A genome-wide association study of Chinese and English language abilities in Hong Kong Chinese children

Yu-Ping Lin[1], Yujia Shi[1], Ruoyu Zhang[1], Xiao Xue[1], Shitao Rao[1,2,3], Kevin Fai-Hong Lui[3,4], Dora Jue PAN[5,6], Urs Maurer[6,7], Richard Kwong-Wai Choy[8], Silvia Paracchini[9], Catherine McBride[6], Hon-Cheong So[1,7,10-14]

[1]School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong

[2] Department of Bioinformatics, Fujian Key Laboratory of Medical Bioinformatics, School of Medical Technology and Engineering, Fujian Medical University, Fuzhou, China;

[3] Key Laboratory of Ministry of Education for Gastrointestinal Cancer, School of Basic Medical Sciences, Fujian Medical University, Fuzhou, China;

[3]Department of Applied Psychology, Lingnan University, Tuen Mun, Hong Kong

[4]Wofoo Joseph Lee Consulting and Counselling Psychology Research Centre, Lingnan University, Tuen Mun, Hong Kong

[5]School of Humanities and Social Science, The Chinese University of Hong Kong (Shenzhen), Shenzhen

[6]Department of Psychology, The Chinese University of Hong Kong, Hong Kong

[7]Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong SAR, China

[8]Department of Obstetrics and Gynecology, The Chinese University of Hong Kong, Hong Kong

[9] School of Medicine, University of St Andrews, North Haugh KY16 9TF, St Andrews, Scotland

[10]KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research of Common Diseases, Kunming Institute of Zoology and The Chinese University of Hong Kong, China

[11] Department of Psychiatry, The Chinese University of Hong Kong, Hong Kong

[12] CUHK Shenzhen Research Institute, Shenzhen, China

[13]Margaret K.L. Cheung Research Centre for Management of Parkinsonism, The Chinese University of Hong Kong, Shatin, Hong Kong

[14]Hong Kong Branch of the Chinese Academy of Sciences Center for Excellence in Animal Evolution and Genetics, The Chinese University of Hong Kong, Hong Kong SAR, China

Submitted on 30 July 2022

Abstract

Background: Reading and language skills are important and known to be heritable, and dyslexia and developmental language disorder are commonly recognized learning difficulties worldwide. However, the genetic basis underlying these skills remains poorly understood. In particular, most previous genetic studies were performed on Westerners. To our knowledge, few or no previous genome-wide association studies (GWAS) have been conducted on literacy skills in Chinese as a native language or English as a second language (ESL) in a Chinese population.

Methods: We conducted GWAS and related bioinformatics analyses on 34 reading/language-related phenotypes in Hong Kong Chinese bilingual children (including both twins and singletons; *N*=1046). We performed association tests at the single-variant, gene, pathway levels, and transcriptome-wide association studies (TWAS) to explore how imputed expression changes might affect the phenotypes. In addition, we tested genetic overlap of these cognitive traits with other neuropsychiatric disorders, as well as cognitive performance (CP) and educational attainment (EA) using polygenic risk score (PRS) analysis.

Results: Totally 9 independent loci (LD-clumped at r^2=0.01) reached genome-wide significance (p<5e-08) (filtered by imputation quality metric Rsq>0.3 and having at least 2 correlated SNPs(r^2>0.5) with p<1e-3). The loci were associated with a range of language/literacy traits such as Chinese vocabulary, character and word reading, dictation and rapid digit naming, as well as English lexical decision. Several SNPs from these loci mapped to genes that were reported to be associated with intelligence, EA other neuropsychiatric phenotypes, such as *MANEA*, *TNR, PLXNC1* and *SHTN1*. We also revealed significantly enriched genes and pathways based on SNP-based analysis. In PRS analysis, EA and CP showed significant polygenic overlap with a variety of language traits, especially English literacy skills. ADHD PRS showed a significant association with English vocabulary score.

Conclusions: This study revealed numerous genetic loci that may be associated with Chinese and English abilities in a group of Chinese bilingual children. Further studies are warranted to replicate the findings and elucidate the mechanisms involved.

Introduction

Literacy and language skills are important for academic development in children. Learning difficulties (e.g., dyslexia) are common and may affect one's school performance, leading to poorer work attainment and socioeconomic status, as well as decreased general well-being [1]. Multiple cognitive and language skills often serve as a strong foundation for literacy and language development; these include working memory, rapid naming, and vocabulary knowledge [2]. Different factors of environmental and genetic origins may also affect children's literacy-related and language-related skills across languages. Family, twin, and adoption studies have provided strong evidence that these complex cognitive and language traits and academic performance in young children are heritable [3 4 5 6 7] and also highly polygenic [8 9]. However, the exact genes/variants involved in these traits are still not well understood, probably due to the complexity of the phenotypes and difficulty in gathering sufficient samples.

In recent years, several GWAS studies have been conducted on reading and language abilities in European populations. Several studies have focused on developmental dyslexia (DD) or high/low reading ability as a binary outcome, adopting a case-control study method [9 10 11 12 13 14]. Such study design may enable a larger sample size to be collected, but also has its shortcomings. Language and literacy skills cover a broad range of phenotypes, and dyslexia is also a highly heterogenous condition. The focus on a single binary outcome may limit our understanding into the biological mechanisms underlying different domains of language abilities. Other studies have investigated reading and language abilities as continuous traits [8 15 16 17 18 19 20].

However, given the relatively high heritability of literacy and language skills [21 22], the genetic variants discovered thus far are still far from explaining the full genetic basis of these complex traits. In addition, most previous GWAS were conducted in European populations. However, the genetic architecture of language phenotypes may be different across ancestries, and some of the variants may be more readily discovered in other populations due to differences in allele frequency or LD (linkage disequilibrium) structure.

In addition, to our knowledge, no previous GWAS have been published on Chinese children's literacy and language skills in native Chinese and English as a second language (ESL). We note that in one recent GWAS on dyslexia [9], several associated loci were also replicated in the Chinese Reading Study of reading accuracy and fluency; yet the primary GWAS was conducted predominantly on population of European ancestry. Given possible differences in mechanisms underlying Chinese and English literacy and language skills, it is essential to study the genetic variants underlying Chinese literacy phenotypes. While some previous studies have investigated the genetics of cognitive and language phenotypes, most have only focused on a limited number or domain of phenotypes (e.g., rapid naming, word reading).

In view of the above limitations, here we conducted GWAS and related bioinformatics analyses on a comprehensive panel of 34 literacy/language-related phenotypes in a Hong Kong Chinese population. The wide coverage enables a systematic and unbiased analysis of a variety of literacy and language-related phenotypes. Since this is among the first study of Chinese- and ESL-related phenotypes in a Chinese population, and the genetic bases of such phenotypes are

still largely unknown, it is our objective to explore a wider range of traits to maximize the chance of discovery, and to provide a starting point and also important reference for future studies.

Briefly, in this work, we investigated how genetics is associated with individual differences in Chinese and English reading and writing. We performed association tests at the single-variant, gene, and pathway levels, and employed transcriptome-wide association studies (TWAS) to explore how genotype-imputed expression changes affect the phenotypes. In addition, we tested potential associations between these complex cognitive traits with other neuropsychiatric disorders, as well as cognitive performance and educational attainment by polygenic risk score (PRS) analysis. To the best of our knowledge, this is the first GWAS conducted on a comprehensive range of Chinese-language and ESL-related phenotypes in a Chinese population.

Subject and methods

1. Participants

The participants in our study were Hong Kong Chinese-English bilingual twins and singletons, with Chinese (Cantonese) as their native language. Children aged between 3 to 11 were recruited through kindergarten and primary schools in Hong Kong. A total of 1048 children were recruited for this genetic study, including 274 MZ subjects (137 pairs), 350 DZ subjects (175 pairs) and 424 singletons. Zygosity determination on twin pairs was based on the genotyped small tandem repeat (STR) markers using Quantitative Fluorescence Polymerase Chain Reaction (QF-PCR)[23]. Singleton children were selected from the same schools as those twin pairs. Parental written consent for all the participants was obtained before testing. Children completed a series of cognitive and literacy-related tasks in Chinese and English either in a laboratory setting, their school, or their home by trained research assistants. For details of the tasks and phenotypes, please refer to the supplementary text. Briefly, a total of 34 phenotypes were included (Table 1), covering a wide range of literacy- and language-related skills.

All tasks were finished in a given order that had been predetermined. Please refer to the supplementary text for a detailed description of each studied phenotype. A correlation matrix of all phenotypes is presented in Figure S1.

2. Genotype quality control (QC) and imputation

Three groups of subjects, including monozygotic (MZ) twins, dizygotic twins (DZ), and singletons, were genotyped. Based on previous studies [24], reducing the MZ pairs to singletons leads to a loss of statistical power. It has also been shown that including both MZ twins in the genetic analysis does not lead to an inflation of type I error (when relatedness is accounted for) but can improve power [24]. We therefore followed ref [24] and included both MZ twins in our GWAS. Monozygosity was confirmed by QF-PCR as described above, and only one member of each MZ pair was genotyped. The other MZ twin was assumed to share identical genotypes.

Quality control (QC) was performed by PLINK-1.9 on each dataset separately before merging. We removed those SNPs which deviated from Hardy–Weinberg equilibrium (HWE, $P < 1E\text{-}5$), with Minor Allele Frequency (MAF) < 1%, missingness per individual (MIND) > 10%, and missingness per marker (GENO) > 10%. After QC, 911178 SNPs and

1046 individuals were kept for further analysis, including 274 MZ subjects (59 male pairs, 78 female pairs), 349 DZ subjects (39 male pairs, 37 female pairs, 1 member of a female pair and 98 opposite-sex pairs), as well as 423 singletons (218 males, 205 females).

Following QC, variant-level imputation was performed by the Michigan Imputation Server based on "Mininac" [25]. The imputation was based on the reference panel 1000 Genomes (1000G) Phase 3 v5, as previous studies reported satisfactory performance of imputation in Chinese based on the 1000G panel [26]. The imputed data were converted into a binary dosage file by the program "DosageConverter" (https://genome.sph.umich.edu/wiki/DosageConvertor). Imputed variants with INFO score (R-squared) > 0.3 (12,475,316 SNPs) were retained.

3. Genome-wide association study

A genome-wide association study (GWAS) of all phenotypes was conducted through a univariate linear mixed model in GEMMA (http://github.com/genetic-statistics/GEMMA). We included age and sex as fixed-effects covariates. The genetic relationship matrix (GRM) was included as a random effect to account for relatedness between subjects. This approach also controls for population stratification. We tested for the association of allelic dosages with phenotypes as described above. We considered p<5e-8 as the genome-wide significance threshold. Although multiple phenotypes were studied, our primary objective was to explore and prioritize genetic variants for further studies, and a further Bonferroni correction to penalize the number of phenotypes tested may be too conservative for this purpose. Instead, we employed the false discovery rate (FDR) approach to control for multiple testing. FDR controls the expected *proportion* of false positives among the results declared to be significant. This approach has been argued to be a more reasonable methodology as it 'adaptively' considers the data instead of imposing a direct penalty for the number of hypotheses tested, and the FDR approach has also been widely used in genomic studies [27].

To identify independent significant risk loci, we employed PLINK-1.9 to perform LD-clumping with r^2=0.01 and distance = 1000kb, using 1000 Genomes East Asian sample as reference. SNP-to-Gene mapping was done using Bioconductor's software package 'biomaRt' (version 2.48.2) on R-4.0.3.

Histograms of all phenotypes are shown in Figure S2; some of the phenotypes were normally distributed though some were not. Nevertheless, in large sample sizes with few covariates, violation of the normality assumption often does not affect the validity of results [28]. There is no clear consensus on whether transformations (such as the rank-based inverse normal transformation, RINT) should be performed on (non-normal) phenotypes in GWAS. For example, Beasley et al. [29] reported that RINT does not necessarily control type I error and may lead to reduced statistical power, while another study [30] showed improved performance of the RINT approach. Intuitively, the untransformed approach keeps the original value of the phenotype and does not lead to loss of information, and is more interpretable. Here we performed analysis on both RINT-transformed [30] and non-transformed phenotypes for all traits under study. As described below, on inspection of the QQ-plots, most traits have very similar distributions of p-values, except for four phenotypes. We primarily present our results of the non-transformed phenotypes except for the latter four which were RINT-transformed.

4. Gene-based test and pathway enrichment analysis

4.1 Gene-based analysis with MAGMA

Gene-based analysis has been considered more powerful than SNP-based analysis performed in GWAS [31]. We utilized MAGMA (Multi-marker Analysis of GenoMic Annotation) v1.06 to conduct gene-based association tests with GWAS summary statistics of our phenotypes [13]. Briefly, MAGMA considers the aggregate effects of all variants in each gene to produce a gene-based test statistic. We employed the FDR procedure [32] to control for multiple testing. In our gene-based study and the following analyses, results with FDR < 0.05 are regarded as significant, while those with 0.05 < FDR < 0.2 are considered to be suggestive associations.

4.2 Pathway analysis with GAUSS

We subsequently performed pathway enrichment tests with a powerful subset-based gene-set analysis method called GAUSS (Gene-set analysis Association Using Spare Signal) [33], based on gene-based association results obtained by MAGMA. We utilized two collections of gene-sets derived from the Molecular Signature Database (MsigDB v6.2) [34]. The first is a collection of curated pathways (C2) which include canonical pathways such as KEGG, BioCarta, REACTOME, as well as chemical and genetic perturbations; the other is gene-ontology (GO) gene-sets (C5), which include biological processes, molecular processes, and cellular processes. Please refer to https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp for details. If a significant association with a pathway is found, GAUSS also identifies the core subset (CS) of genes within the pathway that is driving the association. FDR was used to control for multiple testing.

4.3 Transcriptome-wide association studies with S-Predixcan & S-Multixcan

We have also employed other approaches to compute gene-based association results. MAGMA is a widely used approach, but it does not consider the functional impact of SNPs (e.g., impact on expression). S-PrediXcan is another gene-based analysis approach which *imputes* gene expression changes in relevant tissues due to genetic variations, using reference eQTL datasets such as GTEx. This approach is also known as the transcriptome-wide association study (TWAS) [35]. Here we considered 13 brain regions, including the amygdala, anterior cingulate cortex (BA24), caudate basal ganglia, cerebellar hemisphere, cerebellum, cortex, frontal cortex (BA9), hippocampus, hypothalamus, nucleus accumbens (basal ganglia), putamen (basal ganglia), spinal cord (cervical c-1) and substantia nigra. To increase statistical power to identify candidate genes, we integrated the joint effects of expression changes across multiple tissues in a secondary analysis by 'S-MultiXcan' [36]. S-MultiXcan combines evidence across tissues using multiple regression (fitting predicted expression as independent variables), which also takes into account the correlation structure.

5. Polygenic risk score analysis

To evaluate the genetic overlap of the studied phenotypes with other neuropsychiatric traits, we also performed a PRS analysis. A PRS analysis for individuals aggregates the joint effect of multiple genetic variants, weighted by the effect size from GWAS summary statistics data. PRS were generated by PLINK 1.9 across 11 P-value thresholds (pthres)={1e-06, 1e-05, 1e-04, 0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5,0.05}(multiple testing corrected by FDR) [37], LD-clumped at r^2=0.1 within a distance of 1000 kb.

The neuropsychiatric traits for constructing PRS included educational attainment (EA) (N= 1,131,881) [38], cognitive performance (CP; N=257,841; derived from scores of verbal-numerical reasoning from the UK Biobank and neuropsychological test results from the COGENT Consortium, details described in [38]), autism spectrum disorders (ASD; N=10610)[39], attention deficit hyperactivity disorder (ADHD), (ADHD; N=55374) [40], schizophrenia (SCZ; N=105318)[41], bipolar disorder (BP; N= 416053)[42] and major depressive disorder (MDD; N=18759) [43]. GWAS summary statistics were downloaded from the Social Science Genetic Association Consortium (SSGAC) (https://www.thessgac.org/), Psychiatric Genomics Consortium (PGC) (https://www.med.unc.edu/pgc) and The Integrative Psychiatric Research project (iPSYCH) (https://ipsych.au.dk/downloads/).

We employed linear mixed models in GEMMA to test for associations between PRS and phenotypes. The model was adjusted for age and sex as fixed effects. GRM was fit as a random effect, accounting for both relatedness and population stratification [44].

Results

Single-variant associations

Quantile-quantile plots (QQ-plots) with lambda (λ) were constructed for each trait with and without RINT transformation. We found that the QQ-plots were very similar for most phenotypes with and without the transformation, except for four [Backward digit span (BDS_Total), Chinese Vocabulary - Receptive Vocabulary (CVA_Total), Chinese digit rapid naming (CDRAN_Mean) and English digit rapid naming (EDRAN_Mean)] (see Figure S3 and Figure S4). For these 4 traits, subsequent analyses were based on the RINT-transformed values, and all four traits showed no evidence of inflated false positives after the transformation based on the updated QQ-plots. Manhattan plot for all traits are shown in Figure S5.

In SNP-based analysis, a total of 9 independent loci (LD-clumped at r^2 threshold=0.01) reached genome-wide (GW) significance (p < 5e-08) with imputation quality score (Rsq)>0.3 and at least 2 correlated SNPs (r^2>0.5) with p<1e-3 (Table 2). The check for correlated significant SNPs was performed to further reduce the risk of false positives, and was done using the default settings of LD-clumping in PLINK. The loci were associated with a variety of language/literacy traits such as Chinese vocabulary, character and word reading, dictation and digit rapid naming, as well as English lexicon decision. Note that 3 loci were each associated with two (correlated) phenotypes, including rs77868538 which was associated with both CVK_Total and CVD_total (correlation (r)=0.97), rs182977703 which was associated with both COM_Norm and CDICT_Total (r=0.37), and rs4865143 which was associated with CWR_total and CVB_total (r=0.63).

We also provide the full list of all GW significant SNPs (with Rsq>0.3) in Table S1 (Table S1.1: all GW-significant hits; Table S1.2: LD-clumped significant hits; Table S1.3: all GWAS results with FDR<0.1). In addition, we also searched the top-listed genes in GWAS catalog for associations with other phenotypes (especially neuropsychiatric traits) in previous studies. Please refer to Table S10 for details.

We detected the largest number of genome-wide significant SNPs with English Lexical Decision (ELD) (see Table S1.2). The most significant association was observed for rs6905617 (C/A, MAF = 0.35, p = 3.29E-09) with ELD; the SNP is located close to *MANEA* (-382.1 kb) and *MANEA-AS1* (-364.7kb). As for Chinese-related traits, the strongest single-variant association with CVK (Chinese Vocabulary – Knowledge) was detected at rs77868538 (T/C, MAF = 0.02548, p = 3.33E-09), an imputed variant located within *TNR* (Tenascin R). In addition to CVK, we also discovered one significant locus for CCR, CWR, CDICT, CDRAN, COM, CVB and CVD respectively (Table 2).

We also calculated the lambda-GC (genomic inflation factor) for each untransformed trait and there was no evidence of inflation (Table S7; largest lambda-GC=1.0255, 29/34 traits showed lambda-GC<1.02).

Association analyses between genetically predicted expression and phenotypes

We evaluated the association between genetically regulated expression (GRex) and phenotypes across multiple brain regions by S-Predixcan. We used pre-computed weights provided by the authors (available at https://predictdb.org/), derived from an elastic net regression model with transcriptome reference data from GTEx(v7). The most significant associations were observed for *DUS3L,* which showed significant associations (FDR < 0.05) with EWR.Total in four brain regions including amygdala (Z = -4.81, P = 4.18E-02), caudate basal (Z = -4.72, P = 4.18E-02), cerebellar hemisphere (Z = -4.69, P = 4.18E-02) and putamen (Z = -4.60, P = 4.77E-02) (see Table S2.1). The top 20 association results from S-PrediXcan are presented in Table 3 (see also Table S2 with top 100 associations).

Furthermore, we also employed S-MulTiXcan to improve power by combining evidence of differential expression across all brain regions. We observed 185 significant gene-level associations (with FDR<0.05) by this approach and identified the best representative brain region (the region showing the strongest single-tissue association). The top 20 results are presented in Table 4 and fuller results in Table S3. We highlight a few findings here. The most significant S-Multixcan association was observed for gene *HSD3B7* with EVA_total (Hydroxy-Delta-5-Steroid Dehydrogenase, 3 Beta- And Steroid Delta-Isomerase 7; best brain region, Brain_Cortex; P = 1.12E-21). *HSD3B7* was also associated with other English literacy phenotypes, such as EVB, EVK, EVD, and EWR. For Chinese literacy skills, the most significant association was observed for gene *SEMA6C* (Semaphorin 6C; Brain_Cerebellar_Hemisphere; Z mean = -1.81; P = 1.98E-16) with CVB_Total.

Gene-based tests using MAGMA

We also conducted gene-based analyses using MAGMA, which aggregates SNP-level associations into a gene-level statistic. This approach considers the statistical significance of SNP-based test but does not explicitly consider how SNPs affect expression levels. The top 20 significant results from MAGMA are presented in Table 5 and full results are given in Table S4. We shall highlight several genes within the top-10 list here.

The most significant association was observed for *KCNC1* (potassium voltage-gated channel subfamily C member 1) with PureC_total (Z=6.03, FDR corrected p=1.49E-5). For English-related phenotypes, the most significant association was identified for gene *CATSPERD* (cation channel sperm associated auxiliary subunit delta) with EWR_Total (Z = 5.1632; FDR corrected p = 2.22E-03); the same gene was also associated with EVB_Total (Z = 5.0327; FDR corrected p = 4.40E-03). Two genes showed associations with EIS_Total, namely *SLC2A12* (solute carrier family 2 member 12;

Z = 5.1581; FDR corrected *p* = 2.27E-03) and *RSPH1* (radial spoke head component 1; Z = 5.009; FDR-corrected *p* = 2.49E-03).

As for Chinese literacy skills, *GTF3C1* (general transcription factor IIIC subunit 1) was associated with CVD_Total (Z = 5.405; FDR corrected *p* = 5.90E-04) and CVK_Total (Z=5.09 ; FDR corrected *p* =3.03E-3); *MAPK10* (mitogen-activated protein kinase 10) was associated with CVB_Total (Z = 5.0938; FDR corrected *p* = 3.20E-03). As for morphosyntactic skills in Chinese, the genes *SMKR1*(small lysine rich protein 1; Z = 5.0475; FDR corrected *p* = 3.25E-03) and *RFX8* (regulatory factor X8; Z = 4.9575; FDR corrected *p* = 3.25E-03) were associated with MS_Total.

Pathway enrichment analysis

To reveal relevant functional pathways, we conducted a self-contained gene-set analysis in GAUSS, testing 10679 canonical pathway and gene ontology (GO) gene sets from the MSigDB database. Full results with FDR<0.2 are shown in Tables S5.1 and S5.2. Table 6 and Table 7 summarize the top 20 pathway and GO analyses results. We also present the top two pathways and GO terms enrichment for every trait in Tables S5.3 and S5.4.

In pathway-based enrichment analysis of Chinese comprehension skills, the strongest association was observed for WO_Total with the Reactome RNA polymerase III transcription pathway (FDR corrected *p* = 1.60E-04). The second most significant association was observed for EWR_Total with the 'Deregulation of CDK5 in Alzheimers Disease' pathway (BioCarta) (FDR corrected *p* = 1.62E-03). Other pathways with the top five included the P2Y receptors (associated with CVK_total) and kinesins pathways (associated with BDS_total). GAUSS has also identified a collection of corresponding core genes (CS) for each pathway (please refer to Table S5).

In gene ontology (GO) enrichment analysis, the most significant enrichment was observed between CDICT_Total and sphingolipid-medicates signaling pathway (FDR corrected *p* = 4.07E-05). Other GO gene-sets within the top 5 (with respect to lowest p-values) included glycerophospholipid catabolic process, proton-transporting V-type ATPase complex, alcohol transmembrane transporter activity and divalent inorganic anion homeostasis. They were associated with PureC_total, CWR_norm, RC_MC and PureC_total, respectively. With regards to English literacy skills, we found that the GO gene-set 'ATP hydrolysis coupled cation transmembrane transport' (FDR corrected *p* = 1.31E-02) was the strongest association (with EWR_total). GAUSS selected 14 core genes for the gene set, in which one of them, *BLOC1S4*, was individually and significantly associated with EWR_Total (see TableS5.2).

Look-up of loci reported in two previous GWAS of dyslexia and literacy traits

Among the 42 (*p* < 5e-08) and 161 loci (*p* < 1e-06) previously reported for dyslexia [9] and literacy phenotypes [8] respectively, none showed genome-wide significance in our analyses. To evaluate the *aggregate* effects of previously reported genes, we also performed a gene-set analysis, in which we considered the genome-wide significant genes from the two previous GWAS [8 9] as candidate 'gene-sets'. An association was found between the literacy skills gene-set (derived from ref [8]) with CVD_Total and CVK_Total (FDR < 0.05). Full results are reported in Table S8.

Genetic overlap with neuropsychiatric phenotypes, cognitive performance (CP), and education attainment (EA)

Here we briefly describe several significant or suggestive findings (with FDR-corrected p<= 0.1) in PRS analysis. The ADHD polygenic score was associated with EVD_Total (FDR-corrected p= 0.026) at p-value threshold (pthres) = *1e-06*. We also observed significant associations with PRS of EA. The strongest association was observed at pthres = 1e-02 with EIS_total (FDR-corrected p = 0.0035); EA PRS was also associated with EDICT, EMA, EVA, EVB, EVK and EWR (FDR corrected p<0.1) at one or more pthres levels. All the associations were in the positive direction. The CP PRS was positively associated with CVD, EDC, EVD, EVK, EWR and MS with FDR-corrected p<0.1, at one or more pthres levels.

For other neuropsychiatric phenotypes, SCZ PRS was associated with BDS_Total across multiple p-value thresholds. Also, we observed an association of MDD PRS with CVD and CVK, and an association of BP PRS with DS at FDR<=0.1. We did not observe significant associations otherwise and no associations were observed with ASD PRS.

We present in Figure 1 the results of a polygenic score analysis on different traits at the best pthres cutoff. Full PRS results across all pthres are reported in Table S6.

Discussion

Overview

In this study, we attempted to uncover the genetic basis of a comprehensive range of cognitive, literacy-related, and language-related phenotypes in Chinese and English. To gain robust insights into the genetic architecture of the above phenotypes, we carried out a GWAS within a group of Hong Kong children, mainly developing twins with Cantonese as their native language and English as their second language. To the best of our knowledge, this is the first GWAS to explore the genetic basis of a comprehensive set of literacy- and language-related traits in both Chinese and English in a Chinese population. Compared to the previous GWAS on language traits (see introduction), this study also covers the widest range of phenotypes, enabling a finer resolution into the genetic architecture of language abilities.

One distinct feature of this study is that we selected the subjects drawn from a large longitudinal project in Hong Kong, a city with a unique linguistic background due to its geographical location and political history [45]. As such, our study represents the first attempt to assess the genetics of language and literacy skills of bilingual (Chinese and English) children systematically.

Genes associated with literacy- or language-related phenotypes

For English literacy skills, the most significant association was observed for a SNP close to *MANEA* and *MANEA-AS1* (rs6905617) with English lexical decision. Interestingly, by a search of the GWAS catalog, we found that a variant in *MANEA* showed tentative association with general cognitive ability in a previous GWAS (*p*=5e-6) [46] ; genetic variants in *MANEA-AS1* may also be associated with behavioral inhibition [47]. Regarding Chinese literacy skills, the SNP rs16848469 located in *TNR* (tenascin-R) was associated with CVD_Total and CVK_Total. *TNR* is primarily expressed in the central nervous system and plays a key role in human brain development by its involvement in axon growth and path finding[48]. Interestingly, variants in *TNR* have been reported to be associated with cognitive performance [49] [46]. In

addition, a recent study [50] revealed that *TNR* showed a significant interaction (at p<5e-8) with total testosterone in affecting fluid intelligence in healthy adults. In addition, a recent GWAS on ADHD identified a variant in *TNR* as the top association achieving genome-wide significance [51]. There were also case reports that deletions in *TNR* may be associated with intellectual disability and neurodevelopmental disorder [52] [53]. Moreover, animal studies showed that TNR-deficient mice had severe memory and coordination deficits [54].

In addition, we also observed that the SNP rs182977703 was associated with COM_Norm and CDICT_Total, which lies within the gene *SHTN1* (Shootin 1). The gene is involved in the generation of internal asymmetric signals necessary for neuronal polarization [55] . Increased expression of *SHTN1* in hippocampal neurons may result in its accumulation in multiple neurites and formation of surplus axons[56]. Clinically, neurodevelopmental disorders such as intellectual disability (ID) may result from defects in neuronal polarity and migration [57] [58] . For example, the association of Shootin 1 with ID was supported by a whole-genome transcriptome analysis on ID patients [59]. Another gene of interest in *PLXNC1*; variants in this gene have been reported to be associated (at p<1e-5) with multiple neuropsychiatric phenotypes such as major depression [60], Lewy body dementia [61], brain shape (segment 15 and 79) [62] and neuroticism[63].

Several gene-based tests reached a significance level after FDR correction for reading and spelling measures. The most significant gene from MAGMA was *KCNC1*, which encodes a subunit of the KV3 voltage-gated K^+ channels. Mutations in this gene were associated with a range of neurological disorders including epilepsy and also intellectual disability and cognitive decline in some patients [64] [65] [66]. In terms of Chinese literacy skills, the most significant association signal was observed for gene *GTF3C1* (General Transcription Factor lllC Subunit 1) with CVD_Total. *GTF3C1* has been widely investigated on its interactive connections to other genes; for example, it is involved in networks pathologically related to neurodegenerative and Alzheimer's disease [67] [68] [69]. It has been shown that *GTF3C1* is involved in regulation of rearrangement of neuronal nuclear architecture following neuronal excitation [70]. Of note, the nuclear architecture plays an important role in neural development and function [71]. *CHL1* was another gene implicated from S-PrediXcan analysis, and variants in this gene have been reported to be associated with education attainment [49] and also mathematics abilities [49].

In addition, our results showed that *SLC2A12* appears to be associated with English comprehension skills. *SLC2A12* encodes GLUT12, a glucose transporter. It has been reported that amyloid-beta increases GLUT12 protein expression in the brain in mouse models, implicating an important role of this transporter in Alzheimer disease [72] and cognitive functioning.

Polygenic score analysis

We discovered that some language traits were associated with PRS of psychiatric disorders, cognitive performance and educational attainment. We found that, for example, PRS for educational attainment and cognitive performance (derived from external GWAS data) were associated with various literacy traits. Our results were consistent with previous studies that have demonstrated shared genetic factors among childhood intelligence, educational attainment, and literacy skills.

For example, Luciano et al. (2017) [73] showed that PRS of word reading, general reading and spelling, as well as non-word repetition, were positively associated with educational attainment (college/university degree versus none), income and verbal-numerical cognitive test results. Moreover, in a GWAS by Price et al.,[74] substantial genetic overlap was found between word reading and number of years of education ($R^2 = 0.07$, $P = 4.91 \times 10^{-48}$) and intelligence score ($R^2 = 0.18$, $P = 7.25 \times 10^{-181}$) in a population-based sample. In another recent study by Gialluisi et al. [14], risk of developmental dyslexia was also significantly associated with PRS of EA and intelligence. Combined with our current findings, these results provide evidence to support a partially shared genetic etiology among literacy skills, cognitive measures, and educational outcomes.

Strengths and limitations

There are several strengths of our study. First of all, to the best of our knowledge, this is the first GWAS to investigate the genetic basis of a wide range of both Chinese and English literacy- and language-related skills in a Chinese population. Importantly, as reading and language comprehension are highly complex traits, here we performed detailed phenotyping to decipher the genetic basis of various different domains of these skills. On the other hand, previous studies largely followed another research strategy by focusing on a limited range of language phenotypes or binary outcomes. While it is also possible to only focus on a few selected phenotypes, for example, those with higher heritability (or by other criteria), such choice of phenotype may inevitably be somewhat arbitrary, and one may still discover variants of biological relevance for a trait with lower heritability. In addition, the SNP-based heritability, or the extent to which common variants contribute to a trait, is unknown for most phenotypes studied here. To enable a more comprehensive and unbiased examination of the genetic architecture of language/literacy-related traits, we have included a wide range of phenotypes in the current study. We also employed the FDR approach to account for multiple testing.

To gain deeper insights into the biological basis of the studied traits, we not only performed standard SNP-based tests but also gene-based (MAGMA, S-PrediXcan, S-MulTiXcan) and pathway-based analysis (GAUSS). This 'multi-level' approach helps to bridge the gap between SNP associations and biological mechanisms, thus enhancing our knowledge and understanding of reading and language. In addition to studying the associations between phenotypes and genetic factors, we performed PRS analysis to study the overlap of included phenotypes with other neuropsychiatric traits, which could provide insight into the genetic architecture of language-related traits.

Our study also has a few limitations. Our study is based on a Hong Kong Chinese sample (under a bilingual environment). It remains uncertain whether the genetic findings from the current study can be generalized to other populations. Further studies in other populations with different genetic and language backgrounds may be warranted. In a similar vein, the GWAS summary statistics of CP, EA and other psychiatric disorders were primarily derived from Europeans (due to lack of relevant data from Chinese populations), which may also attenuate the genetic overlap with the studied phenotypes in a Chinese population. Nevertheless, some studies (on other complex traits) have shown that genetic variants and PRS from Europeans may still be transferrable to Chinese [75] [76]. Also, here we employed the 1000-Genomes as reference for imputation, following the findings from Lin et al. [26] that satisfactory imputation performance in Chinese was achieved using the 1000G panel. In Lin et al.'s report, the mean imputation r^2 in two Chinese cohorts were at or above ~0.7 with MAF>1%, and were even better for higher MAF. At the time of this analysis, most established imputation servers (e.g. Michigan Imputation Server) does not contain Chinese-specific reference panels. Note that we

also reported the imputation quality score (r^2) for all reported variants for easy reference, and have removed variants with low imputation quality (r^2<0.3).

In this study, we performed extensive and deep phenotyping covering most domains of Chinese and English literacy- and language-related skills. This GWAS covers the widest range of language phenotypes to date. However, admittedly, our sample size is relatively moderate and statistical power may be insufficient to detect variants of small effects. In addition, given that we only performed genetic analysis in a single sample and some phenotypes were studied for the first time (e.g. most phenotypes on Chinese reading/comprehension), we emphasize that further replications in other samples are required. As for the genetic analyses, this study focused on the contribution of common variants; rare variant association was not our focus and may require further sequencing studies. In addition, while we have performed further gene-based and pathway-based bioinformatics analyses, the findings are based on statistical associations and will require further experimental validations.

In summary, we have conducted the first GWAS on a comprehensive range of phenotypes on both Chinese and English abilities in a HK Chinese population. We discovered a number of novel genetic loci that may underlie these traits, and revealed genes and pathways that may be associated. We believe our work will be an important starting point and reference for further studies into the biological and genetic basis of language abilities, and ultimately such knowledge will be useful for the development of better treatment for children with specific reading disabilities.

Figure Legends

Figure 1

Results of polygenic risk score (PRS) analysis on the 34 language-related phenotypes analyzed in this study, with PRS constructed from external GWAS data of different neuropsychiatric disorders/traits (training set). The following neuropsychiatric disorders/traits were included: attention deficit hyperactivity disorder (ADHD), autism spectrum disorders (ASD), Education attainment (EA), cognitive performance (CP), schizophrenia (SCZ), bipolar disorder (BP) and major depressive disorder (MDD).

In the heatmap, for each PRS analysis, we select the result with the lowest FDR-adjusted p-value (p.adjust), and show the regression coefficient in the graph.

PT: the optimal p-value threshold at which the most significant association was observed

Acknowledgements

This study was partially supported by a Collaborative Research Fund (CRF) (C4054-17W) from the Research Grants Council. HCS was also supported by the Lo Kwee Seong Biomedical Research Fund.

Competing interests

The authors have declared no competing interest.

Supplementary Information

Supplementary information is available at the journal's website and at https://drive.google.com/drive/folders/1AFZZrGI5zmijUo8M6KEtENC_P-ziMzIp?usp=sharing

References

1. Schelbe L, Pryce J, Petscher Y, et al. Dyslexia in the Context of Social Work: Screening and Early Intervention: *https://doi.org/101177/10443894211042323*. Published online October 25, 2021. doi:10.1177/10443894211042323

2. Cui J, Georgiou GK, Zhang Y, Li Y, Shu H, Zhou X. Examining the relationship between rapid automatized naming and arithmetic fluency in Chinese kindergarten children. *J Exp Child Psychol*. 2017;154:146-163. doi:10.1016/J.JECP.2016.10.008

3. Haworth CMA, Meaburn EL, Harlaar N, Plomin R. Reading and Generalist Genes. *Mind, Brain Educ*. 2007;1(4):173. doi:10.1111/J.1751-228X.2007.00018.X

4. R P, Y K. Generalist genes and learning disabilities. *Psychol Bull*. 2005;131(4):592-617. doi:10.1037/0033-2909.131.4.592

5. Andreola C, Mascheretti S, Belotti R, et al. The heritability of reading and reading-related neurocognitive components: A multi-level meta-analysis. *Neurosci Biobehav Rev*. 2021;121:175-200. doi:10.1016/J.NEUBIOREV.2020.11.016

6. Barbeira A, Shah KP, Torres JM, et al. MetaXcan: Summary Statistics Based Gene-Level Association Method Infers Accurate PrediXcan Results. doi:10.1101/045260

7. Erbeli F, Rice M, Paracchini S. Insights into Dyslexia Genetics Research from the Last Two Decades. *Brain Sci 2022, Vol 12, Page 27*. 2021;12(1):27. doi:10.3390/BRAINSCI12010027

8. Eising E, Mirza-Schreiber N, Zeeuw EL de, et al. Genome-wide association analyses of individual differences in quantitatively assessed reading- and language-related skills in up to 34,000 people. *bioRxiv*. 2021;16:2021.11.04.466897. doi:10.1101/2021.11.04.466897

9. Doust C, Fontanillas P, Eising E, et al. Discovery of 42 Genome-Wide Significant Loci Associated with Dyslexia. *medRxiv*. Published online August 22, 2021:2021.08.20.21262334. doi:10.1101/2021.08.20.21262334

10. Meaburn EL, Harlaar N, Craig IW, Schalkwyk LC, Plomin R. Quantitative trait locus association scan of early reading disability and ability using pooled DNA and 100K SNP microarrays in a sample of 5760 children. *Mol Psychiatry 2008 137*. 2007;13(7):729-740. doi:10.1038/sj.mp.4002063

11. Field LL, Shumansky K, Ryan J, Truong D, Swiergala E, Kaplan BJ. Dense-map genome scan for dyslexia supports loci at 4q13, 16p12, 17q22; suggests novel locus at 7q36. *Genes, Brain Behav*. 2013;12(1):56-69. doi:10.1111/GBB.12003

12. Eicher JD, Powers NR, Miller LL, et al. Genome-wide association study of shared components of reading disability and language impairment. *Genes Brain Behav*. 2013;12(8):792. doi:10.1111/GBB.12085

13. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol*. 2015;11(4):1004219. doi:10.1371/journal.pcbi.1004219

14. Gialluisi A, Andlauer TFM, Mirza-Schreiber N, et al. Genome-wide association study reveals new insights into the heritability and genetic correlates of developmental dyslexia. *Mol Psychiatry*. Published online October 14, 2020:1-14. doi:10.1038/s41380-020-00898-x

15. Gialluisi A, Newbury DF, Wilcutt EG, et al. Genome-wide screening for DNA variants associated with reading and language traits. *Genes Brain Behav*. 2014;13(7):686. doi:10.1111/GBB.12158

16. Truong DT, Adams AK, Paniagua S, et al. Original article: Multivariate genome-wide association study of rapid automatised naming and rapid alternating stimulus in Hispanic American and African–American youth. *J Med Genet*. 2019;56(8):557. doi:10.1136/JMEDGENET-2018-105874

17. Luciano M, Evans DM, Hansell NK, et al. A genome-wide association study for reading and language abilities in two population cohorts. *Genes Brain Behav*. 2013;12(6):645. doi:10.1111/GBB.12053

18. Price KM, Wigg KG, Feng Y, et al. Genome-wide association study of word reading: Overlap with risk genes for neurodevelopmental disorders. *Genes, Brain Behav*. 2020;19(6):e12648. doi:10.1111/GBB.12648

19. Gialluisi A, Andlauer TFM, Mirza-Schreiber N, et al. Genome-wide association scan identifies new variants associated with a cognitive predictor of dyslexia. *Transl Psychiatry 2019 91*. 2019;9(1):1-15. doi:10.1038/s41398-019-0402-0

20. Gialluisi A, Newbury DF, Wilcutt EG, et al. Genome-wide screening for DNA variants associated with reading and language traits. *Genes, Brain Behav*. 2014;13(7):686-701. doi:10.1111/GBB.12158

21. Lancaster HS, Dinu V, Li J, Gruen JR, Consortium TGr. Genetic and Demographic Predictors of Latent Reading Ability in Two Cohorts. *medRxiv*. Published online August 31, 2021:2021.08.24.21262573. doi:10.1101/2021.08.24.21262573

22. Doust C, Gordon SD, Garden N, et al. The Association of Dyslexia and Developmental Speech and Language Disorder Candidate Genes with Reading and Language Abilities in Adults. *Twin Res Hum Genet*. 2020;23(1):23-32. doi:10.1017/THG.2020.7

23. Yang MJ, Tzeng CH, Tseng JY, Huang CY. Determination of twin zygosity using a commercially available STR analysis of 15 unlinked loci and the gender-determining marker amelogenin - A preliminary report. *Hum Reprod*. 2006;21(8). doi:10.1093/humrep/del133

24. Minic\ua CC, Boomsma DI, Vink JM, Dolan C V. MZ twin pairs or MZ singletons in population family-based GWAS? More power in pairs. *Mol Psychiatry*. 2014;19(11):1154-1155.

25. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284-1287. doi:10.1038/ng.3656

26. Lin Y, Liu L, Yang S, et al. Genotype imputation for Han Chinese population using Haplotype Reference Consortium as reference. *Hum Genet*. 2018;137(6-7):431-436. doi:10.1007/S00439-018-1894-Z/FIGURES/3

27. So HC, Sham PC. Multiple testing and power calculations in genetic association studies. *Cold Spring Harb Protoc*. 2011;2011(1). doi:10.1101/PDB.TOP95

28. Schmidt AF, Finan C. Linear regression and the normality assumption. *J Clin Epidemiol*. 2018;98:146-151. doi:10.1016/J.JCLINEPI.2017.12.006

29. Beasley TM, Erickson S, Allison DB. Rank-based inverse normal transformations are increasingly used, but are they merited? *Behav Genet*. 2009;39(5):580-595. doi:10.1007/S10519-009-9281-0

30. McCaw ZR, Lane JM, Saxena R, Redline S, Lin X. Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. *Biometrics*. 2020;76(4):1262-1272. doi:10.1111/BIOM.13214

31. Wang L, Jia P, Wolfinger RD, Chen X, Zhao Z. Gene set analysis of genome-wide association studies: Methodological issues and perspectives. *Genomics*. 2011;98(1):1-8. doi:10.1016/j.ygeno.2011.04.006

32. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B*. 1995;57(1):289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

33. Dutta D, VandeHaar P, Scott LJ, Boehnke M, Lee S. A powerful subset-based gene-set analysis method identifies novel associations and improves interpretation in UK Biobank. *bioRxiv*. Published online 2019. doi:10.1101/799791

34. Liberzon A. A description of the molecular signatures database (MSigDB) web site. *Methods Mol Biol*. 2014;1150:153-160. doi:10.1007/978-1-4939-0512-6_9

35. Wainberg M, Sinnott-Armstrong N, Mancuso N, et al. Opportunities and challenges for transcriptome-wide association studies. *Nat Genet 2019 514*. 2019;51(4):592-599. doi:10.1038/s41588-019-0385-z

36. Barbeira AN, Pividori MD, Zheng J, Wheeler HE, Nicolae DL, Im HK. Integrating predicted transcriptome from multiple tissues improves association detection. *PLoS Genet*. 2019;15(1):e1007889. doi:10.1371/journal.pgen.1007889

37. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics*. 2015;31(9):1466-1468. doi:10.1093/bioinformatics/btu848

38. Lee JJ, Wedow R, Okbay A, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*. 2018;50(8):1112-1121. doi:10.1038/s41588-018-0147-3

39. Serdarevic F, Tiemeier H, Jansen PR, et al. Polygenic Risk Scores for Developmental Disorders, Neuromotor Functioning During Infancy, and Autistic Traits in Childhood. *Biol Psychiatry*. 2020;87(2):132-138. doi:10.1016/j.biopsych.2019.06.006

40. Martins-Silva T, Vaz J dos S, Hutz MH, et al. Assessing causality in the association between attention-deficit/hyperactivity disorder and obesity: a Mendelian randomization study. *Int J Obes*. 2019;43(12):2500-2508. doi:10.1038/s41366-019-0346-8

41. Pardiñas AF, Holmans P, Pocklington AJ, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet*. 2018;50(3):381-389. doi:10.1038/s41588-018-0059-2

42. Ruderfer DM, Ripke S, McQuillin A, et al. Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell*. 2018;173(7):1705-1715.e16. doi:10.1016/j.cell.2018.05.046

43. Sullivan PF, Daly MJ, Ripke S, et al. A mega-Analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry*. 2013;18(4):497-511. doi:10.1038/mp.2012.21

44. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 2012;44(7):821-824. doi:10.1038/ng.2310

45. L Wong SW, Suk-Han Ho C, McBride C, Wing-Yin Chow B, Miu Yee Waye M. Less is More in Hong Kong: Investigation of Biscriptal and Trilingual Development Among Chinese Twins in a (Relatively) Small City. *Twin Res Hum Genet*. 2021;20:2016. doi:10.1017/thg.2016.90

46. Davies G, Lam M, Harris SE, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun*. 2018;9(1). doi:10.1038/S41467-018-04362-X

47. McGue M, Zhang Y, Miller MB, et al. A genome-wide association study of behavioral disinhibition. *Behav Genet*. 2013;43(5):363-373. doi:10.1007/S10519-013-9606-X

48. Leprini A, Gherzi R, Siri A, Querzé G, Viti F, Zardi L. The Human Tenascin-R Gene *. *J Biol Chem*. 1996;271(49):31251-31254. doi:10.1074/JBC.271.49.31251

49. Lee JJ, Wedow R, Okbay A, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*. 2018;50(8):1112-1121. doi:10.1038/s41588-018-0147-3

50. Liang X, Cheng SQ, Ye J, et al. Evaluating the genetic effects of sex hormone traits on the development of mental traits: a polygenic score analysis and gene-environment-wide interaction study in UK Biobank cohort. *Mol Brain*. 2021;14(1). doi:10.1186/s13041-020-00718-x

51. Hawi Z, Yates H, Pinar A, et al. A case–control genome-wide association study of ADHD discovers a novel association with the tenascin R (TNR) gene. *Transl Psychiatry 2018 81*. 2018;8(1):1-8. doi:10.1038/s41398-018-0329-x

52. Dufresne D, Hamdan FF, Rosenfeld JA, et al. Homozygous deletion of Tenascin-R in a patient with intellectual disability. *J Med Genet*. 2012;49(7):451-454. doi:10.1136/JMEDGENET-2012-100831

53. Wagner M, Lévy J, Jung-Klawitter S, et al. Loss of TNR causes a nonprogressive neurodevelopmental disorder with spasticity and transient opisthotonus. *Genet Med 2020 226*. 2020;22(6):1061-1068. doi:10.1038/s41436-020-0768-7

54. Montag-Sallaz M, Montag D. Severe cognitive and motor coordination deficits </br>in Tenascin-R-deficient mice. *Genes, Brain Behav*. 2003;2(1):20-31. doi:10.1034/J.1601-183X.2003.00003.X

55. Toriyama M, Shimada T, Kim KB, et al. Shootin1: a protein involved in the organization of an asymmetric signal for neuronal polarization. *J Cell Biol*. 2006;175(1):147. doi:10.1083/JCB.200604160

56. Qiu J. Shootin 1 for the axon. *Nat Rev Neurosci 2006 712*. 2006;7(12):906-906. doi:10.1038/nrn2045

57. de la Torre-Ubieta L, Bonni A. Transcriptional Regulation of Neuronal Polarity and Morphogenesis in the Mammalian Brain. *Neuron*. 2011;72(1):22-40. doi:10.1016/J.NEURON.2011.09.018

58. Hakanen J, Ruiz-Reig N, Tissir F. Linking cell polarity to cortical development and malformations. *Front Cell Neurosci*. 2019;13:244. doi:10.3389/FNCEL.2019.00244/BIBTEX

59. InanlooRahatloo K, Peymani F, Kahrizi K, Najmabadi H. Whole-Transcriptome Analysis Reveals Dysregulation of Actin-Cytoskeleton Pathway in Intellectual Disability Patients. *Neuroscience*. 2019;404:423-444. doi:10.1016/J.NEUROSCIENCE.2019.01.029

60. Hall LS, Adams MJ, Arnau-Soler A, et al. Genome-wide meta-analyses of stratified depression in Generation Scotland and UK Biobank. *Transl Psychiatry 2017 81*. 2018;8(1):1-12. doi:10.1038/s41398-017-0034-1

61. Chibnik LB, White CC, Mukherjee S, et al. Susceptibility to neurofibrillary tangles: role of the PTPRD locus and limited pleiotropy with other neuropathologies. *Mol Psychiatry 2018 236*. 2017;23(6):1521-1529. doi:10.1038/mp.2017.20

62. Naqvi S, Sleyp Y, Hoskens H, et al. Shared heritability of human face and brain shape. *Nat Genet 2021 536*. 2021;53(6):830-839. doi:10.1038/s41588-021-00827-w

63. Wendt FR, Pathak GA, Lencz T, Krystal JH, Gelernter J, Polimanti R. Multivariate genome-wide analysis of education, socioeconomic status and brain phenome. *Nat Hum Behav 2020 54*. 2020;5(4):482-496. doi:10.1038/s41562-020-00980-y

64. Muona M, Berkovic SF, Dibbens LM, et al. A recurrent de novo mutation in KCNC1 causes progressive myoclonus epilepsy. *Nat Genet*. 2015;47(1):39. doi:10.1038/NG.3144

65. Park J, Koko M, Hedrich UBS, et al. KCNC1-related disorders: new de novo variants expand the phenotypic spectrum. *Ann Clin Transl Neurol*. 2019;6(7):1319. doi:10.1002/ACN3.50799

66. Poirier K, Viot G, Lombardi L, Jauny C, Billuart P, Bienvenu T. Loss of Function of KCNC1 is associated with intellectual disability without seizures. *Eur J Hum Genet 2017 255*. 2017;25(5):560-564. doi:10.1038/ejhg.2017.3

67. Recabarren D, Alarcón M. Gene networks in neurodegenerative disorders. *Life Sci*. 2017;183:83-97. doi:10.1016/j.lfs.2017.06.009

68. Ray M, Ruan J, Zhang W. Variations in the transcriptome of Alzheimer's disease reveal molecular networks involved in cardiovascular diseases. *Genome Biol*. 2008;9(10):R148. doi:10.1186/GB-2008-9-10-R148

69. Korbolina EE, Ershov NI, Bryzgalov LO, Kolosova NG. Application of quantitative trait locus mapping and transcriptomics to studies of the senescence-accelerated phenotype in rats. *BMC Genomics 2014 1512*. 2014;15(12):1-17. doi:10.1186/1471-2164-15-S12-S3

70. Crepaldi L, Policarpi C, Coatti A, et al. Binding of TFIIIC to SINE Elements Controls the Relocation of Activity-Dependent Neuronal Genes to Transcription Factories. *PLOS Genet*. 2013;9(8):e1003699. doi:10.1371/JOURNAL.PGEN.1003699

71. Alexander JM, Lomvardas S. Nuclear architecture as an epigenetic regulator of neural development and function. *Neuroscience*. 2014;264:39-50. doi:10.1016/J.NEUROSCIENCE.2014.01.044

72. Gil-Iturbe E, Solas M, Cuadrado-Tejedo M, et al. GLUT12 Expression in Brain of Mouse Models of Alzheimer's Disease. *Mol Neurobiol 2019 572*. 2019;57(2):798-805. doi:10.1007/S12035-019-01743-1

73. Luciano M, Hagenaars SP, Cox SR, et al. Single Nucleotide Polymorphisms Associated with Reading Ability Show Connection to Socio-Economic Outcomes. *Behav Genet*. 2017;47(5):469-479. doi:10.1007/s10519-017-9859-x

74. Price KM, Wigg KG, Feng Y, et al. Genome-wide association study of word reading: Overlap with risk genes for neurodevelopmental disorders. *Genes, Brain Behav*. 2020;19(6):e12648. doi:10.1111/gbb.12648

75. Ho WK, Tan MM, Mavaddat N, et al. European polygenic risk score for prediction of breast cancer shows similar performance in Asian women. *Nat Commun 2020 111*. 2020;11(1):1-11. doi:10.1038/s41467-020-17680-w

76. Lam M, Chen CY, Li Z, et al. Comparative genetic architectures of schizophrenia in East Asian andEuropean populations. *Nat Genet*. 2019;51(12):1670. doi:10.1038/S41588-019-0512-X

Tables

Table 1. Overview of phenotypes included in the study

Variable	Variable Label
BDS_Total	Backward Digit Span
CCR_Total	Chinese Character Reading
CDC_Total	Chinese Delayed Copying
CDICT_Total	Chinese Dictation
CDRAN_Mean	Chinese Digit Rapid Naming
CLD_Total	Chinese Lexical Decision
COM_Score	Chinese 1 Min Word Reading Adjusted Total Score
COM_Norm	Chinese 1 Min Word Reading Scaled Score
CVA_Total	Chinese Vocabulary - Receptive Vocabulary (10 items)
CVB_Total	Chinese Vocabulary - Expressive Vocabulary (12 items)
CVD_Total	Chinese Vocabulary - Vocabulary Definition (26 items)
CVK_Total	Chinese Vocabulary Knowledge (48 items; sum of CVA, CVB and CVK)
CWR_Total	Chinese Word Reading Raw Score
CWR_Norm	Chinese Word Reading Scaled Score
DS_Total	Chinese Discourse Skills
EDC_Total	English Delayed Copying
EDICT_Total	English Dictation
EDRAN_Mean	English Digit Rapid Naming
EIS_Total	English Invented Spelling
ELD_Total	English Lexical Decision
ELRAN_Mean	English Letter Rapid Naming
EMA_Total	English Morphological Awareness - Written Test
EVA_Total	English Vocabulary - Receptive Vocabulary (15 items)
EVB_Total	English Vocabulary - Expressive Vocabulary (15 items)
EVD_Total	English Vocabulary - Vocabulary Definition (15 items)
EVK_Total	English Vocabulary Knowledge (45 items; sum of EVA, EVB and EVK)
EWR_Total	English Word Reading Total Score
MS_Total	Morphosyntax
PairC_Total	Pair Cancellation
PureC_Total	Pure Copying of Unfamiliar Scripts
RC_MC	Reading Comprehension - Multiple Choice
RC_OE	Reading Comprehension - Open End
RC_Total	Reading Comprehension - Total
WO_Total	Chinese Word Order

Table 2. Results of the SNP-based association analysis

Phenotype	CHR	BP	SNP	A1	A2	P	MAF	Rsq	Genotyped	Closest gene	S0001	FDR-adjust *P*
ELD_Total	6	95643248	rs6905617	C	A	3.29E-09	0.352	0.52	Imputed	*MANEA-AS1(-364.7kb)*	43	3.19E-03
CVK_Total	1	175436068	rs77868538	T	C	3.33E-09	0.028	0.99	Imputed	*TNR(0)*	4	1.32E-02
CVD_Total	1	175436068	rs77868538	T	C	7.44E-09	0.028	0.99	Imputed	*TNR(0)*	4	2.87E-02
ELD_Total	1	238306117	rs370871011	T	G	9.34E-09	0.011	0.64	Imputed	*LOC100130331(+214.5kb)*	2	6.26E-03
CCR_Total	9	115640979	rs56024259	G	A	1.53E-08	0.124	0.98	Imputed	*SLC46A2(-0.22kb)*	7	3.97E-02
ELD_Total	5	32421181	rs925214	T	G	1.64E-08	0.030	0.86	Imputed	*ZFR(0)*	37	8.44E-03
CDRAN_Mean	12	94529190	rs3847795	A	C	1.73E-08	0.173	0.94	Imputed	*PLXNC1(-13.31kb)*	4	7.60E-02
ELD_Total	1	202430424	rs12049280	G	T	3.17E-08	0.021	0.66	Imputed	*PPP1R12B(0)*	5	1.52E-02
CWR_Total	4	57573275	rs4865143	T	C	3.61E-08	0.071	0.80	Imputed	*HOPX(+25.4kb)*	15	1.34E-01
COM_Norm	10	118659171	rs182977703	G	T	4.39E-08	0.016	0.60	Imputed	*SHTN1 (0)*	4	1.49E-01
CDICT_Total	10	118659171	rs182977703	G	T	4.44E-08	0.016	0.60	Imputed	*SHTN1 (0)*	2	1.52E-01
CVB_Total	4	57573275	rs4865143	T	C	4.97E-08	0.071	0.80	Imputed	*HOPX(+25.4kb)*	27	9.92E-02

For full results please refer to Table S1. Results are sorted by P-value. MAF, minor allele frequency; Rsq, R-squared (imputation quality metric); BP, base pair (position of the SNP); S0001, number of clumped SNPs (SNPs in LD) with p<1e-3. Only SNPs with S0001 >= 2 are shown. FDR-adjust P, false-discovery rate-adjusted P-value by the Benjamini-Hochberg method.

Table 3. Top 20 S-Predixcan results after correction of multiple testing

Phenotype[a]	Tissue_name	Gene	Zscore	*P*	FDR-adjust *P*[b]
EWR_Total	Brain_Amygdala	*DUS3L*	4.81	1.52E-06	4.18E-02
EWR_Total	Brain_Caudate_basal_ganglia	*DUS3L*	4.72	2.35E-06	4.18E-02
EWR_Total	Brain_Putamen_basal_ganglia	*DUS3L*	4.69	2.76E-06	4.18E-02
EWR_Total	Brain_Cerebellar_Hemisphere	*DUS3L*	4.6	4.20E-06	4.77E-02
EWR_Total	Brain_Hypothalamus	*AC005523.3*	4.37	1.23E-05	1.12E-01
EMA_Total	Brain_Frontal_Cortex_BA9	*ZNF585B*	-4.67	3.07E-06	1.30E-01
CVB_Total	Brain_Cerebellum	*BNIPL*	4.58	4.70E-06	2.13E-01
EWR_Total	Brain_Frontal_Cortex_BA9	*DUS3L*	4.13	3.60E-05	2.72E-01
RC_MC	Brain_Cortex	*RP11-508N22.12*	-4.52	6.18E-06	2.81E-01
EWR_Total	Brain_Nucleus_accumbens_basal_ganglia	*DUS3L*	3.99	6.58E-05	4.27E-01
EDC_Total	Brain_Cerebellum	*GTF3C5*	4.41	1.03E-05	4.66E-01
ELD_Total	Brain_Cerebellum	*FAM86B2*	-4.37	1.24E-05	5.62E-01
EMA_Total	Brain_Cerebellum	*KIAA0355*	4.12	3.80E-05	5.79E-01
EMA_Total	Brain_Substantia_nigra	*CHL1*	4.1	4.11E-05	5.79E-01
EMA_Total	Brain_Cerebellar_Hemisphere	*TSEN15*	-3.81	1.41E-04	7.48E-01
EMA_Total	Brain_Hippocampus	*HNRNPCP1*	-3.84	1.25E-04	7.48E-01
EMA_Total	Brain_Nucleus_accumbens_basal_ganglia	*RP11-521C20.2*	-3.92	8.98E-05	7.48E-01
EMA_Total	Brain_Putamen_basal_ganglia	*RASA4*	-3.91	9.22E-05	7.48E-01
EMA_Total	Brain_Spinal_cord_cervical_c-1	*C20orf202*	-3.84	1.22E-04	7.48E-01
EVA_Total	Brain_Amygdala	*RP11-178F10.3*	-3.94	8.18E-05	8.33E-01

[a] Please refer to Table 1 for abbreviations of the phenotype

[b] FDR-adjust *P* : Calculated by the R.program p.adjust using Benjamini-Hochberg procedure (BH).

Table 4. Top 20 S-Multixcan results after correction of multiple testing

Phenotype[a]	T_i_best[b]	Gene	P_i_Best[c]	FDR.adjust *P*[d]
EVA_Total	Brain_Cortex	*HSD3B7*	1.71E-03	1.62E-17
CVB_Total	Brain_Cerebellar_Hemisphere	*SEMA6C*	3.77E-04	2.85E-12
EVK_Total	Brain_Caudate_basal_ganglia	*HSD3B7*	6.02E-03	1.61E-11
EVB_Total	Brain_Cortex	*HSD3B7*	1.16E-02	2.50E-10
EWR_Total	Brain_Cortex	*HSD3B7*	1.40E-02	7.49E-08
ELRAN_Mean	Brain_Nucleus_accumbens_basal_ganglia	*BAK1P1*	3.91E-03	8.93E-08
EDICT_Total	Brain_Anterior_cingulate_cortex_BA24	*MYO6*	3.35E-04	9.85E-08
EVD_Total	Brain_Caudate_basal_ganglia	*HSD3B7*	9.38E-03	1.32E-07
CLD_Total	Brain_Caudate_basal_ganglia	*OXCT2P1*	3.87E-04	3.45E-07
COM_Score	Brain_Cerebellum	*RBM8A*	8.38E-02	4.40E-07
ELRAN_Mean	Brain_Nucleus_accumbens_basal_ganglia	*CYP2E1*	7.40E-03	1.79E-06
EVB_Total	Brain_Anterior_cingulate_cortex_BA24	*BCO2*	5.28E-04	6.25E-06
CDC_Total	Brain_Nucleus_accumbens_basal_ganglia	*ZNF565*	3.67E-02	6.48E-06
EWR_Total	Brain_Anterior_cingulate_cortex_BA24	*MYO6*	9.68E-04	8.21E-06
ELRAN_Mean	Brain_Nucleus_accumbens_basal_ganglia	*ZNF565*	2.24E-02	1.33E-05
EVA_Total	Brain_Caudate_basal_ganglia	*AC110781.3*	5.07E-02	1.47E-05
EIS_Total	Brain_Cerebellum	*EXOSC5*	7.48E-03	1.61E-05
MS_Total	Brain_Anterior_cingulate_cortex_BA24	*CTB-161M19.1*	3.66E-02	1.69E-05
COM_Norm	Brain_Cerebellum	*RBM8A*	1.50E-01	2.42E-05
EDICT_Total	Brain_Cortex	*HSD3B7*	4.66E-02	2.78E-05

[a] Please refer to Table 1 for abbreviations of the phenotype

[b] T_i_Best : name of best single-tissue S-Predixcan association.

[c] P_i_Best : best p-value of single tissue S-Predixcan association.

[d] FDR-adjust *P* : FDR-adjusted p-value of the *overall* p-value output by S-Multixcan. FDR was calculated by the R program p.adjust using the Benjamini-Hochberg procedure (BH).

Table 5. Top 20 gene-based results (Magma) after correction of multiple testing

Phenotype[a]	Description	Gene	CHR	ZSTAT	P	FDR.adjust *P*[b]
PureC_Total	potassium voltage-gated channel subfamily C member 1	*KCNC1*	11	6.03	8.18E-10	1.49E-05
CVD_Total	general transcription factor IIIC subunit 1	*GTF3C1*	16	5.41	3.24E-08	5.90E-04
EWR_Total	cation channel sperm associated auxiliary subunit delta	*CATSPERD*	19	5.16	1.22E-07	2.22E-03
EIS_Total	solute carrier family 2 member 12	*SLC2A12*	6	5.16	1.25E-07	2.27E-03
EIS_Total	radial spoke head component 1	*RSPH1*	21	5.01	2.74E-07	2.49E-03
CVB_Total	mitogen-activated protein kinase 10	*MAPK10*	4	5.09	1.76E-07	3.20E-03
MS_Total	regulatory factor X8	*RFX8*	2	4.96	3.57E-07	3.25E-03
MS_Total	small lysine rich protein 1	*SMKR1*	7	5.05	2.24E-07	3.25E-03
CVK_Total	general transcription factor IIIC subunit 1	*GTF3C1*	16	5.09	1.81E-07	3.30E-03
EVB_Total	cation channel sperm associated auxiliary subunit delta	*CATSPERD*	19	5.03	2.42E-07	4.40E-03
CVB_Total	BCL2 interacting protein like	*BNIPL*	1	4.86	5.87E-07	5.34E-03
EVB_Total	cilia and flagella associated protein 65	*CFAP65*	2	4.84	6.46E-07	5.89E-03
BDS_Total	transmembrane serine protease 13	*TMPRSS13*	11	4.96	3.48E-07	6.33E-03
EVK_Total	cilia and flagella associated protein 65	*CFAP65*	2	4.83	6.95E-07	1.27E-02
EWR_Total	caveolae associated protein 2	*CAVIN2*	2	4.46	4.19E-06	1.39E-02
EWR_Total	Morf4 family associated protein 1 like 1	*MRFAP1L1*	4	4.49	3.57E-06	1.39E-02
EWR_Total	biogenesis of lysosomal organelles complex 1 subunit 4	*BLOC1S4*	4	4.44	4.57E-06	1.39E-02
EWR_Total	proline rich 22	*PRR22*	19	4.54	2.81E-06	1.39E-02
EWR_Total	dihydrouridine synthase 3 like	*DUS3L*	19	4.57	2.43E-06	1.39E-02
EDRAN_Mean	ankyrin repeat domain 50	*ANKRD50*	4	4.80	7.76E-07	1.41E-02
CDC_Total	frizzled class receptor 7	*FZD7*	2	4.80	8.08E-07	1.47E-02
CDRAN_Mean	calcium and integrin binding family member 2	*CIB2*	15	4.79	8.25E-07	1.50E-02

[a] Please refer to Table 1 for abbreviations of the phenotype

[b] FDR-adjust *P*: Calculated by the R.program p.adjust using Benjamini-Hochberg procedure (BH).

Table 6. Top 20 GO enrichment results (GAUSS) after correction of multiple testing

GeneSet	Pvalue	Phenotype	FDR.adjust *P*[a]
GO_SPHINGOLIPID_MEDIATED_SIGNALING_PATHWAY	6.88E-09	CDICT_Total	4.07E-05
GO_GLYCEROPHOSPHOLIPID_CATABOLIC_PROCESS	6.40E-08	PureC_Total	3.78E-04
GO_PROTON_TRANSPORTING_V_TYPE_ATPASE_COMPLEX	1.20E-07	CWR_Norm	7.13E-04
GO_ALCOHOL_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	2.38E-07	RC_MC	1.41E-03
GO_DIVALENT_INORGANIC_ANION_HOMEOSTASIS	5.74E-07	PureC_Total	1.70E-03
GO_CELLULAR_ANION_HOMEOSTASIS	2.25E-06	PureC_Total	4.44E-03
GO_BIOACTIVE_LIPID_RECEPTOR_ACTIVITY	2.13E-06	CDICT_Total	6.29E-03
GO_ATP_HYDROLYSIS_COUPLED_TRANSMEMBRANE_TRANSPORT	2.22E-06	EWR_Total	1.31E-02
GO_LYMPHANGIOGENESIS	7.35E-06	CDICT_Total	1.45E-02
GO_ORGANIC_HYDROXY_COMPOUND_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	4.90E-06	RC_MC	1.45E-02
GO_POSITIVE_REGULATION_OF_VASODILATION	2.00E-05	PureC_Total	1.48E-02
GO_POSITIVE_REGULATION_OF_B_CELL_DIFFERENTIATION	2.00E-05	PureC_Total	1.48E-02
GO_POSITIVE_REGULATION_OF_BLOOD_CIRCULATION	2.00E-05	PureC_Total	1.48E-02
GO_NEURON_PROJECTION_GUIDANCE	2.00E-05	PureC_Total	1.48E-02
GO_POLYSACCHARIDE_BINDING	2.00E-05	PureC_Total	1.48E-02
GO_MONOVALENT_INORGANIC_ANION_HOMEOSTASIS	3.00E-05	PureC_Total	1.97E-02
GO_REGULATION_OF_MITOCHONDRIAL_FISSION	1.53E-05	CDICT_Total	2.27E-02
GO_RESPONSE_TO_NERVE_GROWTH_FACTOR	1.08E-05	EWR_Total	2.43E-02
GO_PROTON_TRANSPORTING_TWO_SECTOR_ATPASE_COMPLEX_CATALYTIC_DOMAIN	1.64E-05	EWR_Total	2.43E-02
GO_PROTON_TRANSPORTING_V_TYPE_ATPASE_COMPLEX	1.35E-05	EWR_Total	2.43E-02

Please refer to Table 1 for abbreviations of the phenotypes. Full descriptions of each gene-set can be found by looking up the pathway names at https://www.gsea-msigdb.org/gsea/msigdb/.

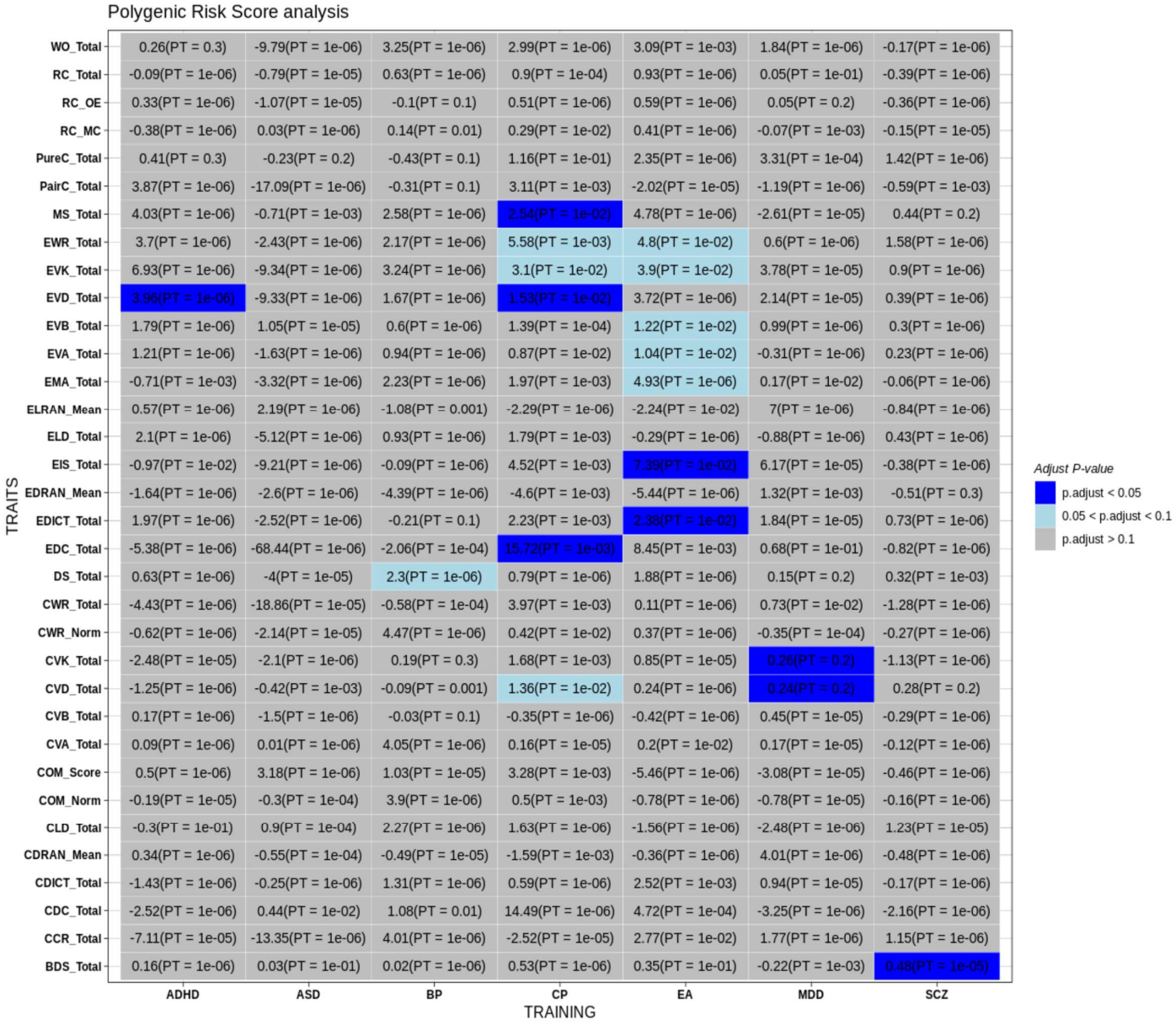
[a] FDR-adjust *P*: Calculated by the R.program p.adjust using Benjamini-Hochberg procedure (BH).

Table 7. Top 20 Pathway enrichment results (GAUSS) after correction of multiple testing

GeneSet	Pvalue	Phenotype	FDR adjust *P*[a]
REACTOME_RNA_POL_III_TRANSCRIPTION	3.36E-08	WO_Total	1.60E-04
BIOCARTA_P35ALZHEIMERS_PATHWAY	3.41E-07	EWR_Total	1.62E-03
REACTOME_P2Y_RECEPTORS	3.94E-07	CVK_Total	1.88E-03
REACTOME_KINESINS	7.07E-07	BDS_Total	3.37E-03
STOSSI_RESPONSE_TO_ESTRADIOL	3.04E-06	RC_MC	1.45E-02
IGLESIAS_E2F_TARGETS_DN	4.29E-06	CWR_Norm	2.04E-02
REACTOME_P2Y_RECEPTORS	5.25E-06	CVD_Total	2.50E-02
PID_S1P_META_PATHWAY	9.02E-06	CDICT_Total	3.88E-02
GOLUB_ALL_VS_AML_DN	1.63E-05	CDICT_Total	3.88E-02
BIOCARTA_AKAPCENTROSOME_PATHWAY	2.00E-05	CCR_Total	4.76E-02
BANDRES_RESPONSE_TO_CARMUSTIN_MGMT_48HR_UP	2.00E-05	CCR_Total	4.76E-02
LIM_MAMMARY_LUMINAL_PROGENITOR_UP	2.00E-05	EWR_Total	4.76E-02
BIOCARTA_BLYMPHOCYTE_PATHWAY	1.17E-05	CDC_Total	5.55E-02
BANDRES_RESPONSE_TO_CARMUSTIN_WITHOUT_MGMT_48HR_UP	4.00E-05	CCR_Total	6.35E-02
KEGG_PRIMARY_BILE_ACID_BIOSYNTHESIS	3.00E-05	EVB_Total	6.35E-02
RODRIGUES_NTN1_AND_DCC_TARGETS	4.00E-05	EVB_Total	6.35E-02
TERAMOTO_OPN_TARGETS_CLUSTER_7	4.00E-05	EVB_Total	6.35E-02
KEGG_PROGESTERONE_MEDIATED_OOCYTE_MATURATION	3.00E-05	CVA_Total	7.14E-02
DUTERTRE_ESTRADIOL_RESPONSE_6HR_UP	2.00E-05	CVA_Total	7.14E-02
SA_FAS_SIGNALING	9.00E-05	CDRAN_Mean	7.14E-02

Please refer to Table 1 for abbreviations of the phenotypes. Full descriptions of each gene-set can be found by looking up the pathway names at https://www.gsea-msigdb.org/gsea/msigdb/.

[a] FDR-adjust *P*: Calculated by the R.program p.adjust using Benjamini-Hochberg procedure (BH).



TRAITS	ADHD	ASD	BP	CP	EA	MDD	SCZ
WO_Total	0.26(PT = 0.3)	-9.79(PT = 1e-06)	3.25(PT = 1e-06)	2.99(PT = 1e-06)	3.09(PT = 1e-03)	1.84(PT = 1e-06)	-0.17(PT = 1e-06)
RC_Total	-0.09(PT = 1e-06)	-0.79(PT = 1e-05)	0.63(PT = 1e-06)	0.9(PT = 1e-04)	0.93(PT = 1e-06)	0.05(PT = 1e-01)	-0.39(PT = 1e-06)
RC_OE	0.33(PT = 1e-06)	-1.07(PT = 1e-05)	-0.1(PT = 0.1)	0.51(PT = 1e-06)	0.59(PT = 1e-06)	0.05(PT = 0.2)	-0.36(PT = 1e-06)
RC_MC	-0.38(PT = 1e-06)	0.03(PT = 1e-06)	0.14(PT = 0.01)	0.29(PT = 1e-02)	0.41(PT = 1e-06)	-0.07(PT = 1e-03)	-0.15(PT = 1e-05)
PureC_Total	0.41(PT = 0.3)	-0.23(PT = 0.2)	-0.43(PT = 0.1)	1.16(PT = 1e-01)	2.35(PT = 1e-06)	3.31(PT = 1e-04)	1.42(PT = 1e-06)
PairC_Total	3.87(PT = 1e-06)	-17.09(PT = 1e-06)	-0.31(PT = 0.1)	3.11(PT = 1e-03)	-2.02(PT = 1e-05)	-1.19(PT = 1e-06)	-0.59(PT = 1e-03)
MS_Total	4.03(PT = 1e-06)	-0.71(PT = 1e-03)	2.58(PT = 1e-06)	2.54(PT = 1e-02)	4.78(PT = 1e-06)	-2.61(PT = 1e-05)	0.44(PT = 0.2)
EWR_Total	3.7(PT = 1e-06)	-2.43(PT = 1e-06)	2.17(PT = 1e-06)	5.58(PT = 1e-03)	4.8(PT = 1e-02)	0.6(PT = 1e-06)	1.58(PT = 1e-06)
EVK_Total	6.93(PT = 1e-06)	-9.34(PT = 1e-06)	3.24(PT = 1e-06)	3.1(PT = 1e-02)	3.9(PT = 1e-02)	3.78(PT = 1e-05)	0.9(PT = 1e-06)
EVD_Total	3.96(PT = 1e-06)	-9.33(PT = 1e-06)	1.67(PT = 1e-06)	1.53(PT = 1e-02)	3.72(PT = 1e-06)	2.14(PT = 1e-05)	0.39(PT = 1e-06)
EVB_Total	1.79(PT = 1e-06)	1.05(PT = 1e-05)	0.6(PT = 1e-06)	1.39(PT = 1e-04)	1.22(PT = 1e-02)	0.99(PT = 1e-06)	0.3(PT = 1e-06)
EVA_Total	1.21(PT = 1e-06)	-1.63(PT = 1e-06)	0.94(PT = 1e-06)	0.87(PT = 1e-02)	1.04(PT = 1e-02)	-0.31(PT = 1e-06)	0.23(PT = 1e-06)
EMA_Total	-0.71(PT = 1e-03)	-3.32(PT = 1e-06)	2.23(PT = 1e-06)	1.97(PT = 1e-03)	4.93(PT = 1e-06)	0.17(PT = 1e-02)	-0.06(PT = 1e-06)
ELRAN_Mean	0.57(PT = 1e-06)	2.19(PT = 1e-06)	-1.08(PT = 0.001)	-2.29(PT = 1e-06)	-2.24(PT = 1e-02)	7(PT = 1e-06)	-0.84(PT = 1e-06)
ELD_Total	2.1(PT = 1e-06)	-5.12(PT = 1e-06)	0.93(PT = 1e-06)	1.79(PT = 1e-03)	-0.29(PT = 1e-06)	-0.88(PT = 1e-06)	0.43(PT = 1e-06)
ELS_Total	-0.97(PT = 1e-02)	-9.21(PT = 1e-06)	-0.09(PT = 1e-06)	4.52(PT = 1e-03)	7.39(PT = 1e-02)	6.17(PT = 1e-05)	-0.38(PT = 1e-06)
EDRAN_Mean	-1.64(PT = 1e-06)	-2.6(PT = 1e-06)	-4.39(PT = 1e-06)	-4.6(PT = 1e-03)	-5.44(PT = 1e-06)	1.32(PT = 1e-03)	-0.51(PT = 0.3)
EDICT_Total	1.97(PT = 1e-06)	-2.52(PT = 1e-06)	-0.21(PT = 0.1)	2.23(PT = 1e-03)	2.38(PT = 1e-02)	1.84(PT = 1e-05)	0.73(PT = 1e-06)
EDC_Total	-5.38(PT = 1e-06)	-68.44(PT = 1e-06)	-2.06(PT = 1e-04)	15.72(PT = 1e-03)	8.45(PT = 1e-03)	0.68(PT = 1e-01)	-0.82(PT = 1e-06)
DS_Total	0.63(PT = 1e-06)	-4(PT = 1e-05)	2.3(PT = 1e-06)	0.79(PT = 1e-06)	1.88(PT = 1e-06)	0.15(PT = 0.2)	0.32(PT = 1e-03)
CWR_Total	-4.43(PT = 1e-06)	-18.86(PT = 1e-05)	-0.58(PT = 1e-04)	3.97(PT = 1e-03)	0.11(PT = 1e-06)	0.73(PT = 1e-02)	-1.28(PT = 1e-06)
CWR_Norm	-0.62(PT = 1e-06)	-2.14(PT = 1e-05)	4.47(PT = 1e-06)	0.42(PT = 1e-02)	0.37(PT = 1e-06)	-0.35(PT = 1e-04)	-0.27(PT = 1e-06)
CVK_Total	-2.48(PT = 1e-05)	-2.1(PT = 1e-06)	0.19(PT = 0.3)	1.68(PT = 1e-03)	0.85(PT = 1e-05)	0.26(PT = 0.2)	-1.13(PT = 1e-06)
CVD_Total	-1.25(PT = 1e-06)	-0.42(PT = 1e-03)	-0.09(PT = 0.001)	1.36(PT = 1e-02)	0.24(PT = 1e-06)	0.24(PT = 0.2)	0.28(PT = 0.2)
CVB_Total	0.17(PT = 1e-06)	-1.5(PT = 1e-06)	-0.03(PT = 0.1)	-0.35(PT = 1e-06)	-0.42(PT = 1e-06)	0.45(PT = 1e-05)	-0.29(PT = 1e-06)
CVA_Total	0.09(PT = 1e-06)	0.01(PT = 1e-06)	4.05(PT = 1e-06)	0.16(PT = 1e-05)	0.2(PT = 1e-02)	0.17(PT = 1e-05)	-0.12(PT = 1e-06)
COM_Score	0.5(PT = 1e-06)	3.18(PT = 1e-06)	1.03(PT = 1e-05)	3.28(PT = 1e-03)	-5.46(PT = 1e-06)	-3.08(PT = 1e-05)	-0.46(PT = 1e-06)
COM_Norm	-0.19(PT = 1e-05)	-0.3(PT = 1e-04)	3.9(PT = 1e-06)	0.5(PT = 1e-03)	-0.78(PT = 1e-06)	-0.78(PT = 1e-05)	-0.16(PT = 1e-06)
CLD_Total	-0.3(PT = 1e-01)	0.9(PT = 1e-04)	2.27(PT = 1e-06)	1.63(PT = 1e-06)	-1.56(PT = 1e-06)	-2.48(PT = 1e-06)	1.23(PT = 1e-05)
CDRAN_Mean	0.34(PT = 1e-06)	-0.55(PT = 1e-04)	-0.49(PT = 1e-05)	-1.59(PT = 1e-03)	-0.36(PT = 1e-06)	4.01(PT = 1e-06)	-0.48(PT = 1e-06)
CDICT_Total	-1.43(PT = 1e-06)	-0.25(PT = 1e-06)	1.31(PT = 1e-06)	0.59(PT = 1e-06)	2.52(PT = 1e-03)	0.94(PT = 1e-05)	-0.17(PT = 1e-06)
CDC_Total	-2.52(PT = 1e-06)	0.44(PT = 1e-02)	1.08(PT = 0.01)	14.49(PT = 1e-06)	4.72(PT = 1e-04)	-3.25(PT = 1e-06)	-2.16(PT = 1e-06)
CCR_Total	-7.11(PT = 1e-05)	-13.35(PT = 1e-06)	4.01(PT = 1e-06)	-2.52(PT = 1e-05)	2.77(PT = 1e-02)	1.77(PT = 1e-06)	1.15(PT = 1e-06)
BDS_Total	0.16(PT = 1e-06)	0.03(PT = 1e-01)	0.02(PT = 1e-06)	0.53(PT = 1e-06)	0.35(PT = 1e-01)	-0.22(PT = 1e-03)	0.48(PT = 1e-05)