Modulation of Genetic Associations with Serum Urate Levels by Body-Mass-Index in Humans

Huffman, Jennifer E.

2015-03-26


http://hdl.handle.net/10138/160337
https://doi.org/10.1371/journal.pone.0119752

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.
Modulation of Genetic Associations with Serum Urate Levels by Body-Mass-Index in Humans


1 Medical Research Council (MRC) Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine (IGMM), University of Edinburgh, Edinburgh, United Kingdom, 2 Institute of Genetic Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany, 3 Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany, 4 King's College London, St. Thomas' Hospital Campus, London, United Kingdom, 5 Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland, 6 Swiss Institute of Bioinformatics, Lausanne, Switzerland, 7 William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom, 8 Institute for Maternal and Child Health—Istituto Di Ricovero e Cura a Carattere Scientifico (IRCCS) “Burlo Garofolo”, Trieste, Italy, 9 University of Trieste, Trieste, Italy, 10 Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Milan, Italy, 11 Department of Psychology, The University of Edinburgh, Edinburgh, United Kingdom, 12 Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh, United Kingdom, 13 Estonian Genome Center, University of Tartu, Tartu, Estonia, 14 Broad Institute, Cambridge, MA, United States of America, 15 Children’s Hospital Boston, Boston, MA, United States of America, 16 Department of Chronic Disease Prevention, National Institute for Health and Welfare (THL), Helsinki, Finland, 17 Department of Cardiovascular Medicine, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, 18 Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States of America, 19 Clinical Research Branch, National Institute on Aging, Baltimore, MD, United States of America, 20 Member of Netherlands Consortium for Healthy Aging (NCHA) sponsored by Netherlands Genomics Initiative (NGI), Leiden, The Netherlands.
Abstract

We tested for interactions between body mass index (BMI) and common genetic variants affecting serum urate levels, genome-wide, in up to 42569 participants. Both stratified genome-wide association (GWAS) analyses, in lean, overweight and obese individuals, and regression-type analyses in a non BMI-stratified overall sample were performed. The former did not uncover any novel locus with a major main effect, but supported modulation of effects for some known and potentially new urate loci. The latter highlighted a SNP at RBFOX3 reaching genome-wide significant level (effect size 0.014, 95% CI 0.008-0.02,
Introduction

Epidemiological studies have associated hyper- and hypo-uricemia with multiple common diseases and conditions in humans [1]; hyperuricemia clusters with all metabolic syndrome components and is a causal risk factor for gout development. To date, 28 loci have been identified and replicated accounting for about 7% of the inter-individual variation in age and sex adjusted serum urate (SU) levels [2]. The top two loci, which account for about half of the genetic variance explained so far, have been noted to display marked gender differences in their effect [3–6], while other urate loci not [2,7]. Variants in the solute carrier SLC2A9 (also known as GLUT9) gene have doubled the effect on SU in women (0.40 standard deviation (sd) in [7]) than that observed in men, and variants in the transporter ABCG2 gene have a stronger effect in men than in women (0.22 sd versus 0.14 sd in [7]).

Body mass index (BMI) is strongly positively correlated with SU levels in population-based studies (phenotypic correlations ranging from 0.27 to 0.44 [8–12]), and the relationship is approximately linear ([12] and S1 Fig.). Obesity is the strongest modifiable risk factor for hyperuricemia and gout [13]. We investigated here to what extent the genetic variants affecting SU are modulated by BMI. The fact that the genetic variants with the largest effect on SU levels are observed in genes encoding for ion transport proteins provides a biological rationale, since the activity of those transporters may be directly or indirectly affected by the metabolic changes associated with BMI increase, e.g. by levels of serum phosphate and hepatic ATP both reported to be inversely correlated with BMI [14,15]. Additionally, many of the newly discovered urate loci are in genes concerned with regulation of energy metabolism and glucose flux which are affected by BMI status. In 2008, a study had suggested that SLC2A9 variants’ effects on SU may be stronger in severely obese individuals (defined as BMI > 40), with a stronger modulating BMI effect in men than in women [9], while a recent publication suggests the opposite, in a predominantly women study [16]. Both these studies had modest sample sizes, calling for a larger study to be carried out.

Here, we performed a genome-wide investigation for genetic variants influencing serum urate levels in a BMI-dependent fashion, primarily by analysing genome-wide association study (GWAS) stratified by BMI. Stratified analyses are best suited when main effects are very different in magnitude or direction between strata and if the environment factor measured on a continuous scale is not acting linearly. In a discovery set, totalling 41,832 participants, GWAS for SU were performed after stratifying subjects by BMI status categorized into three levels: lean (BMI < 25 kg/m²), overweight (25 ≤ BMI ≤ 30) and obese (BMI > 30 kg/m²). This
through the Geestkracht programme of the Dutch Scientific Organization (ZON-MW, grant number 10-000-1002) and matching funds from participating universities and mental health care organizations. Genotyping in NESDA was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (I 480-05-003) along with a supplement from the Dutch Brain Foundation.

NSPHS– The Northern Swedish Population Health Study (NSPHS) was funded by the Swedish Medical Research Council (project number K2007-66X-20270-01-3), and the Foundation for Strategic Research (SSF). The NSPHS as part of EUROSPAN (European Special Populations Research Network) was also supported by European Commission FP6 STRP grant number 01947 (LSHG-CT-2006-01947). This work was also supported by the Swedish Society for Medical Research (Åsa Johansson). Ogliastra Genetic Park (OGP)–This work was supported by grants from the Italian Ministry of Education, University and Research (MIUR) no.5571/DSPAR/2002 and (FIRB) D. M. no. 718/Ric/2005. ORCADES–ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society, the Medical Research Council Human Genetics Unit and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). PROCARDIS–The Procardis consortium was funded by EC Sixth Framework Programme (LSHM-CT-2007-037273) and Astrazeneca AB. The authors acknowledge support from the Wellcome Trust core grant award 075491/Z/04. RS-I, RS-II–The GWA study was funded by the Netherlands Organisation of Scientific Research (NWO) inns (nr. 175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (O14-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Consortium for Healthy Aging (NCHA) project nr. 050-060-810. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. German MediGRID and Services@MediGRID part of the German D-Grid, provided access to their grid resources. SHIP–The Study of Health in Pomerania (SHIP) is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the the Geestkracht programme of the Dutch Scientific Organization (ZON-MW, grant number 10-000-1002) and matching funds from participating universities and mental health care organizations. Genotyping in NESDA was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (I 480-05-003) along with a supplement from the Dutch Brain Foundation.

**Materials and Methods**

**Study subjects**

The discovery BMI-stratified genome-wide association study meta-analyses (GWAMA) combined data from 22 population cohorts encompassing 42741 individuals with measured circulating urate levels and BMI. With six additional follow-up studies, all were studies of European descent participants that contributed to the Global Urate and Gout consortium (GUGC) and have thus been previously described in detail [2]. The study-specific descriptions are reported in S1 Table, in effect a subset of the GUGC publication.

Two extra studies, the Rotterdam study (described in S1 Table as also a GUGC participant) and a New-Zealand study of individuals from Polynesian descent [17] only contributed to the replication for the CLK4 locus. Sample sizes for the different sub-analyses performed and urate summary statistics for all studies with break down per BMI and gender stratum are detailed in S2 Table.

**Genotype collection**

Genome-wide SNP genotyping was undertaken by each cohort using various platforms as previously described [2] and reported in S3 Table. Imputation of allele dosage of SNPs typed in the HapMap CEU population was performed using either MACH or IMPUTE with parameters and pre-imputation filters specified in S3 Table.

**Statistical analysis**

**BMI-stratified main effect GWAMA**. Combined-gender and gender-separate association analyses were performed as described in Kolz et al. [7] within three body mass index (BMI) categories (nine sub-analyses performed in total): lean (BMI<25), overweight (25≤BMI≤30) and obese (BMI>30). Urate level (mg/dl) was adjusted for age, sex, and if required, ancestry principal components. Medications were not taken into account. Residuals were standardised using a z-score and used as response variable. Genome-wide association analyses were performed using imputed allele doses as predictor variable in linear models, and each study submitted regression summary statistics for meta-analysis. Studies with related individuals used a linear mixed model that additionally accounts for a polygenic random effect (e.g a score test mmscore [18] implemented in the GenABEL package [19]). Softwares used by the different studies to implement association testing are reported in S3 Table.

The results from all individual GWA scans were combined into a fixed-effects meta-analysis using inverse variance weighting, implemented in the MetABEL R package [15]. From individual-study analysis, SNPs with minor allele frequency less than 1% or low imputation quality (assessed by the metrics r2hat (MACH) <0.3 or. info (IMPUTE) <0.4) were excluded. The QQ plots for association statistics from each study were visualised in R. This highlighted that two many results from the PROCARDIS–women dataset departed from the null hypothesis distribution, and this subset was removed from the final meta-analysis as driving many significant results if non-excluded. Study-specific genomic control inflation factors are reported in S4 Table. In the meta-analyses, each individual study results were adjusted using the inflation factors; the overall meta-analysis effects’ standard errors and p-value reported were not
further corrected. The overall inflation factor for the nine stratified GWAMA were 1.0167 (lean-combined-gender), 1.0069 (lean-women), 1.0120 (lean-men), 1.0362 (overweight-combined-gender), 1.0167 (overweight-women), 1.0232 (overweight-men), 1.0157 (obese-combined-gender), 1.0180 (obese-women) and 1.0157 (obese-men). The conventional genome-wide significance threshold of 5x10^-8 was used. To avoid results driven by one or two populations that are likely to be spurious, meta-analysis results for the lower allele frequency variants (MAF <5%) are reported only if at least four populations contributed and if the contribution of any single study as calculated by the R package “meta” (http://cran.r-project.org/) was not greater than 30%. Annotation to known GWAS hits in the vicinity (window of 150 kb centred on index SNP) of novel potential urate loci was made using the NHGRI GWAS catalogue [20], 29-10-2013 update.

Main effect gene-based test. A gene-based test for SU association in the BMI-stratified GWAMA was conducted using the VEGAS software. Briefly, this method assigns SNPs to genes (+/- 50kb of 5’ and 3’ UTRs) and combines the association P-values accounting for linkage disequilibrium between markers assigned to the same gene. Analyses were conducted for each of the nine BMI/gender categories GWAMA results. As 17,787 genes are tested, the Bonferroni-corrected threshold for significance is set at 2.8 10^-6.

Replication of the differential effect of the CLK4 variant rs7711186 was sought in six independent studies of individuals of European descent, totalling 1259 individuals, in which the marker was either genotyped or well imputed and, as exploratory foray, in a small sample of individuals of Polynesian descent (N = 290) with prevalent obesity.

Testing for differences in main effect between BMI strata. The meta-analysed SNP main effects on SU were compared between all pairwise BMI categories (lean-obese, lean-overweight, and overweight-obese) using a t-test. Test statistics were calculated using the statistic \( t = (\beta_{bmicat1} - \beta_{bmicat2})/\sqrt{SE_{bmicat1}^2 + SE_{bmicat2}^2 - 2r(SE_{bmicat1}, SE_{bmicat2})} \), with \( \beta_{bmicat} \) and \( SE_{bmicat} \) the meta-analysed SNP effect-estimates and their corresponding standard errors, and \( r \) the Spearman rank correlation coefficient between meta-analyzed beta-estimates, in each of the BMI categories compared, across all SNPs. Under the null hypothesis of no difference in effect sizes between BMI strata, the t statistic is expected to follow a Student’s t distribution.

Interaction effect GWAMA

Discovery studies. Combined-gender and sex-stratified SNP by BMI interaction analyses were also performed in participating discovery studies using linear regression methods. Urate residuals were generated using the same covariates and standardisation as described for the stratified main effect GWAS. For studies with related individuals, relatedness was accounted for by fitting ancestry principal components (PCs) derived from the genomic relationship matrix rather than fitting it in full within a mixed model for the association test as the iterative processes used for parameter estimations of the mixed models often did not converge in a pilot run using family-based populations. The number of PCs to account for, varying from study to study and best determined by examination of scree plots (point to which additional PCs all contribute the same percentage of genetic variation), was left to the decision of each study analyst. Each study GWAS was performed on imputed genotype dose using the following model: \( z(\text{residual}) = \mu + \beta_1 \text{BMI} + \beta_2 \text{SNP} + \beta_1 \text{BMI} \times \text{SNP} + \epsilon \), with BMI as continuous variable, \( z(\text{residual}) \) the serum urate level adjusted for age, sex (in the combined gender analysis) and ancestry principal components expressed as z-score ((individual trait value minus population mean)/population standard deviation), \( \beta \) the regression coefficients for the fitted effects, \( \epsilon \) the error term with normal probability distribution.
Softwares used by the different studies to implement association testing are reported in S2 Table. Coefficients estimates for the main effect ($\beta_1$ and $\beta_2$) were not reported for studies that used Quicktest, as this later only reported the interaction term ($\beta_{12}$). Meta-analyses of the interaction effects ($\beta_{12}$) coefficients were carried out using MetABEL as described for the stratified main effect GWAS, with a higher MAF cut-off (5%) for each individual study. To avoid results predominantly driven by one population that are likely to be spurious, meta-analysis results with individual study contribution greater than 50% as calculated by the meta R package were filtered out.

As the individual studies genomic control inflation factors ($\lambda$) for these analyses were often high (S3 Table), only the studies with a $\lambda$ less than 1.2 were analysed and sensitivity analyses with a reduced set of studies with $\lambda$ less than 1.05 were also performed. The overall inflation factors for the GWAMA of interaction terms with the studies with a $\lambda$ less than 1.2 corrected using genomic control were 0.992 in the combined-gender, 1.011 in the women and 1.024 in the men analyses.

Follow-up set. A small number of studies were available for follow-up of the linear interaction analysis, totalling 9298 participants (INGI-Cilento, OGP Talana, NESDA, INCIFE, INGI-FVG and AGES). All follow-up studies analyses were carried out in the combined-gender data-set only and use the "model-robust method" that is implemented in the ProbABEL and Quicktest packages. Application of the model-robust method in principle leads to lower genomic control inflation for the interaction term [21]. To increase sample size in the follow-up, the CoLaus study (N = 5411) was added as a follow-up rather than discovery set for the regression based interaction term analysis. One study (INCIEPE-N = 940) had high $\lambda$ for both main and interaction effects (S3 Table), and was not included in the meta-analysis.

Meta-analyses of the interaction effects ($\beta_{12}$) coefficients were carried out using MetABEL as previously described for the discovery cohorts. The overall inflation coefficient for this follow-up meta-analysis was 1.018 and 1.006 for the combined discovery and follow-up studies interaction term meta-analysis.

Pathway Analysis

The pathway analysis was carried out using a SNP-based circular permutation method implemented in an extension of the R package "genomicper" (http://cran.r-project.org/) available upon request to the package’s authors. After lift over to build37, SNPs were annotated to genes when they were located within gene regions using annotations from the NCBI Gene database (http://www.ncbi.nlm.nih.gov/gene; build.37.1) and the SNPs (and associated GWAMA-p-values) were ordered according to their location in the genome. Pathways (n = 229) were downloaded using KEGG.db (http://www.genome.jp/kegg/) and the SNPs and genes assigned to the pathways. SNPs with GWAMA p-values less or equal to 5% were considered associated with trait and associated SNPs within a pathway counted. This count was compared to the distribution of counts obtained from 10,000 circular permutations of the SNPs’ GWAMA association p-values with respect to the SNPs locations. In circular genomic permutation the genome is considered circular and ordered from chromosome 1 to 22 and restarting at chromosome 1 [22]. Each permutation is akin to the spinning of a wheel with the whole starting set of SNP labels and locations fixed at the outside of the wheel and the SNPs’ GWAMA p-values on the rotating wheel. As the SNPs’ p-values rotate to the same degree, they retain patterns of correlation similar to those in the original data. The empirical p-value for the trait-pathway association was calculated from the ratio of the total number of permutations with more significant SNPs than the non-permuted set divided by the total number of permutations performed in the analysis [22].
Ethics Statements

Participants gave written informed consent to each original study. All studies received approval from their local ethics committees as listed. S1 Table and protocols comply with the tenets of the Declaration of Helsinki.

Results

BMI stratified urate GWAS

All 22 participating studies had previously contributed to non-stratified SU analyses [2] and study-specific information is reported in S1 and S3 Tables. All study participants were of European ancestry and displayed BMI distribution typical of that of populations that adopted westernized diet and culture, with more than half of the participants overweight or obese (Table 1). The smallest stratum analysed comprised 4,613 individuals (obese-men category), the largest 17,078 (overweight-all category). Individual study SU descriptive statistics are reported in S2 Table. The median population mean SU per stratum analysed was, as expected from the wealth of epidemiological data, higher in males than females and increasing from the lean to the obese group (Table 1).

The stratification process did not yield any novel genome-wide significant signal at the SNP level \( (P < 5 \times 10^{-8}) \) and all but three \((LRRC16, SLC16A9\) and \(RREB1)\) of the eleven loci reported in two earlier, non-stratified, SU genome-wide association meta-analyses (GWAMA) of size roughly comparable to the present analyses [7,23], reached genome-wide significance in at least one of the nine strata (Table 2). All other loci encompassing SNP variant(s) with an association P-value below the suggestive threshold of \(10^{-5}\) in any of the nine meta-analyses are listed in S5 Table. Three of these suggestive loci, \(AICF\) (lean-combined-gender), \(HLF\) (obese-combined-gender) and \(NRG4\) (obese-men) are among the 18 novel, validated and replicated, urate loci in a large recent SU GWAMA (N > 140,000 individuals, a subset of which is analysed here) [2]. No functional link with urate homeostasis is obvious from the genes within the other suggestive signals apart potentially for \(SLC28A1\) (lean-men category), encoding a sodium/nucleoside co-transporter present in kidney. \(MYO18D\) and \(ADAMST17\) (both in lean stratum signals) have been previously listed as suggestive loci for serum urate levels in a small study of African American participants [24].

The gene-based association test implemented in the statistical package VEGAS revealed one novel locus, \(CLK4\), reaching the gene-based genome-wide significance in the obese-men stratum only \( (P\text{-value} = 2 \times 10^{-6}, \text{just below the Bonferroni corrected gene-based threshold of} \ 2.8 \times 10^{-6})\). However, this effect was not reproduced in a replication set (S2 Fig.).

A complete list of top associated genes in the gene based analysis is reported in S6 Table down to the suggestive threshold for gene-based association of \(10^{-4}\). Most encompassed known

Table 1. Statistics for the discovery studies mean serum urate levels.

<table>
<thead>
<tr>
<th></th>
<th>Lean (BMI &lt; 25 kg/m²)</th>
<th>Overweight (BMI 25–30kg/m²)</th>
<th>Obese (BMI &gt; 30 kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median; Min-Max N</td>
<td>median; Min-Max N</td>
<td>median; Min-Max N</td>
</tr>
<tr>
<td>All</td>
<td>4.58; 4.18–5.71 14504</td>
<td>5.36; 4.75–6.19 17078</td>
<td>5.81; 4.98–6.64 9445</td>
</tr>
<tr>
<td>Men</td>
<td>5.5; 4.86–6.19 5529</td>
<td>5.9; 5.2–6.72 10058</td>
<td>6.53; 5.31–7.23 4613</td>
</tr>
<tr>
<td>Women</td>
<td>4.1; 3.73–4.85 9753</td>
<td>4.6; 4.14–5.41 7189</td>
<td>5.22; 4.63–6.19 4690</td>
</tr>
</tbody>
</table>

Median, Minimum and Maximum values for the mean serum urate (SU) concentrations (mg/dl) amongst the twenty two studies used in the BMI and gender stratified meta-analyses are displayed. N represents the total number of participants analysed in each category.

doi:10.1371/journal.pone.0119752.t001
urate loci, including two of the recently reported novel urate-associated loci [2]: A1CF, an essential component of the apolipoprotein B mRNA editing machinery, which is suggestive in the lean-combined sex stratum and MLXIPL, a carbohydrate-responsive element-binding protein, in the overweight-combined sex stratum.

**Effect size variation across BMI strata for genome-wide significant effects**

Some modulation of effect sizes depending on BMI status is suggested by close inspection of the most strongly associated SNPs in each stratum ([Table 2](#tab2)). For example, a GCKR SNP, rs780094, reached genome-wide significance in the obese-combined-gender stratum but no SNP within that locus reached even the suggestive threshold of association (10^−5) in the lean-combined-gender stratum despite the larger number of individuals in the latter.

We formally tested the differences in SU effect sizes across BMI strata pairwise for the variants that reached the genome-wide significance threshold in at least one BMI stratum in this study ([Table 2](#tab2)), discarding SLC2A9 and ABCG2 comparisons in the combined gender analysis as the proportion of male and female is not the same across BMI categories and the effect sizes of the variants are sex-sensitive. Taking a Bonferroni corrected significance threshold for the number of independent SNPs analysed in different settings (0.05/14/3 = 0.0012), only one locus, ABCG2, showed a statistically significant difference in effect size between obese and lean men ([Table 2](#tab2)) and the trend between BMI categories and effect on SU level seemed linear ([Fig. 1A](#fig1A)). The magnitude of the effect on urate for the ABCG2 index SNP was more than halved in the obese category compared to the lean category (effect of rs2231142 allelic substitution: 95% CI 0.257 to 0.389 in lean men versus 95% CI 0.069 to 0.213 in obese male) making the magnitude of effect in obese men similar to that seen in women (95% CI 0.125 to 0.221 in lean women). SNPs at three additional loci reached nominal significance ([Fig. 1B-D](#fig1B-D)).

**Effect size variation across BMI strata genome-wide**

The same tests were also done genome-wide to investigate potential BMI-sensitive SNPs of opposite effect between strata. QQ plots for those analyses ([S3 Fig.](#s3fig)) showed no evidence for an excess of false positive results (genomic inflation factors ranged from 1.004 to 1.016). The most significant effect-differences (P_{diff} < 10^{-5}) for all nine comparisons, after quality control for low frequency variants, are reported in [S7 Table](#s7table) together with results for the 28 urate loci known to date, none of which reaching a P_{diff} lower than 10^{-5}. The lowest P-values were from the lean-obese and lean-overweight comparisons, all in loci not previously associated with urate and displaying different direction of effects in the lean and obese/overweight strata ([Fig. 2](#fig2) and [S4 Fig.](#s4fig)). The variant rs1829975, intergenic in RBMSI-TANK, a region that has been associated with several obesity related traits [25,26,27], reached the genome-wide significance threshold (P_{diff} < 5 \times 10^{-8}) in the men lean-overweight contrast. The second most significant difference, P_{diff} = 9.13 \times 10^{-8}, was also in the men lean-overweight contrast for a variant 5’ of the gene TSPYL5, a gene coding the testis specific Y-encoded-like protein 5 that has been recently suggested to regulate estradiol produced by adipocytes [28]. The most significant loci for the lean-obese comparisons were intergenic ARL5B-PLXDC2 and LASS3 for the men (P_{diff} respectively, 1.1 \times 10^{-7} and 2.2 \times 10^{-7}) and RBFOX3 for women and combined gender (P_{diff} = 4 \times 10^{-7}) which had suggestive main effect in the obese women stratum.

**Interaction effect in linear regression models**

To see whether a simple linear modelling of the BMI by SNP interaction (see [methods](#methods)) would uncover the same loci as the stratified analysis, interaction term analyses in linear models were
Table 2. Loci significantly associated with serum urate within any BMI stratum analysed and mean effect sizes across strata.

<table>
<thead>
<tr>
<th>Locus</th>
<th>SNP</th>
<th>A1</th>
<th>A2</th>
<th>A1 A2</th>
<th>β_{lean}</th>
<th>s.e.</th>
<th>P</th>
<th>β_{overweight}</th>
<th>s.e.</th>
<th>P</th>
<th>β_{obese}</th>
<th>s.e.</th>
<th>P</th>
<th>P-value 2-sided test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined-gender</td>
<td>Lean (BMI &lt; 25 kg/m²)</td>
<td>N = 14504</td>
<td>N = 17078</td>
<td>N = 9445</td>
<td>lean</td>
<td>ov</td>
<td>ob</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC2A9</td>
<td>rs7680126</td>
<td>G</td>
<td>A</td>
<td>0.22</td>
<td>-0.341</td>
<td>0.014</td>
<td>2.36E-134</td>
<td>0.22</td>
<td>-0.298</td>
<td>0.018</td>
<td>6.38E-02</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>ABCG2</td>
<td>rs2231142</td>
<td>G</td>
<td>T</td>
<td>0.89</td>
<td>-0.086</td>
<td>0.012</td>
<td>2.98E-13</td>
<td>0.46</td>
<td>-0.086</td>
<td>0.011</td>
<td>3.79E-15</td>
<td>0.46</td>
<td>-0.051</td>
<td>0.015</td>
</tr>
<tr>
<td>SLC22A11</td>
<td>rs2078267</td>
<td>C</td>
<td>T</td>
<td>0.49</td>
<td>-0.064</td>
<td>0.012</td>
<td>3.16E-09</td>
<td>0.064</td>
<td>-0.064</td>
<td>0.012</td>
<td>3.16E-09</td>
<td>0.064</td>
<td>-0.064</td>
<td>0.012</td>
</tr>
<tr>
<td>ABCG2</td>
<td>rs2199936</td>
<td>G</td>
<td>A</td>
<td>0.90</td>
<td>-0.323</td>
<td>0.033</td>
<td>1.06E-22</td>
<td>0.89</td>
<td>-0.243</td>
<td>0.025</td>
<td>1.17E-22</td>
<td>0.89</td>
<td>-0.141</td>
<td>0.036</td>
</tr>
<tr>
<td>SLC22A12;NRXN2</td>
<td>rs10897518</td>
<td>C</td>
<td>T</td>
<td>0.32</td>
<td>-0.060</td>
<td>0.012</td>
<td>1.49E-06</td>
<td>0.32</td>
<td>0.089</td>
<td>0.012</td>
<td>9.21E-14</td>
<td>0.32</td>
<td>0.089</td>
<td>0.012</td>
</tr>
<tr>
<td>SLC17A1;SLC17A4</td>
<td>rs1165209</td>
<td>G</td>
<td>A</td>
<td>0.46</td>
<td>-0.085</td>
<td>0.012</td>
<td>2.98E-13</td>
<td>0.46</td>
<td>-0.086</td>
<td>0.011</td>
<td>3.79E-15</td>
<td>0.46</td>
<td>-0.051</td>
<td>0.015</td>
</tr>
<tr>
<td>SLC22A12;NRXN2</td>
<td>rs505802</td>
<td>C</td>
<td>T</td>
<td>0.31</td>
<td>0.065</td>
<td>0.021</td>
<td>1.95E-03</td>
<td>0.31</td>
<td>0.065</td>
<td>0.021</td>
<td>1.95E-03</td>
<td>0.31</td>
<td>0.065</td>
<td>0.021</td>
</tr>
<tr>
<td>GCKR</td>
<td>rs780094</td>
<td>C</td>
<td>T</td>
<td>0.59</td>
<td>-0.043</td>
<td>0.012</td>
<td>2.3E-04</td>
<td>0.59</td>
<td>-0.050</td>
<td>0.011</td>
<td>5.64E-06</td>
<td>0.60</td>
<td>-0.085</td>
<td>0.015</td>
</tr>
<tr>
<td>Women</td>
<td>N = 9753</td>
<td>N = 7189</td>
<td>N = 4690</td>
<td>lean</td>
<td>ov</td>
<td>ob</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC2A9</td>
<td>rs13129697</td>
<td>G</td>
<td>T</td>
<td>0.28</td>
<td>-0.403</td>
<td>0.015</td>
<td>7.57E-150</td>
<td>0.28</td>
<td>-0.413</td>
<td>0.018</td>
<td>2.66E-111</td>
<td>0.28</td>
<td>-0.376</td>
<td>0.023</td>
</tr>
<tr>
<td>ABCG2</td>
<td>rs2231142</td>
<td>G</td>
<td>T</td>
<td>0.89</td>
<td>-0.173</td>
<td>0.022</td>
<td>5.56E-13</td>
<td>0.89</td>
<td>-0.190</td>
<td>0.028</td>
<td>2.39E-11</td>
<td>0.89</td>
<td>-0.114</td>
<td>0.037</td>
</tr>
<tr>
<td>SLC17A1;SLC17A4</td>
<td>rs1165209</td>
<td>G</td>
<td>A</td>
<td>0.46</td>
<td>-0.096</td>
<td>0.014</td>
<td>1.20E-11</td>
<td>0.46</td>
<td>-0.087</td>
<td>0.017</td>
<td>1.89E-07</td>
<td>0.46</td>
<td>-0.049</td>
<td>0.021</td>
</tr>
<tr>
<td>Men</td>
<td>N = 5529</td>
<td>N = 1058</td>
<td>N = 4613</td>
<td>lean</td>
<td>ov</td>
<td>ob</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC2A9</td>
<td>rs16890979</td>
<td>G</td>
<td>T</td>
<td>0.75</td>
<td>0.254</td>
<td>0.022</td>
<td>1.15E-29</td>
<td>0.76</td>
<td>0.183</td>
<td>0.017</td>
<td>4.47E-26</td>
<td>0.76</td>
<td>0.176</td>
<td>0.026</td>
</tr>
<tr>
<td>ABCG2</td>
<td>rs10805346</td>
<td>C</td>
<td>T</td>
<td>0.44</td>
<td>-0.161</td>
<td>0.019</td>
<td>5.12E-17</td>
<td>0.43</td>
<td>-0.170</td>
<td>0.015</td>
<td>1.45E-31</td>
<td>0.44</td>
<td>-0.164</td>
<td>0.022</td>
</tr>
<tr>
<td>SLC2A9</td>
<td>rs2199936</td>
<td>G</td>
<td>A</td>
<td>0.90</td>
<td>-0.323</td>
<td>0.033</td>
<td>1.06E-22</td>
<td>0.89</td>
<td>-0.243</td>
<td>0.025</td>
<td>1.17E-22</td>
<td>0.89</td>
<td>-0.141</td>
<td>0.036</td>
</tr>
<tr>
<td>GCKR</td>
<td>rs780094</td>
<td>C</td>
<td>T</td>
<td>0.59</td>
<td>-0.043</td>
<td>0.012</td>
<td>2.3E-04</td>
<td>0.59</td>
<td>-0.050</td>
<td>0.011</td>
<td>5.64E-06</td>
<td>0.60</td>
<td>-0.085</td>
<td>0.015</td>
</tr>
</tbody>
</table>

aFor the same locus, the index SNP (i.e. with the lowest P-value) may vary across stratum and when not in high LD with each other (pairwise r^2 less than 0.5 using the SNAP proxy search tool HapMap2 rel22 data, http://www.broadinstitute.org/mpg/snap/) index SNPs are displayed separately.

A1, allele for which effect (β) is reported, A2, alternate allele, fqA1 weighted average effect-allele frequency across the combined discovery studies. Mean effect sizes (β) are inverse-variance weighted estimates; s.e. standard error of the effect estimate. Effect differences were tested using a 2-sided t test. NA: non applied as the proportion of male and female is not the same across BMI categories and the variants’ effect sizes sex-sensitive. P-value (P) reaching genome-wide significance threshold are indicated in bold.

Abbreviations ov and ob stand for overweight and obese respectively.

doi:10.1371/journal.pone.0119752.t002
conducted in a subset of the discovery studies. Only those with an inflation factor less than 1.2 were combined in a meta-analysis. Two common variants, one intergenic EROL1B-EDARADD and one in the RBFOX3 gene, displayed P-values just below the genome wide significance for a BMI SNP interaction in the combined-sex analysis (rs10802528 P_{inter} = 7.78 x 10^{-8} and rs898534 P_{inter} = 9 x 10^{-8}, Table 3 and full list of most significant results in S8 Table). SNPs at these loci also displayed suggestive interaction in the women-only analysis. The ERO1LB-EDARADD locus remained suggestive in a sensitivity analysis with only the combined sex studies with the lowest genomic inflation analysed (lambda < 1.05, S9 Table), while index SNP rs898534 in RBFOX3's P-value drops to 1.5 x 10^{-4}. For a fair comparison, the tests for difference of main effects between BMI strata presented in S7 Table were recalculated using the exact subset of studies for which BMI SNP term results were analysed (S9 Table) and led to similar conclusions. Noticeably, while the top loci in the lean versus obese comparisons come up as top loci in the linear fitting of an interaction term (S8 and S9 Tables), none of the loci ranking high in the lean versus overweight strata reached suggestive significance in the linear modelling despite the strongest P_{diff} P-values, suggesting a non-linear mode of action for those.

We attempted replication of the linear interaction seen in the combined-sex analysis in a replication set consisting of six studies. Model-robust estimates of effects’ standard errors were calculated to avoid inflated \( \lambda_{GC} \) statistic commonly seen when using classical regression approaches [21]. Top results for this replication set and the combined samples are reported in...
### A. RBMS1-TANK: rs1829975 (C); MAF=18%; Pdiff = 4.71 10^{-6}

<table>
<thead>
<tr>
<th>Study</th>
<th>Men-lean (N=5529)</th>
<th>β</th>
<th>95%−CI</th>
<th>Men-overweight (N=4613)</th>
<th>β</th>
<th>95%−CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CROATIA-Vis</td>
<td>0.33 [-0.19, 0.65]</td>
<td></td>
<td></td>
<td>0.02 [-0.28, 0.33]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORCADES</td>
<td>0.48 [0.11, 0.85]</td>
<td></td>
<td></td>
<td>-0.15 [-0.39, 0.08]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLAUS</td>
<td>0.15 [-0.04, 0.35]</td>
<td></td>
<td></td>
<td>-0.06 [-0.22, 0.10]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERF</td>
<td>-0.08 [-0.46, 0.36]</td>
<td></td>
<td></td>
<td>0.07 [-0.23, 0.37]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INGI_CARL</td>
<td>0.35 [-0.17, 0.87]</td>
<td></td>
<td></td>
<td>0.08 [-0.57, 0.72]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INGI_ValBorbera</td>
<td>0.23 [-0.05, 0.51]</td>
<td></td>
<td></td>
<td>-0.24 [-0.46, 0.00]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2K</td>
<td>-0.01 [-0.20, 0.17]</td>
<td></td>
<td></td>
<td>0.02 [-0.13, 0.18]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBC1996</td>
<td>0.32 [-0.24, 0.87]</td>
<td></td>
<td></td>
<td>-0.12 [-0.40, 0.16]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KORAF3</td>
<td>0.46 [-0.03, 0.96]</td>
<td></td>
<td></td>
<td>0.01 [-0.28, 0.31]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KORAF4</td>
<td>0.06 [-0.40, 0.53]</td>
<td></td>
<td></td>
<td>-0.14 [-0.42, 0.13]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CROATIA-Korcula</td>
<td>0.00 [-0.55, 0.54]</td>
<td></td>
<td></td>
<td>0.00 [-0.27, 0.27]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROCARDIS</td>
<td>0.09 [-0.06, 0.25]</td>
<td></td>
<td></td>
<td>-0.09 [-0.20, 0.03]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSPHS</td>
<td>-0.08 [-0.36, 0.20]</td>
<td></td>
<td></td>
<td>-0.13 [-0.41, 0.15]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHIP</td>
<td>0.17 [0.00, 0.34]</td>
<td></td>
<td></td>
<td>-0.05 [-0.19, 0.09]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MICROS</td>
<td>-0.02 [-0.25, 0.22]</td>
<td></td>
<td></td>
<td>-0.06 [-0.33, 0.22]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCCS</td>
<td>0.38 [0.09, 0.69]</td>
<td></td>
<td></td>
<td>-0.28 [-0.51, -0.04]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRIGHT</td>
<td>0.33 [-0.19, 0.85]</td>
<td></td>
<td></td>
<td>-0.12 [-0.35, 0.12]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIC</td>
<td>0.18 [0.05, 0.32]</td>
<td></td>
<td></td>
<td>-0.06 [-0.33, 0.21]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>InChianti</td>
<td>-0.31 [-0.62, 0.01]</td>
<td></td>
<td></td>
<td>0.00 [-0.37, 0.27]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLSA</td>
<td>-0.10 [-0.47, 0.27]</td>
<td></td>
<td></td>
<td>-0.52 [-0.87, -0.16]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGCUT</td>
<td>0.06 [-0.20, 0.33]</td>
<td></td>
<td></td>
<td>-0.16 [-0.46, 0.14]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### B. TSPYL5; rs16895559(C); MAF=5.3%; Pdiff = 9.13 10^{-8}

<table>
<thead>
<tr>
<th>Study</th>
<th>Men-lean (N=5529)</th>
<th>β</th>
<th>95%−CI</th>
<th>Men-overweight (N=4613)</th>
<th>β</th>
<th>95%−CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CROATIA-Vis</td>
<td>-0.36 [-0.90, 0.18]</td>
<td></td>
<td></td>
<td>0.42 [-0.02, 0.86]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORCADES</td>
<td>0.06 [-0.87, 0.98]</td>
<td></td>
<td></td>
<td>0.43 [0.03, 0.83]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLAUS</td>
<td>-0.20 [-0.40, 0.01]</td>
<td></td>
<td></td>
<td>0.07 [-0.12, 0.25]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERF</td>
<td>-0.38 [-1.08, 0.32]</td>
<td></td>
<td></td>
<td>0.05 [0.48, 0.57]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INGI_CARL</td>
<td>0.12 [-0.58, 0.84]</td>
<td></td>
<td></td>
<td>0.04 [-0.88, 0.97]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INGI_ValBorbera</td>
<td>-0.05 [-0.63, 0.53]</td>
<td></td>
<td></td>
<td>-0.23 [-0.92, 0.45]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2K</td>
<td>0.12 [-0.46, 0.21]</td>
<td></td>
<td></td>
<td>-0.23 [-0.45, -0.03]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBC1996</td>
<td>-0.18 [-1.01, 0.64]</td>
<td></td>
<td></td>
<td>0.19 [0.00, 0.38]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KORAF3</td>
<td>0.03 [-0.66, 0.60]</td>
<td></td>
<td></td>
<td>0.07 [-0.24, 0.36]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KORAF4</td>
<td>0.15 [-0.31, 0.62]</td>
<td></td>
<td></td>
<td>0.12 [-0.28, 0.52]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CROATIA-Korcula</td>
<td>-0.37 [-1.32, 0.58]</td>
<td></td>
<td></td>
<td>0.29 [0.07, 0.51]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROCARDIS</td>
<td>-0.30 [-0.54, -0.06]</td>
<td></td>
<td></td>
<td>-0.10 [-0.59, 0.38]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSPHS</td>
<td>-0.30 [-0.79, 0.19]</td>
<td></td>
<td></td>
<td>-0.03 [-0.21, 0.15]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHIP</td>
<td>0.16 [-0.37, 0.06]</td>
<td></td>
<td></td>
<td>-0.16 [-0.72, 0.39]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MICROS</td>
<td>-0.32 [-0.75, 0.10]</td>
<td></td>
<td></td>
<td>-0.06 [-0.56, 0.44]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCCS</td>
<td>0.16 [-0.29, 0.61]</td>
<td></td>
<td></td>
<td>0.10 [-0.24, 0.35]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRIGHT</td>
<td>-0.04 [-1.35, 0.98]</td>
<td></td>
<td></td>
<td>0.10 [-0.05, 0.24]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIC</td>
<td>-0.24 [-0.43, -0.05]</td>
<td></td>
<td></td>
<td>0.16 [-0.18, 0.50]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>InChianti</td>
<td>-0.27 [-0.60, 0.08]</td>
<td></td>
<td></td>
<td>0.20 [-0.46, 0.86]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLSA</td>
<td>0.17 [-0.50, 0.84]</td>
<td></td>
<td></td>
<td>0.09 [-0.29, 0.48]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGCUT</td>
<td>-0.04 [-0.52, 0.44]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig 2. Forest plots of effect sizes within BMI stratum for variants with the two most significant mean effect size differences between BMI stratum. A. RBMS1-TANK locus and B. TSPYL5 locus. The overall inverse—variance-weighted mean effect per BMI stratum is calculated assuming fixed effect across studies and represented by a lozenge, associated P-value displayed as P. Measure of heterogeneity between studies is reported (I-squared) with associated P-value for significance (p). Pdiff is the test of difference in mean-effect size P-value. For study abbreviations and references, see S1 Table.

doi:10.1371/journal.pone.0119752.g002
Both RBFOX3 and ERO1L-EDARADD SNPs showed consistent direction of interaction effect between discovery and follow-up sets and a low level of heterogeneity across studies and RBFOX3 index SNP reached genome-wide significance in the combined dataset (Table 3).

Pathway analysis

We used a recently developed pathway analysis method where pathway associations are tested following circular permutations of all the GWAS SNPs P-values [22] and compare enriched KEGG defined pathways in all nine strata. Results (S10 Table) did not uncover any pathway reaching the genome-wide significance threshold defined by a strict Bonferroni correction using 229 pathways and nine analyses (P = 2.43 x 10^{-5}) but this threshold is very conservative given that many pathways are interconnected or overlapping and the combined and sex separate analyses are not independent. The most significant pathways were the ribosome pathway (P = 3 x 10^{-4}) in overweight women, glycosaminoglycan degradation in obese men (P = 6 x 10^{-4}) and N-glycan biosynthesis in lean women (P = 6 x 10^{-4}).

N-glycan biosynthesis (KEGG pathway hsa00510- N = 43 genes) is particularly compelling as its ranking amongst associated pathways is stable through the variable sample sized analyses for the same BMI stratum: it ranks top in all the lean meta-analyses (rank = 1 in combined-gender and women, rank = 21 in men), while it is medium-ranked in all the overweight analyses (rank = 68 in combined-gender, rank = 69 in women and rank = 103 in men) and amongst the lowest ranks in all the obese strata (rank = 217 in combined-gender, rank = 223 women and rank = 218 men). The list of genes out of the 43 genes in this pathway with at least one SNP nominally significantly associated with urate levels (P < 0.05) in either BMI stratum in the combined-gender analyses are listed in S11 Table.

Discussion

No novel locus with a genome-wide significant main effect on SU was uncovered when performing GWAS within the three BMI strata investigated, suggesting that changes in BMI do not switch on a yet unknown major urate locus. However, many loci reached suggestive level of SU association in a BMI dependent fashion and/or displayed suggestive difference in main effects across BMI categories that may collectively account for a substantial amount of BMI-sensitive SU variation.

One weakness of this study is its relatively modest size. Gene by environment (GxE) detection requires a larger sample-size than that required for the detection of main effects of comparable magnitude [29] (a rule of thumb proposed for case control design is a four time larger study [30]). Data from over 200,000 individuals were required to confirm the attenuation of FTO obesity risk genotype by physical activity [31] with the reported interaction term.
significant, \( P_{\text{inter}} = 0.001 \), because only one candidate gene was tested. Few scans for GxE interaction have been performed genome-wide to date [32–37]. A stratification strategy was used to uncover novel women-specific genetic effects in waist-related phenotypes with strong statistical support using a very large dataset [36] and gave support for stronger effects of the known to date type 2 diabetes genetic risks variants in lean compared to obese individuals [35]. Other studies have reported modest (\( P_{\text{inter}} \) at best \( 10^{-4} \)) interaction effect after testing for a joint effect of the main SNP effect and interaction term with the significant results driven by the main SNP effect [32,34]. Joint effect meta-analysis (JMA) was implemented fairly recently [38] and best suited when both main and interaction effects are present.

Our study was additionally challenged by using a crude readout, BMI, where similar measures can reflect very different physiological status, e.g. high BMI could correspond to high visceral fat deposition as well as low visceral fat deposition but high muscle mass. It would certainly benefit from more specific measures of environmental exposures for example, of diet (fructose, fat content, alcohol intake) or amount of physical activity or of metabolic status of the subject.

Despite these limitations, this is the largest investigation of the interplay between genetic variants influencing urate and BMI status to date and it provides novel, biologically supported, hypotheses that warrant further investigations.

Of the known urate loci, there was weak statistical evidence for modulation of SLC2A9 variant effects by BMI and no support from previous reports [9,16] of a consistent BMI modulating effect. By contrast, statistically significant change was observed for ABCG2 in men, with a fan-shaped interaction pattern and diminution (by half) of the genetic variant effect size in obese compared to lean men on average. The ATP-binding cassette transporter ABCG2 has been established as a high capacity urate transporter, is expressed in renal proximal tubules, liver and intestines, and the hyperuricemia causal Q141K mutation has been shown to reduce urate transport rates [39]. Surrounding lipids, ATP concentrations, cholesterol and bile acids have been shown to modulate activity of ABCG2 in vitro [40]. Interestingly, BMI-dependent effects of Q141K on urate response to acute fructose exposure have been recently reported [41]. A stronger effect of GCKR variant in the obese strata was only suggestive but it is well supported by the equivalent doubling in the lowering effect reported for the GCKR pleiotropic rs780094 T allele for fasting insulin and glucose in high-BMI participants compared to low-BMI participants, supplementary Table 2 of [32]. It is also consistent with the finding that adjustment for triglyceride (TG) level as potential mediator/confounder attenuates GCKR rs780094 variant urate association [10].

RBFOX3 and EROL1B were the top loci showing interaction with BMI status using linear models (with RBFOX3 index SNP reaching genome-wide significance in the combined discovery and look-up studies GWAMA). Both loci displayed the strongest evidence of a significant difference in SNP main effect when the lean and obese stratified samples were compared (S7 Table), analyses in which no individual study showed a high inflation factor or high heterogeneity across studies, supporting genuine interaction with BMI and in a linear fashion. We noted that, in contrast to those, the two top ranking loci from the stratified analyses comparisons (both for men lean-overweight contrasts) were not significantly interacting with BMI when using a linear model of interaction, and would require replication using the same methodology to be confirmed. RBFOX3 is a neuronal nuclear marker expressed in the Arcuate nucleus in the hypothalamus where orexigenic and anorexigenic neurons reside. Its paralog, RBFOX1, has been proposed as an obesity gene [42]. RBFOX3 (aka HRNBP3) was also selected together with 38 other genes in a gene-centric joint test for significant association with HDL-Cholesterol levels in a dataset combining expression data and GWAS data from independent sources [43]. A metabolic outcome of RBFOX3 knockout in mice (international mouse phenotyping consortium) is decreased circulating alkaline phosphatase, human levels of which correlates with BMI [44] and metabolic syndrome [45], a component of which is
hyperuricemia. EROL1B encoding for the endoplasmic reticulum oxidoreductin 1LB catalyzes the formation of disulfite-bonds in the ER. It represents another good candidate for BMI interaction as it is responsive to the unfolded protein response, a signal triggered by ER stress, levels of which are elevated in state of over-nutrition [46]. ER stress response itself may induce inflammation [47] and has been correlated with increased levels of inflammation marker molecules CRP and IL6 which were both positively correlated with urate levels [48].

The “N-glycan biosynthesis” pathway acting to influence urate levels differentially in lean individuals compared to overweight or obese individuals is intriguing. One of the newly identified urate loci [2], B3GNT4, also acts in a complex capping reaction, of Type II Lactosamine for example, establishing a precedent for a link between glycosylation enzyme variation and urate levels. The glycolysis intermediate Fructose 6P is the main precursor of amino sugar, combining with glutamine to form glucosamine-6-phosphate. Dependence on glutamine for both purine and glycoaminoglycan biosynthesis as illustrated by the inhibition of either pathway by the glutamine analogue antagonist DON [49] also interconnects these pathways.

These links would be important to study further as glucosamine can be prescribed to patients with gout to reduce pain and inflammation but the possibility that it might influence the urate level has not been explored.

Significant changes in N-glycosylation profiles with BMI have been well documented [50,51,52]. Fitting with the urate-association results (S11 Table), core fucosylation (driven by FUT8) was noted to decrease with BMI [52] and transcript levels for the sialyltransferase gene ST6GALT2 to be highly stimulated by the pro-inflammatory cytokines IL6 and IL8 [53] that are potentially elevated in the systemic low-grade inflammation that characterises obesity [54]. It is possible that in obese individuals flux towards O-GlcNacylation rather than towards N-glycan biosynthesis is more prominent, possibly following ER stress. O-GlcNacylation has been proposed as a nutrient sensor activated by glucose availability and correlates with insulin resistance, a common hallmark of obesity [55].

Metabolic pathways are highly inter-connected and their dys-regulation underlies many diseases. Accounting for body mass index in analyses provides a tool to link pathways to both obesity and urate homeostasis.

Supporting Information

S1 Fig. Scatter plots of BMI and serum urate in men and women from two populations used in this study. A CROATIA-Vis and B.ORCADES. Residuals from a mixed linear model adjusting serum urate (SU) levels for age and accounting for relatedness are plotted against each other. As noted in [11] the linear fit is stronger amongst women. (TIF)

S2 Fig. Forest plots for rs7711186 CLK4 variant effect size in the male and female obese stratum in replication datasets together with those of a SLC2A9 variant as positive control. In the discovery dataset, rs7711186 (C allele) was suggestively associated with urate in the men-obese stratum, differentially (overall effect size = 0.21, se = 0.04). Look-up in a small Polynesian study (NZL-Poly) where obesity is prominent is added under the overall meta-analysis value for the replication studies, all of European ancestry (represented by lozenge). * For this Polynesian study only the SLC2A9 variant rs11942223, in LD (r² = 0.6) with variant rs13129697, was available and used in the figure. (TIF)

S3 Fig. QQ plots for difference in SU effect statistics in all nine comparisons performed: lean versus overweight, lean versus obese and overweight versus obese in combined-gender
(ALL) or sex-stratified (MEN, WOMEN) samples. The ordered observed squared t statistic are plotted against the ordered expected statistics of the null, chi2, distribution, where 
\[ t = \left( \frac{\beta_{\text{bmicat1}} \cdot \beta_{\text{bmicat2}}}{\sqrt{\text{SE}_{\text{bmicat1}}^2 + \text{SE}_{\text{bmicat2}}^2}} \right) \cdot 2r(\text{SE}_{\text{bmicat1}}, \text{SE}_{\text{bmicat2}}), \]
with \( \beta_{\text{bmicat}} \) and SE bmi-cat the meta-analysis weighted beta-estimates and their corresponding standard errors and r the Spearman rank correlation coefficient between meta-analyzed beta-estimates in the BMI categories compared across all SNPs. Inflation coefficients, \( \lambda_{GC} \), are reported for each plot in the left upper corner.

S4 Fig. Forest plots of effect sizes within BMI stratum for variants showing the most significant mean effect size differences (associated P-value, \( P_{\text{diff}} \)) between BMI stratum genome-wide, in the combined-gender (all) strata. The overall inverse—variance-weighted mean effect per BMI stratum is calculated assuming fixed effect across studies and represented by a loz-enge, associated P-value displayed as P. Measure of heterogeneity between studies is reported (I-squared) with associated P-value for significance (p). For study abbreviations and references, see S1 Table.

S1 Table. Study description for each study site.

S2 Table. Individual study summary statistics for serum urate levels (SU) within the nine BMI/gender categories analysed. SU unit is in mg/dl, sd stands for standard deviation, N is the number of individuals with BMI and SU measures.

S3 Table. Study-specific genotyping, imputation information and analysis softwares.

S4 Table. List of inflation factors (\( \lambda \)) for each sub-analysis at individual study level. Inflation factors were calculated after filtering out poorly imputed variants and low frequency variants (MAF < 1% for main effect analysis in BMI-stratified GWAS (\( \lambda^* \)), MAF < 5% for SNP*BMI interaction term analysis (\( \lambda^{**} \)). NA flags analysis not performed. +++ indicates that model-ro-bust regression method was used.

S5 Table. List of loci encompassing SNP(s) with SU association suggestive P-value (5 x 10^{-8} = \( P < 10^{-5} \)) in the nine stratified GWAMA performed. Only the information pertaining to the SNP with the lowest P-value (index SNP) is listed. Lower allele frequency variants (1%<MAF <5%) are reported if the meta-analysis included at least four populations and if the contribution of any single study, as calculated by the meta R package, was lower than 30%. A1, allele for which effect (\( \beta \)) is reported; A2 alternate allele, frq(A1) weighted average effect-allele frequency across studies. Associations reported in the vicinity of the urate index SNP (in a 150kb region centred on the SNP) in the NHGRI GWAS catalogue (29_10_2013 update) are listed; highlighted red, the ones with same index SNP or index SNP in high to moderate linkage disequilibrium (r2 >0.4).

S6 Table. List of significant and suggestive loci (P-value < 10^{-4}) from the nine BMI strati-fied GWAMA in the gene-based association test implemented in VEGAS. Novel loci are shaded in grey. In bold, gene reaching genome-wide significant association with serum urate levels (\( P < 2.10^{-6} \)).
S7 Table. List of loci with SNP(s) displaying the strongest evidence of SU mean effect size difference across BMI strata in the discovery studies. Effect differences were tested using a t test. All loci with SNP displaying a $P_{\text{diff}} < 10^{-5}$ are listed with representative index SNP of lowest P-value (the total number of SNPs with suggestive P-value per loci is listed in N suggestivSNPs column). Low MAF SNPs were filtered as in S5 Table. Additionally, Pdiff values for the 28 known urate loci [2] are listed with index SNP from the published data. Locus in bold indicates that the difference in effect size between BMI strata reached genome-wide significance. P-value in bold for the known urate loci are those reaching the nominal threshold of 0.05. Locus with asterix had index SNP with main effect reaching suggestive level of association ($P < 10^{-5}$) in the BMI stratified urate GWAMA analysis (S5 Table).

(XLS)

S8 Table. List of loci with suggestive ($P_{\text{inter}} < 10^{-5}$) SNPxBMI interaction term using regression based method. Studies with inflation factor greater than 1.2 were not included in the analysis. For the combined-gender analysis, the CoLaus study was analysed as a replication study to balance discovery and replication sets. SNP with low MAF (< 5%) were excluded prior to meta-analysis. Results for the discovery, replication and combined sets are presented. Locus in bold indicates a genome-wide significant interaction effect. Shaded are loci common with S9 Table (list of loci with suggestive difference in urate main effects between BMI stratified GWAMA)

(XLS)

S9 Table. Results obtained as in S7 Table when analysis is restricted to the subset of studies (N = 16) used for BMI by SNP interaction testing using a regression-based method and with markers of MAF > 5% for direct comparison. Shaded are loci displaying suggestive association in linear interaction model(listed in S8 Table).

(XLS)

S10 Table. Results from the Pathway analysis tool implemented in the genomicper R package in the nine stratified urate GWAMA performed.

(XLS)

S11 Table. List of genes in the KEGGs N-glycan biosynthesis pathway, hsa00510, harbouring at least one SNP with a serum urate GWAMA P-value (P) nominally significant in one of the three combined-gender BMI categories analysed. N-glycan biosynthesis step coded 1 = N-glycan lipid-linked oligosaccharide precursor synthesis 2 = high mannose oligosaccharide to an Asparagine residue transfer and N-glycan trimming and branching 3 = more elaborate capping reactions

(XLS)

Acknowledgments

Each study thanks the general practitioners, clinicians, team members and participants for their invaluable contributions. Specific additional acknowledgments are listed below.

BRIGHT: The BRIGHT study is extremely grateful to the BRIGHT nursing team and we would also to thank Charles Mein, Richard Dobson, Chris Wallace and the Barts and The London Genome Centre staff. This is part of the translational studies supported by Barts and the London and the Leicester National Institute for Health Research Cardiovascular Biomedical Research Units. This work forms part of the CHAPTER e-health centre research program.

CoLaus: The authors thank Dawn Waterworth and Vincent Mooser, Co-Pis of the CoLaus study. Special thanks to Sven Bergmann, Jacqui Beckmann, Yolande Barreau, Mathieu
Firmann, Vladimir Mayor, Anne-Lise Bastian, Binasa Ramic, Martine Moranville, Martine Baumer, Marcy Sagette, Jeanne Ecoffey and Sylvie Mermoud for data collection.

**CROATIA-Korcula:** We acknowledge Peter Lichtner and the Helmholtz Zentrum München genotyping staff (Munich, Germany) for the SNP genotyping.

**CROATIA-Vis:** We acknowledge Prof. Pavao Rudan, Dr. Nina Smolej-Narancic, Dr. Branka Janicijevic and their colleagues at the Institute for Anthropological Research 10000 Zagreb, Croatia, for participants’ recruitment and field work, the Wellcome Trust Clinical Research Facility (Edinburgh) for performing the initial SNP genotyping.

**ERF:** We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection.

**EGCUT:** We acknowledge EGCUT personnel, especially Mr. V. Soo. Data analyzes were carried out in part in the High Performance Computing Center of University of Tartu.

**INGI_CILENTO:** We thank the populations of Cilento for their participation in the study.

**INGI-Carlabanto and INGI-FVG:** We thank Laura Esposito and Angela D’Eustacchio for technical support. We are very grateful to the municipal administrators for their collaboration on the project and for logistic support. We would like to thank all participants to this study.

**INGI-Val Borbera:** We acknowledge the inhabitants of the Val Borbera who accepted to participate in the study and the collaboration of the local administrators and of ASL-A1, Novi Ligure (A1). We also acknowledge the support of Prof. Clara Camaschella and Federico Caligaris-Cappio (U. Vita e Salute-HSR) for clinical data collection and of Fiammetta Viganò for technical help.

**MICROS:** For the MICROS study, we thank the primary care practitioners Raffaela Stocker, Stefan Waldner, Toni Pizzecco, Josef Plangger, Ugo Marcadent, and the personnel of the Hospital of Silandro (Department of Laboratory Medicine) for their participation and collaboration in the research project.

**NSPHS:** The authors are grateful for the contribution of district nurse Svea Hennix for data collection and Inger Jonasson for logistics and coordination of the health survey. Finally, the authors thank all the community participants for their interest and willingness to contribute to the study.

**Ogliastra Genetic Park (OGP):** We thank the Ogliastra population and all the individuals who participated in this study. We are very grateful to the municipal administrators for their collaboration to the project and for economic and logistic support.

**ORCADES:** DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

**RS-I, RS-II:** We thank Pascal Arp, Mila Jhamai, Dr Michael Moorhouse, Marijn Verkerk, and Sander Bervoets for their help in creating the GWAS database. The authors are very grateful to the participants and staff from the Rotterdam Study, the participating general practitioners and the pharmacists. We would like to thank Dr. Tobias A. Knoch, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid.

**SHIP:** The SHIP authors are grateful to Mario Stanke for the opportunity to use his Server Cluster for SNP Imputation.

**TwinsUK:** SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR.
NZPoly: We would like to thank Marilyn Merriman, Ruth Topless, Gabrielle Sexton, Jill James and Roddi Laurence for assistance in recruitment and laboratory work.

**Author Contributions**

Conceived and designed the experiments: JEH VV. Performed the experiments: JH EA ATeu MM KK TJ ZK NP GP LML TH PS AGo ML TT ADehghan DR GM AVS IMN LP APG LB TE JFP VV. Analyzed the data: JH EA ATeu MM KK TJ ZK NP GP LML TH PS AGo ML TT ADehghan DR GM AVS IMN LP APG LB TE JFP VV. Contributed reagents/materials/analysis tools: PN AsJ AAH OP SEH FM SHW ATen ATin EM AGr GKC JC PD'A SU PV GW SC IK KF MV JM CM ET CB RS ADöring ER KS AHo AGU MW H-EW GD AJG ND LS JHS MK RN MN CIS KB SMF ET Anj VS CSnS ChE MbBu RM NK SK SS SR SC NS GH TN PBM CHa AK CaH CC. Wrote the paper: JH EA ATeu MM KK TJ ZK NP GP LML TH PS AGo ML TT ADehghan DR GM AVS IMN LP APG LB PN AsJ AAH OP TE JFP SEH FM SHW ATen ATeu EM AGr GKC JC PD'A SU PV GW SC IK KF MV JEM CM ET CB RS ADöring ER KS AHo AGU MW H-EW GD AJG ND LS JHS MK RN MN CIS KB SMF ET Anj VS CSnS ChE MbBu RM NK SK SS SR SC NS GH TN PBM CHa OHF TRM VG Mi Pi BWP HS AM MCI PPP CMvD LF GG IJD MGD JFW PG UG TDS AFw CaH HW MPe Mo WHLK Mca DT HV CG AK VV. Design and/or management of individual study: NH HC IR OHF TRM VG Mi Pi BWP HS AM MCI PPP CMvD LF GG IJD MGD JFW PG UG TDS AFw CaH HW MPe Mo WHLK Mca DT HV CG.

**References**


