A *WFS1* variant disrupting acceptor splice site uncovers the impact of alternative splicing on ER-stress independent β cell apoptosis in a patient with Wolfram syndrome.

Raniero Chimienti†[*1,2], Silvia Torchio†[*1,2], Valentina Zamarian[1], Gabriel Siracusano[1,2], Laura Monaco[1,2], Marta Tiffany Lombardo[1], Fabio Manenti[1], Federica Cuozzo[1], Greta Rossi[2,3], Paola Carrera[4,5], Valeria Sordi[1], Vania Broccoli[3,6], Riccardo Bonfanti[7], Giorgio Casari[2,8], Giulio Frontino[#7], Lorenzo Piemonti[#*1,2]

[1]Unit of Beta Cell Biology, Diabetes Research Institute, IRCCS San Raffaele Hospital, Milan, Italy
[2]Vita-Salute San Raffaele University, Milan, Italy
[3]Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Milan, Italy
[4]Unit of Genomics for Human Disease Diagnosis, IRCCS San Raffaele Scientific Institute, Milan, Italy
[5]Laboratory of Molecular Genetics, IRCCS San Raffaele Scientific Institute, Milan, Italy
[6]National Research Council of Italy, Institute of Neuroscience, Milan, Italy
[7]Department of Pediatrics, IRCCS San Raffaele Hospital, Milan, Italy
[8]Unit of Genome-Phenome Relationship, IRCCS San Raffaele Scientific Institute, Milan, Italy

*Correspondence: Prof. Lorenzo Piemonti, Diabetes Research Institute (DRI), IRCCS San Raffaele Hospital, Via Olgettina 60, 20132, Milan, Italy (piemonti.lorenzo@hsr.it); Dr. Raniero Chimienti, PhD, Diabetes Research Institute (DRI), IRCCS San Raffaele Hospital, Via Olgettina 60, 20132, Milan, Italy (chimienti.raniero@hsr.it); Dr. Silvia Torchio, PhD, Diabetes Research Institute (DRI), IRCCS San Raffaele Hospital, Via Olgettina 60, 20132, Milan, Italy [current address: Genetics and Developmental Biology (UMR3215 / U934), Institut Curie, 11-13 rue Pierre et Marie Curie, 75005, Paris, France] (silvia.torchio@curie.fr).
†The authors wish it to be known that, in their opinion, the first two authors (Raniero Chimienti and Silvia Torchio) should be regarded as joint First Authors.
[#]The authors wish it to be known that, in their opinion, the last two authors (Giulio Frontino and Lorenzo Piemonti) should be regarded as joint Last Authors.

Abstract

Wolfram Syndrome 1 is an inherited condition manifesting in childhood-onset diabetes mellitus and progressive optic nerve atrophy, variable hearing impairment/deafness, diabetes insipidus, neurologic

defects and other psychiatric abnormalities. The causative gene is *WFS1*, which encodes for a master regulator of several cellular responses, named Wolframin. As over 200 mutations have been reported in association with a great clinical variability, a convincing genotype-phenotype correlation is crucial to deal with disease severity and identify effective therapy. Herein, we investigate a patient carrying the *WFS1* mutations c.316-1G>A and c.757A>T. By using iPSC-derived pancreatic β cells from this patient, we demonstrated that the allele carrying the acceptor splice site (ASS) mutation c.316-1G>A originates premature termination codon (PTC)-containing splicing variants, and two ORF-conserving mRNAs leading to N-terminally truncated polypeptides. We found that degradation of PTC-carrying transcripts is regulated by nonsense mediated decay (NMD) and inflammatory stress-induced changes in NMD-related genes result in over-expression of these aberrant *WFS1* mRNAs, predisposing β cells to unfolded protein response-independent apoptosis. Following Cas9-mediated recovering of ASS, we retrieved the canonical transcriptional landscape and rescued normal phenotype in patient-derived β cells. Overall, our study provides a model for the characterization of *WFS1* mutations, uncovering new therapeutic targets for this rare disease.

Introduction

Wolfram Syndrome 1 (WS1) is a progressive neurodegenerative disorder mainly characterized by co-morbidity of early-onset diabetes mellitus and optic nerve atrophy, in association or not with other endocrinological, urological and neurological abnormalities, with symptoms of highly variable onset, progression and severity (1, 2). Disease manifestations are attributable to pathogenetic variants of the *WFS1* gene, spanning 33.4kb of genomic DNA within the short arm of chromosome 4 (4p16). *WFS1* consists of eight exons, of which the first one is noncoding. The 3.6kb *WFS1* mRNA encodes for an 890 amino acids-long polypeptide, named Wolframin, having an apparent molecular weight of 100kDa. The protein displays subcellular localization to the endoplasmic reticulum (ER), where it plays a pivotal role in the control of the unfolded protein response (UPR), Ca^{++} homeostasis, endoplasmic reticulum (ER)-mitochondrial cross-talk and autophagy (3–6).

Wolframin structure and its interactome have not been entirely elucidated, but it has been reported to have protein-protein interaction domains at both N- and C-terminal, separated by nine transmembrane domains (TMDs) which allow the insertion in the ER membrane (7). Furthermore, the N-terminal domain is involved in the stabilization of tetrameric form, which is thought to be the most common and functional conformations of Wolframin (8). The 5'-untraslated region (UTR) of *WFS1* mRNA exhibits heterogeneity due to naturally occurring alternative splicing events (9, 10). More precisely, alternative splicing at the acceptor site of exon 2 causes a 4-bp deletion, enabling the differentiation of isoform 1 (aligned with the initial published sequence) from isoform 2, the latter starting four bases

before the canonical ATG translation codon (9). Despite both isoforms 1 and 2 generating the same protein, their biological significance remains unclear (11, 12).

To date, over 200 disease-causing mutations of *WFS1* have been identified along the entire sequence with no evidence of mutational hotspots (https://www.ncbi.nlm.nih.gov/clinvar/?term=WFS1[gene], updated November 2023). As different *WFS1* variants mirror a wide range of clinical phenotypes, several genotype-phenotype association studies have been performed to aid clinicians in predicting more accurate prognoses and pave the way for personalized treatments (13–18). However, works investigating the molecular effects of specific *WFS1* variants on therapeutics are rare (19) and those proposing an in-depth characterization of the transcriptional/translational outcomes stemming from the disruption of splice sites and/or introduction of premature termination codon (PTC) have not been performed yet.

In our work, taking advantage of both induced pluripotent stem cells (iPSCs) and CRISPR/Cas9 technologies, we sought to decipher the consequences at the transcriptional and translational level of two recently discovered pathogenic *WFS1* mutations: c.316-1G>A, affecting the acceptor splice site (ASS) upstream exon 4, and c.757A>T, introducing a PTC in exon 7 (20, 21). Finally, we explored the maturation of *WFS1* mRNA in patient-derived iPSCs and after their differentiation into pancreatic β cells, evaluating the impact of *WFS1* PTC-carrying transcript accumulation upon exogenous stress and inflammation-dependent NMD inhibition on survival of the insulin-producing β cells.

Methods and Materials

PBMC isolation, reprogramming and cell culture

Healthy donor and patient-derived peripheral blood mononuclear cells (PBMCs) were collected using Ficoll-Paque separation method on whole blood. WFS1 iPSCs were generated by reprogramming the $CD34^+$ fraction from the patient-derived PBMCs using the CytoTune-iPS 2.0 Sendai Reprogramming Kit (ThermoFisher) according to the manufacturer's indications. Reprogrammed cells were initially cultivated on a mouse embryonic fibroblast (MEF) layer. Emerging iPSC colonies were individually picked and expanded in mTeSR1 medium (Stemcell Tech.) onto a Vitronectin (Vitronectin [VTN-N] Recombinant Human Protein, Truncated, ThermoFisher) coating at 500 ng/cm^2. Stable clones were switched into the Essential 8 Flex medium (ThermoFisher) for long term cell culturing. Passages were performed every 3-4 days following dissociation with 0.5 mM EDTA (Ambion). The iPSC lines were routinely tested for mycoplasma contamination by using MycoAlert™ Mycoplasma Detection Kit by Lonza, according to the manufacturer's instructions. Karyotyping and aCGH analysis were performed by ISENET Biobanking service unit in Milan, Italy. All WFS1-derived iPSC clones and their gene edited counterpart used in the present study had normal karyotype. Control iPSCs with WT genotype

97 (CGTRCiB10) were already available and routinely employed in the laboratory. This WT iPSC line
98 was used as unrelated healthy control and its in-depth characterization is reported in our previously
99 publications (22, 23). Differentiation into pancreatic β cells was performed in adhesion following *in*
100 *vitro* protocol published by our group (24) on a subclone of WT iPSCs (#05), two clones of WFS1
101 iPSCs (#42 and #43) and three clones of $WFS1^{wt/316\text{-}1G>A}$ iPSCs (#B10, #C6 and #F4). At the end of
102 differentiation (day 24), iPSC-derived β cells were detached using 0.5 mM EDTA and displaced in
103 35 mm or 60 mm petri dishes on orbital shaker at 55 rpm to aggregate into clusters (pseudo-islets).
104 The pseudo-islets were maintained in suspension culture at 55 rpm until their use, using the following
105 medium: CMRL 1066 (Mediatech) supplemented with 10% fetal bovine serum (FBS) (Lonza), 1%
106 Penicillin/Streptomycin (Pen/Strep), 1% L-Glutamine (Euroclone), 10 μM Alk5i II, 1 μM T3, 10 mM
107 Nicotinamide (Sigma), 10 μM H1152 (Euroclone), 100 U/ml DNase I (Sigma). Skin fibroblasts from
108 WS1 patient with genotype W189X/W189X and unrelated healthy control were gently provided by
109 Dr. Vania Broccoli's lab at IRCCS San Raffaele Hospital. Fibroblasts were kept in the Dulbecco's
110 modified Eagle's medium (DMEM) with high glucose, supplemented with 10% FBS, 1% Pen/Strep,
111 1% L-Glutamine. Passages were performed every week following detaching with Trypsin (Lonza).

112 **RNA extraction, retrotranscription and PCR/RT-qPCR**

113 Total RNA was extracted by using mirVana Isolation Kit (Ambion) and quantified by BioTek Epoch
114 Microplate Spectrophotometer (Agilent). About 1 μg of TURBO™ DNAse (Thermo Fisher)-treated
115 RNA was reverse transcribed using SuperScript™ IV First-Strand Synthesis System (Thermo Fisher),
116 with oligo(dT) primers, according to the manufacturer's instructions. For the identification of *WFS1*
117 alternative splicing isoforms, 10 ng of cDNA was pre-amplified using *WFS1* primers listed in **Table**
118 **1** (0.4 μM fw/rev) and the 1.25 U DreamTaq HS DNA Polymerase (Thermo Fisher), with 3.5 mM
119 $MgCl_2$ and 0.2 mM dNTPs, in 50 uL final volume reaction. Pre-amplification consisted of following
120 cycling conditions: 1 x 95°C 3', 15 x (95°C 30'', 60°C 30'', 72°C 45''), 1 x 72°C 5'. A total of 1 uL
121 of pre-amp reaction was used as template for further amplification. Second PCR was performed using
122 1.25 U DreamTaq HS DNA Polymerase, the same primers of pre-amp PCR (0.4 μM each), 2.0 mM
123 $MgCl_2$ and 0.2 mM dNTPs, with the following cycling conditions: 1 x 95°C 3', 35 x (95°C 30'', 60°C
124 30'', 72°C 45''), 1 x 72°C 5'. The PCR products were run on analytical pre-cast E-Gel™ Agarose
125 Gels with SYBR™ Safe DNA Gel Stain, 2% (ThermoFisher), using the GeneRuler 100 bp Plus DNA
126 Ladder (Thermo Fisher) as size markers. The PCR for screening of Cas9-edited clones and for XBP-
127 1 splicing assay was performed using 1.25 U DreamTaq HS DNA Polymerase, 2.0 mM $MgCl_2$ and
128 0.2 mM dNTPs, with the following cycling conditions: 1 x 95°C 3', 40 x (95°C 30'', 60°C 30'', 72°C
129 45''), 1 x 72°C 5'. The spliced/unspliced (s/u) XBP-1 ratio was measured by gel-electrophoresis and
130 densitometric analysis. For RT-qPCR experiments, PowerUp™ SYBR™ Green Master Mix was used

131 according to the manufacturer's protocol, and amplification with annealing temperature of 60°C was
132 performed on a 7900HT Fast Real-Time PCR System (Applied Biosystem) with Fast 96-Well Block
133 Module. Expression level of analyzed genes were normalized applying the 2-ΔCt method, using the
134 GAPDH as housekeeping gene. Complete list of primers is reported in **Table 1**.

135 **PCR product purification, cloning and Sanger sequencing**

136 WFS1-derived PCR products obtained following amplification with Ex3fw-Ex5rev primer pair were
137 separated on 2% agarose extraction gel prepared dissolving low-melt-temperature agarose in 1X TBE.
138 Both 324bp- and 180bp-long amplicons were purified separately by NucleoSpin PCR & Gel Clean-
139 up (Macherey Nagel) and cloned into pCR4-TOPO TA vector using TOPO™ TA Cloning™ Kits for
140 Sequencing (Thermo Fisher), following manufacturer's instructions. The pCR4-TOPO construct was
141 transformed into One Shot TOP10 (Thermo Fisher) chemically competent cells, then using LB-agar
142 plates containing 100 µg/mL ampicillin as selection medium. The bacterial colonies were picked and
143 analyzed for presence of insert by following amplification with the T3/T7 primers, and agarose gel
144 electrophoresis. Positive transformants were expanded in LB added with 100 µg/mL ampicillin and
145 pCR4-TOPO plasmid containing the PCR products was purified using NucleoSpin Plasmid, Mini kit
146 for plasmid DNA (Macherey Nagel). The Sanger sequencing was performed from plasmid by the EZ-
147 Seq service at Macrogen Europe. Electropherograms were visualized and sequence analyzed by using
148 SnapGene v4.3.2 (Dotmatics).

149 **Next Generation Sequencing**

150 Following the patient's referral to a geneticist for further examination, Next Generation Sequencing
151 (NGS) was performed on a panel of 31 monogenic diabetes-related genes: ABCC8 (NM_000352),
152 AIRE (NM_000383), BLK (NM_001715), CEL (NM_001807), CISD2 (NM_001008388), EIF2AK3
153 (NM_004836), FOXP3 (NM_014009), GATA4 (NM_002052), GATA6 (NM_005257), GCK
154 (NM_000162), GLIS3 (NM_001042413), HNF1A (NM_000545), HNF1B (NM_000458), HNF4A
155 (NM_000457), IER3IP1 (NM_016097), INS (NM_000207), ISL1 (NM_002202), KCNJ11
156 (NM_000525), KLF11 (NM_003597), MNX1 (NM_005515), NEUROD1 (NM_002500),
157 NEUROG3 (NM_020999), PAX4 (NM_006193), PAX6 (NM_000280), PDX1 (NM_000209),
158 PTF1A (NM_178161), RFX6 (NM_173560), SIRT1 (NM_012238), SLC19A2 (NM_006996),
159 SLC29A3 (NM_018344), SLC2A2 (NM_000340), WFS1 (NM_006005). Enrichment of fragments
160 was performed with the TruSight One Sequencing kit (Illumina). Sequencing of coding regions and
161 exon-intron junctions was done on NextSeq Illumina platform. Diagnosis of Wolfram Syndrome 1
162 was confirmed when mutations in *WFS1* gene were reported in the medical record of the patients,
163 after comparison with variants deposited in the following databases: NCBI dbSNP, 1000 Genomes,

164 dbNSFP, ClinVar, LOVD. The deep sequencing analysis on PCR products obtained by amplifying
165 *WFS1* sequence comprised between exon 3 and exon 5 was performed to confirm the sequence of the
166 alternative splice variants deriving from c.316-1G>A-carrying allele in the WFS1 cells. PCR-derived
167 fragments were purified by using NucleoSpin PCR & Gel Clean-up (Macherey Nagel). Libraries were
168 obtained following the TruSeq Nano WGS (Illumina) protocol, deriving 250nt-long $4x10^6$ fragments
169 per sample. Paired-end sequencing was performed through MiSeq_500_v2 Illumina platform on two
170 WFS1 iPSCs clones (#01 and #42) and one WT iPSCs clone (#05). The reads in R1 and R2 FASTQ
171 files generated at the end of the sequencing were analyzed for the base quality. Adapter were removed
172 by using Trimmomatic, v.0.32. Preliminary, explorative alignment was performed by STAR v2.5.3a,
173 mapping reads to the reference human genome hg38, Gencode version 28. Percentages of uniquely
174 mapped reads were about 70-77% for the two WFS1 samples and 85% for the WT sample. More than
175 99,99% of reads were assigned to the *WFS1* gene for all the three samples.

176 **CRISPR/Cas9-mediated gene editing**

177 CRISPR/Cas9-mediated correction of the point mutation located upstream exon 4 of *WFS1* gene was
178 performed in patient-derived iPSCs. GeneArt™ Precision gRNA Synthesis Kit (ThermoFisher) was
179 employed for *in vitro* synthesis of the gRNAs, whose sequences are reported in the **Table 1**. As donor
180 templates a single strand oligodeoxynucleotide (ssODN) homologous to the region encompassing the
181 *WFS1* intron 3 - exon 4 boundary junction was used. 20 µg of Cas9 protein (TrueCut™ Cas9 Protein
182 v2, ThermoFisher) and 4 µg of *in vitro* transcribed (IVT) gRNAs were pre-incubated for 20 minutes
183 at RT to form ribonucleoprotein (RNP) complex. An RNP:ssODN ratio of 1:1.5 was used by adding
184 36 µg of ssODN. Finally, 2 µg of pmaxGFP plasmid (Lonza) was added to the mix as selection marker
185 for cell sorting. For Cas9-mediated gene editing, $1 \cdot 10^5$ iPSCs were electroporated by using the 4D-
186 nucleofector with P3 electroporation solution (Lonza) and setting the CB-150 program, according to
187 the manufacturer's instructions. The electroporated cells were replated and cultured in complete fresh
188 media for 5 days. Single cell cloning was performed following cell sorting of GFP-positive cells.

189 **Cell sorting and FACS analysis**

190 Electroporated cells were analyzed with FACSAria™ Fusion Flow Cytometer (BD) and GFP-positive
191 cells were sorted in Vitronectin-coated 96-well plate for single cell cloning. Sorted cells were cultured
192 in Essential 8 Flex added with 10% added with CloneR™ (Stemcell Tech.) until colonies emerged.
193 Single clone-derived colonies were screened by restriction enzyme digest as described below. FACS
194 analysis was performed on trypsin-treated iPSCs and iPSC-derived β cells after incubation with the
195 Live/Dead Fixable Violet Stain (ThermoFisher) in PBS at 4°C for 30' to exclude dead cells from the
196 analysis. For intracellular staining, cells were fixed by using Cytofix/Cytoperm (Becton Dickinson,

197 BD), then permeabilized with Phosflow™ Perm Buffer III (BD), according to the manufacturer's
198 instructions. Staining with conjugated antibody was performed for 30' at 4°C. The list of antibodies
199 used for FACS analysis is reported in **Table 2**. For the ER-stress induction and immunotoxicity assay,
200 early and late apoptosis were evaluated by using the FITC Annexin V Apoptosis Detection Kit I (BD),
201 following to the manufacturers' instructions. Cells were inspected on the FACSCanto™ II System
202 (BD) or CytoFLEX LX (Beckman Coulter) cytometer, and analyzed by FlowJo™ Software V.10
203 (FlowJo LLC, Ashland, Oregon, USA).

204 **Homologous directed repair screening**

205 Single colonies of growing clones were picked and transferred into direct PCR lysis buffer containing
206 150 mM NaCl, 10 mM Tris-HCl pH7.5, 5% Tween-20 and 0.4 mg/ml Proteinase K (Promega), then
207 incubated for 6 hours at 55°C until complete lysis is achieved. Crude lysate was incubated for 45' at
208 85°C to inactivate Proteinase K. After centrifugation to collect debris, 1 µL of lysate was used per 50
209 µL genotyping PCR reaction with HR fw/rev primers reported in **Table 1**. The PCR product was
210 purified using NucleoSpin PCR & Gel Clean-up (Macherey Nagel) and quantified by BioTek Epoch
211 Microplate Spectrophotometer (Agilent). Up to 1 µg of DNA was digested with 10 U of SacI enzyme
212 (NEB) in NEBuffer r1.1 for 1 hour at 37°C. Digested DNA derived from each clone was resolved on
213 2% agarose gel. Amplified sequence is long 272 bp. Enzymatic digestion of recombined sequence
214 produces two bands of 152 bp and 119 bp. If homologous recombination with donor DNA involves
215 one allele, the 272 bp-long and the 152/119 bp-long pair amplicons are observed (50% chance that
216 the c.316-1G>A-carrying allele has been corrected). If homologous recombination occurs in both
217 *WFS1* alleles, only digested products are displayed. For following experiments, only clones incurred
218 in bi-allelic recombination were selected and evaluated for *WFS1* expression.

219 **Immunofluorescence**

220 For indirect immunofluorescence, cells were cultured on coated Falcon™ Chambered Cell Culture
221 Slides. Cell monolayers were washed with PBS, fixed with paraformaldehyde (4%) in PBS for 20' at
222 RT and then treated with 15 mM glycine for 5' at RT. After blocking and permeabilization with PBS
223 supplemented with 0.4% Triton X-100, 2% BSA, 5% FBS for 45' at RT, cells were washed twice
224 with PBS and then incubated overnight at 4°C with appropriate primary antibody reported in **Table**
225 **2** diluted in PBS supplemented with 2% BSA. Cells were washed 3 times with PBS and then incubated
226 with appropriate secondary antibody diluted in the same dilution buffer as the primary antibody for 1
227 h at RT. After washing, the nuclei were counterstained with Hoechst 33342. Images were acquired
228 using Olympus FluoVIEW FV 3000 (VivaScope Research) confocal microscope and analyzed using
229 the Fiji software v.1.52p (25).

230 **Immunoblot analysis**

231 WT and WFS1 iPSCs, and WT and WFS1-KO fibroblasts were lysed in the M-PER™ Mammalian
232 Protein Extraction Reagent and proteins were quantified by Pierce™ Rapid Gold BCA Protein Assay
233 Kit. Protein extracts were resolved on SDS-PAGE (4-20%) using Novex™ Tris-Glycine Mini Protein
234 Gels and electro-transferred onto a polyvinylidene difluoride (PVDF) membrane. Transfer efficiency
235 was evaluated using Ponceau S (Sigma). After 1 h of incubation in blocking buffer (TBS, 0.1% Tween
236 20, 5% skim milk) at RT, the membranes were probed overnight at 4°C with the primary antibodies
237 reported in **Table 2** diluted in TBS supplemented with 0.1% Tween 20 and 5% skim milk or BSA.
238 Horseradish peroxidase-labeled antibodies (diluted 1:1000 in the same dilution buffer as the primary
239 antibody) were used as secondary antibody, and revelation was performed using the SuperSignal™
240 West Pico PLUS chemiluminescent detection system according to the manufacturer's instructions by
241 ChemiDoc MP (Biorad). Quantification of protein levels was performed using Fiji software. GADPH
242 was used to ensure equal loadings and to normalize protein quantification. All the listed reagents were
243 purchased from ThermoFisher unless otherwise specified.

244 **Non-sense mediated decay inhibition**

245 To determine whether NMD inhibition can increase the PTC-carrying *WFS1* transcript levels, iPSCs
246 were cultured in 6-well dishes and once reached ~70% confluence were treated with 5 μM NMDI-14
247 (MedChemExpress) or with vehicle (DMSO 0.1%) for 16 hours. RNA was collected at the end of the
248 treatment and processed following aforesaid methods.

249 **ER-stress induction and immunotoxicity assay**

250 For the thapsigargin (TG, Sigma)-induced ER-stress experiments iPSCs were treated for 8 hours with
251 100 nM TG, followed by 24 or 48 hours of stress recovery in fresh medium without TG. For iPSC-
252 derived β cells, treatment lasted for a total of 16 hours with 50 nM TG before cell death analysis. The
253 conditioning with inflammatory cytokines was performed on iPSC-derived β cells by adding 50 U/mL
254 rhIL-1β, 1000 U/mL rhIFN-γ and 10 ng/mL rhTNF-α (PeproTech) to the culture medium. Cells were
255 analyzed by FACS for positivity to Annexin V/P.I. at 8, 16, 24 and/or 48 hours, according to protocol
256 described above.

257 **Bioinformatic analysis**

258 Inference of the alternative splicing events potentially occurring after disruption of the ASS at c.316-
259 2 (AG) of *WFS1* coding sequence was conducted following preliminary examination of the cryptic
260 splice sites within the exon 4 sequence (ENSE00000701011) ± 150 bp flanking intronic regions.
261 Analysis of the canonical splice sites, branch points, cis-elements (exonic splicing enhancers - ESE -
262 and exonic splicing silencer - ESS) and trans-acting splicing factor sequences, was performed by the

Human Splice Finder Pro (https://hsf.genomnis.com). To model the most thermodynamically favorite alternative splicing events following the c.316-1G>A transition the MaxEntScan framework (26) based on the maximum entropy principle was used. For the identification of alternative splicing isoforms from deep sequencing data, trimmed reads were initially aligned by STAR v2.5.3a to a custom reference genome that was built using the wild type *WFS1* mRNA and the three alternative splicing isoforms found by Sanger sequencing. Due to higher probability of sequencing errors associated with read length, -outFilterMismatchNmax parameter was set to 10. Reads that not aligned to the custom reference sequences were re-analyzed to identify other alternative splicing isoforms, if any, and used to generate the relative consensus sequence using the EMBOSS Cons software package (https://github.com/kimrutherford/EMBOSS). Multi-alignment to the custom reference genome were visualized with Integrative Genomics Viewer (IGV) v2.5. To infer the tridimensional (3D)-structure of Wolframin N-terminal domain, the entire aminoacidic sequence of the protein was used with Ginzu protocol (27) to predict the functional domains of the protein. The sequence of N-terminal domain (residues 1-402) was then submitted to Robetta server's RoseTTAFold (https://robetta.bakerlab.org/), following *ab initio* modeling pipeline. The structures of mutated isoforms were generated employing Rosetta's comparative modelling (28) and using the *de novo* modelled WT N-terminal structure as reference. The molecular dynamics simulations and conformational change analysis were performed by Gromacs v.2021.7 (29). The N-terminal domain structure was deposited in the ModelArchive (https://www.modelarchive.org). Raw data, procedures, ramachandran plots, inter-residue distance deviation and IDDT scores, and Gromacs configuration files are reported in the accompanying data at https://www.modelarchive.org/doi/10.5452/ma-cg3qd.

Statistical analysis

GraphPad Prism software (9.0.1 version) was employed to perform statistical analyses. Assuming the normal distribution of data, student's unpaired t-test was used for comparison between two groups, while to evaluate the effects of two independent variables on a dependent variable, two-way ANOVA with Dunn-Šídák correction for multiple comparison test was applied. In all cases, a p-value below 0.05 was considered as statistically significant. Data are graphed as mean±standard error of the mean (SEM) or as mean±standard deviation (SD). Error bar meaning and number of replicates are reported in the figure legends.

Results

Analysis of patient's *WFS1* mutations

The patient was diagnosed with WS1 by the Next Generation Sequencing (NGS)-based screening for monogenic diabetes-driving mutations, as extensively described in our previous paper (30). DNA-

296 sequencing patterns were compared to the sequence of *WFS1* transcript variant 1 in GenBank (NCBI
297 Ref. Seq: NM_006005.3; Gene ID: 7466), leading to detection of the following mutations: c.316-
298 1G>A and c.757A>T (**Figure 1A-B**). The c.316-1G>A transition disrupts the acceptor (3') splice site
299 of exon 4 with uncertain outcome (**Figure 1A**) and was classified as likely pathogenic (LP) variant,
300 according to the American College of Medical Genetics and Genomics (ACMG) guidelines (31). The
301 c.757A>T transversion was also classified as a LP mutation and is located on the exon 7, where it
302 gives rise to a premature termination codon (PTC) at the amino acid residue 253 (p.Lys253Ter), as
303 shown in **Figure 1B**. Sanger sequencing on the PCR products from genomic DNA isolated from the
304 family members confirmed that the patient inherited both c.316-1G>A and c.757A>T mutations from
305 the father and the mother, respectively (**Figure 1C**). Since the parents have no symptoms related to
306 the syndrome, we assumed that both mutations do not exert any dominant negative effect and were
307 transmitted as an autosomal recessive trait. By querying the ClinVar archives of NCBI, the Human
308 Gene Mutation Database (HGMD) and the Leiden Open Variation Database (LOVD) we found that
309 splicing site variations of *WFS1* gene account for 2.9% of the total mutations and, among them, only
310 1 out of 3 involves an ASS or its proximal polypyrimidine (PY) tract. In **Figure 1D** we display the
311 genomic position of the mutations carried by our patient and we highlight the previously reported
312 ASS variants mapping to the intervening sequence (IVS)-1, IVS-4 and IVS-7 of the gene, as well (32,
313 33). To infer the effects of disruption of canonical AG ASS, we evaluated the cryptic splice sites
314 along the *WFS1* exon 4 sequence (ENSE00000701011) ± 150 bp flanking intronic regions, by using
315 the sequence analysis tool of the Human Splice Finder (**Figure 2A-C**) and we identified a total of 4
316 ASSs with an HSF score ≥ 75 (**Figure 2A, top** and **Table 3**). Among these, the acceptor splice motif
317 located at 6288996 (CACTACCTGC<u>AG</u>TT; HSF Score: 81.29) might cause a 21 bp deletion
318 (c.316_337del) that preserves the open reading frame (ORF) generating a protein lacking a 7 aa
319 fragment (VGKHYLQ) with apparent molecular weight of 99.5 kDa. None of these alternative splice
320 sites were confirmed using inhomogeneous Markov models as shown by plotting the MaxEntScores
321 ≥ 4 (**Figure 2A, bottom** and **Table 3**). Interestingly, we observed that the exonic splicing auxiliary
322 sequences ratio (ESR) resulted significantly higher (>15) along the cryptic acceptor splice sequence
323 starting at 6289115 (CTTGGCGGAC<u>AG</u>AA; HSF Score 78.06) (**Figure 2B**). This positive ESR ratio
324 was suggestive of an increased activity of exon splicing enhancer proteins as the analysis displayed
325 the presence at the 5' of exon 4 of the recognition motifs of the SC35, SRp40 and 9G8 splicing factors
326 (**Figure 2B**). We also find a total of 6 branch points (BP score > 65) along the sequence comprises
327 between 6288948 and 6289146 that could be used from the cryptic ASSs (**Figure 2C** and **Table 3**).
328 To infer which ASS could be selected by the spliceosome when the disruption of the wild type (WT)
329 site occurs, we applied a method based on Maximum Entropy Principle (26). We expected the

330 abolishment of the WT ASS would drive the splicing machinery to utilize the cryptic sequence at
331 position 6288996. However, since the c.316-1G>A mutation determines the 1 bp-shifting of the
332 canonical AG motif, according to the Maximum Entropy model, this ASS was thermodynamically
333 favoured when compared to the cryptic motif at 6288996 (**Figure 2C-D**). Nevertheless, the generation
334 of the 1 bp-shifted splice site at position c.317 dramatically alters the open reading frame (ORF) and
335 results in a PTC. Anyhow, the maximum entropy-based probabilistic model suggested the exon 4
336 skipping as the most probable outcome (MaxEntScore > 6) (**Figure 2C-D**) causing ORF breakdown
337 and PTC formation as well (**Figure 2D**).

338 **Characterization of the transcriptional outcomes from the c.316-1 G>A allele**

339 To explore the transcriptional landscape arising from ASS disruption in IVS-3, we decided to exploit
340 the stem cell technology. Specifically, we reprogrammed peripheral blood $CD34^+$ cells of the patient
341 into iPSCs, as previously described by our group (24), successfully generating a total of eight iPSC
342 syngeneic clones with normal karyotype and conserved *WFS1* heterozygous mutations. Details on
343 genomic integrity and analysis of pluripotency are reported in **Figure S1**. We designed a PCR strategy
344 to amplify *WFS1* transcripts along exons 3, 4 and 5, by using two pairs of primers as shown in **Figure**
345 **3A**. When the *WFS1* mRNA between exon 3 and 5 was amplified in both WT and WFS1 iPSCs, we
346 highlighted a potential mutation-related alternative splicing isoform as we observed an amplicon of
347 ~180bp in the WFS1 iPSCs (**Figure 3B**). Both cell lines also display the expected amplicon of 324
348 bp, which derives from normal gene in the WT cells, and likely from the c.757A>T-carrying allele in
349 the WFS1 ones. The amplification of the sequence between exons 3 and 4 or between exon 4 and 5
350 generated a single amplicon in WFS1 iPSCs, supporting hypothesis that exon 4 skipping was the most
351 probable event occurring in the c.316-1G>A-carrying pre-mRNA. (**Figure 3B**). To determine the
352 sequences of the observed PCR products, we TA-cloned the WFS1-derived 324bp- and 180bp-long
353 bands and >20 clones were analyzed by Sanger sequencing. We found both natural and alternative
354 splicing variants, that we named according to Human Genome Variant Society (HGVS) nomenclature
355 standard. Among the 324bp-long amplicons we identified two sequences: the WT (deriving from the
356 c.757A>T-carrying allele) and the c.316del, corresponding to the maximum entropy-based prediction
357 of alternative splicing resulting from the 1bp-shifiting of ASS. Analyzing the 180bp-long amplicons
358 we found two different sequences, one resulting from the skipping of the entire exon 4 (c.316_461del)
359 and the other retaining the last four bases of the exon 4 (c.316_457del). Among the identified c.316-
360 1 G>A-carrying allele-deriving isoforms only the c.316_457del preserved the ORF (**Figure 3C**). To
361 exclude that reprogramming could alter the processing of *WFS1* pre-mRNA, as well as to rule out
362 that observed isoforms were a specific characteristic of stem cells, due to a broader expression of
363 splicing enhancer factors, we amplified the WT and WFS1 primary PBMC-derived cDNA by Ex3fw-

364 Ex5rev primer pair, observing a complex pattern of amplicons in patient-derived PBMCs (**Figure**
365 **3D**). Based on these findings, we generated a sequence library from both WT and WFS1-derived
366 amplification products of the Ex3fw-Ex5rev primer pair and performed deep sequencing to identify
367 all the potential alternative splicing isoforms occurring in the mutated cells. We performed paired-
368 end sequencing on the PCR products from one WT subclone and two WFS1 clones. Assessment of
369 read quality shown that R1 (forward) reads have poor quality after 210bp, while R2 (reverse) reads
370 lose accuracy after 180bp. Preliminary STAR alignment to reference human genome hg38 revealed
371 a reduced number of reads mapping to the *WFS1* gene in the affected cell-derived samples. This
372 mirrored the percentage of trimmed paired-end reads per sample used to individuate the alternative
373 splicing isoforms in WFS1 iPSCs (**Figure E**). Interestingly, a total of 4 c.316-1 G>A-carrying allele-
374 deriving alternative splicing isoforms were identified in WFS1 iPSCs, whereas in WT cells the
375 normal, correctly spliced transcript represented more than 99% of the mapped reads (**Figure 3F**).
376 NGS analysis thus confirmed sequences identified by cloning and Sanger sequencing and allowed to
377 identify an additional splice variant: c.271_513del. Among detected transcripts, the c.316del and
378 c.316_461del isoforms resulted in frameshift and PTC generation, whereas the c.316_457del and
379 c.271_513del isoforms conserved the ORF, missing part of the *WFS1* mRNA. Specifically,
380 c.316_457del lost the entire exon 4 with the exception of the last four nucleotides (AGAG), while
381 c.271_513del included a larger deletion comprising 3'-end of exon 3, the entire exon 4 and the first
382 52bp of exon 5 (**Figure 3G** and **Table 4**). Relative quantification of total amount of *WFS1* transcripts
383 revealed that reduction in WFS1 iPSCs was not significant ($p = 0.07$), due to the high variability of
384 *WFS1* expression in WT cells (**Figure 3H**). We speculated that stable low expression of WFS1 total
385 mRNAs in affected iPSCs might result from post-transcriptional regulation by non-sense mediated
386 decay (NMD) of the transcript deriving from the c.757A>T-carrying allele and the 2 out of 4
387 transcripts deriving from the c.316-1 G>A-carrying allele as well. Indeed, the two isoforms c.316del
388 and c.316_461del were significantly upregulated by treatment with NMDI-14, a molecule that
389 disrupts the SMG7-UPF1 interactions and strongly inhibits NMD (**Figure 3I**).

390 **Characterization of the translational outcomes from the c.316-1 G>A allele**

391 The ORF-conserving c.316_457del and c.271_513del isoforms induce an in-frame skipping of 141bp
392 and 240bp, resulting in 47aa (Val106-Arg152) and 80aa (Val91-Thr170) losses, respectively. This
393 would lead to the internally truncated forms of the protein, lacking part of the N-terminal domain but
394 preserving transmembrane and C-terminal domains. To confirm the presence of mutated Wolframin,
395 we performed immunoblot with two antibodies recognizing the N-terminal residues surrounding the
396 Ala43 position and the Lys679-Phe783 C-terminal residues, respectively. The WT and WFS1-KO
397 fibroblasts were used as controls of antibody specificity. As shown in **Figure 4A**, the C-terminal-

398 recognizing Ab confirmed that the Wolframin amount in the WFS1 iPSCs was ~50% lower than the
399 WT counterpart. This data highlighted that residual protein in affected cells derived from the c.316-
400 1 G>A-carrying allele only. Of note, the higher expression variability and the stable levels of
401 Wolframin in WT and WFS1 iPSC clones, respectively, mirrored that observed at transcriptional
402 level. As expected, the N-terminal-recognizing Ab did not detect altered isoforms in WFS1 iPSCs
403 (**Figure 4B**). These results confirm that at least one of the splicing isoforms found in WFS1 iPSCs
404 leads to residual protein production, albeit missing a portion of the N-terminal domain. To uncover
405 the impact of the portions lost upon alternative splicing events, we refined the structure of the N-
406 terminal domain (residues 1-402) by following the *ab initio* modeling pipeline of RoseTTAFold. We
407 deposited the *de novo* model in the ModelArchive Database (doi: 10.5452/ma-cg3qd) and used it to
408 perform homology-based, comparative modeling of the mutated isoforms p.Val106-Arg152del and
409 p.Val91-Thr170del (**Figure 4C-D**). The p.Val106-Arg152del isoform partially preserved the WFS1-
410 Calmodulin (CaM) binding site, whereas the p.Val91-Thr170 isoform lost the entire WFS1-CaM
411 domain (34). Moreover, both isoforms resulted in altered N-terminal structure, potentially affecting
412 interaction with other well-known WFS1 interactors, such as neuronal calcium sensor 1 (NCS1), that
413 is involved in the mitochondria-associated membranes formation and calcium signaling (35). The
414 dynamic simulation of predicted structures revealed higher instability of mutated N-terminal domain
415 (> 0.5 Root Mean Standard Deviation, RMSD) compared to the WT ones (< 0.3 RMSD) after 10 ns
416 into a salt-equilibrated aqueous solution (**Figure 4E**). Finally, a radius of gyration of more than 30Å
417 of the p.Val106-Arg152del isoform suggested a reduced compactness of its N-terminal domain
418 (**Figure 4F**). These *in silico* inferences support hypothesis that the observed residual proteins in
419 WFS1 cells might display reduced binding with cytoplasmic interactors as well as altered capacity to
420 form homotetramers, which are considered the most stable and functional forms of Wolframin in the
421 ER membrane (7, 8).

422 **CRISPR/Cas9-mediated gene correction of the c.316-1 G>A allele**

423 To further confirm that the *WFS1* altered isoforms would derive from the c.316-1 G>A-carrying allele
424 and to obtain a syngeneic allele-specific edited iPSC line to explore the functional impact of mutated
425 transcripts, we performed a CRISPR/Cas9-mediated correction of the 316-1 G>A point mutation. We
426 designed single guide RNA (sgRNA) to target exon 4 of *WFS1* gene and trigger homologous-directed
427 repair (HR) of the upstream intron-exon junction by using a single stranded oligonucleotide (ssODN)
428 as shown in **Figure 5A**. In the ssODN sequence we introduced A>G modification at c.316-1 to restore
429 canonical ASS, and inserted four silent mutations at position p.113, 114, 115 and 116 (indicated in
430 light blue in Figure 5A) to destroy the PAM sequence and the gRNA recognition motif, without
431 altering aminoacidic code, to avoid re-targeting upon homologous recombination. Finally, we also

432 introduced a SacI cut site (GAGCTC) to allow identification of gene engineered cells. Indeed, the
433 gene edited clones were screened by amplifying the targeted sequence, followed by digestion with
434 the SacI restriction enzyme (**Figure 5B**). We obtained a total of six WFS1$^{wt/757A>T}$ iPSC clones in
435 which canonical splicing of *WFS1* mRNA and full-length isoform expression of Wolframin were
436 restored (**Figure 5C-D**). These results were confirmed by RT-qPCR showing the absence of the 4
437 alternative splicing isoforms in WFS1$^{wt/757A>T}$ iPSCs (**Figure 5E**). All gene edited clones maintained
438 normal karyotype and displayed stemness properties as confirmed by expression of the pluripotency
439 markers NANOG, OCT4 and SOX2 (**Figure S3A-D**).

440 **Expression of c.316-1 G>A allele-derived *WFS1* isoforms in insulin-producing β cells**

441 Since pancreatic tissue and, in particular, insulin-producing β cells are affected in WS1, we aimed at
442 exploring the impact of aberrant isoforms in the iPSC-derived β cells (iβ). WFS1 and WFS1$^{wt/757A>T}$
443 iPSCs were differentiated into pancreatic β cells following 24-days *in vitro* differentiation protocol,
444 and we did not observe significant differences in the number of PDX1-, NKX6.1- and INS-positive
445 cells at the end of differentiation (**Figure S3E**). Evaluating the levels of the c.316-1 G>A-carrying
446 allele-deriving mRNA isoforms we did not found any difference in c.316del and c.316_461del
447 transcript expression between iPSCs and iβ, confirming that such isoforms are subjected to NMD also
448 in terminally differentiated cells. Surprisingly, we did not observe any increase in ORF-maintaining
449 isoforms after differentiation; conversely, the iβ displayed a significant reduction of c.271_513del
450 expression, suggesting that other mechanisms might be involved in their post-transcriptional
451 regulation (**Figure 6A**). Given the pivotal role played by Wolframin in maintaining homeostasis and
452 promoting insulin biosynthesis in pancreatic β cells (1), the WFS1$^{wt/757A>T}$ iβ significantly upregulated
453 *WFS1* expression compared to undifferentiated stage. We observed from 20 to 50-fold increase in
454 *WFS1* mRNA levels after 24 days of differentiation, where the protein 3-fold exceeded GAPDH
455 levels, demonstrating the importance of Wolframin in β cells. (**Figure 6B-D**). Contrarywise, WFS1
456 cells did not display significant changes in both mRNA and protein levels after differentiation. As
457 Wolframin levels did not increase in the WFS1 iβ, we hypothesized that the residual protein could
458 not be able to guarantee the Wolframin-related functions in β cell stressogenic conditions. Moreover,
459 we observed the homotetrameric form of Wolframin in WFS1$^{wt/757A>T}$ iβ only, confirming that altered
460 residual protein lacked the portion of the N-terminal domain involved in the tetramer formation in
461 WFS1 cells (**Figure 6C**).

462 **Impact of *WFS1* aberrant isoforms on iβ cell survival upon stress and inflammation**

463 From a functional point of view, WS1 is characterized by aberrant ER stress responses and overall
464 upregulation of ER stress-related genes in basal conditions (3). However, our data are in contrast with

465 those reported in literature, as several ER stress-related genes were similarly expressed in WFS1 and
466 WFS1$^{wt/757A>T}$ cells. Interestingly, thapsigargin (TG) treatment did not significantly trigger ER stress-
467 related pathways in WFS1 cells, as displayed by differential expression of *ATF4*, *HSPA5* and *DDIT3*
468 genes, probably because residual protein activity is sufficient to control ER stress response (**Figure**
469 **7A**). Although TG treatment induces cell death, we observed that recovery from TG-induced
470 apoptosis occurred in the WFS1$^{wt/757A>T}$ but not in WFS1 cells. We treated WFS1 and WFS1$^{wt/757A>T}$
471 for 8 hours, then measured the percentage of the Annexin V-positive cells (as indicator of early stage,
472 reversible apoptosis) at the end of TG treatment (8 hours) and after 24 and 48 hours in fresh medium.
473 After 8 hours of treatment with TG the levels of early apoptosis were similar between WFS1 and
474 WFS1$^{wt/757A>T}$ cells. Conversely, the genetically corrected cells were able to reverse stress-induced
475 apoptosis after 24 hours, while the percentage of Annexin-V-positive WFS1 cells remained around
476 40% up to 48 hours after stress induction (**Figure 7B**). We followed the increase of spliced XBP-1
477 (sXBP-1) over unspliced form (uXBP-1) after the TG treatment and during recovery phase up to 48
478 hours. As the XBP-1 mRNA is induced by ATF6 and spliced by IRE1α, it represents a good indicator
479 of response to ER stress strictly related to IRE1- and ATF6-dependent pathways (36). However,
480 WFS1 cells shown a significant decrease in sXBP-1/uXBP-1 ratio in comparison to WFS1$^{wt/757A>T}$
481 counterpart after 24 hours from TG treatment. Reduced sXBP-1 levels did not correlate with higher
482 apoptotic rate WFS1 cells, supporting hypothesis that cell death may be a consequence of ineffective
483 short ER stress response. (**Figure 7C**). The treatment of WFS1 and WFS1$^{wt/757A>T}$ iβ confirmed the
484 susceptibility to TG-induced cell death of affected cells as well. After 16 hours WFS1 iβ displayed
485 higher levels of early (Annexin V^{+} cells) and late apoptosis (P.I.$^{+}$ cells) compared to WFS1$^{wt/757A>T}$
486 iβ (**Figure 7D**). Furthermore, we observed that an ER stress-independent cell death occurred in the
487 WFS1 iβ after exposure to the inflammatory cytokines IL-1β, TNF-α and IFN-γ. Both WFS1 and
488 WFS1$^{wt/757A>T}$ iβ were treated with the inflammatory cytokines up to 12 hours and early apoptosis
489 was analyzed 48 hours later. Interestingly, the WFS1 iβ displayed significant increase in Annexin-V^{+}
490 cells after exposure to inflammatory cytokines (28,75±3,08% treated vs 12,53±0,78% untreated
491 WFS1 iβ) and overall higher susceptibility to inflammatory signals compared to the gene edited
492 counterpart (28,75±3,08% WFS1 vs 14,77±1,36% WFS1$^{wt/757A>T}$). At the same time point, the
493 percentage of P.I.$^{+}$ WFS1 iβ were over two-fold of magnitude than the WFS1$^{wt/757A>T}$ ones (**Figure**
494 **7E**). Of note, as it has been shown that both ER stress induction and inflammation are able to inhibit
495 NMD and increase PCT-carrying mRNAs (37, 38), we wondered whether this might affect the PTC-
496 carrying *WFS1* isoforms, resulting in toxic accumulation of these aberrant transcripts. We therefore
497 evaluated the expression of *UPF1* and *SMG7* genes, which encode for the main components of NMD
498 complex, and as expected, treatment with stress inducers downregulates *SMG7*, but not *UPF1*, in

499 both WFS1 and WFS1$^{wt/757A>T}$ iβ (**Figure 7F**). Finally, the *SMG7* downmodulation resulted in 8-fold
500 increase of the PTC-carrying splicing isoforms c.316del and c.316_461del, similarly to that observed
501 after treatment with NMDI-14. A slight increase of the ORF-conserving isoforms c.316_457del and
502 c.271_513del was also reported **(Figure 7G)**.

503 # Discussion

504 Over the last few years, the WS1 is assuming the connotation of a spectrum of disorders, mainly due
505 to the strong variability in terms of clinical manifestations and progression (39). Because of the role
506 that Wolframin plays in a wide plethora of cellular processes, exploring the impact of *WFS1* variants
507 on cell function and survival may represent a crucial step for their association with clinical phenotype.
508 Previous attempts to discover genotype and phenotype correlations led to the classification of *WFS1*
509 mutations according to their impact on the protein expression and stability (16, 40, 41). A robust
510 classification of *WFS1* genetic variants by disease onset and spectrum and severity of symptoms has
511 been recently published as well (18). Studying relationship between disease severity score reported
512 by the Wolfram Unified Rating Scale (WURS) and *WFS1* variants, the metanalysis performed by
513 Urano's group highlighted the importance of characterization of the classes of mutations for providing
514 a more accurate prognoses of outcomes and a reliable therapy for patients with Wolfram Syndrome.
515 The derivation of iPSCs from patients followed by their differentiation into disease-relevant cell types
516 offers a great promise for patient-specific *in vitro* disease modeling, as well as for personalized drug
517 screening and autologous cell replacement therapy. This technology, combined with the use of Cas9-
518 based genetic engineering techniques, allows us to overcome the limitations imposed by other models,
519 including murine ones, which are unable to recapitulate the genetic diversity of WS1 patients (42,
520 43). In our work, we aimed at studying a case of WS1 patient characterized by heterozygous mutations
521 in the *WFS1* gene: a transversion introducing a PTC in exon 7 (c.757A>T) and a transition from G to
522 A in the splice acceptor site of exon 4 (c.316-1G>A). The presence of a mutation in a splicing site
523 represents an opportunity to understand the mechanisms that guide the maturation of the pre-mRNA
524 and, in the presence of alternative splicing preserving the ORF, to understand the role of the different
525 functional domains of the protein. Mutations in the splicing sites of *WFS1* are rather rare and the
526 effects of alterations in splicing processes of such gene have not been investigated so far. Commonly,
527 authors reporting a splice site mutation in *WFS1* gene presumed that abolishing of correct splicing
528 events results in a truncated, dysfunctional protein, although its validation at the molecular level was
529 not reported. Indeed, Van Den Ouweland and colleagues reported the homozygous intronic G to A
530 transition in the donor splice site of exon 4, assuming that this mutation abolishes right splicing and
531 results in a truncated protein of 157 amino acids (10). Elli and colleagues described a rare case of
532 segmental paternal uniparental heterodisomy, causing the loss of an unspecified splicing motifs and

533 inducing the skipping of the translation start site; this mutant isoform has been hypothesized to be
534 untranslated (44). Another group instead described a patient with GT-to-AT transition in donor splice
535 site of exon 7, leading to a cryptic splice site activation, 86 bp insertion and severely premature protein
536 truncation (45).
537 Here we managed to describe four novel isoforms of *WFS1* mRNA resulting from alternative splicing
538 due to loss of the canonical ASS upstream the exon 4. In particular, 2 out of the 4 found isoforms
539 were correctly predicted by using of probabilistic model based on the maximum entropy principle.
540 Indeed, we inferred that the two most probable alternative splicing isoforms could be c.316del
541 (deriving from 1-bp shifting of the ASS following the mutation in c.316-1) and c.316_461del (losing
542 of the exon 4 due to the direct skipping to the ASS of exon 5). Otherwise, the other 2 isoforms
543 c.316_457del and c.271_513del were not predicted by the *in silico* analysis, but we found that the
544 sequence surrounding the cryptic uncanonical splicing site generating the c.316_457del isoform
545 resulted particularly enriched in ESEs. We excluded that the alternative splicing events occurring in
546 the investigated patients were a consequence of reprogramming of the somatic cells into iPSCs,
547 refuting previous evidences reporting that PBMCs from the same patient displayed *WFS1* transcript
548 originating from the c.757A>T-carrying allele only (46). Despite designing of primers to amplify
549 *WFS1* mRNA was similar to that described by Panfili and colleagues, our protocol allows us to
550 identify the alternative splicing isoforms deriving from the c.316-1G>A-carrying allele also in
551 patient-derived primary PBMCs. Deep sequencing further confirmed the four isoforms, which we
552 have chosen to classify based on the maintenance of the ORF and, therefore, on their ability to
553 originate a protein, although with altered or reduced function.
554 Interestingly, the two ORF-maintaining isoforms c.316_457del and c.271_513del conserved ability
555 to produce a protein with intact transmembrane and C-terminal domains, thus potentially preserving
556 its subcellular localization and intraluminal interactions. We confirmed expression in WFS1 iPSCs
557 of Wolframin lacking part of N-terminal domain, but still maintaining the C-terminal epitope. *In silico*
558 modeling confirmed that aminoacidic sequence lost by both isoforms disrupts the binding sites with
559 pivotal interactors such as CaM and NCS1, supporting hypothesis that the patient-derived iβ could
560 be particularly impaired in their capability to regulate the calcium signaling and the mitochondria-ER
561 calcium-dependent crosstalk. Conversely, maintenance of C-terminal domain could preserve other
562 functions, such as interaction with ER stress-related protein ATF6 for regulation of the UPR and other
563 cargo proteins involved in trafficking of proinsulin in the pancreatic β cells (3). As a matter of facts,
564 we did not observe significant alterations in ER stress markers upon ER stress induction and of the
565 ATF6 cleaved form, nor we reported impairments in insulin biosynthesis and secretion in comparison
566 to unrelated WT or syngeneic gene edited iβ. However, we found that after differentiation into

567 pancreatic β cells, the *WFS1* gene in affected cells was unable to reach expression levels of the gene
568 corrected clones. This resulted in low levels of residual proteins; therefore, we believe that the
569 conserved functions in the N-terminal lacking protein are not capable to ensure long-term survival of
570 cells carrying such mutations. However, investigating the functional aspects in this patient-derived
571 cellular model could be important both to better elucidate the role of Wolframin and its functional
572 domains in β cells, and to prove the validity of targeted therapies aimed at lengthening the half-life
573 of the residual protein or compensating the missing functions. A similar rationale is highlighted in
574 the cystic fibrosis field, another autosomal recessive disorder, where knowledge acquired from the
575 genotype-phenotype association studies has helped in the identification of effective therapeutic
576 strategies and improved the prognosis of patients (47–49).
577 Our work pointed out another important aspect concerning the physiologic mechanisms involved in
578 the pancreatic β cells: NMD. Indeed, the integrity and accuracy of transcript processing is pivotal for
579 β cells to meet the physiological demands and pathophysiological challenges they face (50).
580 We reported that the PTC-carrying c.316del and c.316_461del isoforms are regulated by NMD, as by
581 inhibiting the UPF1/SMG7 interaction the degradation of such transcripts was prevented. Indeed,
582 NMD orchestrates the Regulated Unproductive Splicing and Translation (RUST) mechanism to
583 eliminate the PTC-containing transcript isoforms generated due to perturbed AS (51). Although the
584 role of RUST in pancreatic β cells is still largely unknown, deregulation of NMD due to islet stress
585 or inflammation influences RUST, leading to accumulation of unproductive transcript isoforms,
586 which are implicates in β cell dysfunction, vulnerability, and death (50). Interestingly, we found that
587 patient-derived cells were particularly susceptible to the TG-induced cell death due to impairments
588 to recover from TG-mediated Sarco-Endoplasmic Reticulum Calcium ATPase (SERCA) inhibition.
589 Time course of changes in fraction of sXBP-1 revealed that TG treatment induces rapid increase of
590 IRE1α-mediated XBP-1 cleavage in both WFS1 and $WFS1^{wt/757A>T}$ cells. However, during recovery
591 phase, s/u XBP-1 ratio significantly decreased in affected cells, suggesting that the ER stress response
592 ends after the removal of the stressful stimulus, but cellular damage persists, still leading to activation
593 of the apoptotic cascade.
594 Moreover, exposure to inflammatory cytokines increased the apoptotic rate in WFS1 iβ compared to
595 its genetically corrected counterpart. A previous study on WS1 patient presented in this work reported
596 a state of chronic inflammation associated to high levels of inflammatory cytokines, including IL-1β,
597 TNF-α, IFN-γ and IL-6 (21). The increased secretion of inflammatory cytokines in the serum of the
598 patient and susceptibility of *WFS1*-deficient β cells to exposure to pro-inflammatory factors could
599 represent a new etiopathogenetic model for WS1-related diabetes. Both inflammation and other
600 exogenous stress could be the cause of NMD inhibition and consequently of accumulation of aberrant

transcripts. Indeed, after exogenous stimulus, we observed a decrease in *SMG7* levels and a critical increase of *WFS1* PTC-containing isoforms. Accumulation of these transcripts could be the cause of β cell dysfunction, improper ER stress response and apoptosis. This evidence is strengthened by the low rate of apoptosis in the corrected cells upon stress, despite the decreasing in NMD activity.

The case study we reported represents a new model to study Wolframin from a functional point of view, highlighting how each single mutation of *WFS1* gene can determine dramatically different outcomes from a clinical point of view. We underline the need for a better classification of mutations in *WFS1* gene, in order to explain the penetrance and expressivity of the clinical characteristics of the disease. Further understanding of mechanisms by which *WFS1* variants functionally affect ER stress or inflammatory responses may aid in the identification of other potential therapeutic targets.

Finally, role of NMD in WS1 has not been considered until now, despite the majority of mutations in the *WFS1* gene introducing a PTC. Incomplete RUST leads to accumulation of *WFS1* unproductive transcripts whose translation into unfolded polypeptides overwhelms ER capacity and consequently may drive to unresolved ER stress. The role of NMD among the different cell contexts may explain susceptibility to *WFS1* mutations of specific cell types. Although further research is needed to validate this hypothesis, the NMD inhibition and PTC-carrying mRNA accumulation in WS1 could represent targetable mechanisms for development of new therapeutic strategies.

Data availability

All data are contained within the manuscript and/or supplementary files. WFS1 N-terminal domain structure has been deposited in ModelArchive (https://www.modelarchive.org) and accessible at the following address: https://www.modelarchive.org/doi/10.5452/ma-cg3qd. All raw data that were not directly included in the manuscript or that have not been deposited in online repositories, are available on request from the corresponding authors.

Acknowledgments

We thank the San Raffaele Scientific Institute Advanced Light and Electron Microscopy BioImaging Center (ALEMBIC), for confocal immunofluorescence images, and San Raffaele Scientific Institute Flow cytometry Resource, Advanced Cytometry Technical Applications Laboratory (FRACTAL), for cell sorting experiments. We also thank the Center for Omics Sciences (COSR) of the San Raffaele Scientific Institute to have provided support in deep sequencing and bioinformatic analysis. We are grateful to San Raffaele Telethon Institute for Gene Therapy (SR-TIGET) and to Angelo Lombardo for the access to BLS2 work areas and use of 4D-Nucleofector System.

Author contributions: study conception and experimental design, R.C., G.F. and L.P.; methodology, R.C. and S.T.; investigation, R.C., S.T., V.Z., G.S., L.M., M.T.L., F.M., F.C.; bioinformatic analysis, R.C.; formal analysis and interpretation of results, R.C., S.T., G.S. and L.P.; genetic testing in clinical practice, P.C.; medical valuation and patient specimen collection, G.F. and R.B.; cell reprogramming support and resources, G.R. and V.B.; stem cell differentiation, V.Z., L.M., M.T.L. and V.S.; writing - original draft, R.C. and S.T.; writing - review and editing, R.C., S.T., G.S., P.C., G.C., G.F. and L.P.; supervision, L.P.

Funding

This study was supported by private family donation financing investigations on Wolfram Syndrome 1 at the Diabetes Research Institute (DRI) of the IRCCS San Raffaele Hospital. We also extend many thanks to the Italian Ministry for University and Research for supporting part of activities through the funds of the National Recovery and Resilience Plan (PNRR-MR1-2022-12375914).

Conflict of interest

The authors have no conflicts of interest to declare.

647 # References

648 1. Abreu,D., Asada,R., Revilla,J.M.P., Lavagnino,Z., Kries,K., Piston,D.W. and Urano,F. (2020) Wolfram
649 syndrome 1 gene regulates pathways maintaining beta-cell health and survival. *Lab Invest*, **100**,
650 849–862.

651 2. Rigoli,L., Bramanti,P., Di Bella,C. and De Luca,F. (2018) Genetic and clinical aspects of Wolfram syndrome
652 1, a severe neurodegenerative disease. *Pediatr Res*, **83**, 921–929.

653 3. Fonseca,S.G., Fukuma,M., Lipson,K.L., Nguyen,L.X., Allen,J.R., Oka,Y. and Urano,F. (2005) WFS1 is a novel
654 component of the unfolded protein response and maintains homeostasis of the endoplasmic
655 reticulum in pancreatic beta-cells. *J Biol Chem*, **280**, 39609–39615.

656 4. Cagalinec,M., Liiv,M., Hodurova,Z., Hickey,M.A., Vaarmann,A., Mandel,M., Zeb,A., Choubey,V., Kuum,M.,
657 Safiulina,D., *et al.* (2016) Role of Mitochondrial Dynamics in Neuronal Development: Mechanism for
658 Wolfram Syndrome. *PLoS Biol*, **14**, e1002511.

659 5. La Morgia,C., Maresca,A., Amore,G., Gramegna,L.L., Carbonelli,M., Scimonelli,E., Danese,A.,
660 Patergnani,S., Caporali,L., Tagliavini,F., *et al.* (2020) Calcium mishandling in absence of primary
661 mitochondrial dysfunction drives cellular pathology in Wolfram Syndrome. *Sci Rep*, **10**, 4785.

662 6. Crouzier,L., Richard,E.M., Diez,C., Alzaeem,H., Denus,M., Cubedo,N., Delaunay,T., Glendenning,E.,
663 Baxendale,S., Liévens,J.-C., *et al.* (2022) Morphological, behavioral and cellular analyses revealed
664 different phenotypes in Wolfram syndrome wfs1a and wfs1b zebrafish mutant lines. *Hum Mol*
665 *Genet*, **31**, 2711–2727.

666 7. Wilf-Yarkoni,A., Shor,O., Fellner,A., Hellmann,M.A., Pras,E., Yonath,H., Shkedi-Rafid,S., Basel-Salmon,L.,
667 Bazak,L., Eliahou,R., *et al.* (2021) Mild Phenotype of Wolfram Syndrome Associated With a
668 Common Pathogenic Variant Is Predicted by a Structural Model of Wolframin. *Neurol Genet*, **7**,
669 e578.

670 8. Hofmann,S., Philbrook,C., Gerbitz,K.-D. and Bauer,M.F. (2003) Wolfram syndrome: structural and
671 functional analyses of mutant and wild-type wolframin, the WFS1 gene product. *Hum Mol Genet*,
672 **12**, 2003–2012.

673 9. Eller,P., Föger,B., Gander,R., Sauper,T., Lechleitner,M., Finkenstedt,G. and Patsch,J.R. (2001) Wolfram
674 syndrome: a clinical and molecular genetic analysis. *Journal of medical genetics*, **38**.

675 10. van den Ouweland,J.M.W., Cryns,K., Pennings,R.J.E., Walraven,I., Janssen,G.M.C., Maassen,J.A.,
676 Veldhuijzen,B.F.E., Arntzenius,A.B., Lindhout,D., Cremers,C.W.R.J., *et al.* (2003) Molecular
677 characterization of WFS1 in patients with Wolfram syndrome. *Journal of Molecular Diagnostics*, **5**,
678 88–95.

679 11. Inoue,H., Tanizawa,Y., Wasson,J., Behn,P., Kalidas,K., Bernal-Mizrachi,E., Mueckler,M., Marshall,H.,
680 Donis-Keller,H., Crock,P., *et al.* (1998) A gene encoding a transmembrane protein is mutated in
681 patients with diabetes mellitus and optic atrophy (Wolfram syndrome). *Nature Genetics*, **20**, 143–
682 148.

683 12. Strom,T.M., Hörtnagel,K., Hofmann,S., Gekeler,F., Scharfe,C., Rabl,W., Gerbitz,K.D. and Meitinger,T.
684 (1998) Diabetes insipidus, diabetes mellitus, optic atrophy and deafness (DIDMOAD) caused by
685 mutations in a novel gene (wolframin) coding for a predicted transmembrane protein. *Human*
686 *Molecular Genetics*, **7**, 2021–2028.

687 13. Rohayem,J., Ehlers,C., Wiedemann,B., Holl,R., Oexle,K., Kordonouri,O., Salzano,G., Meissner,T.,
688 Burger,W., Schober,E., *et al.* (2011) Diabetes and neurodegeneration in Wolfram syndrome: a
689 multicenter study of phenotype and genotype. *Diabetes Care*, **34**, 1503–1510.

690 14. Chaussenot,A., Bannwarth,S., Rouzier,C., Vialettes,B., Mkadem,S.A.E., Chabrol,B., Cano,A., Labauge,P.
691 and Paquis-Flucklinger,V. (2011) Neurologic features and genotype-phenotype correlation in
692 Wolfram syndrome. *Ann Neurol*, **69**, 501–508.

693 15. Cano,A., Rouzier,C., Monnot,S., Chabrol,B., Conrath,J., Lecomte,P., Delobel,B., Boileau,P., Valero,R.,
694 Procaccio,V., *et al.* (2007) Identification of novel mutations in WFS1 and genotype-phenotype
695 correlation in Wolfram syndrome. *Am J Med Genet A*, **143A**, 1605–1612.

696 16. de Heredia,M.L., Clèries,R. and Nunes,V. (2013) Genotypic classification of patients with Wolfram
697 syndrome: insights into the natural history of the disease and correlation with phenotype. *Genet*
698 *Med*, **15**, 497–506.

699 17. Delvecchio,M., Iacoviello,M., Pantaleo,A. and Resta,N. (2021) Clinical Spectrum Associated with
700 Wolfram Syndrome Type 1 and Type 2: A Review on Genotype-Phenotype Correlations. *Int J*
701 *Environ Res Public Health*, **18**, 4796.

702 18. Lee,E.M., Verma,M., Palaniappan,N., Pope,E.M., Lee,S., Blacher,L., Neerumalla,P., An,W., Campbell,T.,
703 Brown,C., *et al.* (2023) Genotype and clinical characteristics of patients with Wolfram syndrome
704 and WFS1-related disorders. *Front Genet*, **14**, 1198171.

705 19. Kitamura,R.A., Maxwell,K.G., Ye,W., Kries,K., Brown,C.M., Augsornworawat,P., Hirsch,Y.,
706 Johansson,M.M., Weiden,T., Ekstein,J., *et al.* (2022) Multidimensional analysis and therapeutic
707 development using patient iPSC–derived disease models of Wolfram syndrome. *JCI Insight*, **7**.

708 20. Squitti,R., Cerchiaro,G., Giovannoni,I., Francalanci,P., Siotto,M., Maffei,P., Ricordi,C. and
709 Rongioletti,M.C. (2019) A case of a mild Wolfram Syndrome with concomitant ATP7B mutation.
710 *CellR4 Repair Replace Regen Reprogram*, **7**, e2735.

711 21. Panfili,E., Mondanelli,G., Orabona,C., Belladonna,M.L., Gargaro,M., Fallarino,F., Orecchini,E.,
712 Prontera,P., Proietti,E., Frontino,G., *et al.* (2021) Novel mutations in the WFS1 gene are associated
713 with Wolfram syndrome and systemic inflammation. *Hum Mol Genet*, **30**, 265–276.

714 22. Pellegrini,S., Manenti,F., Chimienti,R., Nano,R., Ottoboni,L., Ruffini,F., Martino,G., Ravassard,P.,
715 Piemonti,L. and Sordi,V. (2018) Differentiation of Sendai Virus-Reprogrammed iPSC into β Cells,
716 Compared with Human Pancreatic Islets and Immortalized β Cell Line. *Cell Transplantation*, **27**,
717 1548–1560.

718 23. Chimienti,R., Baccega,T., Torchio,S., Lombardo,A., Malnati,M., Correspondence,L.P., Manenti,F.,
719 Pellegrini,S., Cospito,A., Amabile,A., *et al.* (2022) Engineering of immune checkpoints B7-H3 and
720 CD155 enhances immune compatibility of MHC-I−/− iPSCs for β cell replacement. *Cell Reports*, **40**,
721 111423.

722 24. Pellegrini,S., Pipitone,G.B., Cospito,A., Manenti,F., Poggi,G., Lombardo,M.T., Nano,R., Martino,G.,
723 Ferrari,M., Carrera,P., *et al.* (2021) Generation of β Cells from iPSC of a MODY8 Patient with a Novel
724 Mutation in the Carboxyl Ester Lipase (CEL) Gene. *The Journal of Clinical Endocrinology &*
725 *Metabolism*, 10.1210/clinem/dgaa986.

726 25. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C.,
727 Saalfeld,S., Schmid,B., *et al.* (2012) Fiji: An open-source platform for biological-image analysis.
728 *Nature Methods*, **9**, 676–682.

729 26. Yeo,G. and Burge,C.B. (2004) Maximum entropy modeling of short sequence motifs with applications to
730 RNA splicing signals. *J Comput Biol*, **11**, 377–394.

731 27. Kim,D.E., Chivian,D., Malmström,L. and Baker,D. (2005) Automated prediction of domain boundaries in
732 CASP6 targets using Ginzu and RosettaDOM. *Proteins*, **61 Suppl 7**, 193–200.

733 28. Kim,D.E., Chivian,D. and Baker,D. (2004) Protein structure prediction and analysis using the Robetta
734 server. *Nucleic Acids Research*, **32**, W526.

735 29. Van Der Spoel,D., Lindahl,E., Hess,B., Groenhof,G., Mark,A.E. and Berendsen,H.J.C. (2005) GROMACS:
736 fast, flexible, and free. *J Comput Chem*, **26**, 1701–1718.

737 30. Frontino,G., Raouf,T., Canarutto,D., Tirelli,E., Di Tonno,R., Rigamonti,A., Cascavilla,L.M., Baldoli,C.,
738 Scotti,R., Bonfanti,R., *et al.* (2021) Off-label Liraglutide use in children with Wolfram Syndrome type
739 1: extensive characterization of four patients.

740 31. Richards,S., Aziz,N., Bale,S., Bick,D., Das,S., Gastier-Foster,J., Grody,W.W., Hegde,M., Lyon,E., Spector,E.,
741 *et al.* (2015) Standards and Guidelines for the Interpretation of Sequence Variants: A Joint
742 Consensus Recommendation of the American College of Medical Genetics and Genomics and the
743 Association for Molecular Pathology. *Genetics in medicine : official journal of the American College*
744 *of Medical Genetics*, **17**, 405.

745 32. Astuti,D., Sabir,A., Fulton,P., Zatyka,M., Williams,D., Hardy,C., Milan,G., Favaretto,F., Yu-Wai-Man,P.,
746 Rohayem,J., *et al.* (2017) Monogenic diabetes syndromes: Locus-specific databases for Alström,
747 Wolfram, and Thiamine-responsive megaloblastic anemia. *Human Mutation*, **38**, 764.

748 33. Sandhu,M.S., Weedon,M.N., Fawcett,K.A., Wasson,J., Debenham,L., Daly,A., Lango,H., Frayling,T.M.,
749 Neumann,R.J., Sherva,R., *et al.* (2009) Common variants in WFS1 confer risk of type 2 diabetes.
750 *Europe PMC Funders Group*, **39**, 951–953.

751 34. Yurimoto,S., Hatano,N., Tsuchiya,M., Kato,K., Fujimoto,T., Masaki,T., Kobayashi,R. and Tokumitsu,H.
752 (2009) Identification and characterization of wolframin, the product of the wolfram syndrome gene
753 (WFS1), as a novel calmodulin-binding protein. *Biochemistry*, **48**, 3946–3955.

754 35. Angebault,C., Fauconnier,J., Patergnani,S., Rieusset,J., Danese,A., Affortit,C.A., Jagodzinska,J., Mégy,C.,
755 Quiles,M., Cazevieille,C., *et al.* (2018) ER-mitochondria cross-talk is regulated by the Ca2+ sensor
756 NCS1 and is impaired in Wolfram syndrome. *Sci Signal*, **11**, eaaq1380.

757 36. Yoshida,H., Matsui,T., Yamamoto,A., Okada,T. and Mori,K. (2001) XBP1 mRNA Is Induced by ATF6 and
758 Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor. *Cell*, **107**,
759 881–891.

760 37. Li,Z., Vuong,J.K., Zhang,M., Stork,C. and Zheng,S. (2017) Inhibition of nonsense-mediated RNA decay by
761 ER stress. *Rna*, **23**, 378–394.

762 38. Usuki,F., Yamashita,A. and Fujimura,M. (2019) Environmental stresses suppress nonsense-mediated
763 mRNA decay (NMD) and affect cells by stabilizing NMD-targeted gene expression. *Scientific reports*,
764 **9**.

765 39. Urano,F. (2016) Wolfram Syndrome: Diagnosis, Management, and Treatment. *Curr Diab Rep*, **16**, 6.

766 40. Hu,K., Zatyka,M., Astuti,D., Beer,N., Dias,R.P., Kulkarni,A., Ainsworth,J., Wright,B., Majander,A., Yu-Wai-
767 Man,P., *et al.* (2022) WFS1 protein expression correlates with clinical progression of optic atrophy
768 in patients with Wolfram syndrome. *J Med Genet*, **59**, 65–74.

769 41. Majander,A., Jurkute,N., Burté,F., Brock,K., João,C., Huang,H., Neveu,M.M., Chan,C.M., Duncan,H.J.,
770 Kelly,S., *et al.* (2022) WFS1-Associated Optic Neuropathy: Genotype-Phenotype Correlations and
771 Disease Progression. *Am J Ophthalmol*, **241**, 9–27.

772 42. Ishihara,H., Takeda,S., Tamura,A., Takahashi,R., Yamaguchi,S., Takei,D., Yamada,T., Inoue,H., Soga,H.,
773 Katagiri,H., *et al.* (2004) Disruption of the WFS1 gene in mice causes progressive β-cell loss and
774 impaired stimulus - Secretion coupling in insulin secretion. *Human Molecular Genetics*, **13**, 1159–
775 1170.

776 43. Plaas,M., Seppa,K., Reimets,R., Jagomäe,T., Toots,M., Koppel,T., Vallisoo,T., Nigul,M., Heinla,I., Meier,R.,
777 *et al.* (2017) Wfs1-deficient rats develop primary symptoms of Wolfram syndrome: Insulin-
778 dependent diabetes, optic nerve atrophy and medullary degeneration. *Scientific Reports*, **7**, 1–16.

779 44. Elli,F.M., Ghirardello,S., Giavoli,C., Gangi,S., Dioni,L., Crippa,M., Finelli,P., Bergamaschi,S., Mosca,F.,
780 Spada,A., *et al.* (2012) A new structural rearrangement associated to Wolfram syndrome in a child
781 with a partial phenotype. *Gene*, **509**, 168–172.

782 45. Giuliano,F., Bannwarth,S., Monnot,S., Cano,A., Chabrol,B., Vialettes,B., Delobel,B. and Paquis-
783 Flucklinger,V. (2005) Wolfram syndrome in French population: characterization of novel mutations
784 and polymorphisms in the WFS1 gene. *Human mutation*, **25**, 99–100.

785 46. Panfili,E., Mondanelli,G., Orabona,C., Belladonna,M.L., Gargaro,M., Fallarino,F., Orecchini,E.,
786 Prontera,P., Proietti,E., Frontino,G., *et al.* (2021) Novel mutations in the WFS1 gene are associated
787 with Wolfram syndrome and systemic inflammation. *Human Molecular Genetics*, **30**, 265–276.

788 47. Cutting,G.R. (2015) Cystic fibrosis genetics: from molecular understanding to clinical application. *Nat*
789 *Rev Genet*, **16**, 45–56.

790 48. Claustres,M., Thèze,C., des Georges,M., Baux,D., Girodon,E., Bienvenu,T., Audrezet,M.-P.,
791 Dugueperoux,I., Férec,C., Lalau,G., *et al.* (2017) CFTR-France, a national relational patient database
792 for sharing genetic and phenotypic data associated with rare CFTR variants. *Hum Mutat*, **38**, 1297–
793 1315.

794 49. Noel,S., Servel,N., Hatton,A., Golec,A., Rodrat,M., Ng,D.R.S., Li,H., Pranke,I., Hinzpeter,A., Edelman,A., *et*
795 *al.* (2022) Correlating genotype with phenotype using CFTR-mediated whole-cell Cl- currents in
796 human nasal epithelial cells. *J Physiol*, **600**, 1515–1531.

797 50. Ghiasi,S.M. and Rutter,G.A. (2021) Consequences for Pancreatic β-Cell Identity and Function of
798 Unregulated Transcript Processing. *Frontiers in Endocrinology*, **12**.

799 51. McGlincy,N.J. and Smith,C.W.J. (2008) Alternative splicing resulting in nonsense-mediated mRNA decay:
800 what is the meaning of nonsense? *Trends Biochem Sci*, **33**, 385–393.

801

802 **Figure 1.**

803 **(A-B)** Sanger sequencing chromatograms showing the heterozygous c.316-1G>A transition and the
804 c.757A>T transversion in *WFS1* gene of the patient. The modified nucleotides are highlighted in blue
805 and red, respectively, as well as the resulting aminoacid substitution in the protein sequence, if known.
806 **(C)** Genetic pedigree of the patient's family. The first line below each symbol represents generation
807 and identification number. The heterozygous mutations found in the patient's father and mother are
808 reported; the patient's brother carries no mutated allele. **(D)** Schematic representation of *WFS1* gene
809 *(top)* and Wolframin protein *(bottom)*. The exons 1 (non-coding, grey), 2-7 (encoding for the N-
810 terminal domain of wolframin, green boxes) and 8 (encoding for the transmembrane and C-terminal
811 domains of wolframin, orange box) are reported. Genomic position of patient's mutations are
812 highlighted and color-coded as in panels A-B. The previously reported ASS variants are reported as
813 well. The N-terminal domain (green) and transmembrane/C-terminal domains (orange) are showed
814 along with the c.757A>T transversion introducing a premature stop codon.

815 **Figure 2**

816 **(A)** Genomic positions of the natural and cryptic donor (green) and acceptor (red) splice sites along
817 the *WFS1* exon 4 (145bp) and its flanking regions (150bp/each) as predicted by Human Splice Finder.
818 Only the splice sites with an HSF score >75 *(top)* and a MaxEnt score >4 *(bottom)* are plotted. The
819 coordinates indicate the base number counting from the p-arm telomere of chromosome 4 (according
820 to the UCSC Genome Browser on Human GRCh38/hg38 Assembly). **(B)** Exonic splicing ratio (ESR)
821 calculated as ratio between auxiliary splicing signals (ESE and ESS) along the same genomic region.
822 The binding sites for the splicing factors SC35 (dark blue), SRp40 (purple) and 9G8 (light blue) are
823 noted for the region of interest surrounding the cryptic ASSs inside exon 4. **(C)** Position of the natural
824 and alternative branch points (light blue) along the analyzed genomic region. Only the branch points
825 with a score >65 are plotted. **(D)** Results of maximum entropy modelling in presence of the wild type
826 ASS *(left)* and after the c.316-1 G>A transition *(right)*. The two most likely alternatives upon the
827 canonical splice site disruption are reported. ASSs scored by the model are underlined and highlighted
828 in yellow. Exonic bases are in upper case letters; intronic bases in lower case letters. The grey area
829 represents spliced-out sequence. Coordinates refers to the nucleotide position in the coding sequence.

830 **Figure 3**

831 **(A)** Schematic representation of PCR strategy for the amplification of the WFS1 cDNA among exons
832 3, 4 and 5. Colored arrows show the mapping of the four primers along the WFS1 gene sequence.
833 The red cross indicates the mutated canonical ASS. The theoretical length of PCR products obtained
834 by performing amplification with different combinations of the indicated primers is reported. **(B)**

835 Representative PCR results for the indicated primer pairs, performed on WT and five patient-derived
836 clones. The red rectangle highlights the ~180bp-long amplicon obtained with Ex3fw-Ex5rev primer
837 pair. **(C)** Electropherograms from Sanger sequencing of WFS1-deriving PCR products. **(D)** Agarose
838 gel electrophoresis of amplification with Ex3fw-Ex5rev from WT donor- and WFS1 patient-derived
839 PBMCs. The additional amplicons, including the 180bp-length amplicon observed in WFS1 iPSCs,
840 are indicated by red triangles. **(E)** High-trhoughput evaluation of RNA-seq data quality and goodness
841 of alignment: *(top, left)* the *sequence quality* graph, reporting the mean quality value (Phred score)
842 across each base position in the R1 (green) and R2 (orange) reads; *(bottom, left)* the *per sequence*
843 *quality score* graph, reporting the number of reads with average quality scores. Red, yellow and green
844 backgrounds highlight the Phred score ranges of bad, poor and good quality, respectively; *(top, right)*
845 *STAR alignment score*, reporting the number of mapped/unmapped reads per sample; *(bottom, right)*
846 percentage of reads per sample after adapter removal by Trimmomatic. **(F)** Percentage of NGS reads
847 aligned to the wild type or the four alternative splice isoform consensus sequences in WT *(top)* and
848 WFS1 *(bottom)* iPSCs. Alignment data for WFS1 are reported as mean of the two analyzed clones.
849 **(G)** Schematic representation of the *WFS1* alternative splicing variants found by NGS, showing the
850 conserved exonic sequence (green) and spliced-out sequence (grey). Sequences at the bottom of each
851 structure zoom at the donor/acceptor alternative splice sites. **(H)** RT-qPCR of total *WFS1* mRNA, in
852 WT and WFS1, and of the **(I)** four isoforms, in WFS1 iPSCs after treatment with NMDI-14 or vehicle
853 (untreated). Data are plotted as mean±SEM. N = 3; ** $p < 0.001$, **** $p < 0.0001$.

854 **<u>Figure 4</u>**

855 **(A-B)** Immunoblot analysis of Wolframin in WT and WFS1 iPSC clones recognized by polyclonal
856 antibody targeting the C- and N-terminal domains. Relative quantification of Wolframin expression
857 is reported on the right. N = 6. **** $p < 0.0001$. **(C-D)** The three-dimensional models of WT (residues
858 1-402), and mutated p.Val106-Arg152del (residues 1-355) and p.Val91-Thr170del (residues 1-322)
859 N-terminal domains, including the first two transmembrane regions (TMR). Proteins are represented
860 as ribbons. Region lost in mutated isoforms (green), the mutation-induced conformational changes
861 (blue) and the unaltered portion of the domain (red) are highlighted. In the pictures displaying the
862 mutated isoforms, magnifications show the superimposed view of TMR alpha helices in WT (light
863 blue) and p.Val106-Arg152del (red) *(top, right)*, and the superimposed view of WFS1-CaM binding
864 domain (yellow) loss in p.Val91-Thr170del *(bottom, right)*. In p.Val106-Arg152del, the partially
865 conserved WFS1-CaM binding interface is shown as Connolly surface. **(E-F)** Results of molecular
866 dynamics simulation, reporting RMSD of protein backbone atoms and radius of gyration (Rg) of the

867 WT (grey) and mutated (orange/purple) N-terminal domains. Flattening of the RMSD and Rg plots
868 of protein was observed around 10 ns at 300 K.

869 **Figure 5**

870 **(A)** Schematic representation of the c.316-1G>A point mutation correction strategy based on the
871 Cas9/gRNA-mediated targeting of *WFS1* exon 4. Mapping of the ssODN (blue) and the fw/rev primer
872 pair (dark/light orange) used for screening of the homology-directed repair is reported. Magnification
873 highlights the comparison between the pathogenic variant containing the c.316-1G>A mutation with
874 the ssODN sequence re-establishing the natural ASS upstream the exon 4. PAM sequence and gRNA
875 recognition site are emphasized in dark green. Silent mutations and SacI cut site in ssODN sequence
876 are reported in light blue and framed by pink rectangle, respectively. **(B)** Screening by SacI restriction
877 enzyme digestion of GFP-positive iPSC clones. **(C)** Immunoblot using the anti-WFS1 N-terminal
878 antibody to assess re-expression of full-length protein in the gene edited $WFS1^{wt/757A>T}$ iPSCs *(left)*.
879 Relative quantification of Wolframin in WFS1 and $WFS1^{wt/757A>T}$ iPSCs expressed as mean±SEM. N
880 = 3 for WFS1, N = 4 for $WFS1^{wt/757A>T}$ *(right)*. **(D)** RT-qPCR of total *WFS1* mRNA in WFS1 and
881 $WFS1^{wt/757A>T}$ iPSCs reported as mean±SEM. N = 4 for WFS1, N = 5 for $WFS1^{wt/757A>T}$; **(E)** RT-
882 qPCR of alternative splicing isoforms in WFS1 and $WFS1^{wt/757A>T}$ iPSCs expressed as mean±SEM.
883 N = 3. *** $p < 0.001$, **** $p < 0.0001$.

884 **Figure 6**

885 **(A)** Representative immunoblot showing Wolframin protein stained with the anti-WFS1 C-terminal
886 antibody before (day 0) and after (day 24) iPSCs differentiation into β cells. Monomeric form (100
887 kDa) and homotetrameric form (> 250 kDa) are indicated by orange and red triangles, respectively.
888 Signal was acquired at different exposure time to detect homotetramers. **(B)** Relative quantification
889 of Wolframin in WFS1 and $WFS1^{wt/757A>T}$ iPSCs (0) and iβ (24) expressed as mean±SEM. N=3 **(C)**
890 RT-qPCR on total *WFS1* mRNA in WFS1 and $WFS1^{wt/757A>T}$ iPSCs (0) versus iβ (24) reported as
891 mean±SEM. N = 4 for WFS1, N = 5 for $WFS1^{wt/757A>T}$. **(D)** Quantification of the alternative splicing
892 isoforms in iPSCs and iβ expressed as mean±SEM. N = 3. * $p < 0.05$, *** $p < 0.001$, **** $p < 0.0001$.

893 **Figure 7**

894 **(A)** Expression of *ATF4*, *HSPA5*, and *DDIT3* genes in WFS1 and $WFS1^{wt/757A>T}$ iβ, before and after
895 50nM TG treatment. Data are plotted as mean±SEM, N = 4. **(B)** Annexin V-positive cell percentages
896 in WFS1 and $WFS1^{wt/757A>T}$ iPSCs during treatment with 100 nM TG and in the recovery period with
897 fresh medium w/o TG. Data are plotted as mean±SEM. N = 3. **(C)** Time course response of XBP-1
898 splicing upon TG-induced stress in WFS1 and $WFS1^{wt/757A>T}$ iPSCs. Data are plotted as mean±SD. N

899 = 3. **(D-E)** Early and late apoptosis was measured as percentage of Annexin V and P.I.-positive cells,
900 respectively, in WFS1 and WFS1$^{wt/757A>T}$ iβ after 50 nM TG or inflammatory cytokine (50 U/mL IL-
901 1β, 1000 U/mL IFN-γ and 10 ng/mL TNF-α) exposure. Data are plotted as mean±SD. N = 3 for
902 WFS1, N = 4 for WFS1$^{wt/757A>T}$. **(F)** Gene expression of the splicing factor UPF1 and SMG7 in WFS1
903 and WFS1$^{wt/757A>T}$ iβ upon stress induction. Data are expressed as fold change over untreated control
904 *(dotted line)*. Asterisks indicate statistical difference compared to untreated cells. Error bar represents
905 SEM. N = 3. **(G)** Relative quantification of alternative splicing isoforms in WFS1 iβ before and after
906 stress induction. Data are plotted as mean±SEM. N = 3. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ****
907 $p < 0.0001$.

908 **Figure S1**

909 **(A)** Representative image of WFS1 iPSC colony (4x magnification). **(B)** Representative karyotype of
910 stabilized WFS1 iPSCs. **(C)** Immunofluorescence analysis of pluripotency markers (SOX2, NANOG,
911 OCT4 and SSEA4) in WFS1 iPSCs. The nuclei were counterstained with Hoechst. Scale bar 40µm.

912 **Figure S2**

913 **(A)** Representative image of WFS1$^{wt/757A>T}$ iPSC colonies (10x magnification). **(B)** Representative
914 karyotype of stabilized WFS1$^{wt/757A>T}$ iPSCs. **(C)** Immunofluorescence analysis of the pluripotency
915 markers (SOX2, NANOG and OCT4). Nuclei were counterstained with Hoechst. Scale bar 40µm.
916 **(D)** Percentage of cells positive for pancreatic markers (PDX1, NKX6.1 and INS) after 24 days of
917 pancreatic differentiation. Data are plotted as mean±SEM. N = 6 for WFS1, N = 9 for WFS1$^{wt/757A>T}$.

918 **Table 1**

919 List of PCR primers and other oligonucleotides employed in the study.

920 **Table 2**

921 List of antibodies employed in the study and respective mode of use.

922 **Table 3**

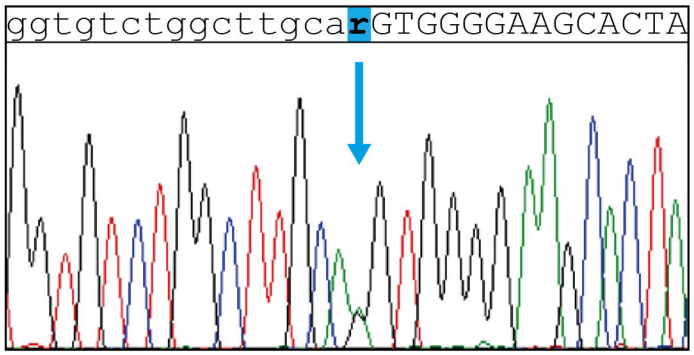
923 List of natural and cryptic donor and acceptor splicing sites, ESE/ESS signals and branch points along
924 the WFS1 exon 4 and its 150bp-long flanking regions.

925 **Table 4**

926 Consensus sequences of alternative splicing variants from the c.316-1G>A-carrying allele as obtained
927 by multi-alignment of NGS reads.

Figure 1

A

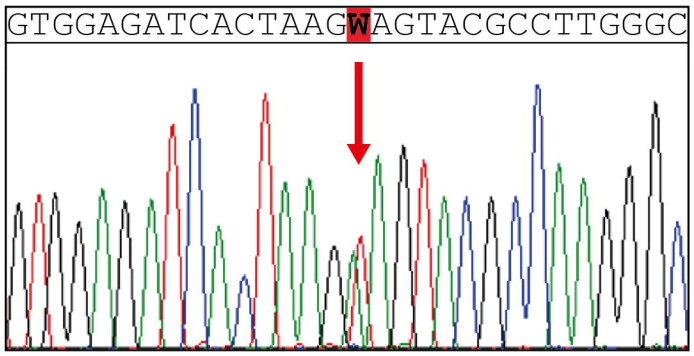


	-	-	-	-	-	106	107	108	109
Amino Acid	-	-	-	-	-	V	G	K	H
Nucleotide	gtg	tct	ggc	ttg	cag	GTG	GGG	AAG	CAC
	-15	-12	-9	-6	-3	316	319	322	325

c.316-1G>A ↓ p.?

	-	-	-	-	-	106	107	108	109
Amino Acid	-	-	-	-	-	V	G	K	H
Nucleotide	gtg	tct	ggc	ttg	caa	GTG	GGG	AAG	CAC
	-15	-12	-9	-6	-3	316	319	322	325

B

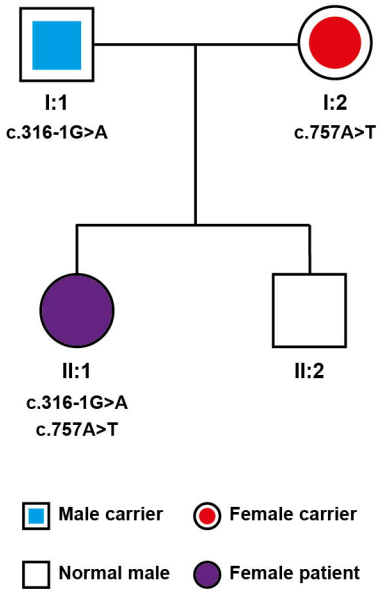


	249	250	251	252	253	254	255	256	257
Amino Acid	E	I	T	K	K	Y	A	K	G
Nucleotide	GAG	ATC	ACT	AAG	AAG	TAC	GCC	AAG	GGC
	645	648	751	754	757	760	763	766	769

c.757A>T ↓ p.253Lys>Ter

	249	250	251	252	253	254	255	256	257
Amino Acid	E	I	T	K	X	Y	A	K	G
Nucleotide	GAG	ATC	ACT	AAG	TAG	TAC	GCC	AAG	GGC
	645	648	751	754	757	760	763	766	769

C



D

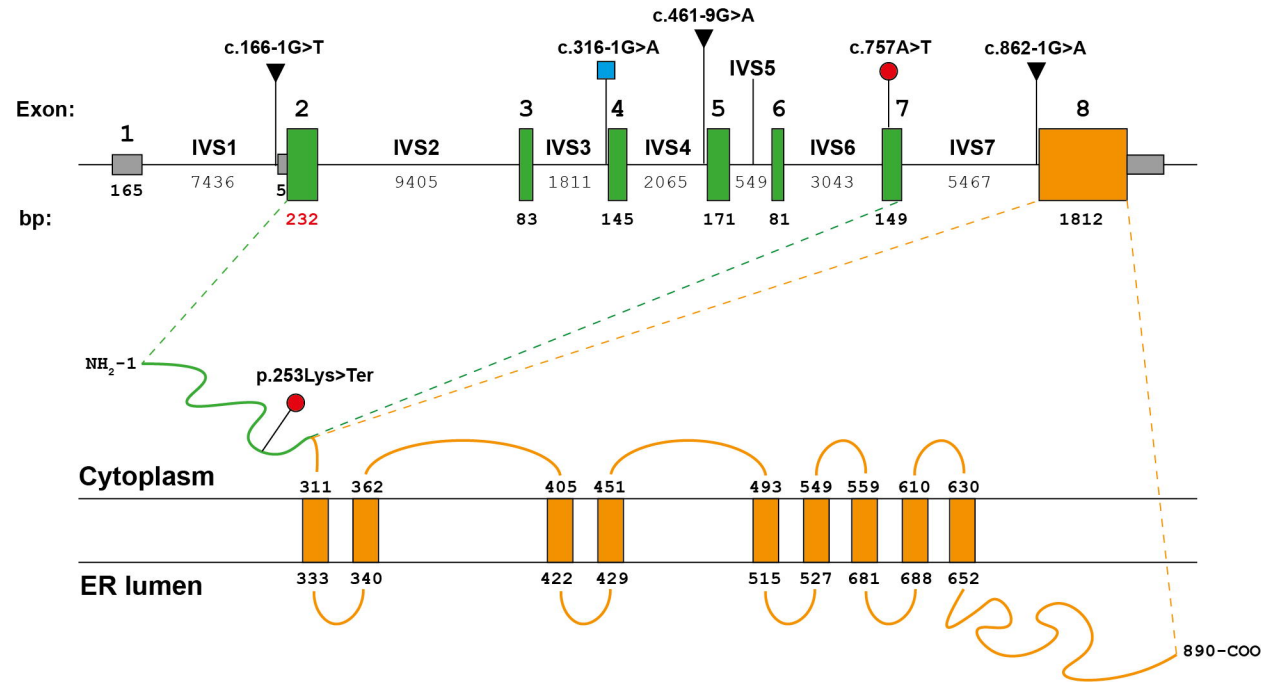


Figure 2

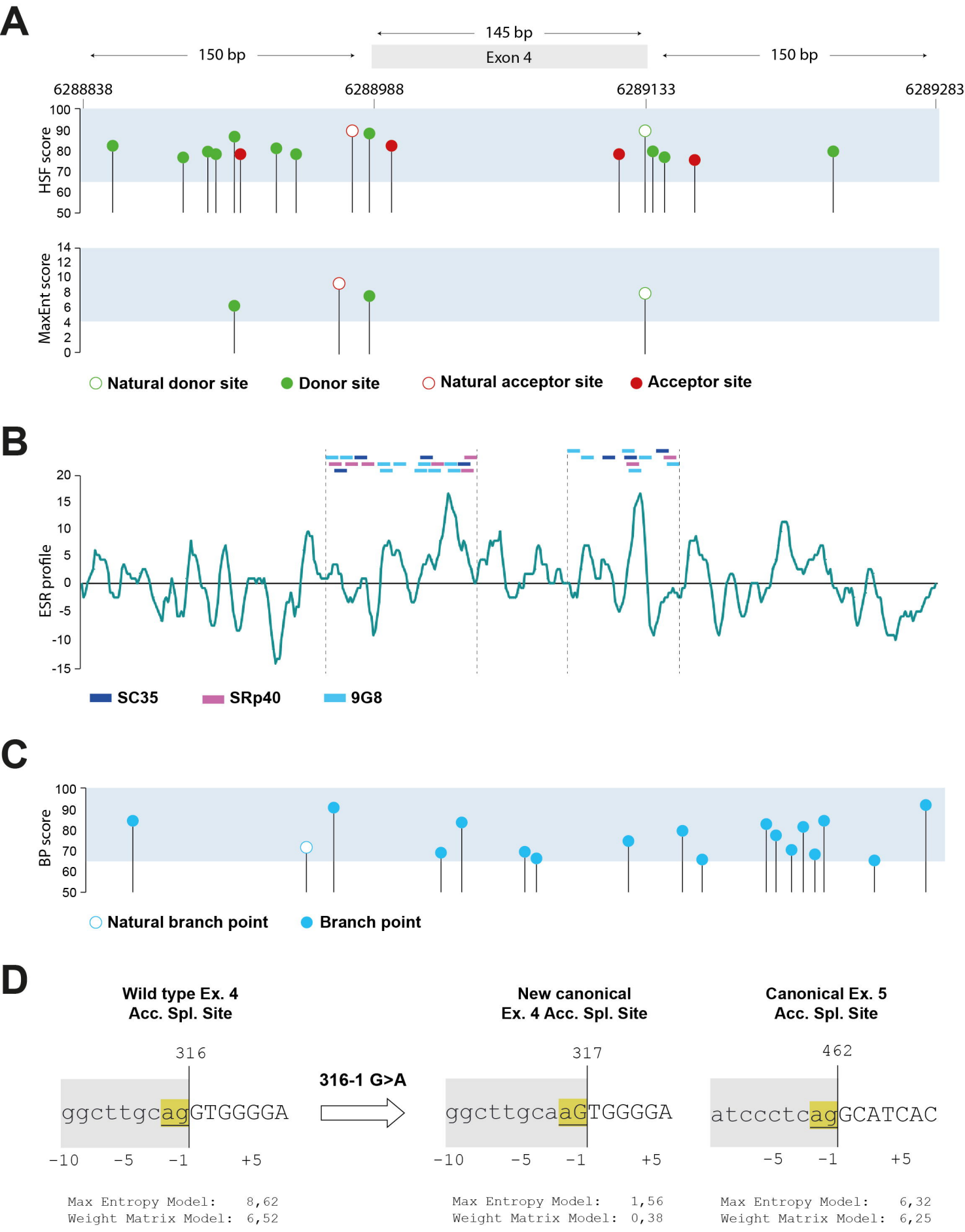

A
145 bp
150 bp
Exon 4
150 bp
6288838
6288988
6289133
6289283
HSF score
MaxEnt score
Natural donor site
Donor site
Natural acceptor site
Acceptor site
B
ESR profile
SC35
SRp40
9G8
C
BP score
Natural branch point
Branch point
D
Wild type Ex. 4
Acc. Spl. Site
316
ggcttgcagGTGGGGA
-10 -5 -1 +5
Max Entropy Model: 8,62
Weight Matrix Model: 6,52
316-1 G>A
New canonical
Ex. 4 Acc. Spl. Site
317
ggcttgcaaGTGGGGA
-10 -5 -1 +5
Max Entropy Model: 1,56
Weight Matrix Model: 0,38
Canonical Ex. 5
Acc. Spl. Site
462
atccctcagGCATCAC
-5 -1 +5
Max Entropy Model: 6,32
Weight Matrix Model: 6,25

Figure 3

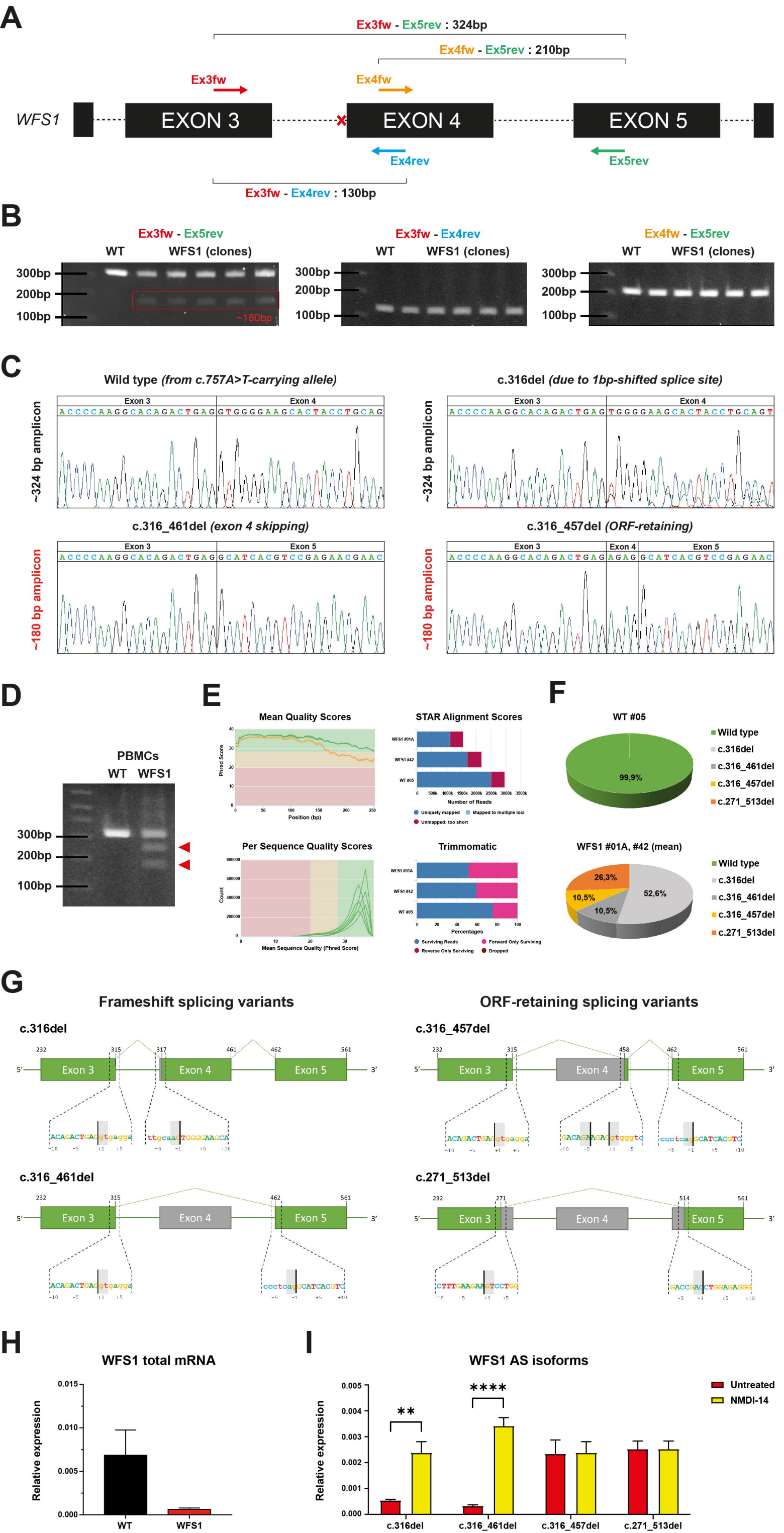

A
Ex3fw - Ex5rev : 324bp
Ex4fw - Ex5rev : 210bp
Ex3fw
Ex4fw
WFS1
EXON 3
EXON 4
EXON 5
Ex4rev
Ex5rev
Ex3fw - Ex4rev : 130bp
B
Ex3fw - Ex5rev
Ex3fw - Ex4rev
Ex4fw - Ex5rev
WT
WFS1 (clones)
300bp
200bp
100bp
~180bp
C
Wild type (from c.757A>T-carrying allele)
c.316del (due to 1bp-shifted splice site)
c.316_461del (exon 4 skipping)
c.316_457del (ORF-retaining)
~324 bp amplicon
~180 bp amplicon
Exon 3
Exon 4
Exon 5
D
PBMCs
WT WFS1
300bp
200bp
100bp
E
Mean Quality Scores
Phred Score
Position (bp)
Per Sequence Quality Scores
Count
Mean Sequence Quality (Phred Score)
STAR Alignment Scores
WFS1 #01A
WFS1 #42
WT #05
Number of Reads
Uniquely mapped
Mapped to multiple loci
Unmapped: too short
Trimmomatic
Percentages
Surviving Reads
Forward Only Surviving
Reverse Only Surviving
Dropped
F
WT #05
99,9%
Wild type
c.316del
c.316_461del
c.316_457del
c.271_513del
WFS1 #01A, #42 (mean)
26,3%
52,6%
10,5%
10,5%
G
Frameshift splicing variants
ORF-retaining splicing variants
c.316del
c.316_457del
c.316_461del
c.271_513del
5′
3′
H
WFS1 total mRNA
Relative expression
WT
WFS1
I
WFS1 AS isoforms
Untreated
NMDI-14
**
****
c.316del
c.316_461del
c.316_457del
c.271_513del

Figure 4

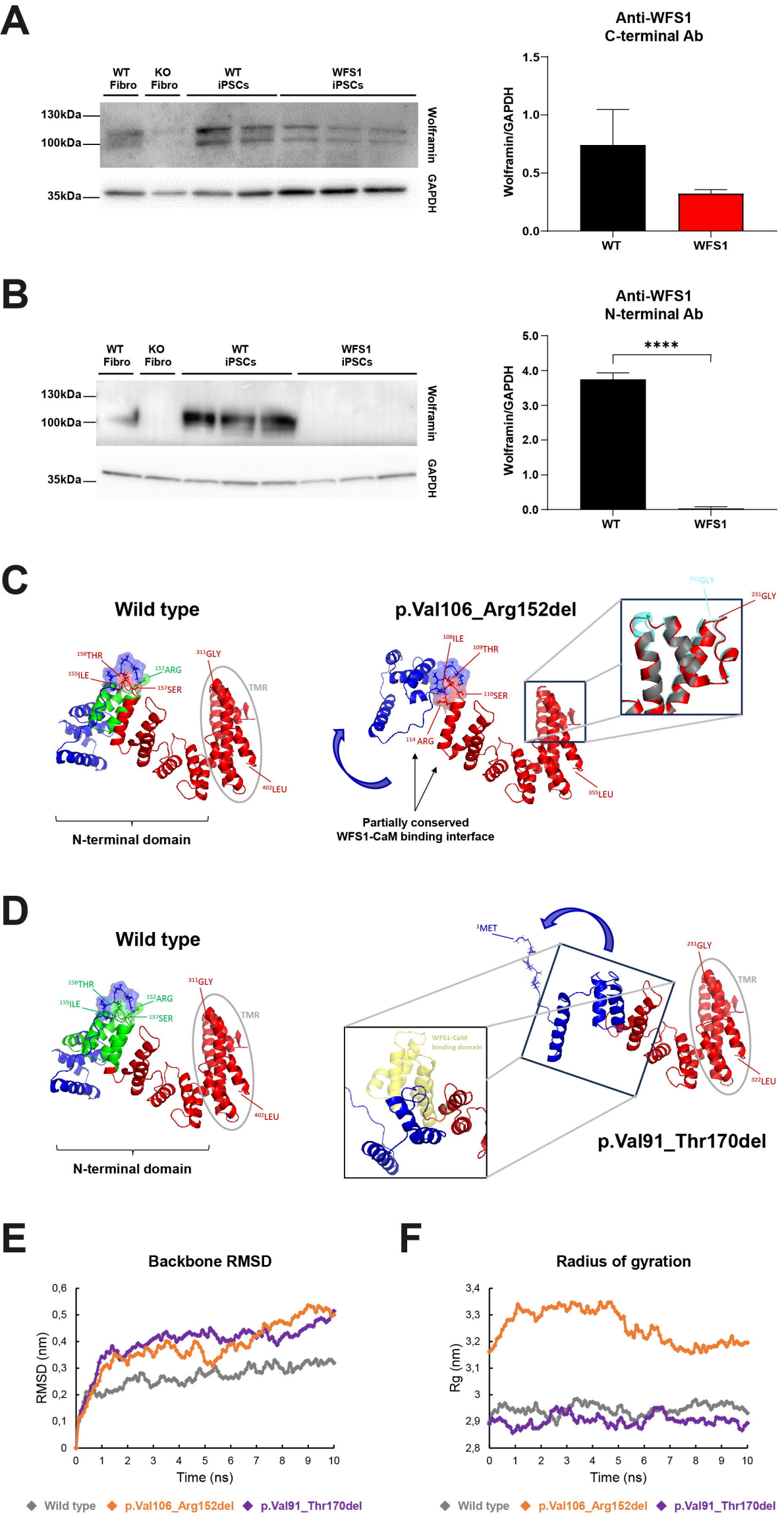

A
WT Fibro
KO Fibro
WT iPSCs
WFS1 iPSCs
130kDa
100kDa
35kDa
Wolframin
GAPDH
Anti-WFS1 C-terminal Ab
Wolframin/GAPDH
1.5
1.0
0.5
0.0
WT
WFS1
B
Anti-WFS1 N-terminal Ab
****
5.0
4.0
3.0
2.0
1.0
0.0
C
Wild type
p.Val106_Arg152del
156THR
155ILE
152ARG
157SER
311GLY
TMR
402LEU
N-terminal domain
108ILE
109THR
110SER
114ARG
355LEU
231GLY
Partially conserved WFS1-CaM binding interface
D
Wild type
1MET
WFS1-CaM binding domain
322LEU
p.Val91_Thr170del
E
Backbone RMSD
RMSD (nm)
Time (ns)
Wild type
p.Val106_Arg152del
p.Val91_Thr170del
F
Radius of gyration
Rg (nm)
Time (ns)
Wild type
p.Val106_Arg152del
p.Val91_Thr170del

Figure 5

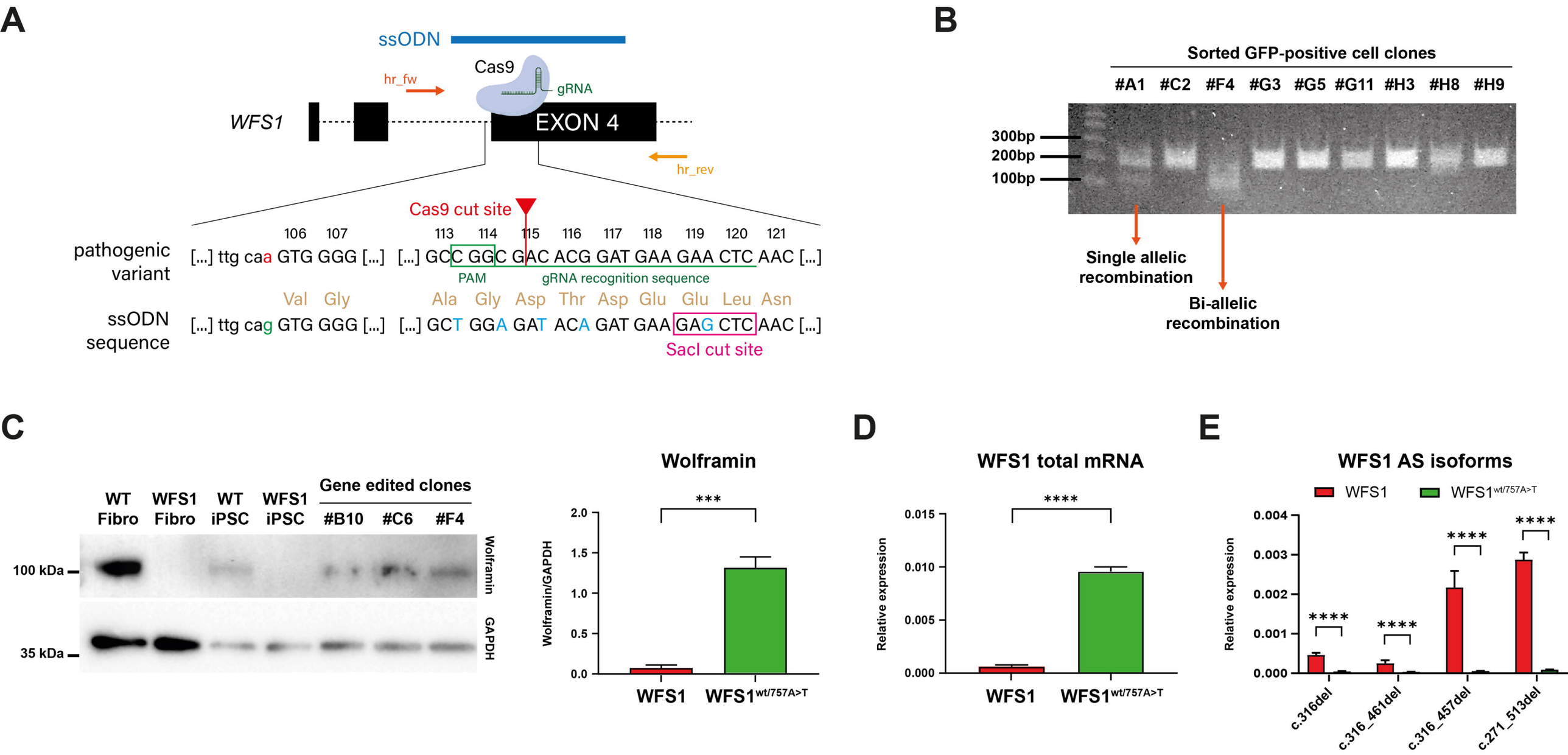

A
ssODN
Cas9
hr_fw
gRNA
WFS1
EXON 4
hr_rev
Cas9 cut site
106 107 113 114 115 116 117 118 119 120 121
pathogenic variant
[...] ttg caa GTG GGG [...] [...] GCC GGC GAC ACG GAT GAA GAA CTC AAC [...]
PAM
gRNA recognition sequence
Val Gly Ala Gly Asp Thr Asp Glu Glu Leu Asn
ssODN sequence
[...] ttg cag GTG GGG [...] [...] GCT GGA GAT ACA GAT GAA GAG CTC AAC [...]
SacI cut site
B
Sorted GFP-positive cell clones
#A1 #C2 #F4 #G3 #G5 #G11 #H3 #H8 #H9
300bp
200bp
100bp
Single allelic recombination
Bi-allelic recombination
C
WT Fibro
WFS1 Fibro
WT iPSC
WFS1 iPSC
Gene edited clones
#B10 #C6 #F4
100 kDa
35 kDa
Wolframin
GAPDH
Wolframin
Wolframin/GAPDH
***
2.0
1.5
1.0
0.5
0.0
WFS1
WFS1wt/757A>T
D
WFS1 total mRNA
Relative expression
****
0.015
0.010
0.005
0.000
WFS1
WFS1wt/757A>T
E
WFS1 AS isoforms
WFS1
WFS1wt/757A>T
Relative expression
0.004
0.003
0.002
0.001
0.000
****
****
****
****
c.316del
c.316_461del
c.316_457del
c.271_513del

Figure 6

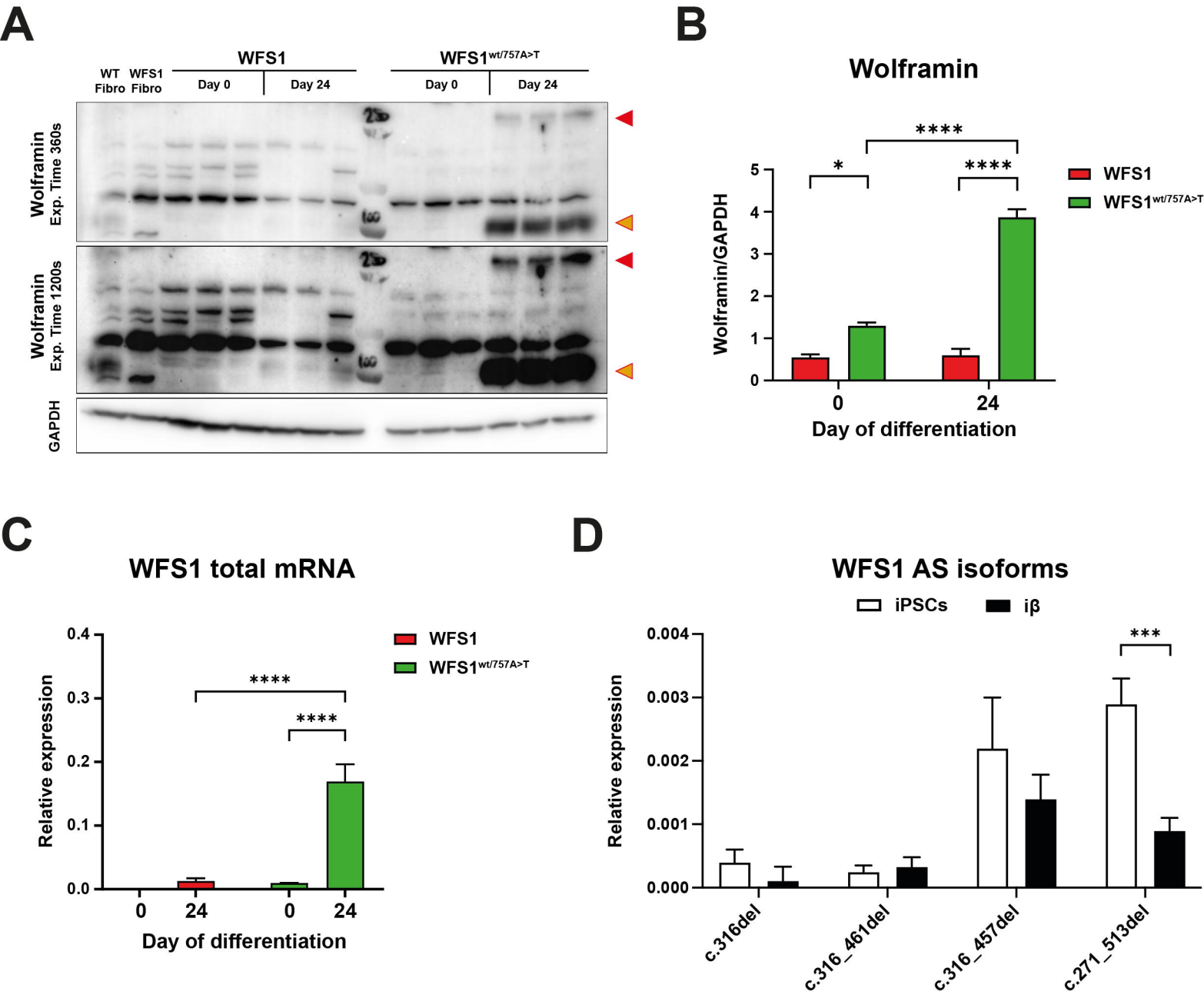

A
WFS1
WFS1wt/757A>T
WT Fibro
WFS1 Fibro
Day 0
Day 24
Day 0
Day 24
Wolframin Exp. Time 360s
Wolframin Exp. Time 1200s
GAPDH
B
Wolframin
Wolframin/GAPDH
Day of differentiation
WFS1
WFS1wt/757A>T
C
WFS1 total mRNA
Relative expression
Day of differentiation
WFS1
WFS1wt/757A>T
D
WFS1 AS isoforms
iPSCs
iβ
Relative expression
c.316del
c.316_461del
c.316_457del
c.271_513del

Figure 7

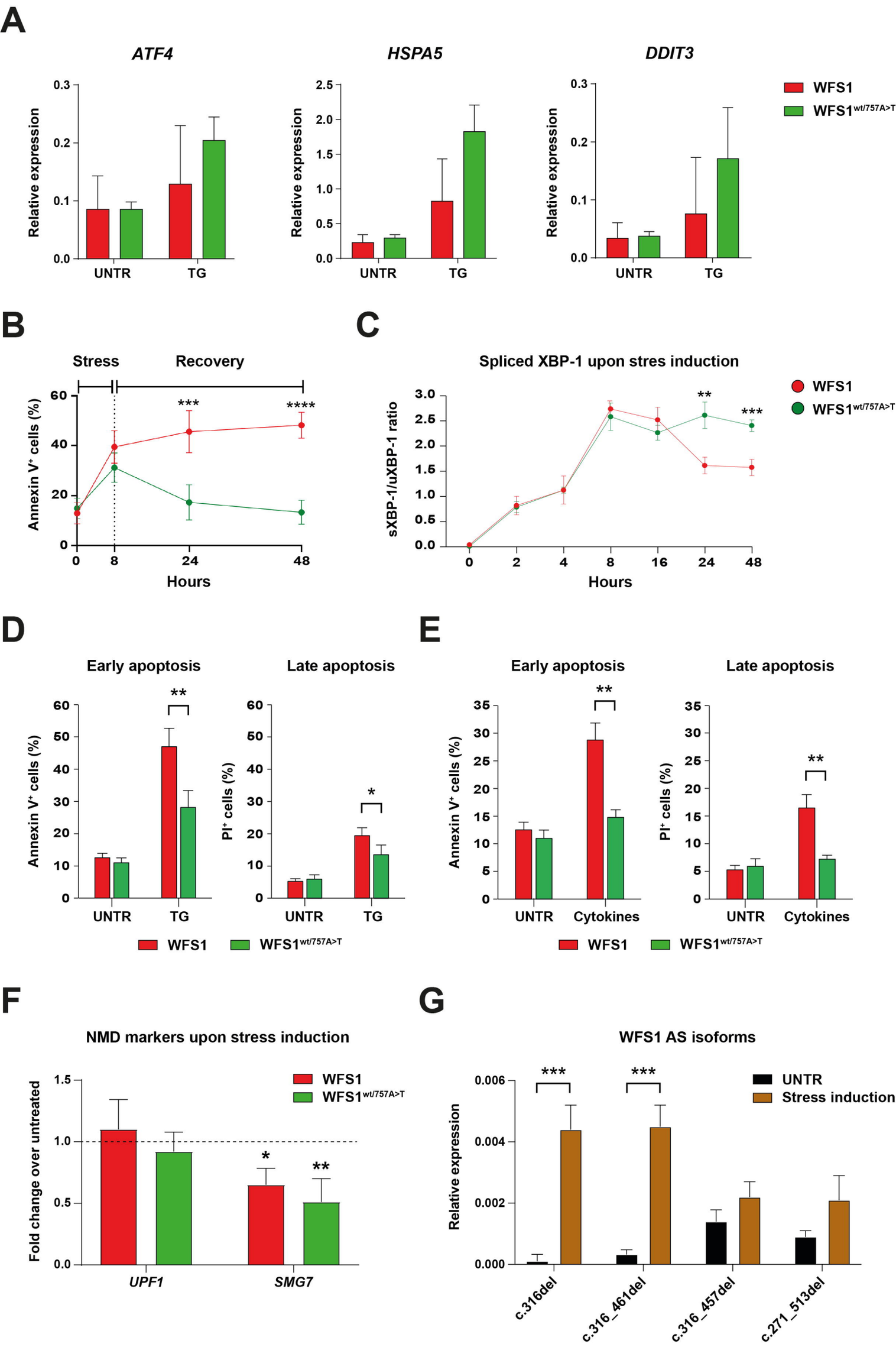

A
ATF4
HSPA5
DDIT3
Relative expression
UNTR
TG
WFS1
WFS1wt/757A>T
B
Stress
Recovery
Annexin V+ cells (%)
Hours
C
Spliced XBP-1 upon stres induction
sXBP-1/uXBP-1 ratio
Hours
D
Early apoptosis
Late apoptosis
Annexin V+ cells (%)
PI+ cells (%)
E
Early apoptosis
Late apoptosis
Cytokines
F
NMD markers upon stress induction
Fold change over untreated
UPF1
SMG7
G
WFS1 AS isoforms
Relative expression
UNTR
Stress induction
c.316del
c.316_461del
c.316_457del
c.271_513del

Table 1

Oligo ID	Sequence
ATF4_fw	GTTCTCCAGCGACAAGGCTA
ATF4_fw	ATCCTGCTTGCTGTTGTTGG
HSPA5_fw	TGTTCAACCAATTATCAGCAAACTC
HSPA5_rev	TTCTGCTGTATCCTCTTCACCAGT
DDIT3_fw	AGAACCAGGAAACGGAAACAGA
DDIT3_rev	TCTCCTTCATGCGCTGCTTT
GAPDH_fw	GTCTCCTCTGACTTCAACAGCG
GAPDH_rev	ACCACCCTGTTGCTGTAGCCAA
XBP-1_fw	CTGCCAGAGATCGAAAGAAGGC
XBP-1_rev	CTCCTGGTTCTCAACTACAAGGC
SMG7_fw	TTTCAGGAGGCAGTGGTGGATG
SMG7_rev	CAAACTCCTCTGGAAGTGGTGTC
UPF1_fw	AACGAGCACCAAGGCATTGGCT
UPF1_rev	GGCTGCTTTGATAGTGCCTTCG
c.316del_fw	CCCAAGGCACAGACTGAGGG
c.316del_rev	CGACATTCTCCAGCAGCTC
c.316del_461del_fw	CCCAAGGCACAGACTGAGGC
c.316del_461del_rev	CGACATTCTCCAGCAGCTC
c.316del_457del_fw	CTGGCCCTGGTGTTAGAGAC
c.316del_457del_rev	TCGGACGTGATGCCTCTCTC
c.271del_513del_fw	AGCAGGAGAGGAGCGAAAG
c.271del_513del_rev	CGCACGGCCCTCTCCAGTTC
WFS1_Ex3fw	GGGCCTACAAAGGGAGACAT
WFS1_Ex4fw	GGCGACACGGATTGAAGAACT
WFS1_Ex4rev	AGTTCTTCATCCGTGTCGCC
WFS1_Ex5rev	CCAGTACATGACCAGGGCTG
WFS1_gRNA_F1	TAATACGACTCACTATAGGAGTTCTTCATCCGTG
WFS1_gRNA_R1	TTCTAGCTCTAAAACGCGACACGGATGAAGAAC

WFS1_hr_fw	CTTCCTCCTCACCCAGCCTG
WFS1_hr_rev	GGGTCTTGGTCACTCACCTT
WFS1_ssODN	GAAAGGAGGTGGGCTGGCAGGGAGCATGGGGTGGGAGAGGGTCGGAGAA TCTGGAGGCTGACTGGTGTCTGGCTTGCAGGTGGGGAAGCACTACCTGCAG TTGGCTGGAGATACAGATGAAGAGCTCAACAGCTGCACCGCTGTGGACTG GCTGGTCCTCGCCGCGAAGCAGGGCCGTCGCGAGGCTGTGAAGCTGCTTC

Table 2

Antibody	Species	Supplier	Cat Nr.	Use	Dilution
Anti-GAPDH [6C5]	Mouse	Abcam	ab8245	WB	1:5000
Alexa Fluor® 647 anti-Insulin [T56-706]	Mouse	BD	565689	FACS	2µl x 10^6 cells
Anti-Nanog	Rabbit	R&D	AF1997	IF	1:100
PE anti-Nkx6.1 [R11-560]	Mouse	BD	563023	FACS	5µl x 10^6 cells
Anti-Oct4 [486]	Mouse	Novus	NBP2-15052	IF	1:100
Alexa Fluor® 488 anti-PDX1 [658A5]	Mouse	BD	562274	FACS	5µl x 10^6 cells
Anti-SSEA4 [MC813-70]	Mouse	Abcam	ab16287	IF	1:500
Anti-Sox2 [245610]	Mouse	R&D	MAB2018	IF	1:100
Anti-WFS1 C-terminal	Sheep	R&D	AF7417	WB	1:200
Anti-WFS1 N-terminal	Rabbit	Cell Signalling	8749	WB	1:1000
HRP-conjugated anti-mouse	Goat	R&D	HAF007	WB	1:1000
HRP-conjugated anti-rabbit	Goat	R&D	HAF008	WB	1:1000
HRP-conjugated anti-sheep	Donkey	R&D	HAF016	WB	1:1000
Alexa Fluor™ Plus 488 anti-mouse	Goat	Thermo Fisher	A32723	IF	1:500
Alexa Fluor™ Plus 488 anti-rabbit	Goat	Thermo Fisher	A32731	IF	1:500
Alexa Fluor™ Plus 488 anti-goat	Donkey	Thermo Fisher	A32814	IF	1:500
Alexa Fluor™ 546 anti-mouse	Goat	Thermo Fisher	A11030	IF	1:500
Alexa Fluor™ 546 anti-goat	Rabbit	Thermo Fisher	A21085	IF	1:500

Table 3

HSF_Signals			
Type	**Position**	**Motif**	**Value**
Donor splice site	6288849	CTGGTGTGA	81,53
Donor splice site	6288851	GGTGTGACC	66,49
Donor splice site	6288886	TGGGTGACA	76,86
Acceptor splice site	6288886	TGGGTGACAAAGGG	69,99
Donor splice site	6288895	AAGGGAAGT	72,72
Donor splice site	6288899	GAAGTGGGT	78,76
Acceptor splice site	6288901	AGTGGGTGAAAGGA	68,74
Donor splice site	6288903	TGGGTGAAA	77,31
Donor splice site	6288913	GAGGTGGGC	85,92
Acceptor splice site	6288916	GTGGGCTGGCAGGG	77,91
Acceptor splice site	6288933	ATGGGGTGGGAGAG	66,2
Donor splice site	6288935	GGGGTGGGA	80,07
Donor splice site	6288945	AGGGTCGGA	77,67
Acceptor splice site	6288952	GAGAATCTGGAGGC	69,92
Donor splice site	6288969	CTGGTGTCT	74,19
Donor splice site	6288971	GGTGTCTGG	66,26
(WT) Acceptor splice site	6288975	TCTGGCTTGCAGGT	88,95
Donor splice site	6288984	CAGGTGGGG	87,37
Acceptor splice site	6288996	CACTACCTGCAGTT	81,29
Donor splice site	6289004	GCAGTTGGC	66,32
Acceptor splice site	6289018	ACACGGATGAAGAA	66,66
Acceptor splice site	6289028	AGAACTCAACAGCT	72,98
Donor splice site	6289059	CTGGTCCTC	67,66
Acceptor splice site	6289065	CTCGCCGCGAAGCA	71,7
Acceptor splice site	6289068	GCCGCGAAGCAGGG	72,8
Acceptor splice site	6289080	GGCCGTCGCGAGGC	68,89
Acceptor splice site	6289089	GAGGCTGTGAAGCT	68,12
Donor splice site	6289092	GCTGTGAAG	67,2
Donor splice site	6289109	CCGGTGCTT	73,55
Acceptor splice site	6289115	CTTGGCGGACAGAA	78,06
(WT) Donor splice site	6289129	GAGGTGGGT	88,85
Donor splice site	6289133	TGGGTCTGT	78,98
Acceptor splice site	6289135	GGTCTGTGTGAGGC	72,8
Donor splice site	6289137	TCTGTGTGA	69,1
Donor splice site	6289139	TGTGTGAGG	76,44
Acceptor splice site	6289141	TGTGAGGCTTAGAA	72,09
Acceptor splice site	6289146	GGCTTAGAACAGCC	73,12
Acceptor splice site	6289155	CAGCCTCTGGAGGG	75,12
Acceptor splice site	6289162	TGGAGGGTTGAGCA	66,07
Donor splice site	6289165	AGGGTTGAG	68,21
Acceptor splice site	6289165	AGGGTTGAGCAGCT	71,81
Donor splice site	6289177	CTTGTAATG	71,43

Donor splice site	6289228	CTGGTGATG	79,1
Donor splice site	6289236	GCTGTTGGG	66,22
Acceptor splice site	6289242	GGGAAATTTCAGTT	74,65
Acceptor splice site	6289256	TCTGTTTTGCTGGT	65,33
Donor splice site	6289265	CTGGTGGCC	71,26
MaxEnt_Signals			
Type	**Position**	**Motif**	**Value**
Donor splice site	6288913	GAGGTGGGC	6,32
(WT) Acceptor splice site	6288967	GACTGGTGTCTGGCTTGCAGGTG	8,62
Donor splice site	6288984	CAGGTGGGG	6,92
(WT) Donor splice site	6289129	GAGGTGGGT	7,07
BP_Signals			
Type	**Position**	**Motif**	**Value**
Branch point	6288857	ACCCCAT	82,59
(WT) Branch point	6288948	GTCGGAG	72.02
Branch point	6288963	GGCTGAC	88,62
Branch point	6289019	CACGGAT	68,98
Branch point	6289030	AACTCAA	82,03
Branch point	6289063	TCCTCGC	68,61
Branch point	6289070	CGCGAAG	65,54
Branch point	6289118	GGCGGAC	73,53
Branch point	6289146	GGCTTAG	79,00
Branch point	6289157	GCCTCTG	65,78
Branch point	6289191	TGCTAAC	81,83
Branch point	6289195	AACTGAA	77,09
Branch point	6289203	AACTAAA	70,20
Branch point	6289210	ATCTTAC	80,32
Branch point	6289215	ACCAAAC	68,60
Branch point	6289220	ACCTAAC	83,69
Branch point	6289247	ATTTCAG	65,58
Branch point	6289274	TTCTCAT	90,98
ESE_Signals			
Name	**Position**	**Motif**	**CV**
ESE_9G8 (ESE Site)	6288839	GTGGAC	77,79
ESE_SRp40 (ESE Site)	6288846	TGCCTGG	79,13
ESE_9G8 (ESE Site)	6288853	TGTGAC	60,94
ESE_SC35 (ESE Site)	6288856	GACCCCAT	77,21
ESE_SC35 (ESE Site)	6288861	CATTTCTG	75,30
ESE_SRp40 (ESE Site)	6288863	TTTCTGC	85,17
ESE_SRp40 (ESE Site)	6288877	TTCCTGG	82,18
ESE_SC35 (ESE Site)	6288882	GGCCTGGG	75,06
ESE_9G8 (ESE Site)	6288888	GGTGAC	82,69
ESE_9G8 (ESE Site)	6288891	GACAAA	66,58
ESE_SRp40 (ESE Site)	6288892	ACAAAGG	84,22
ESE_ASF (ESE Site)	6288893	CAAAGGG	79,12
ESE_ASFB (ESE Site)	6288893	CAAAGGG	79,15

ESE_9G8 (ESE Site)	6288899	GAAGTG	65,31
ESE_9G8 (ESE Site)	6288905	GGTGAA	70,61
ESE_SRp40 (ESE Site)	6288907	TGAAAGG	85,59
ESE_ASF (ESE Site)	6288908	GAAAGGA	74,17
ESE_ASF (ESE Site)	6288911	AGGAGGT	76,09
ESE_9G8 (ESE Site)	6288913	GAGGTG	60,47
ESE_9G8 (ESE Site)	6288916	GTGGGC	61,82
ESE_SC35 (ESE Site)	6288919	GGCTGGCA	80,41
ESE_9G8 (ESE Site)	6288920	GCTGGC	62,42
ESE_SRp40 (ESE Site)	6288922	TGGCAGG	79,37
ESE_ASFB (ESE Site)	6288923	GGCAGGG	72,61
ESE_ASF (ESE Site)	6288923	GGCAGGG	80,75
ESE_9G8 (ESE Site)	6288929	GAGCAT	60,14
ESE_9G8 (ESE Site)	6288941	GGAGAG	71,82
ESE_ASF (ESE Site)	6288942	GAGAGGG	81,92
ESE_9G8 (ESE Site)	6288944	GAGGGT	65,91
ESE_SC35 (ESE Site)	6288947	GGTCGGAG	78,75
ESE_ASF (ESE Site)	6288950	CGGAGAA	77,78
ESE_ASFB (ESE Site)	6288950	CGGAGAA	81,61
ESE_9G8 (ESE Site)	6288951	GGAGAA	74,84
ESE_9G8 (ESE Site)	6288952	GAGAAT	61,35
ESE_9G8 (ESE Site)	6288960	GGAGGC	70,94
ESE_9G8 (ESE Site)	6288961	GAGGCT	62,49
ESE_9G8 (ESE Site)	6288964	GCTGAC	78,39
ESE_ASF (ESE Site)	6288965	CTGACTG	75,80
ESE_ASFB (ESE Site)	6288965	CTGACTG	76,15
ESE_SRp40 (ESE Site)	6288966	TGACTGG	93,50
ESE_SC35 (ESE Site)	6288967	GACTGGTG	84,89
ESE_9G8 (ESE Site)	6288971	GGTGTC	63,29
ESE_SRp40 (ESE Site)	6288973	TGTCTGG	87,21
ESE_SC35 (ESE Site)	6288978	GGCTTGCA	81,20
ESE_SRp40 (ESE Site)	6288981	TTGCAGG	82,42
ESE_ASFB (ESE Site)	6288982	TGCAGGT	72,61
ESE_ASF (ESE Site)	6288982	TGCAGGT	74,99
ESE_SRp55 (ESE Site)	6288982	TGCAGG	75,89
ESE_9G8 (ESE Site)	6288989	GGGGAA	70,00
ESE_9G8 (ESE Site)	6288992	GAAGCA	68,33
ESE_9G8 (ESE Site)	6288996	CACTAC	59,33
ESE_ASFB (ESE Site)	6289002	CTGCAGT	71,77
ESE_9G8 (ESE Site)	6289007	GTTGGC	62,42
ESE_SC35 (ESE Site)	6289010	GGCCGGCG	84,59
ESE_9G8 (ESE Site)	6289011	GCCGGC	66,04
ESE_ASFB (ESE Site)	6289012	CCGGCGA	76,23
ESE_9G8 (ESE Site)	6289014	GGCGAC	86,31
ESE_SRp40 (ESE Site)	6289016	CGACACG	90,98
ESE_ASFB (ESE Site)	6289017	GACACGG	72,00

ESE_ASF (ESE Site)	6289017	GACACGG	86,93
ESE_SRp55 (ESE Site)	6289019	CACGGA	79,54
ESE_ASF (ESE Site)	6289021	CGGATGA	81,63
ESE_ASFB (ESE Site)	6289021	CGGATGA	83,00
ESE_9G8 (ESE Site)	6289023	GATGAA	83,49
PESE (ESE Site)	6289025	TGAAGAAC	62,05
ESE_9G8 (ESE Site)	6289026	GAAGAA	87,72
ESE_Tra2 (ESE Site)	6289027	AAGAA	100,34
ESE_SC35 (ESE Site)	6289029	GAACTCAA	77,46
ESE_SRp40 (ESE Site)	6289031	ACTCAAC	78,83
ESE_SRp40 (ESE Site)	6289034	CAACAGC	79,13
ESE_Tra2 (ESE Site)	6289035	AACAG	83,54
ESE_SRp55 (ESE Site)	6289041	TGCACC	75,76
ESE_ASF (ESE Site)	6289043	CACCGCT	74,06
ESE_ASFB (ESE Site)	6289043	CACCGCT	78,38
ESE_SC35 (ESE Site)	6289043	CACCGCTG	79,42
ESE_ASF (ESE Site)	6289048	CTGTGGA	74,58
ESE_ASFB (ESE Site)	6289048	CTGTGGA	77,00
ESE_SRp55 (ESE Site)	6289049	TGTGGA	78,83
ESE_9G8 (ESE Site)	6289050	GTGGAC	77,79
ESE_SRp40 (ESE Site)	6289052	GGACTGG	78,53
ESE_SC35 (ESE Site)	6289061	GGTCCTCG	81,57
ESE_ASFB (ESE Site)	6289067	CGCCGCG	79,08
ESE_ASFB (ESE Site)	6289069	CCGCGAA	71,00
ESE_ASFB (ESE Site)	6289072	CGAAGCA	72,00
ESE_9G8 (ESE Site)	6289073	GAAGCA	68,33
ESE_ASF (ESE Site)	6289075	AGCAGGG	75,34
ESE_9G8 (ESE Site)	6289077	CAGGGC	59,67
ESE_ASF (ESE Site)	6289079	GGGCCGT	79,30
ESE_SC35 (ESE Site)	6289080	GGCCGTCG	79,61
ESE_9G8 (ESE Site)	6289081	GCCGTC	62,62
ESE_9G8 (ESE Site)	6289089	GAGGCT	62,49
ESE_9G8 (ESE Site)	6289097	GAAGCT	67,32
ESE_SRp55 (ESE Site)	6289102	TGCTTC	81,58
ESE_SC35 (ESE Site)	6289107	CGCCGGTG	76,04
ESE_ASF (ESE Site)	6289107	CGCCGGT	86,40
ESE_ASFB (ESE Site)	6289107	CGCCGGT	93,31
ESE_9G8 (ESE Site)	6289119	GCGGAC	77,79
ESE_ASF (ESE Site)	6289120	CGGACAG	77,08
ESE_ASFB (ESE Site)	6289120	CGGACAG	79,92
ESE_SC35 (ESE Site)	6289121	GGACAGAA	75,86
ESE_SRp40 (ESE Site)	6289123	ACAGAAG	79,90
ESE_ASFB (ESE Site)	6289124	CAGAAGA	80,00
ESE_ASF (ESE Site)	6289124	CAGAAGA	84,25
ESE_9G8 (ESE Site)	6289126	GAAGAG	84,70
ESE_ASF (ESE Site)	6289127	AAGAGGT	78,72

ESE_Tra2 (ESE Site)	6289127	AAGAG	89,41
ESE_9G8 (ESE Site)	6289129	GAGGTG	60,47
ESE_SRp55 (ESE Site)	6289133	TGGGTC	76,72
ESE_SC35 (ESE Site)	6289135	GGTCTGTG	85,81
ESE_SRp55 (ESE Site)	6289139	TGTGTG	75,05
ESE_SRp40 (ESE Site)	6289141	TGTGAGG	78,41
ESE_SRp55 (ESE Site)	6289143	TGAGGC	77,42
ESE_9G8 (ESE Site)	6289144	GAGGCT	62,49
ESE_SRp40 (ESE Site)	6289149	TTAGAAC	79,25
ESE_9G8 (ESE Site)	6289152	GAACAG	62,96
ESE_Tra2 (ESE Site)	6289153	AACAG	83,54
ESE_SC35 (ESE Site)	6289156	AGCCTCTG	84,15
ESE_SRp40 (ESE Site)	6289158	CCTCTGG	89,96
ESE_ASF (ESE Site)	6289159	CTCTGGA	76,04
ESE_ASFB (ESE Site)	6289159	CTCTGGA	79,92
ESE_9G8 (ESE Site)	6289164	GAGGGT	65,91
ESE_9G8 (ESE Site)	6289168	GTTGAG	63,29
ESE_SRp55 (ESE Site)	6289172	AGCAGC	75,18
ESE_SRp55 (ESE Site)	6289184	TGCTGC	76,40
ESE_9G8 (ESE Site)	6289199	GAACAA	65,98
ESE_Tra2 (ESE Site)	6289200	AACAA	94,47
ESE_Tra2 (ESE Site)	6289207	AAAAT	81,36
ESE_SRp40 (ESE Site)	6289216	CCAAACC	81,40
ESE_SRp40 (ESE Site)	6289221	CCTAACG	80,20
ESE_ASF (ESE Site)	6289226	CGCTGGT	77,90
ESE_ASFB (ESE Site)	6289226	CGCTGGT	83,69
ESE_9G8 (ESE Site)	6289230	GGTGAT	69,60
ESE_9G8 (ESE Site)	6289233	GATGCT	63,09
ESE_SC35 (ESE Site)	6289252	AGTTTCTG	79,24
ESE_SRp40 (ESE Site)	6289262	TTGCTGG	80,92
ESE_9G8 (ESE Site)	6289267	GGTGGC	66,72
		ESS_Signals	
Name	**Position**	**Motif**	**CV**
ESS_hnRNPA1 (ESS Site)	6288837	TAGTGG	82,27
Sironi_motif2 (ESS Site)	6288838	AGTGGAC	66,73
Sironi_motif2 (ESS Site)	6288850	TGGTGTG	63,08
Sironi_motif2 (ESS Site)	6288851	GGTGTGA	69,67
Sironi_motif3 (ESS Site)	6288855	TGACCCCA	64,15
Sironi_motif3 (ESS Site)	6288856	GACCCCAT	61,86
Sironi_motif3 (ESS Site)	6288865	TCTGCCCC	73,33
Sironi_motif3 (ESS Site)	6288871	CCTTCCTT	65,62
Sironi_motif3 (ESS Site)	6288875	CCTTCCTG	68,80
Sironi_motif2 (ESS Site)	6288883	GCCTGGG	60,01
Sironi_motif2 (ESS Site)	6288884	CCTGGGT	61,65
Sironi_motif2 (ESS Site)	6288886	TGGGTGA	60,21
Sironi_motif1 (ESS Site)	6288892	ACAAAGGG	64,83

Sironi_motif2 (ESS Site)	6288894	AAAGGGA	62,87
ESS_hnRNPA1 (ESS Site)	6288894	AAAGGG	65,13
ESS_hnRNPA1 (ESS Site)	6288895	AAGGGA	85,13
Sironi_motif2 (ESS Site)	6288898	GGAAGTG	62,31
Sironi_motif2 (ESS Site)	6288899	GAAGTGG	60,55
Sironi_motif1 (ESS Site)	6288899	GAAGTGGG	65,72
ESS_hnRNPA1 (ESS Site)	6288900	AAGTGG	67,03
Sironi_motif2 (ESS Site)	6288901	AGTGGGT	81,60
Sironi_motif2 (ESS Site)	6288903	TGGGTGA	60,21
Sironi_motif1 (ESS Site)	6288907	TGAAAGGA	62,38
Sironi_motif2 (ESS Site)	6288907	TGAAAGG	65,39
ESS_hnRNPA1 (ESS Site)	6288909	AAAGGA	67,98
ESS_hnRNPA1 (ESS Site)	6288910	AAGGAG	73,94
Sironi_motif1 (ESS Site)	6288910	AAGGAGGT	76,76
Sironi_motif2 (ESS Site)	6288912	GGAGGTG	77,46
ESS_hnRNPA1 (ESS Site)	6288913	GAGGTG	74,41
Sironi_motif2 (ESS Site)	6288914	AGGTGGG	73,16
Sironi_motif2 (ESS Site)	6288915	GGTGGGC	87,38
Sironi_motif2 (ESS Site)	6288919	GGCTGGC	66,88
Sironi_motif2 (ESS Site)	6288923	GGCAGGG	69,72
Sironi_motif2 (ESS Site)	6288924	GCAGGGA	66,62
ESS_hnRNPA1 (ESS Site)	6288925	CAGGGA	85,13
Sironi_motif2 (ESS Site)	6288931	GCATGGG	69,50
Sironi_motif1 (ESS Site)	6288932	CATGGGGT	67,99
Sironi_motif2 (ESS Site)	6288932	CATGGGG	73,71
Sironi_motif2 (ESS Site)	6288934	TGGGGTG	71,04
Sironi_motif2 (ESS Site)	6288935	GGGGTGG	67,96
Sironi_motif2 (ESS Site)	6288936	GGGTGGG	76,91
Sironi_motif2 (ESS Site)	6288937	GGTGGGA	86,58
Sironi_motif2 (ESS Site)	6288939	TGGGAGA	60,21
Sironi_motif2 (ESS Site)	6288941	GGAGAGG	77,46
Sironi_motif1 (ESS Site)	6288941	GGAGAGGG	79,52
ESS_hnRNPA1 (ESS Site)	6288942	GAGAGG	74,65
Sironi_motif2 (ESS Site)	6288943	AGAGGGT	78,55
ESS_hnRNPA1 (ESS Site)	6288944	GAGGGT	87,98
Sironi_motif2 (ESS Site)	6288946	GGGTCGG	60,01
Sironi_motif2 (ESS Site)	6288947	GGTCGGA	71,43
ESS_hnRNPA1 (ESS Site)	6288952	GAGAAT	68,22
Sironi_motif2 (ESS Site)	6288957	TCTGGAG	66,68
Sironi_motif1 (ESS Site)	6288958	CTGGAGGC	78,22
Sironi_motif2 (ESS Site)	6288959	TGGAGGC	62,77
Sironi_motif2 (ESS Site)	6288960	GGAGGCT	65,40
ESS_hnRNPA1 (ESS Site)	6288961	GAGGCT	77,98
Sironi_motif2 (ESS Site)	6288966	TGACTGG	65,39
Sironi_motif2 (ESS Site)	6288973	TGTCTGG	68,44
Sironi_motif2 (ESS Site)	6288982	TGCAGGT	60,74

Sironi_motif2 (ESS Site)	6288983	GCAGGTG	60,55
ESS_hnRNPA1 (ESS Site)	6288984	CAGGTG	70,60
Sironi_motif2 (ESS Site)	6288985	AGGTGGG	73,16
Sironi_motif2 (ESS Site)	6288986	GGTGGGG	97,41
Sironi_motif2 (ESS Site)	6288988	TGGGGAA	60,21
Sironi_motif3 (ESS Site)	6288997	ACTACCTG	65,95
Sironi_motif2 (ESS Site)	6289006	AGTTGGC	75,67
ESS_hnRNPA1 (ESS Site)	6289019	CACGGA	65,84
Sironi_motif1 (ESS Site)	6289026	GAAGAACT	63,88
Sironi_motif1 (ESS Site)	6289034	CAACAGCT	73,34
Sironi_motif2 (ESS Site)	6289047	GCTGTGG	63,60
Sironi_motif2 (ESS Site)	6289049	TGTGGAC	73,56
Sironi_motif2 (ESS Site)	6289052	GGACTGG	62,31
Sironi_motif2 (ESS Site)	6289057	GGCTGGT	64,85
Sironi_motif3 (ESS Site)	6289064	CCTCGCCG	67,10
Sironi_motif2 (ESS Site)	6289075	AGCAGGG	65,97
Sironi_motif2 (ESS Site)	6289076	GCAGGGC	67,43
ESS_hnRNPA1 (ESS Site)	6289077	CAGGGC	83,22
Sironi_motif2 (ESS Site)	6289086	CGCGAGG	61,17
Sironi_motif1 (ESS Site)	6289086	CGCGAGGC	75,30
ESS_hnRNPA1 (ESS Site)	6289089	GAGGCT	77,98
Sironi_motif2 (ESS Site)	6289091	GGCTGTG	60,01
Sironi_motif2 (ESS Site)	6289094	TGTGAAG	66,68
Sironi_motif2 (ESS Site)	6289113	TGCTTGG	63,08
Sironi_motif2 (ESS Site)	6289114	GCTTGGC	62,52
Sironi_motif2 (ESS Site)	6289117	TGGCGGA	61,96
Sironi_motif1 (ESS Site)	6289123	ACAGAAGA	64,77
Sironi_motif2 (ESS Site)	6289126	GAAGAGG	60,55
Sironi_motif1 (ESS Site)	6289126	GAAGAGGT	89,73
ESS_hnRNPA1 (ESS Site)	6289127	AAGAGG	70,84
Sironi_motif2 (ESS Site)	6289128	AGAGGTG	73,70
ESS_hnRNPA1 (ESS Site)	6289129	GAGGTG	74,41
Sironi_motif2 (ESS Site)	6289130	AGGTGGG	73,16
Sironi_motif2 (ESS Site)	6289131	GGTGGGT	85,35
Sironi_motif2 (ESS Site)	6289139	TGTGTGA	72,75
Sironi_motif1 (ESS Site)	6289141	TGTGAGGC	65,65
Sironi_motif2 (ESS Site)	6289141	TGTGAGG	83,58
Sironi_motif2 (ESS Site)	6289143	TGAGGCT	68,47
ESS_hnRNPA1 (ESS Site)	6289144	GAGGCT	77,98
ESS_hnRNPA1 (ESS Site)	6289150	TAGAAC	78,70
Sironi_motif2 (ESS Site)	6289160	TCTGGAG	66,68
Sironi_motif1 (ESS Site)	6289161	CTGGAGGG	76,30
Sironi_motif2 (ESS Site)	6289162	TGGAGGG	72,80
Sironi_motif2 (ESS Site)	6289163	GGAGGGT	82,30
ESS_hnRNPA1 (ESS Site)	6289164	GAGGGT	87,98
Sironi_motif2 (ESS Site)	6289167	GGTTGAG	72,55

Sironi_motif2 (ESS Site)	6289170	TGAGCAG	63,63
ESS_hnRNPA1 (ESS Site)	6289181	TAATGC	66,08
Sironi_motif3 (ESS Site)	6289212	CTTACCAA	69,36
ESS_hnRNPA1 (ESS Site)	6289223	TAACGC	66,08
Sironi_motif2 (ESS Site)	6289227	GCTGGTG	63,60
Sironi_motif2 (ESS Site)	6289230	GGTGATG	63,60
Sironi_motif2 (ESS Site)	6289232	TGATGCT	60,51
Sironi_motif2 (ESS Site)	6289238	TGTTGGG	92,53
Sironi_motif2 (ESS Site)	6289239	GTTGGGA	69,67
Sironi_motif2 (ESS Site)	6289263	TGCTGGT	67,93
Sironi_motif2 (ESS Site)	6289264	GCTGGTG	63,60
Sironi_motif1 (ESS Site)	6289265	CTGGTGGC	62,62
Sironi_motif2 (ESS Site)	6289266	TGGTGGC	69,96
Sironi_motif2 (ESS Site)	6289267	GGTGGCC	70,48
Sironi_motif3 (ESS Site)	6289273	CTTCTCAT	62,47

Table 4

WFS1 splice variant	Consensus sequence
NM_006005.3_c.316del	GGGCCTACAAAGGGAGACATGGAAATCCCCTTTGAAGAAGTCCTGGAGAGGGCCAAGGCCGGGGACCCCAAGGCACAGACTG AGTGGGGAAGCACTACCTGCAGTTGGCCGGCGACACGGATGAAGAACTCAACAGCTGCACCGCTGTGGACTGGCTGGTCCTC GCCGCGAAGCAGGGCCGTCGCGAGGCTGTGAAGCTGCTTCGCCGGTGCTTGGCGGACAGAAGAGGCATCACGTCCGAGAACG AACGGGAGGTGAGGCAGCTCTCCTCCGAGACCGACCTGGAGAGGGCCGTGCGCAAGGCAGCCCTGGTCATGTACTGG
NM_006005.3_c.316_461del	GGGCCTACAAAGGGAGACATGGAAATCCCCTTTGAAGAAGTCCTGGAGAGGGCCAAGGCCGGGGACCCCAAGGCACAGACTG AGGCATCACGTCCGAGAACGAACGGGAGGTGAGGCAGCTCTCCTCCGAGACCGACCTGGAGAGGGCCGTGCGCAAGGCAGCC CTGGTCATGTACTGG
NM_006005.3_c.316_457del	GGGCCTACAAAGGGAGACATGGAAATCCCCTTTGAAGAAGTCCTGGAGAGGGCCAAGGCCGGGGACCCCAAGGCACAGACTG AGAGAGGCATCACGTCCGAGAACGAACGGGAGGTGAGGCAGCTCTCCTCCGAGACCGACCTGGAGAGGGCCGTGCGCAAGGC AGCCCTGGTCATGTACTGG
NM_006005.3_c.271_513del	GGGCCTACAAAGGGAGACATGGAAATCCCCTTTGAAGAACTGGAGAGGGCCGTGCGCAAGGCAGCCCTGGTCATGTACTGG