Supplementary Files

Supplementary Figures and Tables can be downloaded from here: https://www.dropbox.com/sh/0933zxr9va1gc5c/AACliTmO7v_dYOVYyOoA_MjEa?dl=0

Online Methods

Participants

Participants were customers of 23andMe, a personal genetics company. The 23andMe cohort has been described in detail elsewhere[1,2]. All participants included in the analyses provided informed consent and answered surveys online according to a human subjects research protocol, which was reviewed and approved by Ethical & Independent Review Services, an AAHRPP-accredited private institutional review board (http://www.eandireview.com). All participants completed the online version of the questionnaire on the 23andMe participant portal. The number of participants and participant overlap is provided in Table 1. Only participants who were primarily of European ancestry (97% European Ancestry) were selected for the analysis using existing methods[3]. Unrelated individuals were selected using a segmental identity-by-descent algorithm[4].

Table 1: Number of participants for each measure.

Measure	All	Female	Male
EQ	46,861	24,543	22,318
SQ-R	51,564	25,501	26,063
SQ-R and EQ overlap	38,889	20,127	18,762

Measures

Two online questionnaires were used in this study. The first, the Empathy Quotient (EQ)[5], is a self-report measure of empathy, and includes items relevant to both cognitive and affective empathy. It comprises 64 questions and has a good test-retest reliability[6]. In this study, participants scored a maximum of 80 and a minimum of 0. The second measure is the

Systemizing Quotient-Revised (SQ-R), which is self-report measure of systemizing drive, or interest in rule-based patterns[7]. There are 75 items on the SQ-R, with a maximum score of 150 and a minimum score of 0.

Genotyping, imputation and quality control

DNA extraction and genotyping were performed on saliva samples by the National Genetic Institute. Participants were genotyped using one of four different genotyping platforms. Participants were genotyped on one of four different platforms – V1, V2, V3 and V4. The V1 and V2 platforms have a total of 560,000 SNPs largely based on the Illumina HumanHap550+ BeadChip. The V3 platform has 950,000 SNPs based on the Illumina OmniExpress+ Beadchip and has custom content to improve the overlap with the V2 platform. The V4 platform is a fully customized array and has about 570,000 SNPs. All samples had a call rate greater than 98.5%. A total of 1,030,430 SNPs (including Insertion/Deletion or InDels) were genotyped across all platforms. Genotyped SNPs were filtered for quality control. Imputation was performed using the September 2013 release of the 1000 Genomes Project.

SNPs present only on platform V1, or in chromosome Y and mitochondrial chromosomes were excluded due to small sample sizes and unreliable genotype calling respectively. Next, using trio data, where available, SNPs that failed a parent offspring transmission test were excluded. SNPs were also excluded if they failed the Hardy-Weinberg Equilibrium Test at $p < 10^{-20}$, or had a genotype rate of less than 90%. Phasing was performed using Beagle4 (V3.3.1) in batches of 8000-9000 individuals with chromosomal segments of no more than 10,000 genotyped SNPs. SNPs were excluded if they were not in Hardy-Weinberg equilibrium ($P < 10^{-20}$), had a call rate less than 95%, or had discrepancies in allele frequency compared to the reference European 1000 Genomes data (chi-squared $P < 10^{-15}$). Imputation was performed using Minimac25 using the September 2013 release of the 1000 Genomes Phase 1 reference haplotypes. phased using Beagle4 (V3.3.1). We restricted the analyses to only SNPs that had a minor allele frequency (MAF) of at least 1%. After quality control, 9,955,952 SNPs were analysed. Genotyping, imputation, and preliminary quality control were performed by 23andMe.

Genetic association

We performed a linear regression assuming an additive model of genetic effects. Age and sex along with the first five ancestry principal components were included as covariates. Additionally, for each trait, we performed a male-only and a female-only linear regression analysis to identify sex-specific loci. Since we were performing three tests for each trait (male-only, female-only, and males and females combined with sex as a covariate), we used a threshold of $P < 1.66 \times 10^{-8}$ to identify significant SNPs for each trait. Leading SNPs in each loci were identified after pruning for LD ($r^2 > 0.8$) using SNAP[8].

Genomic inflation factor, heritability, genetic correlation and functional enrichment

We used Linkage Disequilibrium Score regression coefficient (LDSC) to calculate genomic inflation due to population stratification[9] (https://github.com/bulik/ldsc). The intercept for the SQ-R GWAS was 0.998 and the intercept for the EQ GWAS was 0.993 indicating that there was no unaccounted population stratification. Heritability and genetic correlation was performed using extended methods in LDSC[10]. Summary GWAS data for schizophrenia[11,12], bipolar disorder[11], autism[11], anorexia[13], and depression[11] were downloaded from the Psychiatric Genomics Consortium website (http://www.med.unc.edu/pgc/downloads). Summary GWAS data for educational attainment measured through number of college years[14] was downloaded from the Social Science Genetic Association Consortium website (http://ssgac.org/Data.php). The rationale for choosing these conditions for genetic correlation were two fold. First, summary data were easily available for these conditions with effect sizes and standard errors reported, which is needed for the analyses. Second, LDSC works best when used for studies with sample sizes greater than 4,000. For schizophrenia, we first performed the genetic correlation with the smaller Caucasian only GWAS dataset (N= 17115 cases and controls)[11]. Later, to confirm the genetic correlation, we conducted the analysis with the larger dataset which also includes data from East Asian cohorts[12] (schizophrenia-2 in the results; N=79845 cases and controls). This was possible as there are two large, overlapping publicly available datasets for schizophrenia, which, to our knowledge, is not present for the other conditions tested. The North West European LD scores were used in the analysis, and the intercepts were not constrained as the extent of participant overlap was unknown. Supplementary Table S1 provides the sample sizes for GWAS data used for genetic correlation. Due to the correlation between several of the traits tested, Bonferroni

correction would be too stringent. Accordingly, we used a False Discovery based approach and report significant correlations if they had an FDR q-value below 0.05. We identified enrichment in genomic functional elements for the traits using extended methods in LDSC[15].

Gene-based analysis and sex-difference analysis

Gene based analysis was performed using MetaXcan[16] (https://github.com/hakyimlab/MetaXcan). We used tissue weights provided in the MetaXcan implementation, that uses data from the GTEx project[17]. To identify significant genes for the two traits, we used the non-stratified GWAS as they had the largest sample sizes. For each trait, we ran gene-based analysis for ten neural tissues: anterior cingulate cortex (BA24), caudate basal ganglia, cerebellar hemisphere, cerebellum, cortex, frontal cortex (BA9), hippocampus, hypothalamus, nucleus accumbens basal ganglia, and putamen basal ganglia. We filtered out genes where there were 0 SNPs from our dataset, and genes that correlated poorly with predicted models of gene-expression ($R^2 < 0.01$) as implemented in the software. We used FDR correction to correct for the multiple tests for each trait (correcting for more than 50,000 tests for each trait).

For sex-difference analysis, we ran MetaXcan on the sex-stratified analyses for cortex. We focussed on the cortex as it was relevant for the two traits investigated and we had access to the list of sex-differentially expressed genes from the cortex[18]. To check for overlap, we ran hypergeometric tests. We identified nominally significant genes ($P < 0.05$) in the two sexes separately and checked for overlap among these lists after pruning the background gene-lists to a common set of genes for both the sexes. For the differential gene-expression study, we downloaded the list of differentially expressed genes in the cortex from Werling et al., 2016[18]. We used the two adult datasets available from Supplementary Data 1 from the article. We ran the test for all the four stratified analyses (males : SQ-R and EQ; females SQ-R and EQ) using the first adult dataset, which is the larger of the two, and replicated the significant findings using the second adult dataset. We divided the differentially expressed list of genes into higher-expressed in males (male-expressed) and higher expressed in females (female-expressed) and included all the autosomal genes with a fold-difference greater than 1 in the lists. We looked for enrichment in both the male-expressed and the female-expressed gene lists separately in all four stratified analyses.

References

1. Tung, J. Y. *et al.* Efficient replication of over 180 genetic associations with self-reported medical data. *PLoS One* **6,** e23473 (2011).

2. Do, C. B. *et al.* Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet.* **7,** e1002141 (2011).

3. Eriksson, N. *et al.* Novel associations for hypothyroidism include known autoimmune risk loci. *PLoS One* **7,** e34442 (2012).

4. Henn, B. M. *et al.* Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples. *PLoS One* **7,** e34267 (2012).

5. Baron-Cohen, S. & Wheelwright, S. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* **34,** 163–75 (2004).

6. Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S. & David, A. S. Measuring empathy: reliability and validity of the Empathy Quotient. *Psychol. Med.* **34,** 911–9 (2004).

7. Wheelwright, S. *et al.* Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ). *Brain Res.* **1079,** 47–56 (2006).

8. Johnson, A. D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24,** 2938–9 (2008).

9. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).

10. Bulik-Sullivan, B. *et al. An Atlas of Genetic Correlations across Human Diseases and Traits*. *bioRxiv* (Cold Spring Harbor Labs Journals, 2015). doi:10.1101/014498

11. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381,** 1371–9 (2013).

12. Ripke, S. *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511,** 421–7 (2014).

13. Boraska, V. *et al.* A genome-wide association study of anorexia nervosa. *Mol. Psychiatry* **19,** 1085–94 (2014).

14. Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340,** 1467–71 (2013).

15. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47,** 1228–1235 (2015).

16. Barbeira, A. *et al. MetaXcan: Summary Statistics Based Gene-Level Association Method Infers Accurate PrediXcan Results*. *bioRxiv* (Cold Spring Harbor Labs Journals, 2016). doi:10.1101/045260

17. Ardlie, K. G. *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science (80-.).* **348,** 648–660 (2015).

18. Werling, D. M., Parikshak, N. N. & Geschwind, D. H. Gene expression in human brain implicates sexually dimorphic pathways in autism spectrum disorders. *Nat. Commun.* **7,** 10717 (2016).