Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia

Oriol Dols-Icardo1,2, Alberto García-Redondo3,4, Ricard Rojas-García1,2, Raquel Sánchez-Valle5, Aina Noguera6, Estrella Gómez-Tortosa7, Pau Pastor2,8,9, Isabel Hernández10, Jesús Esteban-Pérez3,4, Marc Suárez-Calvet1,2, Sofía Antón-Aguirre1,2, Guillermo Amer11, Sara Ortega-Cubero2,8,9, Rafael Blesa1,2, Juan Fortea1,2, Daniel Alcolea1,2, Aura Capdevila1, Anna Antonell5, Albert Lladó5, José Luis Muñoz-Blanco12, Jesús S. Mora13, Lucía Galán-Dávila14, Francisco Javier Rodríguez De Rivera15, Alberto Lleo1,2 and Jordi Clarimon1,2,*

1Memory Unit and Neuromuscular Diseases Unit, Neurology Department and Sant Pau Biomedical Research Institute, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, Spain 2Center for Networking Biomedical Research in Neurodegenerative Diseases (CIBERNED), Madrid, Spain 3Department of Neurology, ALS Unit, Instituto de Investigación Biomédica Hospital 12 de Octubre, Madrid, Spain 4Centre for Biomedical Network Research on Rare Diseases (CIBERER), Valencia, Spain 5Department of Neurology, Alzheimer’s Disease and Other Cognitive Disorders Unit, Institute of Neurosciences, IDIBAPS, Hospital Clínic, Barcelona, Spain 6Department of Clinical Analysis, Hospital Universitari Son Espases, Mallorca, Spain 7Department of Neurology, Fundación Jiménez-Díaz, Madrid, Spain 8Neurogenetics Laboratory, Division of Neurosciences, Center for Applied Medical Research, University of Navarra, Pamplona, Spain 9Department of Neurology, Clínica Universidad de Navarra, University of Navarra School of Medicine, Pamplona, Spain 10Memory Clinic of Fundació ACE, Institut Català de Neurociències Aplicades, Barcelona, Spain 11Department of Neurology, Hospital Universitari Son Espases, Mallorca, Spain 12Department of Neurology, ALS Unit, Hospital Universitario Gregorio Marañón, Madrid, Spain 13Department of Neurology, ALS Unit, Hospital Carlos III, Madrid, Spain 14Department of Neurology, ALS Unit, Hospital Clínico Universitario ‘San Carlos’, Madrid, Spain 15Department of Neurology, ALS Unit, Hospital Universitario La Paz, Madrid, Spain

Received August 27, 2013; Revised and Accepted September 16, 2013

Hexanucleotide repeat expansions within the C9orf72 gene are the most important genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). The difficulty of developing a precise method to determine the expansion size has hampered the study of possible correlations between the hexanucleotide repeat number and clinical phenotype. Here we characterize, through a new non-radioactive Southern blot protocol, the expansion size range in a series of 38 ALS and 22 FTD heterozygous carriers of >30 copies of the repeat. Maximum, median and modal hexanucleotide repeat number were higher in ALS patients than in FTD patients (P < 0.05 in all comparisons). A higher median number of repeats correlated with a bigger range of repeat sizes (Spearman’s ρ = 0.743, P = 1.05 × 10⁻¹¹). We did not find any correlation between age of onset or disease duration with the repeat size in neither ALS nor FTD mutation carriers. Clinical presentation (bulbar or spinal) in ALS patients did not correlate either with the repeat length. We finally analyzed two families with affected and unaffected repeat expansion carriers, compared the size of the repeat expansion between two monozygotic (MZ) twins (one affected of ALS and the other unaffected), and examined the expansion size in two different tissues (cerebellum and peripheral blood) belonging to the same FTD patient. The results suggested that the length

*To whom correspondence should be addressed at: Genetics of Neurodegenerative Disorders Unit, IIB-Sant Pau, Neurology Department, Hospital de la Santa Creu i Sant Pau. Sant Antoni M. Claret 167, 08025 Barcelona, Spain. Tel: +34 932919050 x8240; Fax: +34 935565602; Email: jclarimon@santpau.cat

© The Author 2013. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
INTRODUCTION

A non-coding (GGGGCC) hexanucleotide repeat expansion within the first intron of the chromosome 9 open reading frame 72 (C9orf72) gene (GenBank reference NM_001256054.1; MIM# 614260) has been identified as the most frequent cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) (1–3). Typically, a standard repeat-primed PCR (rpPCR) method is used to identify expansion carriers, but this technique cannot size genomic DNA expansions beyond 30 hexanucleotide repeats and therefore is ineffective to estimate larger repeat lengths of this dynamic mutation (1,2). Recently, the expansion size was characterized by Southern blotting through a non-radioactive approach using a hybridizing probe composed of five hexanucleotide repeats (4). This genotyping approach allowed the detection of expansions >275 repeats and showed that DNA isolated from lymphoblast cell lines is not reliable for sizing the C9orf72 expansion. Comparisons of the repeat length across six disease cohorts from UK, including 11 peripheral blood (PB) specimens from FTD patients and 17 PB samples from ALS patients, did not reveal any significant differences (4). A Southern blot protocol using a radioactive probe has also been recently published, but the study did not include any phenotype correlation analysis (5).

In the present study, we show an optimized protocol for Southern blot hybridization using a 954-bp non-radioactive probe that is capable to accurately size the whole range of the C9orf72 hexanucleotide repeat expansion. We also characterized pathogenic expansions through Southern blotting in a series of ALS and FTD patients, and assessed for possible correlations between the repeat expansion length and clinical features of these two neurodegenerative disorders.

RESULTS

Table 1 shows the demographic and clinical characteristics of the analyzed index patients. No differences in the proportion of gender or in the age of onset were found between ALS and FTD patients. In order to allay concerns that our Southern blot protocol might identify nonspecific bands, we first assessed eight DNA samples that did not carry the expansion mutation (Supplementary Material, Fig. S1). The resulting blot clearly demonstrated that our protocol did not detect any unexpected signals above the number of repeats determined by rpPCR genotyping. As presented in the Supplementary Material, Fig. S1, the protocol also showed a high sensitivity for the discrimination of low copy, non-pathogenic number of repeats (see control samples in Lanes VII and VIII carrying 7/25 and 2/19 repeats, respectively). All cases with a pathological repeat expansion in one allele resulted in higher molecular weight signals on the membrane, thus demonstrating the high specificity of the Southern blot protocol (Fig. 1). This pattern was seen in genomic DNA extracted from PB (N = 64) and cerebellum (N = 5).

Figure 2 shows the estimation of hexanucleotide repeat number and size range for all expansion mutation carriers. For statistical analyses, we considered four different metrics of the repeat length that included minimum, maximum, median and modal number of repeats for those patients in which a dense smear fragment appeared beyond the average allele length. None of these metrics correlated with gender or family history of disease in either the entire cohort or the two clinical phenotypes, independently. Likewise, the four expansion size measures did not correlate with age of clinical onset or disease duration (from onset to death) in FTD nor in ALS patients (data not shown). Of note, ALS patients had larger expansions than FTD patients, with significantly higher maximum, median and modal number of hexanucleotide repeats (P = 0.012, 0.02 and 0.01, respectively; Table 1). None of the four estimates were different between ALS cases with bulbar or spinal onset (i.e. median, 1693 ± 755 versus 1615 ± 530, respectively, P = 0.377). The median hexanucleotide repeat number significantly correlated with the range of repeat sizes (Spearman’s ρ = 0.743, P = 1.05 × 10−11; Supplementary Material, Fig. S2), thus indicating that larger expansions presented greater hexanucleotide repeat ranges.

The expansion size was then determined in a family in which the mutation segregated with FTD (Supplementary Material, Fig. S3A). The three affected siblings had a median repeat expansion size of 143, and ranged from 128 to 154. However, expansion mutation carriers from the second generation, composed of two asymptomatic first cousins, carried a median of 120 and 1401 repeats. Besides, the repeat expansion was also assessed in an ALS patient, who started the disease at the age of 57 years old, and his three unaffected sons (two males and one women) who were also mutation carriers (Supplementary Material, Fig. S3B). In this case, the father presented a median of 1898 hexanucleotide repeats, which was slightly larger to the number presented by the three siblings (1516, 1451 and 1127 repeats).

We also compared the repeat expansion size between a 50-year-old index ALS patient with symptoms beginning at the age of 47, and his unaffected MZ twin. All hexanucleotide repeats metrics were higher in the affected ALS patient compared with his sibling (Supplementary Material, Table S1).
Figure 1. Representative Southern blots of small and large repeat expansions. Nine cases carrying repeat expansion (Lanes 2–10) and a control DNA without an expansion (first lane) are shown. Lane 7 (corresponding to case 46 in Fig. 2) shows an example of median and mode determination, represented with a triangle and a round, respectively, within the smear range.

Figure 2. Representation of Southern blot data for ALS ($n = 38$) and FTD ($n = 22$) patients. Individual blot data are shown with approximated hexanucleotide repeat number range (minimum to maximum), median represented with a triangle and modal point as a black dot. DNA samples extracted from PB tissue and cerebellum tissue are represented with a black or gray line, respectively. Cases 13 and 14 represent repeat expansion size parameters of PB and cerebellum tissue from the same individual, respectively (Case 14 is not included in the statistical analysis). Note that Cases 1 and 2 are FTD patients belonging to the FTD family (individuals II and III from the FTD family pedigree represented in Supplementary Material, Fig. S3A) and are not included for statistical purposes.
Finally, the hexanucleotide repeat number was determined in DNA samples from cerebellum and PB tissue belonging to the same FTD patient. Expansion size estimates showed a moderately higher number of hexanucleotide repeats in the cerebellum compared with PB tissue (Supplementary Material, Table S1).

DISCUSSION

We have developed a non-radioactive Southern blotting protocol using a 954-bp probe to characterize the hexanucleotide repeat expansion size range in a series of ALS and FTD patients (Supplementary Material, Fig. S4). Our method allows the detection of hexanucleotide repeats in C9orf72 ranging from 2 to ~4500 repeat units, thus covering the whole expansion size range and enabling the identification of both small and large pathogenic repeat expansions. Our results suggest that ALS patients harbour a higher number of hexanucleotide repeats than FTD patients. This is not in accordance with a recent report in which authors did not find repeat size differences between different neurodegenerative diseases, including ALS and FTD (4). One possibility for this discrepancy is that their methodology, based on a hybridization probe composed of five DIG-labelled hexanucleotide repeats, was not sensitive to detect expansions at the low edge of the mutation spectrum (from 30 to ~275 repeats), and therefore DNA samples with low number of repeats could have been missed. This outcome clearly emphasizes the need of a genotyping technique with enough sensitivity to cover the whole pathogenic expansion size range. Another explanation could be that we have analyzed a greater number of ALS and FTD patients, thus increasing our power to detect differences between both neurodegenerative diseases.

A possible mechanistic explanation for the difference in the repeat length between ALS and FTD in our study might be related to the recently described repeat-associated non-ATG-initiated (RAN) translation that occurs in C9orf72 expansion mutation carriers. This alteration in the translation process leads to the accumulation of insoluble dipeptide-repeat (DPR) protein aggregates in the central nervous system (6,7). Interestingly, the formation of incorrectly translated products seems to occur in a hexanucleotide repeat length-dependent manner. That is, the longer the repeat size the greater quantity of aberrantly translated dipeptides will be produced. Furthermore, longer repeat tracts can express multiple RAN-translated homopolymeric proteins whereas shorter repeats only express a unique DPR (7). Taking all these novel data into account, it could be expected that the length of the repeat could determine, to some extent, the molecular mechanisms driving the disease to either FTD or ALS. Nevertheless, our results should be interpreted cautiously as there is a substantial overlap of hexanucleotide repeat sizes between ALS and FTD index patients, and a threshold to differentiate between ALS and FTD cannot be defined based on our data. For example, index cases number 3, 4 and 26 (Fig. 2), with estimated median allele sizes of 148, 130 and 192 repeats, suffered from either FTD (Patients 3 and 4) or ALS (Patient 26).

Our results also suggest that the hexanucleotide repeat number does not correlate with disease duration or age at onset. This lack of correlation might be due to environmental, genetic or epigenetic factors. In this sense, a recent study has reported that higher levels of DNA methylation near the hexanucleotide repeat significantly correlated with shorter disease duration in ALS patients (8).

We found a striking positive correlation between the median number of hexanucleotide repeats and the range of the observed repeat sizes. This might reflect a greater disruption of replication, repair and/or recombination machineries as the number of repeats increases, thus contributing to genetic instability (9). We also show a family in which the expansion mutation segregates with FTD and present evidence of an unstable hexanucleotide repeat number transference through generations. On the contrary, we also present a family with an ALS affected father and his three healthy sons who carry similar, although slightly lower, expansion sizes. Similar expansion sizes (ranging from 1000 to 1600 repeats) have been previously reported across four family members suffering from FTD (10). These data could explain the clinical heterogeneity between family members that has been described in related individuals carrying this dynamic mutation (11), and clearly indicates the impossibility to predict the expansion size carried by descendants of mutation carriers.

Finally, the fact that we have found subtle discordances in the repeat length between two tissues (PB and cerebellum) from the same patient, and also between MZ twins discordant for the disease, strongly suggests that stochastic expansion events during cell division, which results in somatic and/or germline mosaicism, contributes to the intra- and inter-individual repeat expansion variation.

In conclusion, we have developed a reproducible and optimized protocol of Southern blot hybridization that allows a reliable method to approximate the whole repeat expansion size range in C9orf72 in DNA samples extracted from PB and brain tissue. Our results suggest that larger expansions occur in ALS compared with index FTD patients. We also demonstrate that inter-individual differences in repeat lengths might exist not only between unrelated patients but also between family members, including MZ twins, and different tissues from the same individual. These results could explain the high heterogeneity of the phenotype presented by patients carrying the C9orf72 expansion.

MATERIALS AND METHODS

Samples

DNA samples extracted from PB of 38 index ALS patients and 18 index FTD patients were included in the analysis. Additionally, DNA samples extracted from cerebellum tissue of four patients with pathologically confirmed frontotemporal lobar degeneration were collected. PB-derived DNA samples of two families, one consisting of three FTD affected siblings and two non-affected first cousins, and the other composed of one ALS patient and his four unaffected sons, were also recruited. DNA from the cerebellum of an index FTD patient was also analyzed in order to compare the expansion size with its corresponding PB tissue. Finally, we included DNA from an unaffected monozygotic (MZ) twin of an index ALS patient in order to compare the number of hexanucleotide repeats between identical twins. ALS patients were diagnosed with definite or probable ALS, as defined by the El Escorial research criteria (12). FTD
diagnoses were made according to the international consensus criteria (13,14). All patients were previously identified as heterozygous carriers of a C9orf72 hexanucleotide expansion >30 repeats through an nPCR method (10,15–18). Written informed consent was obtained from all participants and research ethics committees from the respective participating centres approved the study.

**Southern blotting**

A total of 15 μg of gDNA was digested overnight with XhoaI restriction enzyme (New England Biolabs, Ipswich, MA, USA). Electrophoresis of digested DNA samples was carried out in 0.8% agarose gel with 1× tris–borate–EDTA buffer for 22 h at 1.4 V/cm. DNA was transferred to a positively charged nylon membrane (Roche Applied Science) by capillary blotting and was baked at 80°C for 2 h. A 954-bp probe was synthesized through a nested PCR method with FastStart PCR Master Mix (Roche Applied Science) using two PCR rounds and four oligonucleotide primers. The first PCR round was performed with the primers CAACCTGGTAGTATGGTGTAG and CTGAG TTCCAGACTTTGCTACAG. The 1350 bp PCR product was then purified and used as a template for a second PCR using the FastStart PCR Master Mix (Roche Applied Science) and the oligonucleotides CAGAAGTGAGTGCCCGGAGGC and CAGCGAGTACTCTGAGAGCAAG. Digoxigenin (DIG)-dUTP labelling was performed with 2.8 μl of the Vial 2, contained in the DIG probe synthesis kit (Roche Applied Science). Prehybridization was carried out at 48°C for 3 h and hybridization at 45°C overnight. Blots were then washed in 2× standard sodium citrate (SSC) and 0.1% sodium dodecyl sulphate (SDS) at room temperature for 10 min twice. High stringency washes (0.1× SSC and 0.1% SDS) were carried out at 68°C for 10 minutes. Final washes (0.1× SSC) used as a ladder.

**Assessment of the number of repeats**

Hexanucleotide repeat size was estimated by interpolation with a plot of log 10 bp number against migration distance. Minimum, maximum, median and modal sizes of the repeat expansion were assessed and used for statistical analyses. DIG-labelled DNA molecular weight marker III (Roche Applied Science) was used as a ladder.

**Statistical analyses**

Chi-square analyses were performed for categorical data. Mann–Whitney test was used to compare disease phenotypes and clinical outcomes with minimum, maximum, median and modal size of the repeat expansion. Spearman’s correlations were performed between the four expansion size parameters and the clinical features of the disease. Data were analyzed with the Statistical Package for Social Science v19.0 (SPSS Inc., Chicago, IL, USA).

**SUPPLEMENTARY MATERIAL**

Supplementary Material is available at HMG online.

**ACKNOWLEDGEMENTS**

We thank patients, their families and controls for their participation in the study. The authors thank the Neurological Tissue Bank of the Biobanc-Hospital Clinic-IDIBAPS (Barcelona) for providing human brain samples, and specially Dr. Ellen Gelpi for her valuable help.

Conflict of Interest statement. None declared.

**FUNDING**

This work was supported by grants from the Spanish Ministry of Economy and Competitiveness [grants number P11/01311 and P11/00092], CIBERNED and by the Foundation for Applied Medical Research (FIMA) to S.O.-C. and P.P.

**REFERENCES**


