

Functional characterization of a noncanonical intronic PADI6 variant associated with recurrent early embryonic arrest

Ying Liang

Changsha Hospital for Maternal & Child Health Care Affiliated to Hunan Normal University

Bang Liu

Changsha Hospital for Maternal & Child Health Care Affiliated to Hunan Normal University

Yi Fan

Hunan Normal University

Jia-Qi Sun

Second Xiangya Hospital of Central South University

Na Jiang

Second Xiangya Hospital of Central South University

Meng-mei Lv

Changsha Hospital for Maternal & Child Health Care Affiliated to Hunan Normal University

Hua-Lin Huang

Second Xiangya Hospital of Central South University

Dai Zhou

zhoudai@hunnu.edu.cn

Changsha Hospital for Maternal & Child Health Care Affiliated to Hunan Normal University

Research Article

Keywords: Female infertility, PADI6, deep intronic variant, aberrant splicing, early embryonic arrest

Posted Date: September 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10272764/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Early embryonic arrest is an important cause of assisted reproductive technology failure and may involve maternal-effect gene defects. *PADI6*, a component of the subcortical maternal complex, is required for early embryonic development, but the role of noncanonical intronic *PADI6* variants remains insufficiently characterized.

Methods

We studied a woman with recurrent embryo arrest at the 4- to 8-cell stage across three IVF cycles. Whole-exome and Sanger sequencing were performed in the patient and her parents. The candidate variant was assessed using population databases, in silico splice prediction, two minigene reporter systems, western blotting, and immunofluorescence.

Results

A homozygous deep intronic *PADI6* variant, NM_207421.4:c.116 + 43C > T, was identified in the patient; both parents were heterozygous carriers, consistent with autosomal recessive inheritance. The variant was absent from population databases and was predicted by SpliceAI, Fruitfly, and RDDC to activate a cryptic donor splice site. In HEK293T and HeLa cells, two independent minigene assays showed that the mutant construct generated an aberrant transcript containing a 41-bp intronic retention. This insertion was predicted to cause a frameshift and premature termination, resulting in p.Cys40Ter. Western blotting showed expression of a truncated *PADI6*-GFP fusion protein of ~ 30 kDa, whereas the wild-type construct produced a ~ 105-kDa protein. Immunofluorescence suggested altered subcellular localization of the truncated protein.

Conclusion

These findings support the likely pathogenic effect of a noncanonical deep intronic *PADI6* variant in recurrent cleavage-stage embryonic arrest. This study expands the variant spectrum of *PADI6*-associated female infertility and highlights the value of evaluating splice-altering intronic variants in patients with unexplained recurrent embryonic arrest.

1. Introduction

Early embryonic arrest is a significant cause of female infertility and a common factor leading to the failure of assisted reproductive technology (ART) cycles. During *in vitro* fertilization (IVF) and intracytoplasmic sperm injection (ICSI) cycles, approximately 30% to 60% of fertilized oocytes

experience developmental arrest at early stages [1]. These arrested embryos fail to progress to the blastocyst stage and are ultimately discarded. For patients with recurrent ART failure, elucidating the underlying etiology is critical for informing subsequent treatment decisions and genetic counseling.

Evidence suggests that the developmental fate of an embryo is often determined prior to zygotic genome activation (ZGA) [2]. In humans, early embryonic stages rely heavily on maternal transcripts and proteins, with the major wave of ZGA typically occurring at the 4- to 8-cell stage. The subcortical maternal complex (SCMC) is a critical multiprotein structure encoded by maternal-effect genes, playing an essential role in oocyte maturation and early embryo development [3]. Peptidylarginine deiminase 6 (PADI6), a core component of the SCMC, is involved in regulating key processes such as meiosis, translational control, and epigenetic reprogramming [4].

Recent studies have linked PADI6 variants to early embryonic arrest in humans. Since the initial report of human PADI6 variants in 2016, several pathogenic variants have been identified [5–8]. Current evidence indicates that PADI6 variants may lead to embryonic arrest by impeding the ZGA process. The spectrum of PADI6-associated genetic variants continues to expand, and identifying novel variants is of substantial value for optimizing genetic diagnostic workflows.

In the present study, we investigated a patient with recurrent cleavage-stage embryonic arrest to explore the underlying genetic cause. Through genetic analysis, we identified a homozygous noncanonical intronic variant in *PADI6*, c.116 + 43C > T. We then evaluated its potential effect on pre-mRNA splicing using in silico prediction and minigene assays, followed by protein expression and subcellular localization analyses. By integrating clinical, genetic, transcript-level, and protein-level evidence, this study provides supportive evidence for the potential pathogenicity of this splice-altering *PADI6* variant and expands the current understanding of *PADI6*-related female infertility.

2. Materials and Methods

2.1. Study Participants and Clinical Data Collection

A female patient presenting with recurrent early embryonic arrest, along with her parents, was enrolled in this study. Clinical data, including general information, medical history, family history, baseline endocrine profiles, karyotype analysis, semen analysis of the spouse, and ART outcomes, were retrospectively collected. Specific ART parameters such as the ovarian stimulation protocol, gonadotropin dosage, number of oocytes retrieved, number of metaphase II (MII) oocytes, normal fertilization rate, embryonic cleavage, stage of developmental arrest, and blastocyst formation were comprehensively recorded. This study was approved by the Biomedical Research Ethics Committee of Changsha Hospital for Maternal & Child Health Care (Approval No. 2025779). Written informed consent was obtained from all participating family members for genetic testing, functional analysis, and the publication of associated clinical data.

2.2. Assisted Reproductive Technology (ART) Procedures and Embryo Culture

The patient underwent three standard in vitro fertilization (IVF) treatments. The ovarian stimulation protocols utilized included two cycles of the progestin-primed ovarian stimulation (PPOS) protocol and one cycle of the antagonist protocol. Specific medications administered for controlled ovarian stimulation included dydrogesterone (Duphaston; Solvay Pharmaceuticals, Weesp, Netherlands), human menopausal gonadotropin (hMG; Northern Pharmaceutical Co., Ltd., Yantai, China), purified urinary follicle-stimulating hormone (Urofollitropin, Lishenbao; Livzon Pharmaceutical Group Inc., Zhuhai, China), and ganirelix acetate (Orgalutran; Organon, Haarlem, Germany). Upon the dominant follicles reaching the retrieval criteria, human chorionic gonadotropin (hCG; Fengyuan Pharmaceutical Co., Ltd., Maanshan, China) was administered to induce final oocyte maturation. Oocyte retrieval was performed transvaginally under ultrasound guidance 36 to 38 hours following the hCG trigger.

Post-retrieval, oocyte maturation was evaluated, and the number of MII oocytes was documented. Fertilization was achieved via conventional IVF. Pronuclear formation was assessed 16–18 hours post-insemination, with the presence of two pronuclei (2PN) and two polar bodies serving as the standard for normal fertilization. Embryos were cultured in commercial sequential culture media (Vitrolife, Gothenburg, Sweden) within a tri-gas benchtop incubator under specific conditions (37°C, 5% O₂, and 6% CO₂). Embryonic cleavage, cell number, fragmentation rate, developmental arrest, and blastocyst formation were morphologically evaluated on days 2, 3, 5, and/or 6. Embryo grading was conducted according to the routine criteria of our center or the Alpha Scientists in Reproductive Medicine and ESHRE Istanbul consensus (for cleavage-stage embryos) and the Gardner (for blastocysts) criteria.

2.3. Genomic DNA Extraction and Genetic Sequencing

Peripheral venous blood samples from the patient and her parents were collected in EDTA anticoagulant tubes, and genomic DNA was extracted according to standard procedures. Whole-exome sequencing (WES) and subsequent Sanger sequencing for validation of the candidate variant were outsourced to Tsingke Biotechnology Co., Ltd. (Beijing, China).

Briefly, qualified genomic DNA was used for library construction and exome enrichment according to the service provider's protocol. Raw sequencing reads were filtered to remove adapter sequences and low-quality reads, and clean reads were aligned to the human reference genome (GRCh37/hg19) using a BWA/GATK-based workflow. Single-nucleotide variants and small insertions/deletions were called, quality-filtered, and annotated using population and disease-related databases, including dbSNP, the 1000 Genomes Project, gnomAD, ClinVar, OMIM, and relevant literature. Candidate variants were prioritized according to allele frequency, predicted functional consequence, inheritance pattern, gene-disease relevance, and consistency with the patient's phenotype. Noncanonical intronic candidates were further evaluated using splice-prediction tools and functional assays when appropriate. The candidate variant was confirmed by Sanger sequencing in the proband and her parents.

2.4. Runs of Homozygosity (ROH) Analysis

Runs of homozygosity (ROH) analysis was performed using the single nucleotide polymorphism (SNP) data derived from whole-exome sequencing to evaluate genomic homozygosity and potential parental

consanguinity. High-quality autosomal biallelic variants were extracted from the patient's master variant call format (VCF) file after applying strict quality filtering criteria, which included a genotype quality score of 20 or higher, a sequencing read depth of 10 or higher, and a call rate exceeding 95%. ROH tracks across the autosomes were identified using PLINK (version 1.9) via a standard sliding-window approach.

To account for exome capture density variations and potential sequencing errors, the computational parameters for defining an ROH block were specified as follows: a minimum scanning window size of 50 SNPs, a minimum threshold length of 1 Mb for any individual homozygous segment, a maximum of 1 heterozygous call allowed per window, and a maximum of 5 missing genotype calls permitted within a single window. The cumulative total length of autosomal ROH segments was calculated by aggregating all continuous homozygous blocks that met or exceeded the 1 Mb length threshold. The individual genomic coefficient of inbreeding (designated as F_{ROH}) was subsequently calculated as the ratio of the total sum of identified ROH lengths to the total length of the human autosomal genome covered by the exome platform.

2.5. *In Silico* Splicing Prediction and Pathogenicity Assessment

The potential impact of the *PADI6* c.116 + 43C > T variant on pre-mRNA splicing was comprehensively evaluated using three independent *in silico* prediction algorithms based on distinct computational strategies: SpliceAI (a deep-learning-based algorithm), the Fruitfly Splice Predictor (BDGP, a consensus matrix-based algorithm), and the RDDC splicing tool (Tsinghua University platform). For SpliceAI analysis, scores for Acceptor Loss, Donor Loss, Acceptor Gain, and Donor Gain, along with their relative positional changes, were recorded. The Fruitfly tool was utilized to evaluate the potential activation of novel cryptic splice sites or the disruption of native ones. The RDDC platform was employed to further simulate the specific abnormal splice patterns and predict their downstream effects on the reading frame.

Because the c.116 + 43C > T variant is located within intron 1 and outside the canonical donor and acceptor dinucleotides, conventional missense-centric pathogenicity prediction tools (e.g., SIFT and PolyPhen-2) were not applied. Variant interpretation was performed according to the ACMG/AMP guidelines [24]. The evidence considered for this variant included absence from population databases (PM2_supporting), consistency between the patient's phenotype and *PADI6*-associated autosomal recessive early embryonic arrest (PP4), recessive segregation in the family with both parents being heterozygous carriers (PP1_supporting), concordant *in silico* splicing predictions supporting cryptic donor activation (PP3), and functional minigene evidence showing aberrant splicing with a frameshift and predicted premature termination event (PS3_moderate). The strength of functional evidence was applied conservatively because the assays were performed in heterologous cell models rather than patient-derived oocytes or embryos. Based on the combination of one moderate and four supporting criteria, the c.116 + 43C > T variant was classified as likely pathogenic.

2.6. Construction of Minigene and *PADI6* Expression Vectors

To investigate the effect of the *PADI6* c.116 + 43C > T variant on pre-mRNA splicing, wild-type (WT) and mutant mini-gene expression vectors were constructed. Genomic fragments encompassing the target *PADI6* exons and flanking intronic sequences were amplified using genomic DNA from the patient or a healthy control as templates. Vector construction was conducted by Bochu Biotechnology Co., Ltd. (Changsha, China).

Subsequently, based on the aberrant splicing sequences identified from the minigene assay, GFP-tagged *PADI6* WT and mutant expression plasmids were generated. The WT plasmid contained the full-length *PADI6* coding sequence. The mutant plasmid incorporated the abnormal insertion sequence (c.116_117insGTGAGATGCTGGGAGCTCTGCCAGAGGTGGCAGGCAGACAG) to simulate the aberrant transcript resulting from the c.116 + 43C > T variant. The insertion direction, sequence integrity, and targeted variant status of all recombinant plasmids were verified via Sanger sequencing.

2.7. Cell Culture and Transfection

HEK293T and HeLa cells were cultured in high-glucose DMEM (Gibco, MA, USA) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin. Cells were maintained in a humidified incubator at 37°C with 5% CO₂. One day prior to transfection, cells were seeded into 6-well plates to achieve approximately 50% confluence. WT and mutant minigene plasmids were transfected into HEK293T and HeLa cells using Lipofectamine™ 3000 (Thermo Fisher Scientific, MA, USA) according to the manufacturer's protocol. Cells were harvested 48 hours post-transfection for RNA extraction and immunofluorescence assays.

2.8. RNA Extraction, RT-PCR, and Aberrant Splicing Validation

Total RNA was extracted from transfected cells using an RNA extraction kit (Accurate Biology, Changsha, China) and reverse-transcribed into cDNA using a reverse transcription kit (Accurate Biology). RT-PCR was performed utilizing specific primers (**Supplementary Table 1**). PCR products were separated by 3% agarose gel electrophoresis to observe transcriptional band discrepancies between the WT and mutant vectors. Target bands were excised, purified, and subjected to Sanger sequencing to delineate the sequence composition of the aberrant splicing products. Based on the sequencing results, it was determined whether the *PADI6* c.116 + 43C > T variant induced intron retention, exon skipping, or other abnormal splicing events.

2.9. Western Blot Analysis

HEK293T cells transfected with either *PADI6*-WT-GFP or *PADI6*-mut-GFP expression plasmids were harvested 48 hours post-transfection. Total protein was extracted using cell lysis buffer (Beyotime Biotechnology, Shanghai, China), and protein concentrations were determined utilizing the BCA assay

(Beyotime Biotechnology). Equal amounts of protein samples were separated by 10% SDS-PAGE and transferred onto PVDF membranes. Following a 1-hour blockade at room temperature, the membranes were incubated overnight with primary antibodies, including anti-GFP (Abcam, Cambridge, UK) and anti-GAPDH (proMab, Changsha, China).

2.10. Immunofluorescence Analysis

HEK293T cells, seeded on coverslips in 12-well plates, were transfected with PADI6-WT-GFP or PADI6-mut-GFP plasmids. At 48 hours post-transfection, cells were washed with PBS. Due to the presence of the GFP tag on the expression plasmids, GFP fluorescence signals were observed directly. Nuclei were counterstained with DAPI. Following mounting, slides were imaged using a Zeiss LSM 900 confocal fluorescence microscope to assess discrepancies in the subcellular localization between WT and mutant PADI6-GFP fusion proteins. Green fluorescence indicated the GFP-tagged PADI6 protein, whereas blue fluorescence denoted DAPI-stained nuclei.

2.11. Protein Structure Prediction

The three-dimensional structures of the WT PADI6 protein and the p.Cys40Ter truncated protein were predicted utilizing AlphaFold2. The WT protein sequence was retrieved from the UniProt database, while the mutant protein sequence was deduced based on the aberrant splicing results induced by the c.116 + 43C > T variant. PyMOL software was employed for structural visualization and image generation to compare the structural disparities between the WT and truncated proteins.

2.12. Statistical Analysis

As this study primarily focuses on the functional validation of a rare variant, clinical data are presented descriptively. Functional assays were performed in independent experiments, and representative results are shown.

3. Results

3.1. Clinical Characteristics and Assisted Reproductive Outcomes

The patient, a 27-year-old Han Chinese female, presented with a 10-year history of primary infertility without contraception. In April 2016, hysterosalpingography suggested partial bilateral tubal obstruction, after which she underwent laparoscopic tubal recanalization combined with hysteroscopy at an external hospital; however, pregnancy was not achieved, and follicular monitoring was not pursued. In March 2022, a subsequent hysteroscopy revealed a grossly normal uterine cavity, with pathology indicating proliferative endometrium and negative CD138 expression. Semen parameters of her spouse were within the normal reference ranges. The couple subsequently presented to our center for IVF treatment. The patient was otherwise healthy with an unremarkable medical history; notably, her parents were consanguineous, and she had an unmarried, childless brother (Fig. 1A).

Baseline sex hormone evaluation on February 26, 2022, showed levels within normal physiological ranges: follicle-stimulating hormone (FSH) 7.65 IU/L, luteinizing hormone (LH) 5.05 IU/L, estradiol (E2) 24.20 pg/mL, progesterone (P) < 0.50 ng/mL, testosterone (T) 0.10 ng/mL, and prolactin (PRL) 10.77 ng/mL. Peripheral blood karyotype analysis revealed a normal female constitution (46,XX). Semen evaluation of the spouse confirmed parameters consistent with the World Health Organization (WHO) reference standards (volume: 2.8–4.4 mL; concentration: $[102–116] \times 10^6/\text{mL}$; progressive motility: 45%–57%; deformity rate: 92.3%–93.4%). Hysterosalpingography confirmed bilateral tubal patency.

Regarding her ART history, the patient underwent three controlled ovarian stimulation cycles, consisting of two progestin-primed ovarian stimulation (PPOS) cycles and one antagonist cycle. In the first cycle (May 2022), stimulation was initiated with 300 U of human menopausal gonadotropin (hMG), with a total gonadotropin (Gn) dose of 3,300 IU; 9 oocytes were retrieved and 5 normally fertilized 2PN embryos were obtained. However, all embryos failed to progress to blastocysts after 72 hours of in vitro culture, resulting in cycle cancellation due to the lack of transferable embryos. The second cycle (July 2022) utilized highly purified urinary FSH at an initial dose of 300 U, with a total Gn dose of 3,000 IU; 4 oocytes were retrieved, resulting in 2 2PN embryos. These embryos exhibited no further developmental potential at 72 hours and failed to form blastocysts by day 6. The third cycle (August 2022) utilized the antagonist protocol, initiated with 375 U of hMG, with a total Gn dose of 3,000 IU; 15 oocytes were retrieved, from which 9 2PN embryos were obtained. Similar to the previous cycles, embryonic growth ceased around 72 hours post-fertilization, with no normal progression observed up to day 6 (Fig. 1C). Consequently, all embryos were discarded, and no embryos were available for transfer.

3.2. Identification and Multi-Software Splicing Prediction of a Homozygous PADI6 Variant

Following informed consent, WES was performed on peripheral blood DNA from the patient. The analysis identified a homozygous noncanonical intronic variant in PADI6 (NM_207421.4:c.116 + 43C > T). Sanger sequencing showed that both parents were heterozygous carriers of this variant, which was consistent with an autosomal recessive inheritance pattern (Fig. 1B). In addition, runs of homozygosity (ROH) analysis identified extended homozygous segments across the patient's autosomes, totaling 139.47 Mb, consistent with the known parental history of consanguineous marriage (Fig. 1D).

The baseline genetic characteristics and in silico splicing predictions for this PADI6 variant are summarized in Table 1. Population database screening showed that the c.116 + 43C > T variant was not detected in the 1000 Genomes Project or gnomAD database, supporting its rarity. Querying the immediate flanking noncoding region within intron 1 identified several neighboring rare variants, such as c.116 + 41G > A and c.116 + 43C > A, in the general population at very low allele frequencies ranging from 10^{-6} to 10^{-4} (Supplementary Table 2). The absence of the patient's c.116 + 43C > T substitution from these databases, together with its predicted splice effect, supported further functional assessment.

Table 1

Multi-software consensus *in silico* splicing prediction and genetic characteristics of the *PADI6* c.116 + 43C > T variant.

Feature / Computational Tool	Evaluated Parameter / Splicing Event	Score / Value	Predicted Molecular Consequence
Baseline Characteristics			
Genomic Coordinates (GRCh37)	Chr1:17698899 (C > T)	/	Noncanonical intronic variant
Transcript Nomenclature (HGVS)	NM_207421.4:c.116 + 43C > T	/	Intron 1 variant
Minor Allele Frequency (MAF)	gnomAD / 1000 Genomes Project	0.00% (Absent)	Not detected in population databases
Associated Disease (OMIM)	Oocyte/zygote/embryo maturation arrest 16	/	Autosomal recessive female infertility
In Silico Splicing Algorithms			
SpliceAI	Donor Gain (at -2 bp)	0.97	Altered pre-mRNA splicing via cryptic donor activation
	Donor Loss (at -43 bp)	0.34	
	Acceptor Loss (at 160 bp)	0.00	
	Acceptor Gain	0.00	
Fruitfly Splice Predictor (BDGP)	Cryptic donor site activation	0.93	Alternative splice donor site predicted
RDDC Platform (Tsinghua)	Splicing pattern simulation	Positive	41-bp intron 1 retention, translational frameshift, and premature termination (p.Cys40Ter)

Given its noncanonical intronic localization (+ 43 bp within intron 1), conventional missense-centric pathogenicity prediction tools, such as SIFT and PolyPhen-2, were not applicable. To evaluate its potential impact on pre-mRNA processing, three independent computational splicing algorithms were employed. SpliceAI predicted a Donor_gain event with a score of 0.97 (delta position: -2 bp), accompanied by a possible loss of the native donor site (score: 0.34, delta position: -43 bp). The Fruitfly Splice Predictor (BDGP) also suggested activation of a cryptic donor splice site with a score of 0.93. In addition, splicing pattern simulation using the RDDC platform predicted a splice alteration characterized by a 41-bp intronic retention from intron 1, which was expected to introduce a translational frameshift and a downstream premature termination codon (Fig. 2A). Together, these concordant predictions

suggested that the noncanonical c.116 + 43C > T variant may alter PADI6 mRNA splicing, prompting further experimental validation.

3.3. *In Vitro* Transcriptional Evaluation via Minigene Splicing Assays

To investigate the potential impact of the PADI6 c.116 + 43C > T variant on pre-mRNA splicing, we used two independent minigene reporter systems. First, genomic segments encompassing PADI6 exon 1, the variant-containing intron 1, and exon 2 were cloned into the pcDNA3.1 expression vector to establish a local splicing model (Fig. 2B). Following transfection of the wild-type (WT) and mutant vectors into HEK293T and HeLa cells, RT-PCR analysis showed that the WT construct produced a single expected transcript (band a). In contrast, the mutant construct generated an additional larger aberrant transcript (band b) alongside the normal transcript in both cell lines (Fig. 2C). Sanger sequencing of these RT-PCR products showed that, while the WT transcript exhibited normal splicing between exon 1 and exon 2, the aberrant transcript from the mutant construct contained a 41-bp intronic retention immediately downstream of exon 1 (Fig. 2D).

To examine the splicing pattern in an additional multi-exon context, a second minigene assay was performed using the pcMINI-N vector, which incorporates an additional downstream exon (Exon B) (Fig. 2E). Consistent with the initial findings, RT-PCR amplification of the pcMINI-N transcripts showed a single expected band for the WT construct and an additional larger band for the mutant construct across both evaluated cell lines (Fig. 2F). Sanger sequencing of the mutant pcMINI-N transcripts confirmed the same 41-bp intronic retention event (Fig. 2G).

The retained 41-bp sequence was GTGAGATGCTGGGAGCTCTGCCAGAGGTGGCAGGCAGACAG, corresponding to c.116_117ins41bp. This experimental observation was consistent with the computational simulation provided by the RDDC platform as described in Section 3.2, supporting the interpretation that cryptic donor activation was associated with retention of this intronic fragment. This insertion is predicted to introduce a translational frameshift, generating a premature termination codon (PTC) within the retained intronic segment and resulting in a truncated protein variant (p.Cys40Ter).

3.4. Protein Structure and Conservation Predictions

Three-dimensional protein modeling via AlphaFold2 was employed to evaluate the structural consequences of the predicted truncation (Figs. 3A, B). The wild-type PADI6 protein comprises 695 amino acids and contains an N-terminal domain, a C-terminal catalytic domain, and several functional motifs. The p.Cys40Ter alteration is predicted to cause early termination of translation, resulting in a truncated peptide that is expected to lack the C-terminal catalytic domain and multiple major functional regions. Because the C-terminus contains calcium-binding sites and protein-interaction interfaces important for PADI6 activity, this truncation may substantially impair the structural and functional integrity of the protein.

3.5. Protein Expression and Subcellular Localization of Wild-Type and Mutant PADI6

To examine potential protein-level consequences of the aberrant transcript, expression plasmids containing either the wild-type sequence (PADI6-WT) or the mutant sequence with the 41-bp insertion (PADI6-mut), both tagged with GFP, were transfected into HEK293T cells. Western blot analysis showed that the PADI6-mut construct produced a truncated fusion protein migrating at approximately 30 kDa, whereas the PADI6-WT construct produced a full-length fusion protein at approximately 105 kDa (Fig. 3C). These observations support that the 41-bp insertion can produce a truncated protein lacking major catalytic regions in this expression system.

Immunofluorescence analyses were performed to assess the spatial distribution of these proteins. The wild-type PADI6-GFP fusion protein showed predominant cytoplasmic localization. In comparison, the mutant PADI6-GFP fusion protein showed a more diffuse distribution throughout the cell, including the nucleus (Fig. 3D). These findings suggest that the predicted truncation is associated with altered subcellular localization, which may affect the ability of PADI6 to localize to its physiological site of action.

3.6. Review of Reported PADI6 Variants

Several PADI6 variants have been reported in the literature, including coding variants and previously reported splice-site variants distributed across multiple exons or intronic regions. Reported patients with biallelic PADI6 variants generally present with early embryonic arrest, often at the 4- to 8-cell stage. A comparative summary of the variant characteristics identified in this study relative to previously reported PADI6 variants is provided in Table 2.

Table 2
Reported PADI6 variants associated with early embryonic arrest (FS, frameshift)

Gene	Amino Acid Alteration (HGVS)	Nucleotide Alteration (HGVS)	Variant Type	Exon / Region
<i>PADI6</i>	p.Asn598Ser[4]	c.1793A > G	Missense	Exon14
	p.Arg682Gly[4]	c.2045G > A	Missense	Exon16
	p.Pro289Leu[4]	c.866C > T	Missense	Exon8
	p.Pro632Leu[4]	c.1895C > T	Missense	Exon15
	p.Leu375Phefs*13[4]	c.1124dupT	FS ins.	Exon10
	p.Ser508Glnfs*5[4]	c.1521dupC	FS ins.	Exon12
	p.Thr373Pro[4]	c.1117A > C	Missense	Exon10
	p.Arg570Cys[4]	c.1708C > T	Missense	Exon14
	p.Ser278Profs*59[4]	c.831_832del	FS del.	Exon7
	p.Cys163Arg[4]	c.487T > C	Missense	Exon5
	p.Trp475*[4]	c.1425G > A	Nonsense	Exon11
	p.Glu586Lys[4]	c.1756G > A	Missense	Exon14
	p.Ser336Arg[4]	c.1006A > C	Missense	Exon9
	p.Leu237Alafs*24[4]	c.707dupT	FS ins.	Exon7
	Splicing variant [4]	c.294 + 5G > A	Splicing	Intron
	p.Ile416Thr[4]	c.1247T > C	Missense	Exon11
	p.Ser3Glnfs*65[4]	c.5dupT	FS ins.	Exon1
	p.Cys666Leufs*53[4]	c.1996dupT	FS ins.	Exon16
	p.S3Q fs*65[6]	/	FS ins.	Exon1
	p.T187N fs*48[6]	/	FS ins.	/
	p.A220E[6]	/	Missense	/
	p.L237A fs*24[6]	/	FS ins.	/
	c.294 + 5G > A[6]	c.294 + 5G > A	Splicing	Intron
	p.S336R[6]	/	Missense	/
	p.D349R fs*28[6]	/	FS ins.	/
	p.R352C[6]	/	Missense	/

Gene	Amino Acid Alteration (HGVS)	Nucleotide Alteration (HGVS)	Variant Type	Exon / Region
	p.I416T[6]	/	Missense	/
	p.E586K[6]	/	Missense	/
	p.Q589K[6]	/	Missense	/
	p.P632T[6]	/	Missense	/
	p.C666L fs*53[6]	/	FS ins.	/
	p.H211Q[5]	c.633T > A	Missense	Exon6
	p.P433L[6]	/	Missense	/
	p.G540R[5]	c.1618G > A	Missense	Exon13
	p.E670G fs*48[5]	c.2009_2010del	FS del.	Exon16
	p.Gln381*[5]	c.1141C > T	Nonsense	Exon10
	p.Gln324*[5]	c.970C > T	Nonsense	Exon9
	p.S278P fs*59[8]	c.831_832delGT	FS del.	Exon7

4. Discussion

4.1. Identification of a deep intronic *PADI6* variant in a patient with recurrent early embryonic arrest

In this study, we investigated a woman with recurrent early embryonic arrest after apparently normal fertilization across three independent IVF cycles. The embryos repeatedly failed to progress beyond the cleavage stage, mainly around the 4- to 8-cell stage, which coincides with an important developmental window involving maternal-to-zygotic transition and zygotic genome activation [2, 20, 21]. This phenotype is generally consistent with previously reported clinical manifestations associated with biallelic variants in maternal-effect genes, particularly *PADI6* [5–13].

Through genetic analysis, we identified a homozygous deep intronic variant in *PADI6*, NM_207421.4:c.116 + 43C > T. Sanger sequencing showed that both parents were heterozygous carriers, supporting an autosomal recessive inheritance pattern. The patient also showed extended runs of homozygosity, which was compatible with the family history of parental consanguinity. Together with the absence of this variant from population databases and the patient's highly recurrent embryonic arrest phenotype, these findings support a possible causal relationship between the *PADI6* c.116 + 43C > T variant and the observed reproductive phenotype.

The major contribution of this study is the characterization of a noncanonical deep intronic *PADI6* variant with functional evidence. Most *PADI6* variants reported to date are located within coding regions or canonical splice-site regions and include missense, nonsense, frameshift, and canonical splicing variants [5–13]. In contrast, the c.116 + 43C > T variant is located outside the canonical splice donor and acceptor sites. Such variants may be overlooked or given lower priority during routine WES-based variant interpretation. Therefore, our findings may help expand the mutational spectrum of *PADI6*-related early embryonic arrest and highlight the potential importance of noncoding splice-altering variants in reproductive genetic diagnosis.

4.2. Functional evidence supporting aberrant splicing caused by the c.116 + 43C > T variant

The interpretation of deep intronic variants is often challenging because they do not directly alter the coding sequence and may not be captured by conventional missense-focused prediction tools. Their pathogenic potential usually depends on whether they affect pre-mRNA splicing, for example by activating cryptic splice sites, introducing pseudoexons, or altering splicing regulatory elements. Computational splice prediction tools, including deep-learning-based algorithms such as SpliceAI, have improved the ability to identify candidate splice-disrupting variants, but experimental validation remains important for variant interpretation [22, 24]. In the present study, multiple splicing prediction tools suggested that the *PADI6* c.116 + 43C > T variant could activate a cryptic donor splice site within intron 1.

To examine this possibility experimentally, we performed minigene assays using two independent reporter systems. In both the pcDNA3.1-based and pcMINI-N-based minigene models, the mutant construct generated an additional larger transcript in HEK293T and HeLa cells, whereas the wild-type construct mainly produced the expected normally spliced product. Sanger sequencing confirmed that the aberrant transcript contained a 41-bp intronic retention immediately downstream of exon 1. This retained sequence was consistent with the predicted abnormal splicing pattern and was expected to introduce a frameshift followed by a premature termination codon, resulting in p.Cys40Ter.

The use of two minigene systems provides supportive evidence for the splicing effect of this variant. The pcDNA3.1 construct directly tested the local exon 1–intron 1–exon 2 splicing context, whereas the pcMINI-N construct provided an additional multi-exon splicing model. The observation of the same 41-bp retention in both systems and in two cell lines suggests that the c.116 + 43C > T substitution is likely to alter *PADI6* pre-mRNA splicing through cryptic donor activation. Nevertheless, because these assays were performed in heterologous cell systems, further confirmation in patient-derived oocytes or embryos would be valuable if suitable materials become available.

4.3. Potential functional consequences of the predicted *PADI6* truncation

The aberrant transcript generated by the c.116 + 43C > T variant is predicted to produce a severely truncated *PADI6* protein, p.Cys40Ter. Full-length *PADI6* contains 695 amino acids and belongs to the

peptidyl arginine deiminase family, whose members contain conserved structural regions required for normal protein function [14]. By contrast, the predicted truncated product retains only a very short N-terminal fragment and lacks the C-terminal catalytic domain as well as other functional regions. This structural consequence is compatible with a loss-of-function mechanism.

Our protein-level experiments further support this possibility. Western blot analysis showed that the mutant *PADI6*-GFP construct produced a smaller fusion protein of approximately 30 kDa, whereas the wild-type construct produced a full-length fusion protein of approximately 105 kDa. This result indicates that the transcript containing the 41-bp insertion can be translated into a truncated protein in the experimental system used in this study. Although transcripts harboring premature termination codons may be degraded by nonsense-mediated mRNA decay, the detection of the truncated fusion protein suggests that at least part of the mutant transcript may escape complete degradation under the experimental conditions used here [23].

Immunofluorescence analysis also showed a difference in subcellular distribution between wild-type and mutant *PADI6*. Wild-type *PADI6*-GFP was predominantly localized in the cytoplasm, whereas the truncated mutant protein showed a more diffuse distribution throughout the cell, including the nucleoplasm. This altered localization may be relevant because *PADI6* is mainly expressed in oocytes and early embryos and has been implicated in cytoplasmic lattice organization, ribosomal storage, translational regulation, and early embryonic development [3, 14–17]. However, we did not directly examine SCMC assembly or interactions between mutant *PADI6* and other SCMC components. Therefore, although the observed truncation and mislocalization are consistent with impaired *PADI6* function, the exact downstream molecular consequences remain to be clarified.

4.4. Possible link between *PADI6* dysfunction and cleavage-stage embryonic arrest

The recurrent arrest of the patient's embryos at the 4- to 8-cell stage may reflect disruption of processes required during the maternal-to-zygotic transition. Before zygotic genome activation, early embryonic development depends largely on maternal RNAs and proteins accumulated during oogenesis [2, 20, 21]. Maternal-effect gene defects can therefore allow fertilization to occur but impair subsequent cleavage-stage development.

PADI6 is considered an important maternal-effect gene and has been linked to early embryonic development in both human genetic studies and animal models [5–13, 15, 16]. Previous studies suggest that *PADI6* deficiency may not necessarily prevent fertilization but can compromise early embryo development, potentially through effects on cytoplasmic lattice formation, ribosomal storage, translational regulation, and possibly zygotic genome activation [15, 16, 19]. The clinical pattern in the present patient—normal fertilization followed by recurrent cleavage-stage arrest and absence of transferable embryos—is compatible with this biological framework.

Based on the available evidence, we propose that the c.116 + 43C > T variant may contribute to embryonic arrest mainly through loss of functional *PADI6*. The variant alters splicing, introduces a premature termination codon, and is associated with production of a truncated protein with abnormal subcellular localization in vitro. These changes may impair the maternal molecular machinery required for early embryo development. However, this proposed mechanism should be interpreted with caution, as we were unable to directly examine *PADI6* expression, SCMC organization, or ZGA-related transcripts in patient-derived oocytes or embryos.

4.5. Diagnostic implications for noncanonical intronic variants in reproductive genetics

This study also has potential implications for genetic diagnosis in recurrent ART failure. In patients with repeated embryonic arrest despite normal fertilization, especially when the arrest occurs at a reproducible developmental stage, variants in maternal-effect genes should be considered. In addition to *PADI6*, genes such as *TLE6*, *NLRP2*, *NLRP5*, *OOEP*, and *KHDC3L* have been associated with early embryonic developmental failure or related reproductive phenotypes [18, 25–28]. A phenotype-guided genetic testing strategy may therefore be helpful for patients with unexplained recurrent preimplantation developmental failure.

Our findings also illustrate a limitation of exome-centered analysis. Standard WES interpretation usually prioritizes coding variants and canonical splice-site changes, whereas deep intronic or noncanonical splice-altering variants may be missed, filtered out, or classified as variants of uncertain significance without functional validation. Computational splice prediction tools can be useful for identifying candidate splice-disrupting variants, but prediction alone is usually insufficient for confident interpretation [22]. Functional assays, such as minigene analysis, can provide additional evidence for evaluating suspected splice-altering variants under ACMG/AMP-based interpretation frameworks [24].

For clinical practice, this case suggests that noncanonical intronic regions may deserve attention when the phenotype is highly suggestive and no convincing coding variant is identified. In selected patients, reanalysis of intronic flanking regions, targeted sequencing of candidate genes, transcript-level assays, or whole-genome sequencing may improve diagnostic yield. This approach may be particularly relevant in consanguineous families, where homozygous rare variants in recessive maternal-effect genes may be more likely.

4.6. Limitations and future perspectives

Several limitations should be acknowledged. First, this study is based on a single affected patient and her family. Although the phenotype, inheritance pattern, variant rarity, and functional data are mutually supportive, additional unrelated cases with the same or similar *PADI6* deep intronic variants would further strengthen the genotype–phenotype correlation.

Second, the functional assays were performed in heterologous cell models. HEK293T and HeLa cells are useful for validating splice alterations and assessing protein expression or localization, but they cannot

fully reproduce the specific transcriptomic, post-transcriptional, and subcellular environment of human oocytes and early embryos. Therefore, the results should be interpreted as supportive functional evidence rather than a complete replication of the in vivo situation.

Third, because patient-derived oocytes and arrested embryos were not available for molecular analysis, we could not directly assess endogenous *PADI6* transcripts, *PADI6* protein expression, SCMC assembly, or ZGA-related molecular changes in the patient's reproductive cells. In addition, although the mutant construct produced a truncated protein in vitro, the relative contribution of nonsense-mediated mRNA decay versus truncated protein production in patient oocytes remains uncertain [23]. Finally, we did not perform biochemical assays to examine the deiminase activity of the truncated fragment or its interaction with SCMC-associated proteins such as *NLRP5*, *TLE6*, *OOEP*, or *KHDC3L*.

Future studies using patient-derived biosamples, oocyte or embryo transcriptomic analysis, protein-interaction assays, or appropriate model systems may further clarify how deep intronic *PADI6* variants affect early human embryogenesis. Despite these limitations, the present study provides genetic, computational, transcript-level, and protein-level evidence supporting the likely pathogenic effect of the *PADI6* c.116 + 43C > T variant. These findings add to the current understanding of *PADI6*-related female infertility and support the value of integrating splice prediction with experimental validation when interpreting noncanonical intronic variants in reproductive genetics.

5. Conclusion

In summary, we identified a homozygous deep intronic *PADI6* variant, c.116 + 43C > T, in a patient with recurrent early embryonic arrest. Minigene assays indicated that this variant altered pre-mRNA splicing by inducing a 41-bp intronic retention, leading to a predicted premature termination event. Subsequent protein-level assays suggested that the mutant transcript could produce a truncated *PADI6* protein with altered subcellular localization. These findings support the likely pathogenic effect of this noncanonical intronic variant and expand the variant spectrum of *PADI6*-associated female infertility. More broadly, our study suggests that splice-altering intronic variants should be considered during genetic evaluation of patients with unexplained recurrent cleavage-stage embryonic arrest.

Declarations

Competing interests

The authors declare no competing interests.

Funding

This work was funded by the Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0531500, 2024ZD0531503), the National Natural Science Foundation of China (81501248, 82201771), the Natural Science Foundation of Changsha (kq2502312), the Natural Science Foundation of Hunan Province (2024JJ6083), and the Open Project of Key Laboratory of Reproductive Health Diseases Research and Translation (Hainan Medical University), Ministry of Education/The Open Project of Hainan Provincial Key Laboratory for Human Reproductive Medicine and Genetic Research/The Open Project of Hainan Provincial Clinical Research Center for Thalassemia (HNSZLAB202502, HNSZLAB202608), and the Health Research Project of Hunan Provincial Health Commission (Grant No. W20243143).

Author Contribution

YL: Investigation, Methodology, Validation, Visualization, Writing – original draft. BL: Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. YF: Data curation, Investigation, Resources. JQS: Data curation, Investigation, Resources. NJ: Data curation, Investigation, Resources. MML: Formal analysis, Software, Validation. HLH: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. DZ: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. All authors contributed to the article and approved the submitted version.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Hardy K, Stark J, Winston RML (2003) Maintenance of the inner cell mass in human blastocysts from fragmented embryos. *Biol Reprod* 68(4):1165–1169. 10.1095/biolreprod.102.010090
2. Tadros W, Lipshitz HD (2009) The maternal-to-zygotic transition: a play in two acts. *Development* 136(18):3033–3042. 10.1242/dev.033183
3. Bebbere D, Masala L, Albertini DF, Ledda S (2016) The subcortical maternal complex: multiple functions for one biological structure? *J Assist Reprod Genet* 33(11):1431–1438. 10.1007/s10815-016-0778-8
4. Ozturk S (2023) Genetic variants underlying developmental arrests in human preimplantation embryos. *Mol Hum Reprod* 29(7):gaad019. 10.1093/molehr/gaad019
5. Xu Y, Shi Y, Fu J, Yu M, Feng R, Sang Q et al (2016) Mutations in PADI6 cause female infertility characterized by early embryonic arrest. *Am J Hum Genet* 99(3):744–752. 10.1016/j.ajhg.2016.06.024

6. Zheng W, Chen L, Dai J, Dai C, Guo J, Lu C et al (2020) New biallelic mutations in PADI6 cause recurrent preimplantation embryonic arrest characterized by direct cleavage. *J Assist Reprod Genet* 37(1):205–212. 10.1007/s10815-019-01635-2
7. Liu J, Tan Z, He J, Jin T, Han Y, Hu L, Huang S (2021) Two novel mutations in PADI6 and TLE6 genes cause female infertility due to arrest in embryonic development. *J Assist Reprod Genet* 38(6):1551–1559. 10.1007/s10815-021-02194-1
8. Dong J, Fu J, Yan Z, Li L, Qiu Y, Zeng Y et al (2022) Novel biallelic mutations in PADI6 in patients with early embryonic arrest. *J Hum Genet* 67(5):285–293. 10.1038/s10038-021-01004-5
9. Xu Y, Qian Y, Liu Y, Wang Q, Wang R, Zhou Y et al (2022) Novel homozygous PADI6 variants in infertile females with early embryonic arrest. *Front Cell Dev Biology* 10:819667. 10.3389/fcell.2022.819667
10. Wang X, Zhu H, He Y, Li Y (2022) A novel homozygous mutation in the PADI6 gene causes early embryo arrest. *Reproductive Health* 19(1):181. 10.1186/s12978-022-01495-7
11. Zhang T, Liu P, Yao G et al (2023) A complex heterozygous mutation in PADI6 causes early embryo arrest: a case report. *Front Genet* 14:1104085. 10.3389/fgene.2022.1104085
12. Cao G, Zhu X, Lin Y, Fang J, Shen X, Wang S, Kong N (2023) A novel homozygous variant in PADI6 is associated with human cleavage-stage embryonic arrest. *Front Genet* 14:1243230. 10.3389/fgene.2023.1243230
13. Zhou J, Mao R, Gao L, Wang M, Long R, Wang X, Li Z, Jin L, Zhu L (2024) Novel variants in PADI6 genes cause female infertility due to early embryo arrest. *J Assist Reprod Genet* 41(12):3327–3336. 10.1007/s10815-024-03332-1
14. Williams JPC, Walport LJ (2023) PADI6: what we know about the elusive fifth member of the peptidyl arginine deiminase family. *Philosophical Trans Royal Soc B: Biol Sci* 378(1890):20220242. 10.1098/rstb.2022.0242
15. Esposito G, Vitale AM, Leijten FPJ, Strik AM, Koonen-Reemst AMCB, Yurttas P et al (2007) Peptidylarginine deiminase 6 is essential for oocyte cytoskeletal sheet formation and female fertility. *Mol Cell Endocrinol* 273(1–2):25–31. 10.1016/j.mce.2007.05.005
16. Yurttas P, Vitale AM, Fitzhenry RJ, Cohen-Gould L, Wu W, Gossen JA, Coonrod SA (2008) Role for PADI6 and the cytoplasmic lattices in ribosomal storage in oocytes and translational control in the early mouse embryo. *Development* 135(15):2627–2636. 10.1242/dev.016329
17. Kim B, Kan R, Anguish L, Nelson LM, Coonrod SA (2010) Potential role for MATER in cytoplasmic lattice formation in murine oocytes. *PLoS ONE* 5(9):e12587. 10.1371/journal.pone.0012587
18. Mahadevan S, Sathappan V, Utama B, Lorenzo I, Kaskar K, Van den Veyver IB (2017) Maternally expressed NLRP2 links the subcortical maternal complex to fertility, embryogenesis and epigenetic reprogramming. *Sci Rep* 7:44667. 10.1038/srep44667
19. Giaccari C, Cecere F, Argenziano L, Pagano A, Galvao A, Acampora D et al (2024) A maternal-effect Padi6 variant causes nuclear and cytoplasmic abnormalities in oocytes, as well as failure of

- epigenetic reprogramming and zygotic genome activation in embryos. *Genes Dev* 38(3–4):131–150. 10.1101/gad.351238.123
20. Yan L, Yang M, Guo H, Yang L, Wu J, Li R et al (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* 20(9):1131–1139. 10.1038/nsmb.2660
21. Asami M, Lam BYH, Ma MK, Rainbow K, Braun S, VerMilyea MD et al (2022) Human embryonic genome activation initiates at the one-cell stage. *Cell Stem Cell* 29(2):209–216e4. 10.1016/j.stem.2021.11.012
22. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI et al (2019) Predicting splicing from primary sequence with deep learning. *Cell* 176(3):535–548e24. 10.1016/j.cell.2018.12.015
23. Lejeune F (2017) Nonsense-mediated mRNA decay at the crossroads of many cellular pathways. *BMB Rep* 50(4):175–185. 10.5483/BMBRep.2017.50.4.050
24. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J et al (2015) Standards and guidelines for the interpretation of sequence variants. *Genet Sci* 17(5):405–424. 10.1038/gim.2015.30
25. Alazami AM, Awad SM, Coskun S, Al-Hassan S, Hijazi H, Abdulwahab FM et al (2015) TLE6 mutation causes the earliest known human embryonic lethality. *Genome Biol* 16:240. 10.1186/s13059-015-0792-0
26. Mu J, Wang W, Chen B, Wu L, Li B, Mao X et al (2019) Mutations in NLRP2 and NLRP5 cause female infertility characterised by early embryonic arrest. *J Med Genet* 56(7):471–480. 10.1136/jmedgenet-2018-105936
27. Feng R, Sang Q, Kuang Y, Sun X, Yan Z, Zhang S et al (2016) Mutations in TUBB8 and human oocyte meiotic arrest. *N Engl J Med* 374(3):223–232. 10.1056/NEJMoa1510791
28. Tong X, Jin J, Hu Z, Zhang Y, Fan HY, Zhang YL et al (2022) Mutations in OOEP and NLRP5 identified in infertile patients with early embryonic arrest. *Hum Mutat* 43(12):1909–1920. 10.1002/humu.24448

Figures

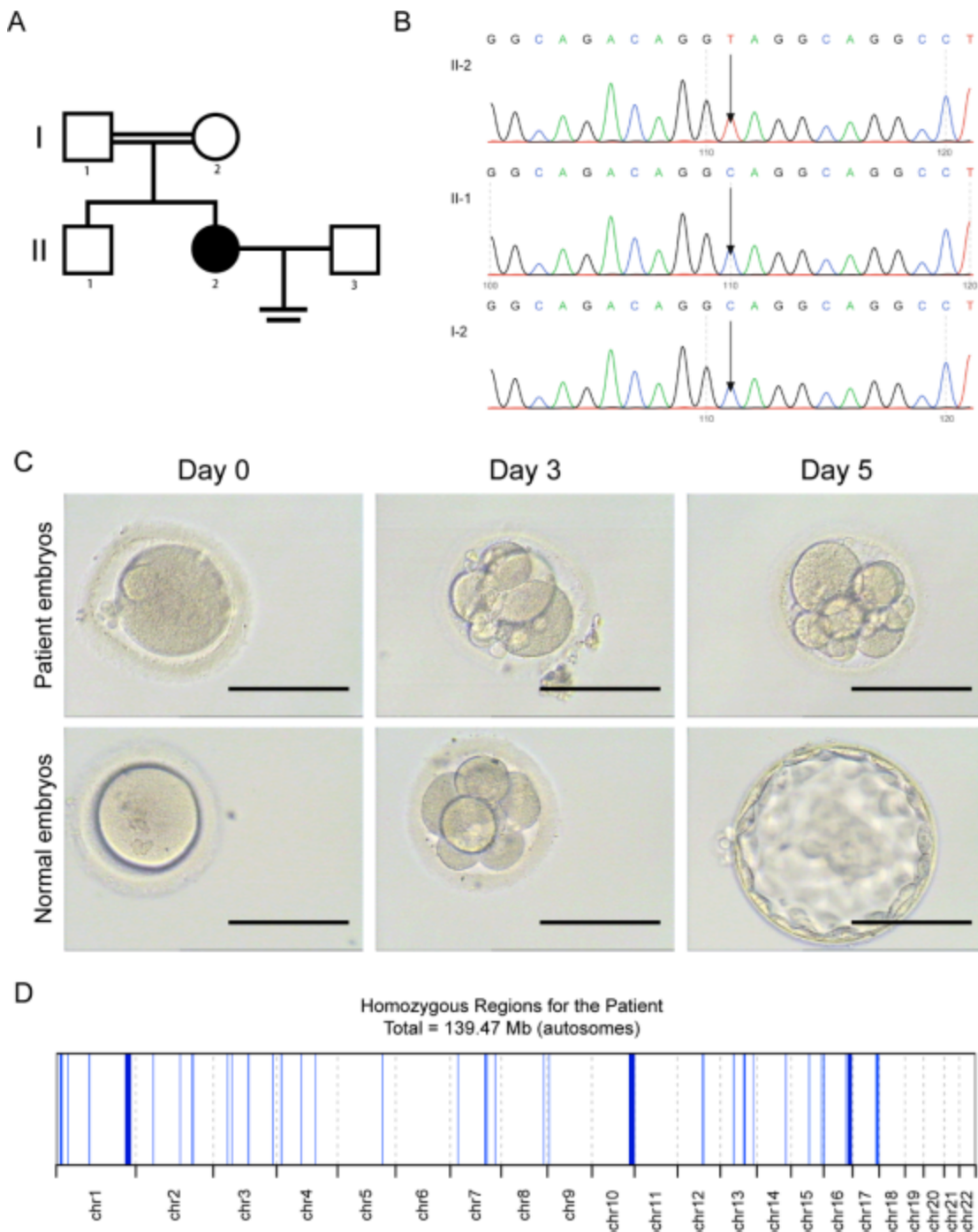


Figure 1

Pedigree analysis, Sanger sequencing validation, embryonic development phenotype, and runs of homozygosity (ROH) analysis. (A) Pedigree of the investigated family. The black arrow indicates the proband; filled symbols represent affected individuals; double horizontal lines indicate parental consanguinity. (B) Sanger sequencing chromatograms showing the PADI6 c.116+43C>T variant. The proband harbors the homozygous variant, whereas both parents are heterozygous carriers, supporting an autosomal recessive mode of inheritance. Vertical arrows indicate the variant site. (C) In vitro developmental presentation of patient-derived embryos compared with control embryos. Patient-derived embryos entered early cleavage stages after fertilization but failed to continue development and showed

arrest at the 4- to 8-cell stage, with no blastocyst formation observed by day 5. Control embryos progressed to form expandable blastocysts. Scale bar: 100 μ m. (D) Autosomal runs of homozygosity (ROH) analysis of the patient. The total length of autosomal homozygous regions is 139.47 Mb, showing extended homozygous blocks across the genome, consistent with the parental consanguineous background.

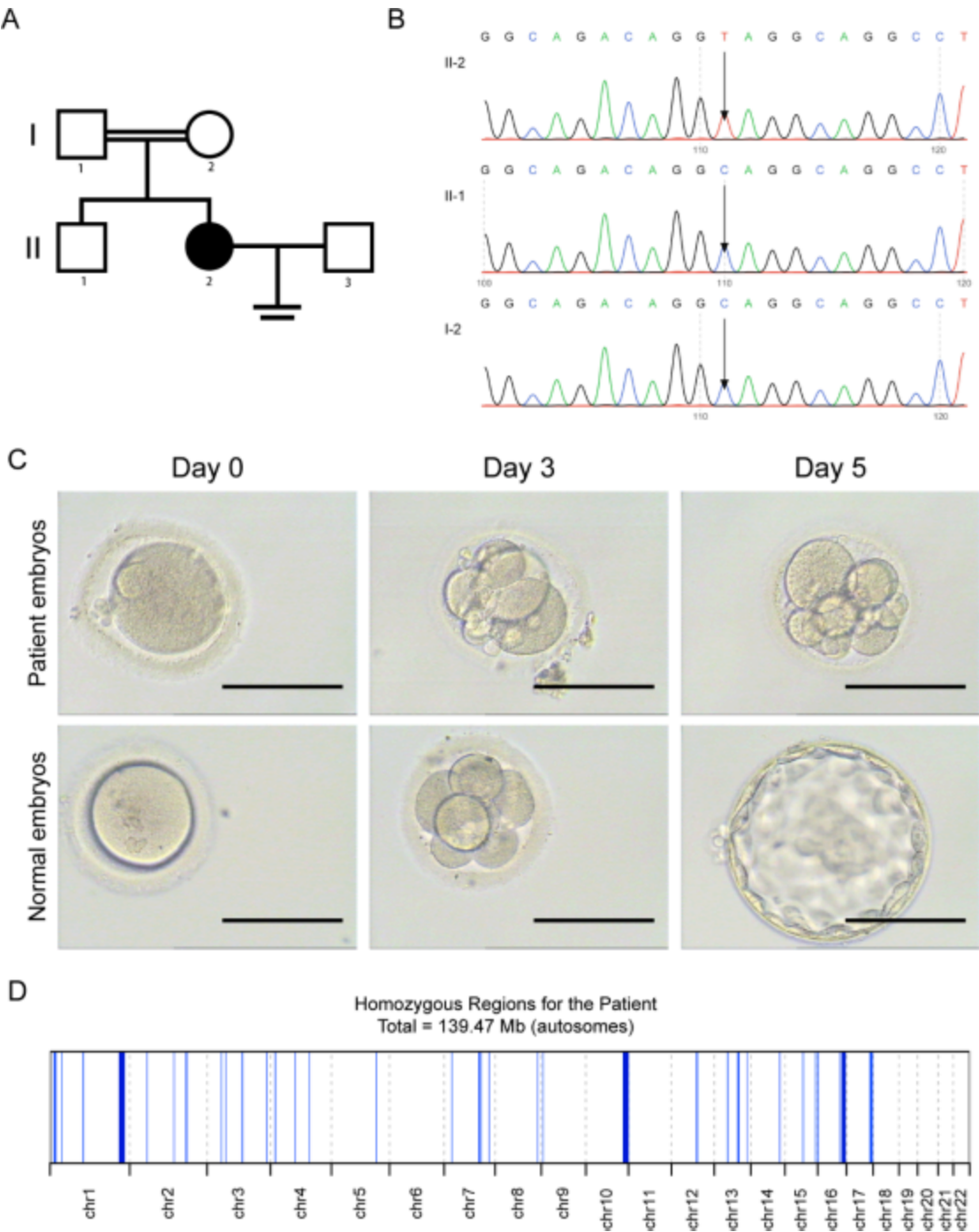


Figure 1

Pedigree analysis, Sanger sequencing validation, embryonic development phenotype, and runs of homozygosity (ROH) analysis. (A) Pedigree of the investigated family. The black arrow indicates the proband; filled symbols represent affected individuals; double horizontal lines indicate parental

consanguinity. (B) Sanger sequencing chromatograms showing the PADI6 c.116+43C>T variant. The proband harbors the homozygous variant, whereas both parents are heterozygous carriers, supporting an autosomal recessive mode of inheritance. Vertical arrows indicate the variant site. (C) In vitro developmental presentation of patient-derived embryos compared with control embryos. Patient-derived embryos entered early cleavage stages after fertilization but failed to continue development and showed arrest at the 4- to 8-cell stage, with no blastocyst formation observed by day 5. Control embryos progressed to form expandable blastocysts. Scale bar: 100 μ m. (D) Autosomal runs of homozygosity (ROH) analysis of the patient. The total length of autosomal homozygous regions is 139.47 Mb, showing extended homozygous blocks across the genome, consistent with the parental consanguineous background.

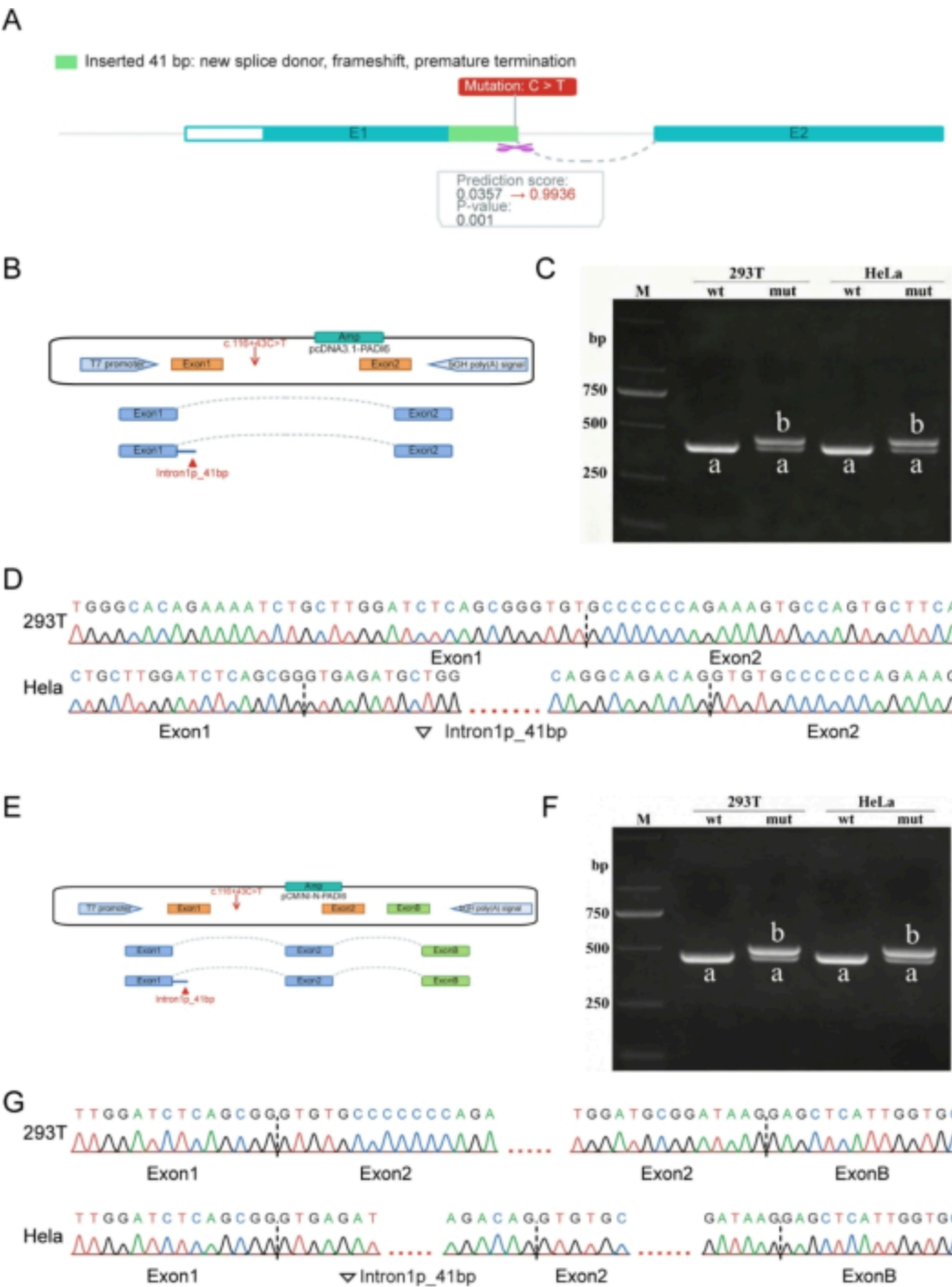


Figure 2

In silico splicing pattern simulation and in vitro minigene characterization of the PADI6 c.116+43C>T variant.

(A) Computational splicing pattern simulation of the c.116+43C>T variant using the RDDC platform. The noncanonical intronic substitution is predicted to activate a cryptic donor splice site (prediction score increasing from 0.0357 to 0.9936, $P = 0.001$), resulting in a 41-bp intronic insertion, a translational frameshift, and downstream premature translation termination. (B) Schematic diagram of the pcDNA3.1-minigene vector construction along with the predicted splicing patterns for wild-type (WT) and mutant constructs. (C) Agarose gel electrophoresis of RT-PCR products following transient transfection of pcDNA3.1-minigene vectors into HEK293T and HeLa cells. Band a represents the normal splicing product; band b denotes the larger aberrant splicing product generated by the mutant construct. (D) Sanger sequencing profiles of the pcDNA3.1 RT-PCR products. Chromatograms showed that the mutant construct generated a transcript containing a 41-bp intronic retention derived from PADI6 intron 1. (E) Schematic diagram of the pcMINI-N-minigene multi-exon vector construction incorporating an additional downstream exon (Exon B) and the predicted splicing patterns for WT and mutant models. (F) Agarose gel electrophoresis of pcMINI-N RT-PCR products in HEK293T and HeLa cells, showing the presence of a larger aberrant splicing band in the mutant groups. (G) Sanger sequencing analysis of pcMINI-N RT-PCR products, confirming the same 41-bp intronic retention (c.116_117ins41bp), which is predicted to introduce a translational frameshift and a premature termination codon (p.Cys40Ter). Abbreviations: WT, wild-type; mut, mutant; M, DNA marker.

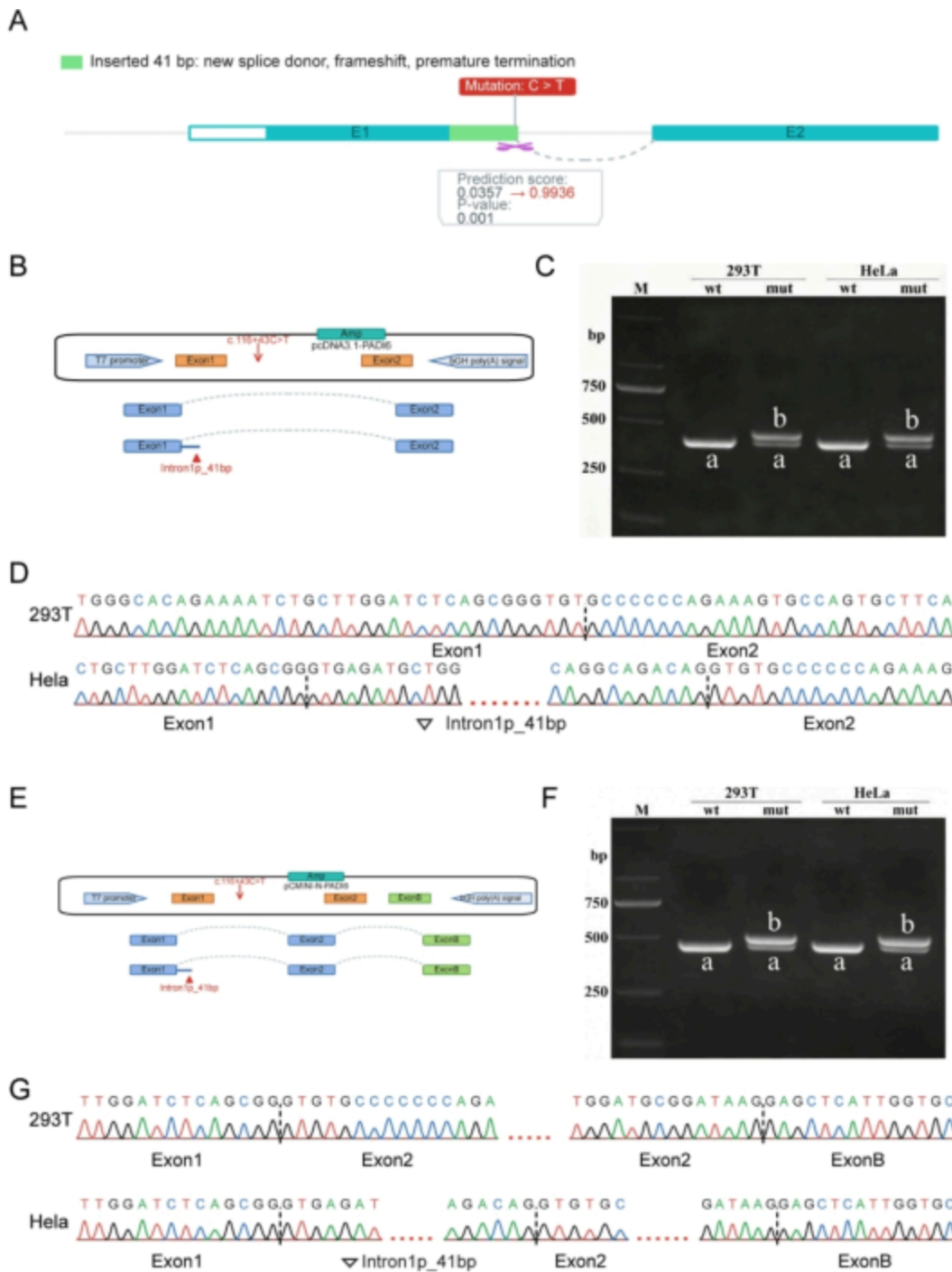


Figure 2

In silico splicing pattern simulation and in vitro minigene characterization of the PADI6 c.116+43C>T variant.

(A) Computational splicing pattern simulation of the c.116+43C>T variant using the RDDC platform. The noncanonical intronic substitution is predicted to activate a cryptic donor splice site (prediction score increasing from 0.0357 to 0.9936, P = 0.001), resulting in a 41-bp intronic insertion, a translational frameshift, and downstream premature translation termination. (B) Schematic diagram of the pcDNA3.1-minigene vector construction along with the predicted splicing patterns for wild-type (WT) and mutant

constructs. (C) Agarose gel electrophoresis of RT-PCR products following transient transfection of pcDNA3.1-minigene vectors into HEK293T and HeLa cells. Band a represents the normal splicing product; band b denotes the larger aberrant splicing product generated by the mutant construct. (D) Sanger sequencing profiles of the pcDNA3.1 RT-PCR products. Chromatograms showed that the mutant construct generated a transcript containing a 41-bp intronic retention derived from PADI6 intron 1. (E) Schematic diagram of the pcMINI-N-minigene multi-exon vector construction incorporating an additional downstream exon (Exon B) and the predicted splicing patterns for WT and mutant models. (F) Agarose gel electrophoresis of pcMINI-N RT-PCR products in HEK293T and HeLa cells, showing the presence of a larger aberrant splicing band in the mutant groups. (G) Sanger sequencing analysis of pcMINI-N RT-PCR products, confirming the same 41-bp intronic retention (c.116_117ins41bp), which is predicted to introduce a translational frameshift and a premature termination codon (p.Cys40Ter). Abbreviations: WT, wild-type; mut, mutant; M, DNA marker.

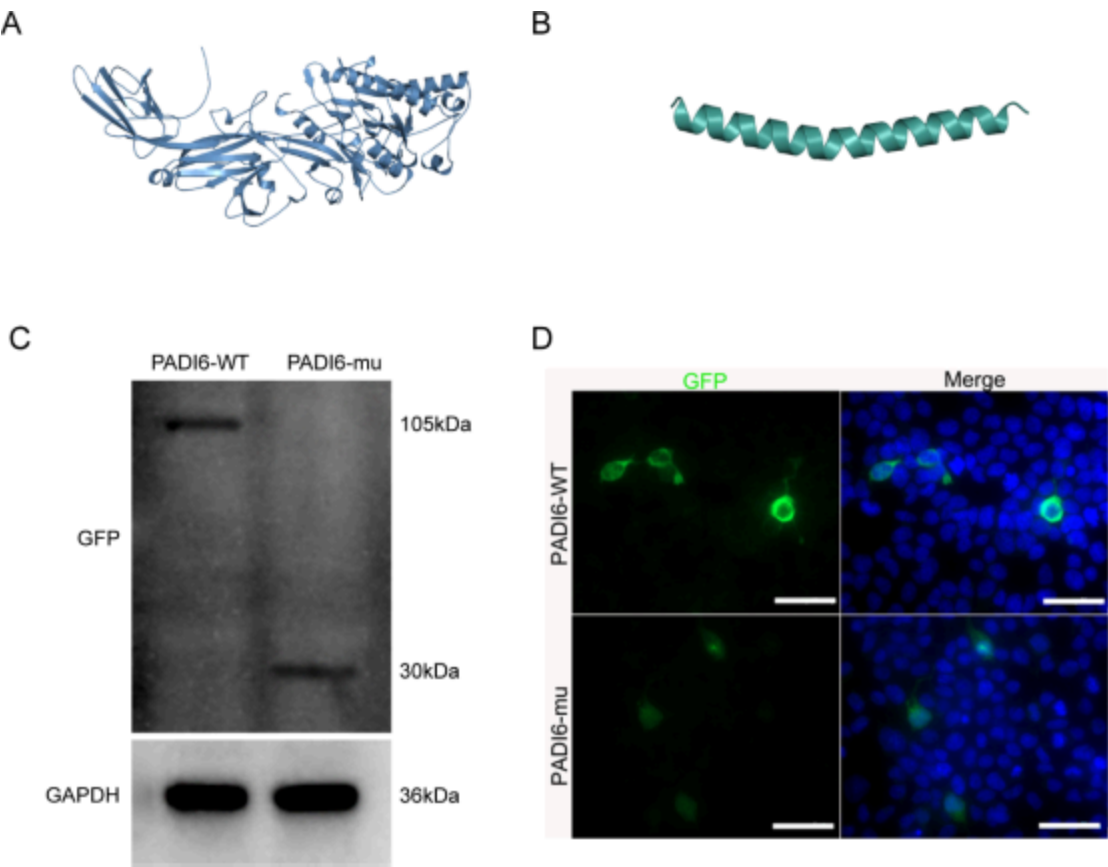


Figure 3

Three-dimensional protein structure predictions, expression analysis, and subcellular localization of wild-type and mutant PADI6.(A) Three-dimensional structure of the wild-type PADI6 protein simulated by AlphaFold2. The full-length wild-type PADI6 contains 695 amino acids. (B) Predicted structural modeling of the p.Cys40Ter truncated protein, which is expected to retain only 39 amino acids at the N-terminus and lack major functional regions, including the catalytic domain. (C) Western blot evaluation of PADI6-WT-GFP and PADI6-mut-GFP fusion protein expression levels in HEK293T cells. The wild-type construct produced a full-length fusion protein migrating at approximately 105 kDa, whereas the mutant construct

produced a smaller truncated band at approximately 30 kDa. GAPDH was used as the internal loading control. (D) Immunofluorescence imaging assessing the spatial distribution of PADI6-WT-GFP and PADI6-mut-GFP fusion proteins in HEK293T cells. Green signals indicate GFP-tagged target proteins, while blue signals indicate DAPI-stained nuclei. Wild-type PADI6 showed predominant cytoplasmic localization, whereas the truncated mutant fusion protein showed a more diffuse distribution in both the cytoplasm and nucleoplasm. Scale bar: 50 μ m. Abbreviations: WT, wild-type; mu, mutant.

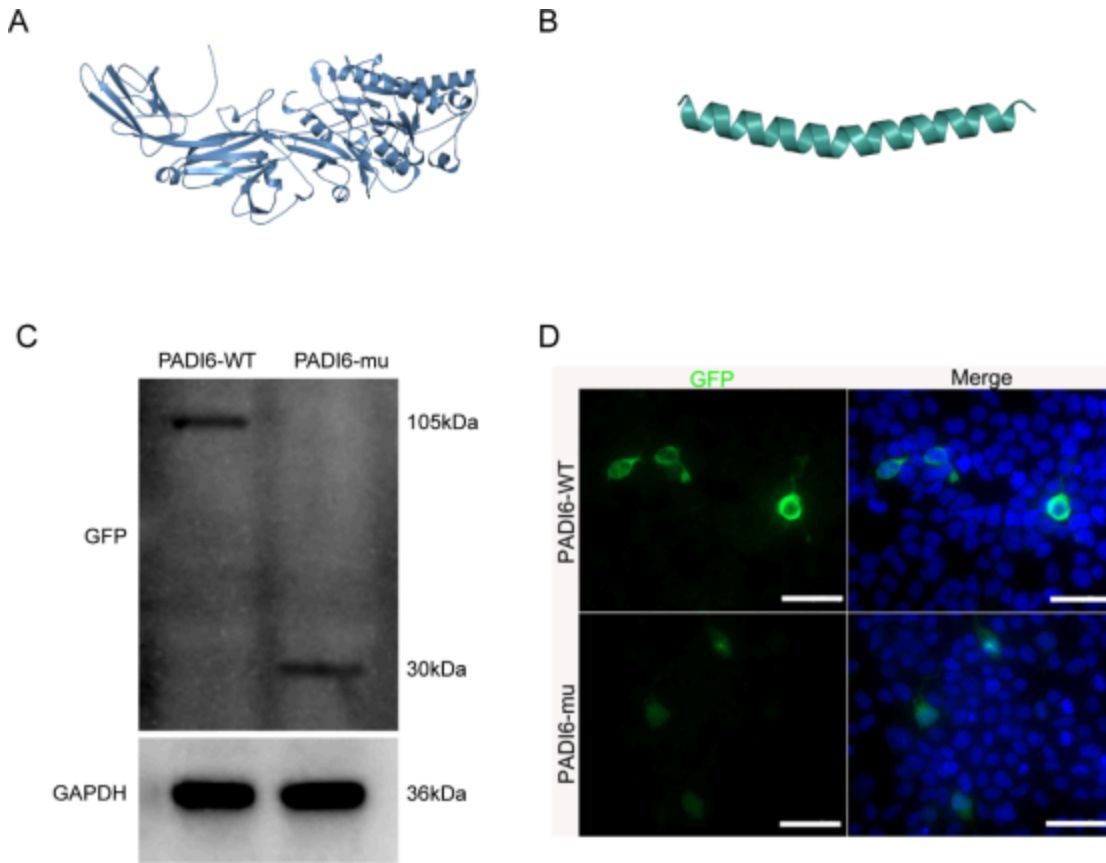


Figure 3

Three-dimensional protein structure predictions, expression analysis, and subcellular localization of wild-type and mutant PADI6. (A) Three-dimensional structure of the wild-type PADI6 protein simulated by AlphaFold2. The full-length wild-type PADI6 contains 695 amino acids. (B) Predicted structural modeling of the p.Cys40Ter truncated protein, which is expected to retain only 39 amino acids at the N-terminus and lack major functional regions, including the catalytic domain. (C) Western blot evaluation of PADI6-WT-GFP and PADI6-mut-GFP fusion protein expression levels in HEK293T cells. The wild-type construct produced a full-length fusion protein migrating at approximately 105 kDa, whereas the mutant construct produced a smaller truncated band at approximately 30 kDa. GAPDH was used as the internal loading control. (D) Immunofluorescence imaging assessing the spatial distribution of PADI6-WT-GFP and PADI6-mut-GFP fusion proteins in HEK293T cells. Green signals indicate GFP-tagged target proteins, while blue signals indicate DAPI-stained nuclei. Wild-type PADI6 showed predominant cytoplasmic localization, whereas the truncated mutant fusion protein showed a more diffuse distribution in both the cytoplasm and nucleoplasm. Scale bar: 50 μ m. Abbreviations: WT, wild-type; mu, mutant.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table2..docx](#)
- [Table2..docx](#)
- [Table1..docx](#)
- [Table1..docx](#)
- [SupplementaryTables.docx](#)
- [SupplementaryTables.docx](#)