**README Gene mapping and pathway analysis in FUMA**

To identify shared genes and biological pathways, we used publicly available GWAS summary statistics for body mass index (BMI)(1), WHR adjusted for BMI(2) (WHRadjBMI), and Alzheimer’s disease(3) (AD), and mapped SNPs to genes, and genes to biological pathways in FUMA(4).

1. **GWAS summary statistics:**

**BMI and WHR from the GIANT consortium:** <https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files>

2018 GIANT and UK BioBank Meta-analysis

BMI and Height GIANT and UK BioBank Meta-analysis Summary Statistics: Updated Meta-analysis Locke et al + UKBiobank 2018 GZIP

WHR GIANT and UK BioBank Meta-analysis Summary Statistics: whradjbmi.giant-ukbb.meta-analysis.combined.23May2018.HapMap2\_only.txt.gz

**AD from the IGAP consortium:**

<https://www.niagads.org/datasets/ng00075>

1. **Gene mapping and shared genes**

**SNP2GENE function**

The SNP2GENE function in FUMA(4) uses summary statistics from genome-wide association studies (GWASs) to map SNPs to functional genes. Using the default settings, the application first identified independent (r2 cut-off 0.6, based on European ancestry samples in the 1000G phase 3(5) reference panel) lead SNPs, i.e. with a GWAS significance of p<5x10-8, and candidate SNPs in linkage disequilibrium (r2 cut-off 0.1, based on European ancestry samples in the 1000G phase 3(5) reference panel) with the independent lead SNPs with a GWAS significance of p<0.05. The lead and candidate SNPs were then mapped to protein coding genes (excluding the MHC region), based on positional, eQTL, and chromatin interaction mapping. The positional mapping retrieves genes with a physical distance of ≤10kb of the relevant SNPs. The eQTL mapping was set to map SNPs to genes likely influencing expression levels up to 1Mb (cis-eQTLs) based on GTEx v8 data and all tissue types, using significant (false discovery rate (FDR) p<0.05) SNP-gene pairs. Chromatin interaction mapping using built in Hi-C data of 21 tissue/cell types from GSE87112, using a FDR threshold of p<10-6. Lead and candidate SNPs are mapped to any genes whose promoter region (defined as 250 base pairs upstream or 50 base pairs downstream of the transcription start site) in the interacting region.

The SNP2GENE function was carried out separately on GWAS summary statistics for AD, BMI, and WHRadjBMI.

Settings

Setting specifications are provided in the Settings folder: SNP2GENE\_settings\_AD\_FUMA\_job77927.txt

SNP2GENE\_settings\_BMI\_FUMA\_job77883.txt SNP2GENE\_settings\_WHRadjBMI\_FUMA\_job129712.txt

Output

SNPs involved in AD were mapped to 295 genes, BMI to 8663 genes, and WHRadjBMI to 5589 genes, available as public results at the FUMA homepage, under the following links and accession numbers:

* AD: ID 220, title: AD\_Kunkle\_2019, link: <https://fuma.ctglab.nl/browse/220>
* BMI: ID 221, title: BMI\_Yengo\_2018, link: <https://fuma.ctglab.nl/browse/221>
* WHRadjBMI: ID 222, title: WHRadjBMI\_Pulit\_2019, link: <https://fuma.ctglab.nl/browse/222>

The gene lists are also available in the output folder: genes\_AD.txt; genes\_BMI.txt; genes\_WHR.txt

**Shared genes**

Output lists of mapped genes from FUMA are processed in SAS to obtain lists of overlapping genes between BMI and AD or WHRadjBMI and AD.

SAS code: Programs\ Forte\_aim2\_genetic\_overlap\_20200428.sas and Programs\ Forte\_aim2\_genetic\_overlap\_WHR\_20200524.sas

This resulted in 175 genes shared between AD and BMI, and 65 genes between AD and WHRadjBMI.

Output

The resulting lists of shared genes are available in the output folder: BMI\_AD\_mapped\_genes.txt; WHR\_AD\_mapped\_genes.txt

1. **Pathway analysis**

**GENE2FUNCTION**

The lists of shared genes (Output\ BMI\_AD\_mapped\_genes.txt and Output\ WHR\_AD\_mapped\_genes.txt) were used in pathway analyses, using the GENE2FUNCTION application in FUMA with default settings. Protein coding genes were selected as background (n= 20,260 genes). In GENE2FUNCTION, hypergeometric tests are used to test if the selected genes (here those in common to AD and BMI or to AD and WHRadjBMI) are overrepresented in pre-defined gene sets, after multiple testing corrections.

Settings

Setting specifications are provided in the Settings folder:

Settings\_BMI\_AD\_mapped\_FUMA\_gene2func38627.txt Settings\_WHR\_AD\_mapped\_FUMA\_gene2func59301.txt

Output

We visually inspected the output plots for GO biological processes in MsigDB c5 to identify potentially mediating biological pathways. The results highlighted inflammation and lipid metabolism as common pathways for AD and BMI, and lipid metabolism for AD and WHRadjBMI.

The plots are available in the output folder: BMI\_AD\_GO\_bp\_FUMA\_gene2func38627.pdf; WHRadjBMI\_AD\_GO\_bp\_FUMA\_gene2func59301.pdf

1. **References**

1. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ∼700000 individuals of European ancestry. Hum Mol Genet. 2018.

2. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet. 2019;28(1):166-74.

3. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nat Genet. 2019;51(3):414-30.

4. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826.

5. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.