



OPEN

Unveiling genetic variants for age-related sarcopenia by conducting a genome-wide association study on Korean cohorts

Heejin Jin^{1,10}, Hyun Ju Yoo^{2,10}, Ye An Kim³, Ji Hyun Lee³, Young Lee⁴, Seung-hyun Kwon⁴, Young Joo Seo⁵, Seung Hun Lee⁶, Jung-Min Koh⁶, Yunmi Ji⁷, Ah Ra Do⁷, Sungho Won^{1,8,9} & Je Hyun Seo⁴✉

Sarcopenia is an age-related disorder characterised by a progressive decrease in skeletal muscle mass. As the genetic biomarkers for sarcopenia are not yet well characterised, this study aimed to investigate the genetic variations related to sarcopenia in a relatively aged cohort, using genome-wide association study (GWAS) meta-analyses of lean body mass (LBM) in 6961 subjects. Two Korean cohorts were analysed, and subgroup GWAS was conducted for appendicular skeletal muscle mass (ASM) and skeletal muscle index. The effects of significant single nucleotide polymorphisms (SNPs) on gene expression were also investigated using multiple expression quantitative trait loci datasets, differentially expressed gene analysis, and gene ontology analyses. Novel genetic biomarkers were identified for LBM (rs1187118; rs3768582) and ASM (rs6772958). Their related genes, including *RPS10*, *NUDT3*, *NCF2*, *SMG7*, and *ARPC5*, were differently expressed in skeletal muscle tissue, while *GPD1L* was not. Furthermore, the 'mRNA destabilisation' biological process was enriched for sarcopenia. Our study identified *RPS10*, *NUDT3*, and *GPD1L* as significant genetic biomarkers for sarcopenia. These genetic loci were related to lipid and energy metabolism, suggesting that genes involved in metabolic dysregulation may lead to the pathogenesis of age-related sarcopenia.

Sarcopenia is the age-related loss of skeletal muscle mass and strength, accompanied by functional impairment. As such, it is associated with disability, poor quality of life, and increased mortality¹⁻³. Considering the difficulties posed by frailty, and the healthcare costs associated with age-related conditions, such as sarcopenia^{4,5}, it is necessary to identify meaningful disease phenotypes and biomarkers. Several studies have suggested various criteria for defining sarcopenia⁶⁻¹⁰. Muscle mass is thought to be an important factor for the diagnosis of sarcopenia; of the parameters related to muscle mass, lean body mass (LBM) is frequently used to predict sarcopenia. In addition, the Asian working group for sarcopenia 2019 (AWGS 2019) recently reached the consensus that skeletal muscle index (SMI) and appendicular skeletal muscle (ASM) may also be reliable parameters¹¹. In this respect, in addition to LBM, it is necessary to analyse SMI and ASM to understand the complex aetiology of sarcopenia,

¹Institute of Health and Environment, Seoul National University, Seoul, South Korea. ²Department of Convergence Medicine, Asan Institute for Life Sciences, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea. ³Division of Endocrinology, Department of Internal Medicine, Veterans Health Service Medical Center, Seoul, South Korea. ⁴Veterans Medical Research Institute, Veterans Health Service Medical Center, Jinhwangdo-ro 61-gil 53, Gangdong-gu, Seoul 05368, South Korea. ⁵Department of Anesthesiology and Pain Medicine, Hangeang Sacred Heart Hospital, Hallym University College of Medicine, Seoul, South Korea. ⁶Division of Endocrinology and Metabolism, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea. ⁷Interdisciplinary Program of Bioinformatics, College of National Sciences, Seoul National University, Seoul, South Korea. ⁸Department of Public Health Science, Institute of Health & Environment, Seoul National University, Seoul, South Korea. ⁹RexSoft Corps, Seoul, South Korea. ¹⁰These authors contributed equally: Heejin Jin and Hyun Ju Yoo. ✉email: jazmin2@naver.com

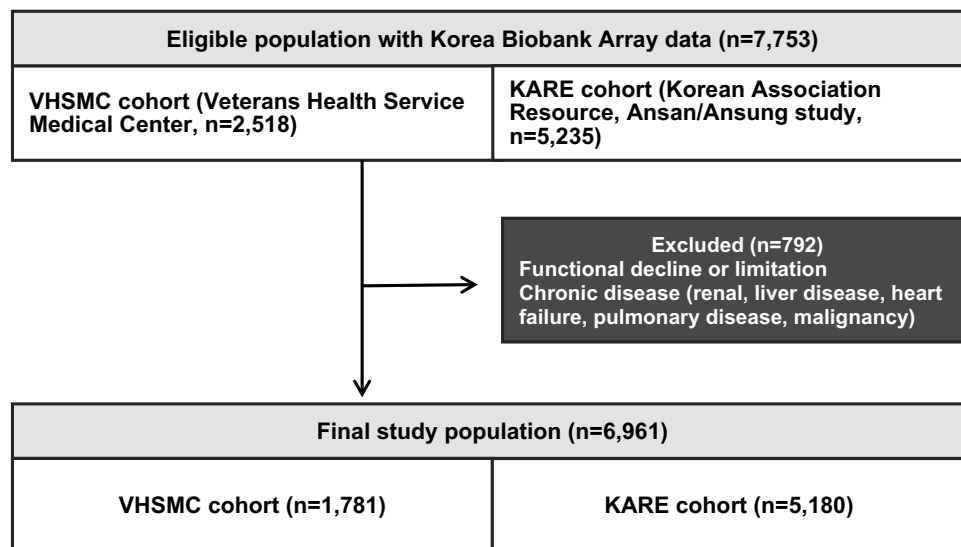


Figure 1. Schematic of study population.

which can be attributed to a variety of factors, including oxidative stress, inflammation, mitochondrial dysregulation, and genetic factors^{12,13}.

Muscle mass has a genetic trait phenotype, with a heritability estimate of over 50%¹⁴. Studies have investigated the genetic factors of LBM using the associations of single nucleotide polymorphisms (SNPs)^{15–19}. Moreover, as osteoporosis and sarcopenia share a common risk factor (ageing), several studies have conducted joint genome-wide association study (GWAS) analyses on overlapping genetic variants^{20,21}. Notably, a GWAS into osteoporosis revealed 64 loci. In contrast, fewer loci were identified by a GWAS into muscle-related phenotypes, thus, providing fewer biological insights into pathways regarding sarcopenia²⁰. To address this issue, a large GWAS meta-analysis was conducted using 20 cohorts of European ancestry, identifying a set of five loci (*HSD17B11*, *VCAN*, *ADAMTSL3*, *IRS1* and *FTO*) for total LBM, and SNPs related to *IRS1*, *ADAMTSL3*, and *VCAN* for appendicular LBM²². However, as this study analysed the entire cohort, irrespective of age, its findings regarding genes associated with sarcopenia as a senile disease were limited. In addition, the GWAS was based on European ancestry, and little is known regarding genetic determinants in elderly East Asians. Still further, few genetic studies have utilised a new index for sarcopenia (released by the AWGS in 2019) to investigate ASM or SMI¹¹. Thus, a need exists for the investigation of genetic components associated with sarcopenia using multiple cohorts comprising elderly East Asians. The current study conducted a GWAS meta-analysis on sarcopenia phenotypes using Korean relatively aged cohorts, combining the Veterans Health Service Medical Center (VHSMC) and Korean Association Resource (KARE) cohorts.

Results

Characteristics of the study participants. A total of 7753 eligible subjects were included in this study (2518 subjects from the VHSMC cohort and 5235 from the KARE cohorts). However, 792 were excluded due to the exclusion criteria (Fig. 1), leaving a remainder of 6961 participants (1781 subjects from the VHSMC cohort and 5180 from the KARE cohort) that were included in analyses. The mean age of the VHSMC cohort was higher than that of the KARE cohort (69.10 ± 7.83 years vs 62.79 ± 8.33 years, $P < 0.001$, Table 1). No significant difference was observed in mean height (1.59 ± 0.08 m in VHSMC vs 1.59 ± 0.09 m in KARE, $P = 1.000$) between the two cohorts. The mean weight (63.24 ± 10.51 kg in VHSMC vs 62.64 ± 10.37 kg in KARE, $P = 0.037$) and BMI (24.74 ± 3.21 kg/m² in VHSMC vs 24.53 ± 3.15 kg/m² in KARE, $P = 0.016$) were statistically different between the cohorts. The LBM of the VHSMC cohort was lower than that of the KARE cohort (40.10 ± 7.83 kg vs 42.03 ± 8.27 kg, respectively, $P < 0.001$), whereas the body fat mass (BFM) of the VHSMC cohort was higher than that of the KARE cohort (20.60 ± 6.23 kg vs 18.28 ± 5.85 kg, respectively, $P < 0.001$). Descriptive statistics for subgroups according to sex are also presented in Table 1. The mean values of SMI and ASM, which could only be calculated for the VHSMC cohort, were 6.77 ± 1.00 kg/m² and 17.49 ± 4.06 kg, respectively.

GWAS meta-analysis of lean body mass and body fat mass. A total of 2,360,975 SNPs were used for the GWAS meta-analysis of LBM and BFM. Quantile–quantile (Q–Q) and Manhattan plots for LBM are shown in Fig. 2. The Q–Q plot revealed no evidence of test statistic inflation (variance inflation factor [VIF] = 1.044). The top ten variants for LBM are listed in Table 2; two of which were genome-wide significant loci. The most significant variant was rs11871118 (effect = 0.720, standard error [SE] = 0.117, $P = 1.09 \times 10^{-9}$, *HetPVal* = 0.199) near *Glutamate Metabotropic Receptor 4 (GRM4)* and *High Mobility Group AT-Hook 1 (HMGAI1)*, followed by rs3768582 (effect = 0.554, SE = 0.100, $P = 4.09 \times 10^{-8}$, *HetPVal* = 0.537) near *Neutrophil Cytosolic Factor 2 (NCF2)*. The remaining eight variants are presented as candidate loci in Table 2. The Q–Q and Manhattan plots for BFM are shown in Supplementary Fig. S1. The Q–Q plot revealed no evidence of test statistic inflation (VIF = 1.037).

Cohort	Entire			Male			Female		
	VHSMC	KARE	P	VHSMC	KARE	P	VHSMC	KARE	P
Numbers	1781	5180		712	2449		1069	2731	
Sex (male)	712 (40%)	2449 (47%)	0.001 ^a	N/A	N/A	N/A	N/A	N/A	N/A
Age (years)	69.10 ± 7.83	62.79 ± 8.33	< 0.001 ^b	71.11 ± 7.34	62.20 ± 8.03	< 0.001 ^b	67.77 ± 7.86	63.32 ± 8.56	< 0.001 ^b
Height (m)	1.59 ± 0.08	1.59 ± 0.09	1.000 ^b	1.67 ± 0.05	1.66 ± 0.06	< 0.001 ^b	1.55 ± 0.05	1.53 ± 0.06	< 0.001 ^b
Weight (kg)	63.24 ± 10.51	62.64 ± 10.37	0.037 ^b	69.95 ± 9.48	67.79 ± 9.68	< 0.001 ^b	58.77 ± 8.62	58.00 ± 8.63	0.013 ^b
BMI (kg/m ²)	24.74 ± 3.21	24.53 ± 3.15	0.016 ^b	24.99 ± 3.02	24.32 ± 2.94	< 0.001 ^b	24.57 ± 3.32	24.72 ± 3.33	0.211 ^b
LBM (kg)	40.24 ± 7.60	42.03 ± 8.27	< 0.001 ^b	47.69 ± 6.23	48.86 ± 6.01	< 0.001 ^b	35.90 ± 4.21	35.90 ± 4.21	< 0.001 ^b
BFM (kg)	20.60 ± 6.23	18.28 ± 5.85	< 0.001 ^b	20.60 ± 6.23	16.32 ± 5.32	< 0.001 ^b	20.04 ± 5.75	20.04 ± 5.75	< 0.001 ^b
SMI (kg/m ²)	6.77 ± 1.00	N/A	N/A	7.67 ± 0.72	N/A	N/A	6.18 ± 0.65	N/A	N/A
ASM (kg)	17.49 ± 4.06	N/A	N/A	21.50 ± 2.80	N/A	N/A	14.82 ± 2.10	N/A	N/A

Table 1. Baseline characteristics of study populations. ASM, appendicular skeletal muscle; BMI, body mass index; LBM, lean body mass; BFM, body fat mass; SMI, skeletal muscle mass index; VHSMC, Veterans Health Service Medical Center; KARE, Korean Association Resource; N/A, not available. ^aChi-square test, ^bt-test.

The GWAS meta-analysis for BFM showed no genome-wide significant loci and the variant with the smallest *P*-value was rs1592269 (effect = 0.753, SE = 0.148, *P* = 3.43 × 10⁻⁷, *HetPVal* = 0.379) near *GRM4* and *HMGAI1*. The top ten candidate loci associated with BFM with *P*-values < 1.00 × 10⁻⁵ are listed in Supplementary Table S1. As the GWAS results for the LBM and BFM phenotypes exhibited similar loci (*GRM4* and *HMGAI1*), linkage disequilibrium (LD) analysis was performed. A high (*r*² = 0.935) LD between rs1187118 and rs1592269 was observed, indicating a relatively high dependency.

GWAS of appendicular skeletal muscle and skeletal muscle index. A total of 2,804,834 SNPs were used for GWAS analyses of ASM and SMI, using only the VHSMC cohort. The Q-Q and Manhattan plots for ASM are shown in Fig. 3; the Q-Q plot did not exhibit evidence of test statistic inflation (VIF = 1.031). The top ten variants for ASM are listed in Table 3; the only significant variant was a genome-wide locus: rs6772958 (effect = - 0.456, SE = 0.081, *P* = 2.30 × 10⁻⁸) near *zinc finger protein 860 (ZNF860)* and *Glycerol-3-Phosphate Dehydrogenase 1 Like (GPD1L)*. The Q-Q and Manhattan plots for SMI are shown in Supplementary Fig. S2 and revealed no evidence of test statistic inflation (VIF = 1.034). However, the GWAS for SMI exhibited genome-wide significant loci; the variant with the smallest *P*-value was rs6772958 (effect = - 0.121, SE = 0.023, *P* = 1.72 × 10⁻⁷, *HetPVal* = 0.379) near *ZNF860* and *GPD1L*. The top ten candidate loci with *P*-values < 1.00 × 10⁻⁵ are suggested in Supplementary Table S2.

Regional analysis and functional annotation. For the genome-wide significant variants of each phenotype (LBM and ASM), the regional plots with the lead SNPs are displayed in Figs. 4 and 5. The first phenotype of interest was LBM. The most genome-wide significant SNP, rs1187118 eQTL analyses from the GTEx Project (V7), showed that *Ribosomal Protein S10 (RPS10)* was highly expressed in skin sun-exposed lower leg tissue (*P* = 1.40 × 10⁻⁷). Its LD variant eQTL association for *Nudix Hydrolase 3 (NUDT3)* was also found in the skeletal muscle tissue (*P* = 4.30 × 10⁻²¹). The second genome-wide significant SNP, rs3768582 eQTL analyses, showed that *NCF2*, *SMG7 (SMG7 Nonsense-Mediated mRNA Decay Factor)* and *Actin Related Protein 2/3 Complex Subunit 5 (ARPC5)* were highly expressed in the artery (*P* = 1.50 × 10⁻⁹), heart (*P* = 2.60 × 10⁻⁵), and cultured fibroblast tissue (*P* = 7.30 × 10⁻⁵), respectively. In the differentially expressed gene (DEG) analysis with GSE38718, compared with the young group, *RPS10* (*P* = 8.00 × 10⁻⁴), *NUDT3* (*P* = 1.19 × 10⁻³), *NCF2* (*P* = 1.26 × 10⁻²), *SMG7* (*P* = 1.03 × 10⁻³) and *ARPC5* (*P* = 4.26 × 10⁻²) were more expressed in the elderly group (Table 4).

The second phenotype of interest was ASM. The only genome-wide significant SNP, rs6772958 eQTL analysis, showed that *GPD1L* was highly expressed in thyroid tissue (*P* = 8.10 × 10⁻¹⁵). However, the DEG analysis showed that *GPD1L* expression was not significant (*P* = 0.159) in the transcriptome study (GSE38718).

In addition, gene ontology (GO) analyses of biological processes revealed that the term '*mRNA destabilisation (GO: 0061157)*' (FDR-adjusted *P* = 0.090) was enriched, which is involved in skeletal muscles related genes. The term contains a pathway of alpha-ketoglutarate-dependent dioxygenase FTO (U6 small nuclear RNA [2'-O-methyladenosine-N(6)-]-demethylase FTO), which is involved in the regulation of fat mass, adipogenesis, and body weight. Thus, it contributes to the regulation of body size and body fat accumulation²³.

Discussion

This study discovered novel genetic biomarkers of LBM (rs1187118) and ASM (rs6772958) from the VHSMC and KARE cohorts, which comprise relatively aged (mean age: 69.10 vs. 62.79, respectively) Koreans. Their related genes for LBM, such as *RPS10*, *NUDT3*, *NCF2*, *SMG7*, and *ARPC5*, were expressed in skeletal muscle tissue. In addition, in the biological process, the term '*mRNA destabilisation (GO: 0061157)*' (FDR-adjusted *P* = 0.090) was enriched for sarcopenia. This process contains alpha-ketoglutarate-dependent dioxygenase FTO. These results suggest that the pathogenesis of sarcopenia requires further investigation using a metabolic pathway linked to *mRNA*.

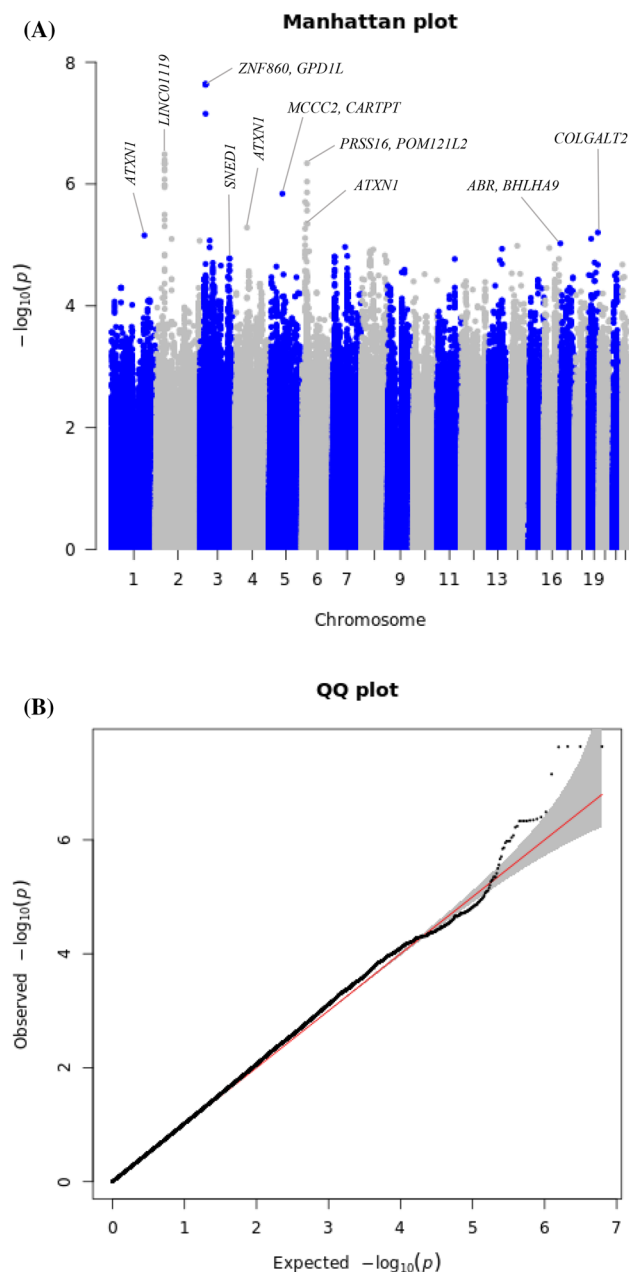


Figure 2. Manhattan and quantile–quantile plot for lean body mass in the meta-analysis. **(A)** Manhattan plot of the P -values in the genome-wide association study (GWAS) meta-analysis for lean body mass. **(B)** Quantile–quantile (Q–Q) plot showing expected vs. observed $[-\log_{10}(P)]$ values. The expected line is shown in red and confidence bands are shown in grey.

The aetiology of sarcopenia is complex and includes oxidative stress, inflammation, inadequate diets, a sedentary lifestyle, and genetic factors¹³. A previous study on genetic markers for sarcopenia identified the loci near *FTO*, *ESR1*, *NOS3*, *KLF5*, and *HLA-DQA1* to be associated with physical phenotypes, such as low handgrip strength and decreased LBM^{24–26}. Nonetheless, these identified loci can only explain a small portion of phenotypic variations; thus, additional genetic loci should be identified. A recent large meta-analysis of the Cohorts for Heart and Ageing Research in Genome Epidemiology (CHARGE) Consortium and various other cohorts identified only a few loci, such as *FTO* and *VCAN* for LBM²². Therefore, assuming that identifying genetic variants for sarcopenia is challenging, we conducted GWAS analysis on a cohort comprising elderly subjects. The findings revealed that several genetic variants related to metabolism could be of importance in determining the pathogenesis of sarcopenia. Previous sarcopenia GWAS for European descendants showed association with *FTO*^{22,27,28} and several loci, including *TGFA* and *HLA-DRB1*²⁹.

Our meta-analysis for both LBM and BFM showed significant differences in the intergenic area of *GRM4* and *HMGAI*, with a high LD between rs1187118 and rs1592269. *HMGAI* is overexpressed in adipose tissue, impairs adipogenesis, and prevents diet-induced obesity, and insulin resistance³⁰. The top loci for LBM and BFM

Chr	SNP	Position	Allele	MAF	Independent study					Meta-analysis			Mapped genes
					Cohort	Effect	SE	MAF	P	Effect	SE	P (HetPVal)	
6	rs1187118	34169020	A/T	0.16 ^a , 0.11 ^b	VHSMC	0.488	0.208	0.16	1.89×10^{-2}	0.720	0.117	1.09×10^{-9} (0.1995)	GRM4, HMGA1 (intergenic)
					KARE	0.831	0.143	0.17	7.92×10^{-9}				
1	rs3768582	183553870	C/T	0.27 ^a , 0.35 ^b	VHSMC	0.637	0.177	0.23	3.42×10^{-4}	0.554	0.100	4.09×10^{-8} (0.5376)	NCF2 (intronic)
					KARE	0.516	0.122	0.27	2.63×10^{-5}				
15	rs12441313	59015488	G/T	0.33 ^a , 0.35 ^b	VHSMC	0.473	0.161	0.32	3.35×10^{-2}	0.475	0.094	4.74×10^{-7} (0.8894)	ADAM10 (intronic)
					KARE	0.477	0.116	0.33	4.19×10^{-5}				
22	rs4822064	42279506	G/C	0.05 ^a , 0.06 ^b	VHSMC	0.663	0.350	0.05	5.91×10^{-2}	0.974	0.199	1.05×10^{-6} (0.3081)	SREBF2 (intronic)
					KARE	1.124	0.243	0.05	3.89×10^{-6}				
6	rs7453329	58721,863	C/T	0.41 ^a , 0.37 ^b	VHSMC	0.436	0.154	0.41	4.81×10^{-3}	0.426	0.090	2.39×10^{-6} (0.8283)	LINC00680-GUSBP4, NONE (intergenic)
					KARE	0.421	0.111	0.41	1.51×10^{-4}				
7	rs77502935	157290819	C/T	0.11 ^a , 0.10 ^b	VHSMC	0.430	0.241	0.12	7.50×10^{-2}	0.651	0.138	2.56×10^{-6} (0.3043)	LOC101927914 (ncRNA_intronic)
					KARE	0.762	0.170	0.11	7.71×10^{-6}				
8	rs12156192	40763707	C/A	0.05 ^a , 0.04 ^b	VHSMC	-0.663	0.300	0.07	2.75×10^{-2}	-0.805	0.174	3.83×10^{-6} (0.6457)	ZMAT4,SFRP1 (intergenic)
					KARE	-0.879	0.215	0.06	4.37×10^{-5}				
3	rs12497693	72400695	C/T	0.40 ^a , 0.34 ^b	VHSMC	0.624	0.155	0.39	6.15×10^{-5}	0.414	0.089	4.33×10^{-6} (0.0831)	LINC00870, RYBP (intergenic)
					KARE	0.311	0.109	0.40	4.52×10^{-3}				
3	rs73872711	141110287	C/G	0.28 ^a , 0.27 ^b	VHSMC	0.401	0.166	0.30	1.61×10^{-2}	0.437	0.096	6.11×10^{-6} (0.8804)	ZBTB38 (intronic)
					KARE	0.457	0.119	0.29	1.26×10^{-4}				
11	rs111357357	3640972	C/T	0.16 ^a , 0.13 ^b	VHSMC	0.421	0.208	0.16	4.30×10^{-2}	0.533	0.119	8.29×10^{-6} (0.5698)	LOC115995503, TRPC2 (intergenic)
					KARE	0.589	0.146	0.17	6.00×10^{-5}				

Table 2. Results of GWAS meta-analysis for lean body mass (leading SNPs, top 10). Chr, chromosome; SNP, single nucleotide polymorphism; MAF, minor allele frequency; SE, standard error; Mapped Genes from ANNOVAR; GWAS, genome-wide association study; VHSMC, Veterans Health Service Medical Center; KARE, Korean Association Resource. ^aKref, Korean reference data; ^bGnomAD Genome Aggregation Database (east Asian).

were similar, and those of ASM and SMI were similar since the parameter of LBM was calculated from body weight minus BFM, and SMI was calculated from ASM/height². Hence, it would be useful to calculate the correlation and genetic correlation for each parameter. In VHSMC cohorts, the correlation and genetic correlation were 0.078 and 0.078 between LBM and BFM, respectively, whereas those between SMI and ASM were 0.948 and 0.948, respectively. In KARE cohorts, the correlation was -0.02 between LBM and BFM with the genetic correlation being 0.349.

The eQTL analysis for muscle mass using GTEx datasets showed that *RPS10*, *NUDT3*, *NCF2*, *SMG7*, and *ARPC5* were differentially expressed in the muscle tissue for sarcopenia. However, this finding requires further validation. As the regional locations of *HMGA1*, *RPS10*, and *SIMM29* were in the upper stream of *NUDT3*, and may represent a regulatory function for the association of *NUDT3* with sarcopenia, further focus should be directed towards *NUDT3*. A previous study by Singh et al. suggested that *NUDT3* was a candidate target-locus, and emphasised the need for real-world validation using transcriptome-wide association study (TWAS) approaches that combine GWAS and eQTL summary data²⁴. In the current study, *NUDT3* was found to be related to LBM in an elderly cohort. *NUDT3* belongs to the *MutT* or *Nudix* protein families, which act as homeostatic checkpoints at important stages in inositol phosphate metabolic pathways. These pathways, such as phosphatidylinositol-3-OH-myoinositol and glycerophospholipid metabolism, from the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (<https://www.kegg.jp/pathway/map00564>)³¹ may, therefore, be related to LBM. For these reasons, it is necessary to understand the metabolic aspects of sarcopenia. A study into DEGs in skeletal muscle tissues from patients with cachexia³² showed that *NCF2* was identified from signal pathways related to inflammation. These findings are consistent with the findings of the present study. *SMG7* encodes a protein that is essential for nonsense-mediated mRNA decay, which is related to body height and BMI-adjusted waist circumference from a GWAS catalogue (<https://www.ebi.ac.uk/gwas/>). As *SMG7* is linked with telomerase reverse transcriptase (TERT), sarcopenia may be related to muscle cell senescence via microRNA-195³³. In addition, *ARPC5* encodes the actin related protein2/3 complex, which exhibited a negative fold change in expression related to the cytoskeleton in muscle tissue³⁴. As these findings may represent secondary changes, or may be postulated from bioinformatics analysis, further studies are needed.

Additionally, the present study found that *GPD1L* is a significant genetic marker for ASM and SMI in the VHSMC cohort. Although this is a novel finding for GWAS using ASM as a parameter for sarcopenia, it requires further validation by studies from several cohorts. A tissue-based study into rat muscle identified *GPD1L* as a candidate locus for sarcopenia³⁵. These findings were also observed in a previous study that investigated the sarcopenic muscle tissue of elderly women³⁶, in which *GPD1L* was found to be downregulated via cytoplasmic

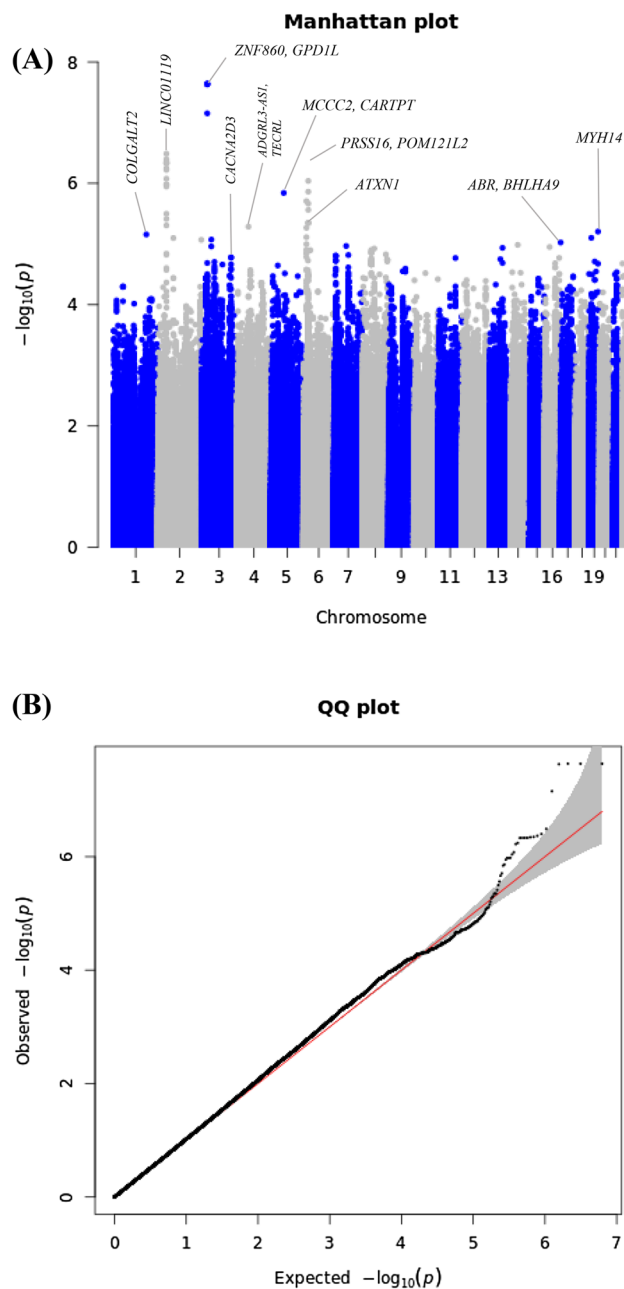


Figure 3. Manhattan and quantile–quantile plot for appendicular skeletal muscle in the genome-wide association analysis. **(A)** Manhattan plot of the P -values in the genome-wide association study (GWAS) meta-analysis for appendicular skeletal muscle. **(B)** Quantile–quantile (Q–Q) plot showing expected vs. observed $[-\log_{10}(P)$ values]. The expected line is shown in red and confidence bands are shown in grey.

energy metabolism. In addition, a systemic genetic approach identified that *GPD1L* and its molecular mechanism for obesity in human adipose tissue were associated with energy metabolism³⁷. *GPD1L* expression was found to be negatively correlated with microRNA-210 (miR-210) levels, and was consistently downregulated in obese subjects³⁷. They hypothesised that the decreased miR-210 levels increased *GPD1L*, thus inhibiting hypoxic transcription factor-1 α (HIF-1 α) activity. A previous study into the circulating miRNAs in plasma revealed that miR-210 is significantly downregulated in elderly patients with sarcopenia, compared to patients without sarcopenia³⁸. Combined with results of previous studies^{37,38}, the findings presented here suggest that *GPD1L* could be a genetic biomarker for sarcopenia, based on both miR-210 and HIF-1 α pathways. Hence, an additional biomarker for sarcopenia may be postulated from this metabolic research. Recent studies into plasma biomarkers for sarcopenia have identified higher levels of amino acids and lower levels of phosphatidylcholines (PCs) and lysophosphatidylcholine (lysoPC)^{39,40}. The association between *GPD1L* and PCs or lysoPC and sarcopenia may involve (1) dysregulation of *GPD1L* related to decreased PCs and lysoPC from previous lipid biomarkers^{39,40},

Chr	SNP	Position	Allele	MAF	Effect	SE	P	Mapped genes
3	rs6772958	32049091	A/G	0.39 ^a , 0.39 ^b , 0.45 ^c	- 0.456	0.081	2.30 × 10 ⁻⁸	ZNF860, GPD1L (intergenic)
2	rs1466024	47065406	T/C	0.38 ^a , 0.37 ^b , 0.37 ^c	0.430	0.008	3.24 × 10 ⁻⁷	LINC01119 (ncRNA_intronic)
6	rs4713078	27227782	T/G	0.10 ^a , 0.11 ^b , 0.07 ^c	0.728	0.147	9.13 × 10 ⁻⁷	PRSS16, POM121L2 (intergenic)
5	rs16870827	70976503	A/G	0.13 ^a , 0.11 ^b , 0.22 ^c	- 0.567	0.117	1.44 × 10 ⁻⁶	MCCC2, CARTPT (intergenic)
6	rs2146753	16694501	A/G	0.34 ^a , N/A, 0.34 ^c	- 0.401	0.084	1.97 × 10 ⁻⁶	ATXN1 (intronic)
4	rs142722457	64547838	A/G	0.06 ^a , 0.06 ^b , 0.04 ^c	- 0.792	0.173	5.19 × 10 ⁻⁶	ADGRL3-AS1, TECRL (intergenic)
19	rs1947424	50750497	C/T	0.15 ^a , 0.15 ^b , 0.17 ^c	- 0.481	0.106	6.24 × 10 ⁻⁶	MYH14 (intronic)
1	rs12089250	183983087	C/T	0.37 ^a , 0.38 ^b , 0.43 ^c	0.375	0.083	6.97 × 10 ⁻⁶	COLGALT2 (intronic)
3	rs11130441	55008440	C/T	0.42 ^a , 0.42 ^b , 0.42 ^c	- 0.360	0.080	8.46 × 10 ⁻⁶	CACNA2D3 (intronic)
17	rs59567306	1143024	A/G	0.06 ^a , 0.08 ^b , 0.12 ^c	- 0.744	0.167	9.44 × 10 ⁻⁶	ABR, BHLHA9 (intergenic)

Table 3. Results of GWAS for appendicular skeletal muscle (leading SNPs, top 10). Chr, chromosome; SNP, single nucleotide polymorphism; MAF, minor allele frequency; SE, standard error; Mapped Genes from ANNOVAR; GWAS, genome-wide association study. ^aVHSMC, Veterans Health Service Medical Center; ^bKref, Korean reference data; ^cGnomAD, Genome Aggregation Database (east Asian).

or (2) an increase in the glycerol-3 phosphate pathway inducing changes in glycolysis via *GPD1L*. However, the results of the present study can only be used to suggest a genetic hypothesis; thus, further follow-up studies are needed.

Analysis of the enriched biological processes identified via GO analysis of the cohorts revealed that alpha-ketoglutarate-dependent dioxygenase *FTO* is related to sarcopenia. This finding is consistent with those of a previous study on the influences of *FTO* and muscle phenotypes²⁷. In addition, alpha-ketoglutarate is a component of the tricarboxylic acid cycle, which is related to the HIF-1 α pathway. This evidence suggests that a simultaneous understanding of both genes and gene-metabolic pathways is necessary to understand the pathogenesis of sarcopenia.

One of the primary strengths of this study is the utilisation of a relatively elderly cohort sample, which provides a better sarcopenic phenotype. Here, *NUDT3* and *RPS10* were replicated using a real cohort, which was an approach suggested by a previous study using the TWAS of muscle tissue²⁴. Furthermore, our study focused on East Asian subjects, which have not been fully evaluated, unlike other ethnic groups. In this regard, we conducted phenome-wide association studies (pheWAS) using the “Common metabolic disease knowledge portal” (<https://hugeamp.org>), indicating that SNPs such as rs1187118, rs3768582, and rs6772958 are related to metabolic conditions such as waist-hip ratio, lipid metabolism, and body fat percentage in the European population (Supplementary Table S3).

Nevertheless, certain limitations were noted in this study. First, although novel signals for LBM and ASM were discovered with genome-wide significance, our results were based on bioinformatics analysis and, therefore, must be replicated in other Asian cohorts or multi-ethnic samples. A large number of samples for phenotypes, such as ASM and SMI, will improve the study’s validity. Hence, further studies, including replication or meta-analysis, are needed in other cohorts of the Asian population. Moreover, the number of SNPs (2,804,834) in the VHSMC cohort was limited as we set the imputation accuracy to 0.9. These points should be considered in the interpretation of the results. Second, the difference in ageing biology between sexes further hinders the identification of meaningful biomarkers for age-related conditions. Although GWAS analysis was conducted according to sex, the results did not show significant loci with genome-wide significance. It is expected that a metabolite-GWAS, considering sex as a factor, could help address this problem. Third, bioinformatics analysis revealed that genetic variants and metabolic pathways were related to sarcopenia, however, the causality of this hypothesis requires further investigation. Moreover, previous studies on genetic variants in sarcopenia have shown that these variants may be associated with the effects of genetic, metabolic, and environmental factors^{22,27,28}. Fourth, we used bioelectrical impedance analysis (BIA) for LBM and BFM examinations, as it is a non-invasive method for measuring body composition. However, dual-energy X-ray absorptiometry (DXA) is the standard method for muscle mass. BIA and DXA have different limitations for studies using body composition measurements. A previous study that compared these two methods found that BIA overestimated ASM compared to DXA⁴¹. In addition, BIA devices differed in the two cohorts (InBody 3.0 for KARE cohort, InBody770 for VHSMC cohort), which may be a confounding factor. In a technical review of BIA for people with high body fat, InBody 3.0 tended to be lower, with a difference of about 2% in an extreme case (unpublished data). A previous study showed that different BIA devices were reliable by high intraclass correlation coefficients and low standard errors⁴². Since the focus of our study was on muscle mass rather than fat mass, and we analysed each cohort using different PCs, differences associated with the BIA device between the two cohorts would not significantly influence the LBM values and analysis results presented in this study. However, it is necessary to consider these when interpreting research results.

In conclusion, sarcopenia can result in adverse outcomes, such as an increased risk of falls, a decreased quality of life, and mortality. Thus, it is necessary to identify a biomarker for this condition. Here, the loci near genes such as *RPS10*, *NCF2*, *SMG7*, *ARPC5*, and *NUDT3* were identified to be significant biomarkers for LBM. In addition, the loci near *GPD1L* were identified as significant biomarkers for ASM and SMI, which serve as novel

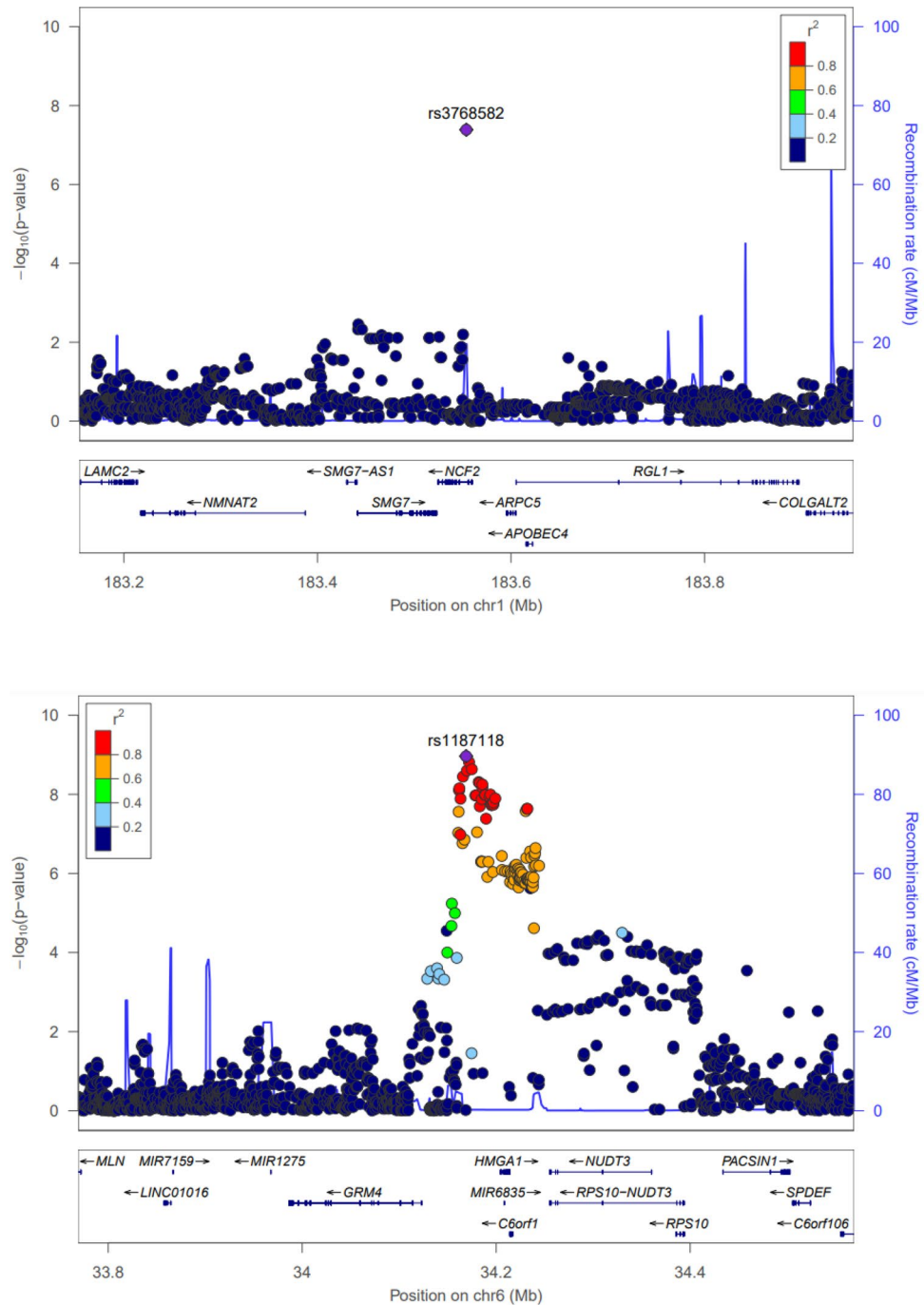


Figure 4. LocusZoom plot of genome-wide significantly associated SNPs for lean body mass.

index for sarcopenia. These genes are related to metabolism pathways, such as glycerophospholipid pathways, energy metabolic pathway, the inositol phosphate and HIF-1 α pathways, and alpha-ketoglutarate-dependent dioxygenase FTO. Further studies are required to evaluate the aetiology of sarcopenia.

Methods

Study subjects. Schematic plots of the analytical study design are shown in Fig. 1. Data were obtained from two cohorts: the VHSMC (n = 2518) and KARE (Ansan/Ansung study: from Korean Genome and Epidemiology Cohort, n = 5235) cohorts. Each cohort has its own distinct characteristics. The VHSMC cohort is a hospital-based elderly cohort that includes many patients with various diseases. The KARE cohort is a nationwide representative cohort for genome research in Korea; it is a longitudinal cohort of the Ansan and Ansung communities in Korea. This study included subjects from the KARE cohort and VHSMC cohort consisting of micro array data. Patients who had functional declines or limitations, or who had chronic diseases that may affect primary

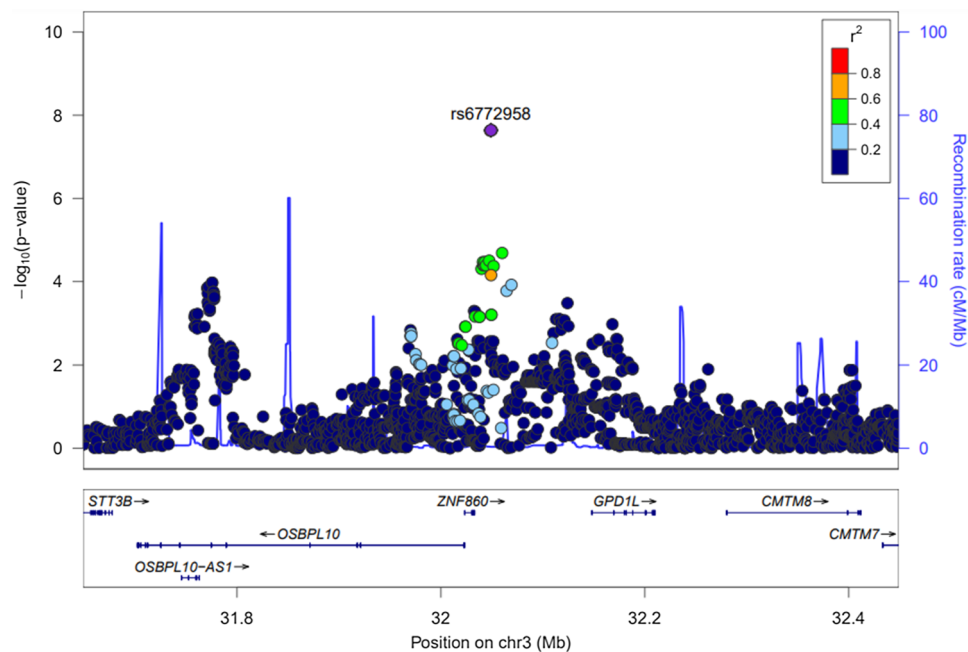


Figure 5. LocusZoom plot of genome-wide significantly associated SNPs for appendicular skeletal muscle.

Data set	Platform	Sample size	Gene	log FC	AveExpr	P-value	B
GSE38718	GPL570	Total: 22 Young: 14 Old: 8 Biceps brachii muscle	<i>RPS10</i>	-0.752	5.026	8.00×10^{-4}	-0.556
			<i>NUDT3</i>	-0.528	7.284	1.19×10^{-3}	-0.904
			<i>NCF2</i>	-0.774	4.457	1.26×10^{-2}	-3.074
			<i>SMG7</i>	-0.609	5.886	1.03×10^{-3}	-0.766
			<i>ARPC5</i>	-0.185	4.672	4.26×10^{-2}	-4.148
			<i>GPD1L</i>	-0.259	10.513	0.159	-5.238

Table 4. Results of differentially expressed genes from Gene Expression Omnibus (GEO) databases. logFC = estimate of the log₂-fold-change corresponding to the effect or contrast; AveExpr = average log₂-expression for the probe over all arrays and channels; B = log-odds that the gene is differentially expressed.

sarcopenia according to AWGS 2019¹¹, were excluded. After exclusion, 6961 participants were enrolled across both cohorts (Fig. 1). The institutional review boards of the Veterans Health Service Medical Center approved this study protocol and informed consent waiver (IRB No. 2020-02-015 and IRB No. 2021-05-005), since this study was performed in a retrospective manner, and the study was conducted in compliance with the Helsinki Declaration. The committee of VHS Biobank (VBP-2020-03) and the National Biobank of Korea (KBN-2021-041) approved the use of bioresources for this study.

Muscle mass measurement. BIA measurements were performed using InBody 770 (Biospace Co., LTD, Seoul, Korea) in the VHSMC cohort and using InBody 3.0 (Biospace Co., LTD, Seoul, Korea) in the KARE cohort. Each subject stood on the footplate and held both of the hand electrodes. The screen automatically displayed measurements of LBM (kg), skeletal muscle mass (kg), BFM (kg), and body fat percentage (%). LBM and BFM data were available for both cohorts and were used as initial phenotypes for analysis. Subgroup analysis was conducted using ASM or SMI, which were derived from BIA; these data were available only for the VHSMC cohort. The parameters were defined according to the consensus of the AWGS 2019¹¹.

Genotyping and imputation. Genomic DNA was separated from venous blood samples, and 100 ng DNA was genotyped using Korea Biobank Array Affymetrix Axiom 1.1 (Affymetrix, Santa Clara, CA), which was designed by the Korean National Institute of Health⁴³. Genotypes were identified with a K-medoid clustering-based algorithm to minimise the batch effect⁴⁴. The PLINK (version 1.9, Boston, MA)⁴⁵ and ONETOOL⁴⁶ software packages were used for quality control procedures and association analyses. Samples matching any of the following criteria were excluded: (1) sex inconsistencies or (2) a call rate of up to 97%. SNPs were filtered if the call rate was lower than the Hardy-Weinberg equilibrium (HWE) test ($P < 1 \times 10^{-5}$). The genotype imputation was conducted using the Michigan imputation server (<https://imputationserver.sph.umich.edu>). Only

'non-European' or 'mixed' populations from Haplotype Reference Consortium release v1.1⁴⁷ were used for reference purposes. Pre-phasing and imputation were performed using Eagle v2.4⁴⁸ and Minimac4⁴⁹, respectively. After the imputation processes, imputed SNPs were removed if the R-squared (i.e., imputation accuracy) was less than 0.9 or there were duplicated SNPs, missing genotype rates were more extensive than 0.05, *P*-values for HWE were less than 1×10^{-5} , or minor allele frequencies (MAFs) were less than 0.05. The MAF was compared with a reference such as Korean reference data (Kref) (<http://coda.nih.gov>) or the Genome Aggregation Database (GnomAD) with East Asian subjects (<https://gnomad.broadinstitute.org/>). Finally, 2422 subjects (and their 2,804,834 SNPs) from the VHSMC cohort and 5235 subjects (and their 3,423,819 SNPs) from the KARE cohort were used for analysis.

Statistical analyses. Baseline characteristics of the study population are presented herein as means with standard deviation (SD) for continuous variables and numbers, and as proportions for categorical variables. Genome-wide analyses were conducted using a linear model; PLINK was used within each cohort. Age, sex, and ten principal component scores were included as covariates. Meta-analyses of the VHSMC and KARE cohorts were performed using the METAL software (<http://csg.sph.umich.edu/abecasis/meta>). Cochran's Q-test for heterogeneity was conducted; its *P*-value was marked with 'HetPVal'⁵⁰, where *HetPVal* < 0.05 indicates heterogeneity between two datasets⁵¹. The dense regional association result of each GWAS was plotted using the LocusZoom software⁵². The threshold for statistical significance in this model was $P < 5.0 \times 10^{-8}$, which is conventionally considered to reflect genome-wide significance.

Functional annotation analyses. Expression Quantitative trait (eQTL) studies were performed using the Genotype-Tissue Expression (GTEx) dataset (<https://gtexportal.org/home/>), which provides a variety of human tissues from donors using the densely genotyped data to assess genetic variations within their genomes. Genes related to metabolites were analysed using KEGG pathway analysis³¹. Associated genes were further investigated for DEGs in the skeletal muscles of subjects 19 to 28 and 65 to 76 years of age from the Gene Expression Omnibus (GEO) dataset (GSE38718)⁵³. In addition, biological process, cellular component, and molecular function GO analyses were performed using gene set enrichment analysis. The Benjamini–Hochberg false discovery rate (FDR)-adjusted 0.1 significance level was applied for multiple hypothesis test corrections⁵⁴.

Ethics declarations. The institutional review boards of the Veterans Health Service Medical Center approved this study protocol and informed consent waiver (IRB No. 2020-02-015 for *VHSMC cohort* and IRB No. 2021-05-005 for *KARE cohort*) since this study was performed in retrospective manner, and the study was conducted in compliance with the Helsinki Declaration.

Consent to participate. Informed consent waiver was approved by the institutional review boards of the Veterans Health Service Medical Center since this study was performed in a retrospective manner and anonymised and de-identified data were used for the analyses. The KARE cohort and VHSMC cohort obtained the informed consents from participants.

Data availability

The data supporting the findings of this study are available upon reasonable request.

Received: 1 September 2021; Accepted: 22 February 2022

Published online: 03 March 2022

References

- Shafiq, G. *et al.* Prevalence of sarcopenia in the world: A systematic review and meta-analysis of general population studies. *J. Diabetes Metab. Disord.* **16**, 21. <https://doi.org/10.1186/s40200-017-0302-x> (2017).
- Tanimoto, Y. *et al.* Association between sarcopenia and higher-level functional capacity in daily living in community-dwelling elderly subjects in Japan. *Arch. Gerontol. Geriatr.* **55**, e9–13. <https://doi.org/10.1016/j.archger.2012.06.015> (2012).
- Cesari, M. *et al.* Skeletal muscle and mortality results from the InCHIANTI study. *J. Gerontol. A Biol. Sci. Med. Sci.* **64**, 377–384. <https://doi.org/10.1093/gerona/gln031> (2009).
- 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. <https://doi.org/10.1111/j.1532-5415.2004.52014.x> (2004).
- McNamee, P., Bond, J., Buck, D., Resource Implications Study of the Medical Research Council Cognitive, F. & Ageing, S. Costs of dementia in England and Wales in the 21st century. *Br. J. Psychiatry* **179**, 261–266. <https://doi.org/10.1192/bjp.179.3.261> (2001).
- Chen, L. K. *et al.* Recent advances in sarcopenia research in Asia: 2016 update from the Asian Working Group for Sarcopenia. *J. Am. Med. Dir. Assoc.* **17**(767), e761–e767. <https://doi.org/10.1016/j.jamda.2016.05.016> (2016).
- Chen, L. K. *et al.* Sarcopenia in Asia: Consensus report of the Asian Working Group for Sarcopenia. *J. Am. Med. Dir. Assoc.* **15**, 95–101. <https://doi.org/10.1016/j.jamda.2013.11.025> (2014).
- Morley, J. E. *et al.* Sarcopenia with limited mobility: An international consensus. *J. Am. Med. Dir. Assoc.* **12**, 403–409. <https://doi.org/10.1016/j.jamda.2011.04.014> (2011).
- Fielding, R. A. *et al.* Sarcopenia: An undiagnosed condition in older adults. Current consensus definition: Prevalence, etiology, and consequence. International working group on sarcopenia. *J. Am. Med. Dir. Assoc.* **12**, 249–256. <https://doi.org/10.1016/j.jamda.2011.01.003> (2011).
- Cruz-Jentoft, A. J. *et al.* Sarcopenia: European consensus on definition and diagnosis: Report of the European Working Group on Sarcopenia in Older People. *Age Ageing* **39**, 412–423. <https://doi.org/10.1093/ageing/afq034> (2010).
- Chen, L. K. *et al.* Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. *J. Am. Med. Dir. Assoc.* **21**, 300–307.e302. <https://doi.org/10.1016/j.jamda.2019.12.012> (2020).
- Roubenoff, R. Sarcopenia: Effects on body composition and function. *J. Gerontol. A Biol. Sci. Med. Sci.* **58**, 1012–1017. <https://doi.org/10.1093/gerona/58.11.m1012> (2003).

13. Rolland, Y. *et al.* Sarcopenia: Its assessment, etiology, pathogenesis, consequences and future perspectives. *J. Nutr. Health Aging* **12**, 433–450. <https://doi.org/10.1007/BF02982704> (2008).
14. 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. <https://doi.org/10.1359/jbmr.1997.12.12.2076> (1997).
15. 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. <https://doi.org/10.1016/j.ajhg.2009.02.004> (2009).
16. 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. <https://doi.org/10.1038/jhg.2011.125> (2012).
17. 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. <https://doi.org/10.1007/s00439-012-1236-5> (2013).
18. 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. <https://doi.org/10.1111/accel.12228> (2014).
19. Ran, S. *et al.* Genome-wide association study identified copy number variants important for appendicular lean mass. *PLoS ONE* **9**, e89776. <https://doi.org/10.1371/journal.pone.0089776> (2014).
20. Trajanoska, K., Rivadeneira, F., Kiel, D. P. & Karasik, D. Genetics of bone and muscle interactions in humans. *Curr. Osteoporos. Rep.* **17**, 86–95. <https://doi.org/10.1007/s11914-019-00505-1> (2019).
21. Urano, T. & Inoue, S. Recent genetic discoveries in osteoporosis, sarcopenia and obesity. *Endocr. J.* **62**, 475–484. <https://doi.org/10.1507/endocrj.EJ15-0154> (2015).
22. Zillikens, M. C. *et al.* Large meta-analysis of genome-wide association studies identifies five loci for lean body mass. *Nat. Commun.* **8**, 80. <https://doi.org/10.1038/s41467-017-00031-7> (2017).
23. Han, Z. *et al.* Crystal structure of the FTO protein reveals basis for its substrate specificity. *Nature* **464**, 1205–1209. <https://doi.org/10.1038/nature08921> (2010).
24. Singh, A. N. & Gasman, B. Disentangling the genetics of sarcopenia: Prioritization of NUDT3 and KLF5 as genes for lean mass & HLA-DQB1-AS1 for hand grip strength with the associated enhancing SNPs & a scoring system. *BMC Med. Genet.* **21**, 40. <https://doi.org/10.1186/s12881-020-0977-6> (2020).
25. Khanal, P. *et al.* Prevalence and association of single nucleotide polymorphisms with sarcopenia in older women depends on definition. *Sci. Rep.* **10**, 2913. <https://doi.org/10.1038/s41598-020-59722-9> (2020).
26. Jones, G. *et al.* Sarcopenia and variation in the human leukocyte antigen complex. *J. Gerontol. A Biol. Sci. Med. Sci.* **75**, 301–308. <https://doi.org/10.1093/gerona/glz042> (2020).
27. Heffernan, S. M. *et al.* Fat mass and obesity associated (FTO) gene influences skeletal muscle phenotypes in non-resistance trained males and elite rugby playing position. *BMC Genet.* **18**, 4. <https://doi.org/10.1186/s12863-017-0470-1> (2017).
28. Hebbbar, P. *et al.* FTO variant rs1421085 associates with increased body weight, soft lean mass, and total body water through interaction with ghrelin and apolipoproteins in Arab population. *Front. Genet.* **10**, 1411. <https://doi.org/10.3389/fgene.2019.01411> (2019).
29. Jones, G. *et al.* Genome-wide meta-analysis of muscle weakness identifies 15 susceptibility loci in older men and women. *Nat. Commun.* **12**, 654. <https://doi.org/10.1038/s41467-021-20918-w> (2021).
30. Arce-Cerezo, A. *et al.* HMGA1 overexpression in adipose tissue impairs adipogenesis and prevents diet-induced obesity and insulin resistance. *Sci. Rep.* **5**, 14487. <https://doi.org/10.1038/srep14487> (2015).
31. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551. <https://doi.org/10.1093/nar/gkaa970> (2021).
32. Narasimhan, A., Greiner, R., Bathe, O. F., Baracos, V. & Damaraju, S. Differentially expressed alternatively spliced genes in skeletal muscle from cancer patients with cachexia. *J. Cachexia Sarcopenia Muscle* **9**, 60–70. <https://doi.org/10.1002/jcsm.12235> (2018).
33. Fochi, S. *et al.* Regulation of microRNAs in satellite cell renewal, muscle function, sarcopenia and the role of exercise. *Int. J. Mol. Sci.* **21**, 6732. <https://doi.org/10.3390/ijms21186732> (2020).
34. Bolotta, A. *et al.* Skeletal muscle gene expression in long-term endurance and resistance trained elderly. *Int. J. Mol. Sci.* **21**, 3988. <https://doi.org/10.3390/ijms21113988> (2020).
35. Chaves, D. F. *et al.* Comparative proteomic analysis of the aging soleus and extensor digitorum longus rat muscles using TMT labeling and mass spectrometry. *J. Proteome Res.* **12**, 4532–4546. <https://doi.org/10.1021/pr400644x> (2013).
36. Gueugneau, M. *et al.* Proteomics of muscle chronological ageing in post-menopausal women. *BMC Genomics* **15**, 1165. <https://doi.org/10.1186/1471-2164-15-1165> (2014).
37. He, H. *et al.* A systems genetics approach identified GPD1L and its molecular mechanism for obesity in human adipose tissue. *Sci. Rep.* **7**, 1799. <https://doi.org/10.1038/s41598-017-01517-6> (2017).
38. He, N. *et al.* Circulating microRNAs in plasma decrease in response to sarcopenia in the elderly. *Front. Genet.* **11**, 167. <https://doi.org/10.3389/fgene.2020.00167> (2020).
39. Moaddel, R. *et al.* Plasma biomarkers of poor muscle quality in older men and women from the Baltimore longitudinal study of aging. *J. Gerontol. A Biol. Sci. Med. Sci.* **71**, 1266–1272. <https://doi.org/10.1093/gerona/glw046> (2016).
40. Gonzalez-Freire, M. *et al.* Targeted metabolomics shows low plasma lysophosphatidylcholine 18:2 predicts greater decline of gait speed in older adults: The Baltimore longitudinal study of aging. *J. Gerontol. A Biol. Sci. Med. Sci.* **74**, 62–67. <https://doi.org/10.1093/gerona/gly100> (2019).
41. Lee, S. Y. *et al.* Comparison between dual-energy X-ray absorptiometry and bioelectrical impedance analyses for accuracy in measuring whole body muscle mass and appendicular skeletal muscle mass. *Nutrients* **10**, 738. <https://doi.org/10.3390/nu10060738> (2018).
42. McLester, C. N., Nickerson, B. S., Kliszczewicz, B. M. & McLester, J. R. Reliability and agreement of various inbody body composition analyzers as compared to dual-energy X-ray absorptiometry in healthy men and women. *J. Clin. Densitom* **23**, 443–450. <https://doi.org/10.1016/j.jocd.2018.10.008> (2020).
43. Moon, S. *et al.* The Korea biobank array: Design and identification of coding variants associated with blood biochemical traits. *Sci. Rep.* **9**, 1382. <https://doi.org/10.1038/s41598-018-37832-9> (2019).
44. Seo, S. *et al.* SNP genotype calling and quality control for multi-batch-based studies. *Genes Genomics* **41**, 927–939 (2019).
45. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
46. Song, Y. E. *et al.* ONETOOL for the analysis of family-based big data. *Bioinformatics* **34**, 2851–2853 (2018).
47. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283. <https://doi.org/10.1038/ng.3643> (2016).
48. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448. <https://doi.org/10.1038/ng.3679> (2016).
49. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287. <https://doi.org/10.1038/ng.3656> (2016).
50. Cochran, W. G. The combination of estimates from different experiments. *Biometrics* **10**, 101–129 (1954).
51. Ioannidis, J. P. Interpretation of tests of heterogeneity and bias in meta-analysis. *J. Eval. Clin. Pract.* **14**, 951–957 (2008).
52. Pruim, R. J. *et al.* LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337. <https://doi.org/10.1093/bioinformatics/btq419> (2010).

53. Raue, U. *et al.* Transcriptome signature of resistance exercise adaptations: Mixed muscle and fiber type specific profiles in young and old adults. *J. Appl. Physiol.* **1985**(112), 1625–1636. <https://doi.org/10.1152/jappphysiol.00435.2011> (2012).
54. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **57**, 289–300 (1995).

Acknowledgements

This study was supported by a VHS Medical Center Research Grant (grant no.: VHSMC20042) and an ASAN medical Center grant (2019IP0789).

KARE cohort (KARE, Korean Association Resource; Ansan/Ansung study): This study was conducted with the bioresources from National Biobank of Korea, the Korea Disease Control and Prevention Agency, Republic of Korea (KBN-2021-041). VHSMC (Veterans Health Service Medical Center) cohort: This study was conducted with bioresources from the Veterans Medical Research Institute Biobank, Republic of Korea (VBP-2020-03). Rex Soft provided technical support for Korean-Chip data quality control and supporting imputation analysis. KEGG Database Project (Kyoto Encyclopedia of Genes and Genomes): Kanehisa Laboratories for Glycerophospholipid metabolism (hsa00564, permission: 220027).

Author contributions

H.J., H.J.Y., S.W., and J.H.S. participated in the study concept and design. Y.A.K., J.H.L., Y.L., S-H.K., Y.J.S, S.H.L., J-M.K. and J.H.S. contributed to acquisition, analysis, or interpretation of data. H.J., Y.L., Y.J., A.R.D., S.W. and J.H.S. carried out statistical analysis, administrative, technical, or material support. H.J., H.J.Y., S.W. and J.H.S. wrote and revise the manuscript. J-M.K. and J.H.S. did study supervision. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-07567-9>.

Correspondence and requests for materials should be addressed to J.H.S.

Reprints and permissions information is available at www.nature.com/reprints.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022