

The American Journal of Human Genetics

Supplemental Data

***THOC2* Mutations Implicate mRNA-Export Pathway**

in X-Linked Intellectual Disability

Raman Kumar, Mark A. Corbett, Bregje W.M. van Bon, Joshua A. Woenig, Lloyd Weir, Evelyn Douglas, Kathryn L. Friend, Alison Gardner, Marie Shaw, Lachlan A. Jolly, Chuan Tan, Matthew F. Hunter, Anna Hackett, Michael Field, Elizabeth E. Palmer, Melanie Leffler, Carolyn Rogers, Jackie Boyle, Melanie Bienek, Corinna Jensen, Griet Van Buggenhout, Hilde Van Esch, Katrin Hoffmann, Martine Raynaud, Huiying Zhao, Robin Reed, Hao Hu, Stefan A. Haas, Eric Haan, Vera M. Kalscheuer, and Jozef Gecz

Figure S1

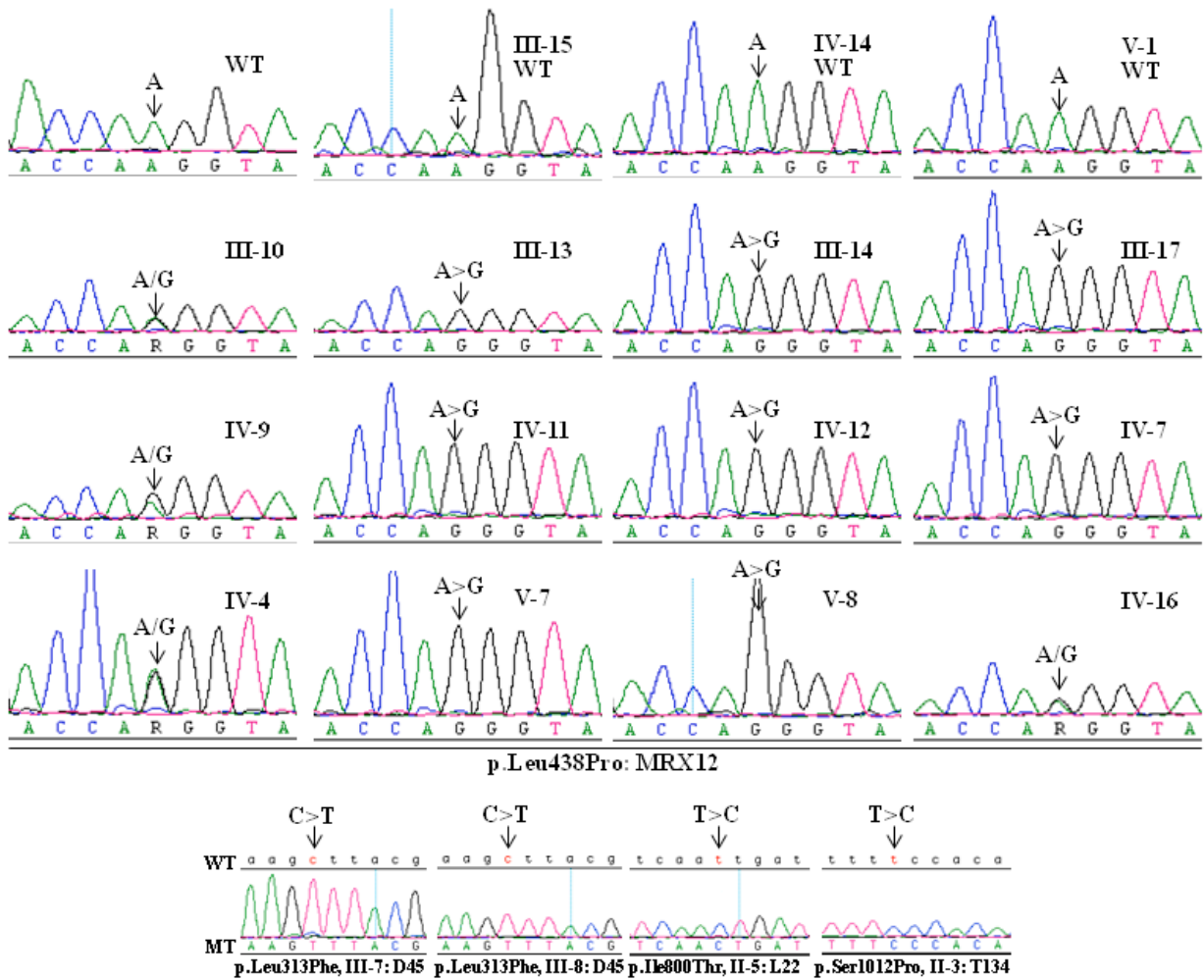


Figure S1. Sanger sequencing chromatograms of representative affected individuals from each family are shown. Top panel. MRX12 family sequence chromatograms of the complementary strand of *THOC2* are shown. Bottom panel. Representative chromatograms are shown from two individuals from family D45 and one each from families L22 and T134, respectively. WT = wild type allele. Primers used for PCR and sequencing are listed in Table S3.

Figure S2

p.Leu438Pro: MRX12

Mutant	NMFCY P GPHLS	443
Homo	NMFCY L GPHLS	443
Mus	NMFCY L GPHLS	443
Rattus	NMFCY L GPHLS	443
Pan	NMFCY L GPHLS	443
Macaca	NMFCY L GPHLS	443
Canis	NMFCY L GPHLS	443
Equus	NMFCY L GPHLS	443
Sus	NMFCY L GPHLS	443
Bos	NMFCY L GPHLS	443
Loxo	HMFCY L GPHLS	443
Gallus	NMLCY L GPHLS	440
Xenopus	TMLCY L GPHLS	439
Drosophila	PMANV L GPAMH	503

p.Leu313Phe: D45

Mutant	QIVRK F TMVVL	318
Homo	QIVRK L TMVVL	318
Mus	QIVRK L TMVVL	318
Rattus	QIVRK L TMVVL	318
Pan	QIVRK L TMVVL	318
Macaca	QIVRK L TMVVL	318
Canis	QIVRK L TMVVL	318
Equus	QIVRK L TMVVL	318
Sus	QIVRK L TMVVL	318
Bos	QIVRK L TMVVL	318
Loxo	QIVRK L TMVVL	318
Gallus	QIVRK L TMVVL	315
Xenopus	QIVRK L TMVVL	315
Drosophila	EMVRK L NLIQT	372

p.Ile800Thr: L22

Mutant	KRVPS T DVLCN	805
Homo	KRVPS I DVLCN	805
Mus	KRVPS I DVLCN	805
Rattus	KRVPS I DVLCN	805
Pan	KRVPS I DVLCN	805
Macaca	KRVPS I DVLCN	805
Canis	KRVPS I DVLCN	805
Equus	KRVPS I DVLCN	805
Sus	KRVPS I DVLCN	805
Bos	KRVPS I DVLCN	805
Loxo	KRVPS I DVLCN	805
Gallus	KRVPS I DVLCN	802
Xenopus	KRVPS I DVLCN	801
Drosophila	ERLPS I ITMLR	867

p.Ser1012Pro: T134

Mutant	KTPNF P TLLCY	1017
Homo	KTPNF S TLLCY	1017
Mus	KTPNF S TLLCY	1017
Rattus	KTPNF S TLLCY	1017
Pan	KTPNF S TLLCY	1017
Macaca	KTPNF S TLLCY	1017
Canis	KTPNF S TLLCY	1017
Equus	KTPNF S TLLCY	1017
Sus	KTPNF S TLLCY	1017
Bos	KTPNF S TLLCY	1017
Loxo	KTPNF S TLLCY	1017
Gallus	KTPNF S TLLCY	1014
Xenopus	KTPNF S TLLCY	1013
Drosophila	KTSNF S TLLCY	1085

Figure S2. Mutated THOC2 amino acid residues are highly conserved. Amino acid sequence alignments flanking the mutated residues (red) at the conserved positions (green) among different orthologs are shown. Sequence alignment of variant amino acids across different species was performed using EBI CLUSTALW server.

Figure S3

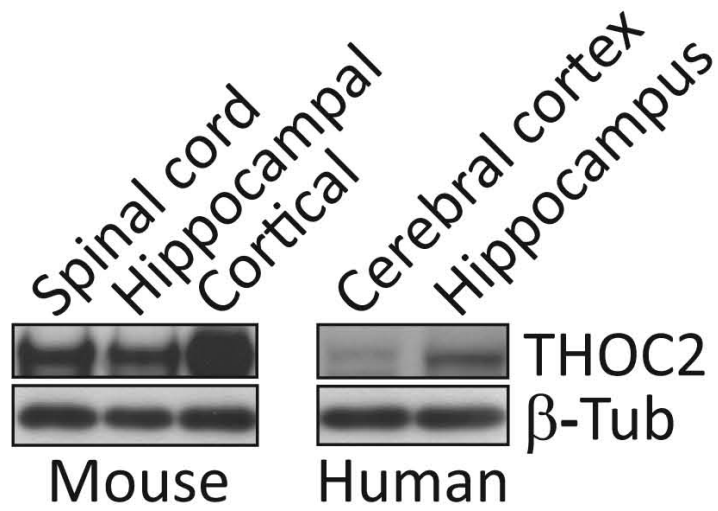


Figure S3. THOC2 expression in mouse hippocampal and cortical neurons and human cerebral cortex and hippocampus.

Figure S4

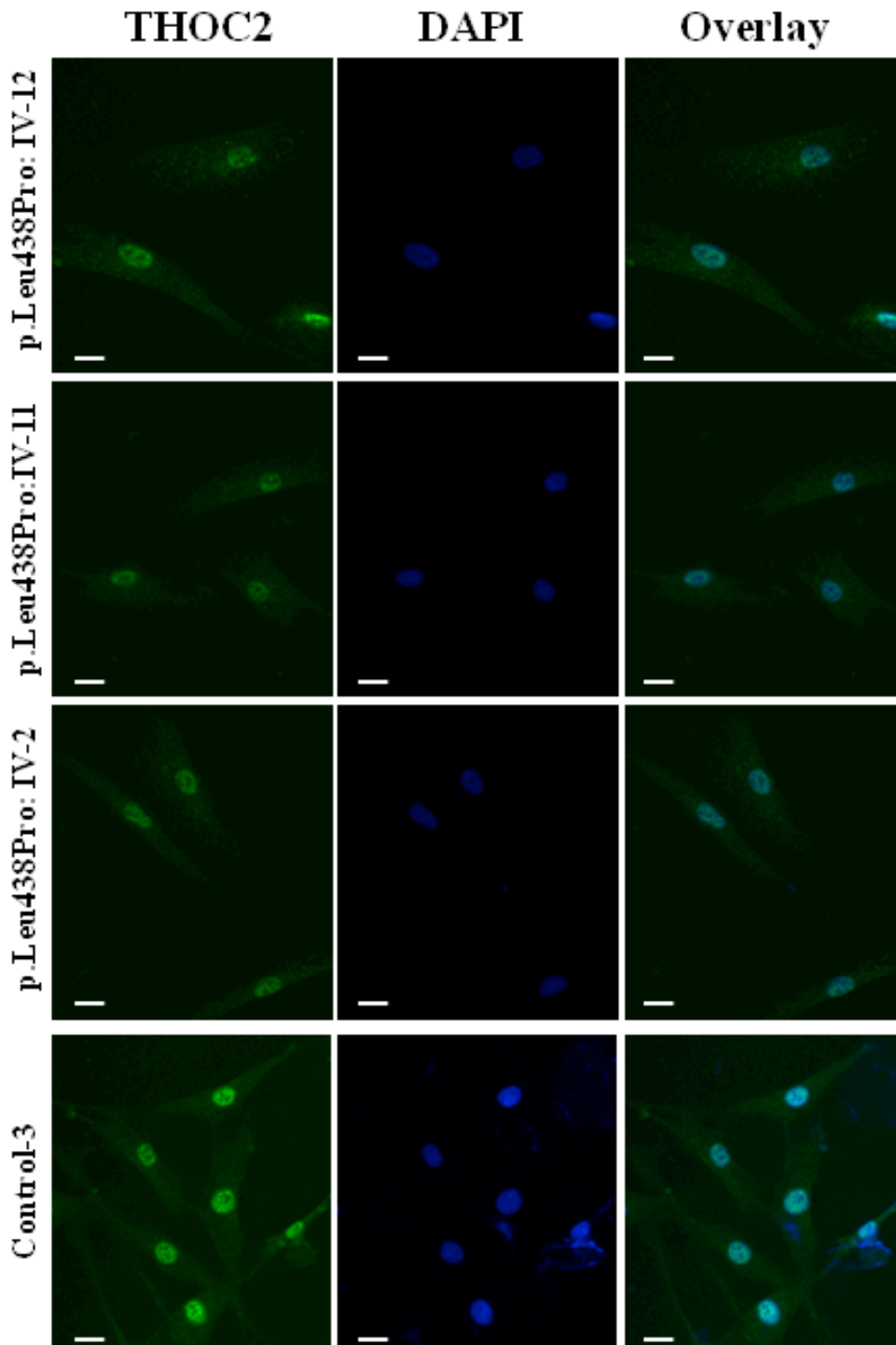


Figure S4. THOC2 immunofluorescence staining is significantly reduced in fibroblasts with the THOC2 p.Leu438Pro mutation compared with a bright speckled staining in the control fibroblasts. Scale bars = 20 μ m. Linked to Figure 3.

Figure S5

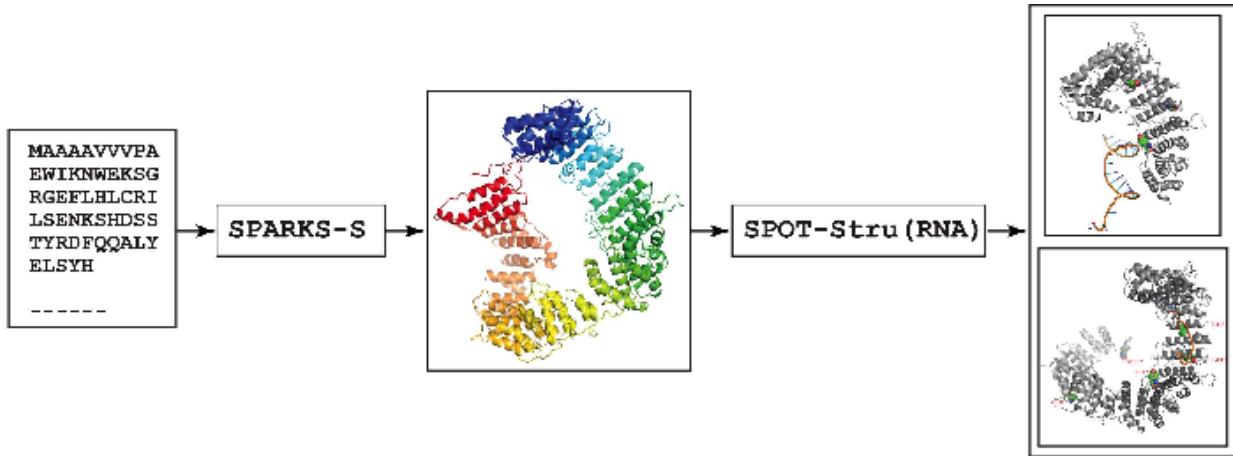


Figure S5. The workflow used for predicting the THOC2-RNA interaction. The 3-dimensional THOC2 structure was predicted from the protein sequence using SPARKS-X (ref 25 of the manuscript) that employs experimentally determined protein structure as templates to build structural models. Whole Protein Data Bank (PDB) was considered as a template library and scanned by matching sequence profiles and structural profiles of query and known structure proteins. Here, the sequence profiles include the information generated from multiple sequence alignment. The structural profiles include actual properties for templates and predicted properties for query protein. The structural properties contain secondary structure, solvent accessible surface area, and main-chain dihedral angles.

Supplementary Tables

Table S1. Pathogenicity predictions by different methods

Table S2. Summary of clinical features

Table S3. Primer sequences and Taqman probes.