Molecular characterization of recessively inherited ataxic and neuropathic disorders in consanguineous Pakistani families

Faiza Aslam[1,4], Muhammad Wajid[1], Amina Iftikhar Butt[1], Elizabeth Wohler[2], Go Hun Seo[3], Weizhen Ji[4], Saquib A. Lakhani[4], Nara Sobreira[2], Sadaf Naz*[1]

[1]School of Biological Sciences, University of the Punjab, Lahore, Pakistan. [2]McKusick-Nathans Department of Genetic Medicine, John Hopkins University, Baltimore, MD, USA, [3]3billion inc, Seoul, South Korea,[4]Pediatric Genomic Discovery Program, Department of Pediatrics, Yale University School of Medicine, New Haven, CT, USA.

F.A and S.A.L: Current address: Department of Pediatrics, Cedars Sinai Guerin Children's, Los Angeles, CA.

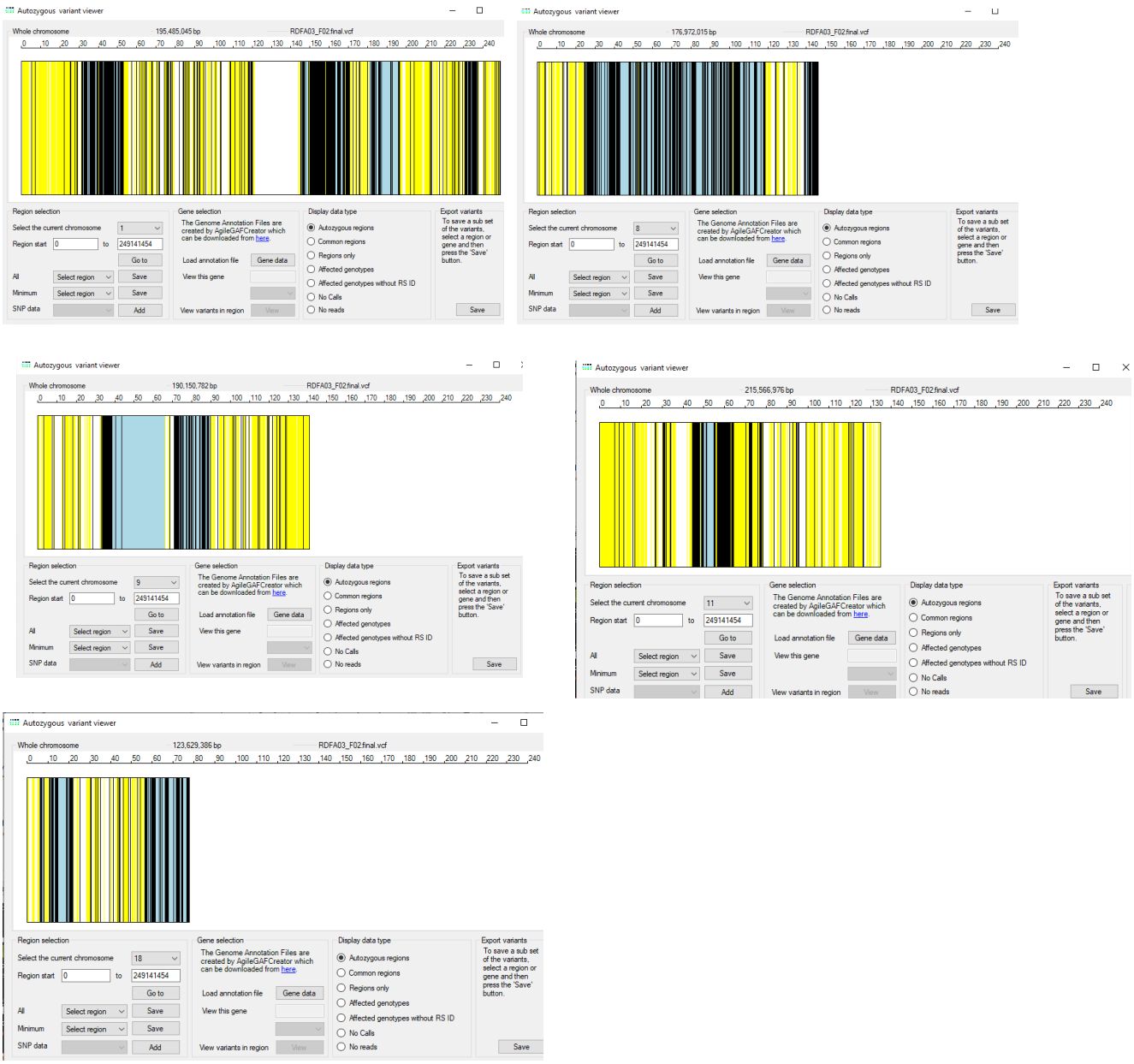
Correspondence to

Sadaf Naz, School of Biological Sciences, University of the Punjab, Lahore, Pakistan,

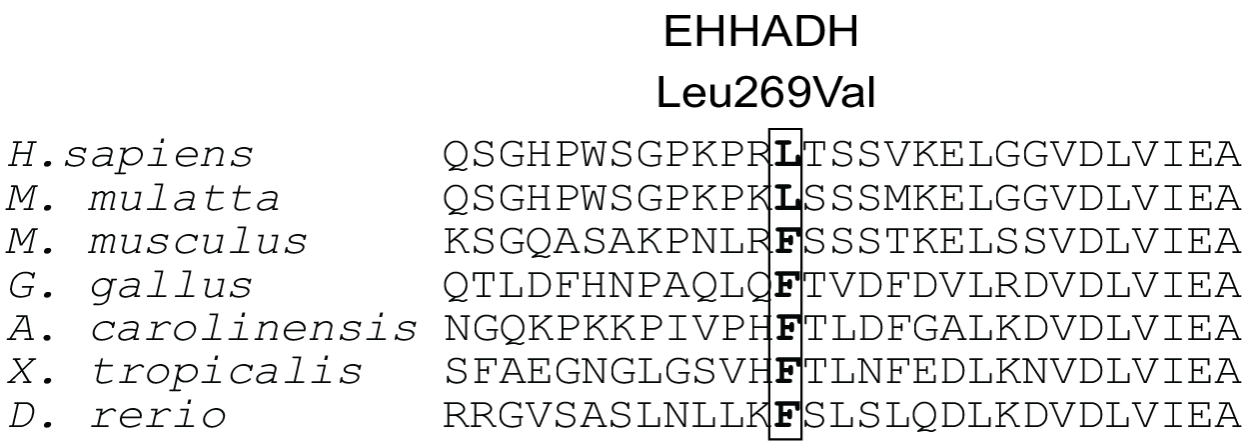
naz.sbs@pu.edu.pk

Supplementary Figures

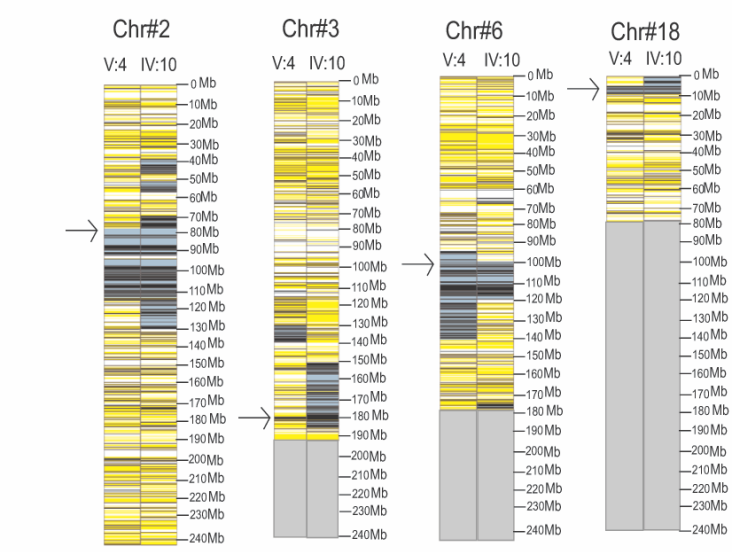
Supplementary Figures 1 to 4



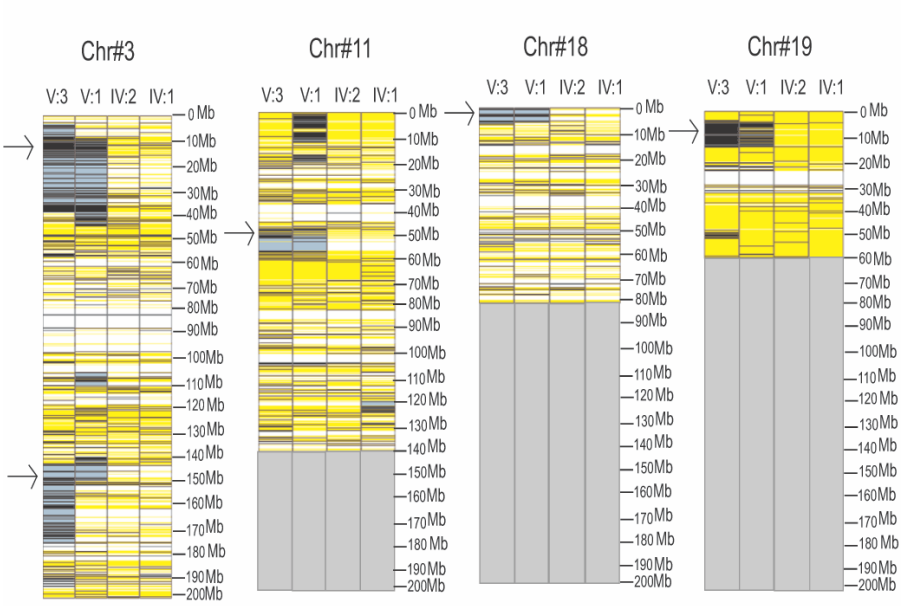
Supplementary Figure. S1. Autozygome of family RDFA03 patient: AgileVCFMapper displays of data for the chromosomes number 1, 8, 9, 11 and 18 for individual V:I. Positions of the regions are provided in the Mb above the figure in the form of a scale. Black/dark gray regions are homozygous, yellow show heterozygous calls, white lines are regions missed during sequencing. There are multiple regions of homozygosity on each chromosome while that on chromosome 8 covers a very large area.



Supplementary Figure. S2. Partial alignment of EHHADH: Clustal O analysis revealed that the Leu269 amino acid residue in EHHADH is not conserved in the vertebrate species. Leucine is replaced with a phenylalanine in all non-primate species.



Supplementary Figure. S3. Autozygome of family RDFA07: AgileVCFMapper displays of chromosomes 2, 3, 6 and 18 showing regions of homozygosity shared by the two affected individuals V:4 and IV:10. Sizes of the regions are provided on maps of all chromosomes and arrows indicate the homozygous regions shared by the affected individuals. Black/dark gray lines represent homozygous regions, yellow depict the heterozygous regions and the white areas shows the regions not captured during sequencing. The large light gray boxes at the lower regions on chromosomes 3, 6 and 18 indicate that there is no data since these chromosomes do not extend that far. The 2Mb region of homozygosity on chromosome 3 corresponds to the position of *EHHADH*.



Supplementary Figure. S4. Autozygome of family RDFA08: AgileVCFMapper display of data for individual chromosomes. Shared regions of homozygosity of the two siblings are visible on chromosomes 3, 11, 18 and 19 which are heterozygous in the unaffected parents. Arrows represent the shared homozygous regions and the position of the regions in Mb are provided at the right of each figure. Black/dark gray regions are homozygous, yellow show heterozygous calls, white lines are regions missed during sequencing and light gray boxes indicate the ends of chromosomes.

Supplementary Tables

Tables S1 to Table S9

Table S1. Primers used for PCR amplification prior to Sanger sequencing

Primer ID		Primer sequences	Product size (bp)
MFSD8	F	AATTTTGTTTTACCAAAGAAATGCT	489
	R	TGTAACAACAAACGCTTAAAATGT	
*MFSD8*Ot	F	AAGAGTGTTGTGTGGTCTGGGAAATGA	860
	R	GTGTGGCTTGTTTCCTTTTCCTTTGTTT	
*MFSD8*In	F	GGCTTCAACCCCAAGAGCAGCAATT**A**	388
	R	AGAACAAGCTGTGTTATATAATGGCATCA**C**	527
GDAP1a3ex	F	CCTTAAGGGTGAGACCACTGA	511
	R	GCTAAATGCTTGGATGAGCTG	
GDAP1ex	F	GGTTAGGGAATGCTTCATGG	587
	R	GCAAACTCAACAACCCTATTCC	
AFG3L2	F	CCGGGACAGTGTCAACTTCT	473
	R	TCCAGGCTGCTGTGTAGATTT	
SETX	F	AGACCAGCAATTCGTGAAGT	400
	R	TGGCAGCTTTGAAAGAAGGAG	
FLNA	F	GGCCATTTTTCTTCACATGC	484
	R	CGGGGGACTACAGCATTCTA	
EHHADH	F	CAATGGTAGTGGGGGAAGAG	547
	R	TGTCACATAAGCCTTCCATCC	
AIMP2	F	TCAAAGGTCGACTTCGTGAC	488
	R	CCGGCCTGGATAAATCTTT	
*MMACHC*Ot	F	CATGTGACAGGACTGCCTCCCATCCAGAA	371
	R	GGGCTCCCTCGGACAAGGTCATAACTCCC	
*MMACHC*In	F	CAGCTACATGGGCTGCTGTCTGGGA**G**	167
	R	CACCCCAACCGACGCCCCAAGATA**T**	257

The primers indicated by "Ot" are for use with allele specific primers, denoted by "In". All four primers are used in one reaction as per tetra-primer PCR protocol and yield a constant band as well as the specific allele products. The bold lettering in primers indicates the allele specific nucleotides while the underlined bases are those which were deliberately mismatched to increase the specificity of the primers. All other primers were used for PCR amplifications, followed by Sanger sequencing of the products.

Table S2. Homozygous gene variants shortlisted for patients of family RDFA05

Gene	*Position (hg19)	Transcript, cDNA change and Effect	gnomAD AF	ACMG classification, OMIM, Predictions	Comments
CCDC188	chr22:20138138	NM_001365892.1 c.372A>G, p.Ser124Ser, synonymous	0.00001387	VUS OMIM# N/A REVEL N/A FATHMM N/A	Discarded since the variant was predicted benign by spliceAI and other software
MFSD8	**Chr4:128851901**	**NM_152778.3, c.935T>C, p.Ile312Thr, missense**	**0.0000238**	**VUS OMIM# 611124 REVEL 0.4 FATHMM 0.32**	**Segregated. ClinVar VCV000211495.11 PMID: 39108195**
TXNRD2	Chr22:19868236	NM_006440.5, c.1091G>A, p.Arg364Gln, missense	0.0000606	VUS OMIM#606448 REVEL 0.17 FATHMM 0.41	Discarded since it is a VUS for primary dilated cardiomyopathy ClinVar VCV000572761.6
USP11	ChrX:47101842	NM_004651.3, c.1538G>A, p.Arg513Gln, missense	0	VUS OMIM# 300050 REVEL 0.2 FATHMM 2.02	Discarded since multiple programs predicted the variant to be benign and it is also present in control population as hemizygous variant.

*Human Genome reference GRCh37, ACMG = American college of medical genetics and genomics, OMIM = Online Mendelian inheritance in man REVEL = Rare exome variant ensemble learner, (Higher scores are more deleterious), FATHMM= Functional analysis through hidden Markov models, Lower scores are more deleterious, VUS=Variant of unknown significance, N/A Not applicable or not available

Table S3. Homozygous gene variants shortlisted for the patient of family RDFA09

Gene	*Position (hg19)	Transcript, cDNA change and Effect	gnomAD AF	ACMG classification, OMIM, Predictions	Comments
AFG3L2	**Chr18:12337348**	**NM_006796.2, c.2167G>A, p.Val723Met, missense**	**0.0001749**	**VUS OMIM# 604581 REVEL 0.58 FATHMM -2.04**	**Segregated. ClinVar VCV000214062.16, likely pathogenic, PMID: 31111429**
ALCAM	Chr3:105252544	NM_001627.4, c.547+10A>G, splice variant	0.000039	VUS, OMIM# 601662 REVEL N/A FATHMM N/A	Discarded as the variant was predicted to be benign by SpliceAI and other software
ARHGEF40	Chr14:21548870	NM_018071.5, c.2425A>G, p.Met809Val, missense	0.000264	VUS, OMIM# 610018 REVEL 0.01 FATHMM 4.48	Discarded as the variant was predicted benign by multiple software
ARG1	Chr6:131904991	NM_000045.3, c.912C>T, p.Phe304Phe, synonymous	0.000629	B, OMIM# 608313 REVEL N/A FATHMM N/A	Discarded as the ClinVar VCV000513240.13 significance is benign for arginase deficiency 1
CLSPN	Chr1:36215381	NM_022111.4, c.2060T>G, p.Ile687Arg, missense	0.000039	VUS, OMIM# 605434, REVEL 0.15 FATHMM 1.92	Discarded as the variant was predicted benign by multiple software
COL6A5	Chr3:130159257	NM_153264.6, c.6075A>C, p.Thr2025Thr, synonymous	0	VUS OMIM# 611916 REVEL N/A FATHMM N/A	Discarded as variant was predicted benign by SpliceAI and other software
DNAJC13	Chr3:132207195	NM_015268.4, c.3321C>T, p.Tyr1107Tyr, synonymous	0.000339	VUSOMIM# 614334 REVEL N/A	Discarded as the variant was predicted benign by

				FATHMM N/A	SpliceAI and other software
GBP5	Chr1:89730606	NM_001134486.2, c.912T>C, p.Ser304Ser, synonymous	0.000167	VUS OMIM# 611467 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
H6PD	Chr1:9322153	NM_004285.3, c.781C>T, p.Arg261Cys, missense	0.000289	LP OMIM# 138090 REVEL 0.94 FATHMM -5.57	ClinVar VCV002361435.1 lists this as VUS, for cortisone reductase deficiency 1 OMIM# 604391. The phenotype which includes obesity does not match the patient condition, who was very slim. The amino acid affected by the variant is not conserved in evolution.
IMPACT	Chr18:22007950	NM_018439.4, c.104A>G, p.Lys35Arg, missense	0.000054	VUS OMIM# 615319 REVEL 0.07 FATHMM 1.91	Discarded as the variant was predicted benign by multiple software
KIAA0825	Chr4:93739457	NM_001145678.1, c.2704T>C, p.Leu902Leu, synonymous	N/A	VUS OMIM# 617266 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
MACF1	Chr1:39852860	NM_012090.5, c.8160C>T, p.Ala2720Ala, synonymous	0.000383	VUS OMIM# 608271 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
MYH6	Chr14:23859570	NM_002471.3,	0.000079	VUS OMIM# 160710	Discarded as the ClinVar VCV000228890.12

		c.3428G>A, p.Arg1143Gln, missense		REVEL 0.52 FATHMM -1.21	VUS, Phenotype of cardiomyopathy hypertrophic 14, OMIM# 613251 does not match the patient condition.
NAPRT	Chr8:144657229	NM_145201.6, c.1481G>C, p.Arg494Thr, missense	0.000384	VUS OMIM# 611552 REVEL 0.43 FATHMM 0.36	Discarded as the variant is predicted uncertain by most software, the affected amino acid was not conserved in evolution
NT5C1A	Chr1:40131861	NM_032526.2, c.183G>A, p.Leu61Leu, synonymous	0.000032	VUS OMIM# 610525 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
PHF20L1	Chr8:133851793	NM_016018.4, c.2353C>G, p.Leu785Val, missense	0.000201	VUS OMIM# 620050 REVEL 0.3 FATHMM 0.15	Discarded as the variant was predicted uncertain by most software and the affected amino acid was not conserved in evolution. Currently, the gene is not associated with any disease.
SCMH1	Chr1:41494263	NM_001031694.2, c.1850G>A, p.Arg617His, missense	0.000107	VUS OMIM# 616396 REVEL 0.47 FATHMM 0.58	Although predicted deleterious by a few software, the affected amino acid was not conserved in evolution. Currently, the gene is not associated with any disease.
TERF1	Chr8:73951437	NM_017489.3, c.1126C>G, p.Arg376Gly, missense	0.000172	Variant of uncertain significance OMIM# 600951 REVEL 0.02	Discarded as the variant was predicted benign by multiple programs

				FATHMM N/A	
YIPF6	ChrX:67731686	NM_173834.4, c.58-5C>A, intronic splice	N/A	VUS OMIM# 300996 REVEL N/A FATHMM N/A	Discarded as the variant is predicted benign by SpliceAI and other software

*Human Genome reference GRCh37, ACMG = American college of medical genetics and genomics, OMIM = Online Mendelian inheritance in man REVEL = Rare exome variant ensemble learner, (Higher scores are more deleterious), FATHMM= Functional analysis through hidden Markov models, Lower scores are more deleterious, VUS=Variant of unknown significance, N/A Not applicable or not available

Table S4. Homozygous gene variants shortlisted for patients of family RDFA11

Gene	*Position (hg19)	Transcript, cDNA change and Effect	gnomAD AF	ACMG classification, OMIM, Predictions	Comments
SETX	**Chr9:135202351**	**NM_015046.7 c.4633-4636delAGTG, p.Ser1545AlafsTer25, frameshift**	**0**	**OMIM#608465 FATHMM N/A**	**Segregated PMID: 19744353**
GTF3C5	Chr9:135917601	NM_012087.4 c.281G>A, p.Arg94Gln, missense	0.0001	OMIM# 604890 REVEL 0.132 FATHMM 0.72	Discarded as the variant was predicted benign or tolerated by multiple software
CD44	Chr11:35160854	(NM_000610.4) c.4G>A, p.Asp2Asn, missense	0	OMIM#107269 REVEL 0.032 FATHMM 2.72	Discarded as the variant was predicted benign by multiple software

*Human Genome reference GRCh37, ACMG = American college of medical genetics and genomics, OMIM = Online Mendelian inheritance in man REVEL = Rare exome variant ensemble learner, (Higher scores are more deleterious), FATHMM= Functional analysis through hidden Markov models, Lower scores are more deleterious, VUS=Variant of unknown significance, N/A Not applicable or not available

Table S5. Homozygous gene variants shortlisted for one patient of family RDFA03

Gene	*Position (hg19)	Transcript, cDNA change and Effect	gnomAD AF	ACMG classification, OMIM, Predictions	Comments
CPZ	Chr4:8607786	NM_001014447.3, c.780C>A, p.Tyr260Ter, nonsense	0	VUS OMIM# 603105 REVEL N/A FATHMM N/A	Discarded as the variant did not segregate (unaffected individuals of the family were also homozygous for the variant)
GDAP1	**Chr8:75276365**	**NM_018972.3, c.840delC, p.Tyr280Ter, nonsense**	**0**	**LP OMIM# 606598 REVEL N/A FATHMM N/A**	**Segregated with the phenotype ClinVar VCV000637127.3 PMID: 25231362**
KIRREL1	Chr1:157963526	NM_018240.7, c.52+8C>T, intronic splice	0	VUS OMIM# 607428 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
LRRC27	Chr10:134162562	NM_001143759.1, c.1021+1G>A, canonical splice site	0	VUS OMIM# N/A REVEL N/A FATHMM N/A	Discarded as the variant affects splicing of an alternative exon of a small isoform. The longer transcript NM_030626.3 is not affected. Variants of this gene are not associated with any disease.
PEX2	Chr8:77895774	NM_000318.3, c.641A>G, p.Gln214Arg, missense	0	VUS OMIM# 170993 REVEL 0.26 FATHMM -1.55	Discarded as the variant was predicted benign by multiple software
STAG2	ChrX:123217267	NM_001042750.1, c.2925-4G>A, intronic splice	0	VUS OMIM# 300826 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software

*Human Genome reference GRCh37, ACMG = American college of medical genetics and genomics, OMIM = Online Mendelian inheritance in man REVEL = Rare exome variant ensemble learner, (Higher scores are more deleterious), FATHMM= Functional analysis through hidden

Markov models, Lower scores are more deleterious, VUS=Variant of unknown significance, N/A Not applicable or not available

Table S6. Homozygous gene variants shortlisted for the patients of family RDFA04

Gene	*Position (hg19)	Transcript, cDNA change and Effect	gnomAD AF	ACMG classification, OMIM, Predictions	Comments
AIMP2	Chr7:6049009	NM_006303.4, c.15G>C, p.Gln5His, missense	0.0000798103	VUS OMIM# 600859 REVEL 0.11 FATHMM 1.14	Discarded as the variant was predicted benign by multiple software
GDAP1	**Chr8:75272408**	**NM_018972.3, c.347T>G, p.Met116Arg, missense**	**0.000008122**	**P OMIM# 606598 REVEL 0.56 FATHMM -5.24**	**Segregated with the phenotype ClinVar VCV000038411.49 PMID: 15377708**
HIVEP1	Chr6:12123763	NM_002114.4, c.3735A>T, p.Arg1245Arg, synonymous	0.000200563	LB OMIM# 194540 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software and designated likely benign in ClinVar RCV000976528
KRBA2	Chr17:8273619	NM_213597.3, c.312G>A, p.Glu104Glu, synonymous	0.0002052346	VUS OMIM# N/A REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software

*Human Genome reference GRCh37, ACMG = American college of medical genetics and genomics, OMIM = Online Mendelian inheritance in man REVEL = Rare exome variant ensemble learner, (Higher scores are more deleterious), FATHMM= Functional analysis through hidden Markov models, Lower scores are more deleterious, VUS=Variant of unknown significance, N/A Not applicable or not available

Table S7. Homozygous gene variants shortlisted for patients of family RDFA07

Gene	*Position (hg19)	Transcript, cDNA change and Effect	gnomAD AF	** ACMG classification, OMIM, Predictions	Comments
EHHADH	Chr3:184911093	NM_001966.4, c.1093T>G, p.Leu365Val, missense	0.0006347	VUS OMIM# 607037 REVEL 0.15, FATHMM -1.2	Segregated with the phenotype, but the dominant disorder Fanconi renotubular syndrome 3 (OMIM# 615605) does not correspond to the observed phenotype of our patients. The amino acid is not conserved in evolution.
FLNA	ChrX:153581700	NM_001110556.2, c.5986C>T, p.Pro1996Ser, missense	0	VUS OMIM# 300017 REVEL 0.46 FATHMM 0.87	Discarded as the variant did not segregate with the disease in the family. The unaffected individuals were also homozygous
LAMA1	Chr18:7011323	NM_005559.4, c.3663G>A, p.Leu1221Leu, synonymous	0.0000893	VUS OMIM# 150320 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
TMEM131	Chr2:98377305	NM_015348.2, c.4962C>T, p.Asn1654Asn, synonymous	0.0000160	VUS OMIM# 615659 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software

*Human Genome reference GRCh37, ACMG = American college of medical genetics and genomics, OMIM = Online Mendelian inheritance in man REVEL = Rare exome variant ensemble learner, (Higher scores are more deleterious), FATHMM= Functional analysis through hidden Markov models, Lower scores are more deleterious, VUS=Variant of unknown significance, N/A Not applicable or not available

Table S8. Homozygous gene variants shortlisted for two patients of family RDFA08

Gene	*Position (hg19)	Transcript, cDNA change and Effect	gnomAD AF	** ACMG classification, OMIM, Predictions	Comments
EPM2AIP1	chr3:37033984	NM_014805.3 c.585C>T, p.Ile195Ile, synonymous	0.000095529	VUS OMIM# 607911 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
XIRP1	chr3:39229244	NM_194293.4 c.1693C>T, p.Arg565Trp, missense	0.00003185	VUS OMIM# 609777 REVEL 0.144 FATHMM 3.47	Discarded as the variant was predicted benign by multiple software
FUT5	chr19:5866638	NM_002034.2 c.1099C>T, p.Arg367Cys, missense	0.00019100	VUS OMIM# 136835 REVEL 0.08 FATHMM 1.67	Discarded as the variant was predicted benign by multiple software
ZNF823	chr19:11833032	NM_001080493.4 c.1317T>G, p.Thr439Thr, synonymous	0.000063702	VUS OMIM# N/A REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
MAGEA6	chrX:151869997	NM_175868.2 c.687G>A, p.Glu229Glu, synonymous	0	LB OMIM# 300176 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software
FAM120C	chrX:54177771	NM_017848.6 c.1065C>T, p.Asp355Asp, synonymous	0.0000458968	LB OMIM# 300741 REVEL N/A FATHMM N/A	Discarded as the variant was predicted benign by SpliceAI and other software

*Human Genome reference GRCh37, ACMG = American college of medical genetics and genomics, OMIM = Online Mendelian inheritance in man REVEL = Rare exome variant ensemble learner, (Higher scores are more deleterious), FATHMM= Functional analysis through hidden Markov models, Lower scores are more deleterious, VUS=Variant of unknown significance, N/A Not applicable or not available

Table S9. Homozygous variants prioritized for the affected individual V:7 in family RDFA08

Gene	*Position (hg19)	Transcript, cDNA change and Effect	gnomAD AF	** ACMG classification, OMIM, Predictions	Comments
ALS2	**Chr2:202588157**	**NM_020919.4 c.3520A>T, p.Lys1174Ter**	**0.000004008**	**P, OMIM# 607225 REVEL N/A FATHMM N/A**	**Prioritized since it is a known pathogenic variant. ClinVar VCV000645923.10 PMID: 11586298**
ZSWIM6	Chr5:60628619	NM_020928.2 c.520_522del, p.Ala184del	0.002819	VUS OMIM# 617865 REVEL N/A FATHMM N/A	Discarded as the variant is located in a stretch of 11 Ala residues and is a region with many common deletions and insertion variants of different lengths
ATN1	Chr12:7045892	NM_001007026.2 c.1462_1476del, p.Gln498_Gln502del	0	VUS OMIM# 618494 REVEL N/A FATHMM N/A	Discarded as the variant is located in a stretch of 19 Gln residues and is a region with many deletions and insertion variants of different lengths
DMD	ChrX:31222190	NM_004015.3 c.491G>A, p.Arg164His	0.00001638	VUS OMIM#310200 REVEL 0.78 FATHMM -0.7	Discarded as the variant is present as hemizygous in 4 individuals in gnomADv4 ClinVar VCV000582901.6

*Human Genome reference GRCh37, ACMG = American college of medical genetics and genomics, OMIM = Online Mendelian inheritance in man REVEL = Rare exome variant ensemble learner, (Higher scores are more deleterious), FATHMM= Functional analysis through hidden Markov models, Lower scores are more deleterious, VUS=Variant of unknown significance, N/A Not applicable or not available