Pleiotropy of genetic variants on obesity and smoking phenotypes: Results from the Oncoarray Project of The International Lung Cancer Consortium

Tao Wang¹ *, Jee-Young Moon¹, Yiquan Wang¹,², Christopher I. Amos³, Rayjane J. Hung⁴, Adonius Tardon⁵, Angeline Andrew⁶, Chu Chen⁷, David C. Christiani⁸, Demetrios Albanés⁹, Erik H. F. M. van der HeijdenISCO, Lazlo Le Marchand¹⁰, Gadi Rennert¹¹, Gary Goodman¹², Geoffrey Liu¹³, James D. Mckay¹⁴, Jian-Min Yuan¹⁵, John K. Field¹⁶, Jonas Manjer¹⁷, Kjell Granqvist¹⁸, Lambertus A. Kiemeneu¹⁹, Loïc Le Marchand¹⁰, M. Dawn Teare¹⁹, Matthew B. Schabath²⁰, Mattias Johansson²¹, Melinda C. Aldrich²², Michael Davies²³, Mikael Johansson²⁴, Ming-Sound Tsao²⁵, Neil Caporaso²⁶, Philip Lazarus²⁷, Stephen Lam²⁸, Stig E. Bojesen²⁹, Susanne Arnold²⁹, Xifeng Wu³⁰, Xuchen Zong³¹, Yun-Chul Hong³², Gloria Y. F. Ho³³,³⁴,³⁵

¹ Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, New York, United States of America; ² Department of Epidemiology & Biostatistics, School of public health, Peking University Health Science Center, Beijing, China; ³ Community and Family Medicine, Geisel School of Medicine, Dartmouth College, Hanover, New Hampshire, United States of America; ⁴ Lunenfeld-Tanenbaum Research Institute, Sinai Health System; Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; ⁵ IUOPA. University of Oviedo and CIBERESP. Oviedo, Spain; ⁶ Norris Cotton Cancer Center, Hanover, New Hampshire, United States of America; ⁷ Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America; ⁸ Harvard School of Public Health, Boston, Massachusetts, United States of America; ⁹ National Cancer Institute, Bethesda, United States of America; ¹⁰ Radboud university medical center, Nijmegen, Netherlands; ¹¹ Catalan Institute of Oncology (ICO), Barcelona, Spain; ¹² Carmel Medical Center, Haifa, Israel; ¹³ International Agency for Research on Cancer (IARC), Lyon, France; ¹⁴ University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania, United States of America; ¹⁵ Royal College Lung Cancer Research Programme, Department of Molecular & Clinical Cancer Medicine, The University of Liverpool, Liverpool, UK; ¹⁶ Department of surgery, Unit for breast surgery, Lund University, Malmö, Skåne University Hospital Malmö, Malmö, Sweden; ¹⁷ Department of Medical Biosciences, Umeå University, Umeå, Sweden; ¹⁸ University of Hawaii Cancer Center, Honolulu, Hawaii, United States of America; ¹⁹ University Of Sheffield, Sheffield, South Yorkshire, United Kingdom; ²⁰ Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, United States of America; ²¹ Department of Thoracic Surgery, Division of Epidemiology, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America; ²² Princess Margaret Cancer Center, Toronto, Ontario, Canada; ²³ Washington State University College of Pharmacy, Washington, United States of America; ²⁴ British Columbia Cancer Agency, Vancouver, British Columbia, Canada; ²⁵ Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen, Denmark; ²⁶ Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Copenhagen, Denmark; ²⁷ Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; ²⁸ Markey Cancer Center, Lexington, Kentucky, United States of America; ²⁹ The University of Texas MD Anderson Cancer Center, Texas, Houston, United States of America; ³⁰ Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea; ³¹ Merinoff Center for Patient-Oriented Research, The Feinstein Institute for Medical Research, New York, United States of America; ³² Epidemiology and Research, Northwell Health, New York, United States of America; ³³ Hofstra Northwell School of Medicine, New York, United States of America

* tao.wang@einstein.yu.edu (TW); gh01@northwell.edu (GYFH)

OPEN ACCESS


Editor: David Meyre, McMaster University, Canada

Received: July 4, 2017
Accepted: September 16, 2017
Published: September 28, 2017

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

Data Availability Statement: Genotype data analyzed in this study are available through dbGAP. The accession number is phs001273.v2.p1.

Funding: This study was funded by National Cancer Institution (https://www.cancer.gov, R21 CA202529, TWG YFH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.
Abstract

Obesity and cigarette smoking are correlated through complex relationships. Common genetic causes may contribute to these correlations. In this study, we selected 241 loci potentially associated with body mass index (BMI) based on the Genetic Investigation of ANthropometric Traits (GIANT) consortium data and calculated a BMI genetic risk score (BMI-GRS) for 17,037 individuals of European descent from the Oncoarray Project of the International Lung Cancer Consortium (ILCCO). Smokers had a significantly higher BMI-GRS than never-smokers (p = 0.016 and 0.010 before and after adjustment for BMI, respectively). The BMI-GRS was also positively correlated with pack-years of smoking (p < 0.001) in smokers. Based on causal network inference analyses, seven and five of 241 SNPs were classified to pleiotropic models for BMI/smoking status and BMI/pack-years, respectively. Among them, three and four SNPs associated with smoking status and pack-years (p<0.05), respectively, were followed up in the ever-smoking data of the Tobacco, Alcohol and Genetics (TAG) consortium. Among these seven candidate SNPs, one SNP (rs11030104, BDNF) achieved statistical significance after Bonferroni correction for multiple testing, and three suggestive SNPs (rs13021737, TMEM18; rs11583200, ELAVL4; and rs6990042, SGCZ) achieved a nominal statistical significance. Our results suggest that there is a common genetic component between BMI and smoking, and pleiotropy analysis can be useful to identify novel genetic loci of complex phenotypes.

Introduction

Both obesity and cigarette smoking are risk factors for many human diseases, including multiple cancers.[1–4] There are complex sources of correlations between smoking behavior and obesity.[5,6] In general, current smokers tend to have a lower body mass index (BMI) than never-smokers, while smoking cessation is associated with weight gain.[7–9] The reasons for the association between BMI and smoking status may involve smoking-induced appetite suppression via neural pathways [10] and increased energy expenditure via energy-regulating hormonal feedback loops.[11,12] On the other hand, heavy smokers tend to have a greater BMI than light smokers; an observation that is seemingly contradictory to the metabolic effects of smoking,[7,13] but may be partially attributed to the unhealthy behaviors associated with heavy smoking. Another reason for the correlation between smoking behavior and obesity is that there may be common underlying biological causes. There is growing evidence suggesting that obesity may be partially due to addiction to food.[14,15] One plausible common mechanism for obesity and smoking is brain reward effects arising from neuronal activity within the dopamine system.[16] In any case, the reasons for the relationship between BMI and smoking behavior remain uncertain.

Shared genetic susceptibility may offer another explanation for the correlation between obesity and smoking. Both smoking and obesity have significant genetic components. In the past, large-scale genome-wide association studies (GWAS) on obesity or variables related to smoking characteristics (e.g., smoking status, age started smoking, and pack-years of smoking, etc) have successfully identified multiple loci associated with these phenotypes.[17–25] Yet, total variation in obesity or smoking traits explained by these GWAS loci is still limited.[20,26–29] The remaining genetic variants still need to be identified. It was estimated that the genetic correlation between smoking status and BMI was 0.20.[30] In a previous study in
Iceland, the genetic risk score (GRS) of 32 common variants identified in GWAS of BMI was associated with smoking initiation and the number of cigarettes smoked per day (CPD), suggesting that smoking and BMI may share common genetic components.[31] However, this study in Iceland only observed correlations of BMI associated SNPs with smoking variables, without accounting for possible causal relationships between SNPs, BMI and smoking variables. We hypothesized that analyzing pleiotropic effects on BMI and smoking behavior may discover novel genetic loci, otherwise undiscovered in GWAS with stringent genome-wide significance, which in turn would further elucidate genetic architectures underlying both smoking behavior and obesity. In this study, leveraging existing genotyping, BMI, and smoking data from a lung cancer consortium, we confirmed the association between the BMI-GRS and smoking-related variables with adjustment for BMI and important covariates, and used causal network inference to identify potential genetic loci with pleiotropic effects on both BMI and smoking-related phenotypes.

**Materials and methods**

**Study population**

The International Lung Cancer Consortium (ILCCO) was established in 2004 with the goal of sharing comparable research data and maximizing research efficiency (http://ilcco.iarc.fr). To further characterize cancer genetic architecture of common cancers, a custom OncoArray (http://oncoarray.dartmouth.edu) genotyping chip that includes 550K markers was designed to genotype samples in collaboration with other cancer consortia under The National Cancer Institute (NCI) initiative on the Genetic Associations and Mechanisms in Oncology (GAME-ON). In this study, we analyzed OncoArray genotypic data of 36,000 subjects of European descent in ILCCO; among them, 17,037 provided individual epidemiological data and were of European descent.

**OncoArray genotyping, quality control and imputation**

The GAME-ON OncoArray chip was previously described.[32] In brief, it includes a GWAS backbone and a customized panel for dense mapping of known susceptibility regions, rare variants from sequencing experiments, pharmacogenetic markers and cancer related traits including smoking and BMI. The genotyping quality control of Oncoarray data was previously described.[33] After filtering out SNPs by success rate and genotype distribution deviation from the expected by Hardy-Weinberg equilibrium, 517,482 SNPs were available for analysis. Standard quality control procedures were used to exclude underperforming samples (2,408), unexpected duplicated or related samples (2,411), samples with sex error (316) and non-Caucasians (8,240). After quality control, 17,037 subjects with full information on both BMI and smoking status, and other important covariates (age, sex, study sites, and lung cancer status) were kept for analysis. Genotype data were imputed by the GAME-ON data coordinating center for all scans for over 10 million SNPs using data from the 1000 Genomes Project (Phase 3, October 2014) as reference.[34,35]. The data were imputed in a two-stage procedure using SHAPEIT [36] to derive phased genotypes, and IMPUTEv2 to perform imputation of the phased data. [35] Genotypes were aligned to the positive strand in both imputation and actual genotyping.

**SNP selection and derivation of the BMI-GRS**

We first identified a large set of 4,961 SNPs associated with BMI with p<10^{-5} based on results from the Genetic Investigation of ANthropometric Traits (GIANT) consortium, a large
collaborative GWAS on human body size and shape. We then pruned SNPs by applying a threshold value of $r^2 = 0.2$ and requiring selected SNPs at least 500Kb apart to reduce redundancy and obtained a subset of 241 independent SNPs that were at least 500Kb apart (S1 Table). To calculate the BMI-GRS, each SNP was recoded as 0, 1, or 2 according to the number of risk alleles (BMI increasing alleles). The BMI-GRS was calculated using the equation: $GRS = \text{weight}_1 \times \text{SNP}_1 + \text{weight}_2 \times \text{SNP}_2 + \ldots + \text{weight}_n \times \text{SNP}_n$, where $n$ is the total number of SNPs. Both un-weighted and weighted BMI-GRSs were calculated in which the weight is 1 for all SNPs and the weight is the $\beta$ coefficient of each individual SNP on BMI derived from GIANT for the un-weighted and weighted GRS, respectively. The results of un-weighted and weighted GRS were largely similar and we presented un-weighted BMI-GRS in the Results. To examine the robustness of the results based on the BMI-GRS of our selected SNPs, we also calculated the BMI-GRS based on 97 BMI-associated SNPs which reached genome-wide significant levels ($P < 5 \times 10^{-8}$) in the GIANT BMI GWAS including up to 322,154 European descents and 17,072 non-European descents [37].

Statistical analysis
Age, sex, smoking statuses, pack-years, BMI, and BMI categories were compared between lung cancer statuses by student t-test and Chi-square test for categorical variables. All statistical tests are two-sided. The analyses were performed using R (v2.6).

Association of BMI with smoking phenotypes
Linear regression model was applied for comparison of BMI among individuals with different smoking categories (never-smokers, current smokers and ex-smokers) with adjustment for age, sex, and study sites. Adjusted means of BMI of individuals with different smoking categories and their 95% confidence intervals (CIs) were calculated using the \textit{lm} function in the statistical software R with a fixed intercept of zero. Additional stratification analyses by lung cancer status were performed.

Association of the BMI-GRS with BMI and smoking phenotypes
Linear regression was also used to compare the BMI-GRS between different BMI categories (underweight, <18.5; Normal, 18.5–24.9; Overweight, 25.0–29.9; and Obese, $\geq 30$) by adjusting for age, sex, study sites, and top four genetic principal components. Although our analyses were performed only for participants of European descent, study sites and top four genetic principle components generated using common SNPs were included in the regression models for BMI-GRS in order to further limit the effects of any possible cryptic population stratification that might cause inflation of test statistics. Trend tests were performed by analyzing the BMI categories as a continuous variable in the regression model. A similar regression analysis was also performed to compare the BMI-GRS between individuals with different smoking categories (never-smokers, current-smokers, and ex-smokers). Partial correlation coefficients between the BMI-GRS and pack-years of smoking were estimated by Pearson correlation coefficients of their residuals from linear regression models after adjusting for age, sex, study sites, and top four genetic principal components. Additional stratification analyses were performed by lung cancer status.

Identifying candidate pleiotropic SNPs for BMI and smoking phenotypes
We used a causal network inference model to identify possible pleiotropic SNPs for both BMI/pack-years and BMI/smoking status (smokers versus non-smokers), respectively.[38] We
describe the approach here using BMI and pack-years as an example. A similar approach was used to identify pleiotropic SNPs for BMI and remaining smoking traits. Specifically, we modeled 12 possible directed acyclic graphs (DAGs) of the genotype value of a SNP on BMI and/or pack-years (Fig 1). We classified these DAGs into four categories: (1) the SNP did not have effects on either BMI or pack-years, (2) the SNP had direct effects on BMI, but not pack-years, (3) the SNP had direct effects on pack-years, but not BMI, and (4) the SNP had pleiotropic effects on both BMI and pack-years.

Based on a given DAG, we fit two linear regression models for BMI and pack-years, respectively, with adjustment for sex, age, study sites, and genetic principal components (PCs). For example, two linear regression models for the DAG of SNP → pack-years → BMI (DAG 8 with genetic effects on pack-years only) are

\[
\text{BMI} \sim \text{Age} + \text{Sex} + \text{Study Sites} + \text{PCs} + \text{Pack years},
\]

\[
\text{Pack years} \sim \text{Age} + \text{Sex} + \text{Study Sites} + \text{PCs} + \text{SNP}.
\]

To identify the model that was the most supported by the data, we calculated AIC for each DAG

\[-2 \log \text{lik (Regression Model 1)} - 2 \log \text{lik (Regression Model 2)} + 2^\text{number of edges}.
\]

We then compared the minimum AIC values of four categories. SNPs with at least 2 of the minimum AIC value of category 4 (model 10, 11, or 12) less than other categories were further examined for their association with pack-years. Similar analyses were performed for smoking status using logistic regression. Those SNPs that achieved a nominal statistical significance (p<0.05) were considered as candidate pleiotropic SNPs, and further validated for their associations with ever-smoking using the independent database from TAG (The Tobacco, Alcohol and Genetics) consortium. (https://www.med.unc.edu/pgc/results-and-downloads)
Results

Characteristics of the study population

In our analysis, 17,037 subjects of European descent from 17 study sites had full information on both BMI and smoking status, and other important covariates (age, sex, study sites, and lung cancer status). As expected, compared to the controls, lung cancer cases were older, had a higher proportion of smokers, and were slightly leaner (Table 1).

Association of BMI and smoking variables

We compared BMI levels among never-smokers, ex-smokers, and current smokers. As expected and as compared to never-smokers, ex-smokers had a significantly higher BMI (difference from never-smokers = 0.39 kg/m\(^2\), \(p = 2.64 \times 10^{-4}\)), while current smokers were leaner (difference from never-smokers = -1.08 kg/m\(^2\), \(p = 8.20 \times 10^{-24}\)) after adjustment for age, sex, and study sites. Such differences in BMI by smoking status were similar for cases and controls (Fig 2).

BMI and pack-years of smoking were positively correlated in both current smokers and ex-smokers after adjustment for age, sex, and study sites, and the correlations were stronger in ex-smokers than those in current-smokers (Table 2). Specifically, the partial coefficient of pack-years of smoking and BMI was 0.054 (95%CI 0.027–0.075) and 0.112 (95%CI 0.088–0.136) for current smokers and ex-smokers, respectively. The correlations between BMI and pack-years were similar for cases and controls.

Association of the BMI-GRS with BMI

We first confirmed if the BMI-GRS based on 241 SNPs identified in GIANT was associated with BMI in the OncoArray Project population. Comparing the BMI-GRS of individuals in different BMI categories with adjustment for age, sex, study sites, genetic principal components, smoking types, and pack-years (Fig 3), we found that the BMI-GRS significantly increased from the categories underweight (BMI < 18.5), to normal weight (BMI 18.5–24.9), to

---

### Table 1. Characteristics of 17,037 European-descent subjects in the OncoArray Project and epidemiologic data.

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>9,633</td>
<td>7,404</td>
<td></td>
</tr>
<tr>
<td>Male (%)</td>
<td>5,461 (56.7)</td>
<td>42,71 (57.7)</td>
<td>0.199</td>
</tr>
<tr>
<td>Age (sd)</td>
<td>65.2 (10.2)</td>
<td>61.1 (10.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Smoking type (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never-smokers</td>
<td>1,101 (11.4)</td>
<td>2,353 (31.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ex-smokers</td>
<td>3,934 (40.8)</td>
<td>2,782 (37.6)</td>
<td></td>
</tr>
<tr>
<td>Current smokers</td>
<td>4,598 (47.7)</td>
<td>2,269 (30.6)</td>
<td></td>
</tr>
<tr>
<td>Pack-years of smoking among smokers (sd)</td>
<td>47.8 (31.3)</td>
<td>33.1 (26.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI, kg/m(^2) (sd)</td>
<td>26.3 (4.9)</td>
<td>26.9 (4.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI categories (kg/m(^2))</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Under weight (&lt;18.5)</td>
<td>268 (2.8)</td>
<td>66 (0.9)</td>
<td></td>
</tr>
<tr>
<td>Normal (18.5–24.9)</td>
<td>3,862 (40.1)</td>
<td>2,708 (36.6)</td>
<td></td>
</tr>
<tr>
<td>Over weight (25–29.9)</td>
<td>3,684 (38.2)</td>
<td>3,154 (42.6)</td>
<td></td>
</tr>
<tr>
<td>Obese (≥30)</td>
<td>1,819 (18.9)</td>
<td>1,476 (19.9)</td>
<td></td>
</tr>
</tbody>
</table>

The basic characteristics of the subjects were described as mean (sd) for continuous variables, and number (proportion, %) for category variables. The p-values were obtained by student t-test for continuous variables and Chi-square test for category variables.

https://doi.org/10.1371/journal.pone.0185660.t001
overweight (BMI 25–29.9), and to obese (BMI ≥30) with $p_{\text{trend}} = 8.40 \times 10^{-74}$. Similar associations between the BMI-GRS and BMI categories were found in cases and controls ($p_{\text{trend}} = 1.75 \times 10^{-39}$ and $p_{\text{trend}} = 5.96 \times 10^{-37}$ for cases and controls, respectively).

**Association of the BMI-GRS with smoking phenotypes**

The BMI-GRS of smokers that include both ex-smokers and current smokers, was first compared with that of never-smokers. Smokers had a significantly higher BMI-GRS than never-smokers.

![Figure 2](https://doi.org/10.1371/journal.pone.0185660.g002)

**Table 2. Partial correlations between BMI and pack-years of smoking by smoking status.**

<table>
<thead>
<tr>
<th>Smoking Status</th>
<th>Current Smokers</th>
<th>Ex-smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Coef</td>
</tr>
<tr>
<td>All*</td>
<td>6,577</td>
<td>0.054</td>
</tr>
<tr>
<td>Cases**</td>
<td>4,396</td>
<td>0.052</td>
</tr>
<tr>
<td>Controls**</td>
<td>2,181</td>
<td>0.072</td>
</tr>
</tbody>
</table>

A table for the partial correlation coefficients between BMI and pack-years in smokers.

* For all subjects, the analysis was adjusted for age, sex, study sites and disease status.

** For cases and controls, the analyses were adjusted for age, sex, and study sites.

https://doi.org/10.1371/journal.pone.0185660.t002
smokers (regression coefficient of ex-smokers: 0.516, p-value = 0.016) before adjustment for BMI. After further adjustment for BMI, the association between the BMI-GRS and smoking was strengthened (regression coefficient 0.545, p-value = 0.010). Ex-smokers and current smokers were then compared separately with that of never-smokers. Ex-smokers and current smokers had a similar BMI-GRS (regression coefficient 0.059, p-value = 0.750), but both of them had a significantly higher BMI-GRS than never-smokers (regression coefficient of ex-smokers: 0.490, p-value = 0.032; regression coefficient of current smokers: 0.549, p-value = 0.021) before adjustment for BMI. After further adjustment for BMI, the association between the BMI-GRS and current smoking was strengthened (regression coefficient 0.838, p-value = 3.9 × 10^{-4}), while the association between the BMI-GRS and ex-smoking was somewhat attenuated (regression coefficient 0.321, p-value = 0.157). The association patterns based on different BMI-GRSs were largely consistent (S2 Table). The results were also similar when analyses were stratified by lung cancer status.

There was also a significantly positive association between the BMI-GRS and pack-years of smoking among smokers (Table 3). The associations were similar in cases and controls. After stratification by smoking status, the association between the BMI-GRS and pack-years tended to be stronger in current smokers (correlation coefficient 0.024, p = 0.049) than in ex-smokers.
The above analyses suggested that among the 241 SNPs composed of the BMI-GRS (or in linkage disequilibrium with the 241 SNPs), there may be some pleiotropic SNPs that have direct effects on smoking and BMI. To identify the pleiotropic loci (DAGs 10, 11, or 12 in Fig 1), we first used network inference to determine the possible causal models of 241 SNPs. In total, five SNPs were classified into category four to be pleiotropic for both BMI and pack-years of smoking, and seven SNPs were classified into category four to be pleiotropic for both BMI and smoking by network inference. We then examined the associations of each of these SNPs with BMI and pack-years of smoking (or smoking status) with adjustment for age, sex, study sites, genetic principal components and lung cancer disease status. (Fig 4). There were four and three SNPs associated with pack-years of smoking and smoking status with p < 0.05, respectively (Table 4). The SNPs classified as Category 4 and associated with smoking status or pack-years with a nominal significance (p < 0.05) were considered as candidate SNPs of pleiotropy. The associations of these candidate pleiotropic SNPs with BMI were quite similar with and without adjustment for pack-years or smoking status (S4 Table). We validated these SNPs using data from the TAG consortium. Of the total of seven candidate pleiotropic SNPs for BMI and pack-years or smoking status, rs11030104 (BDNF) was associated with ever-smoking in TAG data after Bonferroni correction (p = 0.002), and rs13021737 (TMEM18), rs11583200 (ELAVL4) and rs6990042 (GCZ) achieved a nominal significance of 0.05 (p-values were 0.018, 0.008, and 0.043, respectively). Another interesting SNP that achieved a nominal significance in TAG data (p = 0.020) was rs12016871 (MTIF3), but it did not achieve statistical significance with smoking status in the OncoArray Project dataset (p = 0.161).

**Identifying pleiotropic SNPs for both BMI and smoking status**

The above analyses suggested that among the 241 SNPs composed of the BMI-GRS (or in linkage disequilibrium with the 241 SNPs), there may be some pleiotropic SNPs that have direct effects on smoking and BMI. To identify the pleiotropic loci (DAGs 10, 11, or 12 in Fig 1), we first used network inference to determine the possible causal models of 241 SNPs. In total, five SNPs were classified into category four to be pleiotropic for both BMI and pack-years of smoking, and seven SNPs were classified into category four to be pleiotropic for both BMI and smoking by network inference. We then examined the associations of each of these SNPs with BMI and pack-years of smoking (or smoking status) with adjustment for age, sex, study sites, genetic principal components and lung cancer disease status. (Fig 4). There were four and three SNPs associated with pack-years of smoking and smoking status with p < 0.05, respectively (Table 4). The SNPs classified as Category 4 and associated with smoking status or pack-years with a nominal significance (p < 0.05) were considered as candidate SNPs of pleiotropy. The associations of these candidate pleiotropic SNPs with BMI were quite similar with and without adjustment for pack-years or smoking status (S4 Table). We validated these SNPs using data from the TAG consortium. Of the total of seven candidate pleiotropic SNPs for BMI and pack-years or smoking status, rs11030104 (BDNF) was associated with ever-smoking in TAG data after Bonferroni correction (p = 0.002), and rs13021737 (TMEM18), rs11583200 (ELAVL4) and rs6990042 (GCZ) achieved a nominal significance of 0.05 (p-values were 0.018, 0.008, and 0.043, respectively). Another interesting SNP that achieved a nominal significance in TAG data (p = 0.020) was rs12016871 (MTIF3), but it did not achieve statistical significance with smoking status in the OncoArray Project dataset (p = 0.161).

**Discussion**

In summary, we calculated the BMI-GRS for subjects who had OncoArray data of ILCCO using 241 common SNPs potentially associated with BMI and demonstrated that the BMI-GRS was associated with increased propensity to smoke as well as elevated pack-years after adjusting for the potential confounding effects of BMI. These results were consistent with those from a previous study in Iceland in which the GRS of 32 SNPs identified in GWAS was found to be...
The observed associations between the BMI-GRS and smoking variables could not be due to confounding of BMI, because the association of the BMI-GRS with smoking variables remained statistically significant after adjustment for BMI. Moreover, the BMI-GRS was positively associated with current smoking, which was opposite to what would be expected if the association between BMI-GRS and current smoking was due to the confounding effects of BMI, as current smoking and BMI was negatively correlated. Instead, the associations between BMI-GRS and smoking indicate that some loci that composed the BMI-GRS may directly contribute to smoking behavior, and may have pleiotropic effects on both BMI and smoking variables.

Using causal network inference, we identified 4 loci that may have pleiotropic effects on BMI and pack-years of smoking and 3 loci with potential pleiotropic effects on BMI and smoking status. Among them, one locus (BDNF) achieved a statistical significance after Bonferroni correction (p<0.007), and three loci (TMEM18, ELAVL4, and SGCZ) achieved a nominal significance (p<0.05), in ever-smoking data from the TAG consortium. The result of BDNF (brain derived neurotrophic factor) locus on chromosome 11 was consistent with prior studies that
Table 4. SNPs with pleiotropic effects on BMI and smoking variables among 241 SNPs composed of the BMI-GRS.

<table>
<thead>
<tr>
<th>SNP</th>
<th>chr</th>
<th>position</th>
<th>Gene</th>
<th>OncoArray Project</th>
<th>GIANT</th>
<th>TAG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Association direction with smoking phenotype</td>
<td>p-value association with smoking variable</td>
<td>Association direction with BMI</td>
</tr>
<tr>
<td>rs13021737</td>
<td>2</td>
<td>632348</td>
<td>TMEM18</td>
<td>Positive</td>
<td>3.0e-4</td>
<td>Positive</td>
</tr>
<tr>
<td>rs152843</td>
<td>2</td>
<td>181550962</td>
<td>AC009478.1</td>
<td>Positive</td>
<td>6.4e-3</td>
<td>Positive</td>
</tr>
<tr>
<td>rs11583200</td>
<td>1</td>
<td>50559820</td>
<td>ELAVL4</td>
<td>Positive</td>
<td>0.011</td>
<td>Positive</td>
</tr>
<tr>
<td>rs3888190</td>
<td>16</td>
<td>28889486</td>
<td>ATP2A1</td>
<td>Positive</td>
<td>0.034</td>
<td>Positive</td>
</tr>
<tr>
<td>rs11165643</td>
<td>1</td>
<td>96924097</td>
<td>PTBP2</td>
<td>Positive</td>
<td>0.133</td>
<td>Positive</td>
</tr>
</tbody>
</table>

SNPs pleiotropic for smoking status and BMI

<table>
<thead>
<tr>
<th>SNP</th>
<th>chr</th>
<th>position</th>
<th>Gene</th>
<th>OncoArray Project</th>
<th>GIANT</th>
<th>TAG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Association direction with smoking phenotype</td>
<td>p-value association with smoking variable</td>
<td>Association direction with BMI</td>
</tr>
<tr>
<td>rs11030104</td>
<td>11</td>
<td>27684517</td>
<td>BDNF</td>
<td>Positive</td>
<td>2.4e-4</td>
<td>Positive</td>
</tr>
<tr>
<td>rs6990042</td>
<td>8</td>
<td>14173974</td>
<td>SGCZ</td>
<td>Positive</td>
<td>0.031</td>
<td>Positive</td>
</tr>
<tr>
<td>rs9275595</td>
<td>6</td>
<td>32681355</td>
<td>XXbac-BPG254F23.7</td>
<td>Positive</td>
<td>0.043</td>
<td>Positive</td>
</tr>
<tr>
<td>rs7550711</td>
<td>1</td>
<td>110082886</td>
<td>GPR61</td>
<td>Positive</td>
<td>0.094</td>
<td>Positive</td>
</tr>
<tr>
<td>rs929641</td>
<td>2</td>
<td>58792377</td>
<td>LINC01122</td>
<td>Positive</td>
<td>0.136</td>
<td>Positive</td>
</tr>
<tr>
<td>rs12016871</td>
<td>13</td>
<td>28017782</td>
<td>MTIF3</td>
<td>Negative</td>
<td>0.161</td>
<td>Positive</td>
</tr>
<tr>
<td>rs1220375</td>
<td>10</td>
<td>104901491</td>
<td>NT5C2</td>
<td>Positive</td>
<td>0.168</td>
<td>Positive</td>
</tr>
</tbody>
</table>

The SNPs presented in this table were classified as Category four (DAGs of 10, 11, 12 in Fig 1). The association directions and p-values of these SNPs with BMI in GIANT and with ever-smoking in TAG are also presented. SNPs in the shadow were statistically significant for the association with pack-years of smoking or smoking status in the OncoArray Project population (p<0.05) and with ever-smoking data of TAG consortium with a nominal significance of p<0.05.

https://doi.org/10.1371/journal.pone.0185660.t004
had shown strong associations of this locus with BMI \[18,39\] and various smoking phenotypes.\[17\] Evidence from epidemiological studies \[40\] and animal studies \[41\] also indicate associations of \textit{BDNF} gene with other substance abuse related disorders, eating disorders, and schizophrenia. The protein \textit{BDNF} belongs to a neurotrophin family growth factors \[42\] and is the most abundant of the neurotrophins in the brain with high concentrations in the hippocampus and cerebral cortex.\[43,44\] \textit{BDNF} expression in the brain is regulated by the serotonergic\[45\] and the dopaminergic\[46\] neurotransmitter systems which are known to be involved in nicotine use, addictive behaviors, mood and food intake. \[47–50\]

The \textit{TMEM18} (transmembrane protein 18) locus is another known GWAS locus of BMI. \[19\] The Icelandic study that examined 32 GWAS loci of BMI had found significant associations of \textit{TMEM18} with both smoking initiation and cigarettes per day, an observation that was consistent with what we found \[31\]. The function of \textit{TMEM18} is largely unknown. \textit{TMEM18} is highly expressed in neural tissue and has been hypothesized to play a role in energy homeostasis via neural pathways controlling food intake \[51\].

To our knowledge, both \textit{ELAVL4} (ELAV Like RNA Binding Protein 4) and \textit{SCGZ} (Sarcoglycan zeta) loci have not been associated with smoking behavior in GWAS. We examined the GTEx database, and both \textit{ELAVL4} and \textit{SCGZ} are highly expressed in multiple brain tissues (Fig 5). The \textit{ELAVL4} gene is known to be associated with hallucinogen abuse, paraneoplastic neurologic disorders, and Parkinson disease \[52\]. Although there was suggestive evidence of association between \textit{SCGZ locus and BMI}, it has not been considered as the GWAS BMI locus; \[20\] however, a previous copy number variation (CNV) analysis in two African American populations had identified a CNV overlapping with \textit{SCGZ} gene region to be significantly associated with BMI.\[53\] Previously, \textit{SCGZ} and other genes involved in cell adhesion processes were linked to addiction vulnerability.\[54\] Cell adhesion mechanisms are central for properly establishing and regulating neuronal connections during development and can play major roles in mnemonic processes in adults \[55–57\]. In addition to reward processes, there are growing bodies of data implicating that cell-adhesion and related memory-like processes play important roles in substance dependence.\[54,55,57,58\]

Future studies of identification of pleiotropic genes on both BMI and smoking phenotypes may focus on pathways of those candidate loci, in particular \textit{BDNF} gene. Among 241 SNPs, there was one SNP (rs3800229) in the locus of \textit{FOXO3} that can be inactivated by signaling pathways activated by neurotrophins (such as \textit{BDNF}).\[59,60\] This SNP was associated with pack-years in the OncoArray Project data and ever-smoking in TAG data with a nominal significance of \(p<0.05\) (data not shown), but this SNP did not achieve the cut-off to be classified into the pleiotropic category. Nevertheless, our finding on associations of \textit{BDNF} suggests the regulatory pathway of \textit{BDNF} and its other target loci may play a role in both smoking behavior and BMI.

In general, genetic variants, BMI and smoking phenotypes are in complex relationships. In addition to pleiotropic effects of genetic variants on BMI and smoking phenotypes, there may also be interactions between genetic variants and smoking behavior on BMI. For example, a recent study identified several novel BMI loci by accounting for SNP-smoking interactions. \[61\] In the presence of such interaction, one would also expect associations between a SNP and smoking status in BMI-based ascertained samples, although the SNPs is not associated with smoking status in the general population. A future study fully accounting for these relationships may reveal additional novel loci of obesity and smoking phenotypes.

In summary, we identified four potential loci that may have pleiotropic effects on BMI and smoking traits. All four potential pleiotropic loci on BMI and smoking phenotypes are expressed in the human brain, and prior experimental evidence indicates that these genes are involved in relevant complex brain functions, e.g. brain’s reward circuitry and neural cell adhesion mechanisms. The biological functions of these genes support our findings. Future studies
of confirmation of these loci may suggest targets for searching new drugs for controlling smoking and eating behaviors. Sequencing these genes and other genes in relevant pathways may be helpful for identifying functional variants that have pleiotropic effects on both BMI and smoking behavior.

**Supporting information**

S1 Table. 241 selected SNPs and the AIC values of different DAGs and minimum AIC values of different categories.

(DOCX)
S2 Table. The comparison of associations between different BMI-GRSs and smoking categories (n = 17,037).

S3 Table. The comparison of partial correlations between different BMI-GRSs and pack-years of smoking in smokers.

S4 Table. The comparison of associations between seven candidate pleiotropic SNPs and BMI before and after adjustment for smoking phenotypes.

Acknowledgments

This work was supported by an R21 grant (CA202529, Wang/Ho).

URLs


Author Contributions

Conceptualization: Tao Wang, Yun-Chul Hong, Gloria Y. F. Ho.


Formal analysis: Tao Wang, Jee-Young Moon, Yiqun Wu.

Funding acquisition: Gloria Y. F. Ho.


Methodology: Tao Wang, Jee-Young Moon, Gloria Y. F. Ho.

Supervision: Gloria Y. F. Ho.

Writing – original draft: Tao Wang.

Writing – review & editing: Yiqun Wu, Christopher I. Amos, Rayjean J. Hung, Adonina Tardon, Angeline Andrew, Chu Chen, David C. Christiani, Demetrios Albanes, Gloria Y. F. Ho.
References


