


BRAIN COMMUNICATIONS

The role of critical immune genes in brain disorders: insights from neuroimaging immunogenetics

 Beilei Bian,¹ Baptiste Couvy-Duchesne,^{1,2} Naomi R. Wray^{1,3} and Allan F. McRae¹

Genetic variants in the human leukocyte antigen and killer cell immunoglobulin-like receptor regions have been associated with many brain-related diseases, but how they shape brain structure and function remains unclear. To identify the genetic variants in *HLA* and *KIR* genes associated with human brain phenotypes, we performed a genetic association study of ~30 000 European unrelated individuals using brain MRI phenotypes generated by the UK Biobank (UKB). We identified 15 *HLA* alleles in *HLA* class I and class II genes significantly associated with at least one brain MRI-based phenotypes ($P < 5 \times 10^{-8}$). These associations converged on several main haplotypes within the *HLA*. In particular, the human leukocyte antigen alleles within an ancestral haplotype 8.1 were associated with multiple MRI measures, including grey matter volume, cortical thickness (TH) and diffusion MRI (dMRI) metrics. These alleles have been strongly associated with schizophrenia. Additionally, associations were identified between *HLA-DRB1*04~DQA1*03:01~DQB1*03:02* and isotropic volume fraction of diffusion MRI in multiple white matter tracts. This haplotype has been reported to be associated with Parkinson's disease. These findings suggest shared genetic associations between brain MRI biomarkers and brain-related diseases. Additionally, we identified 169 associations between the complement component 4 (*C4*) gene and imaging phenotypes. We found that *C4* gene copy number was associated with cortical TH and dMRI metrics. No *KIR* gene copy numbers were associated with image-derived phenotypes at genome-wide threshold. To address the multiple testing burden in the phenome-wide association study, we performed a multi-trait association analysis using trait-based association test that uses extended Simes procedure and identified MRI image-specific associations. This study contributes to insight into how critical immune genes affect brain-related traits as well as the development of neurological and neuropsychiatric disorders.

1 Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD 4072, Australia

2 Paris Brain Institute, CNRS, INRIA, Paris, France

3 Queensland Brain Institute, The University of Queensland, Brisbane, QLD, Australia

Correspondence to: Allan F. McRae

The University of Queensland

Brisbane,

QLD 4072, Australia

E-mail: a.mcrae@uq.edu.au

Keywords: human leukocyte antigen; complement component 4; killer cell immunoglobulin-like receptor; neuroimaging; genetic association

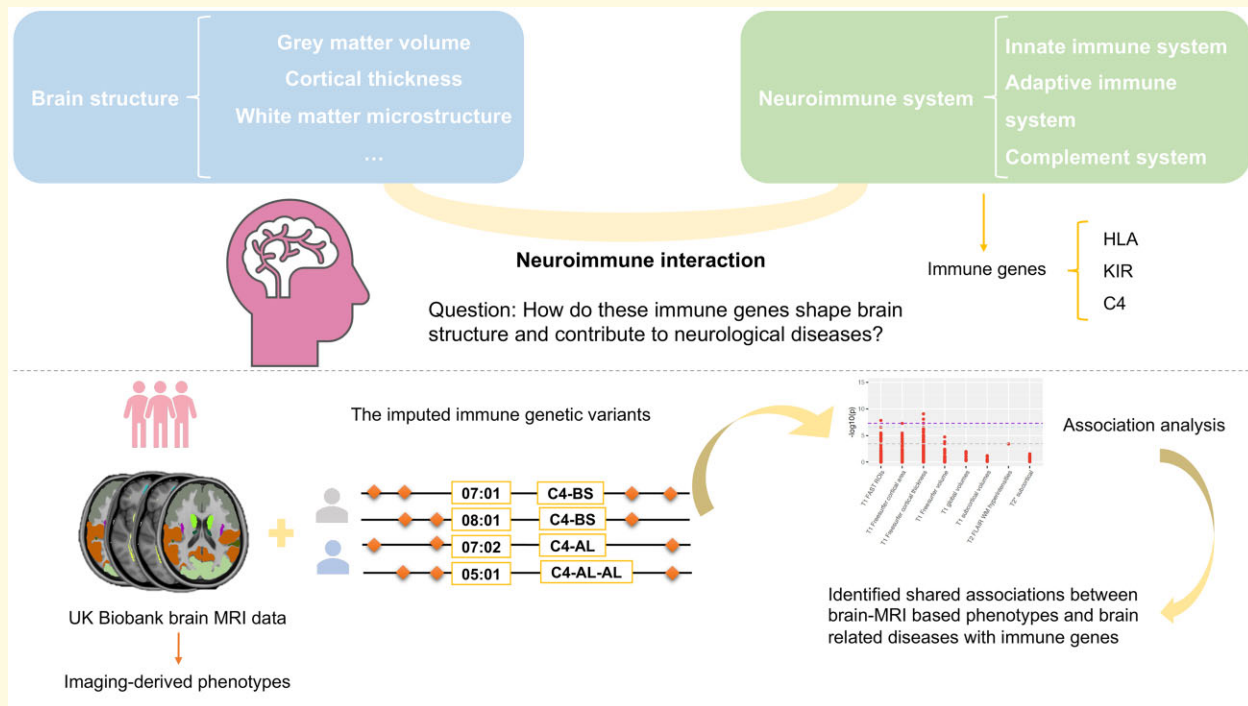
Abbreviations: AH8.1 = ancestral haplotype 8.1; C4 = complement component 4; HLA = human leukocyte antigen; IDP = image-derived phenotype; KIR = killer cell immunoglobulin-like receptor; MHC = major histocompatibility complex; PheWAS = phenome-wide association study; SNP = single nucleotide polymorphism; TATES = trait-based association test that uses extended Simes procedure

Received August 11, 2021. Revised December 08, 2021. Accepted March 28, 2022. Advance access publication March 31, 2022

© The Author(s) 2022. Published by Oxford University Press on behalf of the Guarantors of Brain.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Graphical Abstract



Introduction

Brain MRI has been widely used for detecting the abnormalities of brain structure and function and contributes to the diagnostic process for different brain disorders.^{1–4} Previous studies have demonstrated that abnormalities in different brain regions are linked to different diseases. For instance, schizophrenia is associated with cortical grey matter deficits,⁵ and abnormalities of cerebral structure, such as the frontal pole and temporal pole, have been reported to be associated with bipolar disorder.⁶ With the advances in brain imaging, investigating white matter differences provides new opportunities to understand neurological diseases. Diffusion tensor imaging (DTI) studies have shown that the microstructure of white matter plays a role in Alzheimer's disease, multiple sclerosis, amyotrophic lateral sclerosis and Parkinson's disease.^{7–10} These findings support that MRI data can be used as intermediate phenotypes to discover new biomarkers for the clinical diseases and potentially elucidate disease pathways.

The large-scale datasets collected by the ENIGMA Consortium¹¹ and the UK Biobank¹² offer great opportunities to explore the genetic architecture of human brain structure and function. Genome-wide association studies (GWASs) conducted using UK Biobank data have identified a large number of genetic variants and genes associated with MRI measures.¹³ Further multi-phenotype association tests showed many of these identified genetic regions are also associated with mental illnesses and neurodegenerative

diseases. Another notable study investigated the genetic contribution to variation between individuals in cerebral cortex traits, reporting a negative genetic correlation between surface area (SA) and cortical TH and that these traits are shaped by largely distinct loci.¹⁴

The major histocompatibility complex (MHC), also known as human leukocyte antigen (HLA) in human, and killer cell immunoglobulin-like receptor (KIR) have been associated with many neurological diseases.¹⁵ For instance, the *C4* gene in the HLA region has been associated with schizophrenia risk.¹⁶ Parkinson's disease has been associated with *HLA* class I and class II loci and haplotypes.^{17,18} *HLA-DR* and *HLA-DQ* genes were found to be associated with Alzheimer's disease,^{19,20} while *HLA-DRB1*15:01* was reported as the highest risk allele to be associated with MS.^{21,22} As far as we know, there is no genomic region as consistently associated with many brain-related diseases as the MHC region. However, association analyses of brain MRI phenotypes have provided only simple single nucleotide polymorphism (SNP)-based analysis, which missed the complexity of the MHC region. Here, we study imputed HLA haplotypes. Compared to the role of *HLA* in neurological diseases which has been studied over the past decades, the impact of variation in the KIR region has only been studied more recently. The *KIR* gene family is also a crucial component in regulation of the immune system, which is located on chromosome 19. This complex locus is characterized by high polymorphism in terms of copy number and haplotypes.²³ *KIR* genes encode a family of glycoproteins expressed on natural killer (NK) cells and a small number of subtypes of

T cells. Most of the *KIR* genes encode inhibitory receptors which protect healthy cells from destruction by NK cells. Some of them are activating genes that can help remove the antigens like viruses and cancer. The combination of KIR and HLA class I ligands determines NK cell activation and inhibition, which may affect the outcome of infections.²⁴ Few studies have investigated the association between KIR receptor/HLA ligand pairs and brain-related traits.

The goal of this study is to investigate the relationship between the *HLA* and *KIR* loci and brain MRI-based phenotypes (Fig. 1). We conducted an association study using ~30 000 unrelated individuals of European ancestry for who we had genotyping and brain MRI data in the UK Biobank. We tested the associations between 1422 image-derived phenotypes (IDPs) (Supplementary Table 1) and the imputed *HLA*, *KIR*, *C4* genes as well as KIR/HLA ligand pairs and the IDPs. We identified a large number of associations between immune genes and IDPs in the European population. Our study sheds light on how the immune genes may shape human brain development and contribute to neurological diseases.

Materials and methods

Brain imaging phenotypes data

We used IDPs released in January 2020 by UK Biobank. The UK Biobank acquired six MRI modalities (T1, fluid-attenuated inversion recovery (T2_FLAIR), susceptibility weighted imaging (SWI), dMRI, resting-state functional (rfMRI) and task-based functional (tfMRI)), consisting of structural, diffusion and functional imaging. The IDPs were generated from the raw images by using an imaging-processing pipeline developed by UK Biobank.²⁵ We did not consider the rfMRI edges due to the low SNP-based heritability¹³ ($h^2 = 0.054$, averaged over 1485 edges), which likely reflects a low test-retest reliability. In addition, we also included IDPs (volumes of subcortical structures, SA and cortical TH) estimated by FreeSurfer using Desikan–Killiany–Tourville atlas and Destrieux (a2009s) atlas, respectively. We organized these IDPs into nine groups (Supplementary Table 1) as suggested in the published GWAS paper.¹³ The basic quality control was performed by the UK Biobank in order to exclude the problematic MRI images.²⁵ The description of imaging-processing protocol was summarized in the UK Biobank brain imaging documentation (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf).

Approximately 40 000 participants in UK Biobank have brain imaging phenotypes. Some of the IDPs have extreme distributions that may violate linear regression assumptions (the normality of the residuals). We calculated the skewness for each IDP (Supplementary Table 2). The skewness of each IDP was quite similar to that calculated in the previous published GWAS paper.¹³ To make sure the distribution of each IDP meet the linear regression assumptions, we used z-score calculated by rank-based inverse normal transformation (RINT) instead of raw IDP for each association analysis.

To assess the robustness of the observed associations, we compared the HLA associations results between raw and transformed data using four IDPs with distinct distributions (Supplementary Fig. 1 and Supplementary Table 3). We observed similar results so we used RINT to carry out the association study. In our association analysis for each IDP, we ignored the individuals who had missing values.

Genetic data

We used the raw genotype data released in July 2017 by UK Biobank. Genotyping was performed using Affymetrix UK BiLEVE Axiom array and the Affymetrix UK Biobank Axiom array. There are 456 426 individuals of European ancestry in UK Biobank V2 genotype data. We selected unrelated individuals (gcta²⁶ –rel-cut-off 0.05) who have both genotype data and imaging-derived phenotypes. We removed the withdrawn individuals in UKB cohort from our study. This left a subset of 28 647 individuals identified for inclusion in the association study. We then extracted the specific regions covering *MHC* and *KIR* loci, respectively. These genotype data were used for *C4* and *KIR* genes imputations.

Imputation

Imputed HLA alleles

HLA alleles were imputed by UK Biobank using HLA*IMP:02.²⁷ The HLA genes were imputed to a multi-population reference panel. The UK Biobank imputations include all 11 *HLA* loci with all possible four-digit alleles. For each locus, the two possible alleles were assigned according to the posterior probabilities. We used the allelic dosage data provided by UK Biobank for the association analysis. *HLA* alleles with MAF < 0.01 were removed from downstream analysis.

C4 imputation

We used Beagle 5.1 to impute *C4* alleles against the reference panel²⁸ publicly available in dbGaP under accession number phs000672.v1.p1 (see Supplementary Note). We first extracted those SNPs in MHC region (24Mb–34Mb on chromosome 6) using PLINK1.9²⁹ and converted the plink format file into variant call format (VCF) file. We kept SNPs with genotype missingness < 0.05, HWE *P*-value < 10^{-6} and minor allele count > 5. Then, conform-gt tool was used to modify the VCF file to make sure the SNPs matched with the reference panel. There were 2991 SNPs overlapped with the reference panel and were used for *C4* imputation. We extracted the DS field in the resultant VCF file and used it as allelic dosage data. All possible *C4* structural types were imputed but only the common alleles were used in the association analyses. The resultant *C4* alleles were encoded using <H_n_a_b_h_s> where n is the total number of *C4* genes, a is the number of *C4A* genes, b is the number of *C4B* genes and h is the number of HERV copies, respectively. The ‘s’ here is used to distinguish the haplotypes which have same combination of <a, b, h> but different flanking SNP haplotypes. We then converted the encoded

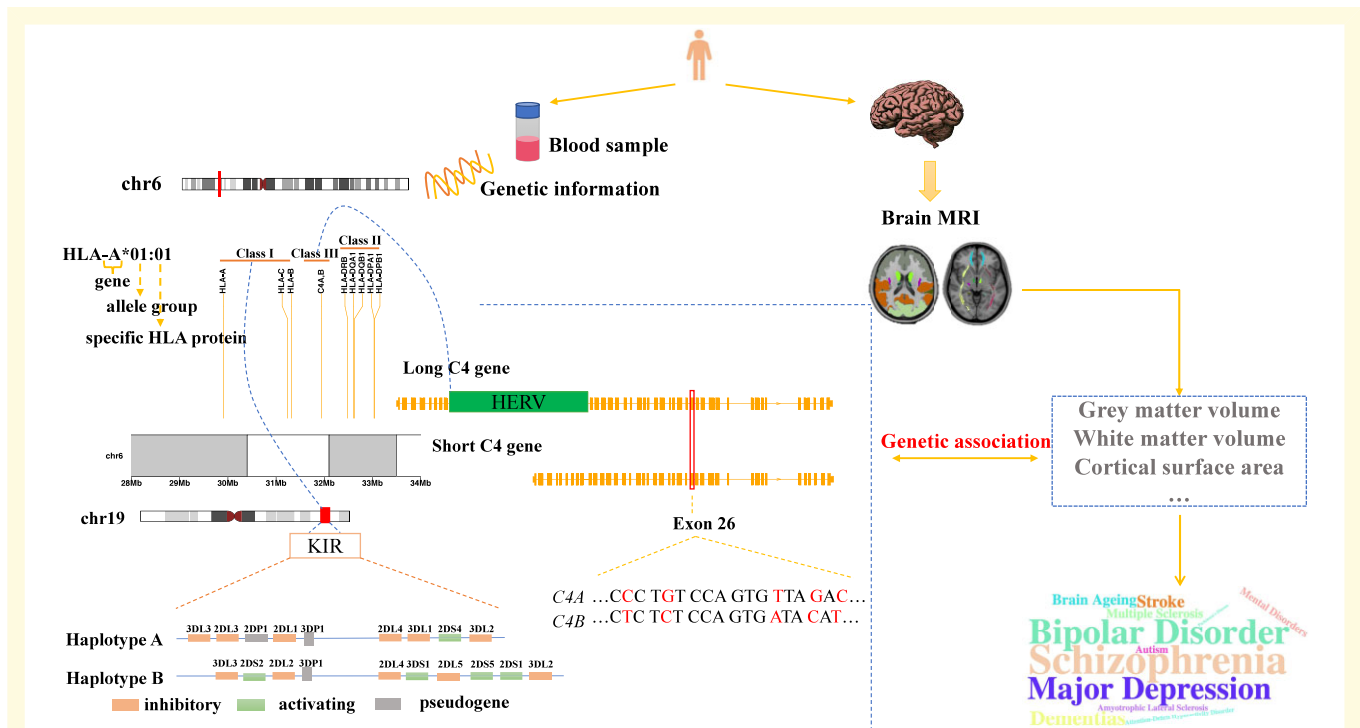


Figure 1 The overview of the study. We identified 28 647 unrelated individuals of European descents from the UK Biobank.

These participants have both genotype data and brain MRI data. The brain imaging phenotypes were used as intermediate phenotypes. The genetic associations of the classical *HLA* genes, complement component 4 gene, *KIRs* as well as the interactions between *HLA* class I genes and *KIRs* were tested with 1422 IDPs. *HLA* gene family, located on chromosome 6p21.3, can be broadly divided into three classes: *HLA-A*, *HLA-B* and *HLA-C* are *HLA* class I genes. *HLA-DQA1*, *HLA-DQB1*, *HLA-DRB*, *HLA-DPA1* and *HLA-DPB1* are *HLA* class II genes. *C4* gene is an *HLA* class III gene, which has two isotypes, *C4A* and *C4B*. The long and short form of *C4* gene can be distinguished by the presence or absence of the *HERV* sequence between exon 9 and exon 10. The difference between *C4A* and *C4B* lies in four amino acids encoded by exon 26. The *KIR* locus is located on the chromosome 19q13.4 and is characterized by a high degree of polymorphism in terms of copy number and haplotypes. The haplotypes can be broadly classified into haplotype A and haplotype B. Gene content in each haplotype is shown. The *HLA* class I molecules are the ligands for the *KIR* receptors.

results to possible *C4* structural alleles. For example, $\langle H_1_0_1_0 \rangle$ means this allele carries 1 *C4* gene and 1 *C4B* gene. The corresponding structural allele is C4-BS. But $\langle H_2_1_1_1 \rangle$ could be AL-BS, AS-BL, BL-AS or BS-AL. The haplotypes do not attempt to distinguish the order of A, B, L and S. We chose the same names as previous publication¹⁶ for such haplotypes. The frequency and haplotype name of each *C4* allele are listed in Supplementary Table 4. The *C4A* expression level was predicted as $C4A_{expression} = (0.47 * AL) + (0.47 * AS) + (0.2 * BL)$.¹⁶

KIR imputation

We imputed *KIR* genes using *KIR*IMP*.²³ *KIR*IMP* provides a UK reference panel with 301 SNPs and 479 haplotypes. We extracted chr19:53Mb-58Mb using PLINK 1.9.²⁹ 221 SNPs overlapped with the reference panel on the UKB array. To gain more accuracy for *KIR* imputation, we uploaded the data to Michigan Imputation Server³⁰ and imputed this region against the HRC panel.³¹ There were 267 SNPs overlapped with the *KIR* reference panel according to SNP positions. We compared the allele frequencies between our data and the reference data and manually removed those ambiguous SNPs. Finally, we identified 258 SNPs that

have similar allele frequencies with the reference data, which can be used for imputation. As a sanity check, we compared gene and haplotype frequencies with other studies. Similar frequencies of *KIR* variants were observed across multiple studies (Supplementary Tables 5 and 6). The imputed *KIR* variants, including 17 *KIR* loci, fine-scaled haplotypes as well as general haplotypes were used for downstream association analysis.

Confounding factors

We followed the published GWAS paper which suggested three categories of confounding factors in association analysis:¹³ (i) imaging confound; (ii) genetic ancestry; and (iii) height, weight and BMI. For the imaging confounds, age, sex, age², age*sex, age²*sex, site, date, date² were included as the covariates. In addition, we also included the location of the head (*x*-position, *y*-position, *z*-position), head motion from both resting fMRI and task-based fMRI and volumetric scaling factor for head size in linear regression. To account for genetic ancestry confounds, we included 40 genetic principal components computed by UKB as covariates in our analyses. Some body measures could also be confounders

for IDPs. Here, we included three traditional confounds height, weight and BMI as covariates. To assess the potential collider bias in our analysis, we also performed association tests without the body measures (height, weight and BMI) as covariates with a selected phenotype (Supplementary Notes, Supplementary Fig. 2).

Statistical analyses

Associations of HLA alleles with IDPs

All the confounding factors and HLA genotype dosage data were regressed to each IDP. No missing values were found with age, sex, site, date, height and weight. Other missing covariate values of imaging confounds were substituted by the mean value of the covariate. Each HLA allele was tested using multiple linear regression.

$$IDP_{z\text{-score}} = \beta_0 + \beta_1 HLA_{\text{dosage}} + \beta X + \varepsilon \quad (1)$$

where $IDP_{z\text{-score}}$ is the z-score of IDP transformed using RINT, β_0 is the intercept of the linear model, β_1 is the effect size of the HLA allele, HLA_{dosage} is the dosage of each HLA allele, β is the coefficients matrix of the covariates, X is the confounding factor matrix and ε is the residual of the model.

Associations of C4 haplotypes with IDPs

We only tested the five most common C4 haplotypes out of the 29 imputed C4 structural haplotypes in our association study. Each haplotype name with the corresponding C4 imputation result is listed in Supplementary Table 4. We summed the dosage data according to haplotype name and used this result for association analysis.

$$IDP_{z\text{-score}} = \beta_0 + \beta_1 C4_{\text{haplotype}} + \beta X + \varepsilon \quad (2)$$

where terms are as in Equation (1), except for $C4_{\text{haplotype}}$ indicates the dosage of each C4 haplotype.

Associations of C4 copy number with IDPs

We counted the copy number of C4A, C4B, C4L, C4S gene as well as the total copy number of C4 gene according to the C4 imputation result. C4A expression level was predicted as described in previous study.¹⁶ We then regressed each C4 gene copy number and C4A expression level with other covariates to each z-score of IDP and conducted association analysis by using linear regression.

$$IDP_{z\text{-score}} = \beta_0 + \beta_1 (\#C4 \text{ gene or } C4A \text{ expression}) + \beta X + \varepsilon \quad (3)$$

where terms are as in Equation (1), except for # C4 gene indicates the number of C4 gene of different isotypes C4A and C4B and structural types C4L and C4S as well as total copy number of C4, C4A expression indicates the predicted C4A expression level.

Associations of KIR gene with IDPs

KIR*IMP imputed 19 genetic variants in total, consisting of general haplotypes, fine-scaled haplotypes as well as 17 KIR genes copy number. We converted each haplotype into a single variable and tested them independently. For each KIR haplotype,

$$IDP_{z\text{-score}} = \beta_0 + \beta_1 KIR_{\text{haplotype}} + \beta X + \varepsilon \quad (4)$$

where terms are as in Equation (1), except for $KIR_{\text{haplotype}}$ is a variable counted the number of the specific haplotype each individual carried.

For each KIR gene copy number,

$$IDP_{z\text{-score}} = \beta_0 + \beta_1 KIR_{\text{copy-number}} + \beta X + \varepsilon \quad (5)$$

where terms are as in Equation (1), except for $KIR_{\text{copy-number}}$ is the copy-number of KIR gene.

Associations of epistatic interaction between KIR and HLA ligands with IDPs

KIR genes and their HLA ligands were listed in Supplementary Table 7. These ligand-receptor pairs have been previously validated experimentally.³² For the subgroup of HLA, we summed the HLA dosage data over four-digit HLA alleles belonging to each subgroup. We tested epistatic associations by introducing the interaction term KIR*HLA into the linear regression. The epistatic association analysis was carried out following the models used in a recent KIR/HLA epistasis study in ankylosing spondylitis disease.³² Briefly, we used dominant (0= gene absent, 1= one or more gene copy) and recessive (0= dosage less than two, 1= homozygosity at locus) models for KIR genes, respectively. HLA alleles were treated as having dominant effect.

$$IDP_{z\text{-score}} = \beta_0 + \beta_1 KIR_{\text{dominant/recessive}} + \beta_2 HLA_{\text{dominant}} + \beta_3 KIR_{\text{dominant/recessive}} * HLA_{\text{dominant}} + \beta X + \varepsilon \quad (6)$$

where $KIR_{\text{dominant/recessive}}$ is the KIR status under dominant or recessive model, HLA_{dominant} is the HLA status under the dominant model, $KIR_{\text{dominant/recessive}} * HLA_{\text{dominant}}$ is the interaction term under KIR dominant or recessive model.

Correction for multiple testing

For single variant association analysis, we used a conservative P value threshold of 2.5×10^{-7} ($0.05/(137*1422)$) which accounted for the number of both genetic variants and IDPs tested in this study. We also considered a relaxed threshold of 3.6×10^{-4} ($0.05/137$), which only accounted for the number of genetic variants tested for each IDP. For KIR/HLA epistatic associations, we adjusted 0.05 significant threshold with the number of KIR/HLA pairs and IDPs ($P < 3.1 \times 10^{-7}$ ($0.05/(112*1422)$)). The conservative thresholds were very close to the genome-wide significant threshold, so we also report the associations at the threshold of

5×10^{-8} . These thresholds are conservative considering the correlation across IDPs and the strong linkage disequilibrium (LD) between neighbouring loci in HLA and KIR region, but these associations results should be interpreted carefully.

Multi-trait association analysis

The multi-trait association analysis was carried out using TATES.³³ We did not choose popular methods such as MTAG³⁴ and Genomic SEM³⁵ for the multivariate analysis considering that they are built on LD score regression,³⁶ which excluded the MHC region. TATES is more flexible and allows users to use *P*-values as input and detects both genetic variants that are common to multiple phenotypes and genetic variants that are specific to a single phenotype. Thus, it is more appropriate for this study. In multi-trait association analysis, the IDPs were grouped into 19 classes (Supplementary Table 8). We only considered the additive effect between HLA, C4 and KIR and each IDP group. In brief, we first calculated the phenotypic correlation matrix after correcting for all the covariates used in univariate association analysis. We then corrected for the correlation matrix and acquired TATES *P*-values for each IDP group.

Data availability

This study makes use of individual-level genotype and brain MRI phenotype data from UK Biobank Resource (application number: 12505). UK Biobank: <https://www.ukbiobank.ac.uk>.

Results

Summary

Imputation of four-digit HLA alleles from SNPs was carried out using HLA*IMP:02²⁷ by the UK Biobank team (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/HLA_impuration.pdf). The C4 haplotypes were imputed using a mixed ancestry reference panel,²⁸ with the C4 copy numbers calculated from the imputed haplotypes. The KIRs including 17 KIR genes, KIR general haplotypes (A/B) as well as fine-scale haplotypes were imputed using KIR*IMP.²³ We performed the genetic association analysis between the common genetic variants and IDPs using an additive genetic model. We considered three *P*-value thresholds with Bonferroni correction (see Materials and methods). At the most stringent threshold, the genome-wide threshold (5×10^{-8}), we identified 299 associations between single genetic variants and IDPs (Supplementary Tables 9–11). The detailed association results at the other two thresholds, which accounted for both the number of variants and IDPs tested and the number of variants tested only, are shown in Supplementary Tables 12–18. The number of the significant associations at different thresholds is summarized in Supplementary Table 19. We observed shared associations

across IDPs. The QQ-plots of the HLA associations with all IDPs and each category of IDPs are displayed in Supplementary Fig. 3, showing some inflation of test statistics. To demonstrate the inflation was not due to other confounding factors, we performed standard GWAS on five selected IDPs (Supplementary Note). These GWAS QQ-plots (Supplementary Fig. 4) and LD score regression intercepts (Supplementary Table 20) for these five IDPs show no other bias exists. These results indicate that the observed inflation in test statistics in our analysis is due to the correlation across HLA alleles as well as IDPs.

Significant HLA alleles associated with IDPs

We identified 15 HLA alleles associated with at least one IDP ($P < 5 \times 10^{-8}$), containing 130 associations. Using the genome-wide threshold, we did not find any significant associations with fMRI. This was consistent with the low SNP-based heritability of fMRI phenotypes¹³ which might be due to the low test-retest reliability of fMRI measurements.³⁷ The PheWAS plot (Fig. 2 and Supplementary Figs. 5 and 6) shows the associations of 15 significant HLA alleles with structural and dMRI IDPs. The most significant associations of these 15 HLA alleles with associated IDPs are given in Table 1.

Overall, we observed some similarities in association patterns in terms of both *P* values and signs of effect across some HLA alleles. For example, we found HLA-A*01:01, HLA-C*07:01, HLA-B*08:01, HLA-DRB1*03:01/DRB3*01:01, HLA-DQB1*02:01 and HLA-DQA1*05:01 were associated with numerous structural and dMRI IDPs, including the regional grey matter volume, cortical TH, fractional anisotropy (FA), intra-cellular volume fraction (ICVF) and mean diffusivity (MD) (Fig. 2 and Supplementary Figs. 5 and 6). HLA-DRB1*04:01/04, HLA-DQA1*03:01, HLA-DQB1*03:02 and HLA-DRB4*01:03 were mainly associated with isotropic volume fraction (ISOVF) of dMRI which indicates free water volume fraction. In addition, HLA-B*07:02 and HLA-C*07:02 were associated with cortical SA in multiple regions at the nominal threshold, while strongly associated with orientation dispersion (OD) in cerebral peduncle and T2* in the hippocampus. HLA-C*05:01 was strongly associated with ICVF and OD in the cerebral peduncle (Supplementary Table 15).

We observed distinct HLA alleles affect cortical SA and TH. For instance, HLA-B*08:01 was associated with TH in several regions at the nominal threshold, while HLA-B*07:02 was associated with SA (Supplementary Table 15). The spatial distribution of $-\log_{10}(P)$ values and effect sizes of HLA-B*08:01 and HLA-B*07:02 within the whole brain indicates the association strengths and effects on SA and TH (Fig. 3; Supplementary Fig. 7).

We investigated if the HLA haplotypes found here to be associated with IDPs have been reported to be associated with brain-related diseases. For example, HLA-B*08:01, HLA-DRB1*03:01 and HLA-DQB1*02 are associated

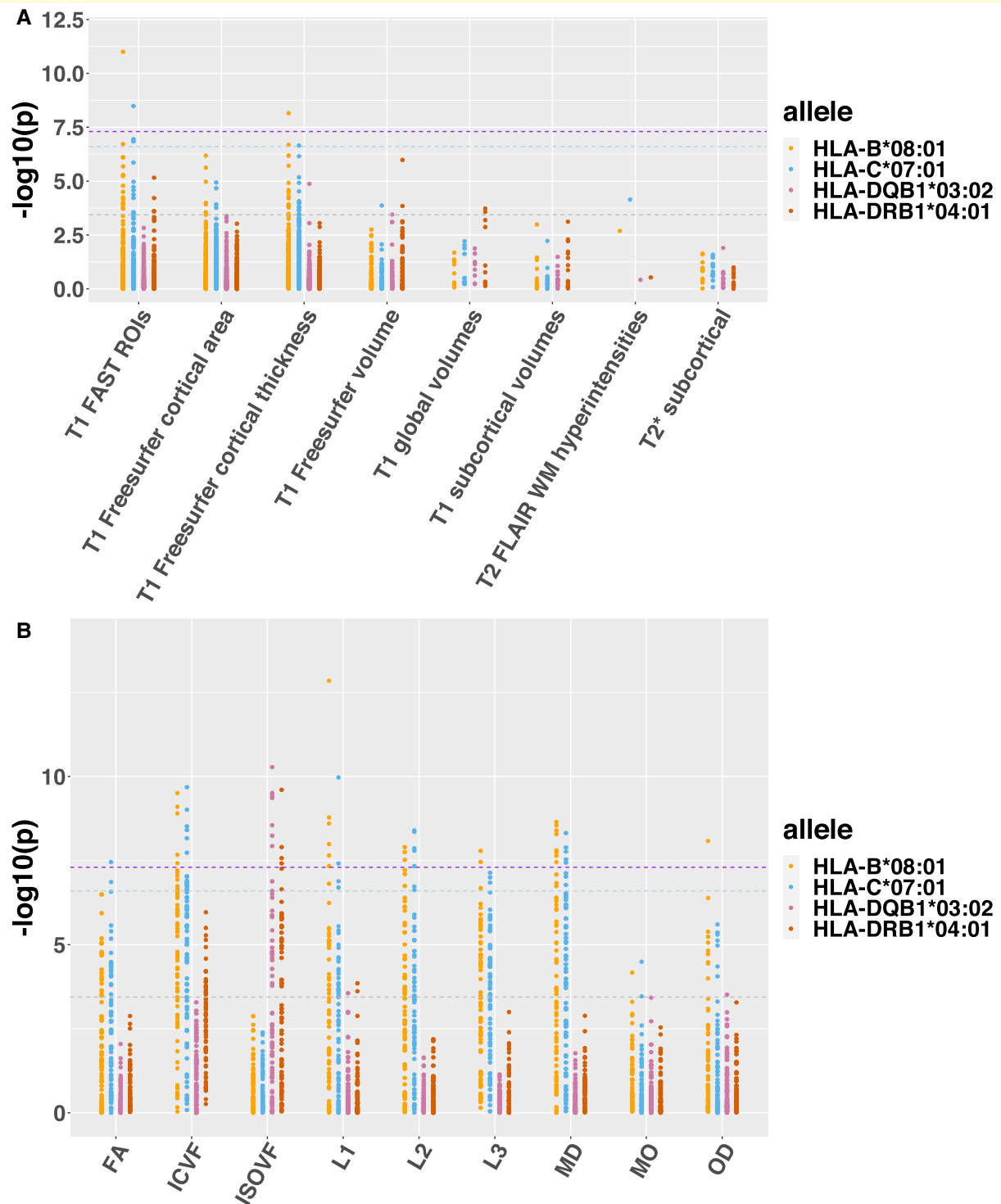


Figure 2 PheWAS plots of selected HLA alleles across IDPs. The IDPs were generally split into structural and dMRI groups. Structural IDPs were divided into eight subgroups, including T1 FAST region of interest (ROIs), T1 freesurfer cortical area, T1 freesurfer cortical TH, T1 freesurfer volume, T1 global volumes, T1 subcortical volumes, T2 FLAIR WM hyperintensities and T2* subcortical. dMRI IDPs were divided into nine subgroups, including FA, ICVF, ISOVF, L1, L2, L3, MD, MO and OD. Different subgroups of IDPs are given along the x-axis. The strength of association ($-\log_{10}(P$ value)) for individual IDP within each subgroup is given on the y-axis. The purple dash line indicates the genome-wide significant $-\log_{10}(P$ value) threshold of 7.3. The blue dash line indicates the conservative $-\log_{10}(P$ value) threshold of 6.6 accounts for the number of loci (137) as well as total IDPs (1422) tested. The grey dash line indicates a nominal significance $-\log_{10}(P$ value) threshold of 3.44 correcting only for the number of the tested loci (137). **(A)** The PheWAS plot of the association between structural IDPs and HLA-B*08:01, HLA-C*07:01, HLA-DQB1*03:02 and HLA-DRB1*04:01. **(B)** The PheWAS plot of the association between dMRI IDPs and HLA-B*08:01, HLA-C*07:01, HLA-DQB1*03:02 and HLA-DRB1*04:01.

Table 1 Summary of the most significant HLA alleles with the associated IDPs

HLA allele name	Top association with structural MRI ($P < 5 \times 10^{-8}$)	Top association with dMRI ($P < 5 \times 10^{-8}$)	Direction of effect (structural MRI/dMRI)
HLA-A*01:01	None	Mean MD in posterior thalamic radiation on FA skeleton (left)	None/–
HLA-B*07:02	None	Mean L1 in cingulum hippocampus on FA skeleton (left)	None/–
HLA-B*08:01	Volume of grey matter in thalamus (left)	Mean L1 in superior longitudinal fasciculus on FA skeleton (left)	–/–
HLA-C*05:01	None	Mean ICVF in cerebral peduncle on FA skeleton (right)	None/–
HLA-C*07:01	Volume of grey matter in thalamus (left)	Mean L1 in superior longitudinal fasciculus on FA skeleton (left)	–/–
HLA-C*07:02	None	Mean OD in cerebral peduncle on FA skeleton (left)	None/+
HLA-DRB3*01:01	Mean TH of insula (left hemisphere)	None	+/None
HLA-DRB4*01:03	None	Weighted-mean ISOVF in tract superior longitudinal fasciculus (right)	None/+
HLA-DRB1*03:01	Mean TH of insula (left hemisphere)	Mean L2 in posterior thalamic radiation on FA skeleton (left)	+/–
HLA-DRB1*04:01	None	Weighted-mean ISOVF in tract inferior fronto-occipital fasciculus (left)	None/+
HLA-DRB1*04:04	None	Weighted-mean ISOVF in tract superior longitudinal fasciculus (right)	None/+
HLA-DQB1*02:01	Volume of grey matter in Thalamus (left)	Mean L2 in posterior thalamic radiation on FA skeleton (left)	–/–
HLA-DQB1*03:02	None	Weighted-mean ISOVF in tract superior longitudinal fasciculus (right)	None/+
HLA-DQA1*03:01	None	Weighted-mean ISOVF in tract superior longitudinal fasciculus (right)	None/+
HLA-DQA1*05:01	Area of S-pericallosal (right hemisphere)	None	–/none

L1,2 = in a diffusion tensor fit, the strength of diffusion along the principal axes of the ellipse (eigen values).

with schizophrenia¹⁶ and depression.³⁸ These alleles were found to be associated with a number of structural IDPs. Another example is the association between *HLA-B*07:02* and Alzheimer's disease.³⁹ The significant association was found between *HLA-B*07:02* and T2* in the hippocampus (essential for memory and a biomarker of Alzheimer's disease). Additionally, we observed distinct *HLA* alleles associated with different diffusion metrics (Fig. 2B and Supplementary Fig. 6). For example, *HLA-C*07:01* (Fig. 4A) and other alleles in LD (Fig. 5) were mainly associated with FA of forceps minor, ICVF of posterior thalamic radiation, MD of cingulate gyrus tract (genetically correlated with major depression⁴⁰). This long haplotype contains both MHC class I and II genes. MHC I molecules have been shown to be expressed on neurons. These proteins are crucial for synapses pruning,⁴¹ early brain development⁴² as well as ageing.⁴³ In contrast, *HLA-DRB1*04:01/04*, *HLA-DQA1*03:01*, *HLA-DQB1*03:02* and *HLA-DRB4*01:03* were only associated with ISOVF across multiple tracts, including superior longitudinal fasciculus, inferior fronto-occipital fasciculus and inferior longitudinal fasciculus (Fig. 4B). Interestingly, these alleles have been repeatedly found to be associated with Parkinson's disease.^{17,18,44} Taken together, these shared associations suggest that MHC I and MHC II molecules might be involved in multiple biological processes in the brain.

To test whether other SNPs in the MHC region were associated with IDPs, we selected five IDPs that were found to be significantly associated with distinct HLA alleles

(Supplementary Table 21). We first performed the genetic association study with the SNPs in the MHC region (chr6: 29Mb-34Mb, covering *HLA-A*, *HLA-B*, *HLA-C*, *HLA-DQ*, *HLA-DP*, *HLA-DR* genes), identifying significant associations with SNPs located near or at the classical HLA genes. To further examine additional independent associations, we tested all the SNPs in the MHC region conditioning on all common classical four-digit *HLA* alleles of the associating *HLA* genes. For all five IDPs, we did not find any significant SNP after fitting the *HLA* allele ($P < 5 \times 10^{-8}$) (Supplementary Fig. 8), indicating the SNPs based association signals are not independent of the *HLA* alleles.

Similarity across HLA alleles

Due to the complex linkage disequilibrium (LD) in MHC region, we observed several sets of *HLA* alleles that have similar association patterns across IDPs. To further investigate the similarity between those significant alleles, we calculated the squared correlation coefficient using the genotype count number of the *HLA* alleles (Fig. 5). From the correlation, we found *HLA-A*01:01*, *HLA-B*08:01*, *HLA-C*07:01*, *HLA-DQA1*05:01*, *HLA-DRB1*03:01/DRB3*01:01* and *HLA-DQB1*02:01* exhibit strong LD. This observation was quite similar with the association patterns (Fig. 2 and Supplementary Figs. 5 and 6). According to previous MHC haplotype studies,^{45,46} *HLA-A*01:01~HLA-C*07:01~HLA-B*08:01~HLA-DQA1*05:01~HLA-DRB1*03:01~HLA-DQB1*02:01* is a common ancestral haplotype (also

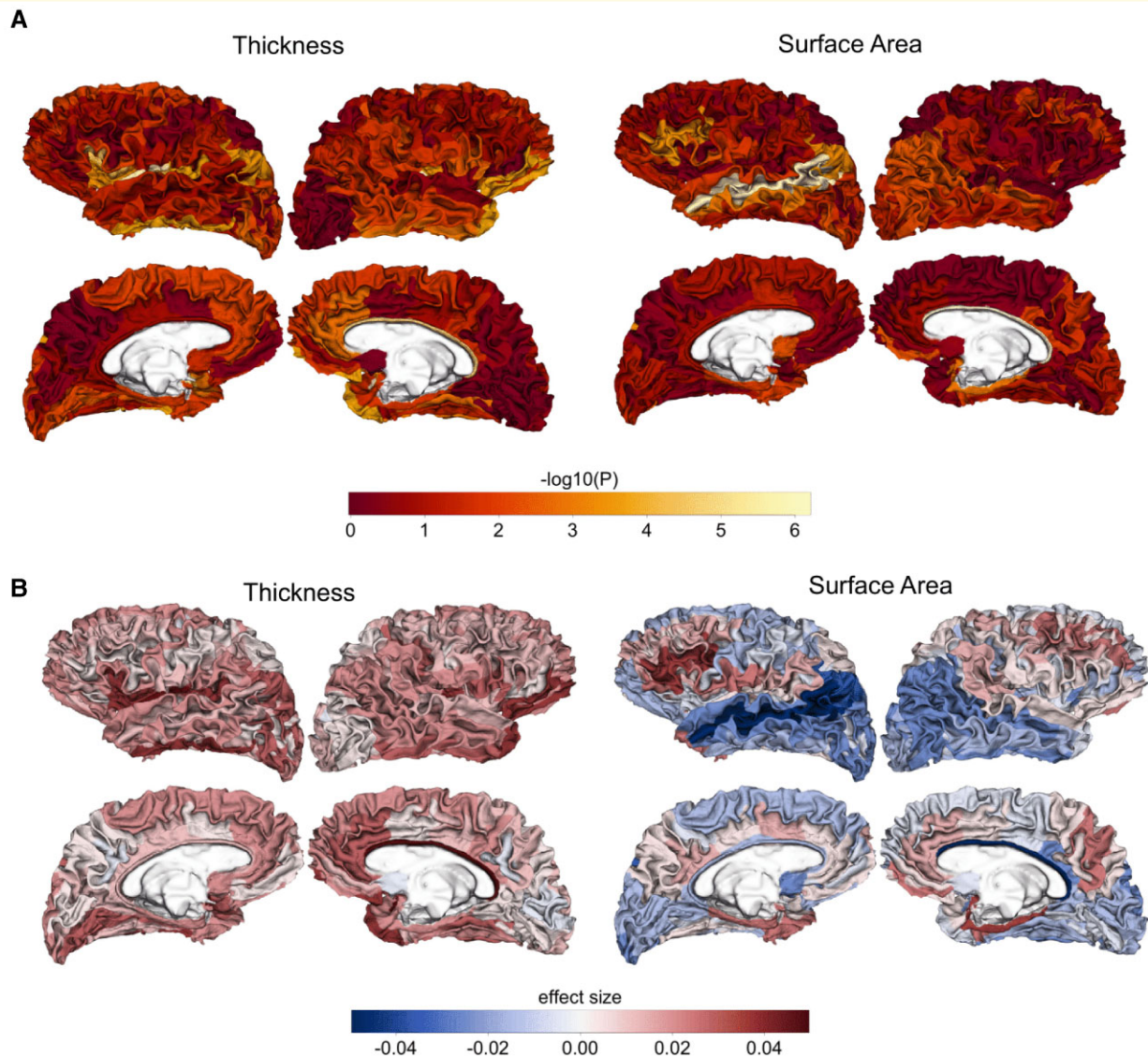


Figure 3 Spatial maps of the associations between HLA-B*08:01 and cortical SA and TH. (A) Spatial maps of the association between HLA-B*08:01 and SA and TH within whole brain. $-\log_{10}(P)$ values are shown on freesurfer a2009s atlas with 74 regions included. **(B)** Spatial maps of the effect size of HLA-B*08:01 on TH and SA within whole brain.

known as AH8.1, frequency = 0.09 in our study) in the European population. Similarly, *HLA-C*07:02~HLA-B*07:02* is a common HLA-C-B haplotype (frequency = 0.15 in our study). *HLA-DQA1*03:01~HLA-DQB1*03:02~HLA-DRB1*04:01/04~DRB4*01:03* is a class II haplotype.

C4-BS was associated with numerous IDPs

The PheWAS plot for the associations of each common C4 haplotype to IDPs is presented in Fig. 6A and C. Five C4 haplotypes including C4-BS, C4-AL, C4-AL-AL, C4-AL-BL and C4-AL-BS were tested with IDPs. We identified a large number of associations between C4-BS and IDPs (Supplementary

Table 10). The shared associations of C4-BS and the *HLA* alleles in AH8.1 across IDPs were due to the LD between them (Fig. 5). To be specific, the strongest association with structural IDPs was observed between C4-BS and grey matter volume in left thalamus ($P = 1.18 \times 10^{-9}$, $\beta = -0.068$). For the associations between C4-BS and dMRI IDPs, we found significant positive associations with FA, ICVF and OD but negative associations with MD across multiple white matter tracts. No significant associations were found with ISOVF.

The associations of C4 gene copy number with IDPs

The two distinct isoforms of human *C4* gene, *C4A* and *C4B*, vary in copy number and structure, with the expression of

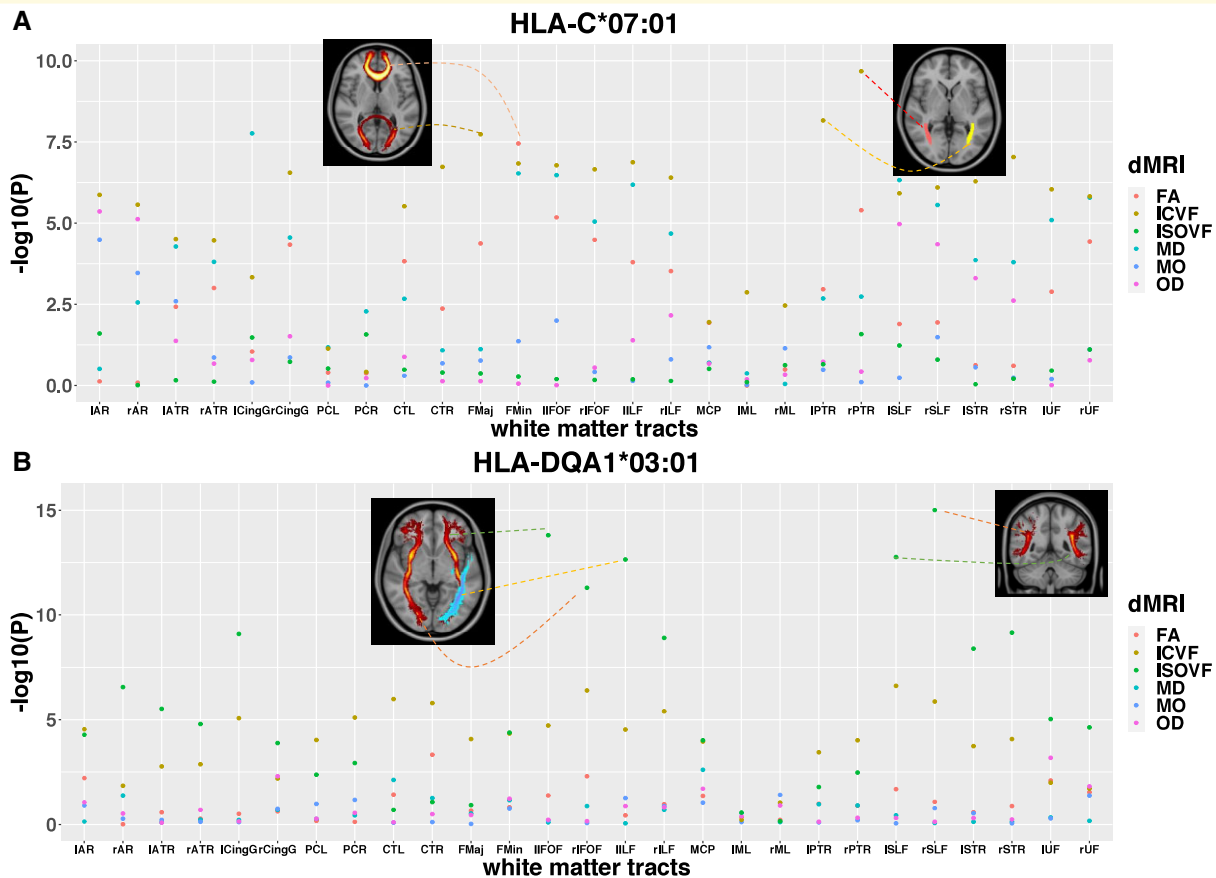


Figure 4 The associations of HLA-C*07:01 and HLA-DQA1*03:01 with dMRI metrics across white matter tracts in the brain. **(A)** The associations of HLA-C*07:01 with dMRI metrics across white matter tracts in the brain. **(B)** The associations of HLA-DQA1*03:01 with dMRI metrics across white matter tracts in the brain. The x-axis indicates the name of each white matter tract. The y-axis indicates the $-\log_{10}(P\text{-value})$. The selected significant associated white matter tracts are displayed on the standard MNI152 T1 image. AR, acoustic radiation; ATR, anterior thalamic radiation; CingG: cingulum gyrus; PCL/PCR, parahippocampal part of cingulum (left/right); CTL/CTR, corticospinal tract (left/right); FMaj, forceps major; FMin, forceps minor; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; MCP, middle cerebellar peduncle; ML, medial lemniscus; MO, diffusion tensor mode; PTR, posterior thalamic radiation; SLF, superior longitudinal fasciculus; STR, superior thalamic radiation; UF, uncinate fasciculus; l, left; r, right.

each gene being proportional to gene copy number.¹³ The most significant associations across structural IDPs were observed between the total *C4* copy number and the mean TH of the left insula ($P = 1.99 \times 10^{-9}$) and long insula gyrus ($P = 2.78 \times 10^{-8}$), which is a small region and functionally regulates neuro-immune homeostasis.⁴⁷ At the nominal significance threshold, we observed higher *C4A*, *C4L* and total *C4* copy number were associated with thinner regional cortical TH and decrease of grey matter in multiple regions, including frontal medial cortex and temporal fusiform cortex. In terms of the dMRI IDPs, we found higher *C4A*, *C4L* and total *C4* copy number were associated with lower FA, ICVF and OD and higher MD of multiple major white matter tracts, including posterior thalamic radiation, sagittal stratum and superior longitudinal fasciculus (Fig. 6B and D; Supplementary Table 17). No significant associations were found with ISOVF and MO. Considering *C4A* overexpression, but not *C4B*, increased synapse pruning,⁴⁸ we also tested the associations of predicted *C4A* expression with IDPs. The predicted *C4A* expression level was calculated

from the imputed haplotypes as previous work.¹⁶ Similar association patterns were observed with the associations to *C4A* copy number (Supplementary Fig. 9). These findings suggest *C4* expression level may not only affect grey matter but white matter microstructure within the brain.

Limited associations were found between KIR and KIR/HLA pairs and IDPs

To identify the association between *KIRs* and IDPs, we examined the copy number of 17 *KIR* genes as well as the *KIR* haplotypes. No associations passed the genome-wide threshold. Considering the glycoprotein products encoded by the *KIR* genes are the receptors of the HLA class I ligands, we also tested the epistatic effects of *KIR/HLA* on IDPs (see ‘Materials and methods’ section). The most significant association was between *KIR3DL1/HLA-B*27:02* and the volume of fifth ventricle under the dominant model

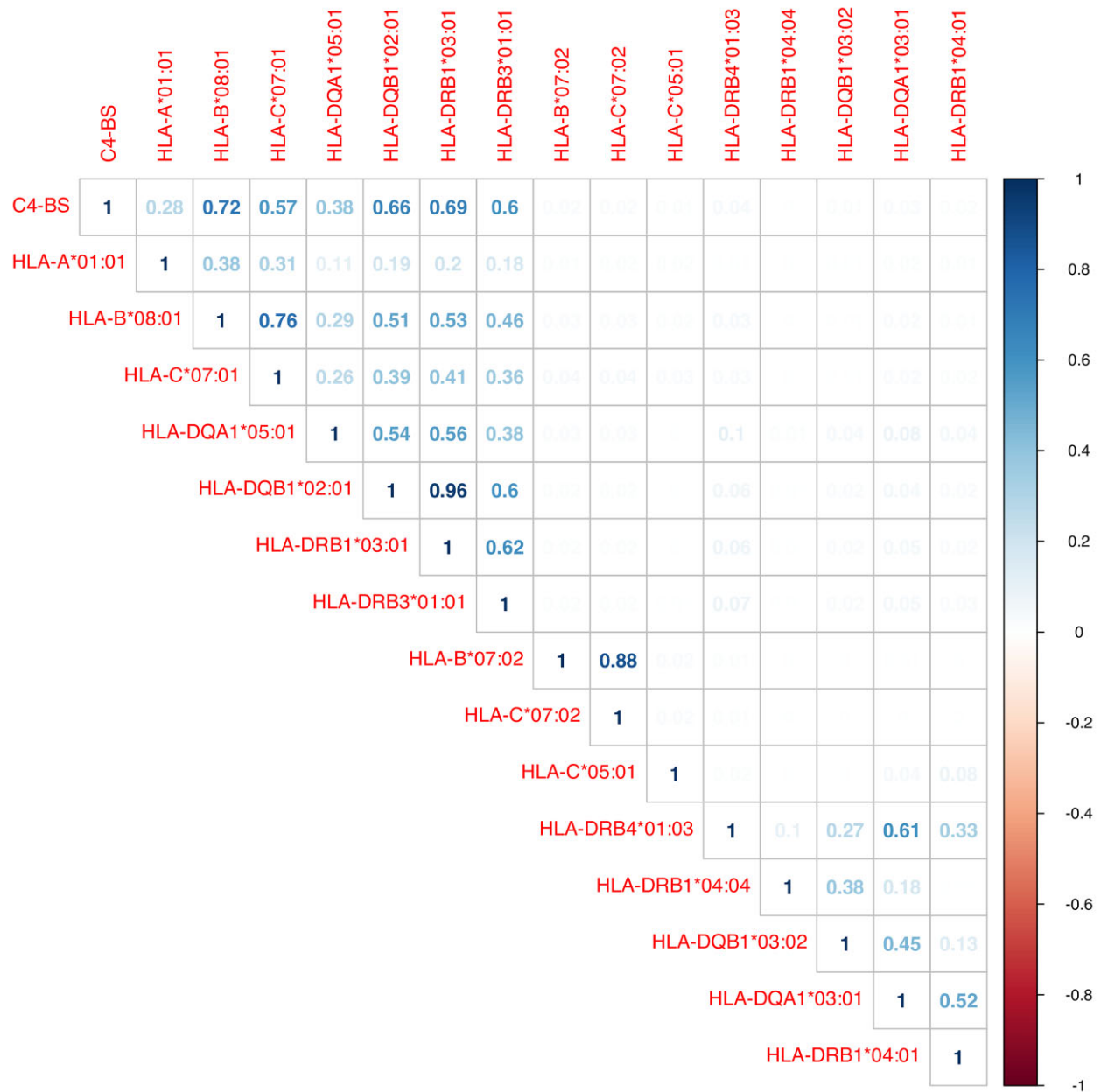


Figure 5 Estimated LD across significant HLA alleles and C4-BS. Only the upper triangular of the matrix is displayed. The rows and columns indicate the selected genetic variants. Dark blue indicates strong LD. The number in each cell indicates the LD measure between two genetic variants which is the squared correlation based on the genotype allele counts.

($P = 4.21 \times 10^{-9}$). The detailed results are in [Supplementary Note](#) and [Supplementary Tables 18, 22–23](#).

Multi-trait association analysis

The IDPs derived from the same MRI image or brain biomarker show phenotypic correlations ([Supplementary Fig. 10](#)). To increase the statistical power of detecting image-specific associations and address the phenotype redundancy problem, we performed multi-trait association analysis using TATES.³³ We identified 90 associations surviving a

Bonferroni correction for the number of immune variants and IDP groups ($0.05/(137 \times 19) = 1.9 \times 10^{-5}$), and a further 67 associations between single genetic variants and 19 IDP groups when correcting for only the number of genetic variants ($P < 3.6 \times 10^{-4}$; [Supplementary Table 24](#)). This strategy can avoid multiple testing burden especially using the large number of IDPs and identify additional associations. For example, *HLA-DRB1*01:01* was found to be associated with brain subcortical volumes, while *HLA-C*07:02*, *HLA-B*07:02* and *HLA-C*05:01* were significantly associated with T2 star metrics. As expected, we observed

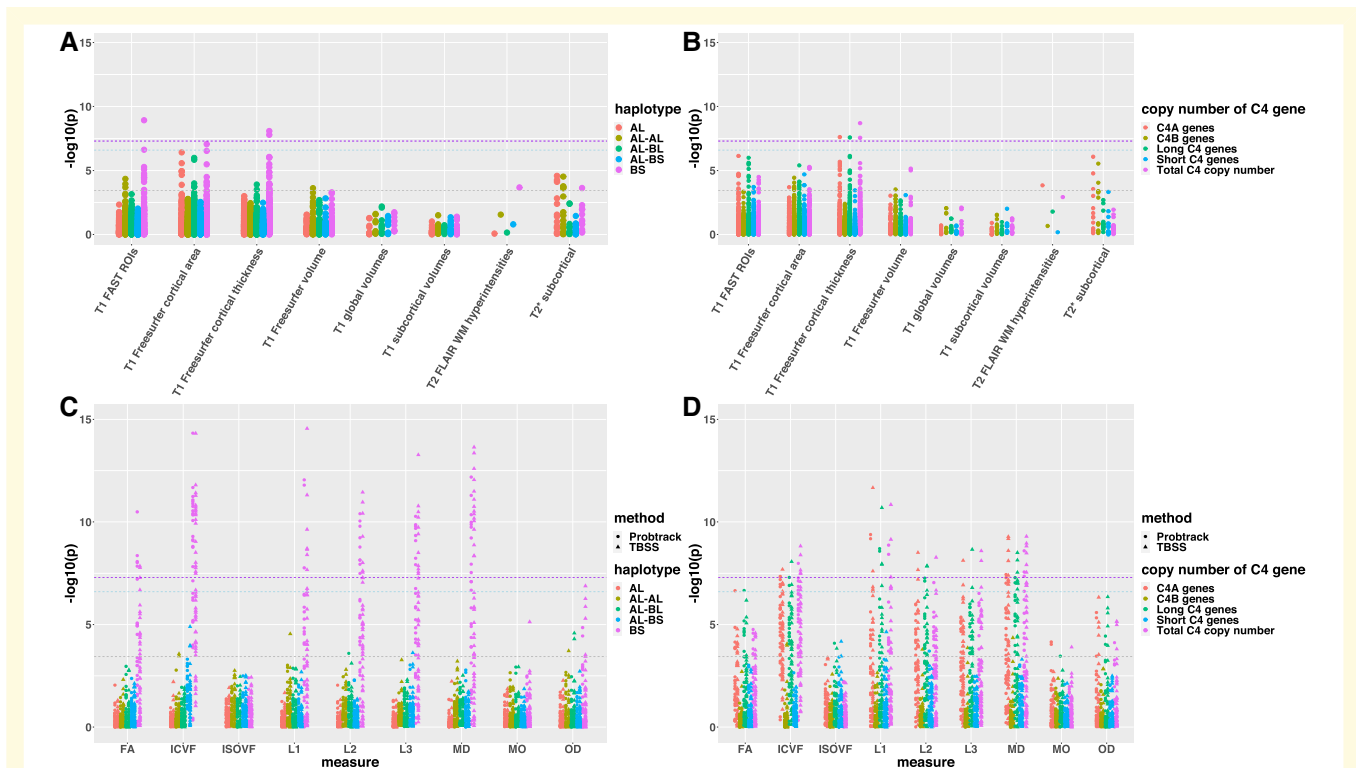


Figure 6 The associations of common *C4* haplotypes and *C4* copy number with structural and diffusion IDPs. The IDP groups are arranged on the x-axis as in Fig. 2. The strength of association ($-\log_{10}(P \text{ value})$) for individual IDP within each subgroup is plotted along the y-axis. The purple dash line indicates the genome-wide significant $-\log_{10}(P \text{ value})$ threshold of 7.3. The blue dash line indicates the conservative $-\log_{10}(P \text{ value})$ threshold of 6.6 accounting for the number of loci (137) as well as total IDPs (1422) tested. The grey dash line indicates a nominal significance $-\log_{10}(P \text{ value})$ threshold of 3.44 correcting only for the number of the tested loci (137). Different structural haplotypes, *C4* copy numbers are shown in different colours. Two statistical approaches (Probrtrack and Tract-Based Spatial Statistics (TBSS)) for estimating dMRI measurements are distinguished by circle and triangle points. **(A)** The PheWAS plot of the association between distinct *C4* haplotypes and structural IDPs. **(B)** The PheWAS plot of the association between *C4A*, *C4B*, *C4L*, *C4S* and total *C4* copy numbers and structural IDPs. **(C)** The PheWAS plot of the association between distinct *C4* haplotypes and diffusion IDPs. **(D)** The PheWAS plot of the association between *C4A*, *C4B*, *C4L*, *C4S* and total *C4* copy numbers and diffusion IDPs.

distinct genetic variants associated with different classes of IDPs. The *HLA* alleles in AH8.1 haplotype and *C4* copy number variation were found to be associated with multiple classes of IDPs, such as grey matter volumes, cortical TH, dMRI FA, ICVF and MD. In contrast, *HLA-DRB1*01:01* and *HLA-DQA1*01:01* were only associated with brain subcortical volumes. This suggested some of the genetic variants affect multiple groups of IDPs, whereas some are specific to single IDP group.

Discussion

We conducted an association analysis of brain imaging phenotypes with the key immune genes, including *HLA*, *C4*, *KIR* as well as the interactions between *HLA* and *KIR*. The *HLA* locus has been associated with a number of neuropsychiatric and neurological diseases, such as schizophrenia, Parkinson's disease, MS and Alzheimer's disease. Our study complements genetic and neuroimaging studies of disorders of the brain in that it generates new insight into the relationship between immune related genes and brain MRI data

which can help accelerate the discovery of new brain biomarkers.

Here, we identified a large number of associations between human immune-related genes and brain MRI-based phenotypes. We report 299 associations ($P < 5 \times 10^{-8}$) between immune related genes and IDPs, suggesting the immune-related genetic loci shape brain structure. These genetic associations were mainly identified with structural and dMRI phenotypes. Limited associations were observed with the fMRI data. The *KIR* locus was associated with a few phenotypes including white matter tracts in cerebrum ($5 \times 10^{-8} < P < 3.6 \times 10^{-4}$). Only a small number of epistatic associations were identified between *HLA/KIR* interactions and IDPs at the nominal significance threshold. Larger samples are needed to gain sufficient power in future study.

Our results can help elucidate how the immune related variants impact specific brain structure, some of which have previously been implicated in disorders of the brain. Notably, *HLA-B*08:01*, *HLA-DRB1*03:01* and *HLA-DQB1*02* which are a part of the AH8.1 haplotype, relating to regional grey matter volumes, cortical area and

TH in our study have been found to be associated with schizophrenia risk.¹⁶ These three alleles were also the most significant HLA allele that associated with major depression with modest effect size.³⁸ Parkinson's disease was reported to be associated with *HLA-DRB1*04* and *HLA-DQA1*03:01*.^{17,18,44} Our findings could help unveil the mechanisms through which the HLA haplotypes contribute to neuropsychiatric and neurological diseases.

The *HLA* alleles within the AH8.1 haplotype were associated with most dMRI metrics except the ISOVF measures. In contrast, *HLA-DRB1*04*, *HLA-DQB1*03:02*, *HLA-DQA1*03:01* and *HLA-DRB4* were only significantly associated with ISOVF in multiple white matter tracts (Fig. 2B; Supplementary Fig 6H–J). ISOVF is thought to measure the extra-cellular water diffusion.⁴⁹ One possible interpretation is that the MHC class II molecules present the exogenous antigens to the CD4⁺ cells (T helper cells) compared to the MHC class I molecules present the intracellular peptides to the CD8⁺ cells (killer T cells).⁵⁰ However, AH8.1 haplotype includes both class I and class II genes that exhibit strong LD. Thus, it is difficult to identify the causal variants.

The AH8.1 has been associated with many autoimmune disorders.⁵¹ Type 1 diabetes was reported to be associated with the component alleles in AH8.1, with *HLA-DQB1*02:01* and *HLA-DRB1*03:01* increased the risk of type 1 diabetes. These alleles shared associations with lupus, coeliac disease and sicca syndrome.⁵² The most significant associations of *HLA-DQB1*02:01* and *HLA-DRB1*03:01* were found with a decreased grey matter volume in the left thalamus. However, whether the autoimmune disorders share the association with brain morphology especially in some specific brain regions is largely unknown.

The higher *C4A* copy number has been found to be associated with higher schizophrenia risk, while the C4-BS allele is a protective factor of schizophrenia.¹⁶ A recent study further demonstrated the overexpression of *C4A* gene promotes the excessive synaptic loss and abnormal behaviour in mice.⁴⁸ Interestingly, they found that the loss of *C4* does not affect the normal developmental synaptic pruning.⁴⁸ At the nominal level of the significance threshold, C4-BS was associated with the increased cortical TH and grey matter volume across different brain regions especially the prefrontal cortex. The deficits of grey matter volume in prefrontal cortex were reported for some psychiatric disorders such as schizophrenia,⁵ bipolar disorder^{53,54} and major depression.⁵⁴ Our findings may suggest C4-BS is a protective factor of the development of these diseases. On the other hand, we observed C4-BS was associated with higher FA, ICVF and OD but lower MD across multiple white matter tracts. Previous literature demonstrated the associations of ageing with microstructural connections within the white-matter quantified by using dMRI.⁵⁵ Their results showed ageing was significantly associated with lower FA, ICVF and OD, with higher MD and ISOVF. Combined with our results, it suggests that C4-BS might be a protective factor of brain

ageing. The strong LD between C4-BS and AH8.1 makes it difficult to find out the causal factor. However, the evidence that C4 does not play a role in normal development of cortical synapses⁴⁸ suggested other mechanisms exist. One possible mechanism is through the MHC class I molecules which have been demonstrated to play a role in synaptic pruning during brain development.^{41,56} Previous findings implicated the expression and function of MHC I molecules in the regulation of synaptic plasticity, as well as their role in the aged brain.⁴² Future studies are needed to further understand each specific gene function, when and where they play a role. This requires looking at dynamic changes of brain spatial-temporal molecular characteristics using integrative genomics data.

We observed *C4A*, *C4L* and total *C4* copy numbers to be associated with many IDPs. For example, higher *C4A* copy number was associated with thinner cortical TH in some regions (Supplementary Table 17). Higher *C4A* copy number correlates with higher *C4A* expression.¹⁶ This may suggest that increased *C4A* expression plays a role in shaping the grey matter structure. In line with the findings of *C4* in schizophrenia, the associating thinner regional cortical TH might indicate increased expression of *C4A* as a disease mechanism. In addition to the associations with grey matter, the *C4A* copy number was also associated with the measures in dMRI across multiple white matter tracts. The alterations of the structure of white matter have been associated with many neurological diseases.^{57–59} But whether *C4* affects the risk of other diseases like Parkinson's disease, amyotrophic lateral sclerosis has not been investigated yet using large-scale genetic data. Our results provide new insights into the potential role of *C4* on brain endophenotypes.

The KIR complex encoding the receptors of the MHC class I ligands is understudied in neurological diseases. These receptors including stimulatory and inhibitory normally work along with their corresponding MHC class I molecules. The first neurological disease found to be associated with KIR/HLA pairs was MS.⁶⁰ Using the genetic epistatic model, we found only a small number of associations reached the nominal threshold. It is still not clear whether the interactions between *KIR* and *HLA* have influences on brain-related traits. Further exploration using clinical diseases as phenotypes is a direction of future study.

It is worth noting that the shared genetic associations between brain-MRI phenotypes and clinical diseases might be driven by the same brain regions. For example, we identified AH8.1 haplotype and predicted *C4A* gene expression level to be associated with brain imaging phenotypes, including grey matter volume in thalamus, cortical TH of insula and white matter structure in the superior longitudinal fasciculus. These genetic variants and brain regions were also reported to be associated with schizophrenia.^{16,61–63} Interestingly, insula is thought to play a role in brain-immune interaction.^{47,64} Loss of grey matter in insula was found to be associated with schizophrenia.⁶⁵ *HLA-A*03:01*~*HLA*

-B*07:02 haplotype was found to be associated with Alzheimer's disease. In our study, we identified the association between HLA-B*07:02 and T2* in hippocampus, which is affected at the early stage in Alzheimer's disease. Overall, our study contributes to identifying important brain regions associated with both immune genes and brain related diseases.

Our work still has several limitations. These association results were based on the imputed genetic variants. The HLA alleles can be imputed accurately given that many associating HLA alleles have been replicated well. However, the accuracy of the KIR imputation can be sensitive to the overlapping SNPs between the genotyping data and the reference panel. We controlled for several known or putative imaging confounders in the association analysis, following some recent UKB analyses, though other unknown of unmeasured confounders may exist. In addition, the long-range LD patterns in the HLA region make it difficult to fine-map the causal variants. A comprehensive genetic map as well as functional annotations is needed to improve the understanding of the disease mechanisms caused by this complex locus. We did not replicate these associations due to the lack of brain imaging data from other independent cohorts. Lastly, only European descents were included in our study because of the limited representation of participants from different ancestry backgrounds in UK Biobank imaging study. Despite these limitations, our study generates new insights into the immunogenetics of human brain structure.

Acknowledgements

We are grateful for the access to the UK Biobank Resource under Application Number 12505. We thank Julia Sidorenko for UKB data extraction and Dr. Longda Jiang for helpful discussion.

Funding

This work was supported by a C.J. Martin Fellowship to B.C.-D. (APP1161356), an National Health and Medical Research Council (NHMRC) Fellowship to N.R.W. (1173790, 1113400) and an Australian Research Council (ARC) Future Fellowship to A.F.M. (FT200100837).

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain Communications* online.

References

- Johnson KA, Fox NC, Sperling RA, Klunk WE. Brain imaging in Alzheimer disease. *Cold Spring Harb Perspect Med.* 2012;2(4):a006213–a006213.
- Pearlson GD, Marsh L. Structural brain imaging in schizophrenia: A selective review. *Biological Psychiatry.* 1999;46:627–649.
- Menke RA, Scholz J, Miller KL, et al. MRI characteristics of the substantia nigra in Parkinson's disease: A combined quantitative T1 and DTI study. *Neuroimage.* 2009;47(2):435–441.
- Baron-Cohen S, Ring HA, Wheelwright S, et al. Social intelligence in the normal and autistic brain: An fMRI study. *Eur J Neurosci.* 1999; 11(6):1891–1898.
- Cannon TD, Thompson PM, Van Erp TGM, et al. Cortex mapping reveals regionally specific patterns of genetic and disease-specific gray-matter deficits in twins discordant for schizophrenia. *Proc Natl Acad Sci U S A.* 2002;99(5):3228–3233.
- Nugent AC, Milham MP, Bain EE, et al. Cortical abnormalities in bipolar disorder investigated with MRI and voxel-based morphometry. *Neuroimage.* 2006;30(2):485–497.
- De Groot M, Ikram MA, Akoudad S, et al. Tract-specific white matter degeneration in aging: The Rotterdam study. *Alzheimers Dement.* 2015;11(3):321–330.
- Filippi M, Cercignani M, Inglese M, Horsfield MA, Comi G. Diffusion tensor magnetic resonance imaging in multiple sclerosis. *Neurology.* 2001;56(3):304–311.
- Ellis CM, Simmons A, Jones DK, et al. Diffusion tensor MRI assesses corticospinal tract damage in ALS. *Neurology.* 1999;53(5):1051.
- Chan LL, Rumpel H, Yap K, et al. Case control study of diffusion tensor imaging in Parkinson's disease. *J Neurol Neurosurg Psychiatry.* 2007;78(12):1383–1386.
- Thompson PM, Stein JL, Medland SE, et al. The ENIGMA consortium: Large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging Behav.* 2014;8(2):153–182.
- Bycroft C, Freeman C, Petkova D, et al. The UK biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562(7726): 203–209.
- Elliott LT, Sharp K, Alfaro-Almagro F, et al. Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature.* 2018;562(7726):210–216.
- Grasby KL, Jahanshad N, Painter JN, et al. The genetic architecture of the human cerebral cortex. *Science.* 2020;367(6484):eaay6690.
- Misra MK, Damotte V, Hollenbach JA. The immunogenetics of neurological disease. *Immunology.* 2018;153(4):399–414.
- Sekar A, Bialas AR, De Rivera H, et al. Schizophrenia risk from complex variation of complement component 4. *Nature.* 2016; 530(7589):177–183.
- Wissemann WT, Hill-Burns EM, Zabetian CP, et al. Association of parkinson disease with structural and regulatory variants in the hla region. *Am J Hum Genet.* 2013;93(5):984–993.
- Hollenbach JA, Norman PJ, Creary LE, et al. A specific amino acid motif of HLA-DRB1 mediates risk and interacts with smoking history in Parkinson's disease. *Proc Natl Acad Sci U S A.* 2019; 116(15):7419–7424.
- Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat Genet.* 2019;51(3): 414–430.
- Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet.* 2019;51(3):404–413.
- Lincoln MR, Montpetit A, Cader MZ, et al. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nat Genet.* 2005;37(10):1108–1112.
- Patsopoulos NA, Baranzini SE, Santaniello A, et al. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science.* 2019;365(6460):eaav7188.

23. Vukcevic D, Traherne JA, Næss S, *et al.* Imputation of KIR types from SNP variation data. *Am J Hum Genet.* 2015;97(4):593–607.
24. Parham P. MHC class I molecules and KIRS in human history, health and survival. *Nat Rev Immunol.* 2005;5(3):201–214.
25. Alfaro-Almagro F, Jenkinson M, Bangerter NK, *et al.* Image processing and quality control for the first 10 000 brain imaging datasets from UK biobank. *Neuroimage.* 2018;166:400–424.
26. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011;88:76–82.
27. Diltthey A, Leslie S, Moutsianas L, *et al.* Multi-population classical HLA type imputation. *PLoS Comput Biol.* 2013;9(2):e1002877.
28. Kamitaki N, Sekar A, Handsaker RE, *et al.* Complement genes contribute sex-biased vulnerability in diverse disorders. *Nature* 2020; 582(7813):577–581.
29. Purcell S, Neale B, Todd-Brown K, *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81(3):559–575.
30. Das S, Forer L, Schönherr S, *et al.* Next-generation genotype imputation service and methods. *Nat Genet.* 2016;48(10):1284–1287.
31. McCarthy S, Das S, Kretzschmar W, *et al.* A reference panel of 64 976 haplotypes for genotype imputation. *Nat Genet.* 2016;48(10):1279–1283.
32. Hanson AL, Vukcevic D, Leslie S, *et al.* Epistatic interactions between killer immunoglobulin-like receptors and human leukocyte antigen ligands are associated with ankylosing spondylitis. *PLoS Genet.* 2020;16(8):e1008906.
33. van der Sluis S, Posthuma D, Dolan C V. TATES: Efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLoS Genet.* 2013;9(1):e1003235.
34. Turley P, Walters RK, Maghziyan O, *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet.* 2018;50(2):229–237.
35. Grotzinger AD, Rhemtulla M, de Vlaming R, *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav.* 2019;3(5):513–525.
36. Bulik-Sullivan B, Loh PR, Finucane HK, *et al.* LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet.* 2015;47(3):291–295.
37. Elliott ML, Knodt AR, Ireland D, *et al.* What is the test-retest reliability of common task-functional MRI Measures? New empirical evidence and a meta-analysis. *Psychol Sci.* 2020;31(7):792–806.
38. Glanville KP, Coleman JRI, Hanscombe KB, *et al.* Classical human leukocyte antigen alleles and C4 haplotypes are not significantly associated with depression. *Biol Psychiatry.* 2020;87(5):419–430.
39. Steele NZR, Carr JS, Bonham LW, *et al.* Fine-mapping of the human leukocyte antigen locus as a risk factor for Alzheimer disease: A case-control study. *PLoS Med.* 2017;14(3):e1002272.
40. Zhao B, Li T, Yang Y, *et al.* Common genetic variation influencing human white matter microstructure. *Science.* 2021;372(6548):eabf3736.
41. Lee H, Brott BK, Kirkby LA, *et al.* Synapse elimination and learning rules co-regulated by MHC class I H2-Db. *Nature.* 2014;509(7499):195–200.
42. Elmer BM, McAllister AK. Major histocompatibility complex class I proteins in brain development and plasticity. *Trends Neurosci.* 2012;35(11):660–670.
43. Lazarczyk MJ, Kemmler JE, Eyford BA, *et al.* Major histocompatibility complex class I proteins are critical for maintaining neuronal structural complexity in the aging brain. *Sci Rep.* 2016; 6(1):26199.
44. Yu E, Ambati A, Stolp Andersen M, *et al.* Fine mapping of the HLA locus in Parkinson's disease in Europeans. *medRxiv.* 2020;7(1):84.
45. Horton R, Gibson R, Coggill P, *et al.* Variation analysis and gene annotation of eight MHC haplotypes: The MHC haplotype project. *Immunogenetics.* 2008;60(1):1–18.
46. Degli-Esposti MA, Leaver AL, Christiansen FT, Witt CS, Abraham LJ, Dawkins RL. Ancestral haplotypes: Conserved population MHC haplotypes. *Hum Immunol.* 1992;34(4):242–252.
47. Koren T, Yifa R, Amer M, *et al.* Insular cortex neurons encode and retrieve specific immune responses. *Cell.* 2021;184(24):5902–5915.e17.
48. Yilmaz M, Yalcin E, Presumey J, *et al.* Overexpression of schizophrenia susceptibility factor human complement C4A promotes excessive synaptic loss and behavioral changes in mice. *Nat Neurosci.* 2021;24(2):214–224.
49. Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage.* 2012;61(4):1000–1016.
50. Neefjes J, Jongsma MLM, Paul P, Bakke O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol.* 2011;11(12):823–836.
51. Price P, Witt C, Allcock R, *et al.* The genetic basis for the association of the 8.1 ancestral haplotype (A1, B8, DR3) with multiple immunopathological diseases. *Immunol Rev.* 1999;167(1):257–274.
52. Karnes JH, Bastarache L, Shaffer CM, *et al.* Phenome-wide scanning identifies multiple diseases and disease severity phenotypes associated with HLA variants. *Sci Transl Med.* 2017;9(389):1–14.
53. Selvaraj S, Arnone D, Job D, *et al.* Grey matter differences in bipolar disorder: A meta-analysis of voxel-based morphometry studies. *Bipolar Disord.* 2012;14(2):135–145.
54. Wise T, Radua J, Via E, *et al.* Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: Evidence from voxel-based meta-analysis. *Mol Psychiatry.* 2017; 22(10):1455–1463.
55. Cox SR, Ritchie SJ, Tucker-Drob EM, *et al.* Ageing and brain white matter structure in 3513 UK Biobank participants. *Nat Commun.* 2016;7:1–13.
56. Huh GS, Boulanger LM, Du H, Riquelme PA, Brotz TM, Shatz CJ. Functional requirement for class I MHC in CNS development and plasticity. *Science.* 2000;290(5499):2155–2159.
57. Koshiyama D, Fukunaga M, Okada N, *et al.* White matter microstructural alterations across four major psychiatric disorders: Mega-analysis study in 2937 individuals. *Mol Psychiatry.* 2020; 25(4):883–895.
58. Hattori T, Orimo S, Aoki S, *et al.* Cognitive status correlates with white matter alteration in Parkinson's disease. *Hum Brain. Mapp* 2012;33(3):727–739.
59. Damoiseaux JS, Smith SM, Witter MP, *et al.* White matter tract integrity in aging and Alzheimer's disease. *Hum Brain. Mapp* 2009; 30(4):1051–1059.
60. Lorentzen ÅR, Karlsen TH, Olsson M, *et al.* Killer immunoglobulin-like receptor ligand HLA-Bw4 protects against multiple sclerosis. *Ann Neurol* 2009;65(6):658–666.
61. Wylie KP, Tregellas JR. The role of the insula in schizophrenia. *Schizophr Res.* 2010;123(2-3):93–104.
62. Pergola G, Selvaggi P, Trizio S, Bertolino A, Blasi G. The role of the thalamus in schizophrenia from a neuroimaging perspective. *Neurosci Biobehav Rev.* 2015;54:57–75.
63. Karlsgodt KH, van Erp TGM, Poldrack RA, Bearden CE, Nuechterlein KH, Cannon TD. Diffusion tensor imaging of the superior longitudinal fasciculus and working memory in recent-onset schizophrenia. *Biol Psychiatry.* 2008;63(5):512–518.
64. Schiller M, Ben-Shaan TL, Rolls A. Neuronal regulation of immunity: Why, how and where? *Nat Rev Immunol.* 2021;21(1):20–36.
65. Wright IC, Ellison ZR, Sharma T, Friston KJ, Murray RM, McGuire PK. Mapping of grey matter changes in schizophrenia. *Schizophr Res.* 1999;35(1):1–14.