Genetic correlation between amyotrophic lateral sclerosis and schizophrenia

C. Robert Cloninger
Washington University School of Medicine in St. Louis
Dragan M. Svrakic
Washington University School of Medicine in St. Louis
et al

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation
https://digitalcommons.wustl.edu/open_access_pubs/5803

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Genetic correlation between amyotrophic lateral sclerosis and schizophrenia


We have previously shown higher-than-expected rates of schizophrenia in relatives of patients with amyotrophic lateral sclerosis (ALS), suggesting an aetiological relationship between the diseases. Here, we investigate the genetic relationship between ALS and schizophrenia using genome-wide association study data from over 100,000 unique individuals. Using linkage disequilibrium score regression, we estimate the genetic correlation between ALS and schizophrenia to be 14.3% (7.05–21.6; \(P = 1 \times 10^{-4}\)) with schizophrenia polygenic risk scores explaining up to 0.12% of the variance in ALS (\(P = 8.4 \times 10^{-7}\)). A modest increase in comorbidity of ALS and schizophrenia is expected given these findings (odds ratio 1.08–1.26) but this would require very large studies to observe epidemiologically.

We identify five potential novel ALS-associated loci using conditional false discovery rate analysis. It is likely that shared neurobiological mechanisms between these two disorders will engender novel hypotheses in future preclinical and clinical studies.
A myotrophic lateral sclerosis (ALS) is a late-onset neurodegenerative condition characterized by progressive loss of upper and lower motor neurons, leading to death from respiratory failure in 70% of patients within 3 years of symptom onset. Although ALS is often described as a primarily motor-system disease, extramotor involvement occurs in up to 50% of cases, with prominent executive and behavioural impairment, and behavioural variant frontotemporal dementia (FTD) in up to 14% of cases. A neuropsychiatric prodrome has been described in some people with ALS–FTD, and higher rates of schizophrenia and suicide have been reported in first and second degree relatives of those with ALS, particularly in kindreds associated with the C9orf72 hexanucleotide repeat expansion. These clinical and epidemiological observations suggest that ALS and schizophrenia may share heritability.

ALS and schizophrenia both have high heritability estimates (0.65 and 0.64, respectively)\(^5\); however the underlying genetic architectures of these heritable components appear to differ. Analysis of large genome-wide association study (GWAS) datasets has implicated over 100 independent risk loci for schizophrenia\(^3\) and estimated that a substantial proportion (23%) of the variance in underlying liability for schizophrenia is due to additive polygenic risk (many risk-increasing alleles of low individual effect combining to cause disease) conferred by common genetic variants.\(^6\) This proportion, the single nucleotide polymorphism (SNP)-based heritability, is lower in ALS (8.2%), in which fewer than ten risk loci have been identified by GWAS.\(^7\) Nevertheless, both diseases have polygenic components, but the extent to which they overlap has not been investigated.

Recently, methods to investigate overlap between polygenic traits using GWAS data have been developed.\(^8\)–\(^10\) These methods assess either pleiotropy (identical genetic variants influencing both traits) or genetic correlation (identical alleles influencing both traits). Genetic correlation is related to heritability; for both measures, binary traits such as ALS and schizophrenia are typically modelled as extremes of an underlying continuous scale of liability to develop the trait. If two binary traits are genetically correlated, their liabilities covary, and this covariance is determined by both traits having identical risk alleles at overlapping risk loci. Studies of pleiotropy and genetic correlation have provided insights into the overlapping genetics of numerous traits and disorders, although none to date has implicated shared polygenic risk between neurodegenerative and neuropsychiatric disease. Here, we apply several techniques to identify and dissect the polygenic overlap between ALS and schizophrenia. We provide evidence for genetic correlation between the two disorders which is unlikely to be driven by diagnostic misclassification and we demonstrate a lack of polygenic overlap between ALS and other neuropsychiatric and neurological conditions, which could be due to limited power given the smaller cohort sizes for these studies.

Results
Genetic correlation between ALS and schizophrenia. To investigate the polygenic overlap between ALS and schizophrenia, we used individual-level and summary data from GWAS for ALS (36,052 individuals) and schizophrenia\(^3\) (79,845 individuals). At least 5,582 control individuals were common to both datasets, but for some cohorts included in the schizophrenia dataset this could not be ascertained so this number is likely to be higher. For ALS, we used summary data from both mixed linear model association testing\(^1\) and meta-analysis of cohort-level logistic regression\(^2\). We first used linkage disequilibrium (LD) score regression with ALS and schizophrenia summary statistics; this technique models, for polygenic traits, a linear relationship between a SNP’s LD score (the amount of genetic variation that it captures) and its GWAS test statistic.\(^13\) This distinguishes confounding from polygenicity in GWAS inflation and the regression coefficient can be used to estimate the SNP-based heritability \((h^2_s)\) for single traits\(^11\). In the bivariate case, the regression coefficient estimates genetic covariance \((\rho_p)\) for pairs of traits, from which genetic correlation \((r_g)\) is estimated; these estimates are unaffected by sample overlap between traits. Using constrained intercept LD score regression with mixed linear model ALS summary statistics, we estimated the liability-scale SNP-based heritability of ALS to be 8.2% (95% confidence interval \(= 7.2–9.1\); mean \(\chi^2 = 1.13\); all ranges reported below indicate 95% confidence intervals), replicating previous estimates based on alternative methods.\(^7\) Estimates based on ALS meta-analysis summary statistics and free-intercept LD score regression with mixed linear model summary statistics were lower (Supplementary Table 1), resulting in higher genetic correlation estimates (Supplementary Table 2); for this reason, we conservatively use constrained intercept genetic correlation estimates for ALS mixed linear model summary statistics throughout the remainder of this paper. Heritability estimates for permuted ALS data were null (Supplementary Table 1).

LD score regression estimated the genetic correlation between ALS and schizophrenia to be 14.3% \((7.05–21.6\); \(P = 1 \times 10^{-4}\)). Results were similar for a smaller schizophrenia cohort of European ancestry (21,856 individuals)\(^14\), indicating that the inclusion of individuals of Asian ancestry in the schizophrenia cohort did not bias this result (Supplementary Fig. 1). In addition to schizophrenia, we estimated genetic correlation with ALS using GWAS summary statistics for bipolar disorder,\(^15\) major depressive disorder,\(^16\) attention deficit-hyperactivity disorder,\(^17\) autism spectrum disorder,\(^17\) Alzheimer’s disease (Supplementary Note 1),\(^18\) multiple sclerosis\(^19\) and adult height,\(^20\) finding no significant genetic correlation between ALS and any secondary trait other than schizophrenia (Fig. 1; Supplementary Table 2).

Polygenic risk score analysis. We supported the positive genetic correlation between ALS and schizophrenia by analysis of
polygenic risk for schizophrenia in the ALS cohort. Polygenic risk scores (PRS) are per-individual scores based on the sum of alleles associated with one phenotype, weighted by their effect size, measured in an independent target sample of the same or a different phenotype\(^{10}\). PRS calculated on schizophrenia GWAS summary statistics for twelve \(P\)-value thresholds (\(P_T\)) explained up to 0.12\% \((P_T = 0.2, \ P = 8.4 \times 10^{-7}\) of the phenotypic variance in a subset of the individual-level ALS genotype data that had all individuals removed that were known or suspected to be present in the schizophrenia cohort (Fig. 2; Supplementary Table 5). ALS cases had on average higher PRS for schizophrenia compared to healthy controls and harbouring a high schizophrenia PRS for \(P_T = 0.2\) significantly increased the odds of being an ALS patient in our cohort (Fig. 3; Supplementary Table 7), which we consider to be too high to be true genetic correlation of 0%, we estimated the required population-based incident misdiagnosis of ALS as schizophrenia to be 4.86\% (2.47–7.13) to perspective, to achieve 80\% power to detect a significant \((\alpha = 0.05)\) excess of schizophrenia in the ALS cohort as a result of this genetic correlation, the required population-based incident cohort size is 16,448 ALS patients (7,310–66,670).

Modelling misdiagnosis and comorbidity. Using BUHMBOX\(^{21}\), a tool that distinguishes true genetic relationships between diseases (pleiotropy) from spurious relationships resulting from heterogeneous mixing of disease cohorts, we determined that misdiagnosed cases in the schizophrenia cohort (for example, young-onset FTD–ALS) did not drive the genetic correlation estimate between ALS and schizophrenia \((P = 0.94)\). Assuming a true genetic correlation of 0\%, we estimated the required rate of misdiagnosis of ALS as schizophrenia to be 4.86\% (2.47–7.13) to obtain the genetic correlation estimate of 14.3\% (7.05–21.6; Supplementary Table 7), which we consider to be too high to be likely. However, if ALS and schizophrenia are genetically correlated, more comorbidity would be expected than if the genetic correlation was 0\%. Modelling our observed genetic correlation of 14.3\% (7.05–21.6), we estimated the odds ratio for having above-threshold liability for ALS given above-threshold liability for schizophrenia to be 1.17 (1.08–1.26), and the same for schizophrenia given ALS (Supplementary Fig. 4). From a clinical

![Figure 2](https://example.com/figure2.png) **Figure 2** | Analysis of PRS for schizophrenia in a target sample of 10,032 ALS cases and 16,627 healthy controls. P-value thresholds \((P_T)\) for schizophrenia SNPs are shown on the x axis, where the number of SNPs increases with a more lenient \(P_T\). Explained variances (Nagelkerke \(R^2\), shown as a \%) of a generalized linear model including schizophrenia-based PRS versus a baseline model without polygenic scores (blue bars) are shown for each \(P_T\). \(-\log_{10}(P\text{-value})\) represents \(P\)-values from the binomial logistic regression of ALS phenotype on PRS, accounting for LD (Supplementary Table 4) and including sex and significant principal components as covariates (Supplementary Fig. 2). Values are provided in Supplementary Table 5.

![Figure 3](https://example.com/figure3.png) **Figure 3** | Odds ratio for ALS by PRS deciles for schizophrenia. The figure applies to schizophrenia \(P\)-value threshold \((P_T) = 0.2\). The \(P\)-value threshold were converted to ten deciles containing near identical numbers of individuals. Decile 1 contained the lowest scores and decile 10 contained the highest scores, where decile 1 was the reference and decimals 2-10 were dummy variables to contrast to decile 1 for OR calculation. The case:control ratio per decile is indicated with grey bars. Error bars indicate 95\% confidence intervals. Significant differences from decile 1 were determined by logistic regression of ALS phenotype on PRS decile, including sex and principal components as covariates and are indicated by *\(P<0.05\) or ***\(P<0.001\).
Pleiotropic risk loci. We leveraged the genetic correlation between ALS and schizophrenia to discover novel ALS-associated genomic loci by conditional false discovery rate (cFDR) analysis\(^{22}\) (Fig. 4; Supplementary Table 8). Five loci already known to be involved in ALS were identified (corresponding to MOBP, C9orf72, TBK1, SARM1 and UNC13A) along with five potential novel loci at cFDR < 0.01 (CNTN6, TNIP1, PPP2R2D, NCKAP5L and ZNF295-AS1). No gene set was significantly enriched (after Bonferroni correction) in genome-wide cFDR values when analysed using MAGENTA.

Discussion

There is evolving clinical, epidemiological and biological evidence for an association between ALS and psychotic illness, particularly schizophrenia. Genetic evidence of overlap to date has been based primarily on individual genes showing Mendelian inheritance, in particular the C9orf72 hexanucleotide repeat expansion, which is associated with ALS and FTD, and with psychosis in relatives of ALS patients\(^{4}\). In this study, we have replicated SNP-based heritability estimates for ALS and schizophrenia using GWAS summary statistics, and have for the first time demonstrated significant overlap between the polygenic components of both diseases, estimating the genetic correlation to be 14.3%. We have carefully controlled for confounding bias, including population stratification and shared control samples, and have shown through analysis of polygenic risk scores that the overlapping polygenic risk applies to SNPs that are modestly associated with both diseases. Given that our genetic correlation estimate relates to the polygenic components of ALS (\(h^2 = 8.2%\)) and schizophrenia (\(h^2 = 23%\)) and these estimates do not represent all heritability for both diseases, the accuracy of using schizophrenia polygenic risk applies to SNPs that are modestly associated through analysis of polygenic risk scores that the overlapping stratification and shared control samples, and have shown significant overlap between the polygenic components of both diseases, suggesting convergent biological mechanisms between the two diseases.

Although phenotypically heterogeneous, both ALS and schizophrenia are clinically recognizable as syndromes\(^{23,24}\). The common biological mechanisms underlying the association between the two conditions are not well understood, but are likely associated with disruption of cortical networks. Schizophrenia is a polygenic neurodevelopmental disorder characterized by a combination of positive symptoms (hallucinations and delusions), negative symptoms (diminished motivation, blunted affect, reduction in spontaneous speech and poor social functioning) and impairment over a broad range of cognitive abilities\(^{25}\). ALS is a late onset complex genetic disease characterized by a predominantly motor phenotype with recently recognized extra-motor features in 50% of patients, including cognitive impairment\(^{5}\). It has been suggested that the functional effects of risk genes in schizophrenia converge by modulating synaptic plasticity, and influencing the development and stabilization of cortical microcircuitry\(^{26}\). In this context, our identification of CNTN6 (contactin 6, also known as NB-3, a neural adhesion protein important in axon development)\(^{26}\) as a novel pleiotropy-informed ALS-associated locus supports neural network dysregulation as a potential convergent mechanism of disease in ALS and schizophrenia.

No significantly enriched biological pathway or ontological term was identified within genome-wide cFDR values using MAGENTA. Low inflation in ALS GWAS statistics, coupled with a rare variant genetic architecture\(^{4}\), render enrichment-based biological pathway analyses with current sample sizes challenging. Nevertheless, nine further loci were associated with ALS risk at cFDR < 0.01. Of these, MOBP, C9orf72, TBK1, SARM1 and UNC13A have been described previously in ALS and were associated by cFDR analysis in this study owing to their strong association with ALS through GWAS\(^7\). The remaining four loci (TNIP1, PPP2R2D, NCKAP5L and ZNF295-AS1) are novel associations and may represent pleiotropic disease loci. TNIP1 encodes TNFAIP3 interacting protein 1 and is involved in autoimmunity and tissue homeostasis\(^{27}\). The protein product of PPP2R2D is a regulatory subunit of protein phosphatase 2 and has a role in PI3K-Akt signalling and mitosis\(^{28}\). NCKAP5L is a homologue of NCKAP5, encoding NAP5, a proline-rich protein that has previously been implicated in schizophrenia, bipolar disorder and autism\(^{29,30}\). ZNF295-AS1 is a noncoding RNA\(^{31}\). Further investigation into the biological roles of these genes may reveal novel insight into the pathophysiology of certain subtypes of ALS and schizophrenia, and as whole-genome and exome datasets become available in the future for appropriately large ALS case–control cohorts, testing for burden of rare genetic variation across these genes will be particularly instructive, especially given the role that rare variants appear to play in the pathophysiology of ALS\(^7\).

![Figure 4 | Pleiotropy-informed ALS risk loci determined by analysis of cFDR in ALS GWAS P-values given schizophrenia GWAS P-values (cFDR\textsubscript{ALS<0.05}). Each point denotes a SNP; its x axis position corresponds to its chromosomal location and its height indicates the extent of association with ALS by cFDR analysis. The solid line indicates the threshold cFDR = 0.01. Any gene whose role in ALS is already established is in bold. A complete list of all loci at cFDR ≤ 0.05 is provided in Supplementary Table 8.](image)
Our data suggest that other neuropsychiatric conditions (bipolar disorder, autism and major depression) do not share polygenic risk with ALS. This finding contrasts with our recent observations from family aggregation studies and may be unexpected given the extensive genetic correlation between neuropsychiatric conditions. This could relate to statistical power conferred by secondary phenotype cohort sizes, and future studies with larger sample sizes will shed further light on the relationship between ALS and neuropsychiatric disease. It is also possible that the current study underestimates genetic correlations due to the substantial role that rare variants play in the relationship between ALS and neuropsychiatric disease.

A potential criticism of this study is that the polygenic overlap between ALS and schizophrenia could be driven by misdiagnosis, particularly in cases of ALS–FTD, which can present in later life as a psychotic illness and could be misdiagnosed as schizophrenia. This is unlikely, as strict diagnostic criteria are required for inclusion in the schizophrenia GWAS dataset. Furthermore, since core schizophrenia symptoms are usually diagnosed during late adolescence, a misdiagnosis of FTD-onset ALS–FTD as schizophrenia is unlikely. In this study, we found no evidence for misdiagnosis of ALS as schizophrenia (BUHMBOX \(P = 0.94\)) and we estimated that a misdiagnosis of 4.86% of ALS cases would be required to spuriously observe a genetic correlation of 14.3%, which is not likely to occur in clinical practice. We are therefore confident that this genetic correlation estimate reflects a genuine polygenic overlap between the two diseases and is not a feature of cohort ascertainment, but the possibility of some misdiagnosis in either cohort cannot be entirely excluded based on available data.

A positive genetic correlation between ALS and schizophrenia predicts an excess of patients presenting with both diseases. Most neurologists and psychiatrists, however, will not readily acknowledge that these conditions co-occur frequently. Our genetic correlation estimate confers an odds ratio of 1.17 (1.08–1.26) for harbouring above-threshold liability for ALS given schizophrenia (or vice versa) and a lifetime risk of 1:34,336 for both phenotypes together. Thus, a very large incident cohort of 16,448 ALS patients (7,310–66,670), with detailed phenotype and neuropsychiatric disease. The presence of both apparent monogenic C9orf72-driven overlap and polygenic overlap in the aetiology of ALS and schizophrenia suggests the presence of common biological processes, which may relate to disruption of cortical circuitry. As both ALS and schizophrenia are heterogeneous conditions, further genomic, biological and clinical studies are likely to yield novel insights into the pathophysiological processes for both diseases and will provide clinical sub-stratification parameters that could drive novel drug development for both neurodegenerative and psychiatric conditions.

Methods

Study population and genetic data. For ALS, 7,740,343 SNPs genotyped in 12,577 ALS patients and 23,475 healthy controls of European ancestry organized in 27 population- and country-defined strata were used. The schizophrenia dataset comprised GWAS summary statistics for 9,444,230 SNPs originally genotyped in 34,241 patients and 45,604 controls of European and Asian ancestry. For LD score regression, GWAS summary statistics were generated for the ALS cohort using mixed linear model association testing implemented in Genome-wide Complex Trait Analysis11 or logistic regression combined with cross-stratum meta-analysis using METAL12. To evaluate sample overlap for PRS and cFDR analyses, we also obtained individual-level genotype data for 27,647 schizophrenia cases and 33,675 controls from the schizophrenia GWAS (Psychiatric Genomics Consortium) and dbGaP accession number phs000021.v3.p2). Using 88,971 LD-pruned (window size 200 SNPs; shift 20 SNPs; \(r^2 > 0.25\)) SNPs in both datasets, we restricted LD score regression to SNPs included in both the GWAS summary statistics and HapMap phase 3; for \(r^2\) estimation in pairs of traits this was the intersection of SNPs for both traits and HapMap. Because population structure and confounding were highly controlled in the ALS summary statistics by the use of mixed linear model association tests, we constrained the LD score regression intercept to 1 for \(h^2\) estimation in ALS, and we also estimated \(h^2\) with a free intercept. For \(h^2\) estimation in ALS traits and \(r^2\) estimation we used a free parameter. We also estimated \(r^2\) using ALS meta-analysis results with free and constrained intercepts and with permuted data conserving population structure. Briefly, principal component analysis was carried out for each stratum using smartpc42 and the three-dimensional space defined by principal components 1-3 was equally subdivided into 1,000 cubes. Within each cube, case–control labels were randomly swapped and association statistics were re-calculated for the entire stratum using logistic regression. Study-level \(P\)-values were then calculated using inverse variance weighted fixed effect meta-analysis implemented in METAL13. \(h^2\) was estimated for these meta-analysed permuted data using LD score regression (Supplementary Table 1).

Polygenic risk score analysis. We calculated PRS for 10,032 cases and 16,627 healthy controls in the ALS dataset (duplicate and suspected or confirmed replication samples with the schizophrenia dataset removed). Summary statistics for schizophrenia-associated alleles and effect sizes reported in the GWAS summary statistics for 6,843,674 SNPs included in both studies and in the phase 1 integrated release of the 1,000 Genomes Project23. For regression weights13, we restricted LD score calculation to SNPs included in both the GWAS summary statistics and HapMap phase 3; for \(r^2\) estimation in pairs of traits this was the intersection of SNPs for both traits and HapMap. Because population structure and confounding were highly controlled in the ALS summary statistics by the use of mixed linear model association tests, we constrained the LD score regression intercept to 1 for \(h^2\) estimation in ALS, and we also estimated \(h^2\) with a free intercept. For \(h^2\) estimation in ALS traits and \(r^2\) estimation we used a free parameter. We also estimated \(r^2\) using ALS meta-analysis results with free and constrained intercepts and with permuted data conserving population structure. Briefly, principal component analysis was carried out for each stratum using smartpc42 and the three-dimensional space defined by principal components 1-3 was equally subdivided into 1,000 cubes. Within each cube, case–control labels were randomly swapped and association statistics were re-calculated for the entire stratum using logistic regression. Study-level \(P\)-values were then calculated using inverse variance weighted fixed effect meta-analysis implemented in METAL13. \(h^2\) was estimated for these meta-analysed permuted data using LD score regression (Supplementary Table 1).
To estimate explained variance of PRS on the phenotype, a baseline linear relationship including only sex and significant PCs as variables was modelled first:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n}, \]

where \( y \) is the phenotype in the ALS dataset, \( x \) is the intercept of the model with a slope \( \beta \) for each variable \( x \).

Subsequently, a linear model including polygenic scores for each schizophrenia \( P_t \) was calculated:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n} + \beta_{pc_t}x_{pc_t}. \]

A Nagelkerke \( R^2 \) value was obtained for every model and the baseline Nagelkerke \( R^2 \) value was subtracted, resulting in a \( \Delta \) explained variance that describes the contribution of schizophrenia-based PRS to the phenotype in the ALS dataset. PRS analysis was also performed in permuted case-control data (1,000 permutations, conserving case–control ratio) to assess whether the increased \( \Delta \) explained variance was a true signal associated with phenotype. \( \Delta \) explained variances and \( P \)-values were averaged across permutation analyses.

To ensure we did not over- or under-correct for population effects in our model, we tested the inclusion of up to a total of 30 PCs in the model, starting with the PC with the most significant effect on the ALS phenotype (Supplementary Fig. 2). Increasing the number of PCs initially had a large effect on the \( \Delta \) explained variance, but this effect levelled out after 11 PCs. On the basis of this test we are confident that adding the 11 PCs that had a significant effect on the phenotype sufficiently accounted for potential PC-related confounding between ALS and schizophrenia. For the schizophrenia \( P_t \) for which we obtained the highest \( \Delta \) explained variance (0.2), we subdivided observed schizophrenia-based PRS in the ALS cohort into deciles and calculated the odds ratio for being an ALS case in each decile compared to the first decile using a similar generalized linear model:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n} + \beta_{pc_k}x_{pc_k}. \]

Odds ratios and 95% confidence intervals for ALS were derived by calculating the exponential function of the beta estimate of the model for each of the deciles 2–10.

**Diagnostic misclassification.** To distinguish the contribution of misdiagnosis from true genetic pleiotropy we used BUMBOX21 with 417 independent ALS risk alleles in a sample of 27,647 schizophrenia patients for which individual-level genotype data were available. We also estimated the required misdiagnosis rate \( M \) of FTD–ALS as schizophrenia that would lead to the observed genetic correlation using the binomial power equation.

To estimate explained variance of PRS on the phenotype, a baseline linear relationship including only sex and significant PCs as variables was modelled first:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n}, \]

where \( y \) is the phenotype in the ALS dataset, \( x \) is the intercept of the model with a slope \( \beta \) for each variable \( x \).

Subsequently, a linear model including polygenic scores for each schizophrenia \( P_t \) was calculated:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n} + \beta_{pc_t}x_{pc_t}. \]

A Nagelkerke \( R^2 \) value was obtained for every model and the baseline Nagelkerke \( R^2 \) value was subtracted, resulting in a \( \Delta \) explained variance that describes the contribution of schizophrenia-based PRS to the phenotype in the ALS dataset. PRS analysis was also performed in permuted case-control data (1,000 permutations, conserving case–control ratio) to assess whether the increased \( \Delta \) explained variance was a true signal associated with phenotype. \( \Delta \) explained variances and \( P \)-values were averaged across permutation analyses.

To ensure we did not over- or under-correct for population effects in our model, we tested the inclusion of up to a total of 30 PCs in the model, starting with the PC with the most significant effect on the ALS phenotype (Supplementary Fig. 2). Increasing the number of PCs initially had a large effect on the \( \Delta \) explained variance, but this effect levelled out after 11 PCs. On the basis of this test we are confident that adding the 11 PCs that had a significant effect on the phenotype sufficiently accounted for potential PC-related confounding between ALS and schizophrenia. For the schizophrenia \( P_t \) for which we obtained the highest \( \Delta \) explained variance (0.2), we subdivided observed schizophrenia-based PRS in the ALS cohort into deciles and calculated the odds ratio for being an ALS case in each decile compared to the first decile using a similar generalized linear model:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n} + \beta_{pc_k}x_{pc_k}. \]

Odds ratios and 95% confidence intervals for ALS were derived by calculating the exponential function of the beta estimate of the model for each of the deciles 2–10.

**Diagnostic misclassification.** To distinguish the contribution of misdiagnosis from true genetic pleiotropy we used BUMBOX21 with 417 independent ALS risk alleles in a sample of 27,647 schizophrenia patients for which individual-level genotype data were available. We also estimated the required misdiagnosis rate \( M \) of FTD–ALS as schizophrenia that would lead to the observed genetic correlation using the binomial power equation.

To estimate explained variance of PRS on the phenotype, a baseline linear relationship including only sex and significant PCs as variables was modelled first:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n}, \]

where \( y \) is the phenotype in the ALS dataset, \( x \) is the intercept of the model with a slope \( \beta \) for each variable \( x \).

Subsequently, a linear model including polygenic scores for each schizophrenia \( P_t \) was calculated:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n} + \beta_{pc_t}x_{pc_t}. \]

A Nagelkerke \( R^2 \) value was obtained for every model and the baseline Nagelkerke \( R^2 \) value was subtracted, resulting in a \( \Delta \) explained variance that describes the contribution of schizophrenia-based PRS to the phenotype in the ALS dataset. PRS analysis was also performed in permuted case-control data (1,000 permutations, conserving case–control ratio) to assess whether the increased \( \Delta \) explained variance was a true signal associated with phenotype. \( \Delta \) explained variances and \( P \)-values were averaged across permutation analyses.

To ensure we did not over- or under-correct for population effects in our model, we tested the inclusion of up to a total of 30 PCs in the model, starting with the PC with the most significant effect on the ALS phenotype (Supplementary Fig. 2). Increasing the number of PCs initially had a large effect on the \( \Delta \) explained variance, but this effect levelled out after 11 PCs. On the basis of this test we are confident that adding the 11 PCs that had a significant effect on the phenotype sufficiently accounted for potential PC-related confounding between ALS and schizophrenia. For the schizophrenia \( P_t \) for which we obtained the highest \( \Delta \) explained variance (0.2), we subdivided observed schizophrenia-based PRS in the ALS cohort into deciles and calculated the odds ratio for being an ALS case in each decile compared to the first decile using a similar generalized linear model:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n} + \beta_{pc_k}x_{pc_k}. \]

Odds ratios and 95% confidence intervals for ALS were derived by calculating the exponential function of the beta estimate of the model for each of the deciles 2–10.

**Diagnostic misclassification.** To distinguish the contribution of misdiagnosis from true genetic pleiotropy we used BUMBOX21 with 417 independent ALS risk alleles in a sample of 27,647 schizophrenia patients for which individual-level genotype data were available. We also estimated the required misdiagnosis rate \( M \) of FTD–ALS as schizophrenia that would lead to the observed genetic correlation using the binomial power equation.

To estimate explained variance of PRS on the phenotype, a baseline linear relationship including only sex and significant PCs as variables was modelled first:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n}, \]

where \( y \) is the phenotype in the ALS dataset, \( x \) is the intercept of the model with a slope \( \beta \) for each variable \( x \).

Subsequently, a linear model including polygenic scores for each schizophrenia \( P_t \) was calculated:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n} + \beta_{pc_t}x_{pc_t}. \]

A Nagelkerke \( R^2 \) value was obtained for every model and the baseline Nagelkerke \( R^2 \) value was subtracted, resulting in a \( \Delta \) explained variance that describes the contribution of schizophrenia-based PRS to the phenotype in the ALS dataset. PRS analysis was also performed in permuted case-control data (1,000 permutations, conserving case–control ratio) to assess whether the increased \( \Delta \) explained variance was a true signal associated with phenotype. \( \Delta \) explained variances and \( P \)-values were averaged across permutation analyses.

To ensure we did not over- or under-correct for population effects in our model, we tested the inclusion of up to a total of 30 PCs in the model, starting with the PC with the most significant effect on the ALS phenotype (Supplementary Fig. 2). Increasing the number of PCs initially had a large effect on the \( \Delta \) explained variance, but this effect levelled out after 11 PCs. On the basis of this test we are confident that adding the 11 PCs that had a significant effect on the phenotype sufficiently accounted for potential PC-related confounding between ALS and schizophrenia. For the schizophrenia \( P_t \) for which we obtained the highest \( \Delta \) explained variance (0.2), we subdivided observed schizophrenia-based PRS in the ALS cohort into deciles and calculated the odds ratio for being an ALS case in each decile compared to the first decile using a similar generalized linear model:

\[ y = \beta_0 + \beta_{sex}x_{sex} + \sum_{n} \beta_{pc_n}x_{pc_n} + \beta_{pc_k}x_{pc_k}. \]

Odds ratios and 95% confidence intervals for ALS were derived by calculating the exponential function of the beta estimate of the model for each of the deciles 2–10.

**Diagnostic misclassification.** To distinguish the contribution of misdiagnosis from true genetic pleiotropy we used BUMBOX21 with 417 independent ALS risk alleles in a sample of 27,647 schizophrenia patients for which individual-level genotype data were available. We also estimated the required misdiagnosis rate \( M \) of FTD–ALS as schizophrenia that would lead to the observed genetic correlation using the binomial power equation.


Acknowledgements

We acknowledge helpful contributions from Mr Gert van de Vendel in the design and execution of PRS analyses. This study received support from the ALS Association; Fondation Thierry Latran; the Motor Neurone Disease Association of England, Wales and Northern Ireland; Science Foundation Ireland; Health Research Board (Ireland); The Netherlands ALS Foundation (Project MInE, to J.H.V., L.H.v.d.B.); the Netherlands Organisation for Health Research and Development (Vici scheme, L.H.v.d.B.) and ZonMW under the frame of E-Rare-2, the ERA Net for Research on Rare Diseases (PYRAMID). Research leading to these results has received funding from the European Community’s Health Seventh Framework Programme (FP7/2007–2013). A.G. is supported by the Research Foundation KU Leuven (C24/16/045). A.A.-C. received salary support from the National Institute for Health Research (NIHR) Dementia Biomedical Research Unit and Biomedical Research Centre in Mental Health at South London and Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. Samples used in this research were in part obtained from the UK National DNA Bank for MND Research, funded by the MND Association and the Wellcome Trust. We acknowledge sample management undertaken by Biobanking Solutions funded by the Medical Research Council (MRC) at the Centre for Integrated Genomic Medical Research, University of Manchester. This is an EU Joint Programme-Neurodegenerative Disease Research (IPND) Project (STRENGTH, SOPHIA). In addition to those mentioned above, the project is supported through the following funding organizations under the auspices of IPND: UK, Economic and Social Research Council, Italy, Ministry of Health and Ministry of Education, University and Research; France, L’Agence nationale pour la recherche. The work leading up to this publication was funded by the European Community’s Health Seventh Framework Programme (FP7/2007–2013; Grant Agreement Number 259,867). We thank the International Genomics of Alzheimer’s Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i-Select chips was funded by the French National Foundation on Alzheimer’s disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de Lille, Université de Lille 2 and the Lille University Hospital. GERAD was supported by the MRC (Grant No. 5,03,176), The Wellcome Trust (Grant No. 082604/207/Z) and German Federal Ministry of Education and Research: Competence Network Dementia Grant no. 01GI0102, 01GI0711, 01GI0420. CHARITY was partly supported by the NIH/NIA Grant R01 AG033193 and the NIA AG081220 and AGES contract N01-AG-12,100, the NHLBI Grant RO1 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA Grants: U01 AG032984, U24 AG021886, U01 AG06976, and the Alzheimer’s Association Grant ADGC-10-196728. The Project MInE GWAS Consortium included contributions from the PARALS registry, SLALOM group, SLAP registry, FAIS Sequencing Consortium, SLAGEN Consortium and NNIPPS Study Group; The Schizophrenia Working Group of the Psychiatric Genomics Consortium included contributions from the Psychosis Endophenotypes International Consortium and Wellcome Trust Case-control Consortium. Members of these eight consortia are listed in Supplementary Note 2.

Author contributions


Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications.

Competing interests: O.H. has received speaking honoraria from Novartis, Biogen Idec, Sanofi Aventis and Merck-Serono. She has been a member of advisory panels for Biogen Idec, Allergen, Ono Pharmaceuticals, Novartis, Cytokinetiks and Sanofi Aventis. She serves as Editor-in-Chief of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. L.H.v.d.B. serves on scientific advisory boards for Primes Beatrix Spierfonds, Thiery Latran Foundation, Baxalta, Cytokinetics and Biogen, serves on the Editorial Board of the Journal of Neurology, Neurosurgery, and Psychiatry, Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, and Journal of Neuromuscular Diseases. A.A.C. has served on advisory panels for Biogen Idec, Cytokinetiks, GSK, OntronPharma and Mitsubishi-Tanabe, serves on the Editorial Boards of Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration and F1000, and receives royalties for The Brain: A Beginner’s Guide, OneWorld Publications, and Genetics of Complex Human Diseases, Cold Spring Harbor Laboratory Press. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: McLaughlin, R. L. et al. Genetic correlation between amyotrophic lateral sclerosis and schizophrenia. Nat. Commun. 8, 14774 doi: 10.1038/ncomms14774 (2017).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2017
Project MinE GWAS Consortium

Schizophrenia Working Group of the Psychiatric Genomics Consortium

New York, USA. 206Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University in Szczecin, Szczecin, Poland. 207Department of Mental Health and Substance Abuse Services, National Institute of Health and Welfare, Helsinki, Finland. 208Department of Mental Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA. 209Department of Psychiatry, University of Bonn, Bonn, Germany. 210Centre National de la Recherche Scientifique, Laboratoire de Génétique Moléculaire de la Neurotransmission et des Processus Neurodégénératifs, Hôpital de la Pitié Salpêtrière, Paris, France. 211Department of Genomics Mathematics, University of Bonn, Bonn, Germany. 212Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK. 213Division of Psychiatry, University of Edinburgh, Edinburgh, UK. 214Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. 215Massachusetts Mental Health Center Public Psychiatry Division of the Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA. 216Estonian Genome Center, University of Tartu, Tartu, Estonia. 217School of Psychology, University of Newcastle, Newcastle, Australia. 218First Psychiatric Clinic, Medical University, Sofia, Bulgaria. 219Department P, Aarhus University Hospital, Risskov, Denmark. 220Department of Psychiatry, Royal College of Surgeons in Ireland, Ireland. 221King’s College London, London, UK. 222Maastricht University Medical Centre, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht, The Netherlands. 223Institute of Translational Medicine, University Liverpool, UK. 224Max Planck Institute of Psychiatry, Munich, Germany. 225Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. 226Department of Psychiatry and Psychotherapy, Jena University Hospital, Jena, Germany. 227Department of Psychiatry, Queensland Brain Institute and Queensland Centre for Mental Health Research, University of Queensland, Brisbane, Queensland, Australia. 228Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. 229Department of Psychiatry, Trinity College Dublin, Dublin, Ireland. 230El Lilly and Company, Lilly Corporate Center, Indianapolis, Indiana, USA. 231Department of Clinical Sciences, Psychiatry, Umeå University, Umeå, Sweden. 232Detect Early Intervention Service for Psychosis, Blackrock, Dublin, Ireland. 233Centre for Public Health, Institute of Clinical Sciences, Queens University Belfast, Belfast, UK. 234Lawrence Berkeley National Laboratory, University of California at Berkeley, Berkeley, California, USA. 235Institute of Psychiatry at King’s College London, London, UK. 236Melbourne Neuropsychiatry Centre, University of Melbourne & Melbourne Health, Melbourne, Australia. 237Department of Psychiatry, University of Helsinki, Finland. 238Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Helsinki, Finland. 239Medical Faculty, University of Belgrade, Belgrade, Serbia. 240Department of Psychiatry, University of North Carolina, Chapel Hill, North Carolina, USA. 241Institute for Molecular Medicine Finland, FIMM, Helsinki, Finland. 242Department of Epidemiology, Harvard University, Boston, Massachusetts, USA. 243Department of Psychiatry, University of Oxford, Oxford, UK. 244Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, Virginia, USA. 245Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, New York, USA. 246Institute of Psychiatry at King’s College London, London, UK. 247Department of Psychiatry and Psychotherapy, University of Gottingen, Göttingen, Germany. 248Psychiatry and Psychotherapy Clinic, University of Erlangen, Erlangen, Germany. 249Hunter New England Health Service, Newcastle, Australia. 250School of Biomedical Sciences, University of Newcastle, Newcastle, Australia. 251Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA. 252University of Iceland, Landslagi, National University Hospital, Reykjavík, Iceland. 253Department of Psychiatry and Drug Addiction, Tbilisi State Medical University (TSMU), Tbilisi, Georgia. 254Research and Development, Bronx Veterans Affairs Medical Center, New York, New York, USA. 255Welcome Trust Centre for Human Genetics, Oxford, UK. 256deCODE Genetics, Reykjavik, Iceland. 257Department of Clinical Neurology, Medical University of Vienna, Vienna, Austria. 258Lieber Institute for Brain Development, Baltimore, Maryland, USA. 259Department of Medical Genetics, University Medical Centre, Utrecht, The Netherlands. 260Rudolf Magnus Institute of Neuroscience, University Medical Centre Utrecht, Utrecht, The Netherlands. 261Berkeley Healthcare N.H.S. Foundation Trust, Bracknell, UK. 262Section of Psychiatry, University of Verona, Verona, Italy. 263Department of Psychiatry, University of Oulu, Oulu, Finland. 264University Hospital of Oulu, Oulu, Finland. 265Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Ireland. 266Health Research Board, Dublin, Ireland. 267Department of Psychiatry and Clinical Neurosciences, School of Psychiatry and Clinical Neurosciences, Queen Elizabeth I.I. Medical Centre, Perth, Western Australia, Australia. 268Department of Psychological Medicine and Neurology, M.R.C. Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Cardiff, Wales, UK. 269Computational Sciences CoE, Pfizer Worldwide Research and Development, Cambridge, Massachusetts, USA. 270Human Genetics, Genome Institute of Singapore, Singapore. 271University College London, London, UK. 272Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, New York, USA. 273The Hebrew University of Jerusalem, Jerusalem, Israel. 274Neuroscience Discovery and Translational Area, Pharma Research and Early Development, F. Hoffman-La Roche, Basel, Switzerland. 275The Perkins Institute of Medical Research, Perth, Australia. 276LWA Centre for Clinical Research in Neuropsychiatry, Virginia Commonwealth University, Richmond, Virginia, USA. 277The Feinstein Institute for Medical Research, Manhasset, New York, USA. 278The Hofstra NS-LIU School of Medicine, Hempstead, New York, USA. 279The Zucker Hillside Hospital, Glen Oaks, New York, USA. 279Saw Swee Hock School of Public Health, National University of Singapore, Singapore. 280Queensland Centre for Mental Health Research, University of Queensland, Brisbane, Queensland, Australia. 281The Broad Institute of M.I.T. and Harvard, Cambridge, Massachusetts, USA. 282Center for Human Genetic Research and Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA. 283Department of Child and Adolescent Psychiatry, Erasmus University Medical Centre, Rotterdam, The Netherlands. 284Department of Complex Trait Genetics, Neuroscience Campus Amsterdam, VU University Medical Center Amsterdam, Amsterdam, The Netherlands. 285Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University, Amsterdam, The Netherlands. 286University of Aberdeen, Institute of Medical Sciences, Aberdeen, Scotland, UK. 287Department of Psychiatry, Neurology, Neuroscience and Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA. 288Department of Psychiatry, University of Copenhagen, Copenhagen, Denmark.