A Mutation of *COX6A1* Causes a Recessive Axonal or Mixed Form of Charcot-Marie-Tooth Disease

Gen Tamiya,[1,*] Satoshi Makino,[1] Makiko Hayashi,[2] Akiko Abe,[2] Chikahiko Numakura,[2] Masao Ueki,[1] Atsushi Tanaka,[3] Chizuru Ito,[4] Kiyotaka Toshimori,[4] Nobuhiro Ogawa,[5] Tomoya Terashima,[5] Hiroshi Maegawa,[5] Daijiro Yanagisawa,[6] Ikuo Tooyama,[6] Masayoshi Tada,[7] Osamu Onodera,[7] and Kiyoshi Hayasaka[2,*]

Charcot-Marie-Tooth disease (CMT) is the most common inherited neuropathy characterized by clinical and genetic heterogeneity. Although more than 30 loci harboring CMT-causing mutations have been identified, many other genes still remain to be discovered for many affected individuals. For two consanguineous families with CMT (axonal and mixed phenotypes), a parametric linkage analysis using genome-wide SNP chip identified a 4.3 Mb region on 12q24 showing a maximum multipoint LOD score of 4.23. Subsequent whole-genome sequencing study in one of the probands, followed by mutation screening in the two families, revealed a disease-specific 5 bp deletion (c.247−10_247−6delCACTC) in a splicing element (pyrimidine tract) of intron 2 adjacent to the third exon of cytochrome *c* oxidase subunit VIa polypeptide 1 (*COX6A1*), which is a component of mitochondrial respiratory complex IV (cytochrome *c* oxidase [COX]), within the autozygous linkage region. Functional analysis showed that expression of *COX6A1* in peripheral white blood cells from the affected individuals and COX activity in their EB-virus-transformed lymphoblastoid cell lines were significantly reduced. In addition, *Cox6a1*-null mice showed significantly reduced COX activity and neurogenic muscular atrophy leading to a difficulty in walking. Those data indicated that *COX6A1* mutation causes the autosomal-recessive axonal or mixed CMT.

Charcot-Marie-Tooth disease (CMT) is the most common inherited neuromuscular disease characterized by clinical and genetic heterogeneity, which has been traditionally subdivided into two major subgroups, demyelinating and axonal forms. Initially, genes encoding major myelin proteins including peripheral myelin protein 22 (*PMP22* [MIM 601097])[1] and myelin protein zero (*MPZ* [MIM 159440])[2] were identified as genes involved in demyelinating CMT as well as mitofusin 2 (*MFN2* [MIM 608507])[3] in axonal CMT. So far, mutations causing CMT have been identified in more than 30 loci. To identify the genetic background of Japanese CMT, we analyzed the disease-causing mutation in about 350 affected individuals; however, we could not identify the causative mutations in about 50% of demyelinating CMT and 80% of axonal CMT.

We studied two families with affected members from consanguineous marriages at different sites in Japan (Figure 1A). The affected siblings of family 1 initially had slightly reduced median motor nerve conduction velocities (NCVs) and onion bulb formation in the sural nerve at young ages, but they had motor NCVs below 38 m/s when they were aged 30–39 years, indicating mixed CMT (Tables S1 and S2 available online). The affected member of family 2 had slightly reduced median motor NCVs at age 39, indicating axonal phenotype (Tables S3 and S4). Although the two families have no record of affinal connections each other, their affected members share similar disease phenotypes and a deduced mode of inheritance (i.e., recessive). Informed consent was obtained from all subjects, and all procedures were approved by the Research Ethics Committees of Yamagata University Faculty of Medicine and Niigata University School of Medicine. For these families, family 1 with two affected members from the marriage between second cousins and family 2 with one from that between first cousins once removed (Figure 1A), we carried out a parametric linkage study using genome-wide 6.5K SNP chip. A genome-wide SNP genotyping was performed for ten members from the two families with the HumanLinkage V Panel (Illumina). A multipoint linkage analysis was performed with Allegro v.2 and a human genetic map based on NCBI dbSNP Build 123 and identified a significant linkage in a 4.3 Mb region on 12q24 covered by nine SNPs showing multipoint LOD scores of 3.85–4.23 at $\theta = 0.00$ (Figures 1B, 1C, and S1; Table 1).

Subsequent whole-genome sequencing for the proband from family 1 was performed on a Genome Analyzer IIx system (Illumina), followed by annotation, filtering, and assessing potential effects of variants (Tables S5 and S6). It yielded enough data (a mean depth of 49.6×) to detect sequence variants by comparison with a reference sequence (Table S5). Of a total of 3,716,455 variants

[1]Advanced Molecular Epidemiology Research Institute, Faculty of Medicine, Yamagata University, Yamagata 990-9585, Japan; [2]Department of Pediatrics, Faculty of Medicine, Yamagata University, Yamagata 990-9585, Japan; [3]Research Institute of Medical Sciences, Faculty of Medicine, Yamagata University, Yamagata 990-9585, Japan; [4]Department of Reproductive Biology and Medicine, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan; [5]Department of Medicine, Shiga University of Medical Science, Otsu 520-2192, Japan; [6]Molecular Neuroscience Research Center, Shiga University of Medical Science, Otsu 520-2192, Japan; [7]Department of Molecular Neuroscience, Brain Research Institute, Niigata University, Niigata 951-8585, Japan
*Correspondence: gtamiya@genetix-h.com (G.T.), hayasaka@med.id.yamagata-u.ac.jp (K.H.)
http://dx.doi.org/10.1016/j.ajhg.2014.07.013. ©2014 by The American Society of Human Genetics. All rights reserved.

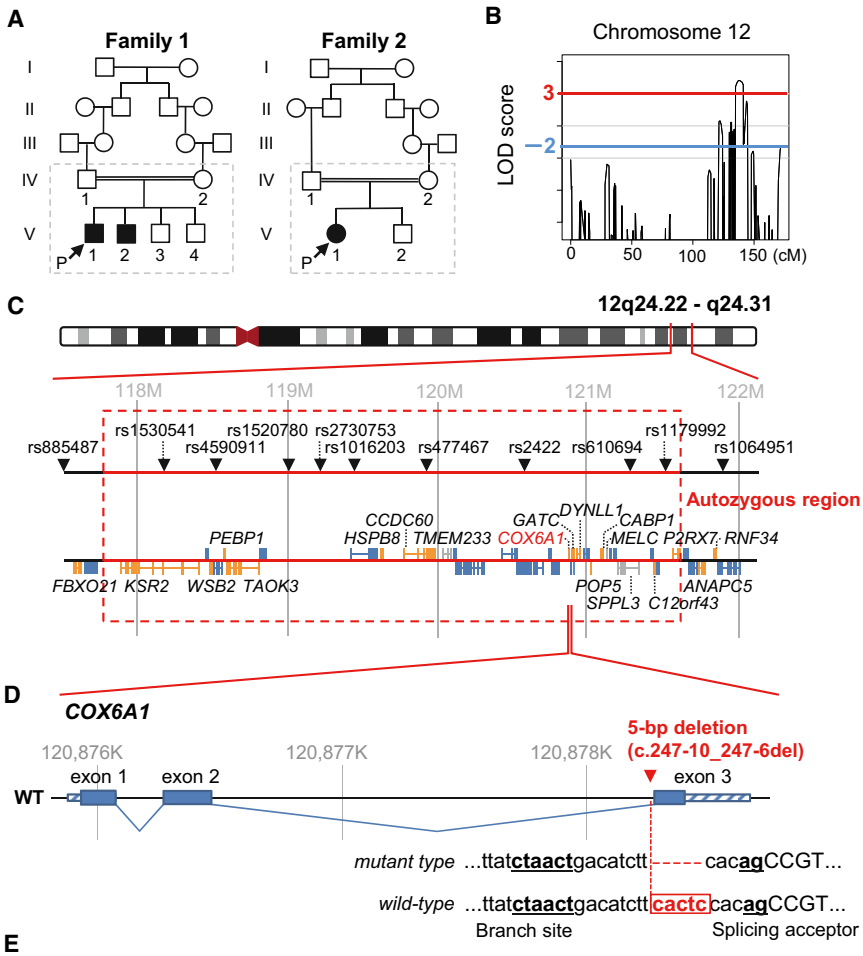

Figure 1. Genomic Analyses in the CMT Families

(A) Pedigrees of two consanguineous CMT families, from which a total of ten members in broken line boxes were sampled.

(B) Multipoint LOD scores on chromosome 12 using these ten members from two families with the HumanLinkage V SNP chip. We extracted genomic DNA from peripheral blood using the QIAamp DNA spin columns (QIAGEN) and quantified them using the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies). Genotyping was performed with the GoldenGate Genotyping Universal-32 kit on a BeadArray Reader System (Illumina) according to the manufacturer's assay guide. Genotype calls were made using the Genotyping module of the GenomeStudio v2009.1 software (Illumina). A multipoint linkage analysis was performed with Allegro v.2[4] and a human genetic map based on NCBI dbSNP Build 123 (see also Figure S1).

(C) Gene map in the significant linkage region on 12q24. These physical coordinates are taken from UCSC Genome Browser build hg19, RefSeq, and dbSNP 138 (UCSC Genome Browser database: 2014 update).

(D) A disease-specific 5 bp deletion (c.247−10_247−6delCACTC) in the pyrimidine tract near to the splicing acceptor near to the third exon of *COX6A1*.

(E) Validation and distribution of the 5 bp deletion among the CMT family members by Sanger-based PCR direct sequencing (see also Table S10).

(Table S7), after filtering out known or heterozygous variants, six were suspected to be deleterious by a majority rule approach using five prediction methods: PolyPhen-2, Grantham, PROVEAN, SIFT, and Mutation Taster (Table 2; see also Table S8 for a relaxed rule). Among them, only a 5 bp deletion (c.247−10_247−6delCACTC; RefSeq accession number NM_004373.3) was located just in a long run of homozygosity spanning 3.7 Mb, between rs10774925 and rs503720, in the significant linkage region on 12q24 (Figures 1C and 1D; Tables 1 and S9; see also Figure S2 and Table S10). In the two families, the 5 bp deletion was cosegregated perfectly with the disease state (Figures 1E and S3). Mutation screening by PCR assay detected the 5 bp deletion in none of 1,452 control chromosomes from diverse ethnic groups (Table S11).

As expected from the fact that this disease-specific 5 bp deletion disrupted an intronic splicing element (pyrimidine tract[11–13]) adjacent to the boundary of the second intron and the third exon of cytochrome *c* oxidase subunit VIa polypeptide 1 (*COX6A1* [MIM 602072]) whose expression is almost ubiquitous excepting skeletal muscles (see Su et al.[10] for human and Figure S4 for mouse), quantitative RT-PCR assay detected the significantly reduced (<1/5 of controls; $p < 0.001$) expression of the normally spliced form of *COX6A1* in peripheral white blood cells from two affected members of family 1 (Figure 2A, see also Table S12). Consistent with these findings, their cultured cell lines showed statistically significant reductions in the *COX6A1* expression ($p < 0.001$; Figure 2B, see also Table S12), cytochrome *c* oxidase (COX) activity (Figure 2C), and total ATP contents (Figure 2D) in these affected individuals relative to controls. Detailed descriptions are found in the legends of figures and tables for all experiments including expression analysis of *COX6A1* mRNA (see also Figure S5 and Table S12), determination of COX activity, and ATP amount. We performed two additional experiments: (1) the qRT-PCR using three

Table 1. Multipoint LOD Scores on 12q24 and Estimated Haplotypes in the Significant Linkage Region

			Family 1												Family 2							
Markers	Positions[a]	LOD Score	IV-1 (Father)		IV-2 (Mother)		V-1 (Affected)		V-2 (Affected)		V-3 (Brother)		V-4 (Brother)		IV-1 (Father)		IV-2 (Mother)		V-1 (Affected)		V-2 (Brother)	
rs885487	117539934	−infinity	2	1	2	1	2	1	2	2	2	1	1	1	2*	2	2*	2	2*	2*	2*	2
rs1530541	118229731	4.0597	2*	2	2*	2	2*	2*	2*	2*	2*	2	2*	2	2*	1	2*	2	2*	2*	2*	1
rs4590911	118562757	4.2433	1*	2	1*	2	1*	1*	1*	1*	1*	2	1*	2	2*	2	2*	2	2*	2*	2*	2
rs1520780	119039725	4.2338	2*	1	2*	2	2*	2*	2*	2*	2*	1	2*	1	2*	2	2*	2	2*	2*	2*	2
rs2730753	119171974	4.2317	2*	1	2*	2	2*	2*	2*	2*	2*	1	2*	1	2*	2	2*	1	2*	2*	2*	2
rs1016203	119423542	4.2249	2*	2	2*	2	2*	2*	2*	2*	2*	2	2*	2	2*	2	2*	1	2*	2*	2*	2
rs477467	119927992	4.1386	2*	2	2*	1	2*	2*	2*	2*	2*	2	2*	2	2*	2	2*	2	2*	2*	2*	2
rs2422	120565188	4.0416	1*	2	1*	1	1*	1*	1*	1*	1*	2	1*	2	2*	2	2*	2	2*	2*	2*	2
5 bp del	120878247		+*	–	+*	–	+*	+*	+*	+*	+*	–	+*	–	+*	–	+*	–	+*	+*	+*	–
rs610694	121304826	3.8859	2*	2	2*	2	2*	2*	2*	2*	2*	2	2*	2	1*	1	1*	2	1*	1*	1*	1
rs1179992	121495432	3.8466	1*	2	1*	2	1*	1*	1*	1*	2*	2	1*	2	2*	2	2*	1	2*	2*	2*	2
rs1064951	121878659	−1.7247	1	2	2	2	1	2	1	2	2	2	2	2	2*	2	2*	2	2*	2*	2*	2

Haplotypes were reconstructed using Allegro v.2 with "HAPLOTYPE" option from each family data including 5 bp deletion (c.247−10_247−6delCACTC). Asterisk indicates the putative disease-associated haplotypes in each family.
[a]Physical coordinate is taken from dbSNP 138 (UCSC Genome Browser, 2014 update).

primers specific to the mutant transcript showed its low level (Figure S6 and Table S12); and (2) the immunoblot showed no band at around 11.0 kDa for the mutant protein (Figure S7). These results suggest that the deletion leads to alternative splicing events triggering a potential nonsense-mediated RNA decay.

Finally, we examined commercially provided *Cox6a1*-null mice (Figure S8 and Table S13). All experimental procedures were performed according to the animal welfare regulations of Yamagata University Faculty of Medicine and Shiga University of Medical Science, and the study protocol was approved by the Animal Subjects Committee of Yamagata University Faculty of Medicine and Shiga University of Medical Science. The null mice showed lack of COX6A1 products (Figure 3A) and remarkable behavioral phenotypes including difficulty in walking (Figure 3B; Movie S1). Histological examinations revealed that the null mice had thinned sciatic nerves (Figure 3C) and neurogenic muscular changes including small angular fiber and small group atrophy (Figure 3D), although no remarkable neurodegeneration was found. In addition, consistent with the findings in the affected individuals, the null mice showed statistically significant reductions in COX activity and ATP contents in liver cells ($p < 0.001$; Figures 4A and 4B) and delayed motor NCV ($p < 0.01$; Figure 4C) relative to control animals.

Table 2. List of Homozygous and Potentially Deleterious Variants Nominated by at Least Three of Five Prediction Methods

		Alleles					Changes	
Chr	Start	Ref	Variant	Gene (MIM ID)	Types	Accession	Nucleotide	Protein
12	120878247	CACTC	–	*COX6A1*	splicing	NM_004373.3	c.247−10_247−6delCACTC	NA
2	114257705	C	–	*FOXD4L1* (MIM 611084)	frame-shift	NM_012184.4	c.876del	p.Gly293Alafs*151
14	24470691	–	A	*DHRS4L2* (MIM 615196)	frame-shift	NM_001193637.1	c.204_205insA	p.His69Thrfs*69
19	20807178	–	A	*ZNF626* (NA)	frame-shift	NM_001076675.2	c.1505dup	p.Ile503Hisfs*95
2	114257443	A	C	*FOXD4L1* (MIM 611084)	nonsynonymous	NM_012184.4	c.610A>C	p.Lys204Gln
8	126443464	G	T	*TRIB1* (MIM 609461)	nonsynonymous	NM_025195.3	c.320G>T	p.Arg107Leu

To assess potential variant effects for nonsynonymous variants, we used PolyPhen-2 v.2.2.2 (HumVar; score > 0.85),[5] Grantham Score from SeattleSeq Annotation v.137 (score > 151), PROVEAN v.1.1.3 (score < −2.5),[6] SIFT (score < 0.05),[7] and Mutation Taster (predicted as "disease causing")[8] and listed potentially deleterious variants voted by at least two prediction methods. Furthermore, we also referred gene expression in human whole brain on BioGPS (GeneAtlas U133A, gcrma data set)[9,10] (see also Table S8).

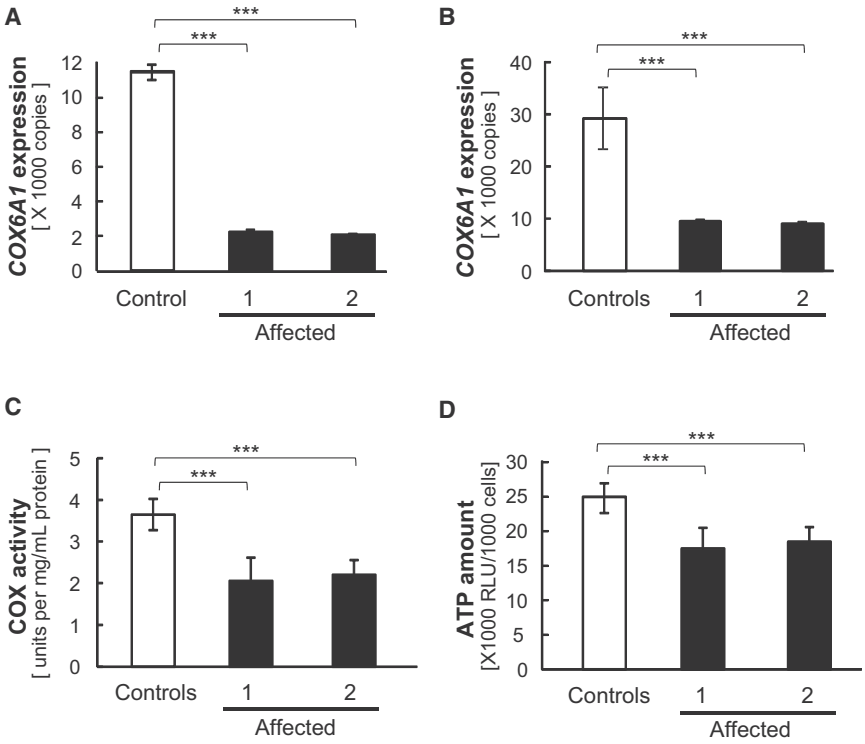

Figure 2. *COX6A1* Expression and COX Activity Levels

(A and B) *COX6A1* expression in fresh whole-blood from a control and the two affected individuals from family 1 (A) and in EBV-transformed B cell lines from four controls and the two affected individuals (B). Lymphoblastoid cells from two affected members in the family 1 were immortalized by infection with the Epstein-Barr virus (VR-1492; American Type Culture Collection). Immortalized cells from the four healthy Japanese individuals (HEV0031, HEV0032, HEV0038, and HEV0041) were provided by RIKEN BioResource Center. Total RNA was extracted using QIAamp RNA Blood Mini Kit or AllPrep DNA/RNA Kit (QIAGEN) according to the manufacturer's instructions with on-column DNase I treatment. After determining RNA concentrations using Quant-iT RiboGreen RNA Assay Kit (Life Technologies), 400 ng of total RNA per 40 μl reaction was used to synthesize cDNA using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies) with random primers. Absolute quantification for *COX6A1* was performed using a custom TaqMan assay (Table S12). PCR products were ligated into pGEM-T easy vector (Promega) and isolated plasmid DNA was then linearized by EcoRI digestion. Before use, plasmid concentration was determined by Quant-iT PicoGreen dsDNA Assay Kit and serial dilutions were performed to generate standard curve. Real-time PCR was conducted using TaqMan Universal Master Mix II (Life Technologies) with a 7500 Fast real-time PCR system (Life Technologies). Each reaction was run in triplicate and contained 2 μl of cDNA template in a final reaction volume of 20 μl and data were analyzed with 7500 Software v.2.0.2.

(C and D) COX activity (C) and ATP amount (D) in mitochondrial fractions from the same four controls and two affected individuals' cell lines. For the determination of COX activity, we used a Cytochrome *c* Oxidase Assay kit (Sigma-Aldrich). Mitochondrial fractions were obtained from cells by homogenization in homogenization buffer (20 mM HEPES-KOH [pH 7.4], 220 mM mannitol, 70 mM sucrose, 1 mM EDTA, 1× protease inhibitor). The determination of COX activity was based on a colorimetric assay that quantifies the oxidation of ferrocytochrome *c* to ferricytochrome *c* via cytochrome *c* oxidase, a reaction that results in a decrease in absorbance at 550 nm. After measurement of absorbance, COX activity was calculated according to manufacturer's instructions. The amount of ATP was measured using a Lumino assay detecting cellular-ATP kit (CA100; Wako Pure Chemical Industries). The collected cells (1,000 cells per well) were used according to manufacturer's instructions. All experiments were triplicated per sample and tested using t test for the difference between controls mean and each affected individual mean. ***$p < 0.001$. The error bars represent the standard deviation.

Haplotype estimation around the linkage region reveals that the 5 bp deletion is harbored in background haplotypes that differ between family 1 and 2 (Tables 1 and S14; Figures S9 and S10), suggesting that this mutation occurred independently on each the haplotype origin. The two independent mutations in human CMT families and the null allele in mice consistently indicate that the 5 bp deletion can cause the disease phenotype through deficiency of *COX6A1* leading to the reduced COX activity in affected individuals' peripheral nerves as well. The *COX6A1* is expressed in the sciatic nerve as well as other tissues including lung, kidney, liver, and brain. Despite its expression in multiple tissues, the affected individuals carrying *COX6A1* mutation show no signs and symptoms other than that of the peripheral nervous system. It is interesting to note that the surfeit 1 (*SURF1* [MIM 185620]) encodes one of assembly factors of COX whose mutations cause demyelinating CMT accompanying with axonal loss[14] leading to multisystem involvement with nystagmus, hearing loss, kyphoscoliosis, and brain MRI abnormalities. In contrast, COX6A1 is small polypeptide required for the stability of holoenzyme. It is likely that the focal symptoms of our affected individuals are due to such narrow function, hypomorphic nature of the mutant allele, and/or the possibility for the residual amount of COX6A1 to warrant sufficient COX activity in other tissues as to maintain bioenergetic compensation. As another possibility, COX6A1 deficiency might be compensated by heart/muscle isoform, cytochrome *c* oxidase subunit VIa polypeptide 2 (COX6A2) in the other tissues or interfere with forming a stable COX assembly especially in the mitochondria of peripheral nervous system. Such mitochondrial respiratory chain deficiency might increase the vulnerability of nerve cells.[15–18] Finally, the tissue specificity might reflect the differential vulnerability to COX6A1 deficiency among them, as shown in many known examples such as TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250 kDa (TAF1),[19] huntintin (HTT),[20] and superoxide dismutase 1, soluble (SOD1).[21] However, a molecular mechanism

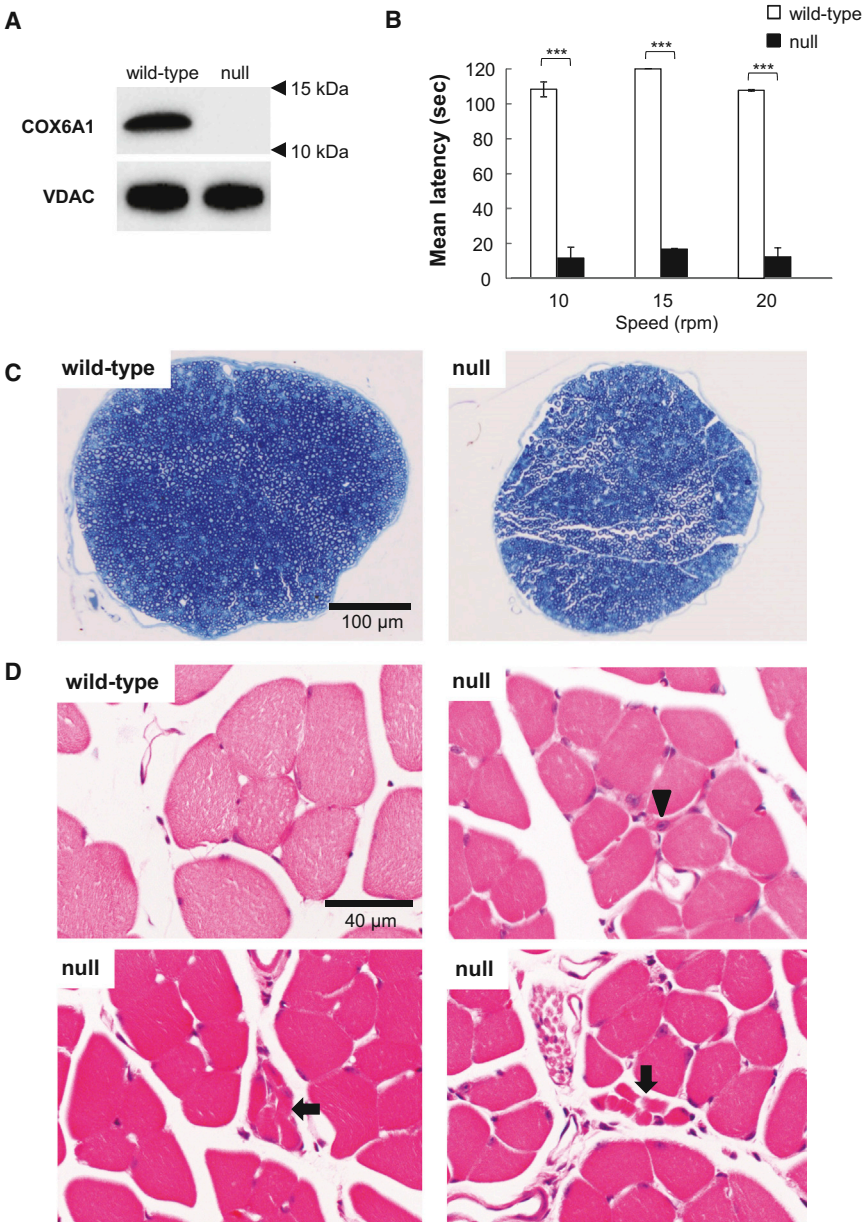

Figure 3. Characterization of *Cox6a1* Knockout Mice

(A) Immunoblot of COX6A1 and voltage-dependent anion channel (VDAC) as control in a wild-type and *Cox6a1* knockout null mice. Mice aged 7–8 weeks, three (one male and two female) knockout and three (two male and one female) wild-type were anesthetized and perfused with 10 mM PBS. Mitochondria fractions were obtained from liver tissue and applied to immunoblot. 20 μg of protein was loaded into a SDS-PAGE and transferred to the polyvinylidene difluoride membrane. The primary antibody for COX6A1 (mouse monoclonal, ab110265; abcam) is diluted 1:1,000 and secondary antibody for anti-COX6A1 (anti-mouse IgG, HRP-conjugated, 315-035-003; Jackson ImmunoResearch) is diluted 1:10,000. All blotting is carried out in 5% skim-milk/TBS solutions at room temperature for 1 hr.

(B) Motor coordination and balance was assessed as the latency to fall in the rotarod. Mice aged 7–8 weeks, four (two male and two female) knockout and four (two male and two female) wild-type were used. Each mouse underwent the same 4 day procedure on a rota-rod (MK-660A; Muromachi Kikai). The first 3 days were used to train the mice (four sessions of 60 s each, walking at 20 rpm). The test sessions were run on the last day. The mice performed two series of three trials (10, 15, and 20 rpm) at each speed, with a 10 min rest period between trials. The latency to fall was recorded with a cut-off at 120 s. The difference between the wild-type and knockout null mice means were tested using t test. ***$p < 0.001$. The error bars represent the standard deviation.

(C and D) Histological examinations by toluidine blue staining sections of mice sciatic nerves at lower magnification (C) and hematoxylin-eosin staining sections of mice lower limb muscles (D). Arrow indicates a smaller number of fibers are involved in small group atrophy and arrowhead indicates small angular fibers despite the limited numbers. Mice aged 7–8 weeks, four (two male and two female) knockout and four (two male and two female) wild-type were anesthetized and perfused with 10 mM PBS, followed by a fixative of 4% paraformaldehyde (for leg muscle) or 2.5% glutaraldehyde (for sciatic nerve) in 0.1 M phosphate buffer. Sciatic nerve specimens were fixed in 2.5% glutaraldehyde in phosphate buffer for 2 hr at room temperature. After post-fixation with 1% OsO_4, the tissues were embedded in epoxy resin. Tissue blocks were sectioned at 1 mm thickness and stained with toluidine blue for light microscopic examination. For the histological analysis of leg muscle, mice tissues were postfixed in 4% paraformaldehyde for 48 hr at 4°C. The muscle tissues were dissected out and then incubated overnight in 10% sucrose in phosphate buffer. After snap freezing with CO_2 gas, tissue blocks were sectioned at 20 μm thickness in a cryostat and stained with hematoxylin and eosin.

between potential dysfunctions of the respiratory chain and the disease phenotype is still unclear. Our findings warrant further mechanistic analysis of structural analysis of COX holoenzyme, COX activity, and assembly in different tissues, including brain, heart, and, as a "positive control," skeletal muscle. To disentangle their high-order functional interplay, the results from our genetic study would warrant further mechanistic analyses of the mitochondrial involvements in the peripheral nervous system.

On the other hand, the multiple mutation events resulting in the same deletion in the pyrimidine tract imply that this site may be a mutational hotspot in human, so that there is a possibility that this deletion would be found across the world, especially from families with consanguineous loop, under the same scenario as Charcot-Marie-Tooth disease type 2F (CMT2F [MIM 606595])[22] or hereditary motor and sensory neuropathy, proximal type (HMSN-P [MIM 604484]).[23,24]

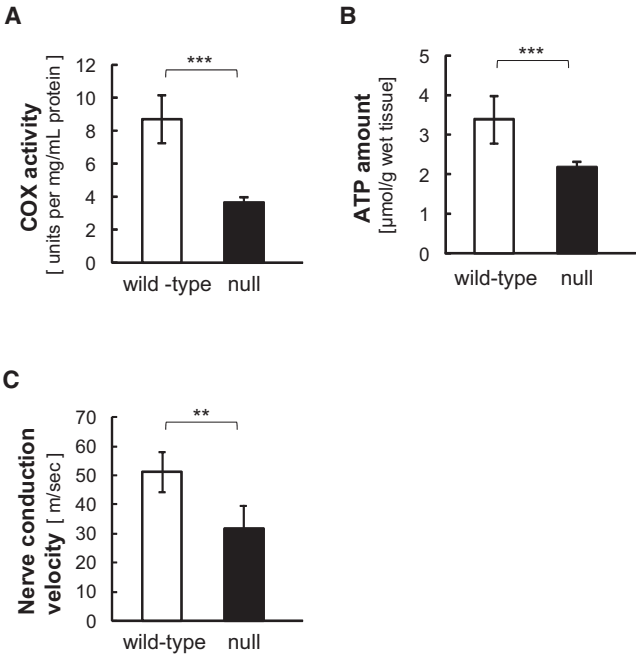

Figure 4. COX Activity and Electrophysiological Analysis in Mice
(A and B) COX activity in mitochondrial fractions from livers (A) and ATP amount in liver homogenates (B) from three wild-type and knockout null mice, respectively. Experiments for COX activity and ATP amount were triplicated per sample and all experiments were tested using t test for the difference between the wild-type and knockout mice means.
(C) Nerve conduction velocity of sciatic nerve from mice. Studies were demonstrated at 7–10 weeks of age, six (three male and three female) wild-type and seven (five male and two female) null mice with Sapphire (Medelec) under anesthesia of pentobarbital sodium (5 mg/kg i.p.). At the right dorsal femoral, sciatic nerve was exposed by opening up overlying skin and electrically stimulated using needle electrodes at the sciatic notch and at the knee joint level under 37°C ± 0.5°C. The compound muscle action potential (CMAP) evoked by the two parts of stimuli were recorded at gastrocnemius. Motor nerve conduction velocity (NCV) was calculated with dividing the distance between the sciatic notch and knee joint level by the delta latency between the two CMAP curves. Asterisks indicate the level of statistical significance (**$p < 0.01$; ***$p < 0.001$). The error bars represent the standard deviation.

Supplemental Data

Supplemental Data include 10 figures, 14 tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2014.07.013.

Acknowledgments

We would like to thank Kaoru Honaga, Michiyuki Kawakami, Akio Kimura, and Keiko Murayama for clinical information. This work was supported by Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan to K.H. (grant number 25461537), A.A. (grant number 25860842), and G.T. (grant number 25129701) and Grant-in-Aid from the Global Center of Excellence program of the Japan Society for the Promotion of Science, "Formation of an International Network for Education and Research of Molecular Epidemiology."

Received: April 6, 2014
Accepted: July 29, 2014
Published: August 21, 2014

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
BioGPS, http://biogps.org/
dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/
MutationTaster, http://www.mutationtaster.org/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
PolyPhen-2, http://www.genetics.bwh.harvard.edu/pph2/
PROVEAN, http://provean.jcvi.org/index.php
RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq
SeattleSeq Annotation 137, http://snp.gs.washington.edu/SeattleSeqAnnotation137/
SIFT, http://sift.bii.a-star.edu.sg/
SNP Genetic Mapping, http://integrin.ucd.ie/cgi-bin/rs2cm.cgi
UCSC Genome Browser, http://genome.ucsc.edu

References

1. Lupski, J.R., de Oca-Luna, R.M., Slaugenhaupt, S., Pentao, L., Guzzetta, V., Trask, B.J., Saucedo-Cardenas, O., Barker, D.F., Killian, J.M., Garcia, C.A., et al. (1991). DNA duplication associated with Charcot-Marie-Tooth disease type 1A. Cell *66*, 219–232.
2. Hayasaka, K., Himoro, M., Sato, W., Takada, G., Uyemura, K., Shimizu, N., Bird, T.D., Conneally, P.M., and Chance, P.F. (1993). Charcot-Marie-Tooth neuropathy type 1B is associated with mutations of the myelin P0 gene. Nat. Genet. *5*, 31–34.
3. Züchner, S., Mersiyanova, I.V., Muglia, M., Bissar-Tadmouri, N., Rochelle, J., Dadali, E.L., Zappia, M., Nelis, E., Patitucci, A., Senderek, J., et al. (2004). Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. Nat. Genet. *36*, 449–451.
4. Gudbjartsson, D.F., Jonasson, K., Frigge, M.L., and Kong, A. (2000). Allegro, a new computer program for multipoint linkage analysis. Nat. Genet. *25*, 12–13.
5. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*, 248–249.
6. Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., and Chan, A.P. (2012). Predicting the functional effect of amino acid substitutions and indels. PLoS ONE *7*, e46688.
7. Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. *4*, 1073–1081.
8. Schwarz, J.M., Rödelsperger, C., Schuelke, M., and Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. Nat. Methods *7*, 575–576.
9. Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C.L., Haase, J., Janes, J., Huss, J.W., 3rd, and Su, A.I.

(2009). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol. *10*, R130.
10. Su, A.I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K.A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. Proc. Natl. Acad. Sci. USA *101*, 6062–6067.
11. Lim, L.P., and Burge, C.B. (2001). A computational analysis of sequence features involved in recognition of short introns. Proc. Natl. Acad. Sci. USA *98*, 11193–11198.
12. Hirose, Y., Chiba, K., Karasugi, T., Nakajima, M., Kawaguchi, Y., Mikami, Y., Furuichi, T., Mio, F., Miyake, A., Miyamoto, T., et al. (2008). A functional polymorphism in THBS2 that affects alternative splicing and MMP binding is associated with lumbar-disc herniation. Am. J. Hum. Genet. *82*, 1122–1129.
13. Raben, N., Nichols, R.C., Martiniuk, F., and Plotz, P.H. (1996). A model of mRNA splicing in adult lysosomal storage disease (glycogenosis type II). Hum. Mol. Genet. *5*, 995–1000.
14. Echaniz-Laguna, A., Ghezzi, D., Chassagne, M., Mayençon, M., Padet, S., Melchionda, L., Rouvet, I., Lannes, B., Bozon, D., Latour, P., et al. (2013). SURF1 deficiency causes demyelinating Charcot-Marie-Tooth disease. Neurology *81*, 1523–1530.
15. Lin, M.T., and Beal, M.F. (2006). Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. Nature *443*, 787–795.
16. Suter, U., and Scherer, S.S. (2003). Disease mechanisms in inherited neuropathies. Nat. Rev. Neurosci. *4*, 714–726.
17. Pitceathly, R.D., Murphy, S.M., Cottenie, E., Chalasani, A., Sweeney, M.G., Woodward, C., Mudanohwo, E.E., Hargreaves, I., Heales, S., Land, J., et al. (2012). Genetic dysfunction of MT-ATP6 causes axonal Charcot-Marie-Tooth disease. Neurology *79*, 1145–1154.
18. Rinaldi, C., Grunseich, C., Sevrioukova, I.F., Schindler, A., Horkayne-Szakaly, I., Lamperti, C., Landouré, G., Kennerson, M.L., Burnett, B.G., Bönnemann, C., et al. (2012). Cowchock syndrome is associated with a mutation in apoptosis-inducing factor. Am. J. Hum. Genet. *91*, 1095–1102.
19. Makino, S., Kaji, R., Ando, S., Tomizawa, M., Yasuno, K., Goto, S., Matsumoto, S., Tabuena, M.D., Maranon, E., Dantes, M., et al. (2007). Reduced neuron-specific expression of the TAF1 gene is associated with X-linked dystonia-parkinsonism. Am. J. Hum. Genet. *80*, 393–406.
20. Sawa, A., Wiegand, G.W., Cooper, J., Margolis, R.L., Sharp, A.H., Lawler, J.F., Jr., Greenamyre, J.T., Snyder, S.H., and Ross, C.A. (1999). Increased apoptosis of Huntington disease lymphoblasts associated with repeat length-dependent mitochondrial depolarization. Nat. Med. *5*, 1194–1198.
21. Pasinelli, P., and Brown, R.H. (2006). Molecular biology of amyotrophic lateral sclerosis: insights from genetics. Nat. Rev. Neurosci. *7*, 710–723.
22. Evgrafov, O.V., Mersiyanova, I., Irobi, J., Van Den Bosch, L., Dierick, I., Leung, C.L., Schagina, O., Verpoorten, N., Van Impe, K., Fedotov, V., et al. (2004). Mutant small heat-shock protein 27 causes axonal Charcot-Marie-Tooth disease and distal hereditary motor neuropathy. Nat. Genet. *36*, 602–606.
23. Maeda, K., Kaji, R., Yasuno, K., Jambaldorj, J., Nodera, H., Takashima, H., Nakagawa, M., Makino, S., and Tamiya, G. (2007). Refinement of a locus for autosomal dominant hereditary motor and sensory neuropathy with proximal dominancy (HMSN-P) and genetic heterogeneity. J. Hum. Genet. *52*, 907–914.
24. Lee, S.S., Lee, H.J., Park, J.M., Hong, Y.B., Park, K.D., Yoo, J.H., Koo, H., Jung, S.C., Park, H.S., Lee, J.H., et al. (2013). Proximal dominant hereditary motor and sensory neuropathy with proximal dominance association with mutation in the TRK-fused gene. JAMA Neurol. *70*, 607–615.

The American Journal of Human Genetics, Volume *95*
Supplemental Data

A Mutation of *COX6A1* Causes a Recessive Axonal or Mixed Form of Charcot-Marie-Tooth Disease

Gen Tamiya, Satoshi Makino, Makiko Hayashi, Akiko Abe, Chikahiko Numakura, Masao Ueki, Atsushi Tanaka, Chizuru Ito, Kiyotaka Toshimori, Nobuhiro Ogawa, Tomoya Terashima, Hiroshi Maegawa, Daijiro Yanagisawa, Ikuo Tooyama, Masayoshi Tada, Osamu Onodera, and Kiyoshi Hayasaka

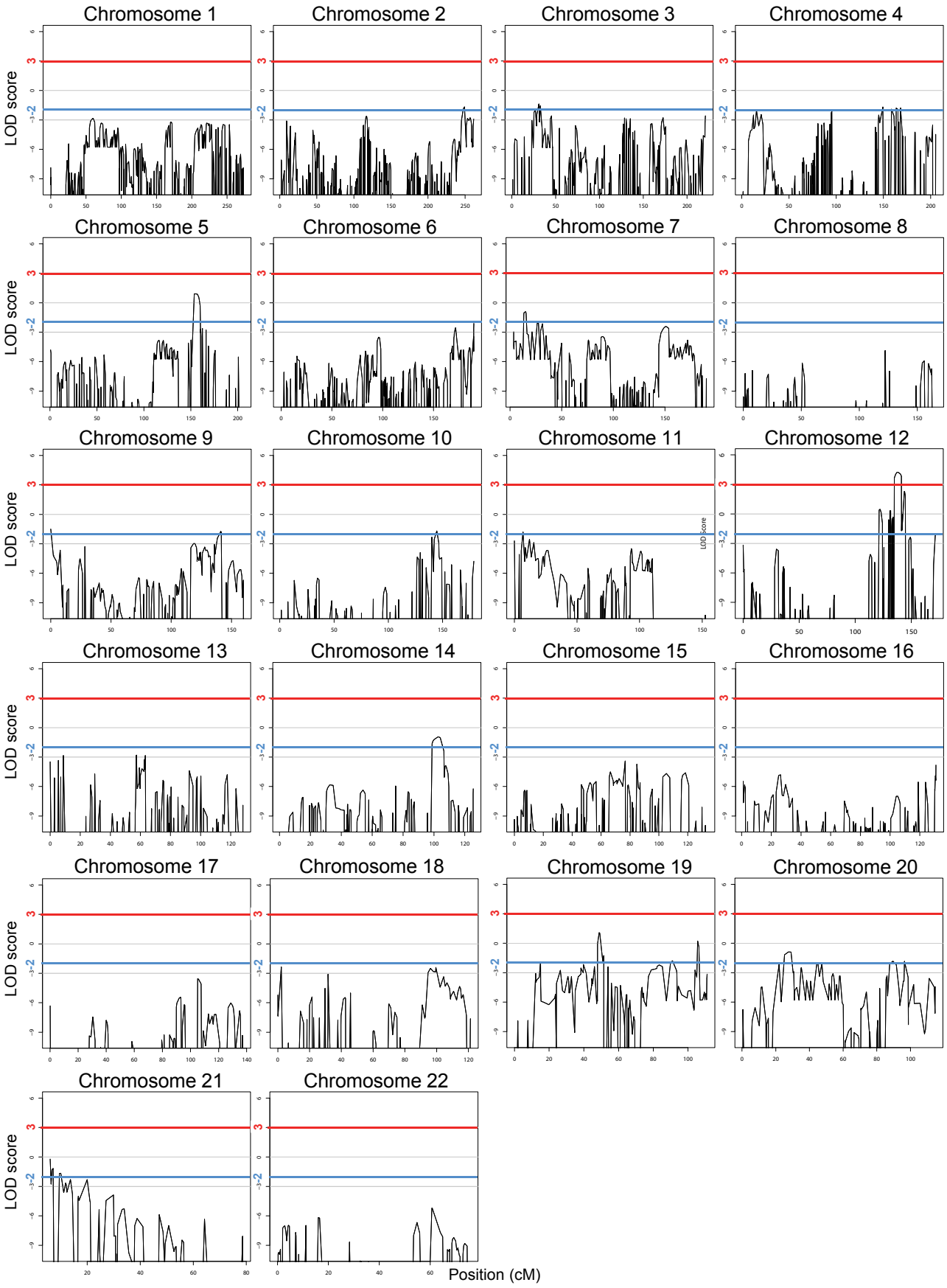

Figure S1. Genome-wide plots of multipoint LOD-scores using the HumanLinkage V SNP panel.

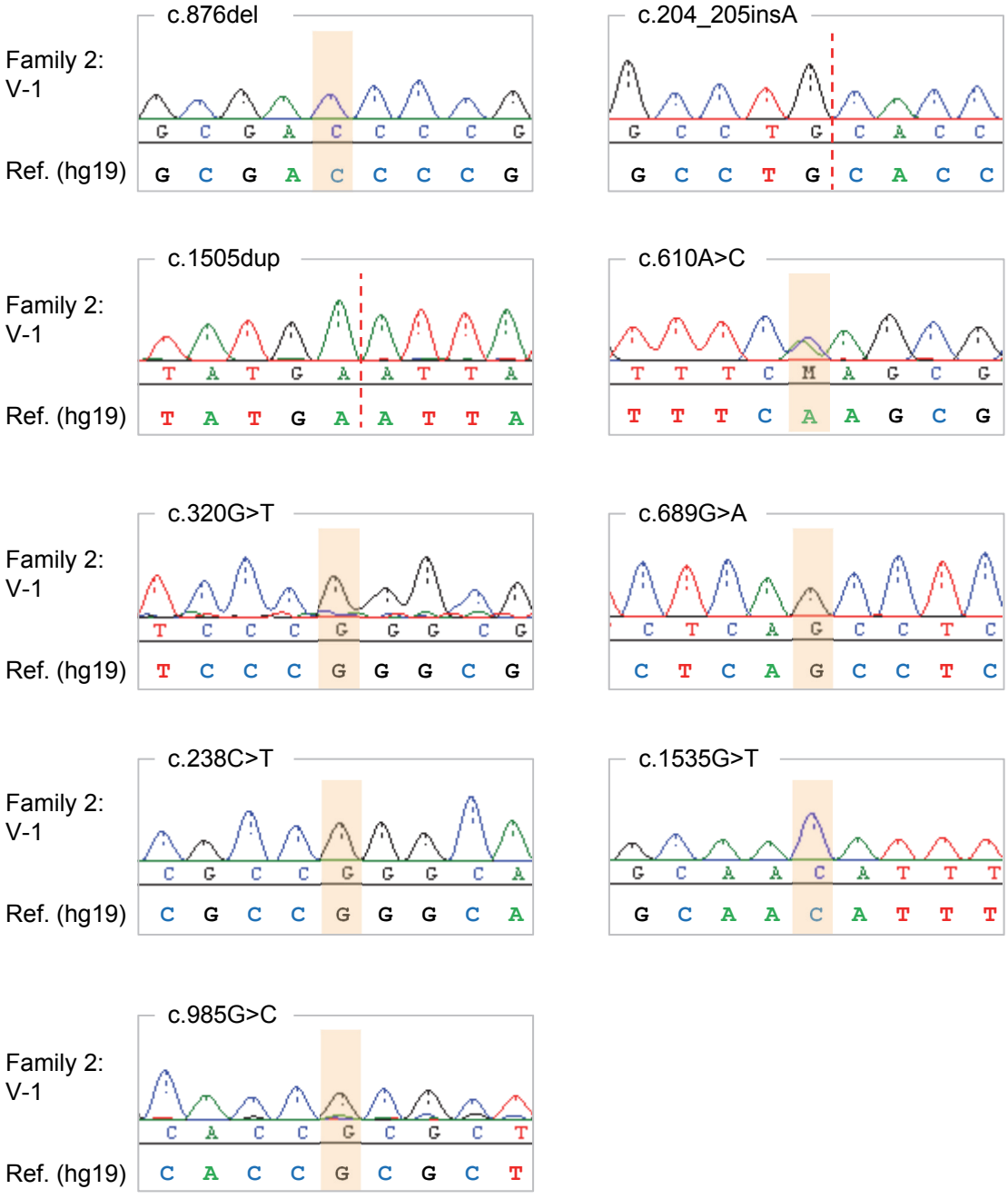

Figure S2. Exclusion of survived variants by genotyping of the proband of family 2.
In the proband of family 2, electropherograms of nine variants listed in Table S8 are shown. We successfully confirmed that these variants except for 5-bp deletion are not carried by the proband of family 2.

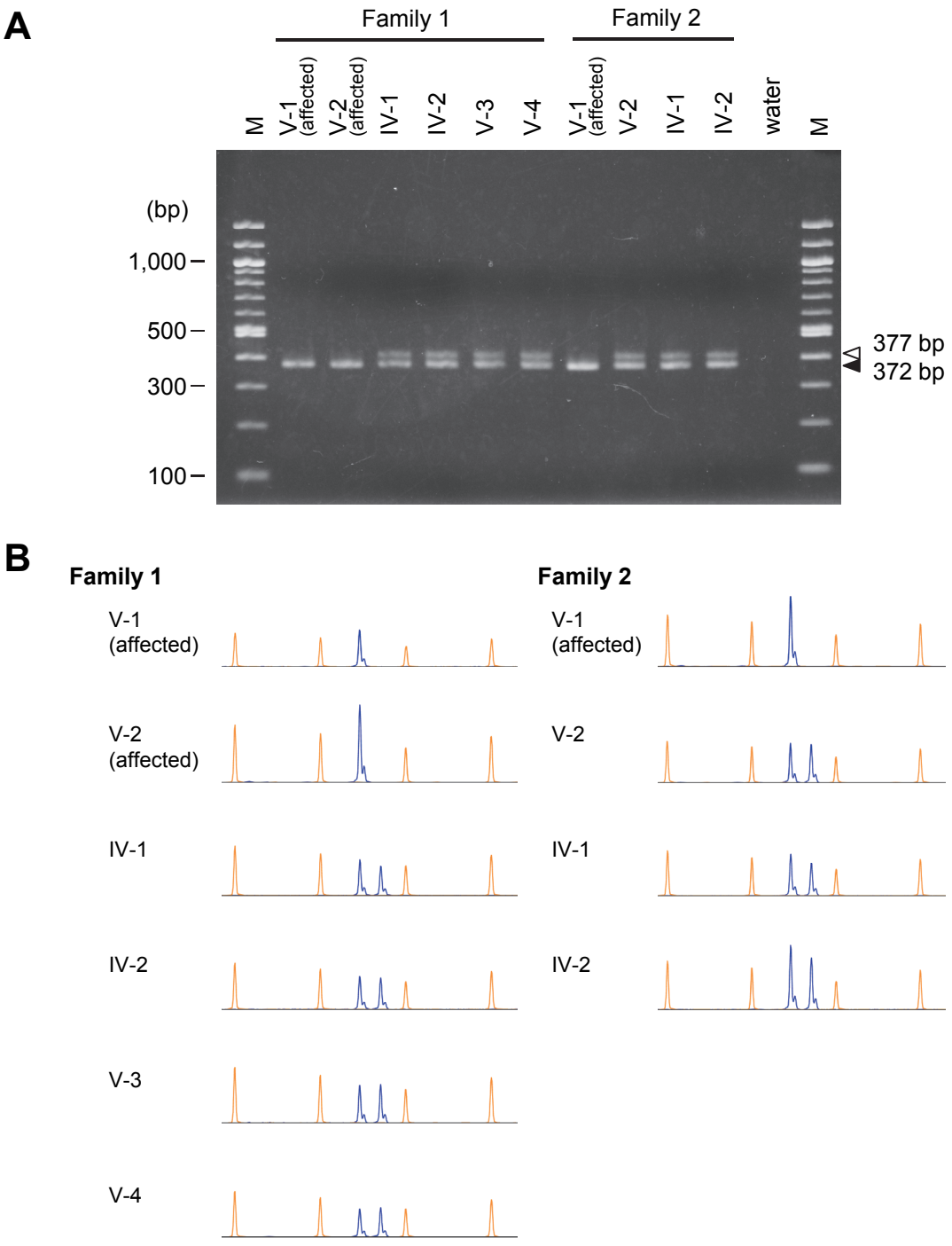

Figure S3. Detection of the 5-bp deletion in *COX6A1*.
The PCR fragments include the 5-bp deletion (372 bp) and wild-type (377 bp) were distinguishable by standard 1.5% agarose gel electrophoresis (A) or GeneScan analysis using 6-FAM labeled primer (B) (see also Figure 1E).

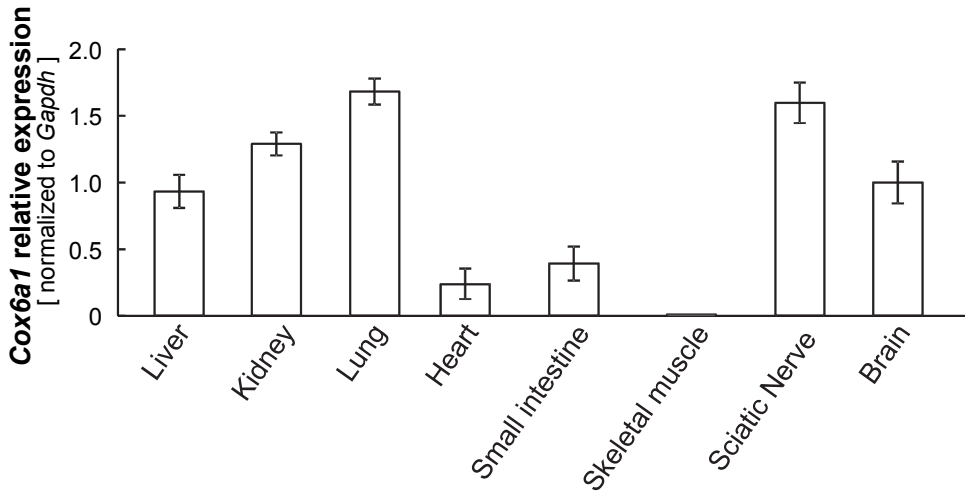

Figure S4. *Cox6a1* expression in mouse tissues.
Cox6a1 expression levels in liver, kidney, lung, heart, small intestine, skeletal muscle, sciatic nerve and brain from three wild-type mice. The error bars represent the standard deviation. Total RNA was extracted using AllPrep DNA/RNA Kit (QIAGEN N.V., The Netherlands) according to the manufacturer's instructions with on-column DNase I treatment. After determining RNA concentrations using Quant-iT RiboGreen RNA Assay Kit (Life Technologies, USA), 50 ng of total RNA per 20 μL reaction was used to synthesize cDNA using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies, USA) with random primers. Relative quantification for *Cox6a1* was performed using a pre-developed TaqMan assay (Mm01612194_m1; Life Technologies, USA). Real-time PCR was conducted using TaqMan Universal Master Mix II (Life Technologies, USA) with a 7500 Fast real-time PCR system (Life Technologies, USA). Each reaction was run in triplicate and contained 2 μL of cDNA template in a final reaction volume of 20 μL and data were analyzed following the Delta-Delta Ct method with 7500 Software v2.0.2. All samples were normalized to *Gapdh* as an endogenous control (4352662; Life Technologies, USA). The error bars represent the standard deviation.

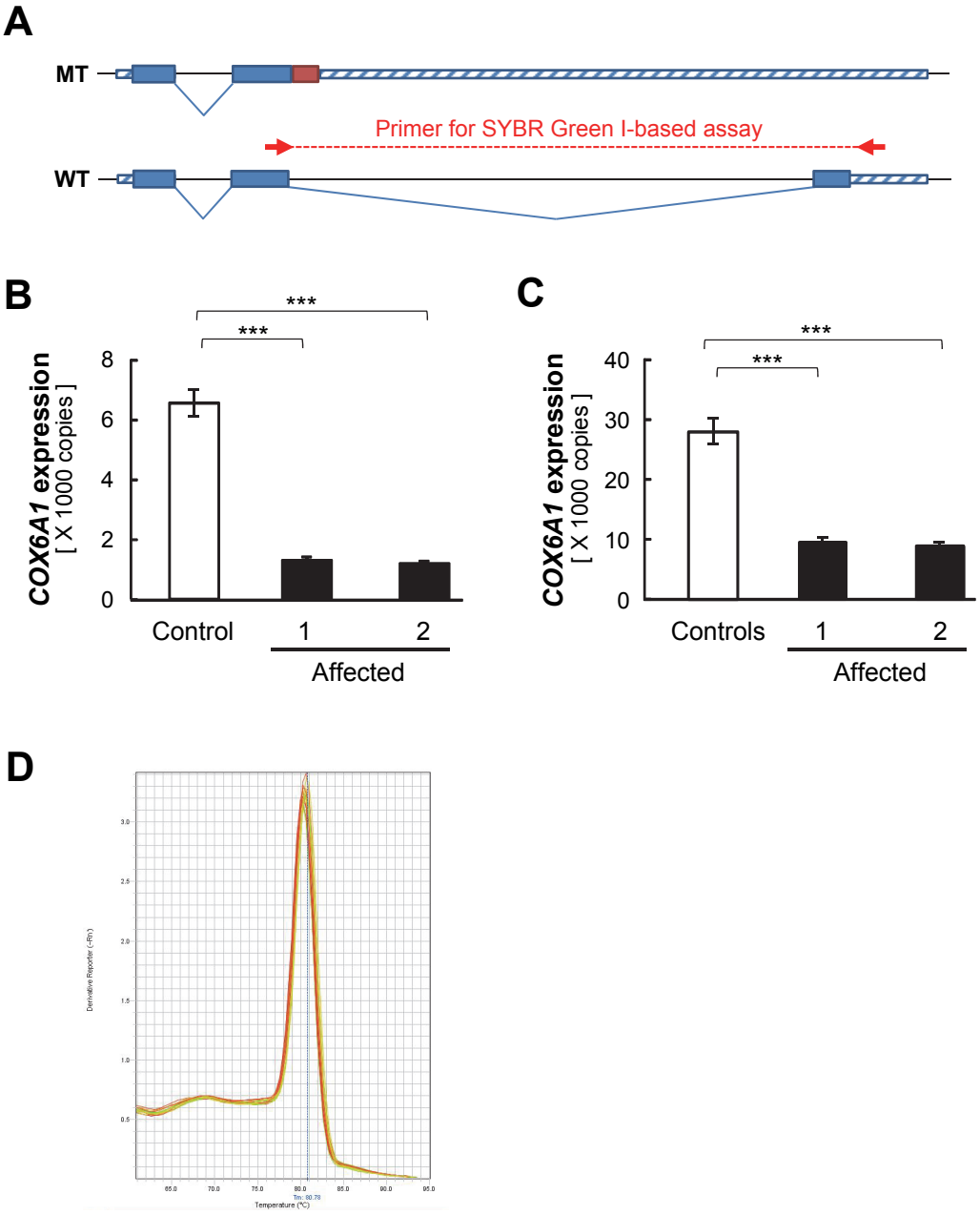

Figure S5. *COX6A1* expression analysis using SYBR Green I-based assay. Primer position (A) and absolute quantification in peripheral white blood cells (B) and EBV transformed B cell lines (C) by real-time RT-PCR. The amplicon reveals a single peak in the melting curve (D). The error bars represent the standard deviation. Three asterisks (***) indicates $p < 0.001$.

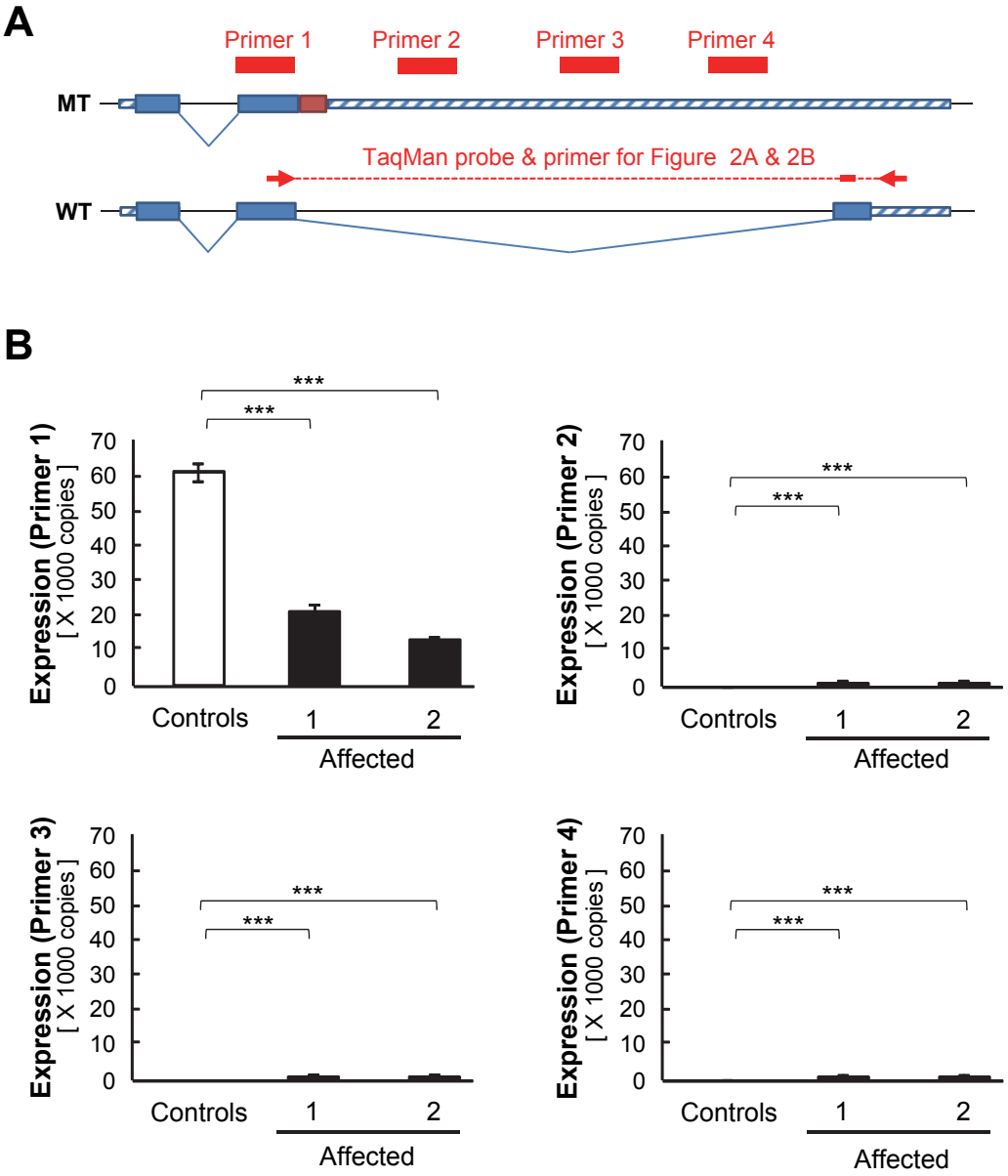

Figure S6. *COX6A1* expression analysis in EBV transformed B cell lines.
Primer position (A) and absolute quantification by real-time RT-PCR (B). To estimate the expression level of both wild-type and mutant-type *COX6A1* in details, the absolute quantification was performed using four primer sets (Table S12) with SYBR Green I fluorescent dye. The standard curves were obtained from PCR products in serial dilution. The error bars represent the standard deviation. Three asterisks (***) indicates $p < 0.001$.

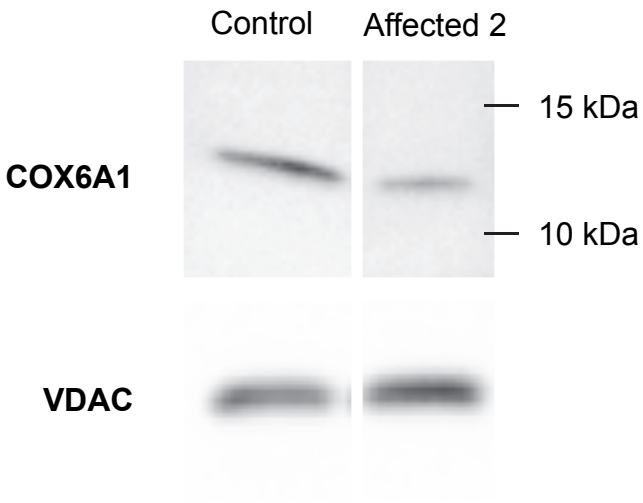

Figure S7. Western blot of COX6A1 and VDAC as control for mitochondria fractions from EBV transformed B cell lines.
The lanes run on different parts of the same blot. The primary antibody for COX6A1 (mouse monoclonal, ab110265; abcam, UK) is diluted 1:1,000. The following abbreviation is used: VDAC, voltage-dependent anion channel.

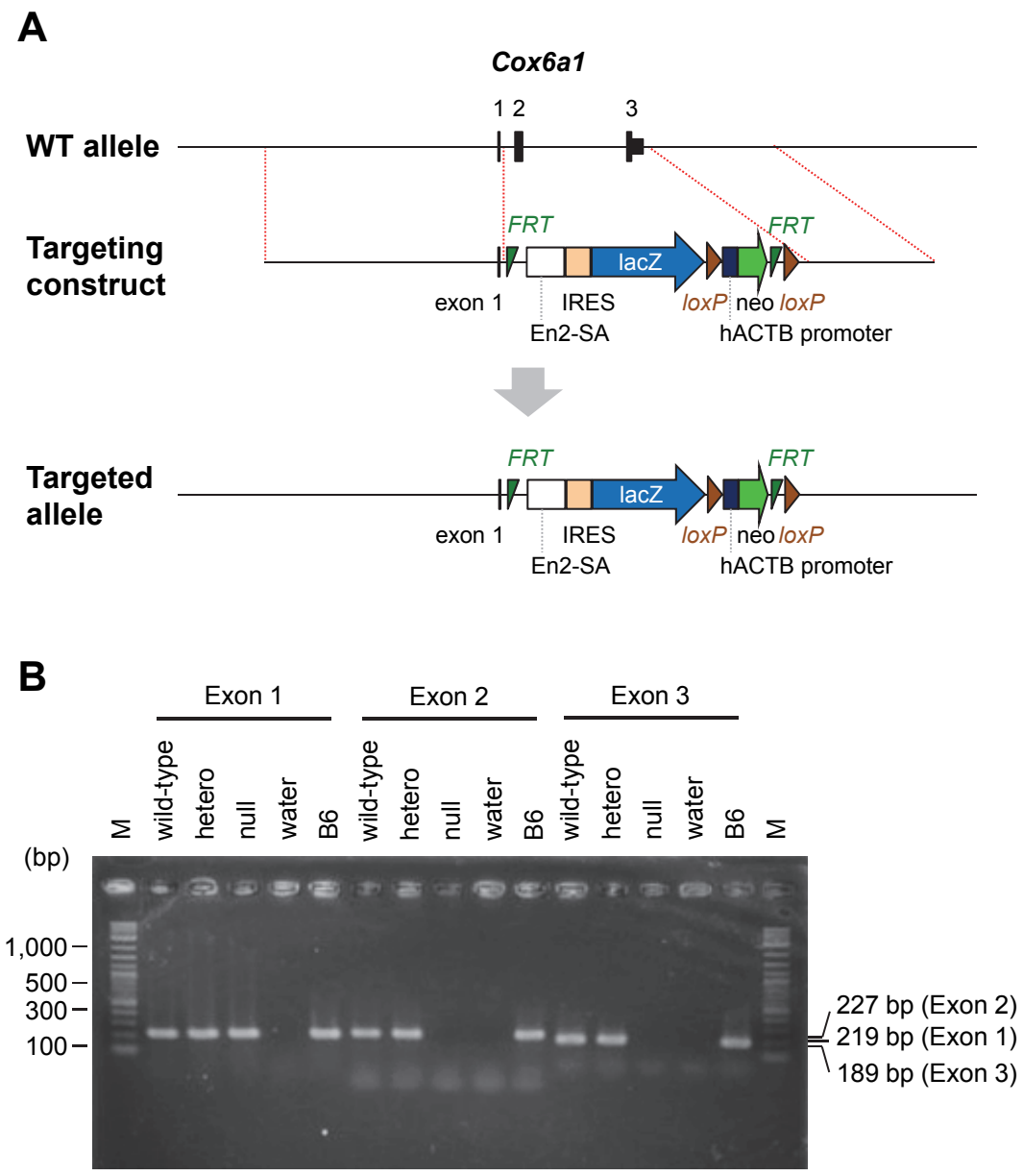

Figure S8. *Cox6a1* knockout mouse.
The knockout construct (A) and PCR genotyping result by agarose gel electrophoresis (B). The following abbreviation is used: B6, mouse background strain C57BL/6. Gene targeting was carried out as part of the International Knockout Mouse Consortium on a C57BL/6N genetic background (Project ID # CSD-74927) [1] and we obtained frozen sperm from Knockout Mouse Project (KOMP) repository, located at the University of California, Davis. Some female C57BL/6JJcl mice (CLEA, Japan) were used as oocyte donors and in vitro fertilization was performed according to the consortium's instruction. To confirm targeted allele, exon-specific PCR were carried out using three different exon-specific primer sets (Table S13). We intercrossed the knockout heterozygous mice to produce homozygous (null) mice.

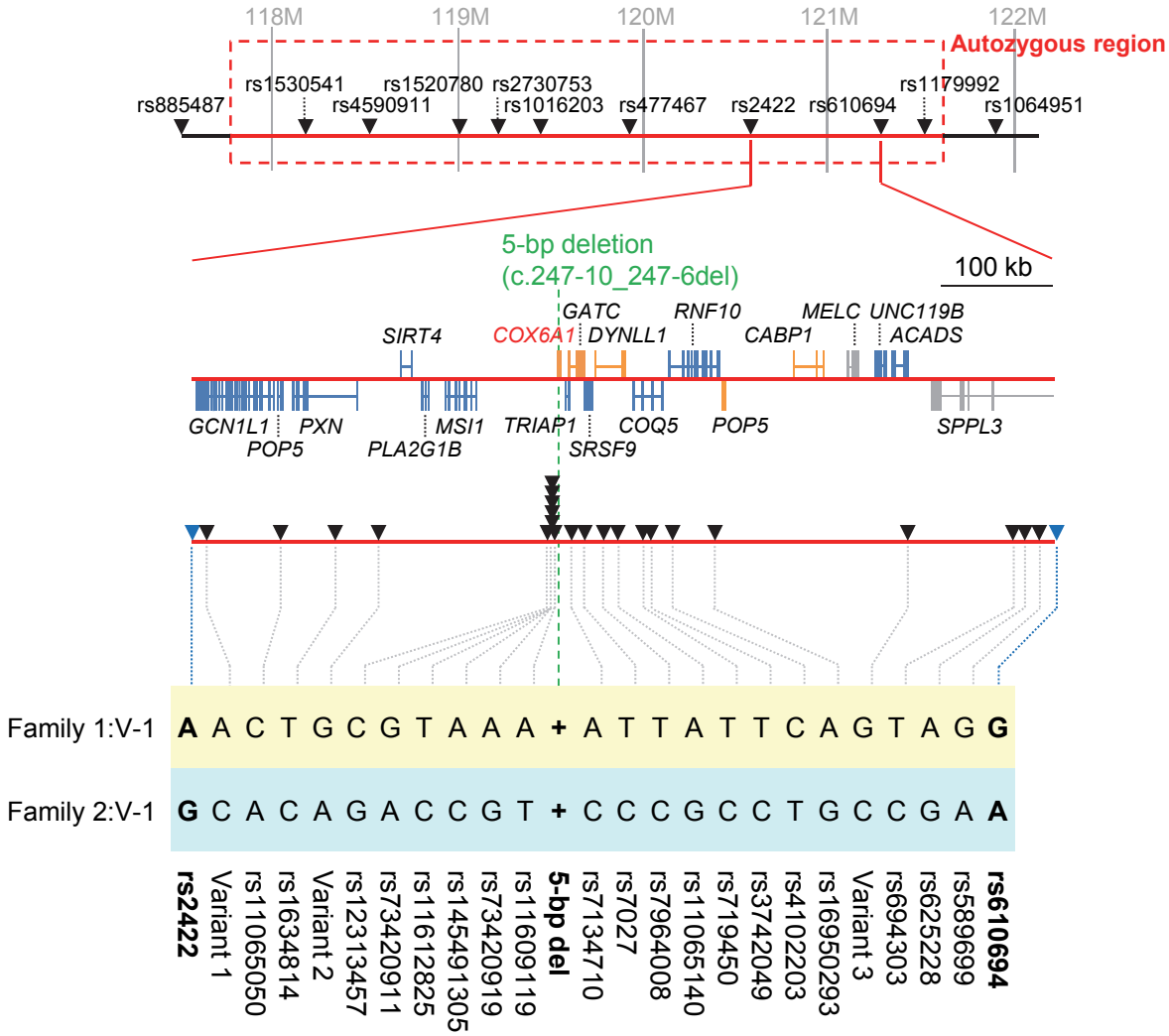

Figure S9. Background haplotype within the interval including the 5-bp deletion. In the 739 kb interval around the 5-bp deletion (between rs2422 and rs610694 from Illumina's linkage panel), 22 SNVs were selected for additional genotyping (see also Table S14). Genotypes were obtained by PCR-direct Sanger sequencing. The PCR fragments were purified by ExoSAP-IT (Affymetrix, USA) and sequenced using the same primers as in the PCR process (Figure S10).

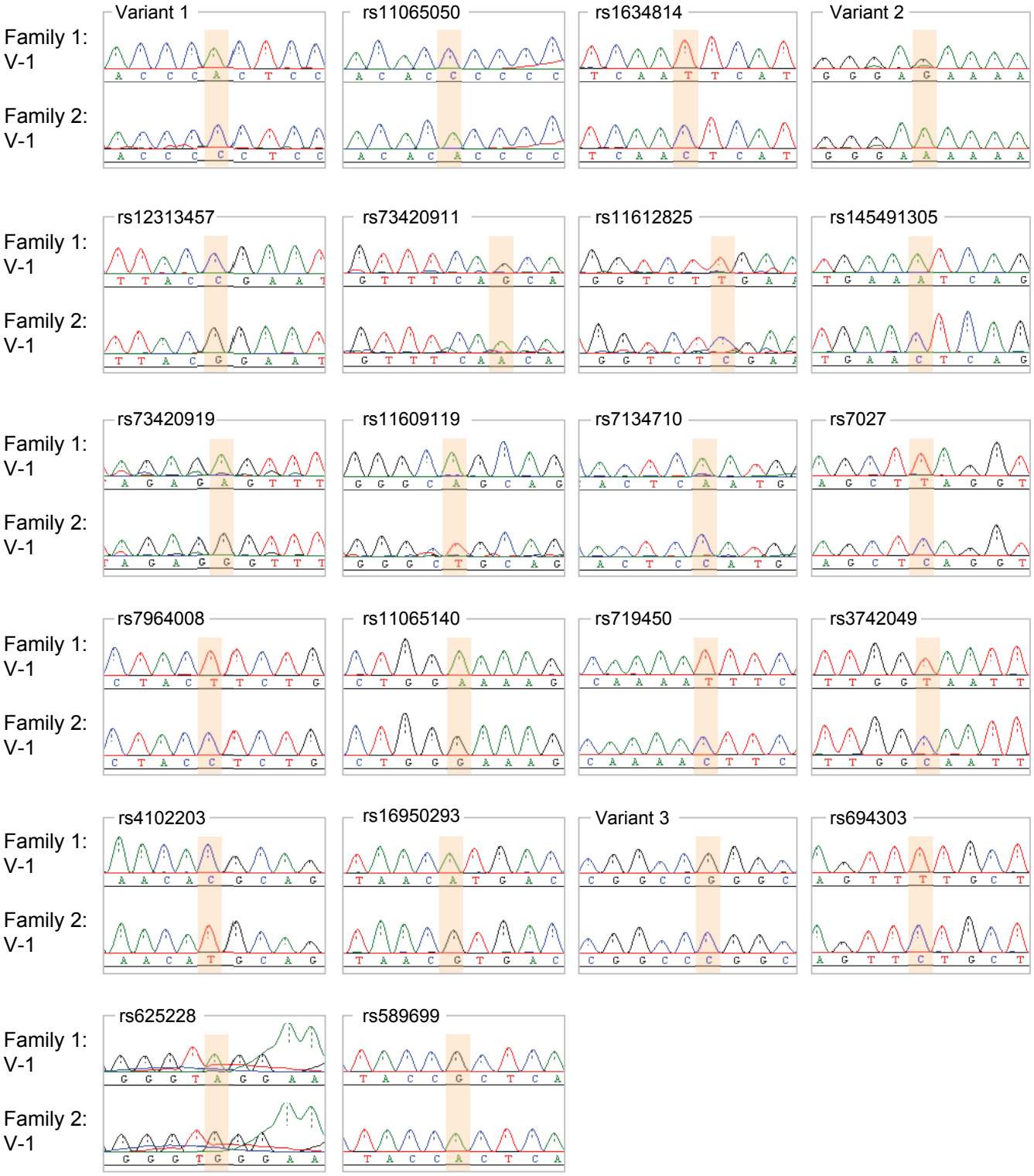

Figure S10. Electropherograms of additionally genotyped single nucleotide variants around the 5-bp deletion.

Around the 5-bp deletion (c.247-10_247-6del), the electropherograms of Sanger-based PCR direct sequencing of 22 SNVs with homozygous alleles are shown.

Table S1. Clinical course of the affected siblings of family 1

Individuals	Parents	Age	Presentation	Examinations			Other Features
				Name	Objects	Results	
Family 1 V-1	Related but healthy parents (second cousins)	Early development	Normal				
		4	Foot drop				
		8		Electrophysiological examination	right median and ulnar nerves	motor NCV 43.4 m/s and 45.0 m/s undetectable sensory nerve action potentials	Evoked action potentials were not detected in the study of motor and sensory tibial nerve.
		9		Biopsy	sural nerve	decreased myelinated fiber density, decrease of myelinated axon caliber and presence of onion bulb formation	
				MRI	brain	mild cerebral atrophy	
		~16	slowly progressed in distal muscle weakness of upper and lower limbs and was wheel-chaired				
		25		Clinical examination		distal muscle wasting of the legs, pes cavus, and clawing of the toes.	Achilles tendon reflexes were absent. Vibration detection, touch and pinprick sensation was reduced in upper and lower distal limbs. He had no cerebral, cerebellar and pyramidal signs and symptoms. He did not have any signs or symptoms of liver and renal diseases with normal laboratory findings including serum albumin, transaminases, electrolytes, urea N and creatinine, and blood lactate and pyruvate.
				Manual muscle strength testing	deltoid and triceps brachii bilaterally	grade 5	
					biceps brachii, and extensors and flexors of wrists bilaterally	grade 4	
					extensor muscles of fingers, iliopsoas, quadriceps femoris and hamstrings bilaterally	grade 2	
					tibialis anterior, gastrocnemius, and extensors and flexors of toes bilaterally	grade 0	
				EMG		proximal and distal lower limb denervation with no evidence of myopathy	
				Electrophysiological examination		delayed motor NCV (35.7 m/s) with low CMAP amplitude and undetectable sensory nerve action potentials (Table S2)	
				Biopsy	sural nerve	decreased myelinated fiber density, decrease of myelinated axon caliber and presence of onion bulb formation	
	V-2 Related but healthy parents (second cousins)			similar clinical manifestation and laboratory findings (Table S2)			

Table S2. Nerve conduction study of the affected siblings of family 1.

Nerve	Motor			F wave		Sensory		
	DL (ms)	CV (m/s)	CMAP (mV)	Lat (ms)	Occur (%)	DL (ms)	CV (m/s)	SNAP (μV)
Affected member V-1								
Rt median	7.6 (N<4.0)	35.7 (N>56)	0.677 (N>3.5)	NE				NE
Rt ulnar	3.3 (N<3.2)	40.7 (N>50)	1.3 (N>2.8)	NE				NE
Rt posterior tibial				NE				NE
Rt Peroneal				NE				NE
Affected member V-2								
Rt median	8.7 (N<4.0)	34.8 (N>56)	0.083 (N>3.5)	NE				NE
Rt ulnar	5.16 (N<3.2)	37.3 (N>50)	0.052 (N>2.8)	NE				NE
Rt posterior tibial				NE				NE
Rt Peroneal				NE				NE

DL: distal latency, CV: conduction velocity, CMAP: compound muscle action potential amplitude, Lat: latency, Occur: occurrence, SNAP: sensory nerve action potential amplitude, Rt: right, N: normal, NE: not evoked

Table S3. Clinical course of the proband of family 2

Individual	Parents	Age	Presentation	Examinations			Other Features
				Name	Objects	Results	
Family 2 V-1	Related parents	Early development	Normal				
		< 5	Unsteady gait				
		5		Clinical examination	upper and lower limbs	steppage gait, pes cavus, and decreased tendon reflexes	She was treated with short leg braces. She had a progression of distal muscle weakness of upper and lower limbs and deterioration of the ankle joint deformities.
				Manual muscle strength testing	distal upper limb with lower limbs	mild decrease	
		8	Tendon achilles lengthening surgery				
		9	Posteromedial release and anterior tibial tendon transfer surgeries				
		39	walked few steps with support	Clinical examination	legs	distal muscle wasting	Deep tendon reflexes were absent in both extremities. Vibration detection, touch and pinprick sensation was reduced in upper and lower distal limbs. She had no cerebral, cerebellar and pyramidal signs and symptoms. She did not have any signs or symptoms of liver and renal diseases.
					toes	clawing	
				Manual muscle strength testing	deltoid, biceps brachii, triceps brachii, iliopsoas and quadriceps femoris bilaterally	grade 5	
					extensors and flexors of wrists, hamstrings and gluteus maximus bilaterally	grade 3	
					extensors and flexors of muscles of fingers, bilaterally	grade 2	
					tibialis anterior, gastrocnemius, peroneus, and extensors and flexors of toes bilaterally	grade 0	
				Electrophysiological examination	right and left median nerves	mildly delayed motor NCV 45.3 and 49.6 m/s, respectively (Table S4)	

Table S4. Nerve conduction study of the proband of family 2.

Nerve	Motor		
	DL (ms)	CV (m/s)	CMAP (mV)
Affected member V-1			
Rt median	3.7 (N<4.0)	45.3 (N>56)	5.0 (N>3.5)
Lt median	3.7	49.6	4.9
Rt ulnar	3.3 (N<3.2)	51.5 (N>50)	4.3 (N>2.8)
Lt ulnar	3.5	41.8	0.8
Rt tibial			NE
Lt tibial			NE

DL: distal latency, CV: conduction velocity, CMAP: compound muscle action potential amplitude, Rt: right, Lt: left, N: normal, NE: not evoked

Table S5. Summary of NGS performance.

Summary	
Reads mapped to genome[a]	1,024,056,995
Data mapped to genome (Gb)	141.8
Average sequenced depth (×)	49.6
Genome covered (%)	96.2

Gb: giga bases

[a]A total of 5 μg genomic DNA sample was subjected to enzymatic fragmentation using the NEB Next dsDNA Fragmentase kit (New England Biolabs, USA) according to the manufacturer's instructions. The resulting size distribution was between 500–700 bp (600 bp peak top) on the Agilent 2100 Bioanalyzer (Agilent Technologies, USA). Then, library preparation was performed using Paired-End DNA Sample Preparation Kit (Illumina, USA). After the check of size distribution and concentration by the Agilent Bioanalyzer 2100, total 5 times of cluster generation were performed from same library using a Paired-End Cluster Generation Kit (v4) on Cluster Station and four cBot Paired-End Cluster Generation Kit (v4) on cBot according to the manufacturer's instructions. 101 bp paired-end runs were performed, using TruSeq SBS kits (v5). Basecalling, mapping to a reference human genome (UCSC NCBI37/hg19) and variant detection were performed using the standard Illumina Pipeline (CASAVA 1.7).

Table S6. Gap[a] filling by Sanger-based PCR direct sequencing.

Start	End	Length	Gene Symbols	Exon	PCR Primers	
					Forward	Reverse
118498754	118498900	147	WSB2	exon 1 5' UTR	CCCGATCTTCCCGATCCTGTCG	GTACCGCCTCCCCGATTTTCCATC
120907136	120907142	7	SRSF9	intron	GAGGCGGACAAGGCTCATTTGGAC	CATCCGCGAGATCGAGCTCAAGAAC
121078583	121078598	16	CABP1	exon 1 coding	GTAGAGGCTGCGCTGTCACATGG	CGGCGGCTGGAACGGAGTC
121078674	121078697	24	CABP1	exon 1 coding	GTAGAGGCTGCGCTGTCACATGG	CGGCGGCTGGAACGGAGTC
121078822	121078910	89	CABP1	exon 1 coding	GTAGAGGCTGCGCTGTCACATGG	CGGCGGCTGGAACGGAGTC
121148179	121148186	8	UNC119B	exon 1 5' UTR	CCTCAGGTTGCTGATGGCAGCTCTAG	CCACCAGAGCTACGTGAGCGTGG
121268251	121268289	39	SPPL3	intron	GAGTAAATTTCACTGTTCACGAACTACACC	CTCCCCAACATATGCATGTACACC
121734510	121734607	98	CAMKK2	intron	CTCCGGGCTTTGTGTGCCATG	CCTCTTCGCTACTTTCGTCTTTGCAGG
121904681	121904758	78	KDM2B	intron	GAGGGTCCTGGCACAACTTTGCG	GAGCGCCGAGGACGACGACTATG
121905015	121905015	1	KDM2B	exon 1 5' UTR	AGCGGCAGCAAAACTTTCTCCTCAT	CCTTCCTGGCTCGCCTCATGC
121905364	121905381	18	KDM2B	intron	ATGCATATGCATGAGGCGAGCCAG	GGAGACTGGGCTCTCAGAGCTTATCCAG
121975839	121975878	40	KDM2B	intron	TTTTCCCTCCAACTTGGCGTAACTCC	GTTCAGTTTCAGCGCGGGCATG

[a]All gaps found on the linkage region were filled by Sanger-based PCR direct sequencing.

Table S7. Summary of variants.

Variant categories[a]	counts
Total number of variants	3,716,455
Frame-shift indel	154
In-frame indel	210
Stop gain	92
Stop loss	15
Nonsynonymous	10,556
Splice site	135
5-UTR	27,123
3-UTR	5,223
Synonymous	11,788
Intron	1,267,264
Intergenic	2,266,228
ncRNA	127,667
Homozygous	1,669,570
Heterozygous	2,046,885

[a]We applied ANNOVAR (version May 8, 2013) to appropriately annotate the genetic variants [2]. This included functional annotation by 'refGene' and 'knownGene' annotation databases, dbSNP (build 135 and 137) identifiers, and 1000 Genomes Project (Apr. 2012) allele frequencies. After the annotation process, functional (Frame-shift indel, In-frame indel, Stop-gain SNV, Stop-loss SNV, non-synonymous SNV, and Splice site variant), novel (Not reported in dbSNP 135 and 137) and homozygous variants were included for further analysis.

Table S8. List of homozygous and potentially deleterious variants voted by at least one prediction methods.

Chr	Start	Alleles		Gene Symbols	Types	Accession	Changes		Functional Predictions					Expression in nervous system
		Reference	Variant				Nucleotide	Protein	PolyPhen-2 (HumVar)	Grantham	PROVEAN	SIFT	Mutation Taster[a] (probability)	
12	120878247	CACTC	-	*COX6A1*	splicing	NM_004373.3	c.247-10_247-6del	NA	splicing	splicing	splicing	splicing	splicing	+
2	114257705	C	-	*FOXD4L1*	frameshift	NM_012184.4	c.876del	p.Gly293Alafs	indel	indel	indel	indel	DC (1)	-
14	24470691	-	A	*DHRS4L2*	frameshift	NM_001193637.1	c.204_205insA	p.His69Thrfs	indel	indel	indel	indel	DC (1)	-
19	20807178	-	A	*ZNF626*	frameshift	NM_001076675.2	c.1505dup	p.Ile503Hisfs	indel	indel	indel	indel	DC (0.999)	NA
2	114257443	A	C	*FOXD4L1*	nonsynonymous	NM_012184.4	c.610A>C	p.Lys204Gln	0.062	53	-3.933	0.021	DC (0.999)	-
8	126443464	G	T	*TRIB1*	nonsynonymous	NM_025195.3	c.320G>T	p.Arg107Leu	0.392	102	-4.452	0.002	DC (0.999)	-
2	73151606	G	A	*EMX1*	nonsynonymous	NM_004097.2	c.689G>A	p.Ser230Asn	0.023	46	-2.144	0.341	DC (0.999)	-
2	97637964	G	A	*FAM178B*	nonsynonymous	NM_001122646.2	c.238C>T	p.Arg80Trp	NA	101	-3.127	0	P (0.999)	NA
4	69403401	C	A	*UGT2B17*	nonsynonymous	NM_001077.3	c.1535G>T	p.Cys512Phe	NA	205	-7.344	0.107	P (0.999)	-
9	133556937	G	C	*PRDM12*	nonsynonymous	NM_021619.2	c.985G>C	p.Ala329Pro	0.113	27	-0.66	0.041	DC (0.973)	+

[a]Prediction by Mutation Taster: DC: Disease Causing; P: Polymorphism

Table S9. Autozygosity in the linkage region on 12q24.

	Position			Allelic state	
	start	end	bp	homozygous	heterozygous
5' flanking region	117,539,934	117,882,412	342,479	281	8
Autozygous region	117,882,413	121,605,073	3,722,661	2,891	0
3' flanking region	121,605,074	121,878,659	273,586	110	382

Table S10. PCR primer sequences for validation of variants by Sanger-based PCR direct sequencing.

Chr	Start	Alleles		Gene Symbols	Accession	Changes		PCR Primers		Related Figures and Tables
		Reference	Variant			Nucleotide	Protein	Forward	Reverse	
12	120878247	CACTC	-	*COX6A1*	NM_004373.3	c.247-10_247-6del	NA	TGGAATTAAAACTGTCCTGTAGC	GGTCCATGTGCAGAGTAAC	Figures 2E & S3, Table S11
2	114257705	C	-	*FOXD4L1*	NM_012184.4	c.876del	p.Gly293Alafs	AAGACGAGGTGGAAGACGAG	ATTGTCCGACAGGCTTGAC	Figure S2
14	24470691	-	A	*DHRS4L2*	NM_001193637.1	c.204_205insA	p.His69Thrfs	GAGCACTGCCCTCTATGTCTAG	GCACAATGTGCACATGTACC	Figure S2
19	20807178	-	A	*ZNF626*	NM_001076675.2	c.1505dup	p.Ile503Hisfs	TCATATCAATTCTTAGTTAGAAATTGAGG	AAGCCTTCAAGCGGTCTTC	Figure S2
2	114257443	A	C	*FOXD4L1*	NM_012184.4	c.610A>C	p.Lys204Gln	AGACGAGGTGGAAGACGAG	CGAGAGCAGTAGGTAGCGAG	Figure S2
8	126443464	G	T	*TRIB1*	NM_025195.3	c.320G>T	p.Arg107Leu	AGACCAGTCTGCAAACTCCA	ATCTACTGATCCGCCCTGTG	Figure S2
2	73151606	G	A	*EMX1*	NM_004097.2	c.689G>A	p.Ser230Asn	GCGTGTCAAGGAATGGAGAG	CTTCCTCCAGGGAACCTG	Figure S2
2	97637964	G	A	*FAM178B*	NM_001122646.2	c.238C>T	p.Arg80Trp	GATGGGTAGCCAAGGAGAAG	GTTCTCTCACCCATTTGCTG	Figure S2
4	69403401	C	A	*UGT2B17*	NM_001077.3	c.1535G>T	p.Cys512Phe	GCTCAGTAACTTTTGTGTGGG	CTGGATCGAGCAGTCTTCTG	Figure S2
9	133556937	G	C	*PRDM12*	NM_021619.2	c.985G>C	p.Ala329Pro	CTACAAGTGCCAGGTGTGCCAGAG	CTCTACAGCAGCAGGGAGACTCGG	Figure S2

Table S11. Screening of 5-bp deletion in *COX6A1* among diverse ethnic populations.

Populations	Number of individuals[a]	PCR product size[b]	
		377 bp (wild type)	372 bp (5-bp deletion)
Caucasian	100	100	0
African American	100	100	0
Amerindian	9	9	0
Cambodian	1	1	0
Druze	1	1	0
Melanesian	2	2	0
Pygmy	5	5	0
Japanese	508	508	0
Total	726	726	0

[a]Genomic DNAs of 218 unrelated individuals from diverse ethnic populations were provided by Coriell Cell Repositories. In addition, genomic DNAs of 508 unrelated Japanese individuals were obtained [3].

[b]After amplification (see also Table S10) by standard PCR in 50 μL reaction volume containing 4 ng of genomic DNA, the PCR fragments include the 5-bp deletion (372 bp) or wild-type (377 bp) were produced. They are distinguishable by standard 1.5% agarose gel electrophoresis.

Table S12. Sequences of PCR primers and TaqMan probe used in the expression analysis of *COX6A1*.

Assay type	Target Molecule	Assay Name	PCR Primers		TaqMan Probe	Related Figures
			Forward	Reverse		
TaqMan	Wild-type	-	CACGAGAGACCCGAGTTCATC	TTTCTGGTCCATGTGCAGAGTAAC	(MGB)ACCAAGCCGTTTCC	2A, 2B
SYBR Green I	Wild-type	-	CATCAGGACCAAGCCGTTTC	TTTCTGGTCCATGTGCAGAGTAAC	-	S5
SYBR Green I	both	Primer 1	GGCGGTAGTTGGTGTGTCCTC	CTTGGTCCTGATGCGGAGATG	-	S6
SYBR Green I	Mutant-type	Primer 2	TTCTCTTTTCGTTATGTGTGCCTTA	GCAGGAGGTAGGATGAATCGG	-	S6
SYBR Green I	Mutant-type	Primer 3	GTGAGCCAAGATCCAGCCAT	TGCCGCTCAAAGACTCAACA	-	S6
SYBR Green I	Mutant-type	Primer 4	TTGTTTGAAGTCTGGGATGGTG	CTGGCTACAGGACAGTTTTAATTCC	-	S6

Table S13. PCR primer sequences for genotyping of *Cox6a1* null mice.

Target	PCR Primers		Amplicon size (bp)
	Forward	Reverse	
exon 1	CTAGGCGCTGACAAGCAG	CGCCATCATGGAACTACAC	219
exon 2	GTGAGCATGCTCAACGTG	CTATCTGCACGGACGGAC	227
exon 3	CCATCCCTTCTGCTTGTAAG	TGCTAAGCACAAGAAACCAG	189

Table S14. Additionally genotyped single nucleotide variants around the 5-bp deletion.

Variants	Positions	Distance from 5-bp del (kb)	PCR Primers		Criteria[a]	dbSNP MAF	Genotypes	
			Forward	Reverse			1:V-1	2:V-1
rs2422	**120565188**	**-313.1**	**-**	**-**	**ILP**	**0.495**	**A**	**G**
Variant 1	120578265	-300.0	CTAGGAATTCAGACAAGGATGC	TTGCAGGCTAAGAGCTAAGTG	NGS	-	A	C
rs11065050	120641900	-236.3	GGATGGAAAATAACATGAATGAG	TGTCTTGGTTTTAAACCTGCTAC	HM3,MAF	0.184	C	A
rs1634814	120685747	-192.5	CTGCAGATGTTTCAAGCAAC	CCTGCTCTTCACCTCTCTTG	HM3	0.371	T	C
Variant 2	120724972	-153.3	TGGTAGTGGGTCTCATCTTTG	TACTACAAGGGCTATCTCAGCAG	NGS	-	G	A
rs12313457	120869349	-8.9	AATTTCTTTACAGCAATGCAAG	AAGGTGGTATATTGGAATGTATTG	1KGP,MAF	0.128	C	G
rs73420911	120872479	-5.8	CTGGTGGTATCTTTTGGAAAG	TGTTAGAATTGTCTTGATCTAGTCC	1KGP,MAF	0.128	G	A
rs11612825	120874200	-4.0	CTTCATGAAACTGCATGACC	TATGCCTCAGATCAAGAGCTC	1KGP,MAF	0.127	T	C
rs145491305	120874512	-3.7	ATGGGAGTTCAACTCTTGTTG	TTCCTCTAAAGACACACCTATGTG	1KGP,MAF	0.128	A	C
rs73420919	120874620	-3.6	ATGGTGCAATCTCAGCTCAC	CCCTTTGAGAATGTAATGAGGAC	1KGP,MAF	0.130	A	G
rs11609119	120877821	-0.4	CATCACCTAGCCCATATAGGAC	CAAGCTCAAGTTAGTTACTCATCATC	1KGP,MAF	0.144	A	T
5-bp del	**120878247**	**0.0**	**-**	**-**	**NGS**	**-**	**+**	**+**
rs7134710	120888510	10.3	TGACTGTAGATTTGTGAAAGTGG	GTCAGGGCTTTTATAGCTAAGATG	1KGP,MAF	0.149	A	C
rs7027	120901266	23.0	CTGCTTTCCTACAAGGAAAGAC	CTAGTGCAGTCGGTCTTCATAC	HM3,MAF	0.177	T	C
rs7964008	120918397	40.2	ACCTCAATCTGCATCCAGAG	TGGAACAAAAGGCATAATGG	HM3,MAF	0.150	T	C
rs11065140	120929541	51.3	GGGGTCTTCAATCAACTCTG	AATGTCGCTGGAGACACTAAG	HM3,MAF	0.159	A	G
rs719450	120954033	75.8	CTTTACAGGAGATGAGAGGTCC	AGACACAGGGGAAGTATTTGTATC	HM3,MAF	0.207	T	C
rs3742049	120954490	76.2	TTCAACCTGCCCATCTATAAG	ATCCAGTTAGCTAACACAGGTACTC	HM3,MAF	0.146	T	C
rs4102203	120976371	98.1	GCCATTTGCTTCCTAGCTAC	TACCAGAATAACAGGAAGAATGC	HM3,MAF	0.074	C	T
rs16950293	121015461	137.2	GGAAAGCTGTAACTCACGAAG	TAACAGAAACACTGGTGATTCTG	HM3,MAF	0.158	A	G
Variant 3	121177337	299.1	CATGGGCTACGTGACAGAG	ACCTCAAGGAAAAGACAGACC	NGS	-	G	C
rs694303	121268624	390.4	GGTGAGGAATTCAGTTCTAAGC	GACCGAAGTTGAATGTTTCAC	HM3,MAF	0.201	T	C
rs625228	121278266	400.0	TTTGGGCCATAAAAAGGAG	CATCTGGCCACAATTCTG	HM3	0.469	A	G
rs589699	121288685	410.4	TCCCTTTTTAAATGTGAGTAACAT(	ATTCATTAGGAATGATAACACCTG	HM3,MAF	0.204	G	A
rs610694	**121304826**	**426.6**	**-**	**-**	**ILP**	**0.494**	**G**	**A**

[a]Variants that satisfy the following criteria were selected for the additional genotyping.
ILP: SNP markers nearest the 5-bp deletion in Illumina's Linkage Panel.
NGS: Single nucleotide variants newly found by our Next Generation Sequencing analysis.
HM3: SNP markers tagging each haplotype in probands of family 1 and 2 from HapMap 3 release 2.
1KGP: SNP markers tagging each haplotype in probands of family 1 and 2 from the 1000 Genomes Project (phase 1 version 3).
MAF: SNP markers with MAF $\lesssim$ 0.2 in dbSNP 138.

Supplemental References

1. Skarnes, W. C., Rosen, B., West, A. P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A. O., Thomas, M., Harrow, J., Cox, T., et al. (2011). A conditional knockout resource for the genome-wide study of mouse gene function. Nature *474*, 337-342.
2. Wang, K., Li, M., Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.
3. Yamagata University Genomic Cohort Consortium. (2014). Pleiotropic Effect of Common Variants at ABO Glycosyltranferase Locus in 9q32 on Plasma Levels of Pancreatic Lipase and Angiotensin Converting Enzyme. PLoS One *9*, e55903.