Universiteit Antwerpen

This item is the archived peer-reviewed author-version of:

Loss-of-function of activity-dependent neuroprotective protein (ADNP) by a splice-acceptor site mutation causes Helsmoortel-Van der Aa syndrome

Reference:
D' Incal Claudio, Annear Dale, Elinck Ellen, van der Smagt Jasper J., Alders Marielle, Dingemans Alexander J.M., Mateiu Ligia, de Vries Bert B.A., Vanden Berghe Wim, Kooy Frank.- Loss-of-function of activity-dependent neuroprotective protein (ADNP) by a splice-acceptor site mutation causes Helsmoortel-Van der Aa syndrome
European journal of human genetics / European Society of Human Genetics - ISSN 1476-5438 - London, Springernature, 32(2024), p. 630-638
Full text (Publisher's DOI): https://doi.org/10.1038/S41431-024-01556-4
To cite this reference: https://hdl.handle.net/10067/2048390151162165141

uantwerpen.be
Institutional repository IRUA

Reseach paper

Loss-of-function of Activity-Dependent Neuroprotective Protein (ADNP) by a Splice-Acceptor Site Mutation causes Helsmoortel-Van der Aa Syndrome.

Claudio Peter D'Incal[1,2], Dale John Annear[1], Ellen Elinck[1], Jasper J. van der Smagt[3], Mariëlle Alders[4], Alexander J. M. Dingemans[5,6], Ligia Mateiu[1], Bert B. A. de Vries[5], Wim Vanden Berghe[2,*] and R. Frank Kooy[1,*]

[1] Cognitive Genetics (CONGET), Centre for Medical Genetics (CMG), Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium Department of Medical Genetics.

[2] Protein Chemistry, Proteomics and Epigenetic Signaling (PPES), Department of Biomedical Sciences, University of Antwerp, Antwerp, Belgium.

[3] Division of Laboratories, Pharmacy and Biomedical Genetics, Section Clinical Genetics, University Medical Center Utrecht, the Netherlands and Rijksuniversiteit Utrecht, Utrecht, the Netherlands

[4] Department of Human Genetics, Amsterdam Reproduction & Development Research Institute, Amsterdam University Medical Centers, University of Amsterdam Amsterdam, The Netherlands

[5] Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands

[6] Department of Artificial Intelligence, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

Abstract

Mutations in *ADNP* result in Helsmoortel-Van der Aa syndrome. Here, we describe the first *de novo* intronic deletion, affecting the splice-acceptor site of the first coding *ADNP* exon in a five-year-old girl with developmental delay and autism. Whereas exome sequencing failed to detect the non-coding deletion, genome-wide CpG methylation analysis revealed an episignature suggestive of a Helsmoortel – Van der Aa syndrome diagnosis. This diagnosis was further supported by PhenoScore, a novel facial recognition software package. Subsequent whole genome sequencing resolved the three-base pair *ADNP* deletion c.[-5-1_-4del] with transcriptome sequencing showing this deletion leads to skipping of exon 4. An N-terminal truncated protein could not be detected in transfection experiments with a mutant expression vector in HEK293T cells, strongly suggesting this is a first confirmed diagnosis exclusively due to haploinsufficiency of the *ADNP* gene. Pathway analysis of the methylome indicated differentially methylated genes involved in brain development, the cytoskeleton, locomotion, behavior, and muscle development. Along the same line, transcriptome analysis identified most of the differentially expressed genes as upregulated, in line with the hypomethylated CpG episignature and confirmed the involvement of the cytoskeleton and muscle development pathways that are also affected in patient cell lines and animal models. In conclusion, this novel mutation for the first time demonstrates that Helsmoortel – Van der Aa syndrome can be caused by a loss of function mutation. Moreover, our study elegantly illustrates the use of EpiSignatures, WGS and Phenoscore as novel complementary diagnostic tools in case a of negative WES result.

*Corresponding authors:

R. Frank Kooy, PhD
Department of Medical Genetics
University of Antwerp
Prins Boudewijnlaan 43/6
2650 Edegem
Belgium

Wim Vanden Berghe, PhD
Epigenetic Signaling lab (PPES)
Department of Biomedical Sciences
University of Antwerp
Universiteitsplein 1
2610 Wilrijk

Tel: + 32 3 275 9760 Tel: +32 3 2652657

Email: frank.kooy@uantwerpen.be Email: wim.vandenberghe@uantwerpen.be

INTRODUCTION

Heterozygous *de novo* mutations in the *ADNP* gene cause Helsmoortel-Van der Aa syndrome (HVDAS; OMIM #615873), a prevalent neurogenetic condition associated with autism, intellectual disability, developmental delays, and multi-organ deficiencies. The ADNP protein is a key transcription factor, involved in neuronal tube closure during embryogenesis[1] and differentially interacts with genes regulating chromatin[2]. It contains nine zinc fingers, a DNA-binding homeobox domain, an ARKS motif as well as a PxVxL sequence. The presence of a nuclear localization signal (NLS) is in line with its nuclear localization in neuronal cells. ADNP can bind the SWI/SNF (BAF) complex[3], is part of the repressive ChAHP complex[4], and is involved in expression regulation of hundreds of additional genes[5].

The mutational mechanism of the Helsmoortel-Van der Aa syndrome is not yet fully understood. Almost all patients are diagnosed with nonsense and frameshift stop mutations in the last coding exon 6 (referred to as exon 5 in earlier nomenclature), of which we have previously demonstrated to escape nonsense-mediated decay[6,7] and thus, theoretically, are able to produce truncated ADNP proteins. Even a much rarer mutation in the penultimate exon 5 was reported to still be translated[8]. Both, a complete loss-of-function of the protein as well as a potential gain-of-function of the mutant proteins, if still produced, are thus theoretically possible causes of the disease. Circumstantial evidence pleads against the loss-of-function hypothesis. Different *ADNP* gene mutations were found to elicit distinct genome-wide epigenetic profiles depending on the localization of the *ADNP* gene mutation. For example, mutations at the extremities of the gene result in an overall hypomethylated CpG pattern, whereas mutations in the central region of the gene result in an overall hypermethylated CpG pattern [9,10]. These methylation patterns are coupled to a differential clinical presentation, with mutations in the central region resulting in a more severe phenotypic presentation[6,11]. Discriminative phenotypic expression depending on the location of the mutation seems not compatible with a simple loss-of-function mechanism. However, we observe evidence against a gain-of-function hypothesis as we have never been able to detect intact mutant ADNP protein expression despite intense experimental efforts[12].

We here present a first case of proven Helsmoortel – Van der Aa syndrome with a complete absence of the ADNP protein. It concerns a five-year-old girl with intellectual disability, autism, and speech and motor

delays, where clinical exome sequencing failed to identify the genetic event. However, through subsequent whole-genome sequencing paralleled with genome-wide CpG methylation analysis and facial recognition software an intronic deletion affecting the splice-acceptor site of the *ADNP* gene was revealed. This case is thus also an elegant illustration of the added value of whole genome sequencing, facial recognition and epigenetic analysis in case of negative whole exome diagnosis[13,14].

MATERIAL AND METHODS

Patient and tissue collection

Trio-open whole exome sequencing (WES) was conducted at UMC Utrecht (Genetics department of Prof. Dr. J.P. van Tintelen). In addition, whole genome sequencing (WGS) in parallel with the human EPIC BeadChip Array (Illumina; California , U.S.) were executed, revealing a heterozygous *de novo* deletion upstream of the ATG start codon of the *ADNP* gene. Clinical information of the five-year-old girl was received from the parents (caregivers) and tending clinicians under informed consent. Upon both consenting parents, blood was drawn and collected in Blood collection tubes, EDTA, BD Vacutainer® with BD Hemogard™ closure (VWR; Pennsylvania, U.S.), and PAXgene® Blood RNA Tube (Qiagen; Hilden, Germany) at the University Hospital of Antwerp (UZA) and approved of the Ethics Committee of the Antwerp University Hospital/University of Antwerp. The mutation was confirmed by Sanger sequencing using the forward primer F: 5'-TCTTGGCCACTGACACAAAG-3' and R: reverse primer 5'-GGAAGGAAGGATGGATGGAT-3'.

Quantifying phenotypic similarity using PhenoScore.

PhenoScore was used to determine whether the individual with the splice-acceptor site mutation suited the molecular diagnosis of Helsmoortel-Van der Aa syndrome[11]. PhenoScore is an artificial intelligence-based phenomics framework that combines state-of-the-art facial recognition technology with analysis of phenotypic data in Human Phenotype Ontology (HPO) terms to quantify phenotypic similarity. Interestingly, depending on the position of the mutation, two clinical subgroups could be discriminated by PhenoScore, differing in disease presentation and severity. These two categories coincide with two partially opposing genome-wide methylation patterns in patients [1–3].. These so called Episignatures are commonly referred to as associated with Class I and Class II mutations, the latter being on average more severely affected. To support a Helsmoortel-Van der Aa

syndrome diagnosis, an HPO score of 0.1 icorrelates with hypomethylation signature (Class I ADNP mutations), whereas a score closer to 1.0 correlates with towards a hypermethylation signature (Class II ADNP mutations).

***In silico* pathogenicity prediction**

Splice variant predictions were performed using Alamut™ splicing software (Sophia Genetics; Lausanne, Switzerland) according to standard procedures (https://www.sophiagenetics.com/).

RNA sequencing, differential splicing, and gene expression analysis.

Total RNA was extracted from whole blood collected PAXgene® Blood RNA Tube (Qiagen; Hilden, Germany) using the PaxGene Blood RNA Kit (Qiagen; Hilden, Germany), according to the manufacturer's protocol. The RNA concentration was estimated with the Qubit™ RNA Broad Range Assay kit (Invitrogen; Massachusetts, U.S.). RNA purity was assessed by determining the 260/280 ratio using the NanoDrop™ 2000/2000c Spectrophotometer (Thermo Scientific™; Massachusetts, U.S.). The RNA integrity was evaluated using the Agilent RNA Screentape Assay on the Agilent 2200 TapeStation System (Agilent; California, U.S.). Samples with the highest RIN score (RIN > 6.5) were sent for total transcriptome sequencing (Novogene; Cambridge, UK). All sequencing data was mapped to the human annotated genome GRCh38.p13 (Ensembl v108). The aligned and sorted bam files after STAR alignment were used for alternative splicing (AS) analysis using rMATS (v4.1.2) in linux, followed up by maser (R package), and visualized as Sashimi plots in Integrative Genome Browser (IGV). Both reads spanning exon junctions and reads covering single exons were used for splicing quantification. We tested for multiple splicing events (p-value < 0.05 and FDR < 0.05), including skipped exon (SE), alternative 5' splice site (A5SS), alternative 3' splice site (A3SS), mutually exclusive exons (MXE), and retained introns (RI). Variant detection (SNPs) was done for each sample following the GATK best practices for RNA-seq data. We quantified ADNP transcript abundance (transcripts per million; TMP) from RNA-seq reads using Salmon (https://combine-lab.github.io/salmon/). Differential gene expression analysis was performed with NOISeq (R package), a non-parametric method for comparison of samples without biological replicates which reports the log2-ratio of the two conditions (M) and the value of the difference between conditions (D). A gene is considered to be differentially expressed if its corresponding M and D values are likely to be higher than in noise (q > 0.95). The genes having an adjusted p-value <0.05, FDR < 0.05, and absolute value of log2FC >=0.5 were considered biologically relevant and further analyzed for functional enrichment (clusterProfiler R package with fGSEA

function for the geneset enrichment analysis and enrichGO for overrepresentation analysis in GO ontologies and KEGG pathways). Additional data visualization was supported by BigOmics, a user-friendly and interactive self-service bioinformatics platform for the in-depth analysis, visualization, and interpretation of transcriptomics data [18]. RT-PCR was used to confirm a selection of genes with log2FC >=0.5 from the RNA sequencing experiment by converting 1 µg of total extracted RNA to cDNA using the SuperScript™ III Reverse Transcriptase kit (Invitrogen™; Massachusetts, U.S.). Primer efficiencies were optimized using a standard dilution curve method on pooled cDNA samples from controls and patients (90% > E > 110%). RT-PCR was performed in triplicate using the CFX384 Touch Real-Time PCR Detection System (BioRad; California, U.S.) with primers listed in **supplementary table T1** using the Takyon™ No ROX SYBR 2X MasterMix (Eurogentec; Seraing, Belgium). Reference gene stability was assessed using the geNorm method in qbase+ (Biogazelle; Ghent, Belgium), after which were selected for normalization. Data analysis was performed in qbase+ (Biogazelle; Ghent, Belgium) with a maximum deviation of 0.5 per triplicate using the stable housekeeping genes *B2M, GAPDH* and *RPL13A*. Statistical analysis was performed in GraphPad Prism 9.3.1 using a 2way ANOVA with Šídák's multiple comparisons test.

Plasmid constructs

The pCMV3 expression vector encoding human wild-type ADNP fused to an N-terminal GFPSpark®-tag was purchased (Sino Biological; Beijing, China). The pCMV3 expression vector encoding human mutant ADNP, incorporating the deletion of exon 4, fused to a C-terminal OFPSpark®-tag was synthesized (Sino Biological; Beijing, China) (**supplementary data S1**). Transformation of One Shot™ TOP10 Chemically Competent *E. coli* cells was performed according to standard procedures (Invitrogen™; Massachusetts, U.S.). DNA was purified using the NucleoSpin Plasmid EasyPure Mini kit (Macherey Nagel; Düren, Germany) according to the manual. Deletion of exon 4 was verified using Sanger sequencing.

Cell culture

HEK293T cells were purchased (ATCC; Virginia, U.S.) at low passage number and cultured in DMEM (Gibco™; Massachusetts, U.S.), supplemented with 10% fetal bovine serum (Gibco™; Massachusetts, U.S.) and 1% penicillin/streptomycin (Gibco™; Massachusetts, U.S.). Cells were grown in a humidified 37%O2/5%CO2 incubator to reach optimal confluency.

Transient Transfection

HEK293T were transfected using 5 µg wild type and mutant ADNP expression vectors using Lipofectamine™ 3000 Transfection Reagent (Invitrogen™; Massachusetts, U.S.) in accordance with the manufacturer's protocol. Transfection efficiency was about 70% in line with the manufacturer's tested performance. Cells were harvested after 24 hours of incubation for western blotting as described by previous methods [12].

Human methylation EPIC BeadChip array and data processing

Total DNA was isolated from whole blood of the affected five-year-old girl and her unaffected parents using the DNeasy Blood and Tissue Kit (Qiagen; Hilden, Germany) according to the manufacturer's instructions. Subsequently, bisulfite conversion of 250 ng isolated DNA was performed using the EZ DNA Methylation Kit (Zymo Research, California , U.S.). To confirm successful bisulfite conversion, a methylation-conserved fragment of the human *SALL3* gene was amplified using the following primers: 5'-GCGCGAGTCGAAGTAGGGC-3' as a forward primer and 5'-ACCCAACGATACCTAATAATAAAACC-3 as a reverse primer with the PyroMark PCR kit (Qiagen; Hilden, Germany). Amplified products were separated on a 1.5% agarose gel stained with GelRed® Nucleic Acid Gel Stain (Biotium; California , U.S.). The TrackIt™ 100 bp DNA Ladder (Invitrogen; Massachusetts, U.S.) will be used as a reference marker. Bisulfite-converted samples were hybridized on the Infinium Human Methylation EPIC BeadChip (Illumina; California , U.S.) as described in the manufacturer's protocol. EPIC chips were analyzed using the Illumina Hi-Scan system, a platform integrating more than 850,000 methylation sites quantitatively across the genome at a single-nucleotide resolution. Methylation analysis was performed in RStudio version 4.3 using the Sesame package version 3.17 to facilitate the analysis of the EPICv2 array [15]. Raw intensity files were first quality checked and probes with a poor signal were removed. Signal correction was performed for dye bias and background signals. β-values were calculated from the probe signals and probe annotation was carried out using the Illumina Infinium MethylationEPIC v2.0 manifest file (https://support.illumina.com/downloads/infinium-methylationepic-v2-0-product-files.html). All annotations (i.e., CpG islands, shelve, and shore regions) are reported based on the GRCh37/hg19 human genome build. Differentially methylation of each CpG locus was performed with the control samples (i.e., the mother and father of the proband) as baseline. Differential methylated regions (DMRs) were statistically tested using a generalized linear model using Euclidean distance to group CpGs and combining p-values for each segment.

Here, the proband was compared to her parents. The difference in methylation was calculated for genes showing biologically relevant methylation differences, i.e., hypomethylated (Δβ-values < -0.1) and hypermethylated (Δβ-values > 0.1) genes. Enriched biological processes and pathways, together with enriched transcription factors (TRRUST) analyses, associated with the significant differential methylated genes, were conducted using Metascape [16].

Targeted pyrosequencing analysis.

Biologically-relevant hypo- and hypermethylated genes were selected for pyrosequencing confirmation. Briefly, the required primers (i.e., forward, reverse, and sequencing primers) were designed using the PyroMaker Assay Design 2.0 software (Qiagen; Hilden, Germany) according to the manufacturer's instructions (**supplementary table T2**). Bisulfite converted DNA fragments were PCR amplified using the PyroMark PCR kit (Qiagen; Hilden, Germany). Successful PCR amplification was assessed by TBE electrophoresis at 1.5% agarose gel, after which the PyroMark Q24 Instrument (Qiagen; Hilden, Germany) was used to perform pyrosequencing. Biotinylated PCR products were immobilized on streptavidin-coated Sepharose beads (GE Healthcare; Illinois, U.S.), captured by the PyroMark vacuum Q24 workstation, washed and denatured. Single-stranded PCR products were subsequently released into a 24-well plate and annealed to the sequencing primer for 5 min at 80°C. After completion of the pyrosequencing run, results were analyzed using the PyroMark Q24 software (Qiagen; Hilden, Germany). Graphical representation was performed with GraphPad Prism version 9.3.1 using an unpaired student T-test assuming equal variances.

RESULTS

Clinical presentation and molecular diagnosis

A five-year-old girl presents with intellectual and developmental delay, severe speech and motor delay, mild facial dysmorphisms, impaired social behavior, and premature primary tooth eruptions [6,19]. In addition, the patient has a diagnosis of autism and exhibits several clinical characteristics associated with the condition, defined by obsessive behavior such as playing with door handles and keys to calm her down. She also has difficulties in distinguishing between what is self-generated and what is perceived as unfamiliar e.g., she randomly hugs unfamiliar individuals. Such a combination of clinical features is compatible with a diagnosis of the Helsmoortel-Van der Aa syndrome (**supplementary table T3**). PhenoScore, a validated artificial

intelligence-based algorithm integrating facial photographs with Human Phenotype Ontology (HPO), supported her facial features as suitable for the ADNP patient group (score: 0.85/0.09), but could not unambiguously confirm a diagnosis of Helsmoortel – Van der Aa syndrome (score: 0.29) (**supplementary data S2**)[11].

Initially, trio-based whole exome sequencing (WES) failed to identify molecular diagnosis. For this reason, whole-genome sequencing (WGS) was performed, but neither leading to any molecular diagnosis. In parallel, genome-wide epigenetic sequencing revealed an ADNP class I episignature, in line with the clinical presentation [1,2]. Subsequent reanalysis of the genome sequencing data revealed a heterozygous *de novo* c.[-5-1_-4del];[=] p.[?];[(=)] variant in the intronic region upstream of the ATG initiation codon of the first coding exon of the *ADNP* gene (exon 4). This intronic deletion starts at nucleotide position -6, corresponding to five base pairs upstream of the first intronic position proximal of the ATG start codon, to position -4 which lies four base pairs downstream of the start codon (referred to as counting position 1), paralleling the trinucleotide 3'-GAA-5' deletion at position chr20(GRCh38):g.50904000_50904002del within the *ADNP* gene. We confirmed the heterozygous 5'-TTC-3' trinucleotide deletion by Sanger sequencing (**supplementary data S2**).

As the intronic trinucleotide deletion can affect the protein, we performed *in silico* splice site prediction of the unique *ADNP* variant using the Sophia Genetics platform (**supplementary data S3**). The computational tool predicted the three base pair deletion to affect the splice acceptor site by combining results from three different splicing prediction algorithms (i.e., MaxEnt, NNSPLICE, and SFF), each with prediction values of 100% indicating a highly likely impact on splicing.

Differential splicing analysis indicates exon 4 skipping as the mutational event underlying the *ADNP* splice-site mutation.

Splice-site affecting mutations, as identified in the patient, can interfere with the normal splicing process, either resulting in intronic retention, exon skipping, or use of alternative cryptic splice sites[20]. To determine the effect of this specific splice site mutation, we applied whole-transcriptome sequencing on whole blood from the affected five-year-old girl and her unaffected parents. Next, we visualized our RNA-sequencing data using IGV, revealing remarkably more reads in proximity of the *ADNP* mutation (3'-TTC-5') for the patient, confirming that the heterozygous splice acceptor-site mutation is incorporated in the mutant transcript (**supplementary data S4**). We visualized the splice junctions in a Sashimi plot using the aligned RNA sequencing

data and tested for alternative splicing events using rMATS. The observation of the spliced junctions in the patient indicates skipping of the entire exon 4 due to the splice acceptor-site mutation ($p < 0.0001$ and FDR < 0.0001), thereby excluding the *ADNP* initiation ATG start site (**figure 1**).

In addition, we utilized the Salmon tool for the quantification of the expression of *ADNP* transcripts **(supplementary table T4)**. *ADNP* contains nine mRNA splice variants, and its transcripts did not show an overall high expression in blood with the highest number of counts around 15 TPM. We strictly filtered the transcripts with p-value < 0.01 and FDR < 0.01 and identified a decreased expression of the canonical *ADNP* transcript *ADNP-204* (ENST00000396032.8) in the patient (TPM = 1.1) in contrast to preserved expression in her healthy parents (TPM = 3.2 for the mother, respectively 7.3 for the father). Moreover, we observed an increased expression of transcript *ADNP-208* (ENST00000645081.1) in the patient (TPM = 15) opposed to her healthy parents (TPM = 3.7). This upregulated transcript *ADNP-208* does not contain exon 4 as observed for the seven of the *ADNP* splice variants, suggestive of a compensatory mechanism. However, quantifying the global *ADNP* mRNA levels resulted in no differential expression in the patient as opposed to her parents. The splice acceptor-site mutation thus prevents translation of the full *ADNP* transcript from the affected allele in the patient as opposed to her healthy parents.

Aberrant ADNP mRNA splicing indicates haploinsufficiency as the mutational mechanism.

To demonstrate potential transcription and/or translation of the mutated allele, we used PCR mutagenesis to delete the skipped exon (ΔEX4) in a recombinant *ADNP* expression vector, C-terminally fused to an OFPSpark®-tag. We overexpressed this mutant construct along with the wild-type GFPSPark® expression vector in a HEK293T cellular system.

The HEK293T overexpression model allowed to quantify possible amounts of *ADNP* mRNA and protein in cellular extracts by time reverse-transcription PCR and western blotting. To discriminate recombinant from endogenous *ADNP* mRNA levels, we designed a primer set at the 3' region of the expression vector, corresponding to the GFPSpark® fusion protein. Here, a significant increase in recombinant wild-type *ADNP* levels was observed after transfection ($p < 0.0001$; ****), whereas no amplification could be detected in the mutants and untransfected controls. Subsequently, we aimed to detect presence of exogenous mutant ΔEX4 *ADNP* mRNA levels by RT-PCR amplification with a primer set at the 3' region, corresponding to the OFPSpark® fusion protein. Here, we detected mutant *ADNP* mRNA in the OFPSpark® condition only ($p < 0.001$; **), with

no amplification in the untransfected control and wild-type condition (**figure 2A**). Taken together, these results demonstrate the presence of wild-type and mutant *ADNP* mRNA in our cellular system, paralleling our transcriptome sequencing experiment of the ADNP patient.

To exclude the presence of a potential partial *ADNP* transcript from the mutated allele, we investigated alternative translation initiation of the mutant *ADNP* transcript. Recently, six alternative out-of-frame initiation codons were predicted in the *ADNP* gene with an alternative ATG start site at position 229 (M229)[13]. An alternative start at methionine on position 229 would for instance result in an N-terminal truncated ADNP protein[14]. At the protein level, we tested whether wild-type and mutant ΔEX4 ADNP were translated using extensively validated C-terminal and N-terminal ADNP antibodies[12]. While we were able to detect exogenous wild-type ADNP levels (175 kDa) in wild-type overexpression lysates together with endogenous ADNP levels (150 kDa) in all tested conditions, N-terminal abbreviated ADNP protein with a predicted molecular weight of 123 kDa could not be detected in mutant overexpression lysates using a validated C-terminal antibody (**figure 2B**). As an additional control experiment, we showed that N-terminal antibody incubation resulted in the detection of endogenous (150 kDa) and exogenous wild-type ADNP (175 kDa) exclusively (**figure 2C**). In conclusion, these findings confirm the absence of an N-terminal truncated protein, indicating haploinsufficiency as the molecular mechanism underlying the ADNP splice-acceptor site mutation.

The splice-acceptor site mutation in the *ADNP* gene results in a hypomethylation episignature.

Intragenic *ADNP* mutations result in genome-wide methylation differences with mutations at the N-terminus and C-terminus of the ADNP protein showing hypomethylation (Class I mutation), whereas mutations affecting the nuclear localization signal, in the middle part of the protein, entail rather hypermethylation (Class II mutation) [2,4]. Therefore, we conducted a genome-wide CpG methylation analysis to investigate the consequences of the intronic splice-acceptor site mutation.

Here, we showed a significant enrichment of 6,299 differentially methylated probes overcoming a 10% methylation difference in the patient opposed to her parents. Specifically, we found 4,122 CpG probes with hypomethylation ($\Delta\beta < -0.1$), while only 2,177 CpG probes were hypermethylated ($\Delta\beta > 0.1$). In addition, the hypomethylated probes were found to be associated with 1,196 genes, whereas the hypermethylated probes were annotated with only 463 genes. Taken together, the patient shows a predominant hypomethylation signature, suggesting an episignature Class I mutation [1,2] (**figure 3A**). The predicted methylation signature could

be confirmed with PhenoScore, which classified our splice-site mutation case as a Class I *ADNP* mutation (score: 0.1) (**supplementary data S2**). For confirmation of specific CpG methylation, we focused on genes related to nervous system development and selected the hypermethylated genes *CNTN1* and *GPX4* together with the hypomethylated genes *YY1* and *MAPT* for pyrosequencing. Here, we could confirm a higher percentage of CpG methylation in the patient for *CNTN1 (13%)* and *GPX4 (13.5%)* compared to the unaffected parents. Respectively, we could also demonstrate a lower percentage of CpG methylation in the patient for *YY1 (11%)* and *MAPT (25.5%)* (**figure 3B**).

The pathways affected in the patient are characteristic of the Helsmoortel – Van der Aa syndrome.

Using our methylome data, we first focused on differentially methylated gene-associated pathways using MetaScape, showing involvement of cell adhesion, brain development, actin filament-based processes, embryonic development, locomotion, behavior, membrane potential, leukocyte differentiation, and muscle structure development (**figure 4A**). Moreover, we also examined enriched transcription factors using the TRUSST database, focusing on co-expressed nodules of differentially methylated CpG, revealing eighteen enriched transcription factors including *SIRT1, KAT2B, KLF2, SP1, HDAC4, SRF*, *NR2F2*, and *SOX9*, either involved in nervous system development, differentiation, apoptosis, hematopoiesis, chromatin remodeling, immune cell differentiation, or congenital heart defects (**figure 4B**).

Next, we applied differential expression analysis on our transcriptome data, where 12,360 genes appeared in peripheral blood of the ADNP patient compared to her unaffected parents using the NOISeq package. In line with the observation of an excess of hypomethylated CpG probes, we observed most of the expressed genes to be upregulated. Using a significance cut-off equivalent to FDR =< 0.05, padj-value < 0.05 and a biologically meaningful effect size ($log_2FC > 0.5$), we found 540 downregulated and 832 upregulated genes with differential expression (**figure 4C**). Gene expression alterations in the ADNP patient were modest with the majority of genes presenting with an absolute log_2FC value < 2. Gene-ontology (GO) enrichment revealed an upregulation of transcription and translation-related processes as well as mitochondrial-related pathways, whereas cell proliferation, cytoskeleton protein binding, and myosin binding were downregulated (**figure 4D**). We confirmed a selected set of genes with RT-PCR, confirming upregulation of the developmental factor *PAX5* ($p < 0.0001$; ****), BAF complex members *BCL11A* ($p < 0.0001$; ****) and *SMARCB1* ($p = 0.03$; *). In addition, we also confirmed downregulation of WNT signaling member *WNT7A* ($p < 0.0001$; ****), cytoskeleton-

associated gene *TUBB1* ($p < 0.0001$; ***), together with lineage-specifying transcription factor *GATA2* ($p = 0.03$; *) (**figure 4E**). Using a combined methylome-transcriptome analysis, we unraveled pathways and disease-associated genes that are characteristic of the Helsmoortel – Van der Aa syndrome.

DISCUSSION

In this study, we report a splice-site acceptor mutation in the *ADNP* gene leading to haploinsufficiency in a patient with the clinical hallmarks of the Helsmoortel – Van der Aa syndrome. In contrast to earlier reports, the patient has a rational and unequivocal diagnosis as demonstrated by whole genome sequencing, the PhenoScore algorithm, the methylation episignature and the downstream disease-related pathways analysis. The clinical presentation of earlier reports of a complete gene deletion or an intragenic inversion leading to allele-specific *ADNP* silencing have been suggestive, but an unambiguous diagnosis had not been unequivocally demonstrated[13,21]. Interestingly, our splice site mutation has a class I episignature, a pattern that is otherwise associated with mutations in the N- and C-terminal regions of the ADNP protein. As N-terminal ADNP mutant proteins are degraded shortly after synthesis at least in a cellular overexpression models, haploinsufficiency is likely the disease mechanism for this specific patient group [5]. Patients with these N-terminal mutations have a distinct hypomethylation episignature, reminiscent of the methylation signature observed in the patient reported in this study [1]. Such an episignature is correlated with an overall milder clinical presentation as opposed to mutations in the central region of the protein that result in a class II epigsignature[1–4]. While the patient is most likely haploinsufficient, it should be stressed here that it cannot be excluded that other or even additional disease mechanisms play a role in a subset of Helmoortel Van der Aa patients . Nevertheless, this is a first reasonable demonstration of haploinsufficiency as a proven cause of Helsmoortel - Van der Aa syndrome.

Our case is also an elegant illustration of the added value of novel diagnostic tools over WES, including WGS, EpiSign and facial recognition imaging. In retrospect, each of these latter techniques was able to point to the correct diagnosis. However, the initial analysis of the WGS data failed to immediately detect the abnormality, presumably because it was intronic. The episignature inspired a reanalysis of the WGS data with a special emphasis on the *ADNP* gene, leading to the detection of the three base pair deletion in the splice acceptor-site with subsequent PhenoScore analysis supporting the molecular diagnosis using facial images of the patient.

However, in order to make a definitive diagnosis, the presence of any form of *ADNP* generated from the mutated allele needed to be excluded. RNA sequencing identified exon 4 skipping as the mutational event. As a previous report postulated an alternative ATG start codons with significant Kozak strength in the *ADNP* gene with Met229 to initiate translation of a shorter ADNP isoform[13]. We investigated protein levels by deletion of this exon in a recombinant expression vector as a model for the human condition. Here, we were able to detect mutant mRNA, as observed in the patient, with failure to initiate in-frame translation, suggesting haploinsufficiency as the mutational mechanism in this patient. We consider the synthesis of a potential N-terminal mutant protein in physiological conditions extremely unlikely, as we were unable to detect any of such potential isoforms in our overexpression system.

We also identified affected cellular pathways that are typified for the Helsmoortel-Van der Aa syndrome. For example, we found genes involved in development, the cytoskeleton, and myosin binding to be dysregulated in the toddler at the methylome and transcriptome level. ADNP was originally discovered to be essential for brain development [1] and is also implicated in regulation of genes crucial for lineage specification [4]. On top, ADNP fulfills a nuclear function as a chromatin remodeler [5], but has also been reported to interact with the microtubule end binding proteins, linking the protein to the cytoskeleton [24]. The patient in this study also presented with severe motor delays with our molecular data showing abnormal myosin functioning. Importantly, myosin light chain (Myl2) dysregulation was reported in the haploinsufficient *Adnp* mouse model, showing motor dysfunctions [25]. Of note, RNA sequencing did not show a reduced expression in *ADNP* mRNA levels, although we provide evidence for presence of mutant mRNA in the patient [7].

In conclusion, we discovered a novel splice acceptor site mutation in the *ADNP* gene, with transcriptome data showing skipping of the first coding exon. As no truncated protein could be detected, postulate haploinsufficiency as the mutational event in this young 5-year old girl.

DATA AVAILABILITY

The datasets generated during this study are available from the corresponding author upon argumented request.

ACKNOWLEDGMENTS

The authors would like to thank the family for participating in this study.

AUTHOR CONTRIBUTIONS

C.P.D. performed the experiments, conceptualized the experimental design, performed data analysis, wrote the manuscript text, and prepared all the figures. D.J.A. provided the bioinformatic analysis of the methylation array. E.E. conducted all the gene expression assays using RT-PCR under supervision of C.D., J.J.V.D.S and M.A. performed the WES, WGS, and EpiSign at early observations of the patient, A.J.M.D. and B.B.A.D.V. provided analysis of the ADNP female using PhenoScore. L.M. provided the bioinformatic analysis of the RNA sequencing. W.V.B. and R.F.K. reviewed and edited the manuscript.. All authors reviewed and approved the final version of manuscript.

FUNDING

R. Frank Kooy acknowledges the support of the Research Fund of the University of Antwerp OEC-Methusalem grant "GENOMED". This work was in part financed by grants from the ERA-NET NEURON "ADNPinMED". This article is also based upon work from COST Action International Nucleome Consortium (INC) CA18127, supported by COST (European Cooperation in Science and Technology) ascribed to Wim Vanden Berghe.

ETHICAL APPROVAL STATEMENT

Written informed consent for investigation of the ADNP mutation in the context of the Helsmoortel-Van der Aa syndrome was obtained from the parents for clinical testing, research use and publication. The protocol was approved by the Ethics Committee of the Antwerp University Hospital, Antwerp, Belgium.

COMPETING INTERESTS

The authors declare no competing interests.

FIGURES

Figure 1. Loss-of-function of ADNP due to an exon skipping event. Sashimi-plot of complete transcriptome sequencing demonstrating ADNP spliced-junctions in the blood of the patient and unaffected parents. The histograms represent read coverage of the exons and arcs indicate the number of junction-spanning reads supporting the exons junction. A physiological junction was absent between exon 4 and exon 5 in the patient, but not in the parents. Nine of the *ADNP* transcription variants are summarized (Ensembl).

Figure 2. The exon 4 skipping event causes ADNP haploinsufficiency. (**A**) Detection of recombinant wild-type (GFPSpark®) and ΔEX4 mutant (OFPSpark®) ADNP mRNA levels in HEK293T overexpression lysates. Expression values

were normalized using the housekeeping genes *B2M, GAPDH, and YHWAZ*. Data was subsequently analyzed using a 2way ANOVA with Šídák's multiple comparisons test and represented as mean with SD. (**B**) ADNP protein expression analysis using a C-terminal antibody. Western blotting showed presence of endogenous wild-type ADNP (150 kDa) in each condition, together with recombinant ADNP-GFPSpark® wild-type (175 kDa). However, the predicted N-terminal abbreviated protein as a consequence of the exon skipping event could be visualized (123 kDa). **(C)** ADNP protein expression analysis using a N-terminal antibody. Western blotting showed presence of endogenous wild-type ADNP (150 kDa) and recombinant ADNP-GFPSpark® wild-type (175 kDa). The N-terminal antibody was included as a negative control. GAPDH was used as a loading control in all western blotting experiments.

Figure 3. The ADNP splice-acceptor site mutation affects genome-wide methylation. (**A**) Genomic scatterplot of genome-wide CpG methylation differences across. The biological impact of methylation is defined by β-coefficient, i.e., CpG methylation of the ADNP toddler – the mean CpG methylation of the parents, with hypomethylation defined as $\Delta\beta < -0.1$, respectively hypermethylation by $\Delta\beta > 0.1$. (**B**) Pyrosequencing confirmation of differentially methylated gene-associated CpGs.

Figure 4. Methylome-transcriptome analysis indicates pathways characteristic for the Helsmoortel-Van der Aa syndrome. (**A**) Metascape pathway analysis showing enriched biological processes and pathways correlated with the differentially methylated CpG sites. The significance was indicated by the -log10(p-value) values at the y-axis. (**B**) TRUSST analysis for enriched transcription factors, associated with co-expressed nodules of differentially methylated genes, revealing twelve enriched transcription factors. (**C**) Volcano plot of DEGs in the patient using the NOIseq package, displaying the significance (-log10q) and effect size (log2FC). The DEG are shown in blue. (**D**) Functional gene set enrichment of GO and BP for differentially expressed genes. Downregulation, dark blue; upregulation, yellow. (**E**) RT-PCR confirmations of differentially expressed genes levels in patient opposed to her unaffected parents. Expression values were normalized using the housekeeping genes *B2M, GAPDH, and RPL13A*. Data was subsequently analyzed using a two-way ANOVA with Šídák's multiple comparisons test and represented as mean with SD.

Supplementary data S1. Identification of a heterozygous *de novo* intronic mutation upstream of the ATG initiation codon in the *ADNP* gene. (**A**) Facial photograph of the five-year-old female ADNP patient. (**B**) Schematic representation of the clinical presentation of the patient, indicated by autism-associated characteristics, intellectual delay, and motor delays together with obsessive behavior. (**C**) PhenoScoring of the patient using the facial image, the phenotypic data and, the PhenoScore, which combines both. Heatmaps are generated using LIME to see which facial areas are most important according to our model. To support a Helsmoortel-Van der Aa syndrome diagnosis, a HPO score of 0.1 indicates similarities

to the control population, whereas a score closer to 1.0 points towards an ADNP diagnosis. To stratify methylation groups, a HPO score of 0.1 is rather associated with a hypomethylation signature (Class I mutation), whereas a score closer to 1.0 points towards a hypermethylation signature (Class II mutation). (**D**) Sanger sequencing electrogram confirming the heterozygous *de novo* mutation c.[-5-1_-4del];[=] p.[?];[(=)] in the *ADNP* gene in the toddler.

REFERENCES

1 Pinhasov A, Mandel S, Torchinsky A, Giladi E, Pittel Z, Goldsweig AM *et al.* Activity-dependent neuroprotective protein: a novel gene essential for brain formation. *Brain Res Dev Brain Res* 2003; **144**: 83–90.

2 Mandel S, Rechavi G, Gozes I. Activity-dependent neuroprotective protein (ADNP) differentially interacts with chromatin to regulate genes essential for embryogenesis. *Dev Biol* 2007; **303**: 814–824.

3 Mandel S, Gozes I. Activity-dependent neuroprotective protein constitutes a novel element in the SWI/SNF chromatin remodeling complex. *J Biol Chem* 2007; **282**: 34448–34456.

4 Ostapcuk V, Mohn F, Carl SH, Basters A, Hess D, Iesmantavicius V *et al.* Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature* 2018; **557**: 739–743.

5 D'Incal CP, Van Rossem KE, De Man K, Konings A, Van Dijck A, Rizzuti L *et al.* Chromatin remodeler Activity-Dependent Neuroprotective Protein (ADNP) contributes to syndromic autism. *Clin Epigenetics* 2023; **15**: 45.

6 Van Dijck A, Vulto-van Silfhout AT, Cappuyns E, van der Werf IM, Mancini GM, Tzschach A *et al.* Clinical presentation of a complex neurodevelopmental disorder caused by mutations in *ADNP*. *Biol Psychiatry* 2019; **85**: 287–297.

7 Helsmoortel C, Vulto-van Silfhout AT, Coe BP, Vandeweyer G, Rooms L, van den Ende J *et al.* A SWI/SNF-related autism syndrome caused by de novo mutations in ADNP. *Nat Genet* 2014; **46**: 380–384.

8 Vandeweyer G, Helsmoortel C, Van Dijck A, Vulto-van Silfhout AT, Coe BP, Bernier R *et al.* The transcriptional regulator ADNP links the BAF (SWI/SNF) complexes with autism. *Am J Med Genet C Semin Med Genet* 2014; **166C**: 315–326.

9 Bend EG, Aref-Eshghi E, Everman DB, Rogers RC, Cathey SS, Prijoles EJ *et al.* Gene domain-specific DNA methylation episignatures highlight distinct molecular entities of ADNP syndrome. *Clin Epigenetics* 2019; **11**: 64.

10 Breen MS, Garg P, Tang L, Mendonca D, Levy T, Barbosa M *et al.* Episignatures Stratifying Helsmoortel-Van Der Aa Syndrome Show Modest Correlation with Phenotype. *Am J Hum Genet* 2020; **107**: 555–563.

11 Dingemans AJM, Hinne M, Truijen KMG, Goltstein L, van Reeuwijk J, de Leeuw N *et al.* PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework. *Nat Genet* 2023; **55**: 1598–1607.

12 D'Incal CP, Cappuyns E, Choukri K, Szrama K, De Man K, Aa NV der *et al.* In Search of the Hidden Protein: Optimization of Detection Strategies for autism-associated Activity-Dependent Neuroprotective Protein (ADNP) mutants. *Res Sq* 2022. doi:10.21203/rs.3.rs-1954095/v1.

13 Georget M, Lejeune E, Buratti J, Servant E, le Guern E, Heron D *et al.* Loss of function of ADNP by an intragenic inversion. *Eur J Hum Genet* 2023; **31**: 967–970.

14 D'Incal CP, Kooy RF. ADNP in reverse gear. *Eur J Hum Genet* 2023; **31**: 849–850.

15 Zhou W, Triche TJ, Laird PW, Shen H. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res* 2018; **46**: e123.

16 Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019; **10**: 1523.

17 Breen MS, Garg P, Tang L, Mendonca D, Levy T, Barbosa M *et al.* Episignatures stratifying ADNP syndrome show modest correlation with phenotype. *BioRxiv* 2020. doi:10.1101/2020.04.01.014902.

18 Akhmedov M, Martinelli A, Geiger R, Kwee I. Omics Playground: a comprehensive self-service platform for visualization, analytics and exploration of Big Omics Data. *NAR Genom Bioinform* 2020; **2**: lqz019.

19 Gozes I, Van Dijck A, Hacohen-Kleiman G, Grigg I, Karmon G, Giladi E *et al.* Premature primary tooth eruption in cognitive/motor-delayed ADNP-mutated children. *Transl Psychiatry* 2017; **7**: e1043.

20 Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet* 2018; **59**: 253–268.

21 Huynh M-T, Boudry-Labis E, Massard A, Thuillier C, Delobel B, Duban-Bedu B *et al.* A heterozygous microdeletion of 20q13.13 encompassing ADNP gene in a child with Helsmoortel-van der Aa syndrome. *Eur J Hum Genet* 2018; **26**: 1497–1501.

22 Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An J-Y *et al.* Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* 2020; **180**: 568–584.e23.

23 van der Sanden BPGH, Schobers G, Corominas Galbany J, Koolen DA, Sinnema M, van Reeuwijk J *et al.* The performance of genome sequencing as a first-tier test for neurodevelopmental disorders. *Eur J Hum Genet* 2023; **31**: 81–88.

24 Ivashko-Pachima Y, Sayas CL, Malishkevich A, Gozes I. ADNP/NAP dramatically increase microtubule end-binding protein-Tau interaction: a novel avenue for protection against tauopathy. *Mol Psychiatry* 2017; **22**: 1335–1344.

25 Kapitansky O, Karmon G, Sragovich S, Hadar A, Shahoha M, Jaljuli I *et al.* Single cell ADNP predictive of human muscle disorders: mouse knockdown results in muscle wasting. *Cells* 2020; **9**. doi:10.3390/cells9102320.

1. Breen MS, Garg P, Tang L, et al. Episignatures Stratifying Helsmoortel-Van Der Aa Syndrome Show Modest Correlation with Phenotype. *Am J Hum Genet*. 2020;107(3):555-563. doi:10.1016/J.AJHG.2020.07.003

2. Bend EG, Aref-Eshghi E, Everman DB, et al. Gene domain-specific DNA methylation episignatures highlight distinct molecular entities of ADNP syndrome. *Clin Epigenetics*. 2019;11(1):1-17. doi:10.1186/S13148-019-0658-5/FIGURES/6

3. Dingemans AJM, Hinne M, Truijen KMG, et al. PhenoScore quantifies phenotypic variation for rare genetic diseases by combining facial analysis with other clinical features using a machine-learning framework. *Nat Genet*. Published online 2023. doi:10.1038/S41588-023-01469-W

4. Breen MS, Garg P, Tang L, et al. Episignatures stratifying ADNP syndrome show modest correlation with

phenotype. *bioRxiv*. Published online April 2, 2020:2020.04.01.014902. doi:10.1101/2020.04.01.014902

5. Cappuyns E, Huyghebaert J, Vandeweyer G, Kooy RF. Mutations in ADNP affect expression and subcellular localization of the protein. *Cell Cycle*. 2018;17(9):1068. doi:10.1080/15384101.2018.1471313

6. D'Incal CP, Cappuyns E, Choukri K, et al. In Search of the Hidden Protein: Optimization of Detection Strategies for autism-associated Activity-Dependent Neuroprotective Protein (ADNP) mutants. *Res Sq*. Published online 2022. doi:10.21203/rs.3.rs-1954095/v1

Figure 1

Sashimi-plot of mRNA sequencing data showing an exon skipping event

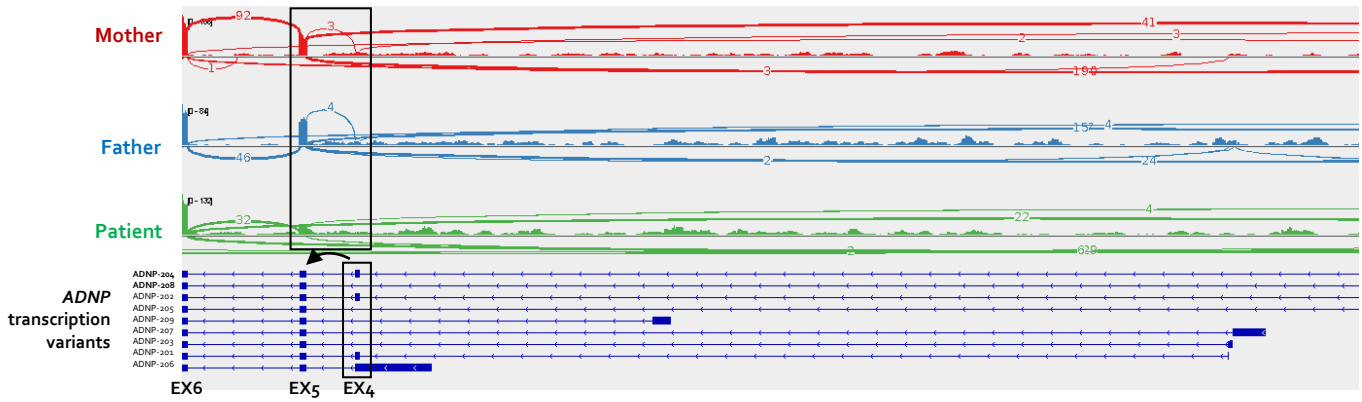
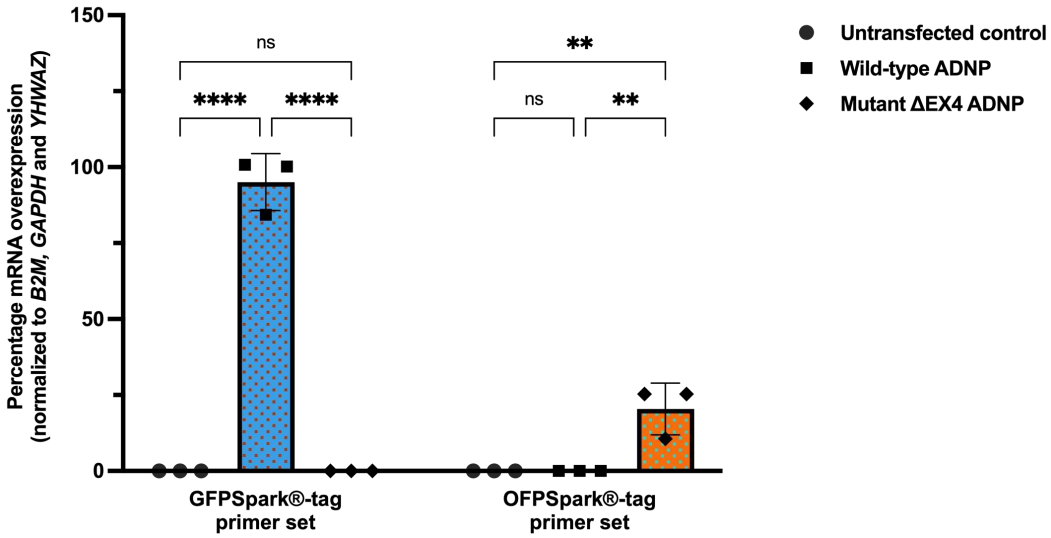
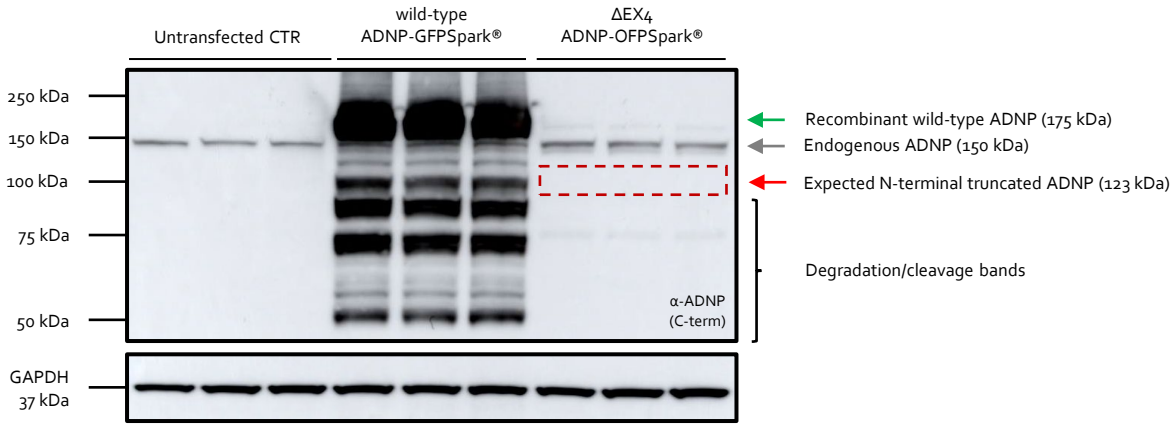


Figure 2

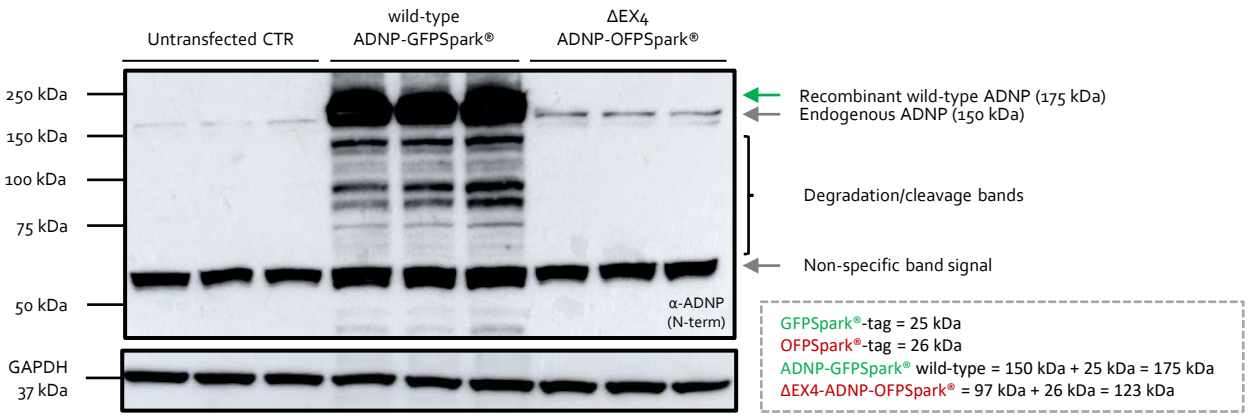
A RT-PCR using expression vector-specific primer sets



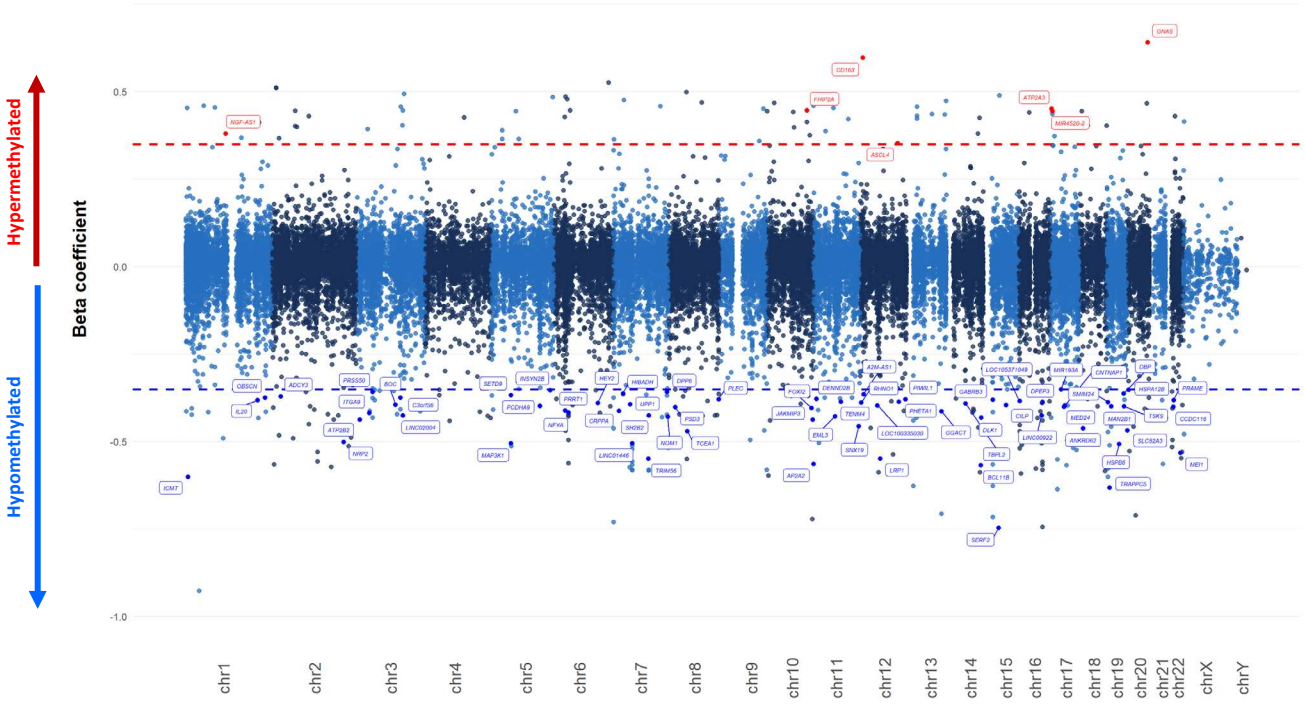
B Immunoblotting using a C-terminal ADNP antibody



C Immunoblotting using an N-terminal ADNP antibody



A Differentially methylated genes in peripheral blood of the ADNP toddler



B Pyrosequencing confirmation of differentially methylated CpGs

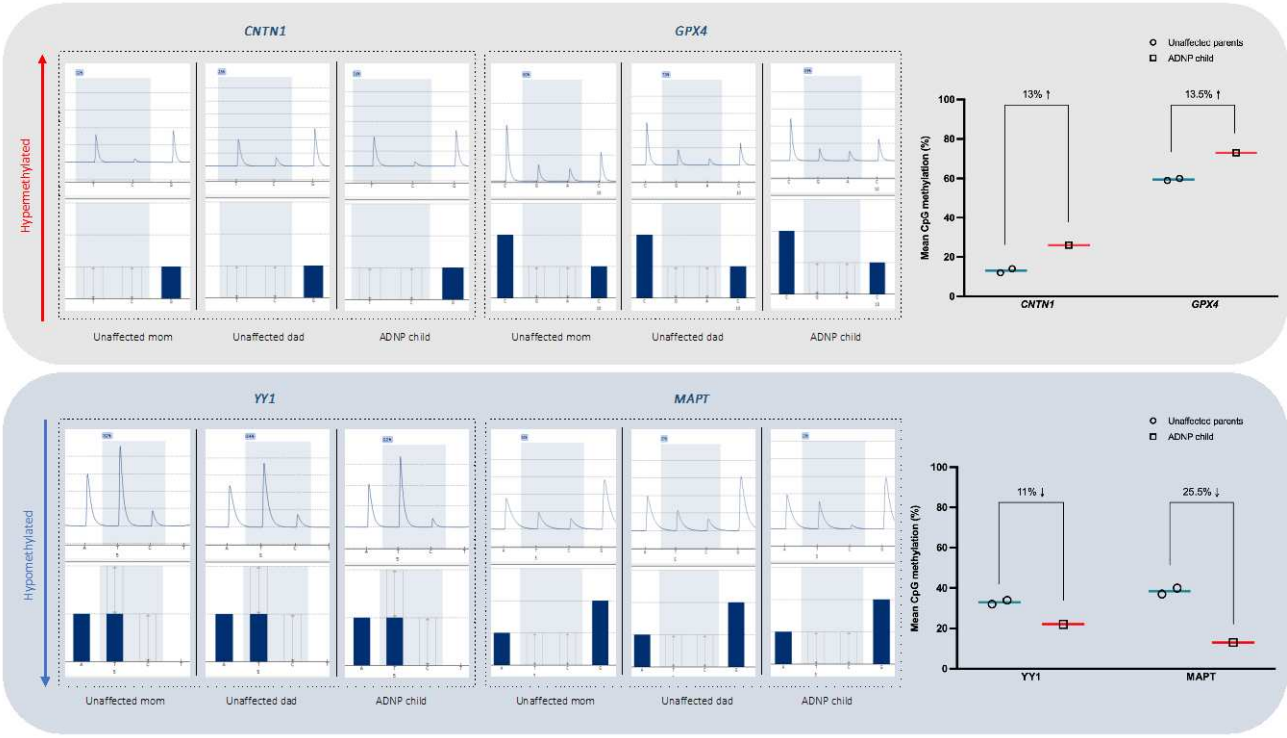


Figure 3

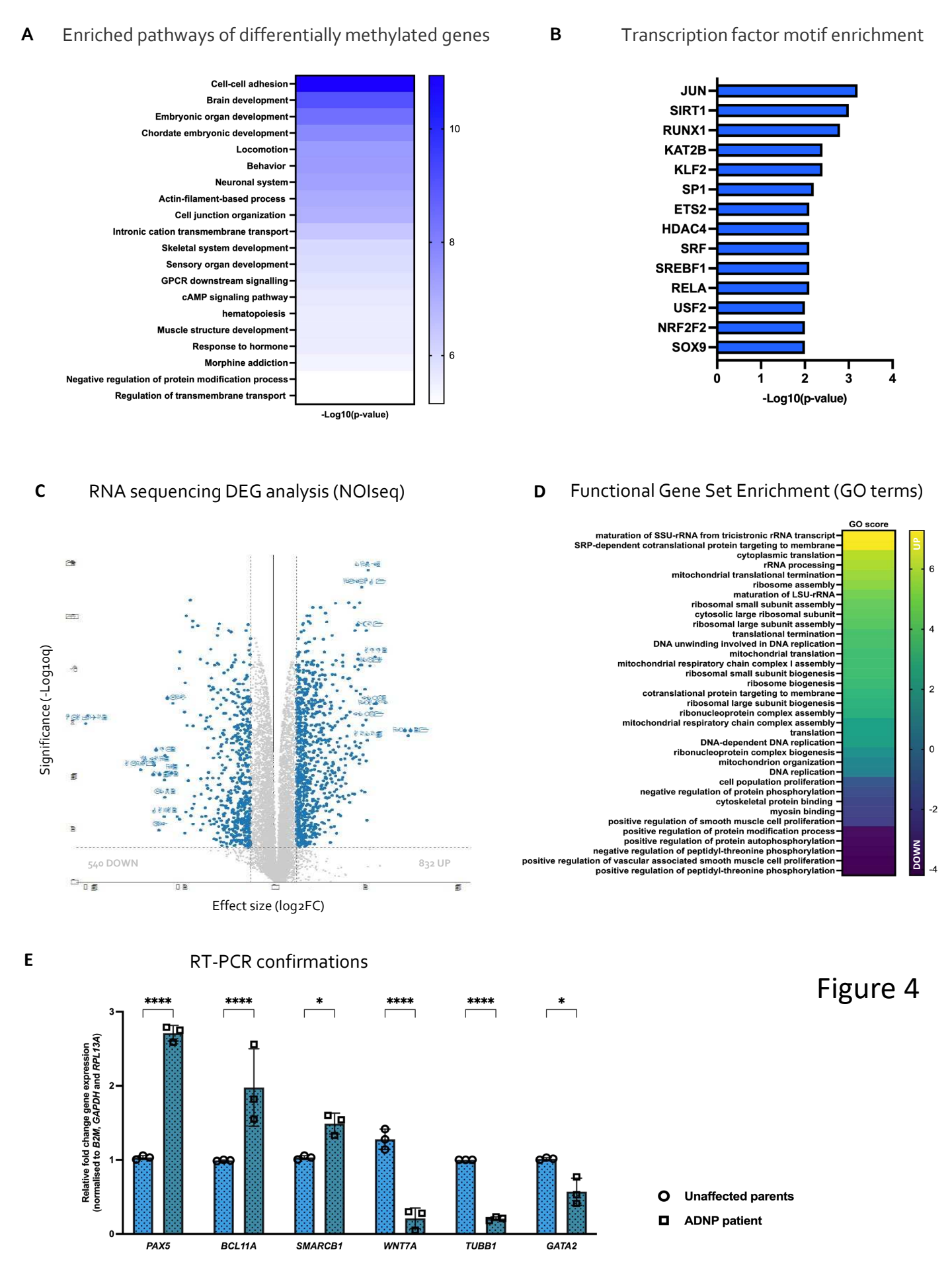
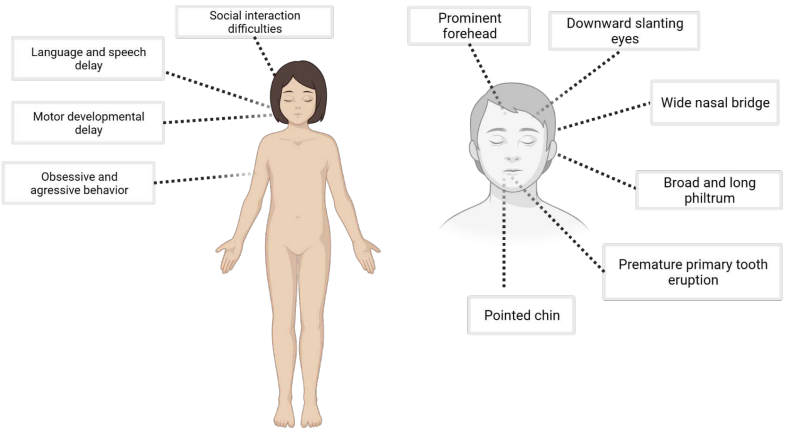

A
Enriched pathways of differentially methylated genes
Cell-cell adhesion
Brain development
Embryonic organ development
Chordate embryonic development
Locomotion
Behavior
Neuronal system
Actin-filament-based process
Cell junction organization
Intronic cation transmembrane transport
Skeletal system development
Sensory organ development
GPCR downstream signalling
cAMP signaling pathway
hematopoiesis
Muscle structure development
Response to hormone
Morphine addiction
Negative regulation of protein modification process
Regulation of transmembrane transport
10
8
6
-Log10(p-value)
B
Transcription factor motif enrichment
JUN
SIRT1
RUNX1
KAT2B
KLF2
SP1
ETS2
HDAC4
SRF
SREBF1
RELA
USF2
NRF2F2
SOX9
0
1
2
3
4
-Log10(p-value)
C
RNA sequencing DEG analysis (NOIseq)
Significance (-Log10q)
540 DOWN
832 UP
Effect size (log2FC)
D
Functional Gene Set Enrichment (GO terms)
GO score
maturation of SSU-rRNA from tricistronic rRNA transcript
SRP-dependent cotranslational protein targeting to membrane
cytoplasmic translation
rRNA processing
mitochondrial translational termination
ribosome assembly
maturation of LSU-rRNA
ribosomal small subunit assembly
cytosolic large ribosomal subunit
ribosomal large subunit assembly
translational termination
DNA unwinding involved in DNA replication
mitochondrial translation
mitochondrial respiratory chain complex I assembly
ribosomal small subunit biogenesis
ribosome biogenesis
cotranslational protein targeting to membrane
ribosomal large subunit biogenesis
ribonucleoprotein complex assembly
mitochondrial respiratory chain complex assembly
translation
DNA-dependent DNA replication
ribonucleoprotein complex biogenesis
mitochondrion organization
DNA replication
cell population proliferation
negative regulation of protein phosphorylation
cytoskeletal protein binding
myosin binding
positive regulation of smooth muscle cell proliferation
positive regulation of protein modification process
positive regulation of protein autophosphorylation
negative regulation of peptidyl-threonine phosphorylation
positive regulation of vascular associated smooth muscle cell proliferation
positive regulation of peptidyl-threonine phosphorylation
UP
DOWN
6
4
2
0
-2
-4
E
RT-PCR confirmations
Relative fold change gene expression (normalised to B2M, GAPDH and RPL13A)
****
****
*
****
****
*
PAX5
BCL11A
SMARCB1
WNT7A
TUBB1
GATA2
Unaffected parents
ADNP patient

Figure 4

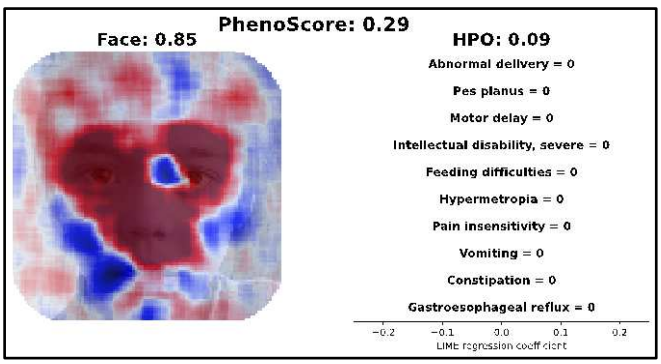
A Facial photograph of the patient

B Clinical manifestation

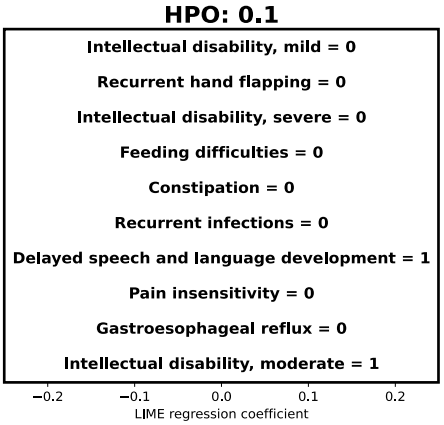


C Genotype-phenotype correlation using PhenoScore

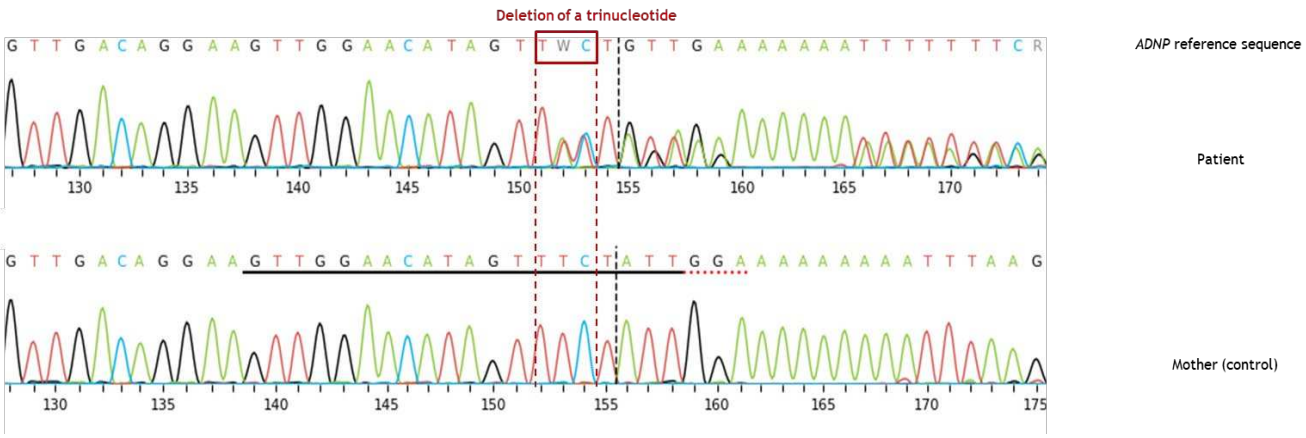
Diagnosis prediction based of facial recognition and HPO terms



Methylation signature prediction based on HPO terms



D Heterozygous *de novo* deletion c.[-5-1_-4del];[=] in the *ADNP* gene



Supplementary S2