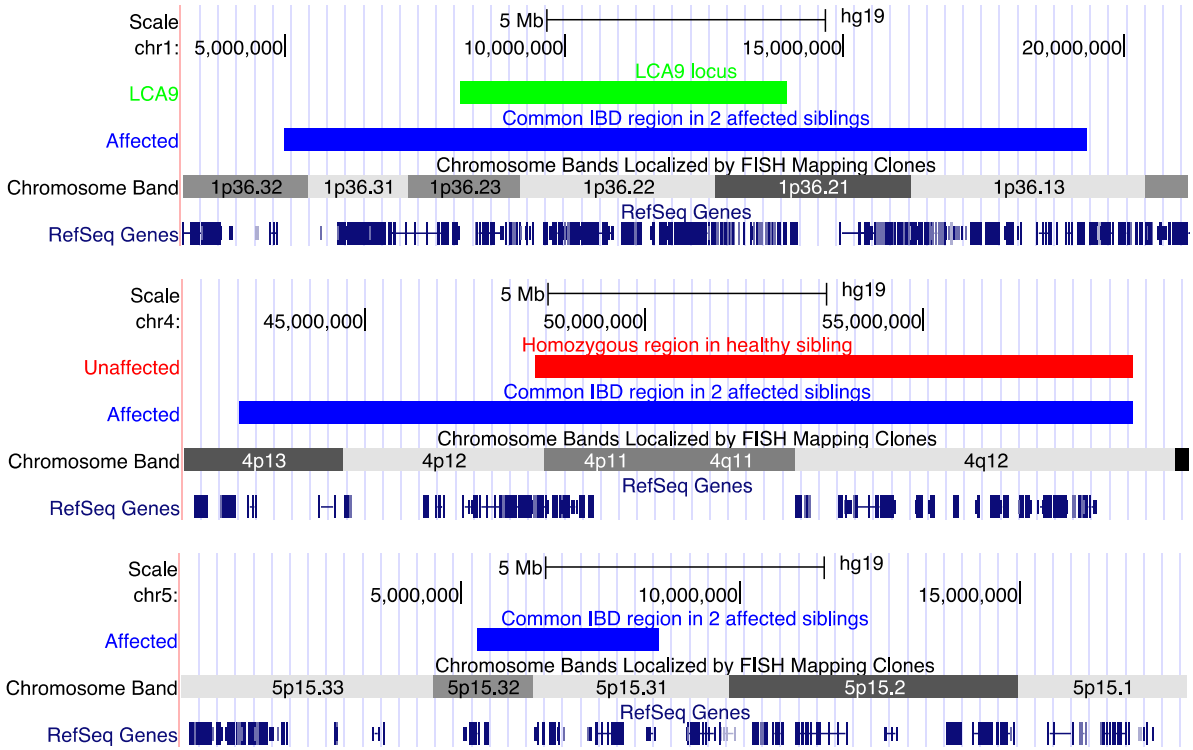
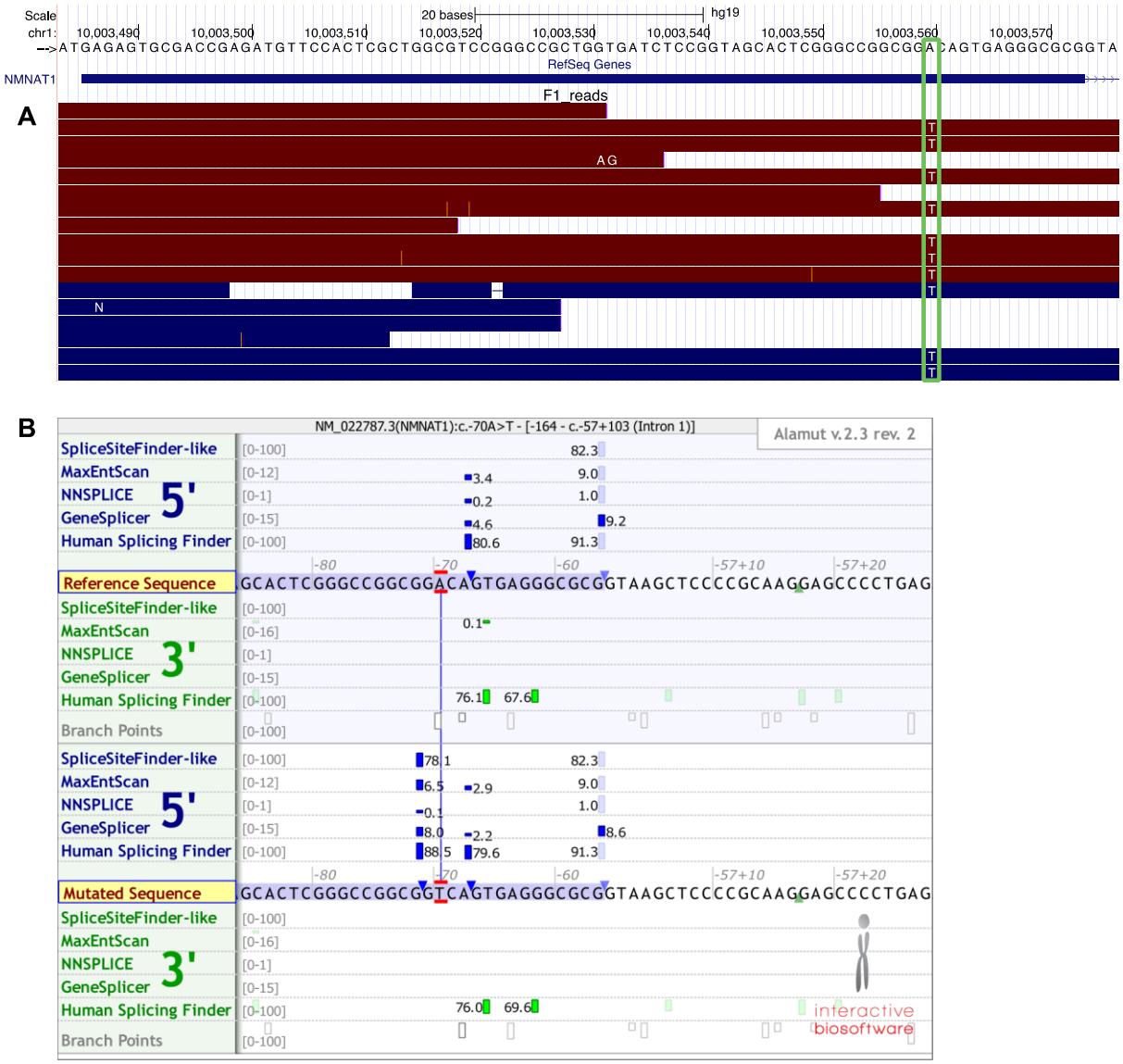
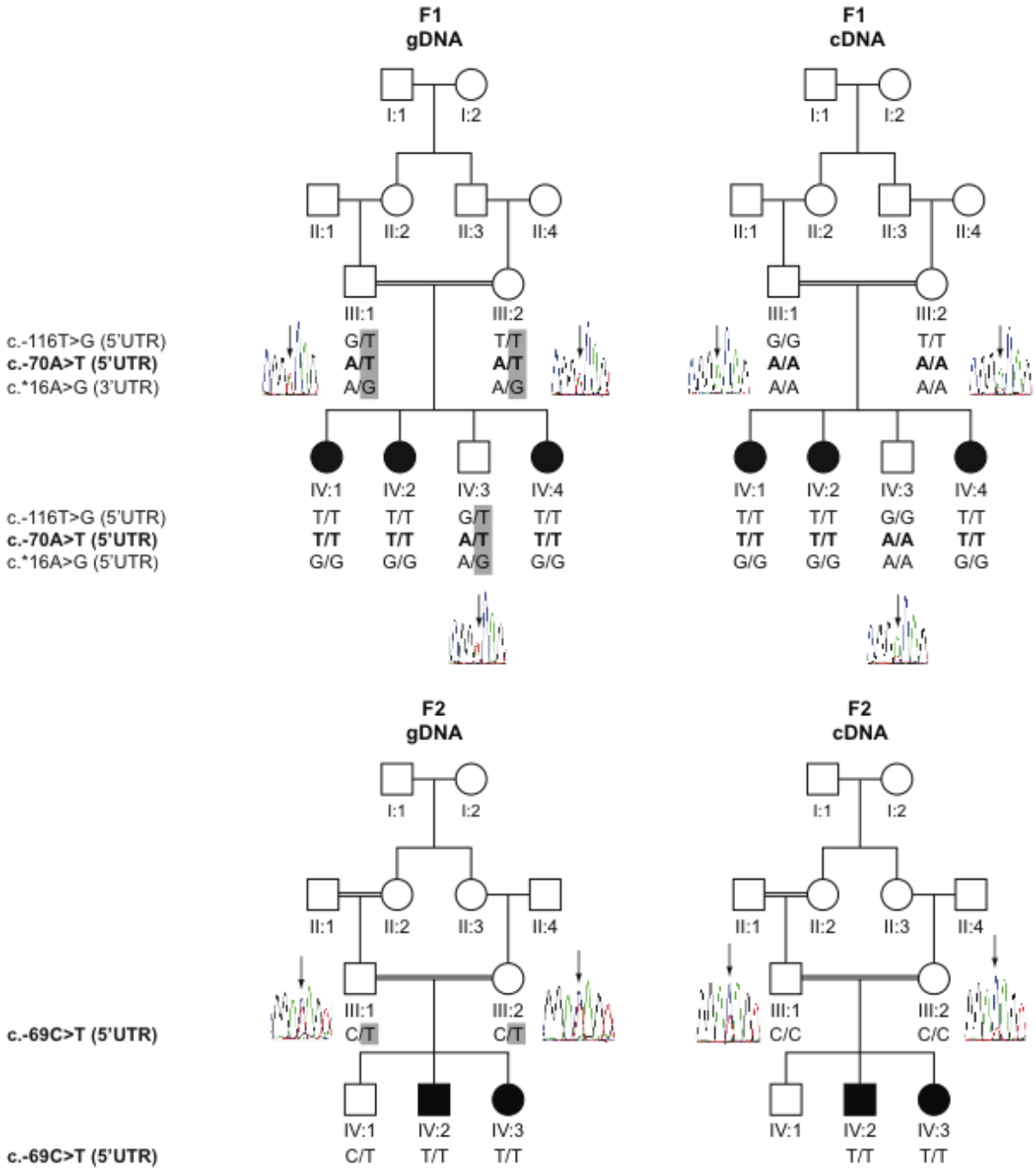
Supporting Information



Supp. Figure S1. UCSC Genome Browser overview of IBD regions identified in F1. Genome-wide identity-by-descent (IBD) mapping identified three regions larger than 3Mb that were common between two affected siblings of F1 (blue bars) and (partially) absent in the healthy sibling (red bar). The region on chromosome 1 harbors the LCA9 locus (green bar).

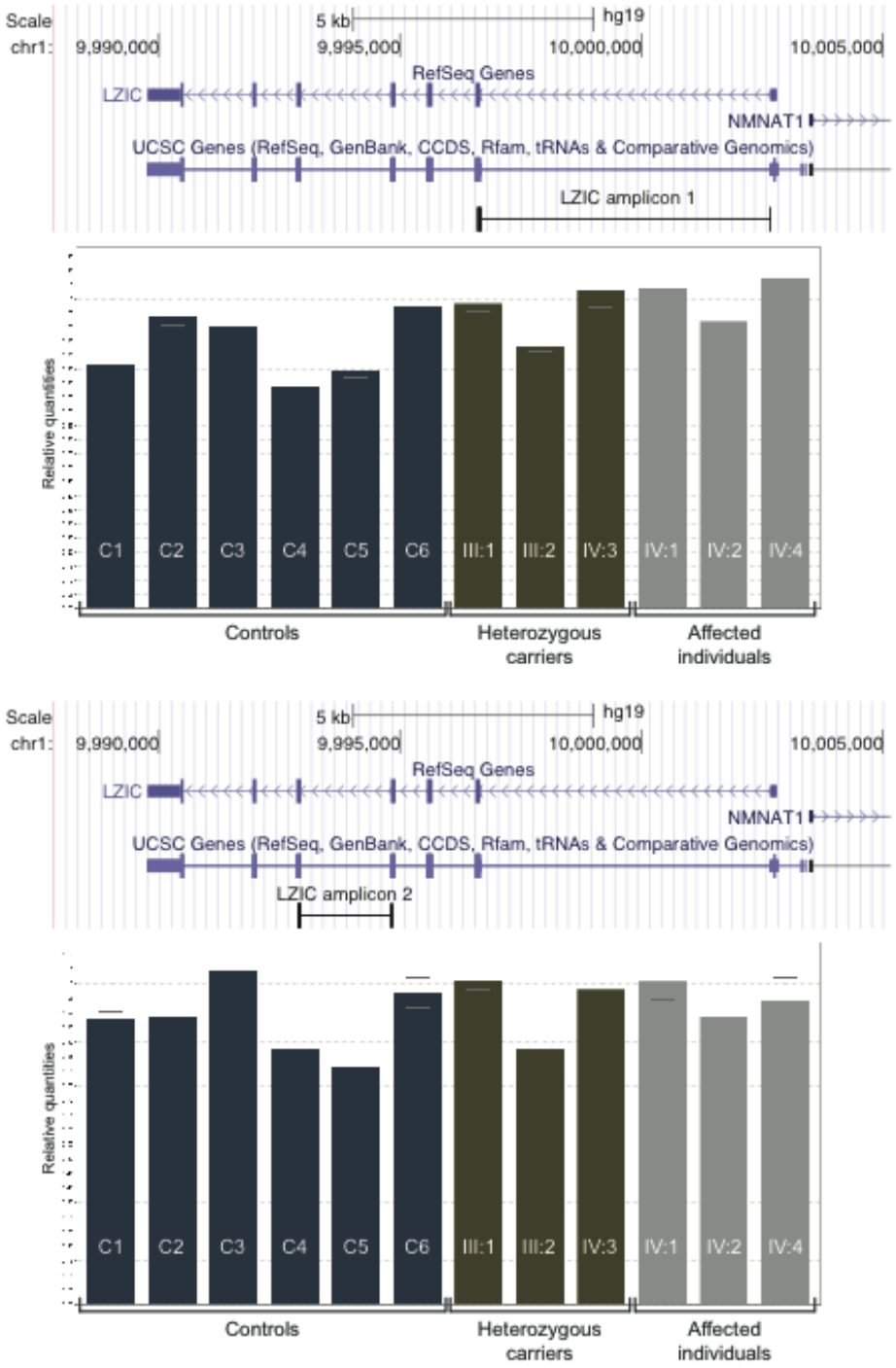


Supp. Figure S2. Massively parallel sequencing reads and splice site predictions for c.-70A>T. **A**) UCSC Genome Browser view of the sequencing reads observed in the proband of F1 covering exon 1 of the *NMNAT1* gene. The homozygous c.-70A>T mutation is indicated with a green box. **B**) Splice site predictions of the Alamut interpretation software (v2.3 – default settings used), predicting the introduction of a novel donor splice site for c.-70A>T (indicated with red bars). This potential new donor splice site is still weaker than the canonical donor splice site of which the strength remains unchanged according to 4/5 prediction programs.

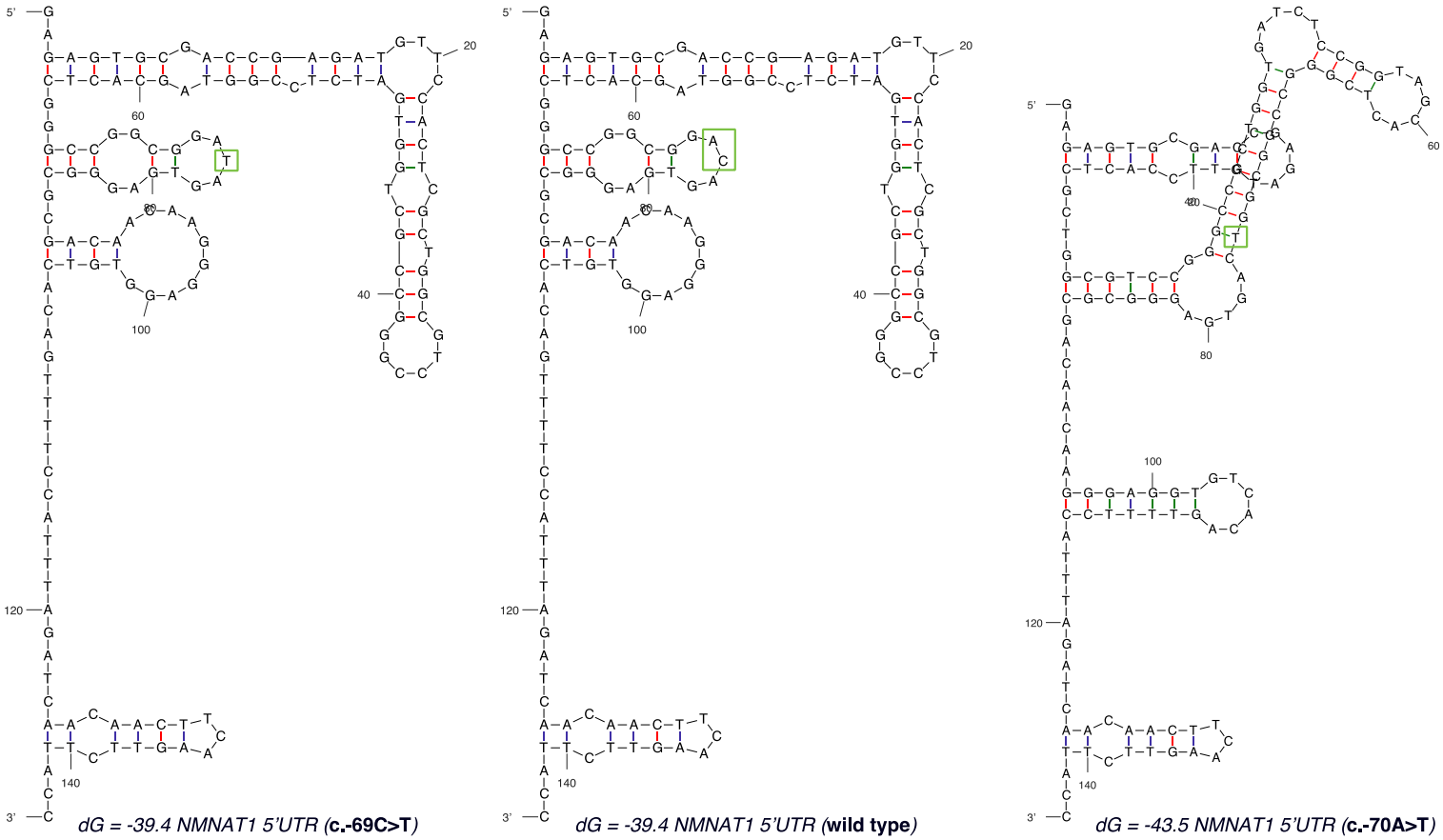


Supp. Figure S3. Segregation analysis of c.-70A>T and c.-69C>T on gDNA and cDNA level in F1 and F2. gDNA and cDNA sequencing in both families showed allele-specific expression at the cDNA level. The 5'UTR mutations and a flanking 5'UTR and 3'UTR SNP are represented at the left. The disease haplotype or allele is highlighted in grey. Haplotype analysis of three *NMNAT1* variants found at the gDNA level displayed allele-specific expression at the cDNA level in the parents and the healthy sib of F1. This result was obtained in cells treated with (shown here) and without puromycine (used to block nonsense-mediated decay), suggesting that the genetic defect has an effect prior to translation. For F2, only untreated lymphocyte cultures

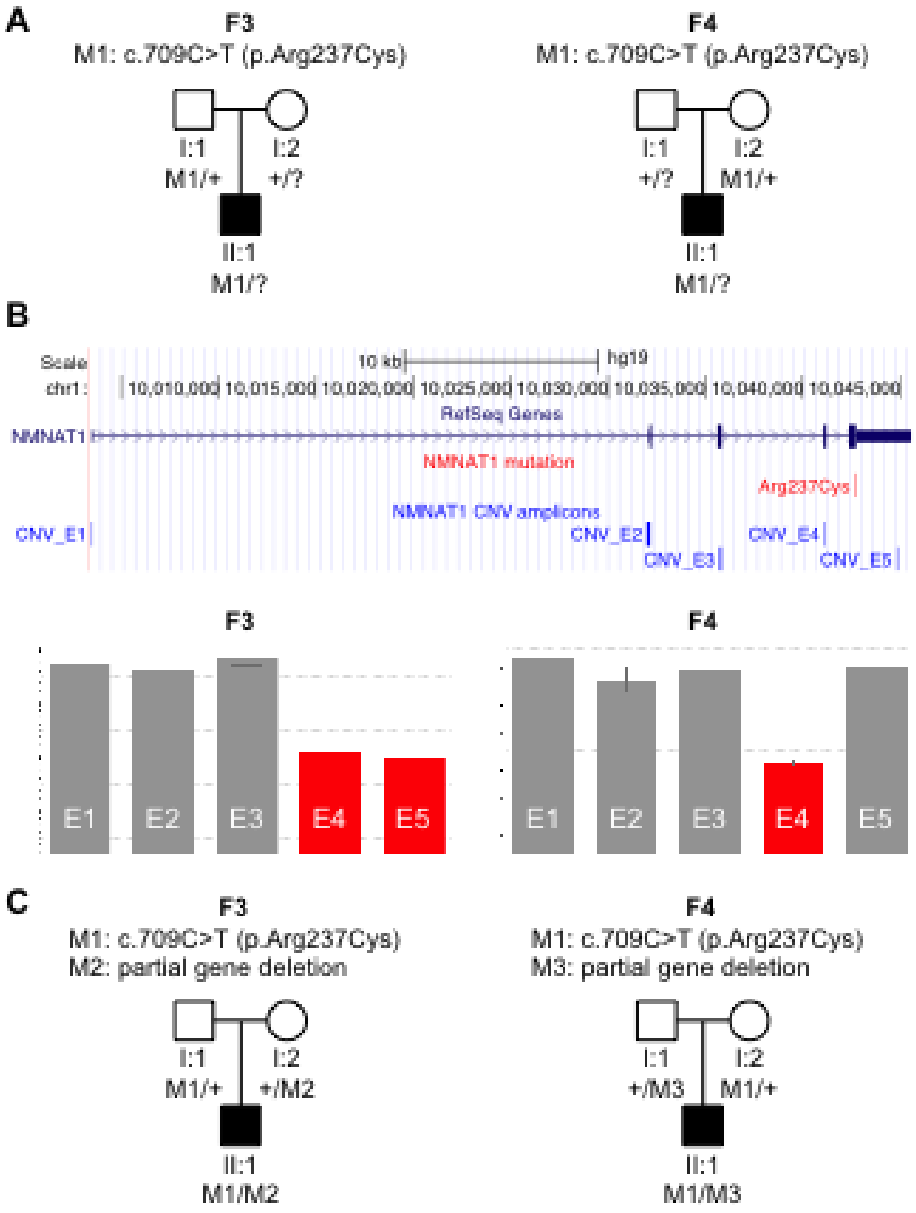
were available. As in F1, cDNA sequencing showed allele-specific expression in both parents. In both families, superposition of low sequence traces could be seen that represent residual mutant mRNA.



Supp. Figure S4. *LZIC* expression analysis in F1. qPCR analysis of two distinct intron-spanning amplicons for the *LZIC* gene, *LZIC* amplicon 1 and *LZIC* amplicon 2, revealed no difference in mRNA expression between the affected individuals of F1 (IV:1, IV:2, IV:4) in comparison with healthy control individuals (C1-C6). The upper panels contain UCSC Genome Browser views of the location of the amplicons, the lower panels show the results of the expression analysis.



Supp. Figure S5. Secondary structure prediction for the wild type and mutated *NMNAT1* 5'UTR sequences. Secondary structure analysis of the *NMNAT1* 5'UTR sequence using UNAFold (Two-state Folding) (default settings used) (Markham and Zuker 2008) revealed a different folding for the c.-70A>T mutation in comparison with the wild type *NMNAT1* 5'UTR. The location of the mutated nucleotides has been indicated with a green box.



Supp. Figure S6. Identification of a heterozygous *NMNAT1* deletion in F3 and F4. **A**) Segregation analysis of the c.709C>T p.(Arg237Cys) mutation in F3 and F4 suggested hemizygosity for this mutation in the probands. **B**) Copy number variation (CNV) analysis using qPCR with one amplicon per exon showed a heterozygous deletion of exon 4 and 5 in the proband of F3 and a heterozygous deletion of exon 4 in the proband of F4. Based on these findings, the 3' end of the deletion in F4 should be located between the p.Arg237Cys mutation and the amplicon for the fifth exon, located in the 3'UTR. **C)** Segregation analysis of the c.709C>T p.(Arg237Cys) mutation (M1) and both partial deletions (M2 and M3), performed using qPCR analysis of one amplicon per exon. Abbreviations used: F: family; M: mutant allele, +: wild-type allele; ?: unknown allele; CNV: copy number variation; E: exon.

Supp. Table S1. Homozygous regions >3 Mb and common between 2 affected siblings of F1

Chr	Centromeric boundary (hg19)	Telomeric boundary (hg19)	Length (bp)	No of SNPs	No of genes	Healthy sibling
4	42,748,945	58,760,043	16,011,098	1,155	109	Partially hom
1	4,984,667	19,332,076	14,347,409	927	272	Het
1	240,611,614	244,938,498	4,326,884	441	52	Hom
5	5,281,887	8,527,844	3,245,957	476	23	Het

The presence of an identical homozygous haplotype in the healthy sibling for the third region (chr1) and a partially overlapping homozygous haplotype for the first region (chr4) resulted in exclusion and reduction of these regions, respectively. Abbreviations used: chr: chromosome; bp: base pairs; no: number; hom: homozygous; het: heterozygous.

Supp. Table S2. Overview of all *NMNAT1* variants identified in the proband of F1

Intron/exon	cDNA (protein)	Coverage	VAF	rs number (MAF)	Splicing
Upstream	c.-173G>C (p.=)	12	100	rs2297843 (C=0.152)	No effect
E1 (5'UTR)	c.-70A>T (p.?)	10x	100	/	New donor site?
I2	c.115+354A>G (p.?)	8x	100	/	No effect
I2	c.116-177T>C (p.?)	19x	100	rs6657844 (T=0.037)	No effect
I3	c.299+176G>A (p.?)	25x	84	rs12082318 (A=0.133)	No effect
I4	c.440-226C>T (p.?)	18x	100	rs11121509 (T=0.133)	New donor site?
E5 (3'UTR)	c.*16A>G (p.?)	43x	98	rs12086725 (G=0.008)	No effect
E5 (3'UTR)	c.*699A>G (p.?)	6x	100	rs6664208 (G=0.132)	No effect
E5 (3'UTR)	c.*2393C>G (p.?)	62x	98	rs4073852 (C=0.221)	New acceptor site?

The *NMNAT1* coding region showed a minimal coverage of 20x. The 5'UTR was covered at least 8x, whereas 82% of bases in the 3'UTR had a minimal coverage of at least 10x. *In silico* predictions for splicing were performed using the Alamut v2.3 interpretation software (default settings). For the c.*16A>G variant, a minor allele frequency of 6.8% was reported in the sub-Saharan African population (Hapmap-YRI). Abbreviations used: I: intron; E: exon; UTR: untranslated region; VAF: variant allele frequency; MAF: minor allele frequency (dbSNP138); ND: no data.

Supp. Table S3. *NMNAT1* coding mutations identified in this study

Family	***Allele 1***				***Allele 2***			
	Exon	**cDNA**	**Protein**	**Reference**	**Exon/intron**	**cDNA**	**Protein**	**Reference**
F5	E5	c.769G>A	p.Glu257Lys	(Chiang et al., 2012; Falk et al., 2012; Koenekoop et al., 2012; Perrault et al., 2012) rs150726175	E4	c.364del	p.Arg122fs	(Perrault et al., 2012)
F6	E5	c.769G>A	p.Glu257Lys	(Chiang et al., 2012; Falk et al., 2012; Koenekoop et al., 2012; Perrault et al., 2012) rs150726175	E5	c.542A>G	p.Tyr181Cys	(Perrault et al., 2012)
F7	E5	c.769G>A	p.Glu257Lys	(Chiang et al., 2012; Falk et	E5	c.679C>T	p.Arg227Trp	novel

Family	***Allele 1***				***Allele 2***			
	Exon	**cDNA**	**Protein**	**Reference**	**Exon/intron**	**cDNA**	**Protein**	**Reference**
				al., 2012; Koenekoop et al., 2012; Perrault et al., 2012) rs150726175				
F8	E5	c.562C>T	p.Arg188Trp	novel	-	-	-	-
F9	E2	c.53A>G	p.Asn18Ser	(Siemiatkowska et al., 2014), UV	-	-	-	-

Coding mutations in *NMNAT1* were identified in three families with LCA (F5, F6 and F7), all sharing the common c.769G>A p.(Glu257Lys) mutation (Chiang et al., 2012; Falk et al., 2012; Koenekoop et al., 2012; Perrault et al., 2012). The probands of F5 and F6 are compound heterozygous for known mutations: c.769G>A p.(Glu257Lys) as the first allele and c.364del p.(Arg122fs) (Perrault et al., 2012) and c.542A>G p.(Tyr181Cys) (Perrault et al., 2012) as the second allele, respectively. The proband of F7 is compound heterozygous for c.769G>A p.(Glu257Lys) and a novel missense mutation, c.679C>T p.(Arg227Trp), which affects a highly conserved amino acid and is predicted to affect protein function by SIFT, Align GVGD and PolyPhen-2 (Alamut v2.3 interpretation software, default settings used). The p.Glu257Lys mutation was also detected in a heterozygous state in 1/179 (0.6%) Belgian controls.

Finally, two probands were found to carry a single heterozygous missense variant: c.562C>T p.(Arg188Trp) (F8) and c.53A>G p.(Asn18Ser) (F9). The Arg188 and Asn18 residues are highly and moderately conserved considering an alignment of sequences from 11 species, respectively. The p.Arg188Trp variant is predicted to affect protein function by PolyPhen-2, SIFT and Align GVGD. The p.Asn18Ser variant is predicted to be probably damaging by PolyPhen-2 but tolerated by SIFT and Align GVGD (Alamut v2.3 interpretation software, default settings used). In addition, this variant was recently described together with p.Asp158His in an LCA

patient (Siemiatkowska et al., 2014). No phenotype data are available for the probands of F8 and F9. In both individuals, Sanger sequencing of the 5'UTR did not reveal a second mutation. In addition, qPCR analysis in F9 did not reveal a CNV. For F8, no DNA was available anymore to perform CNV analysis.

Abbreviations used: F: family; E: exon. Variants were designated as "unclassified variant (UV)" if no consensus was seen in all prediction programs used.

Supp. Table S4. *In silico* evaluation of the *NMNAT1* 5'UTR variants

gDNA/cDNA level	Mechanism	Prediction tool	Results	
			c.-70A>T (F1)	**c.-69C>T (F2)**
cDNA	New uORF	NCBI ORF finder	/	/
gDNA	Transcription factor binding profiles	JASPAR 5.0 (core vertebrata) (Mathelier et al., 2014)	Loss of (score): - **Mafb (5.9)** - **Sox6 (5.4)** - **Sox17 (7.9)** - **SOX10 (5.6)** Gain of (score): - Meis1 (6.2) - Pax2 (4.4) - NFE2L1::MafG (5.1)	Loss of (score): - THAP1 (5.4) - **Mafb (5.9)** - **Sox6 (5.4)** - **Sox17 (7.9)** - **SOX10 (5.6)** - Pax2 (4.6) Gain of (score): - ZNF354C (4.8)
cDNA	Functional RNA motifs and sites	RegRNA 2.0 (Chang et al., 2013)	Gain of: - LXRalpha:RXRalpha	Gain of: - GATA2 - Musashi binding element

Differences between the mutations and the wild type sequence are indicated in bold. Default settings were used. Abbreviations used: uORF: upstream open reading frame; /: no difference with wild type.

Supp. Table S5. G-quadruplexes predictions for the *NMNAT1* 5'UTR mutations

	gDNA/cDNA	**Length**	**QGRS** (Mathelier et al., 2014)	**G-Score**
Wild type	gDNA/cDNA	30	GGGCCGCTGGTGATCTCCGGTAGCACTCGG	19
c.-69C>T	gDNA/cDNA	30	GGGCCGCTGGTGATCTCCGGTAGCACTCGG	19
c.-70A>T	gDNA/cDNA	30	GGGCCGCTGGTGATCTCCGGTAGCACTCGG	19
Wild type	gDNA	20	GGCGG**AC**AGTGAGGGCGCGG	**15**
c.-69C>T	gDNA	20	GGCGGA**T**AGTGAGGGCGCGG	**15**
c.-70A>T	gDNA	20	GGCGG**T**CAGTGAGGGCGCGG	**15**
Wild type	cDNA	29	GG**AC**AGTGAGGGCGCGACAACAAGGGAGG	**12**
c.-69C>T	cDNA	29	GGA**T**AGTGAGGGCGCGACAACAAGGGAGG	**12**
c.-70A>T	cDNA	29	GG**T**CAGTGAGGGCGCGACAACAAGGGAGG	**12**

The location of both mutations in the G-quadruplexes is indicated in bold. Default settings were used.

Supp. Table S6. Clinical findings of the probands of F3 and F4

Individual	**F3**	**F4**
Origin	Japan	Japan
Gender	male	male
Age of diagnosis	3 months	3 months
Nystagmus	+	+
Eyepoking	+	+
Enopthalmos	+	+
ERG (age)	NR (0)	NR (0)
BCVA (age)	0.03*(0)	LP
Refraction (age)	1.31** (0)	7.32** (0)
VF (age)	NE (4)	NE (4)
Color vision	NE (4)	NE (4)
Night blindness	+	+
Photophobia	-	-
Fundus aspect	CD	CD
Strabismus	-	-
Cataract	-	-
Keratoconus	-	-

Visual acuity was measured with Teller acuity cards. Refractometry was performed using a hand held autorefractor (Retinomax2, Nikon Corporation, Tokyo, Japan) after cycloplegia. Fundoscopy was performed after dilatation. Full field ERG was recorded from both eyes respectively after 20 minutes of dark adaptation before recording, under topical anesthesia with after instilling a single drop of 0.4% oxybuprocaine for the contact lens electrodes. Abbreviations used: ERG: electroretinogram; BCVA: best corrected visual acuity; VF: visual field; ODS: both eyes; NR: non recordable; *: Teller acuity cards; LP: light perception; **: average of ODS spherical equivalent; CD: circular degeneration of the macula.

Supp. Table S7. Bioinformatic analysis of the *NMNAT1* deletion breakpoints

Family	Start (hg19)	End (hg19)	Size (bp)	Micro-homology (bp)	Proximal breakpoint region			Distal breakpoint region			Sequence identity between repetitive elements
					Repetitive element	Number of sequence motifs	Number of non-B DNA conformation prediction motifs	Repetitive element	Number of sequence motifs	Number of non-B DNA conformation prediction motifs	
F3	10,039,762	10,056,272	16,510	34	*AluY*	6	1	*AluY*	17	1	92%
F4	10,038,284	10,043,035	4,751	10	*AluSx*	6	0	*AluSc*	12	0	88%

Non-B DNA conformations (with the exception of oligo(G)n tracts) should be located at both sides of the junction or overlapping the junction. Default settings were used. Abbreviations used: F: family; bp: base pairs.

SUPP. REFERENCES

Chang, T. H., H. Y. Huang, J. B. Hsu, S. L. Weng, J. T. Horng and H. D. Huang. 2013. An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. BMC Bioinformatics 14 Suppl 2:S4.

Chiang, P. W., J. Wang, Y. Chen, Q. Fu, J. Zhong, X. Yi, R. Wu, H. Gan, Y. Shi, C. Barnett, D. Wheaton, M. Day, et al. 2012. Exome sequencing identifies NMNAT1 mutations as a cause of Leber congenital amaurosis. Nat Genet 44:972-974.

Falk, M. J., Q. Zhang, E. Nakamaru-Ogiso, C. Kannabiran, Z. Fonseca-Kelly, C. Chakarova, I. Audo, D. S. Mackay, C. Zeitz, A. D. Borman, M. Staniszewska, R. Shukla, et al. 2012. NMNAT1 mutations cause Leber congenital amaurosis. Nat Genet 44:1040-1045.

Koenekoop, R. K., H. Wang, J. Majewski, X. Wang, I. Lopez, H. Ren, Y. Chen, Y. Li, G. A. Fishman, M. Genead, J. Schwartzentruber, N. Solanki, et al. 2012. Mutations in NMNAT1 cause Leber congenital amaurosis and identify a new disease pathway for retinal degeneration. Nat Genet 44:1035-1039.

Markham, N. R. and M. Zuker. 2008. UNAFold: software for nucleic acid folding and hybridization. Methods Mol Biol 453:3-31.

Mathelier, A., X. Zhao, A. W. Zhang, F. Parcy, R. Worsley-Hunt, D. J. Arenillas, S. Buchman, C. Y. Chen, A. Chou, H. Ienasescu, J. Lim, C. Shyr, et al. 2014. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res 42:D142-147.

Perrault, I., S. Hanein, X. Zanlonghi, V. Serre, M. Nicouleau, S. Defoort-Delhemmes, N. Delphin, L. Fares-Taie, S. Gerber, O. Xerri, C. Edelson, A. Goldenberg, et al. 2012. Mutations in NMNAT1 cause Leber congenital amaurosis with early-onset severe macular and optic atrophy. Nat Genet 44:975-977.

Siemiatkowska, A. M., L. I. van den Born, M. M. van Genderen, M. Bertelsen, D. Zobor, K. Rohrschneider, R. A. van Huet, S. Nurohmah, B. J. Klevering, S. Kohl, S. M. Faradz, T. Rosenberg, et al. 2014. Novel compound heterozygous NMNAT1 variants associated with Leber congenital amaurosis. Mol Vis 20:753-759.