Novel Hemizygous Missense Variation in *AFF2* Gene Underlies Fragile XE Syndrome

Iftikhar Ahmed (✉ iftikhar.phdbt60@iiu.edu.pk)

International Islamic University <https://orcid.org/0000-0002-6015-2829>

Gaurav V Harlalka

University of Exeter Medical School

Muhammad Ilyas

International Islamic University Islamabad

Asif Mir

International Islamic University Islamabad

Research article

Keywords: Intellectual Disability, *AFF2* gene, hemizygous mutation, WES, neurological Disorders, Microarray SNP Genotyping

Posted Date: October 5th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-42138/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Background: Fragile XE (FRAXE) is an X-linked recessive condition of intellectual disability affecting 1 in 50,000 new born male. FRAXE is characterized by mild ID, cognitive impairment, speech delay and some cases patients display Autism Spectrum disorder (ASD) like phenotypes. .

Method: In this study, we investigated a family with two male siblings with neurodevelopmental delay Whole exome sequencing analysis (WES) was employed to identify the pathogenic variant. Co-segregation analysis was performed through Sanger sequencing in affected and normal family members.

Results: Two affected Proband of family were diagnosed with intellectual disability. A novel hemizygous variant, c.3348G>T; p.Asp1150Tyr, in *AFF2* gene was identified as the pathogenic cause in affected individuals. It is first novel variant report in *AFF2* gene within Pakistani population.

Conclusion: In this study, novel hemizygous variant, c.3348G>T; p.Asp1150Tyr, in *AFF2* gene was identified. The findings broaden the clinical and genetic spectrum of rare X-linked recessive disorders causing ID.

Background

Cognition is the result of multiple molecular, cellular and biological events in nervous system, and a minor defect in any of the event ultimately leads to cognitive impairment or intellectual disability (ID) [1, 2]. Moreover, it is also believed that developmental delay or intellectual disability is consequence of a defect in synapse formation and plasticity; therefore it is also termed as neurodevelopmental disorder [3, 4].

Prevalence of intellectual disability (ID) in general population is 2–3% [5, 6]. Clinical and genetic heterogeneity of ID pose a great challenge for clinicians and scientists involved in this research [7, 8]. Intellectual disability is linked with about 900 genetic disorders [9, 10].

The syndromic intellectual disability has several types e.g. Joubert syndrome, Kahrizi syndrome, and the majority of them have limited knowledge. So far, non-syndromic intellectual disability (NSID) is least studied and hence only 47 loci and 26 genes have been identified. Comparison between X linked genes and autosomes (chromosomes other than sex chromosomes), there is a large number of genes that control brain and its function are still unidentified. Even in case of non-syndromic intellectual disability, still twenty one loci are identified and their causative genes are remained to be recognized. FRAXE syndrome is a condition in which cognitive function, learning and thinking ability has been affected. Poor writing skills, delayed speech, hyperactivity and short attention span are the common symptoms of people affected with FRAXE syndrome. Autistic behaviour such as intense interest in a particular subject, repetitive behavior, hand flapping are displayed by some affected individuals. In FRAXE syndrome unlike in some other types of ID, cognitive function remains steady and did not decrease with age. [11].

Methods

The study was approved by the Board of Advanced Study Research (BASR) and ethical committee of the International Islamic University Islamabad. DNA was isolated from the collected blood samples using standard protocol. Microarray genotyping was performed using Illumina cytoSNP-12v2.0 arrays in both probands to identify the known genes in family.

Whole Exome Sequence Analysis

Whole exome sequencing was performed commercially on Illumina platform (Macrogen, Seoul, South Korea). WES was performed on one sample in family by using the Agilent SureSelect human all exon V4 exome enrichment kit. Sequence reads aligned to the reference genome UCSC: hg19/ GRCH37 NCBI. GATK tool was used to perform the sequence realignment and base quality recalibration. Read depth of the sequences were determined by using BedTools method [12]. Single nucleotide variations (SNVs), insertion and deletions were detected by using SamTool software [13]. Data annotation was performed with WANNONVAR [14]. Allele specific primer was designed to amplify the target region of DNA using primer3plus tool [15].

Sanger Sequencing

Polymerase chain reaction (PCR) was performed using genomic DNA of the proband and accessible family members for the confirmation of the veracity of the likely causative variant and to assess segregation within the family. The PCR product was purified using Exo-Sap at 37°C for 15 minutes, followed by heating temperature 80°C for 15 minutes. Sequencing reaction was performed using BigDye terminator kit (Applied Biosystems). The purified sequencing reaction product sequencing was performed on ABI 3730xl. Sequences were manually reviewed and compared to reference sequence NM 004595.4 of SMS gene using Codon Code Aligner software.

In-silico Analysis

To predict secondary and three dimensional protein models, Psipred [16] and I-Tasser [17, 18] were used. Predicted models reliability was checked using RAMPAGE [19]. Models were visualized via UCSF Chimera [20]. CASTp server were used to identified proteins pockets on 3D-structure of *AFF2* protein [21, 22]. Meta SNP [23], I-mutant 2.0 [24] and PREDICT SNP [25] used to check the impact of mutation on protein stability. Multiple in-silico tools were used to check the pathogenicity of the mutation. Multiple sequence alignment was performed using Clustal Omega.

Results

Clinical details

The family comprise of six members with two Proband i.e. IV: 2 and IV: 3 (Fig. 1A).

Patient Iv: 2

The Proband (IV: 2) was an 8 year and 3 month old male. He was born at normal gestational age and his Birth weight, birth length and OFC was 2.2 kg, 41.6 cm and 31 cm, respectively at birth. Developmental mile stones were observed at early age. At the age of 14 months, he started walking but unable to speak single word. Proband exhibited clinical manifestations of intellectual disability, memory loss, as well as learning disability. In addition, dysmorphic facial features including prominent nasal tip, large ears was noted.

Patient Iv: 3

The second Proband i.e. IV: 3 was 6 years old male at the time of study. He was born after an uneventful pregnancy and his weight and OFC were 3 kg and 30 cm respectively at birth. Clinical symptoms of both Proband were similar.

Genetic Analysis

The variant NM_002025.4: c.3448G > T; p.Asp1150Tyr; Chr X: 148059506G > T (GRCh37) was identified in the *AFF2* gene (MIM 309548) in the index patient (IV: 2), through whole-exome sequencing analysis. WES raw data was filtered through bioinformatics tools with minor allelic frequency (MAF) less (1%). This variant was then confirmed by Sanger sequencing. Co-segregation analysis of candidate variant in *AFF2* was performed in both affected and normal members. The pathogenic variant, c.3448G > T; p.Asp1150Tyr in *AFF2* was found associated with disease phenotype. Two affected males (IV: 2 and IV: 3) were hemizygous for mutant allele in coding exon 17 of *AFF2* gene where the normal mother III: 1 was heterozygous for the allele and the remaining members IV: 1 and IV: 5 were of wild type (Fig. 1B). This variant leads to a substitution of aspartic acid by tyrosine and at the evolutionary conserved amino acid position (p.Asp1150Tyr) according to the UCSC Human Genome (GRCh37/hg19). The variant is listed in the Genome Aggregation Database (gnomAD v2.1.1) (<https://gnomad.broadinstitute.org/>) in heterozygous form in two females from European (non-Finnish) population and is not reported in males. The variant is absent in south Asian population of gnomAD database.

The pathogenicity of the variant predicated to damaging by FATHMM-MKL with score of (0.88), damaging by DEOGEN2, damaging by PROVEAN with score (-2.75), DANN and GERP score were (0.99), (5.98) respectively. In-silico results strongly indicate that variant (p.Asp1150Tyr) is likely to be disease causing and deleterious to the *AFF2* protein and responsible for disease in family.

Effect Of Mutation On Protein

To study the pathogenicity of p.Asp1150Tyr mutation, structure of *AFF2* protein was predicted by using I-TASSER Tool. The mutated model created by using template 6R80 [26] (Fig. 1C). The result showed that template trigger release of polymerase-II enzyme from promoter proximal pausing site. Paused polymerase-II was playing role in the maintenances of promoter region structure. Affected model shows loss of structure of protein but no change in protein length observed. Loss of function of protein involved in the altering conformation of protein, it also shows that loss of c-terminal helices disturb the function of paused polymerase-II. Therefore the reported mutation p.Asp1150tyr is probably pathogenic for this family.

Multiple sequence alignment revealed conservation of mutated amino acid (p.Asp1150Try) and nearby amino acids in 7 different species. Mutation site denoted with red arrow (Fig. 1D).

Discussion

Fragile X E (FRAXE) is an X-linked syndrome characterized by cognitive impairment, Speech delay and autistic behavior. *AFF2* gene is reported to be involved in the syndrome. The *AFF2* gene is responsible for making a protein which is present in the cell's nucleus but its function is not well known. It has been reported that this protein serves as a transcription factor, which binds to specific regions of DNA and helps to control the activity of some other genes, though these genes are unknown [27]. Current investigation involves a family diagnosed with ID in two members and resulted in identification of novel missense mutation, c.3448G > T; p.Asp1150Tyr, in the coding exon 17 of *AFF2* gene.

Three dimensional protein models revealed conformation changes of amino acids due to mutation. Point mutation altered *AFF2* protein structure although it was a single base change. Figure 1C depicted mutational effect on three dimensional structure of protein. Mutant Protein model showed Loss of function of protein involved in the altering conformation of protein, it also shows that loss of c-terminal helices disturb the function of paused polymerase-II.

Conclusion

Genetic Analysis of the family with FRAXE syndrome identified a novel missense mutation c.3448G > T; p.Asp1150Tyr in the coding exon 17 of *AFF2* gene as the pathogenic variant. The findings highlighted the application of Sanger sequencing as an efficient tool for validation of variant associated with disease. Combinational approach of WES and Sanger sequencing can provide better clinical diagnosis of the ID families.

Abbreviations

ID
Intellectual Disability;
WES

Whole Exome Sequencing;

AFF2

AF4/FRM2 family member 2;

FRAXE

fragile XE syndrome;

DANN

deleterious annotation of genetic variants using neural networks;

GERP

Genomic Evolutionary Rate Profiling; gnomAD:Genome Aggregation Database.

Declarations

Ethical approval and consent to participate

The project was approved from the Board of advanced study and research (BASR) and ethical committee (No: IIU (BI&BT)/FBAS-2018-3591), of International Islamic University Islamabad, Pakistan according to the declaration of Helsinki principle. Written informed consents were obtained from the research participants. Children younger than 16 years of age, written consent to participate obtained from their parents.

Consent for Publication

Written informed consent for publication of this research work was obtained from the research participant and child's < 16 years of age written informed consent was obtained from their parents for the publication of their clinical and genetic data.

Availability of Data

Variant has been submitted to the NCBI database ClinVar, available online by using following Accession number: SCV001423533. All data related to paper available on request from corresponding authors

Competing Interests

Authors declare they have no competing interests.

Funding

The study was supported by higher education commission (HEC) of Pakistan under student scholarship program (award No: 1-8/HEC/HRD/2018/8590, NRPU7028). Microarray/exome sequencing, Sanger sequencing was carried out at RILD Wellcome Wolfson Center, Royal Devon, University of Exeter.

Author Contributions

I.A performed all particle work; compile the manuscript. G.H performs the WES analysis and help in variants selection. M.I performs the bioinformatics analysis, multiple sequencing alignment. A.M designed the project and provided funds for lab work. All authors approved the manuscript.

Acknowledgments

We thank patients and their families for participation in this study. We also thank Professor Andrew H Crosby and Emma Baple, University of Exeter for their help, guidance and providing facilities for this project.

Author Information

Not applicable

Author Details

¹Department of Biological Sciences, International Islamic University Islamabad, 44000, Pakistan.

²College of Medicine and Health, University of Exeter, Barrack Road, EX2 5DW, Exeter, Devon, UK.

³Department of Pharmacology, Rajarshi Shahu College of Pharmacy, Malvihir, Buldana, Maharashtra, India.

References

1. Chelly J, Khelfaoui M, Francis F, Chérif B, Bienvenu T. Genetics and pathophysiology of mental retardation. *Eur J Hum Genet.* 2006;14(6):701–13.
2. Wieczorek D. Autosomal dominant intellectual disability. *medizinische genetik.* 2018;30(3):318–22.
3. Fedulov V, Rex CS, Simmons DA, Palmer L, Gall CM, Lynch G. Evidence that long-term potentiation occurs within individual hippocampal synapses during learning. *J Neurosci.* 2007;27(30):8031–9.
4. Ilyas M, Efthymiou S, Salpietro V, Noureen N, Zafar F, Rauf S, Mir A, Houlden H. Novel variants underlying autosomal recessive intellectual disability in Pakistani consanguineous families. *BMC medical genetics.* 2020;21:1–8.
5. Uyguner O, Kayserili H, Li Y, Karaman B, Nürnberg G, Hennies HC, Becker C, Nürnberg P, Başaran S, Apak M. A new locus for autosomal recessive non-syndromic mental retardation maps to 1p21. 1–p13. 3. *Clin Genet.* 2007;71(3):212–9.
6. McKenzie K, Milton M, Smith G, Ouellette-Kuntz H. Systematic review of the prevalence and incidence of intellectual disabilities: current trends and issues. *Current Developmental Disorders Reports.* 2016;3(2):104–15.
7. Inlow JK, Restifo LL. Molecular and comparative genetics of mental retardation. *Genetics.* 2004;166(2):835–81.

8. Basel-Vanagaite L. Clinical approaches to genetic mental retardation. The Israel Medical Association journal: IMAJ. 2008;10(11):821.
9. Opitz J: **Historiography of the causal analysis of mental retardation**. In: *Speech to the 29th Annual Gatlinburg Conference on Research and Theory in Mental Retardation, Gatlinburg, TN: 1996; 1996*.
10. Castellví-Bel S, Mila M. Genes responsible for nonspecific mental retardation. Mol Genet Metab. 2001;72(2):104–8.
11. Handt M, Epplen A, Hoffjan S, Mese K, Epplen JT, Dekomien G. Point mutation frequency in the FMR1 gene as revealed by fragile X syndrome screening. Mol Cell Probes. 2014;28(5–6):279–83.
12. Quinlan A, Kindlon N. **bedtools: A powerful toolset for genome arithmetic**. 2017.
13. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.
14. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research. 2010;38(16):e164–4.
15. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. Primer3—new capabilities and interfaces. Nucleic acids research. 2012;40(15):e115–5.
16. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics. 2000;16(4):404–5.
17. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinform. 2008;9(1):40.
18. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols. 2010;5(4):725.
19. Lovell S, Davis I, Arendall W, de Bakker P, Word J, Prisant M, Richardson J, Richardson D: **Structure validation by Calpha geometry: phi, psi and Cbeta deviation**. **Proteins 50, 437e450**. In.; 2003.
20. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12.
21. Binkowski TA, Naghibzadeh S, Liang J. CASTp: computed atlas of surface topography of proteins. Nucleic acids research. 2003;31(13):3352–5.
22. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic acids research. 2006;34(suppl_2):W116–8.
23. Capriotti E, Altman RB, Bromberg Y. Collective judgment predicts disease-associated single nucleotide variants. BMC Genomics. 2013;14(S3):2.
24. Emidio C, Piero F, Rita C. **I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure**. *Nucleic acids research*, 33.
25. Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, Brezovsky J, Damborsky J. **PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations**. *PLoS computational biology* 2014, 10(1).

26. Chen Y, Cramer P. Structure of the super-elongation complex subunit AFF4 C-terminal homology domain reveals requirements for AFF homo-and heterodimerization. J Biol Chem. 2019;294(27):10663–73.

27. Bensaid M, Melko M, Bechara EG, Davidovic L, Berretta A, Catania MV, Gecz J, Lalli E, Bardoni B. FRAXE-associated mental retardation protein (FMR2) is an RNA-binding protein with high affinity for G-quartet RNA forming structure. Nucleic acids research. 2009;37(4):1269–79.

Figures

Figure 1

(A) Pedigree of the family. (B) Chromatograms illustrating AFF2 variant C.3448G>A, (p.Asp1150Tyr) in wild type (IV: 1), heterozygous (III: 1) and affected individual (IV: 2) in exon 17 of AFF2 gene. (C) Superimposed normal and mutant AFF2 structure highlighting mutation substitution of Asp (magenta color) with Tyr (red color) at protein position 1150. Protein structure has been shown in green and cornflower blue color representing wild and mutant types respectively. (D). Multiple alignment of mutated protein highly conserved in all species, sequencing alignment also showing conservation with in nearby amino acid in all species.