Briefings in Bioinformatics, 2024, **25(5)**, bbae389
https://doi.org/10.1093/bib/bbae389
Problem Solving Protocol

Genetic determinants of obesity in Korean populations: exploring genome-wide associations and polygenic risk scores

Jinyeon Jo[1], Nayoung Ha[1], Yunmi Ji[2], Ahra Do[2], Je Hyun Seo[3], Bumjo Oh[4], Sungkyoung Choi[5], Eun Kyung Choe[6,7], Woojoo Lee[1,8], Jang Won Son[9,*], Sungho Won[1,2,8,10,*]

[1]Department of Public Health Sciences, Graduate school of Public Health, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, South Korea
[2]Interdisciplinary Program in Bioinformatics, College of Natural Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, South Korea
[3]Veterans Health Service Medical Center, Veterans Medical Research Institute, 53, Jinhwangdo-ro 61-gil, Gangdong-gu, Seoul, 05368, South Korea
[4]Department of Family Medicine, Seoul Metropolitan Government Seoul National University Boramae Medical Center, 20, Boramae-ro 5-gil, Dongjak-gu, Seoul, 07061, South Korea
[5]Department of Applied Mathematics, Hanyang University (ERICA), 55, Hanyang-deahak-ro, Sangnok-gu, Ansan, Gyeonggi-do, 15588, South Korea
[6]Division of Colorectal Surgery, Department of Surgery, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul, 03080, South Korea
[7]Healthcare Research Institute, Seoul National University Hospital Healthcare System Gangnam Center, 39FL, 152, Teheran-ro, Gangnam-gu, Seoul, 06236, South Korea
[8]Institute of Health and Environment, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, South Korea
[9]Division of Endocrinology, Department of Internal Medicine, Bucheon St. Mary's hospital, The Catholic University of Korea, 327, Sosa-ro, Bucheon-si, Gyeonggi-do, Bucheon, 14647, South Korea
[10]RexSoft Corps, Seoul National University Administration Building, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, South Korea

*Corresponding authors. Jang Won Son, Division of Endocrinology, Department of Internal Medicine, Bucheon St. Mary's Hospital, The Catholic University of Korea, Bucheon, South Korea. E-mail: emd76@naver.com; Sungho Won, Department of Public Health Sciences, Graduate School of Public Health, Seoul National University, Seoul, South Korea. E-mail: sunghow@gmail.com

Abstract

East Asian populations exhibit a genetic predisposition to obesity, yet comprehensive research on these traits is limited. We conducted a genome-wide association study (GWAS) with 93,673 Korean subjects to uncover novel genetic loci linked to obesity, examining metrics such as body mass index, waist circumference, body fat ratio, and abdominal fat ratio. Participants were categorized into non-obese, metabolically healthy obese (MHO), and metabolically unhealthy obese (MUO) groups. Using advanced computational methods, we developed a multifaceted polygenic risk scores (PRS) model to predict obesity. Our GWAS identified significant genetic effects with distinct sizes and directions within the MHO and MUO groups compared with the non-obese group. Gene-based and gene-set analyses, along with cluster analysis, revealed heterogeneous patterns of significant genes on chromosomes 3 (MUO group) and 11 (MHO group). In analyses targeting genetic predisposition differences based on metabolic health, odds ratios of high PRS compared with medium PRS showed significant differences between non-obese and MUO, and non-obese and MHO. Similar patterns were seen for low PRS compared with medium PRS. These findings were supported by the estimated genetic correlation (0.89 from bivariate GREML). Regional analyses highlighted significant local genetic correlations on chromosome 11, while single variant approaches suggested widespread pleiotropic effects, especially on chromosome 11. In conclusion, our study identifies specific genetic loci and risks associated with obesity in the Korean population, emphasizing the heterogeneous genetic factors contributing to MHO and MUO.

Keywords: obesity; metabolic healthy obesity; GWAS; gene-based analysis; PRS; subgroup heterogeneity

Introduction

Obesity presents a complex global health challenge that predisposes individuals to a myriad of comorbidities such as type 2 diabetes, cardiovascular disease (CVD), and certain types of cancers [1–7]. The worldwide prevalence of obesity has increased in recent decades, prompting extensive research and public health initiatives. However, obesity characteristics exhibit marked variability among different populations. Notably, East Asians (EAS) tend to manifest obesity-related complications at lower body mass index (BMI) levels than their non-Hispanic white (NHW) counterparts [8]. EAS populations also show inconsistent obesity indicators, with low BMI and high waist circumference (WC) [9, 10]. This divergence underscores the need for a distinct obesity classification standard tailored to EAS populations that differs from the NHW criteria.

High-throughput genotyping and genome-wide association studies (GWASs) have successfully identified specific genetic risk factors for obesity-related traits, treating them as polygenic conditions. Since the publication of the initial GWAS on polygenic obesity in 2007, more than 60 GWAS reports have emerged, identifying more than 1100 single nucleotide polymorphisms (SNPs) associated with obesity-related traits [11–17]. Despite the successful identification of disease susceptibility loci for obesity through GWAS, most efforts have focused on NHW populations,

Received: January 18, 2024. **Revised:** June 24, 2024

primarily BMI, with limited results for EAS [18–21]. Although some progress has been made with the emergence of large-scale biobank data [22–25], comprehensive GWAS specific to EAS and diverse obesity traits or definitions remains limited, except for meta-analyses conducted by the Asian Genetic Epidemiology Network consortium for WC ($N = 53{,}052$) and waist to hip ratio ($N = 48{,}312$) [26–32].

Furthermore, despite the established consensus linking obesity with comorbidities, observational studies have identified a subset of obese individuals with markedly reduced risk. These investigations have highlighted that the concept of metabolically healthy obesity (MHO), which denotes individuals despite being categorized as obese based on BMI, lacks the usual metabolic disturbances often linked to obesity, such as insulin resistance or dyslipidemia [33, 34]. This distinction underscores the necessity of viewing obesity beyond weight and size, emphasizing the significance of metabolic health in a comprehensive assessment [35–39]. Epidemiological studies have consistently indicated the existence of MHO, highlighting the contrasting risks between MHO and metabolically unhealthy obesity (MUO), with a reduced risk in MHO [40–46]. A few genetic investigations have supported the concept of MHO, underscoring the need to explore genetic associations that may vary based on different definitions of obesity [47–49]. However, research in this area remains limited.

The primary objective of this study was to identify novel loci predisposing individuals to obesity by focusing on diverse obesity indicators and obesity itself, within the context of GWAS. A substantial portion of obese individuals likely experience genetic susceptibility resulting from the cumulative impacts of numerous variants, each exerting modest individual effects—a 'polygenic' model akin to other intricate diseases, where most of the inherited susceptibility are attributed to polygenic inheritance involving numerous common genetic variants. The polygenic risk score (PRS) has emerged as a promising tool to address this complexity. To distinguish the genetic heterogeneity between MHO and MUO, we harnessed state-of-the-art computational algorithms for PRS with a cohort of 93,673 Korean subjects. Additionally, we explored the genetic rationale behind heterogeneity factors in obesity.

Results

Descriptive statistics of study populations

Our discovery dataset comprised 85,947 Koreans (51,317 females and 34,630 males). The descriptive statistics for this dataset are presented in Table 1. Table 1 illustrates their respective proportions of these groups in the Korean population: 66.77% non-obese, 15.98% MHO, and 17.25% MUO. The obese group was older and had a higher proportion of males. Additionally, this group exhibited higher BMI, WC, body fat (BF) ratio, and abdominal fat (AF) ratio than the non-obese group. In contrast, in the classification of the obese group based on metabolic health and obesity, the MUO group showed higher levels of obesity than the MHO group and demonstrated unhealthier biochemical levels. This suggests an increasing trend in obesity and metabolic health, in the following order: non-obese, MHO, and MUO. Notably, some indices of metabolic health were healthier in the MHO group. Consistent trends were observed across the three replication datasets, $KoGES_{Affy}$, UKB_{Chi}, and UKB_{NHW}, as shown in Supplementary Tables 1–3. The analysis scheme for study populations is illustrated in Fig. 1.

Genome-wide association studies

GWAS was conducted using the discovery dataset, and QQ and Manhattan plots depicted in Figs 2 and 3. The estimated genetic inflation factors (λ_{GC}) were notably larger than 1, but the LD score intercepts (λ_{LDSC}) were close to 1, which indicated no evidence of inflation. From the GWAS of the continuous obesity traits, we identified 20 genome-wide significant SNPs using linear regression, as presented in Table 2. Rs574367 in the *SEC16B* gene, known to be associated with BMI in the GWAS catalog, met the genome-wide significance level for WC (BETA = 0.04, $P = 7.15 \times 10^{-14}$). We also identified novel disease susceptibility loci associated with obesity. Among the 13 replicated SNPs in Table 3, rs486394, located in the *LINC02702* gene, has been reported to be associated with lipids [50], and our analyses showed its association with obesity ($P = 2.67 \times 10^{-9}$). The remaining 12 SNPs were located within or near the gene region and have been reported to be associated with BMI. They are also involved in the metabolic syndrome.

We also conducted meta-analyses that combined Korean, Japanese, and Chinese populations. QQ and Manhattan plots, illustrated in Supplementary Fig. 1, were generated from the meta-analysis. We identified rs11199833 (BETA = 0.06, $P = 4.03 \times 10^{-8}$) within the *LOC105378523* gene on chromosome 10 as a potential candidate SNP that reached significance in BF. Unfortunately, we were unable to assess the replicability because of the absence of replication datasets.

Gene-based analyses and gene cluster identification

Supplementary Table 4 presents significantly associated genes at the Bonferroni-corrected significance level $\alpha = 7.71 \times 10^{-6}$. Most genes associated with BMI or WC have been previously reported in the GWAS catalog. Table 4 displays the significant genes associated with obesity at $\alpha = 3.47 \times 10^{-6}$. *ADCY3* and *BDNF* showed significant differences in all comparisons between the non-obese and obese groups. Furthermore, the significant genes identified in the comparison between non-obese individuals and those in the combined MHO and MUO groups largely overlapped with those identified in the comparison between the non-obese and MUO groups. In addition to the *ZNF259* gene, the MHO and MUO groups exhibited distinct and significant gene sets. The MUO group was primarily associated with genes located on chromosome 3, whereas the MHO group showed significant expression of genes mainly on chromosome 11.

In the context of the obesity gene-set consisting of 24 genes identified through gene-based analysis, our gene-set analysis from STRING revealed that these proteins exhibit significant connectivity as a group ($P = 6.23 \times 10^{-12}$). This suggests a higher degree of interaction among these genes compared with that expected for a randomly selected set of proteins of the same size and degree distribution from the genome. Additionally, we identified 5 distinct clusters containing 17 genes within the obesity gene set, as shown in Fig. 4, whereas the remaining genes remained un-clustered.

SNP heritability and genetic correlation estimation

SNP heritabilites and genetic correlations for obesity-related measures were estimated using the results presented in Supplementary Table 5. It provides heritability estimates for both Koreans and EAS and illustrates the intra- and inter-class genetic correlations among Koreans and EAS and intra-class trans-ethnic genetic correlations between EAS and NHW. The highest SNP heritability was observed for BMI with $h^2 = 0.21$ ($P = 7.61 \times 10^{-72}$) for Koreans and $h^2 = 0.16$ ($P = 9.29 \times 10^{-140}$) for EAS. All inter-class genetic correlations among the obesity-related traits were significant. Moreover, the intra-class transethnic genetic correlations between EAS and NHW were statistically significant for BMI,

Table 1. Baseline characteristics according to the incident obesity on discovery dataset

Variables	Total (N = 85,947)	Non-obese (N = 51,247)	MHO (N = 12,262)	MUO (N = 13,236)	P	Missing data, N (%)
Age (years)	52.79 ± 9.34	52.86 ± 9.01	53.97 ± 8.72	54.99 ± 8.68	<.001	0 (0%)
Sex, n (%)					<.001	0 (0%)
Female	51,317 (59.71%)	33,523 (65.41%)	7024 (57.28%)	6783 (51.25%)		
Male	34,630 (39.98%)	17,724 (34.59%)	5238 (42.72%)	6453 (48.75%)		
BMI (kg/m^2)	23.92 ± 2.95	22.32 ± 1.80	26.92 ± 1.77	27.42 ± 2.12	<.001	1185 (1.38%)
WC (cm)					<.001	6834 (7.95%)
Female	78.84 ± 8.37	75.52 ± 6.47	85.66 ± 6.60	87.77 ± 7.08		
Male	85.74 ± 7.51	82.08 ± 5.96	90.73 ± 5.68	92.09 ± 6.03		
BF ratio (%)	25.56 ± 6.26	24.02 ± 5.74	27.78 ± 5.94	29.48 ± 5.96	<.001	74,017 (86.12%)
AF ratio (%)	0.88 ± 0.05	0.86 ± 0.04	0.92 ± 0.04	0.93 ± 0.04	<.001	74,017 (86.12%)
SBP (mm Hg)	121.92 ± 15.37	119.49 ± 14.92	121.67 ± 13.03	131.07 ± 15.34	<.001	6328 (7.36%)
DBP (mm Hg)	76.31 ± 10.17	74.68 ± 9.86	76.17 ± 8.95	81.99 ± 10.05	<.001	6328 (7.36%)
Fasting glucose (mg/dL)	95.19 ± 20.00	93.40 ± 18.48	92.99 ± 14.81	104.85 ± 26.00	<.001	7882 (9.17%)
HbA1c (%)	5.72 ± 0.74	5.63 ± 0.65	5.68 ± 0.59	6.06 ± 0.96	<.001	39,840 (46.35%)
Triglyceride (mg/dL)	128.20 ± 87.72	114.91 ± 77.60	107.40 ± 48.66	195.67 ± 112.99	<.001	6401 (7.45%)
High density lipoprotein (mg/dL)					<.001	6353 (7.39%)
Female	54.83 ± 13.32	56.71 ± 13.54	56.46 ± 11.17	44.98 ± 9.32		
Male	48.19 ± 11.67	50.08 ± 12.14	49.83 ± 9.74	42.09 ± 9.51		

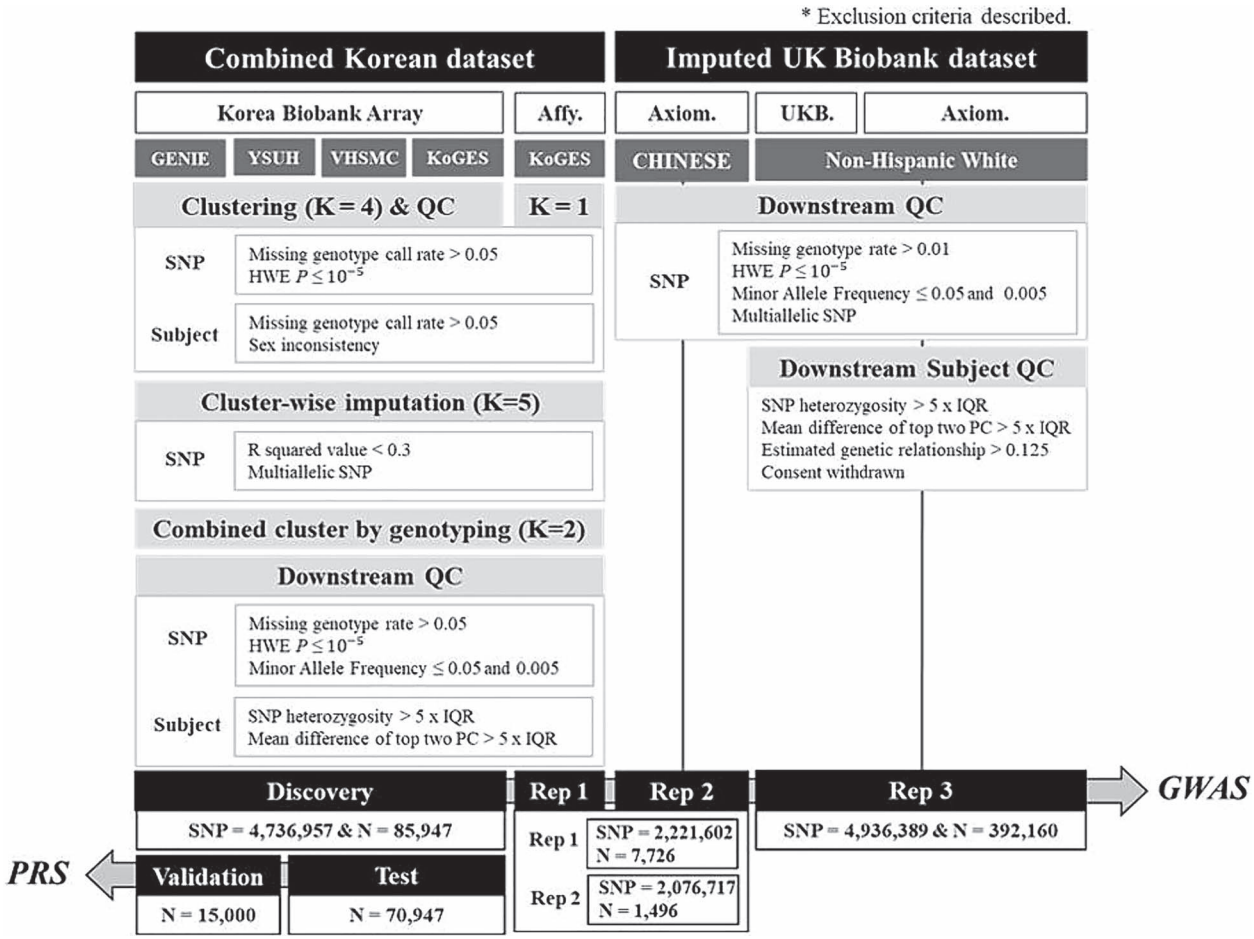

Figure 1. Data processing flowchart.

WC, and BF. All inter-class transethnic genetic correlations were greater than 0.5, with significant correlations observed between BMI and WC ($TGC_{BMI, WC} = 0.68$; $P_{BMI, WC} = 1.11 \times 10^{-16}$), BMI and BF ($TGC_{BMI, BF} = 0.66$; $P_{BF, BMI} = 2.08 \times 10^{-3}$) as well as WC and BF ($TGC_{WC, BF} = 0.55$; $P_{WC, BF} = 2.00 \times 10^{-5}$).

We calculated the SNP heritability for the obesity trait for each group. Obesity comprises three levels, and we estimated SNP heritability using summary statistics. These estimates were obtained by comparing the MHO and MUO groups with the non-obese group. When comparing the MHO and MUO groups with the

Table 2. Summary statistics of GWAS on discovery dataset; significant SNPs identified by GWASs on BMI and WC

CHR: BP	SNP	GENE	REF	ALT	MAF	HWE	MR	TRAIT	BETA	SE	P	$KoGES_{Affy}$	UKB_{Chi}	UKB_{EUR}
1:177873210	rs574367	*SEC16B*	G	T	0.25	0.74	0	BMI	0.06	0.006	3.03×10^{-28}	2.18×10^{-5}	0.82	2.48×10^{-70}
								WC	0.04	0.005	7.15×10^{-14}	2.33×10^{-3}	0.86	4.89×10^{-50}
2:632101	rs11127486	*TMEM18*	T	C	0.09	0.82	0	BMI	−0.07	0.008	2.94×10^{-17}	-	0.70	1.23×10^{-75}
2:25133148	rs6744205	*ADCY3*	T	C	0.45	0.75	0	BMI	0.04	0.005	5.52×10^{-19}	5.22×10^{-4}	-	-
2:54240105	rs6715161	*ACYP2*	C	T	0.01	0.87	0	BMI	−0.17	0.024	2.70×10^{-12}	-	-	9.32×10^{-5}
2:632101	rs11127486	*TMEM18*	T	C	0.09	0.82	0	WC	−0.05	0.008	3.74×10^{-10}	-	0.90	6.57×10^{-59}
3:52893465	rs6798941	*TMEM110*	C	T	0.37	0.5	0	BMI	0.04	0.005	5.39×10^{-17}	5.60×10^{-3}	-	-
4:45182527	rs10938397	*GNPDA2*	A	G	0.28	0.48	0	BMI	0.04	0.005	7.76×10^{-12}	0.07	0.19	6.11×10^{-41}
								WC	0.03	0.005	8.19×10^{-9}	0.04	0.35	8.14×10^{-29}
5:122750847	rs6595447	*CEP120*	T	C	0.32	0.81	0	BMI	−0.03	0.005	1.05×10^{-9}	4.78×10^{-3}	-	-
6:20686878	rs67131976	*CDKAL1*	C	T	0.46	0.03	0	BMI	−0.03	0.005	1.35×10^{-11}	0.05	0.20	1.06×10^{-6}
								WC	−0.02	0.005	4.95×10^{-8}	0.33	0.10	1.06×10^{-4}
7:138817193	rs11525873	*TTC26*	T	C	0.29	0.31	0	BMI	−0.03	0.005	5.40×10^{-9}	-	0.51	9.05×10^{-8}
10:104222963	rs12570201	*MFSD13A*	T	C	0.25	0.45	0	BMI	−0.03	0.005	3.85×10^{-9}	0.28	0.66	5.25×10^{-4}
11:8612000	rs4418812	*STK33*	A	G	0.41	0.14	0	BMI	0.03	0.005	1.32×10^{-8}	0.97	0.47	3.32×10^{-8}
11:27704209	rs34379767	*BDNF-AS*	G	A	0.46	0.01	0	BMI	−0.05	0.005	8.08×10^{-23}	2.75×10^{-3}	0.52	4.13×10^{-45}
16:20037123	rs57705530	*GPR139*	G	A	0.39	0.33	0	BMI	0.03	0.005	4.66×10^{-8}	0.02	0.28	2.44×10^{-8}
16:30093779	rs2278557	*PPP4C*	C	G	0.32	0.6	0	BMI	0.03	0.005	2.39×10^{-11}	7.86×10^{-3}	0.53	4.53×10^{-25}
16:30068354	rs9939774	*ALDOA*	C	T	0.32	0.6	0	WC	0.03	0.005	8.96×10^{-11}	0.04	0.88	2.60×10^{-30}
18:57852587	rs476828	*MC4R*	T	C	0.27	0.53	0	BMI	0.07	0.005	1.65×10^{-34}	0.01	0.75	5.14×10^{-85}

MR. SNP missing rate $KoGES_{Affy}$, UKB_{Chi}, and UKB_{EUR}. Replication *P*-value on replication dataset.

non-obese group, we found a genetic correlation of $GC_{MHO,\,MUO} = 0.99$ ($P = 0$). For the non-obese versus MHO comparison, SNP heritability is $h^2 = 0.12$ ($P = 2.06 \times 10^{-30}$); for the non-obese versus MUO, it is $h^2 = 0.11$ ($P = 1.94 \times 10^{-19}$).

Effects of PRS for BMI on obesity

For PRS analyses, the results from the validation dataset, as detailed in Supplementary Table 6, demonstrate that the PRS for BMI derived from LDpred auto yielded the best model performance. This model was employed to construct the PRS for BMI and to assess its association with obesity in the test data. As indicated in Table 5, for obesity based on metabolic health, the analysis compared the non-obese group to the obese groups comprising MHO and MUO. The PRS associations of a decrease in risk were observed in the L group compared with the M group, and an increase in risk was observed in the H group. However, it is noteworthy that a single PRS for BMI failed to differentiate risk associations between the obese groups. The effects of PRS on obesity are illustrated in Fig. 5A.

Multiple trait PRS association

We constructed a multi-PRS for measuring metabolic syndrome. The best models were selected using the correlation between the trait and PRS based on the results from the validation data presented in Supplementary Table 6. Subsequently, the selected models were applied to the test data. As demonstrated in Supplementary Table 7, the multiple trait PRS model (M2) outperformed the BMI PRS model (M1) on comparisons for obesity in terms of both AIC and area under the ROC curve (AUC). However, when comparing MHO + MUO group (case), which consists of the obese group based on BMI with non-missing metabolic traits, to the non-obese group (control), we did not observe a significant improvement in AUC.

Supplementary Figure 2 visualizes the distributions of group-wise PRS for each trait and PRSsum. For respective PRSs, the BMI PRS was well separated between non-obese and obese groups, but no significant mean difference was observed between MHO and MUO. The remaining PRSs showed slightly increasing genetic risk in the order of non-obese, MHO, and MUO, with statistically significant mean differences between the obese groups. In the case of PRSsum, the trend continued, showing higher genetic risk in MUO compared with MHO. The relationships between PRSsum and obesity are summarized in Table 5. This table reveals the consistency in risk direction across different groups based on the summed multiple trait PRS. The group denoted as H and categorized based on summed multiple trait PRS showed persistent associations across group comparisons with decreased risk levels in obesity. This finding contrasts with the limitations of a single PRS for BMI, which failed to distinguish risk association differences between obese groups despite the reported risk increase in metabolic health in the order of non-obese, MHO, and MUO. However, multiple trait PRS in the context of comparing the MHO group (control) with the MUO group (case) yielded nuanced results. It showed a decreased risk in the L group and an increased risk in the H group compared with the M group, shedding light on the complex interplay of genetic factors in different obesity subgroups. The effects of PRS on obesity are illustrated in Fig. 5B.

Heterogeneity of genetic effect on MUO and MHO

We analyzed the overall, regional, and single genetic effects for MUO and MHO (Table 6). To assess the overall genetic effect, we used the trinomial logistic regression to estimate the regression coefficient of PRS across non-obese, MHO, and MUO. The hypothesis test, which compared models with and without the constraint $H_0: \beta_{MHO} = \beta_{MUO}$, resulted in $P < 2.16 \times 10^{-16}$, thereby rejecting the null hypothesis. For the regression coefficients for the obese groups, the PRSs for BMI, fasting plasma glucose (FPG), triglycerides (TG), and high density lipoprotein (HDL) were significant for both MHO and MUO compared with the non-obese group, with differing directions except for the PRS for BMI. The PRSs for SBP and DBP were partially significant. We calculated the correlations among PRSs, which were presented in Supplementary Table 8. Using bivariate GREML, we calculated genetic correlations ($\rho = 0.89$, $P = 2.14 \times 10^{-33}$), indicating genetic similarity with minor differences.

Table 3. Summary statistics of GWAS on discovery dataset; significant SNPs identified by GWASs on obesity

								Discovery dataset				$KoGES_{Affy}$		UKB_{Chi}		UKB_{EUR}	
CHR: BP	**SNP**	**GENE**	**REF**	**ALT**	**MAF**	**HWE**	**MR**	**LRT P**	**MUL P**	**BETA**	**EQ P**	**REP P**	**BETA**	**REP P**	**BETA**	**REP P**	**BETA**
1:177873210	rs574367	*SEC16B*	G	T	0.25	0.74	0	1.73×10^{-16}	0.40	0.12	0.42	0.02	0.04	0.27	0.53	5.10×10^{-32}	0.09
										0.10			0.13		−0.30		0.08
2:25108197	rs1865689	*ADCY3*	T	C	0.43	0.62	0	3.52×10^{-10}	0.49	0.08	0.49	3.45×10^{-4}	0.12	-	-	-	-
										0.07			0.15	-	-	-	-
3:52893465	rs6798941	*STIMATE*	C	T	0.37	0.5	0	9.60×10^{-11}	0.42	0.07	0.42	0.02	−0.02	-	-	-	-
										0.08			0.11	-	-	-	-
6:20674691	rs9368219	*CDKAL1*	C	T	0.47	0.02	0	3.31×10^{-11}	5.12×10^{-3}	−0.10	4.77×10^{-3}	0.14	−0.04	3.31×10^{-7}	−0.10	0	−0.04
										−0.05			−0.08		0.49		−0.02
8:19845376	rs7841189	*LPL*	C	T	0.12	0.62	0	2.08×10^{-17}	0.00	0.10	2.97×10^{-18}	0.02	−0.10	0.15	0.64	1.86×10^{-74}	0.17
										−0.13			−0.17		0.45		−0.12
11:27703480	rs35038967	*BDNF-AS*	T	A	0.45	0.02	0	1.02×10^{-11}	0.93	−0.08	0.92	0.05	−0.09	0.96	−0.04	5.93×10^{-27}	−0.07
										−0.08			−0.08		−0.05		−0.08
11:116526322	rs486394	*LINC02702*	A	C	0.12	0.56	0	2.67×10^{-9}	3.55×10^{-9}	−0.12	3.61×10^{-9}	0.02	−0.06	-	-	-	-
										0.04			0.15	-	-	-	-
11:116651463	rs1942478	*ZPR1*	T	G	0.22	0.64	0	9.03×10^{-17}	0.00	0.10	1.25×10^{-17}	2.39×10^{-4}	0.16	0.22	0.31	1.08×10^{-113}	0.06
										−0.08			−0.13		−0.36		−0.03
11:116664776	rs1787680	*APOA5*	T	A	0.22	0.06	0	4.85×10^{-13}	6.35×10^{-14}	0.09	6.88×10^{-14}	0.05	0.13	-	-	-	-
										−0.07			−0.04	-	-	-	-
11:116837089	rs78044162	*SIK3*	C	T	0.17	0.64	0	2.25×10^{-12}	1.96×10^{-12}	0.11	1.92×10^{-12}	0.02	0.14	1.88×10^{-3}	−0.08	5.74×10^{-4}	0.06
										−0.05			−0.08		0.41		3.29×10^{-3}
15:68123915	rs4776987	*SKOR1*	A	T	0.37	0.5	0	3.81×10^{-8}	1.19×10^{-8}	0.08	4.12×10^{-3}	0.04	6.17×10^{-4}	-	-	-	-
										0.03			0.10	-	-	-	-
16:56994528	rs17231506	*CETP*	C	T	0.17	0.26	0	1.13×10^{-19}	0.00	0.13	3.96×10^{-20}	-	-	0.54	0.31	3.76×10^{-169}	0.10
										−0.08		-	-		−0.22		−0.07
18:57831468	rs633265	*MC4R*	G	T	0.28	0.66	0	1.07×10^{-18}	0.41	0.10	0.40	0.02	0.02	-	-	-	-
										0.12			0.12	-	-	-	-

MR. SNP missing rate For each SNPs, upper row and lower row represents the effects of MHO group and MUO group compared with the non-obese group, respectively. MUL P. *P*-value from multinomial regression testing for SNP effect on MHO group and MUO group compared with the non-obese group, respectively. EQ P. *P*-value from multinomial regression testing for equal SNP effect between MHO and MUO group. REP P. *P*-value from replication.

Table 4. Significant gene comparison in MHO versus MUO

CHR	GENE	P		
		Non-obese vs. MHO + MUO	Non-obese vs. MHO	Non-obese vs. MUO
2	*ADCY3*	4.96×10^{-13}	1.47×10^{-8}	3.44×10^{-6}
2	*EFR3B*	2.98×10^{-8}	3.77×10^{-7}	3.86×10^{-3}
3	*NT5DC2*	3.61×10^{-8}	6.05×10^{-3}	1.38×10^{-6}
3	*SMIM4*	2.03×10^{-8}	4.38×10^{-3}	8.00×10^{-7}
3	*PBRM1*	2.00×10^{-8}	4.63×10^{-3}	7.76×10^{-7}
3	*GNL3*	2.59×10^{-8}	6.00×10^{-3}	8.08×10^{-7}
3	*GLT8D1*	2.30×10^{-8}	4.51×10^{-3}	1.06×10^{-6}
3	*SPCS1*	2.64×10^{-8}	6.57×10^{-3}	2.41×10^{-6}
3	*NEK4*	1.80×10^{-8}	5.05×10^{-3}	7.27×10^{-7}
3	*ITIH1*	7.56×10^{-9}	5.88×10^{-4}	8.37×10^{-7}
6	*CDKAL1*	2.19×10^{-8}	1.71×10^{-9}	0.05
7	*MLXIPL*	1.64×10^{-3}	3.01×10^{-6}	0.76
8	*LPL*	0.51	1.04×10^{-5}	8.42×10^{-8}
10	*TMEM180*	2.77×10^{-6}	1.43×10^{-6}	0.08
11	*BDNF*	2.50×10^{-12}	6.74×10^{-8}	6.77×10^{-8}
11	*BUD13*	0.37	4.27×10^{-10}	9.00×10^{-6}
11	*ZNF259*	0.32	5.27×10^{-13}	1.13×10^{-7}
11	*APOA4*	0.08	5.30×10^{-7}	0.79
11	*APOA1*	0.58	8.00×10^{-7}	5.39×10^{-4}
11	*SIK3*	0.05	1.22×10^{-8}	0.22
11	*PAFAH1B2*	0.02	1.25×10^{-8}	0.49
11	*PCSK7*	0.04	2.60×10^{-7}	0.19
15	*MAP2K5*	6.45×10^{-5}	1.68×10^{-6}	0.76
15	*SKOR1*	5.48×10^{-6}	1.57×10^{-7}	0.70

Table 5. Effect of PRS on obesity and obesity-related diseases

	BMI PRS				Multiple trait PRS			
Traits	PRS group	Case N (Prev.%)	OR (95% CI)	P	PRS group	Case N (Prev.%)	OR (95% CI)	P
Non-obese	L	1495 (21.43%)	0.54 (0.51, 0.58)	3.33×10^{-87}	L	1898 (27.20%)	0.74 (0.70, 0.79)	4.32×10^{-25}
versus	M (ref.)	18,507 (33.16%)	-	-	M (ref.)	18,566 (33.25%)	-	-
Obese by BMI	H	3353 (48.06%)	1.90 (1.80, 2.00)	4.13×10^{-135}	H	2901 (41.58%)	1.44 (1.36, 1.51)	1.32×10^{-43}
Non-obese	L	1153 (18.01%)	0.65 (0.61, 0.70)	2.04×10^{-35}	L	1368 (21.32%)	0.79 (0.75, 0.85)	1.37×10^{-12}
versus	M (ref.)	12,868 (25.06%)	-	-	M (ref.)	12,889 (25.09%)	-	-
Obese by WC	H	2223 (34.48%)	1.60 (1.51, 1.69)	9.20×10^{-61}	H	1987 (30.97%)	1.34 (1.27, 1.42)	1.17×10^{-23}
Non-obese	L	1341 (21.27%)	0.54 (0.51, 0.58)	4.83×10^{-79}	L	1701 (26.97%)	0.74 (0.70, 0.78)	5.60×10^{-24}
vs. MHO+MUO	M (ref.)	16,699 (33.02%)	-	-	M (ref.)	16,766 (33.14%)	-	-
	H	3066 (48.33%)	1.93 (1.83, 2.03)	1.64×10^{-129}	H	2639 (41.69%)	1.45 (1.38, 1.53)	3.19×10^{-42}
Non-obese	L	637 (11.37%)	0.54 (0.50, 0.59)	7.37×10^{-44}	L	883 (16.09%)	0.80 (0.74, 0.86)	7.88×10^{-9}
versus	M (ref.)	7967 (19.04%)			M	8062 (19.25%)	-	-
MHO	H	1471 (30.97%)	1.93 (1.81, 2.07)	3.70×10^{-84}	(ref.)	1130 (23.44%)	1.29 (1.20, 1.39)	2.23×10^{-12}
Non-obese	L	704 (12.42%)	0.55 (0.50, 0.59)	2.65×10^{-46}	IH	818 (15.08%)	0.68 (0.63, 0.74)	1.75×10^{-21}
versus	M (ref.)	8732 (20.49%)			M (ref.)	8704 (20.47%)	-	-
MUO	H	1595 (32.73%)	1.93 (1.81, 2.06)	1.46×10^{-86}	H	1509 (29.02%)	1.60 (1.50, 1.71)	1.49×10^{-45}
MHO	L	704 (52.5%)	0.99 (0.89, 1.11)	0.90	L	818 (48.09%)	0.85 (0.77, 0.94)	1.19×10^{-3}
versus	M (ref.)	8732 (52.29%)			M (ref.)	8704 (51.91%)	-	-
MUO	H	1595 (52.02%)	1.00 (0.92, 1.08)	0.92	H	1509 (57.18%)	1.25 (1.15, 1.36)	1.36×10^{-7}

Traits MHO vs. MUO. obese group comprising of MHO (control) vs. obese group comprising of MUO (case). PRS group. L, M, and H Low, Medium, and high PRS group of bottom decile, 2nd–9th decile, and top decile. Statistical analyses. OR and *P* odds ratio and significance level of PRS group from logistic model on obesity adjusted by age, sex, PRS group, and PC1–5.

To confirm the regional genetic effect on MUO and MHO, we conducted ρ-HESS analysis, visualized in Supplementary Fig. 3. The region CHR11:116383543-117901740 showed a significant local genetic correlation ($\rho = 1.36 \times 10^{-3}$, $P = 1.94 \times 10^{-5}$), containing 22 genes: *LINC00900*, *LOC101929011*, *BUD13*, *ZPR1*, *APOA5*, *APOA4*, *APOA1*, *APOA1-AS*, *SIK3*, *PAFAH1B2*, *SIDT2*, *LOC100652768*, *PCSK7*, *RNF214*, *BACE1*, *CEP164*, *DSCAML1*, *FXYD2*, *FXYD6*, *TMPRSS13*, *IL10RA*, and *TMPRSS4-AS1*. Among these, three genes (*ZPR1*, *APOA1*, *SIK3*) and five genes (*BUD13*, *APOA1*, *APOA4*, *SIK3*, *PAFAH1B2*) overlapped with significant findings in GWAS and gene-based analyses.

We then examined the single SNP effects on obese groups. We tested whether the 13 genome-wide significant SNPs identified from GWAS have equal effects on MHO and MUO groups using trimodal logistic regression (Table 3). Eight SNPs, including four SNPs located on chromosome 11, showed $P < 0.05$ for the null hypothesis $H_0: \beta_{MHO} = \beta_{MUO}$. Rs9368219 and rs4776987 showed the same direction of effects, while the rest showed opposite

Table 6. Summary of subgroup heterogeneity analyses summary between MHO and MUO groups

Approach	Method	Data level	Result		
Overall	PRS	Individual	$P < 2.16 \times 10^{-16}$, which means reject H_0: $\beta_{MHO} = \beta_{MUO}$ for BMI, SBP, DBP, FPG, TG, and HDL PRS simultaneously		
				MHO vs. non-obese	MUO vs. non-obese
			BMI	$\beta = 0.38$ ($P < 2.16 \times 10^{-16}$)	$\beta = 0.34$ ($P < 2.16 \times 10^{-16}$)
			SBP	$\beta = -0.01$ ($P = 0.42$)	$\beta = 0.07$ ($P = 2.78 \times 10^{-5}$)
			DBP	$\beta = -0.04$ ($P = 0.01$)	$\beta = 0.01$ ($P = 0.58$)
			FPG	$\beta = -0.07$ ($P = 7.38 \times 10^{-9}$)	$\beta = 0.02$ ($P = 0.04$)
			TG	$\beta = -0.11$ ($P < 2.16 \times 10^{-16}$)	$\beta = 0.09$ ($P < 2.16 \times 10^{-16}$)
			HDL	$\beta = 0.12$ ($P < 2.16 \times 10^{-16}$)	$\beta = -0.08$ ($P = 8.65 \times 10^{-5}$)
	Bivariate GREML	Individual	GC = 0.89 ($P = 2.14 \times 10^{-33}$)		
	LDSC	Summary	GC = 0.99 ($P = 0$)		
Regional	ρ-HESS	Individual	Local GC = 1.36×10^{-3} ($P = 1.94 \times 10^{-5}$) in the region of CHR11:116383543–-117901740		
Single SNP	GWAS	Individual	EQ $P < .05$ in Table 3, which means eight SNPs among 13 genome-wide significant SNPs from GWAS have statistically significant heterogenous effects on MHO and MUO groups when testing H_0: $\beta_{MHO} = \beta_{MUO}$ for each SNP in trinomial logistic regression. Eight SNPs, four of which are in chromosome 11, include two SNPs with effect in same direction, and six in opposite direction.		

GC. Genetic correlation

direction, suggesting the presence of a subgroup within the obesity groups.

Discussion

In this study, we aimed to address the genetic components implicated in obesity by conducting GWASs and PRS analyses in the Korean population, which is a representative sample of EAS. We assumed that the genetic effect on obesity differs according to metabolic health conditions and considered four obesity-related traits and obesity based on metabolic health as outcomes to comprehensively evaluate the genetic effect on anthropometric adiposity.

We first conducted GWASs on obesity traits and identified 20 genome-wide significant SNPs in BMI and WC that were also found to be significant in NHW populations. We observed that the SNP heritabilities estimated using GWAS summary statistics ranged from 0.12 to 0.21 and were significantly larger than 0 for all traits, indicating a substantial genetic contribution to their phenotypic variability. When comparing the estimates between the EAS and NHW groups, a portion of the genetic correlations remained significant, albeit at reduced magnitudes, even after accounting for the differences in genetic architecture between these geographical genetic ancestries. Despite the disparate nature of the obesity-related traits investigated in this study (for instance, BMI serving as an anthropometric indicator, and WC as a measure of visceral fat distribution), they appear to have shared genetic influences. Consequently, it is reasonable that SNPs and genes that are significant for one obesity trait may also be significant for other obesity traits. We observed the significance of the *SMIM4* and *SPCS1* gene associations reported in Supplementary Table 4 for WHR [51, 52] in the GWAS catalog to extend to BMI in our analyses.

Contrary to the conventional assumption of homogeneity in obese participants, we observed heterogeneous SNP and gene effects between the MHO and MUO groups. This observation was consistent with epidemiological evidence [33, 53]. Regarding the SNP effects, we specifically identified that rs486394 in *LINC02702* (Long Intergenic Non-Protein Coding RNA 2702) exhibited effects of −0.12 for MHO and 0.04 for MUO. This finding is consistent with recent research indicating that lncRNAs play a role in obesity via their involvement in adipogenesis and lipid metabolism [54]. A genetic correlation of −0.99 between these groups indicates that the same causal SNPs affect both phenotypes, but in opposite directions. Intriguingly, despite these divergent SNP effects, SNP heritability estimates were comparably close in both groups: 0.12 for MHO and 0.11 for MUO, both of which were statistically significant.

On the other hand, significant genes such as *ADCY3* and *BDNF* in Table 4 were observed in all comparisons of the non-obese and obese groups, which are known to function in metabolism or energy balance, as well as the regulation of BMI and body weight. Although the mechanisms of *ZNF259* have not been extensively studied, they have been reported to be associated with the risk of CVD, which serves as the main association that distinguishes MHO from MUO (the MHO group is known to have a lower risk of CVD than the MUO group) [33, 41, 47]. The remaining genes, which are mostly significant SNPs found in GWASs, have been reported to be associated with metabolism and obesity. Considering the limitations that LDSC do not distinguish among the various scenarios depending on the direct/indirect SNP effects or direction of effects (e.g. the SNPs on phenotype 1 affect phenotype 2, or vice versa), and the unknown mechanisms of genes, we cannot ascertain the detailed mechanisms underlying the genetic factors determining obesity classifications based on metabolic health.

However, the obesity gene set listed in Table 4 displayed statistically significant enrichment in interactions compared with a randomly selected gene set. Furthermore, among the five distinct clusters identified in the cluster analysis, two clusters consisted of genes on chromosome 3 that were significantly different in both the obese and MUO groups compared with the non-obese group. The remaining three clusters mainly consisted of genes on chromosome 11 and the three genes *APOA1*, *APOA4*, and *LPL* in the yellow cluster in Fig. 4 were reported to be involved in cholesterol metabolism in the KEGG pathways (hsa04979; FDR = 0.01). While this result does not fully describe the genetic predisposition to

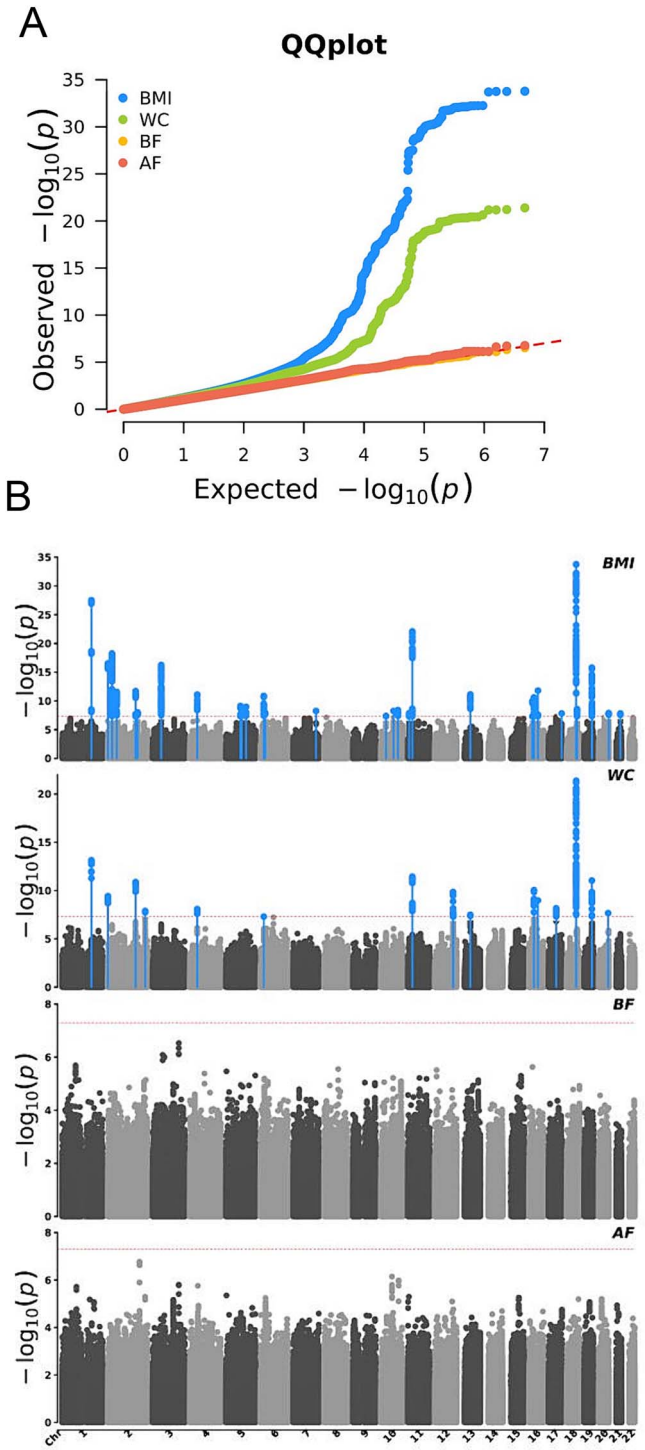

Figure 2. GWAS results on obesity-related traits on discovery dataset (A) QQ plots constructed using the GWAS summary statistics on obesity-related traits (B) Manhattan plots constructed using the GWAS summary statistics on obesity-related traits.

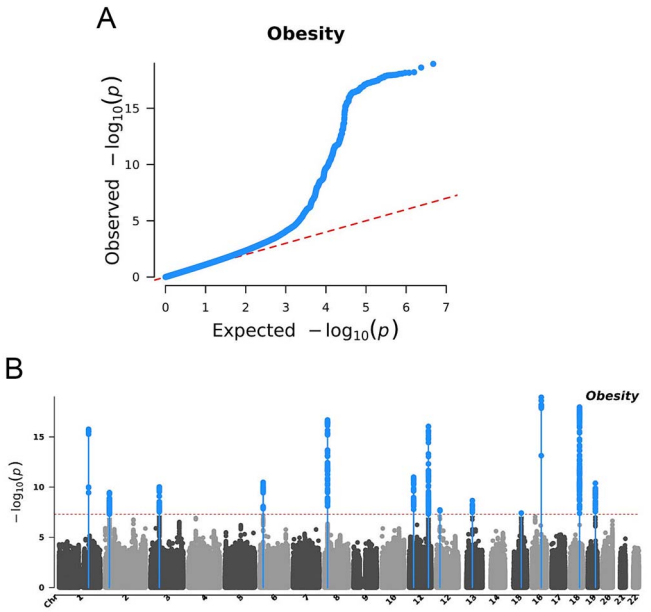

Figure 3. GWAS results on obesity on discovery dataset (A) QQ plots constructed using the GWAS summary statistics on obesity (B) Manhattan plots constructed using the GWAS summary statistics on obesity.

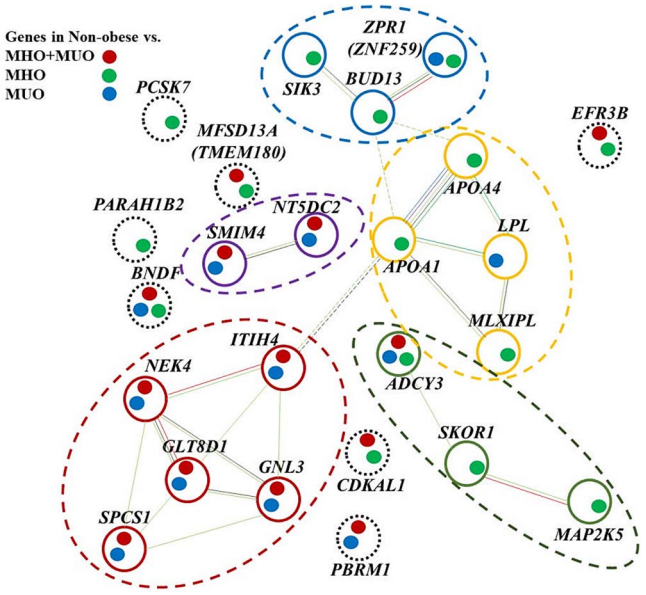

Figure 4. Gene clusters identified in MHO vs. MUO.

obesity based on metabolic health, it suggests the existence of heterogeneous genetic components between the obese groups using pleiotropic effects on metabolic traits. Given that the obesity criteria for metabolic health support the association between both obesity and metabolic traits, this appears to be a case of pleiotropy.

The inherited genetic risk of BMI was evaluated using the PRS for BMI, and we found that individuals with a high PRS for BMI were more likely to be obese. The performance of the PRS for BMI improved with multiple trait PRS. This observation suggests that the shared genetic components among the various obesity-related traits identified in this study contribute to a better understanding of obesity across different criteria. However, when it comes to subgroups within the obese category, although a single PRS for BMI may offer insights into genetic predisposition, it may not fully capture the polygenicity. The complexity of these conditions likely involves a combination of multiple genetic factors beyond BMI alone. In conclusion, the PRS for BMI proves to be valuable in assessing the inherited genetic risk of obesity; however, for subgroups within obese groups, a more comprehensive approach that incorporates multiple trait PRS and accounts for other genetic

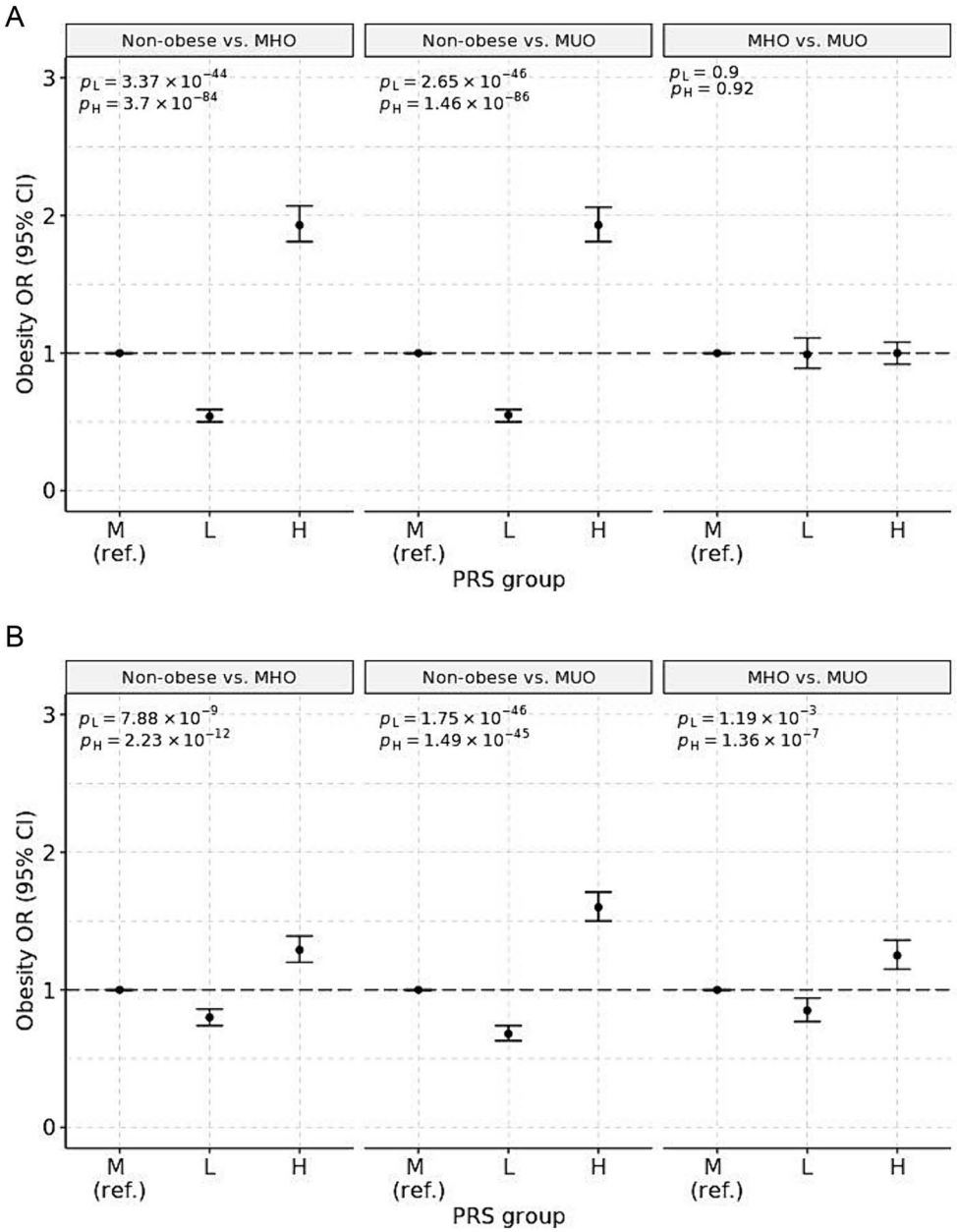

Figure 5. PRS for BMI and multiple trait PRS on obesity (A) PRS for BMI on obesity (B) Multiple trait PRS on obesity.

components may be necessary to gain a deeper understanding of their genetic underpinnings.

The heterogeneity in obese groups identified through GWAS, gene-based analyses, and the PRS trend in obesity classification was quantified by comparing the overall, regional, and single SNP effects between MHO and MUO groups. For the overall genetic effect comparison, we found significant differences in the effects of each PRS on MHO and MUO with small differences in regression coefficients. These differences were further confirmed with genetic correlations (0.89 from bivariate GREML). The regional genetic effect comparison revealed significant local genetic correlations primarily in chromosome 11, with genetic correlation of 1.36×10^{-3}. Single variant approaches indicated more widespread minor heterogeneity across the entire genome. These observations underscore some genetic differences between MHO and MUO, highlighting the complex interplay of genetic factors in obesity. Our study emphasizes the importance of considering both common and distinct genetic components in understanding MHO and MUO.

Although we unveiled novel associations between GWASs and PRS, along with post-hoc analyses, there are certain limitations that should be considered. First, our findings pertained primarily to the Korean population. Although our summary statistics were from the Japanese population, it is important to note that the Japanese and Korean populations are genetically close (GC over 0.9 on BMI). It is possible that some bias may have been introduced because of genetic differences between these populations for some phenotypes. Secondly, the absence of a universally accepted gold standard for defining MHO poses a challenge. In our study, we used moderate criteria for MHO. However, there are no established validation methods for determining the optimal definition. Despite these limitations, our research offers valuable insights into obesity by studying a large Korean population, which is the largest genetic study

conducted on this population to date. Additionally, we provide evidence of heterogeneous genetic effects within obese groups, highlighting the need to consider subpopulations within this category. Overall, these findings contribute to our understanding of genetic predisposition to obesity.

Materials and methods

Study population

We considered several prospective cohorts drawn from patients registered at our hospital. For discovery dataset, we utilized portions of the following four cohorts: the Korean Genome Epidemiology Study (KoGES) cohort ($N = 211,725$) [24] (https://biobank.nih.go.kr/cmm/main/mainPage.do), Gene-Environment of Interaction and phenotype cohort [55] (GENIE, $N = 10,300$), the YonSei University Hospital medical center (YSUH, $N = 6679$), and Veterans Health Service Medical Cent (VHSMC, $N = 2598$). The KOGES cohort is a population-based cohort comprising three subcohorts from the Korean Association REsource ($N = 10,030$), Cardiovascular Disease Association ($N = 28,338$), and Health EXAminee ($N = 173,357$) studies. This cohort was segmented into two groups based on the genotyping platform used: the Korean Biobank Array (KoreanChip) [56] and Affymetrix Genome-Wide Human SNP Array 5.0 or 6.0 (Affymetrix, Santa Clara, CA, USA). For the discovery analyses, we specifically used subjects genotyped using KoreanChip. For the GENIE, YSUH, and VHSMC cohorts, health-related information was collected from patients who visited the hospital and provided informed consent. After excluding participants with missing obesity outcomes, the discovery dataset comprised 85,947 participants and 4,736,957 SNPs.

For replication, we used three datasets: a fraction of the KoGES cohort genotyped with Affymetrix ($N = 8856$), as well as two ethnic groups of Chinese ($N = 1503$) and NHW ($N = 459,259$) from the United Kingdom Biobank (UKB, $N = 502,413$) cohort. These were denoted as $\text{KoGES}_{\text{Affy}}$, UKB_{Chi}, and UKB_{NHW}, respectively. The UKB cohort [57] (https://www.ukbiobank.ac.uk) was a large-scale prospective cohort study that included subjects residing in the UK. The baseline survey for this cohort commenced in 2006, and follow-up is still in progress. All participants were aged 40–69 years at baseline and provided electronically signed consent. Each cohort was designed to examine both genetic and phenotypic data in order to investigate a wide range of risk factors associated with common complex diseases. The final replication datasets for $\text{KoGES}_{\text{Affy}}$, UKB_{Chi}, and UKB_{NHW} included 7726 Koreans with 2,221,602 SNPs, 1496 Chinese individuals with 2,076,717 SNPs, and 392,160 NHW with 4,936,389 SNPs.

Obesity-related measures and operational definition of metabolic health and obesity

Several obesity metrics including BMI, WC, BF, and AF were evaluated. For GWASs, continuous measurements were used for BMI and WC. In the context of PRS analyses, BMI was dichotomized to cases (obese group) if BMI ≥ 25 kg/m^2, while those with lower BMIs were defined as controls. For WC, females with a measurement $\geq$85 cm and males with a measurement $\geq$90 cm were designated as cases [58].

Participants with no missing biochemical observations were further divided into three groups based on obesity as defined by BMI and metabolic health status: non-obese, MHO, and MUO. The EAS with a BMI less than 25 kg/m^2 and NHW with a BMI less than 30 kg/m^2 were classified as non-obese, whereas the others were classified as obese. An individual was considered MHO if they were obese but had fewer than or equal to two of the following metabolic risk conditions: (i) elevated blood pressure [either systolic blood pressure (SBP)/diastolic blood pressure (DBP) $\geq$ 130/85 mm Hg or taking anti-hypertensive medication], (ii) impaired FPG ($\geq$100 mg/dL or a diagnosis of diabetes mellitus, or prescription for antidiabetic medication), (iii) high plasma TG ($\geq$150 mg/dL), and (iv) low HDL cholesterol (<40 mg/dL in men or <50 mg/dL in women) [59]. Based on these criteria, participants were either classified as MHO (obese with fewer than or equal to two metabolic risk factors) or MUO (obese with at least three metabolic risk factors).

Genotyping, quality control, and imputation

KoGES included 81,153 participants. Of these, 72,297 were genotyped using KoreanChip and the remaining 8856 participants were genotyped using Affymetrix. Participants in the GENIE, YSUH, and VHSMC cohorts were genotyped using KoreanChip. The genotypes generated by KoreanChip were called using the K-medoid algorithm [60] to minimize heterogeneities between batches and studies.

For downstream quality control, SNPs were removed if the missing genotype call rates were >0.05 or the Hardy Weinberg Equilibrium (HWE) was $P < 10^{-5}$. Participants were excluded if they had missing genotype call rates >0.05 or there was a sex inconsistency. The remaining SNPs were prephased with Eagle v2.4 [61] and untyped SNPs were imputed using the Northeast Asian Reference Database [62] imputation server (https://nard.macrogen.com/). Imputed SNPs were removed if $R^2 < 0.3$ or the number of alleles $\neq 2$, missing genotype rate > 0.05, minor allele frequency (MAF) < 0.005 (KoreanChip), MAF < 0.05 (Affymetrix), or HWE $P < 10^{-5}$. Participants were excluded if SNP heterozygosity was $>5 \times$ IQR or if the mean difference between the top two principal components (PCs) was $>5 \times$ IQR.

The imputed genotypes were downloaded from the UKB. SNPs were removed if the missing genotype rate > 0.01, P for HWE $< 10^{-6}$, or the number of alleles $\neq 2$. We considered participants of Chinese and European ancestry, with NHW participants identified as having a White, British, Irish, and any other White background. For the Chinese group, we eliminated SNPs with an MAF < 0.05. In the NHW group, we excluded participants if MAF < 0.005 and if SNP heterozygosity $>5 \times$ IQR, estimated genetic relationship >0.125 or mean difference of top two PCs $> 5 \times$ IQR.

All data management and quality control processes were performed using PLINK [63], PLINK2 [64], GCTA [65], and ONETOOL [66].

Genome-wide association studies

In order to identify the loci associated with obesity, we conducted a GWAS on obesity-related variables using the discovery data. Linear regression was used for continuous outcomes, incorporating age, sex, and the first 10 PCs as covariates. For categorical outcomes, binary or multinomial logistic regression was performed using the same covariates as those used in the linear regression. For the multinomial logistic regression, association analyses were conducted using a likelihood ratio test (LRT). Linear and logistic regressions were conducted using PLINK and Rex [67] and the VGAM package in R version 1.1.7. The significance level was set at $\alpha = 5 \times 10^{-8}$. Significant genome-wide SNPs were annotated using ANNOVAR [68]. Furthermore, we performed two meta-analyses. The first involved GWASs for the discovery dataset and replication dataset of $\text{KoGES}_{\text{Affy}}$ by using METAL [69]. The second incorporated all the GWAS results for EAS, including the publicly accessible Biobank Japan (BBJ) summary statistics for BMI ($N = 158,284$).

Gene-based analyses and gene cluster identification

We conducted gene-based analyses of 18,432 genes within the discovery datasets, $KoGES_{Affy}$, and UKB_{Chi} using FUMA [70] with 1000G Phase 3 EAS as a reference population. We performed the same analyses on 14,411 genes using the logistic regression results from the discovery dataset. This allowed us to compare the obese group comprising both the MHO and MUO groups and the MHO and MUO groups with the non-obese group. For UKB_{NHW}, on the other hand, we utilized the 1000G Phase 3 EUR.

Specifically for the analysis comparing MHO, MUO, or both groups with non-obese group, we expanded our analysis by performing gene-set and gene cluster analyses using the Search Tool for the Retrieval of Interacting Genes (STRING version 12.0) [71]. For the gene set enrichment analysis, the whole genome was used as the statistical background. For the gene network clustering analysis, we applied the Markov Cluster Algorithm with the inflation parameter set to three.

SNP heritability and genetic correlation

GWAS results were used in conjunction with LDSC [72] to estimate SNP heritability and genetic correlations within EAS populations. The transient genetic correlation between the EAS and NHW populations was calculated using POPCORN [73]. In cases where there was an overlap among participants used to calculate summary statistics, for instance, regression coefficients between MHO compared with non-obese individuals and between MHO compared with non-obese individuals were adjusted for correlations between summary statistics using Erase Sample Overlap and Relatedness (EraSOR) [74], resulting in 595,220 SNPs and then incorporated into LDSC to adjust correlations between summary statistics.

PRS and prediction model building

The construction of PRS requires summary statistics from the GWAS. We used GWAS data on BMI from the BBJ [75] (http://jenger.riken.jp/en), focusing on SNPs with MAF > 0.005. The remaining 5,925,388 SNPs were used for the PRS calculation. The PRS model has been constructed using various methods, including clumping + thresholding (CT) [76], LDpred with infinitesimal, grid, and auto models [77, 78], lassosum [79], and PRScs [80]. The EAS subpopulation from 1000 Genomes Phase 3 was used as the linkage disequilibrium reference panel. For LDpred grid model, proportions of causal SNP ρ were set at 1%, 3%, 10%, 30%, and 100% with the default value used for the other options.

PRS analyses require both validation and testing. For validation, we randomly selected 15,000 participants from all Koreans genotyped using KoreanChip, while the remaining 70,947 Koreans were used as test data. The validation data underwent 10-fold cross-validation for BMI, and we used a logistic regression model to create a prediction model incorporating baseline age, sex, BMI RPS, and PC1–5 as covariates.

To enhance the accuracy of the prediction model, we considered a multiple trait PRS approach [81]. This approach included PRSs for BMI, SBP, DBP, FPG, TG, and HDL as covariates. The PRS models were constructed using summary statistics downloaded from the BBJ. These multiple trait PRS models were validated using the same method used for the BMI. These multiple trait PRS were used to develop a predictive model for obesity. The best method for the PRS was selected using the correlation between the trait and PRS. The prediction accuracy of the selected model was evaluated using the test data, and the AUCs between the models based on BMI PRS and multiple trait PRS were compared using the DeLong test [82].

Association analyses of PRS

The association between the PRS and obesity was investigated using logistic regression analysis. We considered two distinct PRSs: one for BMI and another PRSsum [83] that combined multiple trait PRS for BMI, SBP, DBP, FPG, TG, and HDL with equal weights. From the constructed PRSs, we compared the PRS among the groups using pairwise T-tests to determine the mean differences between MHO and MUO. PRSs were categorized into three groups: low (L), medium (M), and high (H). Here 'L' represents the lower 10%, 'M' encapsulates the range of 10%–90%, and 'H' denotes the upper 10%.

Heterogeneity of genetic effect on MHO and MUO

We compared the overall, regional, and single genetic effects between MHO and MUO groups. To compare the overall genetic effect, we tested the hypothesis that the effects of PRSs for BMI, SBP, DBP, FPG, TG, and HDL on MHO and MUO were the same compared with the non-obese group (i.e. $\beta_{MHO} = \beta_{MUO}$ for BMI PRS, etc.) using trinomial regression (VGAM package in R version 1.1.7). Additionally, genetic correlations were measured using bivariate GREML analysis [84] implemented in GCTA for individual-level data. Regional genetic effects were compared using local genetic correlation using ρ-HESS [85], which quantifies the correlation between traits in specific genomic regions, with significance determined using Bonferroni-corrected $\alpha = 3.48 \times 10^{-5}$. Lastly, single SNP genetic effects were compared using trinomial regression with restriction $\beta_{MHO} = \beta_{MUO}$. SNPs with $P < 0.05$ were identified as having heterogenous SNP effects.

Key Points

- Conventional genetic studies on obesity face two primary limitations. Firstly, extensive research has predominantly focused on the genetic architectures of European ancestry White populations, neglecting the diverse genetic makeup of other populations. Secondly, despite observational studies revealing subgroups within the obese category—such as metabolically healthy or unhealthy obesity—traditional classifications often oversimplify obesity as either non-obese or obese.
- Our aim was to unravel the underlying genetic components associated with multi-status obesity classifications within the East Asian population. To achieve this, we conducted comprehensive genome-wide association studies (GWAS) and subsequent post-hoc analyses. Additionally, we employed polygenic risk score (PRS) associations using both BMI-centric and combined multiple trait approaches.
- Our study revealed distinct genetic components contributing to the metabolic health gap observed within obese groups, as identified by GWAS and post-hoc analyses. Furthermore, the genetic risks associated with these components were ranked using PRS.

Supplementary data

Supplementary data is available at *Briefings in Bioinformatics* online.

Author contributions

Jang Won Son and Sungho Won contributed to the design of the study and revised the manuscript. Jinyeon Jo analyzed the data and wrote the manuscript. Nayoung Ha, Yunmi Ji, Ahra Do, Je Hyun Seo, Bumjo Oh, Sungkyung Choi, Eun Kyung Choe, and Woojoo Lee reviewed the manuscript.

Conflict of interest: The authors declare that they have no competing interest.

Funding

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI22C0154), the National Research Foundation (NRF) grants (NRF-2021R1A5A1033157) and (RS-2024-00346850) funded by the Korean government (MSIT). Statistical analyses were supported by the National Supercomputing Center with supercomputing resources including technical support (KSC-2022-CRE-0319 and KSC-2023-CRE-0117). This study was conducted with bioresources from National Biobank of Korea, the Korea Disease Control and Prevention Agency, Republic of Korea (NBK-2020-101). The committee of VHS Biobank (VBP-2020-03) approved the use of bioresources for this study. This study was supported by Research Grant (Grant No. KSSO-D-2021002) from Korean Society for the Study of Obesity.

Data availability

The data used in this study are sourced from various repositories, including KoGES (Korean Genome Epidemiology Study), GENIE (Gene–Environment of Interaction and Phenotype Cohort), YSUH (YonSei University Hospital Medical Center), VHSMC (Veterans Health Service Medical Center), and UK Biobank. For researchers interested in accessing the KoGES and UK Biobank data, these datasets are available through institutional data access and can be requested by eligible researchers. Detailed information and access criteria for KoGES can be found at KoGES Data Access (https://biobank.nih.go.kr/cmm/main/mainPage.do), and UK Biobank data access can be requested via UK Biobank (https://www.ukbiobank.ac.uk/). Access to GENIE, YSUH, and WHSMC data is facilitated through collaboration with researchers in the respective hospitals. Additionally, summary statistics from the BioBank Japan (BBJ) project can be downloaded from https://pheweb.jp/. We encourage interested researchers to follow the provided links for more information on data availability and access procedures.

References

1. Afshin A, Forouzanfar MH, Reitsma MB. *et al.* Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med* 2017;**377**:13–27. https://doi.org/10.1056/NEJMoa1614362.
2. Jiang L, Rong J, Wang Y. *et al.* The relationship between body mass index and hip osteoarthritis: a systematic review and meta-analysis. *Joint Bone Spine* 2011;**78**:150–5. https://doi.org/10.1016/j.jbspin.2010.04.011.
3. Jiang L, Tian W, Wang Y. *et al.* Body mass index and susceptibility to knee osteoarthritis: a systematic review and meta-analysis. *Joint Bone Spine* 2012;**79**:291–7. https://doi.org/10.1016/j.jbspin.2011.05.015.
4. Lauby-Secretan B, Scoccianti C, Loomis D. *et al.* Body fatness and cancer — viewpoint of the IARC working group. *N Engl J Med* 2016;**375**:794–8. https://doi.org/10.1056/NEJMsr1606602.
5. Must A, Spadano J, Coakley EH. *et al.* The disease burden associated with overweight and obesity. *JAMA* 1999;**282**:1523–9. https://doi.org/10.1001/jama.282.16.1523.
6. Singh GM, Danaei G, Farzadfar F. *et al.* The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. *PLoS One* 2013;**8**:e65174.
7. Wormser D, Kaptoge S, Angelantonio E. *et al.* Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. *Lancet* 2011;**377**:1085–95. https://doi.org/10.1016/S0140-6736(11)60105-0.
8. Sun C, Kovacs P, Guiu-Jurado E. Genetics of obesity in East Asians. *Front Genet* 2020;**11**:575049. https://doi.org/10.3389/fgene.2020.575049.
9. Deurenberg P, Deurenberg-Yap M, Guricci S. Asians are different from Caucasians and from each other in their body mass index/-body fat per cent relationship. *Obes Rev* 2002;**3**:141–6. https://doi.org/10.1046/j.1467-789X.2002.00065.x.
10. Yoon KH, Lee JH, Kim JW. *et al.* Epidemic obesity and type 2 diabetes in Asia. *Lancet* 2006;**368**:1681–8. https://doi.org/10.1016/S0140-6736(06)69703-1.
11. Buniello A, MacArthur JAL, Cerezo M. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019;**47**:D1005–d1012. https://doi.org/10.1093/nar/gky1120.
12. Locke AE, Kahali B, Berndt SI. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;**518**:197–206. https://doi.org/10.1038/nature14177.
13. Shungin D, Winkler TW, Croteau-Chonka DC. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015;**518**:187–96. https://doi.org/10.1038/nature14132.
14. Maes HHM, Neale MC, Eaves LJ. Genetic and environmental factors in relative body weight and human adiposity. *Behav Genet* 1997;**27**:325–51. https://doi.org/10.1023/A:1025635913927.
15. Magnusson PKE, Rasmussen F. Familial resemblance of body mass index and familial risk of high and low body mass index. A study of young men in Sweden. *Int J Obes Relat Metab Disord* 2002;**26**:1225–31.
16. Rose KM, Newman B, Mayer-Davis EJ. *et al.* Genetic and behavioral determinants of waist-hip ratio and waist circumference in women twins. *Obes Res* 1998;**6**:383–92. https://doi.org/10.1002/j.1550-8528.1998.tb00369.x.
17. Wardle J, Carnell S, Haworth CM. *et al.* Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. *Am J Clin Nutr* 2008;**87**:398–404. https://doi.org/10.1093/ajcn/87.2.398.
18. Justice AE, Winkler TW, Feitosa MF. *et al.* Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nat Commun* 2017;**8**:14977. https://doi.org/10.1038/ncomms14977.
19. Graff M, Scott RA, Justice AE. *et al.* Genome-wide physical activity interactions in adiposity - a meta-analysis of 200,452 adults. *PLoS Genet* 2017;**13**:e1006528. https://doi.org/10.1038/ncomms14977.
20. Hoffmann TJ, Choquet H, Yin J. *et al.* A large multiethnic genome-wide association study of adult body mass index identifies novel loci. *Genetics* 2018;**210**:499–515. https://doi.org/10.1534/genetics.118.301479.
21. Turcot V, Lu Y, Highland HM. *et al.* Protein-altering variants associated with body mass index implicate pathways that control

energy intake and expenditure in obesity. *Nat Genet* 2018;**50**:26–41. https://doi.org/10.1038/s41588-017-0011-x.
22. Chen Z, Chen J, Collins R. *et al.* China Kadoorie biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 2011;**40**:1652–66. https://doi.org/10.1093/ije/dyr120.
23. Nagai A, Hirata M, Kamatani Y. *et al.* Overview of the BioBank Japan project: study design and profile. *J Epidemiol* 2017;**27**:S2–s8. https://doi.org/10.1016/j.je.2016.12.005.
24. Kim Y, Han B-G, KoGES Group. Cohort profile: the Korean genome and epidemiology study (KoGES) consortium. *Int J Epidemiol* 2017;**46**:e20. https://doi.org/10.1093/ije/dyv316.
25. Fan CT, Lin JC, Lee CH. Taiwan biobank: a project aiming to aid Taiwan's transition into a biomedical island. *Pharmacogenomics* 2008;**9**:235–46. https://doi.org/10.2217/14622416.9.2.235.
26. Zhu Z, Li J, Si J. *et al.* A large-scale genome-wide association analysis of lung function in the Chinese population identifies novel loci and highlights shared genetic aetiology with obesity. *Eur Respir J* 2021;**58**:2100199. https://doi.org/10.1183/13993003.00199-2021.
27. Sakaue S, Kanai M, Tanigawa Y. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet* 2021;**53**:1415–24. https://doi.org/10.1038/s41588-021-00931-x.
28. Nam K, Kim J, Lee S. Genome-wide study on 72,298 individuals in Korean biobank data for 76 traits. *Cell Genomics* 2022;**2**:100189. https://doi.org/10.1016/j.xgen.2022.100189.
29. Chen C-Y, Chen T-T, Feng Y-CA. *et al.* Analysis across Taiwan biobank, biobank Japan and UK biobank identifies hundreds of novel loci for 36 quantitative traits. *medRxiv* 2021:21255236. https://doi.org/10.1016/j.xgen.2022.100189.
30. Wen W, Cho YS, Zheng W. *et al.* Meta-analysis identifies common variants associated with body mass index in East Asians. *Nat Genet* 2012;**44**:307–11. https://doi.org/10.1038/ng.1087.
31. Wen W, Zheng W, Okada Y. *et al.* Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum Mol Genet* 2014;**23**:5492–504. https://doi.org/10.1093/hmg/ddu248.
32. Wen W, Kato N, Hwang J-Y. *et al.* Genome-wide association studies in East Asians identify new loci for waist-hip ratio and waist circumference. *Sci Rep* 2016;**6**:17958. https://doi.org/10.1038/srep17958.
33. Blüher M. Metabolically healthy obesity. *Endocr Rev* 2020;**41**. https://doi.org/10.1210/endrev/bnaa004.
34. Tsatsoulis A, Paschou SA. Metabolically healthy obesity: criteria, epidemiology, controversies, and consequences. *Curr Obes Rep* 2020;**9**:109–20. https://doi.org/10.1007/s13679-020-00375-0.
35. Vague J. The degree of masculine differentiation of obesities: a factor determining predisposition to diabetes, atherosclerosis, gout, and uric calculous disease. *Am J Clin Nutr* 1956;**4**:20–34. https://doi.org/10.1093/ajcn/4.1.20.
36. Sims EA. Are there persons who are obese, but metabolically healthy? *Metabolism* 2001;**50**:1499–504. https://doi.org/10.1053/meta.2001.27213.
37. Stefan N, Häring HU, Hu FB. *et al.* Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. *Lancet Diabetes Endocrinol* 2013;**1**:152–62. https://doi.org/10.1016/S2213-8587(13)70062-7.
38. Stefan N, Häring HU, Schulze MB. Metabolically healthy obesity: the low-hanging fruit in obesity treatment? *Lancet Diabetes Endocrinol* 2018;**6**:249–58. https://doi.org/10.1016/S2213-8587(17)30292-9.
39. Stefan N, Kantartzis K, Machann J. *et al.* Identification and characterization of metabolically benign obesity in humans. *Arch Intern Med* 2008;**168**:1609–16. https://doi.org/10.1001/archinte.168.15.1609.
40. Tajik S, Mirzababaei A, Ghaedi E. *et al.* Risk of type 2 diabetes in metabolically healthy people in different categories of body mass index: an updated network meta-analysis of prospective cohort studies. *J Cardiovasc Thorac Res* 2019;**11**:254–63. https://doi.org/10.15171/jcvtr.2019.43.
41. Hinnouho G-M, Czernichow S, Dugravot A. *et al.* Metabolically healthy obesity and the risk of cardiovascular disease and type 2 diabetes: the Whitehall II cohort study. *Eur Heart J* 2014;**36**:551–9. https://doi.org/10.15171/jcvtr.2019.43.
42. Zheng X, Peng R, Xu H. *et al.* The association between metabolic status and risk of cancer among patients with obesity: metabolically healthy obesity vs. metabolically unhealthy obesity. *Front Nutr* 2022;**9**:783660.
43. Murphy N, Cross AJ, Abubakar M. *et al.* A nested case-control study of metabolically defined body size phenotypes and risk of colorectal cancer in the European prospective investigation into cancer and nutrition (EPIC). *PLoS Med* 2016;**13**:e1001988. https://doi.org/10.1371/journal.pmed.1001988.
44. Kim JW, Ahn ST, Oh MM. *et al.* Incidence of prostate cancer according to metabolic health status: a nationwide cohort study. *J Korean Med Sci* 2019;**34**:e49. https://doi.org/10.3346/jkms.2019.34.e49.
45. Kim JW, Ahn ST, Oh MM. *et al.* Increased incidence of bladder cancer with metabolically unhealthy status: analysis from the National Health Checkup database in Korea. *Sci Rep* 2020;**10**:6476. https://doi.org/10.1038/s41598-020-63595-3.
46. Chung HS, Lee JS, Song E. *et al.* Effect of metabolic health and obesity phenotype on the risk of pancreatic cancer: a Nationwide population-based cohort study. *Cancer Epidemiol Biomarkers Prev* 2021;**30**:521–8. https://doi.org/10.1158/1055-9965.EPI-20-1262.
47. Huang LO, Loos RJF, Kilpeläinen TO. Evidence of genetic predisposition for metabolically healthy obesity and metabolically obese normal weight. *Physiol Genomics* 2018;**50**:169–78. https://doi.org/10.1152/physiolgenomics.00044.2017.
48. Park J-M, Park D-H, Song Y. *et al.* Understanding the genetic architecture of the metabolically unhealthy normal weight and metabolically healthy obese phenotypes in a Korean population. *Sci Rep* 2021;**11**:2279. https://doi.org/10.1038/s41598-021-81940-y.
49. Piko P, Llanaj E, Nagy K. *et al.* Genetic background of metabolically healthy and unhealthy obesity phenotypes in Hungarian adult sample population. *Int J Mol Sci* 2023;**24**:5209. https://doi.org/10.3390/ijms24065209.
50. León-Mimila P, Villamil-Ramírez H, Macías-Kauffer LR. *et al.* Genome-wide association study identifies a functional SIDT2 variant associated with HDL-C (high-density lipoprotein cholesterol) levels and premature coronary artery disease. *Arterioscler Thromb Vasc Biol* 2021;**41**:2494–508. https://doi.org/10.1161/ATVBAHA.120.315391.
51. Christakoudi S, Evangelou E, Riboli E. *et al.* GWAS of allometric body-shape indices in UK biobank identifies loci suggesting associations with morphogenesis, organogenesis, adrenal cell renewal and cancer. *Sci Rep* 2021;**11**:10688. https://doi.org/10.1038/s41598-021-89176-6.
52. Pulit SL, Stoneman C, Morris AP. *et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649

individuals of European ancestry. *Hum Mol Genet* 2019;**28**:166–74. https://doi.org/10.1093/hmg/ddy327.
53. Tsatsoulis A, Paschou SA. Metabolically healthy obesity: criteria. *Curr Obes Rep* 2020;**9**:109–20. https://doi.org/10.1007/s13679-020-00375-0.
54. Fontanini M, Cabiati M, Giacomarra M. *et al.* Long non-coding RNAs and obesity: new potential pathogenic biomarkers. *Curr Pharm Des* 2022;**28**:1592–605. https://doi.org/10.2174/1381612828666220211153304.
55. Lee C, Choe EK, Choi JM. *et al.* Health and prevention enhancement (H-PEACE): a retrospective, population-based cohort study conducted at the Seoul National University Hospital Gangnam Center. *BMJ Open* 2018;**8**:e019327. https://doi.org/10.1136/bmjopen-2017-019327.
56. Moon S, Kim YJ, Han S. *et al.* The Korea biobank Array: design and identification of coding variants associated with blood biochemical traits. *Sci Rep* 2019;**9**:1382. https://doi.org/10.1038/s41598-018-37832-9.
57. Bycroft C, Freeman C, Petkova D. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;**562**:203–9. https://doi.org/10.1038/s41586-018-0579-z.
58. Kim BY, Kang SM, Kang JH. *et al.* 2020 Korean Society for the Study of Obesity Guidelines for the Management of Obesity in Korea. *J Obes Metab Syndr* 2021;**30**:81–92. https://doi.org/10.7570/jomes21022.
59. Smith GI, Mittendorfer B, Klein S. Metabolically healthy obesity: facts and fantasies. *J Clin Invest* 2019;**129**:3978–89. https://doi.org/10.1172/JCI129186.
60. Seo S, Park K, Lee JJ. *et al.* SNP genotype calling and quality control for multi-batch-based studies. *Genes Genomics* 2019;**41**:927–39. https://doi.org/10.1007/s13258-019-00827-5.
61. Loh P-R, Danecek P, Palamara PF. *et al.* Reference-based phasing using the haplotype reference consortium panel. *Nat Genet* 2016;**48**:1443–8. https://doi.org/10.1038/ng.3679.
62. Yoo S-K, Kim C-U, Kim HL. *et al.* NARD: whole-genome reference panel of 1779 northeast Asians improves imputation accuracy of rare and low-frequency variants. *Genome Med* 2019;**11**:64. https://doi.org/10.1186/s13073-019-0677-z.
63. Purcell S, Neale B, Todd-Brown K. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559–75. https://doi.org/10.1086/519795.
64. Chang CC, Chow CC, Tellier LC. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 2015;**4**:7. https://doi.org/10.1186/s13742-015-0047-8.
65. Yang J, Lee SH, Goddard ME. *et al.* GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011;**88**:76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.
66. Song YE, Lee S, Park K. *et al.* ONETOOL for the analysis of family-based big data. *Bioinformatics* 2018;**34**:2851–3. https://doi.org/10.1093/bioinformatics/bty180.
67. Lee B, An J, Lee S. *et al.* Rex: R-linked EXcel add-in for statistical analysis of medical and bioinformatics data. *Genes Genomics* 2023;**45**:295–305. https://doi.org/10.1007/s13258-022-01361-7.
68. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;**38**:e164. https://doi.org/10.1093/nar/gkq603.
69. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;**26**:2190–1. https://doi.org/10.1093/bioinformatics/btq340.
70. Watanabe K, Taskesen E, van Bochoven A. *et al.* Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017;**8**:1826. https://doi.org/10.1038/s41467-017-01261-5.
71. Szklarczyk D, Gable AL, Lyon D. *et al.* STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2018;**47**:D607–13. https://doi.org/10.1093/nar/gky1131.
72. Bulik-Sullivan BK, Loh P-R, Finucane HK. *et al.* LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015;**47**:291–5. https://doi.org/10.1038/ng.3211.
73. Brown BC, Ye CJ, Price AL. *et al.* Transethnic genetic-correlation estimates from summary statistics. *Am J Hum Genet* 2016;**99**:76–88. https://doi.org/10.1016/j.ajhg.2016.05.001.
74. Choi SW, Mak TSH, Hoggart CJ. *et al.* EraSOR: a software tool to eliminate inflation caused by sample overlap in polygenic score analyses. *GigaScience* 2023;**12**:giad043. https://doi.org/10.1093/gigascience/giad043.
75. Kanai M, Akiyama M, Takahashi A. *et al.* Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet* 2018;**50**:390–400. https://doi.org/10.1038/s41588-018-0047-6.
76. Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. *Bioinformatics* 2014;**31**:1466–8. https://doi.org/10.1093/bioinformatics/btu848.
77. Vilhjálmsson BJ, Yang J, Finucane HK. *et al.* Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet* 2015;**97**:576–92. https://doi.org/10.1016/j.ajhg.2015.09.001.
78. Privé F, Arbel J, Vilhjálmsson BJ. LDpred2: better, faster, stronger. *Bioinformatics* 2020;**36**:5424–31. https://doi.org/10.1093/bioinformatics/btaa1029.
79. Mak TSH, Porsch RM, Choi SW. *et al.* Polygenic scores via penalized regression on summary statistics. *Genet Epidemiol* 2017;**41**:469–80. https://doi.org/10.1002/gepi.22050.
80. Ge T, Chen C-Y, Ni Y. *et al.* Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun* 2019;**10**:1776. https://doi.org/10.1038/s41467-019-09718-5.
81. Krapohl E, Patel H, Newhouse S. *et al.* Multi-polygenic score approach to trait prediction. *Mol Psychiatry* 2018;**23**:1368–74. https://doi.org/10.1038/mp.2017.163.
82. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**:837–45. https://doi.org/10.2307/2531595.
83. Kurniansyah N, Goodman MO, Kelly TN. *et al.* A multi-ethnic polygenic risk score is associated with hypertension prevalence and progression throughout adulthood. *Nat Commun* 2022;**13**:3549. https://doi.org/10.1038/s41467-022-31080-2.
84. Lee SH, Yang J, Goddard ME. *et al.* Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* 2012;**28**:2540–2. https://doi.org/10.1093/bioinformatics/bts474.
85. Shi H, Mancuso N, Spendlove S. *et al.* Local genetic correlation gives insights into the shared genetic architecture of complex traits. *Am J Hum Genet* 2017;**101**:737–51. https://doi.org/10.1016/j.ajhg.2017.09.022.