Beyond C4: Analysis Of The Complement Gene Pathway Shows Enrichment For IQ In Patients With Psychotic Disorders And Healthy Controls

Authors: Jessica F. Holland[1], Donna Cosgrove[1], Laura Whitton[1], Denise Harold[2,3], Aiden Corvin[2], Michael Gill[2], David O. Mothersill[1], Derek W. Morris[1], Gary Donohoe[1].

1. Cognitive Genetics & Cognitive Therapy Group, The Centre for Neuroimaging, Cognition and Genomics (NICOG), School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Ireland.
2. Neuropsychiatric Genetics Research Group, Department of Psychiatry, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland
3. School of Biotechnology, Dublin City University, Dublin, Ireland

Corresponding Author:
Professor Gary Donohoe
NUI Galway Professor of Psychology
Director, Center for Neuroimaging and Cognitive Genomics.
Rm 1040, The School of Psychology,
National University of Ireland Galway.
Galway, Ireland.

Email: gary.donohoe@nuigalway.ie
Tel: +353 872891485

Word Count
Abstract: 291
Full text: 4463
Tables: 3
Number of Figures: 0
Supplementary Material: 1

Abstract

Introduction: Variation in cognitive performance, which strongly predicts functional outcome in schizophrenia (SZ), has been associated with multiple immune-relevant genetic loci. These loci include complement component 4 (C4A), structural variation at which was recently associated with SZ risk and synaptic pruning during neurodevelopment and cognitive function. Here, we test whether this genetic association with cognition and SZ risk is specific to C4A, or extends more broadly to genes related to the complement system.

Methods: Using a gene-set with an identified role in 'complement' function (excluding C4A), we used MAGMA to test if this gene-set was enriched for genes associated with human intelligence and SZ risk, using genome-wide association summary statistics (IQ; N=269,867, SZ; N=105,318). We followed up this gene-set analysis with a complement gene-set polygenic score (PGS) regression analysis in an independent dataset of patients with psychotic disorders and healthy participants with cognitive and genomic data (N=1000).

Results and conclusions: Enrichment analysis suggested that genes within the complement pathway were significantly enriched for genes associated with IQ, but not SZ. In a gene-based analysis of 90 genes, SERPING1 was the most enriched gene for the phenotype of IQ. In a PGS regression analysis, we found that a complement pathway PGS associated with IQ GWAS statistics also predicted variation in IQ in our independent sample. This association (observed across both patients and controls) remained significant after controlling for the relationship between C4A and cognition. These results suggest a robust association between the complement system and cognitive function, extending beyond structural variation at C4A.

Introduction

SZ is a complex heritable neurodevelopmental disorder, in which level of disability is strongly predicted by deficits in cognitive function (Sekar *et al.*, 2016). Genome-wide association studies (GWAS) have consistently found evidence of association between SZ and variation at or within the major histocompatibility complex (MHC) locus on chromosome 6 (Mcallister, 2014). While both the size and high degree of linkage disequilibrium within the region has made it difficult to disentangle the contribution of this region to SZ risk, Sekar et al.(Sekar *et al.*, 2016) demonstrated that part of the association signal at this locus could be explained by structural variation mapping to C4A. C4A is part of the complement system, an innate immune system pathway comprising a large number of plasma proteins that help antibodies and phagocytes to clear pathogens (Sekar *et al.*, 2016). This study found that structural variation at C4 was associated with (a) variation in RNA expression levels in post-mortem brain samples, (b) increased risk for SZ, and (c) that increased RNA expression predicted by C4 structural variation influenced development of the neural system via a pattern of altered synaptic pruning. Following up these findings, we recently showed that higher predicted C4 RNA expression was associated with poorer performance on measures of memory function in both patients with SZ and healthy participants, and with reduced cortical activation during visual task performance in healthy participants (Donohoe *et al.*, 2018). This, along with evidence that other risk-related genes and proteins within the complement system are also associated with variation in both cognition and brain structure in both patients and healthy individuals (Allswede *et al.*, 2018, Athanasiu *et al.*, 2017, Bralten *et al.*, 2011, Chung *et al.*, 2014, Donohoe *et al.*, 2013, Nettiksimmons *et al.*, 2016, Rose *et al.*, 2013, Song *et al.*, 2014, Zhang *et al.*, 2017), led us to consider whether variation within genes that encode the complement system could, as a whole, influence cognition.

Examples of other complement function related genes that have been associated with both SZ risk and cognition include Cub and Sushi Multiple Domains-1 (CSMD1), complement Factor H (CFH), and complement C3b/C4b receptor 1 (CR1) (Donohoe *et al.*, 2018, Dunkelberger & Song, 2010, Föcking *et al.*, 2019, Gigante *et al.*, 2011). In patients with SZ and health participants, we previously reported that a GWAS-identified SZ risk variant within CSMD1, which encodes for a regulator of complement, was associated with poorer general cognitive ability and episodic memory function (Donohoe *et al.*, 2013). Other variants within CSMD1 gene have also been associated with poorer cognitive performance in large samples of healthy participants (Athanasiu *et al.*, 2017). In a sample of 1,783 patients with SZ and healthy controls, expression of complement factor H (CFH) in the hippocampus was found to be associated with both increased SZ risk and poorer memory function (Zhang *et al.*, 2017). The CR1 gene has also been associated with cognitive impairment in a study of AD (Chung *et al.*, 2014), and entorhinal cortex volume in young healthy adults (Bralten *et al.*, 2011). In addition to these examples, complement gene expression more broadly has been linked to superior frontal cortex thickness in healthy humans (Allswede *et al.*, 2018), which has been linked to general intelligence in previous studies (Jung & Haier, 2007). Furthermore, in a proteomic analysis of the complement signalling pathway using longitudinal population-based data, individuals who went on to experience psychotic like experience were found to have upregulation of multiple complement proteins. (English *et al.*, 2017, Föcking *et al.*, 2019).

Given these lines of evidence, the present study aimed to systematically characterise, for the first time, the association between genetic variation within genes related to complement function and both SZ risk and cognition. To do this, we performed gene-set analysis (GSA) of complement pathway genes, based on prior work in curating complement gene-sets based on publically available data by (Birnbaum *et al.*, 2018, Qian *et al.*, 2019).

The resulting complement gene-set was tested for enrichment using IQ GWAS summary data from Savage et al. (Savage *et al.*, 2018), and SZ GWAS summary data from the Psychiatric Genomics Consortium (PGC) (Ripke *et al.*, 2014). We tested the hypotheses that 1) the complement pathway gene-set (excluding C4) would show significant enrichment for genes associated with both IQ and with SZ. Depending on whether these hypotheses were supported, we sought to further test whether 2) a polygenic score for complement pathway genes based on IQ or SZ GWAS summary statistics would explain variation in cognitive performance in a sample of patients with psychotic disorders and healthy controls.

Methodology

Samples included in enrichment analysis

IQ GWAS dataset (Savage et al 2018; 18): This dataset included GWAS summary statistics on 269,867 individuals from 14 cohorts: UK Biobank (UKB), the Cognitive Genomics Consortium (COGENT), the Rotterdam Study (RS), the Generation R Study (GENR), the Swedish Twin Registry (STR), Spit for Science (S4S), the High-IQ/Health and Retirement Study (HiQ/HRS), the Twins Early Development Study (TEDS), the Danish Twin Registry (DTR), IMAGEN: reinforcement-related behaviour in normal brain function and psychopathology (Schumann *et al.*, 2010), the Brisbane Longitudinal Twin Study (BLTS), the Netherlands Study of Cognition, Environment, and Genes (NESCOG), Genes for Good (GfG), and the Swedish Twin Studies of Aging (STSA) (Savage *et al.*, 2018). Participants ranged from children to older adults, with older samples being screened for cognitive decline to exclude the possibility of dementia affecting performance on neurocognitive tests. This analysis identified 205 genomic loci (harbouring 1,016 genes) related to human intelligence. In each cohort, varying measures of intelligence were subject to a principal components

analysis to create a common latent g factor underlying multiple dimensions of cognition (Savage *et al.*, 2018). For each single nucleotide polymorphism (SNP), computed beta value represented the likelihood of increased 'g' factor scores based on the SNP of interest. The combined summary statistics of these GWAS in relation to the IQ phenotype were used for enrichment analysis.

SZ GWAS dataset (Pardiñas *et al.*, 2018): This recent GWAS for SZ combined existing genome wide association data (Ripke *et al.*, 2014) with genome-wide genotype information for SZ cases from the UK (the CLOZUK sample), amalgamated with control datasets obtained from public repositories to form a meta-analysis of 40,675 SZ cases and 64,643 healthy controls (Pardiñas *et al.*, 2018), excluding overlapping samples. This meta-analysis found 145 loci related to SZ diagnosis, with 50 of these novel discoveries. An independent sample total of 5,762 cases and 154,224 controls was used in analysis to replicate meta-analytic findings. Details of samples can be found in supplementary information on the original paper (Pardiñas *et al.*, 2018, Ripke *et al.*, 2014). Participants in the SZ sample were diagnosed based on clinicians report or using a specific inclusion criteria and research-based assessment. Healthy controls were screened for other psychiatric illness and were within normal IQ range. Summary statistics of this GWAS provided an odds ratio score for likelihood of developing SZ based on risk alleles of each SNP.

Complement gene-set

To test for enrichment of genes related to complement function in IQ and SZ, we based our gene-set list on recent publications targeting complement-related immunity genes (Birnbaum *et al.*, 2018, Qian *et al.*, 2019). The Birnbaum paper (Birnbaum *et al.*, 2018) provided a gene-set related to complement in schizophrenia by collating information from neuroimmunology

and general immunology literature, as well as pathway annotation programs, cross-referenced with multiple gene expression databases (eg. GO, KEGG, IMPORT, IPA, and IMMUNOME) to assemble 34 complement-related genes (Dunkelberger & Song, 2010, Orsini *et al.*, 2014, Ricklin *et al.*, 2010, Sarma & Ward, 2011, Veerhuis *et al.*, 2011). The second publication (Qian *et al.*, 2019) included 32 of the original 34 genes with additional gene lists included from Molecular Signatures Database, the Human Biological Pathway Unification Database and the HUGO Gene Nomenclature Committee (https://pathcards.genecards.org, (http://software.broadinstitute.org/gsea/msigdb/index.jsp, https://www.genenames.org) searching the keyword "complement". After removing the duplicated genes and genes directly encoding for C4 (C4A, C4B, C4BPA, C4BPB, C4_B), 108 genes remained as the candidate genes for further analysis (Supplementary Table 1). As 90 of these genes were available to be tested for enrichment using MAGMA, these 90 genes were brought forward for analysis (listed in table 1).

Enrichment analysis statistical approach

The statistical toolbox MAGMA (https://ctg.cncr.nl/software/magma; (De Leeuw *et al.*, 2015) was used to test for enrichment of the complement gene-set for genes associated with the IQ and SZ phenotypes. Enrichment analysis based on MAGMA consisted of three steps: first, an annotation step to map SNPs onto genes; second, a gene analysis step to compute gene p-values; and three, a gene-level analysis step, carried out both on a gene-wide and a gene-set-wide level. By doing so, MAGMA tests whether multiple genetic markers in a given gene or gene-set (e.g. complement genes) are more strongly associated with SZ or IQ than other genes in the genome.

Samples included in polygenic score (PGS) analysis

In total, 808 cases and 192 healthy participants completed a full neuropsychological assessment battery and had full genome-wide data available. Cases consisted of n=585 clinically stable patients with a diagnosis of SZ and schizoaffective disorder (SZA), and an additional n=223 patients diagnosed with bipolar disorder with psychotic features, major depressive disorder with psychotic features, delusional disorder, or psychosis not otherwise specified, as described elsewhere (Cosgrove *et al.*, 2017, Donohoe *et al.*, 2018, Whitton *et al.*, 2016). Participants were recruited from five sites across Ireland, and ethics approval was obtained from local ethics committees. Written informed consent was obtained from all participants. Inclusion criteria required participants to be clinically stable at the time of cognitive assessment, aged between 18–65 years, had no history of comorbid psychiatric disorder, history of seizures, substance abuse in the preceding 6 months, or prior head injury with loss of consciousness. All participants had Irish ancestry (all four grandparents born in Ireland). Patients were diagnosed by trained psychiatrists using the Structured Clinical Interview for DSM-IV Axis I Diagnosis (First *et al.*, 2002). Additional diagnostic and clinical information ascertained at time of interview, including symptom severity (Andreasen, 1984) and medication dosage were also recorded for analysis.

The healthy control sample was recruited on the basis of response to local media advertisements. Control participants were included if they were aged between 18 and 65 and met the criteria of having no history of major mental health problems, intellectual disability or acquired brain injury based on clinical interview, and no history of substance misuse in the preceding 6 months based on self-report. Control participants were also excluded from the study if they reported having a first-degree relative with a history of psychosis. Both patient and control clinical assessments were conducted in accordance with the relevant ethics

committees' approval from each participating site. All participants provided written informed consent.

Cognitive assessment: General cognitive functioning (IQ) was measured in the sample using selected subtests (Vocabulary, Similarities, Block Design and Matrix Reasoning) from the Wechsler Adult Intelligence Scale, third edition (WAIS-III;(Wechsler, 1997)), deriving a full scale, verbal and performance IQ for each participant. Premorbid IQ was measured using the Wechsler test of adult reading (WTAR; (Holdnack, 2001). Episodic memory was assessed in the Irish samples using the logical memory subtests (immediate and delayed conditions) from the Wechsler Memory Scale, third edition (WMS-III) (Wechsler, 1997), and the paired associations learning task (PAL; stages completed and total errors) from the Cambridge Automated Neuropsychological Test Battery (CANTAB; (Robbins *et al.*, 1994)). As in our previous study of C4, an unrotated principle components analysis based on the four available episodic memory tests was carried out to reduce multiple testing burden. This memory factor explained 72% of variance in memory scores, as described previously (Donohoe *et al.*, 2018). Spatial working memory was assessed in the samples using the WMS-III letter number sequencing task and the spatial working memory task from the Cambridge Automated Neuropsychological Test Battery (CANTAB SWM;(Robbins *et al.*, 1994)).

Genotyping Genotyping was conducted on DNA extracted from whole blood or saliva. Full GWAS data were available for all samples. A proportion of samples (n=575) were genotyped with an Affymetrix 6.0 chip (Santa Clara, CA, USA; as part of the WTCCC2 (Consortium & 2, 2012) and the remainder on the Illumina HumanCoreExome chip (San Diego, CA, USA), granting comparisons between platforms indicate high agreement (Barnes *et al.*, 2005). SNPs were excluded on the basis of MAF (minor allele frequency)<0.1% (i.e rate of second most

common allele occurs at a frequency of 10% or greater in the population), SNP missingness ⩽ 2%, and Hardy–Weinberg equilibrium ⩽ P^{10-6}. Imputation was carried out on these data sets separately using 1000 Genomes Phase I integrated haplotypes (Dec 2013 release) and IMPUTE2 to give ~10 million SNPs genome-wide per sample.

PGS score calculation

Based on the complement set of 90 genes described above, we began by identifying all SNPS within (+/−20 kb) of these gene loci, and then extracting the genotype values for our samples, which had been genotyped using either Affymetrix 6.0 or Illumina HumanCoreExome. PRSice software (https://choishingwan.github.io/PRSice/(Euesden *et al.*, 2014)) was used to perform quality control on data from each sample and SNPs were excluded from further analysis if there was >10% missing genotype data, HardyWeinberg equilibrium <1 × 10−5, or MAF <1%. Next, the lists of SNPs that passed QC for each sample were joined, providing a SNP list common to each sample (n = 14237 SNPs). Based on this list, an effect-size weighted PGS was generated for each individual using PRSice based on a threshold of p<0.05 for SNPs in the IQ and SZ GWAS sets. While multiple p value thresholds can be used, we have previously found a threshold of p = 0.05 to be the most informative in terms of signal-to-noise ratio (Cosgrove *et al.*, 2017).

Statistical Analysis

To estimate the association between computed complement PGS and performance on cognitive tasks, multiple regression analyses were carried out on the whole sample and on the patient sample only using IBM SPSS Statistics Version 25.0 (IBM Corp, Armonk, NY, 2017). Three cognitive domains of IQ, episodic memory and working memory were tested. PGS scores for complement were used as the independent variable, and age and gender were

entered as covariates of no interest. For IQ analysis, only gender was entered as a covariate, as scaled scores were used. To maximize power to detect differences, we carried out our analysis on the full dataset of cases and controls (n=1000). To test the possibility of a particular phenotype population driving results, analysis was also carried out in healthy controls and patients separately.

Results

Enrichment analysis

Using MAGMA to test for the enrichment of the complement gene-set in IQ and SZ GWAS data, the complement gene-set showed enrichment for IQ (beta=0.29, p=0.010; $\alpha = 0.05 / 2 = 0.025$ after bonferroni correction for the two phenotypes tested). By comparison, no enrichment was observed using the SZ data (beta=-0.07, p=0.26). Following up the complement gene-set analysis with a gene-based association analysis, 32 of the 90 genes included showed nominal association for the phenotype of IQ, with 12 of these genes surviving correction for multiple testing, using bonferroni correction for 90 genes ($\alpha = 0.05 / 90 = 0.00055$) of which SERPING1 was the highest associated (see Table 1).

Insert Table 1 here

PGS analysis: Clinical and demographic information

Demographic and clinical information for all participants included in the PGS analysis is outlined in table 2. No association between complement PGS and either age, gender, or years of education were observed, see supplementary table 2. Similarly, based on a principle components analysis described previously by the group (O'gráda *et al.*, 2009), no association was found between complement PGS and positive or negative symptom factor scores. No association between medication dosage and complement PGS was observed.

Insert table 2 here

Association between complement-based polygenic score (PGS) and IQ

As the complement pathway was enriched for IQ-associated genes based on GWAS summary statistics (Savage *et al.*, 2018), polygenic scores were created for the complement gene-set to test for an association with cognition in an independent sample of cases and healthy participants. Based on a regression analysis on the total sample (cases and healthy participants combined), complement PGS significantly predicted variation in measures of Premorbid IQ and verbal IQ (see table 3). The observed direction of effect was that increased complement-based PGS was associated with an increase in IQ in the sample. These findings remained significant when gender was covaried for (premorbid IQ: F change=5.938; Std Beta=0.087, df = 2,772; p = 0.015; verbal IQ: F change=4.037, Std Beta=0.067, df=2,903; p=0.045). If corrected for 3 domains of cognition analysed, however, premorbid IQ only survives multiple testing correction ($\alpha = 0.05 / 3 = 0.016$). When the sample was sub-divided into patients and control groups (considered separately) as a post-hoc analysis, the association between complement PGS and IQ was no longer significant in cases. However, complement PGS was found to influence Full Scale IQ scores in healthy controls only (F change=4.195; Std Beta= 0.147, df = 2,182; p =0.042). Notably, a significant difference in complement PGS between patients and healthy participants was observed, such that patients had lower PGS than healthy participants (F=1.82, df 1,953; p=0.018). No significant associations were observed between complement PGS and memory scores (see table 3).

To determine whether the association between complement PGS and IQ remained significant once variations in the reported C4A structural haplotype were accounted for, we re-ran the analysis by including predicted C4A RNA expression scores as a covariate (score previously used by group (Donohoe *et al.*, 2018), seen to impact on cognition). After the

effects of C4A had been accounted for in this way, complement PGS continued to significantly predict variation in Premorbid IQ (F change=6.057; Std Beta= 0.088, df = 3,771; p = 0.014), and Verbal IQ (F change=4.016, Std Beta= 0.067, df=3,902; p=0.045). For healthy controls only, complement PGS continued to predict variation in full scale IQ scores (F change=4.153; Std Beta= 0.147, df = 3,181; p = 0.043).

***Insert table 3 here**

Discussion

The purpose of this study was to characterise the association between genetic variation within the complement system and performance on measures of cognitive function related to disability in SZ. Following earlier studies from our group and others, demonstrating that multiple individual loci within complement, including C4A, are associated with cognition, we sought to determine whether variation in the complement gene-set as a whole was associated with cognitive performance. To do this, we carried out a series of enrichment analyses, and PGS analysis in independent datasets. We found that (1) the complement gene-set was enriched for association with cognitive function (as measured by IQ) but not SZ risk, with multiple individual genes being associated with IQ, of which SERPING1 was the strongest associated. We further found, based on independent samples of patients and controls, that (2) a complement-based polygenic score for IQ predicted variation on multiple IQ measures, even when the effects of C4A was accounted for. Importantly, (3) these cognitive effects did not appear to differ between patient and healthy participant samples, but in post-hoc analysis of healthy controls only the complement PGS still affected full scale IQ.

Collectively, these findings suggest that whereas individual genetic components of the complement pathway (including C4A as analysed in (Donohoe *et al.*, 2018) and closely

related gene CSMD1 (Donohoe *et al.*, 2013) maybe associated with SZ risk, the complement pathway as a whole is associated with neurodevelopmental processes related to cognition. This is evident in the effects of PGS on IQ in the healthy control population for full scale IQ, and in the whole group irrespective of diagnosis. At a gene-set level, this may suggest that the wider complement pathway is more strongly associated with neurodevelopmental processes important to global cognitive development than with illness processes per se. In the context of the present study and previous studies on individual complement genes carried out by our group (Donohoe *et al.*, 2018, Donohoe *et al.*, 2013), while the role of an individual gene in neurodevelopment may make it particularly important for illness risk (e.g. C4A and SZ), the pathway's broader neurodevelopmental role is largely independent of SZ risk.

Enrichment analysis of complement genes

In addition to observing enrichment to the overall complement pathway for IQ, we found that 32 of the available 90 genes listed in the complement pathway showed gene-based association with IQ. 12 of these genes remained significant after correction for multiple testing, with SERPING1 the highest associated. SERPING1, also known as C1 inhibitor, has been suggested to participate in neuronal stem cell proliferation, and play a role in the three activation arms of the complement system (Gorelik *et al.*, 2017). Knockdown of circulating C1 inhibitor induces neurovascular impairment, glial cell activation, neuroinflammation, and behavioral deficits (Farfara *et al.*, 2019). Using a shRNA mediated knock-out of SERPING1 to silence the gene expression, a recent study suggests knock-out of this gene impairs neuronal migration and affects brain development (Gorelik *et al.*, 2017). Preliminary evidence in a study of complement genes in twins with SZ implicated peripheral expression of C5 and SERPING1 to be associated with cortical thinning in the superior frontal region (Allswede *et al.*, 2018); SERPING1 has also been found to be overexpressed in the amygdala

of patients with SZ (Chang *et al.*, 2017). In animal studies, C1 inhibitor created by SERPING1 attenuated acute neurobehavioral deficits, ischemic volume, and neurodegeneration (Longhi *et al.*, 2009). Therefore, SERPING1 is a gene pivotal to inflammatory response regulation, neural development, and a key contributor to complement cascade activation.

PGS analysis of complement gene-set

In regression analysis in a sample of 1000 individuals with psychosis and healthy controls, increased PGS (i.e. carrying a higher proportion of IQ-associated alleles) was associated with increased scores across several IQ measures (Pre-morbid IQ as measured by the WTAR, Verbal and Full-scale IQ from the WAIS) in the whole sample and healthy controls only. However, we did not have a specific hypothesis around this aspect of cognitive performance, and instead interpret the data to indicate a broad association with each of the measures available, depending of the sensitivity of each measure used. The direction of this association is interesting given multiple previous reports citing upregulation of complement genes (theoretically driving increased neuroinflammation and synaptic pruning) to be associated with cognitive decline (Gigante *et al.*, 2011, Heyer *et al.*, 2013, Shi *et al.*, 2015, Song *et al.*, 2014). In enrichment analysis, genes related to the phenotype of IQ (SERPING1, CD46) tended to be regulators of the complement system, thus our PGS may represent regulatory activities of the complement cascade, inhibiting inappropriate activation of complement and inflammatory processes (Farfara et al., 2019; Cardone et al., 2011; Winston et al.,2019). This finding was in the overall sample of healthy participants and patients, showing an association between complement and cognition irrespective of diagnosis. This is consistent with MAGMA analysis, as complement genes were enriched for the phenotype of IQ in a GWAS of healthy participants, but no enrichment of these genes for the phenotype of SZ was observed. When taking cases and healthy controls separately, effects on these measures of IQ

were no longer significant, but full scale IQ was found to be affected, explaining 2% of variance, a higher proportion than initial findings in patients.

The complement pathway, neurodevelopment, and cognition

The current results are consistent with several studies linking genetic variation within complement to cognitive processes. In human studies of complement genes, CFH- a gene significantly enriched for IQ in MAGMA analysis- has been seen to influence cognitive decline in several different cohorts, including healthy controls, patients with SZ (Zhang *et al.*, 2017), patients with AD (Gezen-Ak *et al.*, 2013), and post-operative patients with heart conditions (Gigante *et al.*, 2011, Heyer *et al.*, 2013). In another study of post-operative cognitive decline, the complement gene CD59 (which was also significantly enriched for IQ in our analysis) was found to be upregulated in patients with neurocognitive decline as measured by assessments of memory, executive function, attention, language, and global cognition (Ramlawi *et al.*, 2007), suggesting that several elements of the complement system influence this cognitive phenotype. Again, this cognitive decline may be the result of upregulated neural inflammation, attributed to various complement components.

In addition to C4, other complement genes have been associated with variation in cognition via biological processes related to synaptic pruning. Chibnik et al. (Chibnik *et al.*, 2011) found an association between variation in the CR1 gene and age related changes in cognition. The CR1 gene is associated with general loss of cognitive function, and with the development of AD, possibly due to alternations in neural circuits via excessive synaptic pruning (Chibnik *et al.*, 2011, Chung *et al.*, 2014). In a sample of healthy controls only, CR1 was also found to contribute to a reduction in entorhinal cortex volume (Bralten *et al.*, 2011). This further supports a theory that complement gene expression plays a role in neural processes throughout development, such as synaptic pruning, which may be particularly disrupted in the development of SZ (Shatz, 2009).

Limitations

The 'complement' gene-set which formed the basis of this study was curated by Birnbaum (Birnbaum *et al.*, 2018), based on multiple independent publicly available databases (GO, KEGG, IMPORT, IPA, and IMMUNOME) combined with a complement gene-set curated by Qian et al (Qian *et al.*, 2019). While this list is an empirically well validated, as with any list, some gene inclusions/exclusions will make it imperfect. For example, CSMD1, a regulator of C4 and a multiple-domain regulator of the complement system was not included in this gene-set (Kraus *et al.*, 2006). Genetic variation at a locus corresponding to CSMD1 has previously been associated with both increased SZ risk, cognitive performance (Donohoe *et al.*, 2013), and variation in brain structure and function (Rose *et al.*, 2013).

Another limitation of the gene-set chosen is that some components of the complement pathway are neuroprotective, and others neurodegenerative, with very diverse functions. The effects of individual complement genes on cognition are likely to be both nuanced and interactive and this may be masked when viewed as a set. For instance, increased C5aR (which is expressed in neural stem cells) is said to increase neurogenesis (Hernandez *et al.*, 2017), whereas increased CR1 is said to eliminate neural connections (Chibnik *et al.*, 2011). In other words, some aspects of the complement system may disrupt neural functioning when expressed at a high level, while others may disrupt outcomes when expressed at a low level. Therefore, having one score for complement components irrespective of direction of affects could cancel out the full extent that complement genes affect cognition. Assuming a linear relationship between genes in the PGS and cognition may be an oversight, as some biological processes could be non-linear, with too little or too much complement activity impairing outcome (Veerhuis *et al.*, 2011). Evidence from proteomic profiling of blood plasma samples in children indicated that a majority of complement proteins were upregulated in those who

go on to develop psychotic disorders, although some downregulated complement proteins were also associated with disease (English *et al.*, 2017). Further investigation will be required to understand the direction of the association observed here- for example, by subdividing variants into those associated with up-regulation or down-regulation of complement expression using proteomic analysis. More studies establishing the effect each individual gene has on cognition could shed light on the functional and phenotypic effects of up-regulation or down-regulation, and therefore future gene-sets could be weighted based on effect, or reversed scores could be used. A final potential limitation is that the complement PGS was created using weightings from a GWAS of IQ, a phenotype whose genetic architecture is likely to differ from that of the memory phenotype used in our cognitive analysis. Using other, more refined GWAS information to create PGS for specific cognitive measures/domains will be important to shed further light on the specific role of complement in cognition.

Future directions

Given that an immune response is principally a response to environmental factors (infection, stress), the inclusion of environmental data relevant to immune function would enhance analysis of immune relevant genetic variation. The multi-factorial polygenic threshold model of SZ posits that a large number of genetic risk factors with distinct, aggregate, small effects and environmental factors that can interact with established genetic risk factors (Nimgaonkar *et al.*, 2017). Similarly, we and others have shown that environmental processes associated with immune function have a significant impact on cognitive development (Rokita *et al.*, 2018). Future studies of SZ risk, cognitive function, and genetic variation related to immune function may benefit from the inclusion of information about exposure to relevant environmental factors (maternal infection, early social adversity)

likely to mediate the relationship between these variables. Finally, structural and functional brain imaging studies of immune genetics related to SZ may also help shed light on how these immune and complement genetics affect neural structure and function.

Conclusion

This study contributes to the growing literature on the role of complement in brain function, SZ pathology and cognition. An association between a previously curated complement gene-set and cognitive function was supported based on both enrichment analysis of GWAS data and polygenic score analysis. This was in the absence of a similar association between complement and SZ risk, despite some of the individual genes within this gene-set previously showing genome wide association with SZ. In terms of the broader genetic architecture of SZ, these findings suggest that while SZ risk variants are likely to be found at neurodevelopmentally relevant loci (genes involved in synaptogenesis). When complement gene-sets are taken as a whole, their relevance is to neurodevelopment, not illness.

Acknowledgements & Funding

This publication is the work of the authors and Jessica Holland and Gary Donohoe will serve as guarantors for the contents of this paper. This research was funded through an Irish Research Council PhD scholarship grant to JH, and European Research Council (grant no.677467) and Science Foundation Ireland (grant no.16/ERCS/3787) funding to GD.

References

Allswede, D.M., Zheutlin, A.B., Chung, Y., Anderson, K., Hultman, C.M., Ingvar, M. & Cannon, T.D. (2018) Complement gene expression correlates with superior frontal cortical thickness in humans. *Neuropsychopharmacology,* **43,** 525.

Andreasen, N. (1984) The scale for the assessment of positive and negative symptoms (SAPS and SANS). Iowa City, IA, USA: Iowa University Press.

Athanasiu, L., Giddaluru, S., Fernandes, C., Christoforou, A., Reinvang, I., Lundervold, A.J., Nilsson, L.-G., Kauppi, K., Adolfsson, R. & Eriksson, E. (2017) A genetic association study of CSMD1 and CSMD2 with cognitive function. *Brain, behavior, and immunity,* **61,** 209-216.

Barnes, M., Freudenberg, J., Thompson, S., Aronow, B. & Pavlidis, P. (2005) Experimental comparison and cross-validation of the Affymetrix and Illumina gene expression analysis platforms. *Nucleic acids research,* **33,** 5914-5923.

Birnbaum, R., Jaffe, A., Chen, Q., Shin, J., Schubert, C.R., O'Donnell, P., Quan, J., Wendland, J.R., Xi, H.S. & Winslow, A.R. (2018) Investigating the neuroimmunogenic architecture of schizophrenia. *Molecular psychiatry,* **23,** 1251.

Bralten, J., Franke, B., Arias-Vásquez, A., Heister, A., Brunner, H.G., Fernández, G. & Rijpkema, M. (2011) CR1 genotype is associated with entorhinal cortex volume in young healthy adults. *Neurobiology of aging,* **32,** 2106. e2107-2106. e2111.

Cardone, J., Le Friec, G. & Kemper, C. (2011) CD46 in innate and adaptive immunity: an update. *Clinical & Experimental Immunology,* **164,** 301-311.

Chang, X., Liu, Y., Hahn, C., Gur, R., Sleiman, P. & Hakonarson, H. (2017) RNA-seq analysis of amygdala tissue reveals characteristic expression profiles in schizophrenia. *Translational psychiatry,* **7,** e1203.

Chibnik, L.B., Shulman, J.M., Leurgans, S.E., Schneider, J.A., Wilson, R.S., Tran, D., Aubin, C., Buchman, A.S., Heward, C.B. & Myers, A.J. (2011) CR1 is associated with amyloid plaque burden and age-related cognitive decline. *Annals of neurology,* **69,** 560-569.

Chung, S.J., Kim, M.-J., Kim, Y.J., Kim, J., You, S., Jang, E.H., Kim, S.Y. & Lee, J.-H. (2014) CR1, ABCA7, and APOE genes affect the features of cognitive impairment in Alzheimer's disease. *Journal of the neurological sciences,* **339,** 91-96.

Consortium, I.S.G. & 2, W.T.C.C.C. (2012) Genome-wide association study implicates HLA-C* 01: 02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biological psychiatry,* **72,** 620-628.

Cosgrove, D., Harold, D., Mothersill, O., Anney, R., Hill, M., Bray, N., Blokland, G., Petryshen, T., Donnelly, P. & Bates, L. (2017) MiR-137-derived polygenic risk: effects on cognitive performance in patients with schizophrenia and controls. *Translational psychiatry,* **7,** e1012.

de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. (2015) MAGMA: generalized gene-set analysis of GWAS data. *PLoS computational biology,* **11,** e1004219.

Donohoe, G., Holland, J., Mothersill, D., McCarthy-Jones, S., Cosgrove, D., Harold, D., Richards, A., Mantripragada, K., Owen, M. & O'Donovan, M. (2018) Genetically predicted complement component 4A expression: effects on memory function and middle temporal lobe activation. *Psychological medicine,* **48,** 1608-1615.

Donohoe, G., Walters, J., Hargreaves, A., Rose, E., Morris, D., Fahey, C., Bellini, S., Cummins, E., Giegling, I. & Hartmann, A. (2013) Neuropsychological effects of the CSMD1

genome-wide associated schizophrenia risk variant rs10503253. *Genes, Brain and Behavior,* **12,** 203-209.

Dunkelberger, J.R. & Song, W.-C. (2010) Complement and its role in innate and adaptive immune responses. *Cell research,* **20,** 34.

English, J.A., Lopez, L.M., O'Gorman, A., Föcking, M., Hryniewiecka, M., Scaife, C., Sabherwal, S., Wynne, K., Dicker, P. & Rutten, B.P. (2017) Blood-Based Protein Changes in Childhood Are Associated With Increased Risk for Later Psychotic Disorder: Evidence From a Nested Case–Control Study of the ALSPAC Longitudinal Birth Cohort. *Schizophrenia bulletin,* **44,** 297-306.

Euesden, J., Lewis, C.M. & O'reilly, P.F. (2014) PRSice: polygenic risk score software. *Bioinformatics,* **31,** 1466-1468.

Farfara, D., Feierman, E., Richards, A., Revenko, A.S., MacLeod, R.A., Norris, E.H. & Strickland, S. (2019) Knockdown of circulating C1 inhibitor induces neurovascular impairment, glial cell activation, neuroinflammation, and behavioral deficits. *Glia*.

First, M.B., Spitzer, R.L., Gibbon, M. & Williams, J.B. (2002) Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition. SCID-I/P.

Föcking, M., Sabherwal, S., Cates, H.M., Scaife, C., Dicker, P., Hryniewiecka, M., Wynne, K., Rutten, B.P., Lewis, G. & Cannon, M. (2019) Complement pathway changes at age 12 are associated with psychotic experiences at age 18 in a longitudinal population-based study: evidence for a role of stress. *Molecular psychiatry***,** 1.

Gezen-Ak, D., Dursun, E., Hanağası, H., Bilgiç, B., Lohman, E., Araz, Ö.S., Atasoy, İ.L., Alaylıoğlu, M., Önal, B. & Gürvit, H. (2013) BDNF, TNFα, HSP90, CFH, and IL-10 serum levels in patients with early or late onset Alzheimer's disease or mild cognitive impairment. *Journal of Alzheimer's Disease,* **37,** 185-195.

Gigante, P.R., Kotchetkov, I.S., Kellner, C.P., Haque, R., Ducruet, A.F., Hwang, B.Y., Solomon, R.A., Heyer, E.J. & Connolly, E.S. (2011) Polymorphisms in complement component 3 (C3F) and complement factor H (Y402H) increase the risk of postoperative neurocognitive dysfunction following carotid endarterectomy. *Journal of Neurology, Neurosurgery & Psychiatry,* **82,** 247-253.

Gorelik, A., Sapir, T., Woodruff, T.M. & Reiner, O. (2017) Serping1/C1 inhibitor affects cortical development in a cell autonomous and non-cell autonomous manner. *Frontiers in cellular neuroscience,* **11,** 169.

Hernandez, M.X., Jiang, S., Cole, T.A., Chu, S.-H., Fonseca, M.I., Fang, M.J., Hohsfield, L.A., Torres, M.D., Green, K.N. & Wetsel, R.A. (2017) Prevention of C5aR1 signaling delays microglial inflammatory polarization, favors clearance pathways and suppresses cognitive loss. *Molecular neurodegeneration,* **12,** 66.

Heyer, E.J., Kellner, C.P., Malone, H.R., Bruce, S.S., Mergeche, J.L., Ward, J.T. & Connolly, E.S. (2013) Complement polymorphisms and cognitive dysfunction after carotid endarterectomy. *Journal of neurosurgery,* **119,** 648-654.

Holdnack, H. (2001) Wechsler test of adult reading: WTAR. *San Antonio, TX: The Psychological Corporation*.

Jung, R.E. & Haier, R.J. (2007) The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behavioral and Brain Sciences,* **30,** 135-154.

Kraus, D.M., Elliott, G.S., Chute, H., Horan, T., Pfenninger, K.H., Sanford, S.D., Foster, S., Scully, S., Welcher, A.A. & Holers, V.M. (2006) CSMD1 is a novel multiple domain complement-regulatory protein highly expressed in the central nervous system and epithelial tissues. *The Journal of Immunology,* **176,** 4419-4430.

Longhi, L., Perego, C., Ortolano, F., Zanier, E.R., Bianchi, P., Stocchetti, N., McIntosh, T.K. & De Simoni, M.G. (2009) C1-inhibitor attenuates neurobehavioral deficits and reduces contusion volume after controlled cortical impact brain injury in mice. *Critical care medicine,* **37,** 659-665.

McAllister, A.K. (2014) Major histocompatibility complex I in brain development and schizophrenia. *Biological psychiatry,* **75,** 262-268.

Nettiksimmons, J., Tranah, G., Evans, D.S., Yokoyama, J.S. & Yaffe, K. (2016) Gene-based aggregate SNP associations between candidate AD genes and cognitive decline. *Age,* **38,** 41.

Nimgaonkar, V.L., Prasad, K.M., Chowdari, K.V., Severance, E.G. & Yolken, R.H. (2017) The complement system: a gateway to gene–environment interactions in schizophrenia pathogenesis. *Molecular psychiatry,* **22,** 1554.

O'Gráda, C., Barry, S., McGlade, N., Behan, C., Haq, F., Hayden, J., O'Donoghue, T., Peel, R., Morris, D.W. & O'Callaghan, E. (2009) Does the ability to sustain attention underlie symptom severity in schizophrenia? *Schizophrenia research,* **107,** 319-323.

Orsini, F., De Blasio, D., Zangari, R., Zanier, E.R. & De Simoni, M.-G. (2014) Versatility of the complement system in neuroinflammation, neurodegeneration and brain homeostasis. *Frontiers in cellular neuroscience,* **8,** 380.

Pardiñas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D. & Hamshere, M.L. (2018) Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature genetics,* **50,** 381.

Qian, D., Liu, H., Wang, X., Ge, J., Luo, S., Patz Jr, E.F., Moorman, P.G., Su, L., Shen, S. & Christiani, D.C. (2019) Potentially functional genetic variants in the complement-related immunity gene-set are associated with non-small cell lung cancer survival. *International journal of cancer,* **144,** 1867-1876.

Ramlawi, B., Otu, H., Rudolph, J.L., Mieno, S., Kohane, I.S., Can, H., Libermann, T.A., Marcantonio, E.R., Bianchi, C. & Sellke, F.W. (2007) Genomic expression pathways associated with brain injury after cardiopulmonary bypass. *The Journal of thoracic and cardiovascular surgery,* **134,** 996-1005. e1004.

Ricklin, D., Hajishengallis, G., Yang, K. & Lambris, J.D. (2010) Complement: a key system for immune surveillance and homeostasis. *Nature immunology,* **11,** 785.

Ripke, S., Neale, B.M., Corvin, A., Walters, J.T., Farh, K.-H., Holmans, P.A., Lee, P., Bulik-Sullivan, B., Collier, D.A. & Huang, H. (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature,* **511,** 421.

Robbins, T.W., James, M., Owen, A.M., Sahakian, B.J., McInnes, L. & Rabbitt, P. (1994) Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia and Geriatric Cognitive Disorders,* **5,** 266-281.

Rokita, K.I., Dauvermann, M.R. & Donohoe, G. (2018) Early life experiences and social cognition in major psychiatric disorders: A systematic review. *European psychiatry,* **53,** 123-133.

Rose, E.J., Morris, D.W., Hargreaves, A., Fahey, C., Greene, C., Garavan, H., Gill, M., Corvin, A. & Donohoe, G. (2013) Neural effects of the CSMD 1 genome-wide associated schizophrenia risk variant rs10503253. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* **162,** 530-537.

Sarma, J.V. & Ward, P.A. (2011) The complement system. *Cell and tissue research,* **343,** 227-235.

Savage, J.E., Jansen, P.R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C.A., Nagel, M., Awasthi, S., Barr, P.B. & Coleman, J.R. (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature genetics,* **50,** 912.

Schumann, G., Loth, E., Banaschewski, T., Barbot, A., Barker, G., Büchel, C., Conrod, P., Dalley, J., Flor, H. & Gallinat, J. (2010) The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Molecular psychiatry,* **15,** 1128.

Sekar, A., Bialas, A.R., De Rivera, H., Davis, A., Hammond, T.R., Kamitaki, N., Tooley, K., Presumey, J., Baum, M. & Van Doren, V. (2016) Schizophrenia risk from complex variation of complement component 4. *Nature,* **530,** 177.

Shatz, C.J. (2009) MHC class I: an unexpected role in neuronal plasticity. *Neuron,* **64,** 40-45.

Shi, Q., Colodner, K.J., Matousek, S.B., Merry, K., Hong, S., Kenison, J.E., Frost, J.L., Le, K.X., Li, S. & Dodart, J.-C. (2015) Complement C3-deficient mice fail to display age-related hippocampal decline. *Journal of Neuroscience,* **35,** 13029-13042.

Song, F., Poljak, A., Kochan, N.A., Raftery, M., Brodaty, H., Smythe, G.A. & Sachdev, P.S. (2014) Plasma protein profiling of Mild Cognitive Impairment and Alzheimer's disease using iTRAQ quantitative proteomics. *Proteome science,* **12,** 5.

Veerhuis, R., Nielsen, H.M. & Tenner, A.J. (2011) Complement in the brain. *Molecular immunology,* **48,** 1592-1603.

Wechsler, D. (1997) Wechsler adult intelligence scale-III. San Antonio, TX: The Psychological Corporation.

Whitton, L., Cosgrove, D., Clarkson, C., Harold, D., Kendall, K., Richards, A., Mantripragada, K., Owen, M.J., O'donovan, M.C. & Walters, J. (2016) Cognitive analysis of schizophrenia risk genes that function as epigenetic regulators of gene expression. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics,* **171,** 1170-1179.

Winston, C.N., Goetzl, E.J., Schwartz, J.B., Elahi, F.M. & Rissman, R.A. (2019) Complement protein levels in plasma astrocyte-derived exosomes are abnormal in conversion from mild cognitive impairment to Alzheimer's disease dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring,* **11,** 61-66.

Zhang, C., Lv, Q., Fan, W., Tang, W. & Yi, Z. (2017) Influence of CFH gene on symptom severity of schizophrenia. *Neuropsychiatric disease and treatment,* **13,** 697.