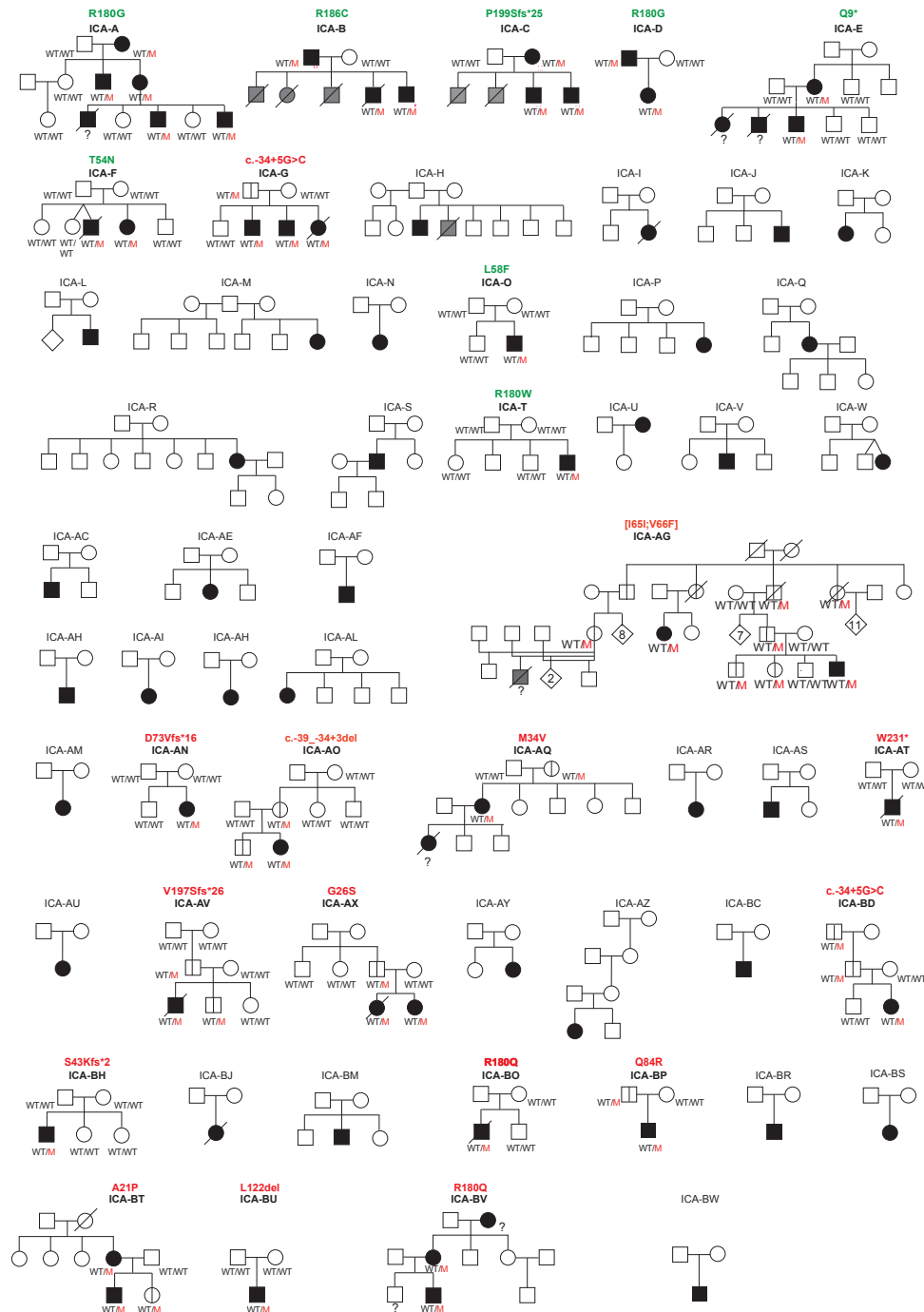


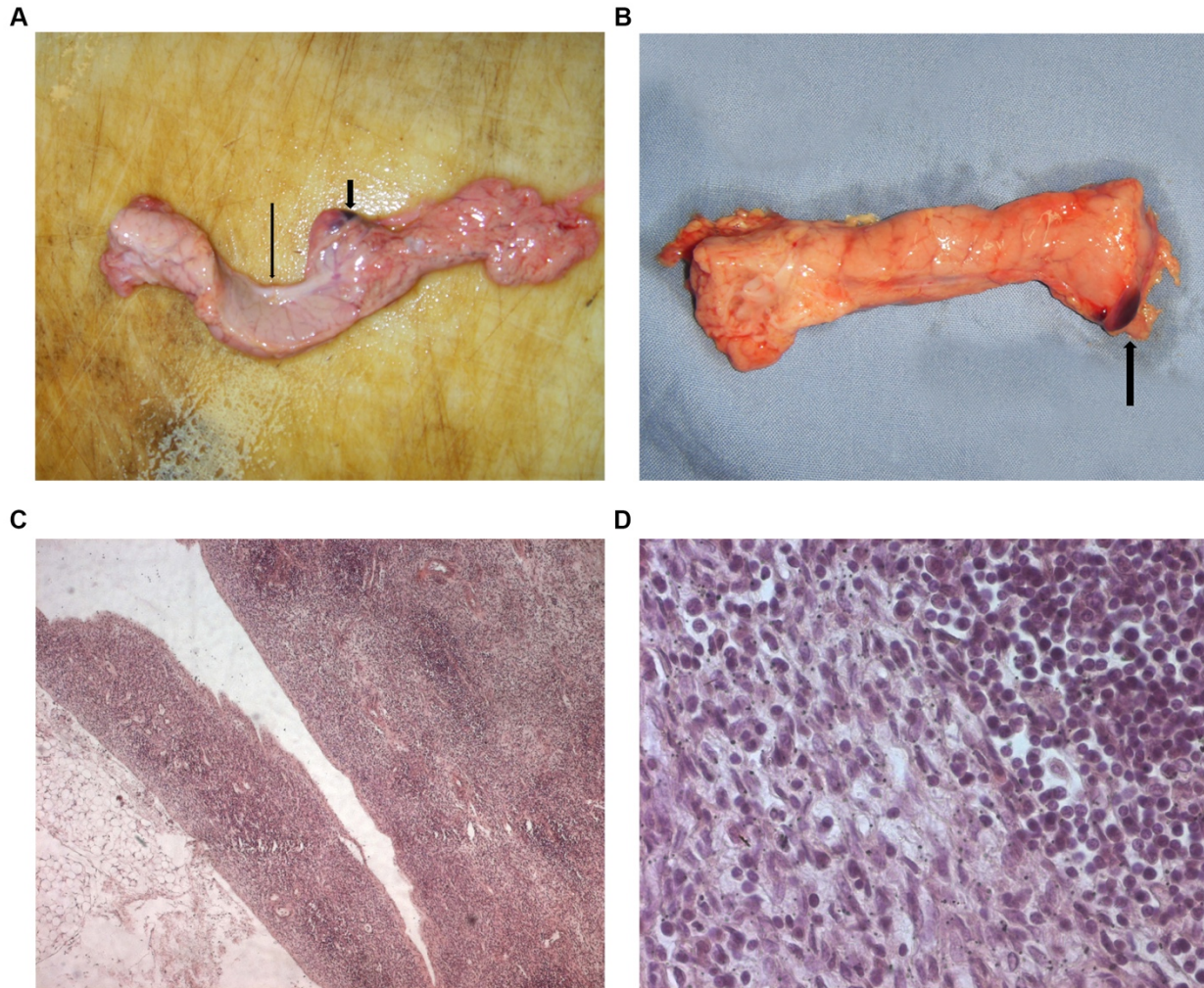
# Figure S1: Pedigrees of families who participated in our ICA research project

Black symbols represent ICA patients. Gray symbols represent probable ICA patients: individuals that died very early from severe infections, with a family history of ICA, and for whom asplenia could not be ascertained. Each kindred is represented by letters (ICA-“letter”), shown above the family tree. ICA-causing *RPSA* mutations are indicated above the kindred code when applicable. Mutations in green indicate it was identified and reported in our 2013 paper. In these kindreds, the genotype for the mutation is indicated. WT, wild type; M, mutation; ?, no gDNA was available.



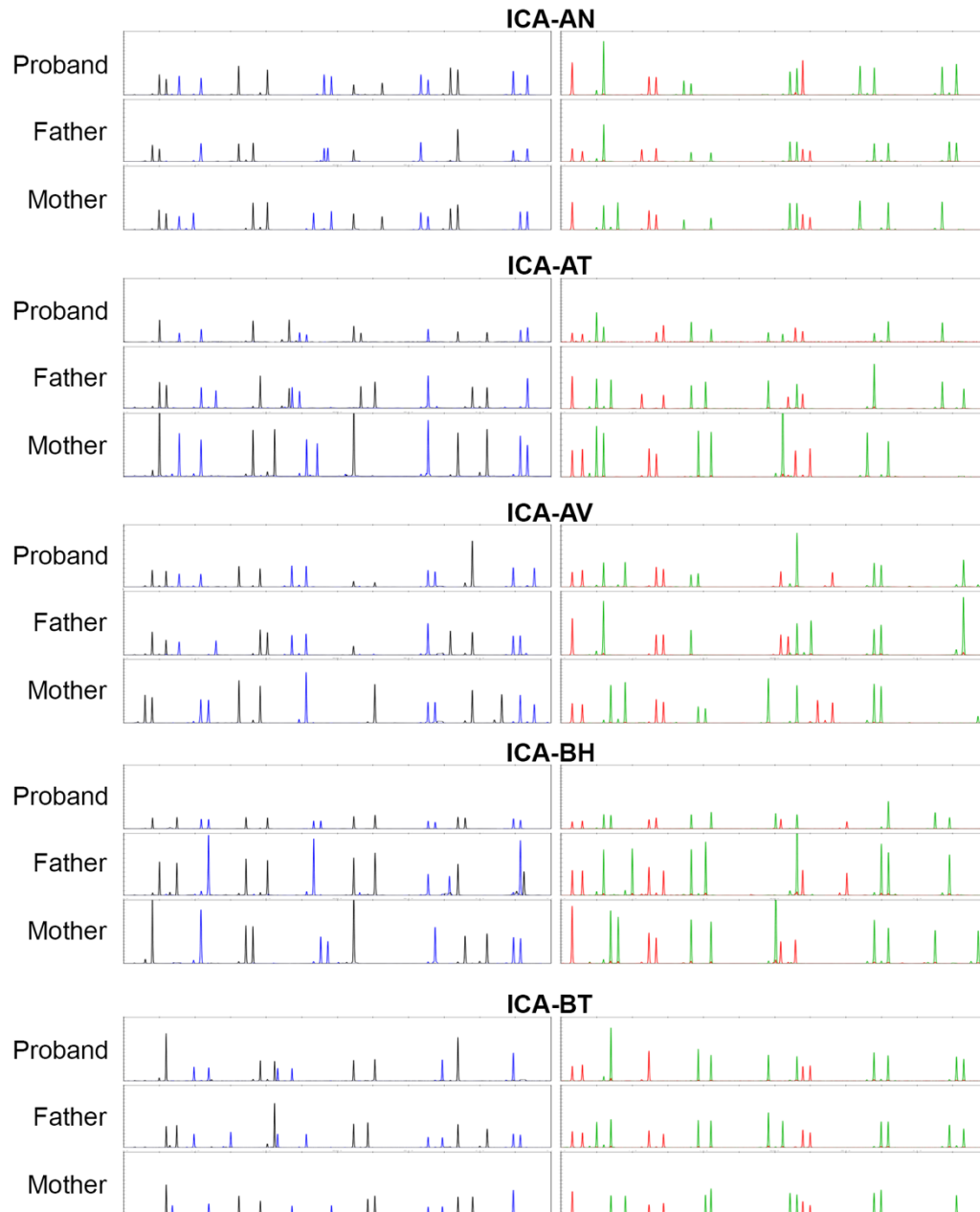
**Figure S2: Post-mortem analyses of ICA patients carrying RPSA mutations reveal rudimentary spleen**

Autopsy of ICA patient from kindred ICA-AT revealed tiny spleen parenchyme next to the pancreas. **A.** narrow arrow indicates what is likely to be the spleen artery. **B.** Arrow points to the hypoplastic spleen nodule. **C.** HE stain x40 **D.** HE stain x400.



### Figure S3: Microsatellite analyses in kindreds ICA-AN, -AT, -AV, -BH and -BT

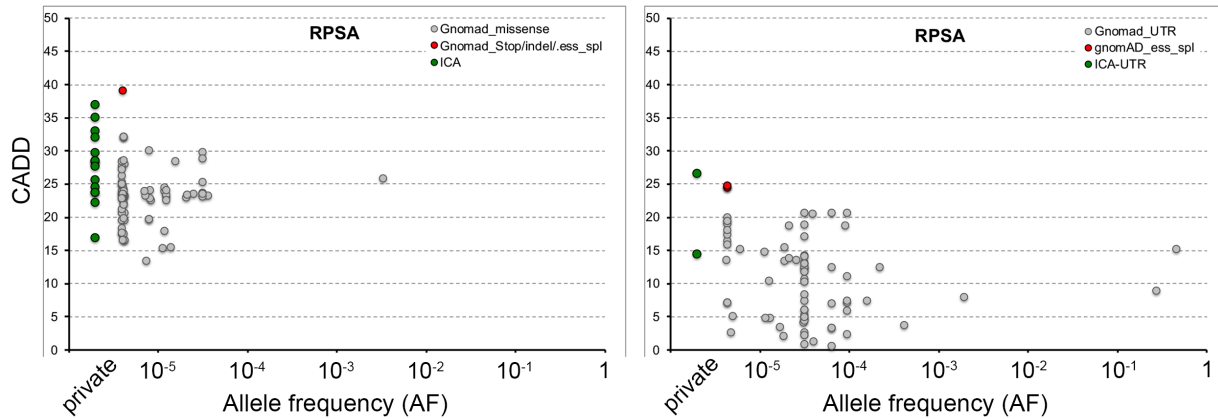
We confirmed whether the samples sequenced in kindreds in which RPSA mutations appeared *de novo* in probands from family ICA-AN, ICA-AT, ICA-AV, ICA-BH and -BT were from the expected parents by analyzing 15 polymorphic microsatellites throughout their genome (proband, father and mother). Snapshot of the analysis using Peak Scanner software after electrophoresis separation on 3730 Genetic Analyzer (Applied Biosystem) is shown. Each microsatellite fragments are identified based on the length of the product (X-axis; i.e. size of the amplicon) and the dye-specific labelled primers (Y-axis). The 4 colors correspond to the four different dyes used (6-Fam™ in blue, VIC® in green, NED™ in yellow, PET® in red).



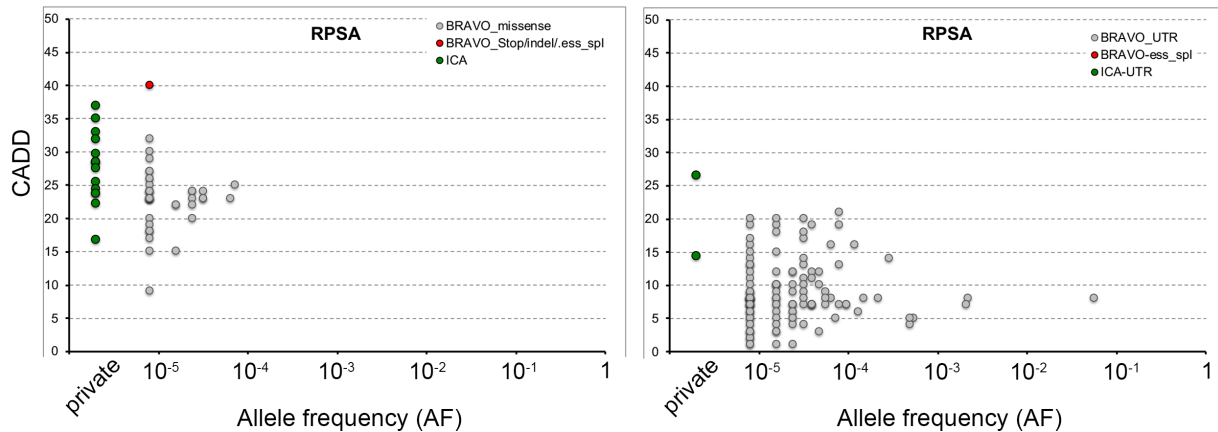
**Figure S4: CADD versus AF plots of variants found in ICA cohort and public database**

Both CADD score and AF or *RPSA* variants reported in gnomAD (panel A) and Bravo (panel B) are plotted. The left panels represent all missense, nonsense, indel and essential splicing variants impacting the coding region of *RPSA*. The right panels, indicate UTR and/or splice variants impacting the UTR of *RPSA*. The green dots indicate the variants found in the ICA cohort, whereas red (non-sense, indel and essential splicing variants) and grey (missense and UTR variants) indicate the variants present in either public database.

A

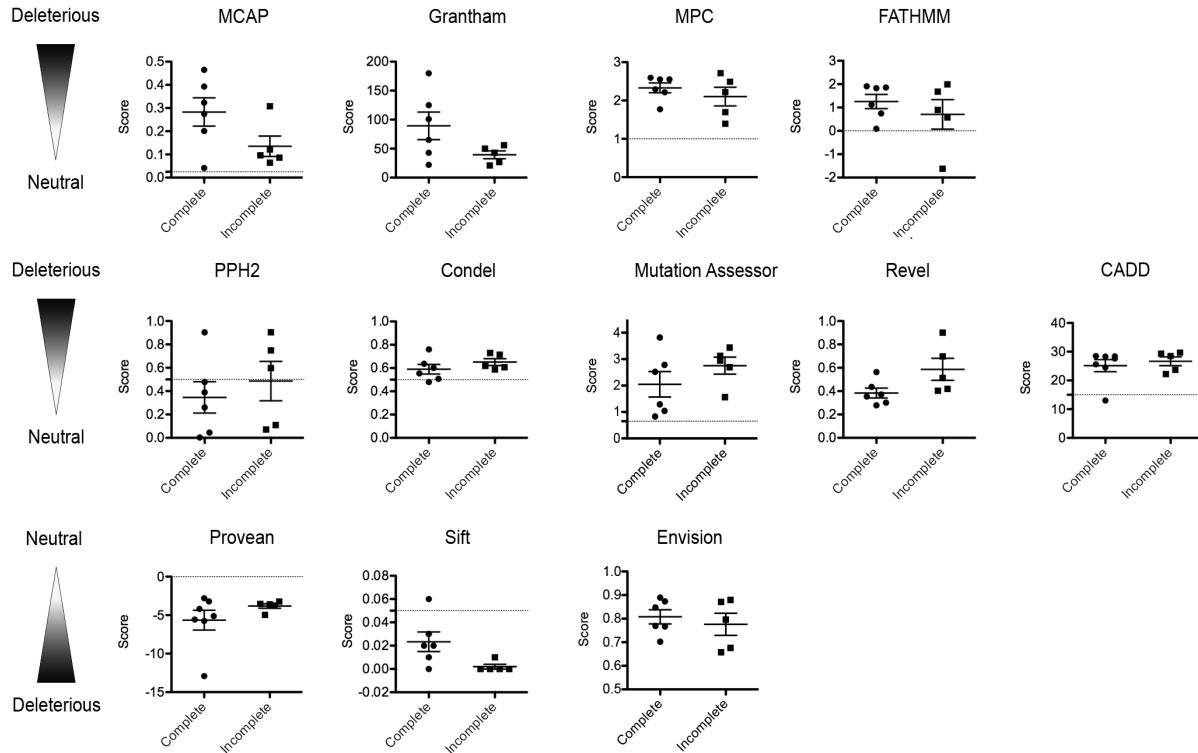


B

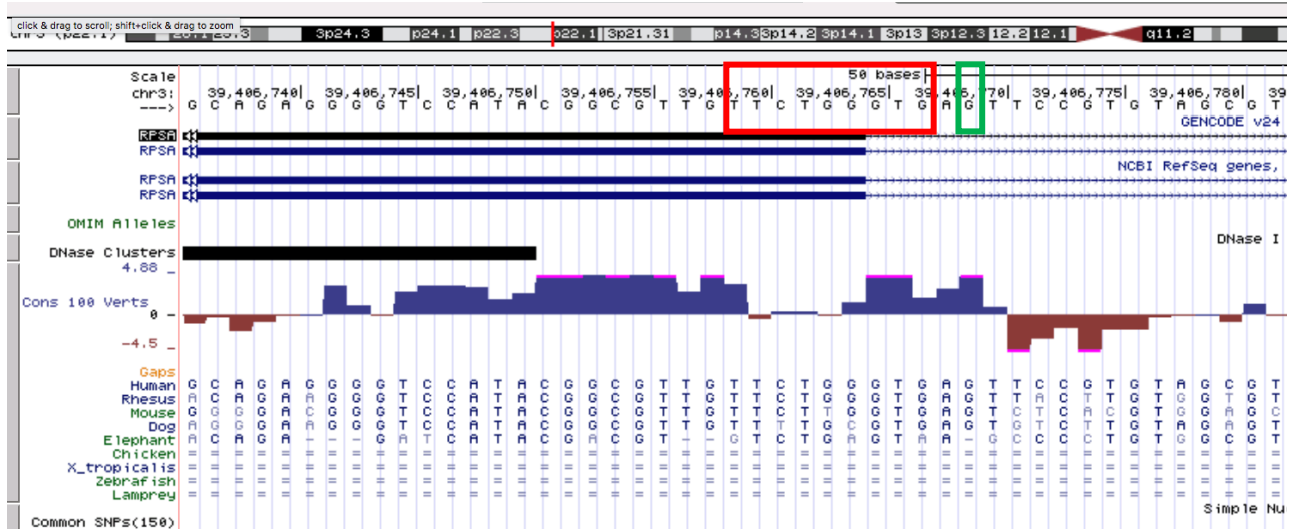


### Figure S5: Prediction of the deleteriousness of missense or indel-frameshift mutations in ICA.

We used twelve software (CADD (1), Condel (2), Envision (3), FATHMM (4), Grantham (5), Mutation Assessor (6), MCAP (7), MPC (8), Polyphen2 (PPH2) (9), Provean (10), Revel (11), and Sift (12)) to predict the impact of the missense and inframe-indel with complete penetrance (p.R180Q, p.R180W, p.R186C, p.L58F, p.R180G, p.T54N, p.L122del) and incomplete penetrance (p.V66F, p.Q84R, p.M34V, p.A21P, p.G26S). Note that the p.L122del was only considered by CADD and Provean.



Green and red boxes indicate the position of nucleotides mutated in ICA patients (c.-34+5G>C and c.-39\_-34+3del respectively). Blue bars in histogram indicate nucleotides with positive phyloP score, brown bars indicate nucleotides with negative PhyloP scores – adapted from UCSC.



**Figure S7: Consequence of the essential splicing mutations on the *RPSA* mRNA**

mRNA sequence from WT, c.-39\_-34+3del and c.-34+5G>C from Exon 1 to Exon 2 have been extracted from RNAseq data. The partial intronic sequence are in lower case. The two new ATGs are indicated in red underlined lowercase; they are not inframe with the original ATG (capital ATG). The codon triplets are indicated by three successive dashes. Deduced amino acid-sequences from the two new ATGs with premature stop codon are indicated.

WT  
c.-39\_-34+3del  
c.-34+5G>C

< -----Exon1----- >  
GCCTGTCTTTTCCGTGCTACCTGCAGAGGGTCCATACGGCGTTGTTCTGG-----agttccgtgtagcgtccctggcgccctccagg  
GCCTGTCTTTTCCGTGCTACCTGCAGAGGGTCCATACGGCGTTGTTCTGGGtgacttccgtgtagcgtccctggcgccctccagg

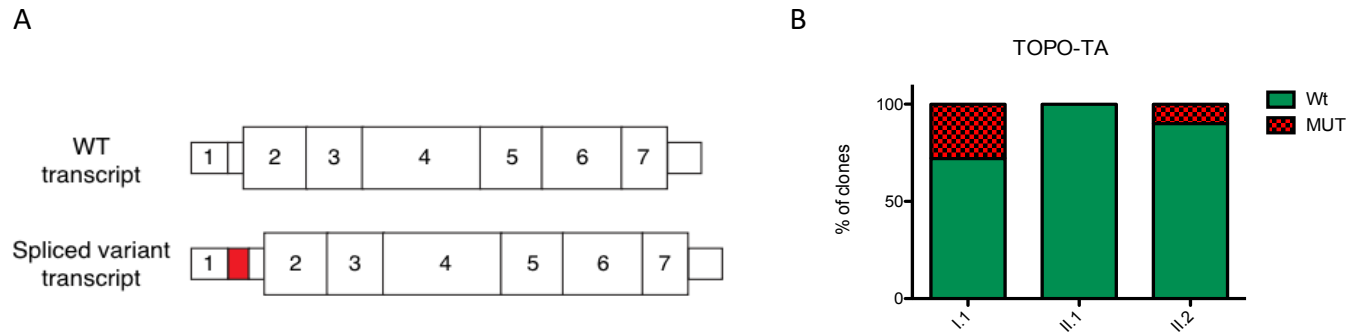
Frame  
WT  
c.-39\_-34+3del  
c.-34+5G>C

< -----Exon2----- >  
-----ATCCCGTCGTAACCTAAAGGGAAATTTTCACAATGTCCGGAGCCCTTGATG  
gctagaaaaatgagcttttcctgtcctaaatgaaggATCCCGTCGTAACCTAAAGGGAAATTTTCACAATGTCCGGAGCCCTTGATG  
gctagaaaaatgagcttttcctgtcctaaatgaaggATCCCGTCGTAACCTAAAGGGAAATTTTCACAATGTCCGGAGCCCTTGATG  
M S F S C S N E G F P S \*  
M K D S R R N L K G N F H N V R S P \*



**Figure S8: mRNA expression in heterozygous c-34+5G>C/WT patient's cells**

(A) A 70 bp insertion is identified by TOPO-TA cloning on cDNA from the EBV-B cells from the father (I.1; WT/M), his two sons (II.1; WT/WT, II.2; WT/M). (B) Quantification of clones carrying WT or mutated alleles in ICA-G I.1 (46 WT (72%) and 18 splice mutants (28%)), in ICA-G II.1 (130 WT (100%)) and in ICA-G II.2 (76 WT (90%) and 8 splice mutants (10%)).



**Additional references:**

1. Kircher M, *et al.* (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46(3):310-315.
2. Gonzalez-Perez A & Lopez-Bigas N (2011) Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet* 88(4):440-449.
3. Gray VE, Hause RJ, Luebeck J, Shendure J, & Fowler DM (2018) Quantitative Missense Variant Effect Prediction Using Large-Scale Mutagenesis Data. *Cell Systems* 6(1):116-124.e113.
4. Shihab HA, *et al.* (2013) Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Human mutation* 34(1):57-65.
5. Grantham R (1974) Amino Acid Difference Formula to Help Explain Protein Evolution. *Science* 185(4154):862.
6. Reva B, Antipin Y, & Sander C (2011) Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* 39(17):e118-e118.
7. Jagadeesh KA, *et al.* (2016) M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nature Genetics* 48:1581.
8. Samocha KE, *et al.* (2017) Regional missense constraint improves variant deleteriousness prediction. *bioRxiv*.
9. Adzhubei IA, *et al.* (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7(4):248-249.
10. Choi Y, Sims GE, Murphy S, Miller JR, & Chan AP (2012) Predicting the functional effect of amino acid substitutions and indels. *Plos One* 7(10):e46688.
11. Ioannidis NM, *et al.* (2016) REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet* 99(4):877-885.
12. Kumar P, Henikoff S, & Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols* 4(7):1073-1081.



**Table S1: Clinical phenotype of RPSA-ICA patients**

Family: Patient <sup>a</sup>	Recruitment	Sex	Age at ICA Diagnosis	Age at 1st infection	Spleen (US/PM)	Howell-Jolly Bodies	CBC	Country of origin	RPSA mut
ICA-A: I.2	Physician	F	Adult	None	Small (US)	NA	normal <sup>d</sup>	UK	p.R180G
ICA-A: II.3	Physician	M	20 y	None	Absent (US)	yes	Anemia (Hb 4.4 g/dL) <sup>d</sup>	UK	p.R180G
ICA-A: II.4	Physician	F	40 y	None	Absent (US)	yes	normal	UK	p.R180G
ICA-A: III.2	Physician	M	34 mo	21 mo	Very small (PM) <sup>b</sup>	NA	N/A	UK	No DNA
ICA-A: III.4	Physician	M	Early childhood	None <sup>c</sup>	Absent (US)	NA	normal	UK	p.R180G
ICA-A: III.6	Physician	M	At birth	None <sup>c</sup>	Absent (US)	NA	normal	UK	p.R180G
ICA-B: I.1	Physician	M	56 y	None	Absent (US)	yes	normal	Senegal	p.R186C
ICA-B: II.4	Physician	M	23 mo	23 mo	Very small (PM) <sup>b</sup>	NA	N/A	Senegal	p.R186C
ICA-B: II.5	Physician	M	At birth	None <sup>c</sup>	Absent (US)	NA	normal	Senegal	p.R186C
ICA-C: I.2	Physician	F	54 y	None	Absent (US)	yes	normal	France	p.P199fs*25
ICA-C: II.1	Physician	M	Early childhood	Fatal	Absent (PM)	NA	N/A	Reunion island	No DNA
ICA-C: II.2	Physician	M	Early childhood	Fatal	Absent (PM)	NA	N/A	Reunion island	No DNA
ICA-C: II.3	Physician	M	Early childhood	None <sup>c</sup>	Small, bilobar (US)	NA	normal	Reunion island	p.P199fs*25
ICA-C: II.4	Physician	M	Early childhood	None <sup>c</sup>	1.86 x 0.5 cm (US)	yes	normal	Reunion island	p.P199fs*25
ICA-D: I.1	Physician	M	49 y	None	Absent (US)	yes	N/A	Italy	p.R180G
ICA-D: II.1	Physician	F	19 mo	None	Absent (US)	yes	lymphocytosis	Italy	p.R180G
ICA-E: II.2	Physician	F	57 y	None	Absent (US)	yes	normal	USA	p.Q9*
ICA-E: III.1	Physician	F	6 mo	6 mo	Absent (PM)	NA	N/A	USA	No DNA
ICA-E: III.2	Physician	M	Early childhood	Infant	Absent (PM)	NA	N/A	USA	No DNA
ICA-E: III.3	Physician	M	7 mo	None <sup>c</sup>	Absent (US)	yes	normal	USA	p.Q9*
ICA-F: II.3	Physician	M	22 mo	22 mo	Very small (PM) <sup>b</sup>	NA	N/A	France	p.T54N
ICA-F: II.4	Physician	F	4 d	None <sup>c</sup>	Absent (US)	yes	normal	France	p.T54N
ICA-O: II.2	Physician	F	20 mo	20 mo	Absent (US)	yes	normal	France	p.L58F
ICA-T: II.4	Physician	M	9 mo	8 mo	Absent (US)	yes	normal	Sweden	p.R180W
ICA-G: II.2	Physician	M	2 y	None <sup>c</sup>	Absent (US)	yes	N/A	UK	c.-34+5G>C
ICA-G: II.3	Physician	F	8 mo	8 mo	Very small (PM) <sup>b</sup>	NA	N/A	UK	c.-34+5G>C
ICA-G: II.4	Physician	M	Early childhood	None <sup>c</sup>	Absent (US)	yes	N/A	UK	c.-34+5G>C
ICA-AG:IV:4	<b>Self-referred</b>	M	Early childhood		NA	NA	N/A	USA	No DNA
ICA-AG: IV:12	<b>Self-referred</b>	M	9 mo	9 mo	Absent (US)	NA		USA	p.[I65I;V66F]
ICA-AG: III:10	<b>Self-referred</b>	F	66 y	None	Absent (US)	NA	Anemia (Hb 5.4 g/dL) <sup>d</sup>	USA	p.[I65I;V66F]
ICA-AN: II:2	Physician	F	Early childhood	Infant	Absent (US)	yes	N/A	Germany	p.D73Vfs*16
ICA-AO: III:2	<b>Self-referred</b>	F	31 y	None	Absent (US/CT)	yes	normal	UK	c. -39_-34+3del
ICA-AQ: III:1	Physician	F	9 mo	9mo	Absent (PM)	NA	N/A	USA	No DNA
ICA-AQ: II:1	Physician	F	34 y	None	Absent (CT)	NA	Thrombocytosis	USA	p.M34V
ICA-AT: II:1	Physician	M	10 mo	10 mo	Absent (PM) <sup>b</sup>	NA	N/A	UK	p.W231*
ICA-AV: III:1	Physician	M	4 y	4 y	N/A	NA	N/A	France	p.V197Sfs*26
ICA-AX: II:1	Physician	F	16mo	16 mo	Absent (PM)	NA	normal	UK	p.G26S
ICA-AX: II:2	Physician	F	2 mo	None <sup>c</sup>	Absent (US)	yes	Transient neutropenia	UK	p.G26S
ICA-BD: III:2	Physician	F	11mo	11 mo	Absent (PM)	NA	N/A	UK	c.-34+5G>C
ICA-BH: II:1	<b>Self-referred</b>	M	8 y	8 y	Very small (PM)	NA	N/A	Honduras/ Nicaragua	p.S43Kfs*2
ICA-BO: II:1	<b>Self-referred</b>	M	7.5 mo	7.5 mo	Absent (PM)	NA	N/A	USA	p.R180Q
ICA-BP: II:1	<b>Self-referred</b>	M	3 mo	3 mo	Absent (US)	Yes	NI	USA	p.Q84R
ICA-BT:III.1	Physician	M	18 mo		Very small (US)	no	normal	Argentina	p.A21P
ICA-BT:II.1	Physician	F	Adult		Very small (US)	NA	normal	Argentina	p.A21P
ICA-BU:II.1	Physician	M	6 mo		Absent	NA	N/A	France	p.L122del
ICA-BV:V.1	Physician	M	3 yo		Absent	Yes	normal	France	p.R180Q
ICA-BV:IV.2	Physician	F	Adult		Small (PM)	Yes	NI	France	p.R180Q

US: ultrasound. PM: post mortem analysis. NA: not assessed. y: year. mo: months, d: days

<sup>a</sup>: the patient code used in **Figure S1** is used to identify patients. The first letter represents the family ID, the first number identifies the generation and the final number the position within that generation.

<sup>b</sup>: It is important to note that the spleens analyzed in the several independent post mortem reports were very similar. There was a spleen, or a spleen anlage, in all cases. However, these rudimentary spleens were much smaller (1-2 cm long) than those of healthy children of the same age (22-34 mo) and would probably not be detectable on US or functional. Most of the organ consisted of scar tissue, also described as a large fibrotic nodule. There were few lymphoid follicles. Finally, hemosiderin deposition was noted in at least three cases (sometimes described as Gandy-Gamna bodies). For family F, the detailed analysis has been reported (11).

<sup>c</sup>: under prophylaxis since early childhood.

<sup>d</sup>: required blood transfusions.

<sup>e</sup>: all mutations are protein level unless otherwise indicated

**Table S2: Nonsense variants, insertions, deletions and essential splicing variants extracted from gnomAD and Bravo databases**

Variants (position hg38, reference allele, alternate allele)	location	Consequence	Transcript Consequence (ENST00000301821)	Protein Consequence	Bravo	gnomAD	Total
3:39407619,A,G	intron	splice acceptor	c.-33-2A>G			1	1
3:39407620,G,C	intron	splice acceptor	c.-33-1G>C			1	1
3:39411677,G,A	CDS	stop gained	c.527G>A	p.Trp176*		1	1
3:39411953,G,GGT	CDS	frameshift	c.685insGT	p.Glu230Valfs*60	1		1
3:39412327,C,T	CDS	stop gained	c.847C>T	p.Gln283*		1	1

**Table S3: Missense variants from gnomAD and Bravo databases**

Variants (position hg38, reference allele, alternate allele)	location	Consequence	Transcript Consequence (ENST00000301821)	Protein Consequence	Bravo	gnomAD	Total
3:39407663,G,A	CDS	missense	c.10G>A	p.Ala4Thr	1		1
3:39407672,G,A	CDS	missense	c.19G>A	p.Val7Ile		1	1
3:39407681,A,G	CDS	missense	c.28A>G	p.Met10Val	1	3	4
3:39407707,C,A	CDS	missense	c.54C>A	p.Phe18Leu		1	1
3:39407733,G,T	CDS	missense	c.80G>T	p.Gly27Val	1		1
3:39407739,A,G	CDS	missense	c.86A>G	p.Asn29Ser		1	1
3:39407741,C,G	CDS	missense	c.88C>G	p.Leu30Val		1	1
3:39407744,G,C	CDS	missense	c.91G>C	p.Asp31His		1	1
3:39407749,C,A	CDS	missense	c.96C>A	p.Phe32Leu	1		1
3:39407750,C,G	CDS	missense	c.97C>G	p.Gln33Glu		1	1
3:39407756,G,A	CDS	missense	c.103G>A	p.Glu35Lys	9		9
3:39407756,G,C	CDS	missense	c.103G>C	p.Glu35Gln	1		1
3:39407760,A,T	CDS	missense	c.107A>T	p.Gln36Leu	1		1
3:39407768,T,C	CDS	missense	c.115T>C	p.Tyr39His	1		1
3:39407781,G,A	CDS	missense	c.128G>A	p.Ser43Asn		1	1
3:39408610,C,G	CDS	missense	c.138C>G	p.Ile46Met		1	1
3:39408614,A,C	CDS	missense	c.142A>C	p.Ile48Leu		1	1
3:39408614,A,G	CDS	missense	c.142A>G	p.Ile48Val	1	3	4
3:39408615,T,C	CDS	missense	c.143T>C	p.Ile48Thr		1	1
3:39408627,A,G	CDS	missense	c.155A>G	p.Lys52Arg	1		1
3:39408639,A,G	CDS	missense	c.167A>G	p.Glu56Gly		2	2
3:39408650,C,G	CDS	missense	c.178C>G	p.Leu60Val		1	1
3:39408657,C,G	CDS	missense	c.185C>G	p.Ala62Gly		1	1
3:39408659,C,G	CDS	missense	c.187C>G	p.Arg63Gly		1	1
3:39408659,C,T	CDS	missense	c.187C>T	p.Arg63Cys		4	4
3:39408663,C,T	CDS	missense	c.191C>T	p.Ala64Val		1	1
3:39408665,A,G	CDS	missense	c.193A>G	p.Ile65Val	1	3	4
3:39408672,C,T	CDS	missense	c.200C>T	p.Ala67Val		1	1
3:39408674,A,G	CDS	missense	c.202A>G	p.Ile68Val		1	1
3:39408686,G,A	CDS	missense	c.214G>A	p.Ala72Thr		1	1
3:39408687,C,T	CDS	missense	c.215C>T	p.Ala72Val	1	1	2
3:39408696,G,C	CDS	missense	c.224G>C	p.Ser75Thr		1	1
3:39408702,T,C	CDS	missense	c.230T>C	p.Ile77Thr	4	3	7
3:39408714,A,G	CDS	missense	c.242A>G	p.Asn81Ser		1	1
3:39410773,C,G	CDS	missense	c.272C>G	p.Ala91Gly	1	1	2
3:39410775,G,A	CDS	missense	c.274G>A	p.Ala92Thr	1		1
3:39410782,C,T	CDS	missense	c.281C>T	p.Thr94Ile		1	1
3:39410796,A,G	CDS	missense	c.295A>G	p.Ile99Val	2	4	6
3:39410836,T,C	CDS	missense	c.335T>C	p.Ile112Thr		1	1
3:39410842,C,T	CDS	missense	c.341C>T	p.Ala114Val		1	1
3:39410850,C,T	CDS	missense	c.349C>T	p.Arg117Trp	3	3	6
3:39410851,G,A	CDS	missense	c.350G>A	p.Arg117Gln		1	1
3:39410859,C,T	CDS	missense	c.358C>T	p.Arg120Trp	3	1	4
3:39410860,G,A	CDS	missense	c.359G>A	p.Arg120Gln		1	1
3:39410893,A,C	CDS	missense	c.392A>C	p.His131Pro		781	781
3:39410905,C,T	CDS	missense	c.404C>T	p.Thr135Met	4	3	7
3:39410914,C,T	CDS	missense	c.413C>T	p.Ser138Phe	1	1	2
3:39410931,A,G	CDS	missense	c.431A>G	p.Thr144Ala	1		1
3:39410934,A,G	CDS	missense	c.433A>G	p.Ile145Val	4	10	14
3:39410937,G,T	CDS	missense	c.436G>T	p.Ala146Ser		1	1
3:39410958,C,T	CDS	missense	c.457C>T	p.Pro153Ser		1	1
3:39410959,C,G	CDS	missense	c.458C>G	p.Pro153Arg		1	1
3:39410961,C,G	CDS	missense	c.460C>G	p.Leu154Val	1		1
3:39410964,C,T	CDS	missense	c.463C>T	p.Arg155Cys		3	3
3:39410965,G,A	CDS	missense	c.464G>A	p.Arg155His		6	6
3:39410968,A,G	CDS	missense	c.467A>G	p.Tyr156Cys	2	2	4
3:39410968,A,T	CDS	missense	c.467A>T	p.Tyr156Phe	1	3	4
3:39410976,A,G	CDS	missense	c.475A>G	p.Ile159Val		2	2

3:39411661,G,C	CDS	missense	c.511G>C	p.Val171Leu		1	1
3:39411680,T,C	CDS	missense	c.530T>C	p.Met177Thr		1	1
3:39411681,G,C	CDS	missense	c.531G>C	p.Met177Ile		2	2
3:39411701,G,A	CDS	missense	c.551G>A	p.Arg184His		1	1
3:39411703,A,G	CDS	missense	c.553A>G	p.Met185Val		1	1
3:39411704,T,C	CDS	missense	c.554T>C	p.Met185Thr		2	2
3:39411709,G,A	CDS	missense	c.559G>A	p.Gly187Ser		1	1
3:39411734,G,C	CDS	missense	c.584G>C	p.Trp195Ser		1	1
3:39411739,G,A	CDS	missense	c.589G>A	p.Val197Ile		5	5
3:39411768,T,A	CDS	missense	c.618T>A	p.Asp206Glu	2	1	3
3:39411920,G,A	CDS	missense	c.652G>A	p.Ala218Thr	1		1
3:39411930,C,T	CDS	missense	c.662C>T	p.Ala221Val	1		1
3:39411942,A,G	CDS	missense	c.674A>G	p.Glu225Gly	1		1
3:39411966,C,T	CDS	missense	c.698C>T	p.Ala233Val		1	1
3:39411974,C,T	CDS	missense	c.706C>T	p.Pro236Ser	1	2	3
3:39411984,C,G	CDS	missense	c.716C>G	p.Thr239Ser		1	1
3:39411989,A,G	CDS	missense	c.721A>G	p.Thr241Ala	1		1
3:39412001,G,C	CDS	missense	c.733G>C	p.Val245Leu		1	1
3:39412019,G,A	CDS	missense	c.751G>A	p.Gly251Ser		1	1
3:39412034,T,C	CDS	missense	c.766T>C	p.Ser256Pro		1	1
3:39412038,T,G	CDS	missense	c.770T>G	p.Val257Gly	3	2	5
3:39412044,T,C	CDS	missense	c.776T>C	p.Ile259Thr	1		1
3:39412053,T,A	CDS	missense	c.785T>A	p.Phe262Tyr		1	1
3:39412285,G,A	CDS	missense	c.805G>A	p.Ala269Thr		1	1
3:39412286,C,T	CDS	missense	c.806C>T	p.Ala269Val		1	1
3:39412291,C,T	CDS	missense	c.811C>T	p.Pro271Ser		1	1
3:39412301,A,G	CDS	missense	c.821A>G	p.Glu274Gly	3		3
3:39412309,T,G	CDS	missense	c.829T>G	p.Ser277Ala		2	2
3:39412312,G,A	CDS	missense	c.832G>A	p.Ala278Thr	8	6	14
3:39412313,C,G	CDS	missense	c.833C>G	p.Ala278Gly		1	1
3:39412321,A,G	CDS	missense	c.841A>G	p.Thr281Ala	3	2	5
3:39412325,C,G	CDS	missense	c.845C>G	p.Ala282Gly	1		1
3:39412331,C,T	CDS	missense	c.851C>T	p.Ala284Val	1		1
3:39412333,A,G	CDS	missense	c.853A>G	p.Thr285Ala		1	1
3:39412354,A,G	CDS	missense	c.874A>G	p.Thr292Ala		1	1
3:39412363,T,C	CDS	missense	c.883T>C	p.Ser295Pro	1		1
3:39412364,C,T	CDS	missense	c.884C>T	p.Ser295Phe		1	1

**Table S4: UTR variants from gnomAD and Bravo databases**

Variants (position hg38, reference allele, alternate allele)	location	Consequence	Transcript Consequence (ENST00000301821)	Protein Consequence	Bravo	gnomAD	Total
3:39406696,A,C	5' UTR	5' UTR	c.-102A>C		1		1
3:39406697,G,A	5' UTR	5' UTR	c.-101G>A		8		8
3:39406697,G,T	5' UTR	5' UTR	c.-101G>T		16	3	19
3:39406698,G,A	5' UTR	5' UTR	c.-100G>A		5	2	7
3:39406701,G,A	5' UTR	5' UTR	c.-97G>A		1	1	2
3:39406707,T,C	5' UTR	5' UTR	c.-91T>C		1		1
3:39406708,C,G	5' UTR	5' UTR	c.-90C>G		1		1
3:39406709,C,G	5' UTR	5' UTR	c.-89C>G		6		6
3:39406711,G,A	5' UTR	5' UTR	c.-87G>A		3		3
3:39406711,G,T	5' UTR	5' UTR	c.-87G>T		10	3	13
3:39406712,C,G	5' UTR	5' UTR	c.-86C>G		1		1
3:39406712,C,T	5' UTR	5' UTR	c.-86C>T		19		19
3:39406713,C,A	5' UTR	5' UTR	c.-85C>A		4		4
3:39406713,C,T	5' UTR	5' UTR	c.-85C>T		24267	8643	32910
3:39406714,G,T	5' UTR	5' UTR	c.-84G>T		4		4
3:39406716,C,T	5' UTR	5' UTR	c.-82C>T		1		1
3:39406717,T,G	5' UTR	5' UTR	c.-81T>G		2	2	4
3:39406717,TG,T	5' UTR	5' UTR	c.-80delG		10	2	12
3:39406719,T,C	5' UTR	5' UTR	c.-79T>C		3	1	4
3:39406720,C,CT	5' UTR	5' UTR	c.-74dupT		10	3	13
3:39406720,C,T	5' UTR	5' UTR	c.-78C>T		6		6
3:39406724,T,TC	5' UTR	5' UTR	c.-72dupC		2		2
3:39406725,C,T	5' UTR	5' UTR	c.-73C>T		3		3
3:39406728,T,A	5' UTR	5' UTR	c.-70T>A		2		2
3:39406731,T,A	5' UTR	5' UTR	c.-67T>A		2		2
3:39406734,C,G	5' UTR	5' UTR	c.-64C>G		12	3	15
3:39406734,C,T	5' UTR	5' UTR	c.-64C>T		27	5	32
3:39406737,C,A	5' UTR	5' UTR	c.-61C>A		1		1
3:39406737,C,T	5' UTR	5' UTR	c.-61C>T		272	61	333
3:39406738,A,G	5' UTR	5' UTR	c.-60A>G		9	1	10
3:39406739,G,A	5' UTR	5' UTR	c.-59G>A		4		4
3:39406740,A,G	5' UTR	5' UTR	c.-58A>G		5	1	6
3:39406741,G,A	5' UTR	5' UTR	c.-57G>A		2		2
3:39406742,G,A	5' UTR	5' UTR	c.-56G>A		4	1	5
3:39406743,G,A	5' UTR	5' UTR	c.-55G>A		3	1	4
3:39406744,G,A	5' UTR	5' UTR	c.-54G>A			1	1
3:39406745,T,G	5' UTR	5' UTR	c.-53T>G		1		1
3:39406745,TC,T	5' UTR	5' UTR	c.-51delC		1		1
3:39406746,C,G	5' UTR	5' UTR	c.-52C>G		5	1	6
3:39406746,C,T	5' UTR	5' UTR	c.-52C>T		4	1	5
3:39406747,C,G	5' UTR	5' UTR	c.-51C>G		1		1
3:39406747,C,T	5' UTR	5' UTR	c.-51C>T		1	1	2
3:39406748,A,C	5' UTR	5' UTR	c.-50A>C		3	1	4
3:39406748,A,G	5' UTR	5' UTR	c.-50A>G		37	7	44
3:39406748,A,T	5' UTR	5' UTR	c.-50A>T		10	2	12
3:39406749,T,C	5' UTR	5' UTR	c.-49T>C		1		1
3:39406750,A,G	5' UTR	5' UTR	c.-48A>G		4	1	5
3:39406751,C,A	5' UTR	5' UTR	c.-47C>A		1		1
3:39406751,C,G	5' UTR	5' UTR	c.-47C>G		4		4
3:39406751,C,T	5' UTR	5' UTR	c.-47C>T		5		5
3:39406752,G,A	5' UTR	5' UTR	c.-46G>A		2	1	3
3:39406753,G,T	5' UTR	5' UTR	c.-45G>T		8		8
3:39406754,C,A	5' UTR	5' UTR	c.-44C>A		10	3	13
3:39406754,C,G	5' UTR	5' UTR	c.-44C>G		4	1	5
3:39406754,C,T	5' UTR	5' UTR	c.-44C>T		4	2	6
3:39406756,T,C	5' UTR	5' UTR	c.-42T>C		1		1
3:39406757,T,C	5' UTR	5' UTR	c.-41T>C		15		15
3:39407624,C,A	5' UTR	5' UTR	c.-30C>A		2		2

3:39407624,C,G	5' UTR	5' UTR	c.-30C>G			1	1
3:39407625,C,T	5' UTR	5' UTR	c.-29C>T			1	1
3:39407626,C,T	5' UTR	5' UTR	c.-28C>T		10	24	34
3:39407626,CG,C	5' UTR	5' UTR	c.-27delG		1	5	6
3:39407627,G,A	5' UTR	5' UTR	c.-27G>A		2	3	5
3:39407629,C,T	5' UTR	5' UTR	c.-25C>T		2	9	11
3:39407630,G,A	5' UTR	5' UTR	c.-24G>A			1	1
3:39407634,C,T	5' UTR	5' UTR	c.-20C>T		1	1	2
3:39407635,T,G	5' UTR	5' UTR	c.-19T>G			1	1
3:39407640,G,A	5' UTR	5' UTR	c.-14G>A			1	1
3:39407641,G,A	5' UTR	5' UTR	c.-13G>A			1	1
3:39407645,A,AC	5' UTR	5' UTR	c.-9_-8insC		1		1
3:39407645,AT,A	5' UTR	5' UTR	c.-5delT		1	5	6
3:39407646,T,C	5' UTR	5' UTR	c.-8T>C		30921	120855	151776
3:39407647,T,C	5' UTR	5' UTR	c.-7T>C		1	5	6
3:39407648,T,G	5' UTR	5' UTR	c.-6T>G			1	1
3:39407650,C,T	5' UTR	5' UTR	c.-4C>T		1		1
3:39407651,A,G	5' UTR	5' UTR	c.-3A>G			1	1
3:39407653,A,G	5' UTR	5' UTR	c.-1A>G			5	5
3:39412372,G,A	3' UTR	3' UTR	c.*4G>A			1	1
3:39412372,G,C	3' UTR	3' UTR	c.*4G>C		2	6	8
3:39412373,T,C	3' UTR	3' UTR	c.*5T>C			3	3
3:39412375,C,G	3' UTR	3' UTR	c.*7C>G			1	1
3:39412375,C,T	3' UTR	3' UTR	c.*7C>T			1	1
3:39412377,T,C	3' UTR	3' UTR	c.*9T>C			3	3
3:39412377,T,G	3' UTR	3' UTR	c.*9T>G		2	3	5
3:39412379,C,T	3' UTR	3' UTR	c.*11C>T		2	7	9
3:39412380,A,G	3' UTR	3' UTR	c.*12A>G			9	9
3:39412384,G,A	3' UTR	3' UTR	c.*16G>A		1	1	2
3:39412387,C,T	3' UTR	3' UTR	c.*19C>T		1		1
3:39412389,T,G	3' UTR	3' UTR	c.*21T>G			1	1
3:39412391,A,G	3' UTR	3' UTR	c.*23A>G		1		1
3:39412392,G,A	3' UTR	3' UTR	c.*24G>A		1		1
3:39412403,A,G	3' UTR	3' UTR	c.*35A>G			4	4
3:39412404,A,G	3' UTR	3' UTR	c.*36A>G			1	1
3:39412405,A,G	3' UTR	3' UTR	c.*37A>G			3	3
3:39412406,T,TGGTTGA	3' UTR	3' UTR	c.*39_*44dupGGTTGA			1	1
3:39412420,T,C	3' UTR	3' UTR	c.*52T>C			1	1
3:39412425,A,C	3' UTR	3' UTR	c.*57A>C		1		1
3:39412429,G,A	3' UTR	3' UTR	c.*61G>A			1	1
3:39412431,T,G	3' UTR	3' UTR	c.*63T>G		1		1
3:39412433,C,A	3' UTR	3' UTR	c.*65C>A			1	1
3:39412433,C,G	3' UTR	3' UTR	c.*65C>G		1		1
3:39412445,T,G	3' UTR	3' UTR	c.*77T>G		4		4
3:39412447,C,T	3' UTR	3' UTR	c.*79C>T			1	1
3:39412458,C,G	3' UTR	3' UTR	c.*90C>G		1		1
3:39412474,A,G	3' UTR	3' UTR	c.*106A>G		1	1	2
3:39412481,C,G	3' UTR	3' UTR	c.*113C>G		3	1	4
3:39412501,A,C	3' UTR	3' UTR	c.*133A>C		1		1
3:39412518,T,C	3' UTR	3' UTR	c.*150T>C		27	3	30
3:39412521,C,T	3' UTR	3' UTR	c.*153C>T		115	13	128
3:39412529,G,A	3' UTR	3' UTR	c.*161G>A		2	2	4
3:39412532,G,A	3' UTR	3' UTR	c.*164G>A		1		1
3:39412536,T,C	3' UTR	3' UTR	c.*168T>C		1		1
3:39412536,T,G	3' UTR	3' UTR	c.*168T>G		3		3
3:39412537,T,A	3' UTR	3' UTR	c.*169T>A		1		1
3:39412538,C,A	3' UTR	3' UTR	c.*170C>A		1	1	2
3:39412540,A,G	3' UTR	3' UTR	c.*172A>G			1	1



**Table S5: Splice variants from gnomAD and Bravo databases**

Variants (position hg38, reference allele, alternate allele)	location	Consequence	Transcript Consequence (ENST00000301821)	Protein Consequence	Bravo	gnomAD	Total
3:39407611,CTT,C	intron	splice region	c.-33-9_-33-8delTT			3	3
3:39407613,T,A	intron	splice region	c.-33-8T>A		1	6	7
3:39407613,T,C	intron	splice region	c.-33-8T>C		1		1
3:39407621,A,T	intron	splice region	c.-33A>T			2	2
3:39407794,A,G	intron	splice region	c.133+8A>G		3	1	4
3:39408595,CCCT,C	intron	splice region	c.134-9_134-7delCTC			1	1
3:39408596,C,G	intron	splice region	c.134-10C>G		1		1
3:39408601,C,T	intron	splice region	c.134-5C>T			1	1
3:39408607,C,T	CDS	splice region	c.135C>T	p.Gly45Gly		2	2
3:39408728,T,C	intron	splice region	c.252+4T>C			1	1
3:39410746,C,T	intron	splice region	c.253-8C>T			1	1
3:39410750,A,C	intron	splice region	c.253-4A>C		11	13	24
3:39410750,A,G	intron	splice region	c.253-4A>G			1	1
3:39411006,A,G	intron	splice region	c.498+7A>G			4	4
3:39411007,T,A	intron	splice region	c.498+8T>A		2	14	16
3:39411641,C,A	intron	splice region	c.499-8C>A		1		1
3:39411651,A,C	CDS	splice region	c.501A>C	p.Gly167Gly		2	2
3:39411780,A,G	intron	splice region	c.627+3A>G			1	1
3:39411781,A,G	intron	splice region	c.627+4A>G		1	4	5
3:39411783,C,T	intron	splice region	c.627+6C>T			1	1
3:39411785,T,C	intron	splice region	c.627+8T>C			2	2
3:39411891,C,T	intron	splice region	c.628-5C>T			1	1
3:39411893,C,G	intron	splice region	c.628-3C>G		1	9	10
3:39411898,T,C	CDS	splice region	c.630T>C	p.Ile210Ile		3	3
3:39412065,T,C	intron	splice region	c.793+4T>C			1	1
3:39412066,G,A	intron	splice region	c.793+5G>A			4	4
3:39412068,A,G	intron	splice region	c.793+7A>G		2		2
3:39412261,ATTCTC,A	intron	splice region	c.794-10_794-6delCTCTT		1	1	2
3:39412267,T,G	intron	splice region	c.794-7T>G			1	1

**Table S6: Prediction of the impact of 5'UTR variants on the splicing site by NNSPLICE 0.9**

	Donor	Acc.	Donor	Acc.	Donor	Acc.	Donor	Acc.	Acc.	Donor	Donor	Acc.	Acc.	Donor	Acc.	Donor	Donor	Donor	position
Position cDNA	137	157	207	260	455	542	583	641	679	884	922	979	1076	1159	1178	1194	1227	1238	
c.-61C>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-61
c.-61C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-61
c.-60A>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-60
c.-59G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-59
c.-58A>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-58
c.-57G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-57
c.-56G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-56
c.-55G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-55
c.-54G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-54
c.-53T>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-53
c.-52C>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-52
c.-52C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-52
c.-51C>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-51
c.-51C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-51
c.-51delC	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-51
c.-50A>C	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-50
c.-50A>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-50
c.-50A>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-50
c.-49T>C	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-49
c.-48A>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-48
c.-47C>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-47
c.-47C>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-47
c.-47C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-47
c.-46G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-46
c.-45G>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-45
c.-44C>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-44
c.-44C>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-44
c.-44C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-44
c.-42T>C	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-42
c.-41T>C	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-41
<b>c.-39_-34+3del</b>		0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-39
<b>c.-34+5G&gt;C</b>		0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-34
c.-33-1G>C	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60		0.58	0.90	0.57	0.96	0.84	0.77	-33
c.-33-2A>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60		0.58	0.90	0.57	0.96	0.84	0.77	-33
c.-30C>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.92	0.58	0.90	0.57	0.96	0.84	0.77	-30
c.-30C>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.94	0.58	0.90	0.57	0.96	0.84	0.77	-30
c.-29C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-29
c.-28C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.87	0.58	0.90	0.57	0.96	0.84	0.77	-28
c.-27delG	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.97	0.58	0.90	0.57	0.96	0.84	0.77	-27
c.-27G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.87	0.58	0.90	0.57	0.96	0.84	0.77	-27
c.-25C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.90	0.58	0.90	0.57	0.96	0.84	0.77	-25
c.-24G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.90	0.58	0.90	0.57	0.96	0.84	0.77	-24
c.-20C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.94	0.58	0.90	0.57	0.96	0.84	0.77	-20
c.-19T>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.87	0.58	0.90	0.57	0.96	0.84	0.77	-19
c.-14G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.94	0.58	0.90	0.57	0.96	0.84	0.77	-14
c.-13G>A	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-13
c.-9_-8ins	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-9
c.-8T>C	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-8
c.-7T>C	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-7
c.-6T>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-6
c.-5delT	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-5
c.-4C>T	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-4
c.-3A>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-3
c.-1A>G	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	-1
WT	0.98	0.85	0.98	0.71	0.99	0.52	0.59	0.98	0.83	0.85	0.60	0.91	0.58	0.90	0.57	0.96	0.84	0.77	

In red, variants that impair the canonical splicing site, in green, variants that do not impair the canonical splicing. The two canonical splicing sites (donor site in position 137 and acceptor site in position 979) are indicated in grey. For reference, the WT sequence has been indicated in the last row. The two variants found in the ICA cohort are labelled in bold and red.

