An *Alu*-mediated duplication in *NMNAT1*, involved in NAD biosynthesis, causes a novel syndrome, SHILCA, affecting multiple tissues and organs

Nicola Bedoni[1,2‡]; Mathieu Quinodoz[1,3,4,5‡]; Michele Pinelli[6,7*]; Gerarda Cappuccio[6,7]; Annalaura Torella[6,8]; Vincenzo Nigro[6,8]; Francesco Testa[9]; Francesca Simonelli[9]; TUDP (Telethon Undiagnosed Disease Program); Marta Corton[10,11]; Susanna Lualdi[12]; Federica Lanza[12]; Giovanni Morana[13]; Carmen Ayuso[10,11]; Maja Di Rocco[12]; Mirella Filocamo[12]; Sandro Banfi[6,8]; Nicola Brunetti-Pierri[6,7]; Andrea Superti-Furga[2‡]; Carlo Rivolta[3,4,5‡*]

[1]Department of Computational Biology, University of Lausanne, Lausanne, Switzerland.
[2]Division of Genetic Medicine, Lausanne University Hospital, Lausanne, Switzerland.
[3]Department of Genetics and Genome Biology, University of Leicester, Leicester, United Kingdom.
[4]Institute of Molecular and Clinical Ophthalmology Basel, Basel, Switzerland.
[5]Department of Ophthalmology, University of Basel, Basel, Switzerland.
[6]Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli (Naples), Italy.
[75]Department of Translational Medicine, Section of Pediatrics, Federico II University, Naples, Italy.
[8]Medical Genetics, Department of Precision Medicine, University of Campania “Luigi Vanvitelli”, Naples, Italy.
[9]Eye Clinic, Multidisciplinary Department of Medical, Surgical and Dental Sciences, University of Campania “Luigi Vanvitelli”, Naples, Italy.
[10]Department of Genetics, Instituto de Investigación Sanitaria – Fundación Jiménez Díaz, University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain.
[11]Center for Biomedical Network Research on Rare Diseases (CIBERER), ISCIII, Madrid, Spain.
[12]Laboratorio di Genetica Molecolare e Biobanche, Istituto G. Gaslini, Genova, Italy.
[13]Neuroradiology Unit, IRCCS Istituto Giannina Gaslini, Genova.

‡Equal contribution. *Correspondence.

Correspondence:

Carlo Rivolta; Institute of Molecular and Clinical Ophthalmology Basel (IOB), Mittlere Strasse 91, 4031 Basel; email: carlo.rivolta@iob.ch; phone: +41-43-215-2795

Abstract

We investigated the genetic origin of the phenotype of three children from two unrelated Italian families presenting with a previously-unrecognized, seemingly autosomal recessive disorder that included a severe form of spondylo-epiphyseal dysplasia, sensorineural hearing loss, intellectual disability, and Leber congenital amaurosis (SHILCA), as well as some brain anomalies that were visible at the MRI. Autozygome-based analysis showed that these children shared a 4.6 Mb region of homozygosity on chromosome 1, with an identical haplotype. Nonetheless, whole-exome sequencing failed to identify any shared rare coding variants, in this region or elsewhere. We then determined the transcriptome of patients' fibroblasts by RNA sequencing, followed by additional whole-genome sequencing experiments. Gene expression analysis revealed a 4-fold downregulation of the gene *NMNAT1*, previously associated with Leber congenital amaurosis (LCA) and residing in the shared autozygous interval. Short- and long-read whole-genome sequencing highlighted a duplication involving 2 out of the 5 exons of *NMNAT1* main isoform (NM_022787.3), leading to the production of aberrant mRNAs. No other pathogenic variants in *NMNAT1* have been previously shown to cause non-syndromic LCA. However, no patient with null biallelic variants has ever been described, and murine *Nmnat1* knockouts show embryonic lethality. We hypothesize that complete absence of NMNAT1 activity is not compatible with life. The rearrangement found in our cases, presumably causing a strong but not complete reduction of enzymatic activity, may therefore result in an intermediate syndromic phenotype, between non-syndromic LCA and lethality.

Introduction

Leber congenital amaurosis (LCA) is a recessive form of hereditary retinal degeneration, with an incidence between 1 in 33,000 and 1 in 100,000 individuals (1-3). The non-syndromic form is caused by mutations in at least 19 genes, according to OMIM (phenotypic series PS204000; accessed on March 2020). Its distinctive feature is the very early age of onset, leading in most cases to retinal blindness within the first years of life. Additional signs include nystagmus and rapidly-progressing macular degeneration leading to central atrophy (4).

Syndromic LCA is much rarer and limited to Senior-Loken and Joubert syndrome (four genes), a condition labelled as “retinal degeneration with or without additional developmental anomalies” (5), as well as a recently-described multisystemic phenotype including LCA, named Liberfarb syndrome (6).

NMNAT1 has been associated with LCA for the first time in 2012 (4, 7-9) and, since then, pathogenic variants in this gene have been identified in several individuals (10-15), accounting for more than 5% of all forms of inherited retinal degenerations (3). NMNAT enzymes (NMNAT1, NMNAT2 and NMNAT3) play a key role in the biosynthesis of nicotinamide adenine dinucleotide (NAD) (16), catalyzing the condensation of nicotinamide mononucleotide (NMN) or nicotinic acid mononucleotide (NaMN) to form NAD^+ or nicotinic acid dinucleotide (NaAD). The NAD^+/NADH coenzyme is involved in hundreds of biochemical reaction and metabolic pathways and is essential for survival of both eukaryotic and prokaryotic cells (17). NMNAT1 is the nuclear form of the nicotinamide mononucleotide adenylyltransferases (NMNATs), whereas forms 2 and 3 exert their function in the Golgi complex and mitochondrion, respectively (18). In line with the essential role of NAD^+/NADH for basic cell metabolism, *Nmnat1* knockout mice display embryonic lethality. In addition, no occurrence of biallelic *null* variants was ever observed in patients carrying homozygous or compound heterozygous *NMNAT1*

mutations among the 48 identified to date (10-14), suggesting that retinal degeneration results from reduced (although not completely absent) function of the NMNAT1 enzyme.

In this study, we identify a likely *Alu*-mediated genomic rearrangement encompassing *NMNAT1*, in subjects carrying a previously-undescribed phenotype characterized by the combination of LCA and additional clinical features involving the ear, the brain, and the skeleton. This rearrangement leads to production of both aberrant and canonical (although at a low level) *NMNAT1* mRNA forms, resulting in syndromic retinal blindness, thus expanding the spectrum of phenotypic abnormalities due to defects in this gene.

Results

Patients and clinical presentations

Three children from two families presented with overlapping clinical features, which were seemingly never described previously as part of the same phenotype (**Fig. 1A**). They all showed early onset retinal degeneration, diagnosed as LCA, sensorineural hearing loss, short stature due to a spondylo-epiphyseal dysplasia, developmental delay, and brain magnetic resonance imaging (MRI) abnormalities including delayed myelinization, white matter hyperintensity, and cerebellar hypoplasia (**Fig. 2**).

P1 and P2 were two siblings, both presenting with developmental delay, vision loss, and ataxia. They were born at term of a normal gestation to healthy non-consanguineous parents, both originating from Southern Italy. The older brother (P1, **Fig. 1A** and **2A**) was noted to have vision defects since the first months of life because of lack of visual tracking and nystagmus. At 5 years of age his Best Corrected Visual Acuity (BCVA) was light perception. Optic disc pallor, marked attenuation of retinal vessels, widespread retinal pigment epithelium dystrophy and macular coloboma were observed at fundus examination. Electroretinogram (ERG) showed scotopic and photopic responses below noise level. At 5 years of age, he developed a complete bilateral cataract. He had hypotonia and was able to sit independently by the age of 12 months; he walked independently at 4 years of life. His gait has always been ataxic and his growth in height was deficient. His brain MRI showed delayed myelination and leukoencephalopathy. His auditory evoked potentials were progressively abnormal bilaterally at 6 months, 16 months, 34 months and at 4 years. He had bilateral acoustic prosthesis. At his latest examination at the age of 14 years his height was 129 cm (-4.2 SD) with an arm span/height ratio of 1.1, his weight was 30 kg (-2.8 SD) and his head circumference was 51.5 cm (-2.0 SD). He had dysmorphic facial features with synophris, large mouth, and irregular dentition. He had scoliosis and his spine X-ray showed biconcave profile of the

vertebral bodies (**Fig. 2F**). On skeletal survey small and irregular epiphyseal changes were also noted.

The younger sister (P2, **Fig. 1A** and **Fig. 2B**) had a vision problem since the first months of life, with lack of visual tracking and nystagmus; she was diagnosed with LCA. In contrast to his brother, who was orthophoric, she had exotropia. At 10 years of age, BCVA was light perception with high hyperopic refractive errors of +8 Diopters. Retinal pigment epithelium dystrophy with macular coloboma was observed at fundus examination (**Fig. 2K**). ERG testing showed scotopic and photopic responses below noise level. She was able to sit independently by the age of 12 months and walked independently at 3 years of life. Her brain MRI showed marked hypomyelination. Auditory evoked potentials showed bilaterally reduced responses and she also was treated with bilateral acoustic prosthesis. At her latest examination at the age of 11 years and 10 months her height was 123.4 cm (-3.7 SD) with an arm span/height ratio of 1.08, her weight was 29.5 Kg (-1.9 SD) and her head circumference was 50.8 cm (-1.9 SD). She had scoliosis and her spine X-ray also showed biconcave profile of the vertebral bodies. Small and irregular epiphyseal changes were also noted (**Fig. 2G**). Similar to P1, she displayed irregular dentition (**Fig. 2J**).

P3 (**Fig. 1A** and **Fig. 2C-D**) was born to consanguineous parents, also originating from Southern Italy. Visual problems were reported in the first year of life, when the diagnosis of LCA was supported by specific abnormalities of ERG and visual evoked potentials. He also had exotropia. Developmental delay was noticed in the first years of life, whereas sensorineural hearing loss was diagnosed and treated with acoustic prosthesis at the age of 5 years. He was evaluated at the age of 13 years, when he presented with growth defect, spondylo-epiphyseal dysplasia (**Fig. 2E-I**), and severe cognitive impairment. Brain MRI images were consistent with hypomyelinating leukoencephalopathy, cerebral and cerebellar atrophy (**Fig. 2L-N**). From the age of 15 years he had progressive loss of motor

skills, attributed to cervical myelopathy due to both reduced caliber of foramen magnum and os odontoideum causing atlanto-occipital instability. Neurosurgical cervico-medullary decompression without stabilization of cervical spine was ineffective and the child developed flaccid quadriplegia.

Autozygome-based analysis on exome sequencing data reveals a shared homozygous interval

Exome sequencing failed to detect candidate genes or DNA variants with properties of rare mutations that were shared by all three affected individuals. Moreover, neither rare insertion/deletion nor changes in copy number were found in any of the exomes from the two families. However, we could identify by autozygome mapping a 4.76 Mb homozygous region on the short arm of chromosome 1 (chr1:6,947,973-11,708,736, hg19) that was common to all three affected children (**Fig. 1B**). Furthermore, we found that the haplotype was identical in the two families, suggesting the presence of a remote common ancestor to the two families. This genomic fragment comprises 40 genes, none of which harbored mutations or rare variants in their coding regions. Among these genes, four are responsible for disorders with some overlap with the phenotype observed in the three affected children: *CAMTA1*, involved in autosomal dominant (AD) cerebellar ataxia with mental retardation (OMIM: 614756); *RERE,* linked to AD neurodevelopmental disorder with or without anomalies of the brain, eyes or heart (OMIM: 616975); *NMNAT1*, involved in autosomal recessive (AR) LCA (OMIM: 608553); and *CTNNBIP1*, with no association with any human disease to date, but known to cause short stature in mice when overexpressed (19). Of those, only *NMNAT1* is associated with an autosomal recessive mode of inheritance, i.e. the pattern of inheritance most likely occurring in the two families (**Fig. 1A**).

In cells from patients, *NMNAT1* has a reduced expression compared to controls

RNA-seq analysis on primary skin fibroblasts showed that only 21 transcripts within the whole transcriptome were significantly differentially expressed (**Supplementary Fig. S1**) in our three subjects compared to 28 healthy controls. Among these 21 transcripts, *NMNAT1* (NM_022787.3), showing a 5.4-fold downregulation (*p*-value $1.3x10^{-4}$), was the only gene residing in the autozygous interval shared by the three individuals (the probability that one of the 21 differentially expressed genes also resides in the interval of interest is 0.04).

Since *NMNAT1* variants were not detected in the coding region by exome sequencing and because there were no evident abnormalities in exome coverage, we Sanger-sequenced *NMNAT1* regulatory regions. However, no significant changes were detected. Moreover, the sequence of the *NMNAT1* CTCF upstream insulator was intact, making unlikely disruptions of the topologically associated domain (TAD) in this region.

RT-PCR reveals patient-exclusive aberrant isoforms of *NMNAT1*

RT-PCR experiments on *NMNAT1* transcripts from patients' cells revealed the presence of aberrant mRNA isoforms. Sanger sequencing showed a heterogeneous population of transcripts: some harbored a partial retention of intron 3, others a duplication of exon 4, and some others a duplication of both exon 4 and part of exon 5, as well as the wild-type (WT) transcript (**Fig. 3A-B**). The part of exon 5 that was retained consisted in the first 278 bases, terminating at the weak donor splice site "TTGgtacca" (putative invariant donor site underlined). Human Splice Finder assigns to this sequence a score of 70 (range 0 to 100), and MAXEntScan a score of 4.88 (range -10 to +10). Additional quantitative RT-PCR with primers that were specific for each isoform and a primer pair binding to all isoforms (total), including the WT product, validated the RNA sequencing results and, above all, provided a hint that this dysregulation of splicing could be due to a

genomic structural anomaly (**Fig. 3C**). Quantification of exclusively WT mRNA was not possible due to the nature of the duplication and lack of sequences unique to this isoform. From the data shown in **Fig. 3C**, we can conclude that WT mRNA is at least 3.4 times lower in cases compared to controls.

Short- and long-read whole-genome sequencing identify a 7.4-kb duplication in *NMNAT1*

Based on RNA data, we sought to analyze the genomic structure of this region by performing both short-read WGS (Illumina) and long-read WGS (Pacific Bioscience). Both technologies allowed us to detect a homozygous genomic 7.4 kb duplication (**Fig. 4A**), involving the two last exons of *NMNAT1* and spanning the beginning of intron 3 to the middle of the 3'-UTR (hg19, chr1:10,036,359-10,043,727; NC_000001.10:g.10036359_10043727del; NM_022787.3:c.299+526_*968del) in all three affected children. Interestingly, the duplicated fragment was flanked by two *Alu* elements, *AluSx* and *AluSx3* (chr1:10,036,063-10,036,363 and chr1:10,043,431-10,043,727, respectively), which might have mediated a tandem duplication event by non-allelic homologous recombination. In addition, genome-wide single nucleotide variants (SNVs) and structural variants (SVs) analysis of these WGS data excluded the presence of additional variants that were shared among the affected children.

Additional screening for this duplication, by PCR, of a Spanish cohort of 76 patients with LCA or early-onset retinal dystrophy identified one individual with the same SV found in the two Italian families. It was present in a compound heterozygous state with the *NMNAT1* missense variant p.Glu257Lys (NM_022787.3:c.769G>A), a pathogenic variant frequently found in LCA cases (4, 7-9, 14, 20-25). This individual, the only child of a Spanish couple, at 3 years of age presented with severe visual loss with light perception in both eyes, a central macular atrophic lesion and fine peripheral mottling on fundoscopy,

moderate exotropia, nystagmus and non-recordable full-field ERG. Aside from these ophthalmic findings, which are fully compatible with the typical *NMNAT1*-associated phenotype, physical and intellectual development were normal until his current age (6 years). Haplotype analysis revealed rare heterozygous SNP genotypes in proximity of the duplication, which were shared with the Italian patients of this study.

Discussion

The disease mechanisms resulting in LCA caused by *NMNAT1* mutations are still unclear. The NMNAT1 enzyme, together with NMNAT2 and NMNAT3, is responsible for the conversion of NaMN and NMN into NAD, an essential metabolite for a wide range of biological processes. Each NMNAT enzyme localizes to a specific subcellular compartment: NMNAT1 is in the nucleus, NMNAT2 in the Golgi, and NMNAT3 in mitochondria (18, 26). The three isozymes are thought to have compartment-specific functions, rather than functional redundancy, since they have differences in their substrate specificity, and because they appear to be simultaneously expressed, at least in HeLa and HEK293 cells (18). Prior to its association with LCA, *NMNAT1* has been studied in *wld*^{s} mice, because of its neuroprotective role in Wallerian degeneration. Interestingly, *Nmnat1* knockout animals die before birth, indicating that the extranuclear NMNAT2 and NMNAT3 isozymes cannot compensate for NMNAT1 loss (27-29). Nevertheless, mouse models carrying *Nmnat1* missense mutations recapitulate the human LCA phenotype (30). Very importantly, among the numerous human pathogenic variants in *NMNAT1* reported to date, biallelic disrupting variants were never identified, indicating that the complete absence of NMNAT1 is probably not compatible with life in human as well. This is further supported by the evidence that all null mutations in LCA patients are always found in a compound heterozygous state with a missense mutation, with a notable exception of a single case from Perrault *et al.* (4), in whom a homozygous stop was identified. In this case, however, the location of the mutation, close to the 3' end of the gene, suggested that this nonsense mutation would escape nonsense-mediated RNA decay (NMD) (31) and therefore result in the production of a NMNAT1 protein partly lacking its C-terminal end and possibly retaining some residual function (4). Similarly, non-coding *NMNAT1* variants were previously found in two LCA patients with homozygous single nucleotide substitutions in the 5'-UTR portion of the gene, causing

decreased *NMNAT1* mRNA levels in lymphocytes (10). In this case as well, it is likely that, since in these individuals the NMNAT1 protein sequence is preserved, residual WT function could be retained, leading exclusively to an LCA phenotype.

The results of the genetic analysis of the three affected children we analyzed, as well as the comparison between genomic and transcriptomic data, led us to hypothesize that their condition, i.e. an intermediate phenotype between lethality and LCA, could be attributed to a structural variant located in the middle of the gene, conferring a strong but not complete reduction of expression of WT *NMNAT1* mRNA. This is further supported by the identification of this same genomic rearrangement in compound heterozygosity with a missense mutation (the most common pathogenic *NMNAT1* missense p.Glu257Lys), in an individual with non-syndromic LCA. Furthermore, the strong resemblance of phenotypes and the common haplotype in the three children are additional arguments suggesting that the genetic cause of this novel syndrome is to be found in the shared interval on chromosome 1. The genomic rearrangement in the *NMNAT1* gene is certainly contributing to the disease manifestation. What needs to be clarified is whether this particular SV in *NMNAT1* is responsible for the whole syndromic phenotype, or if it is resulting exclusively in LCA. Since *NMNAT1* is the only gene in this region showing differential expression in patients vs. controls and no additional impacting rare SNVs or SVs were found to be shared by the three patients at the level of the whole genome, we are inclined to think that this tandem duplication is the sole responsible for this newly recognized syndrome presenting with LCA, sensorineural hearing loss and skeletal dysplasia. This hypothesis is further supported by the evidence that all other genes that were found to be differentially expressed in fibroblasts from patients were either on different chromosomes or on the opposite arm of chromosome 1 with respect to the shared autozygous region (*CRABP2* and *TXNIP*, being ~147 Mb and ~136 Mb away, respectively), making unlikely a rearrangement-mediated regulatory effect with

pathological consequences. Along those same lines, we found no evidences for this duplication to affect the TAD in which it is included, again reducing the chances that other genes, the altered expression of which might have been undetected in fibroblasts, could contribute to the disease.

Long-read WGS allowed us to determine the precise breakpoint of the duplicated *NMNAT1* fragment, as well as the regions flanking this segment. The two *Alu* elements identified, *AluSx* and *AluSx3*, presenting with highly homologous sequences (88% identity at the nucleotide level), are very likely involved in the mutational event at the origin of the duplication, creating an *AluSx-AluSx3* fusion element between the two copies of the duplicated fragment. Such rearrangement is a non-allelic homologous recombination, consisting in a crossing-over between non-equivalent sequences with high degree of identity. To date, *Alu*-mediated recombination has been associated with numerous diseases (32-35), including two LCA cases (10), in compound heterozygosity with other missense changes. Interestingly, despite the first Italian family reported no consanguinity, the presence of this duplication in homozygous state in affected individuals (embedded in a common haplotype) is clear indication that we identified a founder mutation. The close geographical origin of the two families in the same area from Southern Italy, i.e. Campania, supports the hypothesis of common ancestors. The presence of an identical haplotype surrounding this rearrangement in an additional patient from Spain is also indicative of a common and likely remote ancestral genetic event. Furthermore, the GnomAD-SV database (36) reports a similar duplication (chr1:10.036.294-10.043,744) in one heterozygous individual of European origin. This chromosomal rearrangement is not identical to the one detected in the patients analyzed here, but it may have been originated by a rearrangement involving the same *Alu* elements. This finding, unique entry within a total of 21,690 sequenced alleles, supports as well the finding that this

rearrangement is not pathogenic *per se* in a heterozygous state –as it is the case for instance for the parents from families 1 and 2.

The reasons for which different mutations can lead to phenotypes that are retina-restricted vs. syndromic are currently not clear. Since the retina is considered to be one of the most metabolically active tissues of the human body (37), it could be hypothesized that retinal degeneration takes place as a first manifestation of mild *NMNAT1* impairment, whereas stronger deficiencies could be responsible for the appearance of additional and more severe phenotypes (**Supplementary Fig. S2**). Of note, some of the known LCA-associated loci appear to have a spectrum of clinical expression that ranges from isolated LCA or retinal degeneration to a more widespread phenotype including the CNS (with intellectual disability), the ear, and the kidney. This may support the possibility that *NMNAT1* variants may also be responsible for a retina-only as well as retina-plus phenotypes.

Both *in vitro* and *in vivo* mouse studies on the Wallerian degeneration mouse model *wld*^{S}, overexpressing a fusion protein that incorporates *Nmnat1* and the ubiquitination factor *Ube4b*, showed NMNAT1 to have a beneficial effect in neuronal protection after injury (38). Additionally, it has been demonstrated *in vitro* that inactivation of NMNAT1 enzymatic activity in the *wld*^{S} mouse model causes the loss of the neuroprotective effect (38, 39). In contrast, it has been reported that overexpression of *Nmnat1* in WT mice has no neuroprotective effect (27). Altogether, these results suggest that the beneficial outcome of *Nmnat1* expression might be simply linked to its function in the production of NAD and that the retina, as a tissue with a very high metabolism, may be particularly sensitive to NAD deficiency and therefore to *NMNAT1* mutations.

On the other hand, it has been recently shown in mouse that modulation of the enzymatic activity by overexpressing or downregulating *NMNAT1* was not associated with significant alterations of NAD levels (40). Altogether, these data open the possibility

for the presence of additional functions of NMNAT1, other than in NAD biosynthesis, that might be implicated in the disease development of our patients. Nonetheless, we tend to favor the idea that a decrease of nuclear NAD could indeed play a role in the disease mechanism.

This hypothesis is further supported by previous data linking NAD deficiency and disease. Patients with homozygous variants in *HAAO* and *KYNU*, two genes from the kynurenine pathway and responsible for the *de novo* synthesis of NAD from tryptophan, showed phenotypes somehow similar to those identified by us: vertebral segmentation defects, short stature (skeletal dysplasia), sensorineural hearing loss, microcephaly, intellectual disability, as well as cardiac and renal defects (41). In addition, they had reduced levels of NAD, again supporting the possibility that all these phenotypes may have a common link in the biosynthesis of NAD.

In conclusion, we identified the molecular cause of a newly recognized multisystem disorder, SHILCA, in three Italian probands sharing a common haplotype in an interval of chromosome 1. This phenotype is reminiscent of that of another condition we have recently defined, the Liberfarb syndrome (6), with the difference that SHILCA does not include joint laxity. Combination of WGS and RNA-seq led to the discovery of a partial duplication of the *NMNAT1* gene involving two full exons and a part of the 3'-UTR. As for LCA caused by mutations in this gene, the pathogenesis for this new disorder remains unclear and further experiments need to be done in order to fully understand the role of NAD in the context of these disorders.

Materials and Methods

Ethical considerations and patients' material

This study has been conducted in accordance with the tenets of the Declaration of Helsinki and was approved by the Institutional Review Boards of the University of Lausanne and Basel, as well as of the Gaslini Hospital in Genoa, of the Federico II University Hospital in Naples, and of the Fundación Jiménez Diaz University Hospital in Madrid. Following the signature of a written informed consent, also approved by the respective Ethics Committees, cell and nucleic acid samples were obtained for analysis and storage from patients and their family members. More specifically, cell lines originated from the 'Cell line and DNA Biobank from patients affected by Genetic Diseases' located at Gaslini Institute (Genoa, Italy) and from the Telethon Network of Genetic Biobanks (http://biobanknetwork.telethon.it/).

Exome sequencing and autozygome analysis

Exome sequencing was performed on the two children and the two healthy parents from family 1, as well as the affected child from family 2 (**Fig. 1A**), using 2 µg DNA extracted from peripheral white blood cells. Protein-coding DNA regions were captured using the SureSelectXT reagent kit (Agilent) and an Illumina HiSeq 2500 instrument was used for paired-end sequencing. Raw reads were mapped to the human reference genome (hg19/GRCh37) using the Novoalign software (V3.08.00, Novocraft Technologies) and Isaac aligner. Next, Picard (version 2.14.0-SNAPSHOT) was used to remove duplicate reads. Single nucleotide variants and small insertions and deletions were detected using the Genome Analysis Tool Kit (GATK v4.0) software package, using the Best Practice Guidelines identified by the developers (42), Strelka Germline Variant Caller (43), Manta SV caller (44) and Illumina Annotation Engine (https://github.com/Illumina/Nirvana/wiki). The pathogenicity of the detected genetic

variants was assessed after functional annotation through ANNOVAR (45) and in-house databases. Shared regions of homozygosity and common haplotypes were analyzed using an internally-developed software (Quinodoz *et al.*, manuscript in preparation). Variant nomenclature was verified with VariantValidator (46).

Short-read WGS

Whole-genome sequencing (WGS) was performed on patients P1 and P2 and their parents (**Fig. 1A**) with Illumina NovaSeq instrument obtaining an average coverage of 50.6x and average fragment length of 453 bp in the four samples. Downstream analysis of sequencing reads was performed with the same tools used for data from exomes. Variant notation for the duplication was verified with VariantValidator (46).

High molecular weight DNA preparation and long-read WGS

High molecular weight DNA was extracted from patient P1 primary fibroblasts using the Genomic tip 20/G kit (QIAGEN). At about 90% confluence, cells were scraped from two 100 mm dishes, and DNA was extracted in agreement with the instructions provided by the manufacturer. DNA integrity was verified by using a Fragment Analyzer System (Agilent) and assessed using the PROSize software, version 3.0 (Advanced Analytical Technologies Inc.). The average DNA size was calculated to be approximately 45.6 kbp.

Long-read WGS was performed by using a Pacific Bioscience (PacBio) instrument. In order to get enough coverage, 6 SMRT flow-cells were used on a PacBio RS II platform, reaching in the end an approximate coverage of 10x. Reads were assembled to the reference human genome (hg19) with the long-read specific aligner NGMLR (47), and variant calling was performed using the Sniffles software (47).

RNA analysis

Total RNA was extracted from skin-derived fibroblasts of the three patients and of one unrelated control using Direct-zol RNA MiniPrep (ZYMO Research). RNA libraries were prepared by a service provider (Fasteris, Geneva, Switzerland) and single-read RNA sequencing was performed using an Illumina HiSeq 2500 instrument.

Fastq files were analyzed with both the CLCbio Genomics Workbench (Qiagen) and the publicly-available computer pipeline RNAcocktail (48). Raw fastq files of 27 other healthy short term cultivated skin fibroblasts were downloaded from ArrayExpress (accession E-MTAB-4652, ENA study ERP015294) (49) and were used as additional controls for statistical analyses.

cDNA was obtained using random primers and the GoScript reverse transcriptase (Promega). Quantitative real-time RNA (cDNA) analysis was used to measure gene expression levels. Prior to sample measurements, efficiency of each primer pair was tested using a standard curve. Reference controls were selected according to their stability throughout the samples, the most stables being *TBP*, *GAPDH* and *ACTB*. Normalization was done using a geometric mean of the Ct values of the three reference genes. Amplification was performed using the SYBR Green PCR Master Mix (Applied Biosystems). All real-time PCR products were visualized on 1% agarose gel to verify the specificity of primers. Primers used for detection of *NMNAT1* isoforms were: all transcripts, 5'-GGTGGAAGTTGATACATGGGA-3' (F) and 5'-TCTTCCTTCCAGGCCTTTCTA-3' (R); duplicated exon 4, 5'-GGTGGAAGTTGATACATGGGA-3' (F) and 5'-AATTTCTCTTGATGGTGTCTTTTGTT-3' (R); duplicated exon 4 and part of exon 5, 5'-CATTCGCTACTTGACACCATCA-3' (F) and 5'-TCTTCCTTCCAGGCCTTTCTA-3' (R); partial retention of intron 3, 5'-TGGAGCATGTGAGAAAGAGAAATAT-3' (F) and 5'-TTGGCACAGCTTTTGTTTTTGG-3' (R). Primers for 'all transcripts' amplify all isoforms (WT and aberrant forms).

PCR screening for the same duplication in other cohorts of patients

PCR screening for the *NMNAT1* exon 4 and 5 duplication in other patient cohorts was performed on genomic DNA with the following primer pairs: 5'-GGAGGCAGAGGTTGCAGTAA-3' (forward, matching the 3'-UTR), 5'-ATGGAAGCAGCACACGAATCA-3' (reverse, matching intron 3). The PCR product is visible (~300 bp) in presence of the duplication and not visible in WT controls.

Acknowledgements

This work was supported by the Swiss National Science Foundation (grant # 176097, to CR), the PhD Fellowships in Life Science of the University of Lausanne (to MQ), the Fondation Guillaume Gentil in Lausanne (to AS-F), the Fondazione Telethon - Telethon UnDiagnosed Project, TUDP (to MP, SB, VN, NBP, GC), and the Fondazione Roma (to FS and SB). We are grateful to the TIGEM Next Generation Sequencing and Bioinformatics Cores for technical assistance. We would also like to thank the Instituto de Salud Carlos III (ISCIII) from the Spanish Ministry of Health, including CIBERER (06/07/0036), IIS-FJD Biobank PT13/0010/0012, and FIS (PI16/00425) grants (to CA) and the regional government of Madrid, (grant RAREGenomics-CM, B2017/BMD-3721 to CA), all partially supported by European Regional Development Fund; the Spanish National Organization of the Blind (ONCE), the Spanish Fighting Blindness Foundation (FUNDALUCE), and the Ramon Areces Foundation (to CA). MC is supported by the ISCII Miguel Servet Program (CPII17_00006).

Conflict of interest statement.

None declared.

References

1 Allikmets, R. (2004) Leber congenital amaurosis: a genetic paradigm. *Ophthalmic genetics*, **25**, 67-79.

2 Stone, E.M. (2007) Leber congenital amaurosis - a model for efficient genetic testing of heterogeneous disorders: LXIV Edward Jackson Memorial Lecture. *Am J Ophthalmol*, **144**, 791-811.

3 Koenekoop, R.K. (2004) An overview of Leber congenital amaurosis: a model to understand human retinal development. *Surv Ophthalmol*, **49**, 379-398.

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

5 Xu, M., Xie, Y.A., Abouzeid, H., Gordon, C.T., Fiorentino, A., Sun, Z., Lehman, A., Osman, I.S., Dharmat, R., Riveiro-Alvarez, R. *et al.* (2017) Mutations in the Spliceosome Component CWC27 Cause Retinal Degeneration with or without Additional Developmental Anomalies. *Am J Hum Genet*, **100**, 592-604.

6 Peter, V.G., Quinodoz, M., Pinto-Basto, J., Sousa, S.B., Di Gioia, S.A., Soares, G., Ferraz Leal, G., Silva, E.D., Pescini Gobert, R., Miyake, N. *et al.* (2019) The Liberfarb syndrome, a multisystem disorder affecting eye, ear, bone, and brain development, is caused by a founder pathogenic variant in thePISD gene. *Genet Med*, **21**, 2734-2743.

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

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

9 Chiang, P.W., Wang, J., Chen, Y., Fu, Q., Zhong, J., Chen, Y., Yi, X., Wu, R., Gan, H., Shi, Y. *et al.* (2012) Exome sequencing identifies NMNAT1 mutations as a cause of Leber congenital amaurosis. *Nat Genet*, **44**, 972-974.

10 Coppieters, F., Todeschini, A.L., Fujimaki, T., Baert, A., De Bruyne, M., Van Cauwenbergh, C., Verdin, H., Bauwens, M., Ongenaert, M., Kondo, M. *et al.* (2015) Hidden

Genetic Variation in LCA9-Associated Congenital Blindness Explained by 5'UTR Mutations and Copy-Number Variations of NMNAT1. *Hum Mutat*, **36**, 1188-1196.

11 Nash, B.M., Symes, R., Goel, H., Dinger, M.E., Bennetts, B., Grigg, J.R. and Jamieson, R.V. (2018) NMNAT1 variants cause cone and cone-rod dystrophy. *Eur J Hum Genet*, **26**, 428-433.

12 Corton, M., Nishiguchi, K.M., Avila-Fernandez, A., Nikopoulos, K., Riveiro-Alvarez, R., Tatu, S.D., Ayuso, C. and Rivolta, C. (2016) Correction: Exome Sequencing of Index Patients with Retinal Dystrophies as a Tool for Molecular Diagnosis. *PLoS One*, **11**, e0153121.

13 Wang, S., Zhang, Q., Zhang, X., Wang, Z. and Zhao, P. (2016) Clinical and genetic characteristics of Leber congenital amaurosis with novel mutations in known genes based on a Chinese eastern coast Han population. *Graefes Arch Clin Exp Ophthalmol*, **254**, 2227-2238.

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

15 Jin, X., Qu, L.H., Meng, X.H., Xu, H.W. and Yin, Z.Q. (2014) Detecting genetic variations in hereditary retinal dystrophies with next-generation sequencing technology. *Mol Vis*, **20**, 553-560.

16 Magni, G., Amici, A., Emanuelli, M., Raffaelli, N. and Ruggieri, S. (1999) Enzymology of NAD+ synthesis. *Adv Enzymol Relat Areas Mol Biol*, **73**, 135-182, xi.

17 Berger, F., Ramirez-Hernandez, M.H. and Ziegler, M. (2004) The new life of a centenarian: signalling functions of NAD(P). *Trends Biochem Sci*, **29**, 111-118.

18 Berger, F., Lau, C., Dahlmann, M. and Ziegler, M. (2005) Subcellular compartmentation and differential catalytic properties of the three human nicotinamide mononucleotide adenylyltransferase isoforms. *J Biol Chem*, **280**, 36334-36341.

19 Chen, M., Zhu, M., Awad, H., Li, T.F., Sheu, T.J., Boyce, B.F., Chen, D. and O'Keefe, R.J. (2008) Inhibition of beta-catenin signaling causes defects in postnatal cartilage development. *J Cell Sci*, **121**, 1455-1465.

20 Coppieters, F., Van Schil, K., Bauwens, M., Verdin, H., De Jaegher, A., Syx, D., Sante, T., Lefever, S., Abdelmoula, N.B., Depasse, F. *et al.* (2014) Identity-by-descent-guided mutation analysis and exome sequencing in consanguineous families reveals unusual clinical and molecular findings in retinal dystrophy. *Genetics in medicine : official journal of the American College of Medical Genetics*, **16**, 671-680.

21 Ronen, D. and Benvenisty, N. (2012) Genomic stability in reprogramming. *Curr Opin Genet Dev*, **22**, 444-449.

22 Bravo-Gil, N., Mendez-Vidal, C., Romero-Perez, L., Gonzalez-del Pozo, M., Rodriguez-de la Rua, E., Dopazo, J., Borrego, S. and Antinolo, G. (2016) Improving the management of Inherited Retinal Dystrophies by targeted sequencing of a population-specific gene panel. *Sci Rep*, **6**, 23910.

23 Sasaki, Y., Margolin, Z., Borgo, B., Havranek, J.J. and Milbrandt, J. (2015) Characterization of Leber Congenital Amaurosis-associated NMNAT1 Mutants. *J Biol Chem*, **290**, 17228-17238.

24 Boulanger-Scemama, E., El Shamieh, S., Demontant, V., Condroyer, C., Antonio, A., Michiels, C., Boyard, F., Saraiva, J.P., Letexier, M., Souied, E. *et al.* (2015) Next-generation sequencing applied to a large French cone and cone-rod dystrophy cohort: mutation spectrum and new genotype-phenotype correlation. *Orphanet J Rare Dis*, **10**, 85.

25 Siemiatkowska, A.M., Schuurs-Hoeijmakers, J.H., Bosch, D.G., Boonstra, F.N., Riemslag, F.C., Ruiter, M., de Vries, B.B., den Hollander, A.I., Collin, R.W. and Cremers, F.P. (2014) Nonpenetrance of the most frequent autosomal recessive leber congenital amaurosis mutation in NMNAT1. *JAMA Ophthalmol*, **132**, 1002-1004.

26 Ryu, K.W., Nandu, T., Kim, J., Challa, S., DeBerardinis, R.J. and Kraus, W.L. (2018) Metabolic regulation of transcription through compartmentalized NAD(+) biosynthesis. *Science*, **360**.

27 Conforti, L., Fang, G., Beirowski, B., Wang, M.S., Sorci, L., Asress, S., Adalbert, R., Silva, A., Bridge, K., Huang, X.P. *et al.* (2007) NAD(+) and axon degeneration revisited: Nmnat1 cannot substitute for Wld(S) to delay Wallerian degeneration. *Cell Death Differ*, **14**, 116-127.

28 Coleman, M.P. and Freeman, M.R. (2010) Wallerian degeneration, wld(s), and nmnat. *Annu Rev Neurosci*, **33**, 245-267.

29 Avery, M.A., Sheehan, A.E., Kerr, K.S., Wang, J. and Freeman, M.R. (2009) Wld S requires Nmnat1 enzymatic activity and N16-VCP interactions to suppress Wallerian degeneration. *J Cell Biol*, **184**, 501-513.

30 Greenwald, S.H., Charette, J.R., Staniszewska, M., Shi, L.Y., Brown, S.D.M., Stone, L., Liu, Q., Hicks, W.L., Collin, G.B., Bowl, M.R. *et al.* (2016) Mouse Models of NMNAT1-Leber Congenital Amaurosis (LCA9) Recapitulate Key Features of the Human Disease. *Am J Pathol*, **186**, 1925-1938.

31 Hentze, M.W. and Kulozik, A.E. (1999) A perfect message: RNA surveillance and nonsense-mediated decay. *Cell*, **96**, 307-310.

32 Deininger, P.L. and Batzer, M.A. (1999) Alu repeats and human disease. *Mol Genet Metab*, **67**, 183-193.

33 Garland, J., Stephen, J., Class, B., Gruber, A., Ciccone, C., Poliak, A., Hayes, C.P., Singhal, V., Slota, C., Perreault, J. *et al.* (2017) Identification of an Alu element-mediated deletion in the promoter region of GNE in siblings with GNE myopathy. *Mol Genet Genomic Med*, **5**, 410-417.

34 Liccardo, R., De Rosa, M., Rossi, G.B., Rigler, G., Izzo, P. and Duraturo, F. (2018) Characterization of novel, large duplications in the MSH2 gene of three unrelated Lynch syndrome patients. *Cancer Genet*, **221**, 19-24.

35 Geoffroy, V., Stoetzel, C., Scheidecker, S., Schaefer, E., Perrault, I., Bar, S., Kroll, A., Delbarre, M., Antin, M., Leuvrey, A.S. *et al.* (2018) Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140. *Hum Mutat*, **39**, 983-992.

36 Collins, R.L., Brand, H., Karczewski, K.J., Zhao, X., Alfoldi, J., Francioli, L.C., Khera, A.V., Lowther, C., Gauthier, L.D., Wang, H. *et al.* (2020) A structural variation reference for medical and population genetics. *Nature*, **581**, 444-451.

37 Niven, J.E. and Laughlin, S.B. (2008) Energy limitation as a selective pressure on the evolution of sensory systems. *J Exp Biol*, **211**, 1792-1804.

38 Araki, T., Sasaki, Y. and Milbrandt, J. (2004) Increased nuclear NAD biosynthesis and SIRT1 activation prevent axonal degeneration. *Science*, **305**, 1010-1013.

39 Jia, H., Yan, T., Feng, Y., Zeng, C., Shi, X. and Zhai, Q. (2007) Identification of a critical site in Wld(s): essential for Nmnat enzyme activity and axon-protective function. *Neurosci Lett*, **413**, 46-51.

40 Rossi, F., Geiszler, P.C., Meng, W., Barron, M.R., Prior, M., Herd-Smith, A., Loreto, A., Lopez, M.Y., Faas, H., Pardon, M.C. *et al.* (2018) NAD-biosynthetic enzyme NMNAT1 reduces early behavioral impairment in the htau mouse model of tauopathy. *Behav Brain Res*, **339**, 140-152.

41 Shi, H., Enriquez, A., Rapadas, M., Martin, E., Wang, R., Moreau, J., Lim, C.K., Szot, J.O., Ip, E., Hughes, J.N. *et al.* (2017) NAD Deficiency, Congenital Malformations, and Niacin Supplementation. *N Engl J Med*, **377**, 544-552.

42 DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M. *et al.* (2011) A framework for

variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*, **43**, 491-498.

43 Saunders, C.T., Wong, W.S., Swamy, S., Becq, J., Murray, L.J. and Cheetham, R.K. (2012) Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics*, **28**, 1811-1817.

44 Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Kallberg, M., Cox, A.J., Kruglyak, S. and Saunders, C.T. (2016) Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*, **32**, 1220-1222.

45 Wang, K., Li, M. and Hakonarson, H. (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*, **38**, e164.

46 Freeman, P.J., Hart, R.K., Gretton, L.J., Brookes, A.J. and Dalgleish, R. (2018) VariantValidator: Accurate validation, mapping, and formatting of sequence variation descriptions. *Hum Mutat*, **39**, 61-68.

47 Sedlazeck, F.J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A. and Schatz, M.C. (2018) Accurate detection of complex structural variations using single-molecule sequencing. *Nat Methods*, **15**, 461-468.

48 Sahraeian, S.M.E., Mohiyuddin, M., Sebra, R., Tilgner, H., Afshar, P.T., Au, K.F., Bani Asadi, N., Gerstein, M.B., Wong, W.H., Snyder, M.P. *et al.* (2017) Gaining comprehensive biological insight into the transcriptome by performing a broad-spectrum RNA-seq analysis. *Nat Commun*, **8**, 59.

49 Kaisers, W., Boukamp, P., Stark, H.J., Schwender, H., Tigges, J., Krutmann, J. and Schaal, H. (2017) Age, gender and UV-exposition related effects on gene expression in in vivo aged short term cultivated human dermal fibroblasts. *PLoS One*, **12**, e0175657.

Legends to Figures

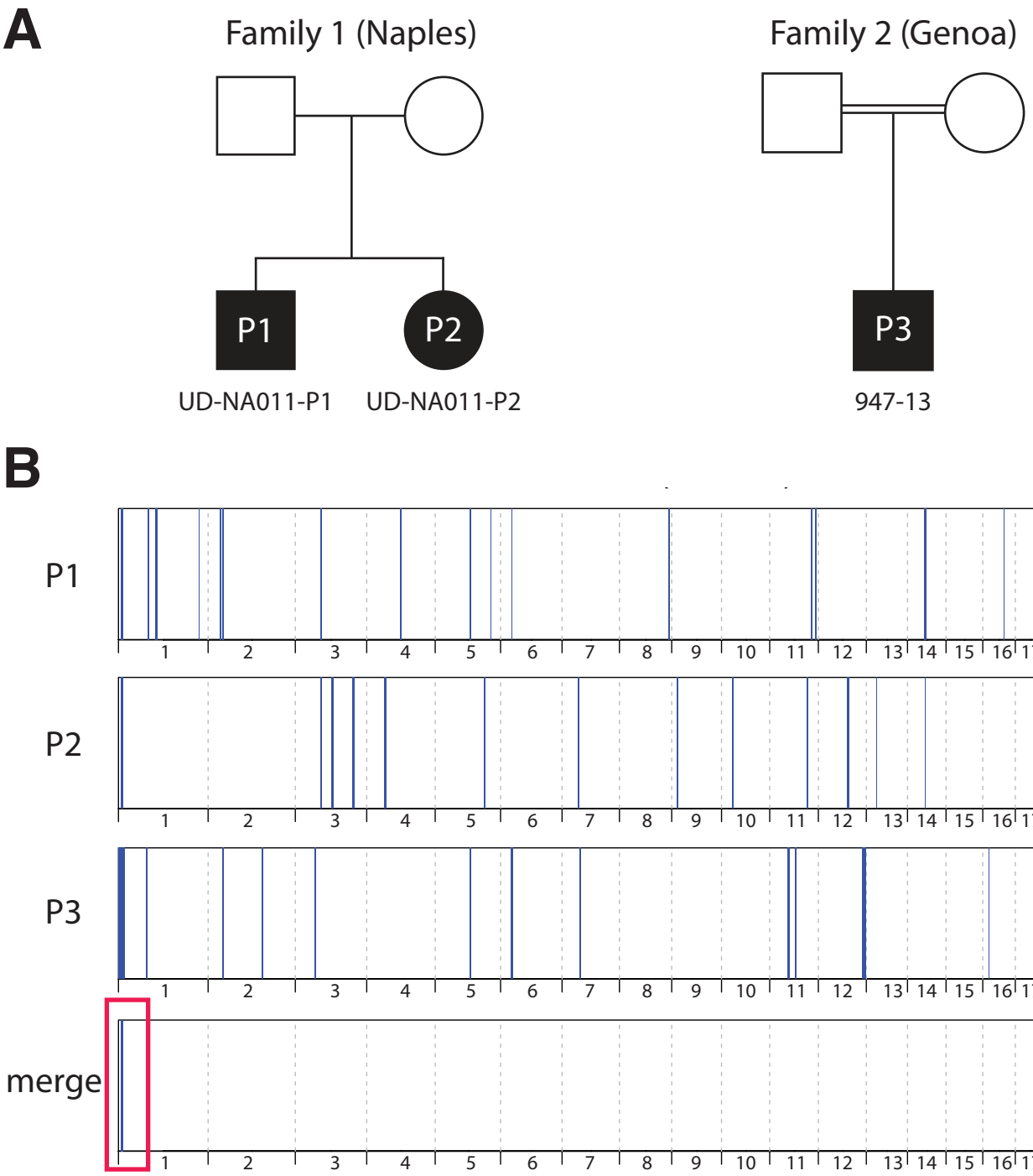
Figure 1: Pedigrees (A) and homozygosity mapping (B) of the three patients assessed in this study. Total autozygous regions, indicated by blue bars, in P1, P2, and P3 span 8.51 Mb, 9.52 Mb, and 29.01 Mb, respectively. The single shared interval (merge), highlighted by the red box, consists in a 4.6 Mb region on chromosome 1.

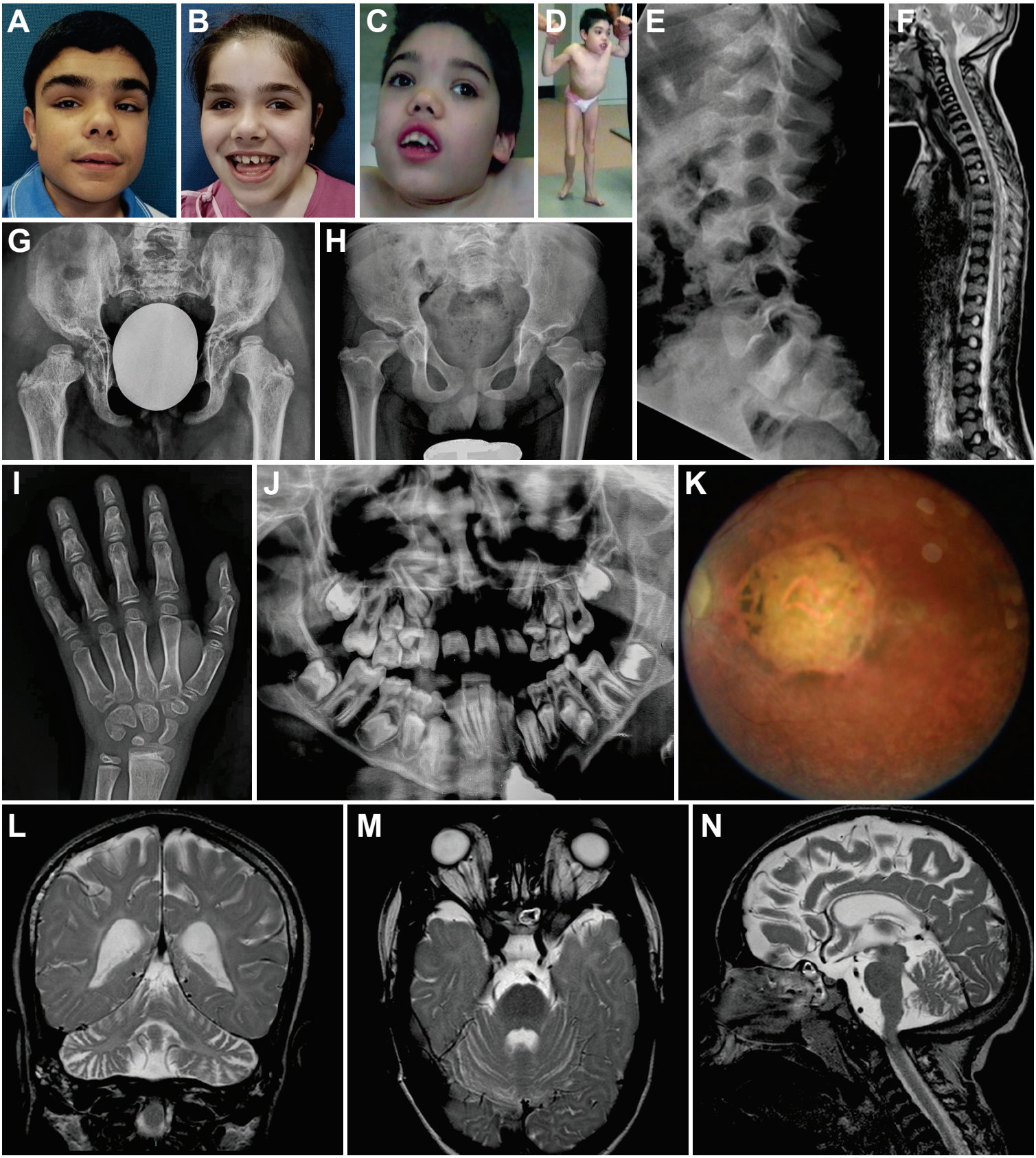
Figure 2: Photographs of patients P1 (A), P2 (B), and P3 (C,D). Facial appearance of the three affected individuals show mild coarsening and a deep nasal bridge. Strabismus is evident in P3, secondary to retinal degeneration. Panel D shows short trunk and signs of lower limb paraparesis. Lateral lumbar spine radiographs of patient P3 (E) and the magnetic resonance imaging (MRI) of patient P1 (F) show the presence of a deep sagittal notch in the center of the vertebral bodies, suggesting that in early childhood there may have been sagittal clefts. The MRI image (panel F) confirms the notching of the vertebral bodies with preservation of a round nucleus pulposus in the intervertebral disks. The cranio-occipital transition does not show signs of stenosis. Pelvic radiographs in panels G (P2, age 6 years) and H (P3, age 11 years) show marked dysplasia of the femoral heads and of the acetabulum. The iliac wings are large and wide. P3's hand x-ray (I) shows mild reduction in the carpal height with dysplasia of the carpal bones. Altogether, these signs are diagnostic for a form of spondylo-epiphyseal dysplasia (SED). Panel J shows a panoramic radiograph of dental abnormalities in P2. Color fundus of the left eye of patient P2 at age 10 (K) shows pale optic disc, marked attenuation of retinal vessels and retinal pigment epithelium dystrophy with macular coloboma. Panels L, M and N show cerebral MRI images of patient P3 at age 13. There is diffused lack of white substance and mild cerebellar atrophy (L, N). We also observe thin optic nerves with a dilated sheath

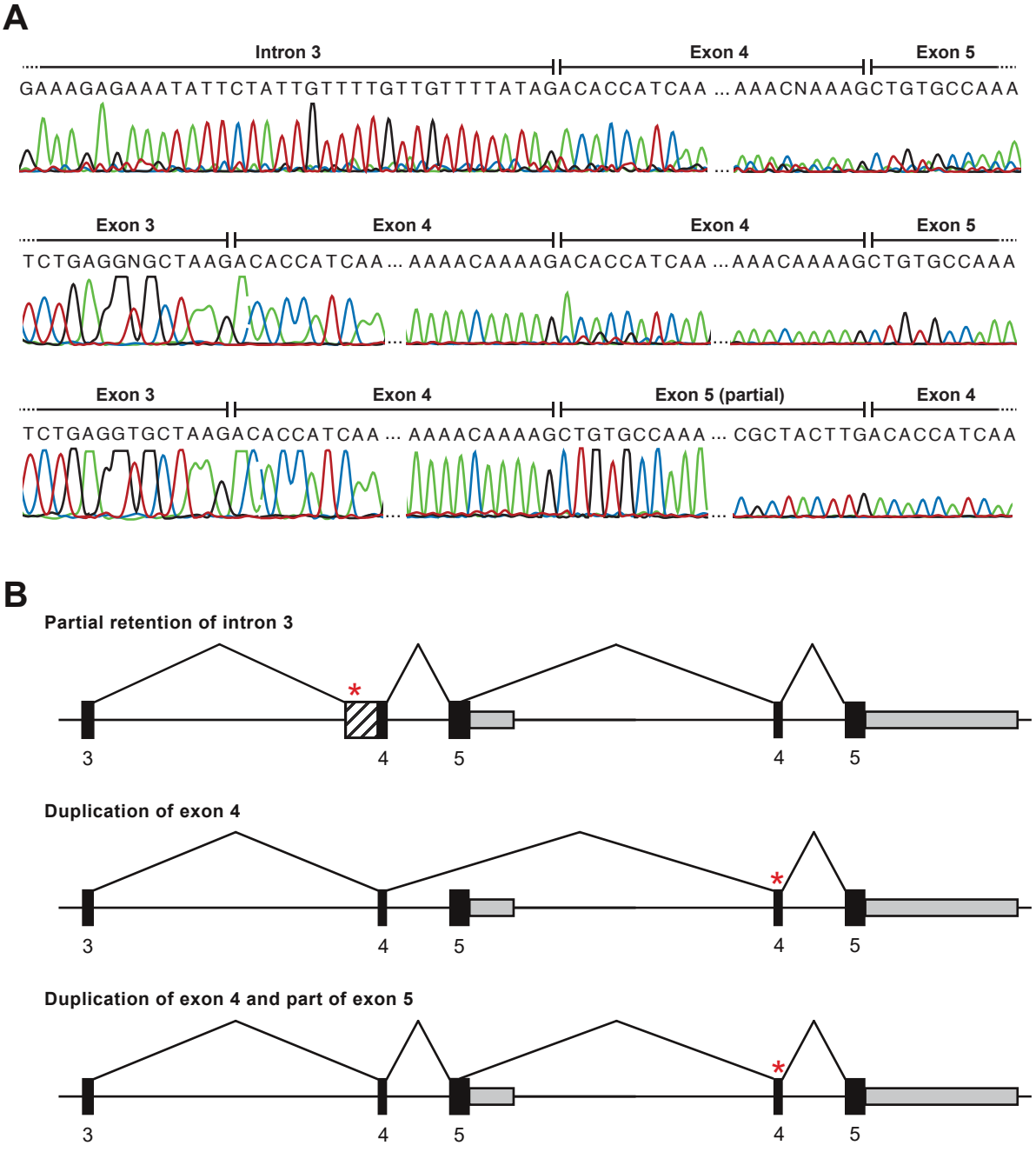
(M) and thinning of the corpus callosum, as well as signs of foramen magnum stenosis (N).

Figure 3: *NMNAT1* splice products from patients' mRNA. (A) Sequences of aberrant mRNAs (cDNAs) and (B) their corresponding schematic representations. Red asterisks show the location of premature stop codons resulting from abnormal splicing. (C) Relative expression levels of the total and aberrant mRNA isoforms, after normalization with respect to three housekeeping genes (*TBP*, *GAPDH* and *ACTB*, N=3 for both patients and controls). Measurements of the total isoforms (first histogram) represent all isoforms (WT and aberrant = all transcripts).

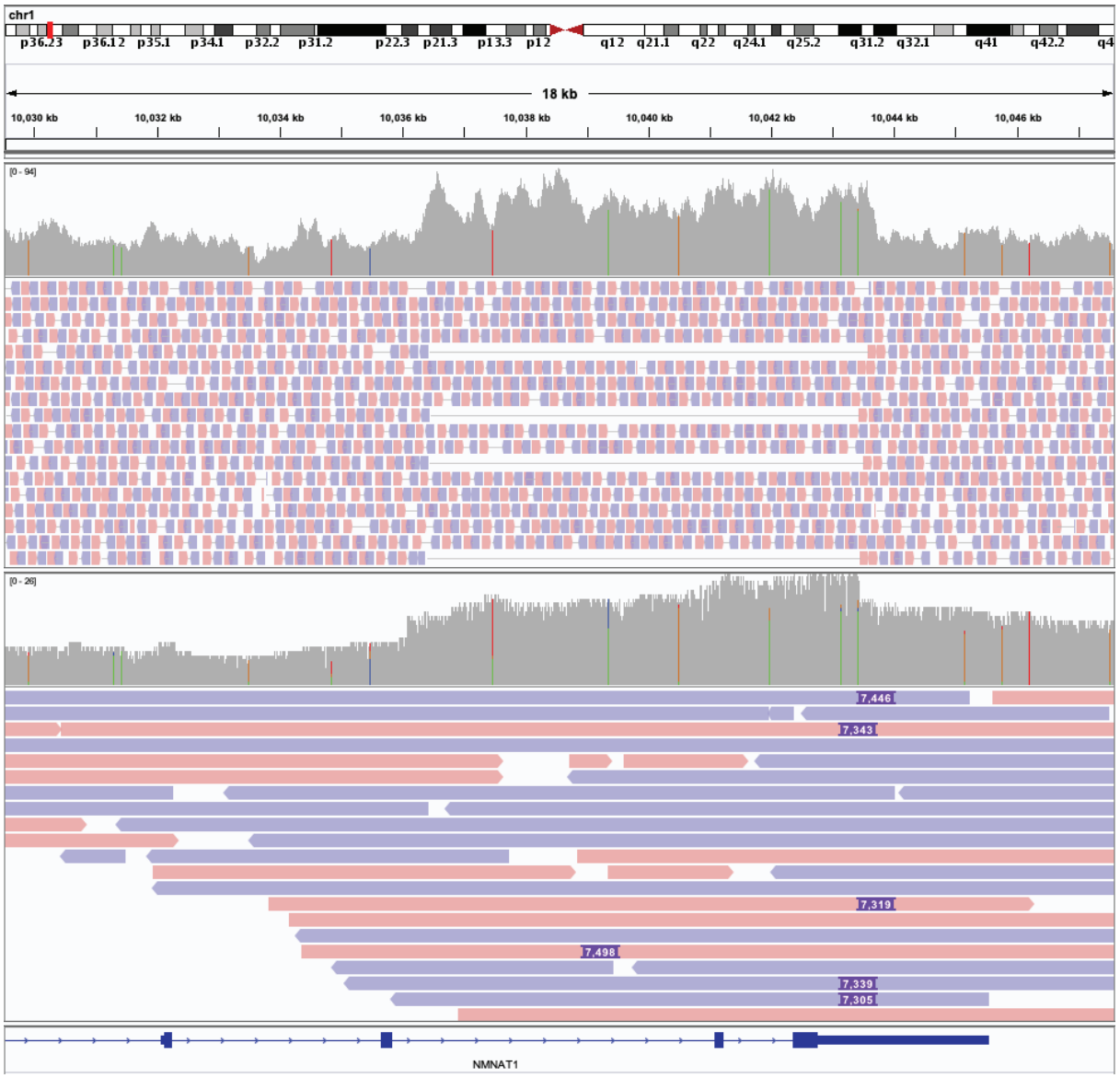
Figure 4: Structure of the partial *NMNAT1* duplication, at the genomic level. (A) IGV visualization of short reads by Illumina (upper half) and long reads generated by PacBio (lower half). Only the reads entirely spanning the duplicated interval show a 7.4 kb insertion. Alignment of the 'insert' with respect to the reference human sequence (dark purple) indicate indeed the presence of a tandem duplication. The orientation of the reads is color-coded (red, forward strand; blue, reverse strand). (B) In-scale schematic diagrams of the reference vs. patients' genomic *NMNAT1* sequences, with the duplicated fragment shown in red. *NMNAT1* exons are numbered.


A
Family 1 (Naples)
Family 2 (Genoa)
P1
P2
P3
UD-NA011-P1
UD-NA011-P2
947-13
B
P1
P2
P3
merge
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22


A
B
C
D
E
F
G
H
I
J
K
L
M
N


A
Intron 3
Exon 4
Exon 5
GAAAGAGAAATATTCTATTGTTTTGTTGTTTTATAGACACCATCAA ... AAACNAAAGCTGTGCCAAA
Exon 3
Exon 4
Exon 4
Exon 5
TCTGAGGNGCTAAGACACCATCAA ... AAAACAAAAGACACCATCAA ... AAACAAAAGCTGTGCCAAA
Exon 3
Exon 4
Exon 5 (partial)
Exon 4
TCTGAGGTGCTAAGACACCATCAA ... AAAACAAAAGCTGTGCCAAA ... CGCTACTTGACACCATCAA
B
Partial retention of intron 3
3
4
5
4
5
Duplication of exon 4
3
4
5
4
5
Duplication of exon 4 and part of exon 5
3
4
5
4
5
C
total
intron 3 retention
exon 4 duplication
exon 4 and exon 5 duplication
relative expression (arbitrary units)
****
Controls
Patients

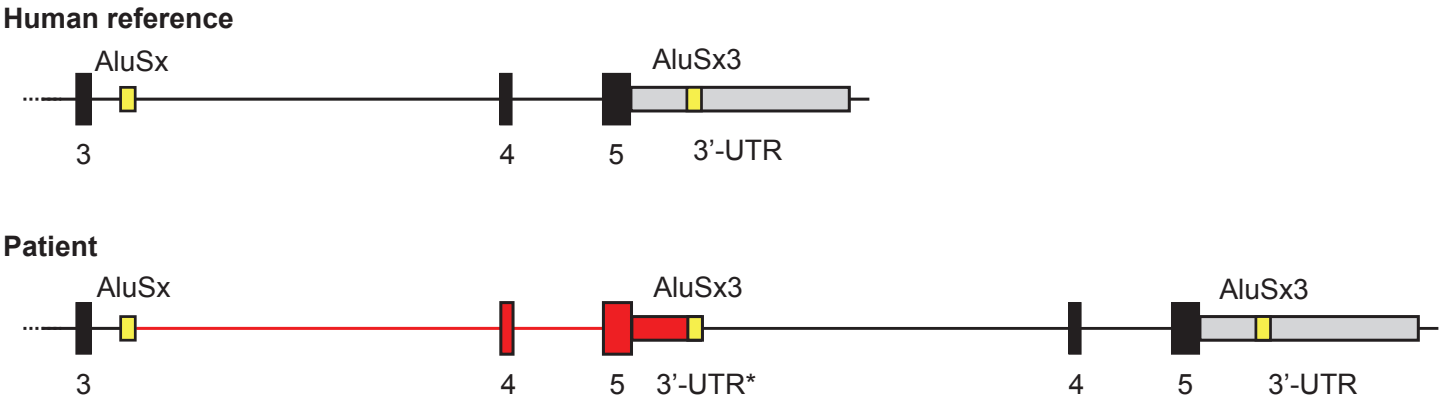
A



Illumina WGS with short reads

PacBio WGS with long reads

B



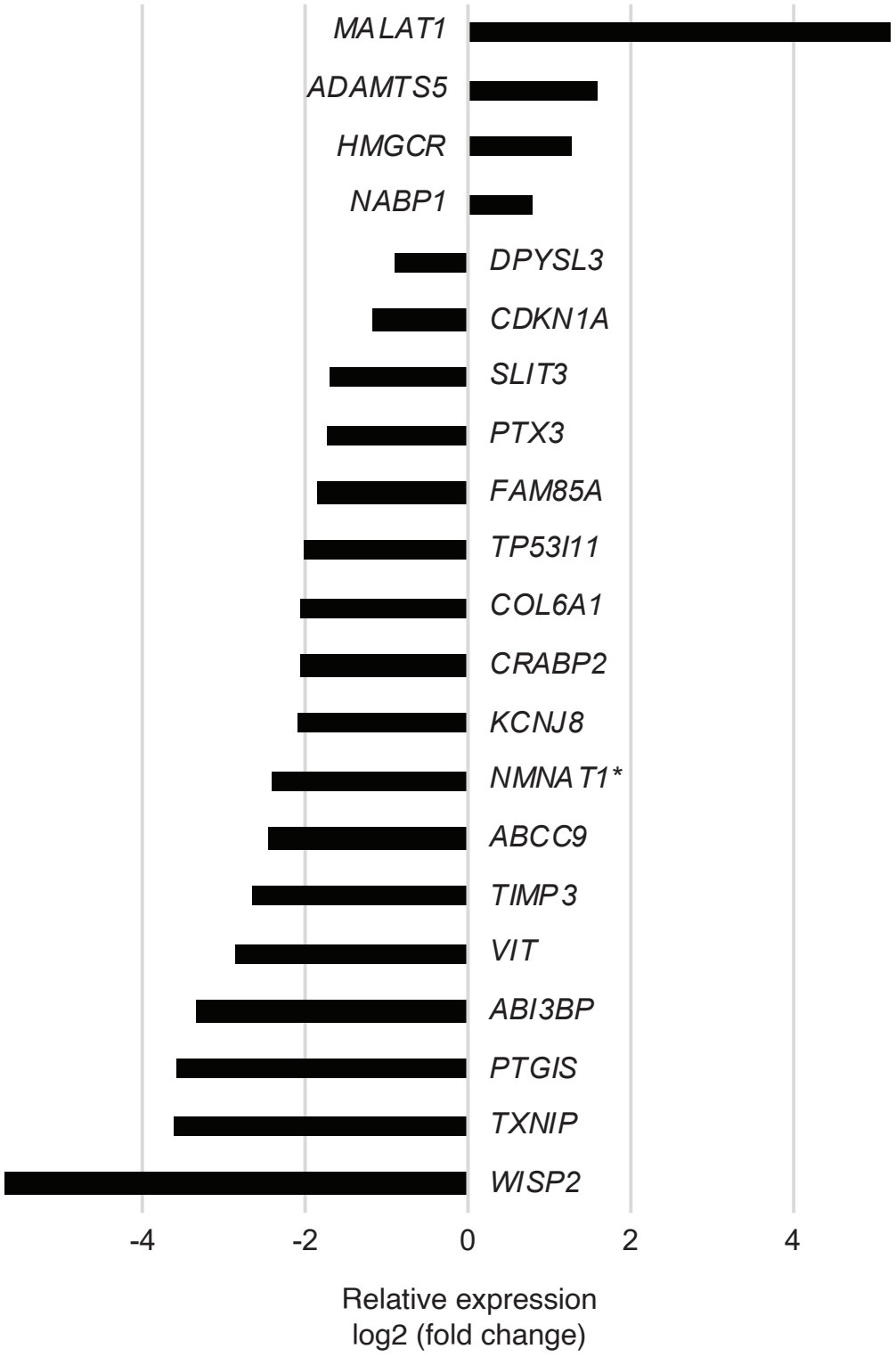


Figure S1: Results of RNA-seq analyses. Twenty-one significantly up- and down-regulated genes in fibroblasts from the three subjects vs. fibroblasts from 28 healthy controls are shown. Among them, *NMNAT1* is the only gene located within the shared homozygous interval (*).

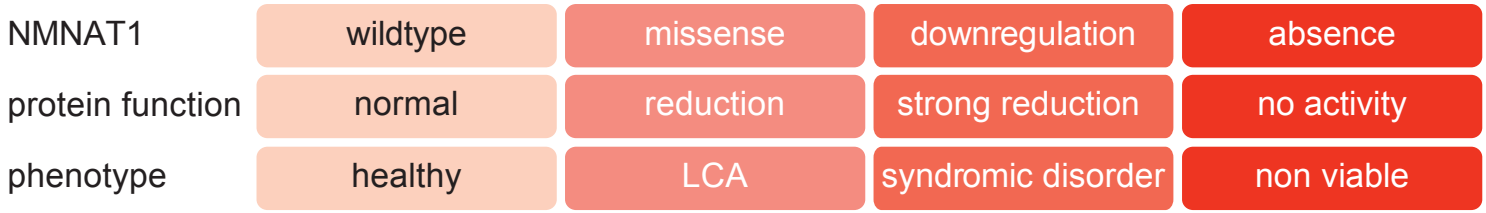


Figure S2: Schematic representation of our working hypothesis about the genotype-phenotype relationship driven by different *NMNAT1* mutations.