-01252 and AstraZeneca. We thank Merck Research Labs for conducting DNA genotyping on RISC samples. Rotterdam III: Rotterdam Study (RS): See discovery. SHIP and SHIP TREND: This work was supported by SHIP, which is part of the Community Medicine Research Network of the University of Greifswald, Germany, by the Federal Ministry of Education and Research (01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania and the network "Greifswald Approach to Individualized Medicine (GANI_MED)" funded by the Federal Ministry of Education and Research (03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the "Center of Knowledge Interchange" program of the Siemens. A.G. and the Cache' Campus program of the InterSystems GmbH. The SHIP authors are grateful to the contribution of Florian Ernst, Anja Wiechert, and Astrid Petersmann in generating the SNP data and to Mario Stanke for the opportunity to use his Server Cluster for SNP

Imputation. Data analyses were further supported by the German Research Foundation (DFG Vo 955/10-1) and the Federal Ministry of Nutrition, Agriculture and Consumer's Safety. SOF: The Study of Osteoporotic Fractures (SOF) is supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576. The National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) provided funding for the SOF ancillary study "GWAS in MrOS and SOF" under the grant number RC2AR058973. Uppsala Longitudinal Study of Adult Men (ULSAM): The study was funded by grants from the Swedish research council (projects 2008-2202 and 2005-8214), the Wallenberg foundation, and ALF/LUA research grants from Uppsala university hospital, Uppsala, Sweden. Andrew P. Morris is a Wellcome Trust Senior Fellow in Basic Biomedical Science, grant number WT098017. CROATIA-VIS (VIS): The CROATIA-Vis study was funded by grants from the Medical Research Council (UK) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). We acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, the Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health. The SNP genotyping for the CROATIA-Vis cohort was performed in the core genotyping laboratory of the Wellcome Trust Clinical Research Facility at the Western General Hospital, Edinburgh, Scotland. Women's Health Initiative (WHI): The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through contracts N01WH22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-32, and 44221. We thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A listing of WHI investigators can be found at <https://www.whi.org/researchers/Documents%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf>. FUSION: This research was supported in part by US National Institutes of Health grants 1-ZIA-HG000024 (to F.S.C.), U01DK062370 (to M.B.), R00DK099240 (to S.C.J.P.), the American Diabetes Association Pathway to Stop Diabetes Grant 1-14-INI-07 (to S.C.J.P.), and Academy of Finland Grants 271961 and 272741 (to M.L.) and 258753 (to H.A.K.). We thank all the subjects for participation and the study personnel for excellent technical assistance. The Pima Indian Study: This study was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH, USA. Studies of a Targeted Risk Reduction Intervention with Defined Exercise (STRRIDE): This study was supported by the National Heart Lung and Blood Institute of the National Institutes of Health, HL57453 (WEK). Gene expression in old and young muscle biopsies: S.M. and T.G. were supported in part by NIH U24AG051129.

Author contributions

Writing group: (Amish) L.M.Y.-A.; (FHS) S.D., Y.-H.H., D.K., D.P.K.; (GOOD) C.O.; (HABC) T.B.H.; (Rotterdam Study II) M.C.Z. Individual study design: (AGES) G.E., M.G., V.G., L.J.L.; (Amish) E.A.S.; (Berlin Aging Study II) L.B., I.D., E.S.-T.; (B-PROOF) R.A.M.D.-R., N.M.v.S.; (CAIFOS) R.L.P.; (CHS) N.L.G., B.M.P.; (CoLaus) P.V., D.W.; (deCODE) G.S., K.S.; (DOPS) B.L.L., L.M.; (EPIC) K.-T.K., R.J.F.L., R.N.L.; (ERF) B.A.O., C.M.v.D.; (Fenland) T.O.K., R.J.F.L., N.J.W.; (FHS) D.K.; (FUSION) M.B., F.S.C., M.R.E., H.A.K., M.L., S.C.J.P., L.J.S., J.T.; (Genmets—controls) A.J., M.P., S.R.; (HABC) S.R.C., T.B.H., S.K., Y.L., A.N., S.S.; (Helsinki Birth Cohort) A.P.; (Johnston County Osteoarthritis Project) L.D., S.S.; (KOGES) N.H.C., J.-Y.L.; (KORA F3/F4) A.D., T.M., J.S.R.; (Lolipop) J.G., J.S.K.; (MrOS Sweden) D.M., C.O.; (MrOS US) D.S.E., C.M.N., E.S. Orwoll; (Old and Young Muscle Bx) A.H., S.M., M.T.; (OPRA) K.A., F.E.Mc.G.; (ORCADES) H.C., S.H.R., J.F.W.; (PEAK 25) K.A., F.E.Mc.G.; (PIMA Indians) L.J.B., R.L.H., W.-C.H., V.M.O.; (PIVUS) E.L., K.M.; (RISC) W.X.; (Rotterdam Study III) C.M.-G.; (Rush Memory and Aging Project) D.A.B.; (SHIP 2/SHIP-TREND) S.S.; (SOF) S.R.C.; (STRRIDE) E.H., M.H., W.E.K.; (Twins UK) N.S.; (ULSAM) H.M., K.M.; (VIS) O.P., I.R. Data collection: (AGES) G.E., V.G., T.L.; (Amish) B.D.M., E.A.S.; (Berlin Aging Study II) I.D., E.S.-T.; (B-PROOF) A.W.E., K.M.A.S.; (CAIFOS) R.L.P.; (CHS) B.M.P., J.A.R., J.I.R.; (CoLaus) P.V.; (deCODE) G.S., U.S.; (DOPS) L.B.H., B.L.L., L.M.; (EPIC) K.-T.K.; (ERF) B.A.O., C.M.v.D.; (Fenland) N.J.W.; (FHS) D.K.; (FUSION) M.R.E., H.A.K., M.L.; (Genmets—controls) M.P., E.T.; (GOOD) J.-O.J., M.L., C.O.; (HABC) S.R.C., T.B.H., S.K., Y.L., A.N., S.S.; (Helsinki Birth Cohort) J.G.E., A.P., K.R.; (Johnston County Osteoarthritis Project) L.D., S.S.; (KOGES) N.H.C., J.-Y.L.; (KORA F3/KORA F4) T.M., J.S.R.; (Lolipop) J.C.C., J.G., J.S.K., J.S., S.-T.T., A.V.; (METSIM) A.S.; (MrOS Sweden) J.E., M.K., Ö.L., D.M., C.O.; (MrOS US) D.S.E., C.M.N., E.S.O.; (Old and Young Muscle Bx) A.H., S.M., M.T.; (OPRA) K.A., F.E.M.; (ORCADES) S.H.R.; (PEAK 25) K.A., F.E.M.; (PIMA Indians) L.J.B., R.L.H., W.-C.H., V.M.O.; (PIVUS) E.L., L.L., K.M.; (RISC) W.X.; (Rotterdam Study I) L.S.; (Rotterdam Study II) J.S.L.K., F.R., M.C.Z.; (Rotterdam Study III) C.M.-G.; (Rush Memory and Aging Project) D.A.B., L.Y.; (SHIP 2/SHIP-TREND) A.A., G.H., T.I., M.M.L., S.S.; (SOF) S.R.C.; (STRRIDE) E.H., K.M.H., M.H.; (Twins UK) N.S., T.D.S.; (ULSAM) H.M., K.M.; (VIS) C.H., O.P., I.R. Genotyping: (AGES) A.V.S.; (Berlin Aging Study II) L.B.; (CAIFOS) J.R.L., R.L.P.; (CHS) J.I.R.; (deCODE) U.T.; (EPIC) I.B., R.J.F.L., R.N.L., J.H.Z.; (ERF) L.B.; (FamHS) I.B.B., M.F.F.; (Fenland) R.J.F.L., J.L., N.J.W.; (FUSION) M.R.E.; (Genmets—controls) E.T.; (GOOD) J.-O., M.L.; (HABC) S.K.; (Helsinki Birth Cohort) J.L., K.R.; (Johnston County Osteoarthritis Project) L.D., Y.L., W.M.; (KOGES) H.J.C., B.-G.H.; (KORA F3/KORA F4) C.G., H.G., T.I., N.K.; (Lolipop) J.G., J.S.K.; (MrOS Sweden) D.M.; (MrOS US) E.S.O.;

(OPRA) K.A., F.E.M.; (PEAK 25) K.A., F.E.M.; (PIVUS) L.L., K.M.; (Rush Memory and Aging Project) A.S.B.; (SHIP 2/SHIP-TREND) R.B.; (Twins UK) I.M.; (ULSAM) H.M., K.M., A.M.; (VIS) I.R. Genotype preparation: (AGES) A.V.S.; (Amish) J.R.O'C., L.M.Y.-A.; (Berlin Aging Study II) L.B.; (CAIFOS) J.R.L., R.L.P.; (CHS) N.L.G., J.I.R.; (deCODE) G.S., U.S., U.T.; (DOPS) L.B.H.; (EPIC) R.N.L., J.H.Z.; (ERF) N.A., L.B., B.A.O., C.M.v.D.; (Fenland) J.L.; (FUSION) M.R.E.; (HABC) S.K.; (Helsinki Birth Cohort) K.R.; (Johnston County Osteoarthritis Project) L.D., J.M.J., Y.L., W.M.; (KOGES) H.J.C., B.-G.H.; (KORA F3/KORA F4) A.D., H.G., T.I.; (Lolipop) J.G., J.S.K.; (MrOS US) E.S.O.; (OPRA) K.A., F.E.M.; (PEAK 25) K.A., F.E.M.; (PIVUS) K.M.; (Rotterdam Study II) J.S.L.K.; (Rush Memory and Aging Project) A.S.B.; (SHIP 2/SHIP-TREND) R.B.; (SOF) S.R.C.; (Twins UK) I.M.; (ULSAM) K.M., A.M.; (VIS) I.R. Phenotype preparation: (AGES) M.G., A.V.S.; (Amish) E.A.S., L.M.Y.-A.; (Berlin Aging Study II) I.D., E.S.-T.; (B-PROOF) A.W.E., K.M.A.S.; (CAIFOS) J.R.L., R.L.P.; (CHS) N.L.G., J.A.R.; (CoLaus) P.V.; (deCODE) G.S., U.S., G.T.; (DOPS) B.L.L., L.M.; (EPIC) K.-T.K., R.J.F.L., R.N.L., J.H.Z.; (ERF) N.A.; (Fenland) T.O.K., R.J.F.L., J.L., N.J.W.; (FHS) D.K.; (FUSION) H.A.K., M.L., R.P.W.; (Genmets—controls) E.T.; (GOOD) J.-O.J., M.L., C.O.; (HABC) S.R.C., S.S.; (Helsinki Birth Cohort) J.G.E., A.P.; (Indiana) M.P.; (Johnston County Osteoarthritis Project) L.D., J.M.J., S.S.; (KOGES) N.H.C., J.-Y.L.; (Lolipop) J.G., J.S.K.; (METSIM) A.S.; (MrOS Sweden) J.E., M.K., Ö.L., D.M., C.O.; (MrOS US) D.S.E., E.S.O.; (Old and Young Muscle Bx) S.M., M.T.; (OPRA) K.A., F.E.Mc.G.; (ORCADES) S.H.R., J.F.W.; (PEAK 25) K.A., F.E.Mc.G.; (PIVUS) L.B., L.L., K.M.; (Rotterdam Study II) J.S.L.K., F.R., M.C.Z.; (Rush Memory and Aging Project) L.Y.; (SHIP 2/SHIP-TREND) G.H., M.M.L., H.V.; (Twins UK) N.S., T.D.S.; (ULSAM) H.M., K.M.; (VIS) C.H., O.P., I.R. Analysis plan development: (Amish) L.M.Y.-A.; (CHS) J.A.C., B.M.P.; (FHS) D.P.K., W.-C.C., S.D., Y.-H.H., D.K.; (FUSION) M.R.E., S.C.J.P., L.J.S.; (GOOD) M.L.; (Johnston County Osteoarthritis Project) S.S.; (KOGES) N.H.C., J.-Y.L.; (Lolipop) J.G., J.S.K.; (MrOS Sweden) D.M.; (MrOS US) J.S.; (Old and Young Muscle Bx) T.G., A.H., S.M.; (OPRA) K.A., F.E.Mc.G.; (ORCADES) H.C., S.H.R., J.F.W.; (PEAK 25) K.A., F.E.Mc.G.; (PIVUS) K.M.; (Rotterdam Study II) F.R., M.C.Z.; (SOF) S.R.C.; (STRRIDE) W.E.K.; (Twins UK) G.L.; (ULSAM) K.M. Analysis plan review: (AGES) A.V.S.; (Amish) L.M.Y.-A.; (CHS) J.A.C., N.L.G., J.I.R.; (CoLaus) M.B., T.J., D.W.; (EPIC) R.J.F.L.; (ERF) N.A., L.B., B.A.O., C.M.v.D.; (Fenland) T.O.K., R.J.F.L.; (FHS) W.-C.C., S.D., Y.-H.H., D.K.; (FUSION) M.B., F.S.C., M.L., J.T.; (GOOD) M.L.; (HABC) S.R.C.; (Indiana) M.J.E.; (Johnston County Osteoarthritis Project) L.D., J.M.J., S.S.; (KOGES) N.H.C., J.-Y.L.; (Lolipop) J.G., J.S.K.; (MrOS Sweden) D.M.; (MrOS US) P.M.C., D.S.E., J.S.; (Old and Young Muscle Bx) T.G., A.H., S.M.; (OPRA) K.A., F.E.Mc.G.; (ORCADES) H.C., S.H.R., J.F.W.; (PEAK 25) K.A., F.E.Mc.G.; (PIVUS) K.M.; (Rotterdam Study I) L.S.; (Rotterdam Study II) J.S.L.K., F.R., M.C.Z.; (SOF) S.R.C.; (VIS) I.R.; (WHI) J.W.-W. Study Data Analysis: (AGES) A.V.S.; (Amish) L.M.Y.-A.; (Berlin Aging Study II) L.B., T.L.; (B-PROOF) A.W.E., K.M.A.S.; (CAIFOS) J.R.L., R.L.P.; (CHS) N.L.G., J.I.R.; (CoLaus) M.B., Z.K., T.J.; (deCODE) U.S., G.T.; (EPIC) J.H.Z.; (ERF) N.A., L.B.; (FamHS) M.F.F.; (Fenland) J.L.; (FHS) W.-C.C., D.P.K., Y.Z.; (FUSION) R.P.W.; (Genmets—controls) V.S., E.T.; (GOOD) J.-O.J., M.L., C.O., L.V.; (HABC) S.R.C., S.K.; (Helsinki Birth Cohort) J.G.E.; (Indiana) M.J.E.; (Johnston County Osteoarthritis Project) J.M.J.; (KOGES) N.H.C.; (KORA F3/KORA F4) A.P.; (Lolipop) J.C.C., J.G., J.S.K., J.S., S.-T.T., A.V., W.Z.; (METSIM) H.C., T.K., J.K., M.L., A.S.; (MrOS Sweden) D.M.; (MrOS US) P.C., D.S.E., J.S.; (Old and Young Muscle Bx) T.G., A.H., S.M.; (OPRA) K.A., F.E.Mc.G.; (ORCADES) H.C., S.H.R., J.F.W.; (PEAK 25) K.A., F.E.Mc.G.; (PIVUS) K.M.; (RISC) M.W., M.N.W.; (Rotterdam Study III) N.C.-O., K.E.; (Rush Memory and Aging Project) P.L.D.J.; (SHIP 2/SHIP-TREND) G.H., M.M.L.; (SOF) G.J.T.; (STRRIDE) E.H., M.H., W.E.K.; (Twins UK) G.L.; (ULSAM) K.M.; (VIS) I.R. Reviewed/interpreted analyses: (AGES) A.V.S.; (Amish) M.F., B.D.M., J.R.O'C., E.A.S., L.M.Y.-A.; (Berlin Aging Study II) L.B., T.L.; (CAIFOS) J.R.L., R.L.P.; (CHS) B.M.P., J.I.R.; (CoLaus) M.B., Z.K., T.J.; (EPIC) R.J.F.L.; (Fenland) T.O.K., R.J.F.L.; (FHS) W.-C.C., S.D., Y.-H.H., D.K., D.P.K.; (FUSION) M.R.E., S.C.J.P., L.J.S.; (GOOD) M.L.; (HABC) S.R.C.; (Helsinki Birth Cohort) J.G.E., E.W.; (Indiana) M.J.E., D.L.K.; (Johnston County Osteoarthritis Project) L.D., J.M.J., S.S.; (KOGES) N.H.C., H.J.C., B.-G.H., J.-Y.L., C.S.S.; (Lolipop) J.G., J.S.K.; (METSIM) J.K., M.L.; (MrOS Sweden) D.M.; (MrOS US) D.S.E.; (OPRA) K.A., F.E.Mc.G.; (ORCADES) H.C., S.H.R., J.F.W.; (PEAK 25) K.A., F.E.Mc.G.; (PIVUS) L.B., E.I., L.L.; (Rotterdam Study I) A.H., A.G.U.; (Rotterdam Study II) J.S.L.K., F.R., M.C.Z.; (ULSAM) C.L., H.M., A.M.; (VIS) I.R.; (WHI) J.W.-W. Meta-analyses: (Amish) L.M.Y.-A.; (FHS) S.D. Pathway/other analyses: (FHS) W.-C.C., M.C., Y.-H.H. Lean mass study design: (AGES) A.V.S.; (Amish) L.M.Y.-A.; (Berlin Aging Study II) I.D., E.S.-T.; (FHS) Y.-H.H., D.K.; (GOOD) M.L.; (HABC) S.R.C.; (MrOS Sweden) D.M.; (OPRA) K.A., F.E.Mc.G.; (PEAK 25) K.A., F.E.Mc.G.; (PIVUS) K.M.; (Rotterdam Study II) F.R., M.C.Z.; (SHIP 2/SHIP-TREND) H.V.; (STRRIDE) W.E.K.; (Twins UK) N.S.; (ULSAM) K.M.; (VIS) I.R. Manuscript preparation: (Amish) L.M.Y.-A.; (FHS) S.D., Y.-H.H., D.K., D.P.K.; (GOOD) C.O.; (HABC) T.B.H.; (Rotterdam Study II) M.C.Z. Manuscript review: (AGES) G.E., M.G., V.G., T.L., L.J.L., A.V.S.; (Amish) M.F., B.D.M., J.R.O'C., E.A.S., L.M.Y.-A.; (Berlin Aging Study II) L.B.; (CAIFOS) J.R.L., R.L.P.; (CHS) J.A.C., N.L.G., B.M.P., J.A.R., J.I.R.; (CoLaus) M.B., Z.K., T.J., P.V., D.W.; (DOPS) L.B.H., B.L.L., L.M.; (EPIC) R.J.F.L.; (ERF) N.A., L.B., B.A.O., C.M.v.D.; (FamHS) M.F.F.; (Fenland) T.O.K., R.J.F.L.; (FHS) W.-C.C., M.C., S.D., Y.-H.H., D.K., Y.Z.; (FUSION) M.B., H.A.K., S.C.J.P., L.J.S.; (Genmets—controls) M.P., E.T.; (GOOD) J.-O.J., M.L., C.O., L.V.; (HABC) S.R.C., T.B.H., Y.L., A.N., S.S.; (Helsinki Birth Cohort) J.G.E., J.L., A.P., K.R., E.W.; (Indiana) M.J.E., D.L.K.; (KOGES) N.H.C., H.J.C., B.-G.H., J.-Y.L., C.S.S.; (KORA) H.E.W.; (Lolipop) W.Z.; (METSIM) H.C., T.K., J.K., M.L., C.S.S., A.S.; (MrOS Sweden) J.E., M.K., Ö.L., D.M., C.O.; (MrOS US) P.M.C., D.S.E., C.M.N., E.S.O.; (Old and Young Muscle Bx) T.G., A.H., S.M., M.T.; (OPRA) K.A., F.E.Mc.G.; (ORCADES)

H.C., S.H.R., J.W.; (PEAK 25) K.A., F.E.Mc.G.; (PIMA Indians) L.J.B., R.L.H., W.-C.H., V.M.O.; (PIVUS) L.B., E.I., L.L., K.M.; (RISC) M.W., M.N.W.; (Rotterdam Study I) A.H., A.G.U.; (Rotterdam Study II) J.S.L.K., F.R.; (Rotterdam Study III) N.C.-O., K.E., C.M.-G.; (Rush Memory and Aging Project) D.A.B., A.S.B., P.L.D.J.; (SHIP 2/SHIP-TREND) H.V.; (SOF) S.R.C., G.J.T.; (STRIDE) E.P.H., M.H., K.M.H., W.E.K.; (Twins UK) N.S., T.D.S., F.M.K.W.; (ULSAM) C.L., H.M., K.M., A.M.; (VIS) I.R.; (WHI) Z.C., P.T., C.A.T., J.W.-W.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-00031-7.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017

M. Carola Zillikens^{1,2}, Serkalem Demissie³, Yi-Hsiang Hsu^{4,5,6}, Laura M. Yerges-Armstrong⁷, Wen-Chi Chou^{4,5,8}, Lisette Stolk^{1,2}, Gregory Livshits^{9,10}, Linda Broer¹¹, Toby Johnson^{12,13,14}, Daniel L. Koller¹⁵, Zoltán Kutalik^{12,13,14}, Jian'an Luan¹⁶, Ida Malkin⁹, Janina S. Ried¹⁷, Albert V. Smith^{18,19}, Gudmar Thorleifsson²⁰, Liesbeth Vandenput²¹, Jing Hua Zhao¹⁶, Weihua Zhang^{22,23}, Ali Aghdassi²⁴, Kristina Åkesson^{25,26}, Najaf Amin¹¹, Leslie J. Baier²⁷, Inês Barroso^{28,29,30}, David A. Bennett³¹, Lars Bertram^{32,33}, Rainer Biffar³⁴, Murielle Bochud¹⁴, Michael Boehnke³⁵, Ingrid B. Borecki^{36,37}, Aron S. Buchman³¹, Liisa Byberg³⁸, Harry Campbell³⁹, Natalia Campos Obanda¹, Jane A. Cauley⁴⁰, Peggy M. Cawthon⁴¹, Henna Cederberg⁴², Zhao Chen⁴³, Nam H. Cho⁴⁴, Hyung Jin Choi^{45,46}, Melina Claussnitzer^{4,5,8,47,48,49}, Francis Collins⁵⁰, Steven R. Cummings⁴¹, Philip L. De Jager^{5,51,52}, Ilja Demuth^{53,54}, Rosalie A. M. Dhonukshe-Rutten⁵⁵, Luda Diatchenko^{56,57}, Gudny Eiriksdottir¹⁸, Anke W. Enneman¹, Mike Erdos⁵⁰, Johan G. Eriksson^{58,59,60,61,62}, Joel Eriksson²¹, Karol Estrada^{1,11}, Daniel S. Evans⁴¹, Mary F. Feitosa³⁶, Mao Fu⁷, Melissa Garcia⁶³, Christian Gieger^{17,64,65}, Thomas Girke^{66,67}, Nicole L. Glazer⁶⁸, Harald Grallert^{17,67,64,69,70,71}, Jagvir Grewal^{23,72}, Bok-Ghee Han⁷³, Robert L. Hanson²⁷, Caroline Hayward⁷⁴, Albert Hofman^{2,11}, Eric P. Hoffman⁷⁵, Georg Homuth⁷⁶, Wen-Chi Hsueh²⁷, Monica J. Hubal^{77,78}, Alan Hubbard⁷⁹, Kim M. Huffman⁸⁰, Lise B. Husted⁸¹, Thomas Illig^{64,82,83}, Erik Ingelsson^{84,85}, Till Ittermann⁸⁶, John-Olov Jansson⁸⁷, Joanne M. Jordan⁸⁸, Antti Jula⁶², Magnus Karlsson⁸⁹, Kay-Tee Khaw⁹⁰, Tuomas O. Kilpeläinen^{16,91,92}, Norman Klopp^{64,83}, Jacqueline S. L. Kloth¹, Heikki A. Koistinen^{93,94,95,96}, William E. Kraus⁹⁷, Stephen Kritchevsky⁹⁸, Teemu Kuulasmaa⁴², Johanna Kuusisto⁴², Markku Laakso⁴², Jari Lahti⁹⁹, Thomas Lang¹⁰⁰, Bente L. Langdahl⁸¹, Lenore J. Launer⁶³, Jong-Young Lee⁷³, Markus M. Lerch²⁴, Joshua R. Lewis^{101,102}, Lars Lind⁸⁴, Cecilia Lindgren¹⁰³, Yongmei Liu¹⁰⁴, Tian Liu^{105,106}, Youfang Liu⁸⁸, Östen Ljunggren⁸⁴, Mattias Lorentzon²¹, Robert N. Luben⁹⁰, William Maixner⁵⁷, Fiona E. McGuigan²⁵, Carolina Medina-Gomez^{1,11}, Thomas Meitinger^{48,107}, Håkan Melhus⁸⁴, Dan Mellström²¹, Simon Melov^{108,109}, Karl Michaëlsson³⁸, Braxton D. Mitchell^{7,110}, Andrew P. Morris^{103,111}, Leif Mosekilde⁸¹, Anne Newman¹¹², Carrie M. Nielson¹¹³, Jeffrey R. O'Connell⁷, Ben A. Oostra^{114,115}, Eric S. Orwoll¹¹³, Aarno Palotie^{116,117}, Stephan Parker¹¹⁸, Munro Peacock¹¹⁹, Markus Perola^{62,116,120,121}, Annette Peters^{17,64}, Ozren Polasek¹²², Richard L. Prince^{101,123}, Katri Räikkönen⁹⁹, Stuart H. Ralston¹²⁴, Samuli Ripatti^{116,125,126}, John A. Robbins¹²⁷, Jerome I. Rotter¹²⁸, Igor Rudan³⁹, Veikko Salomaa⁶², Suzanne Satterfield¹²⁹, Eric E. Schadt¹³⁰, Sabine Schipf⁸⁶, Laura Scott³⁵, Joban Sehmi^{23,72}, Jian Shen¹¹³, Chan Soo Shin⁴⁵, Gunnar Sigurdsson^{19,131}, Shad Smith¹³², Nicole Soranzo¹²⁶, Alena Stančáková⁴², Elisabeth Steinhagen-Thiessen⁵³, Elizabeth A. Streeten^{7,133}, Unnur Styrkarsdottir²⁰, Karin M. A. Swart¹³⁴, Sian-Tsung Tan^{23,72}, Mark A. Tarnopolsky¹³⁵, Patricia Thompson¹³⁶, Cynthia A. Thomson⁴³, Unnur Thorsteinsdottir^{19,20},

Emmi Tikkanen^{62,116,124}, Gregory J. Tranah⁴¹, Jaakko Tuomilehto^{61,137,138,139}, Natasja M. van Schoor¹³⁴, Arjun Verma²³, Peter Vollenweider¹⁴⁰, Henry Völzke⁸⁶, Jean Wactawski-Wende¹⁴¹, Mark Walker¹⁴², Michael N. Weedon¹⁴³, Ryan Welch³⁵, H.-Erich Wichmann^{17,144,145}, Elisabeth Widen¹¹⁶, Frances M. K. Williams¹⁰, James F. Wilson^{39,74}, Nicole C. Wright¹⁴⁶, Weijia Xie¹⁴³, Lei Yu³¹, Yanhua Zhou³, John C. Chambers^{22,23,147,148}, Angela Döring^{17,149}, Cornelia M. van Duijn^{11,115}, Michael J. Econs¹⁵⁰, Vilmundur Gudnason^{18,19}, Jaspal S. Kooner^{23,72,148}, Bruce M. Psaty^{151,152}, Timothy D. Spector¹⁰, Kari Stefansson^{19,20}, Fernando Rivadeneira^{1,2,11}, André G. Uitterlinden^{1,2,11}, Nicholas J. Wareham¹⁶, Vicky Ossowski²⁷, Dawn Waterworth¹⁵³, Ruth J. F. Loos^{16,154,155,156,157}, David Karasik^{4,5,158}, Tamara B. Harris⁶³, Claes Ohlsson²¹ & Douglas P. Kiel^{4,5,49}

¹Department of Internal Medicine, Erasmus MC, Rotterdam 3000, The Netherlands. ²Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA), Leiden 2593, The Netherlands. ³Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA. ⁴Hebrew SeniorLife, Institute for Aging Research, Roslindale, MA 02131, USA. ⁵Harvard Medical School, Boston, MA 02115, USA. ⁶Molecular and Integrative Physiological Sciences Program, Harvard School of Public Health, Boston, MA 02115, USA. ⁷Program in Personalized and Genomic Medicine, and Department of Medicine, Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, MD 21201, USA. ⁸Broad Institute, Cambridge, MA 02142, USA. ⁹Sackler Faculty of Medicine, Department of Anatomy and Anthropology, Tel Aviv University, Tel Aviv 6997801, Israel. ¹⁰Department of Twin Research and Genetic Epidemiology, King's College London, St Thomas' Campus, London WC2R 2LS, UK. ¹¹Department of Epidemiology, Erasmus MC, Rotterdam 3000, The Netherlands. ¹²Department of Medical Genetics, University of Lausanne, Lausanne 1011, Switzerland. ¹³Swiss Institute of Bioinformatics, Lausanne 1015, Switzerland. ¹⁴Centre Hospitalier Universitaire (CHUV), University Institute for Social and Preventive Medicine, Lausanne 1010, Switzerland. ¹⁵Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202, USA. ¹⁶MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK. ¹⁷Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg 85764, Germany. ¹⁸Icelandic Heart Association, Kopavogur 201, Iceland. ¹⁹Faculty of Medicine, University of Iceland, Reykjavik 101, Iceland. ²⁰deCODE Genetics, Reykjavik 101, Iceland. ²¹Department of Internal Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg SE-405 30, Sweden. ²²Department Epidemiology and Biostatistics, School of Public Health, Imperial College, London SW7 2AZ, UK. ²³Cardiology Department, Ealing Hospital NHS Trust, Middlesex UB1 3HW, UK. ²⁴Department of Medicine A, University of Greifswald, Greifswald 17489, Germany. ²⁵Department of Clinical Sciences, Lund University, Malmö 22362, Sweden. ²⁶Department of Orthopedics, Skåne University Hospital, Malmö S-205 02, Sweden. ²⁷Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, NIH, Phoenix, AZ 85014, USA. ²⁸Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK. ²⁹NIHR Cambridge Biomedical Research Centre, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK. ³⁰Institute of Metabolic Science, Addenbrooke's Hospital, University of Cambridge Metabolic Research Laboratories, Cambridge CB2 0QQ, UK. ³¹Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL 60612, USA. ³²Lübeck Interdisciplinary Platform for Genome Analytics, Institutes of Neurogenetics and Experimental & Integrative Genomics, University of Lübeck, Lübeck 23562, Germany. ³³School of Public Health, Faculty of Medicine, Imperial College London, London W6 8RP, UK. ³⁴Centre of Oral Health, Department of Prosthetic Dentistry, Gerodontology and Biomaterials, University of Greifswald, Greifswald 17489, Germany. ³⁵Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109, USA. ³⁶Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St Louis, MO 63110, USA. ³⁷Division of Biostatistics, Washington University School of Medicine, St Louis, MO 63110, USA. ³⁸Department of Surgical Sciences, Uppsala University, Uppsala 75185, Sweden. ³⁹Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, Scotland EH8 9AG, UK. ⁴⁰Department of Epidemiology Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA. ⁴¹California Pacific Medical Center Research Institute, San Francisco, CA 94107, USA. ⁴²Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio 70210, Finland. ⁴³Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ 85714, USA. ⁴⁴Department of Preventive Medicine, Ajou University School of Medicine, Youngtong-Gu, Suwon 16499, Korea. ⁴⁵Department of Internal Medicine, Seoul National University College of Medicine, Seoul 03080, Korea. ⁴⁶Department of Internal Medicine, Chungbuk National University Hospital, Cheongju Si, Korea. ⁴⁷Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA 02139, USA. ⁴⁸Institute of Human Genetics, MRI, Technische Universität München, Munich 81675, Germany. ⁴⁹Beth Israel Deaconess Medical Center, Boston, MA 02215, USA. ⁵⁰Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, Bethesda, MD 20892, USA. ⁵¹Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Boston, MA 02115, USA. ⁵²Program in Medical and Population Genetics, Broad Institute, Cambridge, MA 02142, USA. ⁵³Lipid Clinic at the Interdisciplinary Metabolism Center, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin 13353, Germany. ⁵⁴Institute of Medical and Human Genetics, Charité - Universitätsmedizin Berlin, Berlin 13353, Germany. ⁵⁵Department of Human Nutrition, Wageningen University, PO Box 17, Wageningen 6700 AA, The Netherlands. ⁵⁶Alan Edwards Centre for Research on Pain, McGill University, Montreal, H3A 0G1, Canada. ⁵⁷Regional Center for Neurosensory Disorders, School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. ⁵⁸Department of General Practice and Primary Health Care, University of Helsinki, Helsinki 00014, Finland. ⁵⁹Unit of General Practice, Helsinki University Central Hospital, Helsinki 00014, Finland. ⁶⁰Folkhalsan Research Centre, Helsinki 00250, Finland. ⁶¹Vasa Central Hospital, Vasa 65130, Finland. ⁶²National Institute for Health and Welfare, Helsinki 00271, Finland. ⁶³Laboratory of Epidemiology and Population Sciences, Intramural Research Program, National Institute for Aging, Bethesda, MD 20892, USA. ⁶⁴Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg 85764, Germany. ⁶⁵Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg 85764, Germany. ⁶⁶Institute for Integrative Genome Biology, University of California, Riverside, CA 92521, USA. ⁶⁷Department of Botany and Plant Sciences, University of California, Riverside, CA 92521, USA. ⁶⁸Departments of Medicine and Epidemiology, Boston University School of Medicine and Public Health, Boston, MA 02118, USA. ⁶⁹German Center for Diabetes Research (DZD), Neuherberg, Germany. ⁷⁰CCG Type 2 Diabetes, Helmholtz Zentrum München, Neuherberg 85764, Germany. ⁷¹CCG Nutrigenomics and Type 2 Diabetes, Helmholtz Zentrum München, Neuherberg 85764, Germany.

Germany. ⁷²National Heart and Lung Institute, Imperial College London, London SW3 6LY, UK. ⁷³Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do 28159, Korea. ⁷⁴MRC Human Genetics Unit, IGMM, University of Edinburgh, Edinburgh, Scotland EH4 2XU, UK. ⁷⁵Department of Pharmaceutical Sciences, SUNY Binghamton, Binghamton, NY 13902, USA. ⁷⁶Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald 17487, Germany. ⁷⁷Department of Exercise and Nutrition Sciences, George Washington University, Washington, DC 20052, USA. ⁷⁸Research Center for Genetic Medicine, Children's National Medical Center, Washington, DC 20052, USA. ⁷⁹Division of Biostatistics, School of Public Health, University of California, Berkeley, CA 94720, USA. ⁸⁰Division of Rheumatology, Department of Medicine, Duke Molecular Physiology Institute, Duke University School of Medicine, Durham, NC 27710, USA. ⁸¹Endocrinology and Internal Medicine, Aarhus University Hospital, Aarhus DK 8000, Denmark. ⁸²Department of Human Genetics, Hannover Medical School, Hannover 30625, Germany. ⁸³Hannover Unified Biobank, Hannover Medical School, Hannover 30625, Germany. ⁸⁴Department of Medical Sciences, Uppsala University, Uppsala 75185, Sweden. ⁸⁵Division of Cardiovascular Medicine, Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA. ⁸⁶Institute for Community Medicine, University of Greifswald, Greifswald 17489, Germany. ⁸⁷Department of Physiology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg SE 405 30, Sweden. ⁸⁸Thurston Arthritis Research Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27517, USA. ⁸⁹Department of Clinical Sciences and Orthopaedics, Lund University, Skåne University Hospital SUS, Malmö 22362, Sweden. ⁹⁰Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. ⁹¹The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, University of Copenhagen, Copenhagen 2100, Denmark. ⁹²Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ⁹³Department of Medicine, University of Helsinki and Helsinki University Central Hospital, Helsinki 00029, Finland. ⁹⁴Endocrinology, Abdominal Center, University of Helsinki and Helsinki University Central Hospital, Helsinki 00029, Finland. ⁹⁵Department of Health, National Institute for Health and Welfare, Helsinki 00271, Finland. ⁹⁶Minerva Foundation Institute for Medical Research, Helsinki 00290, Finland. ⁹⁷Division of Cardiology, Department of Medicine, Duke Molecular Physiology Institute, Duke University School of Medicine, Durham, NC 27710, USA. ⁹⁸Sticht Center on Aging, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA. ⁹⁹Institute of Behavioural Sciences, University of Helsinki, Helsinki FI00014, Finland. ¹⁰⁰University of California San Francisco, San Francisco, CA 94143, USA. ¹⁰¹School of Medicine and Pharmacology, University of Western Australia, Perth 6009, Australia. ¹⁰²Centre for Kidney Research, School of Public Health, University of Sydney, Sydney 2006, Australia. ¹⁰³Wellcome Trust Centre for Human Genetics, Oxford University, Oxford OX3 7BN, UK. ¹⁰⁴Department of Epidemiology and Prevention, Wake Forest School of Medicine, Winston-Salem, NC 27517, USA. ¹⁰⁵Max Planck Institute for Molecular Genetics, Berlin 14195, Germany. ¹⁰⁶Max Planck Institute for Human Development, Berlin 14195, Germany. ¹⁰⁷Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg 85764, Germany. ¹⁰⁸Buck Institute for Research on Aging, Novato, CA 94945, USA. ¹⁰⁹Leonard Davis School of Gerontology, University of Southern California, LA, CA 90089, USA. ¹¹⁰Geriatrics Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, MD 21201, USA. ¹¹¹Institute of Translational Medicine, University of Liverpool, Liverpool L69 3BX, UK. ¹¹²Center for Aging and Population Health, University of Pittsburgh, Pittsburgh, PA 15261, USA. ¹¹³Oregon Health & Science University, Portland, OR 97239, USA. ¹¹⁴Department of Clinical Genetics, Erasmus MC, Rotterdam 300 CA, The Netherlands. ¹¹⁵Centre for Medical Systems Biology and Netherlands Consortium on Healthy Aging, Leiden RC2300, The Netherlands. ¹¹⁶Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki 00251, Finland. ¹¹⁷Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki FI00014, Finland. ¹¹⁸Human Genetics and Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA. ¹¹⁹Department of Medicine, Indiana University School of Medicine, Indianapolis, IN 46202, USA. ¹²⁰Diabetes and Obesity Research Program, University of Helsinki, Helsinki FI00014, Finland. ¹²¹Estonian Genome Center, University of Tartu, Tartu, Estonia. ¹²²Faculty of Medicine, Department of Public Health, University of Split, Split 21000, Croatia. ¹²³Department of Endocrinology and Diabetes, Sir Charles Gardiner Hospital, Perth 6009, Australia. ¹²⁴Molecular Medicine Centre, MRC Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, Scotland EH4 2XU, UK. ¹²⁵Hjelt Institute, University of Helsinki, Helsinki, Finland. ¹²⁶Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK. ¹²⁷Department of Medicine, University of California at Davis, Sacramento, CA 95817, USA. ¹²⁸Institute for Translational Genomic and Population Sciences, Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor UCLA Medical Center, Torrance, CA 90502, USA. ¹²⁹Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN 38163, USA. ¹³⁰Department of Genetics and Genomic Science, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ¹³¹Department of Endocrinology and Metabolism, Landspítali, The National University Hospital of Iceland, Reykjavik 101, Iceland. ¹³²Center for Translational Pain Medicine, Department of Anesthesiology, Duke University Medical Center, Durham, NC 27710, USA. ¹³³Geriatric Research and Education Clinical Center (GRECC) - Veterans Administration Medical Center, Baltimore, MD 21201, USA. ¹³⁴Department of Epidemiology and Biostatistics, and the EMGO Institute, VU University Medical Center, Amsterdam BT1081, The Netherlands. ¹³⁵Department of Medicine, McMaster University Medical Center, Hamilton, ON, Canada L8N 3Z5. ¹³⁶Department of Pathology, Stony Brook School of Medicine, Stony Brook, NY 11794, USA. ¹³⁷Department of Neuroscience and Preventive Medicine, Danube-University Krems, Krems 3500, Austria. ¹³⁸Diabetes Research Group, King Abdulaziz University, Jeddah 12589, Saudi Arabia. ¹³⁹Dasman Diabetes Institute, Dasman 15462, Kuwait. ¹⁴⁰Department of Medicine and Internal Medicine, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne CH-1011, Switzerland. ¹⁴¹Department of Epidemiology and Environmental Health, University at Buffalo, State University of New York, Buffalo, NY 14214, USA. ¹⁴²Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne NE2 4HH, UK. ¹⁴³Genetics of Complex Traits, University of Exeter Medical School, Exeter EX1 2LU, UK. ¹⁴⁴Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig-Maximilians-Universität, Munich 81377, Germany. ¹⁴⁵Institute of Medical Statistics and Epidemiology, Technical University, Munich 81675, Germany. ¹⁴⁶Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA. ¹⁴⁷NIHR Cardiovascular Biomedical Research Unit, Royal Brompton and Harefield NHS Foundation Trust and Imperial College, London SW3 6NP, UK. ¹⁴⁸Imperial College Healthcare NHS Trust, London W2 1NY, UK. ¹⁴⁹Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg 85764, Germany. ¹⁵⁰Department of Medicine and Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202, USA. ¹⁵¹Departments of Medicine, Epidemiology, and Health Services, Cardiovascular Health Research Unit, University of Washington, Seattle, WA 98101, USA. ¹⁵²Kaiser Permanente Washington Health Research Institute, Washington, Seattle, WA 98101, USA. ¹⁵³Medical Genetics, GlaxoSmithKline, Philadelphia, PA 19112, USA. ¹⁵⁴The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ¹⁵⁵Institute of Child Health and Development, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ¹⁵⁶The Genetics of Obesity and Related Traits Program, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ¹⁵⁷Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ¹⁵⁸Faculty of Medicine in the Galilee, Bar-Ilan University, Safed 1311502, Israel. M. Carola Zillikens, Serkalem Demissie, Yi-Hsiang Hsu and Laura M. Yerges-Armstrong contributed equally to this work. David Karasik, Tamara B. Harris, Claes Ohlsson and Douglas P. Kiel jointly supervised this work.