Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Origin of the Spinocerebellar Ataxia Type 7 (Sca7) Mutations in the Spanish Population: Evidence for a Founder Effect

Aroa López-Alcojor
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Fiona Blanco-Kelly
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Fermina López-Grondona
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
María José Trujillo-Tiebas
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Maria Pilar Martin-Gutierrez
Hospital Universitario Fundación Jiménez Díaz
Yolanda Benítez
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Ana María Álvarez
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Beatriz Isidoro
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Olga Zurita
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Borja González
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Jesús Gallego
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Pablo Mínguez
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Marta Cortón
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Carmen Ayuso
Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid
Almudena Ávila-Fernández
AAvila@quironsalud.es

Instituto de Investigación Sanitaria-Fundación Jiménez Díaz, Universidad Autónoma de Madrid

Research Article

Keywords: Spinocerebellar ataxia, cone-rod dystrophies, founder effect, haplotypes, triplet repeat expansion

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10969308/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Spinocerebellar ataxia type 7 is a rare, progressive, autosomal dominant neurodegenerative disorder caused by a pathogenic CAG repeat expansion in exon 3 of the *ATXN7* gene, characterized by cerebellar ataxia and retinal degeneration. Its prevalence varies markedly across geographical regions, which suggests potential founder effects.

Objective

to investigate the origin of pathogenic spinocerebellar ataxia type 7 expansions in Spain by a comprehensive clinical and genetic overview of a cohort of 12 Spanish families.

Methods

the cohort was clinically described and the patients were selected and molecular genotyped using a dual approach involving direct triplet repeat-primed PCR analysis or bioinformatics screening of Next Generation Sequencing data. Haplotypes were constructed using four microsatellite markers and one intragenic SNP. Statistical analyses were conducted using R software.

Results

reveal high phenotypic heterogeneity of spinocerebellar ataxia type 7, with and age of onset and expansion sizes (35 to > 100 repeats) varying widely. Haplotype analysis identified a common allele among eight individuals from five unrelated families in the same geographical area, as well as one patient from a different region.

Conclusion

these findings strongly support the existence of a founder effect for spinocerebellar ataxia type 7 within the Spanish population.

INTRODUCTION

Spinocerebellar ataxia type 7 (SCA7; OMIM: 164500) is a progressive autosomal dominant neurodegenerative disorder characterised by the presentation of cerebellar ataxia (resulting from degeneration of the cerebellum, brainstem, and spinal cord) and cone or cone-rod dystrophy (CRD)[1,2]. Clinical diagnosis of SCA7 relies on standardised neurological assessment scales[3] and additional ophthalmological evaluation, including visual acuity testing, multimodal imaging and visual field testing.

This disease is caused by a pathogenic CAG repeat expansion in exon 3 of the *ATXN7* gene (chr.3p14.1; NM_000333), which encodes the protein ataxin-7[4]. In the general population, the most common allele contains 10 CAG repeats[5]. Alleles containing a CAG triplet expansion can be classified based on stability and penetrance: normal (< 27 repeats), unstable intermediate (27–33), incomplete penetrance (34–36), and fully pathogenic (> 36)[6]. Carriers of intermediate alleles are prone to meiotic instability, increasing the risk for expansion into pathogenic ranges, which could lead to SCA7 in their offspring[7].

Ataxin-7 is a nuclear transcription factor and core component of two transcriptional coactivator complexes: the STAGA complex and the USP22 complex, involved in histone acetylation and deubiquitination, respectively[8]. The expansion of a polyglutamine (polyQ) tract within ataxin-7 disrupts its subcellular distribution between the nucleus and the cytoplasm[1]. Accumulation of mutant ataxin-7 in the nucleus has been associated with cytoplasmic microtubule destabilisation, leading to progressive neurodegeneration[9]. In addition, studies in SCA7 mouse models and in vitro showed that mutant ataxin-7 selectively disrupts transcription mediated by the cone-rod homeobox protein (CRX), resulting in cone-rod dystrophy[10].

In individuals affected by SCA7, phenotypic severity is strongly correlated with both the age at onset and the length of the expansion. Specifically, there is an inverse correlation between repeat length and age at onset: longer expansions are associated with an earlier disease debut, accelerated neurological deterioration and more severe ophthalmological involvement[11,12]. In contrast, adult-onset cases typically involve smaller expansion sizes and a slower progressive neurological course. In these patients, visual impairment characterised by loss of central vision, impaired colour vision and reduced electroretinogram (ERG) responses, may occur before, alongside or after the onset of ataxia[13].

A hallmark feature of SCA7 is genetic anticipation, a phenomenon where the disease manifests earlier and with greater severity in successive generations. This anticipation is significantly more pronounced in paternal transmissions compared to maternal ones[14]. The molecular basis for this bias resides in the instability of trinucleotide expansions during male spermatogenesis, which frequently leads to repeats elongation, thus more prominent symptoms when the condition is inherited from the father[15]. This sex-specific difference is attributed to the biological differences in gametogenesis[15,16].

The prevalence of SCA7 exhibits significant inter-population variability with geographical enrichments, such as those in the Mexican state of Veracruz[11], which underscores a potential critical role of founder effects in the regional prevalence. This variation suggests that recurrent de novo expansions occurring on at-risk genetic backgrounds, as those derived from a founder allele in the intermediate range, may explain the shared haplotype observed among families within the same geographical area[17]. Nevertheless, the SCA7 genetic landscape can be heterogeneous; as evidenced by Stevanin et al. (1999) in the Italian population, a single geographical region may harbour multiple pools of at-risk chromosomes carrying SCA7 intermediate alleles, resulting in a diversity of haplotypes[18]. Therefore, while founder effects are prominent, multiple mutational events on different ancestral backgrounds also contribute to the regional prevalence of SCA7. Consequently, investigating the haplotype architecture of different cohorts is essential to determine whether

the local prevalence of SCA7 stems from a singular ancestral event or from multiple independent mutational origins.

The present study aims to provide a comprehensive clinical and molecular characterization of a cohort of Spanish families affected by SCA7, evaluating expansion size distribution, genotype-phenotype correlations, and patterns of familial transmission, as well as to investigate the ancestral origin of the pathogenic SCA7 alleles within this population through haplotype analysis, in order to determine whether a common founder effect underlies the prevalence of the SCA7 in Spain.

PATIENTS AND METHODS

Cohort description and sample processing

A total of 19 patients from 12 unrelated families of Spanish origin, with a prior diagnosis of SCA7 or family history of the disease, were recruited between 1990 and 2026. Participants were identified from the inherited retinal diseases (IRD) cohort (IRD, n = 1991 families), including patients with a clinical diagnose of macular dystrophy, cone and cone-rod dystrophy, and the spinocerebellar ataxia (SCA, n = 1917 families) cohort at Research Institute Fundación Jiménez Díaz University Hospital (FJD) in Madrid, Spain.

Clinical data, including age of onset, ophthalmological and neurological symptoms, and family history, were retrospectively ascertained from medical records. For those patients not followed up at our site (n = 7), medical data were obtained through review of outside hospital records. Although previous genetic results were available, CAG expansion sizes were re-evaluated using high-resolution molecular techniques to ensure data consistency.

For molecular analysis, peripheral blood samples were collected, and genomic DNA (gDNA) was isolated using the EZ1 Advanced XL extraction system (QIAGEN, Aarhus, Denmark). The integrity of the extracted gDNA was assessed by agarose gel electrophoresis, while purity and concentration were quantified using the DS11-FX system (DeNovix Inc., Wilmington, NC, USA) and Qubit ds-DNA assays (Thermo Fisher Scientific, Waltham, MA, USA), respectively.

Molecular genotyping and CAG expansion analysis

SCA7 cohort

ATXN7 genotyping of the 19 patients was performed using a dual approach involving direct triplet repeat-primed PCR (TP-PCR) analysis or bioinformatics screening of available Next Generation Sequencing (NGS) data: clinical and whole exome sequencing (CES and WES).

TP-PCR was used as first-tier diagnostic approach in 19 SCA7 patients with clinical suspicion of SCA7 based on the presence of neurological symptoms and retinal involvement or familiar background of the disease, and also to re-evaluate all SCA7 patients diagnosed prior to this study.

TP-PCR reactions were performed with the SCA7 primers pairs (forward tagged with 6-FAM) previously used by Dorschner[19], and QIAGEN Multiplex PCR Plus Kit as previously described by the manufacturer (QIAGEN; Venlo, Netherlands). Fluorescently labelled amplicons were resolved by automated capillary electrophoresis on a 3730xl Genetic Analyzer and analyzed using GeneMapper v.5.0 software (Thermo Fisher Scientific, Waltham, USA). The number of CAG repeats was calculated using the European Molecular Quality Network (EMQN)-validated formula that corrects fragment size (in bp) obtained from the electropherogram based on the 3730xl Genetic Analyzer:

$$(CAG)\,n = \frac{bp - 126}{3} - 7$$

Alleles were categorized as described by Sequeiros, Seneca and Martindale[20].

Uncharacterized cases from IRD cohort

Additionally, pathogenic expansions in the *ATXN7* gene were retrospectively analysed in the available NGS data (CES (Sophia Genetics, Boston, MA, USA)) from 417 unsolved cases from the IRD cohort, which includes cases from macular diseases, cone dystrophies and cone rod dystrophies. CES panel targets a total of 4828 genes and regulatory regions. *ExpansionHunter* software[21] was used to analyse this region. Variants were filtered and prioritised in a search for pathogenic alleles (> 36 repeats).

Haplotype Analysis

To investigate SCA7 founder effects, haplotyping was performed using four microsatellite markers (D3S1600, D3S1287, D3S3635, and D3S3644) located on the short arm of chromosome 3, and one intragenic SNP (rs3774729)[18].

The STR markers were located within 2Mb of the *ATXN7* gene: the telomeric marker D3S1600 and three centromeric markers (D3S1287, D3S3635, and D3S3644) as shown in Supplementary Fig. 1 and Supplementary Table 1. Primers pairs used to microsatellite amplification were obtained from GeneLoc (Weizmann Institute; https://geneloc.weizmann.ac.il/index.shtml) and specificity and quality were assessed in silico using the UCSC Genome Browser PCR tool (https://genome.ucsc.edu/cgi-bin/hgPcr). Microsatellite analysis was performed using PCR conditions optimized for each marker using a touchdown-style protocol using a reverse and a forward one, labelled at the 5’ end with the 6-FAM fluorochrome. Initial denaturation (95°C, 12 min) was followed by two cycling phases: 10 cycles with a 95°C denaturation, and 20 cycles with a reduced denaturation temperature of 89°C to increase specificity (detailed temperatures in Supplementary Table 2). Fluorescently labelled amplicons, also with the 6-FAM fluorochrome, were resolved by automated capillary electrophoresis on a 3730xl Genetic Analyzer and analyzed using GeneMapper v.5.0 software.

SNP genotyping was performed by direct Sanger sequencing using specific primers designed for the intragenic G > A change (rs3774729) using the NCBI Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Following PCR amplification, amplicons were purified using the ExoProStar™ PCR and Sequence Reaction Clean-Up Kit (Cytiva; Thermo Fisher Scientific) and

sequenced on both strands using the Big Dye™ Terminator Cycle v1.1 Sequencing Kit (Thermo Fisher Scientific). Sequencing products were resolved on a 3130xl Genetic Analyzer.

Statistical analysis

The relationship between CAG repeat expansion length and the age of onset for both ataxia and presence of ophthalmological symptoms (CRD on the graph) were evaluated using Spearman's rank correlation coefficient. All statistical analyses and data visualization were performed using R software[22].

RESULTS

Description of the cohort

A total of 19 patients (11 males and 8 females) from 12 unrelated families were genetically and clinically characterized with SCA7. Within the IRD cohort, the prevalence of SCA7 was 0.60% (12/1991 families) whereas for syndromic IRD it was 1.5% (12/794 families). In the SCA cohort, the prevalence was 0.62% (12/1917 families). Notably, SCA7 accounted for 13% of the 92 families characterized with ataxia at our hospital (12/92 families). All of patients from IRD cohort who were suspected of SCA7 due to their familiar background or the dual presence of ophthalmological symptoms and ataxia and were finally genetically characterized using a direct *ATXN7* genotyping by TP-PCR. Retrospective bioinformatics screening of 417 unsolved IRD cases yielded no additional pathogenic or incomplete penetrance expansions. Detailed clinical manifestations and family pedigrees are summarized in Table 1 and Fig. 1, respectively.

Table 1
Clinical and genetic data of the Spanish SCA7 cohort.

Family	Individual	Sex	YOB	CAGn SCA7	Inheritance		Origin	Ataxia	AO	CRD	AO
F1	II:7	M	1943	42 (P)		aM	Guadalajara	Yes	47	Yes	39
	III:10	M	1972	44 (P)	2	aM		Yes	25	Yes	NA
F2	III:1	M	1956	44 (P)		aM		Yes	15	Yes	40
F3	II:4	F	1942	43 (P)		NA		Yes	56	Yes	56
	III:4	M	1978	43 (P)	0	pM		Yes	16	Yes	19
F4	II:1	M	NA	43 (P)		aP		Yes	NA	NA	NA
F5	III:2	F	NA	40 (P)		aM		Yes	NA	NA	NA
	III:3	M	1956	43 (P)		aM		Yes	28	Yes	28
F6	III:2	F	1979	44 (P)		aM		Yes	26	Yes	37
F7	I:1	F	1959	35 (IP)		NA		Yes	66	No	
	II:1	M	1991	35 (IP)	0	pM		No		No	
	III:1	F	2024	> 100 (P)	65	pP		Yes	1	Yes	1
F8	III:5	F	1962	48 (P)		aP	Madrid	Yes	33	Yes	33
F9	II:1	M	1942	45 (P)		NA		Yes	NA	NA	NA
F10	I:1	F	NA	53 (P)		NA		Yes	NA	NA	NA
	II:1	M	NA	64 (P)	11	pM		Yes	NA	NA	NA
F11	I:1	M	1985	46 (P)		NA		Yes	NA	NA	NA
	II:1	M	2012	60 (P)	14	pP		Yes	NA	NA	NA

Family	Individual	Sex	YOB	CAGn SCA7	Inheritance	Origin	Ataxia	AO	CRD	AO
F12	II:1	F	1905	50 (P)	NA	Burgos	Yes	32	NA	NA

Families are grouped by geographical origin. YOB: year of birth; CAGn: number of CAG repeats, Allele Class: P (pathogenic, > 36), IP (incomplete penetrance, 34–36); Inheritance: pattern of transmission and intergenerational expansion size change in CAG repeats (p/aM/P: proven/alleged maternal/paternal), CRD: cone-rod dystrophy/ophthalmological symptoms, AO: Age of onset, NA: not available.

This cohort exhibited a significant geographical cluster as 58.3% (7/12) of the families (Families 1–7) originated from a small single village in Guadalajara, Spain. The remaining families were traced to Madrid, Spain (Families 8–11) and Burgos, Spain (Family 12). The precise ancestral background for Families 8–12 remains undetermined, limiting further inferences regarding their specific regional origins. Pedigrees and clinical data indicate high penetrance and vertical transmission across up to four generations, which is consistent with the autosomal dominant nature of the SCA7. Ten families showed a multi-generational transmission pattern with neurological symptoms consistent with SCA7 appearing in at least two successive generations (Fig. 1).

Genetic and clinical characteristics of the cohort

Molecular analysis confirmed the presence of pathogenic or incomplete penetrance CAG expansions in all families. The distribution of expansion sizes for both alleles is detailed in Supplementary Table 3. All affected individuals were found to be heterozygous, carrying one normal allele (most frequently 10 CAG repeats) and one expanded allele. Pathogenic alleles widely ranged from 35 to > 100 repeats (mean = 48.53 repeats; median = 44), reflecting the high molecular heterogeneity of the cohort. Specifically, two individuals carried an allele of 35 repeats, placing them in the incomplete penetrance range, while the rest of the cohort carried fully penetrant alleles (> 36 repeats).

Molecular findings revealed also significant intergenerational instability of SCA7 allele (Supplementary Fig. 3). Remarkably, genetic anticipation was observed in four families (F3, 7, 10 and 11), as intergenerational increases were observed in 66.6% (4/6) of the documented transmissions, ranging from + 2 to + 65 repeats (Table 1). The Family 7 illustrates the extreme instability of the *ATXN7* locus. A paternal transmission from an asymptomatic father (F7, II:1, 35 repeats) to his daughter (F7, III:1) resulted in a massive expansion of repeats (> 100), manifesting as a devastating severe infantile-onset phenotype at 1 year-old with both neurological and ophthalmological symptoms.

Clinical data from at least one visit were available for 11 of the 19 SCA7 patients (8 index cases and 3 affected relatives) (Table 1). The mean age at disease onset was 31.4 (± 18.37) years. While neurological symptoms were near-universal (18/19 patients) in the SCA7 cohort, ophthalmological ones were reported in just 47% (9/19 patients). Notably, case II:7 from Family 1 was the only patient to present ophthalmological symptoms before the onset of neurological manifestations. The ages at onset of the first neurological symptoms ranged from 1 to 66 years (data available for 11 patients), occurring before adulthood in only 3

patients with early-onset ataxia (3/11, 27%). For visual symptoms, the onset ages ranged from 1 to 56 years (data available for 8 patients), occurring before adulthood in only 2 patients with early-onset retinal dystrophy (2/8, 25%).

Family 1: a three-generation family comprised of six affected individuals, with medical records available for two cases (II:7 and III:10). The patient II:7 from the 2nd-generation, carried a 42 CAG repeat expansion. Clinical onset began with ophthalmological symptoms at 39 years of age, followed by the development of ataxic gait at age 47. Further examination revealed dysarthria, altered reflexes and cerebellar atrophy. His unaffected son (III:12) carried a normal allele of 10 CAG repeats. Patient III:10, from the 3rd generation, carries a larger expansion of 44 CAG repeats, and exhibited a significantly earlier onset at age 25 years, manifesting both retinal degeneration and ataxic gait. He has an affected sister with early-onset visual impairment at age 10 who was unavailable for genetic testing at our site (III:6).

Family 2: a three-generation family that illustrates the clinical variability associated with incomplete penetrance alleles. The index case III:1 carried a 44 CAG repeat expansion. He exhibited a very early neurological onset with ataxic gait at 15 years, followed by visual impairment at age 40 years. His mother (II:2) showed a much later clinical debut, with ataxic gait at 73 years, highlighting a significant phenotypic shift between generations.

Family 3: a two-generation family which shows genetic anticipation. The index cases II:4 is a female carrying a 43 CAG expansion repeats and developed both ophthalmological symptoms and ataxia at age 56 years, presenting ataxic gait, stiffness, and altered reflexes. Her son III:4 inherited the maternal allele of 43 repeats, which resulted in a dramatically earlier onset, with ataxic symptoms and cerebellar atrophy and ophthalmological symptoms at the age of 16 and 19 years, respectively.

Family 4: a suspected two-generation family in which genetic testing confirms SCA7 despite that clinical data were limited. The index case II:1 carried 43 CAG repeats and was referred with clinical diagnosis of ataxia. His aunt, brother and nephew (not followed up at our hospital) were reported to have lost their vision. His asymptomatic father (I:2) carried 29 CAG repeats.

Family 5: several individuals from this four-generation family exhibited both neurological and ophthalmological symptoms without a formal SCA7 diagnosis. The index case III:3 carried 43 CAG repeats, presumably inherited from his mother, presented with ataxia, altered reflexes, hypotonia, dysphagia, dysarthria, ophthalmoplegia and photosensitivity, at the age of 28 years. His sister III:2 carried 40 repeats and suffered from ataxia, but no further clinical data are available. The son of III:3 (IV:1) was reported to develop both visual impairment and ataxic symptoms at the age of 3 and 5 years, respectively, dying shortly after. Although DNA sampling was unavailable for this patient, this clinical history is highly suggestive of a massive intergenerational expansion.

Family 6: the index case of this three-generation family (III:2) carried 44 CAG repeats and developed ataxic gait at age 26 years and ophthalmological symptoms at 37 years. A maternal inheritance pattern was suggested, as her mother (II:2) presented with similar symptoms with a later onset (age 50 years), although no DNA sample was available for confirmation.

Family 7: this three-generation family is a very illustrative example of extreme meiotic instability during paternal transmission. The index case III:1, a one-year-old female, carried a massive expansion of > 100 CAG repeats. She presented with a severe phenotype, characterized by ophthalmological symptoms, nystagmus, hypotonia and early cerebellar atrophy. Her father II:1 is an asymptomatic carrier of 35 repeats, which he inherited from his mother I:1. Patient I:1 also carried 35 repeats but developed late-onset speech impairment at age 66. The expansion from 35 to > 100 repeats in a single generation underscores the high risk of paternal transmission in SCA7.

Family 8: the index case III:5 of this four-generation family carried a 48 CAG repeat expansion. She initially presented with visual impairment at the age of 33 years and had been diagnosed with macular dystrophy before the neurological symptoms (ataxic gait) in her family (father, aunt, and grandmother) pointed towards SCA7. She had a brother with ataxic symptoms, but further clinical data and genetic testing are unavailable (III:7). Her deceased affected son (IV:1) was reported to carry 58 repeats, further illustrating the tendency for elongation of the expansion in this lineage. No DNA sample was available to study the proband's son in our centre to confirm his genetic result.

Families 9, 10 & 11: these families were clinically diagnosed in other centres after presenting with SCA7 characteristic symptoms. Index case II:1 from Family 9 is an apparently sporadic case and carried 45 CAG repeats. Family 10 showed anticipation between the index case I:1 (53 repeats) and her son II:1 (64 repeats). Family 11 is a two-generation lineage with clear paternal instability, with the expansion increasing from 46 repeats (external data) in the father (I:1) to 60 repeats in the son (II:1).

Family 12: represents a second sporadic case with unknown family history of the disease (II:1). She developed neurological symptoms at 32 years and was clinically diagnosed with ataxia in a different geographical area from the rest of families in our cohort (Burgos, Spain). Genetic testing performed at our centre confirmed the molecular diagnosis associated with a 50 CAG repeat.

Haplotype analysis

To investigate a potential common ancestral origin for the *ATXN7* expansions in the Spanish cohort, haplotypes were reconstructed using four microsatellite markers (STRs) and the intragenic SNP (rs3774729) (Supplementary Fig. 1 and Supplementary Table 1). Analysis was performed on the expanded alleles of index patients and available relatives from the 12 families (Fig. 2). Accordingly, haplotypes could not be assembled for Families 1, and 6. Family 12 did not provide information to the analysis, as no common allele was observed with the other families.

Genotype-Phenotype correlation

Since different symptoms onset associated with *ATXN7* expansions were observed in our cohort, a genotype-phenotype analysis was performed with the aim of identifying potential correlations between CAG repeat length and the onset of neurological symptoms and visual loss. Therefore, the relationship between the CAG repeat expansion size and the age at symptoms onset was modelled using exponential fits in the affected cases with available data (n = 11 for ataxia and n = 8 for ophthalmological manifestations).

Statistical analysis using Spearman's rank correlation coefficients revealed distinct patterns for neurological and ophthalmological manifestations, as shown in Fig. 3. Spearman correlation coefficients were negative for both variables, indicating an inverse relationship between expansion size and age at onset. However, the negative correlation for ataxia onset (r= -0.56; Fig. 3a) is stronger than for visual impairment onset (r = -0.37; Fig. 3b), which is moderate, suggesting that ophthalmological onset may be influenced by additional modifiers.

DISCUSSION

Spinocerebellar ataxia type 7 (SCA7) is an autosomal dominant neurodegenerative disorder caused by a CAG repeat expansion exceeding 36 units in the *ATXN7* gene[8,23]. This condition is clinically distinguished from other SCAs by the hallmark presence of retinal degeneration alongside with progressive neurological symptoms[24]. In the present study, SCA7 accounted for 13% of molecularly characterized index cases (12 out of 92) at our hospital. This prevalence is similar to the previously value, ranging from 7.5 to 13.8%, reported by the EUROSCA register[25], and higher than the ones reported in Spain, as the 10% reported before[26], or in other populations, as the 7.4% identified in a Mexican cohort[17]. This enrichment is probably driven by a distinct geographical cluster in Guadalajara, Spain, where 58.3% of the families in our study originated. Such a high concentration of a rare neurodegenerative disorder in a specific region strongly suggests an underlying genetic link among these families.

SCA7 has been previously reported in families of diverse ethnic and geographical backgrounds, including Europe, Africa, the Middle East, Asia and America[27], which often points toward the existence of founder effects. Investigating these events through haplotype analyses allows for the identification of shared ancestral lineages. However, the reduction of genetic diversity in humans due to bottleneck effects remains challenging, as it typically requires a comparison with the original population prior to the event[28]. In this study, the identification of several apparently unrelated SCA7 families with ancestry tracing back to a small Spanish region led us to investigate a shared mutational origin.

The haplotype reconstruction provided strong molecular evidence of a founder effect in central Spain. Six families, representing 47.4% of patients in the cohort, shared a conserved core haplotype: 184 (D3S1600) – A (rs3774729) – 253 (D3S1287) – 102 (D3S3635) – 165 (D3S3644). The identification of this genetic signature in Family 9 directly links them to the Guadalajara ancestral pool, suggesting a shared lineage despite their current geographical dispersion in Madrid. Despite these findings, current data is insufficient to precisely estimate when this pathogenic allele first arose in Spain, as documented family histories only span two to three generations. It remains an intriguing question whether SCA7 expansion in this cluster results from a single ancestor, as this could provide further insight into the mechanisms underlying the mutation process.

Conversely, the genetic heterogeneity observed in the remaining four families analysed by haplotyping, which carry divergent haplotypes, suggested that SCA7 in Spain result of a predominant founder effect in the central region alongside multiple independent mutational events.

SCA7 is a heterogeneous illness in which the magnitude of the expansion dictates clinical severity and age of onset[29]. Every patient in our cohort harboured expanded alleles in the pathogenic range (> 36 repeats),

except for two individuals from one families (Family 7), who carried 35 repeats, corresponding to alleles with incomplete penetrance. Our patients exhibited a wide spectrum of symptoms, with ataxic gait being the presenting symptom in 7 cases and ophthalmological symptoms in 9 cases. Among all patients, 8 individuals reported altered reflexes and 4 showed cerebellar atrophy on MRI. As previously reported[11,16], significant intrafamilial variability in phenotypic expression was observed in our cohort. For instance, the Family 3 proband (III:4) displayed a broader variety of symptoms than his mother (II:4).

Consistent with the phenomenon of anticipation, earlier clinical presentation was associated with longer expansion and more severe phenotype[30]. Our comprehensive genotype-phenotype correlation analysis supported a negative correlation between CAG repeat length and age at onset. Earlier generations displayed fewer repeats and milder manifestations than subsequent ones, with subjects ranging between 35 and 64 repeats, with the exception of a deceased first-born child (Family 7, III:1) carried more than 100 repeats. This correlation was particularly evident in Family 3, where the son (III:4) developed symptoms at age of 16 with the same expansion size as his mother (onset at age 56). Both carried pathogenic alleles of 43 repeats. Even though our data is limited, it supports this negative correlation, given that onset and severity of these symptoms seem to be driven by the expansion size.

However, the moderate correlation for visual onset suggests that the ophthalmological debut may be influenced by additional factors. The differences between the correlation coefficients suggests that while expansion size is the primary driver for both phenotypes, the age of onset for CRD may be subject to a greater influence from additional modifiers compared to the neurological phenotype. For example, the proband in Family 2 (III:1), carrying a 44 CAG repeats allele, presented with severe ataxic symptoms as early as age 15, but ophthalmological symptoms did not appear until age 40. In contrast, the proband in Family 6 (III:2) and the nephew in Family 1 (III:10) had the same size expansion as well, but developed neurological symptoms at a later age. In Family 3, both the proband (II:4) and her son (III:3) carried a 43 CAG repeat expansion, but he presented with the disease years earlier. These findings might suggest that, in addition to meiotic instability occurring during parental transmission, somatic instability may also exist among different cell types[31], and alongside the expansion size, this could help to explain the differences in phenotype.

Our data underscores the extreme instability of the *ATXN7* locus, particularly during paternal transmission. In five families (4, 7, 8, 11 and 12), the expanded allele was inherited paternally. The most striking case was Family 7, where a paternal “jump” from an asymptomatic 35 repeat carrier state to a massive expansion of > 100 repeats resulted in an infantile phenotype (onset < 1 year) with altered reflexes, cerebellar atrophy, and visual impairment. These findings have critical implications for genetic counselling. First, incomplete penetrance alleles in the 35-repeat range can lead to vastly different outcomes, from lifelong asymptomatic status, late-onset mild ataxia to catastrophic expansions in the next generation. Furthermore, the transition from an intermediate or reduced-penetrance allele to a fully pathogenic range can occur in a single generation, as seen in Family 7.

SUMMARY

In summary, this study characterizes one of the largest SCA7 clusters in Europe, driven by a founder effect in Guadalajara. Additionally, expanding the screening in this region will be essential to accurately map the

historical spread of this ancestral mutation.

Declarations

Competing interests

The authors declare that they have no competing interests.

Ethics approval

The study protocol was approved by the FJD Research Ethics Committee (reference PIC017-24_FJD) and adhered to the tenets of the Declaration of Helsinki.

Informed consent

Written informed consent was obtained from all participants prior to their inclusion in the study.

Authors' contribution statement

A.L.-A., C.A., M.C. and A.A.-F. conceived and designed the study, acquired data, and contributed to the interpretation of the results. A.L.-A. also drafted the manuscript. F.B.-K., F.L.-G., M.P.M.-G. and M.J.T.-T. contributed to data collection and patient diagnosis and revised the manuscript. Y.B. and P.M. performed the bioinformatics analyses and revised the manuscript. A.M.A., B.I., O.Z., B.G. and J.G. contributed to the design and performance of the experiments. All authors approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Funding statement

This work was supported by Fundación Conchita Rábago de Jiménez Díaz, and by research grants awarded to Dr. Ayuso ("*Un estudio exhaustivo de las distrofias de retina hereditarias: mejora del diagnóstico molecular, enfoques terapéuticos, ensayos clínicos y medida del impacto en los pacientes*"; ISCIII, PI22/00321) and to Dr. Ávila-Fernández ("*Desarrollo de kits diagnósticos de enfermedades neurológicas de causas genéticas y producidas por expansión de repeticiones*"; State Research Agency and the Ministry of Science and Innovation, CPP2022-009538).

Author Contribution

A.L.-A., C.A., M.C. and A.A.-F. conceived and designed the study, acquired data, and contributed to the interpretation of the results. A.L.-A. also drafted the manuscript. F.B.-K., F.L.-G., M.P.M.-G. and M.J.T.-T. contributed to data collection and patient diagnosis and revised the manuscript. Y.B. and P.M. performed the bioinformatics analyses and revised the manuscript. A.M.A., B.I., O.Z., B.G. and J.G. contributed to the design

and performance of the experiments. All authors approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Acknowledgement

Statistics analyses were performed by Dr. Mahíllo Fernández, from the Biostatistics and Epidemiology Unit, Health Research Institute Fundación Jiménez Díaz, University Hospital Fundación Jiménez Díaz (Madrid, Spain). DNA samples were obtained from the Fundación Jiménez Díaz Hospital Biobank (PT13/0010/0012) (Madrid, Spain). The authors acknowledge all clinicians for the recruitment of clinical data, and thank the generous support from all patients and their families for consenting to participate in the study.

Data Availability

The data analyzed during this study are available within this article and its Supplementary Information files.

References

1. Goswami R, Bello AI, Bean J, Costanzo KM, Omer B, Cornelio-Parra D, et al. The molecular basis of spinocerebellar ataxia type 7. Front Neurosci. 2022;16:818757. 10.3389/fnins.2022.818757.
2. Pounraja VK, Sul JH, Herman J, O`Keeffe S, Rajagopal V, Bai X, et al. Population-scale repeat expansions elucidate disease risk and brain atrophy. Nature. 2026;653:796–808. 10.1038/s41586-026-10345-6.
3. Azevedo PB, Rocha AG, Keim LMN, Lavinsky D, Furtado GV, de Mattos EP, et al. Ophthalmological and neurologic manifestations in pre-clinical and clinical phases of spinocerebellar ataxia type 7. Cerebellum. 2019;18(3):388–96. 10.1007/s12311-019-1004-3.
4. Zou X, Yao F, Li F, Wu S, Li H, Sun Z, et al. Clinical characterisation and the improved molecular diagnosis of autosomal dominant cone-rod dystrophy in patients with SCA7. Mol Vis. 2021;27:221–32.
5. Matilla-Dueñas A, Corral-Juan M, Volpini V, Sánchez I. The spinocerebellar ataxias: clinical aspects and molecular genetics. Adv Exp Med Biol. 2012;724:351–74. 10.1007/978-1-4614-0653-2_27.
6. La Spada AR. Spinocerebellar ataxia type 7. In: Adam MP, Feldman J, Mirzaa GM, editors. GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993-. Updated 2020 Jul 23. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1256/
7. Gousse G, Patural H, Touraine R, Chabrier S, Rolland E, Antoine JC, et al. Lethal form of spinocerebellar ataxia type 7 with early onset in childhood. Arch Pediatr. 2018;25(1):42–4. 10.1016/j.arcped.2017.09.029.
8. Marianelli BF, Rezende Filho FM, Salles MV, Pedroso JL, Barsottini OGP, Sallum JMF. Retinal degeneration in spinocerebellar ataxia type 7: an overview of the current knowledge. Arq Bras Oftalmol. 2025;88(4):e20240248. 10.5935/0004-2749.2024-0248.
9. Nakamura Y, Tagawa K, Oka T, Sasabe T, Ito H, Shiwaku H, et al. Ataxin-7 associates with microtubules and stabilizes the cytoskeletal network. Hum Mol Genet. 2012;21(5):1099–110. 10.1093/hmg/ddr539.

10. Nassisi M, Coarelli G, Blanchard B, Dubec-Fleury C, Drine K, Kitic N, et al. *ATXN7*-related cone-rod dystrophy: the integrated functional evaluation of the cerebellum (CERMOI) study. JAMA Ophthalmol. 2024;142(4):301–8. 10.1001/jamaophthalmol.2024.0001.
11. Cerecedo-Zapata CM, Tapia-Guerrero YS, Ramírez-González JA, Meza-Dorantes A, Tercero-Pérez KN, Cortés H, et al. Current overview of spinocerebellar ataxia type 7 in Mexican population: challenges in specialized care for a rare disease. Int J Mol Sci. 2024;25(19):10750. 10.3390/ijms251910750.
12. Martin J, Van Regemorter N, Del-Favero J, Löfgren A, Van Broeckhoven C. Spinocerebellar ataxia type 7 (SCA7): Correlations between phenotype and genotype in one large Belgian family. J Neurol Sci. 1999;168(1):37–46. 10.1016/S0022-510X(99)00176-8.
13. Park JY, Joo K, Woo SJ. Ophthalmic manifestations and genetics of the polyglutamine autosomal dominant spinocerebellar ataxias: A review. Front Neurosci. 2020;14:892. 10.3389/fnins.2020.00892.
14. Karam A, Trottier Y. Molecular mechanisms and therapeutic strategies in spinocerebellar ataxia type 7. Adv Exp Med Biol. 2018;1049:197–218. 10.1007/978-3-319-71779-1_9.
15. Trang H, Stanley SY, Thorner P, Faghfoury H, Schulze A, Hawkins C, et al. Massive CAG repeat expansion and somatic instability in maternally transmitted infantile spinocerebellar ataxia type 7. JAMA Neurol. 2015;72(2):219–23. 10.1001/jamaneurol.2014.1902.
16. Johansson J, Forsgren L, Sandgren O, Brice A, Holmgren G, Holmberg M. Expanded CAG repeats in Swedish spinocerebellar ataxia type 7 (SCA7) patients: effect of CAG repeat length on the clinical manifestation. Hum Mol Genet. 1998;7(2):171–6. 10.1093/hmg/7.2.171.
17. García-Velázquez LE, Canizales-Quinteros S, Romero-Hidalgo S, Ochoa-Morales A, Martínez-Ruano L, Márquez-Luna C, et al. Founder effect and ancestral origin of the spinocerebellar ataxia type 7 (SCA7) mutation in Mexican families. Neurogenetics. 2014;15(1):13–7. 10.1007/s10048-013-0387-4.
18. Stevanin G, David G, Dürr A, Giunti P, Benomar A, Abada-Bendib M et al. Multiple origins of the spinocerebellar ataxia 7 (SCA7) mutation revealed by linkage disequilibrium studies with closely flanking markers, including an intragenic polymorphism. Eur J Hum Genet 1999;7(8):889–96. doi:10.1038/sj.ejhg.5200392.
19. Dorschner MO, Barden D, Stephens K. Diagnosis of five spinocerebellar ataxia disorders by multiplex amplification and capillary electrophoresis. J Mol Diagn. 2002;4(2):108–13. 10.1016/S1525-1578(10)60689-7.
20. Sequeiros J, Seneca S, Martindale J. Consensus and controversies in best practices for molecular genetic testing of spinocerebellar ataxias. Eur J Hum Genet. 2010;18(11):1188–95. 10.1038/ejhg.2010.10.
21. Dolzhenko E, Deshpande V, Schlesinger F, Krusche P, Petrovski R, Chen S, et al. ExpansionHunter: A sequence-graph-based tool to analyze variation in short tandem repeat regions. Bioinformatics. 2019;35(22):4754–6. 10.1093/bioinformatics/btz431.
22. Posit Team. RStudio: Integrated Development Environment for R. Boston (MA): Posit Software, PBC; 2025.
23. Martínez-Laso J, Giménez-Pardo A, Jiménez E, Gallego J, Cervera I, Ayuso C, et al. Intergenerational de novo spinocerebellar ataxia 7 CAG enlargement in a Moroccan family detected by triplet primed PCR. J Genet Disord Genet Rep. 2017;6(4). 10.4172/2327-5790.1000163.

24. Aleman TS, Cideciyan AV, Volpe NJ, Stevanin G, Brice A, Jacobson SG. Spinocerebellar ataxia type 7 (SCA7) shows a cone-rod dystrophy phenotype. Exp Eye Res. 2002;74(6):737–45. 10.1006/exer.2002.1169.
25. De Mattei F, Ferrandes F, Gallone S, Canosa A, Calvo A, Chiò A, et al. Epidemiology of spinocerebellar ataxias in Europe. Cerebellum. 2024;23(3):1176–83. 10.1007/S12311-023-01600-x.
26. Infante J, Combarros O, Volpini V, Corral J, Llorca J, Berciano J. Autosomal dominant cerebellar ataxias in Spain: molecular and clinical correlations, prevalence estimation and survival analysis. Acta Neurol Scand. 2005;111(6):391–9. 10.1111/j.1600-0404.2005.00400.x.
27. Jonasson J, Juvonen V, Sistonen P, Ignatius J, Johansson D, Björck EJ et al. Evidence for a common spinocerebellar ataxia type 7 (SCA7) founder mutation in Scandinavia. Eur J Hum Genet 2000;8(12):918–22. doi:10.1038/sj.ejhg.5200557.
28. Rodríguez-Labrada R, Martins AC, Magaña JJ, Vazquez-Mojena Y, Medrano-Montero J, Fernández-Ruíz J, et al. Founder effects of spinocerebellar ataxias in the American continents and the Caribbean. Cerebellum. 2020;19(3):446–58. 10.1007/s12311-020-01109-7.
29. Horton LC, Frosch MP, Vangel MG, Weigel-DiFranco C, Berson EL, Schmahmann JD. Spinocerebellar ataxia type 7: clinical course, phenotype-genotype correlations, and neuropathology. Cerebellum. 2013;12(2):176–93. 10.1007/s12311-012-0412-4.
30. David G, Dürr A, Stevanin G, Cancel G, Abbas N, Benomar A, et al. Molecular and clinical correlations in autosomal dominant cerebellar ataxia with progressive macular dystrophy (SCA7). Hum Mol Genet. 1998;7(2):165–70. 10.1093/hmg/7.2.165.
31. Gu W, Wang Y, Liu X, Zhou B, Zhou Y, Wang G. Molecular and clinical study of spinocerebellar ataxia type 7 in Chinese kindreds. Arch Neurol. 2000;57(10):1513–8. 10.1001/archneur.57.10.1513.

Figures

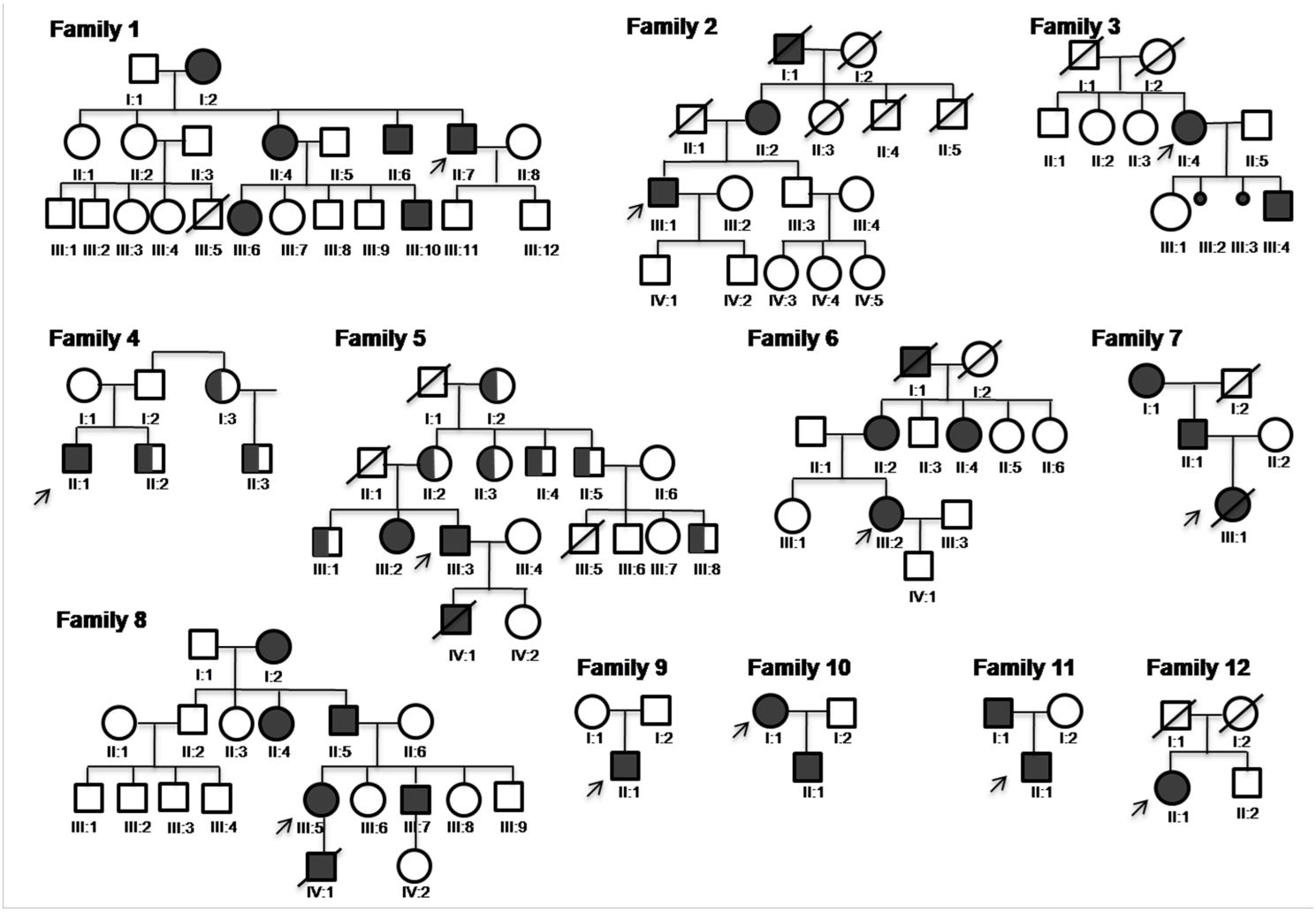


Figure 1

SCA7 families. Pedigrees are sorted by geographical origin. Half-filled symbols in Family 4 and 5 trees indicate individuals who presented neurological or ophthalmological symptoms. Index probands II:1-Family 9 and II:1-Family 12 presented as sporadic cases with no known family history of ataxia nor retinal dystrophy.

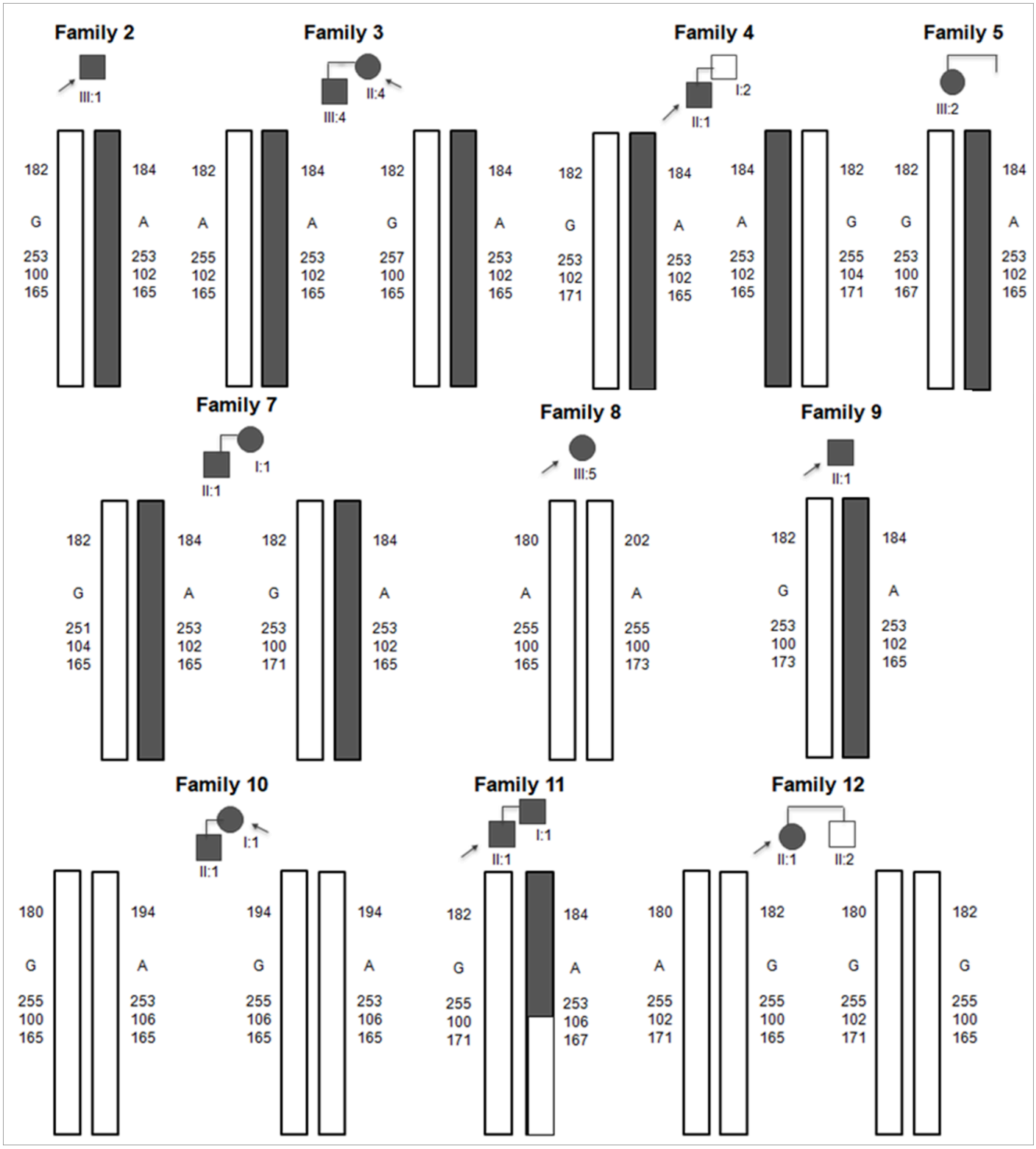


Figure 2

Haplotypes. Shared haplotypes for the altered chromosome are indicated in black colour. Phase was inferred, as segregation analysis was not possible in all cases.

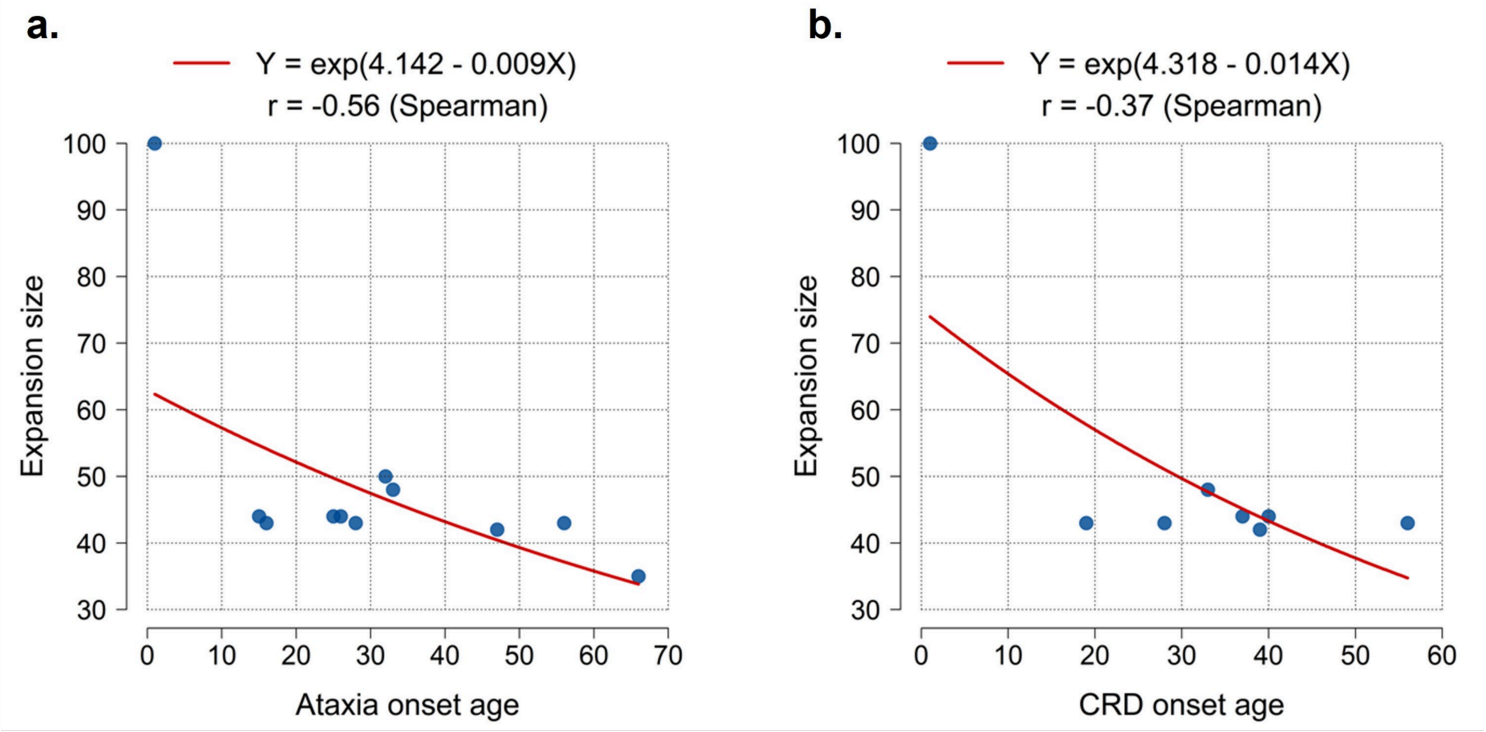


Figure 3

Correlation between age at onset and CAG repeat length. Scatter plots displaying the Spearman correlation between the age of onset (X-axis) for ataxia (a) and ophthalmological problems/cone-rod dystrophy (b) and CAG expansion size (Y-axis). Each point represents an individual patient from the FJD cohort. The line represents the exponential fit of the data.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1LopezAlcojoretal.2026.xlsx
- SupplementaryInformationLopezAlcojor.docx