ADHD GWAS Results, June 2017 Release

Introduction

These are the GWAS results files from the meta-analysis of ADHD by the Psychiatric Genomics Consortium (PGC) and the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH) released in June 2017.

Citation for studies using these data:
Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E., Belliveau, R., Bybjerg-Grauholm, J., Bækved-Hansen, M., Cerrato, F., Chambert, K., Churchhouse, C., Dumont, A., Eriksson, N., Gandal, M., Goldstein, J., Grove, J., Hansen, C., Hauberg, M., Hollegaard, M., Howrigan, D.P., Huang, H., Maller, J., Martin, A.R., Moran, J., Pallesen, J., Palmer, D.S., Pedersen, C.B., Pedersen, M.G., Poterba, T., Poulsen, J.B., Ripke, S., Robinson, E.B., Satterstrom, F.K., Stevens, C., Turley, P., Won, H., ADHD Working Group of the Psychiatric Genomics Consortium (PGC), Early Lifecourse & Genetic Epidemiology (EAGLE) Consortium, 23andMe Research Team, Andreassen, O.A., Burton, C., Boomsma, D., Cormand, B., Dalsgaard, S., Franke, B., Gelernter, J., Geschwind, D., Hakonarson, H., Haavik, J., Kranzler, H., Kuntsi, J., Langley, K., Lesch, K.-P., Middeldorp, C., Reif, A., Rohde, L.A., Roussos, P., Schachar, R., Sklar, P., Sonuga-Barke, E., Sullivan, P.F., Thapar, A., Tung, J., Waldman, I., Nordentoft, M., Hougaard, D.M., Werge, T., Mors, O., Mortensen, P.B., Daly, M.J., Faraone, S.V., Børglum, A.D., & Neale, B.M. Discovery of the first genome-wide significant risk loci for ADHD. Preprint at http://biorxiv.org/content/early/2017/06/03/145581 (2017).

Disclaimer

These data are provided "as is", and without warranty, for scientific and educational use only. Ifyou download these data, you acknowledge that these data will be used only for non-commercial research purposes; that the investigator is in compliance with all applicable state, local, and federal laws or regulations and institutional policies regarding human subjects and genetics research; that secondary distribution of the data without registration by secondary parties is prohibited; and that the investigator will cite the appropriate PGC publication in any communications or publications arising directly or indirectly from these data.

Methods

See the manuscript for full details. Briefly:

Genotype array data for 20,183 ADHD cases and 35,191 controls were collected from 12 cohorts.
These samples included a population-based cohort of 14,584 cases and 22,492 controls from Denmark collected by the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), and 11 European, North American and Chinese cohorts aggregated by the Psychiatric Genomics Consortium (PGC). ADHD cases in iPSYCH were identified from a national research register and diagnosed by psychiatrists at a psychiatric hospital according to ICD10 (F90.0), and genotyped using Illumina PsychChip. Designs for the PGC cohorts has been described previously.
Prior to analysis, stringent quality control procedures were performed on the genotyped markers and individuals in each cohort using the bioinformatics pipeline Ricopili (available at https://github.com/Nealelab/ricopili). In order to avoid potential study effects the 11 PGC samples and the 23 genotyping batches within iPSYCH were each processed separately. Subjects and SNPs were included in the analyses based on the following quality control parameters: SNP call rate > 0.95 (before sample removal), subject call rate > 0.98 (> 0.95 for

the iPSYCH samples), autosomal heterozygosity deviation (| F_{het} | < 0.2), SNP call rate > 0.98 (after sample removal), difference in SNP missingness between cases and controls < 0.02, and SNP Hardy-Weinberg equilibrium (HWE) ($P > 10^{-6}$ in controls or $P > 10^{-10}$ in cases). Related individuals were removed, and genetic outliers were excluded based on principal component analysis. Non-genotyped markers were imputed using the 1000 Genomes Project Phase 3 reference panel. For trio cohorts, pseudocontrols were defined from phased haplotypes prior to imputation.

GWAS was conducted in each cohort using logistic regression with the imputed additive genotype dosages. Principal components were included as covariates to correct for population stratification, along with relevant study-specific covariates where applicable, and variants with imputation INFO score < 0.8 or minor allele frequency (MAF) < 0.01 were excluded. The GWAS were then metaanalyzed using an inverse-variance weighted fixed effects model. Association results were considered only for variants with an effective sample size greater than 70% of the full metaanalysis, leaving 8,047,421 variants in the final meta-analysis. A meta-analysis restricted to European-ancestry individuals (19,099 cases, 34,194 controls) was also performed to facilitate secondary analyses.

File Description

adhd_jun2017.gz: Full ADHD GWAS meta-analysis (20,183 cases, 35,191 controls)
adhd_eur_jun2017.gz: European ancestry meta-analysis (19,099 cases, 34,194 controls)

CHR	Chromosome (hg19)
SNP	Marker name
BP	Base pair location (hg19)
A1	Reference allele for OR (may or may not be minor allele)
A2	Alternative allele
INFO	Imputation information score
OR	Odds ratio for the effect of the A1 allele
SE	Standard error of the log(OR)
P	P-value for association test in the meta-analysis

Additional Notes

- For long insertion/deletion variants, the A1/A2 alleles are truncated to the first 13 bases
  with a specification of the remaining length (e.g. AACACACACACAC+16)
- For multiallelic variants, “m” is appended to the marker name for different alternative alleles in order to insure that the marker name is unique.
- The reported imputation INFO score is a weighted average across the cohorts contributing to meta-analysis for that variant
- Allele frequencies and case/control counts per variant are currently omitted from public release for data privacy. For inquiries about accessing this data, please contact the ADHD Data Access Committee (DAC) representative (pgc.dac.add@gmail.com).

Non-Compete Request for Pre-Publication Summary Statistics

Consistent with the responsibilities of resource users under 2003 Fort Lauderdale principles and PGC data sharing policies, we request that investigators do not compete with the following analyses before publication of the submitted paper:

(i) analyses in the submitted paper, and
(ii) ongoing cross-disorder/trait studies between ADHD and the significantly correlated traits identified in the paper (illustrated in Figure 3 of the paper, which is inserted below).

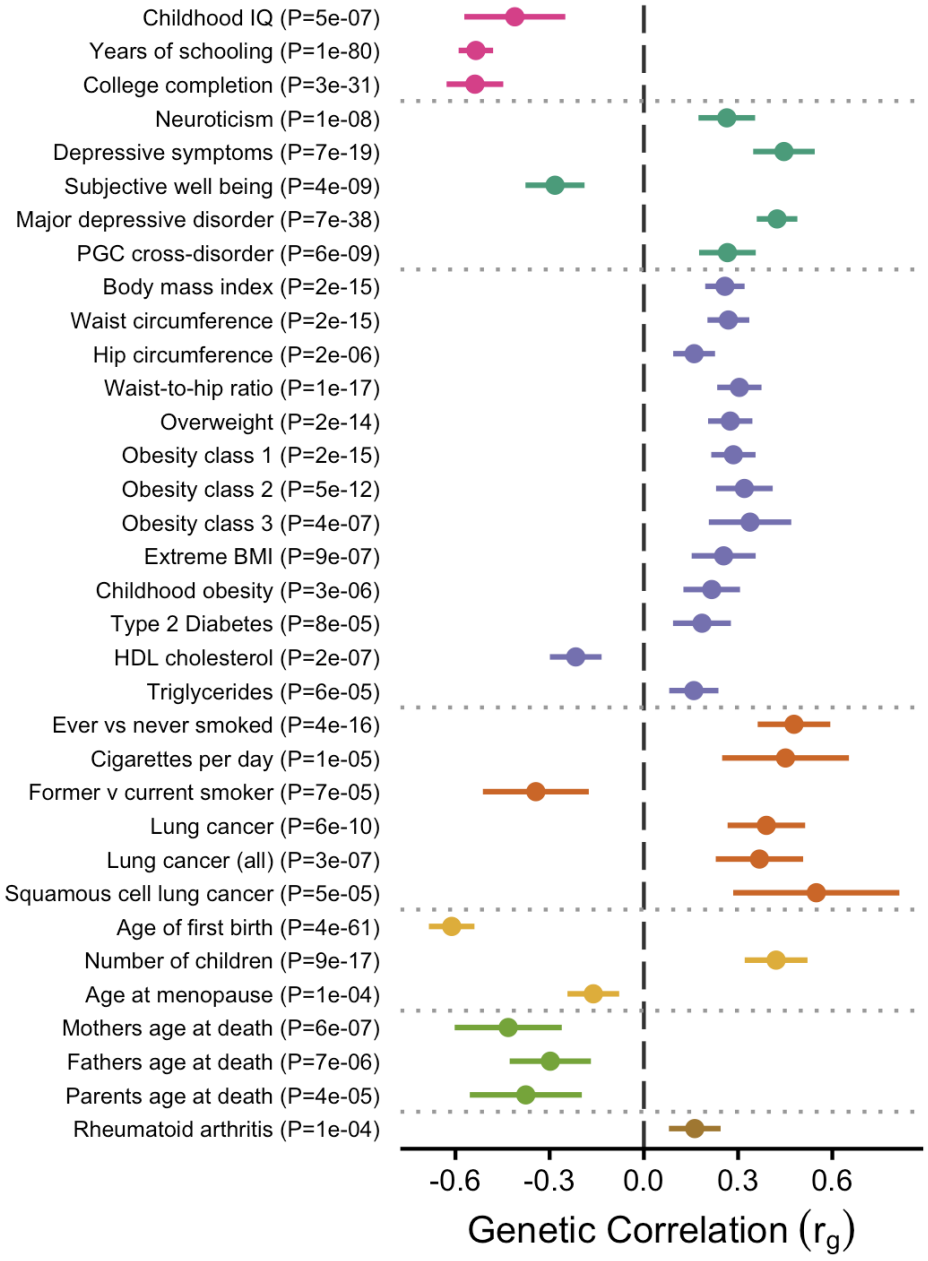

Figure 3. Significant genetic correlations between ADHD and other traits reveal overlap of genetic risk factors for ADHD across several groups of traits (grouping indicated by a horizontal line): educational, psychiatric/personality, weight (and possible weight related traits), smoking behavior/smoking-related cancer, reproductive traits and parental longevity. In total 220 traits were tested. Two significant educational phenotypes are omitted due to substantial overlap with years of schooling. Error bars indicate 95% confidence limits.

Data Use Agreement

1. Investigators acknowledge that these data are provided on an "as-is" basis, without warranty of any type, expressed or implied, including but not limited to any warranty as to their performance, merchantability, or fitness for any particular purpose;
2. Investigators will use these results for scientific research and educational use only.
3. The downloaded results can be shared among collaborators but the reposting or public distribution of the result file is prohibited;
4. Investigators certify that they are in compliance with all applicable local, state, and federal laws or regulations and institutional policies regarding human subjects and genetics research;
5. Investigators will cite the appropriate publication in any communications or publications arising directly or indirectly from these data;
6. Investigators will never attempt to identify any participant who contributed to these data;
7. Investigators may not use these data to develop any type of risk or predictive test for an unborn individual;
8. For any risk or predictive test for a child or adult, investigators must acknowledge that this is an experimental use of these data and that essentially all psychiatric disorders have important non-genetic etiological components;
9. When these data are made available prior to publication, investigators agree to respect and not compete with the scientific priorities of the iPSYCH team according to the Fort Lauderdale principles.

Experience has taught us that the appropriate use of these data requires considerable attention to detail, prior experience, and technical skill. Errors are easy to make. If investigators use these data, any and all consequences are entirely their responsibility.