

Supplementary Material

Gene	mRNA reference number	Chr.	Genomic Position	Nucleotide change	Amino acid change	ExAC allele frequency (%)	gnomAD allele frequency (%)	Phenotype MIM number (OMIM)	MGI phenotype
<i>LRPPRC</i>	NM_133259.3	2	44,190,789	c.1426A>G	p.(Thr476Ala)	0.1129	0.09423	220111	1919666
<i>ATAD1</i>	NM_032810.3	10	89,514,459	c.1070_1071del	p.(His357Argfs*15)	0.0008266	0.00122	–	1915229
<i>RNLS</i>	NM_001031709.2	10	90,341,351	c.340A>G	p.(Ile114Val)	0.002481	0.003371	–	1915045
<i>ZMYM5</i>	NM_001142684.1	13	20,426,139	c.180_182del	p.(Val61del)	0	0	–	–
<i>CDH8</i>	NM_001796.4	16	61,859,025	c.726T>C	p.Val242=	0.01903	0.03224	–	107434

Supplementary Table 1 Identified rare homozygous candidate variants shared by the two affected siblings and absent in the mother in the homozygous state. –: not available; Chr.: chromosome; MGI: Mouse Genome Informatics; OMIM: Online Mendelian Inheritance in Man.

		Patient 1	Patient 2	Patient 3	Mother	Father
1	<i>LRPPRC</i>	G/G	A/A	G/G	A/G	A/G
2	<i>ATAD1</i>	del/del	del/del	del/del	del/WT	del/WT
3	<i>RNLS</i>	G/G	G/G	G/G	A/G	A/G
4	<i>ZMYM5</i>	del/del	del/WT	del/del	del/WT	del/WT
5	<i>CDH8</i>	C/C	C/C	C/C	T/C	T/C

Supplementary Table 2 Segregation analysis of the five homozygous candidate variants within the family. Genotypes interfering with the expected segregation are highlighted in bold excluding these variants to underlie the phenotype in the three affected siblings. del: deletion; WT: wild-type.

Variants	GERP++	CADD	REVEL	M-CAP
ATAD1: c.1070_1071del/p.(His357Argfs*15)	5.48	35	–	–
RNLS: c.340A>G/p.(Ile114Val)	1.95	10.28	0.058	0.013
CDH8: c.726T>C/p.Val242=	6.02	14.23	–	–

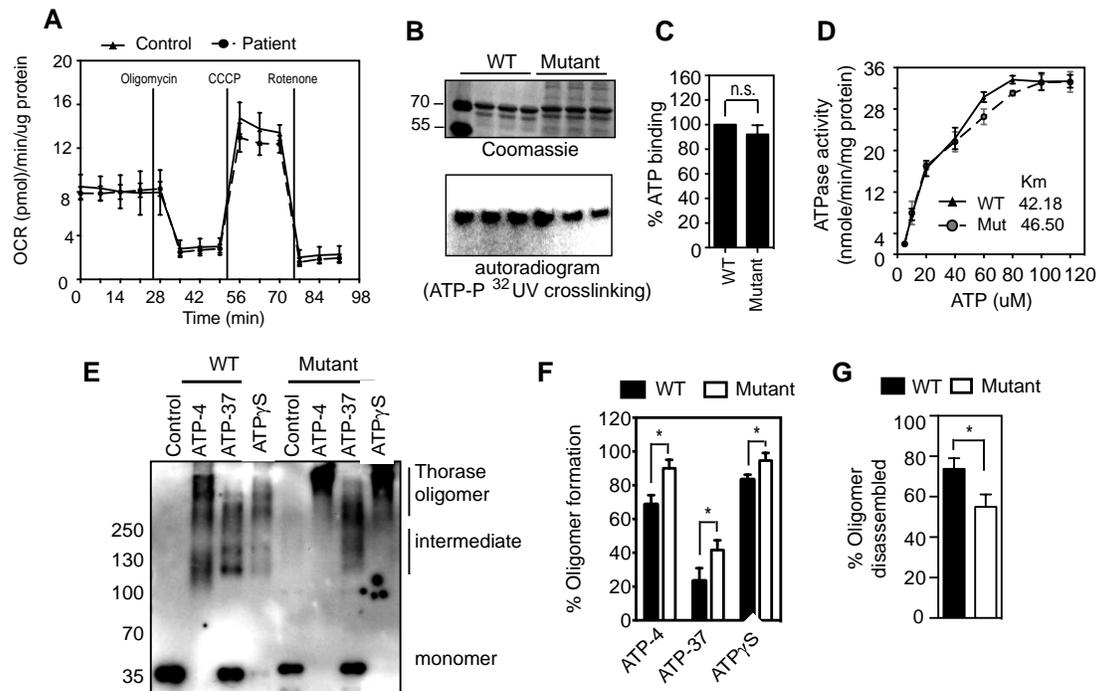
Supplementary Table 3 *In silico* pathogenicity prediction of variants in *ATAD1*, *RNLS*, and *CDH8*. The functional impact of the identified variants were predicted by the GERP++, Combined Annotation Dependent Depletion (CADD), Rare Exome Variant Ensemble Learner (REVEL) and Mendelian Clinically Applicable Pathogenicity (M-CAP) scoring systems. GERP++ scores measure evolutionary constraint across mammalian species at the single-nucleotide level and have previously been shown to correlate with mutational deleteriousness and to be useful in the identification of highly penetrant mutations (Cooper *et al.*, 2005, Davydov *et al.*, 2010). CADD is a framework that integrates multiple annotations into one metric by contrasting variants that survived natural selection with simulated mutations. Reported CADD score is a phred-like rank score based upon the rank of that variant’s score among all possible single nucleotide variants of hg19, with 10 corresponding to the top 10%, 20 to the top 1%, and 30 to the top 0.1%. The larger the score the more likely the variant has deleterious effects; the score range observed here is strongly supportive of pathogenicity, with all observed variants ranking above ~99% of all variants in a typical genome and scoring similarly to variants reported in ClinVar as pathogenic (~85% of which score >15) (Kircher *et al.*, 2014). REVEL is an ensemble method predicting the pathogenicity of missense variants with a strength for distinguishing pathogenic from rare neutral variants with a score ranging from 0-1. The higher the score the more likely the variant is pathogenic (Ioannidis *et al.*, 2016). M-CAP is a clinical pathogenicity classifier which is trained to not misclassify more than 5% of pathogenic variants while reducing the number of variants of uncertain significant. The authors recommend a threshold of >0.025 to classify a variant as pathogenic (Jagadeesh *et al.*, 2016). –: not applicable.

	Splice View		NetGene2		BDGP	
	wild-type	mutant	wild-type	mutant	wild-type	mutant
<i>RNLS</i>: c.340A>G	DSS: 83 ASS: 78	DSS: 83 ASS: 78	DSS: 0.63 ASS: 0.25	DSS: 0.63 ASS: 0.25	DSS: 0.63 ASS: 0.73	DSS: 0.63 ASS: 0.73
<i>CDH8</i>: c.726T>C	DSS: 79 ASS: 91	DSS: 79 ASS: 91	DSS: 0.81 ASS: 0.95	DSS: 0.81 ASS: 0.95	DSS: 0.97 ASS: 0.93	DSS: 0.97 ASS: 0.93

Supplementary Table 4. Splice site prediction for the nucleotide changes in *RNLS* and *CDH8*. Donor and acceptor splice site prediction scores were calculated for wild type and mutated sequences by using the programs Splice View, NetGene2, and Berkeley Drosophila Genome Project (BDGP). The prediction of exonic cryptic donor or acceptor splice sites was not observed by any of the three programs. High and low scores indicate strong and weak splices sites, respectively. ASS: acceptor splice site; DSS: donor splice site.

Species	aa	Alignment
<i>H. sapiens</i>	357	S K D A A F Q N V L T H V C L D
<i>P. troglodytes</i>	357	S K D A A F Q N V L T H V C L D
<i>M. mulatta</i>	357	S K D A A F Q N V L T H V C L D
<i>F. catus</i>	357	S K D A A F Q N V L T H V C L D
<i>M. musculus</i>	357	S K D A A F Q N V L T H V C L D
<i>G. gallus</i>	358	S K N A T L H D D L M H V I I D
<i>T. rubripes</i>	375 M L P V V P E V A L D
<i>D. rerio</i>	361	S K S A G V H E A F M Q V A L D
<i>X. tropicalis</i>	358	S K S A T N Q N V L M H V S L D
Mutated	357	S K D A A F Q N V L T R L F R L R V K I I C T V Q

Supplementary Figure 1 Amino acid sequence alignment of the human ATAD1 C-terminus (aa 357-361; NP_116199.2) with orthologs. The predicted amino acid sequence resulting from the frameshift mutation and leading to an extended C-terminus with 14 ATAD1-unrelated amino acids (bold) is shown below (Mutated). Multiple alignment was gathered from <http://mutationtaster.org/>. Conserved residues among different species affected by the mutation are grey shaded. aa: amino acid.



Supplementary Figure 2 (A) Oxygen consumption rate (OCR) in patient and control fibroblasts. The data was normalized to the amount of protein in each well and represents average from three independent experiments (n= 6 wells for each experiments). (B) UV light-induced cross-linking of radiolabeled [α -P³²]ATP to purified proteins. The cross-linked proteins were separated on 10% SDS-PAGE. Upper panel is Coomassie stain of the protein input and lower panel is a representative of autoradiogram of [α -P³²]ATP bound to proteins. (C) Percent ATP bound to the Thorase in (B). (D) Analysis of ATP hydrolysis (ATPase) activities of purified Thorase. The Km (mM) values are shown on the graph (mean \pm standard error of the mean [SEM] of experiments performed in triplicate. n.s, not significant. $p > 0.10$, ANOVA with Tukey-Kramer post-hoc test when compared with wildtype). (E) Immunoblot analyses of Thorase wildtype and mutant^{His357Argfs*15} oligomer formation. The samples were cross-linked by glutaraldehyde in the presence ATP or ATP γ S. The ATP-treated samples were treated at 4°C for binding (ATP-4) or at 37°C (ATP-37) for ATP hydrolysis. (F) The graph represents normalized percent oligomer formed for (E). (G) The percentage of Thorase oligomer disassembled at 4°C (mean \pm standard error of the mean [SEM] of three experiments performed in triplicate. n = 3, * $p < 0.10$, n.s $p > 0.10$, ANOVA with Tukey-Kramer post-hoc test compared with wildtype, Power: $1 - \beta$ err prob = 1.0).

Supplemental References

- Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A, et al. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Research*. 2005; 15: 901-13.
- Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol*. 2010; 6: e1001025.
- Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *American Journal of Human Genetics*. 2016; 99: 877-85.
- Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN, et al. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat Genet*. 2016; 48: 1581-6.
- Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014; 46: 310-5.