Supplementary Methods

Whole exome sequencing

For the two individuals in this family, sequence reads were converted from Illumina fastq format to fastq files that were compatible with the Sanger sequencing specification. Sequencing reads were aligned with the human reference sequence (hg19) using Burroughs Wheeler Alignment (BWA) version 0.7.17 and default parameters [1]. Alignments were converted from SAM format to BAM files with SamTools [2]. Duplicate reads were removed from BAM files and a base quality score was recalibrated using GATK 4.1.2.0 [3]. Genotype calling was executed with the GATK 4.1.2.0 Haplotype Caller. Filtering was performed to remove variants with QD (variant confidence/quality by depth) <2.0, FS (phased-scaled p-value using Fisher's exact test to detect strand bias) >60.0, MQ (RMS mapping quality) < 40.0, MQRankSum (Z-score from Wilcoxon rank sum test Alt vs. Ref read mapping qualities <−12.5, and ReadPosRankSum <−8.0, SOR > 4.0).

Output files from GATK were converted to VCF format. All sets of single nucleotide variant (SNV) and indel variant calls from GATK were annotated using the ANNOVAR 2019Mar20 program to identify exonic or splicing variants with their allele frequencies and functional information [4]. Locations and genotypes of variants for each individual were identified to localize subsets of variants on autosomal chromosomes. Annotations were assigned to all variants using the ANNOVAR 2019Mar20 program on the basis of their frequency in the National Heart, Lung and Blood Institute Exome Sequencing Project (ESP6500), the 1000 Genomes Project, The Genome Aggregation Database (gnomAD), and 2 Japanese databases, which are 4.7KJPN in jMorp by Tohoku Medical Megabank Organization (ToMMo) and Human Genetic Variation Database (HGVD) [6-10]. Impacts of variants on coding features were examined against the following three databases: ensGene tracks from Ensembl Genes; knownGene tracks from the University of California, Santa Cruz Known Genes; and refGene tracks from the National Center for Biotechnology Information Reference Sequence Database [4, 5]. The candidate gene variants were presumed to be a potential heterozygous mutation because of autosomal dominant inheritance in the family. Candidate variants were selected if they met the following criteria: 1) minor allele frequency of <0.1% in the total population in 1000 Genomes Project, ESP6500, gnomAD, 4.5KJPN and HGVD; 2) PASS in variant filtering; 3) either nonsynonymous, splice site, insertion/deletion, stop gain/loss, or unknown variants in at least one database; and 4) an exonic or a splice site in at least one database. American College of Medical Genetics and Genomics (ACMG) standards and guidelines were used to determine the pathogenicity of variants [6]. Copy number variations for each sample were analyzed using the CLC Genomics Workbench version 20.0 (Arhus, Denmark) with the default parameters.

Confirmation of the patient's variant using Sanger sequencing

Validation of the patient's variant (ch20:39316901G>T) was performed using genomic polymerase chain reaction (PCR) and Sanger sequencing. PCR was conducted using PrimeSTAR MAX DNA polymerase (Takara, Tokyo, Japan) in accordance with the manufacturer's instructions. The primers for genomic PCR were as follows: forward primer, 5'-CGTCCCCAGACAAAGGCTTG-3'; reverse primer, 5'-TCACCTCGTCCTTGGTGAAG-3' [7]. For Sanger sequencing, BigDye Terminator Ready Reaction Mix (v3.1) (Thermo Fisher Scientific, Waltham, MA, USA) was used and samples were purified using the BigDye XTermnator purification kit (Thermo Fisher Scientific).

Supplementary References

1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009 Jul 15;25(14):1754-60.
2. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078-9.
3. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010 Sep;20(9):1297-303.
4. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010 Sep;38(16):e164.
5. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012 Nov 1;491(7422):56-65.
6. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015 May;17(5):405-24.
7. Sato Y, Tsukaguchi H, Morita H, Higasa K, Tran MTN, Hamada M, et al. A mutation in transcription factor MAFB causes Focal Segmental Glomerulosclerosis with Duane Retraction Syndrome. Kidney Int. 2018 Aug;94(2):396-407.

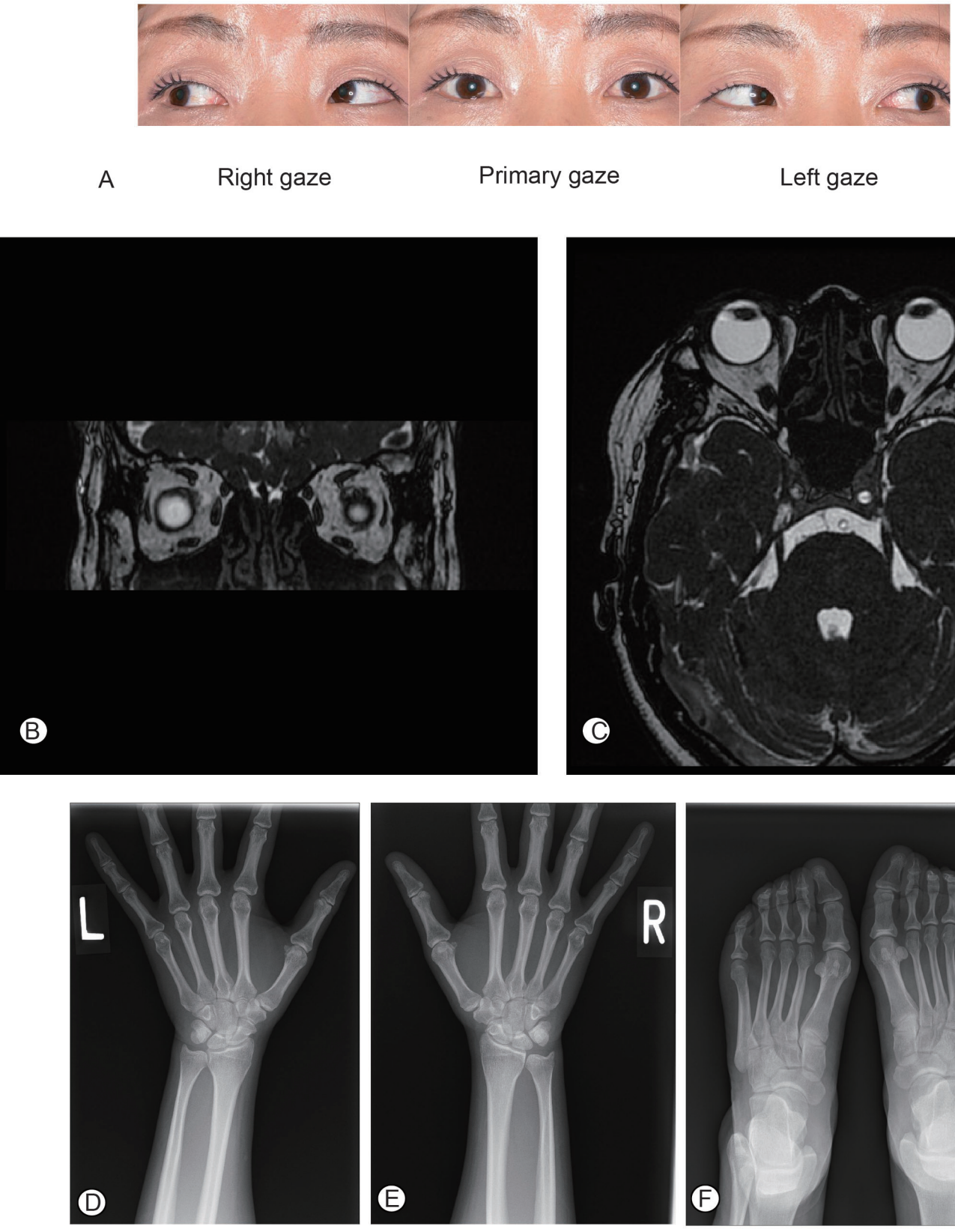

Figure S1: Extrarenal manifestations in patient II-1. (A) Eye movement in Patient II-1.

(B) Coronal T1-weighted magnetic resonance image of extraocular muscles.

(C) Axial T2-weighted magnetic resonance image of both abducens nerves.

(D–F) X-ray images of upper and lower extremities.

Table S1. Whole exome sequence data summary

	raw data (bp)	trimmed data (bp)	%mapped reads (paired)	called variant number per sample	rare variant number (MAF<0.1%) per sample	II-1 unique rare variant number (MAF<0.1)
II-1	7,827,949,200	7,824,272,402	99.09%	199,585	314	144
II-2	9,898,802,200	9,894,297,627	99.10%	211,342	314	

Table S2. The 117 candidate genes with 144 rare variants (MAF<0.1%) unique to patient II-1

Gene Name	Gene Name	Gene Name	Gene Name	Gene Name
PLEKHN1	C3orf37	MTUS1	NXPH4	P2RX1
KIAA1751	MUC4	FGFR1	PIP4K2C	MPRIP
ACOT7	CENPC1	PABPC1	METAP2	TMEM97
ZBTB48	TXNDC15	GRHL2	POLR3B	CEP112
EXOSC10	UIMC1	ARHGAP39	OAS2	EMILIN2
ESPNP	MAML1	ZBTB5	SKA3	WDR18
PDE4DIP	ATXN1	DKFZp434P0216	NUPL1	APC2
CGN	HLA-H	BRD3	NHLRC3	AX748210
FLG2	MUC21	PITRM1	LPAR6	STXBP2
PAQR6	GPR116	SLC39A12	RCBTB1	MUC16
IQGAP3	MCM3	PARD3	ATP7B	ZNF788
BCAN	SYNE1	LIPF	PCCA	HSH2D
CD1E	HGC6.3	CEP55	LINC00346	ZNF850
GUK1	C7orf26	ADRB1	LRR1	MGC2752
GCKR;C2orf16	NFE2L3	CPXM2	NOXRED1	SLC23A2
CCDC148	BBS9	TRIM22	SNW1	SNAP25
SCN9A	ABCA13	PARVA	abParts	FRG1B
CCDC141	ZNF273	OTOG	ELL3	MAFB
ALS2	POM121	PHF21A	WDR76	SLPI
GLB1L	CACNA2D1	TBC1D10C	ZFAND6	APCDD1L
ASIC4	AKR1B1	GDPD4	RPS15A	CRYAA
ALPP	LUZP6	MRE11A	FBRS	
DYNC1LI1	TCRBV12S2	PRMT8	SLC5A2	
UBA3	TCRBV20S1A1N2	LARP4	C16orf58	

Table S3. Name list of genes with different copy number between II-1 and II-2.

Gene Name	Gene Name	Gene Name
HERC2P4	IGHV3OR16-15	FCGR2B
RP11-56L13.7	IGHV3OR16-6	FCGR2C
TP53TG3D	RP11-19N8.2	ZNF705C
RP11-17M15.4	IGHV3OR16-10	FAM66A
ACTR3BP3	IGHV3OR16-8	RP11-351I21.7
RP11-626K17.5	NBPF8	RP11-351I21.6
ABCD1P3	AL592284.1	ZDHHC11
TP53TG3	PPIAL4B	RP4-576H24.4
RP11-586K12.8	RP11-640M9.2	SIRPB1
HERC2P5	CROCC	CCL4L1
IGHV2OR16-5	HSPA6	TBC1D3C
SLC6A10P	RP11-25K21.6	TBC1D3H
RP11-989E6.3	FCGR3A	CCL3L1