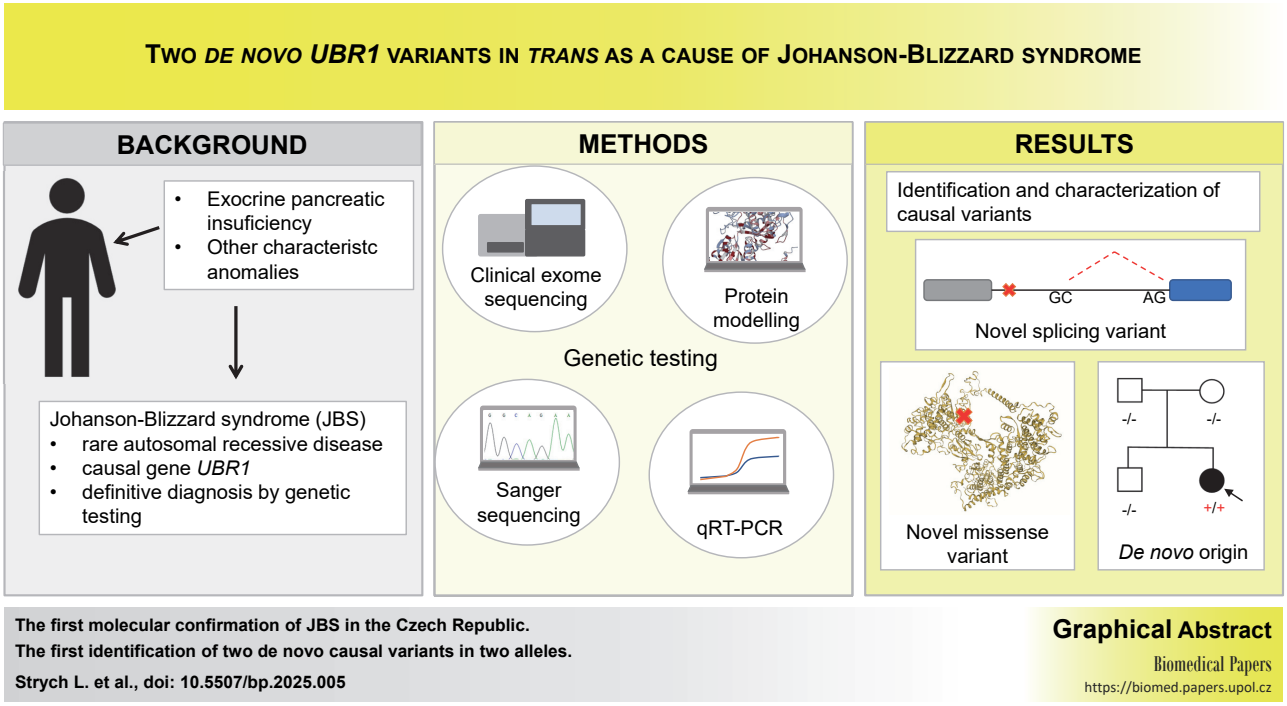
Two *de novo UBR1* variants in *trans* as a cause of Johanson-Blizzard syndrome

Lukas Strych[1], Tomas Zavoral[1], Pavla Komrskova[1], Tomas Vanecek[2,3], Ivan Subrt[1]

Aims/Background. Johanson-Blizzard syndrome (JBS) is a rare autosomal recessive disease caused by pathogenic variants in the *UBR1* gene. JBS is usually suspected based on characteristic anomalies, but only genetic testing provides a definitive diagnosis. Since most variants are inherited from the parents, we aimed to identify the causal variants in a Czech proband with clinically suspected JBS and perform segregation analysis.
Methods. A proband with clinically suspected JBS underwent clinical exome sequencing (CES). Sanger sequencing was used for the validation, characterization, and segregation of variants in the family. The variants were also characterized using quantitative real-time PCR (qPCR) and *in silico* analysis.
Results. Using CES in the proband, we identified two novel causal variants in the *UBR1* gene, c.3482A>C and c.3509+6T>C. Although the variants were found in *trans*, neither was detected in the parents. Sanger sequencing of the cDNA revealed that the novel variant c.3509+6T>C caused activation of the non-canonical GC donor splice site. The inclusion of 70 bp of the intronic sequence generated a frameshift and a premature termination codon leading to nonsense-mediated decay, as detected by qPCR. *In silico* protein structural analysis showed that the novel missense variant c.3482A>C in the zinc-stabilized domain RING-H2 altered a highly conserved zinc-coordinating histidine by proline.
Conclusion. To the best of our knowledge, we report the first molecular confirmation of JBS in the Czech Republic and the first identification of two *de novo* causal variants in two alleles. Our findings also expand the spectrum of pathogenic variants in the *UBR1* gene.



Key words: Johanson-Blizzard syndrome, *UBR1* gene, novel variant, *de novo*

Received: December 9, 2024; Revised: January 27, 2025; Accepted: January 28, 2025; Available online: February 5, 2025
https://doi.org/10.5507/bp.2025.005
© 2025 The Authors; https://creativecommons.org/licenses/by/4.0/

[1]*Department of Medical Genetics, Faculty of Medicine in Pilsen, Charles University and University Hospital Pilsen, Pilsen, Czech Republic*
[2]*Sikl's Department of Pathology, University Hospital Pilsen, Pilsen, Czech Republic*
[3]*Biopticka laborator s.r.o., Pilsen, Czech Republic*
Corresponding author: Lukas Strych, e-mail: Lukas.Strych@lfp.cuni.cz

INTRODUCTION

Johanson-Blizzard syndrome (JBS) (MIM #243800) is a rare autosomal recessive disease characterized by pancreatic exocrine insufficiency and other abnormalities comprising nasal wing hypoplasia/aplasia, oligodontia/hypodontia of permanent teeth, hearing impairment, cognitive impairment, short stature, scalp defects, hypothyroidism, microcephaly, heart defects, imperforate anus and genital malformations[1-3]. The estimated birth prevalence of JBS in Europe is approximately 1:250 000 (ref.[2]).

JBS is caused by pathogenic variants in the *UBR1* gene encoding an E3 ubiquitin (Ub) ligase of N-degron pathways (formerly "N-end rule pathways") that are evolutionarily conserved ubiquitin-dependent proteolytic systems[2,4,5]. In N-degron pathways, E3 Ub ligases (also called N-recognins) recognize proteins containing specific N-terminal degradation signals (termed N-degrons), conjugate a Ub, and thus mark them for subsequent proteasome degradation[4]. More specifically, UBR1 is one of at least four E3 Ub ligases responsible for the Arg/N-degron pathway[6]. UBR1 recognizes two types of degrons: type-1 (basic) N-degrons and type-2 (bulky hydrophobic) N-degrons[7]. Although the role of UBR1 in protein degradation has been well characterized, the impact of impaired protein degradation on the development of abnormalities in JBS patients remains to be elucidated.

The *UBR1* gene contains 47 exons that encode a large protein with a total length of 1749 amino acids. It contains several domains, including two zinc finger domains[8], the UBR box involved in recognizing type-1 N-degrons[9], and the RING-H2 domain involved in E3 ubiquitin ligase activity[10]. In addition to these domains, UBR1 contains the N-domain (also known as the ClpS domain) required for recognition of type-2 N-degrons[9], and the C-terminal UBR-specific autoinhibitory domain thought to be involved in the regulation of UBR1 activity[11]. Moreover, new key structural elements were recently identified and characterized in the yeast UBR1 and thus may also be present in the mammalian UBR1 (ref.[12]).

To date, the largest database of *UBR1* variants (http://databases.lovd.nl/shared/variants/UBR1) contains 75 likely pathogenic/pathogenic variants identified in 78 individuals with JBS. A review of the clinical and molecular data of 60 patients with JBS in 2014 showed significant genotype-phenotype correlations[1]. Pathogenic variants that create premature stop codons such as nonsense, frameshift, and splice site variants, leading to a complete lack of UBR1 protein, were associated with the full JBS phenotype. In contrast, splice site variants leading to in-

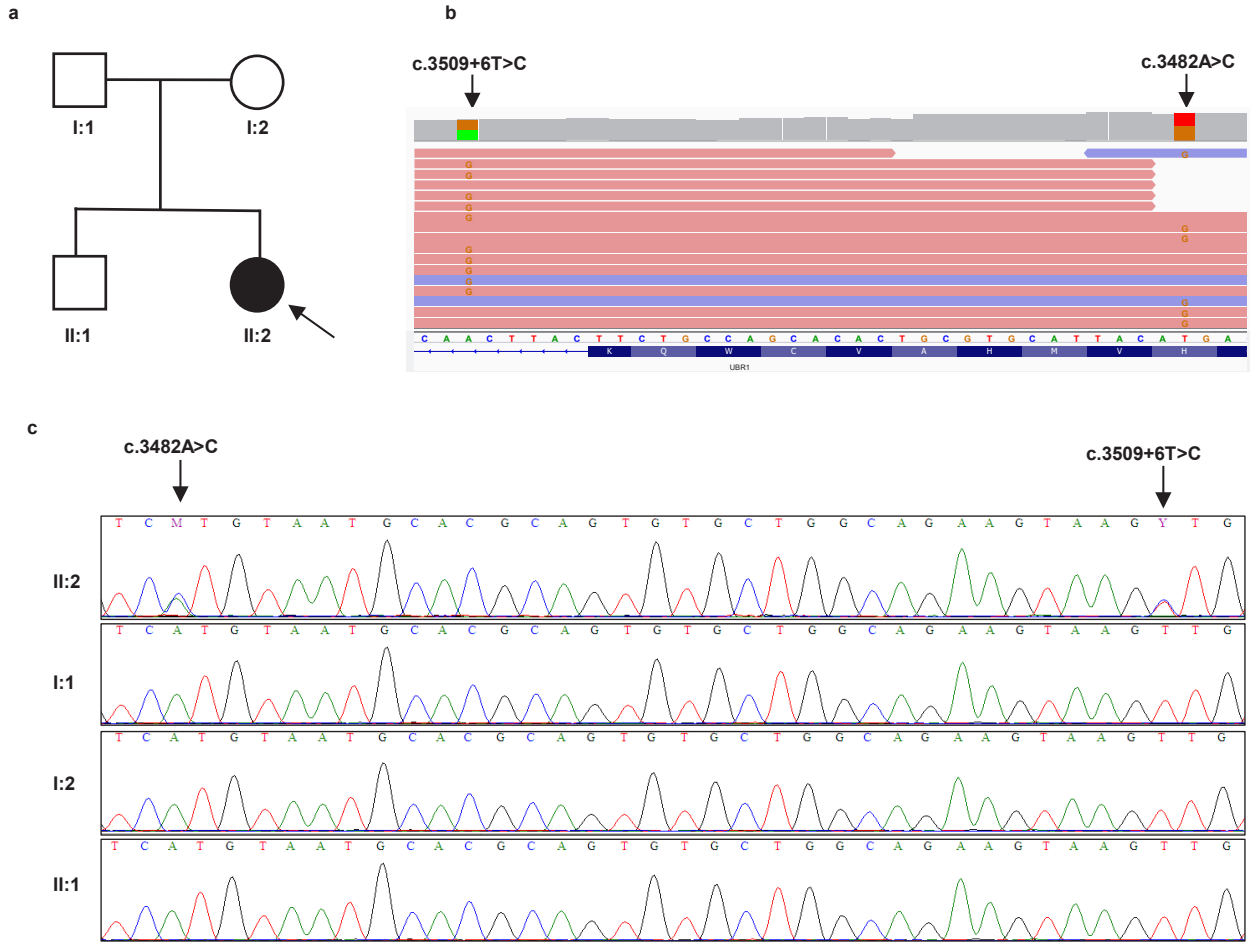


Fig. 1. Identification of heterozygous *UBR1* variants in the family.
a. Pedigree of the family. **b.** IGV image of the CES data in the proband showing the *trans* position of the variants. Note that the orientation of *UBR1* is the opposite orientation in the genome. **c.** Sanger sequencing showing both variants in the proband and none in the parents and brother.

frame deletions, small in-frame deletions, and missense variants correlated with milder physical abnormalities and normal intelligence. The residual protein function of missense *UBR1* variants was demonstrated by functional analysis using yeast as a model organism[13]. No prominent mutation hotspots were identified, and the majority of the reported variants were nonrecurring[1].

Although JBS is usually diagnosed/suspected based on rare early onset of exocrine pancreatic insufficiency and accompanying characteristic anomalies, the diagnosis is confirmed only by molecular testing of *UBR1*, the only diagnostic tool that provides a definitive diagnosis[1]. Molecular testing of *UBR1* was conventionally performed using Sanger sequencing of all 47 *UBR1* exons and flanking intronic regions[1,14].

Nowadays, it seems more promising to use next-generation sequencing with analysis based on the patient phenotype. In addition to possible benefits for the proband, molecular confirmation of JBS is highly valuable for parents planning future pregnancies. The generation of *de novo* variants in the human genome is extremely rare, with approximately 60–100 new variants per individual[15]. Thus, the probability that both parents are carriers is significantly higher than the occurrence of at least one *de novo* variant in the proband. In that case, parents with identified variants can be offered preimplantation diagnostics and prenatal testing during genetic counselling[14]. Only molecular testing enables early diagnosis because prenatal ultrasound scanning may not detect any physical abnormalities[1].

We report a rare case of a proband with clinically suspected JBS in which molecular testing confirmed the diagnosis by identification and characterization of two *de novo* pathogenic variants in *UBR1*.

MATERIALS AND METHODS

Ethics approval and consent to participate

Ethics approval was obtained from the Ethics Committee of the University Hospital and the Faculty of Medicine, Charles University in Pilsen. Written informed consent was obtained from the family members. All methods were carried out in accordance with relevant guidelines and regulations.

Patient clinical phenotype

The proband, 22 years old at her last follow-up, was referred for genetic counseling at 4 weeks of age. She was born to healthy, unrelated parents of Czech origin (Fig. 1a). At birth, she presented with a scalp defect in the parieto-occipital region, hypoplasia of the nasal wings, and pancreatic insufficiency. Cytogenetic analysis revealed no chromosomal aberration. Due to the typical clinical picture, the diagnosis of Johanson-Blizzard syndrome was suspected. Subsequently, she developed hearing loss, oligodontia, and short stature. At the last follow-up, she studied computer graphics at a special school for individuals with hearing impairment. A cochlear implant is planned to correct the hearing loss. She has an older, healthy brother.

Genomic DNA extraction, total RNA extraction, and cDNA synthesis

Genomic DNA was isolated from the peripheral blood of the proband and her family members using the Gentra Puregene Blood Kit (Qiagen, Hilden, Germany) according to a standard protocol. RNA was isolated from the peripheral blood of the proband using Tempus™ Spin RNA Isolation Kit (Thermo Fisher Scientific, Waltham, MA) and reverse-transcribed using High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's instructions.

Clinical exome sequencing and data analysis

Clinical exome sequencing and data analysis were performed on the proband's sample. Clinical exome library preparation was performed using SOPHiA Clinical Exome Solution v2 (SOPHiA Genetics, Lausanne, Switzerland) and high-throughput sequencing was generated on a MiSeq (Illumina, San Diego, CA, USA) in the 300 bp paired-end mode according to the manufacturer's instructions. For data analysis, sequence reads were aligned to the UCSC human reference genome (GRCh37/hg19 assembly). Our in-house pipelines include FastQC v0.11.8, Bowtie2 v.2.3.5 (ref.[16]), Picard v2.21.6, SAMtools 1.10 (ref.[17]), Freebayes v1.3.1, and bedtools v2.29.2 (ref.[18]). The Variant Call Format file was uploaded to the Franklin Analysis platform (https://franklin.genoox.com) for variant annotation, classification, and filtration. Variants were classified according to ACMG guidelines[19] in line with the ClinGen SVI general recommendations for using ACMG/AMP criteria.

To identify causal variants, we considered only pathogenic, likely pathogenic, and leaning pathogenic variants of unknown significance in the *UBR1* gene with minor allele frequency <0.05 both in the general population and in-house database. Variants also have to be with >10 coverage and >100 quality. In order to determine the pathogenicity of potential causative variants, missense prediction tools (Polyphen-2 (ref.[20]), AlphaMissense (ref.[21]), MutationTaster2021 (ref.[22]), Mutpred2 (ref.[23]), DANN (ref.[24]), SIFT (ref.[25]), FATHMM (ref.[26])) and splice site prediction tools (SpliceAI (ref.[27]), Pangolin (ref.[28]), dbscSNV Ada (ref.[29]), and dbscSNV RF (ref.[29])) were used. The HGMD (ref.[30]), ClinVar (ref.[31]), and LOVD (ref.[32]) databases were also used to evaluate the pathogenicity of the variant. Candidate variants were manually checked using the Integrative Genomics Viewer (IGV) (ref.[33]).

Polymerase chain reaction (PCR) and Sanger sequencing

Variants were confirmed by PCR and bidirectional Sanger sequencing of the genomic DNA of proband and family members. PCR and Sanger sequencing were performed as previously described[34]. Forward 5′– gtagcgcgacggccagtTGGACATTCCTAAGACAGTTTGTG–3′ and reverse 5′–cagggcgcagcgatgacTTATGGAGAGCAAGCA-

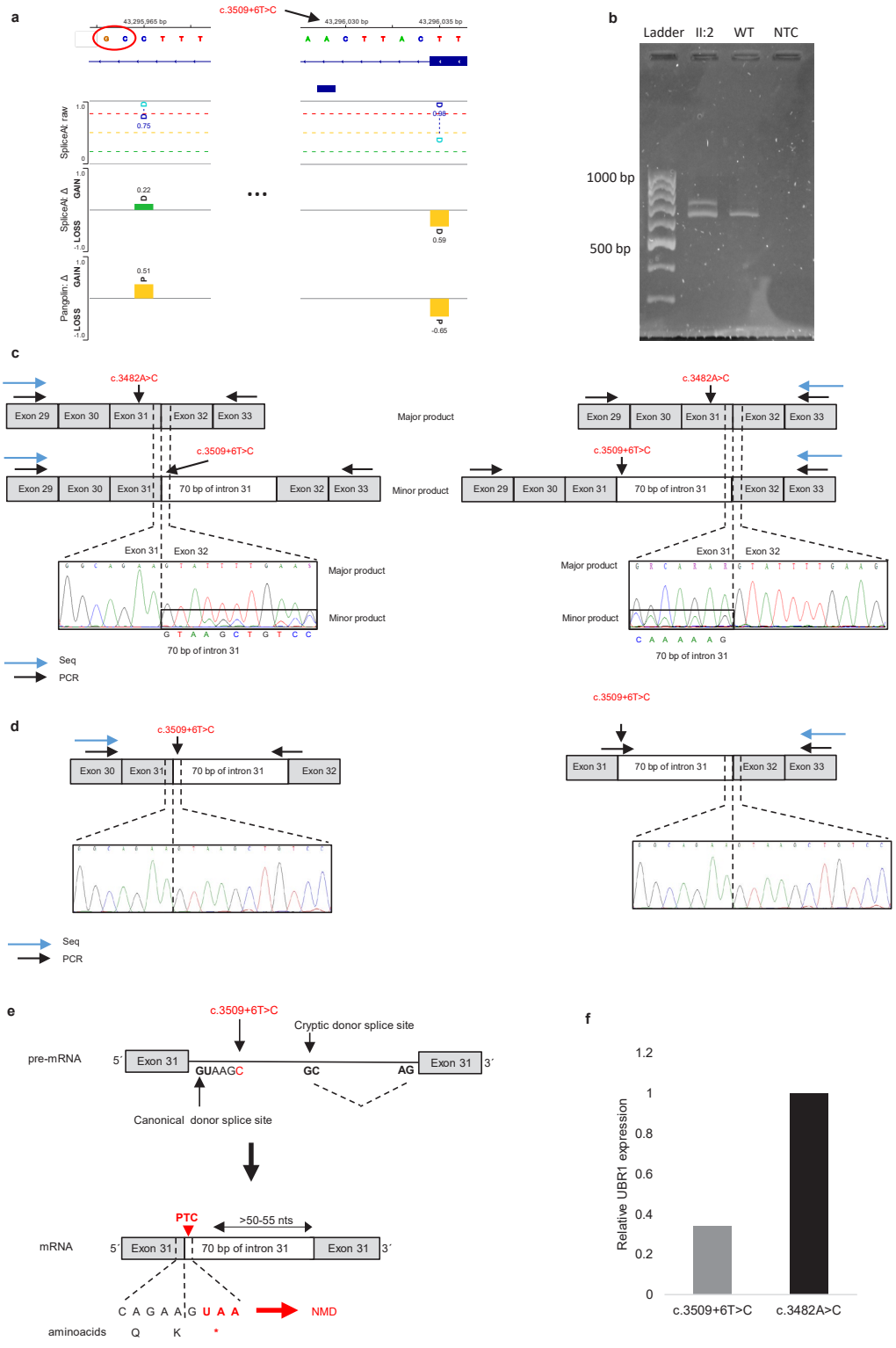


Fig. 2. Characterization of the c.3509+6T>C variant in the *UBR1* gene.
a. Visualization of the SpliceAI and Pangolin results indicating the activation of the cryptic donor splice site GC. Note that the orientation of *UBR1* is the opposite orientation in the genome. **b.** Agarose gel electrophoresis of cDNA products showing an aberrantly spliced longer PCR product in the proband (II.2) and not in the healthy control (WT). **c.** Sanger sequencing of cDNA products in the proband showing a weak signal of the 70 bp intronic sequence after exon 31 in forward-direction sequencing and before exon 32 in reverse-direction sequencing. In the scheme, PCR primers (PCR) for amplification and sequencing primer (Seq) are indicated by arrows. **d.** Bidirectional Sanger sequencing of aberrantly spliced transcript. In the scheme, PCR primers (PCR) for amplification and sequencing primer (Seq) are indicated by arrows. **e.** Illustration of variant-induced cryptic splice site activation leading to partial intron retention, frameshift, and premature termination codon (PTC) located > 50 to 55 nucleotides upstream of the 3' most splice-generated exon-exon junction triggering NMD. **f.** qPCR data showing the relative expression level of aberrantly spliced transcript bearing the c.3509+6T>C variant compared to wild-type transcript bearing the c.3482A>C variant in the proband.

CAAAGGA-3′ primers containing a universal sequencing tail (forward: 5′-gtagcgcgacggccagt, reverse: 5′-cagggcgcagcgatgac-3′) were used. The impact of the variant on RNA splicing was evaluated by PCR and Sanger sequencing of proband's cDNA. PCR and Sanger sequencing of cDNA was performed using the forward 5′gtagcgcgacggccagtGGCTACATCGCCAGAAGATCA-3′ and reverse 5′-cagggcgcagcgatgacCCGTGCCAGGGTCAAAAGT-3′ primers. For Sanger sequencing of aberrantly spliced transcript in the forward direction, the primers specific to aberrantly spliced transcript were as follows: forward 5′-gtagcgcgacggccagtGTGCTGACGTGCATCCTTTG-3′ and reverse 5′-cagggcgcagcgatgacACCTTTTTGTGGCCTTGAGTG-3′. For Sanger sequencing of aberrantly spliced transcript in the reverse direction, the primers specific to aberrantly spliced transcript were as follows 5′-gtagcgcgacggccagtCAGTGTGCTGGCAGAAGTAAG-3′ and 5′-cagggcgcagcgatgacCCGTGCCAGGGTCAAAAGT-3′. Agarose gel was used to verify the amplification products.

PCR and capillary electrophoresis

The analysis of STR loci was performed via AmpFlSTR® Identifiler PCR kit and ABI Prism 3130xl analyzer (Thermo Fisher, Waltham, MA, USA) according to the manufacturer's instructions. The probability of paternity/maternity was counted by Bayes theorem (*a priori* probability 50%) with allelic frequencies for the Czech population[35].

Quantitative real-time PCR

The allele-specific expression level of the *UBR1* gene in the proband's cDNA was detected by quantitative real-time PCR (qPCR) using ABsolute QPCR Mix, no ROX (Thermo Fisher Scientific, Waltham, MA) and 20× EvaGreen Dye (Biotium, Freemont, CA) on Rotor-Gene Q (Qiagen, Germantown, MD) according to the manufacturer's instructions in three technical replicates. Primers specific to aberrantly spliced and wild-type transcripts were used. The wild-type specific primers were as follows: forward 5′-GCACAGGGGAAAACCCATAGA-3′, and reverse 5′-CAAAATACTTCTGCCAGCACACT-3′. The primers specific to the aberrantly spliced transcript were as follows: forward 5′-ACTGCCTTAACCCAGCACAG-3′, and reverse 5′- AAGGACAGCTTACTTCTGCCA-3′. The reference gene *GAPDH* was amplified using forward 5′-GAGAAGGCTGGGGCTCATTT-3′, and reverse 5′-TAAGCAGTTGGTGGTGCAGG-3′ primers. qPCR data were analyzed with double delta Ct analysis and the expression level was normalized to the expression of *GAPDH* reference gene in the same sample. The relative expression level of the aberrantly spliced transcript to the wild-type transcript is shown in the graph.

Protein modelling

The three-dimensional (3D) structures of wild-type and mutated UBR1 with His1161Pro were modelled using the online software, SWISS-MODELL (ref.[36]). The structures of UBR1 were modelled based on the crystal structures of yeast scUBR1 (PDB: 7MEX, 7MEY). The experimental 3D structure of PDB: 7MEX, 7MEY was explored in RCSB PDB. The visual illustration is displayed in Mol* viewer[37].

RESULTS

Identification of two novel candidate variants in *UBR1*

To identify the pathogenic variants, the proband clinically diagnosed with JBS underwent clinical exome sequencing. After the filtering steps, only two candidate heterozygous variants in *UBR1* were left, NM_174916.3:c.3482A>C located in exon 31 and NM_174916.3:c.3509+6T>C located in the donor splice site of exon 31. These variants were not found in public genetic databases or the literature. According to ACMG guidelines, both variants were classified as variants of unknown significance using PP3 and PM2 criteria. Due to their proximity, we visually checked the NGS data in IGV and discovered that they were in the *trans* position (Fig. 1b). Sanger sequencing confirmed both variants in the proband but did not detect either variant in the parents and brother of the proband (Fig. 1c). However, both maternity and paternity were verified by analyzing STRs loci (data not shown).

Effect of the c.3509+6T>C variant on splicing

To assess the possible deleterious effect of the c.3509+6T>C variant on pre-mRNA splicing, multiple *in silico* prediction tools were used. High scores from SpliceAI (0.58), dbscSNV Ada (0.99), and dbscSNV RF (0.84) given by Franklin indicated a high likelihood of disruption of the canonical donor splice site of exon 31. In addition, SpliceAI and Pangolin on the website (https://spliceailookup.broadinstitute.org) also predicted possible activation of cryptic donor splice GC site 70 bp downstream from the canonical donor splice site (Fig. 2a).

The predicted effect of the c.3509+6T>C variant on pre-mRNA splicing was further investigated by RNA analysis. We isolated RNA from the proband's peripheral blood and performed reverse transcription and PCR amplification of exons 29–33 of the proband's cDNA. Agarose gel electrophoresis of the cDNA products in the proband showed an aberrantly spliced longer product in addition to the normally spliced product found in the healthy control (Fig. 2b). Subsequent Sanger sequencing of the cDNA products in the proband revealed a minor wave of sequences at the junction of exon 31 and exon 32 in both directions, corresponding to 70 bp of the intronic sequence (Fig. 2c). It was further confirmed by bidirectional Sanger sequencing of aberrantly spliced product using primers specific to the aberrantly spliced transcript (Fig. 2d). The inclusion of the 70 bp intronic sequence shifts the open reading frame that introduces a premature termination codon leading to degradation via nonsense-mediated decay (NMD) according to 50–55 nt rule[38]. The diagram for aberrant splicing is shown in Fig. 2e. Elimination of aberrantly spliced transcript via the NMD pathway was confirmed by qPCR measuring the *UBR1* mRNA expression (Fig. 2f). qPCR data showed a

decrease in the amount of aberrantly spliced transcript compared to the wild-type transcript by approximately 65%. Based on our functional study showing the deleterious effect of the variant, we classified the variant as pathogenic (PP3, PM2, PVS1).

In silico analysis of the c.3482A>C variant

The c.3482A>C variant caused the substitution of histidine with proline at position 1161. This position corresponds to the first zinc-coordinating histidine (H1) in a highly conserved C3H2C3 zinc finger motif of the RING-H2 domain (Fig. 3a). Multiple prediction tools predicted deleterious effects on the protein. MutPred2 predicted loss/gain of certain properties of five residues around the substitution site (loss of catalytic site at H1161, gain of disulfide linkage at C1159, altered metal binding, gain of strand and loop). In order to visually assess the possible structural effect of the variant, we predicted the 3D structure of both wild-type and mutant UBR1 based on the crystal structures of yeast scUBR1, using SWISS-MODELL. During homology modelling, we identified that His1161 corresponds to His1297 in the scUBR1. By assessing the structure of scUBR1, we identified that His1297 has already been subjected to targeted mutagenesis[10] (Fig. 3b). It was reported that His1297Ala slightly inhibited the degradation of N-end rule substrates. Additionally, the generated 3D structures showed that His1161 coordinates a second zinc ion in the RING-H2 domain (Fig. 3c) and that proline at this position thus should abolish the coordination of the zinc ion (Fig. 3d). Based on our findings, we classified the c.3482A>C variant as a pathogenic variant (PP3, PM2, PM3, PS3, PM1).

DISCUSSION

In this study, molecular testing confirmed JBS in the clinically suspected proband by the identification of two heterozygous *UBR1* variants in the *trans* position. Neither variant was found in the parents nor the brother, indicating their *de novo* origin. The detection of two germline *de novo* variants in a single gene in a patient has been reported only in several publications[39-46]. In all cases, both variants were identified on a single allele in close proximity, probably due to a single mutation event[39]. Here, we identified two *de novo* variants within/around exon 31 on different alleles. We found no report showing the detection of two causal *de novo* variants in a single gene on different chromosomes in a single patient.

Since we consider any event by which two closely lo-

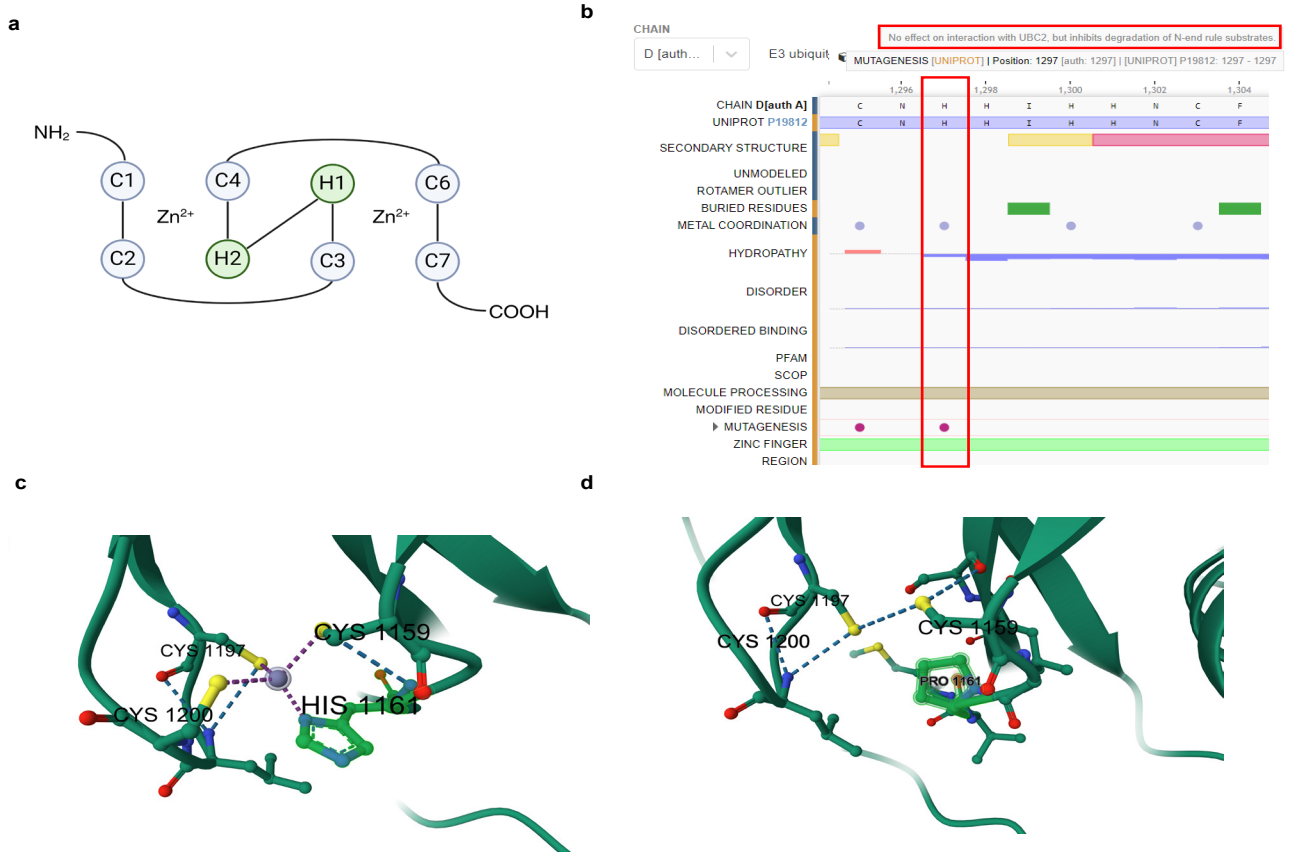


Fig. 3. Characterization of the c.3482A>C variant.
a. Schematic localization of zinc-coordinating cysteines (C) and histidines (H) in the C3H2C3 motif of the RING-H2 domain. **b.** Sequence Annotations Viewer for PDB 7MEX showing the results of performed mutagenesis at position 1297 derived from UniProt. **c.** 3D structural model of wild-type UBR1 zoomed at the position of the second zinc ion in the RING-H2 domain. **d.** 3D structural model of mutant UBR1 zoomed at the position of the His1161Pro variant.

cated variants arising from a single mutation event would be transferred to different alleles highly improbable, we assume that these variants arise independently by two mutation events and that their proximity is due to coincidence. Considering statistics, they probably occurred as paternal and maternal "one-off" variants during sperm and oocyte production[47]. Since it has been assumed to arise in a single sperm and oocyte, recurrence risk is negligible in other children[47]. With a much lower probability, there is a possibility of parental germline mosaicism or early post-zygotic mutation events in the proband[47]. As in most clinical genetic laboratories, these possibilities were not tested due to a lack of biological samples.

The identification of two novel candidate *UBR1* variants in *trans* by CES was consistent with the phenotype and mode of inheritance. They were initially classified as variants of unknown significance, so we decided to confirm their pathogenicity. SpliceAI and newer Pangolin splicing prediction models were visited on the website for more information. Our results demonstrated that these tools also consider non-canonical splice sites, such as GC-AG splice sites, accounting for ~0.8% of human splice sites[48]. The activation of non-canonical splice sites can be easily missed even by widely used splice-site prediction tools such as NNSplice0.9, which considers only GT for the donor site and AG for the acceptor site[49,50]. Since prediction should not be the sole source for assessing pathogenicity[50], we performed functional analysis on the proband's RNA.

Amplification and sequencing of exons 29–33 were performed in case of exon skipping or another non-predicted event. Detection of the aberrantly spliced product confirmed the predicted activation of the cryptic donor splice site but with low signal intensity. Therefore, we further confirmed these findings by aberrant-splicing specific PCR and Sanger sequencing. Extraction from gel is not used in our laboratory. The predicted degradation of the aberrantly spliced transcript by NMD could be observed at the low signal of the 70 bp intronic sequence during the sequencing of exons 29–33. However, Sanger sequencing is not quantitative. We thus validated the degradation by allele-specific qPCR. The amount of aberrantly spliced transcript decreased to 35% of the wild-type transcript, which is consistent with other publications confirming NMD via qPCR (ref.[51,52]). To date, no other splice-site variant-induced activation of the non-canonical GC splice site has been reported in a patient with JBS.

Molecular characterization of the c.3509+6T>C variant also provides additional evidence of the pathogenicity of the c.3482A>C (p.His1161Pro) variant (PM3 criterion). In addition, during protein modelling we identified a source[10] with a functional study showing deleterious effects of Cys→Ser or His→Ala mutations in the C3H2C3 motif in the RING-H2 domain (PS3 criterion). Proline, which has specific physico-chemical properties, should have similar or more adverse effects at this position than alanine. There was no record in the UniProt database of previously performed mutagenesis when searching for the protein UBR1_HUMAN (Q8IWV7), only for UBR1_YEAST(P19812). We performed homology modelling because the RING-H2 domain in UBR1 has not been experimentally characterized. The predicted structure of human UBR1 can be obtained from Alphahold (AlphaFold Model – AF-Q8IWV7-F1) but without ligands (zinc ions). Our 3D models of UBR1 generated by the SWISS-MODEL homology-modelling pipeline illustrated the deleterious impact of the variant on the coordination of the zinc ions. Therefore, we applied the PM1 criterion. Our findings correlate with those of a recent review[53] showing that mutation of any conserved Cys or His involved in zinc coordination should affect the ubiquitination activity of E3 Ub ligases since coordination of two zinc ions in the RING domain is required for E3 Ub ligase activity. Additionally, the crucial role of the first zinc-coordinating histidine in the RING-H2 domain has also been demonstrated in another E3 Ub ligase[54]. In conclusion, our data show that protein modelling, which is not routinely performed, can provide valuable evidence in clinical diagnostic laboratories, as suggested in recent review[55]. To date, only three pathogenic variants have been identified within the highly conserved RING-H2 domain: two nonsense variants (p.Glu1110*, p.Trp1168*) and one missense variant (p.Gln1102Glu).

The presence of one missense variant and one splicing variant in our proband, who exhibited intellectual abilities in the normal range, is consistent with the results of a genotype-phenotype study[1]. In this study, more than half of the patients with at least one "non-truncation" variant had intellectual abilities within the normal range. In contrast, 100% of patients with biallelic truncating variants exhibited some degree of cognitive impairment.

CONCLUSION

To our knowledge, we report the first molecular confirmation of JBS in a proband of Czech origin and the first identification of two *de novo* causal variants in *trans*. We also expanded the mutation spectrum of JBS-causing variants by identification and characterization of the splice site variant c.3509+6T>C inducing removal of a shorter non-canonical GC-AG intron and the missense variant c.3482A>C altering a highly conserved zinc-coordinating histidine in the zinc-stabilized domain RING-H2. Additionally, our data highlight the limitations of several prediction tools in predicting non-canonical splicing, the importance of functional studies, and the benefits of structural modelling.

Acknowledgements: The authors thank the families for their collaboration. The authors acknowledge the National Center for Medical Genomics research infrastructure (LM2023067 funded by MEYS CR) for their support with obtaining scientific data presented in this paper. The study was also supported by the project of the Faculty of Medicine in Pilsen SVV-2024-260654.

Author contributions: LS: literature search, manuscript writing, study design, data collection, data interpretation; TZ: data collection, data interpretation; PK: manuscript

writing, data interpretation; TV: data collection, data interpretation; IS: manuscript writing, genetic counseling, final approval.

Conflict of interest statement: The authors state that there are no conflicts of interest regarding the publication of this article.

REFERENCES

1. Sukalo M, Fiedler A, Guzmán C, Spranger S, Addor M-C, Mcheik JN, Oltra Benavent M, Cobben JM, Gillis LA, Shealy AG, Deshpande C, Bozorgmehr B, Everman DB, Stattin E-L, Liebelt J, Keller K-M, Bertola DR, van Karnebeek CDM, Bergmann C, Liu Z, Düker G, Rezaei N, Alkuraya FS, Oğur G, Alrajoudi A, Venegas-Vega CA, Verbeek NE, Richmond EJ, Kirbiyik Ö, Ranganath P, Singh A, Godbole K, Ali FAM, Alves C, Mayerle J, Lerch MM, Witt H, Zenker M. Mutations in the Human UBR1 Gene and the Associated Phenotypic Spectrum. Hum Mutat 2014;35(5):521-31.
2. Zenker M, Mayerle J, Lerch MM, Tagariello A, Zerres K, Durie PR, Beier M, Hülskamp G, Guzman C, Rehder H, Beemer FA, Hamel B, Vanlieferinghen P, Gershoni-Baruch R, Vieira MW, Dumic M, Auslender R, Gil-da-Silva-Lopes VL, Steinlicht S, Rauh M, Shalev SA, Thiel C, Winterpacht A, Kwon YT, Varshavsky A, Reis A. Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). Nat Genet 2005;37(12):1345-50.
3. Johanson A, Blizzard R. A syndrome of congenital aplasia of the alae nasi, deafness, hypothyroidism, dwarfism, absent permanent teeth, and malabsorption. J Pediatr 1971;79(6):982-7.
4. Varshavsky A. N-degron and C-degron pathways of protein degradation. Proc Natl Acad Sci 2019;116(2):358-66.
5. Bartel B, Wünning I, Varshavsky A. The recognition component of the N-end rule pathway. EMBO J 1990;9(10):3179-89.
6. Tasaki T, Mulder LCF, Iwamatsu A, Lee MJ, Davydov I V., Varshavsky A, Muesing M, Kwon YT. A Family of Mammalian E3 Ubiquitin Ligases That Contain the UBR Box Motif and Recognize N-Degrons. Mol Cell Biol 2005;25(16):7120-36.
7. Varshavsky A. The N-end rule: functions, mysteries, uses. Proc Natl Acad Sci 1996;93(22):12142-9.
8. Kwon YT, Reiss Y, Fried VA, Hershko A, Yoon JK, Gonda DK, Sangan P, Copeland NG, Jenkins NA, Varshavsky A. The mouse and human genes encoding the recognition component of the N-end rule pathway. Proc Natl Acad Sci 1998;95(14):7898-903.
9. Tasaki T, Zakrzewska A, Dudgeon DD, Jiang Y, Lazo JS, Kwon YT. The Substrate Recognition Domains of the N-end Rule Pathway. J Biol Chem 2009;284(3):1884-95.
10. Xie Y. The E2-E3 interaction in the N-end rule pathway: the RING-H2 finger of E3 is required for the synthesis of multiubiquitin chain. EMBO J 1999;18(23):6832-44.
11. Du F, Navarro-Garcia F, Xia Z, Tasaki T, Varshavsky A. Pairs of dipeptides synergistically activate the binding of substrate by ubiquitin ligase through dissociation of its autoinhibitory domain. Proc Natl Acad Sci 2002;99(22):14110-15.
12. Pan M, Zheng Q, Wang T, Liang L, Mao J, Zuo C, Ding R, Ai H, Xie Y, Si D, Yu Y, Liu L, Zhao M. Structural insights into Ubr1-mediated N-degron polyubiquitination. Nature 2021;600(7888):334-8.
13. Hwang C-S, Sukalo M, Batygin O, Addor M-C, Brunner H, Aytes AP, Mayerle J, Song HK, Varshavsky A, Zenker M. Ubiquitin Ligases of the N-End Rule Pathway: Assessment of Mutations in UBR1 That Cause the Johanson-Blizzard Syndrome. PLoS One 2011;6(9):e24925.
14. Sukalo M, Mayerle J, Zenker M. Clinical utility gene card for: Johanson–Blizzard syndrome. Eur J Hum Genet 2014;22(1):152.
15. Gonzaga-Jauregui C, Lupski J. Genomics of rare diseases : understanding rare disease genetics through genomic approaches. Amsterdam: Academic Press; 2021.
16. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9(4):357-9.
17. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25(16):2078-9.
18. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26(6):841-2.
19. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. 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):405-24.
20. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods 2010;7(4):248-9.
21. Cheng J, Novati G, Pan J, Bycroft C, Žemgulytė A, Applebaum T, Pritzel A, Wong LH, Zielinski M, Sargeant T, Schneider RG, Senior AW, Jumper J, Hassabis D, Kohli P, Avsec Ž. Accurate proteome-wide missense variant effect prediction with AlphaMissense. Science 2023;381(6664):eadg7492. doi: 10.1126/science.adg7492
22. Steinhaus R, Proft S, Schuelke M, Cooper DN, Schwarz JM, Seelow D. MutationTaster2021. Nucleic Acids Res 2021;49(W1):W446-W451.
23. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H-J, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD, Radivojac P. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat Commun 2020;11(1):5918.
24. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics 2015;31(5):761-3.
25. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4(7):1073-81.
26. Gough J, Karplus K, Hughey R, Chothia C. Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. J Mol Biol 2001;313(4):903-19.
27. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, Kosmicki JA, Arbelaez J, Cui W, Schwartz GB, Chow ED, Kanterakis E, Gao H, Kia A, Batzoglou S, Sanders SJ, Farh KK-H. Predicting Splicing from Primary Sequence with Deep Learning. Cell 2019;176(3):535-548.e24.
28. Zeng T, Li YI. Predicting RNA splicing from DNA sequence using Pangolin. Genome Biol 2022;23(1):103.
29. Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. Nucleic Acids Res 2014;42(22):13534-44.
30. Stenson PD, Mort M, Ball E V., Chapman M, Evans K, Azevedo L, Hayden M, Heywood S, Millar DS, Phillips AD, Cooper DN. The Human Gene Mutation Database (HGMD®): optimizing its use in a clinical diagnostic or research setting. Hum Genet 2020;139(10):1197-207.
31. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res 2014;42(D1):D980-D985.
32. Fokkema IFAC, Taschner PEM, Schaafsma GCP, Celli J, Laros JFJ, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 2011;32(5):557-63.
33. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 2013;14(2):178-92.
34. Strych L, Černá M, Hejnalová M, Zavoral T, Komrsková P, Tejcová J, Bitar I, Sládková E, Sýkora J, Šubrt I. Targeted long-read sequencing identified a causal structural variant in X-linked nephrogenic diabetes insipidus. BMC Med Genomics 2024;17(1):29.
35. Šimková H, Faltus V, Marvan R, Pexa T, Stenzl V, Brouček J, Hořínek A, Mazura I, Zvárová J. Allele frequency data for 17 short tandem repeats in a Czech population sample. Forensic Sci Int Genet 2009;4(1):e15–e17.
36. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46(W1):W296-W303.
37. Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V, Velankar S, Burley SK, Koča J, Rose AS. Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res 2021;49(W1):W431-W437.
38. Nagy E, Maquat LE. A rule for termination-codon position within

intron-containing genes: when nonsense affects RNA abundance. Trends Biochem Sci 1998;23(6):198-9.
39. Zhu W, Li J, Chen S, Zhang J, Vetrini F, Braxton A, Eng CM, Yang Y, Xia F, Keller KL, Okinaka-Hu L, Lee C, Holder JL, Bi W. Two de novo novel mutations in one SHANK3 allele in a patient with autism and moderate intellectual disability. Am J Med Genet Part A 2018;176(4):973-9.
40. Wang M, Kishnani P, Decker-Phillips M, Kahler SG, Chen YT, Godfrey M. Double mutant fibrillin-1 (FBN1) allele in a patient with neonatal Marfan syndrome. J Med Genet 1996;33(9):760-3.
41. Tessitore A, Sinisi AA, Pasquali D, Cardone M, Vitale D, Bellastella A, Colantuoni V. A Novel Case of Multiple Endocrine Neoplasia Type 2A Associated with Two de Novo Mutations of the RETProtooncogene*. J Clin Endocrinol Metab 1999;84(10):3522-7.
42. Salipante SJ, Benson KF, Luty J, Hadavi V, Kariminejad R, Kariminejad MH, Rezaei N, Horwitz MS. Double de novo mutations ofELA2 in cyclic and severe congenital neutropenia. Hum Mutat 2007;28(9):874-81.
43. Mongan NP, Jääskeläinen J, Green K, Schwabe JW, Shimura N, Dattani M, Hughes IA. Two de Novo Mutations in the AR Gene Cause the Complete Androgen Insensitivity Syndrome in a Pair of Monozygotic Twins. J Clin Endocrinol Metab 2002;87(3):1057-61.
44. Debeer P, Huysmans C, Van de Ven WJM, Fryns J, Devriendt K. Carpal and tarsal synostoses and transverse reduction defects of the toes in two brothers heterozygous for a double de novo NOGGIN mutation. Am J Med Genet Part A 2005;134A(3):318-20.
45. Lundén L, Boxhammer S, Carlsson G, Ellström K, Nordenskjöld M, Lagerstedt-Robinson K, Fadeel B. Double de novo mutations of ELANE (ELA2) in a patient with severe congenital neutropenia requiring high-dose G-CSF therapy. Br J Haematol 2009;147(4):587-90.
46. Stella A, Lastella P, Viggiano L, Bagnulo R, Resta N. Clinical presentation and genetic analyses of neurofibromatosis type 1 in independent patients with monoallelic double de novo closely spaced mutations in the NF1 gene. Hum Mutat 2022;43(10):1354-60.
47. Bernkopf M, Abdullah UB, Bush SJ, Wood KA, Ghaffari S, Giannoulatou E, Koelling N, Maher GJ, Thibaut LM, Williams J, Blair EM, Kelly FB, Bloss A, Burkitt-Wright E, Canham N, Deng AT, Dixit A, Eason J, Elmslie F, Gardham A, Hay E, Holder M, Homfray T, Hurst JA, Johnson D, Jones WD, Kini U, Kivuva E, Kumar A, Lees MM, Leitch HG, Morton JE V., Németh AH, Ramachandrappa S, Saunders K, Shears DJ, Side L, Splitt M, Stewart A, Stewart H, Suri M, Clouston P, Davies RW, Wilkie AOM, Goriely A. Personalized recurrence risk assessment following the birth of a child with a pathogenic de novo mutation. Nat Commun 2023;14(1):853.
48. Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R. Comprehensive splice-site analysis using comparative genomics. Nucleic Acids Res 2006;34(14):3955–67.
49. REESE MG, EECKMAN FH, KULP D, HAUSSLER D. Improved Splice Site Detection in Genie. J Comput Biol 1997;4(3):311-23.
50. Ghosh R, Harrison SM, Rehm HL, Plon SE, Biesecker LG. Updated recommendation for the benign stand-alone ACMG/AMP criterion. Hum Mutat 2018;39(11):1525-30.
51. Pertesi M, Vallée M, Wei X, Revuelta M V., Galia P, Demangel D, Oliver J, Foll M, Chen S, Perrial E, Garderet L, Corre J, Leleu X, Boyle EM, Decaux O, Rodon P, Kolb B, Slama B, Mineur P, Voog E, Le Bris C, Fontan J, Maigre M, Beaumont M, Azais I, Sobol H, Vignon M, Royer B, Perrot A, Fuzibet J-G, Dorvaux V, Anglaret B, Cony-Makhoul P, Berthou C, Desquesnes F, Pegourie B, Leyvraz S, Mosser L, Frenkiel N, Augeul-Meunier K, Leduc I, Leyronnas C, Voillat L, Casassus P, Mathiot C, Cheron N, Paubelle E, Moreau P, Bignon Y, Joly B, Bourquard P, Caillot D, Naman H, Rigaudeau S, Marit G, Macro M, Lambrecht I, Cliquennois M, Vincent L, Helias P, Avet-Loiseau H, Moreno V, Reis RM, Varkonyi J, Kruszewski M, Vangsted AJ, Jurczyszyn A, Zaucha JM, Sainz J, Krawczyk-Kulis M, Wątek M, Pelosini M, Iskierka-Jażdżewska E, Grząśko N, Martinez-Lopez J, Jerez A, Campa D, Buda G, Lesueur F, Dudziński M, García-Sanz R, Nagler A, Rymko M, Jamroziak K, Butrym A, Canzian F, Obazee O, Nilsson B, Klein RJ, Lipkin SM, McKay JD, Dumontet C. Exome sequencing identifies germline variants in DIS3 in familial multiple myeloma. Leukemia 2019;33(9):2324-30.
52. Khan K, Mehmood S, Liu C, Siddiqui M, Ahmad A, Faiz BY, Chioza BA, Baple EA, Ullah MI, Akram Z, Satti HS, Khan R, Harlalka GV, Jameel M, Akram T, Baig SM, Crosby AH, Hassan MJ, Zhang F, Davis EE, Khan TN. A recurrent rare intronic variant in CAPN3 alters mRNA splicing and causes autosomal recessive limb-girdle muscular dystrophy-1 in three Pakistani pedigrees. Am J Med Genet A 2022;188(2):498-508. doi: 10.1002/ajmg.a.62545
53. Garcia-Barcena C, Osinalde N, Ramirez J, Mayor U. How to Inactivate Human Ubiquitin E3 Ligases by Mutation. Front Cell Dev Biol 2020;8:39. doi: 10.3389/fcell.2020.00039
54. Chen A, Wu K, Fuchs SY, Tan P, Gomez C, Pan ZQ. The Conserved RING-H2 Finger of ROC1 Is Required for Ubiquitin Ligation. J Biol Chem 2000;275(20):15432-9.
55. Caswell RC, Gunning AC, Owens MM, Ellard S, Wright CF. Assessing the clinical utility of protein structural analysis in genomic variant classification: experiences from a diagnostic laboratory. Genome Med 2022;14(1):77.