Clinical phenotyping and genetic diagnosis of a large cohort of Sudanese families with hereditary spinocerebellar degeneration

Giovanni Stevanin (giovanni-b.stevanin@inserm.fr)
Univ. Bordeaux, CNRS, INCIA, UMR 5287, F-33000

Ahlam Hamed
Faculty of Medicine, University of Khartoum

Inaam Mohamed
Faculty of Medicine, University of Khartoum

Maha Elseed
Faculty of Medicine, University of Khartoum

Mustafa Salih
King Saud University

Sarah Elsadig

Hassab Elrasoul Siddig
Sudan Medical Council

Ali Nasreldien
Red Cross Memorial Children Hospital (RCWMCH), University of Cape Town

Mohamed Ahmed Abdullah
Faculty of Medicine, University of Khartoum

Maha Elzubair
Faculty of Medicine, University of Khartoum

Farouk Omer
Faculty of Medicine, Omdurman Islamic University

Aisha BAKHIET
Faculty of Medicine, University of Khartoum

Rayan Abubaker
National University Biomedical Research Institute, National University

Fatima Abozar
Faculty of Medicine, University of Khartoum

Rawaa Adil
Faculty of Medicine, University of Khartoum

Sara Emad
Faculty of Medicine, University of Khartoum

Mhammed Alhassan Musallam
Faculty of Medicine, University of Khartoum
Isra Eltazi
Hamad Medical Corporation
Zulfa Omer
Hiba Malik
Mayada Mohamed
Sudan Medical Specialization Board
Ali Elhassan
Sudan Neuroscience Projects, University of Khartoum https://orcid.org/0000-0001-8177-4494
Eman Mohamed
Faculty of Medicine, University of Khartoum
Ahmed Ahmed
Elhami Ahmed
Faculty of Medicine, University of Khartoum
Esraa Eltaraifee
Faculty of Medicine, University of Khartoum
Bidour Hussein
Institute of Endemic Diseases, University of Khartoum
Amal Abd Allah
Faculty of Medicine, University of Khartoum
Lina Mohamed
Hamad Medical Corporation https://orcid.org/0000-0002-1877-9572
Mohamed Nimir
University Hospitals Coventry and Warwickshire NHS Trust
Omnia Tag Elseed
Faculty of Medicine, University of Khartoum
Tasneem Elhassan
Faculty of Medicine, University of Khartoum
Abubakr Elbashier
Faculty of Medicine, University of Khartoum
Esraa Alfadul
Faculty of Medicine, University of Khartoum https://orcid.org/0000-0001-9812-5095
Moneeb Fadul
Faculty of Medicine, University of Khartoum
Khalil Ali
Royal Derby Hospital
Shaimaa Taha
Universal Hospital
Elfatih Bushara

Faculty of Medicine, University of Khartoum
Mutaz Amin
Faculty of Medicine, Al-Neelain University
Mahmoud koko
Muntaser Ibrahim
University of Khartoum https://orcid.org/0000-0003-2242-2143
Ammar Ahmed
Liena Elsayed
College of Medicine, Princess Nourah bint Abdulrahman University
Ashraf Yahia
Karolinska Institutet https://orcid.org/0000-0002-0376-8264

Article

Keywords: Hereditary spinocerebellar degenerative disorders, Clinical presentation, Genetic diagnosis, Sudan, next generation sequencing, exome, variant

Posted Date: November 3rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2219015/v1

Version of Record: A version of this preprint was published at European Journal of Human Genetics on April 3rd, 2023. See the published version at https://doi.org/10.1038/s41431-023-01344-6.

Abstract

Hereditary spinocerebellar degenerative disorders (SCDs) is an umbrella term that covers a group of monogenic conditions that share common pathogenic mechanisms and include spastic paraplegia, spastic ataxia, cerebellar ataxia, and spinocerebellar ataxia. They are often complicated with axonal neuropathy and/or intellectual impairment. More than 200 genes and loci inherited through all modes of Mendelian inheritance are known. Autosomal recessive inheritance predominates in consanguineous communities; however, autosomal dominant and X-linked inheritance can also occur. Sudan is inhabited by genetically diverse populations, yet it has high consanguinity rates. We used next-generation sequencing, genotyping, bioinformatics analysis, and candidate gene approaches to study 90 patients from 38 unrelated Sudanese families segregating multiple forms of SCDs focusing on known human disease-associated genes. We reached the genetic diagnosis in 63% and up to 73% of the studied families when considering variants of unknown significance. Taking into account a series of Sudanese families that we previously analyzed, the combined success rate in the two series reached 52–59% (31–35/59 families). We also highlighted the genetic and clinical heterogeneity of SCDs in Sudan, as we identified no single major gene in our cohort and the potential for discovering novel SCDs genes in this population.

Introduction

Hereditary forms of spastic paraplegia (HSP), cerebellar ataxia, spastic ataxia, the dominant and recessive cerebellar ataxia, and spinocerebellar ataxia with axonal neuropathy are distinct but overlapping clinical entities caused by related mechanisms and encompassing a continuum of phenotypes known as hereditary spinocerebellar degenerative disorders (SCDs) (1–3).

SCDs are characterized clinically by ataxia and/or spasticity complicated, in some cases, by other neurological or extra-neurological manifestations (1, 2). They have more than 220 subtypes that afflict ~ 1:10,000 individuals worldwide with evident phenotypic and genetic heterogeneity and clinical overlap (4, 5). The advent of next-generation sequencing (NGS) has markedly boosted SCDs diagnosis in recent years by the identification of a multitude of causative mutations in a large variety of disease causing genes (4). Dysfunction of mitochondria, channels and metabolisms are the main altered functions by the causative variants in these genes in addition to the abnormal expansions of nucleotide repeats (4).

Sudan is an East-African country with complex genetic and population structures (6). This complexity stemmed from the linguistic and cultural differences between its ethnic groups acting in parallel with other, sometimes opposing, population genetic forces, e.g., consanguinity, admixture, and migration (6–9). For instance, 67% of marriages in some parts of the country are consanguineous (10).

In a previous study, we screened 25 Sudanese families with HSP for mutations in 68 HSP genes using NGS targeted gene panel and reached a genetic diagnosis in 28% of these families (11). In the current study, we investigated 38 Sudanese families with SCDs using a combination of candidate gene

approaches, NGS targeted gene panel screening, and whole-exome sequencing (WES). We documented the studied patients' clinical presentations and compared the diagnostic utility of the approaches in the two studies.

Subjects And Methods

Patients recruitment and interviews

We included a total of 90 patients from 38 Sudanese families in this study with the following inclusion criteria:

1- Patients presenting with symptoms, signs, and/or history suggestive of a hereditary spinocerebellar disorder.

2- Non-genetic causes have been excluded or are extremely unlikely.

3- Participants from the family (patients and at least two healthy controls) or their guardian, in the case of patients below 18 years old or patients with intellectual disabilities, agreed to participate in the study.

4- Sudanese nationals.

5- Multiple cases affected in the same family; or first-degree relatives from a consanguineous marriage in apparently sporadic cases.

Four out of the 38 families were screened in our previous study without reaching a genetic diagnosis (11). Patients and families were interviewed and examined at the Department of Biochemistry, Faculty of Medicine, University of Khartoum, Sudan; the Pediatric Neurology Clinics, Soba University Hospital, Sudan; or the families' residences in the capital of Sudan, Khartoum, or other Sudanese cities. The diagnosis protocol followed the EUROSPA/SPATAX clinical criteria (https://spatax.wordpress.com/downloads/). We collected 2 ml of saliva from the patients and healthy-related controls using Oragene®•DNA (OG-500 and OG-575) kits (DNA Genotek Inc., Ottawa, ON, Canada).

The strategy of genetic studies

In this study, we used multiple genetic approaches, including NGS targeted gene panel screening, WES, candidate gene approach, and array genotyping (Fig. 1.A). Array genotyping was mainly used for homozygosity mapping and detection of copy number variations (CNVs). The presence of CNVs was also tested through coverage analysis in NGS data (gene panel and WES). Twenty-six families were investigated initially using HSP-targeted NGS gene panel (HSP panel) screening. Of these, eleven families were further investigated using WES and eight using WES and array genotyping. Eleven families were directly investigated using WES, without HSP panel screening. Two families were screened for repeat expansion-associated autosomal dominant spinocerebellar ataxias, and one for Friedreich's ataxia repeat expansion.

DNA extraction and quality check

We extracted DNA from saliva following the prepIT®.L2P manual protocol provided by the manufacturer (DNA Genotek). DNA quantity (Absorbance at 260nm) and quality (check of the high molecular weight DNA, absorbance ratio 260/280 and 260/320) were checked using a NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DE, USA), a Qubit® fluorometer (Promega, Madison, WI, USA), and standard agarose gel electrophoresis.

Next-generation panel screening of HSP genes

Fifty µl of patients' DNA solution at a concentration of 50 ng per dl were sent for NGS panel screening at the genotyping and sequencing core facility of the Paris Brain Institute - ICM, Paris, France. A double capture enrichment strategy was used (Roche NimbleGen® SeqCap® Ez, USA). Sequencing was done on the MiSeq® platform (Illumina, CA, USA). Detailed methodology and bioinformatics analysis are available in previous reports (11, 12). We systematically searched for point variations and genomic rearrangements. A minimal overall coverage of 30x was required for interpretation of the data.

Whole-exome sequencing

Twenty µl of DNA solution at a concentration of 20 ng per dl were sent for WES at the genotyping and sequencing core facility of the Paris Brain Institute - ICM, Paris, France. Exons were captured on the genomic DNA using the SeqCap® EZ MedExome Kit (Roche, IN, USA), followed by massively parallel sequencing on a Novaseq® 6000 sequencer (Illumina, CA, USA). Except for aligning reads to the hg37 version of the human genome (NCBI) using Burrows-Wheeler Aligner software, we processed exome data up to the calling of variants using the Genome AnalysisToolkit software (GATK) following the GATK4 best-practice pipeline.

We annotated and prioritized variants using software included in VarAFT annotation and filter tool (13). Data analysis and variants filtration were carried out based on the minor allele frequency, the variant's effect, and *in silico* prediction. We filtered all variants with allele frequencies < 0.0001 in the GnomAD genome database. First, we examined variants with predicted major structural effects; nonsense, stop loss, frameshift, and canonical splice site variants. After checking for loss of function variants, we examined missense variants annotated as pathogenic by Sift and Polyphen software (14, 15) and non-frame-shift variants. To verify that we had not missed strong candidate variants due to our conservative frequency filter, we repeated the analysis using a frequency cut-off of 0.001 in the GnomAD genome database. In this study, we focused the analysis to Online Inheritance in Man (OMIM) disease-related genes (https://www.omim.org/) and recently published HSP-causative genes with strong evidence from the literature. When multiple affected relatives were processed from the same family, they were analyzed together according to the suspected inheritance mode and then individually to take into account possible phenocopies. Genomic rearrangements were tested using PennCNV-1.0.5 (16).

Sanger sequencing

Primers were designed using Primer3 Plus software (17). DNA was amplified on a GeneAmp® PCR System 9700 (Thermo Fisher, MA, USA). We checked the quantity and quality of PCR products, including product size and off-target amplification, using the Caliper®LabChip GX System and its related software (PerkinElmer, MA, USA) according to the manufacturer's protocol. Sanger sequencing was then done at the labs of Eurofins Genomics (Germany) using the Big Dye Chemistry in an ABI3730 automated sequencer (Applied Biosystems, Thermo Fisher Scientific, USA) using the procedures recommended by the manufacturer on the PCR product. Sequencing files (ABI format) were then visualized and analyzed using Sequence Scanner Software® v2.0 (Thermo Fisher Scientific, USA).

Array genotyping

Two hundred to one thousand ng of genomic DNA from participating members of the families F5, F54, F65, F70, F73, F75, F80, F81, and F85 were sent for genotyping at the Pitié-Salpêtrière Post-Genomic Platform (P3S), Paris, France. Genotyping was performed on Illumina Infinium OmniExpress-24vl-3-A1 array, which contained ~ 710,000 SNP markers. Raw data were analyzed at the P3S platform using GenomeStudio™ Software. Runs of homozygosity were performed using version 1.07 of Plink software (18) to prioritize the variants in WES analysis. Candidate pathogenic copy number variants (CNV) were searched using PennCNV-1.0.5 software (16).

Repeats expansion detection

Genomic DNA from patients with clinical presentations and pedigree structures suggestive of dominant spinocerebellar ataxias (F49 and F65) or Friedreich's ataxia (F38) were screened for repeats expansion using specific PCR-based approaches at the genetics departments of the Pitié-Salpêtrière Hospital and University Hospital of Montpellier, France, respectively. From the dominant spinocerebellar ataxias, we screened for pathogenic DNA repeat expansions in the SCA genes *ATXN1* (SCA1), *ATXN2* (SCA2), *ATXN3* (SCA3), *CACNA1A* (SCA6), *ATXN7* (SCA7), *TBP* (SCA17), and *ATN1* (DRPLA) using a multiplex PCR amplification followed by capillary electrophoresis in a 3730 ABI sequencer (Applied Biosystems). The *FRDA* gene-associated repeat was amplified by a repeat-primed PCR approach.

Results

We studied 38 families (90 sampled patients), each including at least one patient manifesting features of SCDs. The studied families originated from multiple regions in Sudan, though the distribution is markedly skewed towards the central parts of the country. More than one-fifth of the families (23.6%) originated from a single state in central Sudan, the River Nile state (Fig. 1.B). The number of affected males and females in our cohort was approximately equal (53% males vs. 47% females). However, the patients' age at examination distribution was less homogenous; most patients were less than 18 years old. The mean and median patients' ages-at-examination were 16.6 +/- 1.5 (SD = 14.1) and 12 years, respectively (Fig. 2.B).

Disease phenotypes

Most patients had an early-onset disease; the mean and median ages-at-onset were 6.5+/- 0.93 (SD = 8.6) and three years, respectively (Fig. 2.A). Most of the patients in our cohort had spasticity (70%). Limb ataxia was noted only in ~ 30% of the patients, while ocular cerebellar signs were noted in 20% (Fig. 2.C). A pure SCDs phenotype was noted in ~ 14% of the patients, nine presented with a pure HSP while four presented with pure cerebellar ataxia. The most common features complicating the SCD phenotype in our patients were skeletal deformities, developmental delay or regression, and intellectual impairment (Fig. 2.D). Table 1 summarizes the clinical presentation and the genetic diagnosis in each family where appropriate. A more comprehensive description and case discussions are provided in the Supplementary material (only the families with a genetic diagnosis) and/or Table S1 (all the families). The diversity of clinical association did not allow us to distinguish a frequent phenotype that could have been analyzed separately as a whole. However, several families presented with similar clinical presentations, such as families F63 and F84, but they finally appeared to segregate mutations in different genes.

Genetic tests' results

We reached a provisional genetic diagnosis in 63% (24/38) of the studied families, 73% (28/38) if including families with variants of uncertain significance (VUS). In most of these families (22/28, 78.5%), diagnosis concerned all patients of the family. However, probands from six families (F38, F41, F85, F54, F70, and F80) did not share the culprit variant with some other affected members (supplementary material), a situation probably due to the high consanguinity rate that concentrated several disease-causing mutations or non-inherited phenocopies in the same sibship.

Inheritance patterns

The pattern of inheritance in most of the provisionally diagnosed families (a total of 28 families) was an autosomal recessive pattern (24/29, 83%; of note, F79 was calculated twice as it segregated two likely causative variants with different patterns of inheritance but both possibly contributing to the phenotype as we reported previously (19)). Most autosomal recessive families were segregating homozygous variants (73%) but compound heterozygous variants were observed in 10% (3/29). Autosomal dominant inheritance was identified in 10% (3/29) of the families, and two families showed X-linked inheritance (Fig. 3).

Genetic variants

We identified 31 causative or likely causative variants in known disease genes in the 28 families of this study (Table 2). One variant NM_024306.4(FA2H):c.674T > C (p.Leu225Pro) was identified twice in families F61 and F68 who shared a related phenotype. Most of the variants we identified were missense variants (12/31, 39%), followed in frequency by splice-site (23%) and frameshift (19%) variants. We identified pathogenic repeat expansions in two families and nonsense variants in three families. Array genotyping didn't detect candidate CNV or chromosomal rearrangements. Table 2 list the genetic variants identified in this study and their ACMG 2015 classes, where appropriate.

Approximately eighty percent (23/29) of the single nucleotide and insertion/deletion variants located in known disease genes were either pathogenic or likely pathogenic, according to the ACMG 2015 guidelines for interpreting sequence variations (20). Additionally, five likely causative variants fitted to the category of VUS but some with convincing evidence of pathogenicity however.

Variants of uncertain significance (VUS)

All the candidate deleterious VUS identified in this cohort segregated with the disease and could fit the categories of pathogenic or likely pathogenic variants if additional evidence is identified in the future. The VUS NM_152778.2(MFSD8):c.753A > G (p.Glu251Glu), identified in family F67, is a synonymous variant but predicted to alter the splicing of *MFSD8* (TraP score 0.96; SpliceAI score 0.7) and cause skipping of exon 8.

The second VUS, NM_001174116.1(DMXL2):c.5020A > C (p.Lys1674Gln), was identified in the two probands from family F66. It is predicted as pathogenic by Sift, Polyphen 2, MutationTaster (21), LRT (22), and Provean (23) and had a CADD score of 28. The variant was not predicted by the Missense3D tool to alter the protein structure however. On the other hand, the variant was predicted to unmask a splice site inside exon 21 which may affect the mRNA stability and must then be explored in patient's cells if expressed in leukocytes or fibroblasts. Pathogenic mutations in the *DMXL2* gene cause the autosomal dominant deafness type 71 (OMIM # 617605), and the autosomal recessive developmental and epileptic encephalopathy type 81 (OMIM # 618663) and polyendocrine-polyneuropathy syndrome (OMIM # 616113) (24–26). We herein, potentially extended the phenotype of *DMXL2* mutations to include complex HSP. Details about the clinical presentations of the patients from family F66 and the previous families with *DMXL2* variants are provided in the Supplementary material.

The variant NM_001145026.2(PTPRQ):c.5893C > A (p.Pro1965Thr) was detected in one adult patient from family F85 who presented with congenital deafness and mutism. This variant was also predicted as deleterious by Sift, Polyphen 2, MutationTaster, and Provean. It was absent in the gnomAD v2.1.1 database. Details about these variants and the associated clinical phenotypes are provided in the Supplementary materials.

Lastly, details about VUS identified in family F79 were provided in a previous report (19).

Discussion

The Sudanese population is characterized by a complex genetic structure and high consanguinity rates (6, 10). The increased homozygosity in our cohort was reflected by the predominance of mono-allelic recessive diseases (73%) and the detection of three established/possible founder variants. Two of these founder variants were in *ADAT3* and *PRUNE1* genes as we reported previously (19, 27). The third possible founder variant, NM_024306.4(FA2H):c.674T > C (p.Leu225Pro), was detected in two unrelated families, F61 and F68, that descended from different tribes in Kordofan province, western Sudan. Nevertheless, we also identified autosomal dominant and X-linked (hemizygous) conditions in several families. Most of

our families originated from the central parts of Sudan. This can be attributed either to differences in the accessibility to the health system and our collaborating clinics or genuine differences in the frequency of genetic diseases between central Sudan populations and other Sudanese populations. We favor the first explanation as other consanguinity-linked genetic diseases, such as sickle cell anemia, are common in non-central parts of the country (10).

All age groups were represented in our cohort, particularly those < 18 years, indicating the degree of care provided to this age group by their families. On the other hand, we have patients with childhood-onset diseases who were first examined after their forties (after decades of disease duration, > 40 years in two patients), epitomizing the long-term odysseys of patients with genetic diseases and underlining the importance of genetic diagnosis for patients and families satisfaction. Also, the percentages of males and females in our cohort were approximately equal, signifying the absence of gender-based inequalities in the accessibility of care and minimizing the contribution of X-linked dominant inheritance to SCDs in our cohort.

Previously, we screened 25 Sudanese families with HSP for mutations in 68 known HSP genes using NGS targeted gene panel (28). We reached a genetic diagnosis in 28% of these cases (28), a diagnosis rate very similar to Portuguese (29) and European (12) patients. This last study, (ref. 12), showed that combining the HSP panel with subsequent WES increased the diagnosis rate up to 50% when focusing on OMIM disease-related genes. WES used to further identify novel genes was shown to give a diagnostic yield of up to 75% (30). In the current study, by using multiple genetic approaches, we identified disease-causing variants in known SCDs genes in 63–73% of the studied families (our overall diagnostic success rate if we consider our previous cohort (ref 11) is 52–59% (31–35/59 families)). Furthermore, extending the analysis to novel genes, we identified variants in novel candidate genes in seven out of the ten remaining families, potentially raising our diagnostic success rate ceiling to 92% instead of 73% (one of those seven novel causative genes has been reported (31) and the others are under validation). According to the results of our two studies, most of the major autosomal recessive SCDs genes are present in Sudan (*SACS*, *SPG11*, *FXN*) and some of the major dominant ones as well (e.g., *SCA3*), but there is no single major gene causing SCDs in Sudan. This might result from the position of Sudan in east Africa, at the frontiers between North Africa, the Middle East, and sub-Saharan Africa.

WES outweighs NGS targeted gene panel in discovering new SCDs genes (4). However, based on our experience with the Sudanese population, and the experience of others, exome sequencing also significantly outweighs NGS-targeted gene panels in diagnosing known SCDs phenotypes, particularly in complex phenotypes (32). Furthermore, WES enables the extension of phenotypes previously associated with mutations in certain genes in contrast to conservative NGS-targeted gene panels that target only the phenotype of interest. For instance, we extended the phenotypes associated with mutations in *CCDC82* and *CCDC88C* in the current Sudanese cohort by using WES. *CCDC82* was reported previously to cause an intellectual disability syndrome (33, 34). We expanded the *CCDC82*-linked phenotype to include spastic paraplegia (19). Later, another report of a patient of Pakistani origin confirmed that spasticity is part of the *CCDC82*-linked syndrome (35). Similarly, we expanded the presentation of mono-allelic mutations in

CCDC88C to include early-onset pure spastic paraplegia (36). Before, mono-allelic gain-of-function *CCDC88C* mutations were only associated with spinocerebellar ataxia SCA40 (37). In this report we also potentially extended the phenotype of *DMXL2*-linked disorders to include complex HSP.

In our opinion, the higher diagnostic success rate of WES overrides its technical difficulties when compared to NGS targeted gene panel upon studying diseases with overlapping phenotypes like SCDs, particularly when considering the increasing technical feasibility of WES (38). However, WES is less efficient for rearrangement detection than panels of genes, usually optimized for such discovery, as discussed (ref, 12). An issue in SCDs is the detection of nucleotide repeat expansions that require independent specific techniques but there are improvements of some algorithm for such quest in WES data and in genome sequencing (39).

In conclusion, up-to-now, SCDs in Sudan are caused by multiple genes; none of them significantly predominate over the others. The use of multiple genetic approaches that included WES enhanced the diagnosis of known SCDs phenotypes and the potential discovery of new SCDs genes.

Declarations

Data availability statement

The data supporting the findings of this study (not including participants' personal information) are available from the corresponding author upon reasonable request. All novel variants have been submitted to ClinVar (submission number SUB12076628).

Acknowledgments

We are grateful to the genotyping and sequencing core facility of the Brain Institute – ICM and the Pitié-Salpêtrière Post-Genomic Platform (P3S), Paris, France for doing the next-generation sequencing and array genotyping, respectively. We thank Nicolas Auger, Claire-Sophie Davoine, Mélanie Papin and Lena Guillot-Noel for their advices and Prof. Michel Koenig and Dr. Guillaume Banneau for their help in nucleotide repeat expansion detection.

Authors' contributions

AY, AEA, LEOE, and GS designed the study. AY; AAAH; INM; MAH; MAS; SME; HES; AEMN; MAA; ME; FYO; AMB; SOMAT; EEB; MK; AEA; and LEOE evaluated the patients. AY, RA1; FA; RA2; SE; MAM; IZME; ZO; HM; MOEM; AAE; EOEM; AKMAA; EAAA; EE; BKH; ASIAA; LS; MN; OMTE; TEAE; AE; ESAA; MF; KFA; MA; and LEOE collected the samples and relevant data and performed the experiments. AY and GS interpreted the results. MEI, AEA, LEOE, and GS supervised the study. AY and GS drafted the manuscript. GS obtained funds for implementing the study procedures. All authors critically revised and approved the final version

of the manuscript and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

This study was financially supported by the European Union (grant n°779257 - Solve-RD from the Horizon 2020 Research and Innovation Program, to GS), the Association Connaitre les Syndromes Cérébelleux (to GS) and by the ASL-HSP-France patient association (to GS). AY was supported by the Ministry of Higher Education, Sudan and the French Embassy, Sudan. MAS was supported by Researchers Supporting Project, King Saud University, Riyadh, Saudi Arabia (project number RSP-2020/38).

Ethics approval and consent to participate

This study was approved by The Ethical Committee of Medical Campus, University of Khartoum, Sudan, and The Ethical Committee of the National University, Sudan (approval number NU-RECG200). All procedures in this study were performed per the 1975 Helsinki declaration and its later amendments. Informed written consent forms for participation were obtained from all participants or their guardians.

Competing interests

GS received a grant from BIOGEN (Cambridge, USA) unrelated to this work, and the other authors declare that they have no competing interests.

References

1. Synofzik M, Schüle R. Overcoming the divide between ataxias and spastic paraplegias: Shared phenotypes, genes, and pathways [Internet]. Vol. 32, Movement Disorders. John Wiley and Sons Inc.; 2017 [cited 2020 Aug 28]. p. 332–45. Available from: /pmc/articles/PMC6287914/?report=abstract
2. Parodi L, Coarelli G, Stevanin G, Brice A, Durr A. Hereditary ataxias and paraparesias: Clinical and genetic update [Internet]. Vol. 31, Current Opinion in Neurology. Lippincott Williams and Wilkins; 2018 [cited 2020 Sep 15]. p. 462–71. Available from: https://pubmed.ncbi.nlm.nih.gov/29847346/
3. Manto M, Gandini J, Feil K, Strupp M. Cerebellar ataxias: An update [Internet]. Vol. 33, Current Opinion in Neurology. Lippincott Williams and Wilkins; 2020 [cited 2020 Nov 23]. p. 150–60. Available from: https://pubmed.ncbi.nlm.nih.gov/31789706/
4. Yahia A, Stevanin G. The History of Gene Hunting in Hereditary Spinocerebellar Degeneration: Lessons From the Past and Future Perspectives. Front Genet [Internet]. 2021 Mar 23 [cited 2021 Apr 17];12. Available from: /pmc/articles/PMC8021710/

5. Ruano L, Melo C, Silva MC, Coutinho P. The Global Epidemiology of Hereditary Ataxia and Spastic Paraplegia: A Systematic Review of Prevalence Studies. Neuroepidemiology [Internet]. 2014 [cited 2018 Oct 15];42(3):174–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24603320
6. Ibrahim ME. Genetic diversity of the Sudanese: insights on origin and implications for health. Hum Mol Genet [Internet]. 2021 Mar 1 [cited 2021 Nov 29];30(R1):R37–41. Available from: https://pubmed.ncbi.nlm.nih.gov/33864377/
7. Hollfelder N, Schlebusch CM, Günther T, Babiker H, Hassan HY, Jakobsson M. Northeast African genomic variation shaped by the continuity of indigenous groups and Eurasian migrations. Tishkoff SA, editor. PLOS Genet [Internet]. 2017 Aug 24 [cited 2020 Mar 22];13(8):e1006976. Available from: https://dx.plos.org/10.1371/journal.pgen.1006976
8. Dobon B, Hassan HY, Laayouni H, Luisi P, Ricaño-Ponce I, Zhernakova A, et al. The genetics of East African populations: a Nilo-Saharan component in the African genetic landscape. Sci Rep [Internet]. 2015 May 28 [cited 2019 Sep 25];5(1):9996. Available from: http://www.nature.com/articles/srep09996
9. Babiker HMA, Schlebusch CM, Hassan HY, Jakobsson M. Genetic variation and population structure of Sudanese populations as indicated by 15 Identifiler sequence-tagged repeat (STR) loci. Investig Genet [Internet]. 2011 May 4 [cited 2021 Mar 6];2(1):12. Available from: http://investigativegenetics.biomedcentral.com/articles/10.1186/2041-2223-2-12
10. Daak AA, Elsamani E, Ali EH, Mohamed FA, Abdel-Rahman ME, Elderdery AY, et al. Sickle cell disease in western Sudan: Genetic epidemiology and predictors of knowledge attitude and practices. Trop Med Int Heal [Internet]. 2016 May 1 [cited 2020 Mar 22];21(5):642–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27028397
11. Elsayed LEO, Mohammed IN, Hamed AAA, Elseed MA, Johnson A, Mairey M, et al. Hereditary spastic paraplegias: Identification of a novel SPG57 variant affecting TFG oligomerization and description of HSP subtypes in Sudan. Eur J Hum Genet [Internet]. 2016;25(1):100–10. Available from: http://dx.doi.org/10.1038/ejhg.2016.108
12. Méreaux J-L, Banneau G, Papin M, Coarelli G, Valter R, Raymond L, et al. Clinical and genetic spectra of 1550 index patients with hereditary spastic paraplegia. Brain [Internet]. 2022 Apr 29 [cited 2022 May 28];145(3). Available from: https://pubmed.ncbi.nlm.nih.gov/34983064/
13. Desvignes J-P, Bartoli M, Delague V, Krahn M, Miltgen M, Béroud C, et al. VarAFT: a variant annotation and filtration system for human next generation sequencing data. Nucleic Acids Res [Internet]. 2018 [cited 2020 Mar 21];46(W1):W545–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29860484
14. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations [Internet]. Vol. 7, Nature Methods. NIH Public Access; 2010 [cited 2020 Sep 29]. p. 248–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2855889/

15. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res [Internet]. 2003 Jul 1 [cited 2020 Sep 29];31(13):3812–4. Available from: /pmc/articles/PMC168916/?report=abstract
16. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SFA, et al. PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res [Internet]. 2007 Nov [cited 2020 Nov 2];17(11):1665–74. Available from: https://pubmed.ncbi.nlm.nih.gov/17921354/
17. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res [Internet]. 2007 Jul [cited 2020 Nov 2];35(SUPPL.2). Available from: https://pubmed.ncbi.nlm.nih.gov/17485472/
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am J Hum Genet [Internet]. 2007 [cited 2021 Dec 2];81(3):559. Available from: /pmc/articles/PMC1950838/
19. Yahia A, Ayed I Ben, Hamed AA, Mohammed IN, Elseed MA, Bakhiet AM, et al. Genetic diagnosis in Sudanese and Tunisian families with syndromic intellectual disability through exome sequencing. Ann Hum Genet [Internet]. 2022 [cited 2022 Mar 11]; Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/ahg.12460
20. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med [Internet]. 2015 May 8 [cited 2020 Dec 19];17(5):405–24. Available from: https://pubmed.ncbi.nlm.nih.gov/25741868/
21. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations [Internet]. Vol. 7, Nature Methods. Nature Publishing Group; 2010 [cited 2020 Sep 29]. p. 575–6. Available from: https://www.nature.com/articles/nmeth0810-575
22. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res [Internet]. 2009 Sep 1 [cited 2020 Dec 21];19(9):1553–61. Available from: http://www.genome.org/cgi/doi/10.1101/gr.092619.109.
23. Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics [Internet]. 2015 Jan 19 [cited 2020 Sep 29];31(16):2745–7. Available from: /pmc/articles/PMC4528627/?report=abstract
24. Tata B, Huijbregts L, Jacquier S, Csaba Z, Genin E, Meyer V, et al. Haploinsufficiency of Dmxl2, Encoding a Synaptic Protein, Causes Infertility Associated with a Loss of GnRH Neurons in Mouse. PLoS Biol [Internet]. 2014 [cited 2021 Jan 3];12(9). Available from: https://pubmed.ncbi.nlm.nih.gov/25248098/
25. Maddirevula S, Alzahrani F, Al-Owain M, Al Muhaizea MA, Kayyali HR, AlHashem A, et al. Autozygome and high throughput confirmation of disease genes candidacy. Genet Med [Internet]. 2019 Mar 1 [cited 2021 Jan 3];21(3):736–42. Available from: https://pubmed.ncbi.nlm.nih.gov/30237576/

26. Chen DY, Liu XF, Lin XJ, Zhang D, Chai YC, Yu DH, et al. A dominant variant in DMXL2 is linked to nonsyndromic hearing loss. Genet Med [Internet]. 2017 May 1 [cited 2021 Jan 3];19(5):553–8. Available from: https://pubmed.ncbi.nlm.nih.gov/27657680/
27. Koko M, Yahia A, Elsayed LE, Hamed AA, Mohammed IN, Elseed MA, et al. An identical-by-descent novel splice-donor variant in PRUNE1 causes a neurodevelopmental syndrome with prominent dystonia in two consanguineous Sudanese families. Ann Hum Genet [Internet]. 2021 Jun 10 [cited 2021 Jun 14];ahg.12437. Available from: https://onlinelibrary.wiley.com/doi/10.1111/ahg.12437
28. Elsayed LEO, Mohammed IN, Hamed AAA, Elseed MA, Johnson A, Mairey M, et al. Hereditary spastic paraplegias: identification of a novel SPG57 variant affecting TFG oligomerization and description of HSP subtypes in Sudan. Eur J Hum Genet [Internet]. 2016 Jan 7 [cited 2018 Jan 15];25(1):100–10. Available from: http://www.nature.com/articles/ejhg2016108
29. Morais S, Raymond L, Mairey M, Coutinho P, Brandão E, Ribeiro P, et al. Massive sequencing of 70 genes reveals a myriad of missing genes or mechanisms to be uncovered in hereditary spastic paraplegias. Eur J Hum Genet [Internet]. 2017 Nov 23 [cited 2018 Jan 15];25(11):1217–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28832565
30. Novarino G, Fenstermaker AG, Zaki MS, Hofree M, Silhavy JL, Heiberg AD, et al. Exome Sequencing Links Corticospinal Motor Neuron Disease to Common Neurodegenerative Disorders. Science (80-) [Internet]. 2014 Jan 31 [cited 2018 Jan 15];343(6170):506–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24482476
31. Yahia A, Elsayed LEO, Valter R, Hamed AAA, Mohammed IN, Elseed MA, et al. Pathogenic Variants in ABHD16A Cause a Novel Psychomotor Developmental Disorder With Spastic Paraplegia. Front Neurol [Internet]. 2021 Aug 20 [cited 2021 Dec 5];12. Available from: https://pubmed.ncbi.nlm.nih.gov/34489854/
32. Coutelier M, Hammer MB, Stevanin G, Monin ML, Davoine CS, Mochel F, et al. Efficacy of exome-targeted capture sequencing to detect mutations in known cerebellar ataxia genes. JAMA Neurol [Internet]. 2018 May 1 [cited 2020 Oct 17];75(5):591–9. Available from: https://pubmed.ncbi.nlm.nih.gov/29482223/
33. Harripaul R, Vasli N, Mikhailov A, Rafiq MA, Mittal K, Windpassinger C, et al. Mapping autosomal recessive intellectual disability: combined microarray and exome sequencing identifies 26 novel candidate genes in 192 consanguineous families. Mol Psychiatry [Internet]. 2018 Apr 1 [cited 2021 Jan 3];23(4):973–84. Available from: https://pubmed.ncbi.nlm.nih.gov/28397838/
34. Riazuddin S, Hussain M, Razzaq A, Iqbal Z, Shahzad M, Polla DL, et al. Exome sequencing of Pakistani consanguineous families identifies 30 novel candidate genes for recessive intellectual disability. Mol Psychiatry [Internet]. 2017 Nov 1 [cited 2021 Jan 3];22(11):1604–14. Available from: https://pubmed.ncbi.nlm.nih.gov/27457812/
35. Bauer G, Buchert R, Haack TB, Harting I, Gutschalk A. CCDC82 frameshift mutation associated with intellectual disability, spastic paraparesis, and dysmorphic features. Clin Genet [Internet]. 2022 Jul 1

[cited 2022 Sep 14];102(1):80–1. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/cge.14135

36. Yahia A, Chen ZS, Ahmed AE, Emad S, Adil R, Abubaker R, et al. A heterozygous mutation in the CCDC88C gene likely causes early-onset pure hereditary spastic paraplegia: a case report. BMC Neurol [Internet]. 2021 Dec 18 [cited 2021 Feb 19];21(1):78. Available from: https://bmcneurol.biomedcentral.com/articles/10.1186/s12883-021-02113-y
37. Tsoi H, Yu ACS, Chen ZS, Ng NKN, Chan AYY, Yuen LYP, et al. A novel missense mutation in CCDC88C activates the JNK pathway and causes a dominant form of spinocerebellar ataxia. J Med Genet. 2014;51(9):590–5.
38. Suwinski P, Ong CK, Ling MHT, Poh YM, Khan AM, Ong HS. Advancing personalized medicine through the application of whole exome sequencing and big data analytics [Internet]. Vol. 10, Frontiers in Genetics. Frontiers Media S.A.; 2019 [cited 2020 Oct 19]. Available from: https://pubmed.ncbi.nlm.nih.gov/30809243/
39. Depienne C, Mandel JL. 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges? Am J Hum Genet [Internet]. 2021 May 6 [cited 2022 Oct 18];108(5):764–85. Available from: http://www.cell.com/article/S0002929721000951/fulltext
40. Yahia A, Elsayed L, Babai A, Salih MA, El-Sadig SM, Amin M, et al. Intra-familial phenotypic heterogeneity in a Sudanese family with DARS2-related leukoencephalopathy, brainstem and spinal cord involvement and lactate elevation: A case report 11 Medical and Health Sciences 1109 Neurosciences 06 Biological Sciences 0604 G. BMC Neurol. 2018;18(1):1–6.

tables

Tables 1 and 2 are available in the Supplementary Files section.

Figures

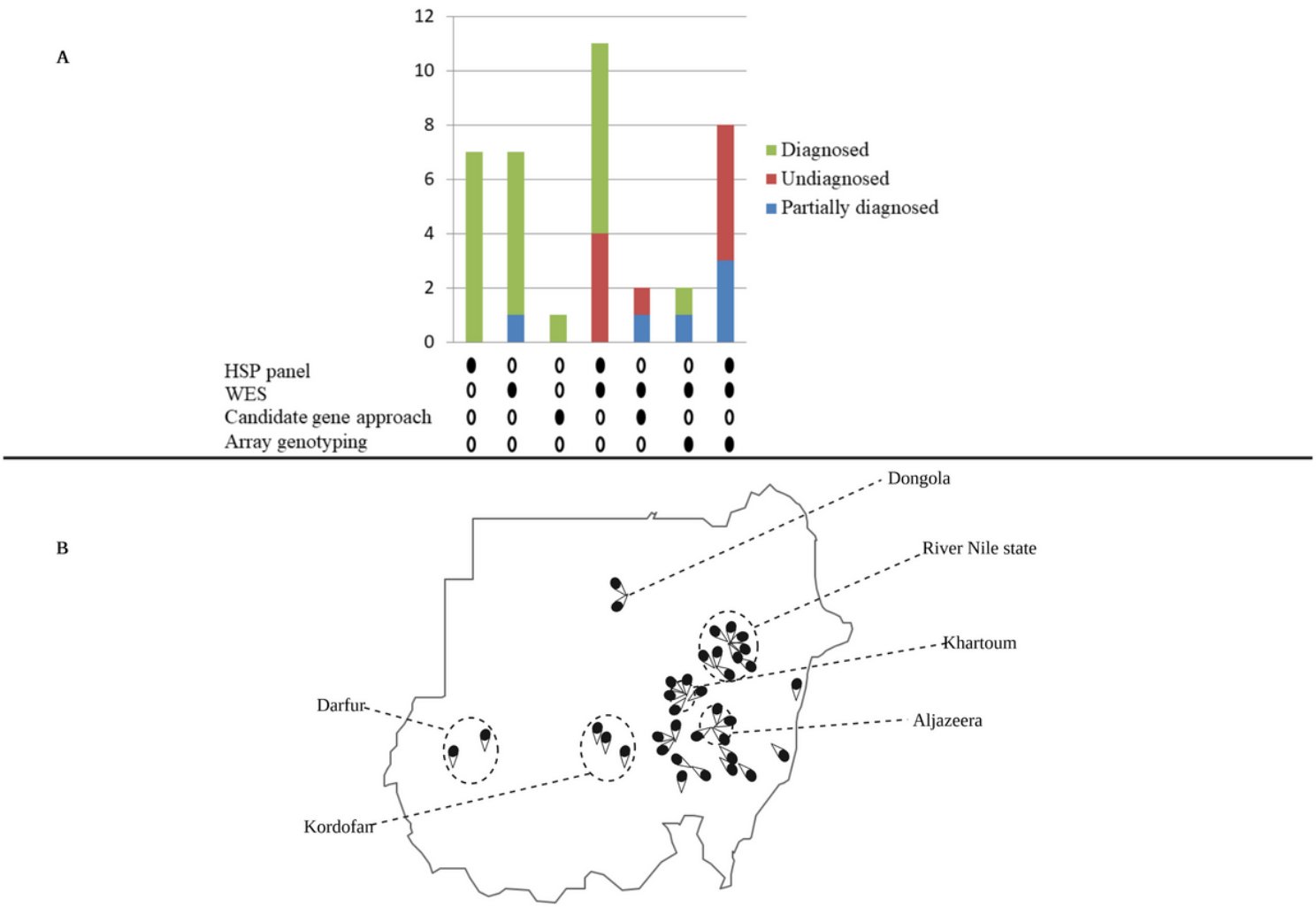

Figure 1

Genetic tools used for investigating our families and their utility (A) and the geographical origin of the families (B). A. The various strategies used include HSP panel, hereditary spastic paraplegia next-generation sequencing targeted gene panel; WES, whole-exome sequencing, candidate gene approach and DNA microarrays. The families partially diagnosed relate to families where only a fraction of the patients was diagnosed. B. The figure shows some regions, states, or cities in Sudan from which the studied families originated (each pin-drop represents a single family).

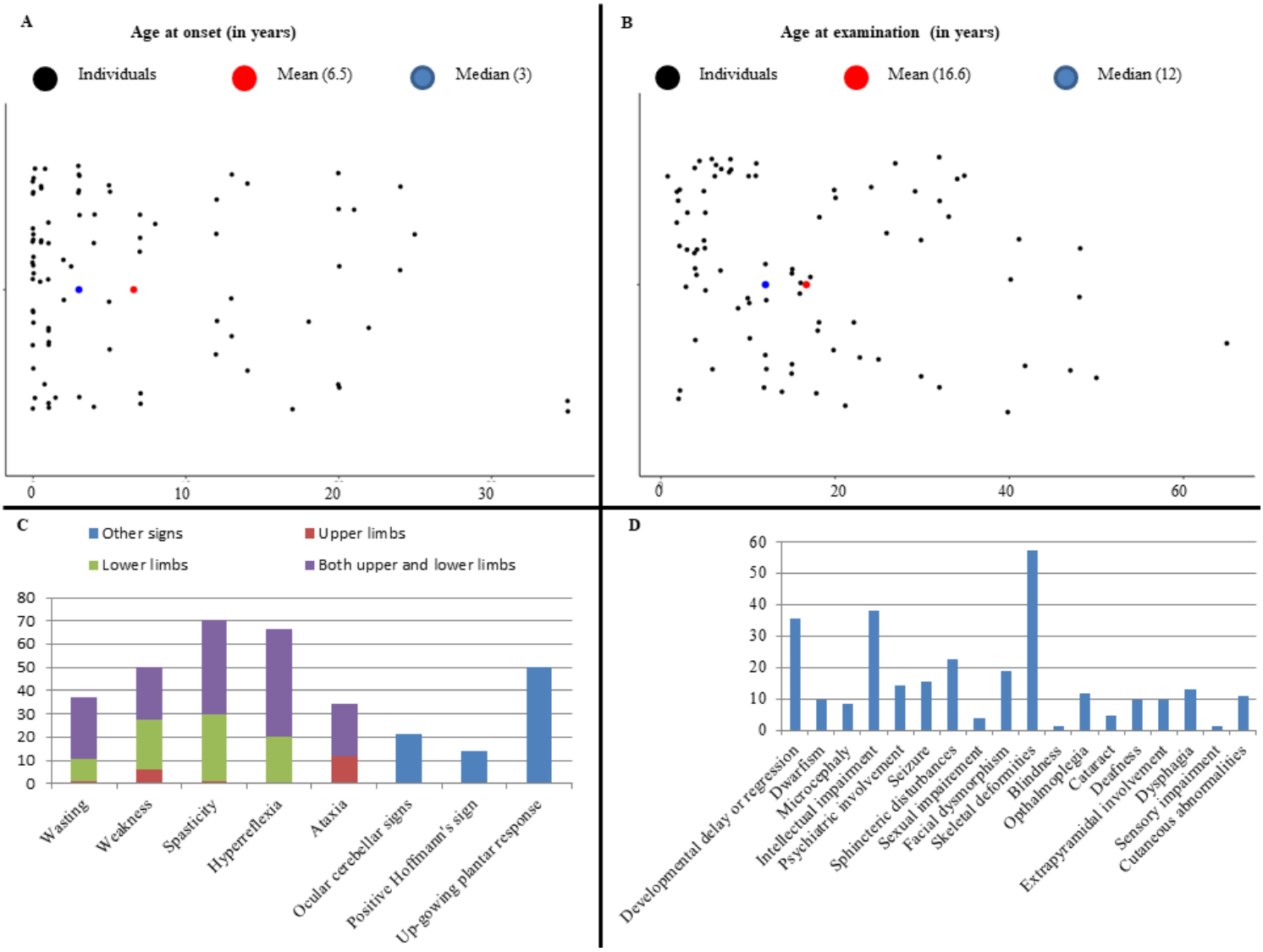

Figure 2

Clinical overview of the cohort. A. Patients' age-at-examination. The mean age-at-examination was 16.6 years. B. Age-at-onset of the SCDs in our patients. The mean age-at-onset was 6.5 years. C. Signs detected during patients' examination. The percentages of patients with pyramidal and cerebellar signs are shown. The majority of our patients presented with pyramidal features. D. Features complicating the SCDs phenotype in our cohort. Skeletal deformities, intellectual impairment, and developmental delay and/or regression are the most common features complicating the SCDs phenotype in our cohort.

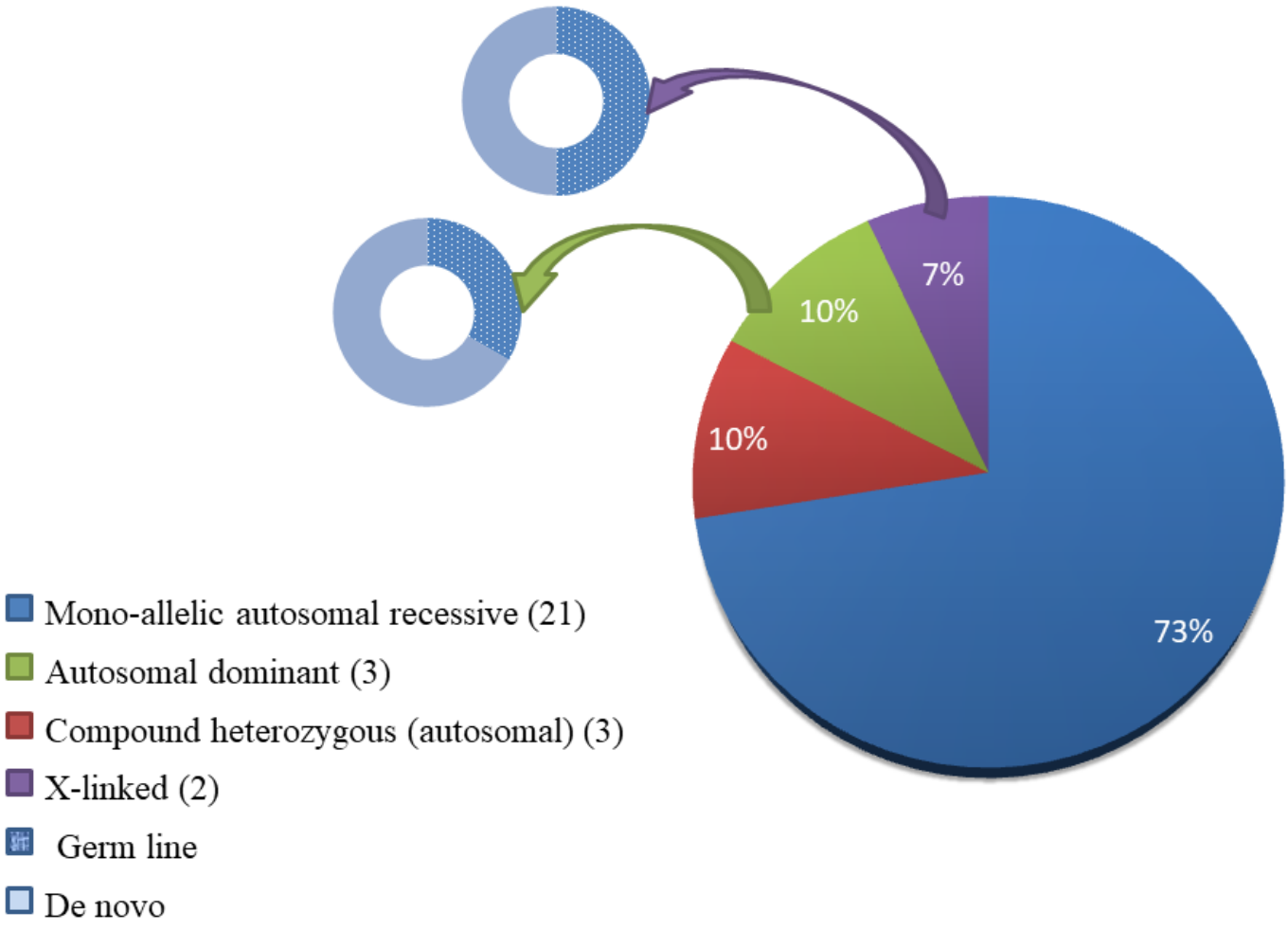

Figure 3

The pattern of inheritance in the families with mutations in known disease genes. Compound heterozygous inheritance is separated from mono-allelic (homozygous) autosomal recessive inheritance to highlight the effect of consanguinity.

Supplementary Files

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

- TableS1.xlsx
- Table1.xlsx
- Table2.xlsx
- SCDsSupplementarymaterialv8.docx