The American Journal of Human Genetics, Volume *99*

Supplemental Data

ARCN1 Mutations Cause a Recognizable Craniofacial Syndrome Due to COPI-Mediated Transport Defects

Kosuke Izumi, Maggie Brett, Eriko Nishi, Séverine Drunat, Ee-Shien Tan, Katsunori Fujiki, Sophie Lebon, Breana Cham, Koji Masuda, Michiko Arakawa, Adeline Jacquinet, Yusuke Yamazumi, Shu-Ting Chen, Alain Verloes, Yuki Okada, Yuki Katou, Tomohiko Nakamura, Tetsu Akiyama, Pierre Gressens, Roger Foo, Sandrine Passemard, Ene-Choo Tan, Vincent El Ghouzzi, and Katsuhiko Shirahige

Supplemental Information

Supplemental Note: Case Reports

Subject 1 was referred to the Genetics Clinic of Nagano Children's Hospital at the age of 1 month with the chief complaints of facial dysmorphism and intrauterine growth retardation. Severe intrauterine growth retardation was noted during pregnancy. Subject 1 was born at 35 weeks gestation and was delivered by normal vaginal delivery. Her birth weight was 1100 gm (<3rd centile: -4.18SD), length was 34 cm (<3rd centile: -3.72SD), and head circumference was 26 cm (<3rd centile: -3.15SD). At birth, severe micrognathia and facial dysmorphism were noted. An echocardiogram revealed the presence of a ventricular septal defect. After birth, Subject 1 developed respiratory distress, requiring ventilator support. Severe micrognathia was considered to be the reason of respiratory distress. Therefore, when she was 6 months old, a tracheostomy was performed. The subsequent development of Subject 1 has been slightly delayed; she started standing with holding at 10 months and sitting independently at 11 months. At 1 year and 10 months, her developmental quotient (DQ) was 65. At the age of 3 years and 1 month, DQ for motor skills, cognition/adaptation, and language/social skills

were 75, 77, and 69, respectively. The overall DQ was 72. At the age of 3 years and 4 months, she recently met the developmental milestones of approximately 2 years and 6 months, since she was able to jump up, draw straight lines, and take off her clothes.

Since birth, she has continued to have short stature and has exhibited failure to thrive (Figure S1A). On physical examination, several facial dysmorphisms, including a prominent forehead, downslanted palpebral fissures, and severe micrognathia, were noted. In addition, shortening of upper arms and legs, and small joint laxity were observed (Figure 1A). Beighton score was 5/9. No skin color or hair abnormalities were identified. She has not shown any signs of cerebellar ataxia. A skeletal survey revealed widening of the metaphysis and a wide femoral neck (Figure 1B). Ophthalmological evaluation revealed astigmatism. Audiology evaluation was unremarkable. Brain MRI was also unremarkable. Currently, she is 3 years and 4 months old, and her height was 80.2cm (-4.1SD), weight was 8.8kg (-3SD), and head circumference was 43.4cm (-3.2SD). The family history of Subject 1 was unremarkable.

Subject 2, a boy with developmental delay and mild autism, was first seen in the Genetics Clinic, KK Women’s & Children’s Hospital, at the age of 1 year and 8

months. Antenatal scans showed intrauterine growth retardation, micrognathia, and possible ambiguous genitalia. Amniocentesis revealed a normal karyotype. Subject 2 was delivered at 34 weeks gestation via cesarean section and had a birth weight of 1360 gm (<3rd centile), birth length of 40cm (<3rd centile), and birth head circumference of 29cm (10th centile). Postnatally, it was observed that he had severe micrognathia and penoscrotal hypospadias and required tracheostomy for respiratory support during infancy (Figure S1B). He also developed severe gastroesophageal reflux, requiring fundoplication and gastrostomy insertion for the first 3 years of life. The development of Subject 2 was slightly delayed; he started walking independently at 13 months and climbing stairs at 18 months. He spoke his first words at 17 months. At 7 years of age, he has moderate-to-severe language difficulties and currently attends special education school. His medical issues include epilepsy and obstructive sleep apnea requiring continuous positive airway pressure support at night. An MRI scan of the brain of Subject 2 was normal and an electroencephalogram confirmed epileptiform activity. On physical examination, it was observed that Subject 2 had micrognathia, scaphocephaly, and hypotelorism, as well as small joint laxity (Figure 1C). Beighton score was 2/7. No

skin color or hair abnormalities identified. He has not shown any signs of cerebellar ataxia. The growth parameters of Subject 2 improved postnatally, and as of 7 years, his height is 121 cm, which is in the 10th centile. His mid-parental height is 172.5 cm, which is 50th centile. Currently he is 10 year and 6 months old, and his current height is 136cm (10th percentile) and weight 33.2kg (25th percentile), head circumference was 51cm (6th centile (-1.53SD)). His ophthalmological issues include occasional divergent squint and myopia. He does not have hearing problems.

Subject 3 (25 years old) was born at 36 weeks gestation from consanguineous parents (cousins): weight 1.450kg (<3rd centile), length 38cm (<3rd centile), head circumference 27.6 cm (<3rd centile). Hypotelorism, micrognathia, short bowed legs, and high arched palate were noticed at birth. Bilateral posterior cataract was diagnosed when he was 6 years old. He had hypertrophic muscular appearance, rhizomelic shortness of upper and lower limbs, large hands and feet, micropenis, and hypoplastic scrotum. A skeletal survey at 7 years of age demonstrated short femoral and humeral diaphysis, thick femoral neck with coxa valga, and advanced wrist bone maturation. The Subject 3's IQ was rated at 65 based on the Wechsler Intelligence

Scale for Children (WISC). At the age of 25 years, his weight was 98kg, his length was 152cm (-3.5 SD), and his head circumference was 50 cm (-5 SD) (Figure 1D). He cannot read or write but can be understood. He is an employee in his father's company and has married a woman from another region.

Subject 4, the daughter of Subject 3, shared common features with Subject 3, which included intrauterine growth retardation followed by postnatal growth failure. Currently, she is 3 years old, and her height was 79.5 cm (-4 SD), her weight was 9.650kg (-3SD) and head circumference was 45 cm (-5 SD). She was also found to have dysmorphic features including microretrognathia, hypotelorism, shortening of upper arms and legs, and muscular hypertrophy (Figure 1E). Subject 4 was born at term: weight 1.400kg (<3rd centile), length 37cm (<3rd centile), head circumference 27.5cm (<3rd centile). She had a cleft palate, and feeding difficulties that required nutritional support until the age of 2 years. She started walking at 2 years of age with ataxia, and she does not speak yet. Her brain MRI showed microcephaly, delayed white mater myelination in T2 weight images, and mild cerebellar atrophy (Figure S1C).

Exome sequencing was performed for Subject 1 as well as her unaffected

biological parents. The G-band karyotype and SNP array results were normal. Exome capture was performed with Agilent SureSelect XT Human All Exon V5 (Agilent, Santa Clara, Calif., USA), and 100-bp or 126-bp paired-end read sequencing was performed by HiSeq 2500 (Illumina, San Diego, Calif., USA). The sequenced reads were aligned to the human genome hg19 using BWA v0.7.12 with BWA-MEM algorithm and '-M' option[1]. The Genome Analysis Toolkit (GATK) v3.3.0 was used for variant discovery[2]. After marking duplicates, local realignment around indels, and base quality score recalibration, SNVs and indels for each sample were called using the HaplotypeCaller. SNVs and indels with a GQ score above 15 were used for further analysis. The effects of variants were predicted by SnpEff v3.6b[3] and annotated by ANNOVAR[4]. Information and statistics regarding sequenced reads, mapping rate, and coverage are summarized in Table S1. Six *de novo* mutations in 4 genes (*ARCN1*, *DCAF15*, *HLA-DRB5*, *FAM58A*) were identified. However, variants in *DCAF15*, *HLA-DRB5,* and *FAM58A* were present in dbSNP, leaving *ARCN1* as the prime candidate gene to account for the subject's phenotype. The mutation in *ARCN1* resulted in a premature stop codon in exon 2 (NM_001655:exon2:c. 260C>A:p.Ser87*). The *ARCN1* mutation was confirmed by

Sanger sequencing (Figure 2A and Figure S2A). Sanger sequencing demonstrated that the *DCAF15* variant was inherited from her unaffected father and the *FAM58A* variants were inherited from her unaffected parents (Figure S3A and S3B).

Subject 2 was independently identified in a cohort of individuals from Singapore with undiagnosed genetic diseases. He had been previously tested and his karyotype and chromosome microarray analysis results were normal. Exome sequencing of Subject 2 and his unaffected parents was carried out by the Genome Institute of Singapore using the NimbleGen SeqCap EZ Human exome V3 kit (Roche, Basel Switzerland) and sequenced on the HiSeq 2000 platform (Illumina, San Diego, USA) using 100-bp paired-end reads. Alignment to the human genome hg19 was carried out using BWA 0.7.5a and the variants were annotated using ANNOVAR. Variants with minor allele frequency (MAF) >0.01 in dbSNP or EVS (Exome Variant Server) were excluded and rare functional variants were prioritized according to *de novo* status, homozygosity, compound heterozygosity and X chromosome-linkage. After filtration, *de novo* heterozygous variants were identified in *ARCN1* and *SYT1*. Both of these variants were confirmed by Sanger sequencing. The *ARCN1* mutation was a

frameshift variant: NM_001655:c.633del; p.Val212Trpfs*15 predicted to create a frameshift in exon 4 starting at codon Val212 and ending in a stop codon 14 positions downstream (Figure 2A and Figure S2B). The *SYT1* (NM_001135805.1:c.697G>A; p.Asp233Asn) variant is a missense variant that substitutes Aspartic acid at codon 233 with Asparagine (Figure S3C). There were no other potentially pathogenic mutations.

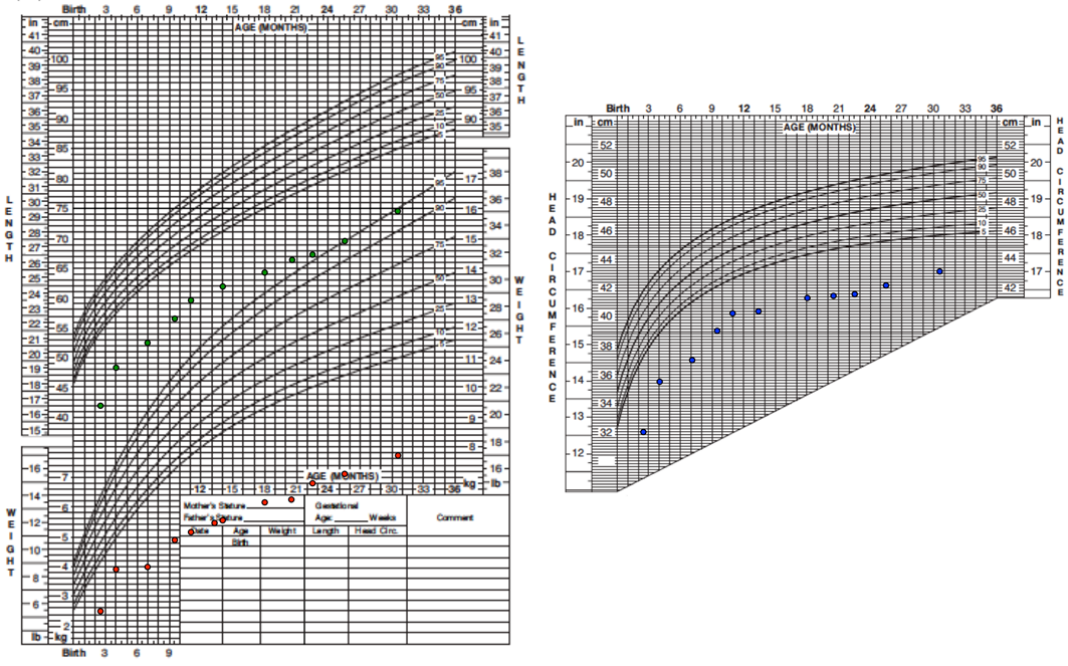
Subjects 3 and 4 were also ascertained independently. Whole-exome sequencing was performed for Subject 3 and his healthy parents by IntegraGen Genomics. Exons of DNA samples were captured using an in-solution enrichment methodology (Agilent SureSelect All Exon V4 + UTRs) and sequenced with an Illumina HiSeq 2000 instrument. SNPs and indels were determined using the CASAVA 1.8 software. Identification of the disease-causing variants was performed with the Exome Resequencing Intelligent Sorter (ERIS, Integragen) program for the annotation and filtering of genetic variants. The following criteria were used for excluding non-pathogenic variants: (1) variants represented with an allele frequency of more than 0.5% in dbSNP 138, the NHLBI Exome Sequencing Project EVS, or the in-house database of Integragen; (2) variants in the 5’ or 3’ UTR; (3) variants with intronic

locations and no prediction of local splice effect; and (4) synonymous variants without prediction of local splice effect. Filtered variants occuring *de novo* were retained for further confirmation and segregation analysis by Sanger sequencing. The *ARCN1* mutation was the only probably damaging mutation that occurred *de novo* in Subject 3 and was transmitted to his affected daughter, Subject 4. The *ARCN1* mutation was NM_001655:c.157_158del, leading to p.Ser53Cysfs*39 (Figure S2C).

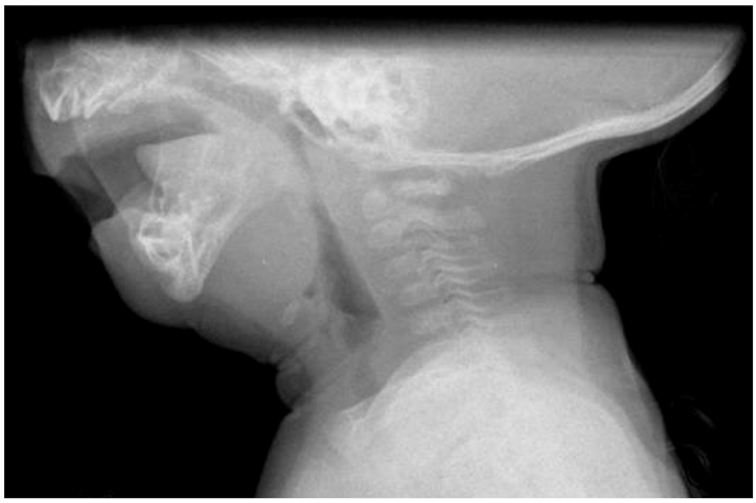
Supplementary Figures

Figure S1. Additional clinical information of the individuals with *ARCN1* mutations. (A) Growth chart of Subject 1. Green dots represent height measurements, red dots represent weight measurements, and blue dots represent head circumference measurements. (B) Radiograph of Subject 2 demonstrating the presence of micrognathia. (C) Brain MRI of Subject 4. Microcephaly with gyral simplification, thin corpus callosum, delayed myelination and mild cerebellar atrophy are demonstrated.

(A)



(B)



(C)

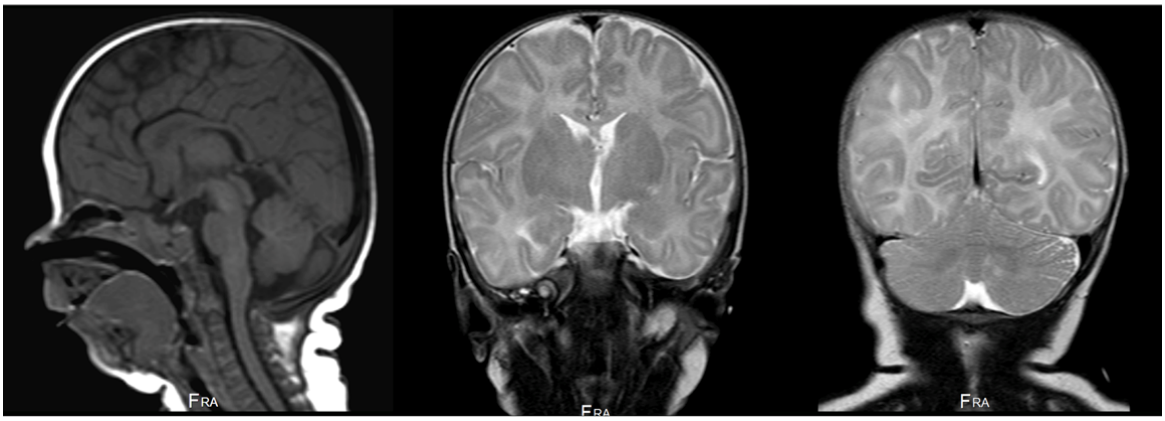
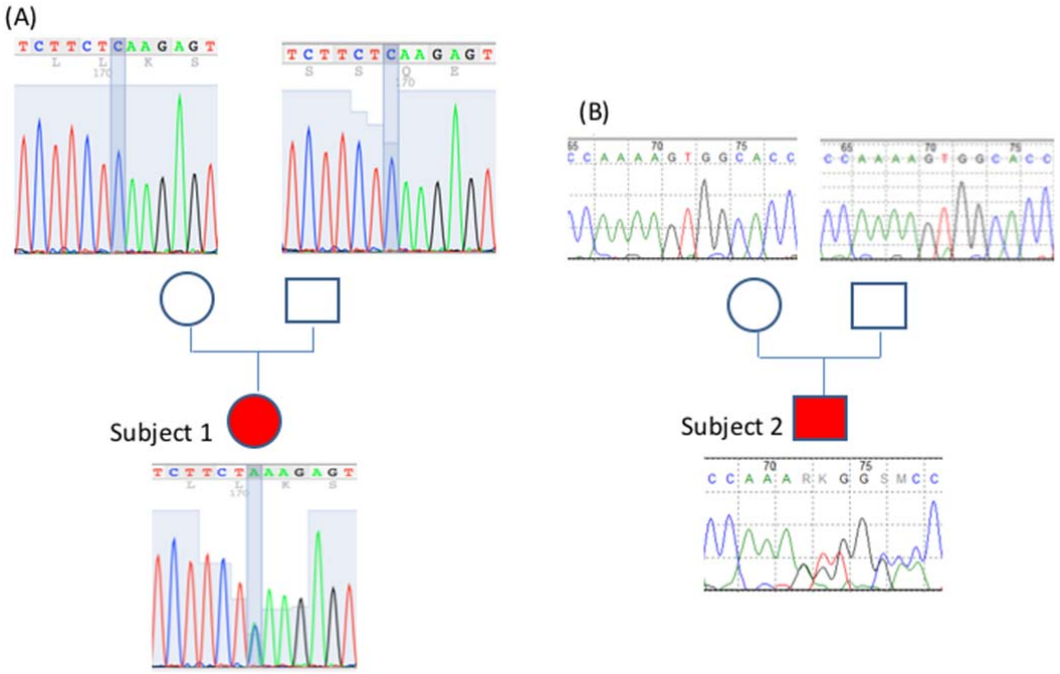
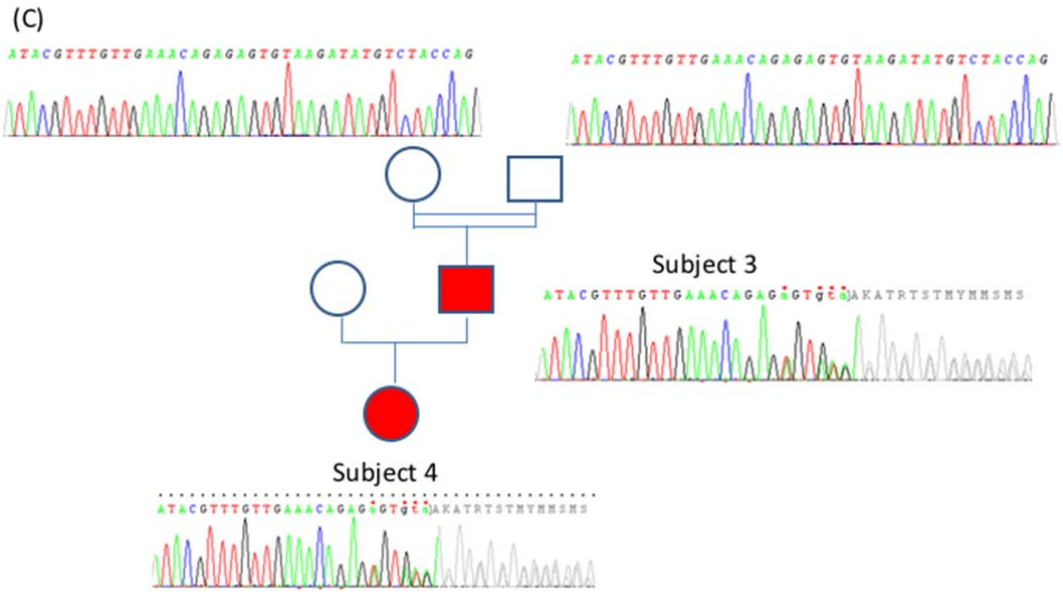


Figure S2. Sanger sequencing chromatogram of *ARCN1* mutations. (A) *ARCN1* mutation in Subject 1. Neither of the parents carried the same mutation. (B) *ARCN1* mutation in Subject 2. Neither of the parents carried the same mutation. (C) *ARCN1* mutation in Subjects 3 and 4. *ARCN1* mutation was not detected in the parents of Subject 3. (D) *ARCN1* mutations introduced by CRISPR/Cas9 in HCT116 cell line clones.




(C)
Subject 3
Subject 4

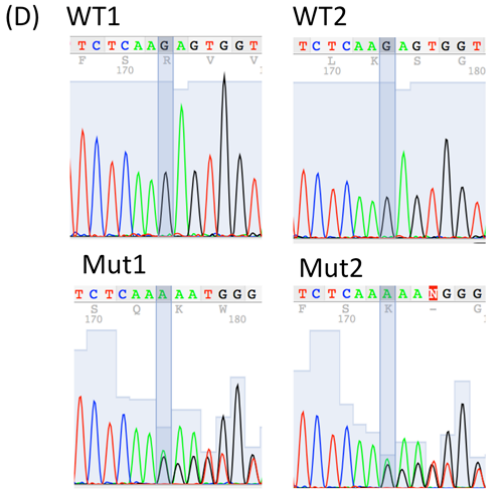
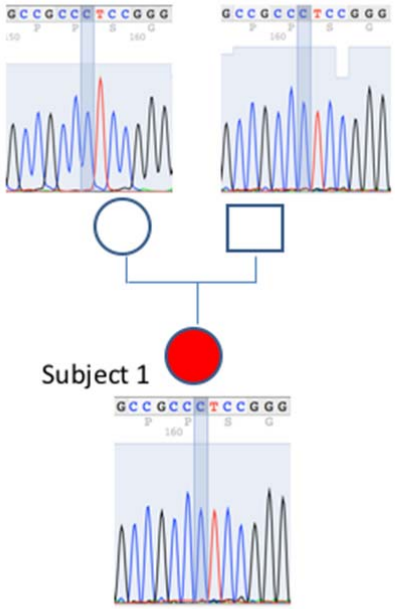

(D)
WT1
WT2
Mut1
Mut2

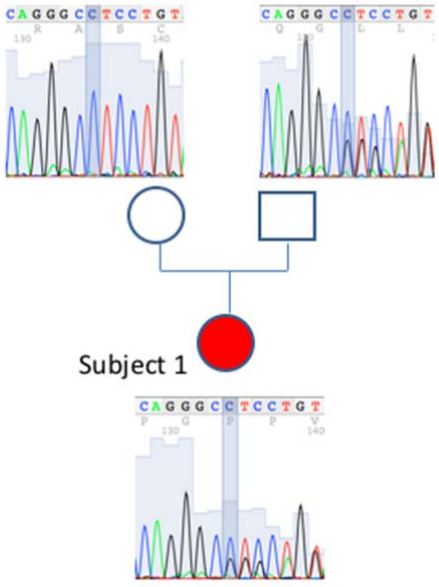
Figure S3. Sanger sequencing confirmation of *FAM58A*, *DCAF15* and *SYT1* variants. (A) The *FAM58A* variant was identified not only in Subject 1, but also in both parents, arguing against a possible role of the *FAM58A* variant in explaining the clinical phenotype of Subject 1. This variant was reported in dbSNP as rs200645137, and its population frequency is between 0.987 and 1.0 in the Human Genetic Variation database (http://www.genome.med.kyoto-u.ac.jp/SnpDB/index.html). (B) The *DCAF15* variant was confirmed in Subject 1, but the same variant was also identified in her father, arguing against its causality. A similar insertion was reported in dbSNP as rs3217681, and its population frequency is 0.407 according to the Human Genetic Variation database. (C) The *SYT1* variant in Subject 2 was confirmed by Sanger sequencing, and neither parent was found to have this variant. This *SYT1* variant represents a novel variant, and has not been previously reported in the public SNV database.

(A) *FAM58A* Sanger sequencing



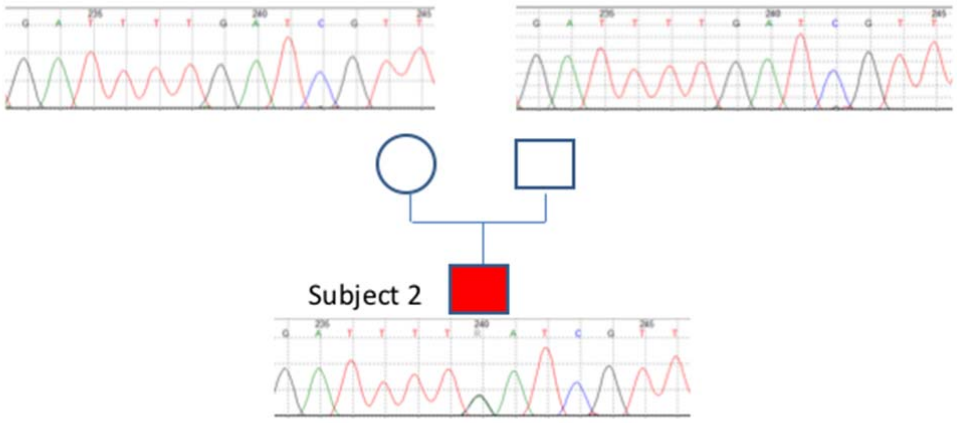
FAM58A: NM_152274: c.16_17ins G

(B) *DCAF15* Sanger sequencing



DCAF15:NM_138353:c.1439_1440insGGTGGGCCCAGGGCGGGCAG:p.E480fs

(C) *SYT1* Sanger sequencing



SYT1: NM_001135805.1:c.697G>A

Figure S4. Quantitative RT-PCR and western blot using *ARCN1*-related syndrome subject-derived skin fibroblast samples. (A) Quantitative RT-PCR. Total RNA was extracted from fibroblasts with the RNeasy Mini Kit according to the manufacturer's instructions (Qiagen, Courtaboeuf, France). Total RNA was subjected to reverse transcription using the iScript cDNA synthesis kit (Bio-rad, Marnes la Coquette, France). qPCR was performed in duplicate for each sample using the SybrGreen Supermix (Bio-rad). The relative expression of *ARCN1* was expressed relative to the expression of *HPRT*. The *ARCN1* expression level was significantly reduced in samples from Subjects 3 and 4 compared to that of the control sample. *** P<0.001, one-way ANOVA followed by a Bonferroni post-test. Error bars represent standard error of the mean. Two replicates were performed. (B) and (C) ARCN1 and BiP western blotting. Whole cell extracts were fractionated by SDS-PAGE and transferred to a polyvinylidene difluoride (PVDF) membrane. After incubation with 10% nonfat milk in TBST (10 mM Tris, pH 8.0, 150 mM NaCl, 0.5% Tween 20) for 1 h, membranes were incubated with antibodies against ARCN1 (COP1D, AV54594, Sigma-Aldrich), BiP (Grp78, ab21685, Abcam) or beta-Actin (MAB1501, Millipore) overnight at 4 °C. Membranes were washed three

times and incubated with horseradish peroxidase-conjugated anti-mouse or anti-rabbit antibodies. Blots were washed with TBST and developed with the ECL system (Amersham Biosciences) according to the manufacturer's protocols. (B) The protein level of ARCN1 was slightly reduced in the Subject 4 sample, compared to that of the control sample. Two replicates demonstrated consistent results. (C) Western blotting with BiP antibody yielded two bands, and lower band representing BiP based on the molecular size (about 78kDa). The BiP protein, which is an ER stress marker, was elevated in the Subject 4 sample, compared to that of the control.

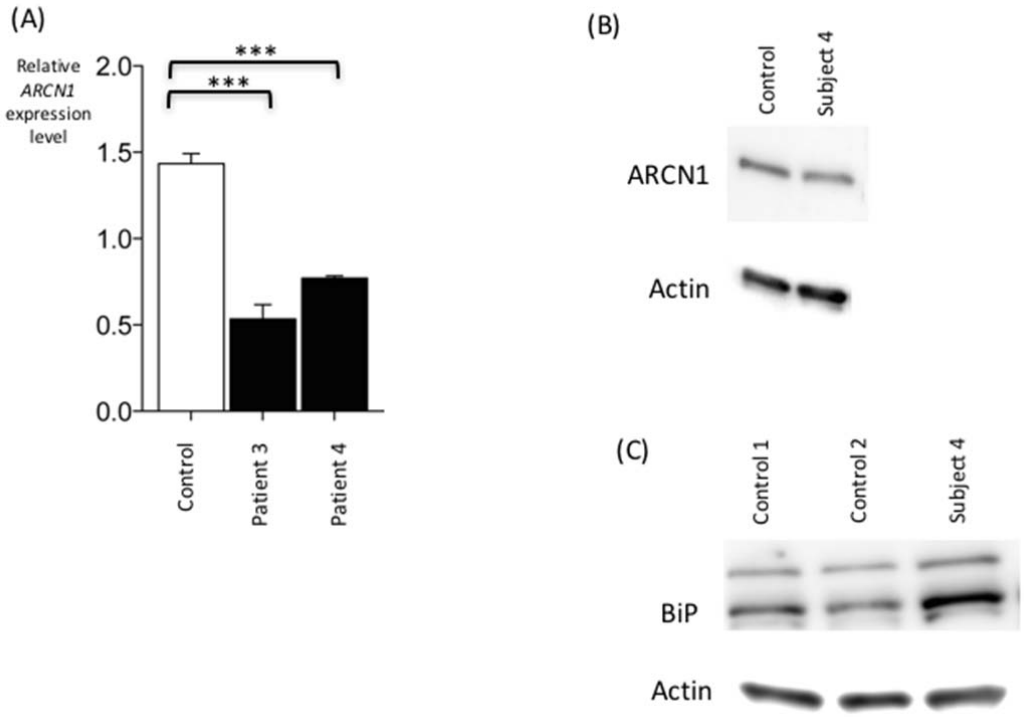


Figure S5. Immunocytochemical analysis. Cultures were fixed for 20 min with 4% paraformaldehyde. For immunocytochemistry, the cells were permeabilized with 0.3% TritonX-100 for 15min, washed, blocked for 1 h with 10% serum in PBS and incubated with primary antibody overnight at 4°C. Primary antibodies used were anti-ARCN1/COP1D (ab96725, Abcam), anti-GM130 (610823, BD Transduction Lab), anti-GIANTIN (ab80864, Abcam), and anti-PDI (MA3-019, Thermofisher). Following washing, cells were treated with secondary antibody for 1 h in the dark at room temperature. All images were captured on a Zeiss microscope IMAGER.Z1, using the Apotome imaging system coupled to AxioCam MRm and the Axiovision Rel.4.8 software. BiP represents an ER stress marker, and GIANTIN and GM130 are Golgi markers. PDI is a marker of ER. Immunofluorescence staining of the Golgi apparatus (GIANTIN and GM130), and that of ER (PDI) as well as COPI vesicle (ARCN1) indicated the swollen appearance of the Golgi stacks in the Subject 4's fibroblast. Accumulation of BiP in the cytosol is demonstrated in the Subject 4's fibroblast compared to that of the control.

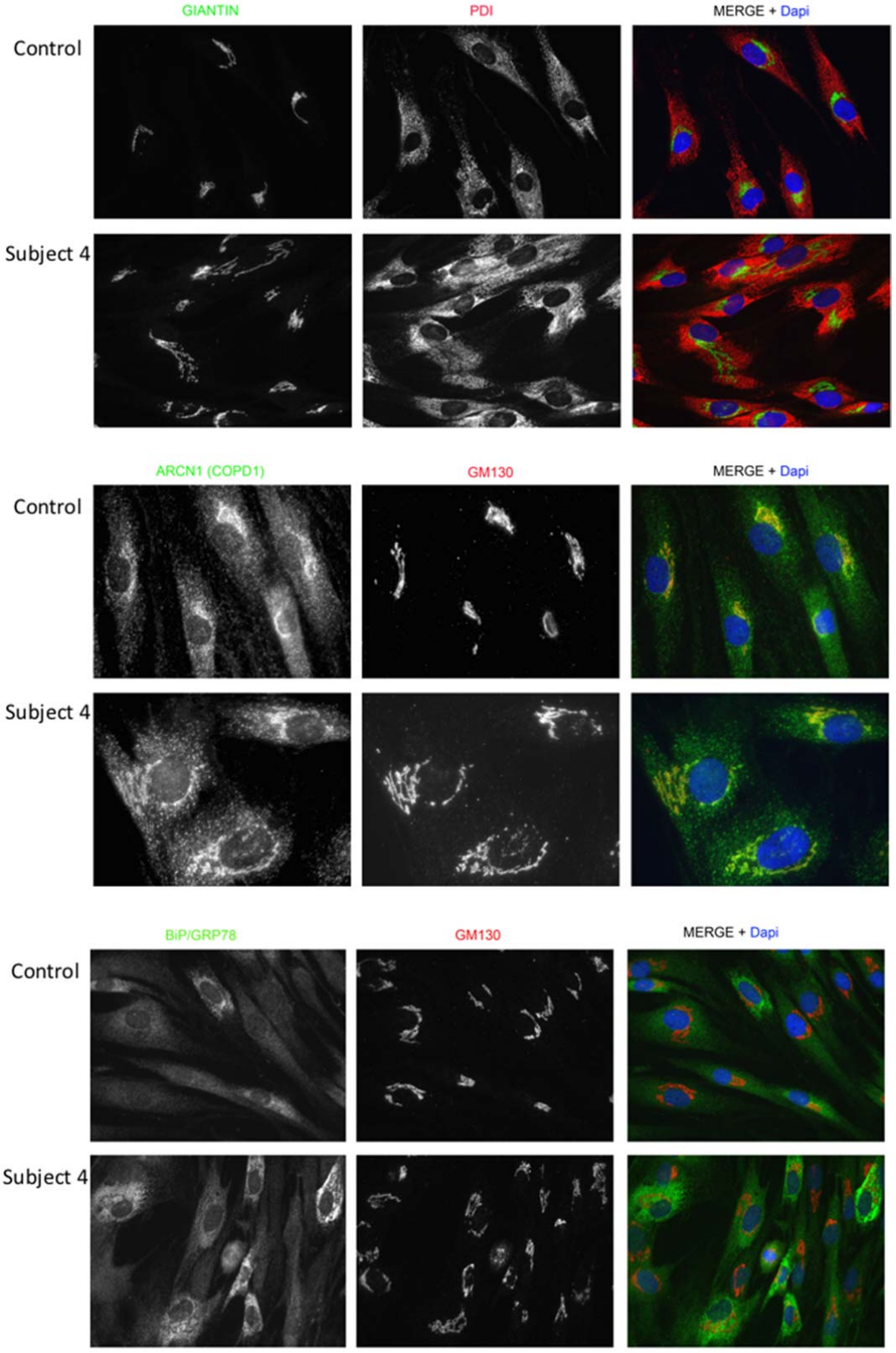

GIANTIN
PDI
MERGE + Dapi
Control
Subject 4
ARCN1 (COPD1)
GM130
MERGE + Dapi
Control
Subject 4
BiP/GRP78
GM130
MERGE + Dapi
Control
Subject 4

Figure S6. ER stress gene expression was elevated with tunicamycin and thapsigargin treatments. Data represents mean ± 2 s.d. *** P<0.001, two-tailed *t* test; n=3 technical replicates per group. Quantitative PCR was run in triplicates, and mean and standard deviation was calculated. Two different concentrations of tunicamycin treatment (2µg/ml (Tuni2) and 5µg/ml (Tuni5)) and thapsigargin treatment (2µM (Tg2) and 5µM (Tg5)) were used. DMSO was added for control. Gene expression level was normalized against *GAPDH* expression. Relative gene expression level to the control sample is demonstrated.

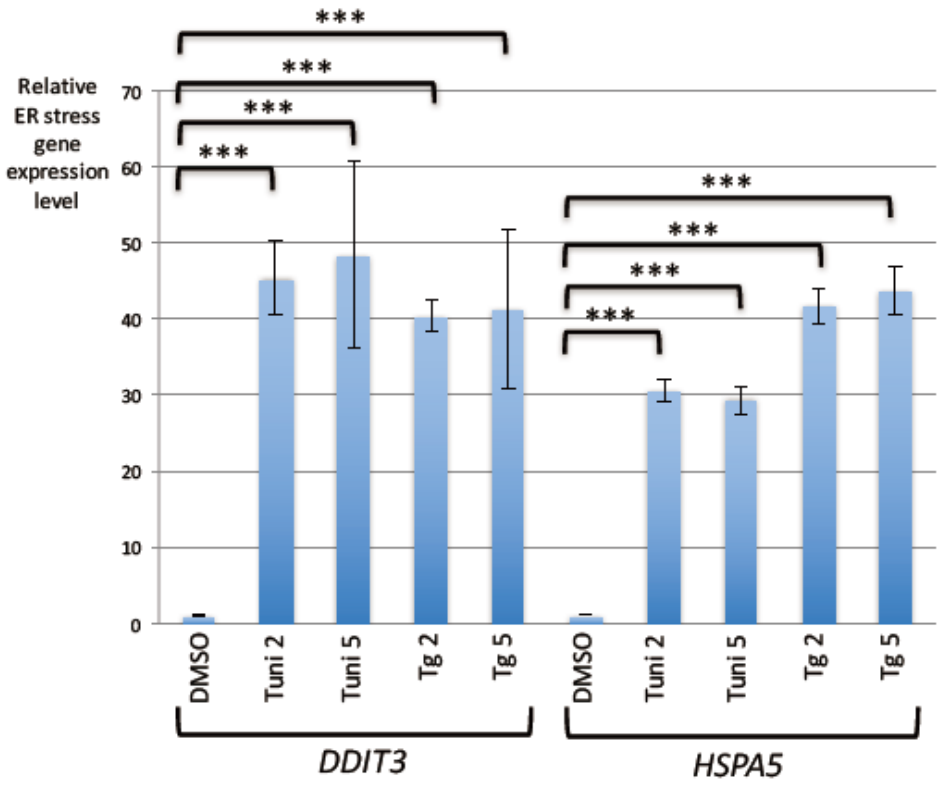


Table S1. Exome statistics

sample name	platform	read length (bp)	read type	method	total reads	mapped reads	mapping rate (%)	mapping tool
Subject 1	Hiseq 2500	100 or 126	paired-end	Exome sequencing	134459385	134274949	99.86	BWA 0.7.12
Subject 1 Father	Hiseq 2500	100 or 126	paired-end	Exome sequencing	138379220	138209817	99.88	BWA 0.7.12
Subject 1 Mother	Hiseq 2500	100 or 126	paired-end	Exome sequencing	136127612	135820259	99.77	BWA 0.7.12
Subject 2	Hiseq 2000	100	paired-end	Exome sequencing	121420265	120842553	99.52	BWA 0.7.5a
Subject 2 Father	Hiseq 2000	100	paired-end	Exome sequencing	112431640	112118195	99.72	BWA 0.7.5a
Subject 2 Mother	Hiseq 2000	100	paired-end	Exome sequencing	111109537	110800648	99.72	BWA 0.7.5a

Subject 3	Hiseq 2000	75	paired-end	Exome sequencing	70427250	68814466	97.71	ELANDv2e
Subject 3 Father	Hiseq 2000	75	paired-end	Exome sequencing	73368174	71746737	97.79	ELANDv2e
Subject 3 Mother	Hiseq 2000	75	paired-end	Exome sequencing	50373665	49265444	97.8	ELANDv2e

Read information for Subjects 1, 2 and 3

sample Name	mean	proportion of bases above 10X coverage (%)	proportion of bases above 20X coverage (%)	proportion of bases above 30X coverage (%)	proportion of bases above 50X coverage (%)
Subject 1	200.67	99.6	99.1	98.3	95.4
Subject 1 Father	208.9	99.7	99.1	98.3	95.4
Subject 1Mother	194.82	99.6	99.1	98.2	95.2
Subject 2	99.96	96	93	89	76
Subject 2 Father	91.96	95	93	88	72
Subject 2 Mother	90.95	95	92	87	71

sample Name	mean depth (X)	proportion of bases above 1X coverage (%)	proportion of bases above 4X coverage (%)	proportion of bases above 10X coverage (%)	proportion of bases above 25X coverage (%)
Subject 3	72	99	97	93	82
Subject 3 Father	77	99	97	93	83
Subject 3 Mother	58	99	96	91	77

Coverage information for Subjects 1,2 and 3

Table S2. Quantitative RT-PCR primers

	F	R
GAPDH	GCACCGTCAAGGCTGAGAAC	TGGTGAAGACGCCAGTGGA
HPRT	GGTGAAAAGGACCCCACG	TCAAGGGCATATCCTACAACA
ARCN1 ex9-10	CGAAATACCCTGGAGTGGTG	GGGGCTGTTTCCATCTACCT
ARCN1 ex1-2	TCCCGTTCCCCAGACCCTA	TTTCCTGCTTTTGTGCAGAC
ATF4	TCAAACCTCATGGGTTCTCC	GCATGGTTTCCAGGTCATCT
DDIT3	GACCTGCAAGAGGTCCTGTC	AGGTGCTTGTGACCTCTGCT
HSPA5	TGAAACTGTGGGAGGTGTCA	TTTGTCAGGGGTCTTTCACC
COL1A1	GACTGGCAACCTCAAGAAGG	GGAGGTCTTGGTGGTTTTGT
COL1A2	GCAACCTGAAAAAGGCTGTC	GGCGTGATGGCTTATTTGTT

ARCN1 ex9-10 primer pair was used for HCT116 cell lines, and *ARCN1* ex1-2 primer pair was used for skin fibroblast cell lines.

Table S3. Comparison of *SYT1* mutation cases

Subject	*SYT1* variant	Variant effect prediction			Clinical phenotype				
		Polyphen2 HumDiv	SIFT	Mutation taster	Developmental delay	Movement disorder	Seizure	Micrognathia	IUGR
Subject 2 (Current report)	c.697G>A; p.Asp233Asn	Probably damaging	Tolerated	Disease causing	Mild	No	Yes	Yes	Yes
Baker et al.(2015) J. Clin. Invest. 125, 1670–1678.	c.1103T>C; p.Ile368Thr	Benign	Tolerated	Disease causing	Profound	Yes	No	No	No
Cafiero et al.(2015) Eur. J. Hum. Genet. 23, 1499–1504.	c.908T>A; p.Met303Lys	Benign	Damaging	Disease causing	Moderate	No	No	No	No

Supplementary References

1. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma. Oxf. Engl. *25*, 1754–1760.
2. 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.
3. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80–92.
4. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.