

## Sex-specific risk loci and modified *MEF2C* expression in ALS

Ross P Byrne<sup>1\*</sup>, Wouter van Rheenen<sup>2</sup>, Tomás DS Gomes<sup>1</sup>, Ciarán M Kelly<sup>3</sup>, Emrah Kaçar<sup>1</sup>, Project MinE ALS GWAS Consortium<sup>4</sup>, International ALS/FTD Genomics Consortium<sup>5</sup>, Ahmad Al Khleifat<sup>6,7</sup>, Alfredo Iacoangeli<sup>6,7,8,9</sup>, Ammar Al-Chalabi<sup>6,10</sup>, Jan H Veldink<sup>2</sup>, Russell L McLaughlin<sup>1</sup>

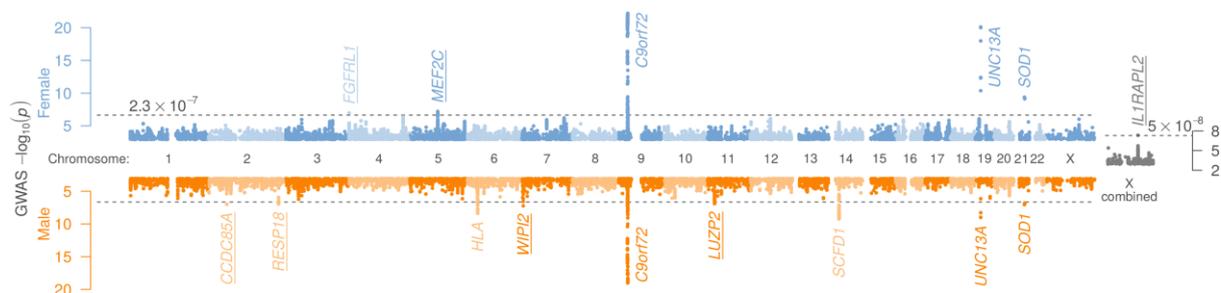
\* To whom correspondence should be addressed:

Ross P Byrne PhD Email: [rbyrne5@tcd.ie](mailto:rbyrne5@tcd.ie)  
Complex Trait Genomics Laboratory  
Smurfit Institute of Genetics  
Trinity College Dublin  
D02 PN40, Republic of Ireland

- 1 Complex Trait Genomics Laboratory, Smurfit Institute of Genetics, Trinity College Dublin, D02 PN40, Ireland
- 2 Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, the Netherlands
- 3 School of Pharmacy and Biomolecular Sciences, RCSI University of Medicine and Health Sciences, Dublin, Ireland
- 4 A full list of authors and their affiliations appears in Supplementary note 1
- 5 A full list of authors and their affiliations appears in Supplementary note 2
- 6 King's College London, Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, London SE5 9RX, United Kingdom
- 7 Perron Institute for Neurological and Translational Science, University of Western Australia Medical School, Perth, WA 6009, Australia
- 8 Department of Biostatistics and Health Informatics, King's College London, London, United Kingdom
- 9 NIHR BRC SLAM NHS Foundation Trust, London, United Kingdom
- 10 King's College Hospital, Denmark Hill, SE5 9RS London, UK.

**1 Significantly more men develop amyotrophic lateral sclerosis (ALS) than women, and heritability is not**  
**2 uniform between male and female transmissions, together suggesting a role for sex in the genetic aetiology of**  
**3 the disease. We therefore performed sex-stratified genome-wide and transcriptome-wide analyses of ALS risk,**  
**4 identifying six novel sex-specific risk loci including *MEF2C*, which shows increased expression in female ALS**  
**5 motor neurones. X-chromosome analysis revealed an additional risk locus at *IL1RAPL2*.**

**6 We calculated that previously-reported<sup>1</sup> differences in like-sexed and unlike-sexed parent-offspring transmission**  
**7 concordance for ALS implies a female-male genetic correlation of 62.8% (95% c.i. 48.0-73.2%), suggesting that some**  
**8 inherited ALS risk factors differ between women and men. To investigate this, we re-analysed the largest genome-**  
**9 wide association study (GWAS) of ALS<sup>2</sup>, splitting autosomal data by sex ( $n_{\text{female}} = 67,957$ ;  $n_{\text{male}} = 65,692$ ;**  
**10 Supplementary table 1), which revealed novel female-specific risk loci at *FGFRL1* and *MEF2C* and male-specific risk**  
**11 loci at *CCDC85A*, *BESPL8*, *WIP1* and *IL13BP2* (Figure 1, Supplementary table 2). A further two previously unappreciated**

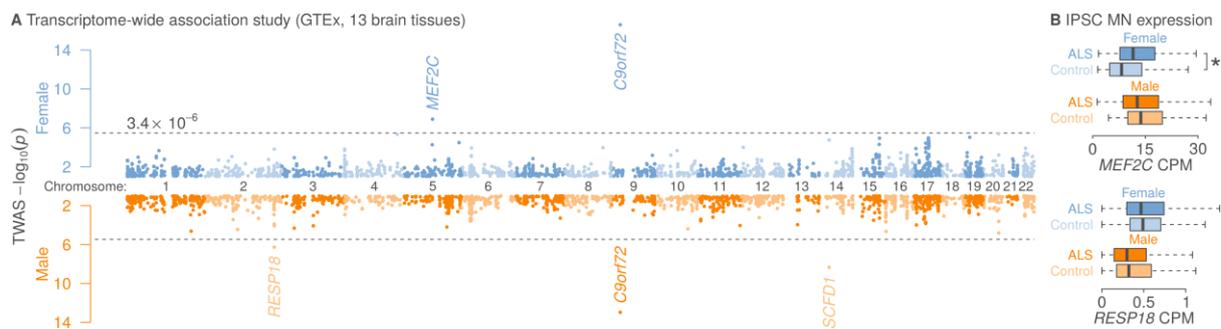


**Figure 1 Sex-stratified genome-wide association study (GWAS) statistics for amyotrophic lateral sclerosis (ALS) risk, including sex-combined X-chromosome association statistics.** For sex-stratified analyses, the p-value threshold to consider a locus as significantly associated with ALS corresponds to a false discovery rate of 1%. Combined X chromosome analysis uses the conventional genome-wide significance threshold of  $5 \times 10^{-8}$ . Significant loci are annotated with their closest gene and for brevity, only loci where  $p < 0.001$  are plotted. For the six novel sex-specific loci (underlined), cohort- and locus-level detail are provided in Supplementary figure 2-7.

12 loci (HLA locus and *SCFD1*) were only significantly associated with ALS risk in males, but did not meet our definition  
 13 of sex-specific ( $p < 2.3 \times 10^{-7}$  in one sex, corresponding to a false discovery rate  $< 1\%$ , and  $p > 0.05$  in the other sex).  
 14 GWAS statistics for both sexes were modestly inflated ( $\lambda_{GC} = 1.03$  in both; Supplementary figure 1) suggesting some  
 15 polygenic risk, which was further supported by univariate LD score regression<sup>3</sup> heritability estimates ( $h_{SNP}^2$  (s.e.) =  
 16 2.8% (0.65) for females and 2.1% (0.42) for males), with no apparent role for confounding in GWAS inflation (LDSC  
 17 intercepts  $< 1$ ). Despite the identification of sex-specific loci, bivariate LD score regression<sup>4</sup> estimated the genetic  
 18 correlation between male and female autosomal GWAS statistics to be 99% (s.e. 16%;  $p = 4.1 \times 10^{-10}$ ), suggesting an  
 19 overall shared polygenic component of ALS risk architecture between the two sexes. X chromosome data were  
 20 available for a subset of individuals ( $n_{female} = 19,576$ ;  $n_{male} = 21,991$ ); these were analysed per sex then resulting  
 21 statistics were meta-analysed, revealing one further novel ALS-associated risk locus at *ILIRAPL2* (Figure 1  
 22 Supplementary table 3).

23 To identify gene expression modified by ALS risk alleles, we performed a transcriptome-wide association study  
 24 (TWAS) of sex-stratified GWAS results with expression data for 13 brain tissues in GTEx version 8. A previously-  
 25 noted<sup>5</sup> association between ALS risk and expression of *C9orf72* was observed in both sexes and modified expression  
 26 of *MEF2C* and *RESP18* were associated with ALS risk in a strongly sex-specific manner (Figure 2A, Supplementary  
 27 table 4). Of the other GWAS risk loci, only *SCFD1* was significantly associated by TWAS analysis in males, but this  
 28 was not sex-specific ( $p_{female} = 1.81 \times 10^{-5}$ ). Notably, independent analysis of motor neurones (MNs) derived from  
 29 induced pluripotent stem cells (IPSCs) from ALS patients and controls<sup>6</sup> revealed *MEF2C* expression to be significantly  
 30 higher in female ALS patients than in female controls ( $p_{female} = 3 \times 10^{-4}$ , Figure 2B), but not significantly different  
 31 between male ALS patients and male controls ( $p_{male} = 0.07$ ).

32 *MEF2C* is a transcription factor involved in development of the mammalian neocortex, and has been associated with  
 33 Alzheimer's and Parkinson's diseases, along with several neuropsychiatric traits. As ALS risk modulated by *MEF2C*  
 34 expression likely involves its transcriptional targets, we reasoned that genetic variation present in target genes may  
 35 also confer risk for ALS. We therefore performed stratified LD score regression<sup>7</sup> to identify partitioned heritability  
 36 enrichment for ALS in a set of 1,055 genes either up- or downregulated in *Mef2c* conditional knockout mice<sup>8,9</sup>,  
 37 revealing 2.4-fold heritability enrichment in human orthologues of downregulated genes in female ALS GWAS  
 38 statistics ( $p = 2.6 \times 10^{-3}$ ) and a statistically insignificant heritability depletion in upregulated targets ( $p = 0.56$ ). No  
 39 significant heritability enrichment was observed in male ALS GWAS statistics (Supplementary table 5).



**Figure 2 Sex-specific analysis of gene expression in amyotrophic lateral sclerosis (ALS).** (A) Transcriptome-wide association study (TWAS) statistics using GWAS results from Figure 1 analysed with expression data from 13 brain tissues in GTEx version 8. Significance threshold is Bonferroni-corrected for 14,706 tests. Genes whose expression are significantly modified by ALS GWAS risk alleles are annotated and for brevity, only loci where  $p < 0.1$  are plotted. (B) For TWAS loci identified as sex-specific ( $p < 3.4 \times 10^{-6}$  in one sex;  $p > 0.05$  in the other), boxplots show sex-stratified expression (counts per million, CPM) in an independent set of motor neurones (MNs) derived from induced pluripotent stem cells (IPSCs) generated from 471 ALS patients and 283 age-matched controls in Answer ALS<sup>6</sup>. \*  $p < 8.3 \times 10^{-3}$  (logistic regression, corrected for 6 tests).

40 It is possible that *MEF2C* transcriptomic results partially explain the sex ratio in ALS<sup>10</sup>: although dysregulation of  
41 *MEF2C* targets has previously been implicated in ALS in sex-agnostic analyses<sup>11</sup>, our data suggest that lower  
42 expression in female motor neurones is protective against ALS, and increased expression to the range observed in  
43 men (who are generally at higher risk) increases risk in women. It is also worth noting that the IPSC MN system is  
44 isolated from endocrine differences between men and women, so sex differences are likely due to cell-autonomous  
45 regulation or sex-divergent epigenetic cellular identity established over the lifetime of the individual. This contrasts  
46 with *RESP18*, whose expression in MNs was not significantly different between cases and controls of either sex,  
47 despite its association by TWAS in men. As *RESP18* encodes a secreted neuroendocrine protein, it is not necessarily  
48 expected that IPSC MN expression in isolation would differ between ALS cases and controls. However, MN  
49 expression was lower in men overall than in women ( $p = 1.09 \times 10^{-9}$ ), and female rats have previously been shown to  
50 express twice the level of RESP18 protein in plasma compared to male rats<sup>12</sup>, consistent with a possible role for sex  
51 in determining its contribution to neurodegeneration. In support of this, *Resp18* knockout impairs locomotor activity  
52 and motor coordination but appears to protect against behavioural impairment in a chemically-induced model of  
53 Parkinson's disease in male mice<sup>13</sup>.

54 Although gene expression changes at remaining novel autosomal GWAS loci (*CCDC85A*, *FGFRL1*, *WIPI2* and  
55 *LUZP2*) were not significantly associated with ALS by TWAS, these genes also have biological roles that may  
56 contribute to candidate disease mechanisms in ALS when their risk alleles are inherited. *CCDC85A* is highly expressed  
57 in the frontal cortex, associating with synaptic function and interactions between neurones and astrocytes. The  
58 *FGFRL1* locus also includes *IDUA*, which has been associated with Alzheimer's disease<sup>14</sup>, and *GAK*, which was  
59 previously associated with Parkinson's disease<sup>15</sup> and implicated as a shared risk locus through colocalization analysis  
60 in sex-agnostic ALS GWAS<sup>2</sup>. *WIPI2* is involved in autophagy (a key process in ALS pathophysiology), and interacts  
61 with ALS-related proteins TBK1 and optineurin. *LUZP2* has recently been associated with age at onset of Alzheimer's  
62 disease<sup>16</sup>. Finally, *ILIRAPL2* was revealed by X chromosome analysis, which has previously been overlooked in ALS  
63 genomics. This encodes a central nervous system-specific protein which is part of the interactome of interleukin-2, a  
64 cytokine currently under clinical trial investigating its potential neuroprotective effect in ALS<sup>17</sup>. Stratification of trial  
65 results by rs17332232 genotype may reveal a pharmacogenomic element to any observed therapeutic benefit.

## 66 Methods

### 67 Pedigree-based genetic correlation estimation

68 Heritability estimates from like-sexed (mother-daughter and father-son) and unlike-sexed (mother-son and father-  
69 daughter) parent-offspring pairings were obtained from Ryan et al (2019)<sup>1</sup> to estimate genetic correlation in ALS  
70 liability between sexes. Falconer (1967)<sup>18</sup> expresses this as

$$71 \quad r_g = \frac{h_{\text{female}} h_{\text{male}} (\text{unlike} - \text{sexed})}{h_{\text{female}} h_{\text{male}} (\text{like} - \text{sexed})},$$

72 where  $r_g$  is the genetic correlation between sexes and  $h_{\text{female}}$  and  $h_{\text{male}}$  are the square roots of heritability in female and  
73 male cohorts, respectively. Confidence intervals were calculated using the upper and lower bounds of the 95%  
74 confidence intervals from the source study in the numerator and denominator, respectively for the upper bound, and  
75 *vice versa* for the lower bound.

### 76 Sex-stratified genome-wide association study

77 Genotype data were obtained from the European component the ALS GWAS presented by van Rheenen et al. (2021)<sup>2</sup>.  
78 Only variants and individuals passing quality control (QC) described in the original publication were retained for  
79 analysis. The source study subdivides data into six strata representing similar genotyping platforms; these same strata  
80 were used and further subdivided into male-only and female-only subsets based on reported sex (Supplementary table  
81 1). One sample from any pair with a coefficient of relatedness greater than 0.125 (calculated over 150k SNPs) between  
82 the male and female subsets of each stratum was removed to avoid potential bias in analyses comparing the two sex-  
83 specific GWAS subsets (e.g. genetic correlation analysis). Samples with relatedness within each stratum were  
84 retained; these are appropriately handled by SAIGE v0.39<sup>19</sup>, which we used to calculate association statistics per  
85 stratum. Firstly, we fit a null logistic mixed model in a leave-one chromosome out scheme per stratum using the first  
86 20 principal components as covariates. For this model we used a subset of high-quality (INFO > 0.95) SNPs which  
87 were LD pruned using PLINK 1.9 (--indep-pairwise 50 25 0.05 and high LD regions removed:  
88 [https://genome.sph.umich.edu/wiki/Regions\\_of\\_high\\_linkage\\_disequilibrium\\_\(LD\)](https://genome.sph.umich.edu/wiki/Regions_of_high_linkage_disequilibrium_(LD))). Next we ran a SNP-wise  
89 logistic mixed model implementing a saddlepoint approximation on all hard called genotypes in a leave one  
90 chromosome out scheme. Resulting sex-specific association statistics were combined across strata using inverse  
91 variance weighted meta-analysis implemented in METAL<sup>20</sup>. Associated loci were classified as sex-specific if  
92 significantly associated with ALS in one sex after stringent correction for multiple testing (1% false discovery rate;  
93 Benjamini-Hochberg) but not even nominally associated in the other sex ( $p > 0.05$ ).

94 To estimate SNP-based heritability we ran univariate LD score regression on sex-specific GWAS summary statistics,  
95 using pre-calculated LD scores from the European subsample of the 1000 Genomes Project and considering only SNPs  
96 included in HapMap phase 3. Heritability estimates were converted to the liability scale using sex-specific lifetime  
97 risk estimates for ALS<sup>1</sup> (1 in 436 for women; 1 in 347 for men) as prevalence. The sum of effective sample sizes (i.e.  
98 the effective sample size of an equivalent GWAS with a 50:50 case control split) was used in place of total N, and  
99 sample prevalence was set at 0.5 following evidence that using total N and sample prevalence downwardly biases

100 heritability estimates when GWAS results from meta-analysis of several cohorts with different ascertainment<sup>21</sup>.  
101 Effective sample size was calculated as  $\frac{4N_{ALS}N_{control}}{N_{ALS}+N_{control}}$  in each stratum and summed across strata.

## 102 X chromosome analysis

103 Previously unanalysed X chromosome data for 41,567 samples (15,146 cases and 26,421 controls totalling 19,576  
104 women and 21,991 men), compiled from two recent ALS GWAS cohorts<sup>2,22</sup>, were analysed to identify genetic  
105 associations between ALS and X-chromosome variants. Only samples passing QC in the base papers with associated  
106 X chromosome data were retained. This dataset was analysed separately from the autosomal GWAS datasets due to  
107 both the different sample composition (X chromosome data were unavailable for many samples) and different QC and  
108 analysis strategies needed to perform robust X chromosome association analysis.

109 Data from the raw cohorts were mapped to GRCh37 using liftOver and for each cohort we performed an initial pre-  
110 imputation QC. Briefly, we split cohorts by sex, then for each we removed: i) SNPs with missingness > 0.02 in either  
111 sex; ii) SNPs with minor allele frequencies (MAF) < 0.01 in either sex; iii) SNPs showing significant deviation from  
112 Hardy-Weinberg equilibrium (HWE;  $p < 1 \times 10^{-4}$ ); iv) strand ambiguous AT or GC SNPs; v) SNPs showing  
113 heterozygous calls on the X-chromosome in men. Cohorts were then merged into 28 genotyping platform- and  
114 country-matched strata which were balanced for case-control ratio following the scheme of the source papers<sup>2,22</sup> and  
115 subject to a second round of QC, removing SNPs per stratum with: i) significant deviation from HWE ( $p < 1 \times 10^{-5}$ );  
116 ii) missingness > 0.1; iii) MAF < 0.01. Finally, one sample from any pair with a coefficient of relatedness greater than  
117 0.125 was removed.

118 For imputation, input VCF files aligned to the Haplotype Reference Consortium (HRC) reference allele were prepared  
119 following the Michigan Imputation Server best practices guidelines. Imputation was then carried out per stratum on  
120 the Michigan imputation server using the HRC (version r1.1 2016) as a reference panel, Eagle 2.4 for phasing and  
121 Minmac 4 for imputation. Following imputation, SNPs with an Rsq value of 0.3 or greater were retained and subjected  
122 to a final round of QC, including removal of SNPs: i) significantly deviating from HWE ( $p < 1 \times 10^{-5}$ ) per stratum; ii)  
123 with missingness > 0.05 in the whole dataset; iii) with MAF < 0.01 in the whole dataset.

124 For association testing, strata were merged using plink, male genotypes were recoded as 0/1 and the software XWAS<sup>23</sup>  
125 was run using a sex-stratified model with 20 principal components calculated from autosomal SNPs as covariates.  
126 Test statistics for males and females were combined across sexes using Fisher's method (`--fisher`), which does  
127 not make assumptions about X inactivation, simplifying interpretation.

## 128 Transcriptomic analyses

129 We carried out a transcriptome-wide association study (TWAS) using FUSION<sup>24</sup> on the summary statistics from the  
130 male and female GWAS for ALS to test for association between gene expression and ALS in 13 brain tissues from  
131 GTEx version 8. For this analysis we used the European subset of the 1000 Genomes Project as an LD reference panel  
132 and precomputed weights from all GTEx samples (<http://gusevlab.org/projects/fusion/>). Resulting p-values per gene  
133 were aggregated across brain tissues using the SCAT method developed by Xiao et al. (2020)<sup>5</sup>.

134 To analyse sex-specific TWAS hits in iPSC-derived motor neurones, raw transcriptomic count data for 773 samples  
135 were downloaded from the Answer ALS data portal (version 6.0; January 2024). Definite ALS cases and controls  
136 were extracted based on clinical metadata, excluding non-ALS MND and asymptomatic carriers. Count data were  
137 normalised and converted to counts per million (CPM) using the edgeR package. Association between ALS and gene  
138 expression for *MEF2C* and *RESPI8* was estimated using a logistic regression model in R and computed in the full  
139 dataset (471 ALS, 283 control) as well as male-stratified (294 ALS, 87 control) and female-stratified (177 ALS, 196  
140 control) subsets of the data. A Bonferonni adjusted p-value threshold of  $P < 0.0083$  was used to correct for multiple  
141 testing (6 tests).

## 142 Heritability enrichment of *MEF2C* targets

143 Stratified LD score regression<sup>7</sup> was used to analyse heritability enrichment in a set of genes<sup>9</sup> differentially expressed  
144 after *Mef2c* conditional knockout in developing mice<sup>8</sup>. This analysis was run using sex-stratified GWAS summary  
145 statistics and combined data from the European-only subset of the 2021 ALS GWAS<sup>2</sup>. An annotation was first made  
146 for SNPs in a 100 kb region surrounding each differentially-expressed gene. LD scores were calculated for this  
147 annotation using the European subset of the 1000 Genomes Project dataset and considering only HapMap 3 SNPs.  
148 The annotation was run against the baseline LD model.

## 149 References

- 150 1. Ryan, M., Heverin, M., McLaughlin, R. L. & Hardiman, O. Lifetime Risk and Heritability of Amyotrophic Lateral  
151 Sclerosis. *JAMA Neurol.* **76**, 1367–1374 (2019).
- 152 2. van Rheenen, W. *et al.* Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15  
153 risk loci with distinct genetic architectures and neuron-specific biology. *Nat. Genet.* **53**, 1636–1648 (2021).
- 154 3. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide  
155 association studies. *Nat. Genet.* **47**, 291–295 (2015).
- 156 4. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–  
157 1241 (2015).
- 158 5. Xiao, L. *et al.* Multiple-Tissue Integrative Transcriptome-Wide Association Studies Discovered New Genes  
159 Associated With Amyotrophic Lateral Sclerosis. *Front. Genet.* **11**, 587243 (2020).
- 160 6. Baxi, E. G. *et al.* Answer ALS, a large-scale resource for sporadic and familial ALS combining clinical and multi-  
161 omics data from induced pluripotent cell lines. *Nat. Neurosci.* **25**, 226–237 (2022).
- 162 7. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary  
163 statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
- 164 8. Harrington, A. J. *et al.* MEF2C regulates cortical inhibitory and excitatory synapses and behaviors relevant to  
165 neurodevelopmental disorders. *eLife* **5**, e20059 (2016).

- 166 9. Cosgrove, D. *et al.* Genes influenced by MEF2C contribute to neurodevelopmental disease via gene expression  
167 changes that affect multiple types of cortical excitatory neurons. *Hum. Mol. Genet.* **30**, 961–970 (2021).
- 168 10. Manjaly, Z. R. *et al.* The sex ratio in amyotrophic lateral sclerosis: A population based study. *Amyotroph. Lateral*  
169 *Scler. Off. Publ. World Fed. Neurol. Res. Group Mot. Neuron Dis.* **11**, 439–442 (2010).
- 170 11. Arosio, A. *et al.* MEF2D and MEF2C pathways disruption in sporadic and familial ALS patients. *Mol. Cell.*  
171 *Neurosci.* **74**, 10–17 (2016).
- 172 12. Darlington, D., Schiller, M., Mains, R. & Eipper, B. The expression of regulated endocrine-specific protein of 18  
173 kDa in peptidergic cells of rat peripheral endocrine tissues and in blood. *J. Endocrinol.* **155**, 329–341 (1997).
- 174 13. Su, J. *et al.* RESP18 deficiency has protective effects in dopaminergic neurons in an MPTP mouse model of  
175 Parkinson’s disease. *Neurochem. Int.* **118**, 195–204 (2018).
- 176 14. Bellenguez, C. *et al.* New insights into the genetic etiology of Alzheimer’s disease and related dementias. *Nat.*  
177 *Genet.* **54**, 412–436 (2022).
- 178 15. Nagle, M. W. *et al.* The 4p16.3 Parkinson Disease Risk Locus Is Associated with GAK Expression and Genes  
179 Involved with the Synaptic Vesicle Membrane. *PLOS ONE* **11**, e0160925 (2016).
- 180 16. Li, Y. *et al.* Identification of novel genes for age-at-onset of Alzheimer’s disease by combining quantitative and  
181 survival trait analyses. *Alzheimers Dement.* **19**, 3148–3157 (2023).
- 182 17. Camu, W. *et al.* Repeated 5-day cycles of low dose aldesleukin in amyotrophic lateral sclerosis (IMODALS): A  
183 phase 2a randomised, double-blind, placebo-controlled trial. *EBioMedicine* **59**, 102844 (2020).
- 184 18. Falconer, D. S. The inheritance of liability to diseases with variable age of onset, with particular reference to  
185 diabetes mellitus. *Ann. Hum. Genet.* **31**, 1–20 (1967).
- 186 19. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic  
187 association studies. *Nat. Genet.* **50**, 1335–1341 (2018).
- 188 20. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans.  
189 *Bioinforma. Oxf. Engl.* **26**, 2190–2191 (2010).
- 190 21. Grotzinger, A. D., Fuente, J. de la, Privé, F., Nivard, M. G. & Tucker-Drob, E. M. Pervasive Downward Bias in  
191 Estimates of Liability-Scale Heritability in Genome-wide Association Study Meta-analysis: A Simple Solution.  
192 *Biol. Psychiatry* **93**, 29–36 (2023).
- 193 22. van Rheenen, W. *et al.* Genome-wide association analyses identify new risk variants and the genetic architecture  
194 of amyotrophic lateral sclerosis. *Nat. Genet.* **48**, 1043–1048 (2016).
- 195 23. Gao, F. *et al.* XWAS: A Software Toolset for Genetic Data Analysis and Association Studies of the X  
196 Chromosome. *J. Hered.* **106**, 666–671 (2015).

197 24. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–  
198 252 (2016).

## 199 Acknowledgements

200 This study was funded by the MND Association of England, Wales and Northern Ireland under grant number 879-  
201 791. Data used in the preparation of this article were obtained from the ANSWER ALS Data Portal (AALS-01184).  
202 For up-to-date information on the study, visit <https://dataportal.answerals.org>. This publication has emanated from  
203 research supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number  
204 21/RC/10294 and co-funded under the European Regional Development Fund and by FutureNeuro industry partners.  
205 R.P.B. currently receives support from the MND Association (979-799). A.I. is funded by South London and  
206 Maudsley NHS Foundation Trust, MND Scotland, Motor Neurone Disease Association, National Institute for Health  
207 and Care Research, Spastic Paraplegia Foundation, Rosetrees Trust, Darby Rimmer MND Foundation, the Medical  
208 Research Council (UKRI), Alzheimer's Research UK and LifeArc. AAK is funded by ALS Association Milton  
209 Safenowitz Research Fellowship (grant number22-PDF- 609.DOI:10.52546/pc.gr.150909.), the Motor Neurone  
210 Disease Association (MNDA) Fellowship (Al Khleifat/Oct21/975-799), The Darby Rimmer Foundation, and The  
211 NIHR Maudsley Biomedical Research Centre. AAC is an NIHR Senior Investigator (NIHR202421). This is an EU  
212 Joint Programme - Neurodegenerative Disease Research (JPND) project. The project is supported through the  
213 following funding organisations under the aegis of JPND - [www.jpnd.eu](http://www.jpnd.eu) (*United Kingdom, Medical Research*  
214 *Council* (MR/L501529/1; MR/R024804/1) *and Economic and Social Research Council* (ES/L008238/1)) and through  
215 the Motor Neurone Disease Association, My Name's Doddie Foundation, and Alan Davidson Foundation. This study  
216 represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical  
217 Research Centre at South London and Maudsley NHS Foundation Trust and King's College London.

## 218 Author contributions

219 R.L.McL. and R.P.B. conceived the study and drafted the manuscript. R.P.B. conducted the analyses with the  
220 assistance of T.D.S, C.M.K. and E.K. W.v.R., A.A-K., A.I., A.A-C. and J.H.V. supplied data for GWAS and XWAS  
221 analyses. All authors provided critical review of the manuscript.

## 222 Data availability

223 The GWAS summary statistics generated in this study will be made publicly available upon publication in the NHGRI-  
224 EBI GWAS Catalog at <https://www.ebi.ac.uk/gwas/> (for female, male and X chromosome meta-analyses,  
225 respectively) and through the Project MinE website (<https://www.projectmine.com/research/download-data/>).  
226 Individual level genotype data used in this study was originally published in van Rheenen *et al* 2021 which includes  
227 the following publicly available datasets: the Wellcome Trust Case Control Consortium (<https://www.wtccc.org.uk/>)  
228 and dbGaP datasets ([phs000101.v3.p1](https://www.ncbi.nlm.nih.gov/bioproject/101), NIH Genome-Wide Association Studies of Amyotrophic Lateral  
229 Sclerosis; [phs000126.v1.p1](https://www.ncbi.nlm.nih.gov/bioproject/126), CIDR: Genome Wide Association Study in Familial Parkinson Disease  
230 (PD); [phs000196.v1.p1](https://www.ncbi.nlm.nih.gov/bioproject/196), Genome-Wide Association Study of Parkinson Disease: Genes and  
231 Environment; [phs000344.v1.p1](https://www.ncbi.nlm.nih.gov/bioproject/344), Genome-Wide Association Study of Amyotrophic Lateral Sclerosis in

232 Finland; [phs000336](#), a Genome-Wide Association Study of Lung Cancer Risk; [phs000346](#), Genome-Wide  
233 Association Study for Bladder Cancer Risk; [phs000789](#), Collaborative Study of Genes, Nutrients and Metabolites  
234 (CSGNM); [phs000206](#), Whole Genome Scan for Pancreatic Cancer Risk in the Pancreatic Cancer Cohort Consortium  
235 and Pancreatic Cancer Case–Control Consortium (PanScan); [phs000297](#), eMERGE Network Study of the Genetic  
236 Determinants of Resistant Hypertension; [phs000652](#), Cohort-Based Genome-Wide Association Study of Glioma  
237 (GliomaScan); [phs000869](#), Barrett’s and Esophageal Adenocarcinoma Genetic Susceptibility Study  
238 (BEAGESS); [phs000812](#), the Breast and Prostate Cancer Cohort Consortium (BPC3) GWAS of Aggressive Prostate  
239 Cancer and ER<sup>-</sup> Breast Cancer; [phs000428](#), Genetics Resource with the HRS; [phs000360.v3](#), eMERGE Network  
240 Genome-Wide Association Study of Red Cell Indices, White Blood Count (WBC) Differential, Diabetic Retinopathy,  
241 Height, Serum Lipid Levels, Specifically Total Cholesterol, HDL (High Density Lipoprotein), LDL (Low Density  
242 Lipoprotein), and Triglycerides, and Autoimmune Hypothyroidism; [phs000893.v1](#), Genome-Wide Association Study  
243 of Endometrial Cancer in the Epidemiology of Endometrial Cancer Consortium (E2C2); [phs000168.v2](#), National  
244 Institute on Aging—Late Onset Alzheimer’s Disease Family Study: Genome-Wide Association Study for  
245 Susceptibility Loci; [phs000092.v1](#), Study of Addiction: Genetics and Environment (SAGE); [phs000864.v1](#), Genomic  
246 Predictors of Combat Stress Vulnerability and Resilience; [phs000170.v2](#), a Genome-Wide Association Study on  
247 Cataract and HDL in the Personalized Medicine Research Project Cohort; [phs000431.v2](#), IgA Nephropathy GWAS  
248 on Individuals of European Ancestry (IGANGWAS2); [phs000237.v1](#), Northwestern NUGene Project: Type 2  
249 Diabetes; [phs000169.v1](#), Whole Genome Association Study of Visceral Adiposity in the Health Aging and Body  
250 Composition (Health ABC) Study; [phs000982.v1](#), Genetic Analysis of Psoriasis and Psoriatic Arthritis: GWAS of  
251 Psoriatic Arthritis; [phs000289.v2](#), National Human Genome Research Institute (NHGRI) GENEVA Genome-Wide  
252 Association Study of Venous Thrombosis (GWAS of VTE); [phs000634.v1](#), National Cancer Institute (NCI) Genome  
253 Wide Association Study (GWAS) of Lung Cancer in Never Smokers; [phs000274.v1](#), Genome-Wide Association  
254 Study of Celiac Disease; [phs001172.v1](#), National Institute of Neurological Disorders and Stroke (NINDS) Parkinson’s  
255 Disease; [phs000389.v1](#), Genetics of Nephropathy—an International Effort (GENIE) GWAS of Diabetic Nephropathy  
256 in the UK GoKinD and All-Ireland Cohorts; [phs000460.v1](#), Genetics of 24 Hour Urine Composition; [phs000138.v2](#),  
257 GWAS for Genetic Determinants of Bone Fragility in European–American Premenopausal Women; [phs000394.v1](#),  
258 Autopsy-Confirmed Parkinson Disease GWAS Consortium (APDGC); [phs000948.v1](#), Genetic Discovery and  
259 Application in a Clinical Setting: Continuing a Partnership (eMERGE Phase II); [phs000630.v1](#), Exome Chip Study  
260 of NIMH Controls; [phs000678.v1](#), a Family-Based Study of Genes and Environment in Young-Onset Breast  
261 Cancer; [phs000351.v1](#), National Cancer Institute Genome-Wide Association Study of Renal Cell  
262 Carcinoma; [phs000314.v1](#), Genetic Associations in Idiopathic Talipes Equinovarus (Clubfoot)—  
263 GAIT; [phs000147.v3](#), Cancer Genetic Markers of Susceptibility (CGEMS) Breast Cancer Genome-wide Association  
264 Study (GWAS)—Primary Scan: Nurses’ Health Study—Additional Cases: Nurses’ Health Study 2; [phs000882.v1](#),  
265 National Cancer Institute (NCI) Prostate Cancer Genome-Wide Association Study for Uncommon Susceptibility Loci  
266 (PEGASUS); [phs000238.v1](#), National Eye Institute Glaucoma Human Genetics Collaboration (NEIGHBOR)  
267 Consortium Glaucoma Genome-Wide Association Study; [phs000397.v1](#), National Institute on Aging (NIA) Long Life  
268 Family Study (LLFS); [phs000421.v1](#), a Genome-Wide Association Study of Fuchs’ Endothelial Corneal Dystrophy  
269 (FECD); [phs000142.v1](#), a Whole Genome Association Scan for Myopia and Glaucoma Endophenotypes using Twin  
270 Studies; [phs000303.v1](#), Genetic Epidemiology of Refractive Error in the KORA (Kooperative Gesundheitsforschung  
271 in der Region Augsburg) Study; [phs000125.v1](#), CIDR: Collaborative Study on the Genetics of Alcoholism Case

272 Control Study; [phs001039.v1](#), International Age-Related Macular Degeneration Genomics Consortium—Exome Chip  
273 Experiment; [phs000187.v1](#), High Density SNP Association Analysis of Melanoma: Case–Control and Outcomes  
274 Investigation; [phs000101.v5](#), Genome-Wide Association Study of Amyotrophic Lateral Sclerosis; [phs002068.v1.p1](#),  
275 Sporadic ALS Australia Systems Genomics Consortium (SALSA-SGC)). Gene expression data used in the preparation  
276 of this article were obtained from the ANSWER ALS Data Portal (AALS-01184). For up-to-date information on the  
277 study, visit <https://dataportal.answerals.org>.

## 278 Conflict of interest statement

279 J.H.V. reports to have sponsored research agreements with Biogen and AstraZeneca. AAK is a consultant for NESTA.  
280 AAC reports consultancies or advisory boards for Amylyx, Apellis, Biogen, Brainstorm, Cytokinetics, GenieUs, GSK,  
281 Lilly, Mitsubishi Tanabe Pharma, Novartis, OrionPharma, Qoralis, and Wave Pharmaceuticals. All other authors have  
282 nothing to declare.

283