Smoking-Interaction Loci Affect Obesity Traits: A Gene-Smoking Stratified Meta-Analysis of 545,131 Europeans

Won-Jun Lee, Ji Eun Lim, Ji-One Kang, Tae-Woong Ha, Hae-Un Jung, Dong Jun Kim, Eun Ju Baek, Han Kyul Kim, Ju Yeon Chung, Bermseok Oh

Keywords
Stratified analysis · Gene-smoking interaction · Body mass index · UK Biobank · GIANT Consortium · Waist circumference · Waist-hip ratio

Abstract

Introduction: Although many studies have investigated the association between smoking and obesity, very few have analyzed how obesity traits are affected by interactions between genetic factors and smoking. Here, we aimed to identify the loci that affect obesity traits via smoking status-related interactions in European samples. Methods: We performed stratified analysis based on the smoking status using both the UK Biobank (UKB) data (N = 334,808) and the Genetic Investigation of ANthropometric Traits (GIANT) data (N = 210,323) to identify gene-smoking interaction for obesity traits. We divided the UKB subjects into two groups, current smokers and nonsmokers, based on the smoking status, and performed genome-wide association study (GWAS) for body mass index (BMI), waist circumference adjusted for BMI (WCadjBMI), and waist-hip ratio adjusted for BMI (WHRadjBMI) in each group. And then we carried out the meta-analysis using both GWAS summary statistics of UKB and GIANT for BMI, WCadjBMI, and WHRadjBMI and computed the stratified p values \( p_{\text{stratified}} \) based on the differences between meta-analyzed estimated beta coefficients with standard errors in each group. Results: We identified four genome-wide significant loci in interactions with the smoking status \( p_{\text{stratified}} < 5 \times 10^{-8} \): rs336396 (INPP4B) and rs12899135 (near CHRNB4) for BMI, rs998584 (near VEGFA) and rs6916318 (near RSPO3) for WHRadjBMI. Moreover, we annotated the biological functions of the SNPs using expression quantitative trait loci (eQTL) and GWAS databases, along with publications, which revealed possible mechanisms underlying the association between the smoking status-related genetic variants and obesity. Conclusions: Our findings suggest that obesity traits can be modified by the smoking status via interactions with genetic variants through various biological pathways.

Won-Jun Lee and Ji Eun Lim contributed equally to this work.

Correspondence to:
Bermseok Oh, ohbs@khu.ac.kr
Introduction

Obesity is associated with various diseases, including type 2 diabetes, coronary heart disease, and stroke [1]. Genetic effects on obesity have been reported, and many genome-wide association studies (GWASs) on obesity [2–4], body mass index (BMI) [3, 5–8], waist circumference [7–9], and waist-hip ratio (WHR) [8–10] have been performed. In addition to genetic factors, epidemiological studies have also identified obesity-related environmental factors, such as alcohol use [11, 12], smoking [13–15], nutrient intake [16], and physical activity [17, 18]. Moreover, several recent studies have investigated the effects of gene-environmental interaction on obesity [19].

Smoking is related to obesity, and various association analyses between smoking and obesity-related traits have been conducted. Typically, smoking is associated with weight loss [13, 14, 20–22]. Previous studies have reported that nicotine causes weight loss by increasing energy expenditure and insulin resistance, as well as reducing appetite and insulin sensitivity [13, 20]. However, smoking increases the waist circumference, WHR, and abdominal and visceral fat [15, 23, 24]. Moreover, former smokers tend to increase in body weight and subcutaneous fat for 1–8 years after smoking cessation; however, most of them do not gain excessive weight, and their average BMIs tend to be similar to those of nonsmokers [14, 20, 21, 25, 26]. Thus, smoking is a complex environmental factor affecting obesity and obesity-related traits, and more accurate information can be obtained about its effects by identifying genes and pathways that interact with smoking.

Several studies have described the interaction between genes and smoking in obesity. A gene-smoking interaction study in Chinese subjects reported the interaction between smoking and genetic factors in association with the BMI, waist circumference adjusted for BMI (WCadjBMI), and WHR adjusted for BMI (WHRadjBMI). Three interactions between smoking and SNPs were identified: two SNPs near the INPP4A and CHRNB4 genes for BMI, and one SNP near the GRIN2A gene for WCadjBMI. A replication analysis using the UK Biobank Phase 1 data (N = 119,644) for the three SNPs, along with the results of a meta-analysis of the GIANT and UK Biobank data, found that the BMI was affected by a significant interaction between smoking and rs12902602 (near CHRNB4).

Here, we performed gene-smoking interaction analyses for three obesity-related traits (BMI, WCadjBMI, and WHRadjBMI) by stratified analysis of the GIANT and UK Biobank data. GWAS analyses of obesity traits were performed separately in the smoker and nonsmoker groups using the genetic and phenotypic data of 502,536 subjects from the UK Biobank released in July 2017, and the meta-analyses were performed using summary statistics from the UK Biobank and GIANT Consortium. The objectives of our study were to identify genetic variants with different effects on obesity-related traits in smokers and nonsmokers and to gain insights into the mechanism by which smoking affects obesity through functional annotation of these genetic variants. The results of this study were expected to provide insights into smoking-genetic factor interactions and their effects on obesity traits.

Materials and Methods

Study Design Overview

We performed a stratified analysis to identify genetic loci that are associated with the current smoking status and affect obesity traits (Fig. 1). We analyzed three obesity traits: BMI, WCadjBMI, and WHRadjBMI. We divided the subjects included in the UK Biobank into smoker and nonsmoker groups and analyzed the genetic effects of SNPs on obesity traits for each group. We also obtained the GWAS summary statistics for smokers and nonsmokers from the GIANT Consortium for a meta-analysis. We computed the stratified \( p \) values (\( p_{\text{stratified}} \)) based on the differences between meta-analyzed estimated effects, assessed by estimated beta coefficients with standard errors using the \( t \)-statistic in each group [31]. Our meta-analysis focused on the results from the European descent study populations.

Cohort Descriptions and Sample Sizes

The UK Biobank was formed to prevent, diagnose, and treat a wide range of serious and life-threatening illnesses and provide health information to approved researchers [32]. The UK Biobank recruited people aged 37–73 years between 2006 and 2010 from across the country. Among 502,536 participants in the UK Biobank data released in July 2017, we used 334,808 individuals for a meta-analysis (33,735 smokers and 301,073 nonsmokers) after excluding individuals based on the following criteria: (1) sex mis-
match and aneuploidy, (2) non-British ancestry, (3) relatives identified, (4) missing smoking status, and (5) missing information on obesity-related traits (Fig. 2a). BMI was calculated as the weight divided by height squared. Waist circumference and WHR were analyzed with adjustment for BMI using residuals from linear regression. We defined the smoking status as current smoker and current nonsmoker [30] because former smokers and never smokers tend to have a similar trend of obesity [21, 26].

Genotyping Analysis

The UK Biobank subjects were genotyped using the UK Biobank Axiom Array [32] that included 820,967 genetic markers. The UK Biobank genetic dataset contained >92 million autosomal variants imputed from the Haplotype Reference Consortium reference panel and a merged reference panel derived from the UK10K data [33] and the 1000 Genomes Project phase III data [32]. SNP imputation was performed with IMPUTE4, a recoded version of the haplotype imputation functionality implemented in IMPUTE2 [34], which chooses a custom reference panel for each study individual in each 1-Mb segment of the genome. Genotype quality control procedures for exclusion were applied to 92 million imputed data as follows: missing genotype call rate >95%, Hardy-Weinberg equilibrium p value <1 × 10⁻⁶, or minor allele frequency <0.01. Consequently, we used 8,894,755 SNPs in our analyses (Fig. 2b).

Fig. 1. Overview of study design and results. We divided the UK Biobank data into smoker and nonsmoker groups and analyzed the genetic effects of SNPs on obesity traits in each group. We also obtained the genetic association data from the smoker and nonsmoker groups using the GIANT summary statistical data for meta-analysis. We computed stratified p values ($p_{stratified}$) to test for differences between the meta-analyzed estimated effects.

Statistical Analysis

We performed the association analyses between obesity traits including BMI, WC and WHR, and smoking status using linear regression model with adjustment for age, age², sex for BMI, and age, age², sex, and BMI for WC and WHR using SPSS (PASW Statistics v23.0). After dividing the total subjects into smoker and nonsmoker groups, associations between genetic factors and obesity traits were evaluated by linear regression analysis in each group. For each analysis, age, age², sex, and 10 principal components were adjusted for the UK Biobank data. The 10 principal components were calculated using the pca option in PLINK v1.90 to account for population stratification [35].

We conducted group-specific meta-analyses to combine the UK Biobank association results and GIANT summary statistical analysis results [30] for smoking-stratified models. Generally, beta estimates and standard errors were meta-analyzed using a basic fixed-effect model. In total, we used 545,131 subjects (79,703 smokers, 465,428 nonsmokers) and 2,375,325 genotype data in our analyses. We used PLINK v1.90 to conduct individual SNP association analyses and meta-analyses [35]. To detect the smoking-stratified effects, we computed stratified p values ($p_{stratified}$) assessed by estimated beta coefficients with standard errors using the t-statistic [31]:

$$t_{stratified} = \frac{\text{Beta}_{smoker} - \text{Beta}_{nonsmoker}}{\sqrt{\text{SE}^2_{smoker} + \text{SE}^2_{nonsmoker} - 2 \times \gamma_{smoking} \times \text{SE}_{smoker} \times \text{SE}_{nonsmoker}}},$$

(1)
In the equation, Beta_{smoker} and Beta_{nonsmoker} are the smoking-specific estimated effects, and SE_{smoker} and SE_{nonsmoker} are the standard errors for each beta coefficient. \( \gamma_{\text{smoking}} \) is the correlation between Beta_{smoker} and Beta_{nonsmoker} computed as the Spearman rank correlation coefficient across all SNPs for each phenotype. The values of \( \gamma_{\text{smoking}} \) for BMI, WCadjBMI, and WHRadjBMI were 0.201, 0.213, and 0.166, respectively. The GWAS significance threshold (\( p_{\text{stratified}} < 5 \times 10^{-8} \)) was considered statistically significant to account for multiple testing. The lead SNPs were selected using the FUMA platform (https://fuma.ctglab.nl), as follows. We first selected independent significant SNPs, as being \( r^2 < 0.1 \) based on linkage disequilibrium (LD) information from 1000 Genome Project phase 3 European samples among the significant SNPs with \( p_{\text{stratified}} < 5 \times 10^{-8} \). Second, we selected only one SNP with the lowest \( p_{\text{stratified}} \) for the lead SNP among these independent SNPs, located closely each other based on less than 500 kb distance. Additionally, we excluded the lead SNPs, which had no other signals with \( p_{\text{stratified}} < 1 \times 10^{-5} \) within 500 kb flanking region. After stratified analyses on each trait, we also assessed the association and 1° of freedom (df) interaction [36–38] for stratified significant variants in the UK Biobank data using PLINK v1.90.

**Investigation of the Biological Function of Significant Loci**

To investigate the biological function and possible effects of significant variants on obesity traits, we used several expression quantitative trait loci (eQTL) databases and a GWAS catalog. We conducted eQTL analysis using the FUMA platform, which integrated several eQTL online sources, including the Broad Institute GTEx portal, eQTLGen, eQTLcatalogue, and BRAINEAC. We also summarized blood cell-related eQTL data and obesity-relevant nonblood cell tissue eQTLs (e.g., adipose and brain tissues). We preferentially used significant eQTL results for sentinel SNPs to identify biological functions and retained proxy SNP (LD \( r^2 > 0.4 \)) eQTL results. Furthermore, we summarized the GWAS results for all high LD proxy SNPs (LD \( r^2 > 0.6 \)) within 500 kb of any lead SNP in the National Human Genome Research Institute – European Bioinformatics Institute GWAS Catalog [40] (https://www.ebi.ac.uk/gwas/).

**Data Plotting**

We generated Manhattan plots for the stratified analysis results of each obesity trait using the R program v3.6.0. Additionally, we used the LocusZoom program to generate regional association plots for significant loci (http://locuszoom.sph.umich.edu/). In the regional plots, we presented a LD using a European sample from the 1000 Genome Project phase I reference panel [39].

---

**Fig. 2.** The diagrams of sample and SNP quality control. **a** Among 502,536 participants in the UK Biobank data released in July 2017, we included 334,808 individuals into the meta-analysis (33,735 smokers, 301,073 nonsmokers) after applying the following exclusion criteria: sex mismatch (14,626 subjects), aneuploidy (470 subjects), non-European ancestry (78,530 subjects), relatives identified (71,763 subjects), missing smoking status (1,174 subjects), and missing any obesity traits (1,155 subjects). **b** Among a total of 93,095,623 SNPs in the UK Biobank dataset, we used 8,894,755 for our analyses of genetic variants after applying the following exclusion criteria: high missing call rate (>5%) (853,482 SNPs), Hardy-Weinberg equilibrium (HWE) of \( p \leq 1 \times 10^{-6} \) (66,697 SNPs), and low minor allele frequency (MAF <0.01) (83,280,329 SNPs).
Table 1. Baseline characteristics of UK Biobank and association of smoking status with obesity traits

| Characteristics | UK Biobank | smoker | nonsmoker |
|-----------------|------------|--------|-----------|
|                 | all        | N (%)/ mean (SD) | smoking effect for obesity traits | N (%)/ mean (SD) | beta (SE) | p value | N (%)/ mean (SD) | beta (SE) | p value |
| Subjects        | 334,808    | 33,735 (10.08) | - | 301,073 (89.92) | - | - | - |
| Male, %         | 154,977 (46.29) | 18,261 (54.13) | - | 136,716 (45.41) | - | - | - |
| Age, years      | 56.87 (7.99) | 55.05 (8.12) | - | 57.07 (7.95) | - | - | - |
| BMI, kg/m²      | 27.39 (4.74) | 27.06 (4.78) | -0.432 (0.027) | 27.43 (4.74) | 4.40×10⁻⁵⁷ | - | - |
| Waist circumferences, cm | 90.31 (13.45) | 91.19 (13.42) | 0.861 (0.033) | 90.21 (13.44) | 9.75×10⁻¹⁵³ | - | - |
| WHR             | 0.872 (0.090) | 0.89 (0.087) | 0.014 (0.0003) | 0.870 (0.090) | 8.9 (10⁻¹³) | - | - |

Data are presented as the number of participants (%) or the mean (SD). The association between smoking status and BMI was analyzed using a linear regression model adjusted for age, age², and sex. The associations between smoking status, waist circumference, or WHR were analyzed using a linear regression model, adjusted for age, age², sex, and BMI. Subjects for the analysis of smoking status were divided into two groups depending on the UK Biobank data (Data field 20116: never or previous smoker: nonsmoker; current smoker: smoker). SD, standard deviation; SE, standard error.
Association of Smoking with Obesity Traits

The subject characteristics of the UK Biobank dataset used in this study are provided in Table 1. We selected 334,808 participants (154,977 men, 46.29%) for the smoking-stratified analysis. The mean age of the subjects was 56.87 ± 7.99 years, and the mean values of the BMI, waist circumference, and WHR were 27.39 ± 4.74 kg/m², 90.31 ± 13.45 cm, and 0.872 ± 0.090, respectively. The association analysis between smoking and obesity traits was performed using a linear regression model adjusted for age, age², and sex for BMI, and age, age², and sex, and BMI for WC and WHR. We found that smoking was significantly associated with BMI (Beta ± SE = −0.432 ± 0.027, p = 4.40 × 10⁻¹⁵³), WC (Beta ± SE = 0.861 ± 0.033, p = 9.75 × 10⁻¹⁵³), and WHR (Beta ±SE = 0.014 ± 0.0003, p < 1 × 10⁻¹⁵³).

Identification of Significant Variants by Smoking-Stratified Analysis

We conducted a stratified analysis using both the UK Biobank and GIANT Consortium data. The GWASs of obesity traits for the UK Biobank were performed in both smokers and nonsmokers, whereas the GWASs of obesity traits for the GIANT Consortium were ascertained from the summary statistics database (https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files). We then performed a meta-analysis of obesity traits between the UK Biobank and the GIANT Consortium for smokers and nonsmokers separately (Fig. 1). Next, we conducted smoking-stratified analyses on three obesity traits (BMI, WCadjBMI, and WHRadjBMI). The results are displayed using Manhattan plots (Fig. 3). We selected genome-wide significant lead SNPs from the results of the stratified analyses using the FUMA program. Our analysis identified two SNPs (rs336396 and rs12899135) for BMI and two SNPs (rs998584 and rs6916318) for WHRadjBMI (p_stratified < 5 × 10⁻⁸) (Table 2; Fig. 3, and online suppl. Fig. S1; for all online suppl. material, see www.karger.com/doi/10.1159/000525749).

Importantly, the effect of the interaction of rs998584 and smoking on WHRadjBMI is a novel finding that has not been reported in previous GIANT investigations [30].

Table 2 and Figure 4 show the genetic effects of the four lead SNPs in each group relative to the smoking status based on the data of the UK Biobank or the GIANT Consortium, along with the corresponding results of the meta-analysis. The significant effect of two SNPs, rs336396 and rs12899135, which were identified by smoking-stratified analysis of the BMI, was restricted to the smoker group.
The rs336396, an intron variant of inositol polyphosphate-4-phosphatase type II B (INPP4B), was marginally associated with BMI only in the smoker group based on the meta-analysis data ($\beta_{\text{smoker}} = 0.058$, $p_{\text{smoker}} = 1.14 \times 10^{-6}$). The rs12899135, near the cholinergic receptor nicotinic beta 4 subunit (CHRN4), showed a significant effect on BMI only in the smoker group of meta-analysis data ($\beta_{\text{smoker}} = -0.049$, $p_{\text{smoker}} = 2.96 \times 10^{-12}$). In contrast, the genetic effects of two SNPs, rs998584 and rs6916318, which were identified by the stratified analysis for WHRadjBMI, were stronger in nonsmokers than in smokers. We detected a significant effect of rs998584, an intergenic variant near vascular endothelial growth factor A (VEGFA), on WHRadjBMI among nonsmokers ($\beta_{\text{nonsmoker}} = 0.033$, $p_{\text{nonsmoker}} = 1.03 \times 10^{-95}$), whereas its effect among smokers was modest ($\beta_{\text{smoker}} = 0.012$, $p_{\text{smoker}} = 1.19 \times 10^{-4}$). We further found that rs6916318, an intergenic variant near R-Spondin 3 (RSPO3), significantly affected the WHRadjBMI among nonsmokers ($\beta_{\text{nonsmoker}} = -0.027$, $p_{\text{nonsmoker}} = 4.75 \times 10^{-68}$), whereas its effect among smokers was modest ($\beta_{\text{smoker}} = -0.009$, $p_{\text{smoker}} = 6.83 \times 10^{-3}$). In contrast, we did not identify SNPs with significant smoking status-related effects on the WCadjBMI.

Further, we performed an eQTL analysis using 14 eQTL databases on the FUMA platform. We identified eQTLs of LD-linked proxy SNPs ($r^2 > 0.3$) for four SNPs, as summarized in online supplementary Table S1. For rs1907137, a proxy SNP of rs336396, we found that rs1907137 was associated with the expression of INPP4B in whole blood (based on the eQTLGen consortium) and artery tissues (based on GTEx v8.0). The expression of several other genes was associated with rs12899135 and its proxy SNPs. We determined that rs12899135 was associated with the expression of disintegrin-like and metalloprotease with thrombospondin motifs 7 (ADAMTS7)-pseudogene 3 (ADAMTS7P3) in the heart, arteries, and fibroblasts (based on GTEx v8.0). The proxy SNPs of rs12899135, rs4886580 ($r^2 = 0.97$), and rs8042849 ($r^2 = 0.37$) were associated with the expression of CHRN4, CHRN3, and CHRN5 in whole blood and the brain, and rs55988292 ($r^2 = 0.50$) was significantly associated with iron-responsive element binding protein 2 (IREB2) expression in whole blood, as well as in adipose and brain tissues. In various tissues, such as arteries and brain tissues, the ADAMTS7 gene expression was also associated with rs4887113, the proxy SNP of rs12899135 ($r^2 = 0.48$). In WHRadjBMI analyses, the vascular endothelial growth factor A (VEGFA) gene was the nearest to rs998584, but we did not detect any association between VEGFA expression and rs998584. However, the proxy SNPs of rs998584 (rs1358980, $r^2 = 0.85$; rs11967262, $r^2 = 0.98$) were associated with the expression of two genes, RNA polymerase I and III subunit C (POLR1C) in B cells and PEX6 in blood tissue. We also determined that rs9491696 ($r^2 = 0.76$), the proxy SNP of rs6916318, was associated with RSPO3 gene expression in various tissues, such as adipose and brain tissues, based on GTEx v8.0, and we found that rs9491702 ($r^2 = 0.63$) was associated with ring finger protein 146 (RNF146) expression in blood, spleen, and skeletal muscle tissues based on eQTLGen and GTEx v8.0.

In addition, we performed an interaction analysis of significant SNPs by linear regression using the UK Biobank data (online suppl. Table S2). We identified a marginally significant interaction between rs12899135 and the smoking status ($P_{\text{SNPint}} = 0.0153$) for BMI, along with another marginally significant interaction between rs6916318 and the smoking status ($P_{\text{SNPint}} = 0.0212$) for WHRadjBMI. The other two SNPs, rs336396 and rs998584, did not have an interaction effect based on the UK Biobank dataset.

**Fig. 4.** Forest plots for gene-smoking interaction loci. Forest plots present the estimated effects (beta and 95% CI) for gene-smoking interaction loci of stratified by smoking status on (a) BMI and (b) WHRadjBMI. CI, confidence interval.
Discussion

In this study, we conducted smoking-stratified meta-analyses on three obesity traits using the UK Biobank and the GIANT Consortium datasets to identify the genetic variants interacting with smoking. We identified four genome-wide significant loci in interactions with the smoking status (two loci for BMI and two for WHRadjBMI).

We conducted a stratified analysis to examine gene-smoking interactions based on differences in the genetic effects on obesity-related traits within each group. Our approach differed from the commonly applied 1df interaction and 2df joint analyses, testing gene-environmental (G-E) interaction using regression model that includes action and 2df joint analyses, testing gene-environmental approach differed from the commonly applied 1df interaction analysis. The 1df interaction analysis, testing the G-E interaction effect, is useful to find strong G-E interaction effect, but requires 4 times more samples than that of the GWAS [36]. The 2df joint analysis, testing both genetic and interaction effect, is efficient to identify genetic variants having marginal genetic effects with moderate G-E interactions, but could identify genetic variants with only genetic main effect without interaction effect. On the other hand, the stratified analysis approach is a powerful statistical tool used in epidemiology that allows the testing of confounding interactions [42]. Moreover, we carried out a meta-analysis of the full dataset of the UK Biobank (N = 334,808) and the GIANT Consortium data (N = 210,323) to assess a large sample of European participants (total N = 545,131). The involvement of the VEGFA locus, which was not associated with significant results in a previous study, was discovered by the meta-analysis in the present study.

In a previous study by Justice et al. [30], the GIANT Consortium (n = 241,258) and the UK Biobank phase 1 (n = 119,644) datasets were used for assessing the gene-smoking interactions by both linear regression analysis and GxE screening after the smoking-adjusted association. Analysis of the GIANT Consortium data by the authors identified gene-smoking interactions among genetic variants related to BMI (rs336396 near INPP4B and rs12902602 near CHRN4B4), WCadjBMI (rs4141488 near GRIN2A), and WHRadjBMI (rs765751 near LYLAL1 and rs7766106 near RSPO3). Among these variants, three of rs12902602, rs765751, and rs7766106 were retained for their interactions in the meta-analysis of the GIANT Consortium and the UK Biobank phase 1. However, our study identified two genetic variants related to BMI (rs336396 near INPP4B and rs12899135 near CHRN4B4) and two genetic variants related to WHRadjBMI (rs998584 near VEGFA and rs6916318 near RSPO3). The two SNPs located at the CHRN4B4 locus, rs12899135 and rs12902602, had a high LD ($r^2 = 0.96$), indicating that these signals were dependent on each other. In addition, the two SNPs located at the RSPO3 locus, rs6916318 and rs7766106, also had a high LD ($r^2 = 0.86$). These results strongly demonstrate the fact that the interaction of these three loci (near INPP4B, CHRN4B4, and RSPO3) with smoking has an effect on obesity traits. Our study did not replicate the effects of rs4141488 (near GRIN2A) and rs765751 (near LYLAL1) on WCadjBMI and WHRadjBMI, respectively, at the genome-wide significance level, although these two SNPs were also marginally significant based on the p values, 1.77 x 10^{-4} and 1.32 x 10^{-4}, respectively (online suppl. Table S3). Additionally, Justice et al. [30] reported seven genetic SNP variants interacting with smoking in sex- or ethnic-specific populations, which did not interact with smoking in the present analysis (online suppl. Table S3).

The SNP rs336396 lies within the INPP4B gene, which encodes inositol polyphosphate 4-phosphatase type II, one of the enzymes involved in phosphatidylinositol signaling pathways. This SNP is associated with INPP4B expression in whole blood (based on the eQTLGen consortium) and artery tissues (based on GTEx v8.0). Genetic variants in INPP4B have been associated with smoking traits [43, 44], lung function, and obesity traits [45]. The BMI-related analysis identified variant rs12899135 near the CHRNA5-CHRNAS-CHRN4B4 gene cluster, which plays an important role in nicotine activity. Nicotine affects energy homeostasis and food consumption by changing the activity of neurons harboring anorexigenic and orexigenic peptides in the brain [46]. A shift in the equilibrium between these peptide groups can lead to reduced food intake and weight loss. As expected, rs12899135 was associated with a decreased BMI only among smokers (Table 2). Likewise, rs6916318 near RSPO3, which was identified as a significant SNP in the WHRadjBMI analysis, can be linked to obesity via regulation by Wnt signaling in adipogenesis. Wnt signaling is suppressed by smoking, but RSPO3 functions as an activator of the Wnt (wingless-type MMTV integration site family)/beta-catenin and Wnt/planar cell polarity signaling pathways. Wnt signaling downregulates adipogenesis [47], and smoking suppresses Wnt signaling [48]. Thus, there is the possibility that the minor allele A of rs6916318 increases the expression of RSPO3 (eQTL Beta = 0.299, $p = 1.51 \times 10^{-12}$ in adipose tissue and eQTL Beta = 0.091, $p = 7.54 \times 10^{-9}$ in the brain), which in turn stimulates Wnt signaling that results in reduced adipogenesis among...
nonsmokers. However, smoking suppresses the elevated Wnt signaling by rs6916318, leading to a diminished influence on adipogenesis. As expected, the decrease in WHRadjBMI associated with rs6916318 was more substantial in nonsmokers than in smokers.

This study also identified rs998584 as a lead SNP that is located downstream of the VEGFA gene, which is essential for both physiological and pathological angiogenesis. A previous study reported that genetic variants of VEGFA are associated with the activity of rheumatoid arthritis by smoking-related interactions, and the association of VEGFA variant with rheumatoid arthritis disease activity occurs only in patients who have never smoked [49]. VEGFA variants also appear to be involved as shared genetic links between obesity traits and asthma [50], suggesting that VEGFA may play a role in the function of obesity as a risk factor for asthma. Other genetic variants of VEGFA are known to be associated with obesity traits based on a study accounting for smoking behavior [30] and physical activity [51], indicating that genetic variants of VEGFA affect obesity via a variety of environmental exposure types and may not be specific to smoking. The mechanism underlying the contribution of VEGFA via environmental exposure may involve hypoxia induced by smoking. The VEGFA level in blood is known to be affected by a variety of other factors, including hypoxia [52–54]. These findings suggest that the gene expression of VEGFA was altered by both rs998584 and smoking, indicating that the effect of VEGFA SNP rs998584 on obesity traits is less pronounced in smokers.

There are several limitations and difficulties in studying gene-environment interactions. First, there is the challenge of statistical power, which remains a major issue in gene-environment interaction studies [55]. To solve this problem, we assessed the UK Biobank and GIANT datasets by meta-analyses to obtain the data from as many subjects as possible. Second, to date, there are several challenges in interpreting the analysis results because the distribution of traits is different for each cohort, the frequency of smoking status is different, and the LD pattern or SNP allele frequency in each cohort is different. Third, we divided subjects into two groups, nonsmokers and current smokers, based on smoking status. However, the previous study reported that intensity of smoking affected biological mechanisms [56]. Basson et al. [57] performed gene-smoking interaction analysis on blood pressure using a number of cigarettes per day (CPD), by dividing subjects into two groups based on the CPD cutoffs of 10, 15, and 20. If this stratification method using quantitative smoking measures such as CPD and pack-years is applied to this study, it is expected that additional genesmoking interactions will be identified. Despite these limitations, we consistently detected certain genetic effects within each smoking-stratified group for at least three genetic variants near INPP4B, CHRNA4, and RSPO3 in this study. Nevertheless, to obtain more accurate results for variants that did not meet the GWS threshold and to solve the bias between each cohort, it will be necessary to use larger sample sizes and replicate the results by performing multiple studies.

We conducted smoking-stratified analyses of three obesity traits using the UK Biobank and GIANT Consortium datasets. We replicated the previously identified gene-smoking interactions associated with CHRNA4, INPP4B, and RSPO3 loci that are related to obesity traits (CHRNA4 and INPP4B affect BMI; RSPO3 affects WHRadjBMI). Additionally, an SNP downstream of VEGFA (rs998584) was identified as a novel genetic variant that has smoking status-related effects on WHRadjBMI. Our findings suggest that obesity traits can be modified by the smoking status via interactions with genetic variants through various biological pathways.

Acknowledgments

This study was conducted using the resources of the UK Biobank (application 48422).

Statement of Ethics

UK Biobank obtained ethics approval from the North West – Haydock Research Ethics Committee, which covers the UK (approval number: 21/NW/0157), and written informed consent from all participants. The present analyses were conducted under the UK Biobank Application No. 48422.

Conflict of Interest Statement

The authors declare no competing interests.

Funding Sources

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2021R1I1A1A01049439 and 2019M3E5D3073365 [the Bio & Medical Technology Development Program]).
Author Contributions

Won-Jun Lee and Ji Eun Lim conceived the original idea for the study and drafted the manuscript. Won-Jun Lee, Tae-Woong Ha, Hae-Un Jung, Dong Jun Kim, Eun Ji Baek, and Ju Yeon Chung conducted the statistical analyses. Bermesek Oh, Ji-One Kang, and Ji Eun Lim provided guidance on study design and statistical analyses. Ji Eun Lim and Han Kyul Kim contributed to the analytical study.

References

1 Grundy SM. Metabolic complications of obesity. *Endocrine*. 2000;13(2):155–65.
2 Singh RK, Kumar P, Mahalingam K. Molecular genetics of human obesity: a comprehensive review. *C R Biol*. 2017;340(2):107–108.
3 Speakman JR, Loos RFJ, O’Rahilly S, Hirschhorn JN, Allison DB. GWAS for BMI: a treasure trove of fundamental insights into the genetic basis of obesity. *Int J Obes*. 2018;42(8):1524–31.
4 Fall T, Ingelsson E. Genome-wide association studies of obesity and metabolic syndrome. *Am J Hum Genet*. 2014;94(1):74–57.
5 Locke AE, Kahali B, Berndt SI, Justice AE, Krämer J, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518(7538):206–210.
6 Akiyama M, Okada Y, Kanai M, Takahashi A, Momozawa Y, Ikeda M, et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat Genet*. 2017;49(10):1458–67.
7 Fox CS, Heard-Costa N, Cupples LA, Dupuis J, Vasan RS, Atwood LD, et al. Genome-wide association to body mass index and waist circumference: the Framingham Heart Study 100K project. *BMC Med Genet*. 2007;8(Suppl 1):S18.
8 Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature*. 2015;518(7538):187–96.
9 Wen W, Kato N, Hwang JY, Guo X, Tabara Y, Kim YJ, et al. Alcohol consumption and waist circumference. *Sci Rep*. 2016;6(1):17958.
10 Pulkstenis P, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, 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(11):1667–74.
11 Traversy G, Chappuis P, Alcohol consumption and obesity: an update. *Curr Obes Rep*. 2015;4(1):122–30.
12 Colditz GA, Giovannucci E, Rimm EB, Stampfer M, Rosner B, Speizer FE, et al. Alcohol intake in relation to diet and obesity in women and men. *Am J Clin Nutr*. 1991;54(1):49–55.
13 Chiolerio A, Faeh D, Paccaud F, Cornuz J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. *Am J Clin Nutr*. 2008;87(4):801–9.
14 Dare S, Mackay DF, Pell JP. Relationship between smoking and obesity: a cross-sectional study of 499 504 middle-aged adults in the UK general population. *PLoS One*. 2015;10(4):e0123579.
15 Canoy D, Wareham N, Luben R, Welch A, Bingham S, Day N, et al. Cigarette smoking and fat distribution in 21,828 British men and women: a population-based study. *Obes Res*. 2005;13(8):1466–75.
16 Swinburn BA, Caterson I, Seidell JC, James WP. Diet, nutrition and the prevention of excess weight gain and obesity. *Public Health Nutr*. 2004;7(1A):123–46.
17 Kim BY, Choi DH, Jung CH, Kang SK, Mok JH, Kim JH, et al. Alcohol consumption and fat distribution in 21,828 British men and women: a population-based study. *Obes Res*. 2005;13(8):1466–75.
18 Reddon R, Gueant JL, Meyre D. The importance of gene-environment interactions in human obesity. *Clin Sci*. 2016;130(18):1571–97.
19 Filizot C, Fernandez Pinilla MC, Fernandez-Cruz A. Smoking cessation and weight gain. *Obes Rev*. 2004;5(2):95–103.
20 Pistelli F, Aquilini F, Carrozzi L. Weight gain after smoking cessation. *Monaldi Arch Chest Dis*. 2009;72(1):81–7.
21 Sneve M, Jorde R. Cross-sectional study on smoking status, cigarettes per day, and their relationship to overweight and obesity among former and current smokers in a national study. *Nicotine Tob Res*. 2004;6(1):122–30.
22 Pistelli F, Aquilini F, Carrozzi L. Weight gain after smoking cessation. *Monaldi Arch Chest Dis*. 2009;72(1):81–7.
23 Kim JH, Shin KW, Yoon YS, Lee Yi, Kim SS, Oh SW, et al. Cigarette smoking increases abdominal and visceral obesity but not overall fatness: an observational study. *PLoS One*. 2012;7(9):e48515.
24 Morris RW, Taylor AE, Fluharty ME, Bjorntorp PH, Asvold BO, Elvestad Gabrielsen M, et al. Heavier smoking may lead to a relative increase in waist circumference: evidence for a causal relationship from a Mendelian randomisation meta-analysis. *The CARTA consortium*. *BMJ Open*. 2015;5(8):e008808.
25 John U, Hanke M, Rumpf HJ, Thyrian JR. Smoking status, cigarettes per day, and their relationship to overweight and obesity among former and current smokers in a national adult general population sample. *Int J Obes*. 2005;29(10):1289–94.
26 Munafò MR, Trilling K, Ben-Shlomo Y. Smoking status and body mass index: a longitudinal study. *Nicotine Tob Res*. 2009;11(6):765–71.
27 Zhang H, Xu T, et al. Smoking modifies the effect of two independent SNPs rs5063 and rs198358 of NPPA on central obesity in the Chinese Han population. *J Genet*. 2018;97(4):987–94.
28 Ahmad S, Zhao W, Renstrom F, Rasheed A, Zaidi M, Samuel M, et al. A novel interaction between the FLJ33534 locus and smoking in obesity: a genome-wide study of 14 131 Pakistani adults. *Int J Obes*. 2016;40(1):186–90.
29 Young KL, Graff M, North KE, Richardson AS, Mohile KL, Lange LA, et al. Interaction of smoking and obesity susceptibility loci on adolescent BMI: the National Longitudinal Study of adolescent to adult health. *BMC Genet*. 2015;16(1):131.
30 Justice AE, Winkler TW, Feitosa MF, Graff M, Fisher VA, Young K, 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.
31 Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, et al. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet*. 2013;9(6):e1003500.
32 Bynoe C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203–9.
33 Huang J, Howie B, McCarthy S, Memari Y, Walter K, Min JL, et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nature communications*. 2015;6(1):8111.
34 Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3*. 2011;1(6):457–70.
35 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–75.
36 Hunter DJ. Gene-environment interactions in human diseases. *Nat Rev Genet*. 2005;6(4):287–98.
37 Nakamura S, Narimatsu H, Sato H, Sho R, Otani K, Kawasaki K, et al. Genome-environment interactions in obesity: implication for future applications in preventive medicine. *Hum Genet*. 2016;134(3):241–7.
38 Tyrrell I, Wood AR, Ames RM, Yaghootkar H, Beaumont RN, Jones SE, et al. Gene-obsen- sogenic environment interactions in the UK Biobank study. *Int J Epidemiol*. 2017;46(2):559–75.
39 The 1000 Genomes Project Consortium; Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467(7319):1061–73.
40 Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A. 2009; 106(23):9362–7.

41 Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. Hum Hered. 2007;63(2):111–9.

42 Ni A, Satagopan JM. Estimating additive interaction effect in stratified two-phase case-control design. Hum Hered. 2019;84(2):90–108.

43 Liu M, Jiang Y, Wedow R, Li Y, Bresal DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet. 2019;51(2):237–44.

44 Bresal DM, Jiang Y, Hughey JM, Turcot V, Zhan X, Gong J, et al. Exome chip meta-analysis fine maps causal variants and elucidates the genetic architecture of rare coding variants in smoking and alcohol use. Biol Psychiatry. 2019;85(11):946–55.

45 Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging polygenic functional enrichment to improve GWAS power. Am J Hum Genet. 2019;104(1):65–75.

46 Stojakovic A, Espinosa EP, Farhad OT, Lutfy K. Effects of nicotine on homeostatic and hedonic components of food intake. J Endocrinol. 2017;235(1):R13–31.

47 Christodoulides C, Lagathu C, Sethi JK, Vidal-Puig A. Adipogenesis and WNT signalling. Trends Endocrinol Metab. 2009;20(1):16–24.

48 Baarsma HA, Konigshoff M. "WNT-er is coming": WNT signalling in chronic lung diseases. Thorax. 2017;72(8):746–59.

49 Chen Y, Dawes PT, Mattey DL. Polymorphism in the vascular endothelial growth factor (VEGFA) gene is associated with serum VEGF-A level and disease activity in rheumatoid arthritis: differential effect of cigarette smoking. Cytokine. 2012;58(3):390–7.

50 Zhu Z, Guo Y, Shi H, Liu CL, Panganiban RA, Chung W, et al. Shared genetic and environmental links between obesity-related traits and asthma subtypes in UK Biobank. J Allergy Clin Immunol. 2020;145(2):537–49.

51 Graff M, Scott RA, Justice AE, Young KL, Feitosa MF, Barata L, et al. Genome-wide physical activity interactions in adiposity: a meta-analysis of 200,452 adults. PLoS Genet. 2017;13(4):e1006528.

52 Ferrara N. Molecular and biological properties of vascular endothelial growth factor. J Mol Med. 1999;77(7):527–43.

53 Koukourakis MI, Papazoglou D, Giatromanolaki A, Bougioukas G, Maltezos E, Sivridis E. VEGF gene sequence variation defines VEGF gene expression status and angiogenic activity in non-small cell lung cancer. Lung Cancer. 2004;46(3):293–8.

54 Prior SJ, Hagberg JM, Paton CM, Douglass LW, Brown MD, McLenithan JC, et al. DNA sequence variation in the promoter region of the VEGF gene impacts VEGF gene expression and maximal oxygen consumption. Am J Physiol Heart Circ Physiol. 2006;290(5):H1848–55.

55 Aschard H, Lutz S, Maus B, Duell EJ, Fingerlin TE, Chatterjee N, et al. Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. Hum Genet. 2012;131(10):1591–613.

56 Poussin C, Laurent A, Peitsch MC, Hoeng J, De Leon H. Systems biology reveals cigarette smoke-induced concentration-dependent direct and indirect mechanisms that promote monocyte-endothelial cell adhesion. Toxicol Sci. 2015;147(2):370–85.

57 Basson J, Sung YJ, de las Fuentes L, Schwander K, Cupples LA, Rao DC, et al. Influence of smoking status and intensity on discovery of blood pressure loci through gene-smoking interactions. Genet Epidemiol. 2015;39(6):480–8.