

Supplementary Information

Identification of a novel non-coding deletion in Allan-Herndon-Dudley syndrome by HiFi long-read genome sequencing

Jihoon G. Yoon, Seungbok Lee, Soojin Park, Se Song Jang, Jaeso Cho, Man Jin Kim, Soo Yeon Kim, Woo Joong Kim, Jin Sook Lee, and Jong-Hee Chae

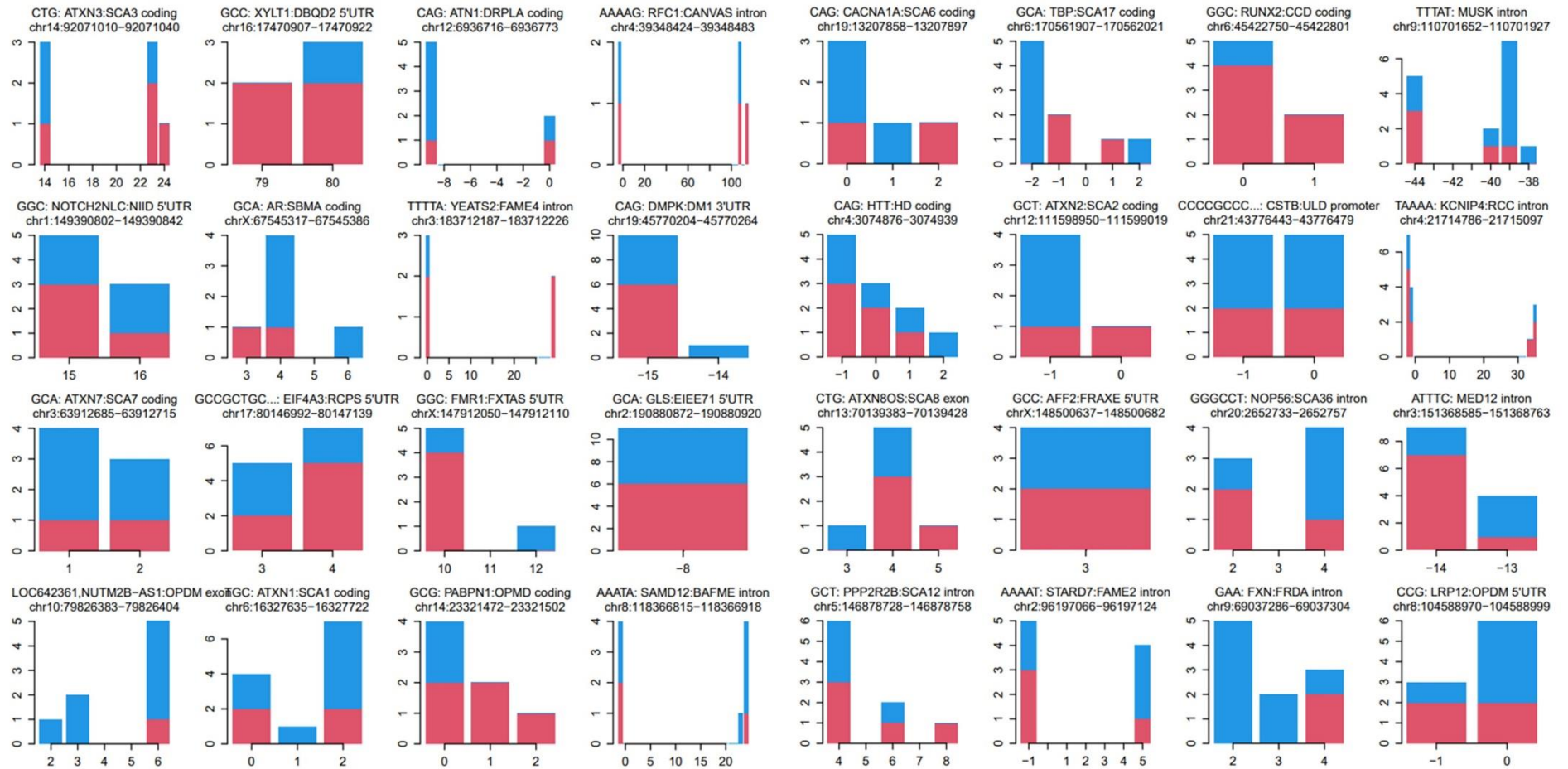
< Table of Contents >

Figure S1. Analysis of known pathogenic short-tandem repeat locus	2
Figure S2. Assessment of the <i>SLC16A2</i> intronic deletion's effect on mRNA splicing	6
Figure S3. In silico identification of RORA and RORC transcription factor binding sites in conserved genomic regions	7
Figure S4. Identification of overlapping deletions in gnomAD SV	8
Table S1. Catalog of homozygous and hemizygous structural variations in OMIM genes observed in siblings	9

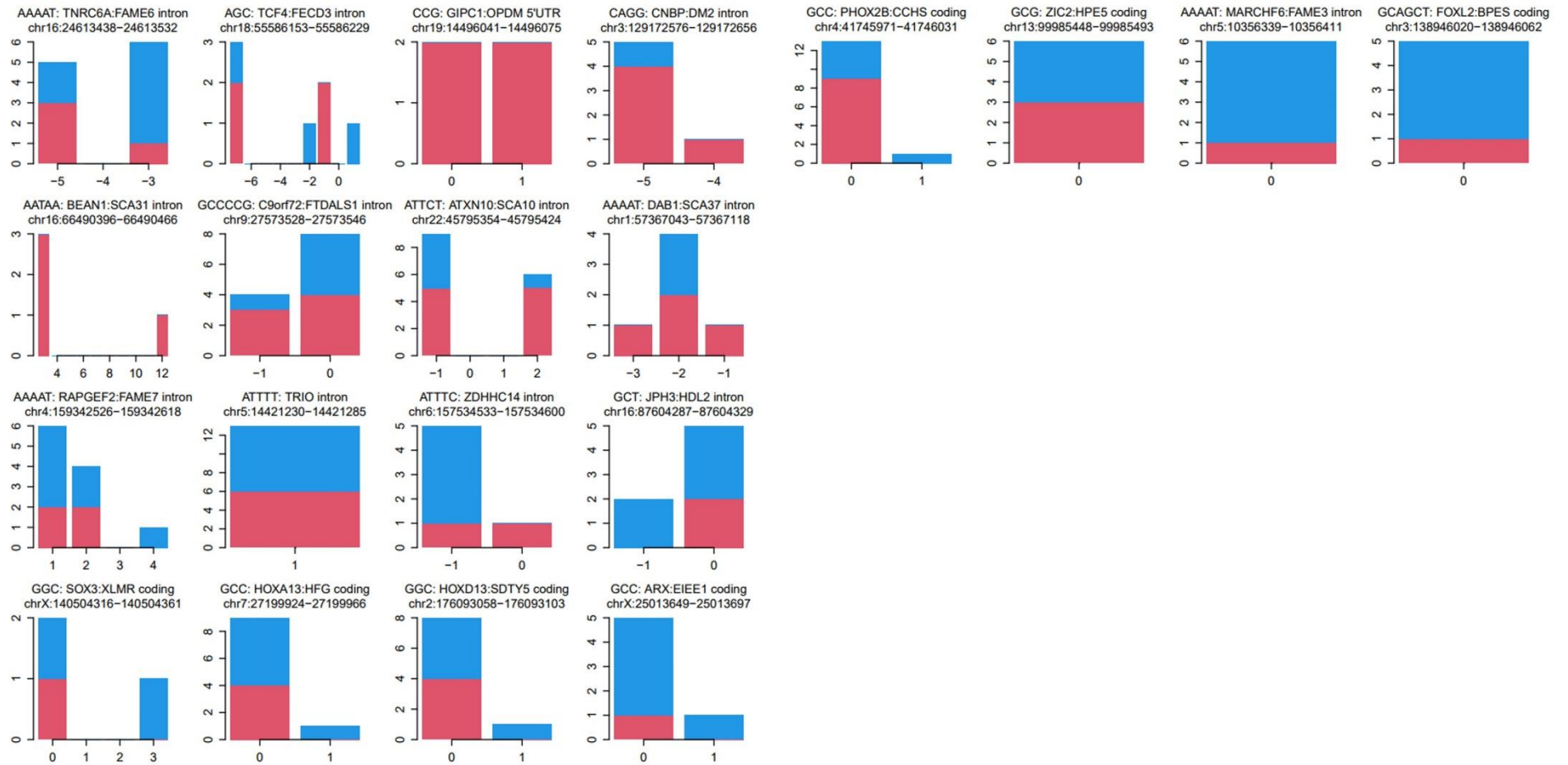
Figure S1 | Genotyping analysis of pathogenic short-tandem repeat loci in siblings.

This figure presents the genotyping results from long-read sequencing for 52 known pathogenic short-tandem repeat (STR) sites in two siblings. Results for the second child (II:2, **A**) and the third child (II:3, **B**) shows no pathogenic repeat expansions in the known STR disorders. The bars display the number of repeat units (y-axis) for each allele observed at that locus, with the x-axis indicating the specific allele sizes detected. Red and blue bars represent reads with forward and reverse strands, respectively.

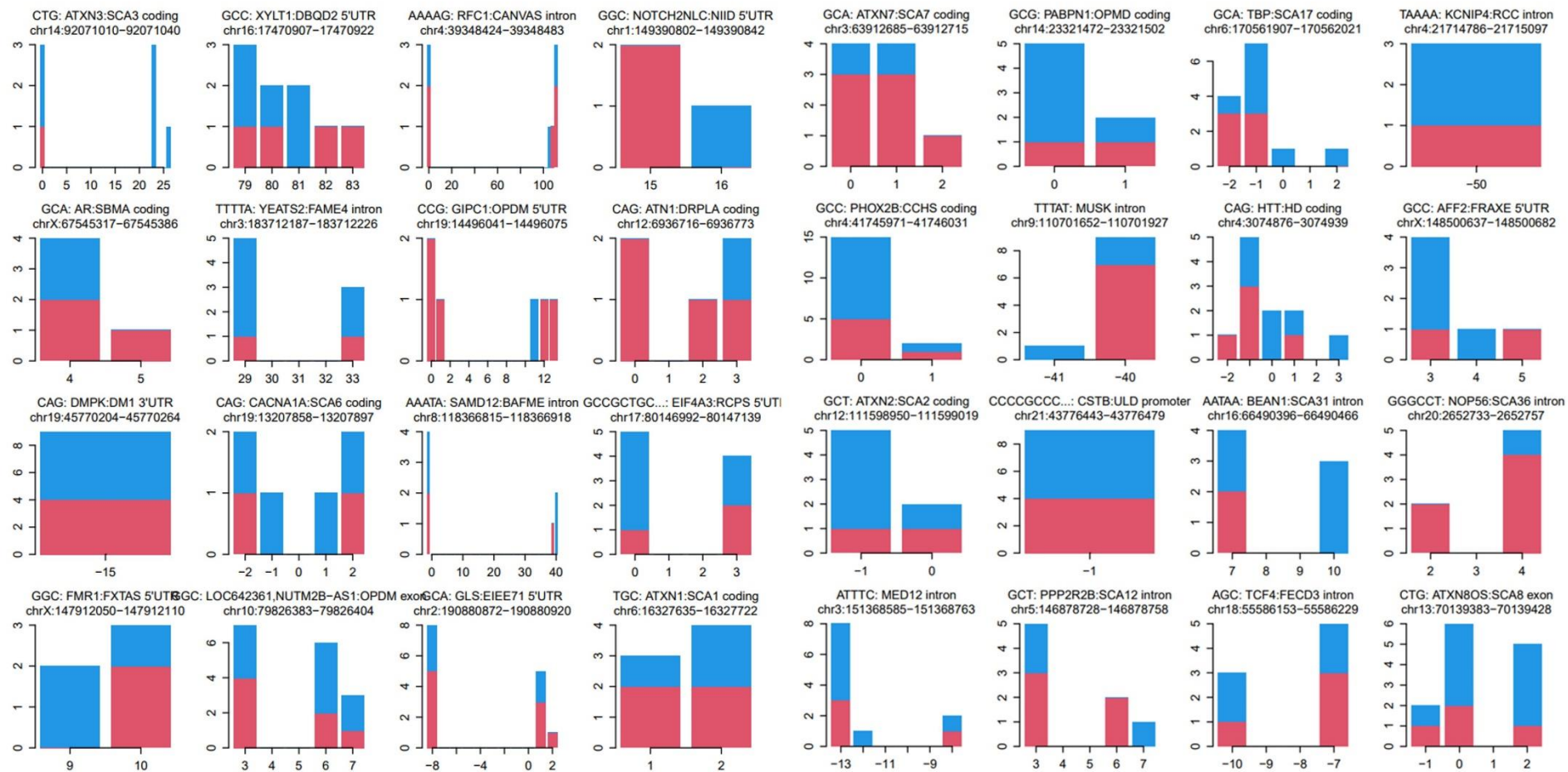
A



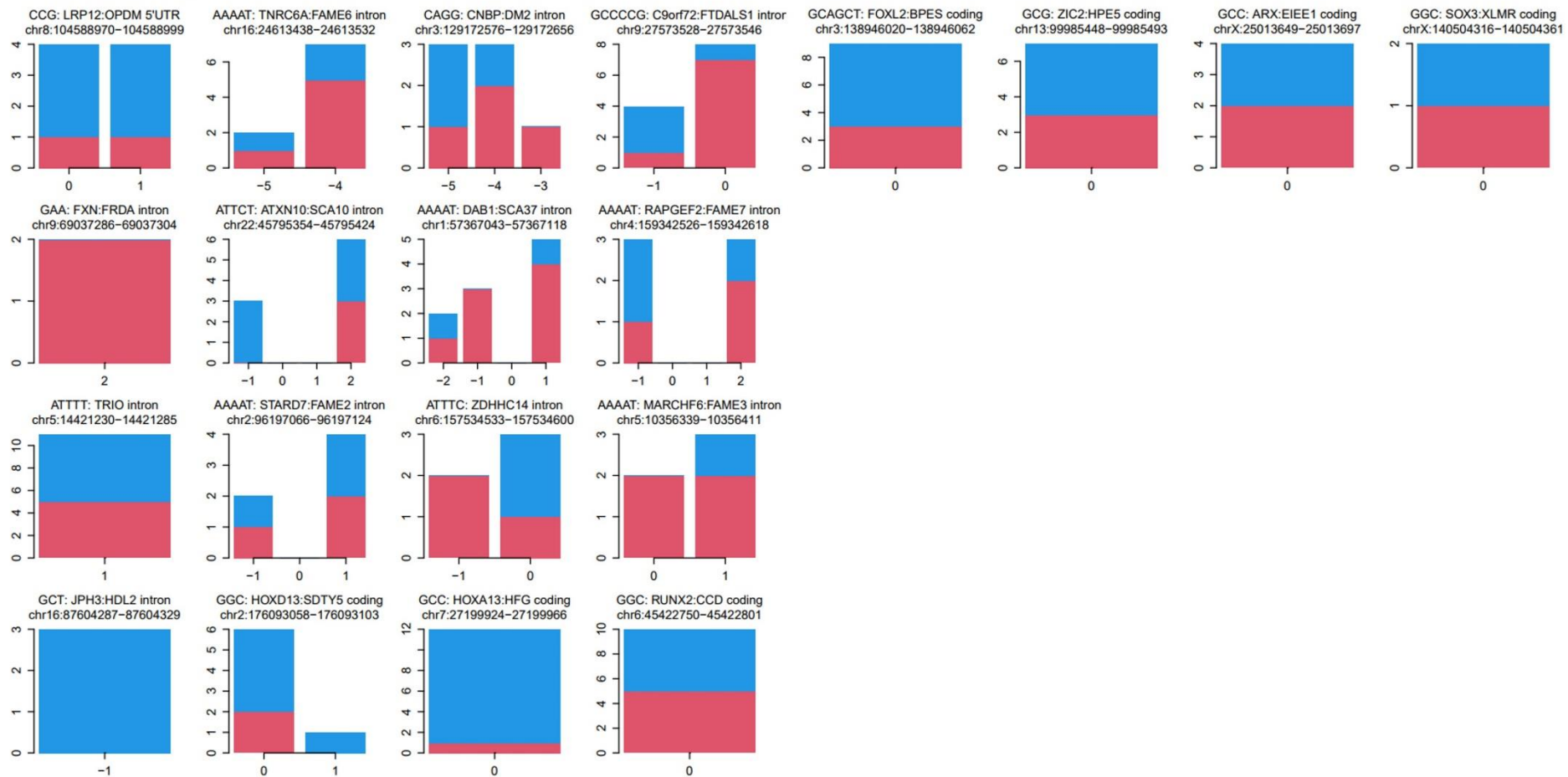
A (continued)



B



B (continued)



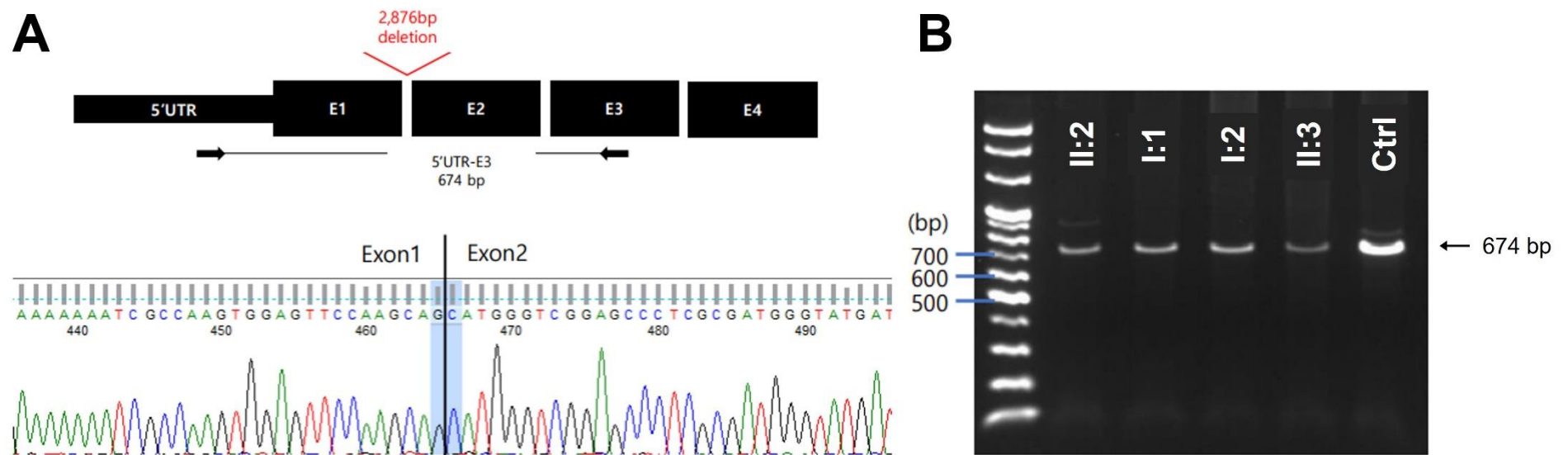


Figure S2 | Assessment of the *SLC16A2* intronic deletion's effect on mRNA splicing.

A. This panel displays a schematic of the *SLC16A2* gene indicating the location of the 2,876 bp intronic deletion and the exons involved. The chromatogram beneath illustrates the boundary between Exon 1 and Exon 2, with no visible signs of splicing anomalies. **B.** This panel depicts an agarose gel electrophoresis of RT-PCR products from blood-derived mRNA, comparing the splicing patterns of siblings II:2 and II:3 with a control. The expected 674 bp band is present in all samples, consistent across both affected siblings and the control, indicating the absence of aberrant splicing events between Exon 1 and Exon 2 attributable to the intronic deletion. This suggests that the intronic deletion does not disrupt the normal splicing of the *SLC16A2* gene in the blood samples tested.

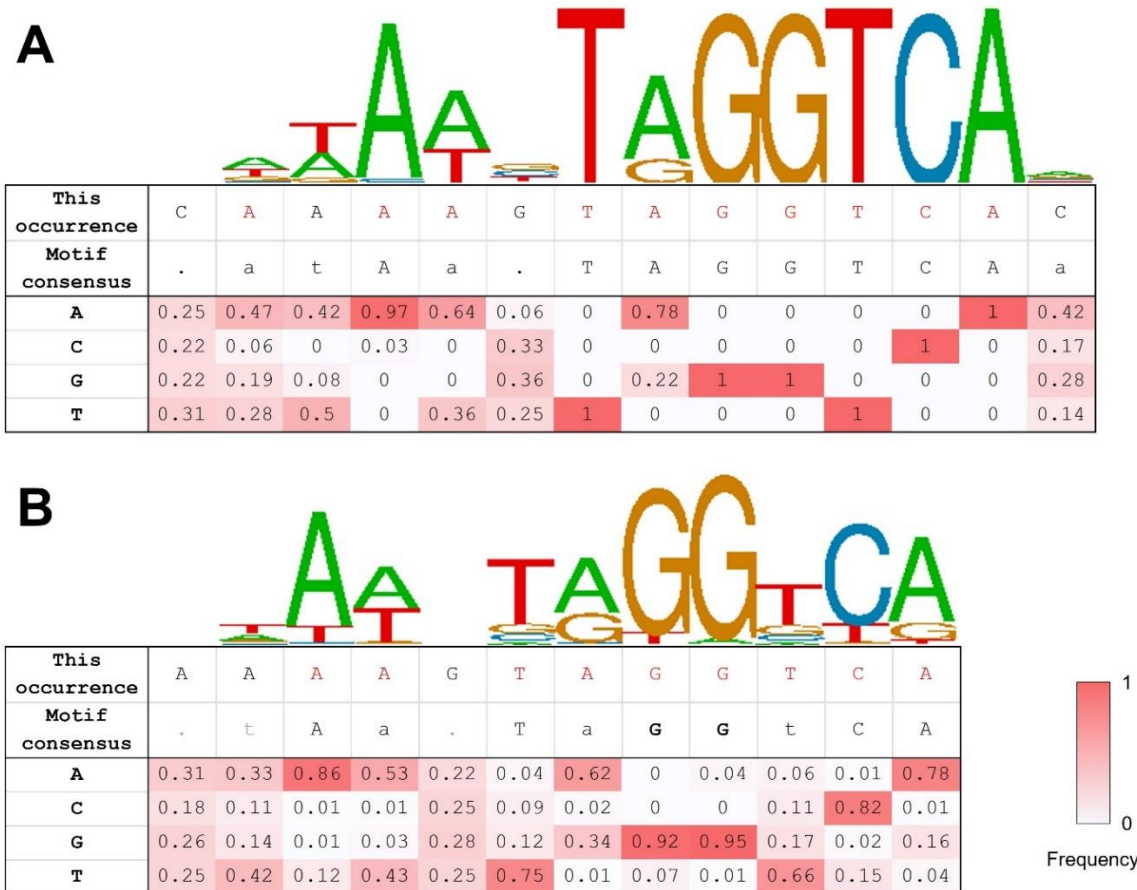


Figure S3 | In silico identification of RORA and RORC transcription factor binding sites in conserved genomic regions.

Sequence motifs derived from the JASPAR database representing predicted binding sites for (A) RORA and (B) RORC within evolutionarily conserved regions of the deleted region. Both motifs are visualized using sequence logos, where the height of each letter correlates with the nucleotide frequency at that position, and the overall height of the stack indicates the sequence conservation. The accompanying tables provide the nucleotide occurrence frequency, with the consensus sequence noted below each table. Red shading highlights nucleotides with a high frequency of occurrence, underscoring their potential importance in transcription factor affinity and specificity. **A.** RORA (RAR-related orphan nuclear receptor alpha), a key regulator implicated in cerebellar development and linked to the mendelian disorder characterized by intellectual developmental disorder with or without epilepsy or cerebellar ataxia (OMIM #618060). The depicted motif corresponds to the chrX:74461608-74461621 region (JASPAR ID: MA0072.1), with a prediction score of 621, indicating a high likelihood of RORA binding. **B.** RORC (RAR-related orphan nuclear receptor gamma), the motif represents the chrX:74461609-74461620 region (JASPAR ID: MA1151.1), with a prediction score of 632, suggesting a probable RORC binding site.

SLC16A2 solute carrier family 16 member 2

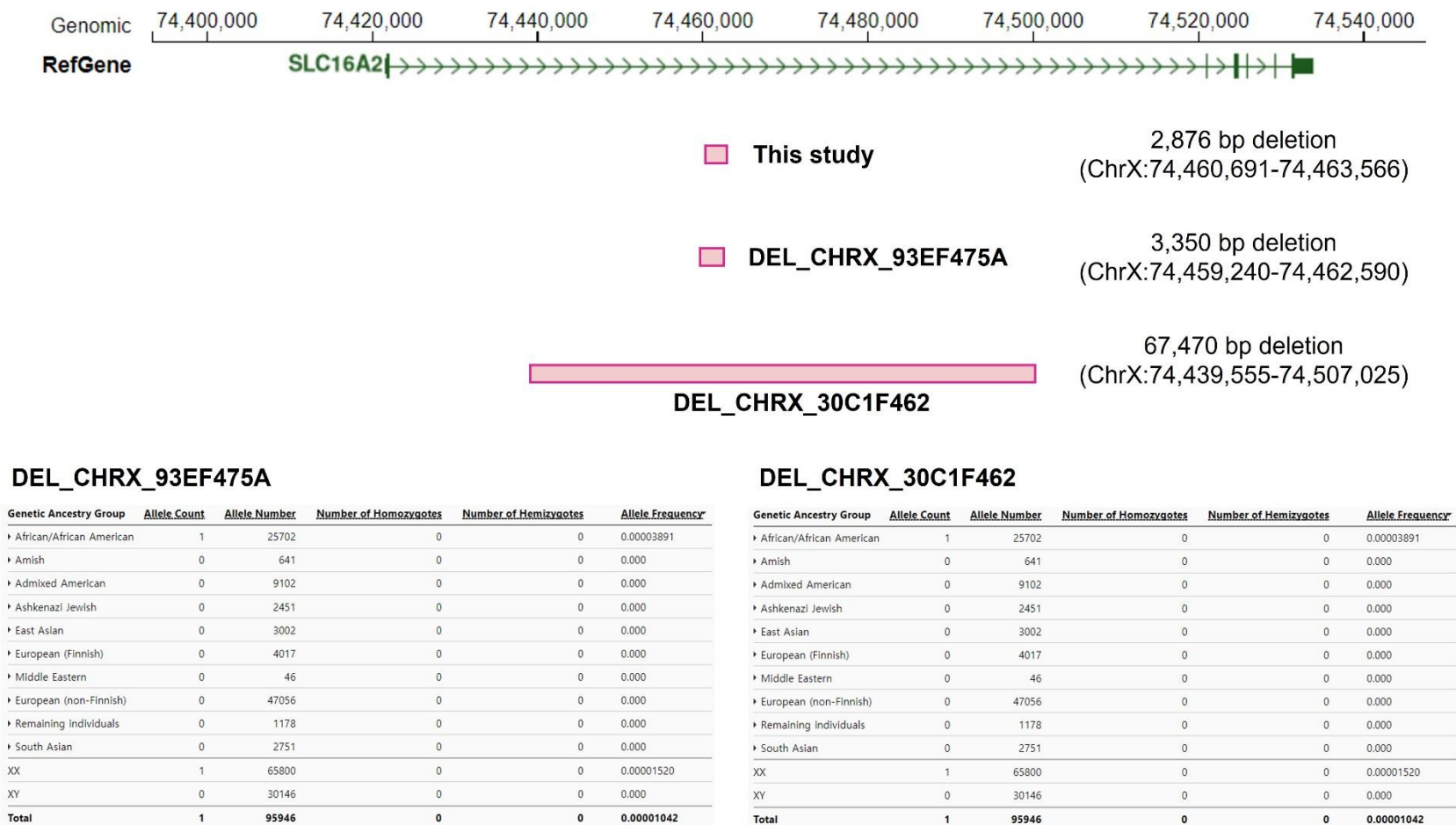


Figure S4 | Identification of overlapping deletions in gnomAD SV.

This figure illustrates two distinct overlapping deletions, along with the intronic deletion found in our study, as cataloged in the Genome Aggregation Database Structural Variant database (gnomAD SV v4.0, available at <https://gnomad.broadinstitute.org/>). Both deletions were in singletons and identified in female carriers of African descent. The data indicate an absence of hemizygous deletion overlapping with the deletion found in our study within the healthy male population of 15,073 individuals reported in the gnomAD SV database.

Table S1 | Catalog of homozygous and hemizygous structural variations in OMIM genes observed in siblings

Chr	Start	End	Length (bp)	SV type	DP (II:2)	DP (II:3)	Gene	OMIM	MOI
3	190380130	190380131	70	INS	7	11	<i>CLDN16</i>	Hypomagnesemia 3, renal (#603959)	AR
16	8711712	8711713	55	INS	4	5	<i>ABAT</i>	GABA-transaminase deficiency (#137150)	AR
17	893648	893794	146	DEL	5	5	<i>NXN</i>	Robinow syndrome, autosomal recessive 2 (#612895)	AR
X	47148831	47148893	62	DEL	11	4	<i>RBM10</i>	TARP syndrome (#300080)	XLR
X	74460690	74463566	2,876	DEL	13	11	<i>SLC16A2</i>	Allan-Herndon-Dudley syndrome (#300095)	XLR

Abbreviations: Chr, chromosome; DP, depth; OMIM, Online Mendelian Inheritance in Man; MOI, mode of inheritance; AR, autosomal recessive; XLR, X-linked recessive; INS, insertion; DEL, deletion;