American Journal of Medical Genetics Part B: Neuropsychiatric Genetics

WILEY

RESEARCH ARTICLE OPEN ACCESS

The Role of SLC39A8.p.(Ala391Thr) in Schizophrenia Symptom Severity and Cognitive Ability: Cross-Sectional Studies of Schizophrenia and the General UK Population

Sophie E. Smart[1] | Sophie E. Legge[1] | Eilidh Fenner[1] | Antonio F. Pardiñas[1] | Grace Woolway[1] | Amy J. Lynham[1] | Valentina Escott-Price[1] | Jeremy Hall[2] | Lawrence Wilkinson[1] | Peter Holmans[1] | Michael C. O'Donovan[1] | Michael J. Owen[1,2] | James T.R. Walters[1]

[1]Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK | [2]Neuroscience and Mental Health Innovation Institute, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK

Correspondence: James T.R. Walters (WaltersJT@cardiff.ac.uk)

Received: 23 January 2025 | **Revised:** 28 March 2025 | **Accepted:** 29 April 2025

Funding: This study was funded by a grant from Takeda Pharmaceuticals Ltd. to Cardiff University. Takeda Pharmaceuticals Ltd. was not involved in the design of the study or the interpretation of the results. S.E.L. was supported by a grant from NIMH (Award R01MH124873) and a Medical Research Council Centre programme grant MR/P005748/1. E.F. was supported by a Wellcome Trust Integrative Neuroscience PhD Studentship (108891/B/15/Z/WT). A.F.P. was supported by the Academy of Medical Sciences Springboard award (SBF005/1083). S.E.S., A.F.P., M.C.O. and J.T.R.W. was supported by funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 964874.

Keywords: metal ion transporters | psychosis | rs13107325 | whole-exome sequencing | ZIP8

ABSTRACT

The missense SNP NC_000004.12:g.102267552C>T (also known as SLC39A8.p.(Ala391Thr), rs13107325) in *SLC39A8* encodes a zinc transporter. This SNP has been linked to schizophrenia and is the likely causal variant for one of the genome-wide association loci associated with the disorder. Using regression analyses, we tested whether the schizophrenia-risk allele at p.(Ala391Thr) was associated with schizophrenia-related phenotypes, including positive, negative, and disorganized symptoms, cognitive ability, educational attainment, and age of psychosis onset, within three schizophrenia cohorts (combined $N = 1232$) and, with equivalent phenotypes, in a sample of population controls (UK Biobank, $N = 355{,}069$). We also used the population controls to test for associations with rare protein-truncating and deleterious missense variants within *SLC39A8*. Within the schizophrenia cohorts, after correction for multiple testing, p.(Ala391Thr) was not significantly associated with any schizophrenia-related phenotypes. In the unaffected participants from the UK Biobank, the schizophrenia-risk allele at p.(Ala391Thr) was associated with significantly poorer cognitive ability and fluid intelligence, a lower probability of obtaining GCSEs or a degree-level qualification, and fewer years in education. There was no association between p.(Ala391Thr) and self-reported psychotic experiences in this cohort. Rare variants in *SLC39A8* were nominally associated with poorer cognitive ability, but these associations did not survive correction for multiple testing. The schizophrenia-risk allele was associated with poorer cognitive ability, but not psychotic experiences, in a volunteer sample drawn from the general population. We found no evidence that p.(Ala391Thr) was associated with symptom severity in schizophrenia. To understand the impact of rare variants in *SLC39A8* on cognitive impairment, larger independent samples are required.

1 | Introduction

The gene *SLC39A8* encodes metal cation symporter ZIP8, which enables the movement of at least five trace elements (manganese, zinc, iron, selenium, and cobalt) from outside the cell into the cytosol (Nebert and Liu 2019; Ng et al. 2015). *SLC39A8* contains a missense single nucleotide polymorphism (SNP), NC_000004.12:g.102267552C>T (also known as SLC39A8.p.(Ala391Thr) and rs13107325), in which alternative alleles encode alanine (Ala; C allele) or threonine (Thr; T allele) on chromosome 4 (Costas 2018). In vitro, the minor allele (T) leads to reduced uptake of manganese and cadmium (Fujishiro et al. 2022), the latter a toxic environmental pollutant, and decreased synaptic uptake of zinc (Tseng et al. 2021).

p.(Ala391Thr) is associated with an increased risk of schizophrenia, a heterogeneous disorder whose presentation includes delusions, hallucinations, behavioral disturbance, social withdrawal, and cognitive impairment. In the most recent genome-wide association study (GWAS), of 76,755 schizophrenia cases and 243,649 controls, the minor allele was associated with increased risk of schizophrenia (odds ratio = 1.17, $p = 1.92 \times 10^{-21}$, minor allele frequency in schizophrenia cases [MAF_{cases}] = 6.9%, $MAF_{controls}$ = 6.0%; multi-ancestry meta-analysis) (Trubetskoy et al. 2022). Furthermore, fine-mapping suggested that p.(Ala391Thr) was likely to be the causal variant underpinning the association with the genomic region containing *SLC39A8*, the SNP having a posterior probability of being causally associated with schizophrenia of 99.20%. Only nine out of 255 loci were fine-mapped to one causal variant (Trubetskoy et al. 2022); only two of these were missense variants, and only p.(Ala391Thr) had a Combined Annotation-Dependent Depletion (CADD) score of > 20 (CADD = 23.80) making it one of the 1% most deleterious variants in the genome (the other missense variant had a CADD score > 10 [CADD = 15.39] making it one of the 10% most deleterious variants in the genome).

In general population samples, including the UK Biobank, p.(Ala391Thr) has not been associated with reported psychotic experiences ($Z = 0.99$, $p = 0.889$) (Legge et al. 2019), but the schizophrenia risk allele has been associated with lower intelligence ($Z = -9.49$, $p = 2.23E\text{-}21$) (Savage et al. 2018), fewer years in education (Beta = −0.02, $p = 1.08E\text{-}13$) (Okbay et al. 2022), and, using a GWAS-by-subtraction design, with the cognitive aspect of educational attainment (noncognitive, $Z = 2.78$, $p = 0.005$; cognitive, $Z = -9.87$, $p = 5.59E\text{-}23$) (Demange et al. 2021). However, these studies use a GWA design and apply strong correction for multiple testing, which excludes variants with weaker effects. To date, no studies have examined the relationship between p.(Ala391Thr) and the clinical presentation or cognitive ability of people with schizophrenia.

Our primary aim was to test whether schizophrenia-risk alleles at p.(Ala391Thr) are associated with phenotypes that capture the clinical presentation of schizophrenia. We hypothesized that the schizophrenia-risk allele (T) at p.(Ala391Thr) would be associated with a more severe clinical presentation, poorer cognitive ability and educational attainment, and younger age of onset. We expanded the analysis to test for similar phenotypic associations in a large sample of the general population (UK Biobank unaffected controls). Given that there is some evidence that *SLC39A8* is also constrained for loss of function and missense variants (Genome Aggregation Database [gnomAD] v2.1.1; Karczewski et al. 2020), we used the UK Biobank to test for phenotypic associations with rare protein-truncating and deleterious missense variants within *SLC39A8*.

2 | Methods

This study followed the guidelines outlined in the STrengthening the REporting of Genetic Association Studies (STREGA) (Little et al. 2009), an extension of the STROBE Statement (checklist in the Supporting Information).

We analyzed data from three cohorts recruited at Cardiff University and data from the UK Biobank. Full details on how samples were ascertained and how phenotypes were measured, calculated, and standardized can be found in the Supporting Information. A brief overview is provided below.

2.1 | Cardiff Schizophrenia Samples

2.1.1 | Participants

Participants were recruited into three cross-sectional studies, all of which have been previously described: CardiffCOGS (Legge et al. 2020), Cardiff F-series (Williams et al. 2006) and Cardiff Affected-Sibs (also known as SibPairs; Williams et al. 1999) samples. The Cardiff Affected-Sibs sample includes a single affected individual from families with two or more siblings diagnosed with schizophrenia or schizoaffective disorder. All participants were recruited from in-patient, voluntary sector, and community mental health services across the UK and underwent a clinical research interview based on the Schedules for Clinical Assessment in Neuropsychiatry (SCAN), with further information available from clinical records. For these analyses, we only retained participants who met DSM-IV or ICD-10 criteria for a diagnosis of schizophrenia or schizoaffective disorder, depressed type (Dennison et al. 2021).

2.1.2 | Phenotypes

The Scale for the Assessment of Positive Symptoms (SAPS) and the Scale for the Assessment of Negative Symptoms (SANS) were scored on a lifetime worst basis using information from the SCAN interview and lifetime psychiatric clinical records. These records were also used to ascertain age at psychosis onset. The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery was used as a measure of current cognitive ability in the CardiffCOGS sample. As described in Legge et al. (2021), a confirmatory factor analysis (CFA) framework was used to estimate phenotype-derived factor scores from the symptom and cognitive ratings, which we refer to as symptom dimensions. Using all three samples, the best model had three

symptom dimensions: positive symptoms (SAPS global hallucinations and SAPS global delusions), negative symptoms of diminished expressivity (SANS global affective flattening and SANS global alogia), and disorganized symptoms (SAPS global positive formal thought disorder and SANS inappropriate affect). As CardiffCOGS was the only sample to have data on cognitive ability, a second CFA framework, using only this sample, was used to estimate phenotype-derived factor scores from the MATRICS domain scores as well as the symptom ratings. The best model had five dimensions: positive symptoms (as above), negative symptoms of diminished expressivity (as above), disorganized symptoms (as above), negative symptoms of motivation and pleasure (SANS global anhedonia/asociality and SANS global avolition/apathy), and cognitive ability (all MATRICS domains apart from social cognition).

As well as the cognitive domain, which captures current cognitive ability, premorbid IQ was assessed in the CardiffCOGS sample using the National Adult Reading Test (NART). Educational attainment was measured across all three samples using years in education and highest educational qualification (General Certificate of Secondary Education [GCSE]/no GCSE, note GCSEs are taken by most UK pupils upon the completion of compulsory education [from 1972 to 2013 this was at age 16], and degree/no degree).

2.1.3 | Genotypes

Genome-wide SNP data for the three samples were curated and harmonized as part of DRAGON-Data (Lynham et al. 2023) (see Supporting Information). Genotypes for p.(Ala391Thr) (INFO = 0.996) were extracted for each participant. Principal components were calculated using pruned SNPs.

Genetic ancestry probabilities were calculated using Ancestry Informative Markers (AIMs) derived from the Allen Ancient DNA Resource reference panel, linear discriminant analysis (LDA), and biogeographical categories defined by Huddart et al. (2019). AIMs are genetic variants with highly divergent allele frequencies across biogeographical genetic ancestries, which can capture genetic associations between an individual and a particular (sub)continental population (Bulbul et al. 2016). Ancestry groups were determined by assigning individuals to their most probable biogeographical category as inferred by LDA (determined using Youden's index as optimality criterion), with individuals not meeting a probability threshold for any category being assigned to an "admixed" group. The schizophrenia risk allele is thought to have been under recent positive selection in Europeans (Li et al. 2016); in Phase 3 of the 1000 Genomes Project, the schizophrenia allele is present in around 8% of European samples but is reported to be almost absent from other populations (Costas 2018). In our data, only Europeans carried two copies of the schizophrenia risk allele (see Table S1). However, to increase our sample size and improve the generalizability of our findings, we restricted our sample to participants from ancestries where the schizophrenia risk allele was present in at least one individual: African American/Afro-Caribbean (0.48%), Central/South Asian (1.06%), European (98.21%), and Middle Eastern/North African (0.24%) (note the Middle Eastern/North African ancestry is referred to as Near Eastern in Huddart et al. (2019)).

2.2 | UK Biobank

2.2.1 | Participants

Participants were from the UK Biobank (UKBB), a large-scale biomedical database of individuals aged between 40 and 69 who were recruited from across the UK (Sudlow et al. 2015). Participants underwent extensive phenotyping. All UKBB field IDs used in this study are reported in the Supporting Information. For these analyses, we removed participants with a psychotic spectrum disorder (ICD-10 F20-F29, see Supporting Information).

We also restricted our sample to participants who self-reported White British or Irish ethnicity, as is consistent with the wider literature using the UKBB (Chen et al. 2023; Legge et al. 2024). Our phenotypes are derived from data collected from online follow-ups, and non-White populations are more likely to be lost to follow-up (Homman et al. 2021; Lamers et al. 2012). The proportion of self-reported non-White participants in the UKBB drops from 5.4% to 3% for the online follow-up mental health questionnaire (Davis et al. 2020). Methods (Schoeler et al. 2023) and genetically derived variables (e.g., polygenic scores) (Adams et al. 2020; Tyrrell et al. 2021) that could hypothetically be used to control for participation bias have so far been exclusively developed using the self-reported White ethnicity subgroup of the UKBB. In addition, for our rare variant analysis, we do not believe that we could accurately define ultra-rare variants in the non-European subgroups; our previous work has shown that the number of rare variants per participant differs across PC-defined strata associated with self-reported ethnicities in the UKBB, with self-reported White British or Irish participants carrying fewer singletons (Fenner et al. 2023). Nor would we have sufficient statistical power to detect associations between rare coding variants and phenotypes in the non-European subgroups (Chen et al. 2023).

2.2.2 | Phenotypes

An abridged version of the Composite International Diagnostic Interview psychosis module (lifetime version) was used in the online follow-up mental health questionnaire. As in Legge et al. (2019), we derived three overlapping binary variables: (i) any psychotic experience defined as a positive response to any of the four symptom questions; (ii) a distressing psychotic experience, defined as any psychotic experience that was rated as "a bit," "quite," or "very" distressing; and (iii) multiple occurrences of psychotic experiences, defined as any psychotic experience that occurred on more than one occasion. The comparator group for these three variables was comprised of individuals who provided a negative response to all four psychotic experience symptom questions (see Supporting Information). We also looked at delusions of persecution alone (UKBB Field 20,468: "Ever believed in an un-real conspiracy against self") as previous work has suggested that this phenotype in the UKBB is particularly enriched for genetic liability for schizophrenia (Legge et al. 2019).

Cognitive ability was measured using a general intelligence factor, *g*, which is considered a reliable measure of cognitive

ability. As in Fawns-Ritchie and Deary (2020), *g* was calculated using principal component analysis (PCA). Four cognitive tests went into the PCA (numeric memory, reaction time, pairs matching, and trail making test (TMT) B, see Supporting Information and Tables S15–S19 for missingness, skewness, and correlations of all the cognitive tests). The first PC was considered an estimate of *g*. A positive *g* score represents better cognitive performance. Alongside the *g* score, the cognitive test of verbal and numerical reasoning, also referred to as the test for fluid intelligence, was used to measure current cognitive ability. Educational attainment was measured using years in education (for those without a college or university degree) and highest educational qualification (GCSEs/no GCSEs and degree/no degree).

2.2.3 | Genotypes

Genotype data were curated by the UKBB (Bycroft et al. 2018) and as described in Leonenko et al. (2021) (see Supporting Information). The genotypes for p.(Ala391Thr) (INFO = 1.00) were extracted for each participant. Principal components provided by the UKBB were used. To further control for population stratification, which is present within those of white British ancestry in the UKBB (Bycroft et al. 2018), we used the first five principal components to identify a subsample of participants that was relatively genetically homogenous. As in Legge et al. (2019), we computed a Minimum Covariance Determinant (MCD) estimator of location and scatter for each participant, and used these to define a hyper-ellipsoid in a multi-dimensional space that contains the majority of MCD points. We used this hyper-ellipsoid to include participants within the 90th percentile of the MCD distance (see Figure S9).

2.2.4 | Rare Variants

Full methods of rare variant calls are reported in Fenner et al. (2023). In brief, variants were annotated in Hail Team (n.d.) using Ensembl's VEP. Protein-truncating variants (PTVs) were defined as splice acceptor, splice donor, stop-gain or frameshift variants that were annotated as high confidence for causing loss of protein function by Loss-Of-Function Transcript Effect Estimator (LoFTEE; Karczewski et al. 2020). Deleterious missense variants were defined as missense variants with a Rare Exome Variant Ensemble Learner score (REVEL; Ioannidis et al. 2016) > 0.75. We chose this threshold as it has been demonstrated to have high specificity for predicting deleterious variants in the Human Gene Mutation Database (Ioannidis et al. 2016) and is above the recommended threshold of > 0.64 for classifying pathogenic variants (Pejaver et al. 2022). PTVs and missense variants were grouped together into one "damaging" category for analysis. Rare variants were defined as those with allele counts ≤ 5 in the quality controlled sample (N = 179,670) of participants from the 200 k exome sequencing release tranche (UK Biobank Field 32050) reported in Fenner et al. (2023). This allele count equates to a MAF $< 1.4 \times 10^{5}$. We included, as a covariate, the burden of synonymous variants that had allele counts ≤ 5 in the sample reported in Fenner et al. (2023).

3 | Statistical Analysis

Each phenotype was regressed onto p.(Ala391Thr) using linear or logistic regression models. Sex, age at interview, the first five genetic principal components (PCs), and any further PCs which were associated with p.(Ala391Thr) (PC 6 and 9 in the Cardiff F-Series sample) were included as covariates. For the UKBB data, the first 10 genetic PCs and genotype batch were included as covariates. For models where a measure of cognitive ability in the UKBB was used as the outcome, age at interview squared was also included as a covariate (Cornelis et al. 2019); this variable was centered before being transformed to prevent the occurrence of collinearity (Schielzeth 2010). For the Cardiff schizophrenia samples, missing phenotype data were imputed, for each sample separately, using multiple imputation by chained equations (MICE; 100 imputations, 10 iterations in the burn-in period). All regression analyses were run in each of the 100 imputed datasets separately and then pooled using Rubin's rules (Rubin 2004). Model assumptions were checked using the R package "performance" and discussed in the Supporting Information: for models were there was evidence of multicollinearity (variance inflation factor [VIF] ≥ 10), we removed PCs from the model until the VIF for p.(Ala391Thr) was < 10, and for models where there was evidence of heteroscedasticity (Breusch-Pagan Test p-value < 0.01), we recalculated standard errors, confidence intervals, and p-values using Eicker–Huber–White robust "HC2" standard errors. For the Cardiff schizophrenia samples, beta values or log(odds) were meta-analyzed, and weighted using their standard errors, using the R package "metafor" with a fixed-effect inverse-variance weighted model. We corrected for multiple testing of phenotypes in the meta-analysis, and within each sample, using the Benjamini–Hochberg False Discovery Rate (FDR) method and used an alpha level of ≤ 0.05 for the adjusted p-values to determine statistical significance. For the rare variant analysis, each of the phenotypes was regressed against the number of rare variants, and adjusted for the number of rare synonymous variants each person carries in *SLC39A8*, age at interview (and age at interview squared for *g* and FI), sex, exome data PC1-10, and sequencing batch.

4 | Results

Phenotype means and proportions, stratified by p.(Ala391Thr) schizophrenia-risk allele count, are presented for each sample in Tables S5–S7, and S22.

4.1 | Cardiff Schizophrenia Samples

Data were available for 662 participants from CardiffCOGS (mean age at interview [AAI] = 43.31 years, standard deviation [SD] = 12.11; 65.75% male), 422 from Cardiff F-Series (mean AAI = 42.11 years, SD = 14.08; 70.66% male) and 148 from Cardiff SibPairs (mean AAI = 41.68 years, SD = 12.93; 66.21% male). Of the total N = 1232 participants, at p.(Ala391Thr), N = 986 participants carried zero schizophrenia-risk alleles, N = 234 carried one, and N = 12 carried two. The MAF was 10.47%.

The proportion of non-missing data ranged across variables from 69% to 100% (86%–100% CardiffCOGS; 88%–100% Cardiff

TABLE 1 | Fixed-effect meta-analysis of adjusted regression models of p.(Ala391Thr) using the observed/imputed data from the CardiffCOGS ($N = 662$), Cardiff F-Series ($N = 422$), and Cardiff SibPairs ($N = 148$) samples.

	Fixed effect					Heterogeneity			
	Beta/OR	Lower 95% CI	Upper 95% CI	Standard error	*p*	I^2 (%)	H^2	*Q*	*Q p*
Three-factor dimensions									
Positive symptoms	*0.01*	*−0.18*	*0.19*	0.09	0.958	0.00	0.16	0.32	0.854
Negative symptoms of diminished expressivity	*0.02*	*−0.27*	*0.31*	0.15	0.9	0.00	0.00	0.00	0.999
Disorganized symptoms	*0.03*	*−0.19*	*0.24*	0.11	0.818	0.00	0.09	0.17	0.918
Years in education	*−0.09*	*−0.61*	*0.43*	0.27	0.726	0.00	0.23	0.46	0.794
Educational qualification: GCSE	0.82	0.46	1.46	0.3	0.494	0.00	0.02	0.04	0.98
Educational qualification: Degree	1.01	0.46	2.22	0.4	0.972	0.00	0.11	0.22	0.895
Age at psychosis onset	*−0.16*	*−1.26*	*0.94*	0.56	0.776	0.00	0.53	1.06	0.59

Note: Beta coefficients indicated by italic. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) and is used to quantify the magnitude of heterogeneity. H^2 is another metric to quantify the magnitude of heterogeneity and defined as the relative excess in *Q* over its degree of freedom or, in other words, the ratio of variability in effect estimates that is due to heterogeneity rather than sampling error to the sampling variability. *Q* is the chi-squared statistic and assesses whether observed differences in results are compatible with chance alone (a low *p* value provides evidence of heterogeneity of effects, that is, variation in effect estimates beyond chance) but does not quantify the extent of heterogeneity. No False Discovery Rate (FDR) adjusted *p*-values (adjusted for 13 tests) were statistically significant at an alpha level ≤ 0.05.
Abbreviations: CI, confidence interval; OR, odds ratio.

F-Series; 69%–100% Cardiff SibPairs; see Table S4). Overall, missing data imputation performed well; for all variables, imputed data were within the range of the observed data and kernel density estimates of the imputed and observed data showed comparable distributions.

p.(Ala391Thr) was not associated with positive symptoms, negative symptoms, or disorganized symptoms derived in either CFA model (Tables 1 and 2). p.(Ala391Thr) was not associated with current cognitive ability in CardiffCOGS; however, it was nominally associated with a lower NART IQ score, but this association did not survive correction for multiple testing (Table 2). Although I^2 should be interpreted cautiously in meta-analyses of only a few studies (von Hippel 2015), there does not appear to be evidence of between-study heterogeneity after meta-analyzing phenotypes that were present in all three samples (Table 1). When analyzing each sample separately, no associations survived correction for multiple testing (Tables S8–S10) and p.(Ala391Thr) accounted for very little of the variance explained (Tables S11–S13).

4.2 | UK Biobank

Data were available for $N = 355{,}069$ participants from UKBB who did not have a psychotic spectrum disorder (mean AAI = 56.92 years, SD = 7.96; 46.25% male). At p.(Ala391Thr), $N = 304{,}253$ participants carried zero schizophrenia-risk alleles, $N = 48{,}907$ carried one, and $N = 1909$ carried two. The MAF was 7.43%.

The proportion of non-missing data ranged across variables from 15% to 100% (see Table S21). $N = 354{,}509$ indicated that they were willing to attempt the cognitive tests and $N = 116{,}935$ completed the online follow-up mental health questionnaire. As such, 31% of the participants had data on reported psychotic experiences. Fluid intelligence was available for 33% of participants and years in education for 68%. *g* could only be calculated for 15% of participants. 78% and 94% of participants had data on GCSEs and degree level qualifications, respectively.

p.(Ala391Thr) was not associated with reported psychotic experiences. However, the schizophrenia-risk allele was associated with a lower *g* score ($\beta = -0.05$; 95% CI, −0.07 to −0.02; FDR-adjusted *p*-value $= 8.61 \times 10^{-5}$), a lower fluid intelligence score ($\beta = -0.05$; 95% CI, −0.07 to −0.04; FDR-adjusted *p*-value $= 3.35 \times 10^{-10}$), a lower likelihood of obtaining GCSEs (OR = 0.96; 95% CI, 0.94–0.98; FDR-adjusted *p*-value $= 6.21 \times 10^{-5}$) or a degree-level qualification (OR = 0.95; 95% CI, 0.93–0.97; FDR-adjusted *p*-value $= 6.03 \times 10^{-6}$) and fewer years in education ($\beta = -0.03$; 95% CI, −0.05 to 0.00; FDR-adjusted *p*-value = 0.035) (Table 3).

Rare variant calls were available for $N = 134{,}370$ (37.84%) of the participants included in the current study. $N = 46$ carried rare synonymous variants and $N = 13$ carried rare damaging variants ($N = 6$ PTVs and $N = 7$ missense carriers; mean AAI = 55.54 years, SD = 9.92; 46.15% male). Most participants with quality controlled exome sequencing data had missing phenotype data; the proportion of non-missing data ranged across variables from $N = 22{,}476$ (6%; for *g*) to $N = 126{,}333$ (36%; for highest educational qualification, degree).

TABLE 2 | Adjusted regression models of p.(Ala391Thr) using the observed/imputed data and the phenotypes only present in the CardiffCOGS sample ($N = 662$).

	Beta	Lower 95% CI	Upper 95% CI	Standard error	*p*	R^2 (SNP)	R^2 adj. (SNP)	RMSE
Five-factor dimensions								
Positive symptoms	−0.05	−0.14	0.04	0.04	0.264	1.05% (0.20%)	−0.16% (0.05%)	0.469
Negative symptoms of diminished expressivity	0.00	−0.09	0.09	0.05	0.991	3.05% (0.00%)	1.86% (−0.15%)	0.4901
Disorganized symptoms	−0.02	−0.07	0.04	0.03	0.52	0.64% (0.03%)	−0.58% (−0.12%)	0.2972
Negative symptoms of motivation and pleasure	0.02	−0.09	0.13	0.05	0.711	2.98% (0.01%)	1.79% (−0.14%)	0.5811
Cognitive ability	−0.12	−0.26	0.02	0.07	0.095	17.19% (0.11%)	16.18% (−0.02%)	0.7429
NART IQ	−0.21	−0.39	−0.03	0.09	0.021	7.51% (1.12%)	6.38% (0.99%)	0.9593

Note: Adjusted for age at first interview, sex, and principal components (PC) 1–5. No False Discovery Rate (FDR) adjusted *p*-values (adjusted for 13 tests) were statistically significant at an alpha level ≤0.05. (SNP) is the difference between the variance explained by the full model and the variance explained by a covariate-only model and can thus be interpreted as the variance explained by the SNP alone. Tjur's R^2 is an alternative to other pseudo-R^2 values like Nagelkerke's R^2 or Cox-Snell R^2 and can be read like any other (pseudo-)R^2 value. RMSE is the square root of the variance of the residuals and indicates the absolute fit of the model to the data (difference between observed data to model's predicted values). It can be interpreted as the standard deviation of the unexplained variance, and is in the same units as the response variable. Lower values indicate better model fit.
Abbreviation: CI, confidence interval.

Rare variants in *SLC39A8* were nominally associated with lower *g* (N rare variants = 2; $\beta = -1.67$; 95% CI, −2.93 to −0.42; *p*-value = 0.009), lower fluid intelligence (N rare variants = 1; $\beta = -2.31$; 95% CI, −4.25 to −0.38; *p*-value = 0.019), and a lower likelihood of obtaining GCSEs (N rare variants = 10; OR = 0.24; 95% CI, 0.05–0.72; *p*-value = 0.044), but these associations did not survive correction for multiple testing (Table 4). By inclusion in our sample, none of the rare variant carriers had a psychotic spectrum disorder, but we also confirmed that none of the carriers had a diagnosis of Bipolar Disorder (ICD-10 Codes F30 or F31), Intellectual Disability (ICD-10 Codes F70, F71, F72, F78, or F79) or a disorder of glycoprotein metabolism (ICD-10 Code E77) identified from the hospital, death, primary care, and self-report records.

As the schizophrenia-risk allele has been previously associated with responding to questions in the UKBB with "prefer not to answer" and "I don't know" (Mignogna et al. 2023), we tested whether p.(Ala391Thr) was associated with willingness to attempt the cognitive tests or the mental health questionnaire. p.(Ala391Thr) was not associated with attempting to complete the baseline cognitive tests, but the schizophrenia-risk allele was associated with *not* attempting the mental health questionnaire.

5 | Discussion

In this study, we found that the schizophrenia-risk allele at p.(Ala391Thr) was associated with poorer cognitive ability, but not psychotic experiences, in a volunteer sample drawn from the general population without psychotic spectrum disorders. The schizophrenia-risk allele was also nominally associated with lower premorbid IQ in patients with schizophrenia; although larger independent samples are required to confirm this result. Exploratory analysis of rare variants in *SLC39A8* in our subsample of the UKBB suggested that rare variants are nominally associated with poorer cognition but were not decisive due to the low number of rare variant carriers with phenotype data.

Although understanding the potential pathophysiological mechanisms of variants identified by GWAS is challenging, there has been an increasing focus on p.(Ala391Thr) in the context of the etiology of schizophrenia (Costas 2018) as well as the mechanisms of action by which it may impact cognition. p.(Ala391Thr) is thought to lead to synaptic glutamate receptor hypofunction, in part, because of downregulated surface localisation of GluA1, GluA2/3, GluN1, and GluN2A (Tseng et al. 2021). The latter are subunits of the *N*-methyl-d-aspartate (NMDA) receptor and the hypofunctioning of NMDAR has been implicated in the etiology of schizophrenia and specifically impaired learning and memory (Nakazawa and Sapkota 2020). The schizophrenia-risk allele at p.(Ala391Thr) has also been associated with lower manganese levels (Fujishiro et al. 2022; Mealer et al. 2020; Moksnes et al. 2023), an essential trace element transported by ZIP8 and involved in glycosylation, the process by which branched sugar polymers are covalently attached to proteins and lipids (Pradeep et al. 2023). Glycosylation is dysregulated in schizophrenia (Pradeep et al. 2023) and disrupted by the schizophrenia-risk

TABLE 3 | Adjusted regression models of p.(Ala391Thr) using the observed data from the UKBB sample ($N = 355{,}069$).

	N	Beta/OR	Lower 95% CI	Upper 95% CI	Standard error	*p*	Tjur's R^2 (SNP)	R^2 (SNP)	R^2 adj. (SNP)	RMSE
Any psychotic experience	115,422	1	0.92	1.07	0.04	0.903	0.28% (0.00%)			0.211
A distressing psychotic experience	111,936	0.97	0.85	1.09	0.06	0.588	0.29% (0.00%)			0.129
Multiple psychotic experiences	112,958	0.97	0.85	1.09	0.05	0.269	0.25% (0.00%)			0.159
Delusions of persecution	110,712	1.05	0.86	1.28	0.1	0.611	0.25% (0.00%)			0.079
g score[a,b]	53,600	*−0.05*	*−0.07*	*−0.02*	0.01	1.91×10^{-5}*		17.18% (0.03%)	16.99% (0.03%)	0.909
Fluid intelligence[a,b]	53,600	*−0.05*	*−0.07*	*−0.04*	0.01	3.72×10^{-11}*		1.83% (0.04%)	1.73% (0.04%)	0.991
Years in education[b]	240,436	*−0.03*	*−0.05*	*0*	0.01	0.019*		2.82% (0.00%)	2.77% (0.00%)	2.158
Educational qualification: GCSEs	276,995	0.96	0.94	0.98	0.01	2.76×10^{-5}*	4.86% (0.01%)			0.476
Educational qualification: Degree	333,736	0.95	0.93	0.97	0.01	2.01×10^{-6}*	1.66% (0.01%)			0.467

Note: Adjusted for age at first interview, sex, principal components (PC) 1–10, and genotype batch. Beta coefficients indicated by italic. SNP is the difference between the variance explained by the full model and the variance explained by a covariate-only model and can thus be interpreted as the variance explained by the SNP alone. Tjur's R^2 is an alternative to other pseudo-R^2 values like Nagelkerke's R^2 or Cox-Snell R^2 and can be read like any other (pseudo-)R^2 value. RMSE is the square root of the variance of the residuals and indicates the absolute fit of the model to the data (difference between observed data to model's predicted values). It can be interpreted as the standard deviation of the unexplained variance, and is in the same units as the response variable. Lower values indicate better model fit.

Abbreviations: CI, confidence interval; OR, odds ratio.

[a]Additionally adjusted for age at first interview squared.

[b]Recalculated standard errors, confidence intervals, and *p*-values to estimate heteroscedasticity-robust variance.

*False Discovery Rate (FDR) adjusted *p*-value (adjusted for 9 tests) was statistically significant at an alpha level ≤0.05.

TABLE 4 | Adjusted regression models of rare variants using the observed data from the UKBB sample ($N = 355{,}069$).

	N	N rare variants (PTVs/ Missense)	Beta/OR	Lower 95% CI	Upper 95% CI	Standard error	p	Tjur's R^2 (RV)	R^2 (RV)	R^2 adj. (RV)	RMSE
Any psychotic experience	46,651	3/1	0.00	NA	13.29	98.34	0.93	0.21% 0.00%			0.21
A distressing psychotic experience	45,272	3/1	0.00	NA	$3.38 \times 10^{+03}$	160.41	0.955	0.20% (0.00%)			0.13
Multiple psychotic experiences	45,651	3/1	0.00	NA	$1.85 \times 10^{+10}$	255.26	0.97	0.16% (0.00%)			0.157
Delusions of persecution	44,749	3/1	0.00	NA	$6.34 \times 10^{+34}$	693.04	0.988	0.15% (0.00%)			0.076
g score[a]	22,476	0/2	*−1.67*	*−2.93*	*−0.42*	0.64	0.009		17.33% (0.03%)	17.27% (0.02%)	0.908
Fluid intelligence[a]	54,789	0/1	*−2.31*	*−4.25*	*−0.38*	0.99	0.019		1.80% (0.01%)	1.77% (0.01%)	0.988
Years in education	89,533	5/5	*−0.8*	*−2.16*	*0.55*	0.69	0.244		2.26% (0.00%)	2.25% (0.00%)	2.181
Educational qualification: GCSEs	104,192	6/4	0.24	0.05	0.89	0.72	0.044	4.34% (0.00%)			0.473
Educational qualification: Degree	126,333	6/6	0.63	0.14	2.13	0.67	0.489	1.49% (0.00%)			0.471

Note: Adjusted for the number of synonymous variants each person carries in SLC39A8, age at interview, sex, exome data PC1-10, and sequencing batch. Beta coefficients indicated by italic. NA indicates that there were not enough data to reliably calculate confidence intervals. (RV) is the difference between the variance explained by the full model and the variance explained by a covariate-only model and can thus be interpreted as the variance explained by the rare variant allele count alone. Tjur's R^2 is an alternative to other pseudo-R^2 values like Nagelkerke's R^2 or Cox-Snell R^2 and can be read like any other (pseudo-)R^2 value. RMSE is the square root of the variance of the residuals and indicates the absolute fit of the model to the data (difference between observed data to model's predicted values). It can be interpreted as the standard deviation of the unexplained variance, and is in the same units as the response variable. Lower values indicate better model fit.

Abbreviations: CI, confidence interval; OR, odds ratio; PTVs, Protein-truncating variants.

[a]Additionally adjusted for age at first interview squared.

allele at p.(Ala391Thr) (Mealer et al. 2022). Individuals with Congenital Disorders of Glycosylation can present with cognitive impairment (Pradeep et al. 2023), but there is no consistent linear association between lower blood manganese concentrations and poorer cognitive ability (Vollet et al. 2016). Finally, the schizophrenia-risk allele is also associated with impaired zinc uptake and transportation, and decreased cortical dendritic spine density (Li et al. 2022). Developmental synaptic pruning has been postulated as a risk mechanism for schizophrenia either through a loss of balance between synaptogenesis and elimination or abnormal activity-dependent plasticity (Kasai et al. 2010). Decreased dendritic spine density has been observed in individuals with schizophrenia (Radhakrishnan et al. 2021) while dendritic spine plasticity is thought to underlie the cognitive resilience of older adults who, despite having Alzheimer's pathophysiology, have not developed dementia (Boros et al. 2017). People with schizophrenia are 8.5 times more likely to develop dementia than age-matched controls (Liou et al. 2023) and 5.2 times more likely to have dementia listed as a cause of death than age- and sex-matched population controls (John et al. 2018).

Our findings support the idea that p.(Ala391Thr) plays a role in cognition, but we found no evidence to support our hypothesis that the schizophrenia-risk allele is associated with schizophrenia symptom severity. This may be due to limitations in our phenotype definitions and/or to power to detect small effects conferred by single variants (see below). An alternative explanation is that it may reflect postulated differences between the genetic architectures of liability to disorder and those of clinical symptom severity, particularly positive symptom severity. Consistent with that hypothesis, previous studies have generally not found a relationship between positive symptom severity and liability to schizophrenia, using either family history or polygenic risk scores as indices of liability (Owen et al. 2023), both of which capture considerably more variance in overall liability to the disorders than does any single variant. We also note that within people with schizophrenia, cognitive ability is only weakly correlated with the severity of clinical symptoms (Legge et al. 2021), and therefore alleles that influence cognition are not necessarily expected to influence symptom severity. Consistent with that, while there is robust evidence that polygenic scores based on alleles that influence cognitive ability in the general population are associated with cognition in schizophrenia, they are not associated with symptom severity (Legge et al. 2019).

p.(Ala391Thr) has been reported to be a highly pleiotropic variant and is associated, at genome-wide significance levels, with 24.16% (129/534) of the traits curated by Open Targets Genetics (Mountjoy et al. 2021); the schizophrenia-risk allele has been associated with lower HDL cholesterol, lower blood pressure, lower levels of apolipoprotein A1, calcium, aspartate aminotransferase, urate, gamma-glutamyl transferase, and serum albumin, as well as higher body mass index and body fat measures, but a lower risk of hypertension and cardiovascular disease. Notably, other than in schizophrenia, p.(Ala391Thr) has not been associated—in GWAS—with psychiatric (Bipolar Disorder, MDD), neurodevelopmental (Autism, Attention deficit hyperactivity disorder), or neurodegenerative (Alzheimer's Disease, Parkinson's Disease) disorders (see Buniello et al. (2019) and Table S1). In the most recent GWAS of alcohol use disorders, the schizophrenia-risk allele was associated with being a control (Zhou et al. 2023). In phenome-wide association studies (PheWAS) using the UK Biobank (UKBB), p.(Ala391Thr) has been associated with diseases of the esophagus, musculoskeletal conditions, metabolic and digestive biomarkers, blood pressure, and dietary and lifestyle factors including weight gain and drinking alcohol (Mitchell et al. 2019; Moksnes et al. 2023, http://www.nealelab.is/uk-biobank/). In a brain MRI PheWAS of the UKBB (Hermann et al. 2021), the schizophrenia-risk allele was associated with greater putamen gray matter volume, reduced cortical thickness, and reduced white matter integrity, MRI phenotypes which have been identified in participants with schizophrenia. Although pleiotropy has benefits for translational research, for example by cutting across current diagnostic categories, the diversity of phenotypes suggests that *SLC39A8* may not impact schizophrenia and/or cognition through the disruption of a single key biological process; rather, it may influence multiple processes, not necessarily all within the brain.

Studying p.(Ala391Thr) could improve our understanding of the cognitive impairments in people with schizophrenia and aid in the prediction of treatment response. Within studies of cognitive remediation therapy (CRT) in schizophrenia, although on average patients with poorer premorbid IQ and fewer years in education show greater improvement after CRT, some studies have shown the opposite (Vita et al. 2021) and there is emerging literature examining whether genetic variants can account for this differential improvement (Penadés et al. 2020). For example, one study examined *SLC1A2*, a high-affinity glutamate transporter that encodes EAAT2, and found that the minor allele at NC_000011.10:g.35419429T>G (rs4354668), which has previously been associated with lower EAAT2 expression and poorer cognition in healthy controls and patients with schizophrenia, was associated with poorer improvement after CRT (Spangaro et al. 2018). p.(Ala391Thr) could be a candidate for future stratification studies.

6 | Limitations

This is the first study to test the relationship between p.(Ala391Thr) and schizophrenia-related phenotypes in participants with the disorder. However, our cohorts of schizophrenia participants were small, and it may be that larger sample sizes are needed to detect the small effects attributable to a single variant. Given a sample size of $N = 1232$, an alpha level of 0.05, and an $R^2 = 0.03$ (the proportion of variance explained by p.(Ala391Thr) for the *g* score in the UKBB), our power was 16.42%. We also note that there was more missing data in the Cardiff SibPairs sample compared to the other Cardiff schizophrenia samples. Although missing data imputation performed well, using imputed datasets, while recommended over complete-case analysis, may not reflect the real-world variability in these variables because imputed data were within the range of observed data. In addition, our analysis of rare variants in *SLC39A8* may not be representative because of the small number of rare variant carriers driving the associations. Gene-based burden tests usually apply a cut off to exclude genes with a low a number of carriers (Chen et al. 2023), which, in our sample, would have excluded *SLC39A8* from a pipeline for exome-wide

gene-based burden tests. Our rare variant results should therefore be used to inform future studies rather than be interpreted in isolation. It is also possible that our findings will not generalize to participants of non-European ancestry. Previous work has described p.(Ala391Thr) as having an almost negligible MAF in non-European populations (Costas 2018). In our schizophrenia cohorts, only Europeans carried two copies of the schizophrenia risk allele, but three other ancestry groups (African American/Afro-Caribbean, Near Eastern, and Central/South Asian) contained participants with one copy, so we chose to include them in our analysis. Nevertheless, less than 2% of the schizophrenia cohorts were of non-European ancestry, and our sample from the UKBB was comprised solely of people with European ancestry.

Another caveat is that it is unclear whether psychotic experiences measured in the UKBB are a good proxy for psychotic experiences experienced by people with schizophrenia. Although there is a shared genetic liability between psychotic experiences and schizophrenia, the genetic correlation is weak (rg = 0.21; Legge et al. 2019), and psychotic experiences were more strongly correlated with ADHD (rg = 0.24), autism spectrum disorder (rg = 0.39), and major depressive disorder (rg = 0.46) than schizophrenia, suggesting that these phenotypes capture general psychopathology rather than schizophrenia-specific psychosis. There is also the possibility of measurement error in the psychotic experiences phenotypes; firstly, because they are measured retrospectively by self-report, and, secondly, because of biases in those who attempt to complete the online follow-up mental health questionnaire. Legge et al. (2019) found that participants who completed the mental health questionnaire had significantly higher intelligence and lower schizophrenia polygenic risk scores, adding to existing evidence of a "healthy volunteer" selection bias in the UKBB (Fry et al. 2017; Legge et al. 2024) which can affect the results of GWAS (Schoeler et al. 2023). In our data, the schizophrenia-risk allele was associated with *not* attempting the mental health questionnaire. This suggests that the psychotic experiences phenotypes in the UKBB are not completed by participants with a higher genetic load for schizophrenia, and our analyses may be affected by collider bias.

Finally, we note that educational attainment is strongly influenced by environmental factors, such as an individual's personality and the social system in which they were educated, as well as noncognitive genetic variants (Demange et al. 2021). Indeed, sampling differences in geography and time moderate genetic effects on educational attainment (Tropf et al. 2017). As such, educational attainment should be considered a related but distinct phenotype of, or a distant proxy for, cognitive ability.

7 | Conclusions

The schizophrenia-risk allele at p.(Ala391Thr) is associated with poorer cognitive ability in a sample of the general population with European ancestry. Larger and more ancestrally diverse studies of participants with schizophrenia are required to determine whether p.(Ala391Thr) and ZIP8 can aid genomic stratification of treatment response for cognitive impairment in those with the disorder.

Author Contributions

S.E.S. designed and conducted the analysis and drafted the manuscript. S.E.L. conducted the confirmatory factor analysis. E.F. called the rare variants. A.F.P. conducted the genetic ancestry analysis. A.J.L. scored the MATRICS Consensus Cognitive Battery. V.E.-P., E.F., J.H., P.H., S.E.L., M.C.O., M.J.O., A.F.P., G.W., L.W., and J.T.R.W. helped design the analysis and interpret the results. All authors revised and approved the final manuscript.

Acknowledgments

We acknowledge the support of the Supercomputing Wales project, which is part funded by the European Regional Development Fund (ERDF) via Welsh Government.

Ethics Statement

CardiffCOGS was approved by the South East Wales Research Ethics Committee Panel (reference number: 07/WSE03/110) and received HRA approval. All participants provided written informed consent. Multicentre and Local Research Ethics Committee approval was obtained for Cardiff F-Series, and all participants gave written informed consent to participate. For Cardiff SibPairs, written consent was obtained following local ethical approval guidelines. Ethical approval for the curation and development of DRAGON-Data was obtained from Cardiff University's School of Medicine Research Ethics Committee (Ref: 19/72). This research has been conducted using the UK Biobank Resource under Application Number 13310. The scientific protocol of the UK Biobank (https://www.ukbiobank.ac.uk) was reviewed and approved by the North West Multi-centre Ethics Committee, and this work uses data provided by patients and collected by the NHS as part of their care and support.

Conflicts of Interest

J.H., M.C.O., M.J.O., L.W., and J.T.R.W. were investigators on the grant from Takeda Pharmaceuticals Ltd. to Cardiff University. S.E.S. and G.W. were employed on this grant. J.H. is the Chief Medical Officer for MeOmics Precision Medicine Ltd., and M.J.O. is the recipient of a grant from Akrivia Health, but neither of these companies was involved in this study.

Data Availability Statement

To comply with the ethical and regulatory framework under which the Cardiff schizophrenia samples were obtained, access to individual-level data requires a collaboration agreement with Cardiff University; requests to access these datasets should be directed to J.T.R.W. (waltersjt@cardiff.ac.uk) and M.J.O. (owenmj@cardiff.ac.uk). UK Biobank data is available by application to the UK Biobank (www.ukbiobank.ac.uk) https://zenodo.org/records/10027873.

References

Adams, M. J., W. D. Hill, D. M. Howard, et al. 2020. "Factors Associated With Sharing e-Mail Information and Mental Health Survey Participation in Large Population Cohorts." *International Journal of Epidemiology* 49, no. 2: 410–421. https://doi.org/10.1093/ije/dyz134.

Boros, B. D., K. M. Greathouse, E. G. Gentry, et al. 2017. "Dendritic Spines Provide Cognitive Resilience Against Alzheimer's Disease." *Annals of Neurology* 82, no. 4: 602–614. https://doi.org/10.1002/ana.25049.

Bulbul, O., G. Filoglu, T. Zorlu, et al. 2016. "Inference of Biogeographical Ancestry Across Central Regions of Eurasia." *International Journal of Legal Medicine* 130, no. 1: 73–79.

Buniello, A., J. A. L. MacArthur, M. Cerezo, et al. 2019. "The NHGRI-EBI GWAS Catalog of Published Genome-Wide Association Studies, Targeted Arrays and Summary Statistics 2019." *Nucleic Acids Research* 47, no. D1: D1005–D1012.

Bycroft, C., C. Freeman, D. Petkova, et al. 2018. "The UK Biobank Resource With Deep Phenotyping and Genomic Data." *Nature* 562, no. 7726: 203–209. https://doi.org/10.1038/s41586-018-0579-z.

Chen, C.-Y., R. Tian, T. Ge, et al. 2023. "The Impact of Rare Protein Coding Genetic Variation on Adult Cognitive Function." *Nature Genetics* 55, no. 6: 927–938. https://doi.org/10.1038/s41588-023-01398-8.

Cornelis, M. C., Y. Wang, T. Holland, P. Agarwal, S. Weintraub, and M. C. Morris. 2019. "Age and Cognitive Decline in the UK Biobank." *PLoS One* 14, no. 3: e0213948. https://doi.org/10.1371/journal.pone.0213948.

Costas, J. 2018. "The Highly Pleiotropic Gene SLC39A8 as an Opportunity to Gain Insight Into the Molecular Pathogenesis of Schizophrenia." *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics* 177, no. 2: 274–283. https://doi.org/10.1002/ajmg.b.32545.

Davis, K. A. S., J. R. I. Coleman, M. Adams, et al. 2020. "Mental Health in UK Biobank - Development, Implementation and Results From an Online Questionnaire Completed by 157 366 Participants: A Reanalysis." *BJPsych Open* 6, no. 2: e18. https://doi.org/10.1192/bjo.2019.100.

Demange, P. A., M. Malanchini, T. T. Mallard, et al. 2021. "Investigating the Genetic Architecture of Noncognitive Skills Using GWAS-By-Subtraction." *Nature Genetics* 53, no. 1: 35–44. https://doi.org/10.1038/s41588-020-00754-2.

Dennison, C. A., S. E. Legge, L. Hubbard, et al. 2021. "Risk Factors, Clinical Features, and Polygenic Risk Scores in Schizophrenia and Schizoaffective Disorder Depressive-Type." *Schizophrenia Bulletin* 47, no. 5: 1375–1384. https://doi.org/10.1093/schbul/sbab036.

Fawns-Ritchie, C., and I. J. Deary. 2020. "Reliability and Validity of the UK Biobank Cognitive Tests." *PLoS One* 15, no. 4: e0231627. https://doi.org/10.1371/journal.pone.0231627.

Fenner, E., P. Holmans, M. C. O'Donovan, M. J. Owen, J. T. Walters, and E. Rees. 2023. "Rare Coding Variants in Schizophrenia-Associated Genes Affect Generalised Cognition in the UK Biobank." medRxiv. 2023.2008. 2014.23294074.

Fry, A., T. J. Littlejohns, C. Sudlow, et al. 2017. "Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population." *American Journal of Epidemiology* 186, no. 9: 1026–1034. https://doi.org/10.1093/aje/kwx246.

Fujishiro, H., S. Miyamoto, D. Sumi, T. Kambe, and S. Himeno. 2022. "Effects of Individual Amino Acid Mutations of Zinc Transporter ZIP8 on Manganese- and Cadmium-Transporting Activity." *Biochemical and Biophysical Research Communications* 616: 26–32. https://doi.org/10.1016/j.bbrc.2022.05.068.

Hail Team. n.d. "Hail." https://github.com/hail-is/hail.

Hermann, E. R., E. Chambers, D. N. Davis, M. R. Montgomery, D. Lin, and W. Chowanadisai. 2021. "Brain Magnetic Resonance Imaging Phenome-Wide Association Study With Metal Transporter Gene SLC39A8." *Frontiers in Genetics* 12: 647946. https://doi.org/10.3389/fgene.2021.647946.

Homman, L. E., S. E. Smart, F. O'Neill, and J. H. MacCabe. 2021. "Attrition in Longitudinal Studies Among Patients With Schizophrenia and Other Psychoses; Findings From the STRATA Collaboration." *Psychiatry Research* 305: 114211. https://doi.org/10.1016/j.psychres.2021.114211.

Huddart, R., A. E. Fohner, M. Whirl-Carrillo, et al. 2019. "Standardized Biogeographic Grouping System for Annotating Populations in Pharmacogenetic Research." *Clinical Pharmacology and Therapeutics* 105, no. 5: 1256–1262. https://doi.org/10.1002/cpt.1322.

Ioannidis, N. M., J. H. Rothstein, V. Pejaver, et al. 2016. "REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants." *American Journal of Human Genetics* 99, no. 4: 877–885.

John, A., J. McGregor, I. Jones, et al. 2018. "Premature Mortality Among People With Severe Mental Illness—New Evidence From Linked Primary Care Data." *Schizophrenia Research* 199: 154–162. https://doi.org/10.1016/j.schres.2018.04.009.

Karczewski, K. J., L. C. Francioli, G. Tiao, et al. 2020. "The Mutational Constraint Spectrum Quantified From Variation in 141,456 Humans." *Nature* 581, no. 7809: 434–443. https://doi.org/10.1038/s41586-020-2308-7.

Kasai, H., M. Fukuda, S. Watanabe, A. Hayashi-Takagi, and J. Noguchi. 2010. "Structural Dynamics of Dendritic Spines in Memory and Cognition." *Trends in Neurosciences* 33, no. 3: 121–129. https://doi.org/10.1016/j.tins.2010.01.001.

Lamers, F., A. W. Hoogendoorn, J. H. Smit, et al. 2012. "Sociodemographic and Psychiatric Determinants of Attrition in The Netherlands Study of Depression and Anxiety (NESDA)." *Comprehensive Psychiatry* 53, no. 1: 63–70. https://doi.org/10.1016/j.comppsych.2011.01.011.

Legge, S. E., A. G. Cardno, J. Allardyce, et al. 2021. "Associations Between Schizophrenia Polygenic Liability, Symptom Dimensions, and Cognitive Ability in Schizophrenia." *JAMA Psychiatry* 78, no. 10: 1143–1151. https://doi.org/10.1001/jamapsychiatry.2021.1961.

Legge, S. E., C. A. Dennison, A. F. Pardiñas, et al. 2020. "Clinical Indicators of Treatment-Resistant Psychosis." *British Journal of Psychiatry* 216, no. 5: 259–266. https://doi.org/10.1192/bjp.2019.120.

Legge, S. E., H. J. Jones, K. M. Kendall, et al. 2019. "Association of Genetic Liability to Psychotic Experiences With Neuropsychotic Disorders and Traits." *JAMA Psychiatry* 76, no. 12: 1256–1265. https://doi.org/10.1001/jamapsychiatry.2019.2508.

Legge, S. E., A. F. Pardiñas, G. Woolway, et al. 2024. "Genetic and Phenotypic Features of Schizophrenia in the UK Biobank." *JAMA Psychiatry* 81, no. 7: 681–690. https://doi.org/10.1001/jamapsychiatry.2024.0200.

Leonenko, G., E. Baker, J. Stevenson-Hoare, et al. 2021. "Identifying Individuals With High Risk of Alzheimer's Disease Using Polygenic Risk Scores." *Nature Communications* 12, no. 1: 4506. https://doi.org/10.1038/s41467-021-24082-z.

Li, M., D. D. Wu, Y. G. Yao, et al. 2016. "Recent Positive Selection Drives the Expansion of a Schizophrenia Risk Nonsynonymous Variant at SLC39A8 in Europeans." *Schizophrenia Bulletin* 42, no. 1: 178–190. https://doi.org/10.1093/schbul/sbv070.

Li, S., C. Ma, Y. Li, et al. 2022. "The Schizophrenia-Associated Missense Variant rs13107325 Regulates Dendritic Spine Density." *Translational Psychiatry* 12, no. 1: 361. https://doi.org/10.1038/s41398-022-02137-z.

Liou, Y.-J., S.-J. Tsai, Y.-M. Bai, T.-J. Chen, and M.-H. Chen. 2023. "Dementia Risk in Middle-Aged Patients With Schizophrenia, Bipolar Disorder, and Major Depressive Disorder: A Cohort Study of 84,824 Subjects." *European Archives of Psychiatry and Clinical Neuroscience* 273, no. 1: 219–227. https://doi.org/10.1007/s00406-022-01389-6.

Little, J., J. P. T. Higgins, J. P. A. Ioannidis, et al. 2009. "STrengthening the REporting of Genetic Association Studies (STREGA)—An Extension of the STROBE Statement." *Genetic Epidemiology* 33, no. 7: 581–598. https://doi.org/10.1002/gepi.20410.

Lynham, A. J., S. Knott, J. F. G. Underwood, et al. 2023. "DRAGON-Data: A Platform and Protocol for Integrating Genomic and Phenotypic Data Across Large Psychiatric Cohorts." *BJPsych Open* 9, no. 2: e32. https://doi.org/10.1192/bjo.2022.636.

Mealer, R. G., B. G. Jenkins, C. Y. Chen, et al. 2020. "The Schizophrenia Risk Locus in SLC39A8 Alters Brain Metal Transport and Plasma Glycosylation." *Scientific Reports* 10, no. 1: 13162. https://doi.org/10.1038/s41598-020-70108-9.

Mealer, R. G., S. E. Williams, M. Noel, et al. 2022. "The Schizophrenia-Associated Variant in SLC39A8 Alters Protein Glycosylation in the Mouse Brain." *Molecular Psychiatry* 27, no. 3: 1405–1415. https://doi.org/10.1038/s41380-022-01490-1.

Mignogna, G., C. E. Carey, R. Wedow, et al. 2023. "Patterns of Item Nonresponse Behaviour to Survey Questionnaires Are Systematic and Associated With Genetic Loci." *Nature Human Behaviour* 7: 1371–1387. https://doi.org/10.1038/s41562-023-01632-7.

Mitchell, R., B. Elsworth, R. Mitchell, et al. 2019. "MRC IEU UK Biobank GWAS Pipeline Version 2."

Moksnes, M. R., A. F. Hansen, B. N. Wolford, et al. 2023. "New Insights Into the Genetic Etiology of 57 Essential and Non-Essential Trace Elements in Humans." medRxiv, 2023.2004.2025.23289097. https://doi.org/10.1101/2023.04.25.23289097.

Mountjoy, E., E. M. Schmidt, M. Carmona, et al. 2021. "An Open Approach to Systematically Prioritize Causal Variants and Genes at All Published Human GWAS Trait-Associated Loci." *Nature Genetics* 53, no. 11: 1527–1533. https://doi.org/10.1038/s41588-021-00945-5.

Nakazawa, K., and K. Sapkota. 2020. "The Origin of NMDA Receptor Hypofunction in Schizophrenia." *Pharmacology & Therapeutics* 205: 107426. https://doi.org/10.1016/j.pharmthera.2019.107426.

Nebert, D. W., and Z. Liu. 2019. "SLC39A8 Gene Encoding a Metal Ion Transporter: Discovery and Bench to Bedside." *Human Genomics* 13, no. Suppl 1: 51. https://doi.org/10.1186/s40246-019-0233-3.

Ng, E., P. M. Lind, C. Lindgren, et al. 2015. "Genome-Wide Association Study of Toxic Metals and Trace Elements Reveals Novel Associations." *Human Molecular Genetics* 24, no. 16: 4739–4745. https://doi.org/10.1093/hmg/ddv190.

Okbay, A., Y. Wu, N. Wang, et al. 2022. "Polygenic Prediction of Educational Attainment Within and Between Families From Genome-Wide Association Analyses in 3 Million Individuals." *Nature Genetics* 54, no. 4: 437–449. https://doi.org/10.1038/s41588-022-01016-z.

Owen, M. J., S. E. Legge, E. Rees, J. T. R. Walters, and M. C. O'Donovan. 2023. "Genomic Findings in Schizophrenia and Their Implications." *Molecular Psychiatry* 28, no. 9: 3638–3647. https://doi.org/10.1038/s41380-023-02293-8.

Pejaver, V., A. B. Byrne, B. J. Feng, et al. 2022. "Calibration of Computational Tools for Missense Variant Pathogenicity Classification and ClinGen Recommendations for PP3/BP4 Criteria." *American Journal of Human Genetics* 109, no. 12: 2163–2177. https://doi.org/10.1016/j.ajhg.2022.10.013.

Penadés, R., M. Bosia, R. Catalán, et al. 2020. "The Role of Genetics in Cognitive Remediation in Schizophrenia: A Systematic Review." *Schizophrenia Research: Cognition* 19: 100146. https://doi.org/10.1016/j.scog.2019.100146.

Pradeep, P., H. Kang, and B. Lee. 2023. "Glycosylation and Behavioral Symptoms in Neurological Disorders." *Translational Psychiatry* 13, no. 1: 154. https://doi.org/10.1038/s41398-023-02446-x.

Radhakrishnan, R., P. D. Skosnik, M. Ranganathan, et al. 2021. "In Vivo Evidence of Lower Synaptic Vesicle Density in Schizophrenia." *Molecular Psychiatry* 26, no. 12: 7690–7698. https://doi.org/10.1038/s41380-021-01184-0.

Rubin, D. B. 2004. *Multiple Imputation for Nonresponse in Surveys*. Vol. 81. John Wiley & Sons.

Savage, J. E., P. R. Jansen, S. Stringer, et al. 2018. "Genome-Wide Association Meta-Analysis in 269,867 Individuals Identifies New Genetic and Functional Links to Intelligence." *Nature Genetics* 50, no. 7: 912–919. https://doi.org/10.1038/s41588-018-0152-6.

Schielzeth, H. 2010. "Simple Means to Improve the Interpretability of Regression Coefficients." *Methods in Ecology and Evolution* 1, no. 2: 103–113.

Schoeler, T., D. Speed, E. Porcu, N. Pirastu, J.-B. Pingault, and Z. Kutalik. 2023. "Participation Bias in the UK Biobank Distorts Genetic Associations and Downstream Analyses." *Nature Human Behaviour* 7, no. 7: 1216–1227. https://doi.org/10.1038/s41562-023-01579-9.

Spangaro, M., M. Bosia, M. Bechi, et al. 2018. "Neurobiology of Cognitive Remediation in Schizophrenia: Effects of EAAT2 Polymorphism." *Schizophrenia Research* 202: 106–110. https://doi.org/10.1016/j.schres.2018.06.059.

Sudlow, C., J. Gallacher, N. Allen, et al. 2015. "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age." *PLoS Medicine* 12, no. 3: e1001779. https://doi.org/10.1371/journal.pmed.1001779.

Tropf, F. C., S. H. Lee, R. M. Verweij, et al. 2017. "Hidden Heritability due to Heterogeneity Across Seven Populations." *Nature Human Behaviour* 1, no. 10: 757–765. https://doi.org/10.1038/s41562-017-0195-1.

Trubetskoy, V., A. F. Pardiñas, T. Qi, et al. 2022. "Mapping Genomic Loci Implicates Genes and Synaptic Biology in Schizophrenia." *Nature* 604: 502–508. https://doi.org/10.1038/s41586-022-04434-5.

Tseng, W. C., V. Reinhart, T. A. Lanz, et al. 2021. "Schizophrenia-Associated SLC39A8 Polymorphism Is a Loss-of-Function Allele Altering Glutamate Receptor and Innate Immune Signaling." *Translational Psychiatry* 11, no. 1: 136. https://doi.org/10.1038/s41398-021-01262-5.

Tyrrell, J., J. Zheng, R. Beaumont, et al. 2021. "Genetic Predictors of Participation in Optional Components of UK Biobank." *Nature Communications* 12, no. 1: 886. https://doi.org/10.1038/s41467-021-21073-y.

Vita, A., S. Barlati, A. Ceraso, et al. 2021. "Effectiveness, Core Elements, and Moderators of Response of Cognitive Remediation for Schizophrenia: A Systematic Review and Meta-Analysis of Randomized Clinical Trials." *JAMA Psychiatry* 78, no. 8: 848–858. https://doi.org/10.1001/jamapsychiatry.2021.0620.

Vollet, K., E. N. Haynes, and K. N. Dietrich. 2016. "Manganese Exposure and Cognition Across the Lifespan: Contemporary Review and Argument for Biphasic Dose–Response Health Effects." *Current Environmental Health Reports* 3, no. 4: 392–404. https://doi.org/10.1007/s40572-016-0108-x.

von Hippel, P. T. 2015. "The Heterogeneity Statistic I2 Can Be Biased in Small Meta-Analyses." *BMC Medical Research Methodology* 15, no. 1: 35. https://doi.org/10.1186/s12874-015-0024-z.

Williams, N. M., E. K. Green, S. Macgregor, et al. 2006. "Variation at the DAOA/G30 Locus Influences Susceptibility to Major Mood Episodes but Not Psychosis in Schizophrenia and Bipolar Disorder." *Archives of General Psychiatry* 63, no. 4: 366–373. https://doi.org/10.1001/archpsyc.63.4.366.

Williams, N. M., M. I. Rees, P. Holmans, et al. 1999. "A Two-Stage Genome Scan for Schizophrenia Susceptibility Genes in 196 Affected Sibling Pairs." *Human Molecular Genetics* 8, no. 9: 1729–1739. https://doi.org/10.1093/hmg/8.9.1729.

Zhou, H., R. L. Kember, J. D. Deak, et al. 2023. "Multi-Ancestry Study of the Genetics of Problematic Alcohol Use in Over 1 Million Individuals." *Nature Medicine* 29, no. 12: 3184–3192. https://doi.org/10.1038/s41591-023-02653-5.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.