Article:
Saez-Atienzar, S., Souza, C.D.S., Chia, R. et al. (2024) Mechanism-free repurposing of drugs for C9orf72-related ALS/FTD using large-scale genomic data. Cell Genomics, 4 (11). 100679. ISSN: 2666-979X

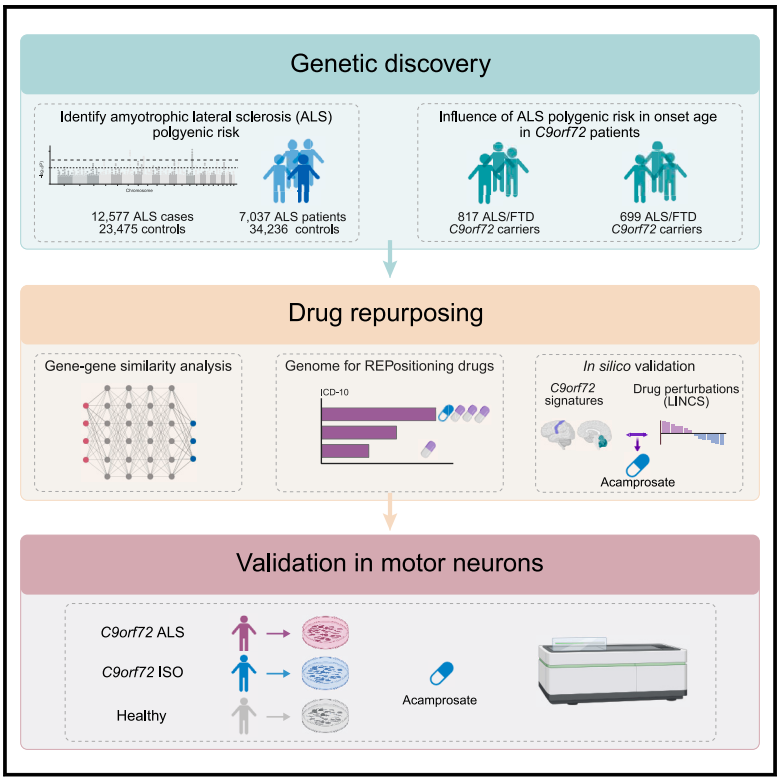
https://doi.org/10.1016/j.xgen.2024.100679

Cell Genomics

Article

Mechanism-free repurposing of drugs for *C9orf72*-related ALS/FTD using large-scale genomic data

Graphical abstract



Highlights

- Repeat expansions in *C9orf72* are the most common genetic cause of ALS and FTD
- The genetic risk of general ALS modifies the age at onset in *C9orf72* cases
- We performed a drug screen based on the genetic variants influencing age at onset
- We identified acamprosate, a GABA analog, as a potential treatment for *C9orf72*

Authors

Sara Saez-Atienzar, Cleide dos Santos Souza, Ruth Chia, ..., Laura Ferraiuolo, Isabella Fogh, Bryan J. Traynor

Correspondence

sara.saezatienzar@osumc.edu (S.S.-A.), traynorb@mail.nih.gov (B.J.T.)

In brief

Saez-Atienzar et al. identify genetic variants influencing the age at onset among patients carrying *C9orf72* repeat expansions. A drug screen based on these variants revealed acamprosate, a GABA analog, as a potentially repurposable treatment for *C9orf72*-related disease. The work underscores the potential of leveraging large-scale genomic data for drug repurposing.

Saez-Atienzar et al., 2024, Cell Genomics *4*, 100679
November 13, 2024 Published by Elsevier Inc.
https://doi.org/10.1016/j.xgen.2024.100679

CellPress

Cell Genomics

CellPress
OPEN ACCESS

Article

Mechanism-free repurposing of drugs for *C9orf72*-related ALS/FTD using large-scale genomic data

Sara Saez-Atienzar,[1,2,80,*] Cleide dos Santos Souza,[3,80] Ruth Chia,[1,80] Selina N. Beal,[3] Ileana Lorenzini,[4] Ruili Huang,[5] Jennifer Levy,[4] Camelia Burciu,[4] Jinhui Ding,[6] J. Raphael Gibbs,[6] Ashley Jones,[7] Ramita Dewan,[1] Viviana Pensato,[8] Silvia Peverelli,[9] Lucia Corrado,[10] Joke J.F.A. van Vugt,[11] Wouter van Rheenen,[11] Ceren Tunca,[12] Elif Bayraktar,[12] Menghang Xia,[5] The International ALS Genomics Consortium,[75] ITALSGEN Consortium,[76] SLAGEN Consortium,[77] Project MinE ALS Sequencing Consortium,[78] Alfredo Iacoangeli,[7,13,14] Aleksey Shatunov,[7] Cinzia Tiloca,[9] Nicola Ticozzi,[9,15] Federico Verde,[9,15] Letizia Mazzini,[16] Kevin Kenna,[11] Ahmad Al Khleifat,[7] Sarah Opie-Martin,[7] Flavia Raggi,[17]

(Author list continued on next page)

[1]Neuromuscular Diseases Research Section, National Institute on Aging, National Institutes of Health (NIH), Bethesda, MD 20892, USA
[2]Department of Neurology, Ohio State University, Columbus, OH 43210, USA
[3]Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield S10 2HQ, UK
[4]Department of Translational Neuroscience, Barrow Neurological Institute, Phoenix, AZ 85013, USA
[5]Division of Pre-clinical Innovation, National Center for Advancing Translational Sciences (NCATS), NIH, Rockville, MD 20850, USA
[6]Computational Biology Group, Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, USA
[7]Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK
[8]Unit of Medical Genetics and Neurogenetics, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[9]Department of Neurology and Laboratory of Neuroscience, Istituto Auxologico Italiano IRCCS, Milan, Italy
[10]Department of Health Sciences, University of Eastern Piedmont, Novara, Italy
[11]Department of Neurology, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands
[12]Neurodegeneration Research Laboratory (NDAL), Research Center for Translational Medicine (KUTTAM), Koç University School of Medicine, Istanbul, Turkey
[13]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK

(Affiliations continued on next page)

SUMMARY

Repeat expansions in the *C9orf72* gene are the most common genetic cause of (ALS) and frontotemporal dementia (FTD). Like other genetic forms of neurodegeneration, pinpointing the precise mechanism(s) by which this mutation leads to neuronal death remains elusive, and this lack of knowledge hampers the development of therapy for *C9orf72*-related disease. We used an agnostic approach based on genomic data (n = 41,273 ALS and healthy samples, and n = 1,516 *C9orf72* carriers) to overcome these bottlenecks. Our drug-repurposing screen, based on gene- and expression-pattern matching and information about the genetic variants influencing onset age among *C9orf72* carriers, identified acamprosate, a γ-aminobutyric acid analog, as a potentially repurposable treatment for patients carrying *C9orf72* repeat expansions. We validated its neuroprotective effect in cell models and showed comparable efficacy to riluzole, the current standard of care. Our work highlights the potential value of genomics in repurposing drugs in situations where the underlying pathomechanisms are inherently complex.

INTRODUCTION

A repeat expansion within the *C9orf72* gene is a common cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), two neurological disorders that result in the deaths of ~17,000 Americans and Europeans annually.[1–4] This genetic disease accounts for 1 in 10 ALS and FTD cases of European descent, and many carriers in the general population have a near-complete chance of manifesting symptoms during their lifetime.[5,6] The exact processes by which this repeat expansion leads to neuronal death are not fully understood, although several mechanisms, such as dipeptide production, RNA

Massimiliano Filosto,[18,19] Stefano Cotti Piccinelli,[18,19] Alessandro Padovani,[19] Stella Gagliardi,[20] Maurizio Inghilleri,[21,22] Alessandra Ferlini,[23] Rosario Vasta,[24] Andrea Calvo,[24,25] Cristina Moglia,[24,25] Antonio Canosa,[24,25,26] Umberto Manera,[24,25] Maurizio Grassano,[24] Jessica Mandrioli,[27,28] Gabriele Mora,[24] Christian Lunetta,[29,30] Raffaella Tanel,[31] Francesca Trojsi,[32] Patrizio Cardinali,[33] Salvatore Gallone,[24] Maura Brunetti,[24] Daniela Galimberti,[34,35] Maria Serpente,[34] Chiara Fenoglio,[34,35] Elio Scarpini,[34] Giacomo P. Comi,[15,36] Stefania Corti,[15,36] Roberto Del Bo,[15,36] Mauro Ceroni,[20,37] Giuseppe Lauria Pinter,[38,39] Franco Taroni,[8] Eleonora Dalla Bella,[38] Enrica Bersano,[38,40]

(Author list continued on next page)

[14]National Institute for Health Research Biomedical Research Centre and Dementia Unit, South London and Maudsley NHS Foundation Trust and King's College London, London, UK
[15]Department of Pathophysiology and Transplantation, "Dino Ferrari" Center, Università degli Studi di Milano, Milan, Italy
[16]Amyotrophic Lateral Sclerosis Center, Department of Neurology "Maggiore della Carità" University Hospital, Novara, Italy
[17]Department of Neurosciences, University of Padova, Padova, Italy
[18]NeMO-Brescia Clinical Center for Neuromuscular Diseases, University of Brescia, Brescia, Italy
[19]Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy
[20]Genomic and Post-Genomic Center, IRCCS Mondino Foundation, Pavia, Italy
[21]Department of Human Neurosciences, Rare Neuromuscular Diseases Centre, Sapienza University, 00185 Rome, Italy
[22]IRCCS Neuromed, Pozzilli, Italy
[23]Unit of Medical Genetics, Department of Medical Science, University of Ferrara, Ferrara, Italy
[24]"Rita Levi Montalcini" Department of Neuroscience, Amyotrophic Lateral Sclerosis Center, University of Turin, Turin, Italy
[25]Azienda Ospedaliero Universitaria Città della Salute e della Scienza, Turin, Italy
[26]Institute of Cognitive Sciences and Technologies, C.N.R., Rome, Italy
[27]Department of Biomedical, Metabolic, and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy
[28]Department of Neurosciences, Azienda Ospedaliero Universitaria di Modena, Modena, Italy
[29]Department of Neurorehabilitation, Istituti Clinici Scientifici Maugeri IRCCS, Institute of Milan, Milan, Italy
[30]NEMO Clinical Center Milano, Fondazione Serena Onlus, Milan, Italy
[31]Operative Unit of Neurology, S. Chiara Hospital, Trento, Italy
[32]Department of Advanced Medical and Surgical Sciences, University of Campania "Luigi Vanvitelli," Naples, Italy
[33]Neurology Unit, AST Fermo, Marche, Italy
[34]Neurodegenerative Diseases Unit, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy
[35]Department of Biomedical, Surgical, and Dental Sciences, University of Milan, Milan, Italy
[36]Neurology Unit, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy
[37]Department of Brain and Behavioural Sciences, University of Pavia, Pavia, Italy
[38]3rd Neurology Unit, Motor Neuron Diseases Center, Fondazione IRCCS Istituto Neurologico "Carlo Besta," Milan, Italy
[39]Department of Medical Biotechnology and Translational Medicine, University of Milan, Milan, Italy
[40]"L. Sacco" Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan, Italy
[41]Social Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology, and Neuroscience (IoPPN), King's College London, London, UK
[42]NIHR BioResource Centre Maudsley, NIHR Maudsley Biomedical Research Centre (BRC) at South London and Maudsley NHS Foundation Trust (SLaM), London, UK
[43]School of Medicine, Dentistry, and Biomedical Sciences, Faculty of Medicine Health and Life Sciences, Queen's University, Belfast, UK
[44]Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, and the NIHR Sheffield Biomedical Research Centre, Sheffield, UK
[45]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology, and Neuroscience (IoPPN), King's College London, London SE5 8AF, UK
[46]NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, London, UK
[47]Health Data Research UK London, University College London, London, UK
[48]Institute of Health Informatics, University College London, London, UK
[49]NIHR Biomedical Research Centre at University College London Hospitals NHS Foundation Trust, London, UK
[50]Department of Anatomy, Physiology, & Genetics, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA
[51]The American Genome Center, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA
[52]Neurodegenerative Diseases Research Section, National Institute of Neurological Disorders and Stroke (NINDS), NIH, Bethesda, MD 20892, USA
[53]Department of Neurology, Johns Hopkins University Medical Center, Baltimore, MD 21287, USA
[54]Department of Clinical Neuroscience, King's College Hospital, London SE5 9RS, UK
[55]Complex Trait Genomics Laboratory, Smurfit Institute of Genetics, Trinity College Dublin, Dublin, Ireland
[56]Academic Unit of Neurology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin, Ireland
[57]Department of Neurology, Beaumont Hospital, Dublin, Ireland
[58]Euan MacDonald Centre for Motor Neurone Disease Research, Edinburgh, UK
[59]UK Dementia Research Institute, University of Edinburgh, Edinburgh, UK
[60]Centre for Neuroregeneration and Medical Research Council Centre for Regenerative Medicine, University of Edinburgh, Edinburgh, UK

(Affiliations continued on next page)

CellPress
OPEN ACCESS

Charles J. Curtis,[41,42] Sang Hyuck Lee,[41,42] Raymond Chung,[41,42] Hamel Patel,[13,42] Karen E. Morrison,[43] Johnathan Cooper-Knock,[44] Pamela J. Shaw,[44] Gerome Breen,[41,42] Richard J.B. Dobson,[45,46,47,48,49] Clifton L. Dalgard,[50,51] The American Genome Center[79], Sonja W. Scholz,[52,53] Ammar Al-Chalabi,[7,54] Leonard H. van den Berg,[11] Russell McLaughlin,[55] Orla Hardiman,[56,57] Cristina Cereda,[20] Gianni Sorarù,[17] Sandra D'Alfonso,[10] Siddharthan Chandran,[58,59] Suvankar Pal,[58,60] Antonia Ratti,[9,61] Cinzia Gellera,[8] Kory Johnson,[62] Tara Doucet-O'Hare,[63] Nicholas Pasternack,[64] Tongguang Wang,[64] Avindra Nath,[64] Gabriele Siciliano,[65] Vincenzo Silani,[9,15] Ayşe Nazlı Başak,[12] Jan H. Veldink,[11] William Camu,[66] Jonathan D. Glass,[67] John E. Landers,[68] Adriano Chiò,[24,25,26] Rita Sattler,[4] Christopher E. Shaw,[69,70,80] Laura Ferraiuolo,[71,80] Isabella Fogh,[69,80] and Bryan J. Traynor[1,53,72,73,74,80,81,*]

[61]Department of Medical Biotechnology and Translational Medicine, Università degli Studi di Milano, Milan, Italy
[62]Bioinformatics Section, Information Technology Program (ITP), Division of Intramural Research (DIR), National Institute of Neurological Disorders & Stroke, NIH, Bethesda, MD 20892, USA
[63]Neuro-oncology Branch, Center for Cancer Research, National Cancer Institute (NCI), NIH, Bethesda, MD 20892, USA
[64]Translational Neuroscience Center, National Institute of Neurological Disorders and Stroke (NINDS), NIH, Bethesda, MD 20892, USA
[65]Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy
[66]ALS Center, CHU Gui de Chauliac, University of Montpellier, Montpellier, France
[67]Department of Neurology, Emory University School of Medicine, Atlanta, GA 30322, USA
[68]Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA
[69]United Kingdom Dementia Research Institute, Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK
[70]Centre for Brain Research, University of Auckland, Auckland, New Zealand
[71]Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield S10 2HQ, UK
[72]Reta Lila Weston Institute, UCL Queen Square Institute of Neurology, University College London, London WC1N 1PJ, UK
[73]National Institute of Neurological Disorders and Stroke (NINDS), NIH, Bethesda, MD 20892, USA
[74]RNA Therapeutics Laboratory, National Center for Advancing Translational Sciences (NCATS), NIH, Rockville, MD 20850, USA
[75]International ALS Genomics Consortium members are listed fully in the supplemental information.
[76]ITALSGEN Consortium members are listed fully in the supplemental information.
[77]SLAGEN Consortium members are listed fully in the supplemental information.
[78]Project MinE ALS Sequencing Consortium members are listed fully in the supplemental information.
[79]The American Genome Center members are listed fully in the supplemental information.
[80]These authors contributed equally
[81]Lead contact
*Correspondence: sara.saezatienzar@osumc.edu (S.S.-A.), traynorb@mail.nih.gov (B.J.T.)
https://doi.org/10.1016/j.xgen.2024.100679

toxicity, disruption of nucleocytoplasmic transport, and haploinsufficiency, have been suggested.[7] This lack of molecular knowledge is a common theme across neurodegenerative disorders, where there is often a complex interplay among multiple pathways and cellular processes.[8–11]

The molecular complexity underlying neurodegenerative diseases also hampers drug discovery; ameliorating a single aspect of a cellular network may not be beneficial as it does not address the other pathological processes co-occurring within the cell.[12] Traditional linear drug-discovery efforts, where millions of compounds are tested against a single target, are likely to fail in the face of such multidimensional conditions, and this disconnect alone may account for the high failure rate observed among clinical trials in neurodegenerative diseases.[13] Genomic and transcriptomic data offer a potential solution, as these data types inherently capture the multifaceted nature of neurological diseases. We can exploit this information to match drugs that restore entire networks and systems, even when the target pathways or the mechanism of action of the drug are not fully understood.[14,15] Compounds supported by genetic evidence are also more likely to succeed in clinical trials and to gain drug approval.[16,17]

In this context, we have used a massive genomic dataset to identify the genetic variants influencing age at onset among patients carrying the *C9orf72* repeat expansion. Building on this genomic information, we have leveraged gene- and expression-pattern matching and pathway modeling to nominate potentially repurposable drugs for *C9orf72*-related ALS/FTD. We focused on onset age because of the wide range observed in this common genetic form of neurodegeneration, spanning from the fourth to the tenth decade of life, representing a natural experiment within the ALS population.[5,18] Genetic factors are known to play a role in this variable age-at-onset presentation,[19] and research in other neurological diseases shows that the phenotypic manifestations of high-risk pathogenic variants are influenced by minor effect variants elsewhere in the genome.[20,21]

Our pipeline nominated acamprosate, an oral medication used to manage alcohol use disorder, as a potential repurposable drug for slowing progression among symptomatic individuals and delaying disease onset among *C9orf72* carriers. *In vitro* data demonstrated a neuroprotective effect of acamprosate in motor neurons derived from these patients. Crucially, our discovery approach evolved from the notion that genomics can nominate specific medications in an agnostic, data-driven manner, even when precise knowledge about disease mechanisms or drug pharmacodynamics is lacking.

RESULTS

Figure 1 shows a graphical representation of our genetic workflow and drug-repurposing pipeline.

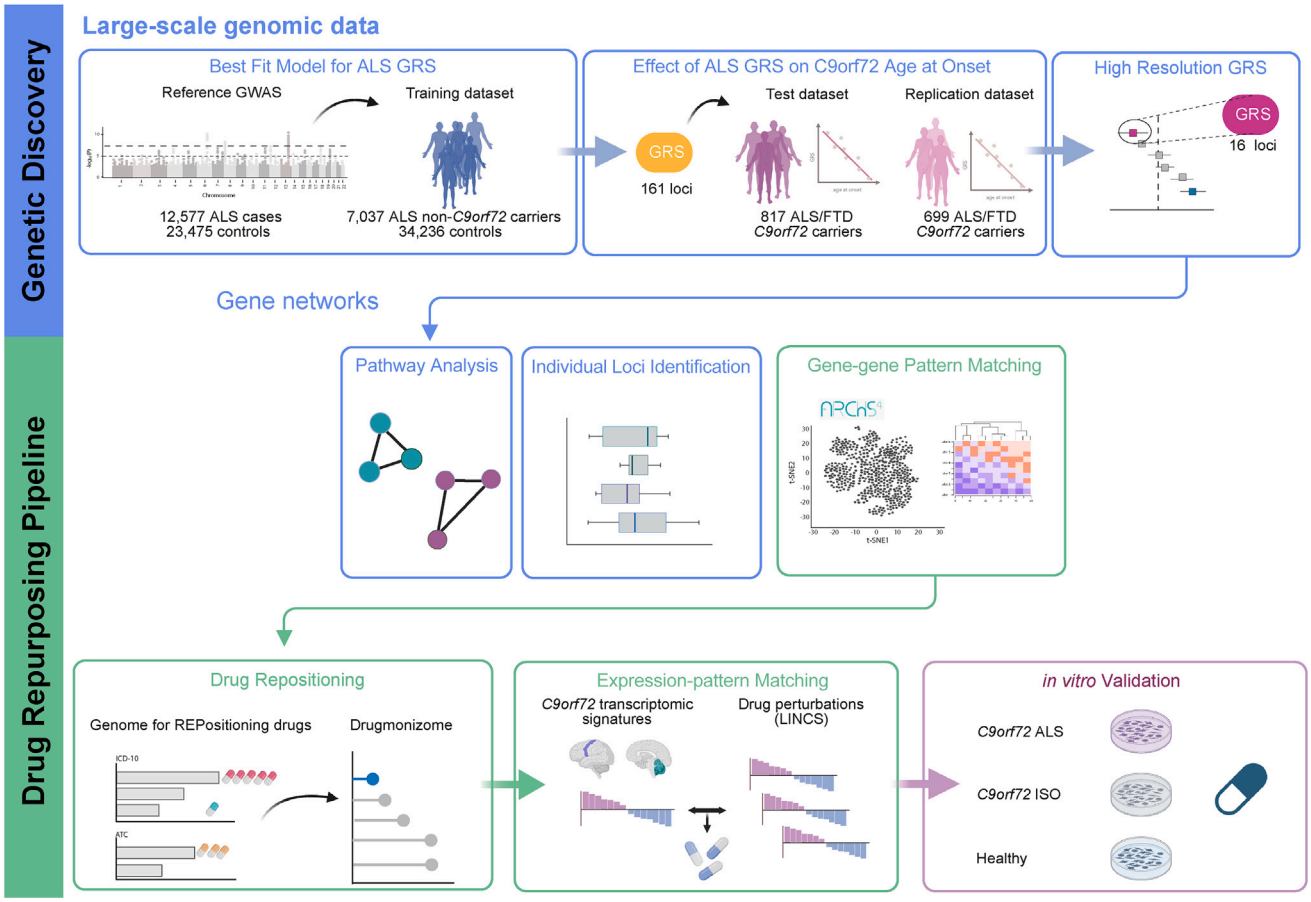


Figure 1. Schematic illustration of the analytical workflow

The genetic risk score for sporadic ALS was generated using large cohorts as the reference and training sets. This general ALS genetic risk score was then calculated for a sizable cohort of *C9orf72* carriers. Follow-up analyses included pathway analysis and the identification of individual loci with a major contribution to the age at onset. Using the information obtained from these genetic analyses, we performed drug repurposing based on gene-gene-pattern matching and expression-pattern matching to identify drugs that may delay symptom onset among *C9orf72* carriers. *In vitro* drug validation confirmed the neuroprotective effect of the drug nominated by this approach.

Sporadic ALS risk variants modify the onset age among *C9orf72*-related ALS/FTD

We built a polygenic risk profile for general ALS using 7,037 cases not carrying the *C9orf72* repeat expansion and 34,236 healthy controls. The best-fit model included 161 SNPs (odds ratio for the model = 1.125, 95% confidence interval [CI] = 1.093–1.156, $p = 1.5 \times 10^{-16}$). Figure 2 and Table S1 describe the genetic variants and mapped genes that make up the polygenic risk score for general ALS.

Next, using the 161 SNPs, we investigated the association of the general ALS genetic risk score with age at onset in a cohort of 1,516 ALS/FTD *C9orf72* expansion carriers. We found the risk of general ALS (i.e., independent of the *C9orf72* repeat expansion) was significantly associated with onset age among the *C9orf72* expansion carriers (meta-analysis $p = 1.5 \times 10^{-3}$, $\beta = -0.781$, 95% CI = −1.263 to −0.299; Figures 3A–3C; Table S2). We observed similar results when analyzing ALS *C9orf72* and FTD *C9orf72* carrier patients alone (Figure S1). In contrast, the general ALS genetic risk score did not influence the age at onset among patients without *C9orf72* repeat expansions ($p = 0.437$, $\beta = 0.115$, 95% CI = −0.175 to 0.404; Figures 3C and S2; Table S2).

A subset of 16 SNPs drives the early-onset age among *C9orf72*-related ALS/FTD

Having established that the general ALS risk influenced the age at onset among *C9orf72* patients, we next determined which of the 161 SNPs in the model was driving the effect.[21] To do this, we performed a leave-one-out analysis to evaluate the contribution of each SNP to the onset age. This enabled us to rank the 161 SNPs and group the variants into 10 deciles (Table S3). We recalculated the genetic risk scores using the SNPs from each decile and used a linear regression model to test these refined genetic risk scores for association with age at onset.[22–27] Figure S3 shows a schematic representation of this approach.

Decile 10 contained the 16 SNPs with the most significant association with earlier onset age (Figures 4A and 4B). A

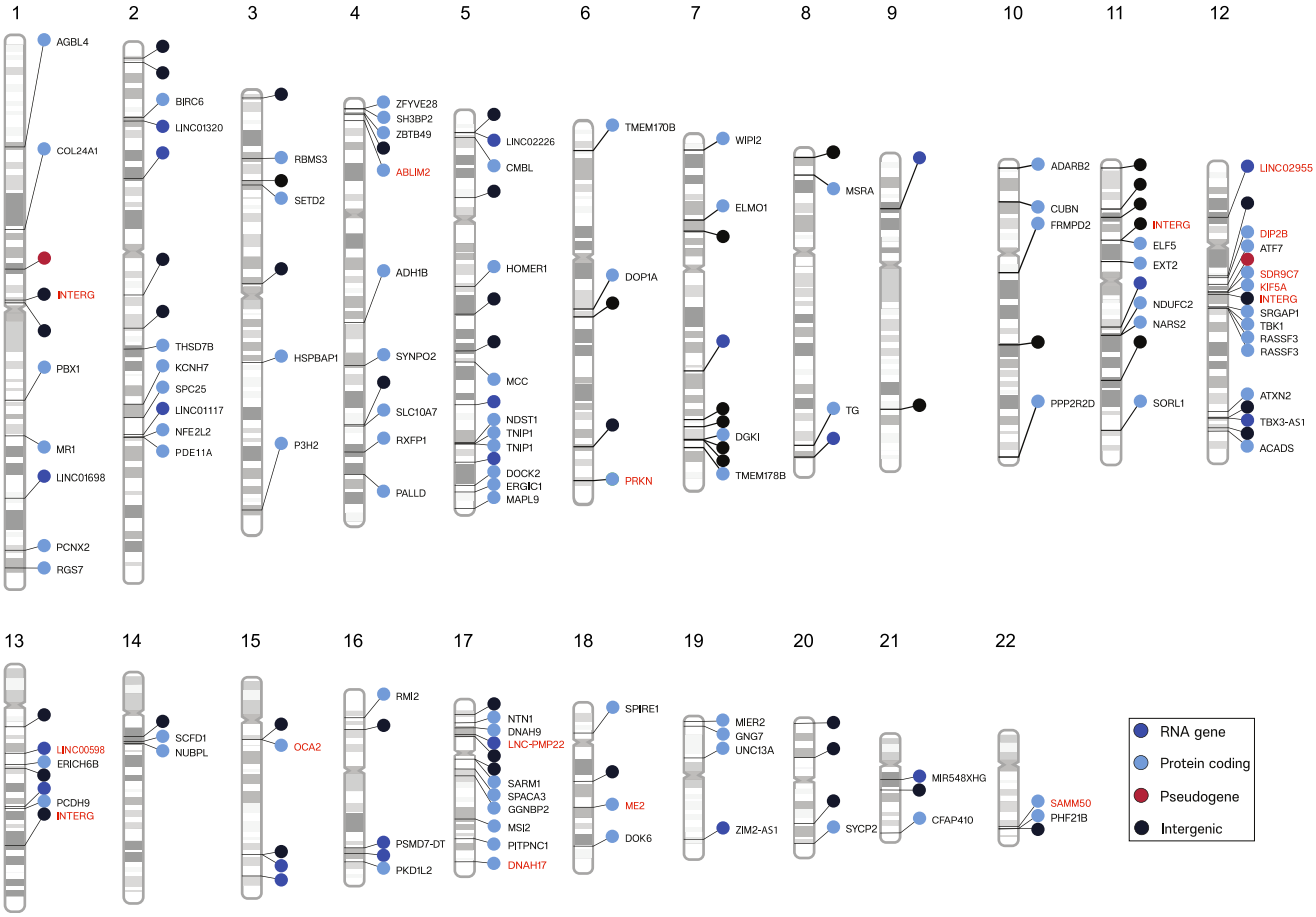


Figure 2. Genetic variants influencing symptom onset age among ALS/FTD patients carrying the *C9orf72* repeat expansion

The ideograms show the 161 ALS genetic risk loci making up the general ALS polygenic risk score. Labels with red text denote the 16 SNPs making up decile 10. The nearest genes to the variants are displayed. The colors of the circles correspond to the gene type: dark blue, RNA gene; light blue, protein-coding gene; red, pseudogene; black, intergenic. The numbers at the top indicate the chromosome. Interg, intergenic.
See also Table S1.

1-SD increase in the genetic risk score of the decile 10 SNPs corresponded to a decrease of 2.17 years (95% CI = 1.52–2.81) in the age at onset, and there was a 4-SD difference between *C9orf72* carriers at the extremes of the genetic risk score distribution. This difference implies that individuals at the highest end of the genetic risk score distribution developed the disease, on average, 8 years (8.68 years, 95% CI = 6.08–11.24; Figure 4B) earlier than their counterparts at the lowest end. In contrast, these loci did not alter the onset age among ALS patients not carrying the *C9orf72* repeat expansion (Figure S4).

Cytoskeletal and axonal transport pathways influence *C9orf72* age at onset

Enrichment analysis of the SNPs in decile 10 revealed that the cytoskeletal and axonal transport pathways influenced the age at onset among patients carrying the *C9orf72* repeat expansion (Figures 4C and S5; Table S4). These data suggested that these biological processes are essential in determining the onset of *C9orf72*-related disease.

The two-step design of our drug-repurposing pipeline

We used our genetic data to perform a drug-repositioning analysis to discover medications that may delay the age at disease onset among *C9orf72* patients. This approach harnesses our insights into the genetic factors influencing the variable age at onset in *C9orf72*-related disease and existing databases containing information on drug effects.[28] A critical advantage of our pipeline is that it identifies drugs with broad, heterogeneous effects for use in complex human diseases like neurodegenerative conditions. Crucially, this approach is not based on a single change within the cell but instead relies on gene-pattern and expression-pattern matching to select medications that correct the disrupted networks.

The pipeline comprises two complementary parts (see Figure 1 for a graphical representation of these steps). In the first step, we performed gene-drug-pattern matching. To do this, we used the *g:SNPense* function of g:Profiler2[29] to identify the genes related to the SNPs in decile 10. These genes were then used as search terms (defined as seed genes in Table S5) in the Geneshot web server (accessed October

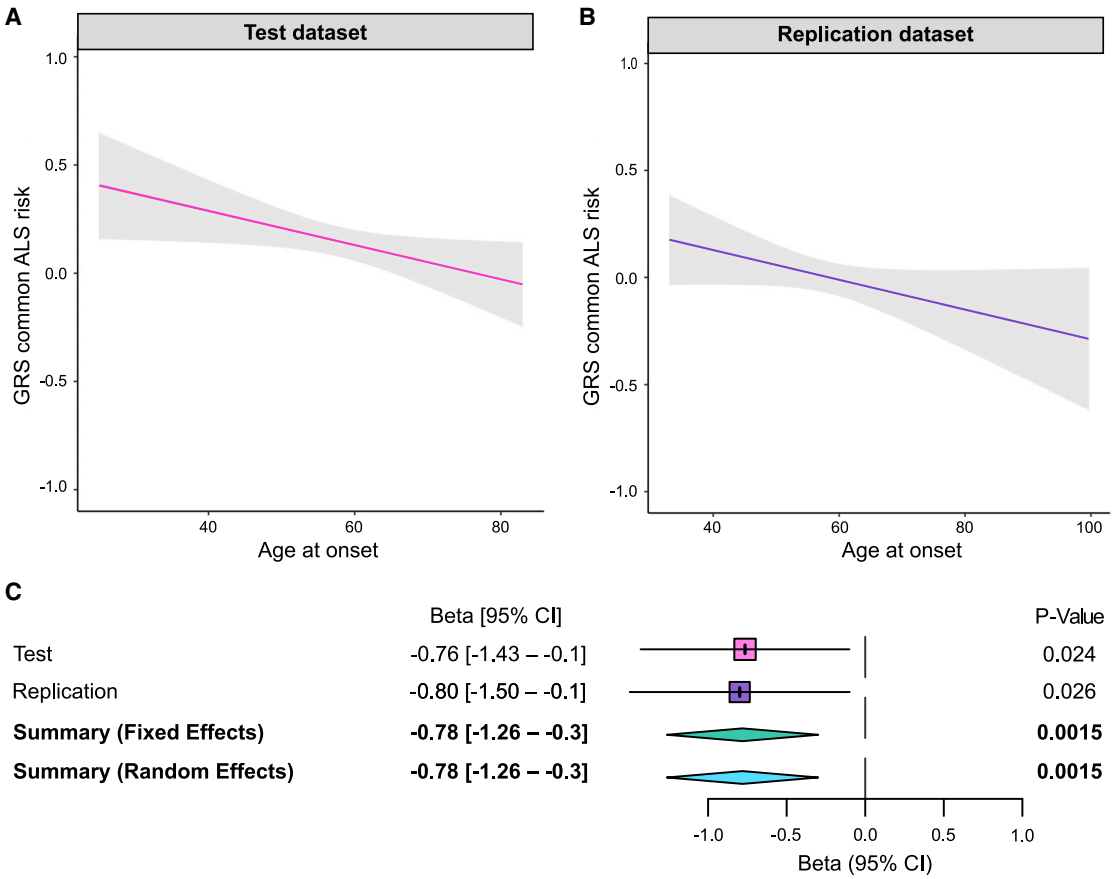


Figure 3. The ALS genetic risk score significantly influences onset age in *C9orf72* carriers

(A and B) The regression lines show the association of the ALS genetic risk scores and age at onset in (A) the test dataset (n = 817 ALS/FTD *C9orf72* carriers, p = 0.024, β = −0.765, 95% CI = −1.429 to −0.101) and (B) the replication dataset (n = 699, p = 0.026, β = −0.799, 95% CI = −1.499 to −0.099). The shadow areas represent the 90% confidence interval of the regression model.

(C) The forest plot shows the results of the meta-analysis of the test and replication datasets ($p = 1.5 \times 10^{-3}$, β = −0.781, 95% CI = −1.263 to −0.299).

See also Figures S1 and S2 and Table S2.

2021) to identify a broader list of functionally related genes.[30] This database identifies genes associated with the search term based on their co-occurrence in publications and gene-gene similarity from human RNA sequencing (RNA-seq) data (ARCHS4).[30] The resulting list of genes plus the seed genes (Table S5) was then used as input for the Genome for REPositioning drugs (GREP) analysis (version 1.0.0), which is a software package designed to identify drugs that target the gene set based on their enrichment in clinical indication categories.[31] The gene-drug software used Fisher's exact test to perform pharmacological enrichment and output the names of drugs associated with the gene set.

In the second step of our pipeline, we used gene expression-pattern matching to refine and narrow the candidate drug list. In essence, this approach identifies compounds that reverse the transcription patterns observed in brain tissue obtained from *C9orf72* patients (n = 44 cases and 76 healthy controls). This method is widely used in drug repositioning to assess how well drugs can counteract disease-related gene expression patterns by comparing their effects to a disease gene signature.[32] To do this, the drug perturbations were queried using the Library of Integrated Network-Based Cellular Signatures database,[33] containing the differential expression analysis of 12,328 genes from 8,140 compound treatments of 30 cell lines. A bidirectional weighted Kolmogorov-Smirnov enrichment statistic test of gene expression ranks in the disease and the expression values of the drug signatures was used to assign a connectivity map (CMap) score to each drug, reflecting the degree to which the drug "flips" the gene expression signature of the disease.

Step 1: Drug prioritization using gene-drug-pattern matching

The genomic-based GREP analysis nominated 52 medications approved for human use that could be repurposed to delay onset among *C9orf72* carriers (Figure 5A; Table S6). To investigate the targets of the nominated medications, we explored the significant biomedical terms associated with these drugs in the Drugmonizome database (accessed October 12, 2022).[34] These therapies were enriched for CNS targets and are typically prescribed for anxiety disorders and epilepsy; exploring their mechanisms

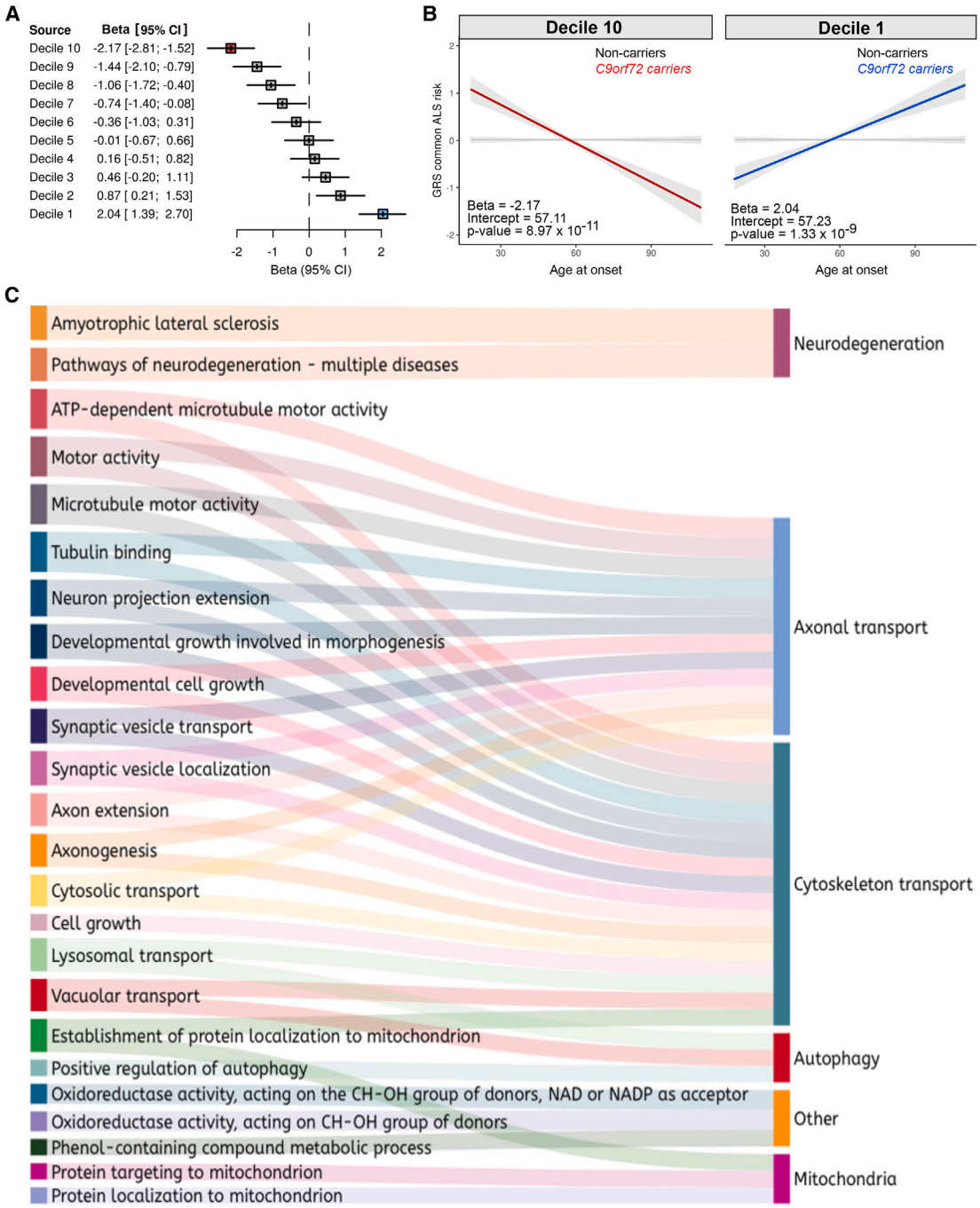


Figure 4. Contribution of individual SNPs to age at symptom onset among *C9orf72* patients

(A) The forest plot shows the effect size of each decile obtained by ranking the 161 individual SNPs based on their effect on age at onset in *C9orf72* patients.

(B) The regression lines show the association between the ALS genetic risk scores and age at onset in 817 ALS/FTD *C9orf72* carriers based on the 16 SNPs of decile 10 (n = 817 ALS/FTD *C9orf72* carriers, $p = 8.97 \times 10^{-11}$, $\beta = -2.17$, 95% CI = −2.81 to −1.52). The shadow area represents the 90% confidence interval of the regression model.

(C) Sankey diagram showing the functional enrichment of decile 10 genes, based on Gene Ontology terms. Only gene lists that contain between 5 and 500 genes were selected for the analysis. The significant threshold was an false discovery rate (FDR)-corrected $p < 0.05$.

See also Figures S3–S5 and Tables S3, S4, and S7.

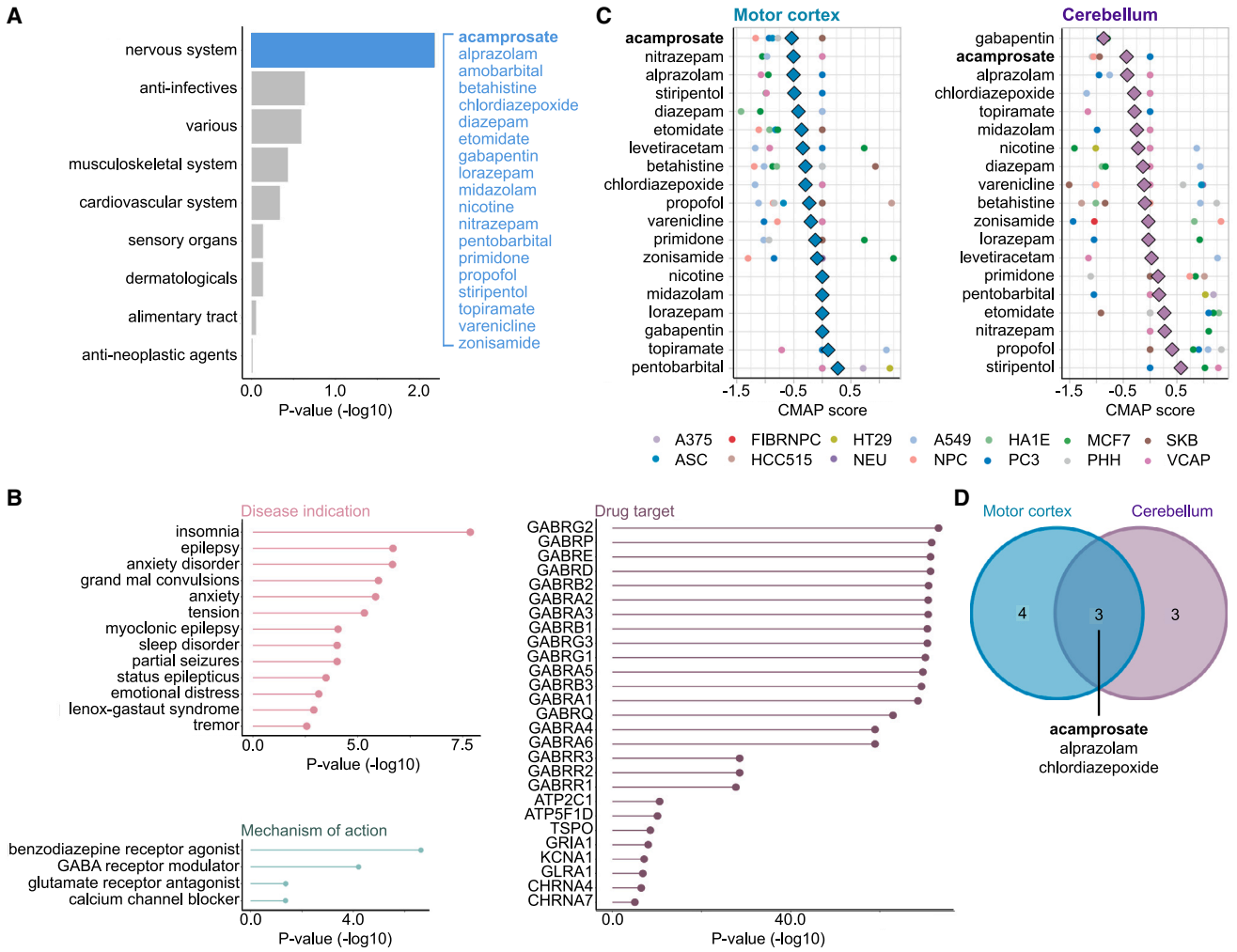


Figure 5. Repurposing drugs to delay onset among *C9orf72* carriers

The figure shows the results obtained from our drug-repositioning pipeline.

(A) Enriched terms derived from the GREP software package, which is based on the Anatomical Therapeutic Chemical classification. Blue indicates the significant "other nervous system drugs" category. Some of the drugs within this category are shown (see Table S6 for a complete list).

(B) Lollipop plots depict drug enrichment analysis for different categories, such as disease indication, gene target, and mechanism of action. Information was obtained from the Drugmonizome database, and the x axis depicts the enrichment corrected *p* value, which uses Bonferroni correction for the disease indication and the gene target plots and FDR for the mechanism of action.

(C) Drug perturbation data were obtained from the LINCS database. The graphs show the CMap scores for the selected drugs across cell types. CMap scores are determined using a bidirectional weighted Kolmogorov-Smirnov enrichment statistic test, which compares gene expression changes in the disease and drug signatures to quantify the extent to which the drug effectively reversed (flipped) the gene expression signature associated with the disease. Lower scores indicate a more substantial potential for therapeutic effectiveness. An average CMap (diamond shape) was calculated using the normalized connectivity score to evaluate the overall effect of each drug across the tested cell lines. Drugs with a reversal potential were selected if (1) they depicted a negative average CMap and (2) they showed a negative or neutral (measured as 0) CMap score for each cell line (circle shape). A375, ASC, FIBRNPC, HCC515, HT29, NEU, A549, NPC, HA1E, PC3, MCF7, PHH, SKB, and VCAP refer to the cell line types available in the LINCS database (see STAR Methods for details).

(D) The Venn diagram shows the drugs that fulfilled these criteria in the motor cortex and cerebellum. Of these, acamprosate was selected for additional *in vitro* validation.

See also Figures S5 and S6.

of action revealed the γ-aminobutyric acid (GABA) receptor as potentially relevant (Figure 5B).

Step 2: Drug prioritization using expression-pattern matching

The drug-disease expression-pattern matching Library of Integrated Network-Based Cellular Signatures (LINCS) analysis found that 3 of our 52 selected drugs demonstrated reversal of the *C9orf72* transcriptomic disease signature across multiple cell lines: acamprosate, chlordiazepoxide, and alprazolam (Figures 5C and 5D). Acamprosate was chosen from this list for the following reasons. First, acamprosate demonstrated neuroprotective properties in a SOD1^{G93A} rat spinal cord model of

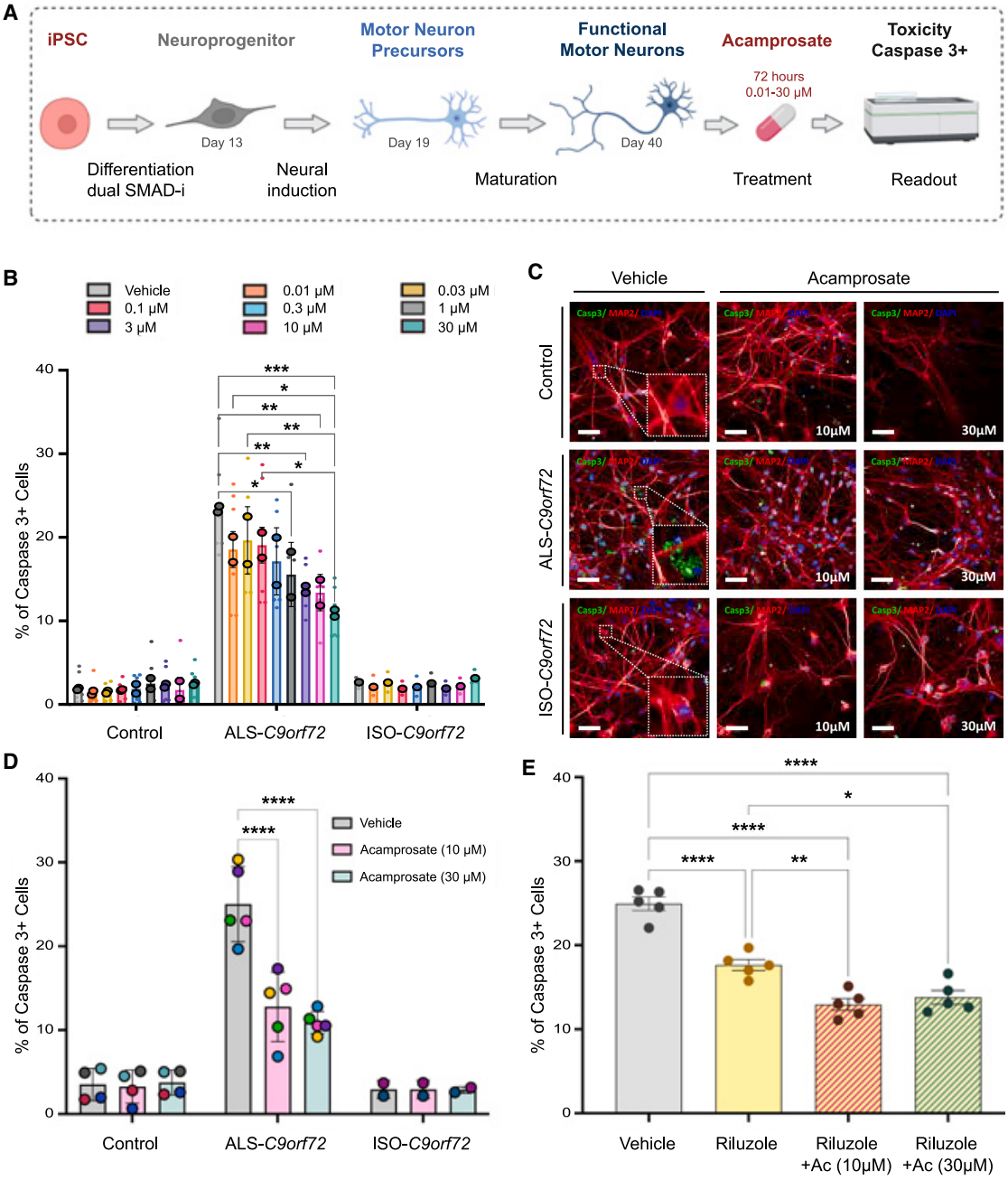


Figure 6. Acamprosate is neuroprotective in iPSC-derived motor neurons from *C9orf72* patients

(A) Schematic representation of the experiments to validate the effect of acamprosate.

(B) The bar graph depicts the percentage of cells showing cleaved caspase-3 (caspase-3^{+} cells) after acamprosate treatment. Minor dots represent biological replicates, averaging from three technical replicates each. In contrast, the bordered dots represent the mean effect in iPSC-derived motor neurons from two healthy donors, two *C9orf72* ALS patients, and one isogenic line. Data are mean ± SD. Comparisons within the control and the ALS-*C9orf72* groups were performed using a two-way ANOVA, with a Tukey's post hoc test ($n = 2$). Only $p < 0.05$ and comparisons with the vehicle group are displayed in the graph.

(C) Representative images of the motor neurons showing cleaved caspase-3 staining (green), MAP2 staining (red), and DAPI (blue). Scale bar, 50 μm. All molecular phenotypes were confirmed in a minimum of 3 technical replicates, and at least 25 fields were randomly selected and scanned per well of a 96-well plate in triplicate.

(D) Acamprosate effective doses (10 and 30 μM) were confirmed in additional iPSC-derived motor neurons, totaling cells from four healthy donors, five *C9orf72* ALS patients, and two isogenic lines. Each point represents the mean effect per cell line. Data are mean ± SD. Comparisons within the control, the ALS-*C9orf72*, and the ISO-*C9orf72* groups were performed using a two-way ANOVA, with Tukey's post hoc test (control, $n = 4$; ALS-*C9orf72*, $n = 5$; ISO-*C9orf72*, $n = 2$). $p < 0.05$ are annotated.

(legend continued on next page)

ALS.[35] Second, and perhaps most important, in contrast to chlordiazepoxide and alprazolam, acamprosate is not associated with sedation or respiratory depression, which are potentially hazardous side effects in the ALS population.[36,37]

Acamprosate is neuroprotective in a patient-derived *C9orf72* cell model

We assessed dose-response curves and potential levels of toxicity of acamprosate in motor neurons derived from induced pluripotent stem cells (iPSCs) obtained from two ALS patients carrying the *C9orf72* repeat expansion, two healthy individuals, and one isogenic control line (see Figure 6A for an outline of the workflow and Figure S6 for a representative immunohistochemistry staining of cell cultures). The drug did not exert a toxic effect on *C9orf72* or healthy motor neurons, even at the higher doses (30 μM for 72 h, viability measured using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assay; Figures S7 and S8). Using cleaved caspase-3, an established biochemical proxy for cell survival in *C9orf72*[38–40] and other genetic forms of ALS,[41] acamprosate strongly reduced cell death in *C9orf72*-derived motor neurons; the percentage of cleaved caspase-3^{+} cells was 1.8-fold lower at 10 μM, and 2.3-fold lower at 30 μM (half-maximal effective concentration = 0.271 μM; Figures 6B, 6C, and S9).

Based on this initial screening data, we selected the highest efficacious doses (10 and 30 μM) for additional testing. We confirmed the efficacy of acamprosate at protecting motor neurons in additional cell lines (five ALS patients carrying the *C9orf72* repeat expansion, four healthy individuals, and two isogenic control lines; Figure 6D). Again, the drug had minimal effects on healthy motor neurons and isogenic controls (Figures 6B and 6C).

Effect of acamprosate is comparable to riluzole, the current standard of care

Next, we compared the efficacy of acamprosate in preventing cell death to riluzole, the most widely prescribed drug for ALS, which extends life expectancy by an average of 3–6 months.[42] As expected, riluzole treatment reduced cell death in *C9orf72* motor neurons (1.4-fold decrease at 10 μM; Figure 6E). Notably, acamprosate exerted an average of 30% higher protection, either alone (Figure 6D) or combined with riluzole, than that observed with riluzole (Figure 6E).

DISCUSSION

We used our genetic and transcriptomic information to nominate medications that could be repurposed to delay symptom onset among *C9orf72* carriers, and our data support further investigation of acamprosate as a potential treatment for this common form of neurodegeneration. Acamprosate is a US Food and Drug Administration (FDA)-approved medication with a favorable safety profile that is prescribed to maintain alcohol abstinence. The drug modulates glutamate receptors,[43] a mechanism overlapping with riluzole, a commonly prescribed medication for ALS patients that slows clinical progression, at least in part, through its anti-glutamate properties.[42,44] Interestingly, acamprosate, alone or in combination with riluzole, exceeded the neuroprotective properties of riluzole in our cell-based model.

Our work supports previous research suggesting that a combination of baclofen and acamprosate may be a plausible therapeutic strategy for ALS; the combination prevented TDP-43 stress granule formation in a human osteosarcoma U2OS cell line overexpressing human TDP-43 and displayed neuroprotective effects in primary motor neurons derived from SOD1^{G93A} rat embryos.[35] However, our work nominating acamprosate was an independent effort based on large-scale human genomic data. It is also the first time that acamprosate has been nominated as a personalized treatment for *C9orf72*-related ALS-FTD. These studies offer evidence that acamprosate may benefit ALS patients, warranting further consideration for the development of acamprosate as a treatment within this population. Our cell-based assays also suggested that a combination of riluzole and acamprosate treatment is more effective than acamprosate alone. Such combination therapy has emerged as a promising approach in neurological diseases, perhaps reflecting the need to ameliorate multiple pathways to produce a clinical benefit.[45]

Drug repositioning has become an attractive option for drug discovery, as using de-risked compounds lowers development costs and shortens time lines.[46] Systematic approaches based on genetic and transcriptomic data integration are also promising.[47,48] An emerging theme in the pharmaceutical industry is that drugs targeting proteins and pathways with genetic evidence are more likely to succeed.[16,17] The effect of acamprosate on *C9orf72*-related disease may be due to its known glutaminergic and GABAergic properties, but this is currently unproven. Instead, our work reinforces the value of genomic data to rapidly nominate drugs for repurposing in situations where the underlying molecular mechanisms of disease are complex and where the mode of action of a drug is unclear. In support of this data-driven, mechanism-free approach, we highlight that acamprosate was first approved for use in alcohol-dependence syndrome 35 years ago, and yet precisely how the drug works in this condition remains speculative. This situation is hardly unique among CNS drugs. Despite decades of patient use, the exact molecular mechanism by which riluzole confers its survival advantage in patients with ALS is unknown. Delaying clinical trials until the mechanism of action of acamprosate is determined may be unnecessary, especially considering the current uncertainty about how the *C9orf72* repeat expansion causes disease and the lack of reliable animal models.

Our findings also partially explain the variable age at onset observed in *C9orf72* patients and confirm that genetic modifiers are crucial in determining the onset age in this patient population.[6,49] Individuals with *C9orf72*-related disease with a high genetic risk of sporadic ALS developed symptoms approximately

(E) The graph shows the percentage of cleaved caspase-3^{+} motor neurons derived from *C9orf72* patients treated with riluzole (10 μM) and riluzole plus acamprosate (10 and 30 μM). Dots represent the mean effect of each line. Data are mean ± SD. Comparisons were performed using a one-way ANOVA with Tukey's post hoc test ($n = 5$). $p < 0.05$ are annotated as follows: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.001$, *****$p < 0.0001$.
See also Figures S6–S9 and Table S10.

3 years earlier than patients with a low genetic risk. Interestingly, the genetic risk attributable to general ALS did not affect onset age among patients not carrying the *C9orf72* repeat expansion, as confirmed recently in a multicenter study.[50] Our findings support the hypothesis that ALS is a multistep process[51–54] and suggest that a personalized approach will be necessary to treat the various genetic forms of ALS and neurodegenerative disorders.

We identified ''neuronal cytoskeleton'' and ''axonal transport'' as the central pathways influencing the age at onset among *C9orf72* carriers. Interestingly, arginine-containing dipeptides generated from *C9orf72* repeat expansions are known to impede microtubule-based motility and axonal transport machinery in cells.[55] This pathogenic mechanism involves physical interactions of the dipeptides with the kinesin-containing motor complexes. Reinforcing this research, we found the rs113247976 variant within this gene to be among the most influential on onset age, with *C9orf72* patients carrying the *KIF5A* variant experiencing symptoms nearly 3.5 years earlier than patients who did not carry this variant (Table S7). These observations indicate that axonal transport is centrally involved in *C9orf72*-related neurodegeneration, and this cellular process may be a credible therapeutic target. Notwithstanding, our data also suggest that multiple cellular processes are disrupted by *C9orf72* mutation, underscoring the importance of our more comprehensive genomic approach to drug repurposing.

Previous research has investigated disease modifiers of *C9orf72*. However, these efforts were focused on individual loci and ignored the collective contribution of the ALS polygenic inheritance.[19,56–58] This knowledge gap was primarily due to the limited availability of large *C9orf72* datasets required for this type of analysis. In contrast, our approach was based on genome-wide research involving 1,516 ALS/FTD *C9orf72* carriers and 7,037 ALS non-carriers, making it one of the largest cohorts reported to date. We also focused on genomic variants outside the *C9orf72* locus, which are likely more amenable to therapeutic intervention with small molecules than the repeat expansion itself.

Our research underscores the critical contribution of common risk factors as modifiers of monogenic forms of ALS.[35] Unfortunately, data were unavailable to allow a similar analysis of other highly penetrant genetic mutations that cause ALS, such as *SOD1* and *TBK1*. However, the onset age among *SOD1* mutation carriers, such as the Ala5Val variant, is much more circumscribed than the variable age penetrance observed among *C9orf72* patients.[59] Nevertheless, our study provides a foundation for future research exploring how the common risk factors of ALS affect other monogenic forms of the disease and leveraging this information to repurpose medications.

In conclusion, we found that the risk of sporadic ALS is a critical determinant of the age at symptom onset among patients carrying the *C9orf72* repeat expansion. We identified various pathways and genes that play pivotal roles in the underlying cellular processes determining onset timing. We integrated our genomic data with transcriptomics to nominate drugs that could be repurposed to delay disease onset, an approach that has previously prioritized medications for chronic conditions such as obesity[60] and osteoporosis.[61] Acamprosate emerged as the primary contender from these analyses, and we subsequently confirmed that this oral, inexpensive medication reduces *C9orf72* motor neuron death in cell-based assays. Future research will explore how acamprosate may be exerting this effect. Nevertheless, our work pinpoints acamprosate for future exploration as a treatment to slow the manifestation of symptoms in this common genetic form of neurodegeneration. Our innovative and multidisciplinary approach could also fill the gap in therapeutic discovery in other complex neurodegenerative conditions, and we have made our computational pipeline publicly available to facilitate such work.

Limitations of the study

Our study has limitations. There was only a 3-year difference in onset age between the patients with the highest and the lowest genetic risk of sporadic ALS, representing a modest portion of the 40-year spread reported for *C9orf72* onset age. Despite this, the magnitude of this effect was comparable to the results reported for the *TMEM106B* locus in *C9orf72* patients.[19] Interestingly, the outcome we observed was independent of this locus. Our capacity to predict age at onset in this group of patients, as well as our ability to assess the contribution of ALS genetic risk on disease progression, will improve as larger datasets with comprehensive phenotypic data become available.

Another limitation of our study is that the underlying mechanism of the neuroprotective effect of acamprosate in *C9orf72*-derived motor neurons remains unknown. Although exploring such mechanisms was beyond the scope of this paper, future research should carefully address this gap, especially considering the implications it may have for the broader ALS patient population. Previously published data suggested that acamprosate may also have a neuroprotective effect on non-*C9orf72* ALS patients. A combination of acamprosate and baclofen (PXT864) has been proposed as a treatment for sporadic ALS. However, no clinical trial has been registered, and the company has discontinued its development, citing financial reasons.[62] We anticipate our work will lead to a renewed interest in the otherwise overlooked acamprosate as a treatment for this uniformly fatal neurodegenerative disease. Such repurposing of FDA-approved drugs for unmet medical needs has strategic advantages that shorten the development time line—preclinical safety testing can be shortened as the safety profile and pharmacokinetic profile of existing drugs are already established. The standard daily dose (1,998 mg) of acamprosate is known to cross the blood-brain barrier and alter brain physiology; patients with alcohol dependence who were treated with acamprosate showed highly significant suppression of glutamate levels in the anterior cingulate gyrus over 4 weeks.[63] These observations provide a quantitative biomarker of target engagement. It also supports the hypothesis that the protective effect observed in our cell-based assays at the relatively low dose of 1 μM will translate into human patients at the well-tolerated standard dose.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Bryan Traynor (traynorb@mail.nih.gov).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The summary statistics of the reference dataset genome-wide association study (GWAS) are publicly available at http://databrowser.projectmine.com/. The individual-level data for the training and test datasets are available on dbGaP (accession nos. phs000101.v5.p1 and phs001963.v1.p1). The data for the replication dataset were from a different study (principal investigator: Christopher Shaw, King's College London) and are available upon reasonable request. RNA-seq data from the motor cortex were obtained from the New York Genome Center, and RNA-seq data from the cerebellum were obtained from GEO (GEO: GSE67196). The figures were created using BioRender and Inkscape. The Sankey diagram was created using Visual Paradigm. The programming code used in this paper is available at https://github.com/sarasaezALS/C9orf72_AAO and https://zenodo.org/records/13259646 (https://doi.org/10.5281/zenodo.13259646) to facilitate the application of this methodology to other disorders. The individual-level polygenic risk scores generated for patients with the *C9orf72* mutation are available at https://zenodo.org/uploads/13769448. The connectivity scores for the drugs evaluated as treatments for *C9orf72*-related disease are available at https://zenodo.org/records/13769483.

CONSORTIA

The members of the International ALS Genomics Consortium are Robert H. Baloh, Robert Bowser, Christopher B. Brady, Alexis Brice, James Broach, William Camu, Ruth Chia, Adriano Chiò, John Cooper-Knock, Daniele Cusi, Jinhui Ding, Carsten Drepper, Vivian E. Drory, Travis L. Dunckley, Eva Feldman, Mary Kay Floeter, Pietro Fratta, Glenn Gerhard, J. Raphael Gibbs, Summer B. Gibson, Jonathan D. Glass, Stephen A. Goutman, John Hardy, Matthew B. Harms, Terry D. Heiman-Patterson, Lilja Jansson, Janine Kirby, Hannu Laaksovirta, John E. Landers, Francesco Landi, Isabelle Le Ber, Serge Lumbroso, Claire Guissart, Daniel JL. MacGowan, Nicholas J. Maragakis, Gabriele Mora, Kevin Mouzat, Liisa Myllykangas, Richard W. Orrell, Lyle W. Ostrow, Stuart Pickering-Brown, Erik P. Pioro, Stefan M. Pulst, John M. Ravits, Alan E. Renton, Wim Robberecht, Ekaterina Rogaeva, Jeffrey D. Rothstein, Erika Salvi, Sonja W. Scholz, Michael Sendtner, Pamela J. Shaw, Katie C. Sidle, Zachary Simmons, David J. Stone, Pentti J. Tienari, Bryan J. Traynor, John Q. Trojanowski, Juan C. Troncoso, Miko Valori, Philip Van Damme, Vivianna M. Van Deerlin, Ludo Van Den Bosch, and Lorne Zinman. The members of the ITALSGEN Consortium are Stefania M. Angelocola, Francesco P. Ausiello, Marco Barberis, Ilaria Bartolomei, Stefania Battistini, Enrica Bersano, Giulia Bisogni, Giuseppe Borghero, Maura Brunetti, Corrado Cabona, Andrea Calvo, Fabrizio Canale, Antonio Canosa, Teresa A. Cantisani, Margherita Capasso, Claudia Caponnetto, Patrizio Cardinali, Paola Carrera, Federico Casale, Adriano Chiò, Tiziana Colletti, Francesca L. Conforti, Amelia Conte, Elisa Conti, Massimo Corbo, Stefania Cuccu, Eleonora Dalla Bella, Eustachio D'Errico, Giovanni DeMarco, Raffaele Dubbioso, Carlo Ferrarese, Pilar M. Ferraro, Massimo Filippi, Nicola Fini, Gianluca Floris, Giuseppe Fuda, Salvatore Gallone, Giulia Gianferrari, Fabio Giannini, Maurizio Grassano, Lucia Greco, Barbara Iazzolino, Alessandro Introna, Vincenzo La Bella, Serena Lattante, Giuseppe Lauria, Rocco Liguori, Giancarlo Logroscino, Francesco O. Logullo, Christian Lunetta, Paola Mandich, Jessica Mandrioli, Umberto Manera, Fiore Manganelli, Giuseppe Marangi, Kalliopi Marinou, Maria Giovanna Marrosu, Ilaria Martinelli, Sonia Messina, Cristina Moglia, Maria Rosaria Monsurrò, Gabriele Mora, Lorena Mosca, Maria R. Murru, Paola Origone, Carla Passaniti, Cristina Petrelli, Antonio Petrucci, Angelo Pirisi, Susanna Pozzi, Maura Pugliatti, Angelo Quattrini, Claudia Ricci, Giulia Riolo, Nilo Riva, Massimo Russo, Mario Sabatelli, Paolina Salamone, Marco Salivetto, Fabrizio Salvi, Marialuisa Santarelli, Luca Sbaiz, Riccardo Sideri, Isabella Simone, Cecilia Simonini, Rossella Spataro, Raffaella Tanel, Gioacchino Tedeschi, Anna Ticca, Antonella Torriello, Stefania Tranquilli, Lucio Tremolizzo, Francesca Trojsi, Rosario Vasta, Veria Vacchiano, Giuseppe Vita, Paolo Volanti, Marcella Zollino, and Elisabetta Zucchi. The members of the SLAGEN Consortium are Vincenzo Silani, Isabella Fogh, Nicola Ticozzi, Antonia Ratti, Cinzia Tiloca, Silvia Peverelli, Cinzia Gellera, Giuseppe Lauria Pinter, Franco Taroni, Viviana Pensato, Barbara Castellotti, Giacomo P. Comi, Stefania Corti, Roberto Del Bo, Cristina Cereda, Mauro Ceroni, Stella Gagliardi, Lucia Corrado, Letizia Mazzini, Gianni Sorarù, Flavia Raggi, Gabriele Siciliano, Costanza Simoncini, Annalisa Lo Gerfo, Massimiliano Filosto, Maurizio Inghilleri, and Alessandra Ferlini. The members of Project MinE ALS Sequencing Consortium are Philip Van Damme, Philippe Corcia, Philippe Couratier, Patrick Vourc'h, Orla Hardiman, Russell McLaughlin, Marc Gotkine, Vivian Drory, Nicola Ticozzi, Vincenzo Silani, Jan H. van den Veldink, Leonard H. Berg, Mamede de Carvalho, Jesus S. Mora Pardina, Monica Povedano, Peter Andersen, Markus Weber, Ayşe Nazlı Başak, Ammar Al-Chalabi, Chris Shaw, Pamela J. Shaw, Karen E. Morrison, John E. Landers, and Jonathan D. Glass. The members of the American Genome Center are Adelani Adeleye, Camille Alba, Dagmar Bacikova, Clifton L. Dalgard, Daniel N. Hupalo, Elisa McGrath Martinez, Anthony R. Soltis, Gauthaman Sukumar, Coralie Viollet, and Matthew D. Wilkerson.

ACKNOWLEDGMENTS

This work was supported in part by the Intramural Research Program of the National Institutes of Health, the National Institute on Aging (ZIA-AG000933, ZIA-AG000934), the National Institute of Neurological Disorders and Stroke (NINDS), the National Center for Advancing Translational Sciences, and by Merck Sharp & Dohme Corp. (a subsidiary of Merck & Co., Kenilworth, NJ). B.J.T. received additional support from the Centers for Disease Control and Prevention, the Muscular Dystrophy Association, Microsoft Research, the Packard Center ALS Research at Johns Hopkins, and the ALS Association. L.F. is supported by the UK Medical Research Council (MR/V027735/1 and MR/V000470/1) and the Motor Neurone Disease Association (MND019801). A.N.B., C. Tunca, and E.B. gratefully acknowledge the use of the services and facilities of the Koç University Research Center for Translational Medicine (KUTTAM), funded by the Presidency of Turkey, Head of Strategy and Budget. R.S. is supported by the Barrow Neurological Foundation. This work was also generated within the European Reference Network for Neuromuscular Diseases (ERN EURO-NMD, https://ern-euro-nmd.eu). This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health. The authors would like to thank the Project MinE GWAS Consortium. This study also used genotype and clinical data from the Wellcome Trust Case Control Consortium and the HyperGenes Consortium. We thank the patients and research subjects who contributed samples for this study, Dr. Ruth Pfeiffer, Dr. Justin Kwan, the Laboratory of Neurogenetics staff, and the NINDS Neurodegeneration Clinic staff for their collegial support. The authors also acknowledge the Target ALS Human Post-mortem Tissue Core, the New York Genome Center

for Genomics of Neurodegenerative Disease, the ALS Association, and the TOW Foundation. Additional acknowledgments may be found in the supplemental information.

AUTHOR CONTRIBUTIONS

S.S.-A. and B.J.T. conceived and designed the study, acquired and analyzed the data, and wrote the paper. C.d.S.S., R. Chia, S.N.B., I.L., R.H., J.L., C.B., J.D., J.R.G., A.J., R.D., V.P., S.P., L.C., J.J.F.A.v.V., W.v.R., C. Tunca, E. Bayraktar, M.X., A.I., A.S., C. Tiloca, N.T., F.V., L.M., K.K., A.A.K., S.O.-M., F.R., M.F., S.C.P., A.P., S. Gagliardi, M.I., A.F., R.V., A. Calvo, C.M., A. Canosa, U.M., M.G., J.M., G.M., C.L., R.T., F.T., P.C., S. Gallone, M.B., D.G., M.S., C.F., E.S., G.P.C., S.C., R.D.B., M.C., G.L.P., F. Taroni, E.D.B., E.B., C.J.C., S.H.L., R. Chung, H.P., K.E.M., J.C.-K., P.J.S., G.B., R.J.B.D., C.L.D., S.W.S., A.A.-C., L.H.v.d.B., R.M., O.H., C.C., G. Sorarù, S.D., S. Chandran, S. Pal, A.R., C.G., K.J., T.D.-O., N.P., T.W., A.N., G. Siciliano, V.S., A.N.B., J.H.V., W.C., J.D.G., J.E.L., A. Chiò, R.S., C.E.S., L.F., and I.F. acquired and analyzed the data and edited the manuscript.

DECLARATION OF INTERESTS

B.J.T. holds patents on clinical testing and therapeutic intervention for the hexanucleotide repeat expansion of *C9orf72*.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human participants
  - Datasets
- METHOD DETAILS
  - SNP array-based genotyping data quality control procedures and imputation
  - Whole-genome sequencing of the test dataset
  - Genetic risk score generation and computation
  - Leave-one-out analysis and decile generation
  - Pathway analysis
  - Drug prioritization via gene-drug pattern matching
  - Drug prioritization via gene expression-pattern matching
  - Drug *in vitro* validation
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2024.100679.

A video abstract is available at https://doi.org/10.1016/j.xgen.2024.100679#mmc2.

Received: April 1, 2024
Revised: July 2, 2024
Accepted: September 22, 2024
Published: October 21, 2024

REFERENCES

1. Larson, T.C., Kaye, W., Mehta, P., and Horton, D.K. (2018). Amyotrophic Lateral Sclerosis Mortality in the United States, 2011-2014. Neuroepidemiology *51*, 96–103. https://doi.org/10.1159/000488891.

2. Logroscino, G., and Piccininni, M. (2019). Amyotrophic Lateral Sclerosis Descriptive Epidemiology: The Origin of Geographic Difference. Neuroepidemiology *52*, 93–103. https://doi.org/10.1159/000493386.

3. Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron *72*, 257–268. https://doi.org/10.1016/j.neuron.2011.09.010.

4. DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron *72*, 245–256. https://doi.org/10.1016/j.neuron.2011.09.011.

5. Majounie, E., Renton, A.E., Mok, K., Dopper, E.G.P., Waite, A., Rollinson, S., Chiò, A., Restagno, G., Nicolaou, N., Simon-Sanchez, J., et al. (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. Lancet Neurol. *11*, 323–330. https://doi.org/10.1016/S1474-4422(12)70043-1.

6. Murphy, N.A., Arthur, K.C., Tienari, P.J., Houlden, H., Chiò, A., and Traynor, B.J. (2017). Age-related penetrance of the C9orf72 repeat expansion. Sci. Rep. *7*, 2116. https://doi.org/10.1038/s41598-017-02364-1.

7. Balendra, R., and Isaacs, A.M. (2018). C9orf72-mediated ALS and FTD: multiple pathways to disease. Nat. Rev. Neurol. *14*, 544–558. https://doi.org/10.1038/s41582-018-0047-2.

8. Sierksma, A., Escott-Price, V., and De Strooper, B. (2020). Translating genetic risk of Alzheimer's disease into mechanistic insight and drug targets. Science *370*, 61–66. https://doi.org/10.1126/science.abb8575.

9. Taymans, J.M., Fell, M., Greenamyre, T., Hirst, W.D., Mamais, A., Padmanabhan, S., Peter, I., Rideout, H., and Thaler, A. (2023). Perspective on the current state of the LRRK2 field. NPJ Parkinsons Dis. *9*, 104. https://doi.org/10.1038/s41531-023-00544-7.

10. Cacabelos, R. (2022). What have we learnt from past failures in Alzheimer's disease drug discovery? Expet Opin. Drug Discov. *17*, 309–323. https://doi.org/10.1080/17460441.2022.2033724.

11. Wilson, D.M., 3rd, Cookson, M.R., Van Den Bosch, L., Zetterberg, H., Holtzman, D.M., and Dewachter, I. (2023). Hallmarks of neurodegenerative diseases. Cell *186*, 693–714. https://doi.org/10.1016/j.cell.2022.12.032.

12. Mead, R.J., Shan, N., Reiser, H.J., Marshall, F., and Shaw, P.J. (2023). Amyotrophic lateral sclerosis: a neurodegenerative disorder poised for successful therapeutic translation. Nat. Rev. Drug Discov. *22*, 185–212. https://doi.org/10.1038/s41573-022-00612-2.

13. Gribkoff, V.K., and Kaczmarek, L.K. (2017). The need for new approaches in CNS drug discovery: Why drugs have failed, and what can be done to improve outcomes. Neuropharmacology *120*, 11–19. https://doi.org/10.1016/j.neuropharm.2016.03.021.

14. Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science *313*, 1929–1935. https://doi.org/10.1126/science.1132939.

15. Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607. https://doi.org/10.1038/nature11003.

16. Ochoa, D., Karim, M., Ghoussaini, M., Hulcoop, D.G., McDonagh, E.M., and Dunham, I. (2022). Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. Nat. Rev. Drug Discov. *21*, 551. https://doi.org/10.1038/d41573-022-00120-3.

17. Hingorani, A.D., Kuan, V., Finan, C., Kruger, F.A., Gaulton, A., Chopade, S., Sofat, R., MacAllister, R.J., Overington, J.P., Hemingway, H., et al. (2019). Improving the odds of drug development success through human genomics: modelling study. Sci. Rep. *9*, 18911. https://doi.org/10.1038/s41598-019-54849-w.

18. Chio, A., Borghero, G., Restagno, G., Mora, G., Drepper, C., Traynor, B.J., Sendtner, M., Brunetti, M., Ossola, I., Calvo, A., et al. (2012). Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. Brain *135*, 784–793. https://doi.org/10.1093/brain/awr366.

19. Gallagher, M.D., Suh, E., Grossman, M., Elman, L., McCluskey, L., Van Swieten, J.C., Al-Sarraj, S., Neumann, M., Gelpi, E., Ghetti, B., et al. (2014). TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. Acta Neuropathol. *127*, 407–418. https://doi.org/10.1007/s00401-013-1239-x.

20. Goodrich, J.K., Singer-Berk, M., Son, R., Sveden, A., Wood, J., England, E., Cole, J.B., Weisburd, B., Watts, N., Caulkins, L., et al. (2021). Determinants of penetrance and variable expressivity in monogenic metabolic conditions across 77,184 exomes. Nat. Commun. *12*, 3505. https://doi.org/10.1038/s41467-021-23556-4.

21. Nalls, M.A., Escott-Price, V., Williams, N.M., Lubbe, S., Keller, M.F., Morris, H.R., and Singleton, A.B.; International Parkinson's Disease Genomics Consortium IPDGC (2015). Genetic risk and age in Parkinson's disease: Continuum not stratum. Mov. Disord. *30*, 850–854. https://doi.org/10.1002/mds.26192.

22. Saez-Atienzar, S., Bandres-Ciga, S., Langston, R.G., Kim, J.J., Choi, S.W., Reynolds, R.H., International ALS Genomics Consortium, ITALSGEN; Dewan, R., Abramzon, Y., Ahmed, S., et al. (2021). Genetic analysis of amyotrophic lateral sclerosis identifies contributing pathways and cell types. Sci. Adv. *7*, eabd9036. https://doi.org/10.1126/sciadv.abd9036.

23. Bandres-Ciga, S., Saez-Atienzar, S., Kim, J.J., Makarious, M.B., Faghri, F., Diez-Fairen, M., Iwaki, H., Leonard, H., Botia, J., Ryten, M., et al. (2020). Large-scale pathway specific polygenic risk and transcriptomic community network analysis identifies novel functional pathways in Parkinson disease. Acta Neuropathol. *140*, 341–358. https://doi.org/10.1007/s00401-020-02181-3.

24. Kim, N.Y., Lee, H., Kim, S., Kim, Y.J., Lee, H., Lee, J., Kwak, S.H., and Lee, S. (2024). The clinical relevance of a polygenic risk score for type 2 diabetes mellitus in the Korean population. Sci. Rep. *14*, 5749. https://doi.org/10.1038/s41598-024-55313-0.

25. Plym, A., Penney, K.L., Kalia, S., Kraft, P., Conti, D.V., Haiman, C., Mucci, L.A., and Kibel, A.S. (2022). Evaluation of a Multiethnic Polygenic Risk Score Model for Prostate Cancer. J. Natl. Cancer Inst. *114*, 771–774. https://doi.org/10.1093/jnci/djab058.

26. Sekimitsu, S., Xiang, D., Smith, S.L., Curran, K., Elze, T., Friedman, D.S., Foster, P.J., Luo, Y., Pasquale, L.R., Peto, T., et al. (2023). Deep Ocular Phenotyping Across Primary Open-Angle Glaucoma Genetic Burden. JAMA Ophthalmol. *141*, 891–899. https://doi.org/10.1001/jamaophthalmol.2023.3645.

27. Wang, F., Ghanzouri, I., Leeper, N.J., Tsao, P.S., and Ross, E.G. (2022). Development of a polygenic risk score to improve detection of peripheral artery disease. Vasc. Med. *27*, 219–227. https://doi.org/10.1177/1358863X211067564.

28. Van Mossevelde, S., van der Zee, J., Gijselinck, I., Sleegers, K., De Bleecker, J., Sieben, A., Vandenberghe, R., Van Langenhove, T., Baets, J., Deryck, O., et al. (2017). Clinical Evidence of Disease Anticipation in Families Segregating a C9orf72 Repeat Expansion. JAMA Neurol. *74*, 445–452. https://doi.org/10.1001/jamaneurol.2016.4847.

29. Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., and Peterson, H. (2023). g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res. *51*, W207–W212. https://doi.org/10.1093/nar/gkad347.

30. Lachmann, A., Schilder, B.M., Wojciechowicz, M.L., Torre, D., Kuleshov, M.V., Keenan, A.B., and Ma'ayan, A. (2019). Geneshot: search engine for ranking genes from arbitrary text queries. Nucleic Acids Res. *47*, W571–W577. https://doi.org/10.1093/nar/gkz393.

31. Sakaue, S., and Okada, Y. (2019). GREP: genome for REPositioning drugs. Bioinformatics *35*, 3821–3823. https://doi.org/10.1093/bioinformatics/btz166.

32. Lussier, Y.A., and Chen, J.L. (2011). The emergence of genome-based drug repositioning. Sci. Transl. Med. *3*, 96ps35. https://doi.org/10.1126/scitranslmed.3001512.

33. Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell *171*, 1437–1452.e17. https://doi.org/10.1016/j.cell.2017.10.049.

34. Kropiwnicki, E., Evangelista, J.E., Stein, D.J., Clarke, D.J.B., Lachmann, A., Kuleshov, M.V., Jeon, M., Jagodnik, K.M., and Ma'ayan, A. (2021). Drugmonizome and Drugmonizome-ML: integration and abstraction of small molecule attributes for drug enrichment analysis and machine learning. Database *2021*, baab017. https://doi.org/10.1093/database/baab017.

35. Boussicault, L., Laffaire, J., Schmitt, P., Rinaudo, P., Callizot, N., Nabirotchkin, S., Hajj, R., and Cohen, D. (2020). Combination of acamprosate and baclofen (PXT864) as a potential new therapy for amyotrophic lateral sclerosis. J. Neurosci. Res. *98*, 2435–2450. https://doi.org/10.1002/jnr.24714.

36. Plosker, G.L. (2015). Acamprosate: A Review of Its Use in Alcohol Dependence. Drugs *75*, 1255–1268. https://doi.org/10.1007/s40265-015-0423-9.

37. Tariq, A., Hill, N.S., Price, L.L., and Ismail, K. (2022). Incidence and Nature of Respiratory Events in Patients Undergoing Bronchoscopy Under Conscious Sedation. J. Bronchology Interv. Pulmonol. *29*, 283–289. https://doi.org/10.1097/LBR.0000000000000837.

38. Dafinca, R., Scaber, J., Ababneh, N., Lalic, T., Weir, G., Christian, H., Vowles, J., Douglas, A.G.L., Fletcher-Jones, A., Browne, C., et al. (2016). C9orf72 Hexanucleotide Expansions Are Associated with Altered Endoplasmic Reticulum Calcium Homeostasis and Stress Granule Formation in Induced Pluripotent Stem Cell-Derived Neurons from Patients with Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. Stem Cell. *34*, 2063–2078. https://doi.org/10.1002/stem.2388.

39. Beckers, J., Tharkeshwar, A.K., Fumagalli, L., Contardo, M., Van Schoor, E., Fazal, R., Thal, D.R., Chandran, S., Mancuso, R., Van Den Bosch, L., and Van Damme, P. (2023). A toxic gain-of-function mechanism in C9orf72 ALS impairs the autophagy-lysosome pathway in neurons. Acta Neuropathol. Commun. *11*, 151. https://doi.org/10.1186/s40478-023-01648-0.

40. Zhang, Y.J., Jansen-West, K., Xu, Y.F., Gendron, T.F., Bieniek, K.F., Lin, W.L., Sasaguri, H., Caulfield, T., Hubbard, J., Daughrity, L., et al. (2014). Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. Acta Neuropathol. *128*, 505–524. https://doi.org/10.1007/s00401-014-1336-5.

41. Hart, M.P., and Gitler, A.D. (2012). ALS-associated ataxin 2 polyQ expansions enhance stress-induced caspase 3 activation and increase TDP-43 pathological modifications. J. Neurosci. *32*, 9133–9142. https://doi.org/10.1523/JNEUROSCI.0996-12.2012.

42. Lacomblez, L., Bensimon, G., Leigh, P.N., Guillet, P., and Meininger, V. (1996). Dose-ranging study of riluzole in amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis/Riluzole Study Group II. Lancet *347*, 1425–1431. https://doi.org/10.1016/s0140-6736(96)91680-3.

43. Kalk, N.J., and Lingford-Hughes, A.R. (2014). The clinical pharmacology of acamprosate. Br. J. Clin. Pharmacol. *77*, 315–323. https://doi.org/10.1111/bcp.12070.

44. Debono, M.W., Le Guern, J., Canton, T., Doble, A., and Pradier, L. (1993). Inhibition by riluzole of electrophysiological responses mediated by rat kainate and NMDA receptors expressed in Xenopus oocytes. Eur. J. Pharmacol. *235*, 283–289. https://doi.org/10.1016/0014-2999(93)90147-a.

45. Cummings, J.L., Tong, G., and Ballard, C. (2019). Treatment Combinations for Alzheimer's Disease: Current and Future Pharmacotherapy Options. J. Alzheimers Dis. *67*, 779–794. https://doi.org/10.3233/JAD-180766.

46. Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., et al. (2019). Drug

repurposing: progress, challenges and recommendations. Nat. Rev. Drug Discov. *18*, 41–58. https://doi.org/10.1038/nrd.2018.168.

47. Jarada, T.N., Rokne, J.G., and Alhajj, R. (2020). A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions. J. Cheminf. *12*, 46. https://doi.org/10.1186/s13321-020-00450-7.

48. Reay, W.R., and Cairns, M.J. (2021). Advancing the use of genome-wide association studies for drug repurposing. Nat. Rev. Genet. *22*, 658–671. https://doi.org/10.1038/s41576-021-00387-z.

49. Beck, J., Poulter, M., Hensman, D., Rohrer, J.D., Mahoney, C.J., Adamson, G., Campbell, T., Uphill, J., Borg, A., Fratta, P., et al. (2013). Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. Am. J. Hum. Genet. *92*, 345–353. https://doi.org/10.1016/j.ajhg.2013.01.011.

50. van Rheenen, W., van der Spek, R.A.A., Bakker, M.K., van Vugt, J.J.F.A., Hop, P.J., Zwamborn, R.A.J., de Klein, N., Westra, H.J., Bakker, O.B., Deelen, P., et al. (2021). Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. Nat. Genet. *53*, 1636–1648. https://doi.org/10.1038/s41588-021-00973-1.

51. Al Khleifat, A., Iacoangeli, A., van Vugt, J.J.F.A., Bowles, H., Moisse, M., Zwamborn, R.A.J., van der Spek, R.A.A., Shatunov, A., Cooper-Knock, J., Topp, S., et al. (2022). Structural variation analysis of 6,500 whole genome sequences in amyotrophic lateral sclerosis. NPJ Genom. Med. *7*, 8. https://doi.org/10.1038/s41525-021-00267-9.

52. Al-Chalabi, A., Calvo, A., Chio, A., Colville, S., Ellis, C.M., Hardiman, O., Heverin, M., Howard, R.S., Huisman, M.H.B., Keren, N., et al. (2014). Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study. Lancet Neurol. *13*, 1108–1113. https://doi.org/10.1016/S1474-4422(14)70219-4.

53. Chio, A., Mazzini, L., D'Alfonso, S., Corrado, L., Canosa, A., Moglia, C., Manera, U., Bersano, E., Brunetti, M., Barberis, M., et al. (2018). The multistep hypothesis of ALS revisited: The role of genetic mutations. Neurology *91*, e635–e642. https://doi.org/10.1212/WNL.0000000000005996.

54. Cady, J., Allred, P., Bali, T., Pestronk, A., Goate, A., Miller, T.M., Mitra, R.D., Ravits, J., Harms, M.B., and Baloh, R.H. (2015). Amyotrophic lateral sclerosis onset is influenced by the burden of rare variants in known amyotrophic lateral sclerosis genes. Ann. Neurol. *77*, 100–113. https://doi.org/10.1002/ana.24306.

55. Fumagalli, L., Young, F.L., Boeynaems, S., De Decker, M., Mehta, A.R., Swijsen, A., Fazal, R., Guo, W., Moisse, M., Beckers, J., et al. (2021). C9orf72-derived arginine-containing dipeptide repeats associate with axonal transport machinery and impede microtubule-based motility. Sci. Adv. *7*, eabg3013. https://doi.org/10.1126/sciadv.abg3013.

56. van Blitterswijk, M., Mullen, B., Wojtas, A., Heckman, M.G., Diehl, N.N., Baker, M.C., DeJesus-Hernandez, M., Brown, P.H., Murray, M.E., Hsiung, G.Y.R., et al. (2014). Genetic modifiers in carriers of repeat expansions in the C9ORF72 gene. Mol. Neurodegener. *9*, 38. https://doi.org/10.1186/1750-1326-9-38.

57. Zhang, M., Ferrari, R., Tartaglia, M.C., Keith, J., Surace, E.I., Wolf, U., Sato, C., Grinberg, M., Liang, Y., Xi, Z., et al. (2018). A C6orf10/LOC101929163 locus is associated with age of onset in C9orf72 carriers. Brain *141*, 2895–2907. https://doi.org/10.1093/brain/awy238.

58. Barbier, M., Camuzat, A., Hachimi, K.E., Guegan, J., Rinaldi, D., Lattante, S., Houot, M., Sánchez-Valle, R., Sabatelli, M., Antonell, A., et al. (2021). SLITRK2, an X-linked modifier of the age at onset in C9orf72 frontotemporal lobar degeneration. Brain *144*, 2798–2811. https://doi.org/10.1093/brain/awab171.

59. Bali, T., Self, W., Liu, J., Siddique, T., Wang, L.H., Bird, T.D., Ratti, E., Atassi, N., Boylan, K.B., Glass, J.D., et al. (2017). Defining SOD1 ALS natural history to guide therapeutic clinical trial design. J. Neurol. Neurosurg. Psychiatry *88*, 99–105. https://doi.org/10.1136/jnnp-2016-313521.

60. Liu, J., Lee, J., Salazar Hernandez, M.A., Mazitschek, R., and Ozcan, U. (2015). Treatment of obesity with celastrol. Cell *161*, 999–1011. https://doi.org/10.1016/j.cell.2015.05.011.

61. Brum, A.M., van de Peppel, J., van der Leije, C.S., Schreuders-Koedam, M., Eijken, M., van der Eerden, B.C.J., and van Leeuwen, J.P.T.M. (2015). Connectivity Map-based discovery of parbendazole reveals targetable human osteogenic pathway. Proc. Natl. Acad. Sci. USA *112*, 12711–12716. https://doi.org/10.1073/pnas.1501597112.

62. Pharnext (2023). In Pharnext refocuses its clinical trial programs on PXT3003, its most promising drug candidate, to optimize financial resources allocation, M.J. Elliott, ed.

63. Umhau, J.C., Momenan, R., Schwandt, M.L., Singley, E., Lifshitz, M., Doty, L., Adams, L.J., Vengeliene, V., Spanagel, R., Zhang, Y., et al. (2010). Effect of acamprosate on magnetic resonance spectroscopy measures of central glutamate in detoxified alcohol-dependent individuals: a randomized controlled experimental medicine study. Arch. Gen. Psychiatr. *67*, 1069–1077. https://doi.org/10.1001/archgenpsychiatry.2010.125.

64. Dewan, R., Chia, R., Ding, J., Hickman, R.A., Stein, T.D., Abramzon, Y., Ahmed, S., Sabir, M.S., Portley, M.K., Tucci, A., et al. (2021). Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis. Neuron *109*, 448–460.e4. https://doi.org/10.1016/j.neuron.2020.11.005.

65. Nicolas, A., Kenna, K.P., Renton, A.E., Ticozzi, N., Faghri, F., Chia, R., Dominov, J.A., Kenna, B.J., Nalls, M.A., Keagle, P., et al. (2018). Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. Neuron *97*, 1267–1288. https://doi.org/10.1016/j.neuron.2018.02.027.

66. Prudencio, M., Belzil, V.V., Batra, R., Ross, C.A., Gendron, T.F., Pregent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E., Desaro, P., et al. (2015). Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. Nat. Neurosci. *18*, 1175–1182. https://doi.org/10.1038/nn.4065.

67. Choi, S.W., and O'Reilly, P.F. (2019). PRSice-2: Polygenic Risk Score software for biobank-scale data. GigaScience *8*, giz082. https://doi.org/10.1093/gigascience/giz082.

68. Duan, Y., Evans, D.S., Miller, R.A., Schork, N.J., Cummings, S.R., and Girke, T. (2020). signatureSearch: environment for gene expression signature searching and functional interpretation. Nucleic Acids Res. *48*, e124. https://doi.org/10.1093/nar/gkaa878.

69. Brooks, B.R. (1994). El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. J. Neurol. Sci. *124*, 96–107. https://doi.org/10.1016/0022-510x(94)90191-0.

70. Neary, D., Snowden, J.S., Gustafson, L., Passant, U., Stuss, D., Black, S., Freedman, M., Kertesz, A., Robert, P.H., Albert, M., et al. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. Neurology *51*, 1546–1554. https://doi.org/10.1212/wnl.51.6.1546.

71. Salvi, E., Kutalik, Z., Glorioso, N., Benaglio, P., Frau, F., Kuznetsova, T., Arima, H., Hoggart, C., Tichet, J., Nikitin, Y.P., et al. (2012). Genomewide association study using a high-density single nucleotide polymorphism array and case-control design identifies a novel essential hypertension susceptibility locus in the promoter region of endothelial NO synthase. Hypertension *59*, 248–255. https://doi.org/10.1161/HYPERTENSIONAHA.111.181990.

72. van Rheenen, W., Shatunov, A., Dekker, A.M., McLaughlin, R.L., Diekstra, F.P., Pulit, S.L., van der Spek, R.A.A., Võsa, U., de Jong, S., Robinson, M.R., et al. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat. Genet. *48*, 1043–1048. https://doi.org/10.1038/ng.3622.

73. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation

genotype imputation service and methods. Nat. Genet. *48*, 1284–1287. https://doi.org/10.1038/ng.3656.

74. Chia, R., Sabir, M.S., Bandres-Ciga, S., Saez-Atienzar, S., Reynolds, R.H., Gustavsson, E., Walton, R.L., Ahmed, S., Viollet, C., Ding, J., et al. (2021). Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. Nat. Genet. *53*, 294–303. https://doi.org/10.1038/s41588-021-00785-3.
75. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.
76. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575. https://doi.org/10.1086/519795.
77. Euesden, J., Lewis, C.M., and O'Reilly, P.F. (2015). PRSice: Polygenic Risk Score software. Bioinformatics *31*, 1466–1468. https://doi.org/10.1093/bioinformatics/btu848.
78. Mehta, P., Kaye, W., Raymond, J., Punjani, R., Larson, T., Cohen, J., Muravov, O., and Horton, K. (2018). Prevalence of Amyotrophic Lateral Sclerosis - United States, 2015. MMWR Morb. Mortal. Wkly. Rep. *67*, 1285–1289. https://doi.org/10.15585/mmwr.mm6746a1.
79. Leonenko, G., Baker, E., Stevenson-Hoare, J., Sierksma, A., Fiers, M., Williams, J., de Strooper, B., and Escott-Price, V. (2021). Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores. Nat. Commun. *12*, 4506. https://doi.org/10.1038/s41467-021-24082-z.
80. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47*, W191–W198. https://doi.org/10.1093/nar/gkz369.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-goat (donkey)	Thermo Fisher	Catalog # A21432, RRID: AB_2535853
Anti-guinea pig (goat)	Thermo Fisher	Catalog # A21450, RRID: AB_2535867
Anti-mouse (donkey)	Thermo Fisher	Catalog # A10037, RRID: AB_2534013
Anti-mouse (donkey)	Thermo Fisher	Catalog # A21202, RRID: AB_141607
Anti-rabbit (donkey)	Thermo Fisher	Catalog # A21206, RRID: AB_2535792
Anti-rabbit (donkey)	Thermo Fisher	Catalog # A10042, RRID: AB_2534017
Beta III tubulin (mouse)	Biolegend	Catalog # 801201, RRID: AB_2728521
Caspase-3 (rabbit)	Merck Millipore	Catalog # AB3623, RRID: AB_91556
ChAT (goat)	Merck Millipore	Catalog # AB144P, RRID: AB_2079751
Islet ½ (rabbit)	Abcam	Catalog # ab109517, RRID: AB_10866454
MAP-2 (guinea pig)	Synaptic systems	Catalog # 188004, RRID: AB_2138181
NeuN (mouse)	Merck Millipore	Catalog # MAB377, RRID: AB_2298772
Chemicals, peptides, and recombinant proteins		
Acamprosate calcium	Sigma-Aldrich	Catalog # A6981
Accutase	StemCell Technologies	Catalog # 07922
All-*trans* retinoic acid	Merck Millipore	Catalog # 554720
B27 supplement	Thermo Fisher	Catalog # 17504044
Brain Derived Neurotrophic Factor (BDNF)	Peprotech	Catalog # 450-02
Camptothecin (CPT)	Cell Signaling Technology	Catalog # 13637
CHIR 99021	Tocris Biosciences	Catalog # 4423
Ciliary Neurotrophic Factor (CNTF)	Peprotech	Catalog # 450-13
Compound E	Merck Millipore	Catalog # 530509
D(−)-2-Amino-5-phosphonopentanoic acid (D-AP5)	Sigma-Aldrich	Catalog # A8054
4,6-diamidino-2-phenylindole (DAPI)	Merck Millipore	Catalog # 508741
Dimethyl Sulfoxide (DMSO)	Merck Millipore	Catalog # 317275
Distilled H_2O	Merck Millipore	Catalog # EM3234
Dorsomorphin homolog 1 (DMH-1)	Tocris Biosciences	Catalog # 4126
GlutaMAX	Thermo Fisher	Catalog # 35050061
Hank's balanced salt solution (HBSS)	Thermo Fisher	Catalog # 14175
Insulin-like Growth Factor-I (IGF-1)	Peprotech	Catalog # 100-11
KnockOut DMEM/F-12	Thermo Fisher	Catalog # A1370801
mTeSR Plus Basal Medium	StemCell Technologies	Catalog # 100-0276
N2 supplement	Thermo Fisher	Catalog # 17502048

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Paraformaldehyde	Thermo Fisher	Catalog # 047392.9M
Penicillin/streptomycin	Merck Millipore	Catalog # 516106
Phosphate buffered saline (PBS)	Merck Millipore	Catalog # 6504
Purmorphamine (PMN)	Merck Millipore	Catalog # 540220
ReLeSR Passaging Reagent	StemCell Technologies	Catalog # 100-0484
Riluzole	Merck Millipore	Catalog # 557324
SB431542	Tocris Biosciences	Catalog # 1614
Triton X-100	Merck Millipore	Catalog # 648464
Vitronectin XF	StemCell Technologies	Catalog # 07180
Y27632 dihydrochloride	Tocris Biosciences	Catalog # 1254
Critical commercial assays		
MTT assay	Thermo Fisher	Catalog #V13154
SNP beadchip genotyping array	Illumina	Catalog #: InfiniumOmni2-5-8v1-4_A1
Deposited data		
ALS GWAS summary statistics	Van Rheenen et al.[50]	https://www.projectmine.com/research/download-data/
The individual-level data for the training and test datasets	Dewan et al.[64]	dbGaP (accession # phs001963.v1.p1), RRID: SCR_002709
The individual-level data for the training and test datasets	Nicolas et al.[65]	dbGaP (accession # phs000101.v5.p1), RRID: SCR_002709
The individual-level control data for the training and test datasets	https://dbgap.ncbi.nlm.nih.gov	dbGaP (accession # phs000001, phs000007, phs000187, phs000196, phs000292, phs000304, phs000315, phs000368, phs000372, phs000394, phs000397, phs000404, phs000421, phs000428, phs000615, phs000675, phs000801, and phs000869), RRID:SCR_002709
The individual-level data used as the replication dataset in this study.	King's College London, unpublished study	Available from study principal investigator (Christopher Shaw, King's College London) upon reasonable request.
The individual-level polygenic risk scores generated for the *C9orf72* patients.	This study	https://zenodo.org/uploads/13769448
LINCS dataset	Subramanian et al.[33]	https://lincsproject.org, RRID: SCR_006454
The LINCS connectivity scores for the drugs evaluated as treatments for *C9orf72*-related disease.	This study	https://zenodo.org/records/13769483
RNA-sequencing data from the motor cortex	New York Genome Center	https://www.nygenome.org
RNA-sequencing data from the cerebellum	Prudencio et al.[66]	GEO (accession # GSE67196), RRID: SCR_005012
Experimental models: Cell lines		
Human: iPS cell line	Cedars-Sinai	CS14iCTR-nxx
Human: iPS cell line	Coriell Biorepository	GM23338, RRID: CVCL_F182
Human: iPS cell line	University of Sheffield	MIFF1, RRID: CVCL_1E69
Human: iPS cell line	Cedars-Sinai	CS02iCTR-NTn1
Human: iPS cell line	University of Sheffield	ALS-183-C9
Human: iPS cell line	University of Sheffield	ALS-78
Human: iPS cell line	Cedars-Sinai	CS28iALS-C9nxx
Human: iPS cell line	Cedars-Sinai	CS29iALS-C9nxx
Human: iPS cell line	Cedars-Sinai	CS52iALS-C9nxx

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human: iPS cell line	Cedars-Sinai	CS29iALS-C9n1.ISOxx
Human: iPS cell line	Cedars-Sinai	CS52iALS-C9n6.ISOxx
Software and algorithms		
Codes and scripts	This paper	https://github.com/sarasaezALS/C9orf72_AAO; https://zenodo.org/records/13259646
Drugmonizome	Kropiwnicki et al.[34]	https://maayanlab.cloud/drugmonizome/#/
ExperimentHub	https://www.bioconductor.org/packages/release/bioc/html/ExperimentHub.html	https://mrcieu.github.io/TwoSampleMR
Geneshot	Lachmann et al.[30]	https://maayanlab.cloud/genesgen, RRID: SCR_017582
g:Profiler2	Kolberg et al.[29]	https://biit.cs.ut.ee/gprofiler/ggos, RRID: SCR_018190
Harmony software	Perkin Elmer	Catalog # hh17000010, RRID: SCR_023543
PLINK (version 1.9)	PLINK Working Group	https://www.cog-genomics.org/plink/1.9/, RRID: SCR_001757
Prism (version 10)	GraphPad Software	https://www.graphpad.com, RRID: SCR_002798
PRSice2	Choi et al.[67]	https://choishingwan.github.ii/PRSice/, RRID: SCR_017057
Python	Python Team	https://www.python.org, RRID: SCR_008394
R (version 4.0.5)	R Core Team	https://www.r-project.org, RRID: SCR_001905
SignatureSearch (version 1.11.0)	Duan et al.[68]	https://bioconductor.org/packpack/release/bioc/vignettes/sigsignatureSea/inst/doc/signatursignatu.html, RRID: SCR_016177

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human participants

Tables S8 and S9 list the source and clinical features of the cohorts used in this study. The ALS patients were diagnosed according to the El Escorial criteria[69] for the training and test datasets, and the FTD patients were analyzed according to the Neary criteria.[70] The *C9orf72* repeat expansions were detected using a repeat-primed polymerase chain reaction (PCR) assay according to an established protocol.[3] Written consent was obtained from all individuals enrolled in this study, and the institutional review board approved the study of the National Institute on Aging (protocol number 03-AG-N329). The *C9orf72* patients of the replication cohort were from a different study conducted at the King's College London and recruited by the SLAGEN Consortium, the University of Edinburgh, the Boğaziçi University, and Project MinE. Written consent was obtained from all individuals at their respective centers (see the *Replication cohort* section for details).

Datasets

Four independent datasets were used in the analysis, as is standard in genetic risk score analysis (see Figure 1 for the analysis workflow). The *reference dataset* consisted of summary statistics from a published GWAS based on 12,577 ALS cases and 23,475 control individuals (Table S9).[65] The allele weights obtained from the *reference dataset* were used to construct the ALS genetic risk score model in the training dataset.

The *training dataset* was composed of 7,037 ALS individuals known not to carry *C9orf72* repeat expansions and 34,235 controls (Table S9). The case samples were previously genotyped in the Laboratory of Neurogenetics, National Institutes of Health, using HumanOmniExpress SNP arrays (version 1.0, Illumina Inc., San Diego, CA).[22,65] The US control samples had been previously genotyped on Illumina SNP arrays as part of other GWAS efforts. These data were downloaded from the dbGaP repository (accession numbers phs000001, phs000007, phs000187, phs000196, phs000292, phs000304, phs000315, phs000368, phs000372, phs000394, phs000397, phs000404, phs000421, phs000428, phs000615, phs000675, phs000801, and phs000869). Additional

SNP array data from the *HYPERGENES* project and the Wellcome Trust Case Control Consortium were included as Italian and UK control subjects.[71] All study participants were of European ancestry, and familial cases were included in the analysis. The individual-level data for the *training dataset* was used as input for the *PRSice2* algorithm.

To assess the effect of the ALS genetic risk score on age at onset in *C9orf72* carriers, we calculate the genetic risk score in the *C9orf72* cohort. This *test dataset* consisted of 817 ALS/FTD cases that carry the *C9orf72* gene (see Table S9 for detailed information). Of these, 666 (81.5% of the cohort) were genotyped on SNP arrays in the Laboratory of Neurogenetics,[65] and 151 (18.4%) underwent whole-genome sequencing.[64] Genotype data for the 161 SNPs making up the model were extracted from the SNP array and whole-genome sequence data.

We replicated our findings in an independent *C9orf72* cohort obtained from a different study conducted at King's College London. This *replication dataset* consisted of 699 ALS/FTD cases known to carry the *C9orf72* repeat expansion and genotyped on SNP arrays using Illumina InfiniumOmni2-5-v1.[72]

METHOD DETAILS

SNP array-based genotyping data quality control procedures and imputation

Standard quality-control procedures were applied to the genotype data of the training dataset (n = 7,037 cases and 34,235 controls) and the test dataset (n = 666 cases) before input into the *PRSice2* algorithm. Briefly, individuals with low call rates (<95%), heterozygosity outliers (F-statistic > −0.15 or < 0.15), and ancestry outliers (+/− 6 standard deviations from means of principal components 1 and 2 of the 1000 Genomes phase 3 Caucasian with European ancestry from Utah (CEU) and Toscani in Siena, Italy (TSI) populations) were excluded. Variants with a missingness rate of >5%, exhibiting deviation from Hardy–Weinberg Equilibrium in controls (p-value < 10^{-6}), and palindromic SNPs were excluded. Cryptically related samples (defined as Pi_hat >0.125) were removed. The remaining sample genotypes were imputed using the Michigan Imputation Server pipeline using Minimac4[73] under default settings with Eagle (version 2.4) phasing based on *Haplotype Reference Consortium* (release 1.1 2016). Samples from the United States, Italy, the United Kingdom, Belgium, and France were imputed as a single group. Those variants were additionally filtered post-imputation to exclude variants with minor allele frequency <0.01, missing call rates >15%, and imputation quality R^2 < 0.3.

The SNP array-based genotyping in the *replication dataset* section describes the quality control and imputation methodology used in the replication dataset. Samples in common between the test and replication datasets were identified using the checksum program *id_geno_checksum.v2* and removed.

Whole-genome sequencing of the test dataset

In the test dataset, one hundred and fifty-one *C9orf72* carriers had previously undergone 150 base pair, paired-end whole-genome sequencing on a HiSeq X Ten sequencer.[64,74] Genotype data for the 161 SNPs making up the model were extracted from this whole-genome sequence data. They were merged with the SNP array genotype data for the remaining 666 *C9orf72* carriers using PLINK (version 1.9).[75,76]

Genetic risk score generation and computation

The genome-wide genetic risk score was calculated using the training dataset based on the weighted allele dosages obtained from the reference dataset as implemented in *PRSice2*.[67,77] This approach allows variants below the typical GWAS significance threshold of 5.0 x 10^{-8} to be included in the analysis, and the model selected a p-value threshold $\leq$ 0.0001 for SNP selection. For the training dataset, 1,000 permutations were used to generate empirical p-value estimates for each GWAS-derived p-value. Each permutation test in the training dataset provided a Nagelkerke's pseudo-R^2 value after adjusting for an estimated ALS prevalence of 5 per 100,000 of the population.[78] To avoid accounting for the *C9orf72* effect, 150 kb upstream and downstream of the *C9orf72* top GWAS variant, rs3849943 (chr9: 9:27,543,382; GRCh37), were removed from the analysis. Sex, age at onset, and principal components one to twenty were included as covariates in the model.

The *–score* command implemented in PLINK (version 1.9)[75] was used to test the general ALS genetic risk score's contribution to the age of symptom onset among the *C9orf72* expansion carriers. Risk allele dosages were counted, giving a dose of two if homozygous for the risk allele, one if heterozygous, and zero if homozygous for the reference allele. Linear regression was used to evaluate the association between the genetic risk scores and the age at onset, as implemented in *R* (version 4.0.3). Sex, disease diagnosis (ALS or FTD), and principal components one to twenty were included as covariates in the model. For replication, 147 out of 161 SNPs were used to build genetic risk scores. Sex and principal component 1 were used as covariates in the model. Genetic risk scores were transformed to *Z* scores based on cases (a *Z* score of one is equivalent to a single standard deviation of increase from the case population mean of the genetic risk score). There was a four-standard deviation difference from the genetic risk score mean between *C9orf72* carriers at the top and the bottom of the genetic risk score distribution. We calculated the maximum age at onset that the genetic risk score can explain by multiplying the slope of the regression (beta) by this standard deviation difference. This approach determined the extent to which the genetic risk score can account for variations in the age at onset. Individuals within the extremes were defined as those within the 3% tails of the genetic risk score distribution. On average, the score at the extremes exceeded ±2 standard deviations from the data mean (Table S2).[79]

Leave-one-out analysis and decile generation

Leave-one-out analyses were performed by iteratively excluding one variant from the ALS genetic risk score (based on 161 predictors) and re-estimating the causal effect on age at onset (Table S3). The 161 variants were then ordered based on the regression coefficient (beta) from the leave-one-out analysis and regrouped in ten deciles. Thus, we generated ten ranked deciles composed of 16 variants each. Decile ten contains the 16 variants with a more significant contribution to early age at onset. Regression analyses to evaluate the contribution of each decile (16 variants) to age at onset were performed using PLINK (version 1.9) as described previously.[75]

Pathway analysis

Functional enrichment analysis was performed using the g:GOSt function of g:Profiler2.[29] Briefly, SNPs from decile one and the *C9orf72* gene name were input for g:Profiler2.[29,80] Enrichment was performed against the biological process and molecular function from the Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. Only gene lists that contain between 5 and 500 genes were selected for the analysis. The significant threshold was an FDR-corrected *p*-value less than 0.05. Pathways containing a single gene were removed from the study.

Drug prioritization via gene-drug pattern matching

The SNPs that compose decile ten (Table S3) were mapped to gene names using the *g:SNPense* function of g:Profiler2.[29] These genes were used as search terms (defined as seed genes in Table S5) to identify functionally related genes based on previous knowledge and gene-gene co-expression data in the *Geneshot* webserver (accessed October 2021).[30] This application identifies genes associated with our search term based on their co-occurrence in publications and gene-gene similarity from human RNA-seq data (ARCHS4) to predict associations between genes and search terms.[30] The resulting list of genes plus the seed genes (Table S5) was then used as input for the GREP analysis (version 1.0.0), a pipeline identifying drugs that can be repurposed to target the gene set based on their enrichment in clinical indication categories.[31]

Finally, *drugmonizome* was used to identify significant biomedical terms within the specified drugs based on their indications (from the SIDER (Side Effect Resource) database, the mechanism of action (MOA, from the DrugRepurposingHub database), and the gene targets (from the Drugbank database) (data accessed 10/12/2022).[34]

Drug prioritization via gene expression-pattern matching

The nominated drugs were validated using the drug-repurposing algorithm called Connectivity map (*CMap*) through the *SignatureSearch* package using the LINCS search method (version 1.8.2).[68] This algorithm uses two inputs: (i) a disease gene expression signature based on a list of the up- and down-regulated genes, and (ii) a drug perturbation dataset composed of differential expression profiles of each gene after drug treatment. A bi-directional weighted Kolmogorov–Smirnov enrichment statistic test of gene expression ranks in the disease and drug signatures was used to assign a weighted normalized connectivity score (referred to as a *CMap* score) to each drug, reflecting the degree to which the drug 'flips' the signature of the disease. The *CMap* scores reflect the similarity between the query drug's signature and those in the database, suggesting that the drug has similar or dissimilar biological effects.

To build the *C9orf72* disease signature in the motor cortex, differentially expressed genes in *C9orf72* cases ($n = 36$) versus controls ($n = 58$) were selected at a Bonferroni-corrected *p*-value <0.05 (see supplemental information for details). This dataset was obtained from the New York Genome Center. To build the *C9orf72* disease signature in the cerebellum, differentially expressed genes in *C9orf72* cases ($n = 8$) versus controls ($n = 8$) were selected at FDR-corrected *p*-value <0.05. This dataset was downloaded as a raw count matrix from GEO (accession number GSE67196).[66] Drug signatures from the *LINCS* database were accessed through *ExperimentHub* (version 2.2.0) in the form of moderated z-scores from differential expression (DE) analysis of 12,328 genes from 8,140 compound treatments of 30 cell lines corresponding to a total of 45,956 signatures. Drug signatures corresponding to the nominated drugs were extracted from the analyzed results and were available for the following cell lines: A375 (LINCS ID = LCL-1235, a human cell line exhibiting epithelial morphology isolated from the skin of a patient with malignant melanoma, provided by the American Type Culture Collection (ATCC)), ASC (LCL-2104, human adipose stem cells, Sciencell Research Laboratories), FIBRNPC (LSC-1021, iPSC), HCC515 (LCL-2084, human cell line isolated from lung adenocarcinoma, Broad Institute), HT29 (LCL-1180, a human cell line with epithelial morphology isolated from a patient with colorectal adenocarcinoma, ATCC), NEU (LDC-1033, differentiated cell), A549 (LCL-1601, a human cell line isolated from the lung of a patient with non-small cell lung carcinoma, ATCC), NPC (LDC-1021, normal stem fibroblast-derived iPSCs), HA1E (LCL-2090, a human cell line isolated from kidney, Broad Institute), PC3 (LCL-1299, a human cell line isolated from a patient with grade IV prostate adenocarcinoma, ATCC), MCF7 (LCL-2138, a human cell line isolated from a patient with breast adenocarcinoma, ATCC), PHH (primary human hepatocyte), SKBR3 (LCL-1475, human cell line isolated from a patient with breast adenocarcinoma, ATCC), and VCAP (LCL-1147, a human cell line that was isolated from a patient with prostate carcinoma, ATCC). Additional information about the cell lines is available at lincsportal.ccs.miami.edu.

Drug *in vitro* validation

Human induced pluripotent stem cells (iPSCs) were maintained in 6-well plates coated with vitronectin XF (10 μg/mL) in the complete mTeSR-Plus medium. Media was replaced every 48 h, and cells were passaged as clumps every four to six days using ReLeSR,

according to the manufacturer's instructions. IPSCs were used between passages 20 and 35, and all iPSCs were cultured in 5% O_2 and 5% CO_2 at 37°C. In this study, we used iPS cells derived from four unaffected controls (CS14iCTR-21nxx, MIFF1, CS02iCTR-NTn1, GM23338), five iPS lines derived from ALS patients harboring *C9orf72* repeat expansions (ALS-183-C9, CS52iALS-C9nxx, ALS-78, CS28iALS-C9nxx, CS29iALS-C9nxx) and two isogenic control lines of CS52iALS-nxx and CS29iALS-C9nxx, respectively (CS52iALS-C9n6.ISOxx, CS29iALS-C9n1.ISOxx (*C9orf72* HRE Corrected)) (Table S10). The cells were fed on alternate days with the neuronal medium until day 40. Cells were previously characterized at day 40 of differentiation and were found to express classical mature motor neuron markers (ChAT, SMI32, Islet 1/2, MAP2, NeuN) (Figure S6 and Table S11).

Drug treatments were applied for 72 h at day 40 of differentiation, and cells were assayed for viability (MTT) or fixed for subsequent immunocytochemistry assays. All the imaging was performed using the Opera Phenix High Content Screening System (PerkinElmer) at ×40 magnification. The supplementary material provides a detailed description of the methodology.

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistics and graphs for the cell line experiments were generated using GraphPad Prism 10 (GraphPad Software, La Jolla, California USA, www.graphpad.com). Comparisons were performed using one-way ANOVA or two-way ANOVA with Tukey's post hoc multiple comparisons test. Three technical replicates per treatment were averaged before plotting, and statistical analysis was performed using the percentage of cleaved caspase-3 mean across cell lines. Plots represent mean ± SD. *p*-values smaller than 0.05 are annotated unless otherwise specified in the figure legend.