Whole Exome Sequencing for the Diagnosis of Rare Genetic Neurodevelopmental Disorders Associated with Cerebellar Atrophy

Engy A. Ashaat (ea.ashaat@nrc.sci.eg)
National Research Centre
Hoda A. Ahmed
National Research Centre
Nesma M. Elaraby
National Research Centre
Alaaeldin Fayez
National Research Centre
Ammal M. Metwally
National Research Centre
Mona K. Mekkawy
National Research Centre
Neveen A. Ashaat
Ain Shams University
Rasha Moheb Elhossini
National Research Centre
Heba Ahmed ElAwady
Fayoum University Hospitals
Randa H. A. Abdelgawad
Ain Shams University
Mona El Gammal
National Research Centre
Mohamed Ahmed Al Kersh
Ain Shams University
Dina Amin Saleh
Ain Shams University

Research Article

Keywords: Neurodegenerative disorders, Novel variants, Cerebellar atrophy, WES

Posted Date: January 19th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2481124/v1

Abstract

Neurodegenerative disorders (NDs) are rare multifactorial disorders characterized by dysfunction and degeneration of synapses, neurons, and glial cells which are essential for movement, coordination, muscle strength, sensation and cognition. It may also be associated with neuroinflammation and oxidative stress. Several genes have been identified underlying the different forms. Herein, we describe seven patients from 6 Egyptian families. The core clinical features of our patients included dysmorphic features, neurodevelopmental delay or regression, gait abnormalities, skeletal deformities, visual impairment, and seizures. Previously unreported clinical phenotypic findings were recorded. Whole-exome sequencing (WES) was performed followed by in silico analysis of the detected genetic variants effect on the protein structure. Three novel variants were identified in three genes *"MFSD8 (CLN7)*, *AGTPBP1*, and *APTX"* and other previously reported three variants have been detected in "*TPP1*, *AGTPBP1* and *PCDHGC4"* genes. In this cohort, we described the detailed unique phenotypic characteristics in view of the identified genetic profile in patients with ND disorders, hence expanding the mutational spectrum of such disorders.

Introduction

Inherited neurodegenerative disorders (NDs) are common in Egypt and constitute a public health burden due to the high rate of consanguinity [1]. NDs is a heterogeneous group of mostly genetically determined diseases that results in progressive loss of neuronal structures or functions in different areas of the central and peripheral nervous system with a resultant loss of the previously acquired motor, sensory and cognitive functions [2]. However, the accretion of new developmental milestones does not exclude the existence of a ND disorder.

On the clinical level, NDs share similar manifestations including visual and hearing impairment, seizures, skeletal deformities, feeding and intellectual difficulties [3]. Therefore, reaching a specific diagnosis could be quite challenging in the pediatric age group especially in resource-limited countries due to several reasons such as "the ability of the clinicians to discriminate between the loss of a previously acquired or a delay in the achievement of specific developmental milestones, lack of expertise and the long list of unaffordable potential investigations including molecular genetic analysis [4]. On the molecular and biochemical level, NDs are characterized by depositions of misfolded, toxic conformations of various proteins, which generally accumulate to form insoluble deposits [5].

Neuronal ceroid lipofuscinosis (NCLs) are a group of clinically and genetically heterogeneous ND lysosomal storage disorders that mostly present by seizures, visual impairment, and a progressive decline in cognitive and motor abilities due to progressive neuronal death [6]. Lysosomal accumulation of autofluorescent lipopigments and proteins in the central nervous system is a key pathological finding of NCLs [7]. This storage process is associated with selective destruction and loss of neurons in brain and retina.

Prior to the revolution in the genetic science, the NCLs classification was based on the age of onset and the ultrastructural electron microscopic findings [7]. Currently, the NCL family includes 14 different subtypes, each having distinct monogenetic defects in genes encoding proteins in the endolysosomal system. The global combined incidence of NCL is estimated to be 1:12,500 to 1: 100,000 [8]. All encountered genes lie on autosomes and in most cases, the disease is inherited in a recessive manner, with the exception of adult onset NCL caused by dominantly inherited mutations in *CLN4/DNAJC5*. Some NCL genes encode lysosomal proteins - enzymes and soluble proteins such as (*CLN1/PPT1*, *CLN2/TPP1*, *CLN5*, *CLN10/CTSD*, *CLN13/CTSF, CLN11/ GRN*,) or membrane proteins (*CLN3*, *CLN7/MFSD8*, *CLN12/ATP13A2*). Two encode endoplasmic reticulum membrane proteins *(CLN6*, *CLN8*), while other NCL proteins are cytoplasmic (*CLN4/DNAJC5*, *CLN14/KCTD7*) that peripherally associate with cellular membranes [9, 10]. Nowadays, the early identification of the molecular NCLs type became very crucial in view of the novel therapeutic approaches for the NCLs and lysosomal storage disorders (LSDs) [11].

Childhood-onset neurodegeneration with cerebellar atrophy (CONDCA; OMIM # 618276) is a rare autosomal recessive (AR) NDs that is caused by variants in the AGTPBP1 gene. AGTPBP1 gene (OMIM *606830) is located on chromosome 9 and encodes cytosolic carboxypeptidase 1 (CCP1), an enzyme involved in deglutamylation of polyglutamylated proteins [12, 13]. All patients with CONDCA had a progressive disease course and presented with impaired intellectual development, poor or absent speech and motor abnormalities. Additional features included microcephaly, strabismus, nystagmus, muscle weakness and atrophy, ataxia, spasticity, tongue fasciculation, and axonal motor neuropathy [14].

Ataxia-ocular apraxia type 1 (AOA1), another NDs was associated with mutations in *APTX* gene and it is the only gene known to be associated with it. AOA1 patients commonly present with cerebellar ataxia with peripheral axonal neuropathy, limitation of ocular movements on command, and hypoalbuminemia. *APTX* gene encodes aprataxin protein that is involved in the repair of DNA damage in cells of various tissues, including the brain, spinal cord, and muscles through its nucleotide-binding activity and its diadenosine polyphosphate hydrolase activity [15].

Neurodevelopmental disorder with poor growth and skeletal anomalies (NEDGS) (OMIM; 619880) is a global NDs with highly variable features characterized by progressive microcephaly, short stature, intellectual disability, seizures, developmental delay and joint anomalies. Mutations in the *PCDHGC4* gene were identified in the affected individuals by Iqba et al in 2021 [16].

In this study, we will review the clinical features, explore the molecular and mutation spectrum in seven Egyptian patients with NDs with an overlapping phenotype. We employed whole exome sequencing (WES) to screen the mutations and investigate the genotypic and phenotypic heterogeneities of molecularly characterized patients. By this, we aim to provide a better understanding of NDs among clinicians especially in resource-limited countries to help them offer appropriate management, prognosis expectations and proper genetic counseling.

Material And Methods

Ethical approval:

The ethical approval was granted by the Medical Research Ethics committee of the National Research Centre (NRC), Cairo, Egypt (No. 932702021) according to the Declaration of Helsinki. Informed consent was obtained from patient's parents.

Patient enrollment and clinical analysis:

The patients were recruited from the Multiple Congenital Anomalies Clinic (MCAs), Clinical Genetics Department, National Research Centre (NRC). Patients either presented primarily at the MCAs or have been referred by the participating pediatric neurologists or ophthalmologists for further evaluation and work up completion.

Patients presented with developmental delay, neurodevelopmental regression (NDR), neurobehavioral disorders, visual or hearing impairment, short stature, abnormal gait, skeletal abnormalities, seizures or family history of early unexplained death with uneventful perinatal or postnatal course were included in this study.

Once patients were identified as having possible NDs, the caregivers were counseled about the possible genetic diagnosis and the required genetic test. The initial evaluation of patients comprised detailed history taking including "family history of a similar condition or other genetic disorders and perinatal history of possible prenatal insult or postnatal complications", neurodevelopmental milestones assessment and thorough physical and neurological examination. Patients with perinatal or postnatal complications such as kernicterus, meningitis, stroke and head trauma were excluded from the study.

Detailed demographic data including age, gender, age at onset of symptoms, mode of delivery and parental consanguinity were reviewed for all the patients. Anthropometric measurements including "head circumference, weight and height/length" were plotted according to the recommendation of the International Biological Program (IBP) **[17]**. A special emphasis on the presence of dysmorphic features, skeletal deformities and review of other body systems was carried out to analyze the phenotypic presentation of the index patients. Physical and neurological examination was carried out for both parents and available siblings. Three generation pedigrees including consanguinity and other affected family members were constructed.

Other ancillary tests:

Skeletal survey including plain X-rays of the "skull, spine, pelvis, short and long bones", brain magnetic resonance imaging (MRI), electroencephalogram (EEG), electroretinogram (ERG), visual evoked potential (VEP) and echocardiography (Echo) was carried out whenever indicated.

Cytogenetic analysis:

Karyotype analysis was performed for all patients to exclude the presence of any chromosomal abnormalities. Chromosomal preparations were done from peripheral blood samples collected on lithium heparin vacutainers, following standard protocols **[18]**.

Molecular analyses:

DNA extraction and WES

Genomic DNA was extracted from peripheral blood samples of all patients and their available family members using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). Quality and quantity of DNA samples of patients were assessed using fluorometric Denovix Qubit™ dsDNA BR Assay Kit (ThermoFisher, Waltham, MA, USA). DNA samples were sequenced by using the Twist Human Core Exome Plus kit (Twist Bioscience, San Francisco, CA, USA) and Illumina NovaSeq 6000 system (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. Libraries were prepared in paired-end mode (2 × 100bp) for an output of 6 GB per sample, and an average coverage of 50X. Sequencing reads were demultiplexed using Illumina Genes 2022, 13, 1056 4 of 24 bcl2fastq (2.20) and adapter sequences were trimmed using Skewer (version 0.2.2) **[19].** The quality of the generated FASTQ files was analyzed with FastQC software (version 0.11.5; Illumina, San Diego, CA, USA). Variant Annotation and Filtration PhenoDB tool were used to annotate Vcf files using ANNOVAR **[20].** Variants were filtered based on the depth of coverage and minor allele frequencies (MAF) (less than 1% MAF) in large population databases, including dbSNP **[21],** 1000 Genomes Project **[22]**, and the Genome Aggregation Database (gnomAD v2.1.1) **[23]**.

Variant Segregation:

Sanger sequencing was used to confirm that prioritized variants segregated consistently among parents and available family members according to the predicted mode of inheritance. We designed primers targeting exons that harbor the filtered variants of interest using Primer3 tool **[24]**. PCR was carried out as previously described. Reactions were sequenced according to manufacturer's recommendation using the Big Dye Termination kit (Applied Biosystems, Waltham, MA, USA), and ABI Prism 3500 Genetic Analyzer (Applied Biosystems, Waltham, MA, USA). Variants were named based on Human Genome Variation Society nomenclature recommendations **[25]**. The standards of the American College of Medical Genetics and Genomics (ACMG) were used to classify the level of variant pathogenicity, i.e., pathogenic, likely pathogenic, variant of unknown significance (VUS), benign, or likely benign **[26]**.

Multi-scale computational analysis:

The possible biological effect of all missense variants was done using multi-scale computational analysis tools considering all probable pathogenicity relevant aspects. Multi scale computational analysis approach supports multiple entry variants for annotation and analysis permitting the closest true pathogenicity implication to be concluded. To explore the functional network among all variants harboring genes in this study, the functional enrichment and protein-protein enrichment analyses were carried out using Gene Mania and STRING servers.

All the described variants were annotated using chromosomal descriptions and were checked by LUMC mutalyzer v. 3.0.4 according to GRCH38 human genome assembly, and all mentioned genes were described according to HGVS nomenclature.

Results

Seven patients descending from six unrelated Egyptian families were studied. All patients were born via caesarian section with an uneventful prenatal course. All families were 1st first degree cousins except one who were 2nd second degree cousins. The detailed family pedigree is shown in **(Fig.1**) . Detailed patients' demographic data and anthropometric measurements are shown in **Table 1** . The core clinical presentation of our patients included dysmorphic features (2 patients), neurodevelopmental delay or regression (7 patients), gait abnormalities (3 patients), skeletal deformities (5 patients), visual impairment (7 patients) and seizures (5 patients). Additional clinical features and diagnostic workup are described in **Table 2**.

Clinical features and phenotyping

Patient 1 presented with global developmental delay (GDD), language impairment, seizures, progressive loss of vision with poor ocular fixation, progressive spastic quadriplegia and scoliosis. Maternal history of recurrent spontaneous abortion was recorded. The ancillary tests- ERG showed retinal dysfunction, EEG showed interictal generalized epileptiform activity and brain MRI showed cerebral and cerebellar atrophy and hypoplastic corpus callosum (**Fig.2- A, B)**.

Patient 2 presented with dysmorphic facial features including "synophrys, short philtrum, thick lips, micrognathia and low set ears", poor ocular fixation and oculomotor apraxia, delayed motor development, spasticity, generalized tonic-clonic seizures, atrial septal defect (ASD), scoliosis and bilateral hip dislocation. Family history of a similarly affected sister was reported. The ancillary tests- EEG was normal and brain MRI showed cerebellar vermis hypoplasia (**Fig.2- C, D)**.

Patient 3 presented with GDD, dysmorphic facial features "course facies, narrow forehead, thick eyebrows, broad bulbous nose, short philtrum, thick lips, and large ears", oculomotor apraxia and seizures. He had multiple skeletal deformities "archnodactyly, bilateral hyperextensibility of the interphalangeal joints, bilateral low inserted thumb, toes camptodactyly, bilateral prominent heel, severe scoliosis and joint contractures" and bilateral fungal infection of both feet (**Fig. 3)**. The ancillary tests- ERG showed retinal dysfunction, EEG was abnormal and brain MRI showed cerebellar atrophy (**Fig 2- E, F).**

Patient 4 presented with intellectual disability, progressive ataxia "started at the age of 5 years", intension tremors and oculomotor apraxia. He had also aortic regurgitation with thickened valve. The ancillary tests-ERG showed retinal dysfunction, VEP showed bilateral optic nerve dysfunction and brain MRI showed cerebral and cerebellar atrophy and hypoplastic corpus callosum **(Fig.2- G , H)**.

Patients 5 and 6 (siblings), both presented with progressive NDR, intellectual disability and seizures (photosensitive epilepsy). They were also noted to have progressive visual decline, poor visual hand motor coordination and abnormal clumsy gait. Patient 5 had focal to bilateral seizures with loss of awareness that were preceded by seeing different colors. His EEG showed a single burst of sharply contoured sharp waves over the right parasagittal area. Patient 6 had scoliosis and seizures "absence and generalized tonic and clonic" that were triggered by intense light. Her EEG showed generalized, fragmented as well as focal rhythmic epileptiform discharges arising mainly from the right hemisphere. Their seizures were initially well controlled on Levetiracetam but later on they showed a refractory course. Both patients had abnormal ERG/VEP as shown in **Table 2.** Both patients had cerebellar atrophy as shown in brain MRI findings in patient 5 (**Fig.2- I, J)**.

Patient 7 presented with GDD, speech difficulties and impaired cognition. He had feeding difficulties, strabismus and oculomotor apraxia. He had left hand preference and poor fine motor skills. He had never been able to walk independently and exhibited bilateral knee contractures and bilateral tightness of the Achilles tendon. The ancillary tests- ERG/VEP showed bilateral optic nerve dysfunction and brain MRI showed cerebellar atrophy (**Fig. 2- K).**

Genetic spectrum of NDs patients

Exome analyses of the seven studied patients descending from six unrelated Egyptian families identified six different disease causing variants in five genes; *TPP I (CLN2) (NM_000391.4), MFSD8* (*CLN7) (*NM_152778.2), *AGTPBP1 (*NM_001330701.2)*, APTX* (NM_175073.3) and *PCDHGC4 (*NM_018928.3) genes as displayed in **Table 2**, **3**. All of these variants were missense except one defined as splicing variant and according to ACMG classification criteria were classified as pathogenic variants. As well, the analysis confirmed that among these, three variants weren't found in dbSNP, 1000G, gnomAD exome, and ExAC databases or in our in-house database of 55 Egyptian exomes.

Sanger sequencing was performed to confirm that prioritized variants segregated consistently among parents and available family members according to the predicted mode of inheritance. The chromatograph for available patients who completed the follow up was described in **(Fig.3).**

Patient 1 had a homozygous splicing variant in *CLN2 (TPPI)* gene (C.1145+2T>G). This variant is predicted to disrupt the highly conserved donor splice site of exon 9. Together with the clinical information and biochemical results, it is classified as pathogenic (class1) according to the ACMG. The result is consistent with the genetic diagnosis of AR NCL type 2.

Patient 2 had a homozygous missense variant in the *PCDHGC4* gene (c.1463C>T; p. (Ala488Val). It is a likely pathogenic variant according to ACMG. Defects in *PCDHGC4* have been associated with NEDGS.

Patient 3 had a novel homozygous missense variant in *AGTPBP1* (c.2650A>C; P.(Thr884Pro) substituting theronine residue for proline at position 884. It is a likely pathogenic variant according to ACMG. Pathogenic variants in *AGTPBP1* gene are associated with AR CONDCA.

Patient 4 had a novel homozygous missense variant in *APTX* gene (c.635C>T; P. (Ala212Val) in exon 6 substituting alanine for valine at position 212. Pathogenic variant has previously been described as disease causing AOA1. It is classified as a variant of uncertain significance (class 3) according to ACMG.

Patients 5 and 6 are affected siblings, both of whom had a novel homozygous missense variant in *CLN7* (*MFSD8*) (c.638C>A; p.(Pro213Gln) substituting proline residue for glutamine at position 213. The homozygous state of this variant has been confirmed by parental targeted testing. It is classified as variant of uncertain significance (class 3) according to ACMG. Pathogenic variants in *MFSD8* gene are associated with AR NCL type 7.

Patient 7 had a homozygous missense variant in *AGTPBP1* gene (c.1534A>G ; p. (Thr512Ala) causing amino acid change from Thr to Ala at position 512. It is classified as a variant of uncertain significance (class 3) according to the recommendations of Centogene and ACMG. Pathogenic variants in *AGTPBP1* gene are associated with AR CONDCA.

Computational analysis implications

All the used computational tools and the corresponding implications were showed in **Table 3**. The functional enrichment analysis was showed in **(Fig. 5).**

Discussion

In this study, we described the detailed phenotypic, radiological and molecular findings of seven Egyptian patients presenting with NDs. All parents were 1st degree cousins, highlighting the impact of consanguineous marriage on the increased rate of AR genetic disorders reported in our country **[27]**.

Genetic diagnostic testing based on exome sequencing revealed three novel variants in"*CLN7 (*NM_152778.2), *AGTPBP1 (*NM_001330701.2)*, APTX* and (NM_175073.3)" genes and three variants have been previously reported in "*TPPI* (*CLN2) (*NM_000391.4), *AGTPBP1 (*NM_001330701.2) and *PCDHGC4 (*NM_018928.3)" genes. These genes are associated with AR NCL type 2, AR NCL type 7, AR CONDCA, AOA1 and NEDGS, respectively **[28]**. Three patients in our cohort were diagnosed with NCL; patient 1 (NCL type 2), patient 5 and 6 - siblings (NCL type 7). To date, 131 variants have been reported in the *CLN2* gene which is distributed over the whole protein structure. These include missense variants (63, 48%) followed by frameshift (21, 16%) and nonsense (17, 13%) variants. Two known common pathogenic variants, c.509–1 G>C and c.622 C>T p.(Arg208*) occur in 60% of affected individuals with NCL2 **[29].** WES analysis revealed a splice site mutation in exon 9 of *CLN2* gene in patient 1. Our patient presented at the age of 3 years with GDD that was followed by progressive visual impairment, motor disability, spasticity and scoliosis. NCL type 2 typically presents with seizures and/or ataxia in the late-infantile period by the age of 2–4 years, often in combination with a history of speech delay, followed by progressive childhood dementia, motor and visual impairment and early death **[30].** Our findings agree with previous reports that studied the clinical characteristics of NCL2 patients **[31, 32, 33]**. However, our patient had progressive spastic quadriplegia and scoliosis so there was a clinical overlap with other conditions such as hereditary spastic paraplegia **[34]**.

A novel (c.638C>A; p.Pro213Gln) in *CLN7* gene in a homozygous state was detected in patients 5 and 6. Stogmannet al. (2009) reported an Egyptian family with late-infantile seizures, deterioration and loss of psychomotor skills one year after the seizures' onset. Additionally, they had aggressive behavior, memory impairment, and language abnormalities with substantial loss of speech function**[35]**. In our study, we have observed marked intra familial disease variability as both patients (P5 and 6) had different age of onset. Also one of the two probands (P6) presented with scoliosis. Several previously reported studies showed inter and intrafamilial phenotypic variability for the same genotype in different forms of NCLs which agrees with our study **[36]**. The clinical heterogeneity may be related to the profoundly different disease mechanisms, the presence of modifier genes, other environmental factors or lifestyle. Modifier genes could influence gene expression levels and post-translational processing **[10]**. The current computational analysis showed that Pro213Gln is highly conserved with potential pathogenicity impact on the transportation function of CLN7 protein accumulating the diseased harmful compound.

Several NCL disease-specific therapies have been identified depending on the unique subtype identified. These therapies range from several options in the CLN2 subtype such as enzyme replacement therapies, gene therapies, stem cell therapies, and pharmacological drugs to no available options in the CLN7 subtype rendering the identification of each type of particular importance **[8]**.

Patient 2 was diagnosed with NEDGS due to a pathogenic homozygous missense variant in the *PCDHGC4* gene (c.1463C>T; p.Ala488Val). This syndrome was first described in a cohort of 19 patients from nine unrelated families originated from Iran, Iraq, Turkey, Morocco, Pakistan, Saudi Arabia, Lebanon, Sudan, and Jordan. It is characterized by the presence of dysmorphic features, neurodevelopmental delay, microcephaly, short stature, seizures, hypotonia, spasticity, strabismus, skeletal anomalies**[16] .**Our patient had similarpresentation apart from cardiac anomalies (ASD) that was not previously reported. Five nonsense, frameshift, or splice site mutations were predicted to result in premature termination and a loss of function, and 3 missense mutations at highly conserved residues were reported. To our knowledge, this is the 2nd study that detects pathogenic variant in *PCDHGC4* gene. Both patients 3 and 7 had novel homozygous mutations in *AGTPBP1*. A gene that was first described by Shashi et al., 2018**[14]**in 13 individuals from 10 unrelated families with abnormal eye movements, GDD, microcephaly, tongue fasciculation, hypotonia, muscle atrophy, feeding difficulties, failure to thrive . All patients had cerebellar atrophy, hence it was subsequently recognized as CONDCA**[37,38].** Another study reported similar phenotypic presentation but without any eye movement abnormalities **[13]**.

To our knowledge, this is the third clinical study reporting patients with CONDCA. Our patients had similar presentation but patient (3) had additional unique phenotypic features "dysmorphic features, seizures and skeletal deformities" that was not previously reported. Patient (7) had areflexia in the lower extremities due to axonal motor neuropathy as confirmed by nerve conduction and electromyography studies possibly due to purkinje cell degeneration and motor neuropathy that has been described in this disorder. Brain MRI in our patients showed cerebellar atrophy which is the hallmark of this disorder that should be carefully interpreted in view of their clinical presentation to avoid confusion with other disorders such as pontocerebellar diseases that are associated with cerebellar hypoplasia**[14]**. The current computational analysis showed that p.Gly884Arg might lead to altered protein conformation inhibiting deglutamylation of tubulin and non-tubulin target proteins.

APTX variant, identified in patient 4, was first described by Aicardi et al. (1988) **[39]** as the cause of a syndrome mimicking ataxia telangiectasia that was named AOA1. Anheim et al. **(2010)** found that AOA1 was the fourth most common cause of AR cerebellar ataxia **[40]**. The p.Pro206Leu mutation was the

most frequent mutation described worldwide **[41]**. Oculomotor apraxia was not a constant finding in all reported patients. Our result was in agreement with the previous studies regarding most of the clinical manifestations, however, our patient had congenital heart disease and poor retinal function that weren't previously reported in *APTX* gene mutations thus expanding its phenotypic spectrum.

To our knowledge, this is the first genetic study of its kind from North Africa "Egypt" exploring the possible molecular defects underlying the overlapping NDs phenotypes. We identified three novel pathogenic mutations and expanded the phenotypic spectrum with newly associated clinical phenotypic findings in the studied patients. In conclusion, this study highlights the importance of genetic testing for patients presenting with ND disorders where phenotypic characterization might not be sufficient for proper classification and early disease identification, especially for potentially treatable ones.

Declarations

Data Availability

The data supporting the findings of this study are available with the corresponding author upon request.

Acknowledgements

The authors would like to thank the patients and their parents for kindly participating in this study. Our gratitude was extended to the late **Professor Mona El Ruby**, our mentor, who passed away before the completion of this manuscript for her participation in the diagnosis and establishing the study design.

Author Contributions

- **Engy A. Ashaat and Nesma M. Elaraby**: designed and conceptualized the study.
- **Hoda A. Ahmed ,Engy A. Ashaat, Dina Amin Saleh**,: wrote the first draft of the manuscript.
- **Ammal M. Metwally ,Neveen A. Ashaat, Dina Amin Saleh:** Referred cases
- **Engy A. Ashaat, Rasha Moheb Elhossini, Heba Ahmed ElAwady, Dina Amin Saleh, Mohamed Ahmed Al Kersh**: examined the patient**,** took the family and medical histories.
- **Randa H. A. Abdelgawad:** performed Ophthalmological examination
- **Nesma M. Elaraby**, **Hoda A. Ahmed :** collected DNA samples, interpreted patient data, and performed Sanger sequencing data analysis.
- **Alaaeldin Fayez:** carried out the bioinformatics analysis.
- **Mona K. Mekkawy:** performed cytogenetic analysis
- **Mohamed Ahmed Al Kersh:** performed the radiological investigation.
- **Mona El Gammal:** revised and editing the final version of manuscript
- **All authors** have read and agreed to the published version of the manuscript.

Funding:

None

Ethics approval

The study was approved by the Medical Research Ethics Committee of the National Research Centre (NRC), Cairo, Egypt. Informed and written consent was obtained from the parents involved in the study.

Conflict of interest

The authors have no conflict of interest to declare.

References

1. Temtamy S, Aglan M (2012) Consanguinity and genetic disorders in Egypt. Middle East Journal of Medical Genetics. 1:12–17. doi: 10.1097/01.
2. Mastrangelo M (2019) Clinical approach to neurodegenerative disorders in childhood: an updated overview [published correction appears in Acta Neurol Belg. 2019 Aug 26. Acta Neurol Belg. 119(4):511–521. https://doi:10.1007/s13760-019-01160-0.
3. Johnston M (2016) Neurodegenerative disorders of childhood. In: Behrman RE, Kliegman RM, Jenson HB, (Eds.); 2016; Nelson textbook of pediatrics (20th edn.)
4. Jan MM (2002) Clinical approach to children with suspected neurodegenerative disorders. Neurosciences (Riyadh). 7:2–6.
5. Stephenson J, Nutma E, van der Valk P, Amor S (2018) Inflammation in CNS neurodegenerative diseases. Immunology.154, 204–219. https://doi:10.1111/imm.12922.
6. Mole SE, Anderson G, Band HA (2019) Clinical challenges and future therapeutic approaches for neuronal ceroid lipofuscinosis. Lancet Neurol. 18(1):107–116. https://doi:10.1016/S1474-4422(18)30368-5.
7. Kozina AA, Okuneva EG, Baryshnikova NV, Kondakova OB, Nikolaeva EA, Fedoniuk ID, et al.(2020) Neuronal ceroidlipofuscinosis in the Russian population: Two novel mutations and the prevalence of heterozygous carriers. Mol Genet Genomic Med; Jul;8(7). https://doi:10.1002/mgg3.1228.

8. Rosenberg JB, Chen A, Kaminsky SM, Crystal RG, Sondhi D (2019) Advances in the Treatment of Neuronal Ceroid Lipofuscinosis. Expert Opin Orphan Drugs. 7(11):473–500. https://doi:10.1080/21678707.2019.1684258.
9. Mole SE, Cotman SL(2015) Genetics of the neuronal ceroid lipofuscinoses (Batten disease).BiochimBiophysActa. Oct;1852(10 Pt B):2237–41.https://doi:10.1016/j.bbadis.2015.05.011.
10. Butz ES, Chandrachud U, Mole SE, Cotman SL (2020) Moving towards a new era of genomics in the neuronal ceroidlipofuscinoses. BiochimBiophysActaMol Basis Dis. 1;1866(9):165571. https://doi:10.1016/j.bbadis.2019.165571.
11. Nelvagal HR, Lange J, Takahashi K, Tarczyluk-Wells MA, Cooper JD (2020) Pathomechanisms in the neuronal ceroid lipofuscinoses. Biochim Biophys Acta Mol Basis Dis. 1866(9):165570. https://doi:10.1016/j.bbadis.2019.165570.
12. Baltanás FC, Berciano MT, Santos E, Lafarga M (2021) The Childhood-Onset Neurodegeneration with Cerebellar Atrophy (CONDCA) Disease Caused by *AGTPBP1* Gene Mutations: The Purkinje Cell Degeneration Mouse as an Animal Model for the Study of this Human Disease. Biomedicines. 4;9(9):1157. https://doi:10.3390/biomedicines9091157.
13. Karakaya M, Paketci C, Altmueller J, Thiele H, Hoelker I, Yis U, Wirth B (2019) Biallelic variant in AGTPBP1 causes infantile lower motor neuron degeneration and cerebellar atrophy. Am J Med Genet A. 179(8):1580–1584.https://doi:10.1002/ajmg.a.61198.
14. Shashi V, Magiera MM, Klein D, Zaki M, Schoch K, RudnikSchöneborn S et al (2018) Loss of tubulin deglutamylase CCP1 causes infantile-onset neurodegeneration. EMBO J. 3;37(23). https://doi:10.15252/embj.2018100540.
15. Mariani LL, Rivaud-Péchoux S, Charles P et al (2017) Comparing ataxias with oculomotor apraxia: a multimodal study of AOA1, AOA2 and AT focusing on video-oculography and alpha-fetoprotein. Sci Rep 7, 15284. https://doi:10.1038/s41598-017-15127-9.
16. Iqbal M, Maroofian R, Çavdarlı B, et al (2021) Biallelic variants in PCDHGC4 cause a novel neurodevelopmental syndrome with progressive microcephaly, seizures, and joint anomalies. Genet Med. 23, 2138–2149. https://doi:10.1038/s41436-021-01260-4.
17. Tanner JM, Hiernaux J, Jarman S (1969) Growth and physique studies. In: Human Biology, A Guide to Field Methods (Weiner JS, Lourie JA, eds). IBP Handbook No. 9.Oxford:Blackwell, International Biological Programme, 2–71.
18. Verma RS, Babu A(1995) Human Chromosomes: Principles and TechniquesMcGraw-Hill, New York. https://doi.org/10.1002/mrd.1996.1080430105
19. Jiang H, Lei R, Ding SW, Zhu S Skewer (2014) a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics. 15:182–None. https://doi:10.1186/1471-2105-15-182.
20. Wang K, Li M, Hakonarson H (2010) ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164. https://doi:10.1093/nar/gkq603.
21. Sherry ST, Ward MH, Kholodov M, Baker J, Phan, L, Smigielski EM, Sirotkin K(2001) dbSNP: The NCBI database of genetic variation. Nucleic Acids Res. 29, 308–311. https://doi:10.1093/nar/29.1.308.
22. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR., et al (2015) A global reference for human genetic variation. Nature 526, 68–74. https://doi:10.1038/nature15393.
23. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi, J, Wang Q C, et al (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. *581*, 434–443. https://doi:10.1038/s41586-020-2308-7.
24. Untergasser A I, Cutcutache T, Koressaar JYe, Faircloth, M. Remm, Rozen S G (2012) Primer3-New capabilities and interfaces. Nucleic Acids Research. 40: e115. https://doi.org/10.1093/nar/gks596.
25. Sen Dunnen JT, Dalgleish R, Maglott DR, Hart R K, Greenblatt MS, McGowan-Jordan J (2016) HGVS recommendations for the description of sequence variants: 2016 Update. Human Mutations. *37*(6), 564–569. https://doi.org/10.1002/humu.22981.
26. Richards S; Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al (2015) 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. 17, 405–423. https://doi:10.1038/gim.2015.30.
27. Rabah M Shawky, Mohammed Y El-Awady, Solaf M Elsayed, Ghada E Hamadan (2011) Consanguineous matings among Egyptian population. Egyptian Journal of Medical Human Genetics. 12(2) 157–163. https://doi.org/10.1016/j.ejmhg.2011.07.001.
28. Ren XT, Wang XH, Ding CH, Shen X, Zhang H, Zhang WH, et al (2019) Next-Generation Sequencing Analysis Reveals Novel Pathogenic Variants in Four Chinese Siblings With Late-Infantile Neuronal CeroidLipofuscinosis. Front Genet. Apr 25;10:370. https://doi:10.3389/fgene.2019.00370.
29. Gardner E, Bailey M, Schulz A, Aristorena M, Miller N, Mole SE. Mutation update: Review of TPP1 gene variants associated with neuronal ceroidlipofuscinosis CLN2 disease. Hum Mutat. 2019; Nov;40(11):1924–1938. https://doi:10.1002/humu.23860
30. Michael Fietz, Moeenaldeen AlSayed, Derek Burke, Jessica Cohen-Pfeffer, Jonathan D. Cooper, Lenka Dvořáková, et al (2016) Diagnosis of neuronal ceroid lipofuscinosis type 2 (CLN2 disease): Expert recommendations for early detection and laboratory diagnosis, Molecular Genetics and Metabolism, 119 (1–2), 160–167. https://doi.org/10.1016/j.ymgme.2016.07.011.
31. MoleS E, Gardner E, Schulz A, Xin WW (2018) Molecular basis of CLN2 disease: A review and classification of TPP1 gene variants reported worldwide. Molecular Genetics and Metabolism. 123(2), S97. https://doi:10.1016/j.ymgme.2017.12.255
32. Nickel, Miriam et al (2016) Natural History of CLN2 Disease: Quantitative Assessment of Disease Characteristics and Rate of Progression. Neuropediatrics. 47. https://doi:10.1055/s-0036-1583730
33. Nickel M, Simonati A, Jacoby D, Lezius S, Kilian D, Van de Graaf B, et al (2018) Disease characteristics and progression in patients with late-infantile neuronal ceroidlipofuscinosis type 2 (CLN2) disease: an observational cohort study. Lancet Child Adolesc Health. 2(8):582–590. https://doi:10.1016/S2352-4642(18)30179-2.

34. Kara E, Tucci A, Manzoni C, Lynch DS, Elpidorou M, Bettencourt C, Chelban V, et al (2016) Genetic and phenotypic characterization of complex hereditary spastic paraplegia. Brain. 139(Pt 7):1904–18. https://doi:10.1093/brain/aww111.
35. Stogmann E, El Tawil S, Wagenstaller J, Gaber, A, Edris, S, Abdelhady, A et al (2009) A novel mutation in the MFSD8 gene in late infantile neuronal ceroidlipofuscinosis. Neurogenetics 10: 73–77.https://doi:10.1007/s10048-008-0153-1.
36. Weisschuh N, Mazzola P, Heinrich T, Haack T, Wissinger B, Tonagel F, Kelbsch C (2020) First submicroscopic inversion of the OPA1 gene identified in dominant optic atrophy—A case report. BMC Med. Genet. 21:236. https://doi:10.1186/s12881-020-01166-z.
37. Sheffer, R., Gur, M., Brooks, R. et al (2019) Biallelic variants in AGTPBP1, involved in tubulin deglutamylation, are associated with cerebellar degeneration and motor neuropathy. Eur J Hum Genet 27, 1419–1426. https://doi.org/10.1038/s41431-019-0400-y.
38. Baltanás FC, Berciano MT, Santos E, Lafarga M (2021) The Childhood-Onset Neurodegeneration with Cerebellar Atrophy (CONDCA) Disease Caused by *AGTPBP1* Gene Mutations: The Purkinje Cell Degeneration Mouse as an Animal Model for the Study of this Human Disease. Biomedicines. 4;9(9):1157. https://doi:10.3390/biomedicines9091157.
39. Aicardi J, Barbosa C, Andermann E, Andermann F, Morcos R, Ghanem Q, et al (1988) Ataxia-ocular motor apraxia: a syndrome mimicking ataxia-telangiectasia. Ann Neurol. 24:497–502.; .: 10.1002/ana.410240404
40. Anheim M, Fleury M, Monga B, Laugel V, Chaigne D, Rodier G, Ginglinger E, et al (2010) Epidemiological, clinical, paraclinical and molecular study of a cohort of 102 patients affected with autosomal recessive progressive cerebellar ataxia from Alsace, Eastern France: implications for clinical management. Neurogenetics 11: 1–12. https://doi.10.1007/s10048-009-0196-y.
41. Ababneh, Nidaa A., Dema Ali, Ban Al-Kurdi, Malik Sallam, Abdulla M. Alzibdeh, Bareqa Salah et al(2020) Identification of APTX disease-causing mutation in two unrelated Jordanian families with cerebellar ataxia and sensitivity to DNA damaging agents. Plos one 15.8. https://doi.org/10.1371/journal.pone.0236808.

Tables

Table 1. Patient's demographic data, anthropometric measurements and similarly affected siblings

Patient	Gender	Age at onset (years)	Current age (years)	Parental consanguinity	Anthropometric measurements	Other affected family members
P1	Male	3	6	1st cousins	- OFC: appropriate - Wt: appropriate - Ht/Lt: appropriate	None
P2	Female	1	1.5	1st cousins	- OFC: microcephaly - Wt: underweight - Ht/Lt: appropriate	1 male & 1 female siblings
P3	Male	2	18	1st cousins	- OFC: appropriate - Wt: underweight - Ht/Lt: N/A (severe contracture deformities and scoliosis)	5 male & 2 female siblings
P4	Male	2	9	1st cousins	- OFC: appropriate - Wt: appropriate - Ht/Lt: appropriate	None
P5	Male	6	12	1st cousins	- OFC: appropriate - Wt: appropriate - Ht/Lt: appropriate	1 female sibling
P6	Female	7	13	1st cousins	-OFC: appropriate -Wt: appropriate -Ht/Lt: appropriate	1 male sibling
P7	male	1.5	5	2nd cousins	- OFC: appropriate - Wt: appropriate - Ht/Lt: appropriate	None

OFC; occipito-frontal circumference, Wt; weight, Ht; height, Lt; length.

Table 2. Patient's clinical features and diagnostic work up

Nucleotide Change Protein Change Gene	Brain MRI	EEG	ECHO	(VEP/ ERG)	Skeletal deformity	Eye manifestation	Neurological examination	Dysmorphic features	Sympto
C.1145+2T>G p.? TPP I(CLN2)	Cerebral & cerebellar atrophy Hyoplastic corpus callosum	Abnormal	Normal	Bilateral poor retinal function	Scoliosis	Poor eye contact and visual tracking	- Spastic quadriplegia - Brisk deep tendon reflexes	No	- GDD - progre Visual l - Seizur
c.1463C>T p. Ala488Val PCDHGC4	Cerebellar vermis hypoplasia	Normal	ASD	NI	Scoliosis Bilateral hip dislocation	oculomotor apraxia	Hypertonia/ brisk reflexes	- Synophrys Short philtrum, thick lips, micrognathia,Low set ears)	- GDD - Visual decline - Seizur - Dyspn exertio
c.2650A>C P.Thr884Pro AGTPBP1	Cerebellar atrophy	Abnormal	Normal	Bilateral poor retinal function	Scoliosis Contracture deformities	- Poor eye contact - Hypometric saccades - Oculomotor apraxia	- Spasticity - Brisk deep tendon reflexes - Contractures	- Course facial features (bulbous broad nose, thick lips, large ears) - Narrow forehead - Thick eyebrows	- GDD - Seizur -Intellec difficult - skelet deform
c.635C>T P.Ala212Val APTX	Cerebral & cerebellar atrophy Hypoplastic corpus callosum	Not done	CHD (Aortic and tricuspid valve regurgitation)	Bilateral optic nerve dysfunction	No	-Hypometric saccades -Bilateral myopic astigmatism Oculomotor apraxia	- Progressive ataxia - Intention tremors - Spasticity of lower extremities - Brisk deep tendon reflexes	No	- Progre NDR - Intelle disabili - Jerky movem - Abnor gait - Dyspn exertio
c.638C>A p. Pro213Gln MFSD8	Cerebellar atrophy	Abnormal	NI	Retinal dystrophy (rod/cone dystrophy) Optic nerve dysfunction	No	Poor vision	- Poor fine motor skills - Clumsy gait - Poor coordination (finger to nose test)	No	-Intellec difficult - NDR - Photos epileps - Abnor gait - Visual decline
c.638C>A p. Pro213Gln MFSD8	Cerebellar atrophy -	Abnormal	NI	Retinal dystrophy Optic nerve dysfunction	Scoliosis	Poor vision	- Poor fine motor skills - Clumsy gait - Poor coordination (finger to nose test)	No	- Intelle disabili - NDR - Photos epileps - Visual decline

									-Abnorr & postu
c.1534A>G p.Thr512Ala *AGTPBP1*	Cerebellar atrophy	NI	Normal	Bilateral optic nerve dysfunction ERG N/A	Knee & Achilles tendon contractures - Bilateral tightness of the Achilles tendon	- Strabismus (wearing eye glasses) - Oculomotor apraxia - Eye lid twitching	- Wheel chair bound (GMFCS level IV - Truncal hypotonia - Weak hand grip and poor fine motor skills - Weakness and spastic lower limbs - Wasting of leg muscles - Arefkexia in lower extremities	No	- Intelle disabili - GDD - Feedir difficult

AOA1; Ataxia-ocular apraxia type 1, CONDCA; childhood-onset neurodegeneration with cerebellar atrophy, NCL; neuronal ceroid lipofuscinosis type 2

, NEDGS; Neurodevelopmental disorder with poor growth and skeletal anomalies.

GDD; global developmental delay, NDR; neurodevelopmental regression, CHD; Congenital heart disease, ASD; atrial septal defect, NI; not indicated.

Table 3: Multi-scale computational analysis results

** All the described variants were annotated using chromosomal descriptions and were checked by LUMC mutalyzer v. 3.0.4 according to GRCH38 human genome assembly.

* Prediction score using SVM and sequence information

Figures

Tool	Score	Reference range	Implications
NC_000004.12(NM_001371596.2):c.638C>A*	**(*MFSD8* gene; p.Pro213Gln; Novel variant)**		
CADD PHRED	24.7	Pathogenic variants with PHRED score < 30	Pro213Gln (P213Q) variant is highly conserved and it falls into the active transporter motif (PF00083), whose sensitivity to deleterious variants is reasonably high. Protein structure and clinical relevance-based prediction scores showed that Pro213Gln led to significant alteration in the protein function with disease-relevant impact. **Finally, the predicted mechanism of action is that Pro213Gln might lead to transporter dysfunction for accumulating harmful compounds.**
SIFT	Deleterious (0.03)	A SIFT score of less than 0.05 is predicted to be deleterious.	
PolyPhen	HumDiv= 0.994 with Sensitivity (TPR)= 0.69; specificity (TNR): 0.97 HumVar=0.822 with Sensitivity (TPR)= 0.74; specificity (TNR): 0.88	(0.0 to 0.15) predicted to be benign. (0.15 to 1.0) possibly damaging. (0.85 to 1.0) more confidently predicted to be damaging.	
BLOSUM62	-1 (Conserved)	Typically +1 or +5 for matches, and -1 or -4 for mismatches	
phastCons30way mammalian including human	0.992 Highly conserved	Ranged from 0-1; higher score means highly conserved	
ClinPred (score)	Deleterious (0.982)	Ranged from 0-1; higher score is likely to be pathogenic	
DEOGEN2 (score)	Deleterious (0.699) representing PFAM log-odd score (PF) = 26.7, and LOR (Log-Odd Ratio) = 22%. Similarity between Prolin (Pro) and Glutamone (Gln) = 49.5%, where the charge similarity is lesser =7.14% of 49.5%)	Ranged from 0-1; higher score is likely to be pathogenic with cutoff=0.5	
MutPred score; based on Q8NHS3-1_human.	0.841 Altered Transmembrane protein (pvalue = 2.1e-04)	It was estimated that a pathogenicity threshold of 0.68 yields a false positive rate of 10% and that of 0.80 yields a false positive rate of 5%.	
PROVEAN score	-3.12	Ranged from -14 to 14, the smaller score is likely has damaging effect. Cutoff for PROVEAN scores was set **to −2.5** for high balanced accuracy	
VEST4 score	0.937	Range from 0 to 1, the highest score causes functional change	
MUpro**	ΔΔG = -0.626 (DECREASE stability with high confidence)	Above zero is increase stability, while less zero is decrease stability	

BayesDel	Deleterious (Supporting) = 0.23	Ranged from -1.29334 to 0.75731. cutoff between deleterious (D) and tolerated (T) is -0.0570105. Based on ClinGen's recommendation, score above 0.5 corresponds with Strong pathogenic evidence, score in the range (0.27, 0.5) Moderate pathognic, score in the range (0.13, 0.27) Supporting pathognic , score in the range (-0.36 ,-0.18) Supporting benign, and score below-0.18 Moderate benign.	
GenoCanyon	Deleterious (1)	Range from 0 to 1, the highest score is deleterious	
fitCons	Deleterious (0.7)	Range from 0 to 1, the highest score score is deleterious	
UniProt mining	Subcellular location; Helical transmembrane (212-232)		
MotifFinder (In Pfam database)	MFS_1 (42-359) **PF07690, Major Facilitator Superfamily** E-value 1.1e-06 Sugar_tr (71-229) **PF00083, Sugar (and other) transporter** E-value 9.5e-11)		

NC_000005.10(NM_018928.3):c.1463C>T (*PCDHGC4*; p.Ala488Val; Reported variant; rs775104626) ClinVar; pathogenic, and citation = 0

CADD PHRED	24.4	Pathogenic variants with PHRED score < 30	Ala488Val (A488V) variant is highly conserved and it falls into the active domain called Codherin domain (PRU00043) that is responsible for cell-cell adhesion. Therefore, sensitivity to deleterious variants is reasonably high. Clinical relevance-based prediction score showed that Ala488Val has disease-relevant impact. Finally, the predicted mechanism of action is that Ala488Val might contribute in cell adhesion dysfunction.
SIFT	0	A SIFT score of less than 0.05 is predicted to be deleterious.	
PolyPhen	HumDiv=0.99 with sensitivity (TPR)= 0.14; specificity (TNR)= 0.99 HumVar= 0.97 with sensitivity (TPR)= 0.60; specificity (TNR)= 0.93	(0.0 to 0.15) predicted to be benign. (0.15 to 1.0) possibly damaging. (0.85 to 1.0) more confidently predicted to be damaging.	
phastCons30way mammalian including human	0.974 Highly conserved	Ranged from 0-1; higher score means highly conserved	
ClinPred (score)	D (0.974)	Ranged from 0-1; higher score is likely to be pathogenic	
PROVEAN score	-3.55	Ranged from -14 to 14, the smaller score is likely has damaging effect. Cutoff for PROVEAN scores was set **to −2.5** for high balanced accuracy	
MUpro*	ΔΔG = -0.4 (DECREASE stability, with moderate confidence)	Above zero is increase stability, while less zero is decrease stability	
UniProt mining	Located in Codherin 5 domain PRU00043		

NC_000009.12(NM_001195248.2):c.635G>T (*APTX*; p.Ser212Ile; Novel variant)

CADD PHRED	23.5	Pathogenic variants with PHRED score < 30	Ser212Ile (S212I) variant is highly conserved and it falls into the active Histidine triad (HIT) domain, whose

SIFT	0	A SIFT score of less than 0.05 is predicted to be deleterious.
PolyPhen	HumDiv= 0.880 with Sensitivity (TPR)= 0.68; specificity (TNR): 0.95 HumVar=0.878 with Sensitivity (TPR)= 0.73; specificity (TNR): 0.87	(0.0 to 0.15) predicted to be benign. (0.15 to 1.0) possibly damaging. (0.85 to 1.0) more confidently predicted to be damaging.
phastCons30way mammalian including human	0.986 Highly conserved	Ranged from 0-1; higher score means highly conserved
BLOSUM62	-2 (Conserved)	Typically +1 or +5 for matches, and -1 or -4 for mismatches
FATHMM (score)	D (-4.04)	
ClinPred (score)	D (0.982)	Ranged from 0-1; higher score is likely to be pathogenic
DEOGEN2 (score)	Deleterious (0.91) representing PFAM log-odd score (PF) = 27.4, and LOR (Log-Odd Ratio) = 19.7%. SNPMuSiC score= 0.43 Hydrophobicity; Ser= -0.80 / Ile= 4.50 Residue Solvent accessibility; Ser= 227.8 / Ile= 272.6 Side chain Solvent accessibility; Ser= 86.34 / Ile= 140.9	Ranged from 0-1; higher score is likely to be pathogenic with cutoff=0.5 The obtained SNPMuSiC score indicates that it is deleterious the solvent accessibility of the wild-type residue is 2.62% which means that the residue is buried and mutant residue is exposed
Hydrogen bonding capacity	(For Ser); with LYS214 and ALA215 (For Ile); with ALA215	
PROVEAN score	-4.89	Ranged from -14 to 14, the smaller score is likely has damaging effect. Cutoff for PROVEAN scores was set to **−2.5** for high balanced accuracy
VEST4 score	0.616	Range from 0 to 1, the highest score causes functional change
Revel	Deleterious (Moderate) (0.83)	Based on ClinGen's recommendation, score above 0.932 corresponds with Strong pathogenic evidence, score in the range (0.773, 0.932) Moderate pathognic, score in the range (0.644, 0.773) Supporting pathognic , score in the range (0.183 ,0.29) Supporting benign, score in the range (0.016, 0.183) Moderate benign, score in the range (0.003, 0.016) Strong benign, and score below 0.003 Very Strong benign.
Varity	Deleterious (0.94)	Score ranges between 0 to 1. Scores less than 0.1 have probability of 0.9 being benign, score of 0.35 has about an equal probability for being benign or deleterious (VUS), score of 0.5 hs a probability of 0.7 of being deleterious, and score of 0.55 a probability of 0.8 being deleterious.
UniProt mining	HIT (Histidine triad) domain (182-287) PROSITE; PS51084; HIT_2	
BayesDel	(Supporting) (0.25)	The range of the score is from -1.29334 to 0.75731. The higher the score, the more likely the variant is

sensitivity to deleterious variants is reasonably high.

HIT domain serve as catabolic enzymes acting on nucleotide-containing substrates, where it is part of binding loop for the alpha-phosphate of purine mononucleotide

Hydrophobicity and solvent accessibility are most chemical properties contributors in Ser212Ile pathogenicity.

Finally, the predicted mechanism of action is that Ser212Ile might lead to binding dysfunction with alpha-phosphate containing purine mononucleotide.

Deleterious		pathogenic	
fitCons	Deleterious (0.71)	Ranged from 0 to 1, the highest score score is deleterious	
MetaLR	D (0.9313)	Ranged from 0 to 1, higher score is likely to be pathogenic	
NC_000009.12(NM_001330701.2):c.2650G>C (*AGTPBP1*; p.Gly884Arg; Novel variant) **REPORTED PREVIOUSLY AS (NM_001330701.2:c.2650G>A; p.Gly884Arg) with ClinVar= NA, and citation=0**			
CADD PHRED	23.5	Pathogenic variants with PHRED score < 30	Gly884Arg (G884R) variant is highly conserved and it falls into crucial loop region for protein structure stabilization. Therefore, sensitivity to deleterious variants is reasonably high. Protein structure and clinical relevance-based prediction scores showed that Gly884Arg led to significant alteration in the bonding capacity with disease-relevant impact. **Finally, the predicted mechanism of action is that Gly884Arg might lead to altered protein conformation inhibiting deglutamylation of tubulin and non-tubulin target proteins.**
SIFT	0	A SIFT score of less than 0.05 is predicted to be deleterious.	
PolyPhen	HumDiv= 1 with Sensitivity (TPR)= 0.80; specificity (TNR): 0.98 HumVar=0.998 with Sensitivity (TPR)= 0.81; specificity (TNR): 0.95	(0.0 to 0.15) predicted to be benign. (0.15 to 1.0) possibly damaging. (0.85 to 1.0) more confidently predicted to be damaging.	
phastCons30way mammalian including human	0.999 Highly conserved	Ranged from 0-1; higher score means highly conserved	
BLOSUM62	-2 (Conserved)	Typically +1 or +5 for matches, and -1 or -4 for mismatches	
ClinPred (score)	D (0.998)	Ranged from 0-1; higher score is likely to be pathogenic	
MutPred score; based on Q9UPW5-1 human	0.787 Loss of Loop (P = 0.05); Altered Transmembrane protein (P = 8.3e-03)	It was estimated that a pathogenicity threshold of 0.68 yields a false positive rate of 10% and that of 0.80 yields a false positive rate of 5%.	
Hydrogen bonding capacity	(For Arg); with LYS638 (For Ile); with LYS638		
Hydrophobic bonding capacity	(For Arg); with LYS638 (For Ile); Null		
PROVEAN score	-6.75	Ranged from -14 to 14, the smaller score is likely has damaging effect. Cutoff for PROVEAN scores was set **to −2.5** for high balanced accuracy	
VEST4 score	0.862	Range from 0 to 1, the highest score causes functional change	
Varity	Deleterious (0.84)	Range from 0 to 1, the highest score is deleterious	
GenoCanyon	Deleterious (1)	Range from 0 to 1, the highest score is deleterious	
fitCons	Deleterious (0.71)	Range from 0 to 1, the highest score is deleterious	
NC_000011.10(NM_000391.4: c.1145+2T>G (*TPP1*; c.1145+2T>G; Reported variant; COSV100196937) ; splice_donor_variant (Int. 9)			

MaxEntScan	Ref= 10.858 Alt=**3.211** Diff=7.647	to diminish splicing (diff > 0) high (alt < 6.2), moderate (6.2 alt 8.5) or low (alt > 8.5) potential of disrupting native splice sites	**This variant has reasonable potential effect to diminish the splicing donor site to be skipped, thus alternative non-functional isoforms might be produced**
BayesDel	Deleterious (Strong) (0.66)	The range of the score is from -1.29334 to 0.75731. The higher the score, the more likely the variant is pathogenic	
SpliceAI	Splice-Altering = 0.99	ranged from 0 to 1 and can be interpreted as the probability of the variant being splice-altering	
dbscSNV Ada	Deleterious (1)	The score can range from 0 to 1, when higher values are more likely of being deleterious.	
dbscSNV RF	Deleterious (0.92)	The score can range from 0 to 1, when higher values are more likely of being deleterious.	
GenoCanyon	Deleterious (1)	The score can range from 0 to 1, when higher values are more likely of being deleterious.	
NC_000009.12(NM_001330701.2):c.1378A>G (*AGTPBP1*; p.Thr460Ala; reported variant rs1375829417) ClinVar; NA, and citation= 0			
CADD PHRED	24.2	Pathogenic variants with PHRED score < 30	Thr460Ala (T460A) variant is highly conserved with decreasing impact on the protein stability. However, it isn't falls into known motifs or domains. Therefore, sensitivity to deleterious variants is reasonably accepted. **Finally, the predicted mechanism of action is that Thr460Ala might lead to altered the protein stability. Hence, the degradable effect is predicted.**
SIFT	0.03	A SIFT score of less than 0.05 is predicted to be deleterious.	
PolyPhen	HumDiv= 0.842 with Sensitivity (TPR)= 0.83; specificity (TNR): 0.93	(0.0 to 0.15) predicted to be benign. (0.15 to 1.0) possibly damaging. (0.85 to 1.0) more confidently predicted to be damaging.	
phastCons30way mammalian including human	0.962 Highly conserved	Ranged from 0-1; higher score means highly conserved	
MutPred score; based on Q9UPW5-1 human	0.372 Conserved	It was estimated that a pathogenicity threshold of 0.68 yields a false positive rate of 10% and that of 0.80 yields a false positive rate of 5%.	
GenoCanyon	Deleterious (1)	Range from 0 to 1, the highest score is deleterious	
fitCons	Deleterious (0.73)	Range from 0 to 1, the highest score is deleterious	
MUpro*	ΔΔG = -0.57 (DECREASE stability with high confidence)	Above zero is increase stability, while less zero is decrease stability	

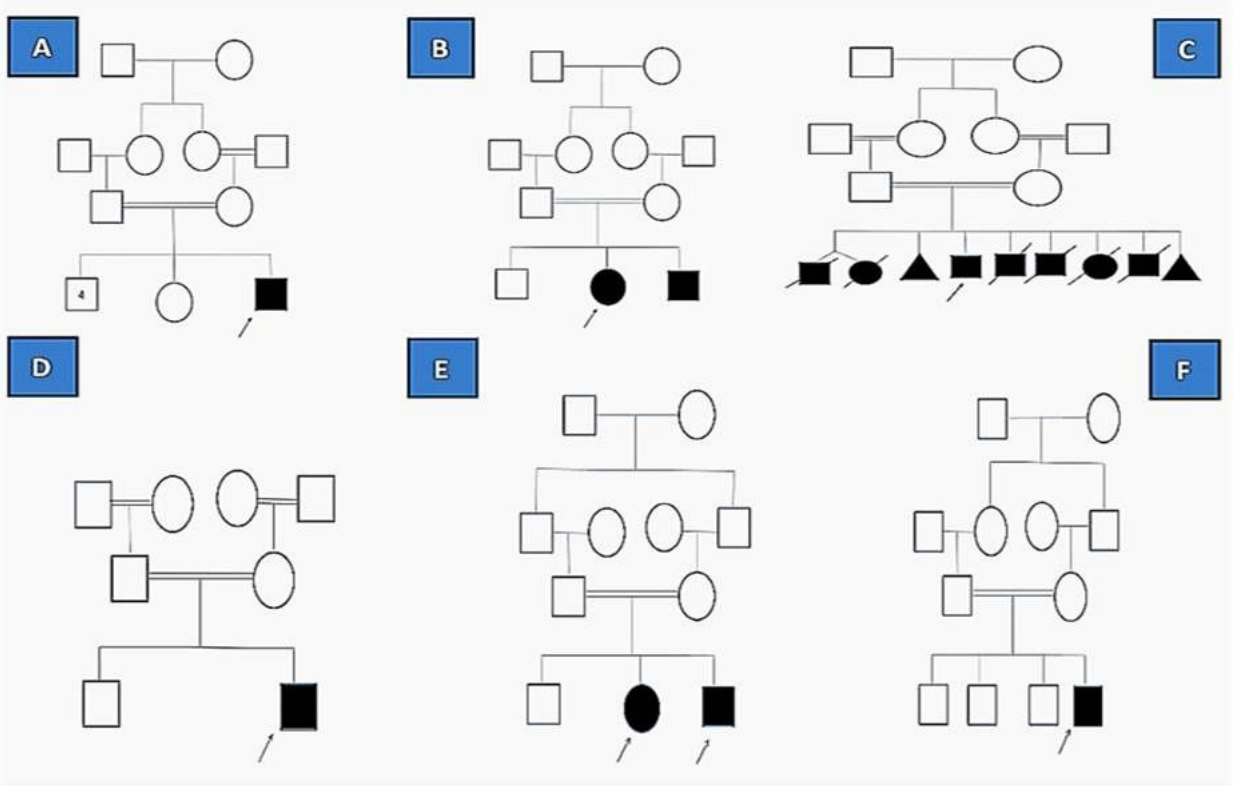

Figure 1

Family pedigrees of the six studied unrelated families.

A:P1 , B:P2, C:P3, D:P4, E:P5&P6 , F:P7

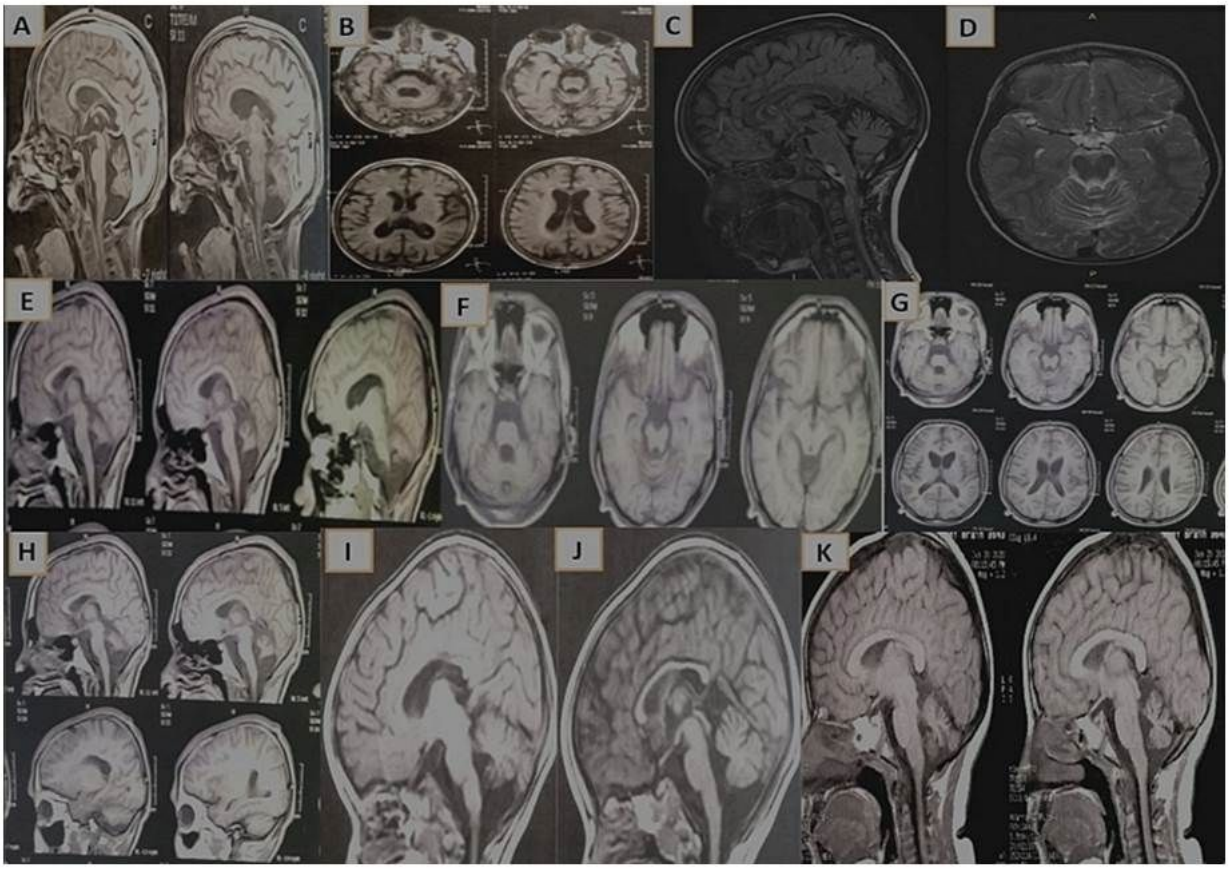

Figure 2

Brain MRI findings of the studied patients. P1 (A, B) cerebral and cerebellar atrophy and hypoplastic corpus callosum; P2 (C, D) cerebellar vermis hypoplasia; P3 (E, F) cerebellar atrophy; P4 (G, H) cerebral and cerebellar atrophy and hypoplastic corpus callosum, P5 (I, J) cerebellar atrophy; P7 (K) cerebellar atrophy

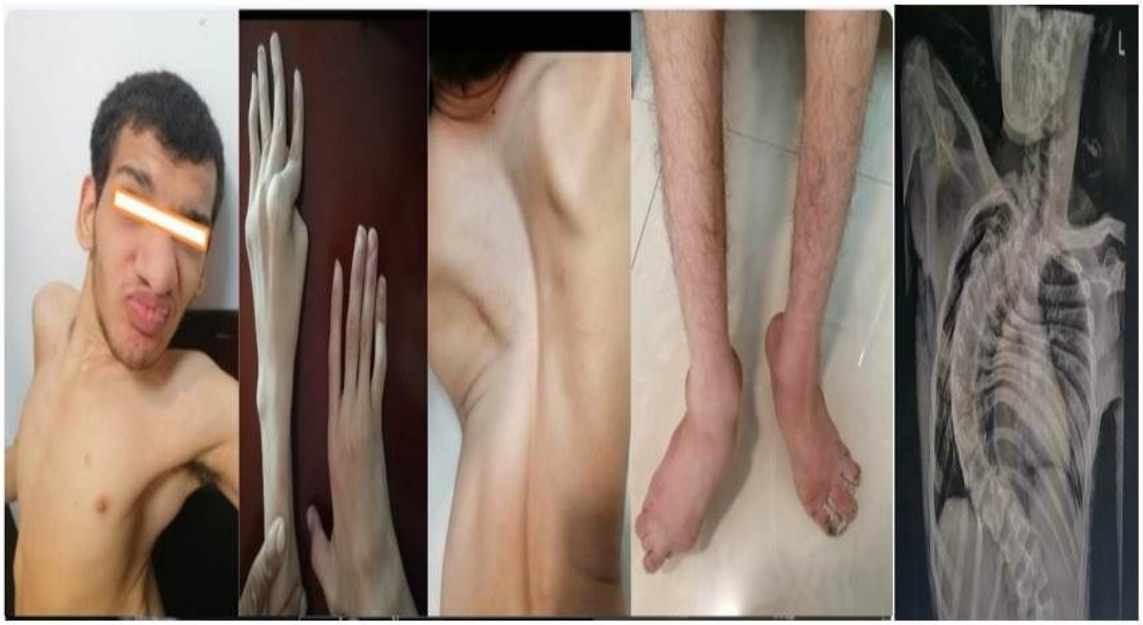

Figure 3

Patient 3- phenotypic dysmorphic features, scoliosis, skeletal deformities and fungal feet infection. X-ray chest PA view showing severe scoliosis and ribs crowding.

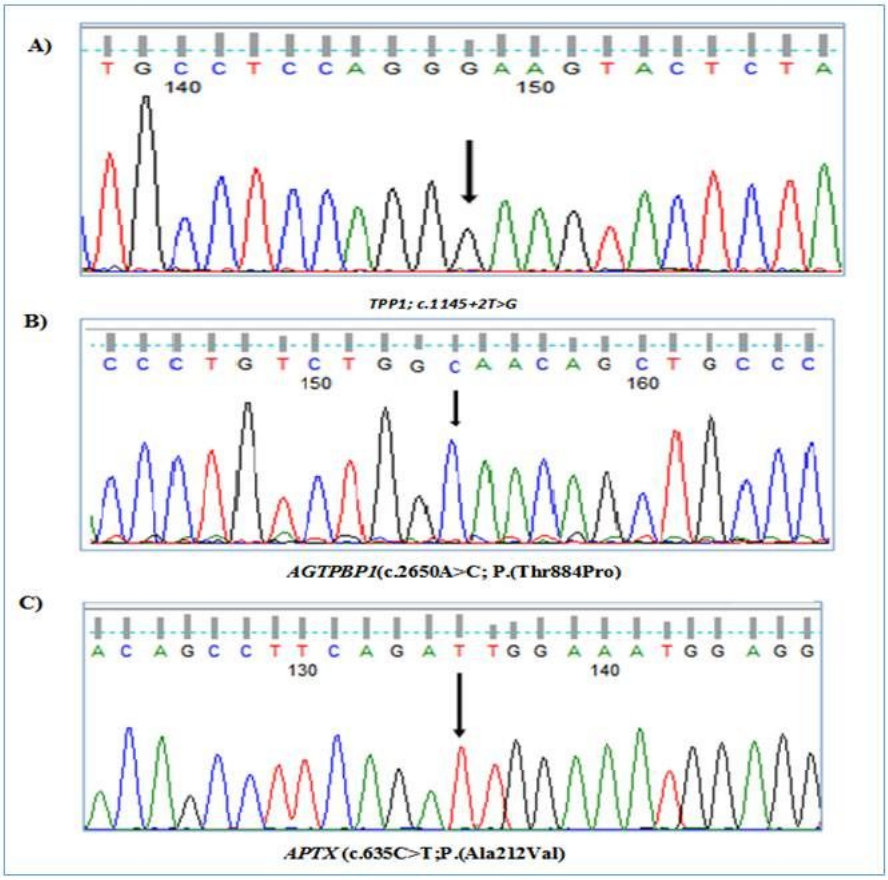

Figure 4

Sanger sequencing chromatograms of three patients (P1, P3, P4).

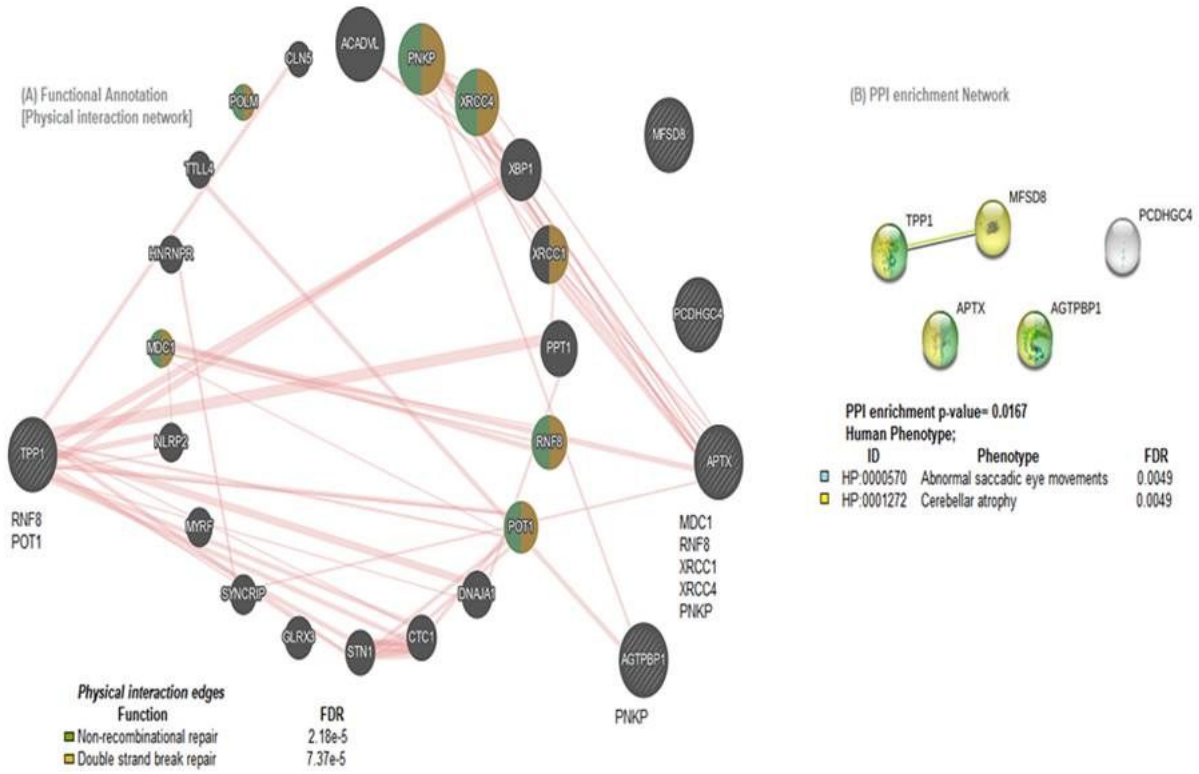

Figure 5

Functional Enrichment analysis representing (A) GeneMANIA results showed the significant functional relationship among the affected gene with false detective rate (FDR) scores, below each gene listed the interacted partner genes physically. The genes with dark green shadow represent set of genes which is shared the non-recombination repair process, while the genes with brown shadow represent set of genes which is shared the double strand break repair process. (B) PPI enrichment network among the affected genes showed that two human phenotypes were significantly enriched in 4 out of 5 analyzed proteins, the highlighted color was corresponded with each phenotype.