Supplemental Material

Examination of the Shared Genetic Basis of Anorexia Nervosa and Obsessive-Compulsive Disorder

Yilmaz *et al.*

Table of Contents

Supplemental Text

Study Samples

AN. All AN data and corresponding controls included in our analyses came from Freeze 1 of the Eating Disorders Working Group of the Psychiatric Genomics Consortium (PGC-ED). PGC-ED comprises two independent cohorts: Genetic Consortium for Anorexia Nervosa (GCAN) and Children's Hospital of Philadelphia/Price Foundation Collaborative Group (CHOP/PFCG). GCAN, founded in 2007, is an international collaboration that represents researchers and clinicians from 16 countries. Detailed description of the samples and methods can be found elsewhere.[1-3] The CHOP/PFCG dataset comprises 1,033 AN cases collected as a part of the Price Foundation International Consortium and 3,733 pediatric controls from CHOP.

OCD. OCD data and corresponding controls came from the TS/OCD Working Group of the PGC (PGC-TS/OCD). The OCD component of the PGC-TS/OCD comprises two OCD consortia. The first, International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC), is an international collaboration of researchers and clinicians from 21 countries in North, Central and South America, Europe, the United Arab Emirates, and South Africa. The second, OCD Genetics Association Study (OCGAS), is a collaboration of clinicians and researchers from six US institutions (Johns Hopkins University, Columbia, Brown, Massachusetts General Hospital, NIMH and UCLA). Samples from IOCDF consisted of cases and ancestry matched controls, while OCGAS included primarily trios, with additional family members, singleton OCD cases and matched controls also included. Detailed descriptions of the samples and methods can be found elsewhere.[4-6] Data from these two consortia were meta-analyzed after duplicate and related samples were removed.

Methods

We first identified all variants that were found in both datasets with identical genomic position, allele 1 and allele 2 listings (a total of 7,461,827). We then took the dataset with fewer variants (in this case, OCD) and for each of the 245,479 variants where we were unable to find an exact variant match in the AN dataset, we flipped allele 1 and 2 and then identified the subset of the 2,124,606 unresolved variants from the AN dataset that have perfect variant with the allele-flipped variant. If doing this resolved any additional variants, the alleles and minor allele frequencies in the OCD dataset for the corresponding variant was flipped to match AN. We did this as opposed to changing OCD to match AN genotypes since the imputation done in AN was done via 1000 genomes phase 3, while for OCD it was done in 1000 genomes phase 1. In all we identified a total of 7,461,827 variants which were in both separate datasets and thus suitable for meta-analysis.

References

1. Boraska V, Franklin CS, Floyd JA, Thornton LM, Huckins LM, Southam L *et al.* A genome-wide association study of anorexia nervosa. *Mol Psychiatry* 2014; **19**(10)**:** 1085-1094.

2. Wang K, Zhang H, Bloss CS, Duvvuri V, Kaye W, Schork NJ *et al.* A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Mol Psychiatry* 2011; **16**(9)**:** 949-959.

3. Duncan L, Yilmaz Z, Gaspar H, Walters R, Goldstein J, Anttila V *et al.* Significant Locus and Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa. *Am J Psychiatry* 2017; **174**(9)**:** 850-858.

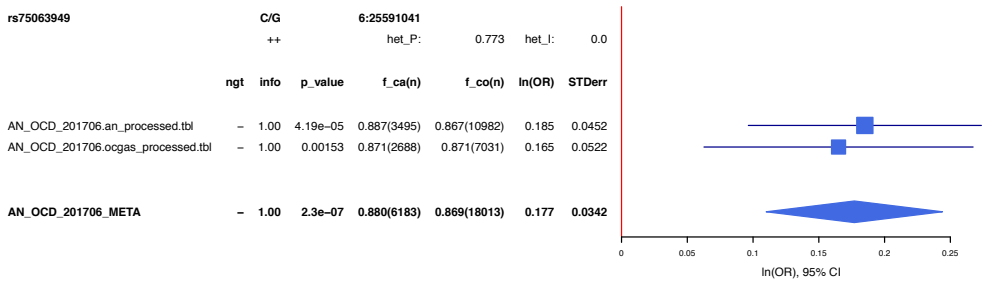
4. Stewart SE, Yu D, Scharf JM, Neale BM, Fagerness JA, Mathews CA *et al.* Genome-wide association study of obsessive-compulsive disorder. *Mol Psychiatry* 2013; **18**(7)**:** 788-798.

5. Mattheisen M, Samuels JF, Wang Y, Greenberg BD, Fyer AJ, McCracken JT *et al.* Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. *Mol Psychiatry* 2014.

6. International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OCGAS). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol Psychiatry* 2017.

Supplemental Figures

(a)



(b)

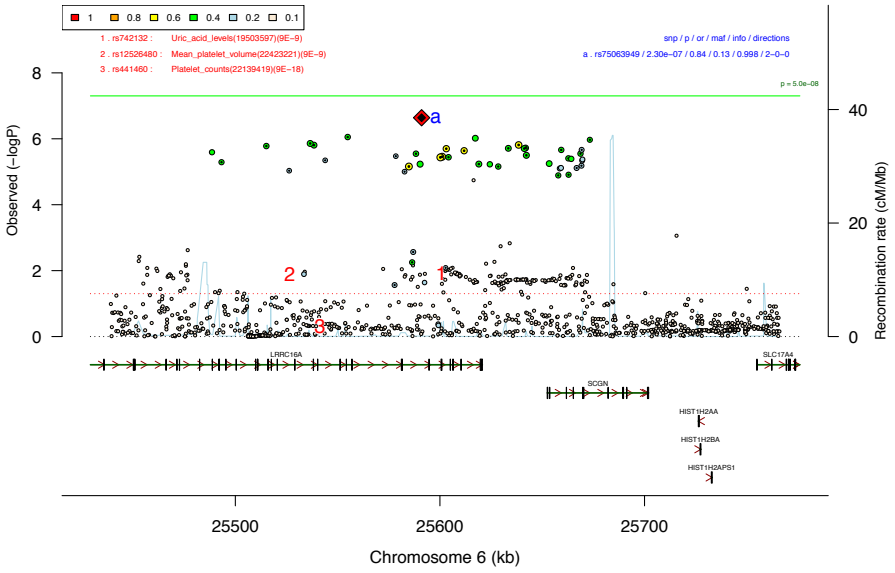
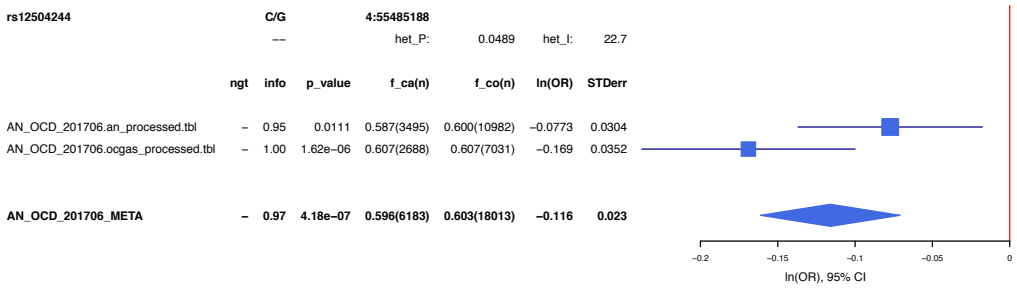


Figure S1. Forest plot **(a)** and regional association plot **(b)** for a locus near the MHC on chr6 that is driven by both AN and OCD.

(a)



(b)

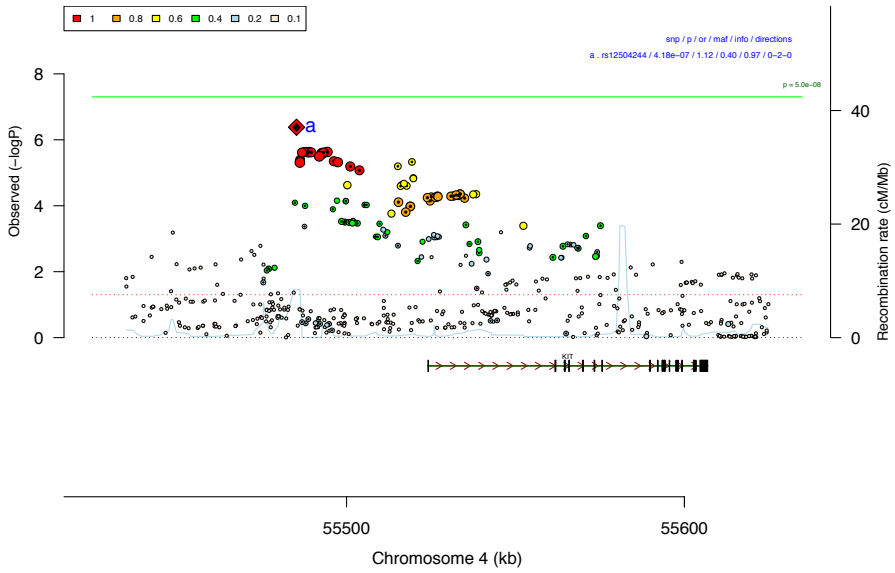
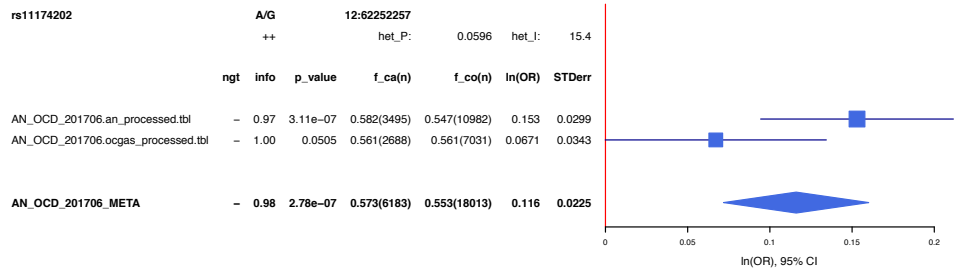


Figure S2. Forest plot **(a)** and regional association plot **(b)** for a locus found upstream of *KIT* on chr4 that is driven by both AN and OCD.

(a)



(b)

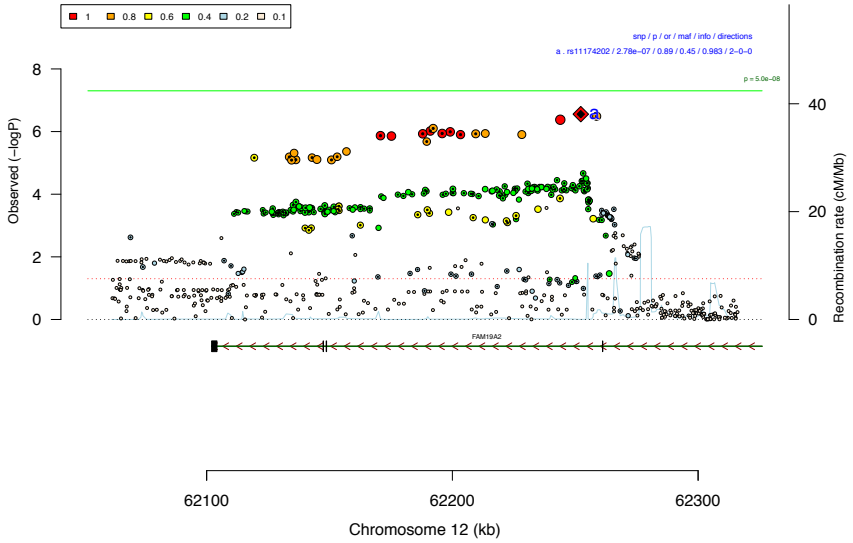


Figure S3. Forest plot **(a)** and regional association plot **(b)** for a locus found within *FAM19A2* on chr12 appears primarily driven by AN.

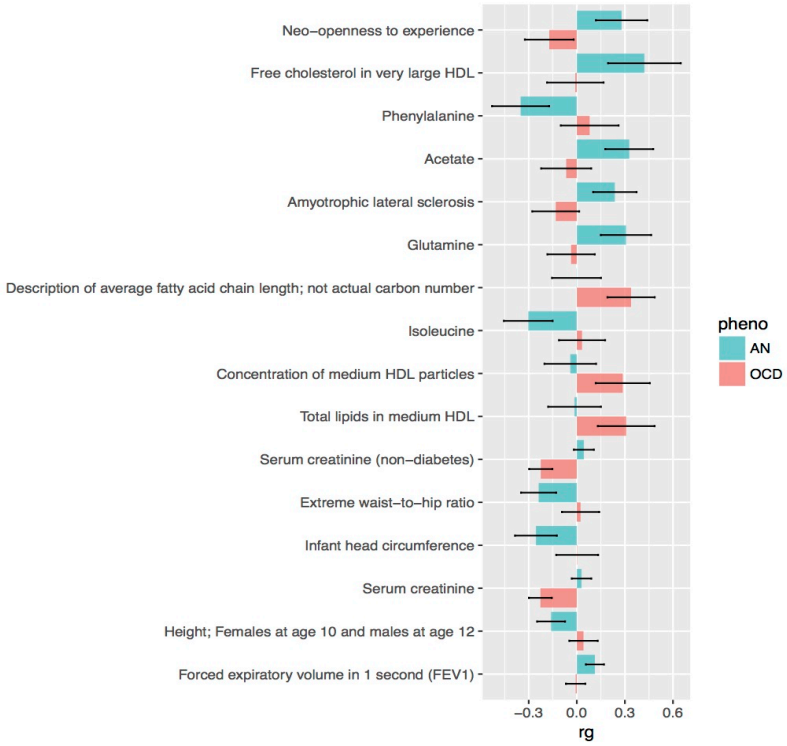


Figure S4. Phenotypes where directions of correlation with AN and OCD point in different directions.

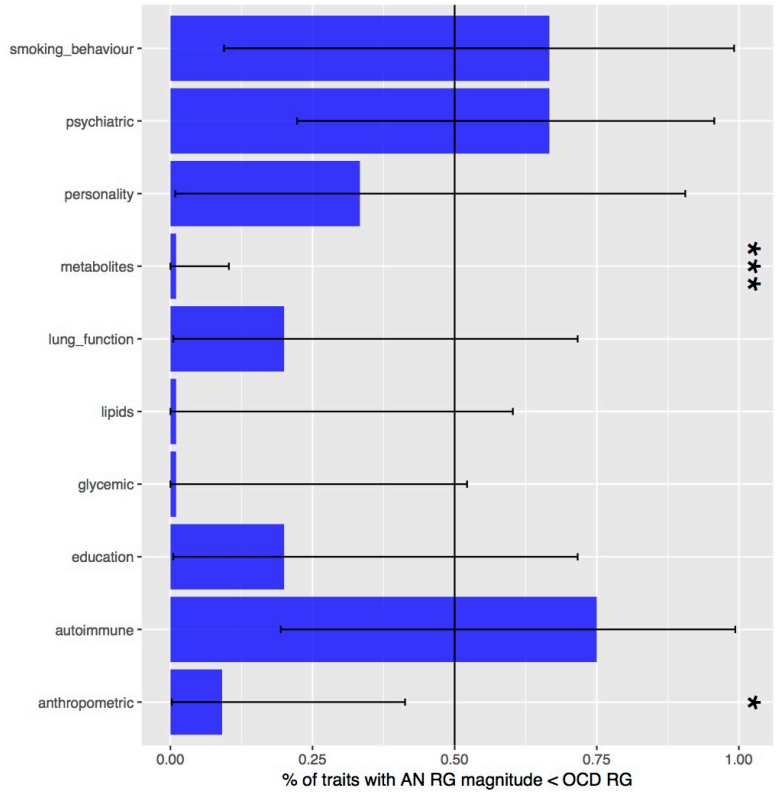


Figure S5. Results of magnitude sign test for each trait category set, testing the null for a set of traits within a category that have the same direction of correlation with AN and with OCD, that the proportion of these traits with a larger magnitude of correlation in AN will equal the proportion with a higher magnitude in OCD. In other words, categories depicted above with values shifted to the left are biased in magnitude towards AN, and those shifted to the right are biased towards OCD. In general, we see bias in magnitude towards AN across categories pertaining to metabolism. For indicators of significance, "***" indicates $p < 0.0001$, while "*" indicates $p < 0.05$.

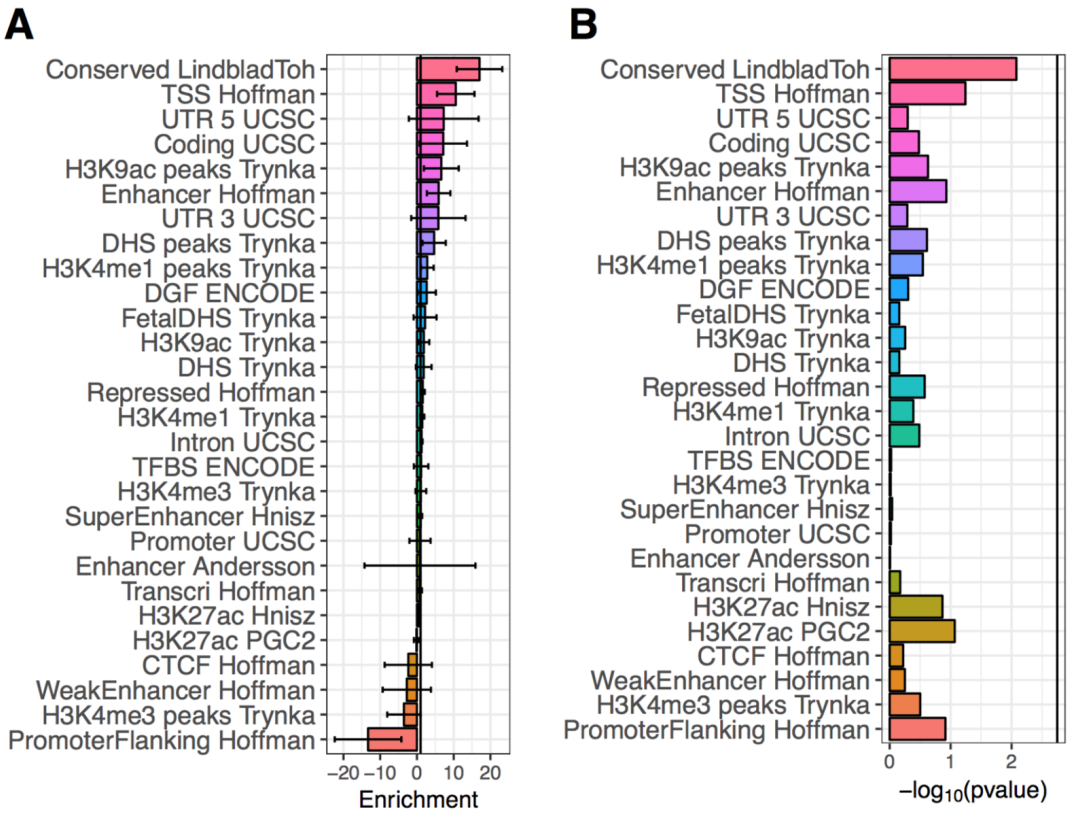


Figure S6. Heritability enrichment of the AN -OCD cross-disorder phenotype partitioned by genomic annotations. The black vertical bar represents the expected heritability enrichment of 1 (**A**) and the Bonferroni significant threshold (**B**).

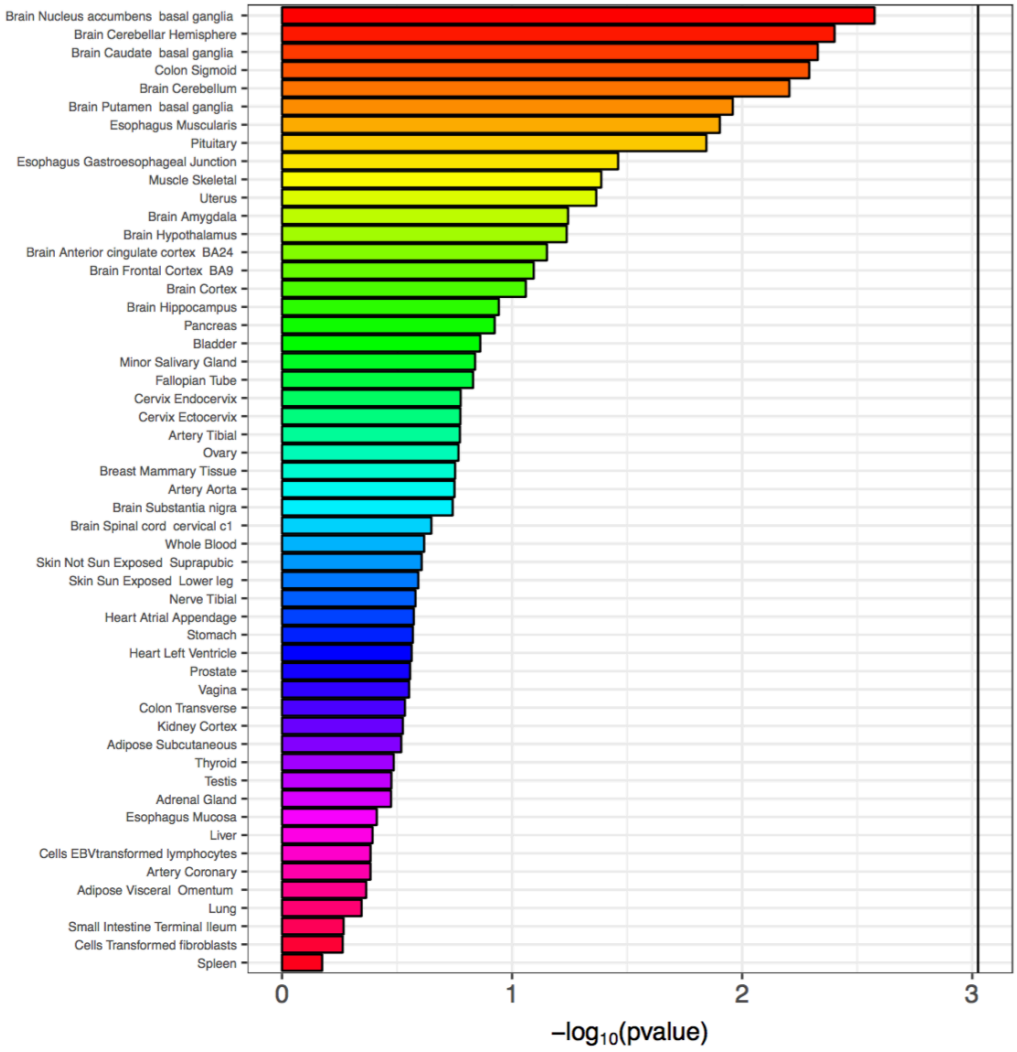


Figure S7. Association between tissue-specific expression and gene-level genetic association to AN-OCD. The black vertical bar represents the Bonferroni significant threshold.

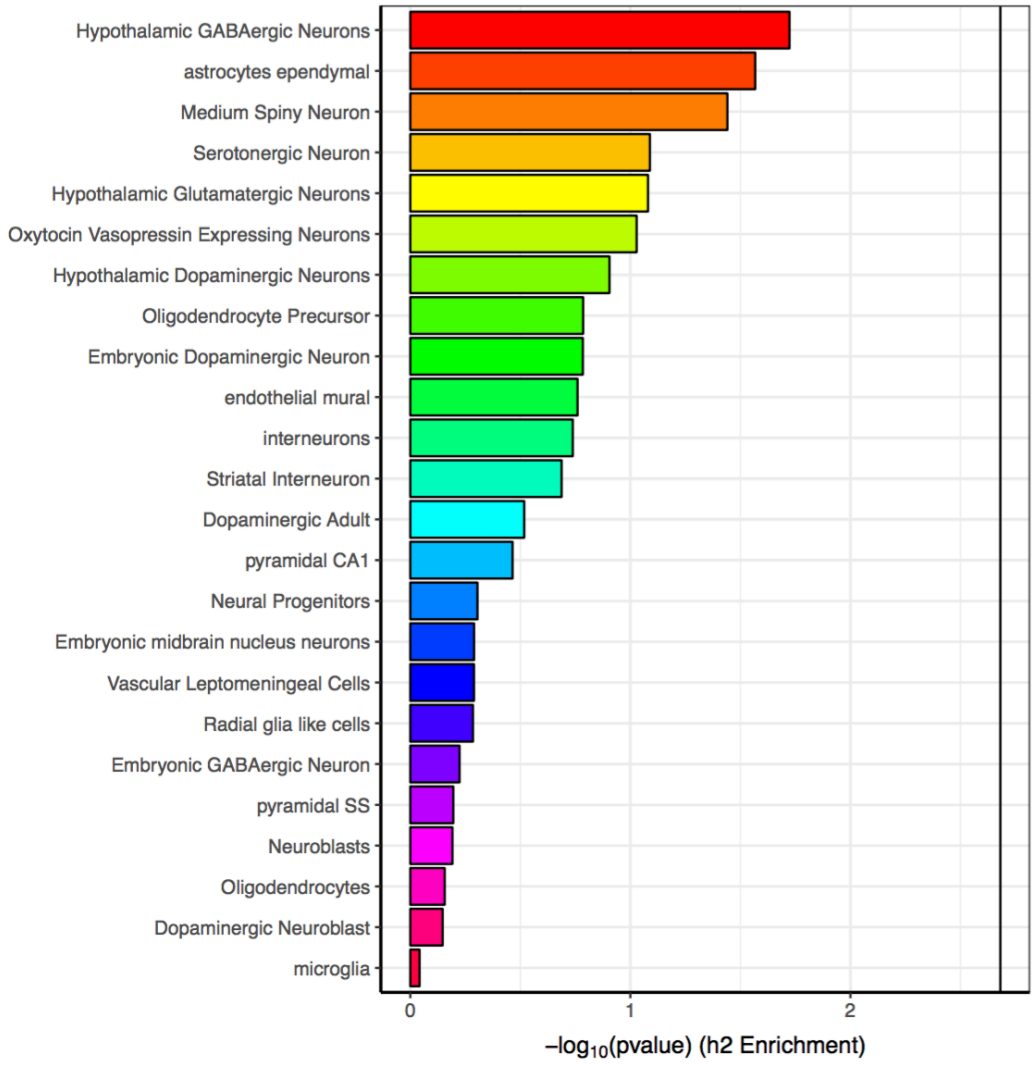


Figure S8. Association between mouse brain cell-type specific expression and gene-level genetic association to AN-OCD using LD score regression. The black vertical bar represents the Bonferroni significant threshold.

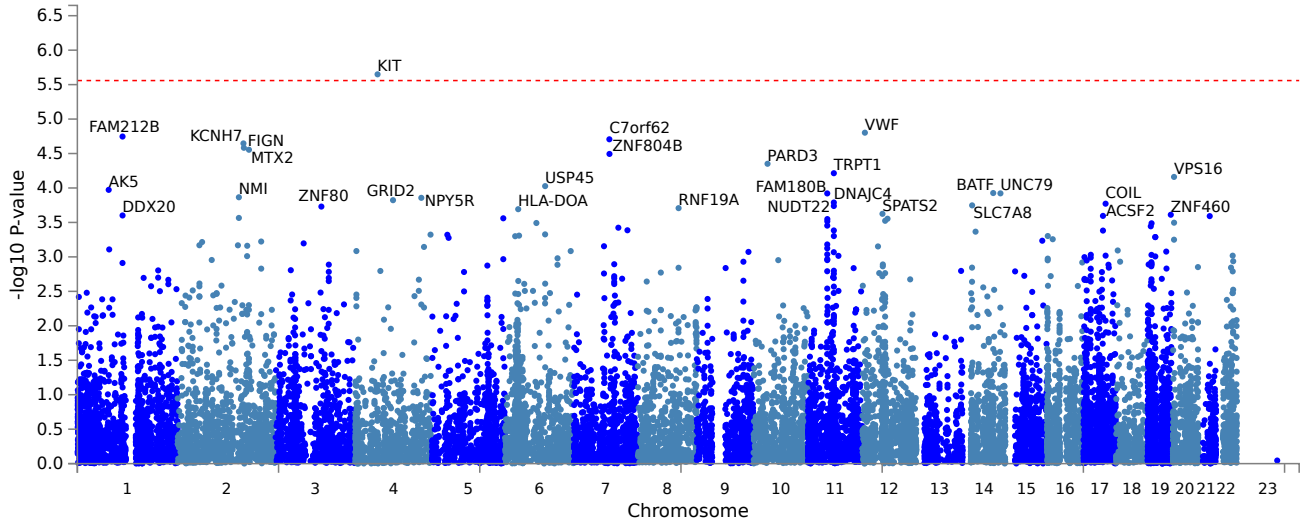


Figure S9. Gene-based association results from MAGMA for AN-OCD. Association was tested using the SNP-wise mean model, in which the sum of -log(SNP *p*-value) for SNPs located within the transcribed region was used as test statistic. All gene-based tests were performed on loci extending from 35kb upstream of transcription start site to 10kb downstream of the transcription end site. MAGMA accounts for gene-size, number of SNPs in a gene and LD between markers when estimating gene-based *p*-values.