

## **Genome and RNA sequencing were essential to reveal cryptic intronic variants associated to defective *ATP6AP1* mRNA processing**

Blai Morales-Romero<sup>a</sup>, Gerard Muñoz-Pujol<sup>a</sup>, Rafael Artuch<sup>b</sup>, Angels García-Cazorla<sup>c</sup>, Mar O'Callaghan<sup>c</sup>, Jolanta Sykut-Cegielska<sup>d</sup>, Jaume Campistol<sup>c</sup>, Pedro Juan Moreno-Lozano<sup>e</sup>, Machteld M. Oud<sup>f</sup>, Ron A. Wevers<sup>g</sup>, Dirk J. Lefeber<sup>g,h</sup>, Anna Esteve-Codina<sup>i</sup>, Vicente A. Yeppez<sup>j,k</sup>, Julien Gagneur<sup>j,k</sup>, Saskia B. Wortmann<sup>l,m</sup>, Holger Prokisch<sup>j,n</sup>, Antonia Ribes<sup>a</sup>, Judit García-Villoria<sup>a,\*</sup>, Frederic Tort<sup>a,\*</sup>

<sup>a</sup>Section of Inborn Errors of Metabolism-IBC, Biochemistry and Molecular Genetics Department, Hospital Clínic de Barcelona, IDIBAPS, CIBERER, ISCIII, Barcelona, Spain.

<sup>b</sup>Clinical Biochemistry Department, Institut de Recerca Sant Joan de Déu, Hospital Sant Joan de Déu, CIBERER, Esplugues de Llobregat, Barcelona, Spain.

<sup>c</sup>Neurology Department, Hospital Sant Joan de Déu, Institut de Recerca Hospital Sant Joan de Déu, CIBERER and MetabERN, Esplugues de Llobregat, Barcelona, Spain.

<sup>d</sup>Department of Inborn Errors of Metabolism and Pediatrics, Institute of Mother and Child, Warsaw, Poland.

<sup>e</sup>Inherited Metabolic Diseases and Muscle Disorders' Research Group, Department of Internal Medicine, Hospital Clinic de Barcelona, IDIBAPS, Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain.

<sup>f</sup>Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands and United for Metabolic Diseases, The Netherlands

<sup>g</sup>Department of Human Genetics, Translational Metabolic Laboratory (TML), Radboud University Medical Center, Nijmegen, The Netherlands, and United for Metabolic Diseases, The Netherlands.

<sup>h</sup>Department of Neurology, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands.

<sup>i</sup>Centro Nacional de Análisis Genómico (CNAG), Parc Científic de Barcelona, Barcelona, Spain.

<sup>j</sup>Institute of Human Genetics, School of Medicine, Technical University of Munich, 81675 Munich, Germany.

<sup>k</sup>TUM School of Computation, Information and Technology, Technical University of Munich, 85748 Garching, Germany.

<sup>l</sup>University Children's Hospital Salzburg, Paracelsus Medical University, Salzburg, Austria.

<sup>m</sup>Amalia Children's Hospital, Department of Pediatrics, Radboudumc, Nijmegen, The Netherlands.

<sup>n</sup>Institute of Neurogenomics, Helmholtz Zentrum München, 85764 Neuherberg, Germany.

E-mail addresses: [bmorales@clinic.cat](mailto:bmorales@clinic.cat) (B. Morales-Romero), [gemunoz@recerca.clinic.cat](mailto:gemunoz@recerca.clinic.cat) (G. Muñoz-Pujol), [rafael.artuch@sjd.es](mailto:rafael.artuch@sjd.es) (R. Artuch), [angeles.garcia@sjd.es](mailto:angeles.garcia@sjd.es) (A. García-Cazorla), [mariammar.ocallaghan@sjd.es](mailto:mariammar.ocallaghan@sjd.es) (M. O'Callaghan), [j.cegielska@wp.pl](mailto:j.cegielska@wp.pl) (J. Sykut-Cegielska), [campistol@sjdhospitalbarcelona.org](mailto:campistol@sjdhospitalbarcelona.org) (J. Campistol), [pjmoreno@clinic.cat](mailto:pjmoreno@clinic.cat) (P.J. Moreno-Lozano), [Machteld.Oud@radboudumc.nl](mailto:Machteld.Oud@radboudumc.nl) (M.M. Oud), [Ron.Wevers@radboudumc.nl](mailto:Ron.Wevers@radboudumc.nl) (R. Wavers), [Dirk.Lefeber@radboudumc.nl](mailto:Dirk.Lefeber@radboudumc.nl) (D. Lefeber), [anna.esteve@cnag.eu](mailto:anna.esteve@cnag.eu) (A. Esteve-Codina), [yepez@in.tum.de](mailto:yepez@in.tum.de) (V.A. Yépez), [gagneur@in.tum.de](mailto:gagneur@in.tum.de) (J. Gagneur), [s.wortmann@salk.at](mailto:s.wortmann@salk.at) (S. B. Wortmann), [prokisch@helmholtz-muenchen.de](mailto:prokisch@helmholtz-muenchen.de) (H. Prokisch), [aribes@clinic.cat](mailto:aribes@clinic.cat) (A. Ribes), [jugarcia@clinic.cat](mailto:jugarcia@clinic.cat) (J. García-Villoria), [ftort@recerca.clinic.cat](mailto:ftort@recerca.clinic.cat) (F. Tort).

**\* These authors contribute equally.**

**Correspondence:**

Frederic Tort

Hospital Clínic de Barcelona, IDIBAPS, CIBERER, ISCIII, Barcelona, Spain.

Carrer de Villarroel, 170, 08036, Barcelona.

E-mail address: [ftort@recerca.clinic.cat](mailto:ftort@recerca.clinic.cat)

Judit García-Villoria

Hospital Clínic de Barcelona, IDIBAPS, CIBERER, ISCIII, Barcelona, Spain.

Carrer de Villarroel, 170, 08036, Barcelona.

E-mail address: [jugarcia@clinic.cat](mailto:jugarcia@clinic.cat)

## ABSTRACT

The diagnosis of Mendelian disorders has notably advanced with integration of whole exome and genome sequencing (WES and WGS) in clinical practice. However, challenges in variant interpretation and uncovered variants by WES still leave a substantial percentage of patients undiagnosed. In this context, integrating RNA sequencing (RNA-seq) improves diagnostic workflows, particularly for WES inconclusive cases. Additionally, functional studies are often necessary to elucidate the impact of prioritized variants on gene expression and protein function.

Our study focused on three unrelated male patients (P1-P3) with ATP6AP1-CDG (congenital disorder of glycosylation), presenting with intellectual disability and varying degrees of hepatopathy, glycosylation defects, and an initially inconclusive diagnosis through WES. Subsequent RNA-seq was pivotal in identifying the underlying genetic causes in P1 and P2, detecting *ATP6AP1* underexpression and aberrant splicing. Molecular studies in fibroblasts confirmed these findings and identified the rare intronic variants c.289-233C>T and c.289-289G>A in P1 and P2, respectively. Trio-WGS also revealed the variant c.289-289G>A in P3, which was a *de novo* change in both patients. Functional assays expressing the mutant alleles in HAP1 cells demonstrated the pathogenic impact of these variants by reproducing the splicing alterations observed in patients.

Our study underscores the role of RNA-seq and WGS in enhancing diagnostic rates for genetic diseases such as CDG, providing new insights into ATP6AP1-CDG molecular bases by identifying the first two deep intronic variants in this X-linked gene. Additionally, our study highlights the need to integrate RNA-seq and WGS, followed by functional validation, in routine diagnostics for a comprehensive evaluation of patients with an unidentified molecular etiology.

***Keywords:***

RNA-seq

Congenital Disorders of Glycosylation

ATP6AP1-CDG

Whole exome sequencing

Whole genome sequencing

Intronic variant

***Abbreviations:***

ALP, alkaline phosphatase; ApoCIII, apolipoprotein CIII; CDG, congenital disorders of glycosylation; cDNA, complementary DNA; CE, cryptic exon; CHX, cycloheximide; CK, creatine kinase; CNV, copy number variant; ESE, exonic splicing enhancer; ESS, exonic splicing silencer; GGT, gamma-glutamyl transferase; HPLC, high performance liquid chromatography; HSF, Human Splicing Finder; IEF, isoelectric focusing; MAE, monoallelic expression; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopy; NMD, nonsense-mediated mRNA decay; PCR, polymerase chain reaction; qRT-PCR, reverse transcription-quantitative PCR; RNA-seq, RNA sequencing; SNV, single nucleotide variant; SRE, splicing-regulatory element; SV, structural variant; VUS, variant of uncertain significance; WES, whole exome sequencing; WGS, whole genome sequencing.

## **CRedit authorship contribution statement**

**Blai Morales-Romero:** Methodology, Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Gerard Muñoz-Pujol:** Methodology, Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Rafael Artuch:** Investigation, Resources, Writing - Review & Editing. **Angels García-Cazorla:** Investigation, Resources, Writing - Review & Editing. **Mar O'Callaghan:** Investigation, Resources, Writing - Review & Editing. **Jolanta Sykut-Cegielska:** Investigation, Resources, Writing - Review & Editing. **Jaume Campistol:** Investigation, Resources, Writing - Review & Editing. **Pedro Juan Moreno-Lozano:** Investigation, Resources, Writing - Review & Editing. **Machteld M. Oud:** Investigation, Resources, Methodology, Writing - Review & Editing. **Ron A. Wevers:** Investigation, Resources, Funding acquisition, Writing - Review & Editing. **Dirk J. Lefeber:** Investigation, Resources, Methodology, Writing - Review & Editing. **Anna Esteve-Codina:** Software, Methodology, Resources, Writing - Review & Editing. **Vicente A. Yopez:** Software, Methodology, Resources, Writing - Review & Editing. **Julien Gagneur:** Software, Funding acquisition, Resources, Writing - Review & Editing. **Saskia B. Wortmann:** Investigation, Resources, Writing - Review & Editing. **Holger Prokisch:** Investigation, Resources, Funding acquisition, Writing - Review & Editing. **Antonia Ribes:** Conceptualization, Validation, Formal analysis, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition. **Judit García-Villoria:** Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing, Visualization, Supervision. **Frederic Tort:** Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition.

**Funding support:** This research was supported by the projects PI19/01310 and PI22/00856, funded by the Instituto de Salud Carlos III (ISCIII) and co-funded by European Union. This research was supported by CIBER - Consorcio Centro de Investigación Biomédica en Red, Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación, and funded by the ER20P2AC737 project. This study was also supported by Generalitat de Catalunya (SLT002/16/00174), Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) (2021: SGR01423), the CERCA Programme/Generalitat de Catalunya. We are grateful for the funds from Stichting Metakids (Grant number UMD-ZOE-2022-012) and UMD (United for Metabolic Diseases) for the Zoemba-international WGS trio analysis in patient 3. This work was supported by the German Bundesministerium für Bildung und Forschung (BMBF) through the ERA PerMed project PerMiM [01KU2016B to VAY and JG].

**Declaration of Competing Interest:** The authors declare no conflict of interest.

## **1. Introduction**

The diagnosis of inherited metabolic disorders, as well as other Mendelian disorders, has experienced profound advancements, primarily due to the implementation of whole exome sequencing (WES) into clinical practice. Nonetheless, an important number of patients still lack a genetic diagnosis, as WES provides a diagnostic rate of approximately 25-50%, depending on the cohort and disease type [1–3]. In the remaining cases, it is presumed that causative variants are located in non-coding or other genomic regions not covered by WES. Alternatively, pathogenic variants may not be identifiable through this approach or are not adequately prioritized. Although whole genome sequencing (WGS) overcomes the limitation of DNA coverage and improves the detection of structural variation, its overall contribution to the diagnostic yield is variable and, in many cases, even discrete [4–6]. Indeed, using WGS, the challenge persists in interpreting a larger number of variants of uncertain significance (VUS) and assessing their functional effects. In this context, RNA sequencing (RNA-seq) has emerged as a valuable tool, improving variant prioritization and providing crucial insights into the impact of variants on gene expression and splicing. It is estimated that RNA-seq analysis in WES-inconclusive patients yields an additional mean diagnosis rate of 15%, ranging from 8-36% depending on the specific disease cohort and the studied tissues [7–15].

Within this evolving landscape, congenital disorders of glycosylation (CDG) have been notably expanded in terms of the number of identified disorders during the current -omics era [16]. CDGs encompass a genetically and clinically heterogeneous group of more than 160 diseases characterized by frequent neurological and multiorgan involvement [17]. These disorders are caused by defects in proteins involved in glycan biosynthesis and metabolism, which can either impact on a single glycosylation pathway, with most CDGs being linked to N-glycosylation defects, or target multiple physiological pathways that

secondarily alter glycosylation. The latter group includes disorders that affect vesicular trafficking and Golgi apparatus homeostasis, impacting on the function of glycosyltransferases, as well as on other enzymes and cellular processes required for glycosylation [18].

Defects in several subunits and accessory proteins of the V-ATPase have been recognized as causing CDG by affecting multiple glycosylation pathways through the deregulation of Golgi pH homeostasis [19]. The *ATP6AP1* gene encodes an accessory subunit of the vacuolar V-ATPase, a proton pump complex involved in luminal acidification of secretory organelles and the extracellular matrix [20]. Disease-causing variants in this gene lead to an X-linked disorder (ATP6AP1-CDG, MIM #300972) associated with variable liver disease, ranging from slight hypertransaminasemia to severe liver cirrhosis. Additionally, most patients exhibit immunodeficiency characterized by recurrent bacterial infections and hypogammaglobulinemia. Some individuals may also manifest neurological symptoms, including intellectual disability, seizures, hypotonia, psychomotor and speech delay, learning difficulties, and behavioral problems. Moreover, connective tissue manifestations such as cutis laxa, dilation of the aortic root, diaphragmatic hernia, and wrinkled skin may also be present [21-33]. Biochemically, low plasma copper and/or ceruloplasmin levels have been reported in most patients, as well as an abnormal type II transferrin glycosylation pattern. Since the initial description in 2016 by Jansen *et al.* involving 11 male patients, the cohort has steadily expanded, leading to an increasingly broadened clinical spectrum. To date, a total of 34 patients from 24 unrelated families have been reported [21–33].

In this study, we identified the underlying genetic alterations in three unrelated patients exhibiting an abnormal type II transferrin glycosylation pattern suggestive of glycosylation deficiency. RNA-seq and WGS were conducted after inconclusive WES

findings, revealing two distinct hemizygous deep intronic variants within the same intron of *ATP6AP1*. We provide functional data demonstrating a shared pathogenic mechanism for both variants, leading to ATP6AP1 deficiency.

## **2. Subjects and Methods**

### **2.1. Patients**

**Patient 1 (P1).** He is a 19-year-old boy, the second child of non-consanguineous healthy parents. Family history of intellectual disability of unknown etiology was reported in a maternal uncle. At the age of 3, he presented with generalized seizures that responded to valproate. At that time, normal psychomotor development and mild cognitive impairment, consistent with speech delay, learning difficulties, and behavioral problems were observed. Physical examination showed facial dysmorphism, including large and low-set ears, and a sharp nose. Brain MRI was normal. Throughout childhood, persistent hypercholesterolemia was noted. Additionally, at 4 years old, an abnormal type II plasma transferrin glycoform profile was detected. By adulthood, elevated plasma cholesterol levels persisted, accompanied by increased alkaline phosphatase and low plasma copper levels with normal ceruloplasmin (**Supplementary Table S1**). Normal hepatic function with slight hepatomegaly was also noted. At the age of 20, there was no evidence of disease progression, although frequent generalized seizures persisted.

**Patient 2 (P2).** He was the first son of non-consanguineous healthy parents with no remarkable familial history. At the age of 11 months, he presented with seizures, psychomotor delay, and hypotonia. His developmental trajectory was marked by profound intellectual disability, manifesting as a lack of verbal communication and extremely limited nonverbal communication skills, coupled with minimal interaction with

the environment. Patient displacement was assisted by a wheelchair due to generalized hypotonia, tetraparesis and clubfeet, although he was able to make isolated steps without aid and stable sitting. Refractory seizures persisted, and he exhibited relative macrocephaly, severe kyphoscoliosis, and dysmorphic traits such as small and low-set ears, gingival hyperplasia, and teeth abnormalities. Brain MRI revealed cortical and cerebellar atrophy with ventricular dilatation and a cerebral hygroma. Recurrent infections were also documented. Laboratory findings included thrombocytopenia and hypercholesterolemia. At 7 years old, an abnormal type II plasma transferrin glycoform profile was detected. During adulthood, a slight hypertransaminasemia remained, along with increased GGT and ALP concentrations and persistent low levels of plasma iron and ferritin (**Supplementary Table S1**). Patient remained clinically stable until his death at the age of 21 due to serious complications arising from a respiratory infection.

**Patient 3 (P3).** He is an 8-year-old Caucasian male, the only child of healthy parents, with an uneventful pre- and perinatal history. In infancy, anxiety and poor feeding due to inability to bite were observed, along with reduced sleep duration. At 3 months old he experienced his first episode of seizures characterized by oculogyric and dystonic movements. Neurologically, global muscular hypotonia with tendon reflexes and dysmorphic features such as elongated head with flat occipital region, low-set ears and broad-set nipples were observed. Brain MRI and MRS revealed no abnormalities. At the age of 5 months an abnormal type II plasma transferrin glycoform profile and mucin O-glycosylation deficiency were detected by apolipoprotein CIII analysis. Neuroconductive hearing loss has been present since infancy, and the patient's psychomotor development is significantly delayed. He began sitting independently at 3.5 years old and can now stand with assistance. Although he occasionally shouts, he does not speak.

Behavioural observations include aggressive tendencies, self-mutilation, self-stimulation, hypersensitivity to touch, and stereotypic movements.

The patient's skin presents with a follicular hemorrhagic rash and hyperkeratosis, particularly on the thighs. The prominent clinical symptom, aside from developmental delay, is drug-resistant epilepsy characterized by complex seizures, occasionally occurring in clusters, and generalized tonic-clonic jerks. Notably, after previous unsuccessful attempts with multiple anti-epileptic drugs, the patient was treated with Sultiame, and has been seizure-free for two years, so they remain on this therapy. Biochemical findings included persistent hypertransaminasemia, hypercholesterolemia, and high ALP and CK. Additionally, low plasma copper and ceruloplasmin levels along with low protein S, and hypogammaglobulinemia were also noticed (**Supplementary Table S1**).

## **2.2. Methods**

### **2.2.1. Whole Exome and Whole Genome Sequencing**

Genomic DNA was extracted from blood using the QIAamp DNA Mini Kit (Qiagen, GmbH, Germany). P1 and P2 WES was performed at the Centre Nacional d'Anàlisi Genòmica (CNAG-CRG, Barcelona, Spain) in an Illumina HiSeq 2000 genome analyzer platform, and data was analyzed using a self-developed pipeline [34]. A trio-WES analysis was performed by sequencing the proband (P1) and his parents, while a single WES was carried out for P2. Sequence reads were mapped to the human genome assembly hg19/GRCh37. Variant calls were analyzed using the URDCAT (<https://rdcat.cnag.crg.eu/>) and RD-Connect (<https://platform.rd-connect.eu/>) genome-phenome analysis platforms.

Trio-WGS was conducted on P3 as part of the ZOEMBA international study, focused on

exome-negative patients suspected of having inherited metabolic diseases. Genomic DNA was sequenced by BGI using a BGISEq500 analyzer. All variants, SNVs/CNVs/SVs, were annotated using an in-house developed pipeline. *De novo* mutations were called as described previously [35] and by using DeNovoCNN v1.1.0. Sequence alterations were reported following the guidelines of the Human Genome Variation Society (HGVS). *ATP6API* variants are referred to the GenBank reference sequence NM\_001183.6.

### ***2.2.2. RNA sequencing***

Total RNA was extracted from P1 and P2 fibroblasts, and RNA-seq was performed at CNAG-CRG as described [36]. The analysis of the aligned data was conducted using DROP to detect aberrantly expressed genes, altered splicing events, and monoallelic expression (MAE) of rare variants [12,37,38]. As controls for the detection of aberrant expression and splicing, the cohort of 269 fibroblasts from patients with Mendelian disorders described in Yopez *et al.* 2022, was utilized [13].

### ***2.2.3. Glycosylation studies***

Plasma transferrin glycoforms in P1 and P2 were analyzed by HPLC, following the method described by Helander *et al.* 2003 [39] and modified by Quintana *et al.* 2009 [40]. For P3, transferrin isoelectric focusing (IEF) was performed as described [41], using an anti-transferrin antibody (Dako, #A0061). Analysis of mucin O-glycosylation in P3 was conducted through apolipoprotein CIII (apoCIII) IEF, as reported previously [41]. Analysis of transferrin glycoforms in P3 was performed by high-resolution mass spectrometry of intact transferrin, after immunopurification from plasma [42].

#### **2.2.4. Cell culture**

Human skin fibroblasts were cultured to confluence at 37 °C in a 5% CO<sub>2</sub> environment in 25 cm<sup>2</sup> flasks, harvested by trypsinization, and either pelleted by centrifugation or reseeded. When necessary, cells were treated with cycloheximide (CHX) (Sigma, St. Louis, MO) as previously described [43].

#### **2.2.5. Western blot analysis**

Fibroblasts' pellets were homogenized using RIPA lysis buffer containing a protease inhibitor cocktail (1862209, Merck, Darmstadt, Germany). Protein extracts were subjected to SDS-PAGE and electroblotted. Proteins were visualized by immunostaining with specific antibodies followed by colorimetric detection using the Opti-4CNTM Substrate Kit (Bio-Rad, Hercules, CA, USA). Used antibodies were anti-LAMP2 (NBP2-22217, Novus Biologicals, Centennial, CO, USA) and anti-ATP5A (ab14748, Abcam, Cambridge, UK).

#### **2.2.6. cDNA analysis**

Total RNA was extracted from patients and control fibroblasts using the RNeasy kit 74104 (Qiagen, Hilden, Germany). Single-stranded complementary DNA (cDNA) was synthesized using oligo(dT) primers and M-MLV Reverse Transcriptase, RNase H Minus, Point Mutant (Promega, Madison, WI, USA) following the manufacturer's protocol. *ATP6AP1* cDNA was then amplified by polymerase chain reaction (PCR) and analyzed by Sanger sequencing. *ATP6AP1* mRNA expression was determined by reverse transcription-quantitative PCR (qRT-PCR) using SYBR green reagent (Life Technologies Ltd., Paisley, UK) and specific primers to amplify the *ATP6AP1* wild-type transcript and *GAPDH* (internal control). Primer sequences and PCR conditions are

available upon request.

### **2.2.7. *In silico* predictions**

The analysis to predict the impact of candidate variants on splicing signals using *in silico* methods was conducted using prediction tools. Exonic splicing enhancers (ESEs) in *ATP6AP1* were predicted using ESEfinder version 3.0 (<http://krainer01.cshl.edu/cgi-bin/tools/ESE3>) [44]. Splicing signals were analyzed with Human Splicing Finder (HSF) v2.28 (Genomins, <https://hsf.genomnis.com/>) [45] and SpliceAI [46].

### **2.2.8. *Minigene* assay**

To determine the effect of the *ATP6AP1* variants, minigene assays were performed. Given the proximity of both variants to the 3' end of intron 2, a genomic DNA fragment comprising exon 3 and parts of introns 2 and 3 was amplified and cloned into an Exontrap Cloning Vector pET01 (MoBiTec GmbH, Göttingen, Germany) to generate the wild-type (pET01-WT) and the two mutant constructs (pET01-MUT1 and pET01-MUT2, carrying the c.289-233C>T and c.289-289G>A variants, respectively). To assess the impact of these variants on mRNA processing, the wild-type and mutant plasmids were transfected into HAP1 cells using Lipofectamine3000 (Promega, Madison, WI, USA) following the manufacturer's protocol. Total RNA was extracted 48 hours after transfection, reverse-transcribed, and amplified by PCR using primers corresponding to pET01 exons. The resulting products were visualized on agarose gels and analyzed by Sanger sequencing. All primers and PCR settings are available upon request.

## **3. Results**

### **3.1. *Identification of deep intronic variants in ATP6AP1***

The WES analysis did not prioritize any potentially disease-causing variants in either known CDG-associated genes or other putative genes involved in the cell glycosylation machinery and secretory pathways. Consequently, RNA-seq was performed in P1 and P2, while trio-WGS was conducted in P3.

The analysis of transcriptomic data using the DROP workflow revealed a significant underexpression of *ATP6AP1* (NM\_001183.6) in both P1 and P2, with this gene being the highest significant outlier in both cases (**Fig. 1A**). Additionally, *ATP6AP1* mRNA levels in these individuals were the lowest among the entire cohort of analyzed transcriptomes, demonstrating an approximate 80% reduction (**Fig. 1B**). Furthermore, the RNA-seq data analysis pipeline identified significant *ATP6AP1* aberrant splicing events in both individuals. Examination of the RNA-seq mapped reads revealed an abnormal incorporation of intron 2 sequences into *ATP6AP1* mRNA in both patients, a phenomenon not observed in controls (**Fig. 1C**). Quantitative PCR of *ATP6AP1* mRNA levels, analyzed in P1 and P2, not only corroborated the dramatic reduction previously detected by RNA-seq but also demonstrated the existence of a remaining pool of properly processed *ATP6AP1* transcripts (**Fig. 1D**).

Furthermore, RNA-seq identified two distinct rare variants within the retained intronic sequences that were not covered by WES (c.289-233C>T in P1, and c.289-289G>A in P2). Interestingly, WGS analysis performed in P3 revealed the presence of the same variant already detected in P2 (c.289-289G>A). These variants had not been previously reported in *ATP6AP1*-CDG patients and were not found in gnomAD v.3.1.2 (<https://gnomad.broadinstitute.org>) or 1000 Genomes Project database (<https://www.ensembl.org>) (last accessed, January 2024). Both variants were classified as VUS according to American College of Medical Genetics and Genomics (ACMG) criteria [47].

Segregation studies showed that mother of P1 was a heterozygous carrier of the c.289-233C>T variant, while the father had the wild-type allele, consistent with the X-linked inheritance pattern expected for this disease. In the case of P2 and P3, neither parent carried the variant c.289-289G>A in leukocyte-derived DNA, suggesting a *de novo* mutation event in both families (**Supplementary Fig. S1**).

### **3.2. ATP6AP1 cDNA analyses corroborated aberrant mRNA processing and expression**

To further characterize the RNA-seq findings, *ATP6AP1* mRNA was analyzed by RT-PCR and Sanger sequencing in P1 and P2 fibroblasts grown with or without cycloheximide (CHX). The results revealed a mixture of correctly and abnormally spliced transcripts in CHX-treated patient fibroblasts (**Fig. 2**). *ATP6AP1* cDNA sequencing confirmed the presence of the c.289-233C>T and c.289-289G>A variants in the abnormal transcripts detected in P1 and P2, respectively. A first abnormal transcript revealed the incorporation of a 138-bp segment corresponding to the second intron of *ATP6AP1* (cryptic exon 1, CE1). This alteration was predicted to generate a frameshift resulting in a premature termination codon (p.Leu97Phefs\*40). A second abnormal transcript exhibited nearly the same exonized sequence as the previous one, sharing the same CE1 donor splice site but incorporating a slightly shorter 122-bp sequence into the *ATP6AP1* mRNA due to the activation of an alternative downstream acceptor splice site (cryptic exon 2, CE2). This alteration was also predicted to generate a frameshift leading to a prematurely truncated protein (p.Leu97Glyfs\*6).

### **3.3. Changes in splicing regulatory elements are predicted to cause cryptic exon activation**

The location of the candidate variants within the activated cryptic exons prompted us to consider the alteration of splicing-regulatory elements (SREs) as a probable cause of the observed *ATP6API* mRNA processing alterations. SREs are cis-acting sequences in mRNA that regulate constitutive and alternative splicing by either enhancing (exonic splicing enhancers, ESE) or suppressing (exonic splicing silencers, ESS) splicing.

To identify potential changes in SREs, we performed an *in silico* analysis of *ATP6API* cDNA sequences using Human Splicing Finder (HSF). This web resource is based on nucleotide-frequency matrices defining consensus motifs recognized by human SR proteins (SF2/ASF, SC35, SRp40 or SRp55), which are involved in RNA splicing. HSF allows the scanning of nucleotide sequences to predict putative ESEs that promote splicing. Additionally, sequences were analyzed by ESEfinder, a tool that also predicts the location of SREs by incorporating several published algorithms.

The analysis by HSF of both substitutions identified changes in SREs for the c.289-289G>A variant, predicting a significant alteration of the ESE/ESS motifs ratio (**Supplementary Fig. S2**). Conversely, no significant impact on SREs was predicted for the c.289-233C>T variant. However, ESEfinder predicted the generation of two novel ESE motifs for the c.289-233C>T variant, with respective scores of 2.502 (threshold 2.383) and 3.883 (threshold 2.670) obtained by the SC35 and SRp40 matrices, respectively. In contrast, no alterations in ESE motifs were predicted by ESEfinder for the c.289-289G>A variant.

SpliceAI delta scores were also calculated for both substitutions: c.289-233C>T showed 0.08 (acceptor gain) and 0.06 (donor gain), whereas c.289-289G>A showed 0.31 (acceptor gain) and 0.15 (donor gain) (delta score threshold >0.2).

#### **3.4. Variants c.289-233C>T and c.289-289G>A are responsible for cryptic exon**

### ***activation in ATP6AP1***

Functional studies were conducted to investigate whether the c.289-233C>T and c.289-289G>A variants were responsible for the mRNA processing defects observed in patients. HAP1 cells were transiently transfected with a minigene vector encompassing the genomic regions of each tested variant, as well as the wild-type allele (**Fig. 3A**). The results demonstrated that HAP1 cells transfected with plasmids carrying the *ATP6AP1* wild-type sequence (pET01-WT) produced the expected products, indicating unaltered splicing events. Conversely, cells expressing the plasmids containing the c.289-233C>T (pET01-MUT1) or c.289-289G>A (pET01-MUT2) variants exhibited higher molecular weight bands in the agarose gel analysis, indicative of splicing alterations (**Fig. 3B**). Sanger sequencing of these products revealed a mixture of properly spliced and two different abnormally processed transcripts that included the same 138- and 122-bp retained sequences observed in patients' fibroblasts (**Supplementary Fig. S3**).

### ***3.5. Deficient protein glycosylation***

Based on the clinical suspicion and as a part of the biochemical screening for inherited metabolic diseases, plasma transferrin glycoforms were analyzed in all three individuals. A defect in transferrin N-glycosylation was observed in all patients. P1 and P2 exhibited an elevation of undersialylated transferrin variants consistent with a type II CDG pattern (**Fig. 4A**). Glycosylation studies in the plasma of P3 also revealed a clear defect in transferrin N-glycosylation and mucin O-glycosylation of apoCIII (**Fig. 4B**), both indicating reduced sialylation. Additional analysis of transferrin N-glycan structural abnormalities in P3 by mass spectrometry corroborated N-glycosylation deficiency by showing both reduced sialylation and galactosylation (**Fig. 4C**), which is consistent with the previously reported *ATP6AP1*-CDG patients [48]. In addition, ApoCIII-0 profile is

also comparable.

The glycosylation defect was further corroborated in P1 and P2 fibroblasts by Western blot analysis of LAMP2, a highly N-glycosylated protein. The immunoblotting pattern revealed lower molecular weight bands in both individuals compared to controls, which corresponded to hypoglycosylated LAMP2 (**Fig. 4D**), confirming the glycosylation defect observed in patients' plasma.

#### **4. Discussion**

The implementation of WES in clinical practice deeply improved the diagnostic rates of Mendelian disorders. However, a substantial number of patients still lack a definitive genetic diagnosis, highlighting the need to incorporate new approaches into the diagnostic workup of individuals with suspected genetic diseases. Our study holds particular significance, not only for revealing three novel ATP6AP1-CDG cases, but also, more notably, because the definitive molecular diagnosis could only be achieved after prioritizing the candidate gene through RNA-seq or WGS. This culminated in conclusively ending the diagnostic odyssey in these families, which lasted nearly two decades in families 1 and 2, and eight years in family 3.

ATP6AP1-CDG is an X-linked disorder caused by *ATP6AP1* pathogenic variants, characterized by a wide range of symptoms, including immunodeficiency, hepatopathy, neurological involvement, and connective tissue manifestations [21–33]. While the three individuals in this report share clinical manifestations previously linked to ATP6AP1-CDG, they add further heterogeneity to the complex clinical picture associated with this disorder.

In the context of the ATP6AP1-CDG phenotypic spectrum, P1 exhibited a mild course of the disease with normal psychomotor development, mild cognitive impairment, and non-

progressive evolution. Although a significant number of *ATP6AP1*-CDG patients have shown normal cognitive development [22–26,28,30,31], most individuals also presented hepatopathy, ranging from hepatomegaly or isolated hypertransaminasemia to cirrhosis and liver failure, along with immunodeficiency leading to recurrent infections [21–33]. However, P1 had only slight hepatomegaly with preserved hepatic function and no signs of immunodeficiency, indicating a milder clinical course of the disease. It is noticeable that P2 and P3 presented a more severe phenotype with profound intellectual disability. P2 showed brain MRI abnormalities along with refractory seizures, psychomotor delay, tetraparesis, hepatopathy, and severe immunodeficiency, resulting in a fatal course of the disease. P3 presented profound developmental delay and drug-resistant epilepsy but had a relatively overall stable clinical course. Remarkably, this patient showed deafness in infancy, a feature previously reported only in a single *ATP6AP1*-CDG family [22]. Remarkably, P3 has remained seizure-free for two years since the initiation of treatment with sultiamide, a carbonic anhydrase (CA) inhibitor. Interestingly, acetazolamide treatment, that is also a CA inhibitor, was found to be effective in the treatment of some neurological symptoms in *PMM2*-CDG patients [49]. However, further investigations will be required to determine whether CA inhibitors could also be effective in other CDGs.

All patients were initially diagnosed of type II CDG based on the observed plasma transferrin N-glycosylation alterations, which were further corroborated by *LAMP2* hypoglycosylation in fibroblasts. Initially, WES did not prioritize any potentially disease-causing variant in glycosylation-related genes. Fortunately, subsequent RNA-seq (in P1 and P2) and WGS (in P3) prioritized intronic variants in *ATP6AP1*. Transcriptomic data revealed aberrantly reduced *ATP6AP1* expression and defective splicing in two of the

patients. Analysis of cDNA confirmed the RNA-seq results and provided a comprehensive characterization of the splicing defect, showing similar alterations in both individuals. Two different abnormal transcripts, together with a remaining pool of properly spliced mRNA, were observed. Both aberrant transcripts presented different-sized and partially overlapping intron retentions carrying the rare intronic variant c.289-233C>T in P1, and the c.289-289G>A in P2. These cryptic exons were located in the second intron of *ATP6API* and were not covered by WES. Segregation studies showed a maternal inheritance in P1 but, interestingly, the same variant was originated *de novo* in P2 and P3. Both aberrant transcripts were exclusively detected in CHX-treated patients' fibroblasts and predicted to disrupt the protein reading frame by generating premature stop codons, leading to nonsense mediated mRNA decay (NMD) degradation, consistent with the low *ATP6API* mRNA levels detected in the patients. Interestingly, the retained intronic sequences are present in non-canonical *ATP6API* transcripts (ENST00000439372 and ENST00000455205, <https://www.ensembl.org/>) predicted to be degraded by NMD (**Supplementary Fig. S4**). Taken together, these findings suggest that this genomic region could be relatively prone to behave as an exon, and the identified variants might exert their pathological effect by enhancing the incorporation of such intronic sequences into *ATP6API* mRNA, promoting the formation of non-functional alternative transcripts. This hypothesis aligns with in-silico analyses predicting changes in SREs that could explain the observed splicing alterations.

To conclusively demonstrate the causative effect of the identified variants, we conducted in vitro functional studies. Transient expression of the c.289-233C>T and the c.289-289G>A variants in HAP1 cells using a minigene system revealed a significant splicing defect caused by both substitutions. Similar to what was observed in patients' fibroblasts, a mixture of aberrantly processed transcripts alongside a residual pool of properly spliced

transcript was detected. This pattern exhibited identical retained intronic sequences, mirroring those seen in the patients' fibroblasts. Therefore, our functional data strongly suggests that the prioritized variants are responsible for the *ATP6AP1* mRNA alterations observed in the patients, and that they may exert their pathogenic effect by altering mRNA processing.

These results, combined with mRNA expression data, prompt us to speculate that *ATP6AP1* function could be partially preserved due to a leaky splicing defect, allowing the production of a residual amount of wild-type transcript. Nonetheless, the pool of intact *ATP6AP1* mRNA does not seem to be sufficient to restore protein function and attenuate clinical presentation, at least in P2 and P3. Additionally, since the levels of wild-type *ATP6AP1* transcript were similar in P1 and P2, it appears that additional, yet unknown modifying factors may be modulating clinical presentation.

Our study underscores the importance of incorporating RNA-seq as an additional and complementary tool to WES and WGS for diagnosing patients with CDG and other Mendelian disorders. Furthermore, the genetic findings illustrate how RNA-seq can contribute to a deeper understanding of the molecular pathomechanisms underlying human disease. It is noteworthy that only a limited number of studies have explored the utility of RNA-seq for the diagnosis of unsolved CDG patients to date [50,51].

## **5. Conclusions**

In summary, our study emphasizes the need of integrating RNA-seq and WGS, together with functional validation, into routine diagnosis. In this regard, the development of appropriate bioinformatic tools is crucial for its reliable implementation in the diagnostic workup. Additionally, our study provides new insights into the molecular basis of *ATP6AP1*-CDG by reporting the first two deep intronic mutations affecting mRNA

processing in this X-linked gene. Furthermore, functional validation studies demonstrated that the pathogenic effect of both variants is exerted through the activation of two overlapping cryptic exons that disrupt *ATP6AP1* transcript.

## **Acknowledgments**

The authors express their gratitude to Dr. Marisa Girós and Dr. Paz Briones, former members of our team, for her invaluable contribution to the biochemical diagnosis of congenital disorders of glycosylation. We thank Metakids and the UMD (United for Metabolic Diseases) for the funds that were released in order to conduct this study. The authors would also like to thank the families for their participation in this study.

## **Data availability**

The data in this study is available in the manuscript, supplementary information and figures. The *ATP6API* identified variants have been submitted to the ClinVar database [Accession numbers: SCV004231892 and SCV004231907].

## References

- [1] Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med*. 2016;18(7):696–704. <https://doi.org/10.1038/gim.2015.148>.
- [2] Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med*. 2018;3:16. <https://doi.org/10.1038/s41525-018-0053-8>.
- [3] Wortmann SB, Oud MM, Alders M, Coene KLM, van der Crabben SN, Feichtinger RG, et al. How to proceed after “negative” exome: A review on genetic diagnostics, limitations, challenges, and emerging new multiomics techniques. *J Inherit Metab Dis*. 2022;45(4):663–681. <https://doi.org/10.1002/jimd.12507>.
- [4] Alfares A, Aloraini T, subaie L Al, Alissa A, Qudsi A Al, Alahmad A, et al. Whole-genome sequencing offers additional but limited clinical utility compared with reanalysis of whole-exome sequencing. *Genet Med*. 2018;20(11):1328–1333. <https://doi.org/10.1038/gim.2018.41>.
- [5] Palmer EE, Sachdev R, Macintosh R, Melo US, Mundlos S, Righetti S, et al. Diagnostic Yield of Whole Genome Sequencing After Nondiagnostic Exome Sequencing or Gene Panel in Developmental and Epileptic Encephalopathies. *Neurology*. 2021;96(13):e1770–e1782. <https://doi.org/10.1212/wnl.0000000000011655>.
- [6] Grether A, Ivanovski I, Russo M, Begemann A, Steindl K, Abela L, et al. The current benefit of genome sequencing compared to exome sequencing in patients with developmental or epileptic encephalopathies. *Mol Genet Genomic Med*. 2023;11(5):e2148. <https://doi.org/10.1002/mgg3.2148>.
- [7] Kremer LS, Bader DM, Mertes C, Kopajtich R, Pichler G, Iuso A, et al. Genetic

diagnosis of Mendelian disorders via RNA sequencing. *Nat Commun.* 2017;8: 15824.  
<https://doi.org/10.1038/ncomms15824>.

[8] Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med.* 2017;9(386):eaal5209. <https://doi.org/10.1126/scitranslmed.aal5209>.

[9] Frésard L, Smail C, Ferraro NM, Teran NA, Li X, Smith KS, et al. Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nat Med.* 2019;25(6):911-919. <https://doi.org/10.1038/s41591-019-0457-8>.

[10] Gonorazky HD, Naumenko S, Ramani AK, Nelakuditi V, Mashouri P, Wang P, et al. Expanding the Boundaries of RNA Sequencing as a Diagnostic Tool for Rare Mendelian Disease. *Am J Hum Genet.* 2019;104(3):466–483.  
<https://doi.org/10.1016/j.ajhg.2019.04.004>.

[11] Stenton SL, Prokisch H. The Clinical Application of RNA Sequencing in Genetic Diagnosis of Mendelian Disorders. *Clin Lab Med.* 2020;40(2):121-133.  
<https://doi.org/10.1016/j.cll.2020.02.004>.

[12] Yépez VA, Mertes C, Müller MF, Klaproth-Andrade D, Wachutka L, Frésard L, et al. Detection of aberrant gene expression events in RNA sequencing data. *Nat Protoc.* 2021;16(2):1276–1296. <https://doi.org/10.1038/s41596-020-00462-5>.

[13] Yépez VA, Gusic M, Kopajtich R, Mertes C, Smith NH, Alston CL, et al. Clinical implementation of RNA sequencing for Mendelian disease diagnostics. *Genome Med.* 2022;14(1):38. <https://doi.org/10.1186/s13073-022-01019-9>.

[14] Murdock DR, Dai H, Burrage LC, Rosenfeld JA, Ketkar S, Müller MF, et al. Transcriptome-directed analysis for Mendelian disease diagnosis overcomes limitations of conventional genomic testing. *J Clin Invest.* 2021;131(1): e141500.  
<https://doi.org/10.1172/jci141500>.

- [15] Dekker J, Schot R, Bongaerts M, de Valk WG, van Veghel-Plandsoen MM, Monfils K, et al. Web-accessible application for identifying pathogenic transcripts with RNA-seq: Increased sensitivity in diagnosis of neurodevelopmental disorders. *Am J Hum Genet.* 2023;110(2):251–272. <https://doi.org/10.1016/j.ajhg.2022.12.015>.
- [16] Sosicka P, Ng BG, Freeze HH. Chemical Therapies for Congenital Disorders of Glycosylation. *ACS Chem Biol.* 2022;17(11):2962–2971. <https://doi.org/10.1021/acscchembio.1c00601>.
- [17] Ferreira CR, Altassan R, Marques-Da-Silva D, Francisco R, Jaeken J, Morava E. Recognizable phenotypes in CDG. *J Inherit Metab Dis.* 2018;41(3):541–553. <https://doi.org/10.1007/s10545-018-0156-5>.
- [18] Lefeber DJ, Freeze HH, Steet R, Kinoshita T. Congenital Disorders of Glycosylation. In: Varki A, Cummings RD, Esko JD, et al., eds. *Essentials of Glycobiology*. 4th ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 2022. Chapter 45. <https://doi.org/10.1101/glycobiology.4e.45>.
- [19] Cannata Serio M, Graham LA, Ashikov A, Larsen LE, Raymond K, Timal S, et al. Mutations in the V-ATPase Assembly Factor VMA21 Cause a Congenital Disorder of Glycosylation With Autophagic Liver Disease. *Hepatology.* 2020;72(6):1968–1986. <https://doi.org/10.1002/hep.31218>.
- [20] J.R. Jansen E, J.M. Martens G. Novel insights into V-ATPase functioning: distinct roles for its accessory subunits ATP6AP1/Ac45 and ATP6AP2/(pro) renin receptor. *Curr Protein Pept Sci.* 2012;13(2):124–133. <https://doi.org/10.2174/138920312800493160>.
- [21] Evers C, Staufner C, Granzow M, Paramasivam N, Hinderhofer K, Kaufmann L, et al. Impact of clinical exomes in neurodevelopmental and neurometabolic disorders. *Mol Genet Metab.* 2017;121(4):297–307. <https://doi.org/10.1016/j.ymgme.2017.06.014>.
- [22] Jansen EJR, Timal S, Ryan M, Ashikov A, Van Scherpenzeel M, Graham LA, et al.

ATP6AP1 deficiency causes an immunodeficiency with hepatopathy, cognitive impairment and abnormal protein glycosylation. *Nat Commun.* 2016; 7:11600. <https://doi.org/10.1038/ncomms11600>.

[23] Witters P, Breckpot J, Foulquier F, Preston G, Jaeken J, Morava E. Expanding the phenotype of metabolic cutis laxa with an additional disorder of N-linked protein glycosylation. *Eur J Hum Genet.* 2018;26(5):618-621. <https://doi.org/10.1038/s41431-017-0044-8>.

[24] Dimitrov B, Himmelreich N, Hipgrave Ederveen AL, Lüchtenborg C, Okun JG, Breuer M, et al. Cutis laxa, exocrine pancreatic insufficiency and altered cellular metabolomics as additional symptoms in a new patient with ATP6AP1-CDG. *Mol Genet Metab.* 2018;123(3):364–374. <https://doi.org/10.1016/j.ymgme.2018.01.008>.

[25] Ondruskova N, Honzik T, Vondrackova A, Stranecky V, Tesarova M, Zeman J, et al. Severe phenotype of ATP6AP1-CDG in two siblings with a novel mutation leading to a differential tissue-specific ATP6AP1 protein pattern, cellular oxidative stress and hepatic copper accumulation. *J Inherit Metab Dis.* 2020;43(4):694–700. <https://doi.org/10.1002/jimd.12237>.

[26] Tvina A, Thomsen A, Palatnik A. Prenatal and postnatal phenotype of a pathologic variant in the ATP6AP1 gene. *Eur J Med Genet.* 2020;63(6):103881. <https://doi.org/10.1016/j.ejmg.2020.103881>.

[27] Lipinski P, Rokicki D, Bogdanska A, Lesiak J, Lefeber DJ, Tylki-Szymanska A. ATP6AP1-CDG: Follow-up and female phenotype. *JIMD Rep.* 2020;53(1):80–82. <https://doi.org/10.1002/jmd2.12104>.

[28] Quelhas D, Martins E, Azevedo L, Bandeira A, Diogo L, Garcia P, et al. Congenital Disorders of Glycosylation in Portugal-Two Decades of Experience. *J Pediatr.* 2021 Apr;231:148-156. <https://doi.org/10.1016/j.jpeds.2020.12.026>.

- [29] Yang X, Lv ZL, Tang Q, Chen XQ, Huang L, Yang MX, et al. Congenital disorder of glycosylation caused by mutation of ATP6AP1 gene (c.1036G>A) in a Chinese infant: A case report. *World J Clin Cases.* 2021;9(26):7876–7885. <https://doi.org/10.12998/wjcc.v9.i26.7876>.
- [30] Barua S, Berger S, Pereira EM, Jobanputra V. Expanding the phenotype of ATP6AP1 deficiency. *Cold Spring Harb Mol Case Stud.* 2022;8(4): a006195. <https://doi.org/10.1101/mcs.a006195>.
- [31] Alharbi H, Daniel EJP, Thies J, Chang I, Goldner DL, Ng BG, et al. Fractionated plasma N-glycan profiling of novel cohort of ATP6AP1-CDG subjects identifies phenotypic association. *J Inherit Metab Dis.* 2023;46(2):300–312. <https://doi.org/10.1002/jimd.12589>.
- [32] Dang Do AN, Chang IJ, Jiang X, Wolfe LA, Ng BG, Lam C, et al. Elevated oxysterol and N-palmitoyl-O-phosphocholineserine levels in congenital disorders of glycosylation. *J Inherit Metab Dis.* 2023;46(2):326–334. <https://doi.org/10.1002/jimd.12595>.
- [33] Semenova N, Shatokhina O, Shchagina O, Kamenec E, Marakhonov A, Degtyareva A, et al. Clinical Presentation of a Patient with a Congenital Disorder of Glycosylation, Type IIs (ATP6AP1), and Liver Transplantation. *Int J Mol Sci.* 2023;24(8):7449. <https://doi.org/10.3390/ijms24087449>.
- [34] Laurie S, Fernandez-Callejo M, Marco-Sola S, Trotta JR, Camps J, Chacón A, et al. From Wet-Lab to Variations: Concordance and Speed of Bioinformatics Pipelines for Whole Genome and Whole Exome Sequencing. *Hum Mutat.* 2016;37(12):1263–1271. <https://doi.org/10.1002/humu.23114>.
- [35] Lelieveld SH, Reijnders MRF, Pfundt R, Yntema HG, Kamsteeg EJ, De Vries P, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci.* 2016;19(9):1194–1196. <https://doi.org/10.1038/nn.4352>.

- [36] Muñoz-Pujol G, Ortigoza-Escobar JD, Paredes-Fuentes AJ, Jou C, Ugarteburu O, Gort L, et al. Leigh syndrome is the main clinical characteristic of PTC3D3 deficiency. *Brain Pathol.* 2023;33(3):e13134. <https://doi.org/10.1111/bpa.13134>.
- [37] Brechtmann F, Mertes C, Matusėvičiūtė A, Yépez VA, Avsec Ž, Herzog M, et al. OUTRIDER: A Statistical Method for Detecting Aberrantly Expressed Genes in RNA Sequencing Data. *Am J Hum Genet.* 2018;103(6):907–917. <https://doi.org/10.1016/j.ajhg.2018.10.025>.
- [38] Mertes C, Scheller IF, Yépez VA, Çelik MH, Liang Y, Kremer LS, et al. Detection of aberrant splicing events in RNA-seq data using FRASER. *Nat Commun.* 2021;12(1):529. <https://doi.org/10.1038/s41467-020-20573-7>.
- [39] Helander A, Husa A, Jeppsson JO. Improved HPLC Method for Carbohydrate-deficient Transferrin in Serum. *Clin Chem.* 2003;49(11):1881–1890. <https://doi.org/10.1373/clinchem.2003.023341>.
- [40] Quintana E, Navarro-Sastre A, Hernández-Pérez JM, García-Villoria J, Montero R, Artuch R, et al. Screening for congenital disorders of glycosylation (CDG): transferrin HPLC versus isoelectric focusing (IEF). *Clin Biochem.* 2009;42(4–5):408–415. <https://doi.org/10.1016/j.clinbiochem.2008.12.013>.
- [41] Linders PTA, Gerretsen ECF, Ashikov A, Vals MA, de Boer R, Revelo NH, et al. Congenital disorder of glycosylation caused by starting site-specific variant in syntaxin-5. *Nat Commun.* 2021;12(1):6227. <https://doi.org/10.1038/s41467-021-26534-y>.
- [42] Van Scherpenzeel M, Steenbergen G, Morava E, Wevers RA, Lefeber DJ. High-resolution mass spectrometry glycoprofiling of intact transferrin for diagnosis and subtype identification in the congenital disorders of glycosylation. *Transl Res.* 2015;166(6):639-649. <https://doi.org/10.1016/j.trsl.2015.07.005>.
- [43] Macías-Vidal J, Gort L, Lluch M, Pineda M, Coll MJ. Nonsense-mediated mRNA

decay process in nine alleles of Niemann-Pick type C patients from Spain. *Mol Genet Metab.* 2009;97(1):60–64. <https://doi.org/10.1016/j.ymgme.2009.01.007>.

[44] Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* 2003;31(13):3568–3571. <https://doi.org/10.1093/nar/gkg616>.

[45] Desmet FO, Hamroun D, Lalande M, Collod-B eroud G, Claustres M, B eroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009;37(9):e67. <https://doi.org/10.1093/nar/gkp215>.

[46] Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell.* 2019;176(3):535-548.e24. <https://doi.org/10.1016/j.cell.2018.12.015>.

[47] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405–424. <https://doi.org/10.1038/gim.2015.30>.

[48] Jansen JC, Cirak S, Van Scherpenzeel M, Timal S, Reunert J, Rust S, et al. CCDC115 Deficiency Causes a Disorder of Golgi Homeostasis with Abnormal Protein Glycosylation. *Am J Hum Genet.* 2016;98(2):310–321. <https://doi.org/10.1016/j.ajhg.2015.12.010>.

[49] Mart nez-Monseny AF, Bolasell M, Callej n-P oo L, Cuadras D, Freniche V, Itzep DC, Gassiot S, et al. AZATAx: Acetazolamide safety and efficacy in cerebellar syndrome in PMM2 congenital disorder of glycosylation (PMM2-CDG). *Ann Neurol.* 2019;85(5):740-751. doi: 10.1002/ana.25457.

[50] Jabato FM, C rdoba-Caballero J, Rojano E, Rom -Mateo C, Sanz P, P rez B, et al.

Gene expression analysis method integration and co-expression module detection applied to rare glucide metabolism disorders using ExpHunterSuite. *Sci Rep.* 2021;11(1):15062. <https://doi.org/10.1038/s41598-021-94343-w>.

[51] Maddirevula S, Kuwahara H, Ewida N, Shamseldin HE, Patel N, Alzahrani F, et al. Analysis of transcript-deleterious variants in Mendelian disorders: implications for RNA-based diagnostics. *Genome Biol.* 2020;21(1):145. <https://doi.org/10.1186/s13059-020-02053-9>.

## FIGURE LEGENDS

**Fig. 1.** RNA-seq and qPCR analyses in fibroblasts from P1 and P2. (A) RNA expression volcano plots showing gene level statistical significance ( $-\log_{10}$  p-value) against z-Score. Red dots highlight the aberrantly expressed genes in each patient, with *ATP6API* being the most aberrantly expressed gene in both P1 and P2. (B) Expression rank plot of *ATP6API* showing P1 and P2 (red dots) as the individuals with the lowest levels of *ATP6API* among the entire cohort of transcriptomes analyzed. (C) *ATP6API* Sashimi plot showing the aberrant incorporation of intronic sequences into *ATP6API* mRNA in patients' fibroblasts (P1 and P2 in red and orange, respectively) compared to controls (in black). The RNA coverage is indicated between brackets. The counts of split-reads spanning the given introns are indicated alongside the lines connecting the exons. Below, the transcript of *ATP6API* is shown according to the RefSeq NM\_001183.6 annotation. (D) Relative *ATP6API* mRNA levels analyzed by qPCR in fibroblasts from P1 and P2, and control individuals. *GAPDH* was used as the internal control. Data are expressed as means and standard errors of the means. Two biological replicates were performed. \*\*, p-value  $\leq 0.01$  for P1 and P2 versus control; RU, relative units.

**Fig. 2.** *ATP6API* cDNA analysis in fibroblasts from patients P1 and P2. (A) RT-PCR followed by agarose gel analysis and sanger sequencing revealed the incorporation of intronic sequences into the *ATP6API* mRNA. These alterations were not observed in fibroblasts from control individuals. The same two abnormally spliced transcripts were detected in both P1 and P2, each one incorporating one of the partially overlapping cryptic exons (CE1 and CE2). Properly spliced transcripts were also detected in both patients. (B) Electropherograms from *ATP6API* cDNA sequencing showing the exon junctions of interest, and the corresponding sequences for both the canonical (top) and the abnormal

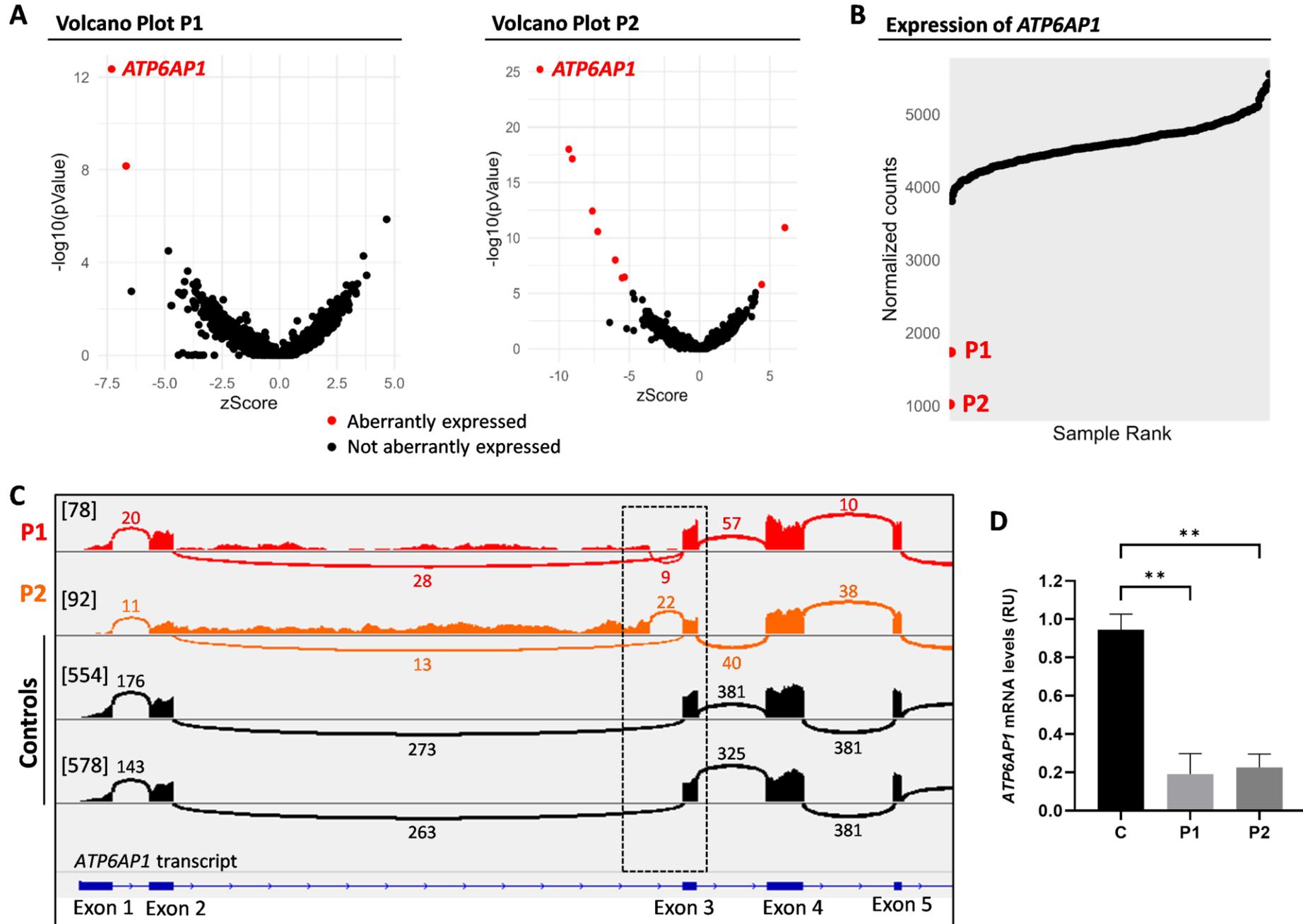
(bottom) *ATP6API* transcripts are shown. The extra 16-bp sequence from CE1 relative to CE2 is underlined. The variant c.289-289G>A from P2 is highlighted with a black square. (C) A schematic representation of the altered *ATP6API* mRNA processing is provided. C, control; P1, patient 1; P2, patient 2; CE1, cryptic exon 1; CE2, cryptic exon 2.

**Fig. 3.** The c.289-233C>T and c.289-289G>A variants lead to altered *ATP6API* mRNA processing. (A) Minigene assays using an Exontrap Cloning Vector pET01 (Mobitec, Göttingen, Germany) incorporating the wild-type or the mutant sequence (carrying the c.289-233C>T or the c.289-289G>A variant) were transfected in HAP1 cells and analyzed by RT-PCR. (B) RT-PCR using vector-specific primers showed defective splicing as a consequence of c.289-233C>T and c.289-289G>A variants. M1: Exontrap exon 1; M2: Exontrap exon 2; *ATP6API* insert: exon 3 and the flanking intronic regions, including the retained sequences from intron 2 (CE); WT: wild-type sequence; MUT: sequence containing the c.289-233C>T or the c.289-289G>A variant; CE, partially overlapping cryptic exons (CE1 and CE2). Red arrows indicate localization of the tested variants.

**Fig. 4.** Glycosylation studies in patients. (A) Plasma transferrin (Trf) glycoform profiles of P1 and P2 obtained through high-performance liquid chromatography (HPLC) using anion-exchange. The vertical axes represent the normalized 470 nm-absorbance percentage of each glycoform relative to the highest peak (tetra-sialoTrf). (B) Table showing quantification results of Trf glycoforms in P1-P3, expressed as the relative percentage of the total Trf signal intensity. P1 and P2 were analyzed by anion-exchange HPLC, while P3 was analyzed by isoelectrofocusing (IEF). An elevation in plasma monosialo- and trisialo Trf was observed in P1 and P2, whereas an increase in disialo-

and trisialoTrf, together with a decrease in tetrasialo Trf, were detected in P3. These findings were consistent with deficient processing of N-glycans, indicating a type II congenital disorder of glycosylation (CDG-II). For P3, quantification of apoCIII glycoforms performed by IEF revealed increased apoCIII-0 and apoCIII-1, consistent with mucine O-glycosylation deficiency. Reference values (R.V.) for plasma Trf and apoCIII glycoforms are provided; N.d., not done. (C) Analysis of Trf glycoforms by high-resolution mass spectrometry of intact Trf, after immunopurification from plasma of P3. Peak numbers in the profiles correspond to the Trf glycoforms as presented in the legend. M: molecular mass; H: Height; R: relative ratio of the indicated Trf glycoform as compared to the normal Trf glycoform (peak 28). Results indicate reduced sialylation and galactosylation. (D) Western blot analysis of LAMP2 in fibroblasts from P1 and P2. The immunoblotting pattern revealed a smear with some additional lower molecular weight bands in both patients when compared to controls. These bands corresponded to hypoglycosylated forms of LAMP2. These findings confirmed the observed glycosylation defect in plasma from P1 and P2. ATP5A was used as a loading control.

# Figure 1



**Figure 2**

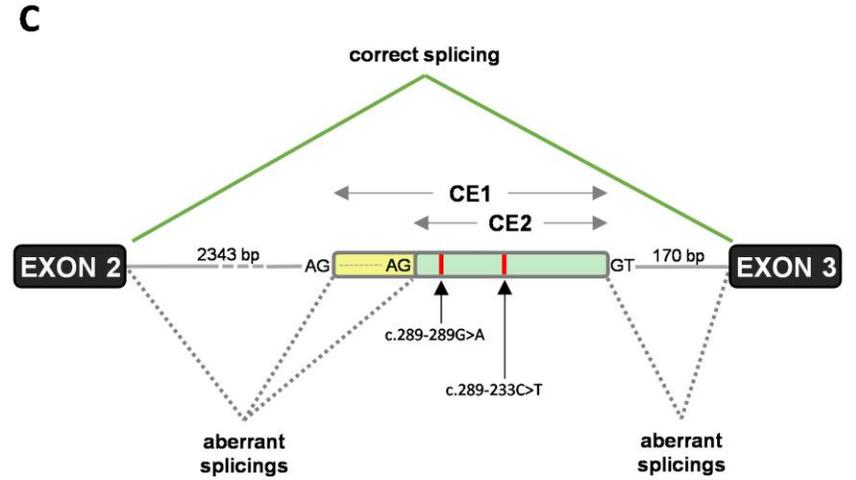
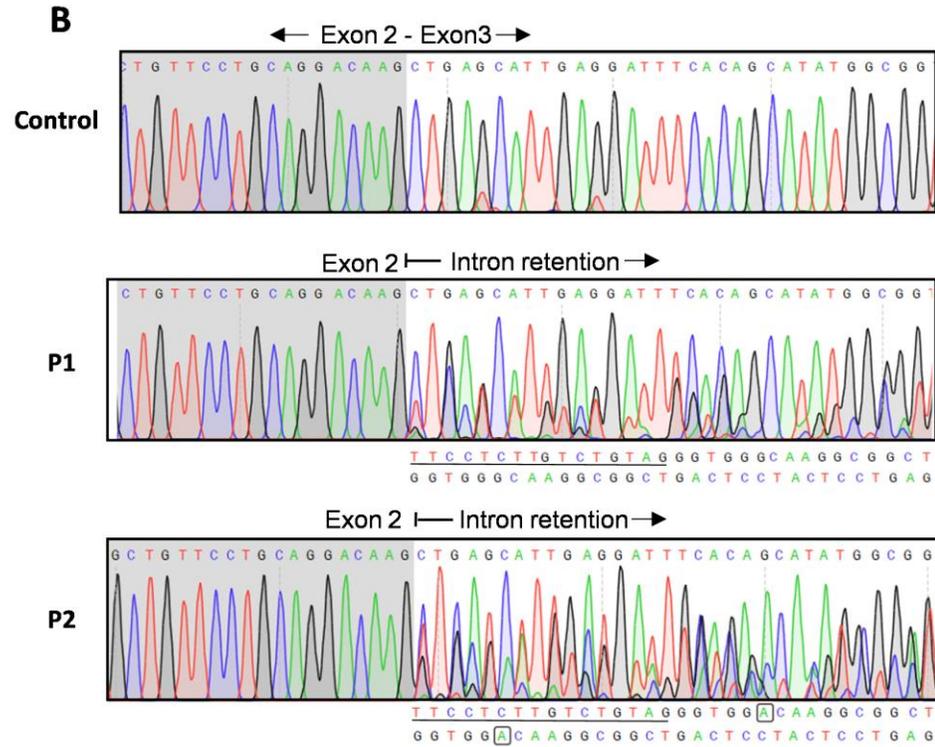
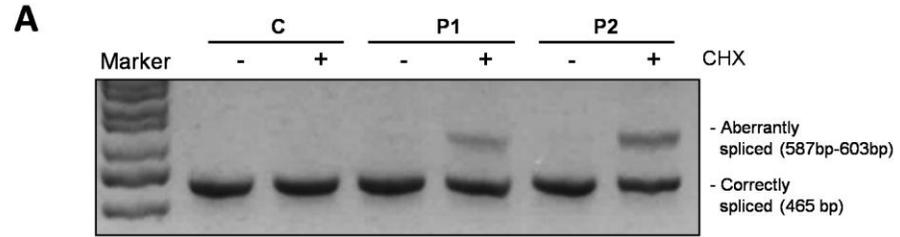
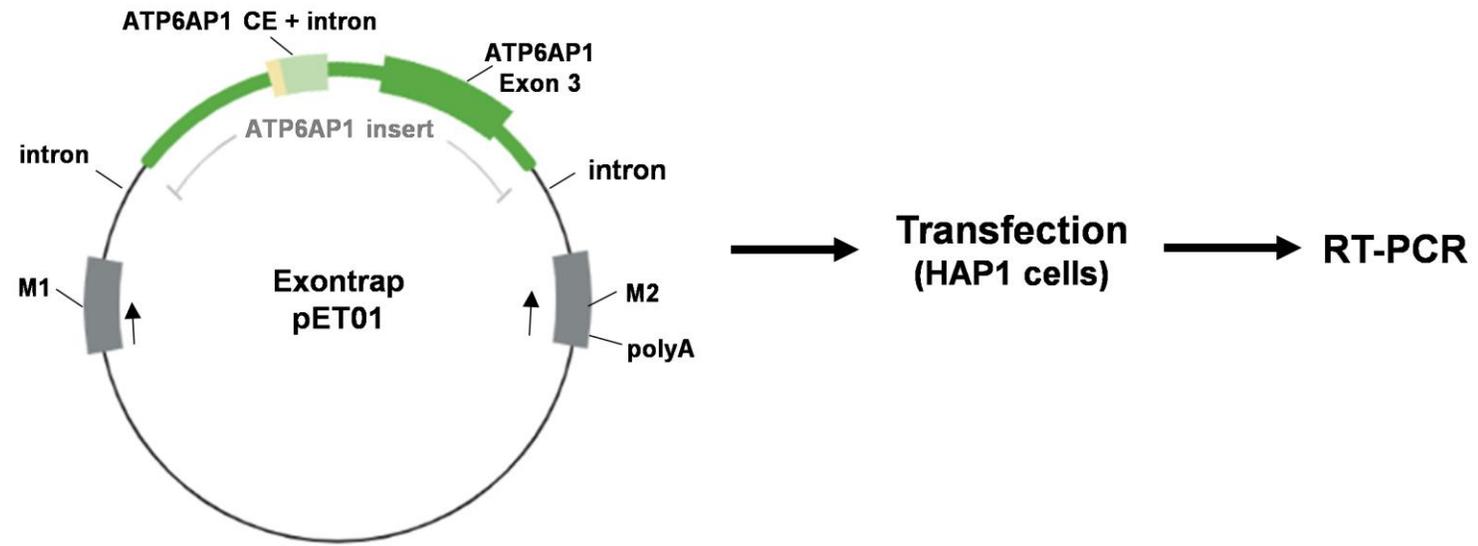
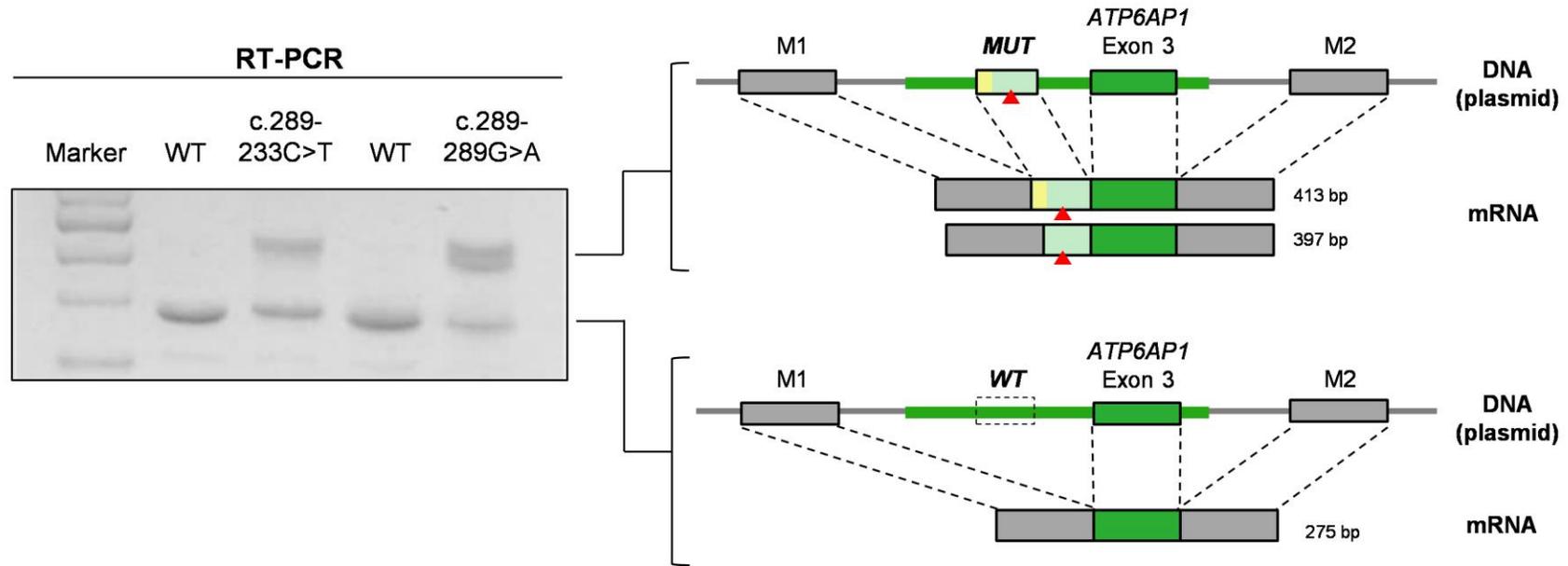


Figure 3

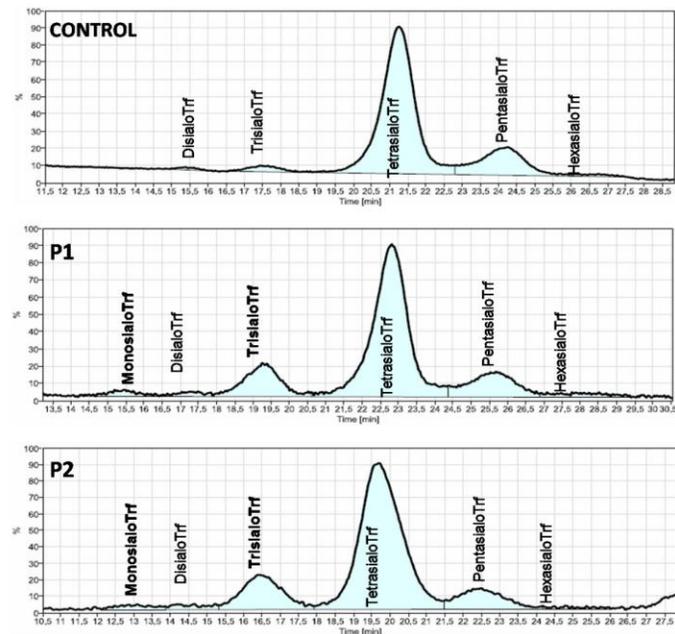
A



B



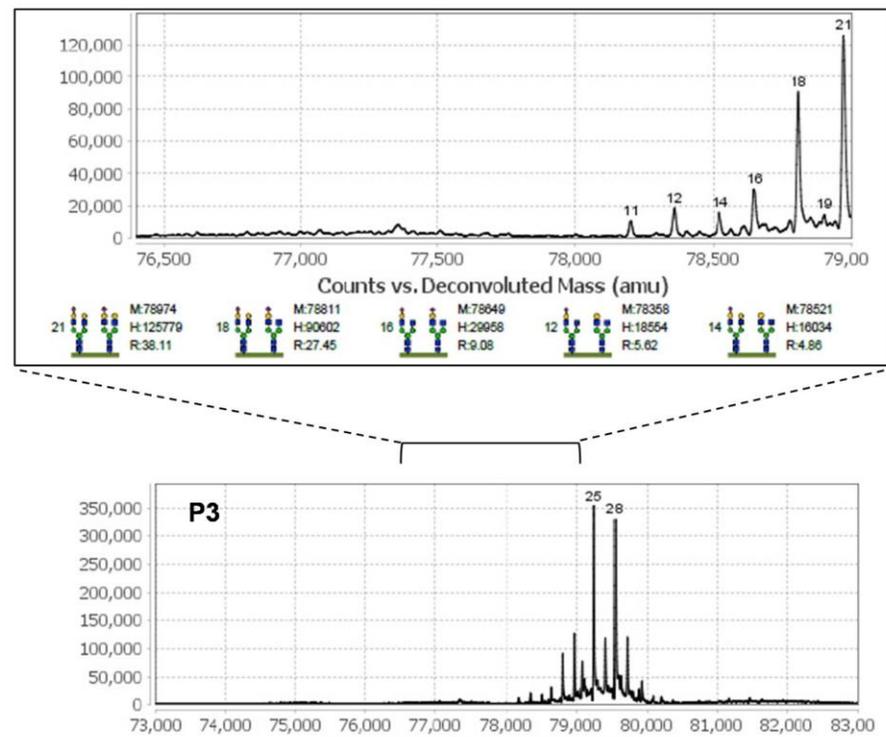
A



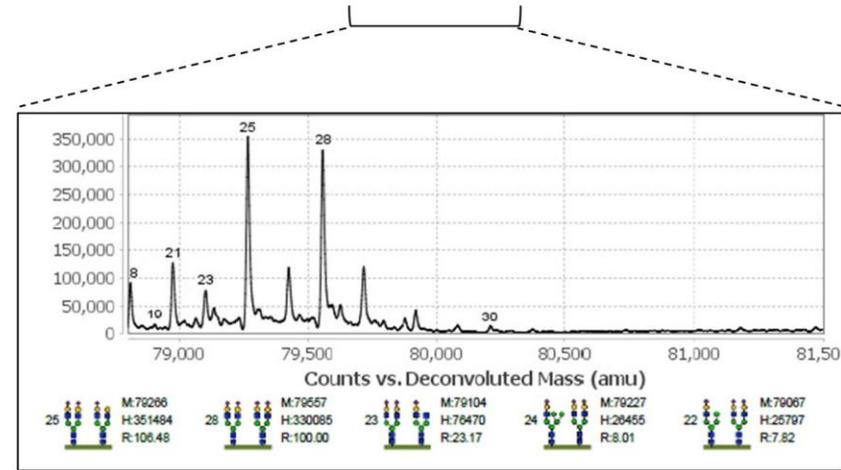
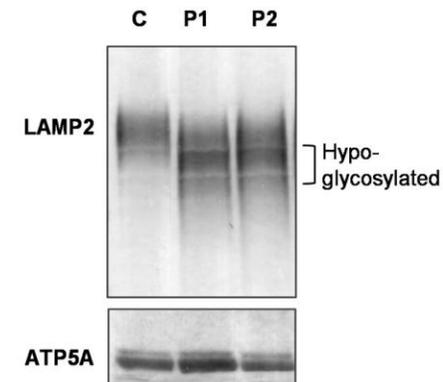
B

|                | P1            | P2            | R.V. (%)    | P3            | R.V. (%)    |
|----------------|---------------|---------------|-------------|---------------|-------------|
| A-sialoTrf     | < 0.5         | < 0.5         | < 0.5       | 1.8           | 0.0 - 3.2   |
| Mono-sialoTrf  | <b>2.5 ↑</b>  | <b>2.1 ↑</b>  | 0.0 - 0.7   | 4.5           | 0.0 - 5.0   |
| Di-sialoTrf    | 1.7           | 2.1           | 0.7 - 2.8   | <b>16.4 ↑</b> | 3.3 - 7.6   |
| Tri-sialoTrf   | <b>14.2 ↑</b> | <b>15.6 ↑</b> | 1.7 - 8.7   | <b>27.3 ↑</b> | 4.9 - 10.6  |
| Tetra-sialoTrf | 64.4          | 68.4          | 60.0 - 84.6 | <b>39.7 ↓</b> | 47.3 - 62.7 |
| Penta-sialoTrf | 17.2          | 11.9          | 10.3 - 25.0 | 9.1           | 18.7 - 31.5 |
| Hexa-sialoTrf  |               |               |             | 1.3           | 3.2 - 7.8   |
| ApoCIII-0      | n.a.          | n.a.          | -           | <b>10.2 ↑</b> | 0.2 - 4.5   |
| ApoCIII-1      | n.a.          | n.a.          | -           | <b>78.8 ↑</b> | 42.7 - 69.8 |
| ApoCIII-2      | n.a.          | n.a.          | -           | <b>11.1 ↓</b> | 26.2 - 56.7 |

C



D





[Click here to access/download](#)

**Supporting File**

[Supplementary Information\\_R1\\_clean\\_copy.docx](#)

