This is the published version of a paper published in *PLoS Genetics*.

Citation for the original published paper (version of record):

Ranking and characterization of established BMI and lipid associated loci as candidates for gene-environment interactions.
*PLoS Genetics*, 13(6): e1006812
https://doi.org/10.1371/journal.pgen.1006812

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-140239
Postdoctoral Fellowship (4.1-2016-00416). TVV is drawn from GIANT and are contained in the Foundation, all grants to PWF. DS is supported by Lung Foundation, and the Novo Nordisk

Characterization of BMI and lipid loci as candidates for G×E interactions
drawn from GIANT and are contained in the following articles “Genome-wide physical activity interactions in adiposity—A meta-analysis of 200,452 adults” (10.1371/journal.pgen.1005628) and “Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits” (10.1038/ncomms14977).

Funding: This research was undertaken as part of a research program supported by the European Commission (CoG-2015-681742_NASCENT), Swedish Research Council (Distinguished Young Researchers Award in Medicine), Swedish Heart-Lung Foundation, and the Novo Nordisk Foundation, all grants to PWF. DS is supported by the Swedish Research Council International Postdoc Fellowship (4.1-2016-00416). TVV is supported by the Novo Nordisk Foundation Postdoctoral Fellowship within Endocrinology/Metabolism at International Elite Research Environments via NNF16OC0020698. TTW was supported by the grants “Bundesministerium für Bildung und Forschung”: BMBF-01ER1206, BMBF-01ER1507. APM is a Wellcome Trust Senior Fellow in Basic Biomedical Science (grant WT098017). LAC acknowledges funding for the Framingham Heart Study. This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine.

The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHArE) project. This work was partially supported by the National Heart, Lung and Blood Institute’s Framingham Heart Study (Contract No. N01-HC-25195 and Contract No. HHSN268201500001I) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. This research was partially supported by grant R01-DK089256 from the National Institute of Diabetes and Digestive and Kidney Diseases (MPIs: L.B. Borecki, LAC, K. North). TOK was supported by the Danish Council for Independent Research (DFF—1333-00124) and Sapere Aude program grant (DFF—1331-00730).

Author summary

Most contemporary studies of gene-environment interactions focus on gene variants that are known to bear strong and reliable associations with the traits of interest. The strategy is intuitive because it helps limit the number of tests performed by focusing on a relatively small number of gene variants. However, this approach is predicated on an implicit assumption that these loci are strong candidates for interactions owing to their established relationships with the index traits. The counter-argument is that, because these loci have highly consistent signals within and between populations that vary by environmental characteristics, the probability that these variants interact with other factors is low. The current analysis tests whether variants with strong marginal effects signals (i.e., those prioritized through conventional genome-wide association analyses) are strong or weak candidates for gene-environment interactions. Here we describe analyses focused on lipids and BMI that test this hypothesis by comparing marginal effect signals with variance effect signals and those derived from explicit genome-wide, gene-environment interaction analyses.
marginal effect signals is a pragmatic data-reduction step that may improve power [3], although this approach risks omitting other promising candidates [4].

In a linear regression setting, the presence of interaction effects drives phenotypic variance heterogeneity by genotype [3,5]. Exploiting variance heterogeneity as a signature of interactions is appealing because, unlike standard approaches for assessing G×E interactions, no explicit information about environmental exposures is needed [6] and multiple exposures can be simultaneously considered.

Here we explored whether loci identified in large-scale genome-wide association studies (GWAS) of blood lipids and body mass index (BMI) are strong candidates for G×E interactions by comparing genome-wide variance heterogeneity P-value distributions generated using Levene’s test against P-value distributions for marginal effects and explicit G×E interaction effects (for smoking and physical activity).

**Results**

We assessed between-genotype variance heterogeneity for up to 1,927,671 directly genotyped or imputed SNPs (HapMap II CEU reference panel [7]) that passed quality control (QC). Meta-analyses of Levene’s test summary statistics [8] were performed for BMI (n=44,211 participants), and blood concentrations of high-density lipoprotein cholesterol (HDL-C) (n≤34,315), low-density lipoprotein cholesterol (LDL-C) (n≤34,180), total cholesterol (TC) (n≤34,318) and triglycerides (TG) (n≤34,110). We then obtained marginal effects results for the same index traits and SNPs from publicly available GWAS summary data from the GIANT (Genetic Investigation of ANthropometric Traits) Consortium [9] and GLGC (Global Lipids Genetics Consortium) [10,11].

We compared the genome-wide marginal effects with between-genotype variance heterogeneity results for each of the five cardiometabolic traits by calculating the association between marginal effects (P_m) and variance heterogeneity (P_v) P-values using the rank-based Spearman correlation (ρ). This was done using a set of 42,710 pruned SNPs produced using the—indeed pairwise command in PLINK (see Materials and Methods) to account for linkage disequilibrium (LD) among variants.

As shown in Table 1 (see also Fig 1A and S1 Table), the Spearman’s ρ for the association between P_m and P_v for all pruned SNPs was of very small magnitude and only statistically significant for BMI. The exclusion of SNPs based on progressively more conservative P_m thresholds (P_m<0.05; P_m<10−4; previously established loci with P_m<5×10−8 in external datasets), saw corresponding improvements in the magnitude of these correlations, which were statistically significant for all traits except TC when focusing on previously established loci. The BMI correlation at the P_m<0.05 threshold, as well as the test of equality with ρ for all SNPs, was statistically significant, suggesting concordance between marginal and variance signals at a nominal level of significance. The odds ratio (OR) for a SNP to have both P_m<0.05 and P_v<0.05 as compared to P_v≥0.05 was 1.33 (95% CI: 1.12, 1.57) for BMI while the 95% CIs of ORs for other traits included 1. On the other hand, the P-value for a non-zero ρ for TG was statistically significant when focusing on the established loci and at P_m<10−4, suggesting concordance between marginal and variance signals at more conservative P_m thresholds.

We further compared P_m with interaction P-values from exposure-specific (smoking and physical activity) genome-wide interaction tests for BMI (P_m); this was only done for BMI owing to the requirement for an adequately powered external dataset (such a dataset was accessible through the GIANT consortium) (Table 2). Marginal effects GWAS were performed by strata of smokers vs. non-smokers and physically active vs. inactive participants (n = 210,316 European-ancestry adults [12]) respectively, and a heterogeneity test [12] was used to...
Table 1. Spearman correlations between marginal effects $P_m$ and heterogeneity of variance from Levene's test $P_v$.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Max Sample Size</th>
<th>All SNPs in analysis</th>
<th>SNPs with $P_m$&lt;0.05</th>
<th>SNPs with $P_m$&lt;10$^{-4}$</th>
<th>Known Loci</th>
<th>Odds ratio (SNPs with $P_m$&lt;0.05 and $P_v$&lt;0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># SNPs Spearman $\rho$ P-value</td>
<td># SNPs Spearman $\rho$ P-value</td>
<td># SNPs Spearman $\rho$ P-value</td>
<td># SNPs Spearman $\rho$ P-value</td>
<td># SNPs Spearman $\rho$ P-value</td>
<td># SNPs Spearman $\rho$ P-value</td>
</tr>
<tr>
<td>TC</td>
<td>34 318 0.001 0.89 0.026 0.22 0.24</td>
<td>126 0.062 0.49 0.50</td>
<td>69 0.188 0.12 0.13</td>
<td>0.97 (0.78–1.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TG</td>
<td>34 110 0.003 0.51 0.006 0.80 0.69</td>
<td>83 0.230 $3.61\times10^{-2}$ $3.87\times10^{-2}$</td>
<td>40 0.401 $1.03\times10^{-2}$ $1.00\times10^{-2}$</td>
<td>1.20 (0.99–1.44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL-C</td>
<td>34 315 0.006 0.24 0.001 0.97 0.77</td>
<td>95 0.074 0.48 0.45</td>
<td>68 0.200 0.10</td>
<td>9.54×10$^{-2}$ 1.12 (0.92–1.35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDL-C</td>
<td>34 180 0.005 0.29 0.013 0.55 0.73</td>
<td>100 0.055 0.59 0.62</td>
<td>53 0.258 $6.18\times10^{-2}$ $6.58\times10^{-2}$</td>
<td>1.06 (0.87–1.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>44 211 0.010 $4.56\times10^{-2}$ 0.066 $3.82\times10^{-3}$ $1.56\times10^{-2}$</td>
<td>68 0.201 9.98×10$^{-2}$ 0.12</td>
<td>71 0.236 $4.76\times10^{-2}$ $6.38\times10^{-2}$</td>
<td>1.33 (1.12–1.57)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BMI: body mass index; HDL-C: low-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; SNP: single nucleotide polymorphism; TC: total cholesterol; TG: triglycerides

https://doi.org/10.1371/journal.pgen.1006812.t001
generate exposure specific $P_{int}$ distributions. Spearman $\rho$ for the pruned set of SNPs in the SNP $\times$ physical activity and the SNP $\times$ smoking analyses were low and not statistically
significant (Table 2). We also compared \( P_{m} \) values and \( P_{v} \) values for BMI. Spearman’s \( \rho \) for the pruned set of SNPs were low and not statistically significant.

We next tested if the number of previously established marginal effect SNPs (\( P_{m} < 5 \times 10^{-8} \)) that were also nominally significant (\( P_{v} < 0.05 \)) for variance heterogeneity was greater than expected by chance (Tables 3 and 4, Fig 1). For 4 out of the 5 index traits, we observed enrichment at the lower end of the \( P_{v} \) distribution (\( P_{v} < 0.05 \)) for the established GWAS-derived lead SNPs. Thus, the nominally significant regions of the \( P_{v} \) distributions were generally enriched for GWAS-derived loci.

We also performed enrichment analyses to test if previously established marginal effects SNPs (\( P_{m} < 5 \times 10^{-8} \)) are enriched for nominally significant (\( P_{int} < 0.05 \)) interactions in the SNP \( \times \) physical activity or SNP \( \times \) Smoking analyses, but no enrichment was observed (Table 3; Fig 1B). By contrast, for the physical activity and smoking interaction tests (using all pruned SNPs), the lower end of the \( P_{int} \) distribution (\( P_{int} < 0.05 \)) was enriched with SNPs that were nominally significant in the Levene’s test analysis (\( P_{v} < 0.05 \)) (Table 4). This enrichment translated into an OR of 1.08 (95% CI: 1.01, 1.14) for a SNP to have \( P_{int} < 0.05 \) given \( P_{v} < 0.05 \) vs.

### Table 2. Spearman correlations between \( P_{int} \) in SNP \( \times \) Physical Activity and SNP \( \times \) Smoking on BMI analyses and marginal effects \( P_{m} \) or heterogeneity of variance from Levene’s test \( P_{v} \)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Max Sample Size</th>
<th>Max Sample Size PA/Smoking</th>
<th>All SNPs</th>
<th>SNPs with ( P_{m} &lt; 0.05 )</th>
<th>Known SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># SNPs</td>
<td>Spearman ( \rho )</td>
<td>( P ) -value</td>
<td>( # ) SNPs</td>
<td>Spearman ( \rho )</td>
</tr>
<tr>
<td>Marginal effects ( P_{m} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA ( \times ) SNP</td>
<td>322,144</td>
<td>180,271</td>
<td>41838</td>
<td>0.001</td>
<td>0.761</td>
</tr>
<tr>
<td>Smoking ( \times ) SNP</td>
<td>322,144</td>
<td>210,306</td>
<td>41371</td>
<td>-0.004</td>
<td>0.429</td>
</tr>
<tr>
<td>Levene’s test for homogeneity of variance ( P_{v} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA ( \times ) SNP</td>
<td>44,211</td>
<td>180,271</td>
<td>41838</td>
<td>0.005</td>
<td>0.35</td>
</tr>
<tr>
<td>Smoking ( \times ) SNP</td>
<td>44,211</td>
<td>210,306</td>
<td>41371</td>
<td>0.004</td>
<td>0.401</td>
</tr>
</tbody>
</table>

PA: physical activity; BMI: body mass index; SNP: single nucleotide polymorphism; \( P_{m} \): Variance (Levene’s) test \( P \) -value; \( P_{v} \): Marginal (linear regression) test \( P \) -value

https://doi.org/10.1371/journal.pgen.1006812.t002

Table 3. Enrichment of variance and gene \( \times \) environment interaction nominally significant results with GWAS-derived loci.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Analysis</th>
<th>Total SNPs/ Observed SNPs with ( P ) &lt; 0.05 (Expected)</th>
<th>( P_{binomial} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>Levene’s</td>
<td>71/10 (3.6)</td>
<td>( 3 \times 10^{-3} )</td>
</tr>
<tr>
<td></td>
<td>SNP ( \times ) PA</td>
<td>71/4 (3.6)</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>SNP ( \times ) Smoking</td>
<td>71/5 (3.6)</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>Average for SNP ( \times ) PA &amp; SNP ( \times ) Smoking</td>
<td>71/2 (3.6)</td>
<td>0.88</td>
</tr>
<tr>
<td>TG</td>
<td>Levene’s</td>
<td>40/9 (2)</td>
<td>( 1 \times 10^{-4} )</td>
</tr>
<tr>
<td>LDL-C</td>
<td>Levene’s</td>
<td>53/8 (2.7)</td>
<td>( 5 \times 10^{-3} )</td>
</tr>
<tr>
<td>HDL-C</td>
<td>Levene’s</td>
<td>68/6 (3.4)</td>
<td>0.12</td>
</tr>
<tr>
<td>TC</td>
<td>Levene’s</td>
<td>69/9 (3.5)</td>
<td>( 7 \times 10^{-3} )</td>
</tr>
</tbody>
</table>

PA: physical activity; BMI: body mass index; GWAS: genome-wide association study; HDL-C: low-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; SNP: single nucleotide polymorphism; TC: total cholesterol; TG: triglycerides

https://doi.org/10.1371/journal.pgen.1006812.t003
Table 4. Enrichment of SNPs with nominally significant $P_{m}$ for test of SNP × Smoking and SNP × Physical Activity interaction for BMI ($P_{m} < 0.05$) by SNPs with nominally significant Levene’s test ($P < 0.05$).

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Total SNPs with $P_{m} &lt; 0.05$</th>
<th>Observed SNPs with $P_{m} &lt; 0.05$ &amp; $P &lt; 0.05$ (Expected)</th>
<th>$P_{\text{binomial}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNP × PA</td>
<td>2142/159 (107.1)</td>
<td>8.52×10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>SNP × Smoking</td>
<td>2351/182 (117.6)</td>
<td>8.63×10⁻⁹</td>
<td></td>
</tr>
</tbody>
</table>

BMI: body mass index; PA: physical activity; SNP: single nucleotide polymorphism; $P_v$: Variance (Levene’s) test $P$-value; $P_m$ = G×E interaction (heterogeneity) test $P$-value; $P_{\text{binomial}}$ = significance of observing $P_m < 0.05$ more than expected by chance

https://doi.org/10.1371/journal.pgen.1006812.t004

$P_v \geq 0.05$ for SNP × physical activity interaction. The corresponding OR for the SNP × smoking interaction test was not significant (OR = 1.02; 95% CI: 0.96, 1.08).

Finally, in the pruned SNP-set we used the Mann–Whitney U test to probe for systematic differences in $P_v$ and $P_m$ ranks. $P$-values were ordered from least significant to most significant, and the lowest 100th centile (i.e. the most significantly associated SNPs) was compared to the remaining 99th percentile for each of the five traits. For BMI, SNPs in the lowest 100th centile of the $P_m$ distribution had markedly higher $P_v$ ranks (i.e. more significant $P_v$) than the remaining SNPs ($P_{\text{Mann–Whitney}} = 1.46×10^{-5}$; Table 5). Even when excluding previously established lead SNPs ($P_{m} < 5×10^{-8}$) for BMI (or SNPs +/-500kb proximal), SNPs from the lowest 100th centile of the $P_m$ rank-ordered distribution had higher $P_v$ ranks than the remaining SNPs ($P_{\text{Mann–Whitney}} = 4.30×10^{-4}$; Table 5). Conversely, no difference in $P_v$ ranks was observed for SNPs from the lowest 100th centile of the $P_m$ rank-ordered distribution for the four blood lipid traits; this may reflect trait-specific G×E effects or differences in statistical power by trait. No differences in $P_v$ ranks between SNPs from the lowest 99th centile of the $P_m$ rank-ordered distribution compared to SNPs from the 98th to 1st centiles of the distribution were observed for any trait ($P_{\text{Mann–Whitney}} > 0.05$; Table 5). Similarly, no difference in $P_m$ ranks was observed for SNPs from the lowest 100th centile of the $P_v$ rank-ordered distribution for any traits ($P_{\text{Mann–Whitney}} > 0.05$; Table 6).

Table 5. Comparison of Levene’s test $P_v$ ranks from different centiles of the $P_m$ rank-ordered distribution for the index traits.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Known SNPs</th>
<th>Min $P_m$ from 100th centile</th>
<th>Max $P_m$ from 100th centile</th>
<th>Median $P_v$ rank for 100th centile</th>
<th>Median $P_v$ rank for 99th-1st centiles</th>
<th>Mann-Whitney $P$-value</th>
<th>Min $P_m$ from 99th centile</th>
<th>Max $P_m$ from 99th centile</th>
<th>Median $P_v$ rank for 99th centile</th>
<th>Median $P_v$ rank for 98th-1st centiles</th>
<th>Mann-Whitney $P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>Included</td>
<td>4.78×10⁻⁹¹</td>
<td>5.82×10⁻³</td>
<td>58.82</td>
<td>49.93</td>
<td>$1.46×10^{-5}$</td>
<td>5.86×10⁻³</td>
<td>1.85×10⁻²</td>
<td>52.79</td>
<td>49.91</td>
<td>0.42</td>
</tr>
<tr>
<td>BMI</td>
<td>Excluded</td>
<td>3.59×10⁻⁶</td>
<td>8.56×10⁻³</td>
<td>55.78</td>
<td>49.95</td>
<td>$4.30×10^{-4}$</td>
<td>8.73×10⁻³</td>
<td>2.18×10⁻²</td>
<td>52.60</td>
<td>49.93</td>
<td>0.36</td>
</tr>
<tr>
<td>HDL-C</td>
<td>Included</td>
<td>3.56×10⁻⁵⁷</td>
<td>6.48×10⁻³</td>
<td>51.49</td>
<td>49.99</td>
<td>0.47</td>
<td>6.48×10⁻³</td>
<td>1.67×10⁻²</td>
<td>50.49</td>
<td>49.98</td>
<td>0.92</td>
</tr>
<tr>
<td>HDL-C</td>
<td>Excluded</td>
<td>6.68×10⁻¹¹</td>
<td>9.94×10⁻³</td>
<td>51.45</td>
<td>49.99</td>
<td>0.77</td>
<td>9.55×10⁻³</td>
<td>2.09×10⁻²</td>
<td>51.06</td>
<td>49.98</td>
<td>0.47</td>
</tr>
<tr>
<td>LDL-C</td>
<td>Included</td>
<td>3.80×10⁻¹⁴³</td>
<td>7.14×10⁻³</td>
<td>53.11</td>
<td>49.98</td>
<td>0.52</td>
<td>7.18×10⁻³</td>
<td>1.75×10⁻²</td>
<td>48.44</td>
<td>49.99</td>
<td>0.85</td>
</tr>
<tr>
<td>LDL-C</td>
<td>Excluded</td>
<td>2.03×10⁻¹¹</td>
<td>9.88×10⁻³</td>
<td>53.42</td>
<td>49.97</td>
<td>0.38</td>
<td>9.90×10⁻³</td>
<td>2.09×10⁻²</td>
<td>48.37</td>
<td>49.99</td>
<td>1.00</td>
</tr>
<tr>
<td>TG</td>
<td>Included</td>
<td>2.23×10⁻¹³</td>
<td>8.18×10⁻³</td>
<td>53.73</td>
<td>49.98</td>
<td>0.32</td>
<td>8.19×10⁻³</td>
<td>1.92×10⁻²</td>
<td>52.42</td>
<td>49.95</td>
<td>0.63</td>
</tr>
<tr>
<td>TG</td>
<td>Excluded</td>
<td>1.00×10⁻¹⁰</td>
<td>1.06×10⁻²</td>
<td>51.27</td>
<td>49.99</td>
<td>0.64</td>
<td>1.06×10⁻²</td>
<td>2.21×10⁻²</td>
<td>53.23</td>
<td>49.95</td>
<td>0.41</td>
</tr>
<tr>
<td>TC</td>
<td>Included</td>
<td>1.41×10⁻¹⁰⁷</td>
<td>5.85×10⁻³</td>
<td>52.03</td>
<td>49.98</td>
<td>0.32</td>
<td>5.87×10⁻³</td>
<td>1.49×10⁻²</td>
<td>51.21</td>
<td>49.97</td>
<td>0.62</td>
</tr>
<tr>
<td>TC</td>
<td>Excluded</td>
<td>3.11×10⁻¹¹</td>
<td>9.14×10⁻³</td>
<td>49.43</td>
<td>50.01</td>
<td>0.66</td>
<td>9.15×10⁻³</td>
<td>1.91×10⁻²</td>
<td>50.12</td>
<td>50.01</td>
<td>0.93</td>
</tr>
</tbody>
</table>

BMI: body mass index; HDL-C: low-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; SNP: single nucleotide polymorphism; TC: total cholesterol; TG: triglycerides; $P_v$: Variance (Levene’s) test $P$-value; $P_m$: marginal (linear regression) test $P$-value

https://doi.org/10.1371/journal.pgen.1006812.t005
Table 6. Comparison of marginal effects $P_m$ ranks from different centiles of the Levene’s test $P_v$ rank-ordered distribution for the index traits.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Known SNPs</th>
<th>Min $P_m$ from 100th centile</th>
<th>Max $P_m$ from 100th centile</th>
<th>Median $P_m$ rank for 100th centile</th>
<th>Median $P_m$ rank for 99th-1st centiles</th>
<th>Mann-Whitney $P$-value</th>
<th>Min $P_v$ from 99th centile</th>
<th>Max $P_v$ from 99th centile</th>
<th>Median $P_v$ rank for 99th centile</th>
<th>Median $P_v$ rank for 98th-1st centiles</th>
<th>Mann-Whitney $P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>Included</td>
<td>2.95×10^{-7}</td>
<td>6.31×10^{-3}</td>
<td>51.28</td>
<td>49.53</td>
<td>0.51</td>
<td>6.33×10^{-3}</td>
<td>1.30×10^{-2}</td>
<td>53.57</td>
<td>49.53</td>
<td>0.13</td>
</tr>
<tr>
<td>BMI</td>
<td>Excluded</td>
<td>2.95×10^{-7}</td>
<td>6.38×10^{-3}</td>
<td>51.40</td>
<td>49.48</td>
<td>0.42</td>
<td>6.38×10^{-3}</td>
<td>1.30×10^{-2}</td>
<td>53.50</td>
<td>49.44</td>
<td>0.17</td>
</tr>
<tr>
<td>HDL-C</td>
<td>Included</td>
<td>2.04×10^{-5}</td>
<td>9.44×10^{-3}</td>
<td>46.28</td>
<td>50.04</td>
<td>0.52</td>
<td>9.45×10^{-3}</td>
<td>1.90×10^{-2}</td>
<td>53.06</td>
<td>50.01</td>
<td>0.44</td>
</tr>
<tr>
<td>HDL-C</td>
<td>Excluded</td>
<td>2.04×10^{-5}</td>
<td>9.45×10^{-3}</td>
<td>46.42</td>
<td>50.05</td>
<td>0.37</td>
<td>9.47×10^{-3}</td>
<td>1.89×10^{-2}</td>
<td>53.37</td>
<td>50.01</td>
<td>0.31</td>
</tr>
<tr>
<td>LDL-C</td>
<td>Included</td>
<td>1.06×10^{-8}</td>
<td>9.12×10^{-3}</td>
<td>52.96</td>
<td>49.98</td>
<td>0.19</td>
<td>9.15×10^{-3}</td>
<td>1.88×10^{-2}</td>
<td>50.78</td>
<td>49.96</td>
<td>0.99</td>
</tr>
<tr>
<td>LDL-C</td>
<td>Excluded</td>
<td>1.44×10^{-5}</td>
<td>9.37×10^{-3}</td>
<td>50.39</td>
<td>49.99</td>
<td>0.64</td>
<td>9.37×10^{-3}</td>
<td>1.92×10^{-2}</td>
<td>51.85</td>
<td>49.97</td>
<td>0.68</td>
</tr>
<tr>
<td>TG</td>
<td>Included</td>
<td>2.45×10^{-6}</td>
<td>8.39×10^{-3}</td>
<td>49.93</td>
<td>50.01</td>
<td>0.60</td>
<td>8.39×10^{-3}</td>
<td>1.78×10^{-2}</td>
<td>51.75</td>
<td>50.01</td>
<td>0.53</td>
</tr>
<tr>
<td>TG</td>
<td>Excluded</td>
<td>2.45×10^{-6}</td>
<td>8.37×10^{-3}</td>
<td>49.23</td>
<td>50.01</td>
<td>0.66</td>
<td>8.39×10^{-3}</td>
<td>1.78×10^{-2}</td>
<td>51.92</td>
<td>50.00</td>
<td>0.51</td>
</tr>
<tr>
<td>TC</td>
<td>Included</td>
<td>3.28×10^{-5}</td>
<td>1.08×10^{-2}</td>
<td>51.61</td>
<td>49.98</td>
<td>0.16</td>
<td>1.08×10^{-2}</td>
<td>2.09×10^{-2}</td>
<td>50.29</td>
<td>49.98</td>
<td>0.92</td>
</tr>
<tr>
<td>TC</td>
<td>Excluded</td>
<td>3.28×10^{-5}</td>
<td>1.10×10^{-2}</td>
<td>51.23</td>
<td>50.00</td>
<td>0.33</td>
<td>1.10×10^{-2}</td>
<td>2.10×10^{-2}</td>
<td>49.92</td>
<td>50.00</td>
<td>0.93</td>
</tr>
</tbody>
</table>

BMI: body mass index; HDL-C: low-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; SNP: single nucleotide polymorphism; TC: total cholesterol; TG: triglycerides; $P_v$: Variance (Levene’s) test $P$-value; $P_m$: marginal (linear regression) test $P$-value

https://doi.org/10.1371/journal.pgen.1006812.1006

To assess whether a trait with a non-normal distribution (e.g. BMI) or strong marginal associations could cause spurious association between the marginal and variance signals, we recapitulated the analysis pipeline (correlation analysis, enrichment analysis, comparisons of rank $P_m$ and $P_v$ values) in simulations described in the Materials and Methods. Careful assessment of results emanating from these simulations did not reveal evidence of type I error inflation caused by the non-normal distribution of an outcome trait nor strong marginal effects. For instance, we extracted correlation $P$-values of $P_m$, $P_v$, and $P_{int}$ generated from 5,000 simulations. QQ-plots of the 5,000 correlation $P$-values, 2,500 binomial $P$-values, and 2,500 Mann-Whitney U test $P$-values revealed no inflation (S1A–S1C Fig, S2A and S2B Fig and S3A and S3B Fig, respectively). Repeating these analyses on subsets of SNPs with low $P_m$ values did not materially change the results.

Discussion

Collectively, our analyses highlight a few variants with genome-wide significant marginal effects that may be strong candidates for G×E interactions owing to their strong concurrent variance heterogeneity $P$-values. For BMI, such SNPs are also overrepresented in the nominally significant part of the $P_v$ distribution. $FTO$ is an excellent example, as it conveys strong marginal effects [13], exhibits high between-genotype heterogeneity here (Tables 2 and 3 and Fig 1B) and elsewhere [5], and reportedly interacts with physical activity, diet and other lifestyle exposures [2,14,15] and is associated with macronutrient intake [16,17].

Although variance heterogeneity tests are potentially powerful screening tools for G×E interactions, like most interaction tests, they may be bias prone. For example, apparent differences in phenotypic variances across genotypes may be caused by scaling, particularly when the phenotypic means also differ substantially [18], such that the per-genotype means and variances for index traits are correlated. However, where necessary we transformed variables, and the correlations between $P_m$ and $P_v$ were generally weak, excluding this as a likely source of bias. Using simulated data, we investigated whether the non-normal distribution of a trait can cause a spurious association between marginal and variance signals, which we show is highly improbable. Through further simulations, we assessed whether SNPs with large marginal effects inflate $P_v$, but observed no inflation, indicating that large genetic marginal effects do
not artificially inflate variance heterogeneity to a meaningful extent, and SNPs with low $P_m$ and low $P_v$-values are thus likely to be strong candidates for G×E interactions, at least in the case of BMI. It might also be that combining populations from ancestral (e.g., hunter-gatherers) and contemporary environments increases variance heterogeneity owing to diversity in population substructure rather than G×E interactions per se [19]. However, this seems unlikely here, as the cohorts examined are from Westernized European-ancestry populations.

There are several additional explanations for between-genotype variance heterogeneity, such as variance misclassification that can occur when the index variant is located within a haplotype containing rare functional variants that convey strong marginal effects [5]. Hence, although variance heterogeneity tests represent a useful data-reduction step, before conclusions are drawn about the presence or absence of G×E interactions, index variants should be validated by testing their interactions with explicit environmental exposures, as we did here with smoking and physical activity. However, genome-wide G×E interactions datasets are not comprised of functionally validated G×E interactions, as no such resource is currently available for human complex traits. This limitation inhibits the extent to which causal effects can be attributed to the top-ranking loci and their interactions with smoking or physical activity.

We conclude that the common approach of prioritizing loci with established genome-wide significant association signals without further discrimination for G×E interaction analyses might be useful, but the efficiency of such analyses could be substantially improved by focusing on variants with low $P$-values for both variance heterogeneity and marginal effects. We provide these rankings here to facilitate this approach.

**Materials and methods**

A detailed project flow-chart is shown in Fig 2.

**Study sample**

We performed a genome-wide search for SNPs whose associations with the following traits are characterized by high between-genotype variance heterogeneity: BMI, TC, TG, HDL-C and LDL-C. The variance heterogeneity analyses were performed using Levene’s test [20] in up to

---

**Fig 2. Data flow-chart.** Three sources of genome-wide results were used: i) meta-analysis of Levene’s test results for between-genotype heterogeneity of phenotypic variances; ii) published results for marginal effects genome-wide association studies undertaken by the GIANT and GLGC consortia; iii) published results for SNP × physical activity and SNP × smoking in BMI (from the GIANT consortium).

https://doi.org/10.1371/journal.pgen.1006812.g002
44,211 participants of European descent from seven population-based cohorts. Descriptions of these cohorts are presented in S2 Table. To minimize bias that might result from unequal sample sizes between SNPs when calculating the correlations between the P-values from the marginal (P_m) and variance heterogeneity (P_v) meta-analyses, we restricted the sample size for analyses to 26,000 participants for BMI and to 24,000 participants for lipid traits (S4 Fig).

Genotyping and imputation
A detailed summary of sample sizes, genotyping platforms, genotype calling algorithms, sample and SNP quality control filters, and analysis software for all participating cohorts are provided in S2 and S3 Tables. For each individual, SNPs were imputed using the CEU reference panel of HapMap II [7] (S2 Table). We excluded SNPs with low imputation quality (below 0.3 for MACH, 0.4 for IMPUTE, and 0.8 for PLINK imputed data), Hardy-Weinberg equilibrium \( P < 10^{-6} \), directly genotyped SNP call rate < 95%, and minor allele frequency (MAF) < 1%.

Selection of SNPs identified through GWAS
We identified SNPs that have been robustly associated (\( P < 5 \times 10^{-8} \)) with the five cardiometabolic traits in European ancestry populations: 77 SNPs associated with BMI discovered by GIANT [9]; and 58 SNPs associated with LDL-C, 71 SNPs associated with HDL-C, 74 SNPs associated with TC, and 40 SNPs associated with TG [10,11] discovered by GLGC.

Variance heterogeneity analyses
We used Levene’s test [20] to identify SNPs that show heterogeneity of phenotypic variances (\( \sigma^2 \)) across the three genotype groups at each SNP locus (\( i = 0, 1, \text{ or } 2 \)). We first log_{10} transformed all five traits followed by a z-score transformation by subtracting the sample mean and dividing by the sample standard deviation (SD), and further Winsorized the z-score values at 4 SD. The transformed phenotype \( Y \) was then used to calculate \( Z \), defined by the absolute deviation of each participant’s phenotype from the sample mean of his or her respective genotype group at a given SNP locus. For each trait, participating cohorts provided the necessary summary statistics for each genotype at each marker [8]. Specifically, the per genotype group counts \( (n_{0i}, n_{1i}, n_{2i}) \), per genotype means \( (\bar{Z}_{0i}, \bar{Z}_{1i}, \bar{Z}_{2i}) \), and per genotype group variances of \( Z \left( \sigma_{0i}^2, \sigma_{1i}^2, \sigma_{2i}^2 \right) \) were centrally collected and meta-analyzed. The minimum number of observations per genotype group required is 30 participants per cohort.

Meta-analyses were performed using the following formula, derived previously [8]:

\[
L = \frac{(N - 3)}{(3 - 1)} \cdot \left( \frac{\sum_{i=0}^{2} \sigma_{zi}^2 \cdot (\sum_{s} \bar{Z}_s \cdot \omega_s)^2 - \left( \sum_{i=0}^{2} \sum_{s} \bar{Z}_s \cdot (\omega_s \cdot \gamma_i) \right)^2}{\sum_{i=0}^{2} \sum_{s} \sigma_{zi}^2 \cdot \omega_s \cdot \gamma_i} \right)
\]

Where \( N \) is the combined sample size, \( \bar{Z}_s \) and \( \sigma_{zi}^2 \) are the sample mean and variance of \( Z \) in the \( i^{th} \) genotype group of the \( s^{th} \) study, respectively. When combining summary-level data to calculate the Levene’s test statistics \( L \), the following natural weights \( \omega_s \) and \( \gamma_i \) were calculated: \( \omega_s = \frac{n_s}{\sum n_s} \) and \( \gamma_i = 1 \), where \( n_s \) the sum of genotype counts in the \( s^{th} \) genotype group across all participating cohorts. These weights are determined by the frequency of the marker amongst the cohorts, such that the sum of both weights is equal to 1, i.e. \( \sum_i \omega_s = 1 \) and \( \sum \gamma_i = 1 \). The meta-analysis Levene’s test \( P \)-value is obtained by comparing \( L \) to an F-distribution with \( df_1 = 2 \) and \( df_2 = N - 3 \).
Comparison between marginal effects and variance heterogeneity P-values

Marginal effects P-values for BMI and the relevant lipid traits were obtained from publically available GWAS summary data from the GIANT [9] and GLGC [10,11] consortia, respectively (all cohorts included here in the Levene’s meta-analysis were also included in the GIANT and GLGC datasets).

To illustrate our findings, we rank-ordered the P-values (from lowest to highest) from both marginal effects and variance effects analyses for all 1,927,671 SNPs so that the lowest P-value for a given trait was assigned a rank equal to the lowest 100th centile. These rank-scaled distributions for \( P_m \) for all five traits are presented in Fig 1.

We calculated Spearman’s correlations for each of the five cardiometabolic traits between \( P_m \) and \( P_v \). This was done using a pruned set of SNPs. Pruning was performed in the TwinGene cohort using the--indep-pairwise 50 5.0 1 command in PLINK [21] by calculating LD (\( r^2 \)) for each pair of SNPs within a window of 50 SNPs, removing one of a pair of SNPs if \( r^2 > 0.1 \); we proceeded by shifting the window 5 SNPs forwards and repeating the procedure. Spearman’s correlations were computed for categories of SNPs: i) all pruned SNPs, ii) the subset of SNPs that was nominally significant (\( P_m < 0.05 \)) in the marginal effects analysis, iii) the subset of SNPs with \( P_m < 10^{-4} \) in the marginal effects analysis, and iv) SNPs that were previously established in conventional marginal effects GWAS meta-analyses (\( P_m < 5 \times 10^{-8} \)). We also compared Spearman’s correlations between these categories of SNPs using the test for equality of two correlations [22].

Next, we performed enrichment analyses to test if there was a higher number of established SNPs in the nominally significant variance P-value (\( P_v < 0.05 \)) distribution than expected by chance under the binominal distribution.

We also tested if there is a difference in \( P_v \) ranks for SNPs from the lowest 100th centile of the \( P_m \) rank-ordered distribution for all five traits and the rest of SNPs in the pruned set of SNPs using the Mann–Whitney U test, including and excluding established SNPs (or SNPs that were +/-500kb from the reported lead SNP). This analysis was repeated for SNPs from the 99th centile vs SNPs from 1st to 98th centiles of the \( P_m \) rank-ordered distribution. The same Mann–Whitney U tests were used to study differences in \( P_m \) ranks for SNPs from the lowest 100th and 99th centiles of the \( P_v \) rank-ordered distribution and the rest of SNPs in the pruned set of SNPs.

All analyses were performed using Stata 12 (StataCorp LP, TX, USA), unless specified otherwise.

SNP × Physical activity and SNP × Smoking interaction analyses for the outcome of BMI

We used now published data from 210,316 European-ancestry adults (from the GIANT consortium) pertaining to marginal effects meta-analyses for BMI that had been performed separately by strata of smoking (45,968 smokers vs. 164,355 non-smokers) [23]. The genetic marginal effect estimates, calculated separately within each of the two strata, were compared using a heterogeneity test [12] to infer the presence or absence of SNP × smoking interaction effects. The same analyses were performed using physical activity as a binary stratifying variable in up to 180,287 European-ancestry adults (42,065 physically active vs. 138,222 physically inactive) [24]. We calculated Spearman correlations between the P-values derived from the marginal effects meta-analysis and the \( P_{int} \) from the interaction effects meta-analysis (i.e., the between-strata heterogeneity test for SNP × smoking and SNP × physical activity interactions from the GIANT consortium); these tests were undertaken for all SNPs and those SNPs that
were nominally significant ($P_m<0.05$) in the marginal effects analysis. We then performed enrichment analyses to test if the numbers of nominally significant ($P_{int}<0.05$) GWAS-derived SNPs from both $SNP \times physical \ activity$ and $SNP \times smoking$ analyses were greater than expected by chance under the binomial distribution. We further calculated the OR of having $P_{int}<0.05$ given $P_v<0.05$ versus $P_v \geq 0.05$ both $SNP \times physical \ activity$ and $SNP \times smoking$ interaction analyses in a pruned set of TwinGene SNPs produced using the—indep-pairwise 50 5 0.8 command in PLINK [21].

Thereafter, we calculated the average rank for each SNP’s ranking on the $P_{int}$ rank-ordered distributions from the $SNP \times smoking$ and $SNP \times physical \ activity$ interaction analyses and performed enrichment analysis using these average ranks with $>95^{th}$ centile instead of $P_{int}<0.05$ as the cut-off.

Simulations

We simulated genetic data for 44,000 individuals from a pruned set of 50,335 SNPs with allele frequencies, effect estimates and $P_m$ values drawn from the GIANT consortium. We generated an outcome trait by summing the products of the simulated allele counts and effect estimates over all SNPs for each individual, and subsequently added a randomly generated non-normal error term such that the trait resembles the observed distribution of the transformed BMI trait used in the main (real data) analyses. We also simulated a fixed binary interacting factor with 30% prevalence. Using this simulated dataset, we calculated $P_m$, $P_v$ and $P_{int}$ values for each SNP and undertook i) pairwise Spearman correlation analyses between $P_m$, $P_v$ and $P_{int}$ values (5,000 simulations), ii) enrichment analysis using binomial tests (2,500 simulations) and iii) Mann-Whitney U tests to determine systematic differences in $P_v$ and $P_{int}$ ranks (2,500 simulations). Following the same pipeline, we created additional simulated datasets narrowing down SNPs to i) those with $P_m$ values from the lowest percentile ($n=504$; highest $P_m=5\times10^{-3}$) and to ii) genome-wide significant SNPs ($n=71$; $P_m<5\times10^{-8}$), and tested the pairwise Spearman correlation for $P_m$, $P_v$ and $P_{int}$ values (1,000 simulations for both sets). Simulations were run using the statistical software R (v. 3.3.2).[25]

Supporting information

S1 Fig. A: Quantile-quantile plot of Spearman correlation test $P$-values for ranks of $P_m$ and $P_v$. Quantile-quantile plot of Spearman correlation test $P$-values for ranks of $P_m$ and $P_v$. The figure illustrates 5,000 Spearman correlation $P$ values testing for correlation between $P_m$ and $P_v$ values drawn from a simulated dataset of 44,000 individuals and 50,335 SNPs. In the figure, distribution under the null hypothesis is represented as a black line while its 95% confidence interval is represented as dashed gray lines. The dashed red line represents the correlation $P$ value obtained from the “real data” analysis presented in the main text. B. Quantile-quantile plot of Spearman correlation test $P$-values for ranks of $P_m$ and $P_{int}$. Quantile-quantile plot of Spearman correlation test $P$-values for ranks of $P_m$ and $P_{int}$. The figure illustrates 5,000 Spearman correlation $P$ values testing for correlation between $P_m$ and $P_{int}$ values drawn from a simulated dataset of 44,000 individuals and 50,335 SNPs. In the figure, distribution under the null hypothesis is represented as a black line while its 95% confidence interval is represented as dashed gray lines. C. Quantile-quantile plot of Spearman correlation test $P$-values for ranks of $P_{int}$ and $P_v$. Quantile-quantile plot of Spearman correlation test $P$-values for ranks of $P_{int}$ and $P_v$. The figure illustrates 5,000 Spearman correlation $P$ values testing for correlation between $P_{int}$ and $P_v$ values drawn from a simulated dataset of 44,000 individuals and 50,335 SNPs. In the figure, distribution under the null hypothesis is represented as a
black line while its 95% confidence interval is represented as dashed gray lines.

**S2 Fig. A. Quantile-quantile plot of binomial test P-values for enrichment of variants with \( P_v < 0.05 \) among variants with \( P_m < 0.05 \).** Quantile-quantile plot of binomial test P-values for enrichment of variants with \( P_v < 0.05 \) among variants with \( P_m < 0.05 \). The figure illustrates 2,500 binomial P values testing for enrichment of variants with \( P_v < 0.05 \) among all variants with \( P_m < 0.05 \). \( P_v \) and \( P_m \) values drawn from a simulated dataset of 44,000 individuals and 50,335 SNPs. In the figure, distribution under the null hypothesis is represented as a black line while its 95% confidence interval is represented as dashed gray lines. The dashed red line represents the correlation P value obtained from the “real data” analysis presented in the main text.

**S3 Fig. A. Quantile-quantile plot of Mann-Whitney U test P-values for systematic differences in \( P_v \), ranks among variants with top ranking and lower ranking \( P_m \) values.** Quantile-quantile plot of Mann-Whitney U test P-values for systematic differences in \( P_v \), ranks among variants with top ranking and lower ranking \( P_m \) values. The figure illustrates 2,500 Mann-Whitney U P values testing for systematic differences in \( P_v \) ranks among those variants with the most significant \( P_m \) values (100th percentile of \( P_m \) distribution) and the remaining variants (1–99 percentile of \( P_m \) distribution). \( P_v \) and \( P_m \) values drawn from a simulated dataset of 44,000 individuals and 50,335 SNPs. In the figure, distribution under the null hypothesis is represented as a black line while its 95% confidence interval is represented as dashed gray lines. The dashed red line represents the correlation P value obtained from the “real data” analysis presented in the main text. **B. Quantile-quantile plot of Mann-Whitney U test P-values for systematic differences in \( P_m \), ranks among variants with top ranking and lower ranking \( P_v \) values.** Quantile-quantile plot of Mann-Whitney U test P-values for systematic differences in \( P_m \) ranks among variants with top ranking and lower ranking \( P_v \) values. The figure illustrates 2,500 Mann-Whitney U P values testing for systematic differences in \( P_m \) ranks among those variants with the most significant \( P_v \) values (100th percentile of \( P_v \) distribution) and the remaining variants (1–99 percentile of \( P_v \) distribution). \( P_v \) and \( P_m \) values drawn from a simulated dataset of 44,000 individuals and 50,335 SNPs. In the figure, distribution under the null hypothesis is represented as a black line while its 95% confidence interval is represented as dashed gray lines. The dashed red line represents the correlation P value obtained from the “real data” analysis presented in the main text.

**S4 Fig. Quantile-quantile plots of Levene’s test P-values for SNP associations with lipid traits and BMI.** Associations between SNPs and BMI (A), LDL (B), HDL (C), TG (D), TC (E) are presented. Only SNPs with \( N \geq 26,000 \) samples for BMI and \( N \geq 24,000 \) for lipid traits are shown. In each sub-figure, distribution under the null hypothesis is represented as a black line while its 95% confidence interval is represented as dashed gray lines.
S1 Table. Detailed results for known BMI, LDL-C, HDL-C, TG and TC loci.
(XLSX)

S2 Table. Study design, number of participants and sample quality control for genome-wide association study cohorts.
(XLSX)

S3 Table. Information on genotyping methods, quality control of SNPs, imputation, and statistical analysis.
(XLSX)

S1 Text. GIANT consortium contributors and their affiliations.
(PDF)

Author Contributions

Conceptualization: DS WQD RM TVV GP PWF.

Data curation: DS WQD.

Formal analysis: DS WQD RM TVV JL EM AM.

Funding acquisition: GP PWF.

Methodology: DS WQD TVV GP PWF.

Project administration: GP PWF.

Resources: APM NGF Cli PKEM NLP GH AYC AEJ MG TWW LMR CLa LAC PMR NJW KKO RJFL DIC EI TOK RAS GP PWF.

Software: DS WQD TVV.

Supervision: GP PWF.

Visualization: DS TVV.

Writing – original draft: DS GP PWF.

Writing – review & editing: DS WQD RM TVV JL EM AM APM NGF Cli PKEM NLP GH AYC AEJ MG TWW LMR CLa LAC PMR NJW KKO RJFL DIC EI TOK RAS GP PWF.

References


