Telomere length analysis in amyotrophic lateral sclerosis using large-scale whole genome sequence data

Ahmad Al Khleifat[1*], Alfredo Iacoangeli[1], Ashley R. Jones[1], Joke J. Van Vugt[2], Matthieu Moisse[3], Aleksey Shatunov[4], Ramona A. Zwamborn[2], Rick A. Van Der Spek[2], Johnathan Cooper-Knock[5], Simon D. Topp[1], Wouter Van Rheenen[2], Brendan Kenna[2], Kristel R. Van Eijk[2], Kevin Kenna[2], Ross Byrne[6], Victoria López Alonso[7], Sarah Opie-Martin[1], Atay Vural[8], Yolanda González[7], Markus Weber[9], Bradley N. Smith[1], Isabella Fogh[1], Vincenzo Silani[10], Karen E. Morrison[11], Richard J. Dobson[1], Michael A. Van Es[2], Russell L. Mclaughlin[6], Adriano Chio[12], Philippe Corcia[13], Mamede De Carvalho[14], Marc Gotkine[15], Monica P. Panades[16], Jesus S. Mora[17], Pamela J. Shaw[5], John E. Landers[18], Jonathan D. Glass[19], Christopher E. Shaw[1], Nazli A. Basak[8], Orla Hardiman[6], Wim Robberecht[3], Philip Van Damme[3], Leonard Van Den Berg[2], Jan Veldink[2], Ammar Al-Chalabi[1*]

[1]Institute of Psychiatry, Psychology & Neuroscience, King's College London, United Kingdom, [2]University Medical Center Utrecht, Netherlands, [3]Leuven Brain Institute, KU Leuven, Belgium, [4]University of Liverpool, United Kingdom, [5]Sheffield Institute for Translational Neuroscience, The University of Sheffield, United Kingdom, [6]Trinity College Dublin, Ireland, [7]Carlos III Health Institute (ISCIII), Spain, [8]Koc University, School of Medicine, Translational Medicine Research Center- NDAL. Istanbul, Turkey, Turkey, [9]Neuromuscular Diseases Unit, Kantonsspital St. Gallen, 9007, St. Gallen, Switzerland, Switzerland, [10]Dino Ferrari Center, University of Milan, Italy, [11]Faculty of Medicine, Health and Life Sciences, Queen's University Belfast, United Kingdom, [12]Department of Neuroscience, School of Medicine, University of Turin, Italy, [13]Faculté de Médecine, Université de Tours, France, [14]NOVA Medical School, Faculty of Medical Sciences of Lisbon, New University of Lisbon, Portugal, [15]Hadassah School of Dental Medicine, Hadassah Medical Center, Israel, [16]Neurology Department, Hospital Universitari de Bellvitge, Barcelona, Spain, Spain, [17]Hospital San Rafael, Madrid, Spain, Spain, [18]Department of Neurology, University of Massachusetts Medical School, United States, [19]Department of Neurology, Center for Neurodegenerative Diseases, Emory University, Atlanta, Georgia USA, United States

Submitted to Journal:
Frontiers in Cellular Neuroscience

Specialty Section:
Cellular Neuropathology

Article type:
Original Research Article

Manuscript ID:
1050596

Received on:
21 Sep 2022

Revised on:
14 Nov 2022

Journal website link:
www.frontiersin.org

Conflict of interest statement

The authors declare a potential conflict of interest and state it below

Competing Interests
AAC is a consultant for Mitsubishi-Tanabe Pharma, GSK, and Chronos Therapeutics, and chief investigator for clinical trials for Cytokinetics and OrionPharma. JHV reports to have sponsored research agreements with Biogen. VS is a consultant for Novartis and Biogen. LHVDB reports grants from Netherlands ALS Foundation, grants from The Netherlands Organization for Health Research and Development (Vici scheme), grants from The European Community's Health Seventh Framework Programme (grant agreement n° 259867 (EuroMOTOR)), grants from The Netherlands Organization for Health Research and Development)the STRENGTH project, funded through the EU Joint Programme – Neurodegenerative Disease Research, JPND), during the conduct of the study; personal fees from Calico, personal fees from Cytokinetics, grants and personal fees from Takeda, non-financial support from Orion, non-financial support from Orphazyme, outside the submitted work; . AC serves on scientific advisory boards for Mitsubishi Tanabe, Roche, Denali Pharma, Cytokinetics, Lilly, and Amylyx and has received a reeseach grant from Biogen. CES reports grants from Avexis , grants from Eli Lilly, grants from Chronos Therapeutics, grants from Vertex Pharmaceuticals, during the conduct of the study; grants from QurAlis, grants from Chronos Therapeutics, grants from Biogen, outside the submitted work. JEL is a member of the scientific advisory board for Cerevel Therapeutics, a consultant for ACI Clinical LLC sponsored by Biogen, Inc. or Ionis Pharmaceuticals, Inc. JEL is also a consultant for Perkins Coie LLP and may provide expert testimony. JEL was supported by funding from NIH/NINDS (R01NS073873 and R56NS073873). AAK, OH, MPP, JSM, PJS, JEL, CES, NB, OH, WR, PVD, NB, KK, BK, HB and LHVB declare no competing interests.

Author contribution statement

Author Contributions
AAC and AAK conceived and planned the study. AAK and AI created the bioinformatics pipeline for analysis. JJFAV ran ExpansionHunter on Project MinE data. MM and RAJZ prepared phenotypic data. AAK and AAC did the statistical analysis and prepared the figures and tables. AAC, JHV, OH, MPP, JSM, PJS, JEL, CES, NB, OH, WR, PVD and LHVB helped in sample collection and provided whole genome sequence data, analysis and intellectual input for data interpretation on behalf of the Project MinE Consortium. JHV, OH, and MM provided intellectual input for data interpretation. AAC and AAK wrote the first draft of the manuscript. All authors reviewed and approved the final manuscript.

Keywords

Amyotrophic Lateral Sclerosis (ALS), Telomere - genetics, Whole genome sequence (WGS), Genomics, Bigdata, MND - motor neuron disorders

Abstract

Word count: 131

Amyotrophic lateral sclerosis (ALS), a neurodegenerative disease of motor neurons, is a complex genetic disease with heritability of 60%. Only about 14% of apparently sporadic ALS is explained by known genetic variation, suggesting that other forms of genetic variation are important, in addition to the known risk factors, male sex and increasing age. Telomeres maintain DNA integrity during cellular replication, differ between sexes, and shorten naturally with age. We find that longer telomeres are associated with ALS and specific phenotypic patterns of disease expression. We see the same pattern of telomere elongation in ALS in brain tissue. The association of longer telomeres with apparently sporadic ALS was also seen in people with familial ALS, supporting the notion that sporadic and familial ALS are not mutually exclusive categories but rather a spectrum.

Contribution to the field

Using a large disease-specific whole genome sequencing dataset, we have shown that longer telomeres are associated with ALS. In keeping with expectations, we also found that mean telomere length was on average longer in females, and in all samples, shortened with increasing age. The association of longer telomeres with apparently sporadic ALS was also seen in FALS, supporting the notion that familial and sporadic ALS are not mutually exclusive categories but rather a spectrum. Furthermore, telomere length was inversely correlated with C9orf72 repeat expansion size. Our findings have the advantage of a large sample size of more than 4,315 cases, compared with previous reports in 50 or fewer, and advance knowledge of telomere length in disease and risk factors for ALS.

Funding statement

AAK is funded by ALS Association Milton Safenowitz Research Fellowship (grant number22-PDF-609.DOI :10.52546/pc.gr.150909.), The Motor Neurone Disease Association (MNDA) Fellowship (Al Khleifat/Oct21/975-799), The Darby Rimmer Foundation, and The NIHR Maudsley Biomedical Research Centre. This project was funded by the MND Association and the Wellcome Trust. This is an EU Joint Programme-Neurodegenerative Disease Research (JPND) project. The project is supported through the following funding organisations under the aegis of JPND - www.jpnd.eu (United Kingdom, Medical Research Council (MR/L501529/1 and MR/R024804/1) and Economic and Social Research Council (ES/L008238/1)). AAC is an NIHR Senior Investigator. C.E.S. and A.A.C. receive salary support from the National Institute for Health Research (NIHR) Dementia Biomedical Research Unit at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The work leading up to this publication was funded by the European Community's Health Seventh Framework Program (FP7/2007–2013; grant agreement number 259867) and Horizon 2020 Program (H2020-PHC-2014-two-stage; grant agreement number 633413). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 772376 - EScORIAL.
The collaboration project is co-funded by the PPP Allowance made available by Health~Holland, Top Sector Life Sciences & Health, to stimulate public-private partnerships.
Project MinE Belgium was supported by a grant from IWT, the Belgian ALS Liga and a grant from Opening the Future Fund (KU Leuven). PVD holds a senior clinical investigatorship of FWO-Vlaanderen and is supported by E. von Behring Chair for Neuromuscular and Neurodegenerative Disorders, the ALS Liga België and the KU Leuven funds "Een Hart voor ALS", "Laeversfonds voor ALS Onderzoek" and the "Valéry Perrier Race against ALS Fund". R.L.McL. is supported by Science Foundation Ireland (17/CDA /4737).
MinE USA is funded by the US ALS Association. We thank the patients and families from the Emory ALS Clinic participating in this research.
Funding was provided by US National Institutes of Health (NIH)/National Institute of Neurological Disorders and Stroke (NINDS) (R01NS073873, J.E.L.) and the American ALS Association (J.E.L.).

Ethics statements

Studies involving animal subjects

Generated Statement: No animal studies are presented in this manuscript.

Studies involving human subjects

Generated Statement: The studies involving human participants were reviewed and approved by Informed consent was obtained from all participants in this project under reference Trent Research Ethics Committee 08/H0405/60.. The patients/participants provided their written informed consent to participate in this study.

Inclusion of identifiable human data

Generated Statement: No potentially identifiable human images or data is presented in this study.

Data availability statement

Generated Statement: The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Telomere length analysis in amyotrophic lateral sclerosis using large-scale whole genome sequence data

Ahmad Al Khleifat[1+], Alfredo Iacoangeli[1,2+], Ashley R Jones[1], Joke J.F.A. van Vugt[3], Matthieu Moisse[4], Aleksey Shatunov[5,6], Ramona A. J. Zwamborn[3], Rick A.A. van der Spek[3], Johnathan Cooper-Knock[7], Simon Topp[1], Wouter van Rheenen[3], Brendan Kenna[3], Kristel R. Van Eijk[3], Kevin Kenna [3], Ross Byrne[8], Victoria López[9], Sarah Opie-Martin[1,] Atay Vural[10], Yolanda Campos González [9], Markus Weber[10], Bradley Smith[1], Isabella Fogh[1], Vincenzo Silani[12] , Karen E Morrison[13], Richard Dobson[2,14] ,Michael A. van Es[3], Russell L. McLaughlin[8], Patrick Vourc'h[15] , Adriano Chio[16,17], Philippe Corcia[15,18], Mamede de Carvalho[19], Marc Gotkine[20], Monica Povedano Panades[21], Jesus S. Mora[22], Pamela J. Shaw[7], John E. Landers[23], Jonathan D. Glass[24], Christopher E Shaw[1,25], Nazli Basak[10], Orla Hardiman[26, 27], Wim Robberecht[4, 28], Philip Van Damme[4,28], Leonard H. van den Berg[3], Jan H. Veldink[3], Ammar Al-Chalabi[1,25] on behalf of the Project MinE Consortium.

1. King's College London, Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, De Crespigny Park, London, SE5 9RX, UK

2. Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, U.K.

3. Department of Neurology, UMC Utrecht Brain Center, Utrecht University, The Netherlands

4. KU Leuven – University of Leuven, Department of Neurosciences, Experimental Neurology; VIB Center for Brain & Disease Research, Laboratory of Neurobiology, Leuven, Belgium

5. North-Eastern Federal University, Institute of Medicine, 58 Belinsky str. Yakutsk, 677000, Russia.

6. University of Liverpool, Dept Molecular and Clinical Pharmacology, Blue Block 1.09, Sherrington Building, Crown St, Liverpool, L693BX, UK.

7.Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK

8. Complex Trait Genomics Laboratory, Smurfit Institute of Genetics, Trinity College Dublin, Ireland

9. Computational Biology Unit. Instituto de Salud Carlos III, Madrid, Spain.

10. Koc University, School of Medicine, Translational Medicine Research Center- NDAL. Istanbul, Turkey

11.Neuromuscular Diseases Unit/ALS Clinic, Kantonsspital St. Gallen, 9007, St. Gallen, Switzerland.

12. Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano and Department of Pathophysiology and Transplantation, "Dino Ferrari" Center, Università degli Studi di Milano, Milano, Italy

13. Faculty of Medicine, Health and Life Sciences, Queen's University Belfast, Belfast, Northern Ireland, UK

14. Institute of Health Informatics, University College London, London, U.K.

15. Centre SLA, CHRU de Tours, Tours, France

16. Rita Levi Montalcini, Department of Neuroscience, ALS Centre, University of Torino, Turin, Italy

17. Azienda Ospedaliera Citta della Salute e della Scienza, Torino, Italy

18. Federation des Centres SLA Tours and Limoges, LITORALS, Tours, France

19. Physiology Institute, Faculty of Medicine, Instituto de Medicina Molecular, University of Lisbon, Lisbon, Portugal

20. Department of Neurology, Hadassah Medical Organization and Faculty of Medicine, Hebrew University of Jerusalem, Israel

21. Neurology Department, Hospital Universitari de Bellvitge, Barcelona, Spain

22. Hospital San Rafael, Madrid, Spain

23. Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA

24. Department of Neurology, Center for Neurodegenerative Diseases, Emory University, Atlanta, Georgia USA

25. King's College Hospital, Denmark Hill, London, SE5 9RS, UK

26.Academic Unit of Neurology, Trinity College Dublin, Trinity Biomedical Sciences Institute, Dublin, Republic of Ireland

27. Department of Neurology, Beaumont Hospital, Dublin, Republic of Ireland

28. Neurology Department, University Hospitals Leuven, Leuven, Belgium

[+]AAK and AI contributed equally

*** Correspondence:**
Professor Ammar Al-Chalabi and Dr Ahmad Al Khleifat.

E-mail: ammar.al-chalabi@kcl.ac.uk Ahmad.al_khleifat@kcl.ac.uk

Keywords: ALS, Telomere, Next generation sequencing, Whole genome sequencing, Variant calling.

Abstract

Amyotrophic lateral sclerosis (ALS), a neurodegenerative disease of motor neurons, is a complex genetic disease with heritability of 60%. Only about 14% of apparently sporadic ALS is explained by known genetic variation, suggesting that other forms of genetic variation are important, in addition to the known risk factors, male sex and increasing age. Telomeres maintain DNA integrity during cellular replication, differ between sexes, and shorten naturally with age. We find that longer telomeres are associated with ALS and specific phenotypic patterns of disease expression. We see the same pattern of telomere elongation in ALS in brain tissue. The association of longer telomeres with apparently sporadic ALS was also seen in people with familial ALS, supporting the notion that sporadic and familial ALS are not mutually exclusive categories but rather a spectrum.

Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease affecting motor neurons in the brain and spinal cord resulting in progressive paralysis and death, within three to five years, typically due to respiratory failure (Brown and Al-Chalabi, 2017; Hardiman *et al.*, 2017). The first symptoms of weakness can occur in the bulbar innervated muscles, manifesting as difficulty with speech or swallowing, or in the spinal innervated muscles, manifesting as limb weakness or breathing difficulty. About 5% may have a frank frontotemporal dementia, and frontotemporal impairment is seen in up to 80% of people by the time King's Stage 4 disease is reached (Roche *et al.*, 2012; Crockford *et al.*, 2018).

The last decade has seen substantial advances in our understanding of the genomic basis of ALS (Hop *et al.*, 2022; Wouter van Rheenen *et al.*, 2021) but a significant proportion of the genetic contribution to risk remains unexplained. This hidden heritability may be harbored in other types of genomic variation as well as in rare variants that may be unique to an affected individual or family(Al-Chalabi *et al.*, 2010; McLaughlin, Vajda and Hardiman, 2015).

Telomeres are repeated TTAGGG nucleotide sequences located at the ends of chromosomes and exist to maintain chromosomal structural integrity during cellular replication. They shorten naturally with age and differ in average length between the sexes(Muzumdar and Atzmon, 2012; Kong, Lee and Wang, 2013; Gardner *et al.*, 2014); age and sex are also risk factors for ALS(McCombe and Henderson, 2010; Al-Chalabi and Hardiman, 2013b; Westeneng *et al.*, 2018). Telomere length is a marker for aging, chromosomal instability and DNA damage, and might therefore be relevant as a risk factor for ALS(Murnane, 2006; Conomos *et al.*, 2012; Xu, Li and Stohr, 2013; Marzec *et al.*, 2015).

Previously, in a pilot study in a UK cohort, we showed that longer telomeres might be associated with ALS when compared to age and sex-matched controls(Al Khleifat *et al.*, 2019). We therefore sought to explore this finding in detail, using whole-genome sequence data from the Project MinE consortium(Van Rheenen *et al.*, 2018), a large international ALS genomics collaboration.

Materials and Methods

Data Sources and data extraction

Blood samples

Samples were from the international Project MinE whole genome sequencing consortium and derived from seven countries: the USA, Ireland, Belgium, the Netherlands, Spain, Turkey, and the United Kingdom(Van Rheenen *et al.*, 2018).

DNA was isolated from venous blood using standard methods. The DNA concentrations were set at 100ng/ul as measured by a fluorimeter with the PicoGreen® dsDNA quantitation assay. DNA integrity was assessed using gel electrophoresis.

Post-mortem samples

Post-mortem motor cortex was from the MRC London Neurodegenerative Diseases Brain Bank based at the Institute of Psychiatry, Psychology & Neuroscience, King's College London. Tissue was flash frozen stored at −80 °C. 100mg tissue blocks were excised. DNA was isolated from the same tissue block and sequenced. The study cohort consisted of 64 people with apparently sporadic ALS and 53 controls with no known neurological disease (controls below Hyperphosphorylated tau (HP-τ) in human brain tissue and BNE\Braak stage 2).

The 100mg tissue blocks were divided to allow DNA purification. For each sample, a 25mg tissue block for DNA was homogenised using a Qiagen PowerLyzer 24 Homogenizer. DNA was purified from the homogenate using the standard protocol from Qiagen's DNeasy Blood and Tissue Mini Kit. DNA was quantified using PicoGreen (Quant-iT™ PicoGreen® dsDNA Reagent, ThermoFisher Scientific) and measured using a Spectromax Gemini XPS (Molecular Devices).

Library preparation and DNA sequencing

Library preparation was performed using the Illumina DNA Sample Preparation HT Kit alongside the Illumina SeqLab DNA PCR-Free Library Prep Guide. Libraries were then quantified using qPCR and

evaluated using gel electrophoresis. All samples were sequenced using Illumina's FastTrack services (San Diego, CA, USA). Some blood-derived DNA samples were sequenced using the Illumina HiSeq 2000 platform. Sequencing was 100bp paired-end performed using PCR-free library preparations and targeted ~40x coverage across each sample. Remaining blood-derived and all brain-derived DNA libraries were clustered onto flow cells using the Illumina cBot System, as per cBot System Guide using Illumina HiSeq X HD Paired End Cluster Kit reagents and sequenced on an Illumina HiSeqX with 151bp paired-end runs using independent flow cell lanes and with a target minimum of 30x average coverage per sample. Binary sequence alignment/map formats (BAM) were generated for each individual. All the genomes were aligned with Isaac (Illumina) to hg19. The details of the Isaac alignment and variant calling pipelines are discussed in Project MinE design (Van Rheenen *et al.*, 2018) and Isaac protocol (Raczy *et al.*, 2013).

Determination of Telomere Length

TelSeq(Ding *et al.*, 2014) was used to quantify telomere length using whole genome sequence data. Telomere lengths were estimated from reads, defined as repeats of more than seven TTAGGG motifs.

Statistical Analysis

The effect of telomere length on ALS risk was tested using a multivariable linear regression model. To account for different sequencing platforms and population stratification, principal components of ancestry, centre and technology platform were included as covariates. To assess the model, Pearson's chi-squared test was used. Because telomere length correlates with age, we performed an additional test to examine the possibility that survival bias could affect the results. To do this, we also performed the analysis restricted to the subgroup of people with ALS onset below the median cohort age (62 years). As brain is composed of neurons which do not divide, as well as glia which do, we expected that the average telomere length in brain would be longer than in blood. Furthermore, nervous tissue is the target of the disease process rather than blood. We therefore additionally tested the effect of age on telomere length in brain tissue.

To determine if telomeres are lengthened in ALS, or simply shorten less rapidly than in controls, we analysed the effect of age on telomere length in each group using multivariable linear regression.

To assess the effect of covariates on telomere length affecting survival, we used Cox regression, controlling for age, sex and site of disease onset (bulbar or spinal), population stratification, principal components, centre and sequencing platforms.

Repeat primed PCR and Expansion Hunter-v2.5.1(Dolzhenko *et al.,* 2017) were used to assay the hexanucleotide repeat expansion in the *C9orf72* gene since this is a known risk factor for ALS and associates with survival.

Statistical tests were performed using IBM SPSS Statistics 24.0 (SPSS Inc., Illinois), and RStudio, R Foundation for Statistical Computing 3.4.1.

Quality Control

Quality control was performed separately on the genotyped data of each population as reported previously(van Rheenen *et al.,* 2016) (Supplementary appendix).

Ethical approval

Informed consent was obtained from all participants in this project under reference Trent Research Ethics Committee 08/H0405/60.

Results

There were 6,580 whole genome sequences (4,515 from people with ALS and 2,065 from controls), reducing to 6,195 (4,315 (95.6%) from people with ALS and 1,880 (91%) from controls) after quality control, with minimum ~25x coverage across each sample. The set was enriched for apparently sporadic ALS (4,236 compared with 79 with familial ALS (FALS)). The male-female ratio was 2:1. Overall, 22 had ALS-frontotemporal dementia (ALS-FTD). Phenotypically, 37 had pure progressive bulbar palsy (PBP) and 68 and progressive muscular atrophy. There were 1,908 sequenced using the HiSeq2000 platform and 4,287 sequenced using the HiSeqX Illumina platform (Table 1). There were 344 people carrying an expanded *C9orf72* hexanucleotide repeat, 334 with ALS and 10 without symptoms.

The mean telomere length in people with ALS was 5.5kb, and in controls, 5.38kb (Figure 1). Multivariable linear regression accounting for sex and age as covariates showed a mean 20% (95% CI 14%, 25%) longer telomere length in people with ALS compared to controls ($p = 1.1 \times 10^{-12}$). Covariate analysis showed that regardless of disease status, females ($p = 2.42 \times 10^{-5}$) and younger people ($p = 1.2 \times 10^{-16}$) had on average longer telomeres (Table 2), confirming the results of earlier studies that telomere length reduces with age and females have on average longer telomeres.

To assess if the observed longer telomere length in apparently sporadic ALS is also seen in familial ALS, we assessed telomere length in 79 people, not included in the main analysis, with a family history of ALS in a first degree relative (FALS). Multivariable linear regression after correcting for age and sex again showed a longer average telomere length in people with FALS than in controls ($p = 2.0 \times 10^{-16}$). To ensure that telomere length analysis was not biased, we excluded the UK, a population we used previously for discovery analysis using a subset of samples, and the association was still observed ($p = 6.6 \times 10^{-9}$).

Examining the effect of age on telomere length in ALS and controls separately, showed that the rate of shortening by age is slower in ALS than in controls, suggesting it is not an active lengthening of telomeres in ALS (0.022%, $p < 0.0001$ vs. 0.012%, $p < 0.0001$) and also arguing against the possibility that telomeres are longer to start with in people who will later develop ALS. The rate of shortening

was different in males and females, with females showing a faster rate of shortening than males (Supplementary Figure 1).

In an analysis exploring survival bias as an explanation for our results, we restricted testing to those younger than the median age (62 years). Multivariable linear regression accounting for sex and age still showed that telomeres were longer in people with ALS compared to controls ($p = 8.12 \times 10^{-12}$) with mean telomere length in people with ALS, 5.8kb, and in controls, 5.5kb. To ensure that telomere length analysis was not biased by population effects, we excluded the UK, a population we used previously for discovery analysis, and using a subset of samples the association was still observed ($p = 6.6 \times 10^{-9}$).

We compared telomere length in 334 people with ALS with *C9orf72* repeat expansion against people with ALS with confirmed non-expanded *C9orf72* status. Multivariable linear regression showed that the telomere was shorter in expansion carriers ($p = 5.0 \times 10^{-4}$) Figure 2A, Table 3. Although ALS *C9orf72* expansion carriers had a shorter telomere length than non-expansion carriers, telomere length was still longer in those carrying a *C9orf72* repeat expansion than controls ($p = 0.001$) Figure 2B.

We therefore assessed the relationship between the estimated number of telomere repeats and estimated number of *C9orf72* repeats using ExpansionHunter in 1589 samples from the UK. Multivariable linear regression showed that the number of telomere repeats is associated negatively with *C9orf72* repeat expansion size ($p = 0.003$) supporting the previous results.

To assess if the relationship between *C9orf72* repeat expansion and telomere repeat size was specific to *C9orf72*-mediated ALS, we ran ExpansionHunter on three ALS genes which also contain disease-associated repeat expansions: *ATXN1*, *ATXN2* and *NIPA1*. There was no difference in telomere length observed between expansion carriers and non-expansion carriers for any of these genes.

Cox regression analysis showed that people with ALS with telomere length less than 5.3Kb had a 10% increase in median survival compared with those with longer telomeres ($p = 5.0 \times 10^{-7}$) after correcting for age, sex, site of onset, *C9orf72* status and principal components of ancestry.

To validate our findings in a different tissue we used 159 *post-mortem* brain samples, 106 from people with apparently sporadic ALS and 53 controls. The male-female ratio was 2:1. The mean telomere length in people with ALS was 6.8kb, and in controls, 6.56Kb, not taking into account gender or age. Multivariable linear regression accounting for these covariates showed that telomere length in people with ALS was longer by mean 29% (95% CI 30%, 55%) compared with controls ($p = 0.03$).

Discussion

Using a large disease-specific whole genome sequencing dataset, we have shown that longer telomeres are associated with ALS, confirming initial findings from a pilot study (Al Khleifat *et al.*, 2019). We were additionally able to show that our findings are likely a result of less rapid shortening of telomeres being associated with ALS, rather than active telomere lengthening in ALS. In keeping with expectations, we also found that mean telomere length was on average longer in females, and in all samples, shortened with increasing age. The association of longer telomeres with apparently sporadic ALS was also seen in FALS, supporting the notion that familial and sporadic ALS are not mutually exclusive categories but rather a spectrum (Al-Chalabi and Hardiman, 2013a; Al-Chalabi *et al.*, 2014; Chiò *et al.*, 2018; Mehta *et al.*, 2018). Furthermore, telomere length was inversely correlated with *C9orf72* repeat expansion size.

Telomere elongation phenomena are well documented but far less well understood than telomere shortening phenomena (Bryan *et al.*, 1995; Cesare and Reddel, 2010; Arora and Azzalin, 2015; Haycock *et al.*, 2017). While telomere shortening is typically seen in cancers, telomere elongation an occur in cancers of the nervous system, and for example, is seen in 25% of primary brain tumours, in glioblastoma multiforme and in 10% of neuroblastomas (Bryan *et al.*, 1997; Hakin-Smith *et al.*, 2003; Henson *et al.*, 2005; Durant, 2012; Boutou *et al.*, 2013). In general, cancers in which cells have long telomeres are resistant to therapy and carry a poor prognosis (Haycock *et al.*, 2017). Telomere elongation as also been associated with schizophrenia (Nieratschker *et al.*, 2013; Zhang *et al.*, 2018), a disorder that genetically overlaps with ALS (McLaughlin *et al.*, 2017). Additionally, longer telomeres are also reported in Parkinson's disease and Lewy body dementia blood and brain(Asghar *et al.*, 2022).

We found that pathologically expanded *C9orf72* repeats are negatively associated with telomere repeat length, so that people with expanded *C9orf72* repeats had shorter telomeres on average. *C9orf72* gene repeat expansion is the most frequent genetic cause of ALS and of the related condition, frontotemporal dementia(Shatunov *et al.*, 2010; Smith *et al.*, 2013; Hardiman *et al.*, 2017; Iacoangeli *et al.*, 2019). A possible explanation for the negative association is in the liability threshold model of disease. Those people who already have a high liability to ALS do not need the additional liability of longer telomeres

and so on average would appear to have shorter telomeres. Alternatively, those in the higher risk group (non-repeat expansion carriers) need a greater contribution from other sources of disease liability and so have longer telomere repeats. Against this explanation is the observation that other gene variants that predispose to ALS risk such as intermediate *ATXN2* repeat expansions, do not show any association with telomere length, and neither do people with familial ALS. The explanation might therefore lie in the *C9orf72* repeat expansion itself. Both telomeres and large *C9orf72* repeats have a tendency to fold into structures called G quadruplexes(Fratta *et al.*, 2012; Grigg, Shumayrikh and Sen, 2014). G quadruplex structures have important roles in DNA replication, recombination and telomere maintenance(Millevoi, Moine and Vagner, 2012; Bryan, 2019). Although there have been several studies of the G quadruplexes formed by *C9orf72* repeat expansion, the relationship between *C9orf72* repeats and other G quadruplexes such as telomeres is not documented at population level and has not been well characterized in vivo(Zhang *et al.*, 2019).

Some genetic variations that contribute to ALS risk also worsen prognosis. This is seen in carriers of the *C9orf72* repeat expansion mutation, those carrying the *UNC13A* homozygous risk genotype, and for some variants of the *SOD1* gene for example. In keeping with that pattern, we have found that longer telomere length is associated with ALS risk as well as with worse prognosis, with those with the shortest telomeres having a 10% increase in survival.

The genetic landscape of ALS is one of some monogenic causes, several gene variations that substantially but not dramatically increase risk, and a polygenic component. For those with a monogenic basis of their disease, the variable and age-dependent penetrance seen is likely because of the contribution of other factors to risk(Al-Chalabi and Hardiman, 2013a; Al-Chalabi *et al.*, 2014; Chiò *et al.*, 2018; Garton *et al.*, 2021)**.** Based on these findings, telomere length could be such a factor.

The main limitation of this study is that we did not directly measure telomere length using Southern blotting, but estimated it using whole genome sequence data. The method we have used, TelSeq, was recently used to estimate telomere length in 75,000 whole genome sequences. Comparing the performance of TelSeq with other bioinformatics tools such as Computel, the estimates of telomere

length were highly correlated between the bioinformatics methods and with Southern blot results, with the advantage that TelSeq had a faster processing time(Taub *et al.*, 2019), although it is not possible to draw conclusions about the exact length of a telomere. Other studies have also shown a good correlation between TelSeq telomere length estimates, Southern blotting and Q-PCR(Ding *et al.*, 2014; Cook *et al.*, 2016). With this in mind, different sequencing technologies might generate different telomere length estimates because of differences in library preparation and platform(Ma, 1995; Aviv *et al.*, 2011). To overcome this potential weakness, we have used the same industry-leading sequencing platform for all samples, as well as designing the study to minimize batch effects by having cases and controls sharing the same sequencing plate. Our study has the advantage of a large sample size of more than 4,500 cases, far larger than for previous reports.

We have shown that despite being an age-related, male-predominant condition, ALS is associated with longer telomere lengths in blood-derived and in brain-derived DNA.

Conflict of Interest

AAC is a consultant for Mitsubishi-Tanabe Pharma, GSK, and Chronos Therapeutics, and chief investigator for clinical trials for Cytokinetics and OrionPharma. JHV reports to have sponsored research agreements with Biogen. VS is a consultant for Novartis and Biogen. LHVDB reports grants from Netherlands ALS Foundation, grants from The Netherlands Organization for Health Research and Development (Vici scheme), grants from The European Community's Health Seventh Framework Programme (grant agreement n° 259867 (EuroMOTOR)), grants from The Netherlands Organization for Health Research and Development)the STRENGTH project, funded through the EU Joint Programme – Neurodegenerative Disease Research, JPND), during the conduct of the study; personal fees from Calico, personal fees from Cytokinetics, grants and personal fees from Takeda, non-financial support from Orion, non-financial support from Orphazyme, outside the submitted work; AAC serves on scientific advisory boards for Mitsubishi Tanabe, Roche, Denali Pharma, Cytokinetics, Lilly, and Amylyx and has received a reeseach grant from Biogen. CES reports grants from Avexis , grants from Eli Lilly, grants from Chronos Therapeutics, grants from Vertex Pharmaceuticals, during the conduct of the study; grants from QurAlis, grants from Chronos Therapeutics, grants from Biogen, outside the submitted work. JEL is a member of the scientific advisory board for Cerevel Therapeutics, a consultant for ACI Clinical LLC sponsored by Biogen, Inc. or Ionis Pharmaceuticals, Inc. JEL is also a consultant for Perkins Coie LLP and may provide expert testimony. JEL was supported by funding from NIH/NINDS (R01NS073873 and R56NS073873). AAK, OH, MPP, JSM, PJS, JEL, CES, NB, OH, WR, PVD, NB, KK, BK, HB and LHVB declare no competing interests.

Author Contributions

AAC and AAK conceived and planned the study. AAK and AI created the bioinformatics pipeline for analysis. JJFAV ran ExpansionHunter on Project MinE data. MM and RAJZ prepared phenotypic data. AAK and AAC did the statistical analysis and prepared the figures and tables. AAC, JHV, OH, MPP, JSM, PJS, JEL, CES, NB, OH, WR, PVD and LHVB helped in sample collection and provided whole genome sequence data, analysis and intellectual input for data interpretation on behalf of the Project MinE Consortium. JHV, OH, and MM provided intellectual input for data interpretation. AAC and AAK wrote the first draft of the manuscript. All authors reviewed and approved the final manuscript.

Funding

AAK is funded by ALS Association Milton Safenowitz Research Fellowship (grant number22-PDF-609.DOI :10.52546/pc.gr.150909.), The Motor Neurone Disease Association (MNDA) Fellowship (Al Khleifat/Oct21/975-799), The Darby Rimmer Foundation, and The NIHR Maudsley Biomedical Research Centre. This project was funded by the MND Association and the Wellcome Trust. This is an EU Joint Programme-Neurodegenerative Disease Research (JPND) project. The project is supported through the following funding organisations under the aegis of JPND - www.jpnd.eu (United Kingdom, Medical Research Council (MR/L501529/1 and MR/R024804/1) and Economic and Social Research Council (ES/L008238/1)). AAC is an NIHR Senior Investigator. C.E.S. and A.A.C. receive salary support from the National Institute for Health Research (NIHR) Dementia Biomedical Research Unit at South London and Maudsley NHS Foundation Trust and King's

College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The work leading up to this publication was funded by the European Community's Health Seventh Framework Program (FP7/2007–2013; grant agreement number 259867) and Horizon 2020 Program (H2020-PHC-2014-two-stage; grant agreement number 633413). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 772376 - EScORIAL. The collaboration project is co-funded by the PPP Allowance made available by Health~Holland, Top Sector Life Sciences & Health, to stimulate public-private partnerships. Project MinE Belgium was supported by a grant from IWT, the Belgian ALS Liga and a grant from Opening the Future Fund (KU Leuven). PVD holds a senior clinical investigatorship of FWO-Vlaanderen and is supported by E. von Behring Chair for Neuromuscular and Neurodegenerative Disorders, the ALS Liga België and the KU Leuven funds "Een Hart voor ALS", "Laeversfonds voor ALS Onderzoek" and the "Valéry Perrier Race against ALS Fund". R.L.McL. is supported by Science Foundation Ireland (17/CDA/4737). MinE USA is funded by the US ALS Association. We thank the patients and families from the Emory ALS Clinic participating in this research. Funding was provided by US National Institutes of Health (NIH)/National Institute of Neurological Disorders and Stroke (NINDS) (R01NS073873, J.E.L.) and the American ALS Association (J.E.L.).

Acknowledgments

Samples used in this research were in part obtained from the UK National DNA Bank for MND Research, funded by the MND Association and the Wellcome Trust. We thank people with MND and their families for their participation in this project. We acknowledge sample management undertaken by Biobanking Solutions funded by the Medical Research Council at the Centre for Integrated Genomic Medical Research, University of Manchester. The authors acknowledge use of the research computing facility at King's College London, *Rosalind* (https://rosalind.kcl.ac.uk), which is delivered in partnership with the National Institute for Health Research (NIHR) Biomedical Research Centres at South London & Maudsley and Guy's & St. Thomas' NHS Foundation Trusts, and part-funded by capital equipment grants from the Maudsley Charity (award 980) and Guy's & St. Thomas' Charity (TR130505). The views expressed are those of the author (s) and not necessarily those of the NHS, the NIHR, King's College London, or the Department of Health and Social Care. National. The authors acknowledge Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The authors acknowledge Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome Trust.

Data Availability Statement

The blood and brain-derived datasets that support the findings in this study are available from The Project MinE consortium public repository. To gain access to the data, an account request must be made to info@projectmine.com. Data access will require the completion of a data access request.

Further information about data access can be found at https://www.projectmine.com/research/data-sharing/

Author Contributions

AAC and AAK conceived and planned the study. AAK and AI created the bioinformatics pipeline for analysis. JJFAV ran ExpansionHunter on Project MinE data. MM and RAJZ prepared phenotypic data. AAK and AAC did the statistical analysis and prepared the figures and tables. AAC, JHV, OH, MPP, JSM, PJS, JEL, CES, NB, OH, WR, PVD and LHVB helped in sample collection and provided whole genome sequence data, analysis and intellectual input for data interpretation on behalf of the Project MinE Consortium. JHV, OH, and MM provided intellectual input for data interpretation. AAC and AAK wrote the first draft of the manuscript. All authors reviewed and approved the final manuscript.

Competing Interests

AAC is a consultant for Mitsubishi-Tanabe Pharma, GSK, and Chronos Therapeutics, and chief investigator for clinical trials for Cytokinetics and OrionPharma. JHV reports to have sponsored research agreements with Biogen. VS is a consultant for Novartis and Biogen. LHVDB reports grants from Netherlands ALS Foundation, grants from The Netherlands Organization for Health Research and Development (Vici scheme), grants from The European Community's Health Seventh Framework Programme (grant agreement n° 259867 (EuroMOTOR)), grants from The Netherlands Organization for Health Research and Development)the STRENGTH project, funded through the EU Joint Programme – Neurodegenerative Disease Research, JPND), during the conduct of the study; personal fees from Calico, personal fees from Cytokinetics, grants and personal fees from Takeda, non-financial support from Orion, non-financial support from Orphazyme, outside the submitted work; AAC serves on scientific advisory boards for Mitsubishi Tanabe, Roche, Denali Pharma, Cytokinetics, Lilly, and Amylyx and has received a reeseach grant from Biogen. CES reports grants from Avexis , grants from Eli Lilly, grants from Chronos Therapeutics, grants from Vertex Pharmaceuticals, during the conduct of the study; grants from QurAlis, grants from Chronos Therapeutics, grants from Biogen, outside the submitted work. JEL is a member of the scientific advisory board for Cerevel Therapeutics, a consultant for ACI Clinical LLC sponsored by Biogen, Inc. or Ionis Pharmaceuticals, Inc. JEL is also a consultant for Perkins Coie LLP and may provide expert testimony. JEL was supported by funding from NIH/NINDS (R01NS073873 and R56NS073873). AAK, OH, MPP, JSM, PJS, JEL, CES, NB, OH, WR, PVD, NB, KK, BK, HB and LHVB declare no competing interests.

References

Al-Chalabi, A. *et al.* (2010) 'An estimate of amyotrophic lateral sclerosis heritability using twin data.', *Journal of neurology, neurosurgery, and psychiatry*, 81(12), pp. 1324–1326. doi: 10.1136/jnnp.2010.207464.

Al-Chalabi, A. *et al.* (2014) 'Analysis of amyotrophic lateral sclerosis as a multistep process: A population-based modelling study', *The Lancet Neurology*, 13(11), pp. 1108–1113. doi: 10.1016/S1474-4422(14)70219-4.

Al-Chalabi, A. and Hardiman, O. (2013a) 'The epidemiology of ALS: a conspiracy of genes, environment and time.', *Nature reviews. Neurology*, 9(11), pp. 617–28. doi: 10.1038/nrneurol.2013.203.

Al-Chalabi, A. and Hardiman, O. (2013b) 'The epidemiology of ALS: A conspiracy of genes, environment and time', *Nature Reviews Neurology*. doi: 10.1038/nrneurol.2013.203.

Arora, R. and Azzalin, C. M. (2015) 'Telomere elongation chooses TERRA ALTernatives', *RNA Biology*, 12(9), pp. 938–941. doi: 10.1080/15476286.2015.1065374.

Asghar, M. *et al.* (2022) 'Mitochondrial biogenesis, telomere length and cellular senescence in Parkinson's disease and Lewy body dementia', *Scientific Reports*. Nature Research, 12(1). doi: 10.1038/s41598-022-22400-z.

Aviv, A. *et al.* (2011) 'Impartial comparative analysis of measurement of leukocyte telomere length/DNA content by Southern blots and qPCR', *Nucleic Acids Research*. Oxford University Press, 39(20), pp. e134–e134. doi: 10.1093/nar/gkr634.

Boutou, E. *et al.* (2013) 'DNA Repair and Telomeres — An Intriguing Relationship', in *New Research Directions in DNA Repair*. doi: 10.5772/56115.

Brown, R. H. and Al-Chalabi, A. (2017) 'Amyotrophic Lateral Sclerosis', *New England Journal of Medicine*, 377(2), pp. 162–172. doi: 10.1056/NEJMra1603471.

Bryan, T. M. *et al.* (1995) 'Telomere elongation in immortal human cells without detectable telomerase activity', *EMBO J*, 14(17), pp. 4240–4248. doi: papers3://publication/uuid/E553EAC0-7D77-471F-A996-D41412778FFB.

Bryan, T. M. *et al.* (1997) 'Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines', *Nature Medicine*. doi: 10.1038/nm1197-1271.

Bryan, T. M. (2019) 'Mechanisms of DNA replication and repair: Insights from the study of G-quadruplexes', *Molecules*. doi: 10.3390/molecules24193439.

Cesare, A. J. and Reddel, R. R. (2010) 'Alternative lengthening of telomeres: Models, mechanisms and implications', *Nature Reviews Genetics*. doi: 10.1038/nrg2763.

Chiò, A. *et al.* (2018) 'The multistep hypothesis of ALS revisited: The role of genetic mutations.', *Neurology*. Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology, p. 10.1212/WNL.0000000000005996. doi: 10.1212/WNL.0000000000005996.

Conomos, D. *et al.* (2012) 'Variant repeats are interspersed throughout the telomeres and recruit nuclear receptors in ALT cells', *Journal of Cell Biology*. doi: 10.1083/jcb.201207189.

Cook, D. E. *et al.* (2016) 'The genetic basis of natural variation in Caenorhabditis elegans telomere length', *Genetics*. doi: 10.1534/genetics.116.191148.

Crockford, C. *et al.* (2018) 'ALS-specific cognitive and behavior changes associated with advancing disease stage in ALS', *Neurology*. doi: 10.1212/WNL.0000000000006317.

Ding, Z. *et al.* (2014) 'Estimating telomere length from whole genome sequence data', *Nucleic Acids Research*, 42(9). doi: 10.1093/nar/gku181.

Dolzhenko, E. *et al.* (2017) 'Detection of long repeat expansions from PCR-free whole-genome sequence data', *Genome Research*. doi: 10.1101/gr.225672.117.

Durant, S. T. (2012) 'Telomerase-independent paths to immortality in predictable cancer sub-types', *Journal of Cancer*. doi: 10.7150/jca.3965.

Fratta, P. *et al.* (2012) 'C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes', *Scientific Reports*. doi: 10.1038/srep01016.

Gardner, M. *et al.* (2014) 'Gender and telomere length: Systematic review and meta-analysis', *Experimental Gerontology*. doi: 10.1016/j.exger.2013.12.004.

Garton, F. C. *et al.* (2021) 'Cardiovascular disease, psychiatric diagnosis and sex differences in the multistep hypothesis

of amyotrophic lateral sclerosis', *European Journal of Neurology*, 28(2). doi: 10.1111/ene.14554.

Grigg, J. C., Shumayrikh, N. and Sen, D. (2014) 'G-quadruplex structures formed by expanded hexanucleotide repeat RNA and DNA from the neurodegenerative disease-linked C9orf72 gene efficiently sequester and activate heme', *PLoS ONE*. doi: 10.1371/journal.pone.0106449.

Hakin-Smith, V. *et al.* (2003) 'Alternative lengthening of telomeres and survival in patients with glioblastoma multiforme', *Lancet*. doi: 10.1016/S0140-6736(03)12681-5.

Hardiman, O. *et al.* (2017) 'Amyotrophic lateral sclerosis', *Nature Reviews Disease Primers*. Nature Publishing Group, 3, p. 17071. doi: 10.1038/nrdp.2017.71.

Haycock, P. C. *et al.* (2017) 'Association between telomere length and risk of cancer and non-neoplastic diseases a mendelian randomization study', *JAMA Oncology*, 3(5), pp. 636–651. doi: 10.1001/jamaoncol.2016.5945.

Henson, J. D. *et al.* (2005) 'A robust assay for alternative lengthening of telomeres in tumors shows the significance of alternative lengthening of telomeres in sarcomas and astrocytomas', *Clinical Cancer Research*.

Hop, P. J. *et al.* (2022) 'Genome-wide study of DNA methylation in Amyotrophic Lateral Sclerosis identifies differentially methylated loci and implicates metabolic, inflammatory and cholesterol pathways', *Philippe Couratier*, 18, p. 37. doi: 10.1101/2021.03.12.21253115.

Iacoangeli, A. *et al.* (2019) 'C9orf72 intermediate expansions of 24–30 repeats are associated with ALS', *Acta Neuropathologica Communications*. doi: 10.1186/s40478-019-0724-4.

Al Khleifat, A. *et al.* (2019) 'Telomere length is greater in ALS than in controls: a whole genome sequencing study', *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. doi: 10.1080/21678421.2019.1586951.

Kong, C. M., Lee, X. W. and Wang, X. (2013) 'Telomere shortening in human diseases', *FEBS Journal*, pp. 3180–3193. doi: 10.1111/febs.12326.

Ma, T. S. (1995) 'Applications and limitations of polymerase chain reaction amplification', *Chest*, 108(5), pp. 1393–1404. doi: 10.1378/chest.108.5.1393.

Marzec, P. *et al.* (2015) 'Nuclear-Receptor-Mediated Telomere Insertion Leads to Genome Instability in ALT Cancers',

Cell. doi: 10.1016/j.cell.2015.01.044.

McCombe, P. A. and Henderson, R. D. (2010) 'Effects of gender in amyotrophic lateral sclerosis', *Gender Medicine*. doi: 10.1016/j.genm.2010.11.010.

McLaughlin, L. R., Vajda, A. and Hardiman, O. (2015) 'Heritability of amyotrophic lateral sclerosis insights from disparate numbers', *JAMA Neurology*. doi: 10.1001/jamaneurol.2014.4049.

McLaughlin, R. L. *et al.* (2017) 'Genetic correlation between amyotrophic lateral sclerosis and schizophrenia', *Nature Communications*. doi: 10.1038/ncomms14774.

Mehta, P. R. *et al.* (2018) 'Younger age of onset in familial amyotrophic lateral sclerosis is a result of pathogenic gene variants, rather than ascertainment bias.', *Journal of neurology, neurosurgery, and psychiatry*. BMJ Publishing Group Ltd, p. jnnp-2018-319089. doi: 10.1136/jnnp-2018-319089.

Millevoi, S., Moine, H. and Vagner, S. (2012) 'G-quadruplexes in RNA biology', *Wiley Interdisciplinary Reviews: RNA*. doi: 10.1002/wrna.1113.

Murnane, J. P. (2006) 'Telomeres and chromosome instability', *DNA Repair*. doi: 10.1016/j.dnarep.2006.05.030.

Muzumdar, R. and Atzmon, G. (2012) 'Telomere Length and Aging', in *Reviews on Selected Topics of Telomere Biology*, pp. 3–30. doi: 10.5772/2329.

Nieratschker, V. *et al.* (2013) 'Longer telomere length in patients with schizophrenia', *Schizophrenia Research*. doi: 10.1016/j.schres.2013.06.043.

Raczy, C. *et al.* (2013) 'Isaac: Ultra-fast whole-genome secondary analysis on Illumina sequencing platforms', *Bioinformatics*. doi: 10.1093/bioinformatics/btt314.

van Rheenen, W. *et al.* (2016) 'Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis', *Nature Genetics*, 48(9), pp. 1043–1048. doi: 10.1038/ng.3622.

Van Rheenen, W. *et al.* (2018) 'Project MinE: study design and pilot analyses of a large-scale whole-genome sequencing study in amyotrophic lateral sclerosis', *European Journal of Human Genetics*, 26(10). doi: 10.1038/s41431-018-0177-4.

559 Roche, J. C. *et al.* (2012) 'A proposed staging system for amyotrophic lateral sclerosis', *Brain*, 135(3), pp. 847–852. doi:
560 10.1093/brain/awr351.

561 Shatunov, A. *et al.* (2010) 'Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other
562 countries: A genome-wide association study', *The Lancet Neurology*. doi: 10.1016/S1474-4422(10)70197-6.

563 Smith, B. N. *et al.* (2013) 'The C9ORF72 expansion mutation is a common cause of ALS+/−FTD in Europe and has a
564 single founder', *European Journal of Human Genetics*. Nature Publishing Group, 21(1), pp. 102–108. doi:
565 10.1038/ejhg.2012.98.

566 Taub, M. A. *et al.* (2019) 'Novel genetic determinants of telomere length from a multi-ethnic analysis of 75,000 whole
567 genome sequences in TOPMed', *bioRxiv*. Cold Spring Harbor Laboratory, p. 749010. doi: 10.1101/749010.

568 Westeneng, H. J. *et al.* (2018) 'Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a
569 personalised prediction model', *The Lancet Neurology*. doi: 10.1016/S1474-4422(18)30089-9.

570 Wouter van Rheenen, A. *et al.* (2021) 'Common and rare variant association analyses in Amyotrophic Lateral Sclerosis
571 identify 15 risk loci with distinct genetic architectures and neuron-specific biology', *Anneke J. van der Kooi*, 17(103), p.
572 129. doi: 10.1101/2021.03.12.21253159.

573 Xu, L., Li, S. and Stohr, B. A. (2013) 'The Role of Telomere Biology in Cancer', *Annual Review of Pathology:*
574 *Mechanisms of Disease*, 8(1), pp. 49–78. doi: 10.1146/annurev-pathol-020712-164030.

575 Zhang, S. *et al.* (2019) 'Evaluation of the selectivity of G-quadruplex ligands in living cells with a small molecule
576 fluorescent probe', *Analytica Chimica Acta: X*. doi: 10.1016/j.acax.2019.100017.

577 Zhang, Y. *et al.* (2018) 'Longer telomeres in elderly schizophrenia are associated with long-term hospitalization in the
578 Japanese population', *Journal of Psychiatric Research*. doi: 10.1016/j.jpsychires.2018.05.014.

579

580

581

Tables

Cohort	Sample	Case	Control	Female	Male
Belgium	548	368	180	209	339
Ireland	403	267	136	161	242
Netherlands	2894	1859	1035	1182	1712
Spain	338	233	105	145	193
Turkey	223	148	75	87	136
United Kingdom	1402	1124	278	603	799
United States	387	316	71	153	234
Total	6195	4315	1880	2540	3655

Table1. Detailed demographic features of the study population.

	Estimate	SD of estimate	p=Value
Age (per year)	-2%	0.02	1.1×10^{-16}
Sex (male vs. female)	-8%	0.03	2.42×10^{-5}
Case-control status (controls vs. cases)	-29%	0.02	1.1×10^{-12}

Table 2. Telomere length comparison between people with ALS and healthy controls using a generalized linear model.

	Estimate	**SD of estimate**	**p=Value**
Age (per year)	-2%	0.02	1.1×10^{-16}
Sex (male vs. female)	-15%	0.03	2.42×10^{-5}
***C9orf72* (expanded)**	-27%	0.02	5.0×10^{-4}

Table 3. Telomere length comparison between 552 people with ALS with a *C9orf72* repeat expansion and 907 people with ALS with normal *C9orf72* repeat length using a multivariable linear regression.

Figure 1.TIFF

Figure 1.

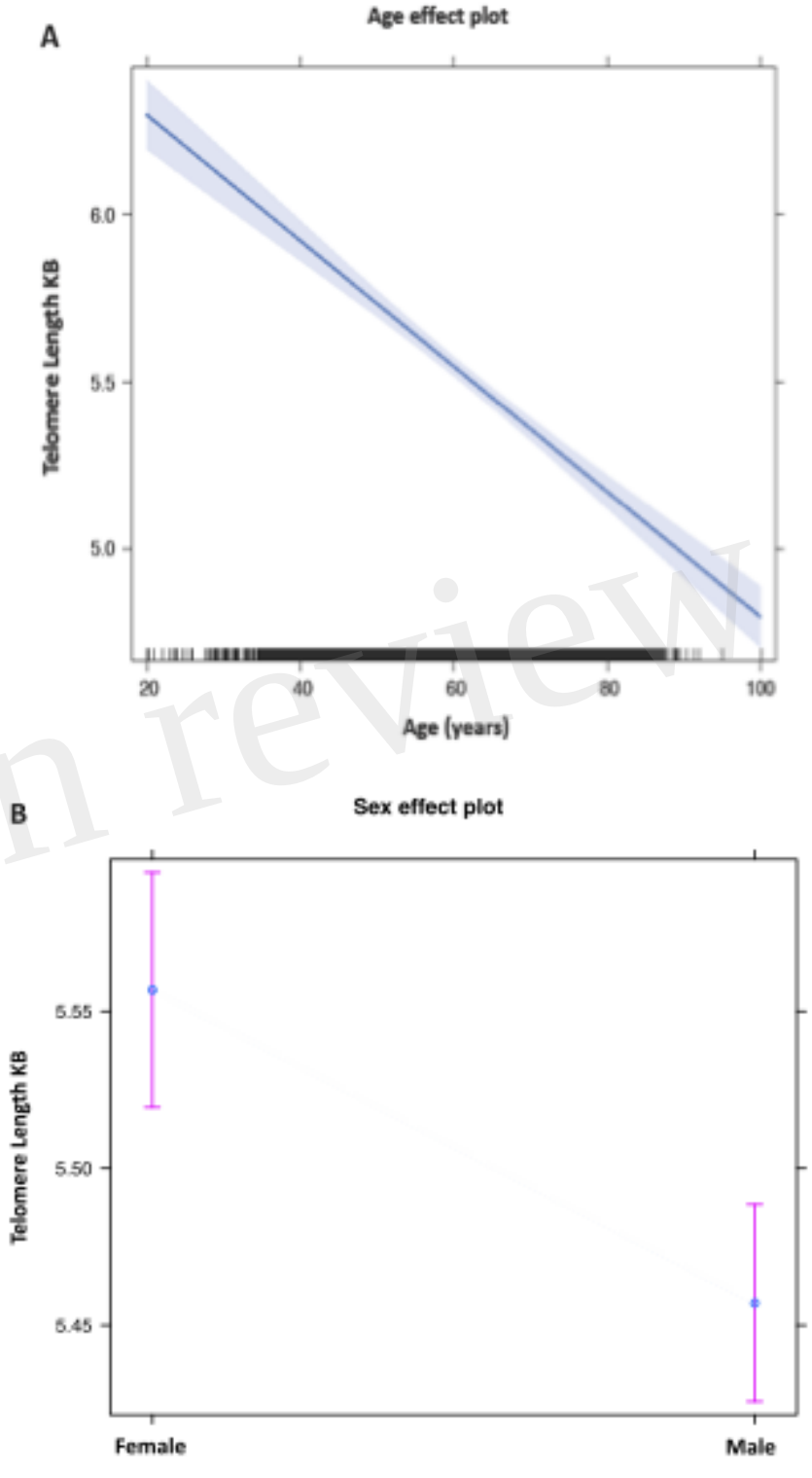

Figure 2.TIFF

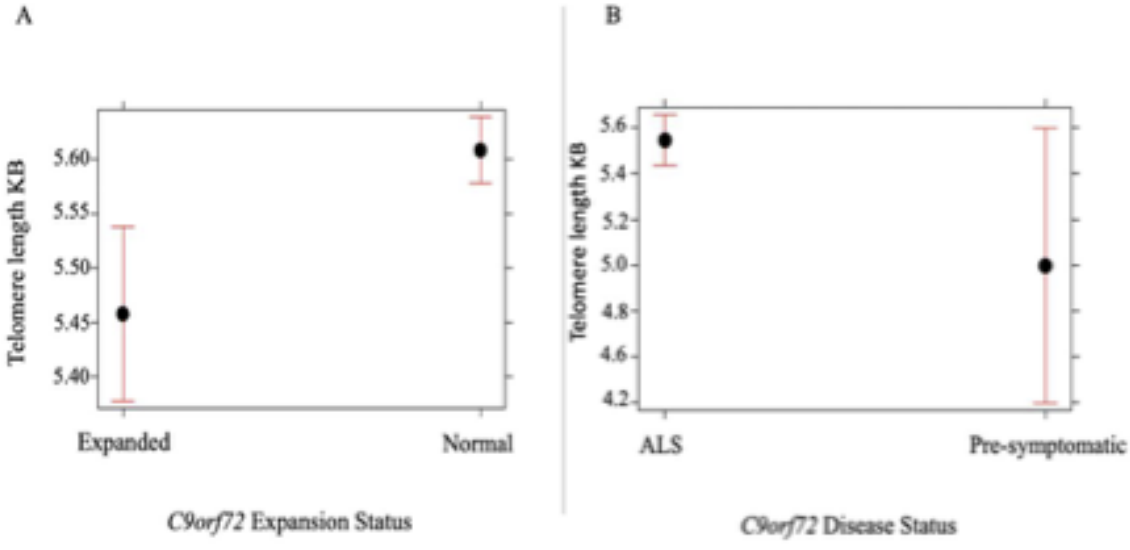

Figure 2. **A-**Telomere length comparison between 334 people with ALS carrying an expanded *C9orf72* repeat against people with ALS not carrying an expansion. Multivariable linear regression shows that the telomere is shorter in those carrying a *C9orf72* expansion compared with age and sex matched disease controls not carrying an expansion ($p = 5.0 \times 10^{-4}$). **B-** Telomere length comparison between 334 people with ALS with *C9orf72* repeat expansion against 10 healthy individuals with confirmed expanded *C9orf72* status. Multivariable linear regression shows that the telomere is longer in people with ALS with *C9orf72* repeat expansion ($p = 0.05$). Telomere length reported in kilobases (kb). Red bars indicate 95% confidence intervals.