

Article

Cerebrospinal Fluid C1-Esterase Inhibitor and Tie-1 Levels Affect Cognitive Performance: Evidence from Proteome-Wide Mendelian Randomization

Loukas Zagkos ^{1,*} , Marie-Joe Dib ², H el ene T. Cronj e ³, Paul Elliott ^{1,4,5}, Abbas Dehghan ^{1,4,5}, Ioanna Tzoulaki ^{1,4,5,6} , Dipender Gill ¹  and Iyas Daghlas ⁷

- ¹ Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London SW7 2BX, UK; p.elliott@imperial.ac.uk (P.E.); a.dehghan@imperial.ac.uk (A.D.); i.tzoulaki@imperial.ac.uk (I.T.); dipender.gill@imperial.ac.uk (D.G.)
- ² Division of Cardiovascular Medicine, Hospital of the University of Pennsylvania, Philadelphia, PA 19104, USA; marie-joe.dib@pennmedicine.upenn.edu
- ³ Department of Public Health, Section of Epidemiology, University of Copenhagen, 1165 Copenhagen, Denmark; toinet.cronje@sund.ku.dk
- ⁴ UK Dementia Research Institute at Imperial College London, Hammersmith Hospital, London W1T 7NF, UK
- ⁵ Medical Research Council Centre for Environment and Health, School of Public Health, Imperial College London, London SW7 2AZ, UK
- ⁶ Centre for Systems Biology, Biomedical Research Foundation of the Academy of Athens, 11527 Athens, Greece
- ⁷ Department of Neurology, University of California, San Francisco, San Francisco, CA 94143, USA; iyas.daghlas@ucsf.edu
- * Correspondence: l.zagkos@imperial.ac.uk

Abstract: Objective: The association of cerebrospinal fluid (CSF) protein levels with cognitive function in the general population remains largely unexplored. We performed Mendelian randomization (MR) analyses to query which CSF proteins may have potential causal effects on cognitive performance. Methods and analysis: Genetic associations with CSF proteins were obtained from a genome-wide association study conducted in up to 835 European-ancestry individuals and for cognitive performance from a meta-analysis of GWAS including 257,841 European-ancestry individuals. We performed Mendelian randomization (MR) analyses to test the effect of randomly allocated variation in 154 genetically predicted CSF protein levels on cognitive performance. Findings were validated by performing colocalization analyses and considering cognition-related phenotypes. Results: Genetically predicted C1-esterase inhibitor levels in the CSF were associated with a better cognitive performance (SD units of cognitive performance per 1 log-relative fluorescence unit (RFU): 0.23, 95% confidence interval: 0.12 to 0.35, $p = 7.91 \times 10^{-5}$), while tyrosine-protein kinase receptor Tie-1 (sTie-1) levels were associated with a worse cognitive performance (-0.43 , -0.62 to -0.23 , $p = 2.08 \times 10^{-5}$). These findings were supported by colocalization analyses and by concordant effects on distinct cognition-related and brain-volume measures. Conclusions: Human genetics supports a role for the C1-esterase inhibitor and sTie-1 in cognitive performance.

Keywords: CSF; cognition; tyrosine-protein kinase receptor; *SERPING1*; *TIE1*



Citation: Zagkos, L.; Dib, M.-J.; Cronj e, H.T.; Elliott, P.; Dehghan, A.; Tzoulaki, I.; Gill, D.; Daghlas, I. Cerebrospinal Fluid C1-Esterase Inhibitor and Tie-1 Levels Affect Cognitive Performance: Evidence from Proteome-Wide Mendelian Randomization. *Genes* **2024**, *15*, 71. <https://doi.org/10.3390/genes15010071>

Academic Editor: Xingguang Luo

Received: 29 November 2023

Revised: 29 December 2023

Accepted: 1 January 2024

Published: 4 January 2024



Copyright:   2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cerebrospinal fluid (CSF) is an ultrafiltrate of plasma produced in the cerebral ventricles and located within the subarachnoid space where it maintains the structural and physiological integrity of the central nervous system (CNS) [1,2]. The composition of CSF, including protein levels, provides information about brain health and may help to stratify neurodegenerative disease risk [3–6]. Given its proximity to nervous system tissue, CSF may be more informative about protein–disease relationships than serum. These proteins may also represent targets for the development of therapeutics that prevent or treat disease [7].

Prior studies have mostly focused on the association of CSF protein levels with the risk of dementia [5,6], while investigations into the relationship between CSF proteins and cognition more broadly have been limited [8]. For example, prior epidemiologic studies investigated the association between CSF biomarkers and cognitive function. The proteins investigated in these studies were curated based on known relationships with AD (e.g., amyloid beta, tau, and neurofilament light chain) [9–11]. Thus, the relationship of the proteome with cognition remains a largely unexplored frontier.

While epidemiologic studies contribute important information to the understanding of the association between biomarkers and disease, associations from conventional epidemiologic analyses can be biased by residual confounding or reverse causality and may therefore not represent causal relationships. Mendelian randomization (MR) is an alternative analytic paradigm that uses human genetic data to estimate less biased causal effects. MR leverages randomly allocated genetic variants as proxies for exposures to estimate the causal effect of an exposure (e.g., levels of a protein in the CSF) on an outcome (e.g., cognitive function) [12,13]. Due to the random and fixed allocation of genetic variants at conception, this approach is less susceptible to potential bias due to confounding from environmental factors or reverse causation. The recent discovery of genetic variants associated with CSF protein levels and the availability of large-scale genetic association studies of cognition-related phenotypes permits the application of MR to examine the relationship between the CSF proteome and cognitive performance [14]. In this work, we applied MR to systematically query the human CSF proteome for proteins with evidence for a causal effect on cognitive performance.

2. Materials and Methods

2.1. Study Design

The overall study design is provided in Figure 1. We conducted *cis*-Mendelian randomization (*cis*-MR) analyses to explore the associations of genetically predicted CSF protein concentration with cognitive performance. To ensure that MR findings were not biased by linkage disequilibrium (LD) with a nearby genetic association signal, further statistical validation was performed using genetic colocalization analyses [15]. For proteins with evidence from colocalization analyses, we conducted complementary MR analyses to assess the association of genetically predicted protein levels with 13 curated cognition-related phenotypes. This study is reported using the Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomization (STROBE-MR) guidelines (Supplementary Table S1).

2.2. Genetic Associations with Protein Levels in CSF and Other Tissues

Yang et al. conducted a genome-wide association study (GWAS) of 713 cerebrospinal fluid (CSF) proteins in 971 individuals recruited from the Washington University School of Medicine in St. Louis, including 249 patients with Alzheimer's disease (AD), 717 cognitively normal controls, and 5 with unknown cognitive function status [14]. The average age of the study participants was 69 years (standard deviation [SD] 9.3 years), 53% were women, and 39% had at least one copy of the *APOE* ϵ 4 allele. Study participants were unrelated and of European ancestry (detailed information available in the GWAS manuscript). After quality control by Yang et al., the total sample size was 835 patients. Of note, the authors demonstrated that genetic associations with protein levels were nearly identical between patients with and without AD and thus the inclusion of AD cases does not bias the genetic associations with protein levels ($r > 0.98$ for genetic associations with proteins in participants with and without AD). CSF was collected via lumbar puncture after an overnight fast. A multiplexed aptamer-based platform developed by SomaLogic Inc. (Boulder, CO, USA) was used to measure the relative concentrations of proteins in the CSF [16]. The protein levels were quantified as relative fluorescence units (RFU). Genetic association estimates with CSF protein levels were adjusted for sex, age, the first two genetic principal components, and the genotyping platform. In additional sensitivity analyses, we used

genetic association data for plasma (n = 529) and brain (n = 380; measured using parietal cortex tissue) proteins from the same study.

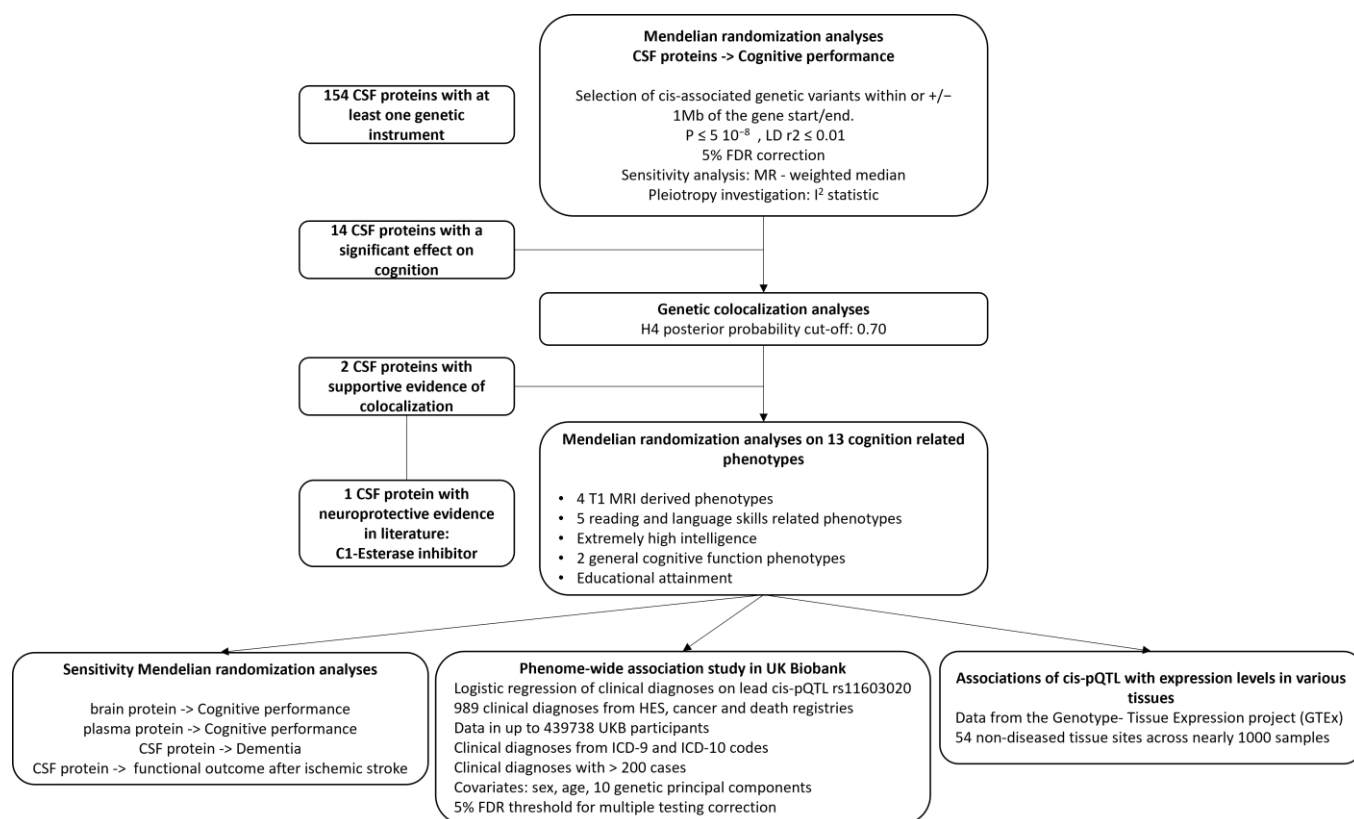


Figure 1. Analysis pipeline. Mendelian randomization (MR) analyses were performed to explore the association of genetically proxied levels of cerebrospinal fluid (CSF) proteins on cognitive performance. After a 5% false discovery rate (FDR) correction, 14 proteins that were significantly associated with cognitive performance were taken forward for genetic colocalization analysis. Two CSF proteins had evidence of genetic colocalization with cognitive performance and were included in an additional MR aimed at assessing their association with 13 complementary cognition-related phenotypes. Because of the literature-based support for the associations we observed with the CSF C1-esterase inhibitor, we conducted additional sensitivity and exploratory analyses including a phenome-wide association study for the lead CSF C1-esterase inhibitor *cis*-protein quantitative trait locus (pQTL) and gene expression association analysis.

2.3. Genetic Associations with the Outcomes

The primary outcome of this MR analysis was cognitive performance. We obtained genetic association estimates for cognitive performance from a Social Science Genetic Association Consortium (SSGAC) meta-analysis conducted by Lee et al. [17], including 257,841 European-ancestry UK Biobank [18] and Cognitive Genomics Consortium (COGENT) participants [19]. In the UK Biobank, cognitive performance was measured in 222,543 participants using a standardized score of a test of verbal–numerical reasoning. The test contained thirteen logic and reasoning questions with a two-minute time limit and was designed as a measure of fluid intelligence. Each respondent took the test up to four times, and the mean of the standardized scores was used, which was then standardized. In COGENT, cognitive performance was quantified as the first unrotated principal component of performance on at least three neuropsychological tests or two intelligence quotient (IQ) test subscales, measuring the overall accuracy or total number of correct responses in 35,298 individuals. In general, the test variable used either the overall accuracy or total

number of correct responses. Adjustments for age, sex, and population stratification were included in the genotype–phenotype association analysis for each cohort.

In secondary analyses, we considered 13 cognition-related phenotypes as outcomes, to explore how proteins might impact different dimensions of cognitive function. Summary statistics were obtained for volumetric brain magnetic resonance imaging (MRI)-derived phenotypes (T1; left and right hippocampal volume, cortical gray matter volume, and total brain volume) in 33,224 UK Biobank participants [20]. Genetic associations with phenotypes related to reading and language skills (word reading, nonword reading, spelling, phoneme awareness, and nonword repetition) in up to 27,180 European-ancestry participants from a published GWAS meta-analysis [21]. Nonword reading and repetition are language assessment tasks where individuals are asked to read and repeat words that do not exist in the language lexicon, which are standard tests that can be used to identify individuals with language difficulties. Genetic association data were also obtained for a binary phenotype of ‘extremely high intelligence’ in 9410 individuals of European ancestry (1238 cases with a mean IQ > 170 and 8172 controls) [22]. For general cognitive function phenotypes, we included reaction time (N = 330,069) and verbal numeric reasoning (N = 168,033) from a GWAS meta-analysis by Davies et al. [23]. Finally, for educational attainment we made use of the meta-analysis by Lee et al. (N = 766,345) [17]. All secondary phenotypes are in standard deviation units apart from the MRI traits, which have mm³ units. There was no sample overlap between the exposure (CSF protein) and any of the outcome GWAS datasets.

2.4. Selection of Genetic Proxies for CSF Protein Levels

For MR analysis, we used *cis* protein quantitative trait loci (pQTLs) as genetic instruments for each CSF protein. Genetic variants were considered *cis* to the gene when they were located within 1 Mb of the gene start or end position. This decision to use *cis* variants was motivated by the notion that such variants are less likely to be biased when used to proxy protein levels compared to variants from throughout the genome (also referred to as *trans* variants; 7). Gene coordinates were determined based on the Genome Reference Consortium Human version 37 (GRCh37) released by the Ensembl genome browser. We selected genome-wide significant pQTLs ($p < 5 \times 10^{-8}$) that were also present in the outcome datasets. To obtain independent genetic instruments for each CSF protein, the variants were clumped using a pair-wise LD $r^2 < 0.01$ (7) from the 1000 genomes project phase 3 European LD reference panel [24]. A distance-based threshold of 1 Mb was used as the clumping window.

2.5. Statistical Analysis

2.5.1. Mendelian Randomization Analysis

We conducted two-sample MR analyses to explore the effect of CSF protein levels on the primary and secondary outcome phenotypes. Genetic associations with the exposures and outcomes were harmonized by aligning effect alleles, using the ‘TwoSampleMR’ v.0.6.0 R package [25]. MR effects were estimated either using the Wald ratio for single variant instruments [25], or the random-effects inverse-variance weighted method for instruments comprising multiple genetic variants [13]. MR effect estimates were considered statistically significant below a false discovery rate (FDR) of 5% to account for the multiple testing of correlated phenotypes. MR effect estimates are reported as the SD change in cognitive performance per 1-log RFU increase in the genetically predicted CSF protein level.

Where more than two genetic instruments were available for a given protein, we performed sensitivity analyses to assess the robustness of MR estimates to bias due to horizontal pleiotropy. Horizontal pleiotropy refers to the association of the genetic instrument with the outcome through pathways independent of the exposure and leads to potential bias in the MR estimate [26]. We first calculated the I^2 statistic for heterogeneity, which is a global indicator of either heterogeneity or pleiotropy [27]. We then used the weighted median method which orders the MR estimates obtained by each genetic instrument by their magnitude, weighted for their precision, and calculates an overall MR estimate based

on the median value [28]. Proteins with consistent MR estimates using the weighted median method were carried forward for colocalization analysis.

2.5.2. Genetic Colocalization Analysis

Due to linkage disequilibrium, genetic variants in close proximity tend to be inherited together and are often correlated. If distinct, correlated causal variants are associated with the exposure and the outcome, this may introduce bias in the *cis*-MR effect estimates as it allows an association between a genetic variant and the outcome via an alternative pathway that does not pass through the exposure (violation of the exchangeability assumption). Genetic colocalization analysis is a Bayesian method that can be used as a sensitivity analysis to identify this type of genetic confounding [15]. This analysis determines whether the genetic associations for any given traits share the same causal variant. We implemented the ‘coloc’ method [29], which works under the assumption that there is no more than one causal variant per trait and compares evidence for different hypotheses in a Bayesian framework. The coloc algorithm returns posterior probabilities for five hypotheses, with the posterior probability H4 corresponding to the hypothesis of a shared causal variant in the genomic region (evidence of colocalization). The posterior probabilities were calculated from the prior probabilities (we used the ‘coloc’ R package (v.5.1.0.1) [30] default settings: $p_1 = 10^{-4}$, $p_2 = 10^{-4}$, $p_{12} = 10^{-5}$). We applied an H4 posterior probability cut-off of 0.70 [31] but considered associations with lower posterior probabilities in exploratory analyses.

2.5.3. Complementary Analysis for C1-Esterase Inhibitor

We performed follow-up analyses to better understand the effect of the C1-esterase inhibitor, which was the protein with the strongest supporting evidence of an association with cognitive performance from our analyses. To assess tissue specificity of the (CSF) pQTL, we assessed the association of genetically proxied levels of the C1-esterase inhibitor in plasma and brain tissue on cognitive performance. Second, given preclinical data suggesting neuroprotective effects of the C1-esterase inhibitor, we explored the associations of genetically predicted C1-esterase inhibitor levels in CSF with Alzheimer’s disease (35,274 AD cases, and 59,163 non-AD controls) [32], and with functional outcomes 3 months after ischemic stroke (N = 6021, modified Rankin scale) [33]. Third, to identify additional disease associations across a wide range of clinical diagnoses, we performed a phenome-wide association study (PheWAS), using the lead *cis*-pQTL of CSF C1-esterase inhibitor, rs11603020, in the UK Biobank. We used the International Classification of Diseases (ICD) versions 9 and 10 [34] to identify cases for the clinical health outcomes from the Hospital Episode Statistics (HES), cancer, and death registries. Clinical diagnoses were mapped to the phecode grouping system, using the ‘PheWAS’ R package v0.99.5.5 [35]. In PheWAS, we performed a series of logistic regressions using the effect allele (C allele) of rs11603020 as the exposure and each clinical diagnosis as an outcome in up to 439,738 UK Biobank participants of European ancestry. Age, sex, and the first 10 genetic principal components were included as covariates. To maintain statistically meaningful calculations, the analysis was limited to health outcomes for which we had sufficient power (N ≥ 200 cases) [36]. A 5% FDR threshold was used to define statistically significant associations. Finally, we explored the association of rs11603020 with gene expression in 54 healthy tissue sites across nearly 1000 samples from the Genotype-Tissue Expression project (GTEx) version 9 [37].

3. Results

A total of 172 genome-wide significant *cis*-pQTLs were used to instrument 154 CSF proteins in MR analysis, of which, 139 proteins had one instrument, 12 proteins had two, and three proteins had three instruments (Supplementary Table S2). Genetically predicted levels of 14 CSF proteins were significantly associated with cognitive performance (Supplementary Table S3), after using a 5% FDR statistical significance threshold ($p < 6 \times 10^{-3}$; Figure 2). We performed colocalization analyses to assess whether these associations are due to a shared causal genetic variant rather than genetic confounding through

a neighboring variant in LD (Supplementary Table S4; Figure 3). Genetic determinants of two CSF proteins (C1-esterase inhibitor and sTie-1) colocalized with variants that determine cognitive performance, with a posterior probability for colocalization > 0.7. For sTie-1, rs839768 was identified to be the shared causal variant with cognitive performance with PPH4 = 0.94, whereas, for the C1-esterase inhibitor, rs4926 (p.Val480Met, a missense variant in *SERPING1*) was identified as the most probable shared causal variant with cognitive performance (PPH4 = 0.82). Although not meeting our prespecified threshold, variants related to LRP8 had suggestive evidence for colocalization with cognitive performance (PPH4 = 0.46) with an inverse association with cognitive performance in MR analyses (Supplementary Table S3). The lead variant used for analysis (rs12031155) was highly correlated (r^2 0.98) with rs5174, a missense variant in LRP8.

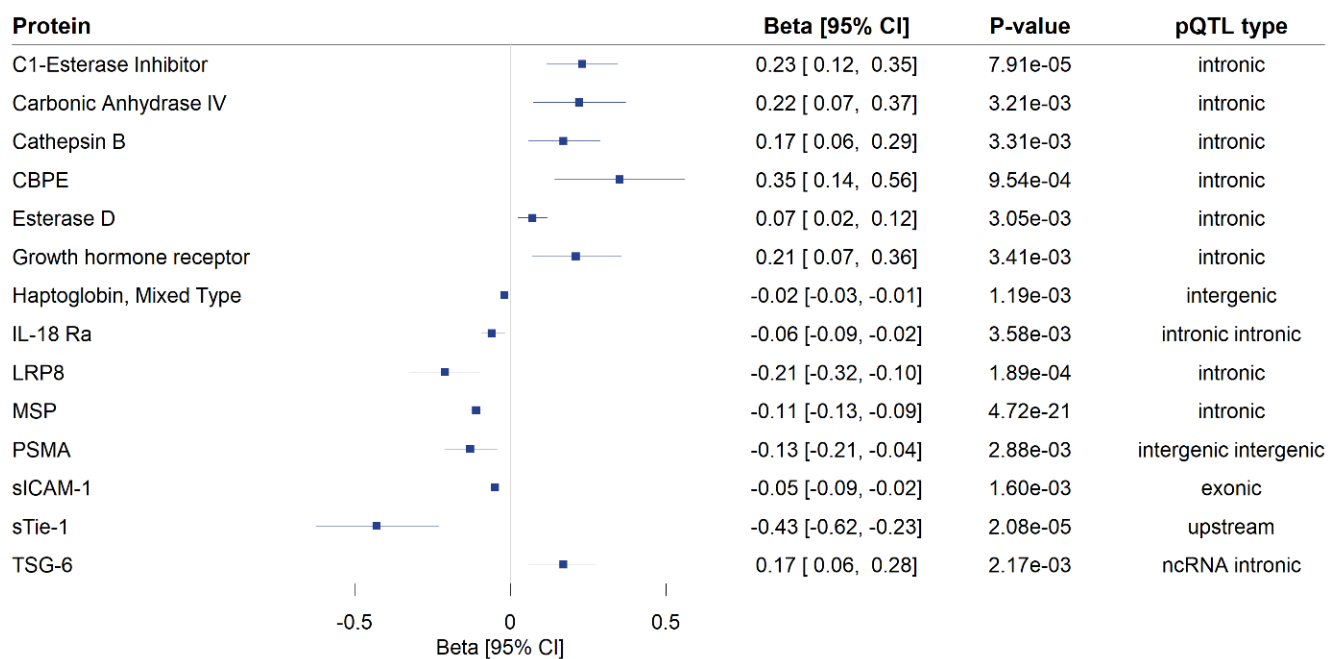


Figure 2. Mendelian randomization findings for the associations of genetically proxied cerebrospinal fluid (CSF) protein levels with cognitive performance. Only significant associations (p FDR < 0.05) are shown. MR estimates are reported as the standard deviation change in cognitive performance per 1 log-relative fluorescence unit (RFU) higher genetically predicted CSF protein level. The type of protein quantitative trait loci used as instruments in the analysis is provided in the fourth column. CI: confidence interval, pQTL: protein quantitative trait locus, ncRNA intronic: intronic non-coding RNA variants, CBPE: carboxypeptidase E, IL-18 Ra: interleukin-18 receptor 1, LRP8: low-density lipoprotein receptor-related protein 8, MSP: hepatocyte growth factor-like protein, PSMA: glutamate carboxypeptidase 2, sICAM-1: intercellular adhesion molecule 1, sTie-1: tyrosine-protein kinase receptor Tie-1, TSG-6: tumor necrosis factor-inducible gene 6 protein.

We conducted two-sample MR to explore the associations of genetically proxied CSF levels of sTie-1 and the C1-esterase inhibitor with phenotypes related to cognitive performance (Figure 4; Supplementary Table S5). Higher genetically predicted CSF C1-esterase inhibitor levels were associated with increased educational attainment, and nominally with a greater peripheral grey volume, faster reaction time, and improved verbal numeric reasoning. Higher genetically predicted levels of CSF sTie-1 were associated with a lower brain volume, lower educational attainment, and worse verbal numeric reasoning, and nominally with the lower left hippocampal volume.

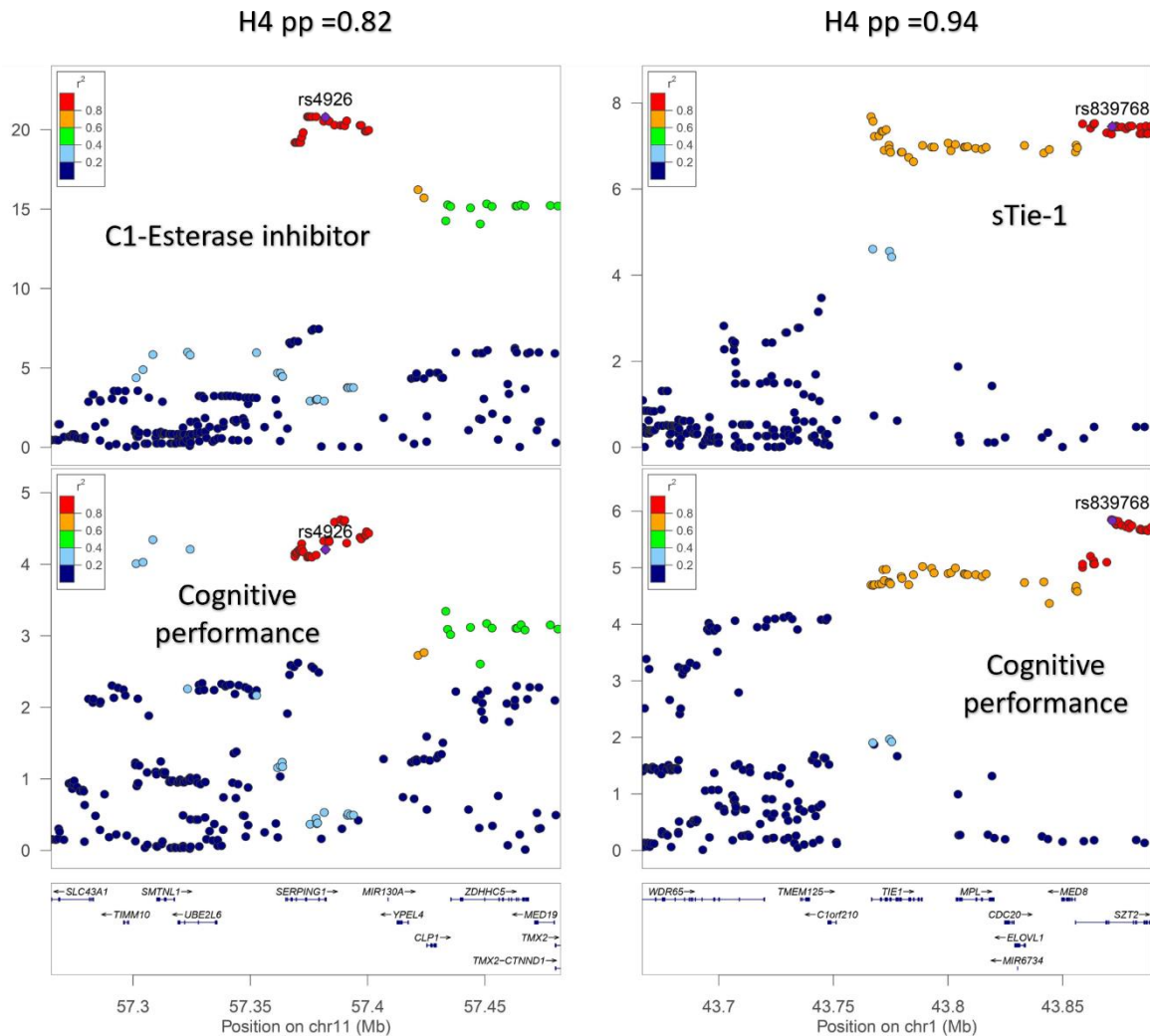


Figure 3. Regional association plots demonstrating colocalization between genetic association signals for cerebrospinal fluid (CSF) protein levels and cognitive performance. (I) Genetic associations with CSF C1-esterase inhibitor concentrations and cognitive performance ($-\log_{10}(P)$) plotted against chromosome position (megabases) for variants *cis* to *SERPING1*. (II) Genetic associations ($-\log_{10}(P)$) plotted against chromosome position (megabases) for variants within the *TIE1* gene or ± 100 kb of the gene start and end position for cerebrospinal fluid sTie-1 levels and cognitive performance. H4 pp: Posterior probability of colocalization.

In follow-up analyses, we found that 1 log-RFU higher genetically predicted plasma C1-esterase was associated with a 0.16 SD higher cognitive performance (beta: 0.16, 95% CI: 0.08 to 0.23, $p = 4.73 \times 10^{-5}$) (Supplementary Table S6). There was no evidence for an association of genetically proxied CSF C1-esterase inhibitor with the risk of AD or with functional outcomes three months after ischemic stroke (Supplementary Table S7). There were no PheWAS associations that met our statistical significance threshold. We identified a nominal association between the instrumented levels of CSF C1-esterase inhibitor and the odds of having a ‘rash and other nonspecific skin eruption’ (odds ratio [OR]: 1.13, 95% CI: 1.05 to 1.22, $p = 8.49 \times 10^{-4}$) in keeping with the known clinical syndrome of C1-esterase inhibitor deficiency (Supplementary Table S8). Finally, using gene-expression data, we found that rs10603020 influences the gene expression of *SERPING1* (the gene coding for the C1-esterase inhibitor) in whole blood ($p = 5.3 \times 10^{-8}$) (Supplementary Table S9).

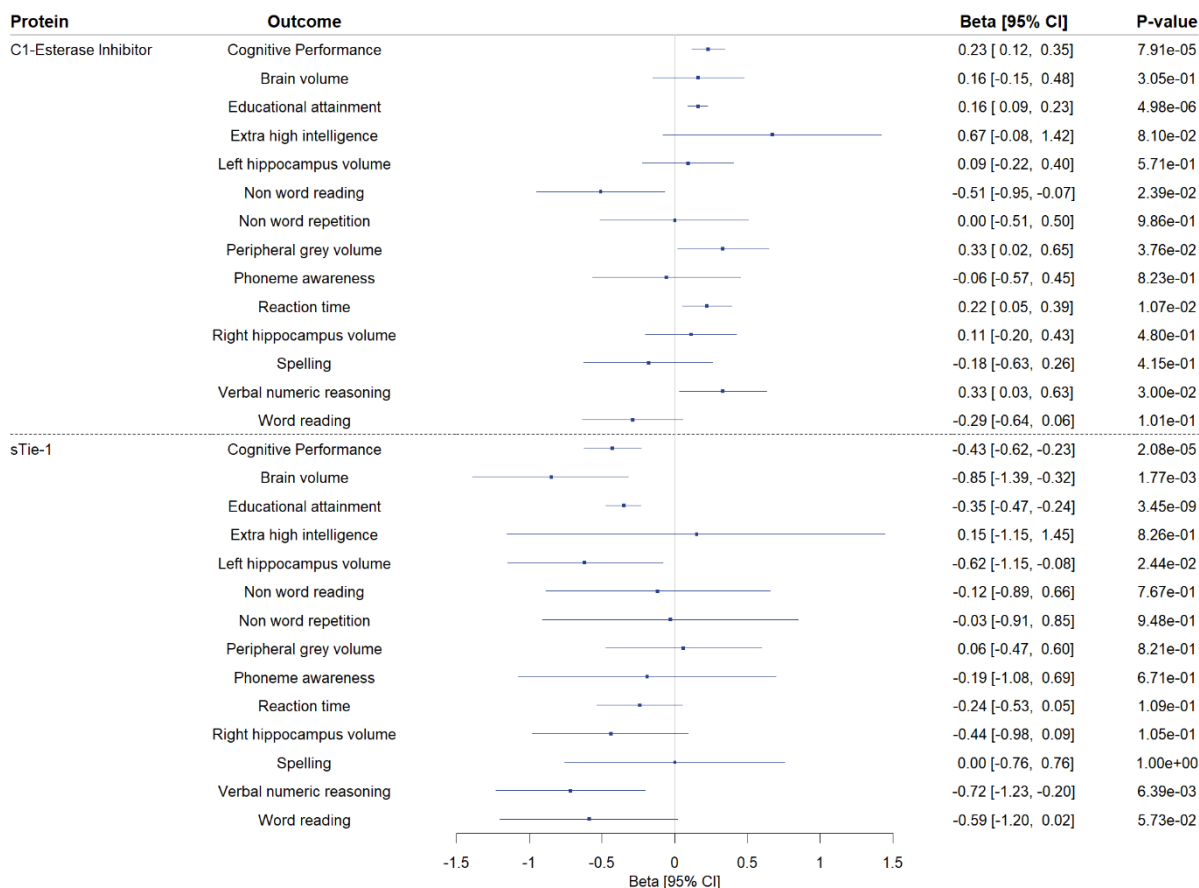


Figure 4. Secondary Mendelian randomization analyses investigating the association of genetically proxied cerebrospinal fluid (CSF) C1-esterase inhibitor and sTie-1 with phenotypes related to cognitive performance. Two-sample Mendelian randomization (MR) findings exploring the association of genetically proxied cerebrospinal fluid C1-esterase inhibitor and sTie-1 levels with cognitive performance and 13 cognitive performance-related phenotypes. MR estimates on outcomes are reported as the standard deviation change per 1 log-RFU higher genetically predicted CSF protein level, except for T1 MRI phenotypes (brain volume, left and right hippocampus volume, peripheral gray matter volume), where MR estimates are reported as mm³ change per 1 log-RFU higher genetically predicted CSF protein level, and for extremely high intelligence, where MR estimates are reported as log-odds change per 1 log-RFU higher genetically predicted CSF protein level. CI: confidence interval.

4. Discussion

We queried the effects of 154 genetically predicted CSF protein levels on cognitive performance using the MR paradigm and found multiple lines of evidence to support the C1-esterase inhibitor and sTie-1 as proteins with relevance to cognitive performance. First, genetically predicted levels of these proteins were associated with cognitive performance in MR analyses. Second, these associations were further supported by colocalization analyses supporting a shared causal variant underlying the association signal between the CSF protein level and cognitive performance. In the case of the C1-esterase inhibitor, the prioritized shared variant was a coding variant in *SERPING1*, and such variants are generally considered to provide stronger evidence for support of gene–phenotype relationships (7). Third, these associations were supported by concordant associations with related cognition and neuroimaging phenotypes.

The C1-esterase inhibitor functions as a serine protease inhibitor involved in regulating the complement system activation to prevent an excessive immune response [38]. The C1-esterase inhibitor also participates in the regulation of the kallikrein–kinin system, which is involved in vascular tone and permeability, as well as the coagulation cascade [39].

These functions underlie its relevance in hereditary angioedema. In terms of brain health, the C1-esterase inhibitor has been investigated using animal models. These predominantly mouse studies have demonstrated that knockdown or inhibition of the C1-esterase inhibitor worsens cognition and neurovascular impairment [39], and the infusion of a recombinant C1-esterase inhibitor mitigates ischemic brain injury [40] and inflammation after traumatic brain injury [41]. Taken together with results from our genetic analyses, these preclinical data further support the hypothesis that the C1-esterase inhibitor may have neuroprotective properties in humans.

In contrast to the C1-esterase inhibitor, sTie-1 is poorly characterised in the literature. This protein has been implicated in various vascular processes such as angiogenesis and vessel wall integrity [42]. These processes have relevance to neural development and cerebrovascular disease and may mediate the association of sTie-1 with cognitive performance. Moreover, the sTie-1 lead *cis*-variant rs3768046 has been reported to be associated with the risk of attention deficit hyperactivity disorder (ADHD) via the dysregulated expression of *TIE1* [43].

Although the colocalization evidence for *LRP8* did not meet our predetermined statistical threshold, it is worth discussing this protein in the context of its known relevance to human brain health. Low-density lipoprotein receptor-related protein 8 (LRP8) is also known as apolipoprotein E receptor 2. The interaction of this receptor with the reelin ligand has been extensively studied for its role in neuronal migration and positioning during cortical development in mice [44]. The rs5174 variant in *LRP8*, which is highly correlated with the pQTL instrument used in the present analysis, has been robustly associated with the GWAS of brain cortical thickness and sulcal depth [45]. Thus, it is likely that this protein is influencing cognition through pathways that are active during neurodevelopment. Moreover, other proteins with significant MR associations but a lack of colocalization evidence are worth exploring further in future studies, such as carbonic anhydrase IV, cathepsin B, or growth hormone.

There are several strengths to highlight from this study. First, we employed the MR paradigm, which is robust against potential biases resulting from confounding and reverse causation. Second, genetic colocalization analyses confirmed that our results were not explained by genetic confounding. Third, multiple sensitivity and complementary analyses yielded supporting evidence for our main conclusion. Fourth, the investigation of protein-cognition associations is relatively unexplored and has potentially wide-reaching implications for the general population. Finally, the genetic association of *LRP8* with cognition demonstrates that this approach can recapitulate known biology.

There are also limitations to consider. First, our results may be biased by pleiotropic effects of the genetic variants on biological pathways that are independent of protein levels. The second key limitation is that the present analysis does not inform us on whether interventions on the levels of these proteins may influence cognitive performance. For instance, the protein levels may only be relevant during certain time windows of neurodevelopment, such as in the case of *LRP8* which is known to influence neuronal development. In such a case, modifying protein levels in adulthood would not be expected to influence cognitive performance. Third, our findings warrant replication by leveraging gold-standard measures or different assays, since in this study a single SomaLogic aptamer assay was used to quantify relative protein abundance. In one possible source of bias in aptamer assays, missense variants that modify the amino acid sequence may consequently affect aptamer–protein binding without affecting the levels of the protein in the sample. Furthermore, aptamer-based protein quantification does not provide absolute protein quantities that are directly translatable to clinical settings. Our results should, therefore, be interpreted merely by their direction of the effect, rather than their effect size. Fourth, our analysis was restricted to individuals of European ancestry, so these findings may not accurately reflect the relationship between the CSF protein levels and cognitive function in diverse populations. Fifth, we did not identify effects of these proteins on the outcome of AD. This suggests that the direct effects of these proteins on AD are minimal, and

any possible benefit on dementia risk would be mediated through mediated effects of improved cognitive reserve. Finally, the similarity of results across various tissue types precludes us from concluding that the CSF is the compartment in which these proteins influence cognition. In the setting of a similar genetic effect on protein abundance in different compartments, we are unable to resolve which of these biological compartments is more relevant to the mechanistic relationship between these proteins and cognition. Additional work is needed to identify the tissue type that is most relevant for these proteins to influence the respective phenotypes.

In conclusion, this MR analysis prioritizes the C1-esterase inhibitor and sTie-1 as proteins with potentially causal effects on cognitive performance. The biological mechanisms relating these proteins to cognition should be further investigated to determine which could be targeted to improve cognition.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes15010071/s1>.

Author Contributions: Conceptualization, D.G. and I.D.; Methodology, L.Z. and I.D.; Formal analysis, L.Z. and M.-J.D.; Investigation, L.Z., H.T.C. and D.G.; Data curation, L.Z.; Writing—original draft, L.Z., M.-J.D. and I.D.; Writing—review & editing, L.Z., M.-J.D., H.T.C., P.E., A.D., I.T., D.G. and I.D.; Visualization, L.Z.; Supervision, D.G. and I.D.; Project administration, I.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the UK Dementia Research Institute at Imperial College, which receives its funding from UK Dementia Research Institute Ltd., funded by the UK Medical Research Council (MRC), Alzheimer’s Society, and Alzheimer’s Research UK. HTC is supported by the Novo Nordic Foundation Challenge Programme: Harnessing the Power of Big Data to Address the Societal Challenge of Aging (NNF17OC0027812). A.D. is funded by a Wellcome Trust seed award (206046/Z/17/Z).

Institutional Review Board Statement: Ethical approval and participant consent were obtained in the original studies. Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Informed Consent Statement: Patient consent was obtained from all subjects in the original GWAS studies.

Data Availability Statement: UK Biobank individual-level data used in this work can be accessed after applying for access at <https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>, accessed on 15 March 2020. All genetic data used in this work were obtained from publicly available sources. Link to datasets: Cerebrospinal fluid and plasma protein levels: <https://www.niagads.org/datasets/ng00102>, accessed on 10 February 2022; Cognitive performance and educational attainment: <https://thessgac.com/papers/3>, accessed on 10 February 2022; T1 MRI-derived phenotypes: https://open.win.ox.ac.uk/ukbiobank/big40/BIG40-IDPs_v4/IDPs.html, accessed on 10 February 2022; Reading and language-related skills phenotypes: <https://www.genlang.org/downloads.html>, accessed on 10 February 2022; Extremely high intelligence: http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST005001-GCST006000/GCST005626/, accessed on 10 February 2022; General cognitive function phenotypes: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5974083/>, accessed on 10 February 2022; Alzheimer’s disease: <https://www.niagads.org/datasets/ng00075>, accessed on 10 February 2022; Functional outcome after ischemic stroke: <https://cd.hugeamp.org/downloads.html>, accessed on 10 February 2022; Genotype–Tissue Expression data (GTEx): <https://www.gtexportal.org/home/>, accessed on 10 February 2022.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sakka, L.; Coll, G.; Chazal, J. Anatomy and physiology of cerebrospinal fluid. *Eur. Ann. Otorhinolaryngol. Head Neck Dis.* **2011**, *128*, 309–316. [[CrossRef](#)] [[PubMed](#)]
2. Spector, R.; Robert Snodgrass, S.; Johanson, C.E. A balanced view of the cerebrospinal fluid composition and functions: Focus on adult humans. *Exp. Neurol.* **2015**, *273*, 57–68. [[CrossRef](#)] [[PubMed](#)]

3. Mollenhauer, B.; Parnetti, L.; Rektorova, I.; Kramberger, M.G.; Pikkarainen, M.; Schulz-Schaeffer, W.J.; Aarsland, D.; Svenningsson, P.; Farotti, L.; Verbeek, M.M.; et al. Biological confounders for the values of cerebrospinal fluid proteins in Parkinson's disease and related disorders. *J. Neurochem.* **2016**, *139* (Suppl. S1), 290–317. [[CrossRef](#)] [[PubMed](#)]
4. Johar, I.; Mollenhauer, B.; Aarsland, D. Cerebrospinal Fluid Biomarkers of Cognitive Decline in Parkinson's Disease. *Int. Rev. Neurobiol.* **2017**, *132*, 275–294. [[PubMed](#)]
5. Blennow, K. Cerebrospinal fluid protein biomarkers for Alzheimer's disease. *NeuroRx* **2004**, *1*, 213–225. [[CrossRef](#)] [[PubMed](#)]
6. Blennow, K.; Dubois, B.; Fagan, A.M.; Lewczuk, P.; de Leon, M.J.; Hampel, H. Clinical utility of cerebrospinal fluid biomarkers in the diagnosis of early Alzheimer's disease. *Alzheimers Dement.* **2015**, *11*, 58–69. [[CrossRef](#)]
7. Daghlas, I. Mendelian randomization as a tool to inform drug development using human genetics. *Camb. Prism. Precis. Med.* **2023**, *1*, e16. [[CrossRef](#)]
8. Schindler, U.; Rush, D.K.; Fielding, S. Nootropic Drugs-Animal-Models for Studying Effects on Cognition. *Drug Dev. Res.* **1984**, *4*, 567–576. [[CrossRef](#)]
9. Olsson, B.; Portelius, E.; Cullen, N.C.; Sandelius, A.; Zetterberg, H.; Andreasson, U.; Höglund, K.; Irwin, D.J.; Grossman, M.; Weintraub, D.; et al. Association of Cerebrospinal Fluid Neurofilament Light Protein Levels with Cognition in Patients with Dementia, Motor Neuron Disease, and Movement Disorders. *JAMA Neurol.* **2019**, *76*, 318–325. [[CrossRef](#)]
10. Kvartsberg, H.; Duits, F.H.; Ingelsson, M.; Andreassen, N.; Ohrfelt, A.; Andersson, K.; Brinkmalm, G.; Lannfelt, L.; Minthon, L.; Hansson, O.; et al. Cerebrospinal fluid levels of the synaptic protein neurogranin correlates with cognitive decline in prodromal Alzheimer's disease. *Alzheimers Dement.* **2015**, *11*, 1180–1190. [[CrossRef](#)]
11. Kern, S.; Syrjanen, J.A.; Blennow, K.; Zetterberg, H.; Skoog, I.; Waern, M.; Hagen, C.E.; Van Harten, A.C.; Knopman, D.S.; Jack, C.R.; et al. Association of Cerebrospinal Fluid Neurofilament Light Protein with Risk of Mild Cognitive Impairment Among Individuals Without Cognitive Impairment. *JAMA Neurol.* **2019**, *76*, 187–193. [[CrossRef](#)] [[PubMed](#)]
12. Davies, N.M.; Holmes, M.V.; Davey Smith, G. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ* **2018**, *362*, k601. [[CrossRef](#)] [[PubMed](#)]
13. Burgess, S.; Butterworth, A.; Thompson, S.G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **2013**, *37*, 658–665. [[CrossRef](#)] [[PubMed](#)]
14. Yang, C.; Farias, F.H.G.; Ibanez, L.; Suhy, A.; Sadler, B.; Fernandez, M.V.; Wang, F.; Bradley, J.L.; Eiffert, B.; Bahena, J.A.; et al. Genomic atlas of the proteome from brain, CSF and plasma prioritizes proteins implicated in neurological disorders. *Nat. Neurosci.* **2021**, *24*, 1302–1312. [[CrossRef](#)] [[PubMed](#)]
15. Zuber, V.; Grinberg, N.F.; Gill, D.; Manipur, I.; Slob, E.A.W.; Patel, A.; Wallace, C.; Burgess, S. Combining evidence from Mendelian randomization and colocalization: Review and comparison of approaches. *Am. J. Hum. Genet.* **2022**, *109*, 767–782. [[CrossRef](#)] [[PubMed](#)]
16. Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E.N.; Carter, J.; Cunningham, V.; Dalby, A.; Eaton, B.; et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS ONE* **2010**, *5*, e15004. [[CrossRef](#)] [[PubMed](#)]
17. Lee, J.J.; Wedow, R.; Okbay, A.; Kong, E.; Maghziyan, O.; Zacher, M.; Nguyen-Viet, T.A.; Bowers, P.; Sidorenko, J.; Karlsson Linnér, R.; et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **2018**, *50*, 1112–1121. [[CrossRef](#)] [[PubMed](#)]
18. Sudlow, C.; Gallacher, J.; Allen, N.; Beral, V.; Burton, P.; Danesh, J.; Downey, P.; Elliott, P.; Green, J.; Landray, M.; et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **2015**, *12*, e1001779. [[CrossRef](#)]
19. Lencz, T.; Knowles, E.; Davies, G.; Guha, S.; Liewald, D.C.; Starr, J.M.; Djurovic, S.; Melle, I.; Sundet, K.; Christoforou, A.; et al. Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: A report from the Cognitive Genomics consortium (COGENT). *Mol. Psychiatry* **2014**, *19*, 168–174. [[CrossRef](#)]
20. Smith, S.M.; Douaud, G.; Chen, W.; Hanayik, T.; Alfaro-Almagro, F.; Sharp, K.; Elliott, L.T. An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nat. Neurosci.* **2021**, *24*, 737–745. [[CrossRef](#)]
21. Eising, E.; Mirza-Schreiber, N.; de Zeeuw, E.L.; Wang, C.A.; Truong, D.T.; Allegrini, A.G.; Shapland, C.Y.; Zhu, G.; Wigg, K.G.; Gerritse, M.L.; et al. Genome-wide analyses of individual differences in quantitatively assessed reading- and language-related skills in up to 34,000 people. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2202764119. [[CrossRef](#)] [[PubMed](#)]
22. Zabaneh, D.; Krapohl, E.; Gaspar, H.A.; Curtis, C.; Lee, S.H.; Patel, H.; Newhouse, S.; Wu, H.M.; Simpson, M.A.; Putallaz, M.; et al. A genome-wide association study for extremely high intelligence. *Mol. Psychiatry* **2018**, *23*, 1226–1232. [[CrossRef](#)] [[PubMed](#)]
23. Davies, G.; Lam, M.; Harris, S.E.; Trampush, J.W.; Luciano, M.; Hill, W.D.; Hagenaars, S.P.; Ritchie, S.J.; Marioni, R.E.; Fawns-Ritchie, C.; et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat. Commun.* **2018**, *9*, 2098. [[CrossRef](#)] [[PubMed](#)]
24. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1092 human genomes. *Nature* **2012**, *491*, 56–65. [[CrossRef](#)] [[PubMed](#)]
25. Hemani, G.; Zheng, J.; Elsworth, B.; Wade, K.H.; Haberland, V.; Baird, D.; Laurin, C.; Burgess, S.; Bowden, J.; Langdon, R.; et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **2018**, *7*, e34408. [[CrossRef](#)] [[PubMed](#)]
26. Greco, M.F.; Minelli, C.; Sheehan, N.A.; Thompson, J.R. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat. Med.* **2015**, *34*, 2926–2940. [[CrossRef](#)] [[PubMed](#)]

27. Burgess, S.; Bowden, J.; Fall, T.; Ingelsson, E.; Thompson, S.G. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology* **2017**, *28*, 30–42. [[CrossRef](#)]
28. Bowden, J.; Davey Smith, G.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **2016**, *40*, 304–314. [[CrossRef](#)]
29. Giambartolomei, C.; Vukcevic, D.; Schadt, E.E.; Franke, L.; Hingorani, A.D.; Wallace, C.; Plagnol, V. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet.* **2014**, *10*, e1004383. [[CrossRef](#)]
30. Wallace, C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet.* **2020**, *16*, e1008720. [[CrossRef](#)]
31. Thom, C.S.; Voight, B.F. Genetic colocalization atlas points to common regulatory sites and genes for hematopoietic traits and hematopoietic contributions to disease phenotypes. *BMC Med. Genom.* **2020**, *13*, 89. [[CrossRef](#)] [[PubMed](#)]
32. Kunkle, B.W.; Grenier-Boley, B.; Sims, R.; Bis, J.C.; Damotte, V.; Naj, A.C.; Boland, A.; Vronskaya, M.; Van Der Lee, S.J.; Amlie-Wolf, A.; et al. Genetic meta-analysis of diagnosed Alzheimer’s disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat. Genet.* **2019**, *51*, 414–430. [[CrossRef](#)] [[PubMed](#)]
33. Soderholm, M.; Pedersen, A.; Lorentzen, E.; Stanne, T.M.; Bevan, S.; Olsson, M.; Cole, J.W.; Fernandez-Cadenas, I.; Hankey, G.J.; Jimenez-Conde, J.; et al. Genome-wide association meta-analysis of functional outcome after ischemic stroke. *Neurology* **2019**, *92*, e1271–e1283. [[CrossRef](#)] [[PubMed](#)]
34. World Health Organization. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*; World Health Organization: Geneva, Switzerland, 1992.
35. Carroll, R.J.; Bastarache, L.; Denny, J.C. R PheWAS: Data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinformatics* **2014**, *30*, 2375–2376. [[CrossRef](#)]
36. Pendergrass, S.A.; Ritchie, M.D. Phenome-Wide Association Studies: Leveraging Comprehensive Phenotypic and Genotypic Data for Discovery. *Curr. Genet. Med. Rep.* **2015**, *3*, 92–100. [[CrossRef](#)] [[PubMed](#)]
37. Carithers, L.J.; Ardlie, K.; Barcus, M.; Branton, P.A.; Britton, A.; Buia, S.A.; Compton, C.C.; DeLuca, D.S.; Peter-Demchok, J.; Gelfand, E.T.; et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv. Biobank.* **2015**, *13*, 311–319. [[CrossRef](#)]
38. Caliezi, C.; Wuillemin, W.A.; Zeerleder, S.; Redondo, M.; Eisele, B.; Hack, C.E. C1-esterase inhibitor: An anti-inflammatory agent and its potential use in the treatment of diseases other than hereditary angioedema. *Pharmacol. Rev.* **2000**, *52*, 91–112.
39. Farfara, D.; Feierman, E.; Richards, A.; Revenko, A.S.; MacLeod, R.A.; Norris, E.H.; Strickland, S. Knockdown of circulating C1 inhibitor induces neurovascular impairment, glial cell activation, neuroinflammation, and behavioral deficits. *Glia* **2019**, *67*, 1359–1373. [[CrossRef](#)]
40. Gesuete, R.; Storini, C.; Fantin, A.; Stravalaci, M.; Zanier, E.R.; Orsini, F.; Vietsch, H.; Mannesse, M.M.; Ziere, B.; Gobbi, M.; et al. Recombinant C1 Inhibitor in Brain Ischemic Injury. *Ann. Neurol.* **2009**, *66*, 332–342. [[CrossRef](#)]
41. Longhi, L.; Perego, C.; Ortolano, F.; Zanier, E.R.; Bianchi, P.; Stocchetti, N.; McIntosh, T.K.; De Simoni, M.G. C1-inhibitor attenuates neurobehavioral deficits and reduces contusion volume after controlled cortical impact brain injury in mice. *Crit. Care Med.* **2009**, *37*, 659–665. [[CrossRef](#)]
42. Korhonen, E.A.; Lampinen, A.; Giri, H.; Anisimov, A.; Kim, M.; Allen, B.; Fang, S.; D’Amico, G.; Sipilä, T.J.; Lohela, M.; et al. Tie1 controls angiopoietin function in vascular remodeling and inflammation. *J. Clin. Investig.* **2016**, *126*, 3495–3510. [[CrossRef](#)] [[PubMed](#)]
43. Chen, X.; Yao, T.; Cai, J.; Zhang, Q.; Li, S.; Li, H.; Fu, X.; Wu, J. A novel cis-regulatory variant modulating TIE1 expression associated with attention deficit hyperactivity disorder in Han Chinese children. *J. Affect. Disord.* **2022**, *300*, 179–188. [[CrossRef](#)] [[PubMed](#)]
44. Honda, T.; Kobayashi, K.; Mikoshiba, K.; Nakajima, K. Regulation of Cortical Neuron Migration by the Reelin Signaling Pathway. *Neurochem. Res.* **2011**, *36*, 1270–1279. [[CrossRef](#)] [[PubMed](#)]
45. van der Meer, D.; Kaufmann, T.; Shadrin, A.A.; Makowski, C.; Frei, O.; Roelfs, D.; Monereo-Sánchez, J.; Linden, D.E.; Rokicki, J.; Alnæs, D.; et al. The genetic architecture of human cortical folding. *Sci. Adv.* **2021**, *7*, eabj9446. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.