

Fasting glucose and fasting insulin sex-specific and sex-dimorphic GWAS meta-analyses
public data release -UPDATED March 2020 - README.PDF

“Sex-dimorphic genetic effects and novel loci for fasting glucose and insulin variability”

Vasiliki Lagou, Reedik Mägi, Jouke-Jan Hottenga *et al.* *Nature Communications* (2021)

These files contain the genome-wide association study (GWAS) meta-analysis results (HapMap II CEU and 1000 Genomes density) for the study presented in the paper above [1]. A set of GWAS meta-analyses was performed for fasting plasma glucose (FG) levels measured in mmol/L, and fasting insulin (FI) concentrations measured in pmol/L. FG and FI were analyzed at the study-level for men and women separately without adjustment for body-mass index. Measures of FG made in whole blood were corrected to plasma level using the correction factor of 1.13 [2, 3]. FI was measured in serum. Similar to previous MAGIC efforts [2, 3], individuals were excluded from the analysis if they had a physician diagnosis of diabetes, were on diabetes treatment (oral or insulin), or had a fasting plasma glucose equal to or greater than 7 mmol/L. Individual studies applied further sample exclusions, including pregnancy, non-fasting individuals and type 1 diabetes. Untransformed FG and natural logarithm transformed FI were analyzed at a study level.

The data represents FG/FI genome-wide association results in i) up to 80,512 individuals from 38 studies genotyped using either Illumina or Affymetrix genome-wide SNP arrays; ii) up to 47,150 individuals from 27 studies genotyped using the iSELECT Metabochip array (~197K SNPs) designed to support efficient large-scale follow-up of putative associations for glycemic and other metabolic and cardiovascular traits; iii) up to 21,173 individuals from 8 studies genotyped for custom variant sets and iv) 4 studies, including up to 13,613 individuals from four family-based studies (sex-combined meta-analyses only).

All participants were of European ancestry, without diabetes and mostly adults, although data from a total of 8,222 adolescents were also included in the meta-analyses (ALSPAC, French Young controls/obese, Leipzig-childhood and NFBC86 studies). Studies with genome-wide arrays undertook imputation of missing genotypes using the HapMap II CEU reference panel.

Each study performed single SNP association for men and women separately (sex-specific). The additive genetic effect of each SNP was estimated using a linear regression model adjusting for age (if applicable), study site (if applicable) and principal components. In case-control studies, the cases and controls were analysed separately. Individual study results were corrected for residual inflation of the test statistics using genomic control (GC) lambda estimates [4]. The GC lambda values were estimated using test statistics from all SNPs for the GWAS. In Metabochip studies, GC values were estimated from test statistics from 5,041 SNPs selected for follow-up of QT-interval associations, as we perceived these to have the lowest likelihood of common architecture of associations with glycemic traits [4]. SNP effect estimates and their standard errors were combined by a fixed effect model with inverse variance weighting using the GWAMA v2.2.3 software [5] within the following two meta-analysis strategies: i) sex-specific, where allelic effect estimates were combined separately within each sex (male-specific or female-specific); ii) sex-dimorphic, where male- and female-specific estimates were combined by allowing for heterogeneity in allelic effects between women and men (chi-squared distribution with two-degrees of freedom) and iii) sex-

combined, where allelic effect estimates from men and women were combined. Studies with highly related individuals (Dundee, FamHS, FHS and Sardinia) were included only in the sex-combined meta-analysis (men and women were analysed together at a study-level and an additional adjustment for sex was made). Studies with highly related individuals (Dundee, FamHS, FHS and Sardinia) were included only in the sex-combined meta-analysis (men and women were analysed together at a study-level and an additional adjustment for sex was made). Additionally, the heterogeneity of allelic effects between sexes was assessed using Cochran's Q-test. Cochran's statistic provides a test of heterogeneity of allelic effects at the j^{th} SNP, and has an approximate chi-squared distribution with N_{j-1} degrees of freedom under the null hypothesis of consistency where N_j denotes the number of studies for which an allelic effect is reported. Both the sex-dimorphic meta-analysis framework and Cochran's Q-test for heterogeneity have been implemented in the GWAMA software. Summary statistics for FG and FI (combined and sex-stratified) were imputed to all-ancestries 1000G reference panel using the SS-imp software v0.5 for summary statistics imputation [6].

Summary statistics are provided for up to 67,506 men / 73,089 women (FG meta-analysis) and 47,806 men / 50,404 women (FI meta-analysis).

HapMap density files

FG_STAGE1_2_3_SEX_GWAS_updated_2019.txt.gz

FI_STAGE1_2_3_SEX_GWAS_updated_2019.txt.gz

For each SNP, we have provided the following information:

- a. snp: rsID of the DNA polymorphism, all variants are aligned to the forward strand of human genome (HapMap II);
- b. effect_allele: allele, for which the genetic effect estimate is provided;
- c. other_allele: other allele of the SNP;
- d. eaf_hapmap_CEU: effect allele frequency reported for the CEU HapMap II individuals;
- e. male_beta: DNA variant effect estimate on FG or FI levels evaluated in males;
- f. male_se: standard error of the effect estimate on FG or FI levels evaluated in males;
- g. male_pvalue: statistical significance of the effect estimate on FG or FI levels evaluated in males;
- h. female_beta: DNA variant effect estimate on FG or FI levels evaluated in females;
- i. female_se: standard error of the effect estimate on FG or FI levels evaluated in females;
- j. female_pvalue: statistical significance of the effect estimate on FG or FI levels evaluated in males;
- k. sex_diff_pvalue: statistical significance of the 2 degrees of freedom test for sex-dimorphism in allelic effects;
- l. sex_het_pvalue: statistical significance of the test for heterogeneity in genetic effects between males and females.

1000Genomes density files

For each SNP, we have provided the following information:

- a. z – Z-score of association for FG or FI;
- b. source – SSIMP for imputed and GWAS for SNPs present in the HapMap density;

- c. rsid – rsID;
- d. a1 – reference allele;
- e. a2 – effect allele;
- f. r2.pred – SS-imp imputation quality;
- g. p-value – P-value of association;
- h. n – sample size;
- i. maf – minor allele frequency;
- j. beta – effect size for FG or FI;
- k. se - standard error of the effect estimate for FG or FI.

When using data from the downloadable meta-analysis results please acknowledge the source of the data as follows: “Data on glycaemic traits have been contributed by MAGIC investigators and have been downloaded from www.magicinvestigators.org” citing the paper.

References

1. Lagou, V. *et al.* Sex-dimorphic genetic effects and novel loci for fasting glucose and insulin variability. *Nat Commun* **12**, 24 (2021).
2. Prokopenko, I. *et al.* Variants in MTNR1B influence fasting glucose levels. *Nat Genet* **41**, 77-81 (2009).
3. Scott, R.A. *et al.* Large-scale association analyses identify new loci influencing glycaemic traits and provide insight into the underlying biological pathways. *Nat Genet* **44**, 991-1005 (2012).
4. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-1004 (1999).
5. Magi, R., Lindgren, C.M. & Morris, A.P. Meta-analysis of sex-specific genome-wide association studies. *Genet Epidemiol* **34**, 846-53 (2010).
6. Rueger, S., McDaid, A. & Kutalik, Z. Evaluation and application of summary statistic imputation to discover new height-associated loci. *PLoS Genet* **14**, e1007371 (2018).