

1 SLC39A8.p.(Ala391Thr) is associated with poorer cognitive ability: a
2 cross-sectional study of schizophrenia and the general UK population

3 Authors

4 Sophie E Smart^{1*}

5 Sophie E. Legge¹

6 Eilidh Fenner¹

7 Antonio F. Pardiñas¹

8 Grace Woolway¹

9 Amy J. Lynham¹

10 Valentina Escott-Price¹

11 Jeremy Hall²

12 Lawrence Wilkinson¹

13 Peter Holmans¹

14 Michael C. O'Donovan¹

15 Michael J. Owen^{1,2}

16 James T.R. Walters¹

17

18 ¹ Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine
19 and Clinical Neurosciences, Cardiff University, Cardiff, UK

20 ² Neuroscience and Mental Health Innovation Institute, Division of Psychological Medicine
21 and Clinical Neurosciences, Cardiff University, Cardiff, UK

22 * Corresponding author: SmartS1@cardiff.ac.uk

23 Abstract

24 The missense SNP NC_000004.12:g.102267552C>T (SLC39A8.p.(Ala391Thr), rs13107325)
25 in *SLC39A8*, which encodes a zinc transporter, has been linked to schizophrenia and is the
26 likely causal variant for one of the genome-wide association loci associated with the disorder.

27

28 We tested whether the schizophrenia-risk allele at p.(Ala391Thr) was associated with
29 schizophrenia-related phenotypes, including positive, negative, and disorganised symptoms,
30 cognitive ability, educational attainment, and age of psychosis onset, within three
31 schizophrenia cohorts (combined N=1,232) and, with equivalent phenotypes, in a sample of
32 population controls (UK Biobank, N=355,069). We used regression analyses controlling for
33 age, sex, and population stratification.

34

35 Within the schizophrenia cohorts, after correction for multiple testing, p.(Ala391Thr) was not
36 significantly associated with any schizophrenia-related phenotypes. In the unaffected
37 participants from the UK Biobank, the schizophrenia-risk allele at p.(Ala391Thr) was
38 associated with significantly poorer cognitive ability and fluid intelligence, a lower
39 probability of obtaining GCSEs or a degree-level qualification, and fewer years in education.

40 There was no association between p.(Ala391Thr) and self-reported psychotic experiences in
41 this cohort.

42

43 The schizophrenia-risk allele was associated with poorer cognitive ability, but not psychotic
44 experiences, in a volunteer sample drawn from of the general population. To determine
45 whether p.(Ala391Thr) is associated with cognitive phenotypes in people with schizophrenia,
46 and to understand the role of p.(Ala391Thr) in the aetiology of cognitive impairment in
47 schizophrenia, larger independent samples are required.

48

49 Keywords: ZIP8, rs13107325, metal ion transporters, psychosis, rare variants

50 Introduction

51

52 The gene *SLC39A8* encodes metal cation symporter ZIP8, which enables the movement of at
53 least five trace elements (manganese, zinc, iron, selenium, and cobalt) from outside the cell
54 into the cytosol. *SLC39* genes are highly conserved and are critical for fundamental life
55 processes throughout the lifespan. *SLC39A8* contains a missense single nucleotide
56 polymorphism (SNP), NC_000004.12:g.102267552C>T (*SLC39A8*.p.(Ala391Thr),
57 rs13107325), in which alternative alleles encode alanine (Ala; C allele) or threonine (Thr; T
58 allele) on chromosome 4 (1). In vitro, the minor allele (T) leads to reduced uptake of
59 manganese and cadmium (2), the latter a toxic environmental pollutant, and decreased
60 synaptic uptake of zinc (3).

61

62 p.(Ala391Thr) is associated with an increased risk of schizophrenia, a heterogeneous disorder
63 whose presentation includes delusions, hallucinations, behavioural disturbance, social
64 withdrawal, and cognitive impairment. In the most recent genome-wide association study
65 (GWAS), of 69,369 schizophrenia cases and 236,642 controls, the minor allele was
66 associated with increased the risk of schizophrenia (odds ratio=1.17, $P=2.9\times 10^{-21}$) (4).
67 Furthermore, fine-mapping suggested that p.(Ala391Thr) was likely to be the causal variant
68 underpinning the association with the genomic region containing *SLC39A8*, the SNP having a
69 posterior probability of being causally associated with schizophrenia of 99.20%. Only nine
70 out of 255 loci were fine-mapped to one causal variant (4), only two of these were missense
71 variants, and only p.(Ala391Thr) had a Combined Annotation-Dependent Depletion (CADD)
72 score of 22 making it one of the 1% most deleterious variants in the genome.

73

74 In non-psychiatric populations, p.(Ala391Thr) has not been associated with reported
75 psychotic experiences ($Z = 0.99$, $P = 0.889$) (5), but the schizophrenia risk allele has been
76 associated with lower intelligence ($Z = -9.49$, $P = 2.23E-21$) (6), fewer years in education
77 ($\text{Beta} = -0.02$, $P = 1.08E-13$) (7), and, using a GWAS-by-subtraction design, with the
78 cognitive aspect of educational attainment (NonCog, $Z = 2.78$, $P = 0.005$; Cog, $Z = -9.87$, $P =$
79 $5.59E-23$) (8). To date, no studies have examined the relationship between p.(Ala391Thr) and
80 the clinical presentation or cognitive ability of people with schizophrenia.

81
82 Our primary aim was to test whether schizophrenia-risk alleles at p.(Ala391Thr) are
83 associated with phenotypes that capture the clinical presentation of schizophrenia. We
84 hypothesised that the schizophrenia-risk allele (T) at p.(Ala391Thr) would be associated with
85 a more severe clinical presentation, poorer cognitive ability and educational attainment, and
86 younger age of onset. We expanded the analysis to test for similar phenotypic associations in
87 a large sample of the general population (UK Biobank unaffected controls). Given that
88 *SLC39A8* is also constrained for loss of function and missense variants (Genome Aggregation
89 Database [gnomAD] v2.1.1; 9), we used the UK Biobank to tested for phenotypic
90 associations with rare protein-truncating and deleterious missense variants within *SLC39A8*.

91 [Methods](#)

92
93 This study followed the guidelines outlined in the STrengthening the REporting of Genetic
94 Association Studies (STREGA) (10), an extension of the STROBE Statement (checklist in
95 the Supplementary Material).

96
97 We analysed data from three cohorts recruited at Cardiff University and data from the UK
98 Biobank. Full details on how samples were ascertained and how phenotypes were measured,

99 calculated, and standardised, can be found in the Supplementary Material. A brief overview
100 is provided below.

101

102 [Cardiff Schizophrenia Samples](#)

103

104 [Participants](#)

105 Participants were recruited into three cross-sectional studies, all of which have been
106 previously described: CardiffCOGS (11), Cardiff F-series (12) and Cardiff Affected-Sibs (13)
107 samples. The Cardiff Affected-Sibs sample includes a single affected individual from
108 families with two or more siblings diagnosed with schizophrenia or schizoaffective disorder.

109 All participants were recruited from in-patient, voluntary sector, and community mental
110 health services across the UK and underwent a clinical research interview based on the
111 Schedules for Clinical Assessment in Neuropsychiatry (SCAN), with further information
112 available from clinical records. For these analyses, we only retained participants who met
113 DSM-IV or ICD-10 criteria for a diagnosis of schizophrenia or schizoaffective disorder,
114 depressed type.

115

116 [Phenotypes](#)

117 The Scale for the Assessment of Positive Symptoms (SAPS) and the Scale for the
118 Assessment of Negative Symptoms (SANS) were scored on a lifetime worst basis using
119 information from the SCAN interview and lifetime psychiatric clinical records. These records
120 were also used to ascertain age at psychosis onset. The Measurement and Treatment Research
121 to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery was used
122 as a measure of current cognitive ability in the CardiffCOGS sample. As described in Legge,
123 Cardno (14), a confirmatory factor analysis (CFA) framework was used to estimate

124 phenotype-derived factor scores from the symptom and cognitive ratings, which we refer to
125 as symptom dimensions. Using all three samples, the best model had three symptom
126 dimensions: positive symptoms (SAPS global hallucinations and SAPS global delusions),
127 negative symptoms of diminished expressivity (SANS global affective flattening and SANS
128 global alogia), and disorganised symptoms (SAPS global positive formal thought disorder
129 and SANS inappropriate affect). As CardiffCOGS was the only sample to have data on
130 cognitive ability, a second CFA framework, using only this sample, was used to estimate
131 phenotype-derived factor scores from the MATRICS domain scores as well as the symptom
132 ratings. The best model had five dimensions: positive symptoms (as above), negative
133 symptoms of diminished expressivity (as above), disorganised symptoms (as above), negative
134 symptoms of motivation and pleasure (SANS global anhedonia/asociality and SANS global
135 avolition/apathy), and cognitive ability (all MATRICS domains apart from social cognition).
136
137 As well as the cognitive domain, which captures current cognitive ability, premorbid IQ was
138 assessed in the CardiffCOGS sample using the National Adult Reading Test (NART).
139 Educational attainment was measured across all three samples using years in education and
140 highest educational qualification (General Certificate of Secondary Education [GCSE]/no
141 GCSE, note GCSEs are taken by most UK pupils upon the completion of compulsory
142 education [from 1972-2013 this was at age 16], and degree/no degree).

143
144 **Genotypes**

145 Genome-wide SNP data for the three samples were curated and harmonised as part of
146 DRAGON-Data (15) (see Supplementary Material). Genotypes for p.(Ala391Thr)
147 (INFO=0.996) were extracted for each participant. Principal components were calculated
148 using pruned SNPs.

149

150 Genetic ancestry probabilities were calculated using Ancestry Informative Markers (AIMs)
151 derived from the Allen Ancient DNA Resource reference panel, linear discriminant analysis
152 (LDA), and biogeographical categories defined by Huddart, Fohner (16). AIMs are genetic
153 variants with highly divergent allele frequencies across biogeographical genetic ancestries,
154 which can capture genetic associations between an individual and a particular
155 (sub)continental population (17). Ancestry groups were determined by assigning individuals
156 to their most probable biogeographical category as inferred by LDA (determined using
157 Youden's index as optimality criterion), with individuals not meeting a probability threshold
158 for any category being assigned to an “admixed” group. The schizophrenia risk allele is
159 thought to have been under recent positive selection in Europeans (18); in Phase 3 of the
160 1000 Genomes Project, the schizophrenia allele is present in around 8% of European samples
161 but is reported to be almost absent from other populations (1). In our data, only Europeans
162 carried two copies of the schizophrenia risk allele. However, to increase our sample size and
163 improve the generalisability of our findings, we restricted our sample to participants from
164 ancestries where the schizophrenia-risk allele was present in at least one individual: African
165 American/Afro-Caribbean (0.48%), Central/South Asian (1.06%), European (98.21%), and
166 Middle Eastern/North African (0.24%) (note the Middle Eastern/North African ancestry is
167 referred to as Near Eastern in Huddart, Fohner (16)).

168

169 [UK Biobank](#)

170

171 [Participants](#)

172 Participants were from the UK Biobank (UKBB), a large-scale biomedical database of
173 individuals aged between 40-69 who were recruited from across the UK (19). Participants

174 underwent extensive phenotyping. All UKBB field IDs used in this study are reported in the
175 Supplementary Material. For these analyses, we removed participants with a psychotic
176 spectrum disorder (ICD-10 F20-F29, see Supplementary Material).

177

178 Phenotypes

179 An abridged version of the Composite International Diagnostic Interview psychosis module
180 (lifetime version) was used in the online follow-up mental health questionnaire. As in Legge,
181 Jones (5), we derived three overlapping binary variables: (i) any psychotic experience defined
182 as a positive response to any of the four symptom questions; (ii) a distressing psychotic
183 experience, defined as any psychotic experience that was rated as “a bit,” “quite,” or “very”
184 distressing; and (iii) multiple occurrences of psychotic experiences, defined as any psychotic
185 experience that occurred on more than one occasion. The comparator group for these three
186 variables was comprised of individuals who provided a negative response to all four
187 psychotic experience symptom questions (see Supplementary Material). We also looked at
188 delusions of persecution alone (UKBB Field 20468: “Ever believed in an un-real conspiracy
189 against self”) as previous work has suggested that this phenotype in the UKBB is particularly
190 enriched for genetic liability for schizophrenia (5).

191

192 Cognitive ability was measured using a general intelligence factor, g , which is considered a
193 reliable measure of cognitive ability. As in Fawns-Ritchie and Deary (20), g was calculated
194 using principal component analysis (PCA). Four cognitive tests went into the PCA (numeric
195 memory, reaction time, pairs matching, and trail making test (TMT) B, see Supplementary
196 Material) and the first PC was considered an estimate of g . A positive g score represents
197 better cognitive performance. Alongside the g score, the cognitive test of verbal and
198 numerical reasoning, also referred to as the test for fluid intelligence, was used to measure

199 current cognitive ability. Educational attainment was measured using years in education (for
200 those without a college or university degree) and highest educational qualification
201 (GCSEs/no GCSEs and degree/no degree).

202

203 Genotypes

204 Genotype data were curated by the UKBB (21) and as described in Leonenko, Baker (22)
205 (see Supplementary Material). The genotypes for p.(Ala391Thr) (INFO=1.00) were extracted
206 for each participant. Principal components provided by the UKBB were used. We restricted
207 our sample to participants who self-reported White British or Irish ancestry then used the first
208 five principal components to identify a subsample of participants that was relatively
209 genetically homogenous. As in Legge, Jones (5), we computed a Minimum Covariance
210 Determinant (MCD) estimator of location and scatter for each participant, and used these to
211 define a hyper-ellipsoid in a multi-dimensional space that contains the majority of MCD
212 points. We used this hyper-ellipsoid to include participants within the 90th percentile of the
213 MCD distance (see Figure 9 in the Supplementary Material).

214

215 Rare Variants

216 Full methods of rare variant calls are reported in Fenner, Holmans (23). In brief, exome
217 sequencing data released for 200,619 UKBB participants in October 2020 were used for the
218 current study. Variants were annotated in Hail using Ensembl's VEP. Protein-truncating
219 variants (PTVs) were defined as splice acceptor, splice donor, stop-gain or frameshift
220 variants that were annotated as high confidence for causing loss of protein function by Loss-
221 Of-Function Transcript Effect Estimator (LoFTEE; 9). Deleterious missense variants were
222 defined as missense variants with a Rare Exome Variant Ensemble Learner score (REVEL;
223 24) > 0.75. PTVs and missense variants were grouped together into one 'damaging' category

224 for analysis. Rare variants were defined as those with allele counts ≤ 5 in the sample reported
225 in Fenner, Holmans (23). We included, as a covariate, the burden of synonymous variants
226 that had allele counts ≤ 5 in the sample reported in Fenner, Holmans (23).

227

228 [Statistical Analysis](#)

229 Each phenotype was regressed onto p.(Ala391Thr) using linear or logistic regression models.
230 Sex, age at interview, the first five genetic principal components (PCs), and any further PCs
231 which were associated with p.(Ala391Thr) (PC 6 and 9 in the Cardiff F-Series sample) were
232 included as covariates. For the UKBB data, the first ten genetic PCs and genotype batch were
233 included as covariates. For models where a measure of cognitive ability in the UKBB was
234 used as the outcome, age at interview squared was also included as a covariate (25), this
235 variable was centred before being transformed to prevent the occurrence of collinearity (26).
236 For the Cardiff schizophrenia samples, missing phenotype data were imputed, for each
237 sample separately, using multiple imputation by chained equations (MICE; 100 imputations,
238 10 iterations in the burn-in period). All regression analyses were run in each of the 100
239 imputed datasets separately and then pooled using Rubin's rules (27). Model assumptions
240 were checked using the R package 'performance' and discussed in the Supplementary
241 Material: for models where there was evidence of multicollinearity (variance inflation factor
242 [VIF] ≥ 10), we removed PCs from the model until the VIF for p.(Ala391Thr) was < 10 , and
243 for models where there was evidence of heteroscedasticity (Breusch-Pagan Test P-value $<$
244 0.01), we recalculated standard errors, confidence intervals, and P-values using Eicker-
245 Huber-White robust "HC2" standard errors. For the Cardiff schizophrenia samples, beta
246 values or log(odds) were meta-analysed, and weighted using their standard errors, using the R
247 package 'metafor' using a fixed-effect inverse-variance weighted model. We corrected for
248 multiple testing of phenotypes in the meta-analysis, and within each sample, using the

249 Benjamini–Hochberg False Discovery Rate (FDR) method and used an alpha level of $\leq .05$
250 for the adjusted p-values to determine statistical significance. For the rare variant analysis,
251 each of the phenotypes was regressed against the number of rare variants, and adjusted for
252 the number of rare synonymous variants each person carries in *SLC39A8*, age at interview
253 (age at interview squared for *g* and FI), sex, exome data PC1-10, and sequencing batch.

254 Results

255 Phenotype means and proportions, stratified by p.(Ala391Thr) schizophrenia-risk allele
256 count, are presented for each sample in the Supplementary Material.

257

258 Cardiff Schizophrenia Samples

259 Data were available for 662 participants from CardiffCOGS (mean age at interview
260 [AAI]=43.31 years, standard deviation [sd]=12.11; 65.75% male), 422 from Cardiff F-Series
261 (mean AAI=42.11 years, sd=14.08; 70.66% male) and 148 from Cardiff SibPairs (mean
262 AAI=41.68 years, sd=12.93; 66.21% male).

263

264 p.(Ala391Thr) was not associated with positive symptoms, negative symptoms, or
265 disorganised symptoms derived in either CFA model (Table 1 and Table 2). p.(Ala391Thr)
266 was not associated with current cognitive ability in CardiffCOGS, however, it was nominally
267 associated with a lower NART IQ score but this association did not survive correction for
268 multiple testing (Table 2). Although I^2 should be interpreted cautiously in meta-analyses of
269 only a few studies (28), there does not appear to be evidence of between-study heterogeneity
270 after meta-analysing phenotypes which were present in all three samples.

271

272 UK Biobank

273 Data were available for N=355,069 participants from UKBB who did not have a psychotic
274 spectrum disorder (mean AAI=56.92 years, sd=7.96; 46.25% male), of which N=354,509
275 indicated that they were willing to attempt the cognitive tests and N=116,935 completed the
276 online follow-up mental health questionnaire.

277

278 p.(Ala391Thr) was not associated with reported psychotic experiences. However, the
279 schizophrenia-risk allele was associated with a lower g score ($\beta = -0.05$; 95% CI, -0.07 to -
280 0.02; FDR-adjusted p-value= 8.61×10^{-5}), a lower fluid intelligence score ($\beta = -0.05$; 95% CI, -
281 0.07 to -0.04; FDR-adjusted p-value= 3.35×10^{-10}), a lower likelihood of obtaining GCSEs
282 (OR=0.96; 95% CI, 0.94-0.98; FDR-adjusted p-value= 6.21×10^{-5}) or a degree-level
283 qualification (OR=0.95; 95% CI, 0.93-0.97; FDR-adjusted p-value= 6.03×10^{-6}) and fewer
284 years in education ($\beta = -0.03$; 95% CI, -0.05 to 0.00; FDR-adjusted p-value=0.035) (Table 3).
285
286 Rare variant calls were available for N=134,370 (37.84%) of the participants included in the
287 current study. N=46 carried rare synonymous variants and N=13 carried rare damaging
288 variants (N=6 PTVs and N=7 missense carriers; mean AAI=55.54 years, sd=9.92; 46.15%
289 male). Most rare variant carriers had missing phenotype data. Rare variants in *SLC39A8* were
290 nominally associated with lower g (N rare variants=2; $\beta = -1.67$; 95% CI, -2.93 to -0.42; P-
291 value=0.009) and lower fluid intelligence (N rare variants=1; $\beta = -2.31$; 95% CI, -4.25 to -
292 0.38; P-value=0.019) but this association did not survive correction for multiple testing
293 (Table 4). None of the rare variant carriers had a diagnosis of Intellectual Disability (ICD-10
294 Codes F70, F71, F72, F78, or F79 identified from the hospital, death, and primary care
295 records).

296 Discussion

297

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

306

307 Although understanding the potential pathophysiological mechanisms of variants identified
308 by GWAS is challenging, there is an increasing focus on p.(Ala391Thr) in the context of the
309 aetiology of schizophrenia (1) and some proposed mechanisms are also thought to play a role
310 in cognition. p.(Ala391Thr) is thought to lead to synaptic glutamate receptor hypofunction, in
311 part, because of downregulated surface localisation of GluA1, GluA2/3, GluN1, and GluN2A
312 (3). The latter are subunits of the *N*-methyl-d-aspartate (NMDA) receptor and the
313 hypofunctioning of NMDAR has been implicated in the aetiology schizophrenia and
314 specifically impaired learning and memory (29). The schizophrenia-risk allele at
315 p.(Ala391Thr) has also been associated with lower manganese levels (2, 30, 31), an essential
316 trace element transported by ZIP8 and involved in glycosylation, the process by which
317 branched sugar polymers are covalently attached to proteins and lipids (32). Glycosylation is
318 dysregulated in schizophrenia (32) and disrupted by the schizophrenia-risk allele at
319 p.(Ala391Thr) (33). Individuals with Congenital Disorders of Glycosylation can present with
320 cognitive impairment (32), but there is no consistent linear association between lower blood

321 manganese concentrations and poorer cognitive ability (34). Finally, the schizophrenia-risk
322 allele is also associated with impaired zinc uptake and transportation, and decreased cortical
323 dendritic spine density (35). Developmental synaptic pruning has been postulated as a risk
324 mechanism for schizophrenia either through a loss of balance between synaptogenesis and
325 elimination or abnormal activity-dependent plasticity (36). Decreased dendritic spine density
326 has been observed in individuals with schizophrenia (37). Dendritic spine plasticity is thought
327 to underlie the cognitive resilience of older adults who, despite having Alzheimer's
328 pathophysiology, have not developed dementia (38).

329

330 p.(Ala391Thr) has been reported to be a highly pleiotropic variant and is associated, at
331 genome-wide significance levels, with 24.16% (129/534) of the traits curated by Open
332 Targets Genetics (39); the schizophrenia-risk allele has been associated with lower HDL
333 cholesterol, lower blood pressure, lower levels of apolipoprotein A1, calcium, aspartate
334 aminotransferase, urate, gamma-glutamyl transferase, and serum albumin, as well as higher
335 body mass index and body fat measures, but a lower risk of hypertension and cardiovascular
336 disease. Notably, in GWAS, other than in schizophrenia, p.(Ala391Thr) has not been
337 associated with any psychiatric disorder (Bipolar Disorder, Major Depressive Disorder) or
338 with neurodevelopmental (Autism, Attention deficit hyperactivity disorder), or
339 neurodegenerative (Alzheimer's Disease, Parkinson's Disease) disorders (see Buniello,
340 MacArthur (40) and Supplementary Table 1). In phenome-wide association studies
341 (PheWAS) using the UK Biobank (UKBB), p.(Ala391Thr) has been associated with diseases
342 of the oesophagus, musculoskeletal conditions, metabolic and digestive biomarkers, blood
343 pressure, and dietary and lifestyle factors including weight gain and drinking alcohol (30,
344 <http://www.nealelab.is/uk-biobank/>, 41). In a brain MRI PheWAS of the UKBB (42), the
345 schizophrenia risk allele was associated with greater putamen grey matter volume, reduced

346 cortical thickness, and reduced white matter integrity, MRI phenotypes which have been
347 identified in participants with schizophrenia. Although pleiotropy has benefits for
348 translational research, for example by cutting across current diagnostic categories, the
349 diversity of phenotypes suggests that *SLC39A8* may not impact schizophrenia or cognition
350 through the disruption of a single key biological process, rather it may influence multiple
351 processes, not necessarily all within the brain.

352

353 As well as increasing our understanding of the aetiology of schizophrenia, studying
354 p.(Ala391Thr) could improve our understanding of the cognitive impairments in people with
355 schizophrenia. Within studies of cognitive remediation therapy (CRT) in schizophrenia,
356 although on average patients with poorer premorbid IQ and fewer years in education show
357 greater improvement after CRT, some studies have shown the opposite (43) and there is
358 emerging literature examining whether genetic variants can account for this differential
359 improvement (44). For example, one study examined *SLC1A2*, a high-affinity glutamate
360 transporter that encodes EAAT2, and found that the minor allele at
361 NC_000011.10:g.35419429T>G (rs4354668), which has previously been associated with
362 lower EAAT2 expression and poorer cognition in healthy controls and patients with
363 schizophrenia, was associated with poorer improvement after CRT (45). p.(Ala391Thr) could
364 be a candidate for future stratification studies.

365

366 [Limitations](#)

367

368 This is the first study to test the relationship between p.(Ala391Thr) and schizophrenia-
369 related phenotypes in participants with the disorder. However, our cohorts of schizophrenia
370 participants were small, and it may be that larger sample sizes are needed to detect the small

371 effects attributable to a single SNP. Given an alpha level of 0.05, and 90% power, to observe
372 a beta coefficient between -0.07 to -0.02, between N=1741 and N=21403 schizophrenia
373 participants would be required. Our analysis of rare variants in *SLC39A8* may not be
374 representative because of the small number of rare variant carriers driving the associations.
375 Gene-based burden tests usually apply a cut off to exclude genes with a low a number of
376 carriers (46), which, in our sample, would have excluded *SLC39A8* from a pipeline for
377 exome-wide gene-based burden tests. Our rare variant results should therefore be used to
378 inform future studies rather than be interpreted in isolation. It is also possible that our
379 findings will not generalise to participants of non-European ancestry. Previously work has
380 described p.(Ala391Thr) as having an almost negligible minor allele frequency in non-
381 European populations (1). In our schizophrenia cohorts, only Europeans carried two copies of
382 the schizophrenia risk allele but three other ancestry groups (African American/Afro-
383 Caribbean, Near Eastern, and Central/South Asian) contained participants with one copy so
384 we chose to include them in our analysis. Nevertheless, less than 2% of the schizophrenia
385 cohorts were of non-European ancestry and our sample from the UKBB was comprised
386 solely of people with European ancestry.
387
388 Another caveat is that it is unclear whether psychotic experiences measured in the UKBB are
389 a good proxy for psychotic experiences experienced by people with schizophrenia. Although
390 there is a shared genetic liability between psychotic experiences and schizophrenia, the
391 genetic correlation is weak ($r_g = 0.21$; 5), and psychotic experiences were more strongly
392 correlated with ADHD ($r_g = 0.24$), autism spectrum disorder ($r_g = 0.39$), and major
393 depressive disorder ($r_g = 0.46$) than schizophrenia, suggesting that these phenotypes capture
394 general psychopathology rather than schizophrenia-specific psychosis. There is also the
395 possibility of measurement error in the psychotic experiences phenotypes; firstly, because

396 they are measured retrospectively by self-report, and, secondly, because of biases in those
397 who attempt to complete the online follow-up mental health questionnaire. Legge, Jones (5)
398 found that participants who completed the mental health questionnaire had significantly
399 higher intelligence and lower schizophrenia polygenic risk scores adding to existing evidence
400 of a “healthy volunteer” selection bias in the UKBB (47, 48) which can affect the results of
401 GWAS (49). As the schizophrenia-risk allele was associated with responding to questions in
402 the UKBB with ‘prefer not to answer’ and ‘I don’t know’ (50), we tested whether
403 p.(Ala391Thr) was associated with willingness to attempt the cognitive tests or the mental
404 health questionnaire. p.(Ala391Thr) was not associated with attempting to complete the
405 baseline cognitive tests, but the schizophrenia-risk allele was associated with *not* attempting
406 the mental health questionnaire (see Supplementary Material). This suggests that the
407 psychotic experiences phenotypes in the UKBB are not completed by participants with a
408 higher genetic load for schizophrenia and our analyses may be affected by collider bias.

409

410 Conclusions

411

412 The schizophrenia-risk allele at p.(Ala391Thr) is associated with poorer cognitive ability in a
413 sample of the general population with European ancestry. Larger and more ancestrally-
414 diverse studies of participants with schizophrenia are required to determine whether there is
415 an association between p.(Ala391Thr) and cognition in schizophrenia, and subsequently
416 determine whether p.(Ala391Thr) and ZIP8 are potential targets for novel therapeutic
417 treatments for cognitive impairment in those with the disorder.

418

419 Data Availability

420 To comply with the ethical and regulatory framework under which the Cardiff Schizophrenia
421 Samples were obtained, access to individual-level data requires a collaboration agreement
422 with Cardiff University; requests to access these datasets should be directed to J.T.R.W.
423 (WaltersJT@cardiff.ac.uk) and M.J.O. (OwenMJ@cardiff.ac.uk). UK Biobank data is
424 available by application to the UK Biobank (www.ukbiobank.ac.uk).

425 Code Availability

426 <https://zenodo.org/records/10027873>

427 References

428

429

- 430 1. Costas J. The highly pleiotropic gene SLC39A8 as an opportunity to gain insight into
431 the molecular pathogenesis of schizophrenia. *Am J Med Genet B Neuropsychiatr Genet.*
432 2018;177(2):274-83.
- 433 2. Fujishiro H, Miyamoto S, Sumi D, Kambe T, Himeno S. Effects of individual amino acid
434 mutations of zinc transporter ZIP8 on manganese- and cadmium-transporting activity.
435 *Biochemical and Biophysical Research Communications.* 2022;616:26-32.
- 436 3. Tseng WC, Reinhart V, Lanz TA, Weber ML, Pang J, Le KXV, et al. Schizophrenia-
437 associated SLC39A8 polymorphism is a loss-of-function allele altering glutamate receptor
438 and innate immune signaling. *Transl Psychiatry.* 2021;11(1):136.
- 439 4. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al.
440 Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature.* 2022.
- 441 5. Legge SE, Jones HJ, Kendall KM, Pardiñas AF, Menzies G, Bracher-Smith M, et al.
442 Association of Genetic Liability to Psychotic Experiences With Neuropsychotic Disorders and
443 Traits. *JAMA Psychiatry.* 2019;76(12):1256-65.
- 444 6. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, et al. Genome-
445 wide association meta-analysis in 269,867 individuals identifies new genetic and functional
446 links to intelligence. *Nature Genetics.* 2018;50(7):912-9.
- 447 7. Okbay A, Wu Y, Wang N, Jayashankar H, Bennett M, Nehzati SM, et al. Polygenic
448 prediction of educational attainment within and between families from genome-wide
449 association analyses in 3 million individuals. *Nat Genet.* 2022;54(4):437-49.
- 450 8. Demange PA, Malanchini M, Mallard TT, Biroli P, Cox SR, Grotzinger AD, et al.
451 Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction.
452 *Nature Genetics.* 2021;53(1):35-44.
- 453 9. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The
454 mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.*
455 2020;581(7809):434-43.
- 456 10. Little J, Higgins JPT, Ioannidis JPA, Moher D, Gagnon F, von Elm E, et al.
457 Strengthening the REporting of Genetic Association Studies (STREGA)—an extension of the
458 STROBE statement. *Genetic Epidemiology.* 2009;33(7):581-98.
- 459 11. Legge SE, Dennison CA, Pardiñas AF, Rees E, Lynham AJ, Hopkins L, et al. Clinical
460 indicators of treatment-resistant psychosis. *The British Journal of Psychiatry.*
461 2020;216(5):259-66.
- 462 12. Williams NM, Green EK, Macgregor S, Dwyer S, Norton N, Williams H, et al. Variation
463 at the DAOA/G30 locus influences susceptibility to major mood episodes but not psychosis
464 in schizophrenia and bipolar disorder. *Arch Gen Psychiatry.* 2006;63(4):366-73.
- 465 13. Williams NM, Rees MI, Holmans P, Norton N, Cardno AG, Jones LA, et al. A Two-
466 Stage Genome Scan for Schizophrenia Susceptibility Genes in 196 Affected Sibling Pairs.
467 *Human Molecular Genetics.* 1999;8(9):1729-39.
- 468 14. Legge SE, Cardno AG, Allardyce J, Dennison C, Hubbard L, Pardiñas AF, et al.
469 Associations Between Schizophrenia Polygenic Liability, Symptom Dimensions, and
470 Cognitive Ability in Schizophrenia. *JAMA Psychiatry.* 2021;78(10):1143-51.

- 471 15. Lynham AJ, Knott S, Underwood JFG, Hubbard L, Agha SS, Bisson JI, et al. DRAGON-
472 Data: a platform and protocol for integrating genomic and phenotypic data across large
473 psychiatric cohorts. *BJPsych Open*. 2023;9(2):e32.
- 474 16. Huddart R, Fohner AE, Whirl-Carrillo M, Wojcik GL, Gignoux CR, Popejoy AB, et al.
475 Standardized Biogeographic Grouping System for Annotating Populations in
476 Pharmacogenetic Research. *Clinical Pharmacology & Therapeutics*. 2019;105(5):1256-62.
- 477 17. Bulbul O, Filoglu G, Zorlu T, Altuncul H, Freire-Aradas A, Söchtig J, et al. Inference of
478 biogeographical ancestry across central regions of Eurasia. *International journal of legal
479 medicine*. 2016;130(1):73-9.
- 480 18. Li M, Wu DD, Yao YG, Huo YX, Liu JW, Su B, et al. Recent Positive Selection Drives the
481 Expansion of a Schizophrenia Risk Nonsynonymous Variant at SLC39A8 in Europeans.
482 *Schizophr Bull*. 2016;42(1):178-90.
- 483 19. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An
484 Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of
485 Middle and Old Age. *PLOS Medicine*. 2015;12(3):e1001779.
- 486 20. Fawns-Ritchie C, Deary IJ. Reliability and validity of the UK Biobank cognitive tests.
487 *PLOS ONE*. 2020;15(4):e0231627.
- 488 21. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank
489 resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203-9.
- 490 22. Leonenko G, Baker E, Stevenson-Hoare J, Sierksma A, Fiers M, Williams J, et al.
491 Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores.
492 *Nature Communications*. 2021;12(1):4506.
- 493 23. Fenner E, Holmans P, O'Donovan MC, Owen MJ, Walters JT, Rees E. Rare coding
494 variants in schizophrenia-associated genes affect generalised cognition in the UK Biobank.
495 *medRxiv*. 2023:2023.08.14.23294074.
- 496 24. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al.
497 REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *The
498 American Journal of Human Genetics*. 2016;99(4):877-85.
- 499 25. Cornelis MC, Wang Y, Holland T, Agarwal P, Weintraub S, Morris MC. Age and
500 cognitive decline in the UK Biobank. *PLOS ONE*. 2019;14(3):e0213948.
- 501 26. Schielzeth H. Simple means to improve the interpretability of regression coefficients.
502 *Methods in Ecology and Evolution*. 2010;1(2):103-13.
- 503 27. Rubin DB. *Multiple imputation for nonresponse in surveys*: John Wiley & Sons; 2004.
- 504 28. von Hippel PT. The heterogeneity statistic I² can be biased in small meta-analyses.
505 *BMC Medical Research Methodology*. 2015;15(1):35.
- 506 29. Nakazawa K, Sapkota K. The origin of NMDA receptor hypofunction in schizophrenia.
507 *Pharmacology & Therapeutics*. 2020;205:107426.
- 508 30. Moksnes MR, Hansen AF, Wolford BN, Thomas LF, Rasheed H, Simić A, et al. New
509 insights into the genetic etiology of 57 essential and non-essential trace elements in
510 humans. *medRxiv*. 2023:2023.04.25.23289097.
- 511 31. Mealer RG, Jenkins BG, Chen CY, Daly MJ, Ge T, Lehoux S, et al. The schizophrenia
512 risk locus in SLC39A8 alters brain metal transport and plasma glycosylation. *Scientific
513 Reports*. 2020;10(1):13162.
- 514 32. Pradeep P, Kang H, Lee B. Glycosylation and behavioral symptoms in neurological
515 disorders. *Translational Psychiatry*. 2023;13(1):154.

- 516 33. Mealer RG, Williams SE, Noel M, Yang B, D'Souza AK, Nakata T, et al. The
517 schizophrenia-associated variant in SLC39A8 alters protein glycosylation in the mouse brain.
518 *Molecular Psychiatry*. 2022;27(3):1405-15.
- 519 34. Vollet K, Haynes EN, Dietrich KN. Manganese Exposure and Cognition Across the
520 Lifespan: Contemporary Review and Argument for Biphasic Dose–Response Health Effects.
521 *Current Environmental Health Reports*. 2016;3(4):392-404.
- 522 35. Li S, Ma C, Li Y, Chen R, Liu Y, Wan LP, et al. The schizophrenia-associated missense
523 variant rs13107325 regulates dendritic spine density. *Translational Psychiatry*.
524 2022;12(1):361.
- 525 36. Kasai H, Fukuda M, Watanabe S, Hayashi-Takagi A, Noguchi J. Structural dynamics of
526 dendritic spines in memory and cognition. *Trends in Neurosciences*. 2010;33(3):121-9.
- 527 37. Radhakrishnan R, Skosnik PD, Ranganathan M, Naganawa M, Toyonaga T, Finnema S,
528 et al. In vivo evidence of lower synaptic vesicle density in schizophrenia. *Molecular*
529 *Psychiatry*. 2021;26(12):7690-8.
- 530 38. Boros BD, Greathouse KM, Gentry EG, Curtis KA, Birchall EL, Gearing M, et al.
531 Dendritic spines provide cognitive resilience against Alzheimer's disease. *Annals of*
532 *Neurology*. 2017;82(4):602-14.
- 533 39. Mountjoy E, Schmidt EM, Carmona M, Schwartzenuber J, Peat G, Miranda A, et al.
534 An open approach to systematically prioritize causal variants and genes at all published
535 human GWAS trait-associated loci. *Nat Genet*. 2021;53(11):1527-33.
- 536 40. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The
537 NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays
538 and summary statistics 2019. *Nucleic acids research*. 2019;47(D1):D1005-D12.
- 539 41. MRC IEU UK Biobank GWAS pipeline version 2 [Internet]. 2019.
- 540 42. Hermann ER, Chambers E, Davis DN, Montgomery MR, Lin D, Chowanadisai W. Brain
541 Magnetic Resonance Imaging Phenome-Wide Association Study With Metal Transporter
542 Gene SLC39A8. *Front Genet*. 2021;12:647946.
- 543 43. Vita A, Barlati S, Ceraso A, Nibbio G, Ariu C, Deste G, et al. Effectiveness, Core
544 Elements, and Moderators of Response of Cognitive Remediation for Schizophrenia: A
545 Systematic Review and Meta-analysis of Randomized Clinical Trials. *JAMA Psychiatry*.
546 2021;78(8):848-58.
- 547 44. Penadés R, Bosia M, Catalán R, Spangaro M, García-Rizo C, Amoretti S, et al. The role
548 of genetics in cognitive remediation in schizophrenia: A systematic review. *Schizophr Res*
549 *Cogn*. 2020;19:100146.
- 550 45. Spangaro M, Bosia M, Bechi M, Buonocore M, Cocchi F, Guglielmino C, et al.
551 Neurobiology of cognitive remediation in schizophrenia: Effects of EAAT2 polymorphism.
552 *Schizophrenia Research*. 2018;202:106-10.
- 553 46. Chen C-Y, Tian R, Ge T, Lam M, Sanchez-Andrade G, Singh T, et al. The impact of rare
554 protein coding genetic variation on adult cognitive function. *Nature Genetics*.
555 2023;55(6):927-38.
- 556 47. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison
557 of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With
558 Those of the General Population. *Am J Epidemiol*. 2017;186(9):1026-34.
- 559 48. Legge SE, Pardiñas AF, Woolway G, Rees E, Cardno AG, Escott-Price V, et al. An
560 Assessment of the Genetic and Phenotypic Features of Schizophrenia in the UK Biobank
561 *JAMA Psychiatry*. Under Review.

- 562 49. Schoeler T, Speed D, Porcu E, Pirastu N, Pingault J-B, Kutalik Z. Participation bias in
563 the UK Biobank distorts genetic associations and downstream analyses. *Nature Human*
564 *Behaviour*. 2023;7(7):1216-27.
- 565 50. Mignogna G, Carey CE, Wedow R, Baya N, Cordioli M, Pirastu N, et al. Patterns of
566 item nonresponse behaviour to survey questionnaires are systematic and associated with
567 genetic loci. *Nature Human Behaviour*. 2023.

568 Acknowledgements

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

571 Funding

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

580 Author Contributions

581 S.E.S. designed and conducted the analysis and drafted the manuscript. S.E.L. conducted the
582 confirmatory factor analysis. E.F. called the rare variants. A.F.P. conducted the genetic
583 ancestry analysis. A.J.L. scored the MATRICS Consensus Cognitive Battery. V.E-P., E.F.,
584 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
585 analysis and interpret the results. All authors revised and approved the final manuscript.

586 Ethical Approval

587 CardiffCOGS was approved by the South East Wales Research Ethics Committee Panel
588 (reference number: 07/ WSE03/110) and received HRA approval. All participants provided
589 written informed consent. Multicentre and Local Research Ethics Committee approval was

590 obtained for Cardiff F-Series, and all participants gave written informed consent to
591 participate. For Cardiff SibPairs written consent was obtained following local ethical
592 approval guidelines. Ethical approval for the curation and development of DRAGON-Data
593 was obtained from Cardiff University's School of Medicine Research Ethics Committee (Ref:
594 19/72). We also used data from the UK Biobank (<https://www.ukbiobank.ac.uk>), the
595 scientific protocol of which has been reviewed and approved by the North West Multi-centre
596 Ethics Committee. Our access to the UK Biobank data was under the project number 13310.

597 Competing Interests

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