Supplementary Methods

DNA sequencing

DNA was extracted from whole blood using the Qiagen Gentra Puregene Blood Kit according to manufacturers protocols and the DNA quantified by Nanodrop, Qubit, and agarose gel electrophoresis. Whole exome sequencing was performed using the SureSelect XT Human All Exon v5 (Agilent, Santa Clara, CA, USA) and Illumina HiSeq 2000/2500 by Otogenetics Corporation to obtain 59 million 100bp paired-end reads for each individual. Reads were mapped to the human GRCh37.p13 reference using bwa (v 0.7.4) (Li and Durbin 2009), duplicates marked with Picard (v 1.96) followed by indel realignment and base quality score recalibration using the Genome Analysis Toolkit (GATK, v 3.4-46). SNP and indel variants were identified using GATK HaplotypeCaller, and GenotypeGVCFs as described (Van der Auwera et al 2013). Mean target coverage was 82 for each sample. 216 high-quality variants shared by both siblings, predicted to affect protein primary structure of 134 ataxia candidate genes, and observed in fewer than 0.1% of the 1000 Genomes, HapMap, genome aggregation database (gnomAD, (Lek et al 2016), accessed August 2017), or our in-house dataset, were considered for disease causality. Sequencing depth for the c830T>C variant was 44 and 58, while the c.1506+1G>A splice variant was covered by 106 and 89 reads, for the proband and her sister, respectively.

RNA sequencing

Blood was collected in PAXgene Blood RNA Tubes and RNA extracted using the Qiagen PAXgene Blood RNA kit according to standard manufacturers' protocols. Samples were quantified using the Agilent 2100 Bioanalyzer and the Qubit 2.0. RNA sequencing was performed using llumina HiSeq HT chemistry to obtain 62

million 100bp paired end reads at the Australian Genome Research Facility (AGRF). Bioinformatic analysis was also performed at AGRF and included read alignment to hg19 with TopHat (version 2.0.14), transcript assembly (Stringtie 1.0.4), and differential expression analysis as described (Trapnell et al 2013).

Aligned RNAseq fragments were assigned to the mother's mutant APOQ8 allele if they included the novel c.1506+1G>A splice site mutation (genomic position 227,172,357) and/or the c.1440C>T (p.Phe480=) polymorphism (rs12593, global MAF 0.35, genomic position 227,172,290); RNAseq fragments including the rs12593 C allele were assigned to her wild-type APOQ8 allele. To exclude fragments potentially derived from contaminating genomic DNA, only RNAseq fragments spliced at their 5'- or 3'-termini were included in this analysis.

PCR conditions

Variants were amplified by PCR and confirmed by Sanger sequencing in both affected children and parents. Both *COQ8A* variants could be amplified in a single reaction using primers: *COQ8A F*: 5' CCGTGGTGACCCCTCTTG 3' and *COQ8A R*: 5' AGGATCCCCACTCCCAGA 3'.

Supplementary References

Lek M, Karczewski KJ, Minikel EV, et al (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536: 285-291.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* 25: 1754-1760.

Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotech* 31: 46-53.

Van der Auwera GA, Carneiro MO, Hartl C, et al (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current protocols in bioinformatics* 43: 11.10.11-33.