BRAIN
ORIGINAL ARTICLE

The BS variant of C4 protects against age-related loss of white matter microstructural integrity

Matthew Traylor,[1,2,3] Rainer Malik,[4] Benno Gesierich[4] and Martin Dichgans[4,5,6]

Age-related loss of white matter microstructural integrity is a major determinant of cognitive decline, dementia and gait disorders. However, the mechanisms and molecular pathways that contribute to this loss of integrity remain elusive.

We performed a genome-wide association study of white matter microstructural integrity as quantified by diffusion MRI metrics (mean diffusivity and fractional anisotropy) in up to 31 128 individuals from UK Biobank (age 45–81 years) based on a two degrees of freedom (2df) test of single nucleotide polymorphism (SNP) and SNP × Age effects.

We identified 18 loci that were associated at genome-wide significance with either mean diffusivity ($n = 16$) or fractional anisotropy ($n = 6$). Among the top loci was a region on chromosome 6 encoding the human major histocompatibility complex (MHC). Variants in the MHC region were strongly associated with both mean diffusivity [best SNP: 6:28866209_TTTTG_T, beta (standard error, SE) = –0.069 (0.009); 2df $P = 6.5 \times 10^{-15}$] and fractional anisotropy [best SNP: rs3129787, beta (SE) = –0.056 (0.008); 2df $P = 3.5 \times 10^{-12}$]. Of the imputed human leukocyte antigen (HLA) alleles and complement component 4 (C4) structural haplotype variants in the human MHC, the strongest association was with the C4-BS variant [for mean diffusivity: beta (SE) = –0.070 (0.010); $P = 2.7 \times 10^{-11}$; for fractional anisotropy: beta (SE) = –0.054 (0.011); $P = 1.6 \times 10^{-7}$]. After conditioning on C4-BS no associations with HLA alleles remained significant. The protective influence of C4-BS was stronger in older participants [age $\geqslant$ 65; interaction $P = 0.0019$ (mean diffusivity), $P = 0.015$ (fractional anisotropy)] and in participants without a history of smoking [interaction $P = 0.00093$ (mean diffusivity), $P = 0.021$ (fractional anisotropy)].

Taken together, our findings demonstrate a role of the complement system and of gene–environment interactions in age-related loss of white matter microstructural integrity.

1 Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, UK
2 The Barts Heart Centre and NIHR Barts Biomedical Research Centre—Barts Health NHS Trust, The William Harvey Research Institute, Queen Mary University London, London, UK
3 Novo Nordisk Research Centre Oxford, Oxford, UK
4 Institute for Stroke and Dementia Research (ISD), University Hospital, LMU Munich, Munich, Germany
5 German Centre for Neurodegenerative Diseases (DZNE, Munich), Munich, Germany
6 Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

Correspondence to: Matthew Traylor, PhD
John Vane Science Centre, William Harvey Research Institute
Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK
E-mail: m.traylor@qmul.ac.uk

Received February 23, 2021. Revised May 12, 2021. Accepted June 14, 2021. Advance access publication August 6, 2021

Correspondence may also be addressed to: Martin Dichgans, MD
Institute for Stroke and Dementia Research, LMU Hospital
Ludwig-Maximilians University, Munich, Germany
E-mail: martin.dichgans@med.uni-muenchen.de

Keywords: genetics; white matter microstructure; complement

Abbreviations: DTI = diffusion tensor MRI; FA = fractional anisotropy; HLA = human leukocyte antigen; LOWMI = loss of white matter microstructural integrity; MD = mean diffusivity; MHC = major histocompatibility complex; SNP = single nucleotide polymorphism; WMH = white matter hyperintensities

Introduction

Age-related loss of white matter microstructural integrity (LOWMI) is a major determinant of cognitive decline, dementia[1,2] and gait disorders[3] and has been linked to both vascular and neurodegenerative pathologies.[4,5] Histological findings include changes in axonal structure and integrity[6] as well as demyelination and inflammation including microglial activation.[7] However, the precise mechanisms and molecular pathways underlying these changes remain elusive.

Insights into white matter microstructure in humans have predominantly arisen from studies of diffusion tensor MRI (DTI) using mean diffusivity (MD) and fractional anisotropy (FA), both of which are sensitive to subtle alterations in white matter microstructure.[8] Previous genome-wide association studies (GWAS) using these metrics in up to 20 000 individuals have identified multiple loci for brain microstructure.[9,10] However, interpreting these findings remains challenging as the respective loci could reflect either age-related or developmental processes. Isolating the loci underlying age-related LOWMI is critical to identify strong candidates that could be taken forward to functional studies. To identify genetic loci that specifically influence age-related LOWMI, and are therefore most informative for mechanisms underlying dementia and cognitive decline, we here perform a genome-wide study of DTI metrics using a two degrees of freedom (2df) test incorporating both single nucleotide polymorphism (SNP) effects and SNP × Age interaction effects followed by filtering on variants showing significant interactions in consistent direction with SNP effects. We further expand this approach to the analysis of gene–environment interactions. On identifying an association of both MD and FA with variants in the extended human major histocompatibility complex (MHC), a gene complex encoding cell-surface proteins involved in innate and adaptive immunity, we focus on this region in further analysis due to its potential to provide novel mechanistic insights.[11]

Genetic variation within the human MHC determines the class II human leukocyte antigens (HLA) A, B and C and class II HLAs DP, DQ and DR. The human MHC class III region encodes complement component 2, complement factor B and complement component 4 (C4) paralogue genes *C4A* and *C4B*,[12] all of which have important roles in the classical complement pathway. The last is a critical constituent of the innate immune system primarily involved in scavenging and elimination of pathogens and antigens. The complement system has an important role in maintaining blood–brain barrier integrity.[13] Moreover, on compromise of the blood–brain barrier, complement factors that are synthesized in the liver can enter the brain, where they activate microglia and induce inflammation.[14,15] In the classical complement pathway, C4 promotes C3 activation, leading to generation of proinflammatory mediators C3a and C5a.[16] The C4 paralogue genes *C4A* and *C4B* differ in amino acid sequence only at one site encoded by exon 26,[17] exist in both long (L) and short (S) forms that differ depending on whether they contain a human endogenous retrovirus sequence in intron 9, and arise due to genetic variation within the human MHC. This structural variation has been associated with disorders in which complement disruption influences innate immunity such as systemic lupus erythematosus and Sjögren's syndrome,[18] and with psychiatric disorders such as schizophrenia, where reduced C4A expression is believed to disrupt synaptic pruning in brain development.[19,20]

Here we show that the BS variant of C4 underlies the observed association of the MHC region with age-related LOWMI, thus demonstrating a role of specific constituents of the complement system. As smoking can activate complement,[21] leads to endothelial dysfunction, disrupts blood–brain barrier integrity,[22] and has been associated with reduced white matter microstructural integrity,[23,24] we further investigate whether history of smoking modifies the association of C4-BS with age-related LOWMI. We find evidence that the protective effect of C4-BS is greatest in those without history of smoking.

Materials and methods

Study dataset

All analyses were performed using the UK Biobank dataset, a population biobank resource including ∼500 000 participants from across the UK, aged between 40 and 69 years at recruitment.[25] The UK Biobank includes clinical and phenotypic information for a broad range of traits and includes MRI data on a subset of participants. This study used the January 2020 release of UK Biobank imaging data on ∼43 000 individuals.[26,27] MRI was performed on two identical Siemens Skyra 3.0 T scanners (Siemens Medical Solutions), running VD13A SP4, with a standard Siemens 32-channel RF receiver head coil. Identical acquisition parameters and careful quality control was used for all scans. We selected individuals described for the diffusion MRI-based phenotypes described next.

Neuroimaging markers of white matter microstructural integrity

To assess age-related LOWMI in UK Biobank participants, we used the DTI parameters FA and MD (Fig. 1A), which measure diffusion properties of water molecules in brain tissue.[28] Increases in water diffusion (MD) and a decrease in directional diffusion along brain tracts (FA) due to white matter microstructural damage occur in both vascular and neurodegenerative pathologies. These DTI parameters were generated by an image-processing pipeline developed and run on behalf of UK Biobank (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf) and were available as

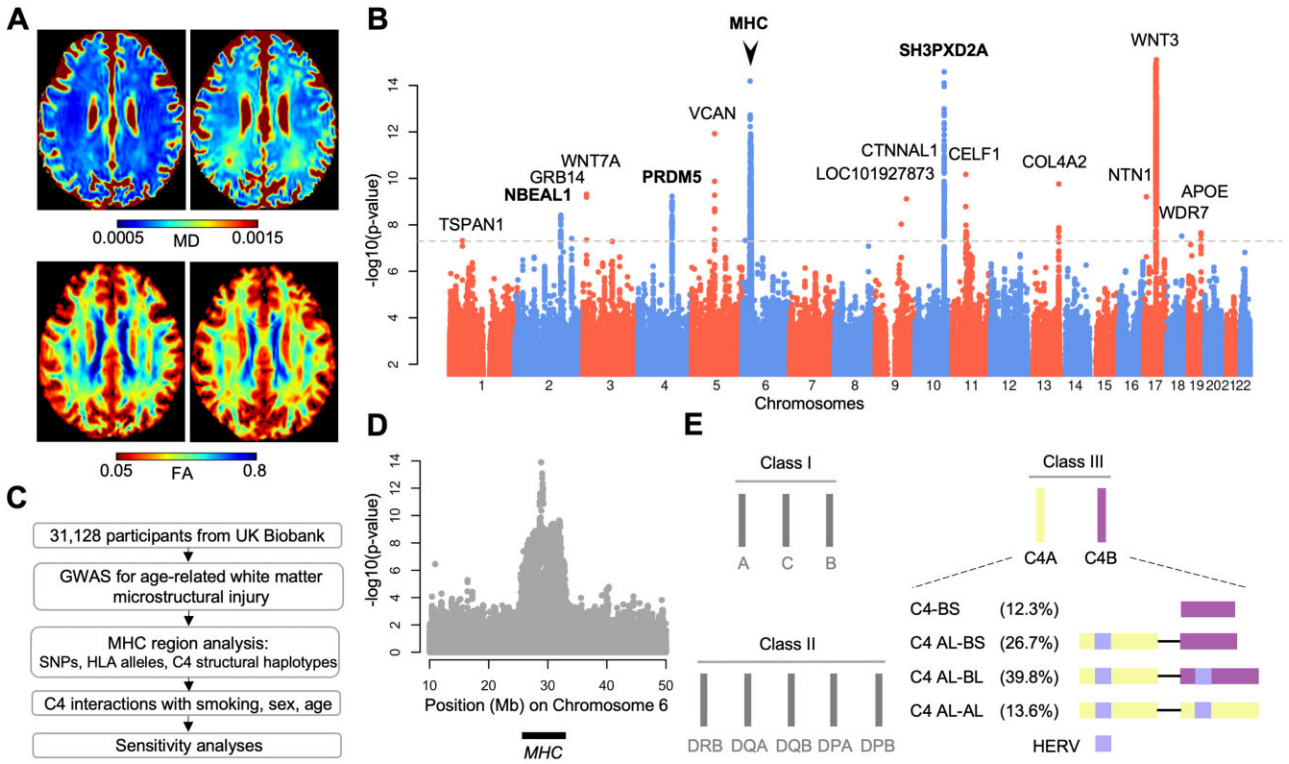

Figure 1 The MHC gene region is a major risk locus for age-related loss of white matter microstructural integrity. (**A**) Exemplar of two individuals with low (*left*) and moderate (*right*) levels of loss of white matter microstructural integrity as measured by MD (*top*) and FA (*bottom*). (**B**) Manhattan plot of $-\log_{10}$(P-value) against genomic position for MD using a 2df test of SNP and SNP × Age effects, with arrow indicating the MHC locus association on chromosome 6 in 31 128 individuals from the UK Biobank. Loci also reaching genome-wide significance for association with FA are shown in bold, the dotted grey line indicates genome-wide significance ($P < 5 \times 10^{-8}$). (**C**) Schematic of study overview. (**D**) Regional plot showing association within the MHC region on chromosome 6. (**E**) Class I and class II HLA alleles, and structural variants of complement component 4 (C4) in class III. The C4 paralogue genes C4A and C4B differ only at one site and exist in both long (L) and short (S) forms that differ based on presence or absence of a human endogenous retrovirus (HERV) sequence in intron 9, and arise due to genetic variation within the human MHC. The most common four of these variants, which comprise 92.4% of those in this UK Biobank population, are shown with their corresponding frequencies.

part of the UK Biobank central analysis for 48 individual white matter regions (FA, fields 25056–25103; and MD, fields 25104–25151).[26,27] To obtain a single global measure of global white matter FA and MD from the DTI images, we performed principal component analysis on centred and scaled data using the R prcomp function on the FA and MD measures of each of the 48 different brain regions analysed, and extracted the first PC. To improve signal to noise ratio, we excluded brain regions with low loadings (eigenvectors $\leqslant 0.1$) and repeated principal component calculation. For each resulting DTI parameter, outliers outside the ±8 standard deviation (SD) range were removed.

Genetic data, HLA alleles and imputation of C4 structural haplotypes

We used genome-wide genetic data imputed to the Haplotype Reference Consortium build by the UK Biobank core analytical team, the quality control and imputation of which have been previously described.[29] HLA alleles were imputed by the UK Biobank core analytical team using HLA*IMP:02,[30] based on multipopulation reference panels. HLA alleles were provided at four-digit resolution in 11 HLA genes: HLA-A, HLA-B and HLA-C in the classical class I MHC and HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DQA1, HLA-DQB1, HLA-DPA1 and HLA-DPB1 in the classical class II MHC. We considered only HLA alleles observed at frequency >1% in the UK Biobank population.

C4 structural haplotype variants were imputed using an established approach based on existing C4 reference panels (https://github.com/freeseek/imputec4).[20] BEAGLE v.4.1 was used to impute C4 as a multi-allelic variant from variants within the MHC region (chr6: 26–34 Mb).[31] Default BEAGLE parameters were used with two key exceptions: we used the GRCh37 PLINK recombination map, and we set the output to include genotype probability for correct downstream probabilistic estimation of C4A and C4B joint dosages. Common C4 structural haplotypes were imputed with good quality in ranges comparable to previous analyses.[32]

Genome-wide association analysis

We performed genome-wide association analysis of autosomes using a linear regression model including SNPs, SNPs × age, age, age-squared, sex and 20 ancestry-informative principal components. The primary association statistic was based on a 2df test of the SNP and SNP × Age terms. To extract only variants showing more prominent effects in ageing, we considered only those showing significant SNP × Age interaction effects ($P < 0.05$) and with consistent SNP and SNP × Age effect directions. Before analysis we excluded individuals who were not classified as white British (field 22006), related individuals with a KING kinship coefficient $\geqslant 0.0884$ (to keep only one individual per group of up to second-degree relationships), those who had withdrawn from the study and those who had history of stroke, multiple sclerosis, or other neurodegenerative disease (self-reported field 20002 codes 1081, 1086, 1491, 1583, 1261, 1262, 1263, 1397; ICD10 fields 41202 and 41204 codes I60, I61, I63, I64, G35, G20, F00-F03, G30-G32, G36, G37). Association analysis was performed using plink v2.00a3LM,

Table 1 Eighteen loci are associated with age-related white matter MD and FA

SNP	CHR:BP	Gene	EA/OA	EAF	SNP Beta (SE); P-value	SNP × age interaction Beta (SE); P-value	2df P-value
MD							
rs68179532	1:46641424	*TSPAN1*	A/G	0.071	–0.076 (0.014); 8.1×10^{-8}	–0.031 (0.014); 0.030	4.8×10^{-8}
rs6707357	2:165014293	*GRB14*	T/C	0.45	–0.038 (0.0073); 2.3×10^{-7}	–0.025 (0.0073); 0.00054	3.8×10^{-9}
rs140244541	2:203808532	*ICA1L-WDR12-CARF-NBEAL1*	A/G	0.16	–0.050 (0.0099); 4.6×10^{-7}	–0.028 (0.0098); 0.0036	3.8×10^{-8}
rs6442409	3:13835087	*WNT7A*	T/C	0.35	–0.047 (0.0076); 1.1×10^{-9}	–0.018 (0.0076); 0.018	4.8×10^{-10}
4:121740591_CCATGGCA_C	4:121740591	*PRDM5*	C/CCATGGCA	0.33	–0.047 (0.0077); 1.4×10^{-9}	–0.019 (0.0077); 0.014	5.9×10^{-10}
rs310528	5:82853209	*VCAN*	A/G	0.23	0.061 (0.0086); 1.3×10^{-12}	0.018 (0.0086); 0.038	1.2×10^{-12}
6:28866209_TTTTG_T	6:28866209	*MHC*	T/TTTTG	0.21	–0.069 (0.0089); 1.1×10^{-14}	–0.022 (0.0089); 0.015	6.5×10^{-15}
rs28361101	9:111775756	*CTNNAL1*	C/G	0.026	–0.14 (0.023); 7.4×10^{-10}	–0.047 (0.023); 0.036	7.6×10^{-10}
rs151005021	9:93178106	*LOC101927873*	A/G	0.45	–0.038 (0.0080); 1.4×10^{-6}	–0.030 (0.0080); 0.00021	9.3×10^{-9}
rs2863994	10:105454043	*SH3PXD2A*	G/T	0.49	–0.050 (0.0073); 6.9×10^{-12}	–0.031 (0.0073); 2.4×10^{-5}	7.9×10^{-15}
rs75203562	11:47556558	*CELF1*	T/C	0.015	0.17 (0.031); 6.1×10^{-8}	0.12 (0.032); 9.0×10^{-5}	6.8×10^{-11}
rs11838776	13:111040681	*COL4A2*	A/G	0.28	0.044 (0.0081); 5.7×10^{-8}	0.032 (0.0081); 7.9×10^{-5}	1.7×10^{-10}
rs199501	17:44862613	*NSF-WNT3*	A/G	0.23	0.066 (0.0085); 6.8×10^{-15}	0.025 (0.0084); 0.0026	7.8×10^{-16}
rs117107596	17:9026070	*NTN1*	G/T	0.010	0.12(0.037); 9.6×10^{-4}	0.19 (0.035); 1.3×10^{-7}	6.2×10^{-10}
rs141200613	18:54658644	*WDR7*	G/A	0.046	0.076 (0.018); 3.2×10^{-5}	0.077 (0.018); 2.3×10^{-5}	3.0×10^{-8}
rs12721046	19:45421254	*APOE*	A/G	0.16	0.054 (0.010); 7.9×10^{-8}	0.027 (0.010); 0.0064	2.2×10^{-8}
FA							
rs76122535	2:203665929	*ICA1L-WDR12-CARF-NBEAL1*	G/C	0.13	–0.069 (0.011); 3.5×10^{-10}	–0.038 (0.011); 0.00038	4.3×10^{-12}
rs776795707	4:121746846	*PRDM5*	G/GTTGT	0.31	–0.047 (0.0081); 8.6×10^{-9}	–0.016 (0.008); 0.047	1.0×10^{-8}
rs13190020	5:65012526	*SGTB-NLN*	A/G	0.34	–0.048 (0.0079); 1.3×10^{-9}	–0.023 (0.0078); 0.0041	1.8×10^{-10}
5:82716332_GA_G	5:82716332	*XRCC4*	G/GA	0.48	0.043 (0.0078); 5.7×10^{-8}	0.018 (0.0079); 0.022	2.6×10^{-8}
rs3129787	6:29048689	*MHC*	T/A	0.29	–0.056 (0.0083); 1.2×10^{-11}	–0.022 (0.0083); 0.0083	3.5×10^{-12}
rs57694670	10:105447782	*SH3PXD2A*	G/A	0.41	–0.043 (0.0077); 1.9×10^{-8}	–0.026 (0.0077); 0.00069	3.7×10^{-10}

Listed are the 18 independent loci reaching genome-wide significance for MD or FA. CHR:BP = chromosome and position on human genome assembly GRCh37/hg19; EA = effect allele; EAF = effect allele frequency; OA = other allele.

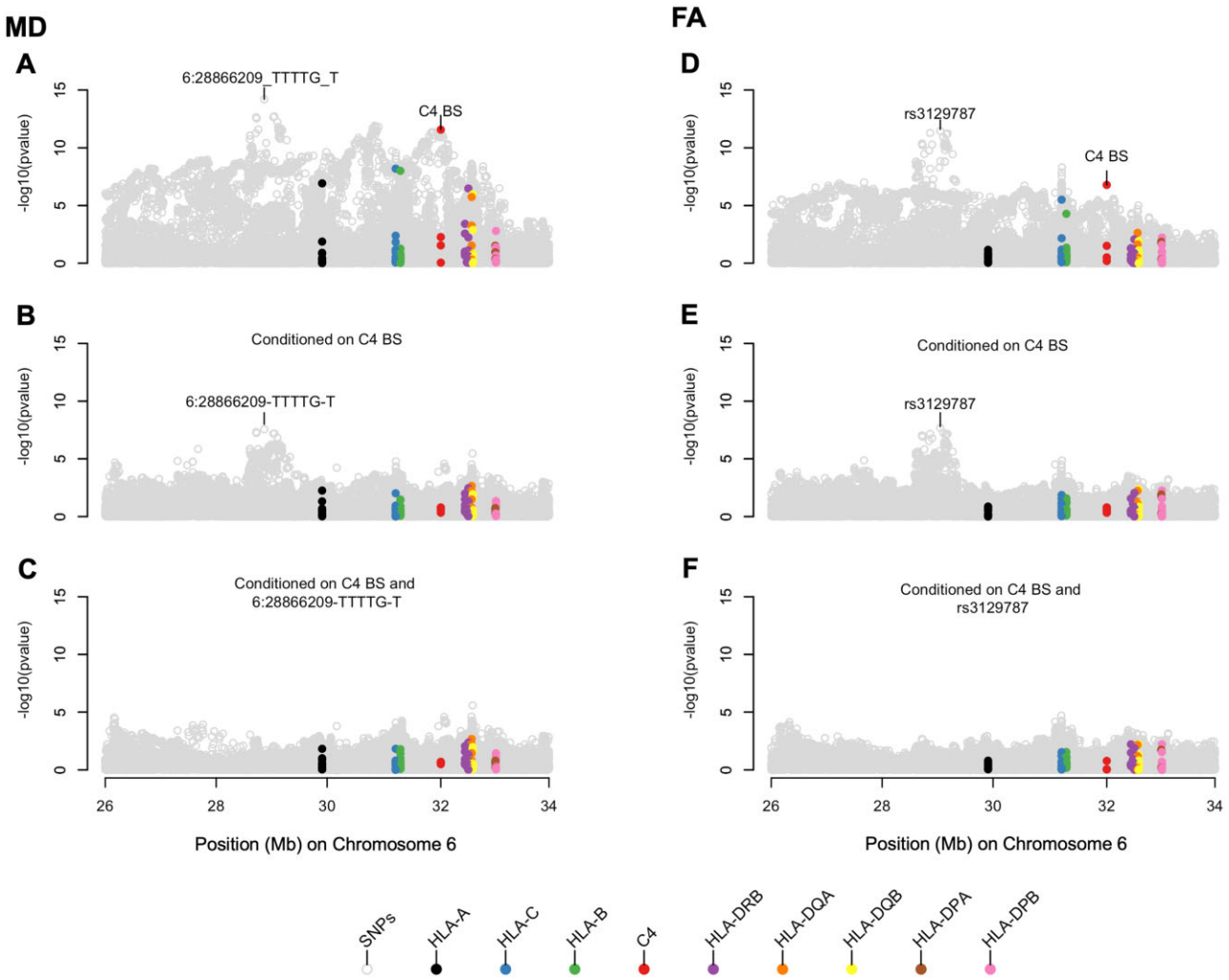

Figure 2 Association with age-related loss of white matter microstructural integrity within the MHC is explained by the C4-BS variant and an additional genetic variant. Association with genetic variants (SNPs, plotted in grey), HLA alleles (plotted in colours indicated by legend below plot) and C4 structural haplotype variants (plotted in red) with mean diffusivity (MD, **A**–**C**) and fractional anisotropy (FA, **D**–**F**) by genomic position in 31 128 individuals from UK Biobank using a 2 degrees of freedom test of SNP and SNP × age effects. (**A**) Primary associations of SNPs, HLA alleles and C4 variants with MD, (**D**) Primary associations of SNPs, HLA alleles and C4 variants with FA. (**B**) Conditional analysis for associations with MD including C4-BS in the model. (C**A**) Conditional analysis for MD including C4-BS and 6:288866209-TTTTG-T in the model. (**E**) Conditional analysis for FA including C4-BS in the model. (**F**) Conditional analysis for FA including C4-BS and rs3129787 in the model.

while other statistical analysis used the R statistical software (v.3.6.2).[33]

Statistical analysis within the human MHC

We assessed the association of SNPs, HLA alleles and C4 structural haplotype variants within the extended human MHC (chromosome 6: 26–34 Mb) with FA, and MD using the same model as above. To assess whether the association within the human MHC could be explained entirely by the association with C4-BS structural haplotype variants, we performed a step-wise conditional analysis including C4-BS in the regression model. If a significant signal remained in the region (at region-wide significance; $P < 3.9 \times 10^{-6}$) after this step, we performed further conditional analysis, including additional SNPs or HLA alleles in the regression model.

Interaction with age, sex and smoking

To investigate the relationship of C4-BS with age, smoking and age we performed interaction analyses. We first used a test for interaction of C4-BS with age within a linear regression model, dichotomizing individuals by median age (65 years), in addition to estimating the association of number of copies of C4-BS with FA and MD, stratified by age dichotomized as before. We investigated whether there was an interaction between a history of smoking and C4-BS for FA and MD. We dichotomized smoking into those with any history of smoking and those without and performed the same analyses as above. As previous analyses have shown differences between males and females in the influence of C4 variants on disease risk, we repeated the same analyses as before, dichotomized by sex.[18]

Association with white matter hyperintensities

To assess the extent that loci for age-related LOWMI were also associated with MRI white matter hyperintensities (WMH), we performed additional analyses in the same UK Biobank dataset. We used total volume of WMH (from T_1 and T_2 FLAIR images, field 25781), which was generated by an image-processing pipeline developed and run on behalf of UK Biobank (https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf).[26,27] WMH was log transformed to approximate a normal distribution, and we used

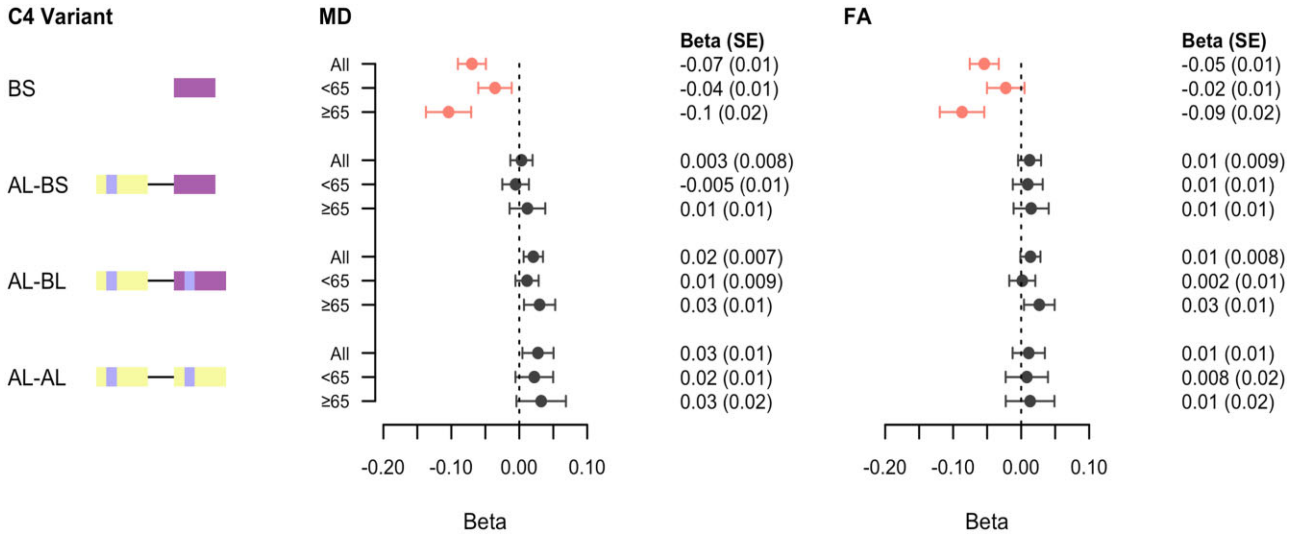

Figure 3 Association of C4 structural variants with age-related loss of white matter microstructural integrity is limited to C4-BS and has greater influence in older age. Association of four common structural variants of C4 with mean diffusivity (MD) and fractional anisotropy (FA) in 31 128 individuals from UK Biobank, with stratification by median age. Error bars show 95% confidence around the effect size estimate. Beta values are for 1 standard deviation change in MD or FA per copy of C4 variant.

total brain volume (field 25010) to account for head size. We performed a GWAS of log transformed WMH including SNPs, SNPs × Age, age, age-squared, sex, brain volume and 20 ancestry-informative principal components as above. The primary association statistic was based on a 2df test of the SNP and SNP × Age terms. We analysed the association of C4-BS with WMH as previously.

Sensitivity analyses

As associations with C4-BS have been demonstrated for schizophrenia,[20] in which DTI alterations have been reported,[34] we assessed whether our observed association could be mediated by neuropsychiatric disease. We extracted individuals corresponding to ICD-10 codes for eight categories of neuropsychiatric disorders: behavioural disorders due to substance abuse (F1), psychotic disorders (F2), mood disorders (F3), neurotic disorders (F4), eating disorders (F5), personality disorders (F6), psychological developmental disorders (F8) and behavioural and emotional disorders (F9); as well as schizophrenia (F20) in isolation. We then (i) tested whether C4-BS was associated with each disorder; (ii) evaluated the association of C4-BS with MD after additionally correcting for each disorder; and (iii) tested the association of each disorder with MD.

Data availability

The data that support the findings of this study are openly available via application to the UK Biobank (https://www.ukbiobank.ac.uk).

Results

The MHC gene region is a major risk locus for age-related loss of white matter microstructural integrity

We undertook a GWAS of age-related LOWMI based on a 2df test including SNP and SNP × Age effects in up to 31 128 individuals from the UK Biobank (age 45–81 years) (Fig. 1A–C and Supplementary Fig. 2) using the DTI markers MD and FA. The linkage disequilibrium score (LDscore) intercept for MD and FA was 1.08 (0.009) and 1.05 (0.007), respectively, while genomic control lambda values were 1.14 and 1.17, giving lambda1000 values of 1.004 and 1.005, respectively (Supplementary Figs 3 and 4).[35,36] Heritability of MD and FA based on LDscore regression was h2 (standard error, SE) = 0.20 (0.027) and 0.24 (0.023), respectively. After filtering on variants showing significant interactions with age ($P < 0.05$) in consistent direction with the main genetic effect, 18 loci were associated at genome-wide significance with either MD ($n = 16$) or FA ($n = 6$) (Table 1, Fig. 1B and Supplementary Figs 1 and 5–25). Four loci (encompassing *NBEAL1*, *PRDM5*, *MHC* and *SH3PXD2A* as the nearest genes) were associated with both MD and FA. Notably, two loci associated with MD harboured genes implicated in Wnt signalling (*WNT7A* and *WNT3*). In addition, several of the loci have previously been implicated in cerebral small vessel disease (*NBEAL1*, *VCAN*, *SH3PXD2A* and *COL4A2*).[9,37,38]

Among the top loci was a region on chromosome 6 encompassing the human MHC region (Fig. 1B and D). Given its potential to provide insights into immune mechanisms underlying age-related LOWMI, we focused on this region in subsequent analyses. To identify the molecular constituents encoded by the human MHC that determine age-related LOWMI we analysed associations with SNPs, HLA alleles and C4 structural haplotype variants in the MHC region (Fig. 1E). For this, we used a region-wide significance threshold of 3.9×10^{-6} to identify variation contributing to the association within the region.[32,39]

Genetic variants in the MHC reached genome-wide significance for both MD [best SNP: 6:28866209_TTTTG_T, beta (SE) = –0.069 (0.009); 2df $P = 6.5 \times 10^{-15}$] and FA [best SNP: rs3129787, beta (SE) = –0.056 (0.008); 2df $P = 3.5 \times 10^{-12}$; Fig. 2]. Of HLA alleles and C4 structural haplotype variants, the strongest association was with C4 structural haplotype variant BS (C4-BS), which was associated with MD at genome-wide significance [beta (SE) = –0.070 (0.010); 2df $P = 2.7 \times 10^{-12}$] and FA at region-wide significance [beta (SE) = –0.054 (0.011); 2df $P = 1.6 \times 10^{-7}$; Fig. 2]. C4-BS was imputed with good quality in the dataset (INFO = 0.75) and had a frequency of 12.3%. Class I HLA alleles were also associated at genome-wide significance with MD [best allele: HLA-C*07:01, beta (SE) = –0.051 (0.0095); 2df $P = 6.4 \times 10^{-9}$], and region-wide significance with FA [best allele: HLA-C*07:01, beta (SE) = –0.044 (0.0098), $P = 3.1 \times 10^{-6}$], but were less strongly associated than C4-BS (Fig. 2 and Supplementary Tables 1 and 2).

C4 structural variants are associated with age-related loss of white matter microstructural integrity

The most common 4 haplotype variants (BS, AL-BS, AL-BL, AL-AL) comprised 92.4% of the haplotypes in our data (Fig. 1E). Following Bonferroni correction, C4-BS was the only common structural variant that associated with DTI measures of age-related LOWMI at region-wide significance (Fig. 3). However, there were associations at nominal significance for MD with C4 AL-BL[beta (SE) = 0.021 (0.0072); 2df $P = 0.0054$] and AL-AL [beta (SE) = 0.027 (0.012); 2df $P = 0.028$] and for FA with C4 AL-BL [beta (SE) = 0.014(0.0075); $P = 0.0089$; Fig. 3].

Conditional analysis reveals an independent association with an additional genetic variant in the MHC region

Next, we performed a step-wise conditional analysis on the C4-BS variant for both MD and FA to determine whether any independent signals remained after including C4-BS in the model. For both MD and FA, no HLA allele reached region-wide significance after conditioning on C4-BS, including those which previously reached genome-wide significance (all P values >0.001, Fig. 2 and Supplementary Tables 1 and 2). However, several SNPs remained genome-wide significant after conditioning on C4-BS for MD [best SNP: 6:28866209_TTTTG_T, MAF = 0.21, beta (SE) = –0.053 (0.0098), 2df $P = 2.7 \times 10^{-8}$] and FA [best SNP: rs3129787, MAF = 0.29, beta (SE) = –0.048 (0.0089), $P = 2.0 \times 10^{-8}$]. After additionally conditioning on the best SNPs for FA and MD, no associations reached region-wide significance, indicating that the most parsimonious model explaining the signal within the MHC comprises C4-BS and an additional genetic variant (6:28866209_TTTTG_T for MD, rs3129787 for FA; Fig. 2 and Supplementary Tables 1 and 2).

The protective influence of C4-BS on age-related LOWMI is greater in older individuals and those without a history of smoking

To characterize the relationship of C4-BS with age, sex and smoking, we calculated the impact of copies of C4-BS on MD and FA stratified by median age, history of smoking and sex, and assessed the evidence of an interaction. As expected, there was a significant interaction of C4-BS with age for MD ($P = 0.0019$) and FA ($P = 0.015$) (Figs 3 and 4), with stronger protective effects of C4-BS in individuals aged $\geqslant 65$ compared to <65. We also found evidence for an interaction between C4-BS and smoking status for MD ($P = 0.00093$) and FA ($P = 0.021$) (Fig. 4); the protective effects of C4-BS on MD and FA were stronger for those without a history of smoking. Conversely, there was no evidence of an interaction with sex with MD ($P = 0.28$) and FA ($P = 0.065$).

C4-BS and other loci for age-related LOWMI are also associated with white matter hyperintensities

We assessed the extent to which loci for age-related LOWMI were also associated with WMH, a radiological marker of cerebral small vessel disease in 31 880 individuals from the UK Biobank. C4-BS was associated with WMH [beta (SE) = –0.036 (0.0098); 2df $P = 6.9 \times 10^{-7}$], and again showed evidence of stronger effects in older individuals (interaction $P = 1.1 \times 10^{-4}$).

Of 16 SNPs associated with MD, 14 were associated with WMH at $P < 0.05$, seven had $P < 1 \times 10^{-4}$ and three (ICA1L-WDR12-CARF-NBEAL1, SH3PXD2A, COL4A2) reached genome-wide significance. These three loci were also associated with lacunar stroke in a recent analysis.[37] Of six SNPs associated with FA, four were associated with WMH at $P < 0.05$, three had $P < 1 \times 10^{-4}$ and one (ICA1L-WDR12-CARF-NBEAL1) was genome-wide significant.

The association of C4-BS with age-related LOWMI is not confounded by neuropsychiatric disease

Given the recently reported association between structurally diverse alleles of C4 genes and schizophrenia and reports showing a reduced white matter integrity in schizophrenia,[20,34] we assessed the potential for the association of C4-BS with MD to be confounded by association with neuropsychiatric disorders. Multiple categories of neuropsychiatric disorders were associated with MD (Supplementary Table 4). However, C4-BS was not associated with any category of neuropsychiatric disorder, or schizophrenia in isolation, indicating that the association of C4-BS with age-related LOWMI is independent of neuropsychiatric disorders.

Discussion

The molecular pathways and mechanisms underlying age-related LOWMI are poorly understood, which limits potential to develop successful treatments. Here we show that individuals carrying the C4-BS structural haplotype variant have improved white matter microstructural integrity, as measured by DTI and reflected in both FA and MD. This effect showed a substantial interaction by age; the protective effects of C4-BS were stronger in older individuals, which is consistent with the observation that C4 protein levels increase considerably in older individuals.[18]

C4-BS carriers have reduced expression of C4A in both serum and brain.[20] Our results therefore imply that reduced C4A expression—or increased C4B to C4A ratio—protects from age-related LOWMI. While the precise mechanism by which this protective effect occurs remain unknown, some speculation is warranted. Recent studies have established an important role of complement in the brain in part through its impact on microglia. C4 has been shown to influence synaptic remodelling, particularly during development,[19,20] and recent studies show that C4A overexpression leads to increased microglial engulfment of synapses.[19] Studies in Alzheimer's disease models further suggest a role of complement in spine elimination in neurodegeneration.[40] Complement receptor 3 (CD11b-CD18) binds complement and is known to mediate spine elimination through binding by fibrinogen,[40] levels of which increase in the brain with the increasing blood–brain barrier permeability seen in ageing.[41] Although these effects have been documented solely in the grey matter, an attractive hypothesis is that C4 has additional impacts on microglia in the white matter,[42] which would be compatible with our findings.

Alternatively, the protective effect of C4-BS might relate to the exacerbating effect of complement activation on tissue injury as is seen in the context of CNS disorders,[43,44] and which has been related to microglial activation, inducing phagocytosis.[45] Studies in C3 knockout mice, and using a C3a-receptor antagonist, have shown that blocking complement can reduce tissue injury in the brain.[46] The expression levels of various components of the classical complement pathway strongly increase during ageing,[47,48] as has specifically been demonstrated for C3 and C4 in human CSF.[18] Hence, it seems possible that the increase in C4 expression with advancing age, in concert with blood–brain barrier permeability, leads to increased complement activation both systemically and through microglial activation as pathogens cross the blood–brain barrier. This in turn could contribute to exacerbation of white matter tissue damage. Functional studies will be required to determine the precise mechanism underlying the LOWMI observed.

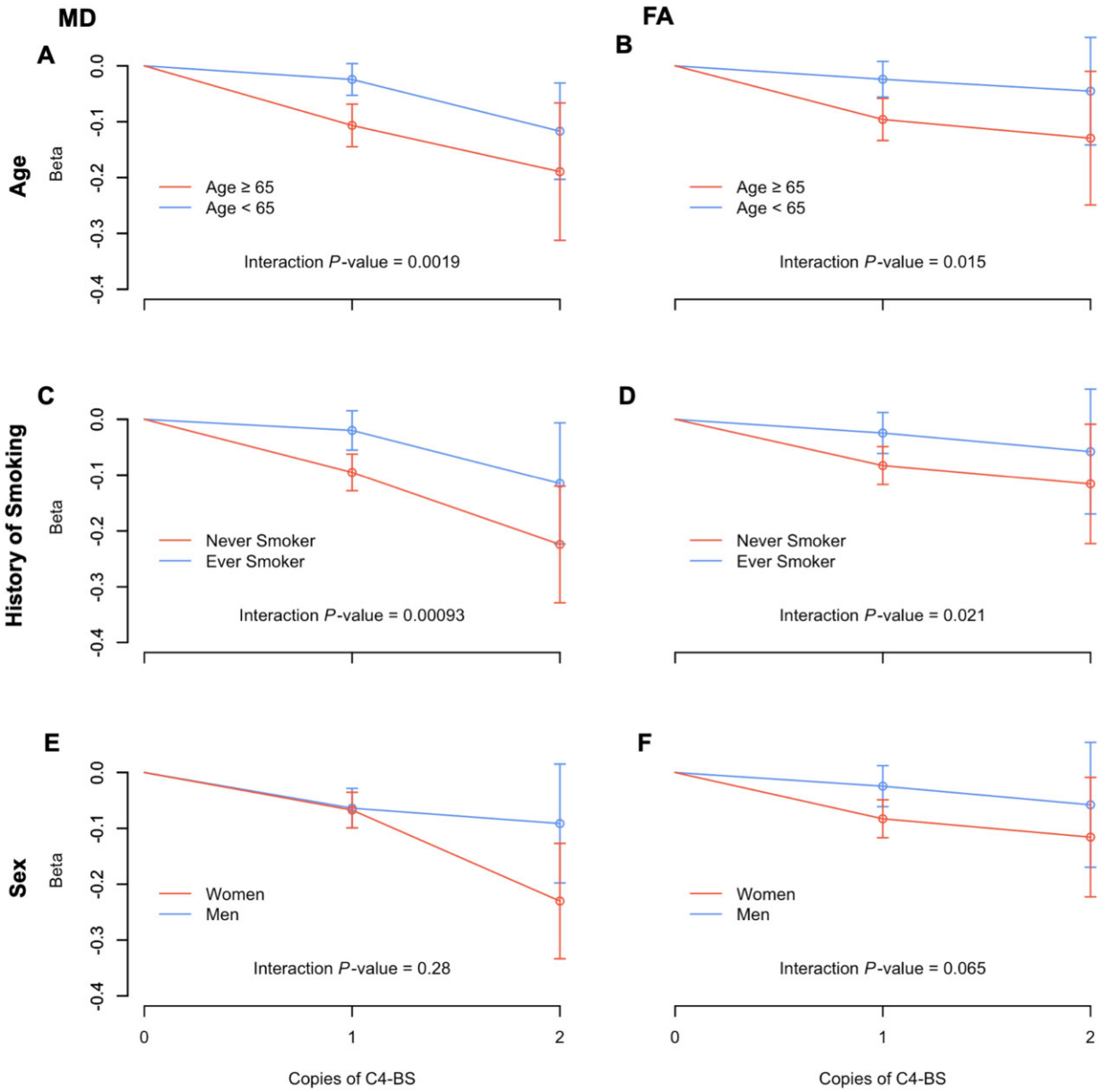

Figure 4 The protective influence of C4-BS on age-related loss of white matter microstructural integrity is greater in older individuals and those without a history of smoking. (**A** and **B**) Association of C4-BS in individuals of age <65 (blue) and age $\geqslant 65$ (red) for MD and FA. (**C** and **D**) Association of C4-BS in individuals with a history of smoking (blue) and those without (red) for MD and FA. (**E** and **F**) Association of C4-BS in males (blue) and females (red) for MD and FA. Error bars show 95% confidence around the effect size estimate. Beta values are for 1 SD change in MD or FA per copy of C4-BS variant.

Notably, we identified an interaction of C4-BS with history of smoking, which emphasizes a role of gene–environment interaction in age-related LOWMI, with the greatest protective effects of C4-BS observed in individuals without smoking history. One explanation for this interaction might be that the proinflammatory effects of smoking,[49] which include complement activation,[21] are amplified in individuals not carrying the C4-BS variant, although the precise mechanisms remain to be determined.

Our study expands knowledge of the phenotypic consequences of C4 structural haplotype variants. Previous studies have implicated a causal role of C4-BS in schizophrenia,[20] where C4-BS variant carriers are at lower risk of disease: a concordant effect with our study. As individuals with neuropsychiatric disease have been reported to have altered DTI measures of white matter microstructure, we investigated whether our findings could be confounded by neuropsychiatric disease. We could find no evidence that this was the case, and the effects of C4-BS were strongest in older individuals in our study, which would be inconsistent with an association with schizophrenia which generally onsets in early adulthood. C4-BS variants have also been associated with systemic lupus erythematosus and Sjögren's syndrome[18]; however, with an opposing effect in that they increase the risk of disease. Future studies should focus on understanding the full phenotypic spectrum of C4-BS variants.

Our results reveal a naturally occurring process by which age-related LOWMI is ameliorated, and therefore highlight a potential therapeutic target. The complement pathway is amenable to therapeutic modulation and has received interest in acute cerebrovascular injury such as stroke.[16,46,50] However, given the opposing effects of C4-BS on Sjögren's syndrome and systemic lupus erythematosus compared to schizophrenia and age-related LOWMI,

therapeutic development would necessitate balancing risk reduction with the adverse effects of increased autoimmune disease susceptibility.

Our GWAS of age-related LOWMI additionally identified other genes of interest. Variants close to *WNT7A* (Wnt family member 7A), known to influence expression of *WNT7A* in the brain,[51] were associated with MD. *WNT7A* has an essential role in the formation and maintenance of the blood–brain barrier,[52,53] and regulates white matter vascularization.[54] Variants in *PRDM5* (PR/SET domain 5) were also associated with both FA and MD. *PRDM5* regulates transcription of extracellular matrix components including proteins with an established role in small vessel disease such as collagen type IV,[55,56] and has been implicated in vascular Ehlers–Danlos syndrome.[57] Finally we found several other genes that have previously been associated with small vessel disease and lacunar stroke in GWAS—*COL4A2*, *SH3PXD2A*, *ICA1L-WDR12-CARF-NBEAL1*—to be associated with LOWMI,[9,37] emphasizing the role of small vessel disease.

This study also has limitations. The UK Biobank is composed of predominantly white British individuals and this analysis focused exclusively on those. Hence our results cannot necessarily be extrapolated to other ethnicities, which differ in MHC characteristics. To our knowledge, no other study with genetic data and measures of LOWMI of comparable size exists, meaning we were not able to attempt replication of our findings. Also, our analyses use imputation of C4 structural haplotypes from genetic information rather than by direct measurement, for example by droplet digital PCR, and our findings should be interpreted with this limitation. The complex genetic background of the HLA region means that categorizing variation on the region can be challenging. Future efforts might benefit from use of long read sequencing technologies.[58] Alterations in white matter microstructure reflect changes in the diffusion properties of water within the white matter and therefore might reflect multiple distinct pathologies. This lack of specificity should be considered when interpreting the results; however, the fact that many of the identified loci also associate with WMH and lacunar stroke suggests that cerebral small vessel disease contributes substantially to the age-related LOWMI observed.

In conclusion, we show that C4-BS—and therefore reduced levels of C4A—is associated with reduced age-related LOWMI Our findings implicate the complement system and highlight the role of gene–environment interactions in modifying risk.

Acknowledgements

This study was conducted using the UK Biobank resource, under UKB project 2532 (UK Biobank Stroke Study).

Funding

M.T. was supported by The Barts Charity and The National Institute of Health Research Barts Biomedical Research Centre. This project has received funding from the European Union's Horizon 2020 research and innovation programme (666881), SVDs@target (to M.D.; 667375), and the D.F.G as part of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy – ID 390857198) and DI 722/16-1.

Competing interests

The authors report no competing interests. Since June 2021, M.T. has been a full time employee of Novo Nordisk.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Power MC, Su D, Wu A, et al. Association of white matter microstructural integrity with cognition and dementia. *Neurobiol Aging*. 2019;83:63–72.
2. O'Brien JT, Thomas A. Vascular dementia. *Lancet*. 2015;15(1):101–706.
3. Verlinden VJ, de Groot M, Cremers LG, et al. Tract-specific white matter microstructure and gait in humans. *Neurobiol Aging*. 2016;43:164–173.
4. Mito R, Raffelt D, Dhollander T, et al. Fibre-specific white matter reductions in Alzheimer's disease and mild cognitive impairment. *Brain*. 2018;141(3):888–902.
5. Wardlaw JM, Smith C, Dichgans M. Small vessel disease: Mechanisms and clinical implications. *Lancet Neurol*. 2019;18(7):684–696.
6. Stahon KE, Bastian C, Griffith S, Kidd GJ, Brunet S, Baltan S. Age-related changes in axonal and mitochondrial ultrastructure and function in white matter. *J Neurosci*. 2016;36(39):9990–10001.
7. Wharton SB, Simpson JE, Brayne C, Ince PG. Age-associated white matter lesions: The MRC cognitive function and ageing study. *Brain Pathol*. 2015;25(1):35–43.
8. de Groot M, Ikram MA, Akoudad S, et al. Tract-specific white matter degeneration in aging: The Rotterdam Study. *Alzheimer's Dement* 2015;11(3):321–330.
9. Persyn E, Hanscombe KB, Howson JMM, Lewis CM, Traylor M, Markus HS. Genome-wide association study of MRI markers of cerebral small vessel disease in 42,310 participants. *Nat Commun*. 2020;11(1):2175.
10. Zhao B, Zhang J, Ibrahim JG, et al. Large-scale GWAS reveals genetic architecture of brain white matter microstructure and genetic overlap with cognitive and mental health traits (n = 17,706). *Mol Psychiatry*. 2021;26(8):3943–3955.
11. Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. *Nat Rev Immunol*. 2018;18(5):325–339.
12. Complete Sequence and Gene Map of a Human Major Histocompatibility Complex. The MHC sequencing consortium. *Nature*. 1999;401(6756):921–923.
13. Alexander JJ. Blood-brain barrier (BBB) and the complement landscape. *Mol Immunol*. 2018;102:26–31.
14. Iadecola C. The pathobiology of vascular dementia. *Neuron*. 2013;80(4):844–866.
15. Crehan H, Hardy J, Pocock J. Blockage of CR1 prevents activation of rodent microglia. *Neurobiol Dis*. 2013;54:139–149.
16. Wagner E, Frank MM. Therapeutic potential of complement modulation. *Nat Rev Drug Discov*. 2010;9(1):43–56.
17. Isenman DE, Young JR. The molecular basis for the difference in immune hemolysis activity of the Chido and Rodgers isotypes of human complement component C4. *J Immunol*. 1984;132(6):3019–3027.
18. Kamitaki N, Sekar A, Handsaker RE, et al.; Schizophrenia Working Group of the Psychiatric Genomics Consortium. Complement genes contribute sex-biased vulnerability in diverse disorders. *Nature*. 2020;582(7813):577–581.
19. 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*. 2020;24(2):214–224.
20. Sekar A, Bialas AR, de Rivera H, et al.; Schizophrenia Working Group of the Psychiatric Genomics Consortium. Schizophrenia

risk from complex variation of complement component 4. *Nature*. 2016;530(7589):177–183.
21. Robbins RA, Nelson KJ, Gossman GL, Koyama S, Rennard SI. Complement activation by cigarette smoke. *Am J Physiol*. 1991; 260(4 Pt 1):L254–9.
22. Sajja RK, Rahman S, Cucullo L. Drugs of abuse and blood-brain barrier endothelial dysfunction: A focus on the role of oxidative stress. *J Cereb Blood Flow Metab*. 2016;36(3):539–554.
23. Gons RA, van Norden AG, de Laat KF, et al. Cigarette smoking is associated with reduced microstructural integrity of cerebral white matter. *Brain*. 2011;134(Pt 7):2116–2124.
24. Cox SR, Lyall DM, Ritchie SJ, et al. Associations between vascular risk factors and brain MRI indices in UK Biobank. *Eur Heart J*. 2019;40(28):2290–2300.
25. Sudlow C, Gallacher J, Allen N, et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12(3):e1001779.
26. Miller KL, Alfaro-Almagro F, Bangerter NK, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat Neurosci*. 2016;19(11):1523–1536.
27. 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.
28. Pasi M, van Uden IW, Tuladhar AM, de Leeuw FE, Pantoni L. White matter microstructural damage on diffusion tensor imaging in cerebral small vessel disease: Clinical consequences. *Stroke*. 2016;47(6):1679–1684.
29. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018; 562(7726):203–209.
30. Dilthey A, Leslie S, Moutsianas L, et al. Multi-population classical HLA type imputation. *PLoS Comput Biol*. 2013;9(2): e1002877.
31. Browning BL, Browning SR. Genotype imputation with millions of reference samples. *Am J Hum Genet*. 2016;98(1):116–126.
32. Glanville KP, Coleman JRI, Hanscombe KB, et al.; Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Classical human leukocyte antigen alleles and c4 haplotypes are not significantly associated with depression. *Biol Psychiatry*. 2020;87(5):419–430.
33. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*. 2015;4:7.
34. Kelly S, Jahanshad N, Zalesky A, et al. Widespread white matter microstructural differences in schizophrenia across 4322 individuals: Results from the ENIGMA Schizophrenia DTI Working Group. *Mol Psychiatry*. 2018;23(5):1261–1269.
35. Bulik-Sullivan BK, Loh PR, Finucane HK, et al.; Schizophrenia Working Group of the Psychiatric Genomics Consortium. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291–295.
36. de Bakker PI, Ferreira MA, Jia X, Neale BM, Raychaudhuri S, Voight BF. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum Mol Genet*. 2008; 17(R2):R122–8.
37. Traylor M, Persyn E, Tomppo L, et al. Genetic basis of lacunar stroke: A pooled analysis of individual patient data and genome-wide association studies. *Lancet Neurol*. 2021;20(5):351–361.
38. Sargurupremraj M, Suzuki H, Jian X, et al.; International Headache Genomics Consortium (IHGC). Cerebral small vessel disease genomics and its implications across the lifespan. *Nat Commun*. 2020;11(1):6285.
39. Lindsey JK, Jones B. Choosing among generalized linear models applied to medical data. *Stat Med*. 1998;17(1):59–68.
40. Merlini M, Rafalski VA, Rios Coronado PE, et al. Fibrinogen induces microglia-mediated spine elimination and cognitive impairment in an Alzheimer's disease model. *Neuron*. 2019;101(6): 1099–1108.e6.
41. Erdő F, Denes L, de Lange E. Age-associated physiological and pathological changes at the blood-brain barrier: A review. *J Cereb Blood Flow Metab*. 2017;37(1):4–24.
42. Safaiyan S, Besson-Girard S, Kaya T, et al. White matter aging drives microglial diversity. *Neuron*. 2021;109(7):1100–1117.e10.
43. Brennan FH, Anderson AJ, Taylor SM, Woodruff TM, Ruitenberg MJ. Complement activation in the injured central nervous system: Another dual-edged sword? *J Neuroinflammation*. 2012;9:137.
44. Orsini F, De Blasio D, Zangari R, Zanier ER, De Simoni MG. Versatility of the complement system in neuroinflammation, neurodegeneration and brain homeostasis. *Front Cell Neurosci*. 2014;8:380.
45. Zhang LY, Pan J, Mamtilahun M, et al. Microglia exacerbate white matter injury via complement C3/C3aR pathway after hypoperfusion. *Theranostics*. 2020;10(1):74–90.
46. Mocco J, Mack WJ, Ducruet AF, et al. Complement component C3 mediates inflammatory injury following focal cerebral ischemia. *Circ Res*. 2006;99(2):209–217.
47. Reichwald J, Danner S, Wiederhold KH, Staufenbiel M. Expression of complement system components during aging and amyloid deposition in APP transgenic mice. *J Neuroinflamm*. 2009;6:35.
48. Stephan AH, Madison DV, Mateos JM, et al. A dramatic increase of C1q protein in the CNS during normal aging. *J Neurosci*. 2013; 33(33):13460–13474.
49. Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease: An update. *J Am Coll Cardiol*. 2004;43(10):1731–1737.
50. Clarke AR, Christophe BR, Khahera A, Sim JL, Connolly ES Jr. Therapeutic modulation of the complement cascade in stroke. *Front Immunol*. 2019;10:1723.
51. Battle A, Brown CD, Engelhardt BE, Montgomery SB. Genetic effects on gene expression across human tissues. *Nature*. 2017; 550(7675):204–213.
52. Cho C, Smallwood PM, Nathans J. Reck and Gpr124 are essential receptor cofactors for Wnt7a/Wnt7b-specific signaling in Mammalian CNS angiogenesis and blood-brain barrier regulation. *Neuron*. 2017;95(5):1056–1073.e5.
53. Wang Y, Cho C, Williams J, et al. Interplay of the Norrin and Wnt7a/Wnt7b signaling systems in blood-brain barrier and blood-retina barrier development and maintenance. *Proc Natl Acad Sci U S A*. 2018;115(50):e11827–e11836.
54. Chavali M, Ulloa-Navas MJ, Pérez-Borredá P, et al. Wnt-dependent oligodendroglial-endothelial interactions regulate white matter vascularization and attenuate injury. *Neuron*. 2020; 108(6):1130–1145.e5.
55. Burkitt Wright EMM, Spencer HL, Daly SB, et al. Mutations in PRDM5 in brittle cornea syndrome identify a pathway regulating extracellular matrix development and maintenance. *Am J Hum Genet*. 2011;88(6):767–777.
56. Galli GG, Honnens de Lichtenberg K, Carrara M, et al. Prdm5 regulates collagen gene transcription by association with RNA polymerase II in developing bone. *PLoS Genet*. 2012;8(5):e1002711.
57. Malfait F. Vascular aspects of the Ehlers-Danlos Syndromes. *Matrix Biol*. 2018;71-72:380–395.
58. Logsdon GA, Vollger MR, Eichler EE. Long-read human genome sequencing and its applications. *Nat Rev Genet*. 2020;21(10):597–614.