intron 11 by molecular sizing of cDNA isolated from LCLs from four cases and three controls.

Quantitative PCR was performed using the THUNDERBIRD SYBR qPCR mix (TOYOBO) and a 7300 Real Time PCR System (Applied Biosystems). The primer sequences for *RNF213* cDNA or *FLJ35220* are described in the Text S1.

Northern blotting

Total RNA was isolated using a QIAamp RNA blood mini kit (Qiagen Inc.). A human adult normal tissue mRNA northern blot I (Biochain) was probed in accordance with the supplier's recommendations. The two probes [*RNF213*_1 (492 bp) and *RNF213*_2 (591 bp)] have been described in the Text S1. The mRNA levels were determined using an Image Analyzer FLA2000 (Fuji Film).

Rapid amplification of cDNA ends (RACE)

The 5′ and 3′ ends of the *RNF213* cDNA were determined by RACE using a GeneRacer kit (Invitrogen) according to the manufacturer's protocol. Details can be found in the Text S1.

Allele-specific mRNA expression assay

Polymorphisms of p.R4810K and a nearby SNP, p.H4557H, in *RNF213* cDNA were used as markers for the SNaPshot assay to measure the expression of the two alleles. Primers for the SNaPshot assay are shown in Table S6.

Western blotting

Cells were lysed in buffer containing 50 mM Tris-HCl pH 8.0, 1% NP-40 and 150 mM NaCl, or in CelLytic M (Sigma) containing a protease inhibitor cocktail. Samples were subjected to immunoblotting using an anti-RNF213 antibody (MyBioSource), anti-HA antibody (mouse 6E2; Cell Signaling Technology) or anti-Myc antibody (mouse 9E10; Santa Cruz).

Immunostaining

The methodology used for immunostaining has been described in detail in the Text S1.

Ubiquitin ligase assay

Plasmids, immunoprecipitation and the self-ubiquitination assay have been described in a previous report [26]. Briefly, cells were lysed in buffer containing 50 mM Tris-HCl pH 8.0, 1% NP-40 and 150 mM NaCl, then centrifuged at 13,000 × *g*. The supernatants were incubated with an anti-HA antibody, and immune complexes were captured using protein G-Sepharose (GE Healthcare). Immunoprecipitates were subjected to immunoblot analysis using an anti-Myc antibody.

Protein preparation and ATPase assay

The region containing the Walker motif (amino acids 2359–2613) was subcloned into the bacterial expression vector pGEX5X-1. The N-terminal glutathione *S*-transferase (GST)-tagged fragment was first purified with GSH beads (GE Healthcare), then was purified using a gel-filtration column (Superdex 200 prep grade; Amersham Pharmacia) in conjunction with the AKTAexplorer system (GE Healthcare). The ATPase assay was performed as described in Text S1.

Effects of p.R4810K or p.D4013N on biochemical function of *RNF213*

Nucleotide changes corresponding to p.R4810K and p.D4013N were introduced into *RNF213* using a site-directed mutagenesis kit (Invitrogen). These mutants were transiently expressed in HEK293 cells as described in the Text S1. Cells were then lysed and subjected to the ubiquitin ligase assay as described above, or subjected to subcellular fractionation as described in the Text S1. To generate a RING finger-deleted mutant of *RNF213*, an exogenous *Eco*RI site was fused to the 3′ terminus just before the RING finger domain, then this site was ligated to an endogenous *Eco*RI site located just after the RING finger, enabling the RING finger domain to be skipped from the *RNF213* cDNA.

Zebrafish model

The Zebrafish experiments were approved by the Osaka University Animal Welfare committee (Osaka University, Japan; approval ID 2197). Embryos and adult fish were raised and maintained under standard laboratory conditions. For *in vivo* experiments, Tg(fli-EGFP)y1 zebrafish [27] were purchased from the Zebrafish International Resource Center. Knockdown of *RNF213-α* and *RNF213-β* expression was achieved by injection of a specific morpholino (MO; Gene Tools) into 1- to 8-cell stage embryos, as described previously [28]. The methods are described in detail in the Text S1.

Results

Linkage analysis and identification of the critical region

Linkage analysis for the eight largest families (Figure 2) revealed a single peak at 17q25.3 with a multipoint LOD score of 8.46 at D17S784 (Figure 3). The largest maximum multipoint LOD score in the 10,000 simulations was 7.25 with a median of 3.56 and a 95% confidence interval between 2.16 and 5.48. Thus, linkage of the eight families to 17q25.3 was highly statistically significant and was unlikely false positive. To evaluate the effect of uncertainty regarding the allele frequency of the disease gene, we conducted a sensitivity analysis by changing the values to 0.000001, 0.001 or 0.01. The sensitivity analysis did not change the maximum multipoint LOD score by more than 1%.

Fine mapping increased the LOD score slightly to 8.52 at D17S784, and revealed several recombination events in the region involving the two flanking markers D17S1806 and rs2280147 (Figure 2). Thus, the core of the locus was a 1.5-Mb region, which harbored 21 genes from *hCG 1776007* to *RPL31P7* (Table S4).

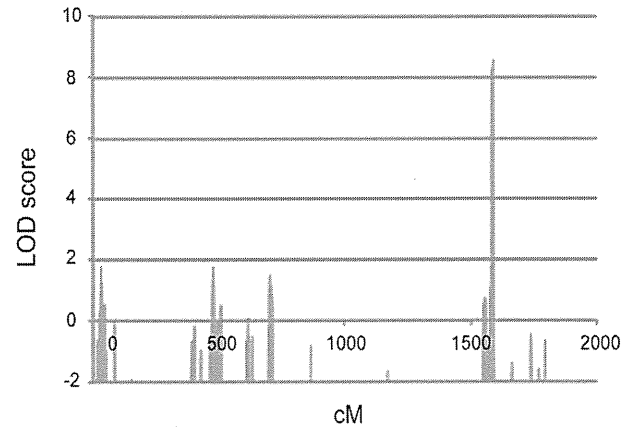


Figure 3. Genome-wide linage analysis uniquely identified a locus on 17q25.3 for the eight families. Genome wide linkage analysis showed a highest LOD score (8.46) at D17S784 in the locus 17q25.3.
doi:10.1371/journal.pone.0022542.g003

The disease haplotype showed complete segregation in the eight families (Figure 2).

Exome analysis

We generated approximately 154 million reads with an on target rate of 59.1%. Twelve billion bases passed the quality assessment and were aligned to the human reference sequence; >98.1% of the bases mapped to the target, with a mean coverage depth >116.8 in each index case. For coverage on the 1.5-Mb region on 17q25.3, we generated approximately 97000 reads with an on target rate of 52.5%. The mean coverage depth was 48.7 with 288 exons, assuring that all genes were adequately covered.

We applied Ng *et al.* 's filtering algorithm [29], with some modifications to make it less stringent because we anticipated genetic heterogeneity and an uncharacterized gene. Additionally, because a reported discordance of phenotype between identical twins indicated a low penetrance [20], [21], we considered as candidates both rare variants that are specific to moyamoya, and more common variants that are more frequently observed in cases than in controls.

We compared our missense (MS)/nonsense (NS)/splicing site (SS)/insertion or deletion (Indel) variants in the eight index cases against dbSNP 131 (Appendix S1) and the exome database of the five Japanese controls, and removed the reported SNPs and SNPs found in the five Japanese controls (Table 1).

The mean numbers of MS/NS/SS/Indel per cases were 6601 from the genome-wide analysis. We next examined the effects on the size of the candidate gene list when analyzing the exomes of the eight index cases in various combinations and examined the potential consequences of genetic heterogeneity, such that only a subset of the exomes of index cases was required to contain new variants in a given gene for it to be considered as a candidate gene (Table 1 and S7). In the second stage of the filtering in which we assumed that any seven of the eight index cases shared the causative gene (Table S7), the numbers of the candidate genes were decreased to two. By filtering two variants in the two genes emerged as candidates (Table 1 and Table S7). These two variants were p.N321S in *PCMTD1* (Chr8: NM_052937) and p.R4810K in *RNF213* (Chr17: NM_020914.4). *PCMTD1* encodes the protein-L-isoaspartate O-methyltransferase domain-containing protein 1(Genecard in Appendix S1) ; *RNF213* encodes ring finger protein 213. We thus considered both *PCMTD1* and *RNF213* as candidate genes.

Confirmation of the exome data by direct sequencing and segregation analysis

In the next step, we conducted sequencing to confirm p.N321S in *PCMTD1* and sequenced the entire exons of *RNF213* with an ORF (NM_020914.4) in 42 index cases by the Sanger method (Table S3). We could not confirm p.N321S in *PCMTD1* in any of the eight index cases by the Sanger method. Thus we discarded *PCMTD1*. In contrast, we could confirm p.R4810K by the Sanger method in 42 index cases but could not find any other unregistered polymorphism in *RNF213*. We genotyped other family members in the 42 families. As shown in Figure S1, p.R4810K was completely segregated in 42 families: all the affected members by RCMJ criteria had p.R4810K, although some carriers were not affected with moyamoya disease, suggesting low penetrance.

Deep sequencing around *RNF213*

To avoid a possible pitfall of exome for searching variants in introns or intergenic regions, we sequenced the entire 260-kb genomic region (UCSC Genome Browser in Appendix S1) around *RNF213* from the 5′ end of *SLC26A11* to the 5′ end of *NPTX1* in the index case of pedigree 11 using BAC clones or direct sequencing. Sequencing the 2.7-kb intron 15 of *RNF213* failed because of the presence of repeats. Southern blotting of the eight index cases and a control (the spouse of individual 2 in pedigree 18) confirmed the absence of insertions or deletions in this intron (data not shown). However, sequencing of the 260-kb genomic region revealed three additional unregistered variants in noncoding regions: ss179362671 (T>C) in intron 15 of *RNF213*, ss179362674 (G>A) in intron 11 of *FLJ35220* and ss179362675 (C>T) at the 3′ end of *NPTX1*, 1868 bp downstream from the 3′UTR.

Exonic rare variants of *RNF213* in Japanese controls

We then sequenced all the exons of *RNF 213* in ten controls and conducted deep sequencing of *RNF213* for a Japanese control. These results and five controls for exome analysis are shown in Table S8. Additionally, 38 chromosomes inherited from parents in 38 index cases, which do not carry p.R4810K, can be identified to

Table 1. Numbers of variants identified based on filtering.

		Index Case of the Family								
Stage	**Criteria**	**Ped 2**	**Ped 10**	**Ped 14**	**Ped 15**	**Ped 17**	**Ped 18**	**Ped 19**	**Ped 20**	**Any eight**
1st Stage	MS/NS/SS/Indel	6638	6741	6512	6668	6561	6361	6822	6507	1038
	Not in dbSNP131	983	1036	963	966	953	1064	1019	934	75
2nd Stage	Not in the five Japanese controls*	583	590	565	611	594	534	624	605	2 : p.N321S in PCMTD1 and p.R4810K in RNF213
	No of genes	377	379	364	372	375	354	380	385	2
	In the 17q25.3	1	1	1	1	3	3	1	2	1
	Genes No of genes	RNF213 1	RNF213 1	RNF213 1	RNF213 1	RNF213 GAA ENPP7 3	RNF213 ENPP7 SLC26A11 3	RNF213 1	RNF213 SLC26A11 2	RNF213 1

Rows show the effect of excluding from consideration of variants found in dbSNP131 and the five Japanese controls. Columns show MS/NS/SS/Indel variants that were observed in each affected individual. The column 10 provided observation that shared by all affected index cases.
*Exome analysis were conducted using the same platform and same experimental conditions for five Japanese controls.
doi:10.1371/journal.pone.0022542.t001

have been transmitted from control parents. Pooling these data, an observed sample size of control chromosomes accounts for 70 (five controls by exome, 10 controls by direct sequencing, one by deep sequencing, and 38 chromosomes inherited from unaffected parents of 38 index cases). Given that the statistical power should be greater than 80%, rare exonic variants with minor allele frequencies of as low as 2% could be detected in the sample size of 70. However, we could not find any unregistered rare variants on control chromosomes.

Possible masked variants in the linkage disequilibrium (LD) region with p.R4810K (ss179362673) and possible CNVs in 17q25.3

Although p.R4810K seems to be a strong candidate for causative variants of moyamoya disease, there also seem to be other possibilities. One of these is the presence of masked variants in the promoters, or intronic or intergenic regions of a bona fide causal gene, which are in strong LD with p.R4810K.

We determined the minor allele frequencies for five unregistered variants found in this study and a previous study [21] in Japanese controls. The frequencies were 0.012 (p1 = 9/768) for ss179362671, 0.014 (p2 = 11/768) for ss179362673, 0.013 (p3 = 10/768) for ss179362674, 0.022 (p4 = 17/768) for ss179362675 and 0.010 (p5 = 8/768) for ss161110142 in 384 Japanese controls, indicating that they are all rare variants. We then genotyped the 42 families. As shown in Figure 4, these four rare variants were transmitted en bloc with p.R4810K. They comprised five subhaplotypes (Figure 5); and p.R4810K was not transmitted alone in any of the families.

Haplotype 2, the most prevalent haplotype, was found in 34 of the 42 families. Because all five variants are rare, it is unlikely that haplotype 2 has been formed by chance ($p = 2^5 \times p1 \times p2 \times p3 \times p4 \times p5 = 1.5 \times 10^{-8}$). Alternatively, a more rational explanation is to assume that haplotype 2 is the founder haplotype. Along this line, haplotypes 3–6 may be explained by historical recombination events. Given historical recombinations, the number of masked variants is very likely to be one and it should be located in the core LD region between ss179362671 and ss179362675 (Figure 5).

Further confirmation of the sequence homogeneities was evaluated by genotyping of 34 additional markers. Transmissions of the 39 SNPs covering *SGSH* to *Raptor* were investigated in the 42 families. The five haplotypes were further classified into subtypes (Figure 5). The typing results confirmed the sequence homogeneity in the core LD region (Figure 5). The genotyping of the SNPs in the *RNF213* regions were completely in accord with the sequencing results in the 42 index cases. Thus, we concluded that the core LD region has been very likely derived from the founder haplotype.

Although we sequenced the coding and non-coding regions in the index case of pedigree 11, we did not detect any unregistered rare variants other than the five rare variants, supporting that p.R4810K or G>A substitution in the intron 11 of *FLJ35220* is a susceptibility variant. In terms of G>A substitution in the intron 11 of *FLJ35220*, we have tested two possibilities, namely that the substitution may induce aberrant splicing or change gene expression levels. Either possibility was discarded (Figure S3). We still tested another possibility. It can sometimes be difficult to detect CNVs by direct sequencing. Therefore, we evaluated the CNVs using a high-density microarray, with a theoretical sensitivity estimated to be 41 kb [30]. However, we did not detect any CNVs in the 17q25.3 region in the three index cases of pedigrees 5, 11 and 18 and a control 2 in pedigree 18 (Figure S4).

Although we cannot theoretically discard the very rare possibility of multiple masked variants having strong LD with p.R4180K by various LD patterns specific to the individual families, these data collectively indicate that it is not a surrogate marker of possible masked variants of another gene. Therefore, *RNF213* is very likely to be a bona fide susceptibility gene for moyamoya disease.

Association of p.R4810K with moyamoya disease in East Asian patients

We next tested the association of *RNF213* p.R4810K with moyamoya disease in East Asian populations. Unrelated cases and index cases were exclusively limited to those who met the diagnostic criteria for definitive moyamoya disease [22]. As shown in Table 2, ss179362673 (p.R4810K in *RNF213*) was significantly associated with moyamoya disease, with a maximum odds ratio (OR) of 338.9 ($P = 10^{-100}$). The association was perfectly replicated in the Korean (OR = 135.6, $P = 10^{-26}$) and Chinese (OR = 14.7, $P = 10^{-4}$) populations (Table 2). Stratification by family histories of moyamoya disease did not substantially change the association of p.R4810K with moyamoya disease (Table S9).

Variant searching in non-p.R4810K East Asian cases and in Caucasian cases

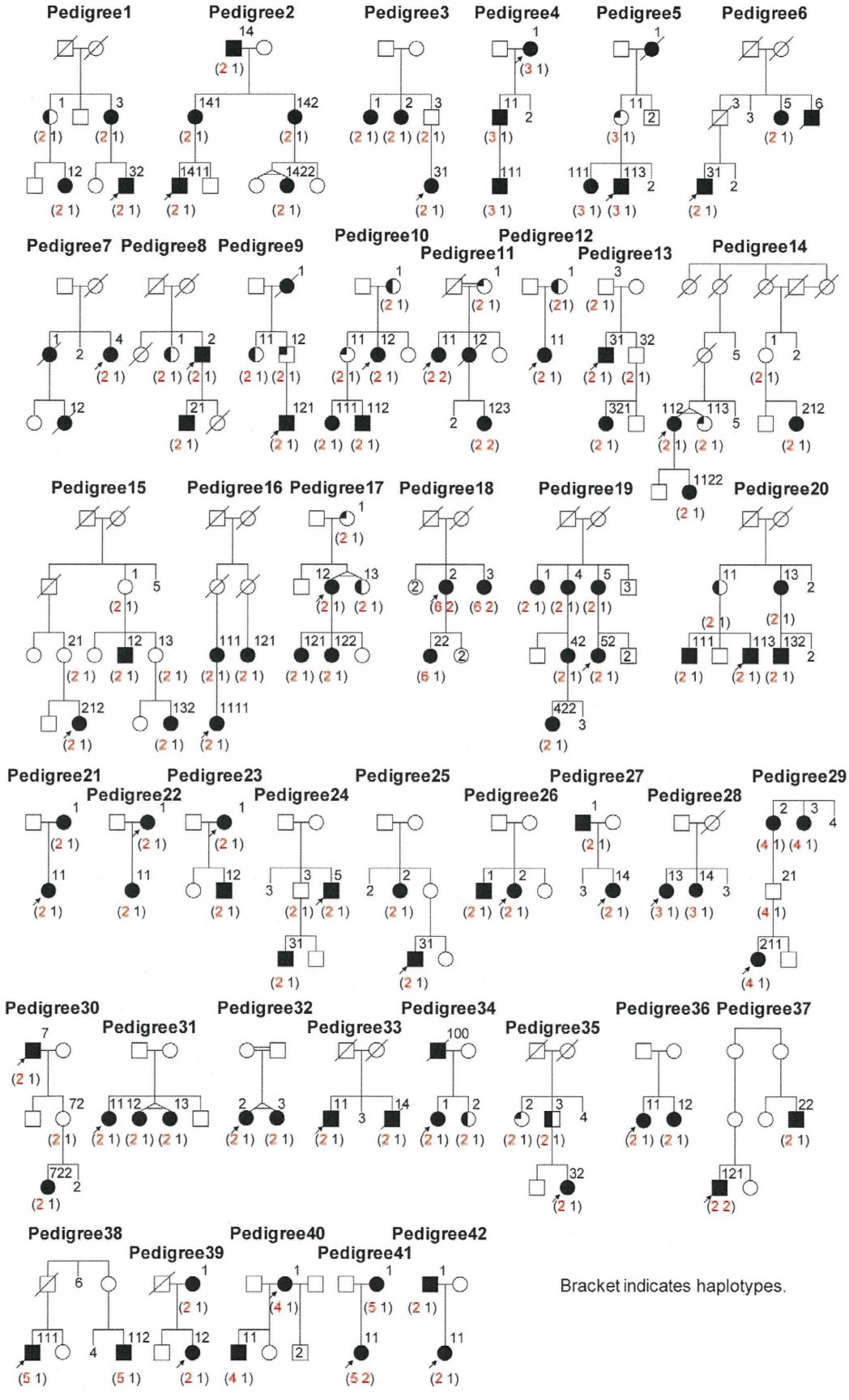
Although p.R4810K was identified in three ethnic populations, the population attributable risks in the Japanese (145/161; 90%) and Korean (30/38; 79%) populations were larger than that in the Chinese population (12/52; 23%). Additionally, we showed by genotyping that the p.R4810K variant was not present among Caucasian cases or controls. These pieces of evidence strongly suggest genetic heterogeneity. We thus sequenced *RNF213* to identify other variants in non-p.R4810K East Asian cases and Caucasian cases.

By direct sequencing, five distinct *RNF213* variants, i.e., p.D4863N, p.E4950D, p.A5021V, p.D5160E and p.E5176G, were identified in seven out of 64 East Asian cases (40 Chinese, 16 Japanese and 8 Korean) (Table 3, Figure S5). None of these variants was found among 757 East Asian controls (Table 3). In Caucasians, four distinct variants, i.e., p.N3962D, p.D4013N, p.R4062Q and p.P4608S, were identified in 4 out of 50 (8%) cases (Table 3): in the index case of the Caucasian pedigree (Figure S2A) and 3 out of 49 other cases (Figure S5). In the Czech family, the p.D4013N (G>A) variant was transmitted to the three affected children and segregated perfectly with moyamoya disease (Figure S2). None of these variants was found among 384 Caucasian controls (Table 3).

In total, our extensive sequencing analysis identified 10 variants in *RNF213* (Table 3). These variants were all located in the 3′ half of the gene (Figure 6). We concluded that *RNF213* is a susceptibility gene for moyamoya disease.

Full-length cDNA cloning of *RNF213*

Because of its large size, *RNF213* was first cloned as five separate fragments, which were then connected using internal restriction enzyme sites, as indicated in Figure S6. The 5′ and 3′ ends of cDNA from the LCL of the index case of pedigree 11 were determined by RACE, which defined the transcriptional start site and the 3′ end of the *RNF213* gene as at nucleotides 78,234,667 and 78,370,086, respectively (UCSC Genome Browser in Appendix S1). Thus, the full length *RNF213* was found to have a 15624-bp ORF and 5431-bp 5′ and 3′UTR. The whole cDNA of *RNF213* [AB537889] is similar in size to the cDNA of *RNF213* [NM_020914.4], but there are differences: it lacks exon 4 and has a 3′UTR 2500 bp longer than that of NM_020914.4 (NCBI in Appendix S1).



Bracket indicates haplotypes.

Figure 4. Transmission patterns of the haplotypes in 42 families. The numbers in the parentheses indicate the haplotypes transmitted in each family. Haplotype 1 symbolizes non-risk haplotypes collectively. Therefore it comprises various haplotypes. Haplotypes can be seen in Figure 5. Several family members that were genotyped were omitted in the pedigree chart for clarity. Those members and their haplotypes were; two siblings (1,1) for pedigree 7; father (1,1) for pedigree 21; father (1,1) for pedigree 22; father (1, 1) and a sibling (1,1) for pedigree 23; father (1,1), mother (2,1) and a sibling (1,1) for pedigree 26; three siblings (1,1), (1,1) and (1,1) for pedigree 27; father (1,1), three siblings (1,1) (1,1) and (1,1) for pedigree 28; father (1,1), mother (2,1) and a child of 2 (1,1) for pedigree 32; three siblings (1,1), (1,1), and (1,1) for pedigree 33; mother (1,1) for pedigree 34; father (1,1) and mother (2,1) for pedigree 36; a sibling (1,1) for pedigree 39; three siblings (1,1), (4,1) and (1,1) for pedigree 40; father (2,1) for pedigree 41; mother (1,1) for pedigree 42.
doi:10.1371/journal.pone.0022542.g004

Two splicing variants of *RNF213* were detected in cDNAs isolated from bone marrow, cerebellum, whole brain, fetal brain, fetal liver, heart, kidney, liver, lung, placenta, skeletal muscle, testis, thymus, spinal cord, artery and HUVECs (Figure 7). The same splicing variants were also detected in LCLs isolated from the five controls and the six cases (data not shown). These two isoforms correspond to the short isoform (270 bp) which skips exon 4 (DDJB, AB537889) and the long isoform (417 bp) which reads exon 4 (NM_020914.4) (Figure 7). The short isoform is the major splicing variant and detected in all tissues examined, but the long isoform is a minor and not found in several tissues.

In silico analysis revealed two well-known domains, a RING finger domain and a Walker motif, which are reported to exhibit ubiquitin ligase activity and ATPase activity, respectively (Figure 6) [31],[32] (Genecard in Appendix S1).

Expression profile of *RNF213* and functional characterization

RNF213 mRNA is expressed in various human tissues (Figure 8A). Northern blotting revealed the expression of full-length *RNF213* in several human tissues (Figure S7A), cultured cell lines and LCLs (Figure S7B). We also detected the expression of the 591-kDa endogenous RNF213 protein in LCLs, HUVECs, CASMCs and HEK293 cells by western blotting using a RNF213-specific antibody. The western blot band size was consistent with that of overexpressed HA-tagged RNF213 (Figure 8B).

Transient transfection of HA-tagged *RNF213* into HeLa cells resulted in expression of the protein throughout the cytosol (Figure 8C), with partial association with the intracellular membrane and cytoskeleton (Figure S8B). The E3 activity of the RING finger domain was confirmed by self-ubiquitination after transfecting full-length *RNF213* cDNA into HEK293 cells (Figure 8D). ATPase activity was detected *in vitro* by free phosphate releasing analysis using a recombinant RNF213 fragment including the Walker motif (Figure 8E). Thus, *RNF213* is a unique protein, and we report here for the first time that it is a single protein that possesses both ubiquitin ligase activity and ATPase activity.

Homology search

A homology search revealed conservation of the arginine residue at position 4810 of *RNF213* in mammals (Table 4) (BLAST in Appendix S1). Three variants: p.R4062Q, which is located in the RING finger domain, p.R4810K and p.E5176G, were conserved among most species (Table 4). Variants p.N3962D, p.P4608S, p.D4863N, p.E4950D, p.A5021V, and p.D5160E were outside the RING finger domain and not conserved across species (Table 4). It should be noted that for five variants (p.N3962D, p.D4013N, p.E4950D, p.A5021V, and p.D5160E), the variant allele was present at the equivalent position in at least one of the species examined.

Biochemical effects of p.R4810K and p.D4013N

We tried to characterize the p.R4810K and p.D4013N allele proteins of *RNF213 in vitro*. We first investigated stability of the p.R4810K variant. The wild-type or p.R4810K variant of RNF213-HA was transiently expressed in HEK293 cells. Cells were lysed and subjected to immunoblotting with an anti-HA antibody. As shown, the molecular stability in cells was not altered (Figure S8A). Next we investigated subcellular localization of the p.R4810K variant of *RNF213*. HEK293 cells transiently expressing the wild-type or p.R4810K variant of RNF213-HA were fractionated into cytosol, membrane/organelle, nucleus, and cytoskeleton using different lysis reagents (ProteoExtract kit; Calbiochem). However, we could not detect alteration of intracellular distribution of p.R4810K (Figure S8B). Next we examined whether self-ubiquitination of the p.R4810K variant or p.D4013N mutant of *RNF213* was altered. HEK293 cells transiently expressing the wild-type or p.R4810K variant or p.D4013N mutant of RNF213-HA and Myc-ubiquitin were lysed and subjected to immunoprecipitation using an anti-HA antibody, followed by immunoblotting using an anti-Myc antibody. Neither of the two variants was found to alter ubiquitin activities (Figure S8C and S8D).

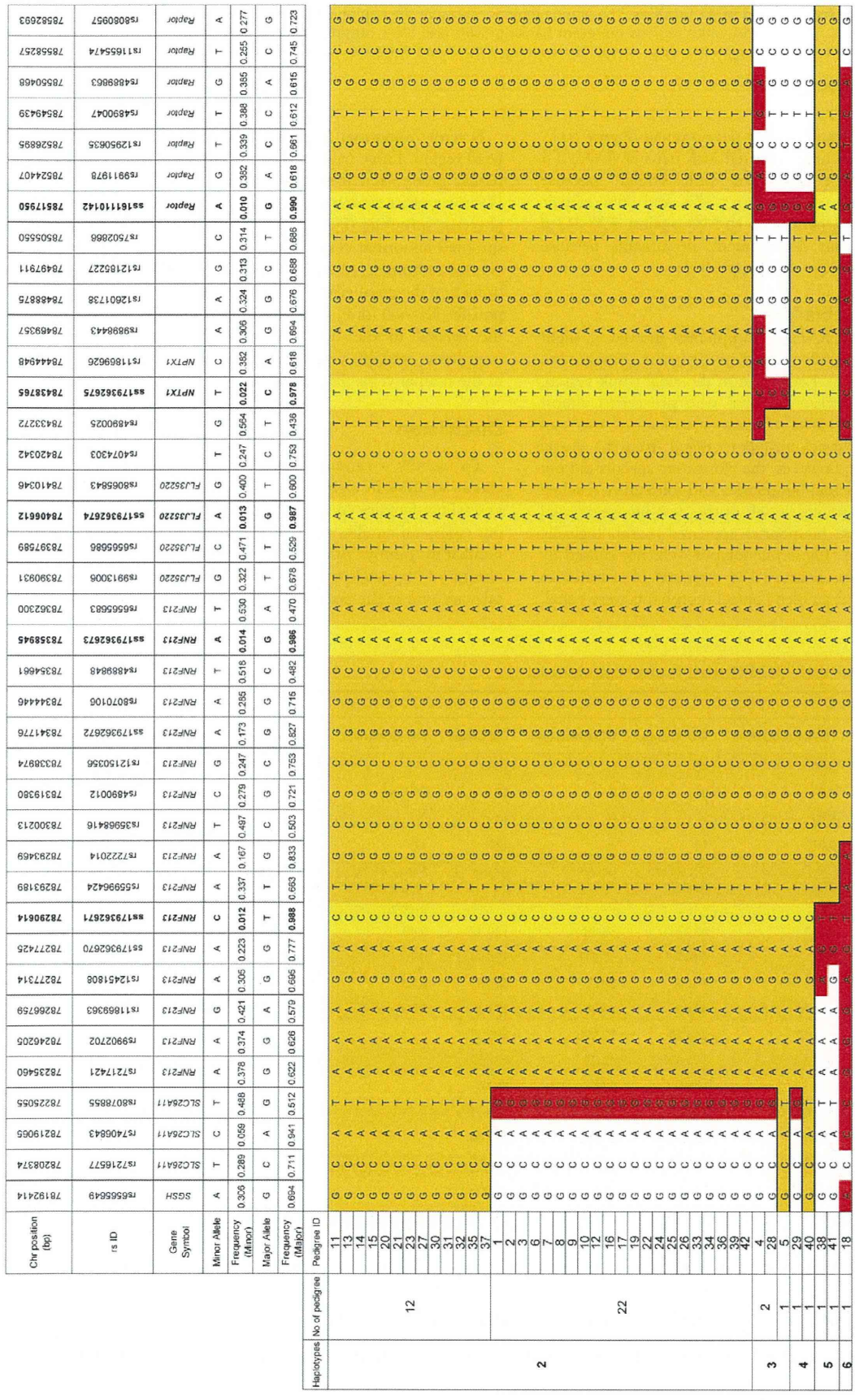
Allele-specific mRNA expression of *RNF213*

To determine whether *RNF213* mRNA from the p.R4810K (G>A) allele is specifically expressed in moyamoya patients, a SNaPshot assay was performed (Figure S9). The allele specific ratios of p.R4810K were between 1.03 and 1.19, suggesting that there is no allele-specific gene expression in moyamoya patients.

RNF213 orthologues in zebrafish

We could not prove detrimental effects of the variants. Therefore we aimed to obtain further insight into the physiological function of *RNF213* by suppressing its gene expression in zebrafish. In zebrafish, two *RNF213* genes, *RNF213-α* and *RNF213-β* are located on different chromosomes as a result of whole-genome duplication. The predicted amino acid sequences indicate that both *RNF213* genes are human *RNF213* orthologues (UCSC Genome Browser in Appendix S1), with nearly perfect conservation in the Walker and RING finger motifs. These genes are highly related to one another at the amino acid level, with *RNF213-α* and *RNF213-β* sharing similar exon structures. However, comparisons of exon-intron boundaries reveal low conservation in these *RNF213* MO target sequences. Therefore we could design a MO that specifically knocks down each *RNF213* gene. Expression analysis by RT-PCR indicated that *RNF213-α* was expressed to a greater extent than *RNF213-β* (Figure S10). *In situ* hybridization analysis was not successful, probably because of weak or scattered expression of each *RNF213*.

To knock down these *RNF213* orthologues, MO nucleotides were designed to specifically target splice sites (sp-MO). Two pairs of *RNF213-α* sp-MO and a *RNF213-β* sp-MO successfully interfered with the splicing of *RNF213* transcripts (Figure S10). To assay vascular development, we used the Tg(fli-EGFP)y1 line, in which endothelial cells are marked by EGFP expression. Injection of both sp-MO pairs against *RNF213-α* (pair 1: *RNF213-α*-MO1-A and MO1-D, pair 2: *RNF213-α*-MO2-A and MO2-D)



Chr position (bp)	rs ID	Gene Symbol	Minor Allele	Frequency (Minor)	Major Allele	Frequency (Major)
78192414	rs6556549	*SGSH*	A	0.306	G	0.694
78208374	rs7216577	*SLC26A11*	T	0.289	C	0.711
78219065	rs7406843	*SLC26A11*	C	0.059	A	0.941
78225055	rs8078855	*SLC26A11*	T	0.488	G	0.512
78235460	rs7217421	*RNF213*	A	0.378	G	0.622
78246205	rs9902702	*RNF213*	A	0.374	G	0.626
78266759	rs11869363	*RNF213*	G	0.421	A	0.579
78277314	rs12451808	*RNF213*	A	0.305	G	0.696
78277425	ss179362670	*RNF213*	A	0.223	G	0.777
78290614	**ss179362671**	***RNF213***	**C**	**0.012**	**T**	**0.988**
78293189	rs55996424	*RNF213*	A	0.337	T	0.663
78293469	rs7222014	*RNF213*	A	0.167	G	0.833
78300213	rs35968416	*RNF213*	T	0.497	C	0.503
78319380	rs4890012	*RNF213*	C	0.279	G	0.721
78338974	rs12150356	*RNF213*	G	0.247	C	0.753
78341776	ss179362672	*RNF213*	A	0.173	G	0.827
78344446	rs8070106	*RNF213*	A	0.285	G	0.715
78354661	rs4889848	*RNF213*	T	0.518	C	0.482
78358945	**ss179362673**	***RNF213***	**A**	**0.014**	**G**	**0.986**
78362300	rs6556583	*RNF213*	T	0.530	A	0.470
78390931	rs9913006	*FLJ35220*	G	0.322	T	0.678
78397589	rs6556586	*FLJ35220*	C	0.471	T	0.529
78406612	**ss179362674**	***FLJ35220***	**A**	**0.013**	**G**	**0.987**
78410346	rs8058643	*FLJ35220*	G	0.400	T	0.600
78420342	rs4074303		T	0.247	C	0.753
78433272	rs4889025		G	0.564	T	0.436
78438765	**ss179362675**	***NPTX1***	**T**	**0.022**	**C**	**0.978**
78444948	rs11869626	*NPTX1*	C	0.382	A	0.618
78469357	rs9898443		C	0.306	G	0.694
78488875	rs12801738		A	0.324	G	0.676
78497911	rs12185227		G	0.313	C	0.688
78505550	rs7502866		C	0.314	T	0.686
78517950	**ss161110142**	***Raptor***	**A**	**0.010**	**G**	**0.990**
78524407	rs9911978	*Raptor*	G	0.382	A	0.618
78526895	rs12950635	*Raptor*	T	0.339	C	0.661
78549439	rs4890047	*Raptor*	T	0.388	C	0.612
78550468	rs4889863	*Raptor*	G	0.385	A	0.615
78558257	rs11655474	*Raptor*	T	0.255	C	0.745
78582693	rs8080957	*Raptor*	A	0.277	G	0.723

Figure 5. Risk haplotypes transmitted in 42 families. The orange regions represent the haplotype of the index case of pedigree 11. The yellow regions indicate rare variants. The red and white regions represent flanking SNPs and SNPs outside of the founder haplotype, respectively. The minimum founder haplotype fell in a a region in a span of 130 kb covering *RNF213* and *FLJ35220*. The physical positions were referred from Build 37.1.
doi:10.1371/journal.pone.0022542.g005

elicited very similar abnormal vascular development (Figure S11). About three quarters of embryos injected with pairs of *RNF213-α* morpholinos (2.5 ng each) presented with similar vascular anomalies by 72 hpf (73% of pair 1, n = 177, and 76% of pair 2, n = 43). By contrast, injection of a sp-MO against *RNF213-β* resulted in normal vascular development, presumably because little *RNF213-β* is expressed *in vivo* (Figure S11).

RNF213 knock-down zebrafish

In bright-field images, *RNF213* morphants showed a slight reduction in body size, a small eye, and a wavy trunk compared to the control (Figure 9A, left). Formation of the axial trunk vessels, the dorsal aorta and posterior cardinal vein proceeded almost normally, indicating that arteriovenous specification was not affected. Intersegmental vessel sprouts emerged from the dorsal aorta at regular positions; however, the elongating sprouts did not track closely to intersegmental boundaries and sometimes reached dorsal longitudinal anastomotic vessels at the next intersegmental boundary, although somite boundaries appeared morphologically normal (Figure 9A, right). These phenotypes were not caused by general embryonic delay, because the number of somites in *RNF213* morphants and scrambled control morphants were equal.

Severely abnormal sprouting vessels were further seen in the head region. Large trunk arteries, such as the lateral dorsal aorta, basilar artery, and primordial channel formed almost normally. However, these vessels were of an irregular diameter and showed aberrant sprouting (Figure 9B). Remarkably, *RNF213* morphants sprouted abnormal vessels from the optic vessels at 60-72 hpf. In control morphants, the inner optic circle (IOC) was formed by a branch of the nasal ciliary artery; even at later stages, up to 7 days post-fertilization (dpf), three branches connected the lateral side of the eyeball to the IOC. However, in two different *RNF213* morphants, multiple vessels sprouted from the IOC and connected to the cranial veins (Figure 9C, red arrow, Figure S11).

Discussion

To identify a causative gene for moyamoya disease, we used exome analysis in eight index cases who met the RCMJ criteria and whose families comprised three generations of patients. Thus, our cases and families can safely be assumed to be prototypical for familial moyamoya disease. For filtering we employed a modified version of Ng *et al.*'s criteria [29]. This modification was made to take account of the possible involvement of uncharacterized genes

Table 2. Association of ss179362673 with moyamoya disease.

Ethnicity	HWE P Cases	HWE P Controls	DD Cases	DD Controls	Dd Cases	Dd Controls	dd Cases	dd Controls	Sample Size
Japanese	0.00	0.00	10	1	135	9	16	374	545
Korean	0.00	0.84	0	0	30	6	8	217	261
Chinese	0.81	0.92	1	0	11	2	40	98	152
Total	0.00	0.01	11	1	176	17	64	689	958
Model	Ethnicity	Chi-Squared ($-\log_{10}P$)	OR (Minor Allele)	Lower CI	Upper CI	Minor Allele Frequency (Cases)	Minor Allele Frequency (Controls)		
Allelic	Japanese	84.63	63.87	33.88	120.42	0.48	0.01		
	Korean	33.11	47.83	18.91	120.93	0.40	0.01		
	Chinese	4.95	14.14	3.13	63.98	0.13	0.01		
	Total	117.73	47.82	29.39	77.81	0.39	0.01		
Additive	Japanese	100.20	244.58	113.98	525.32				
	Korean	26.12	135.63	43.03	427.52				
	Chinese	4.64	13.69	2.86	65.56				
	Total	117.66	97.11	56.15	168.01				
Dominant	Japanese	99.98	338.94	147.82	777.44				
	Korean	26.12	135.63	43.03	427.52				
	Chinese	4.58	14.70	3.05	70.81				
	Total	119.18	111.84	64.01	195.39				
Recessive	Japanese	4.84	25.36	3.09	208.30				
	Korean	0.92	0.36	0.08	1.61				
	Chinese	0.10	0.88	0.33	2.36				
	Total	5.91	32.36	3.99	262.70				

OR: odds ratio; CI: 95% Confidence Interval; HWE: Hardy Weinberg Equilibrium; DD: minor allele homozygote; Dd: heterozygote; dd: major allele homozygote; SE: Standard error.
doi:10.1371/journal.pone.0022542.t002

Table 3. Summary of the variants in *RNF213*.

	Variant	78358945	78360097	78360619	78363034	78367154	78367201	78341560	78341825	78343331	78355371	Total
		G>A	G>A	G>C	C>T	C>G	A>G	A>G	G>A	G>A	C>T	
	Effect*	p.R4810K	p.D4863N	p.E4950D	p.A5021V	p.D5160E	p.E5176G	p.N3962D	p.D4013N	p.R4062Q	p.P4608S	
Japanese	Case (n = 161)	145 (10)	0	0	0	0	0	0	0	0	0	145
	MAF	48.1	0	0	0	0	0	0	0	0	0	48.1
	Control (n = 384)	10 (1)	0	0	0	0	0	ND	ND	ND	ND	10
	MAF	1.4	0	0	0	0	0	ND	ND	ND	ND	1.4
Korean	Case (n = 38)	30 (0)	0	0	0	0	0	0	0	0	0	30
	MAF	39.5	0	0	0	0	0	0	0	0	0	39.5
	Control (n = 223)	6 (0)	0	0	0	0	0	ND	ND	ND	ND	6
	MAF	1.3	0	0	0	0	0	ND	ND	ND	ND	1.3
Chinese	Case (n = 52)	12 (1)	1	2	2	1	1	0	0	0	0	19
	MAF	12.5	1	1.9	1.9	1	1	0	0	0	0	19.3
	Control (n = 150)	2 (0)	0	0	0	0	0	ND	ND	ND	ND	2
	MAF	0.7	0	0	0	0	0	ND	ND	ND	ND	0.7
Total in East Asian	Case (n = 251)	187	1	2	2	1	1	0	0	0	0	194
	%	74.5	0.4	0.8	0.8	0.4	0.4	0	0	0	0	77.3
	Control (n = 757)	18	0	0	0	0	0	ND	ND	ND	ND	18
	%	2.4	0.0	0.0	0.0	0.0	0.0	ND	ND	ND	ND	2.4
Czech	Case (n = 8)	0	0	0	0	0	0	0	1	0	0	1
	MAF	0	0	0	0	0	0	0	6.3	0	0	6.3
	Control (n = 120)	0	ND	ND	ND	ND	ND	0	0	0	0	0
	MAF	0	ND	ND	ND	ND	ND	0	0	0	0	0
German	Case (n = 42)	0	0	0	0	0	0	1	0	1	1	3
	MAF	0	0	0	0	0	0	1.2	0	1.2	1.2	3.6
	Control (n = 164)	0	ND	ND	ND	ND	ND	0	0	0	0	0
	MAF	0	ND	ND	ND	ND	ND	0	0	0	0	0
Caucasian	Control (n = 100)	0	ND	ND	ND	ND	ND	0	0	0	0	0
	MAF	0	ND	ND	ND	ND	ND	0	0	0	0	0
Total in Caucasian	Case (n = 50)	0	0	0	0	0	0	1	1	1	1	4
	%	0	0	0	0	0	0	2	2	2	2	8
	Control (n = 384)	0	ND	ND	ND	ND	ND	0	0	0	0	0
	%	0	ND	ND	ND	ND	ND	0	0	0	0	0

*Based on AB537889.
ND: Not determined.
() :Number of homozygotes.
MAF: Minor allele frequency.
doi:10.1371/journal.pone.0022542.t003

or the susceptibility gene. After applying this filtering process, we obtained two candidate genes, *PCMTD1* and *RNF213*. The former was discarded because it was not replicated by Sanger method in any of the eight index cases. However, p.R4810K in the latter showed complete disease segregation in all 42 families, confirming the previously identified genetic locus 17q25.3 [19]–[21].

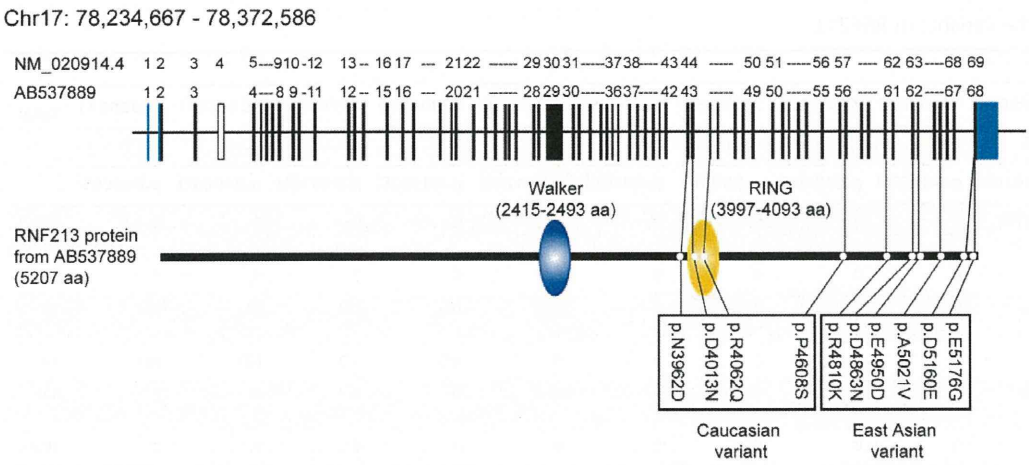


Figure 6. Genomic structure, domains of *RNF213* and variants. Genomic structure was based on DDBJ/EMBL-Bank/GenBank accession number AB537889. Domain structure was obtained by GeneCards.
doi:10.1371/journal.pone.0022542.g006

Furthermore, we showed that p.R4810K was strongly associated with moyamoya disease (OR = 111.8) in all East Asian cases. Finally, we searched for variants in *RNF213* in non-p.R4810K East Asian cases and in Caucasian cases. We found five novel variants in seven Chinese cases and four variants in four Caucasian cases.

Additionally, we identified a founder haplotype and its decayed haplotypes in 42 families. Homogeneity of the haplotypes was confirmed by genotyping of 39 SNPs. These observations strongly suggest that the causative variant should be located within the core LD block region spanning from *RNF213* to *NPTX1*. However, complete sequencing of this region including promoter, intron and intergenic regions in one of the index cases did not show any variants other than p.R4810K or a G>A substitution in intron 11 of *FLJ35220*. The latter variant did not have any effect on splicing or gene expression levels. Furthermore we failed to detect any small Indel in exons in the eight index cases and large CNVs in three index cases in 17q25.3, although we missed small intronic structural abnormalities, which might be a possible cause of moyamoya disease. Taken together, these lines of evidence consistently support a conclusion that *RNF213* is a susceptibility gene for moyamoya disease. Recently a significant association of a SNP in *RNF213* with moyamoya disease in Japanese has been shown [33].

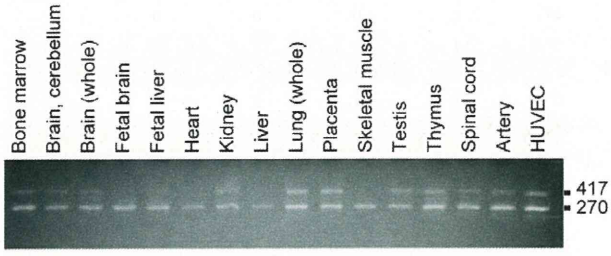


Figure 7. Two isoforms of alternative splicing variants of *RNF213*. We have tested whether exon 4 is read through or not in cDNA isolated from various human tissues and HUVECs. Representative results of human tissue RNAs and HUVEC are shown. A short isoform, which skips the exon 4, has an expected size of 270 bp (AB537889) and a long, which reads exon 4, has an expected size of 417 bp (NM_020914.4).
doi:10.1371/journal.pone.0022542.g007

As we previously reported, the penetrance of autosomal dominant moyamoya disease is low, as illustrated by discordant identical twins or "skipping a generation" [34]. This characteristic of moyamoya disease is in accordance with the notion that *RNF213* is a susceptibility gene for moyamoya disease and reconciles our unexpected observation that this variant was found at an allele frequency of 1% in the Japanese, Korean and Chinese control populations. In addition, the low identification rates of *RNF213* variants in Caucasians compared with East Asians strongly suggests that there is genetic heterogeneity of moyamoya disease between these two populations. In addition, there are obvious differences in the variants themselves between East Asians and Caucasians. Further studies are needed to investigate whether such allelic differences result in the different clinical features.

By cloning the cDNA for this region, we discovered a novel splicing variant for *RNF213* (AB537889), which lacks exon 4 of *RNF213* [NM_020914.4]. This new splicing variant of *RNF213* is a major transcript of *RNF213*. We have experimentally proven that two functional domains, a Walker motif and a RING finger domain, function. Such novel features of *RNF213* hampered our elucidation of its physiological function.

To characterize the physiological role of *RNF213 in vivo*, we investigated the effects of *RNF213* suppression on zebrafish vasculature. *RNF213* knockdown zebrafish showed severely abnormal sprouting vessels in the head region, especially from the optic vessels at 60–72 hpf. In normal embryos, the overall wiring pattern of the major vessels is largely completed by 2–2.5 dpf, despite massive development of smaller-caliber vessels in the head through to 7 dpf [35]. Therefore, the aberrant vessel formation from the IOC seen in *RNF213* morphants at this stage implies a severely impaired program of angiogenesis in the head region. Although abnormal vascular phenotypes in the tail region have been reported after manipulation of genes such as plexin D1 [36], morphants showing aberrant vascular sprouting from the IOC have never been seen [37], suggesting that *RNF213* is involved in a novel signaling pathway in intracranial angiogenesis. As such, we expect that once we identify the pathway, we may be able to elucidate the consequences of vascular remodeling in moyamoya disease [9], [12], [23], [38]. We are aware of a limitation in the current study as a morpholino rescue experiment was not conducted. The rescue experiment is technically difficult to achieve because of the extremely large size of the *RNF213*

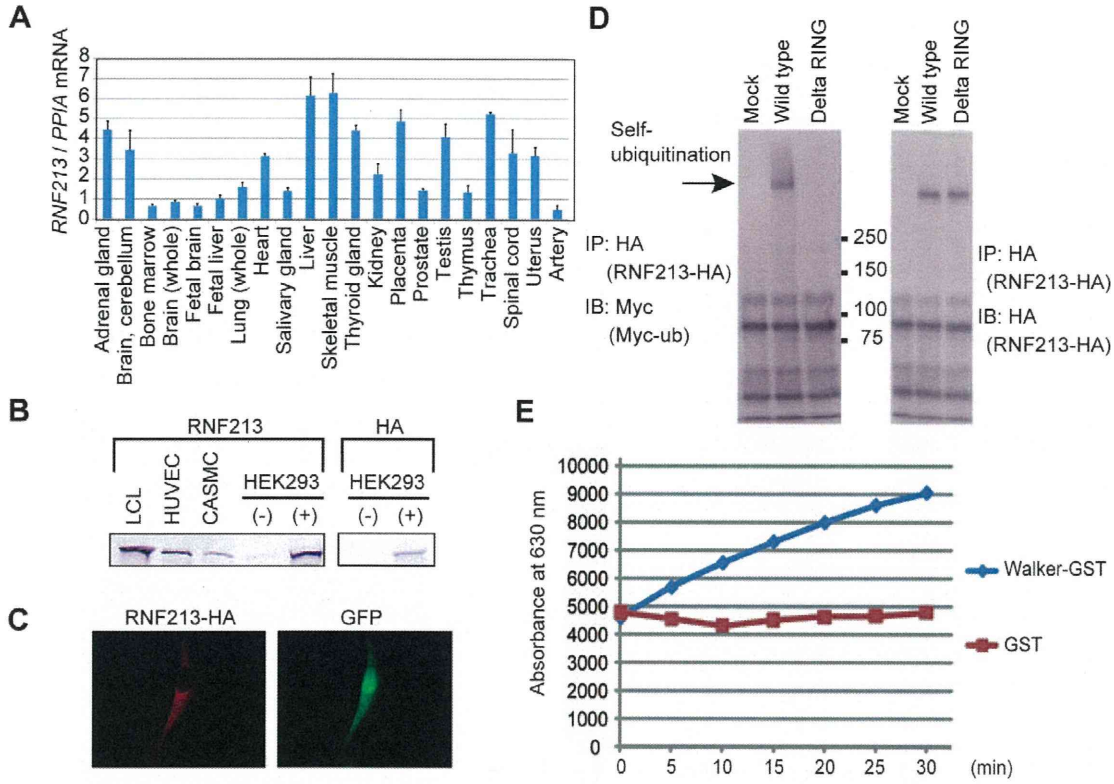


Figure 8. Characterization of the *RNF213* gene and protein. (**A**) *RNF213* mRNA expression. Total RNA from the indicated human tissues was reverse-transcribed to cDNA, and real-time quantitative PCR was performed. (**B**) RNF213 protein expression. LCL, HUVEC, and CASMC were lysed and immunoblotted using an anti-RNF213 antibody. HEK293 cells transiently expressing RNF213-HA (+) or control cells (−) were immunoblotted using anti-RNF213 and anti-HA antibodies. (**C**) Subcellular localization of RNF213. HeLa cells transiently expressing RNF213-HA were stained with an anti-HA antibody. (**D**) Self-ubiquitination of RNF213. HEK293 cells transiently expressing RNF213-HA and Myc-ubiquitin (Myc-ub) were lysed and subjected to immunoprecipitation (IP) using an anti-HA antibody, followed by immunoblotting (IB) using an anti-Myc antibody. As a control, immunoblotting was also performed with an anti-HA antibody. (**E**) ATPase activity of RNF213. Free phosphate released from ATP by the ATPase activity of a recombinant RNF213 fragment (a.a. 2359–2613) tagged with GST was measured using the Malachite Green method. a.a., amino acid. IP, immunoprecipitation. IB, immunoblot.
doi:10.1371/journal.pone.0022542.g008

mRNA. However, the similar and clearly defined vascular phenotypes shown by two different *RNF213* morphants indicate that these phenotypes were derived from the specific deletion of *RNF213* and not from an off-target effect of a morpholino.

A question remains as to how the p.R4810K variant or other nine variants may impair the physiological function of *RNF213* thereby resulting in moyamoya disease. We could not demonstrate that p.R4810K or p.D4013N affects ubiquitin ligase activity or causes other hallmark changes such as mRNA or protein instability. Furthermore, it appears that there was no gene dosage effect because some Japanese controls were homozygous (Table 2). In addition, homology search argued against pathological roles for some variants. It is thus probable that these variants likely perturb RNF213 function by unidentified mechanisms other than through ubiquitin ligase activity or by decreasing protein stability or by mislocalization. At present, we cannot specify the mechanism as haploinsufficiency, dominant negative or gain of function; instead we postulate the involvement of another factor. Requirement of co-factors with *RNF213* might also explain the lower disease prevalence (1 in 10,000) than expected from the relatively high minor allele frequency of approximately 1%. Furthermore, despite similar frequencies of p.R4810K, population attributable risks were different among the three East Asian populations, suggesting existence of environmental factors or another genetic factor. It should be noted, however, that involvement of the additional genetic factor is unlikely because the gene, *RNF213*, was singularly filtered as the common genetic factor by exome. We consider it rational to postulate involvement of environmental factors rather than genetic factors. Several environmental factors may include autoimmune conditions [39], [40], infection [41]–[45] and exposure to radiation [46]. Recently, Bauersachs *et al.* reported that *RNF213* is upregulated in bovine endometrium by pregnancy [47]. In addition, Zhang *et al.* reported that intracellular parasites and cytokines upregulate *RNF213* [48]. These findings support that *RNF213* is upregulated by inflammatory signals generated by interferon or cytokines. Of particular interest is transforming growth factor beta (TGF-β), because it is known to be elevated in the cerebrospinal fluid, blood and arteries of patients with moyamoya disease [23], [38] and a SNP within TGF-β has been reported to be associated with moyamoya disease [49]. Further study on cross-talk of *RNF213* with such conditions that cause endothelial damage, or angiogenesis factors is needed to understand the molecular mechanisms.

Sharing p.R4810K among moyamoya cases urged us to discuss the anthropological history of the founder haplotype carrying p.R4810K. p.R4810K appears to be a neutral variant or an advantageous variant for human survival because it has been maintained in the East Asian population. It is interesting that

Table 4. Homology of *RNF213*.

East Asian variant									
Species		Gene	Accession number	R4810K	D4863N	E4950D	A5021V	D5160E	E5176G
Homo sapiens	Human	*RNF213*		4803 QVEYSSI**R** GFLSKHS 4817	4856 CSTDLDL**D** TEFEILL 4870	4943 EGRETVQ**E** FDLEKIQ 4957	5014 QLQSYSD**A** CEVLSVV 5028	5153 RPQWSLR**D** TLVSYMQ 5167	5169 KESEILP**E** MASQFPE 5183
Homo sapiens	Human	*RNF213*	NP_065965.4	4852 QVEYSSI**R** GFLSKHS 4866	4905 CSTDLDL**D** TEFEILL 4919	4992 EGRETVQ**E** FDLEKIQ 5006	5063 QLQSYSD**A** CEVLSVV 5077	5202 RPQWSLR**D** TLVSYMQ 5216	5218 KESEILP**E** MASQFPE 5232
Pan troglodytes	Chimpanzee	*RNF213*	XP_511726.2	3745 QVEYSSI**R** GFLSKHS 3759	3798 CSTDLDL**D** TEFEILL 3812	3885 EGRETVQ**E** FDLEKIQ 3899	3956 QLQSYSD**A** CEVLSVV 3970	4095 RPQWSLR**D** TLVSYMQ 4109	4111 KESEILP**E** MASQFPE 4125
Mus musculus	Mouse	mCG142721, isoform CRA_a	EDL34702.1	4828 VEYSSI**R** GFIHSHS 4841	4880 CCSDLDL**D** AEFEVIL 4894	4967 QGGETSQ**E** FDLEKIQ 4981	5038 QLQSYSD**A** CEALSII 5052	5175 NPNWSLK**D** TLVSYME 5189	5191 KDSDILS**E** VESQFPE 5205
Rattus norvegicus	Rat	*RNF213*	XP_001081768.2	4828 EVEFSSI**R** SFIHSHHS 4843	4881 CRSDLDL**D** AKFEVIL 4895	4968 QGGETSQ**E** FDLEKIQ 4982	5039 QLQSYSD**A** CEALSIV 5053	5176 NPSWSLK**D** TLVSYME 5190	5192 KDSDVLT**E** VESQFPD 5206
Bos taurus	Cow	similar to *mCG142721*, partial	XP_590465.5	4681 EVEYKSI**R** SFISSH 4694	4734 CSSDLDL**D** TDLEVIL 4748	4821 QGGETLQ**E** FDLEKIQ 4835	4892 QLQSYSD**A** CEALSAT 4906	5029 NPEWSLR**D** TLVSYME 5043	5045 TDSEIPP**E** MESQFPE 5059
Canis familiaris	Dog	*RNF213*	XP_540474.2	2986 EADYQSI**R** SFISSHQ 3000	3039 CSADLDM**D** TNFEVIL 3053	3126 QGKETLQ**E** FDLEKIQ 3140	3197 QLPSYSD*G* CKALSVI 3211	3334 RPEWSLR**D** TLVSYME 3348	3350 KDSEIPP**E** LEYQFPE 3364
Monodelphis domestica	Kangaroo	hypothetical protein LOC100030710	XP_001380151.1	4718 VEYNTI**R** GFL 4727	4770 CDADLSL*E* NEFEILL 4784	4857 KGRETLQ*Q* FDLEKIQ 4871	4928 QLQSYSD**A** CEALSVT 4942	5066 NPKWSLK*E* TLVSYME 5080	5082 KESEIPP**E** VEYQFPE 5096
Taeniopygia guttata	Zebra finch	*RNF213*	XP_002192487.1	4254 EIKHCSI**R** EFLREPH 4268	4307 CDAELSL*E* SRLEVLL 4321	4394 KGGETLQ*D* FDLERIQ 4408	4465 ELQSYSD*V* CDALSLT 4479	4603 KSTWSLK*E* SLLPYLY 4617	4618 KDSELTL**E** LEDTFPD 4632
Danio rerio	Zebrafish	*RNF213*	XP_001921030.2	3888 DLTYKTI**R** EFLQDQK 3902	3941 TEKDLGL**D** ADLQVLL 3955	4028 CGQETLL**E** YDLPKIQ 4042	4099 ELQSYSD*V* CEALSTV 4113	4237 RPDWRLK*H* TVVSYME 4251	4253 KDLDVPP**E** VEEFFPK 4267
Takifugu rubripes	Fugu	*C17of27*	AAL32171.1	3881 AVQTGTI**R** EFLNTQN 3895	3934 CQSDMDH*S* SDFSFLL 3948	4021 RGQESLL**E** YDLAKLQ 4035	4092 ELQSYSE*V* CEALSTL 4106	4229 KPEWSLA*V* TLFSYME 4243	4245 KDLDVSP**E** M-EEFPE 4259
Tetraodon nigroviridis	Spotted green pufferfish	unnamed protein product	CAG00202.1	892 ELS-ATIEEF LNTQN 905	944 CQSDMDL*S* SDFRVLL 958	1031 KGQETLP**E** YDLAKIQ 1045	1102 ELQSYSE*V* CEALSTL 1116	1242 TGASRTR*S* PLTWSAK 1256	1242 TGASRTR*S* PLTWSAK 1256
Caucasian variant									
Species		Gene	Accession number	N3962D	D4013N	R4062Q	P4608S		
Homo sapiens	Human	*RNF213*		3955 LDKCLRE**N** SDVKTHG 3969	4006 DPVCLPC**D** HVHCLRC 4020	4055 IEKHARF**R** QMCNSFF 4069	4601 LINIIKP**P** VRDPKGF 4615		
Homo sapiens	Human	*RNF213*	NP_065965.4	4004 LDKCLRE**N** SDVKTHG 4018	4055 DPVCLPC**D** HVHCLRC 4069	4104 IEKHARF**R** QMCNSFF 4118	4650 LINIIKP**P** VRDPKGF 4664		
Pan troglodytes	Chimpanzee	*RNF213*	XP_511726.2	2901 LDKCLRE**N** SDVKTHG 2915	2952 DPVCLPC**D** HVHCLRC 2966	3001 IEKHARF**R** QMCNSFF 3015	3543 LINIIKP**P** ARDPKGF 3557		
Mus musculus	Mouse	mCG142721, isoform CRA_a	EDL34702.1	3959 LDKCLEE*D* SNLKT 3971	4010 DPVCLPC**D** HVYCLRC 4024	4059 IEKHAQF**R** HMCNSFF 4073	4629 LTVIIKP*W* VQDPQGF 4643		
Rattus norvegicus	Rat	*RNF213*	XP_001081768.2	3959 LDKCLEE*D* SNLKT 3971	4010 DPVCLPC**D** HVYCLPC 4024	4059 IEKHAQF**R** HMCNSFF 4073	4630 LMNIIKP**P** VQDPQGF 4644		
Bos taurus	Cow	similar to *mCG142721*, partial	XP_590465.5	3837 LNKCLLE*D* SDTKTH 3850	3888 DPVCLPC**D** HIFCLRC 3902	3937 IEKHARF**R** QMCNSFF 3951	4483 LMNIIKP**P** VSDPKRF 4497		
Canis familiaris	Dog	*RNF213*	XP_540474.2	2142 LNKYLQD*D* SDIKTYRP 2157	2193 EPVSLPC*G* HVFCLRC 2207	2242 IRKHACL**R** QMCNSFF 2256	2784 VMNIIKP**P** VRDPSSF 2798		

Table 4. Cont.

East Asian variant								
Monodelphis domestica	Kangaroo	hypothetical protein LOC100030710	XP_001380151.1	3868 LGKCLQD**N** SDIKTH 3881	3919 EPVCLPC**D** HVYCQKC 3933	3968 IAKHIQF**R** QMCNSFF 3982	4515 LMKIIKP**P** VRDPESF 4529	
Taeniopygia guttata	Zebra finch	*RNF213*	XP_002192487.1	3441 LAKCFQL*D* SDMKSHP 3455	3492 DPICLPC*N* HVFCHKC 3506	3540 IAKKALF**R** QRCNNFF 3554	4053 LGSMIKP*T* VKNVVSF 4067	
Danio rerio	Zebrafish	*RNF213*	XP_001921030.2	3046 LAQVLEQ*D* SNLKKKK 3060	3097 DPLSLPC**D** HIYCLTC 3111	3146 ISQNASF**R** MRCNAFF 3160	3687 IMQIIKP*A* VVHPDAF 3701	
Takifugu rubripes	Fugu	*C17orf27*	AAL32171.1	3040 LGQILEK**N** SDLKTYE 3054	3091 DPLCLPC**D** HIYCQAC 3105	3140 VNQHARF**R** KQCNAFF 3154	3681 GSAIIKP*V* VHDPGAF 3695	
Tetraodon nigroviridis	Spotted green pufferfish	unnamed protein product	CAG00202.1	65 LGQILEKSSD LKSHQ 79	116 EPLSLPC**D** HIYCLGC 130	165 VNQHAQF**R** KRCNAFF 179	668 VSAIIKPQ VSDPGVF 682	

Orthologue genes were searched by BLAST.
doi:10.1371/journal.pone.0022542.t004

p.R4810K was not found in Caucasians; the relatively high prevalence of p.R4810K among East Asians could account for the higher prevalence of moyamoya disease in East Asians than in Caucasians.

This study has several limitations. First, we cannot provide evidence regarding impairment of physiological function of these variants of *RNF213* in moyamoya disease. Second, the number of subjects used for deep sequencing or CNVs was small. The strengths of our study are that we conducted whole genome-exome analysis and demonstrated strong evidence showing involvement of a single gene, *RFN213*, in moyamoya disease. Other strengths include that we found a founder variant in East Asian cases, an additional nine variants in Chinese and Caucasian cases, and characterized the RNF213 protein biochemically and physiologically. Although further studies are necessary to clarify the biochemical function and pathological role of *RNF213* in moyamoya disease, the discoveries of its association with the disease and its unique roles in angiogenesis may pave a way to early diagnosis and prevention. It should be noted, however, that the majority of the pathological proof awaits further studies.

Supporting Information

Text S1 Supplemental methods, reference and acknowledgments.
(DOC)

Appendix S1 Web Resources.
(DOC)

Figure S1 Pedigree chart. Forty-one Japanese families and one Korean (pedigree 40) family participated in this study. The phenotypes of occlusive lesions are shown. Genotype of p.R4810K (G>A) is shown in brackets.
(TIF)

Figure S2 A variant in a Caucasian family. The index case is the father (CAU_ped1_12), who was born in 1966 and suffered a mild ischemic stroke at the age of 30. His mother died of an ischemic stroke at the age of 35. He has four children from two marriages. The second child from the first marriage (CAU_Ped1_122) was born in 1991 and developed moyamoya disease at the age of 5. The first child of the second marriage (CAU_Ped1_123) was born in 1999 and developed moyamoya disease with symptoms of involuntary movement at the age of 9. The second child, (CAU_Ped1_124) born in 2006, developed moyamoya disease with manifestations of ischemic stroke at the age of 3. Diagnoses of moyamoya disease were made by magnetic resonance imaging (MRI). (A) Caucasian pedigree. The index case (father) and his three affected children carry the p.D4013N variant of *RNF213* (G>A). Genotypes of the variant for each member are shown in the bracket in the pedigree. (B) Sequencing analysis. (C) Genotyping by HpyCH4V.
(TIF)

Figure S3 Effect of G>A substitution in intron 11 of *FLJ35220* on splicing or gene expression. (A) We tested whether exon 11 was read through. A short form, which skips exon 11, had an expected size of 107 bp. A long form, which reads through exon 11, had an expected size of 166 bp (NM_173627.2). M, 100 bp ladder DNA marker. (B) *FLJ35220* mRNA expression in LCLs [controls; JPN1, an unaffected daughter (Ped17_123) of Ped17_12 and an unaffected spouse (Ped18_20) of Ped18_2; cases: Ped11_11, Ped17_12, Ped18_2 and Ped18_22] as determined by real-time quantitative PCR. Data are shown as means ± S.D. of three independent experiments. There is no statistically significant difference between the two groups. Significance was tested using

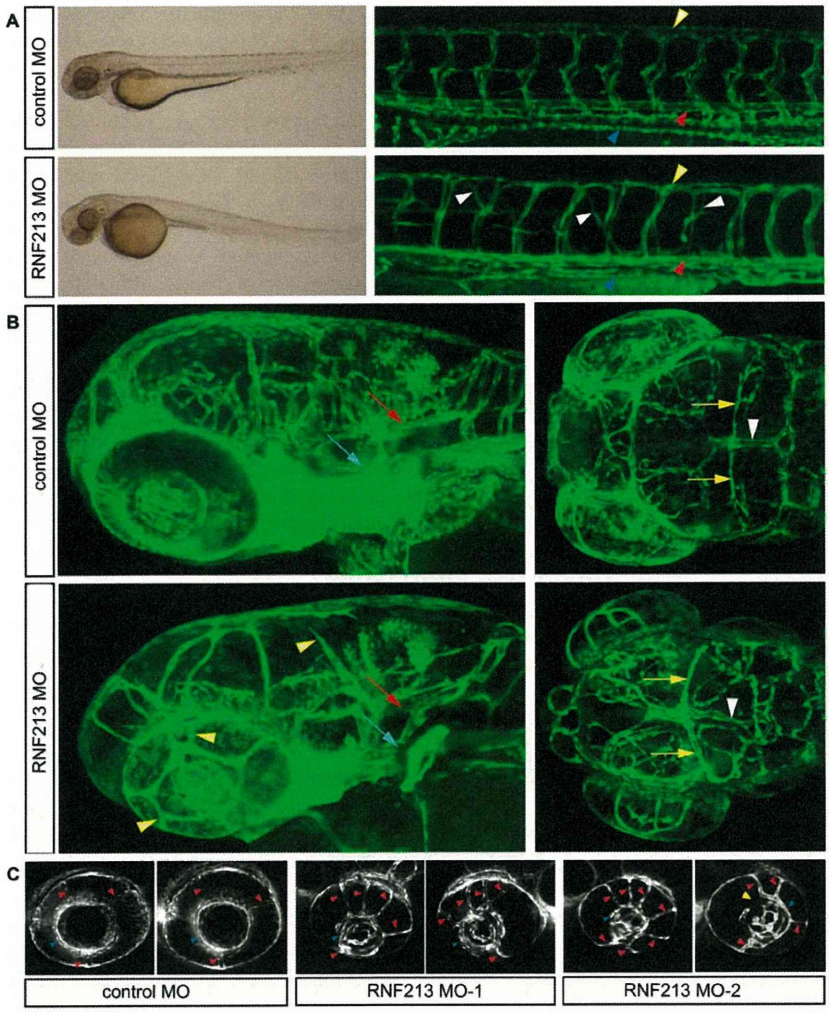


Figure 9. Depletion of *RNF213* causes abnormal vessel sprouting in zebrafish. Tg(fli-EGFP)y1 embryos at 72 h post-fertilization. (**A**) Brightfield image of whole embryos (left) and confocal images of tail vessels (right) of embryos injected with a control (scrambled) or *RNF213* morpholino (MO). Lateral views, dorsal uppermost, anterior to the left. In *RNF213* morphants, abnormal intersegmental vessel sprouting was observed (white arrowheads). Trunk vessels, including the dorsal aorta (red arrowheads), posterior cardinal vein (blue arrowheads) and dorsal longitudinal anastomotic vessels (yellow arrowheads) developed almost normally. (**B**) Confocal images of intracranial vessels. (Left) Lateral views 20 degrees toward the top; dorsal uppermost, anterior to the left. (Right) Dorsal views, head to the left. The trunk artery, including the basilar artery (red arrows), lateral dorsal aorta (blue arrows), mesencephalic artery (yellow arrows) and dorsal longitudinal vein (white arrowheads) developed almost normally in controls and *RNF213* morphants. *RNF213* morphants showed abnormal sprouting vessels (yellow arrowheads) and irregular vessel diameter (white arrowheads). (**C**) Cross-sectional view at the middle of the lens level. In a control morphant, the three branches of the nasal ciliary artery (red arrowheads) drain into the inner optic circle (IOC; blue arrowheads). In two different *RNF213* morphants, multiple aberrant vessels drain into the IOC (red arrowheads), and occasionally part of the IOC was missing (yellow arrowhead).
doi:10.1371/journal.pone.0022542.g009

Student's *t*-test. A $p<0.05$ was considered to be significant. The methods have been fully described in the Text S1.
(TIF)

Figure S4 Copy number analysis for the 1.5-Mb region in 17q25.3. Three index cases of pedigrees 5, 11 and 18 were analyzed. The blue lines represent the copy numbers (log 2 ratio) averaged over 10 SNPs. The copy numbers were compared with control spouse of 2 of pedigree 18.
(TIF)

Figure S5 Sequence chromatograms for the eight novel variants.
(TIF)

Figure S6 Cloning of full-length *RNF213* cDNA. Schematic representation of the internal connecting site of *RNF213* cDNA. RT-PCR with primers set on *RNF213* resulted in amplification of the expected fragments. First, fragments 1, 2, 4, and 5 were cloned into a pcDNA3.1+ vector. Second, fragment 3 was cloned into the vector carrying fragment 2, and fragment 5 was subcloned into the vector carrying fragment 4. Then fragments 2–3 and 4–5 were subcloned into fragment 1 using the restriction enzyme sites indicated.
(TIF)

Figure S7 Northern blotting analysis of *RNF213* mRNA. (A) *RNF213* mRNA expression in the indicated human tissues (heart, brain, liver, pancreas, skeletal muscle, and lung). Arrow indicates *RNF213* mRNA. (B) *RNF213* mRNA expression in cultured human cells (HeLa, HEK293T and LCLs (control: JPN1, Ped18_20 (a spouse of 2) and case: Ped18_2, Ped18_22,

Ped11_11)) using radiolabelled probes corresponding to *RNF213* coding regions. *GAPDH* mRNA expression is shown as a loading control. The position of RNA Millennium markers (Ambion) and positions of the 18S and 28S ribosomal RNAs are indicated on the left.
(TIF)

Figure S8 Characterization of the p.R4810K and p.D4013N allele proteins of *RNF213*. (A) Stability of the p.R4810K variant. Wild-type or p.R4810K variant of RNF213-HA were transiently expressed in HEK293 cells. Cells were lysed and subjected to immunoblotting with an anti-HA antibody. IP, immunoprecipitation. IB, immunoblot. (B) Subcellular localization of the p.R4810K variant of RNF213. HEK293 cells transiently expressing the wild-type or p.R4810K variant of RNF213-HA were fractionated into cytosol, membrane/organelle, nucleus, and cytoskeleton using different lysis reagents (ProteoExtract kit, Calbiochem). (C) Self-ubiquitination of the p.R4810K variant of *RNF213*. HEK293 cells transiently expressing the wild-type or R4810K variant of RNF213-HA and Myc-ubiquitin were lysed and subjected to immunoprecipitation using an anti-HA antibody, followed by immunoblotting using an anti-Myc antibody. (D) Self-ubiquitination of the p.D4013N mutant of *RNF213*. HEK293 cells transiently expressing the wild-type or p.D4013N mutant of RNF213-Flag and Myc-ubiquitin were lysed and subjected to immunoprecipitation using an anti-Flag antibody, followed by immunoblotting using an anti-Myc antibody.
(TIF)

Figure S9 Allele-specific mRNA expression of *RNF213* by labeling and detection of the two alleles of marker SNPs p.R4810K and p.H4557H. (A) Design of allele-specific mRNA expression using SNaPshot assay. Arrows indicate primer positions for amplification. Bold arrows indicate extension primers for each SNP. (B) Genotypes of *RNF213* gene at two SNPs (p.H4557H and p.R4810K) in cases for SNaPshot assay. (C) Allele-specific ratio of *RNF213* mRNA expression in LCLs from SNP heterozygous patients. The common allele/rare allele ratio from cDNA was normalized to that ratio from genomic DNA of the same individual. Data are shown as means ± S.D. of three independent experiments. There was no statistically significant difference between the two alleles for each SNP. Significance was tested by Student's *t*-test. A $p<0.05$ was considered to be significant.
(TIF)

Figure S10 Splicing ablation of *RNF213* transcripts by morpholino injection. RT-PCR showing the defective splicing induced by *RNF213*-α-MO1 and MO2 pairs and a *RNF213*-β-MO. Compared with the results of PCR (using primer pairs *RNF213*-α_2 and 2R, *RNF213*-β_3 and 3R, and *RNF213*-β_1 and 1R) from noninjected embryos, where a single band was generated (lanes 1, 3, and 5, marked "native"), split bands were detected in PCR using embryos injected with a *RNF213*-α-MO1, MO2 pair and a *RNF213*-β-MO (lanes 2, 4, and 6, marked "exon blocked"). M, 1-kb ladder DNA marker. Band indicated by 'genomic' is the amplified genomic sequence. Equal amounts of PCR products and marker were loaded in each lane. Intensity of RT-PCR products indicates that *RNF213*-α is dominantly expressed in vivo.
(TIF)

Figure S11 *RNF213* morphants show multiple sprouting vessels from IOC. Bar graphs showing the number of sprouting vessels from IOC of Tg(fli-EGFP)y1 embryos 72 h post-fertilization. Each group was injected with 2.5 ng morpholinos (MO1, MO2) per embryo indicated in each lane. $n = 20$ per group. Values are means ± SD. $*p<0.01$ versus control scramble morphants. $\#p<0.01$ versus *RNF213*-β-morphants. Neither group of controls nor *RNF213*-β showed any extra sprouting vessels.
(TIF)

Table S1 Demographic feature of familial participants.
(XLS)

Table S2 Summary of demographic and clinical profiles of cases and controls.
(XLS)

Table S3 Primers used for amplification of *RNF213* and *PCMTD1* (Human build 37.1).
(XLS)

Table S4 Positional candidate genes in the 1.5 Mb locus on 17q25.3 (Map Viewer: Build 37.1).
(XLS)

Table S5 Primer sets and restriction enzymes for screening variants in controls.
(XLS)

Table S6 Primers used for SNaPshot assay.
(XLS)

Table S7 Numbers of variants and candidate genes in the 2nd stage for various combination subsets of cases.
(XLS)

Table S8 A summary of the sequencing results for five controls by exome, 10 controls and deep sequencing in a control by the Sanger method.
(XLS)

Table S9 Association of ss179362673 with moyamoya disease with or without family histories by allelic model.
(XLS)

Acknowledgments

We thank Dr. Kayoko Inoue, who died in 2007, for her enthusiastic support. We are also grateful to Drs. Mustuko Minata and Sumiko Inoue, Ms. Michi Hirosawa and Ms. Chihiro Horii (Kyoto University Graduate School of Medicine), Prof. Akira Tsuji (Department of Neurology, The University of Tokyo) Dr. Hiroshi Nanjyo (Division of Pathology, Akita University Hospital) and Prof. Hiroshi Yorifuji and Dr. Tohru Murakami (Department of Anatomy, Gunma University Graduate School of Medicine). We thank Ms. Kazumi Kanamori, Ms. Maki Miyoshi and Ms. Mayumi Kishimoto for technical assistance with the molecular biological and zebrafish studies, respectively. We are also grateful to Dr. Fang Fang (Beijing Children's Hospital, Beijing, China), Drs. Hyun-Seung Kang, Chang Wan Oh (Seoul National University College of Medicine, Seoul, Korea) and Dr. Daniela Berg (Department of Neurology, University of Tuebingen, Germany). We thank the members of the international consortium (Supplemental Acknowledgments) for recruiting patients.

Author Contributions

Conceived and designed the experiments: AK. Performed the experiments: AK WL DM ST HK TH NM SY AT AF. Contributed reagents/materials/analysis tools: NH WL YM HH KK YT RH BK LZ JEK SM. Wrote the paper: AK NH WL DM ST YM HK TH HH SY KHH MK SM KN. Analyzed the genetic data: AK WL NM KHH AT AF. Analyzed the clinical data: NH YM HH KK YT RH BK LZ JEK MK SM. Analyzed the molecular data: DM ST HK TH SY KN. All authors have read and approved submission of the manuscript.

References

1. Takeuchi K, Shimizu K (1957) Hypogenesis of bilateral internal carotid arteries. Brain Nerve 9: 37–43.
2. Suzuki J, Takaku A (1969) Cerebrovascular "moyamoya" disease. Disease showing abnormal net-like vessels in base of brain. Arch Neurol 20: 288–299.
3. Goto Y, Yonekawa Y (1992) Worldwide distribution of moyamoya disease. Neurol Med Chir (Tokyo) 32: 883–886.
4. Kuroda S, Houkin K (2008) Moyamoya disease: current concepts and future perspectives. Lancet Neurol 7: 1056–1066.
5. Miao W, Zhao PL, Zhang YS, Liu HY, Chang Y, et al. (2010) Epidemiological and clinical features of Moyamoya disease in Nanjing, China. Clin Neurol Neurosurg 112: 199–203.
6. Amlie-Lefond C, Bernard TJ, Sebire G, Friedman NR, Heyer GL, et al. (2009) Predictors of cerebral arteriopathy in children with arterial ischemic stroke: results of the International Pediatric Stroke Study. Circulation 119: 1417–1423.
7. Fukui M, Kono S, Sueishi K, Ikezaki K (2000) Moyamoya disease. Neuropathology 20(Suppl): S61–64.
8. Takebayashi S, Matsuo K, Kaneko M (1984) Ultrastructural studies of cerebral arteries and collateral vessels in moyamoya disease. Stroke 15: 728–732.
9. Takagi Y, Kikuta K, Sadamasa N, Nozaki K, Hashimoto N (2006) Caspase-3-dependent apoptosis in middle cerebral arteries in patients with moyamoya disease. Neurosurgery 59: 894–900.
10. Kim SK, Yoo JI, Cho BK, Hong SJ, Kim YK, et al. (2003) Elevation of CRABP-I in the cerebrospinal fluid of patients with Moyamoya disease. Stroke 34: 2835–2841.
11. Malek AM, Connors S, Robertson RL, Folkman J, Scott RM (1997) Elevation of cerebrospinal fluid levels of basic fibroblast growth factor in moyamoya and central nervous system disorders. Pediatr Neurosurg 27: 182–189.
12. Nanba R, Kuroda S, Ishikawa T, Houkin K, Iwasaki Y (2004) Increased expression of hepatocyte growth factor in cerebrospinal fluid and intracranial artery in moyamoya disease. Stroke 35: 2837–2842.
13. Soriano SG, Cowan DB, Proctor MR, Scott RM (2002) Levels of soluble adhesion molecules are elevated in the cerebrospinal fluid of children with moyamoya syndrome. Neurosurgery 50: 544–549.
14. Takahashi A, Sawamura Y, Houkin K, Kamiyama H, Abe H (1993) The cerebrospinal fluid in patients with moyamoya disease (spontaneous occlusion of the circle of Willis) contains high level of basic fibroblast growth factor. Neurosci Lett 160: 214–216.
15. Yoshimoto T, Houkin K, Takahashi A, Abe H (1996) Angiogenic factors in moyamoya disease. Stroke 27: 2160–2165.
16. Ikeda H, Sasaki T, Yoshimoto T, Fukui M, Arinami T (1999) Mapping of a familial moyamoya disease gene to chromosome 3p24.2-p26. Am J Hum Genet 64: 533–537.
17. Inoue TK, Ikezaki K, Sasazuki T, Matsushima T, Fukui M (2000) Linkage analysis of moyamoya disease on chromosome 6. J Child Neurol 15: 179–182.
18. Sakurai K, Horiuchi Y, Ikeda H, Ikezaki K, Yoshimoto T, et al. (2004) A novel susceptibility locus for moyamoya disease on chromosome 8q23. J Hum Genet 49: 278–281.
19. Yamauchi T, Tada M, Houkin K, Tanaka T, Nakamura Y, et al. (2000) Linkage of familial moyamoya disease (spontaneous occlusion of the circle of Willis) to chromosome 17q25. Stroke 31: 930–935.
20. Mineharu Y, Liu W, Inoue K, Matsuura N, Inoue S, et al. (2008) Autosomal dominant moyamoya disease maps to chromosome 17q25.3. Neurology 70: 2357–2363.
21. Liu W, Hashikata H, Inoue K, Matsuura N, Mineharu Y, et al. (2010) A rare Asian founder polymorphism of Raptor may explain the high prevalence of Moyamoya disease among East Asians and its low prevalence among Caucasians. Environ Health Prev Med 15: 94–104.
22. Fukui M (1997) Guidelines for the diagnosis and treatment of spontaneous occlusion of the circle of Willis ('moyamoya' disease). Research Committee on Spontaneous Occlusion of the Circle of Willis (Moyamoya Disease) of the Ministry of Health and Welfare, Japan. Clin Neurol Neurosurg 99(Suppl 2): S238–240.
23. Hojo M, Hoshimaru M, Miyamoto S, Taki W, Nagata I, et al. (1998) Role of transforming growth factor-beta1 in the pathogenesis of moyamoya disease. J Neurosurg 89: 623–629.
24. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58: 1347–1363.
25. Yamada S, Utsunomiya M, Inoue K, Nozaki K, Miyamoto S, et al. (2003) Absence of linkage of familial intracranial aneurysms to 7q11 in highly aggregated Japanese families. Stroke 34: 892–900.
26. Morito D, Hirao K, Oda Y, Hosokawa N, Tokunaga F, et al. (2008) Gp78 cooperates with RMA1 in endoplasmic reticulum-associated degradation of CFTRDeltaF508. Mol Biol Cell 19: 1328–1336.
27. Seguchi O, Takashima S, Yamazaki S, Asakura M, Asano Y, et al. (2007) A cardiac myosin light chain kinase regulates sarcomere assembly in the vertebrate heart. J Clin Invest 117: 2812–2824.
28. Lawson ND, Weinstein BM (2002) In vivo imaging of embryonic vascular development using transgenic zebrafish. Dev Biol 248: 307–318.
29. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, et al. (2009) Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461: 272–276.
30. Coe BP, Ylstra B, Carvalho B, Meijer GA, Macaulay C, et al. (2007) Resolving the resolution of array CGH. Genomics 89: 647–653.
31. Lorick KL, Jensen JP, Fang S, Ong AM, Hatakeyama S, et al. (1999) RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. Proc Natl Acad Sci U S A 96: 11364–11369.
32. Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J 1: 945–951.
33. Kamada F, Aoki Y, Narisawa A, Abe Y, Komatsuzaki S, et al. (2011) A genome-wide association study identifies RNF213 as the first Moyamoya disease gene. J Hum Genet 56: 34–40.
34. Mineharu Y, Takenaka K, Yamakawa H, Inoue K, Ikeda H, et al. (2006) Inheritance pattern of familial moyamoya disease: autosomal dominant mode and genomic imprinting. J Neurol Neurosurg Psychiatry 77: 1025–1029.
35. Isogai S, Horiguchi M, Weinstein BM (2001) The vascular anatomy of the developing zebrafish: an atlas of embryonic and early larval development. Dev Biol 230: 278–301.
36. Torres-Vazquez J, Gitler AD, Fraser SD, Berk JD, Van NP, et al. (2004) Semaphorin-plexin signaling guides patterning of the developing vasculature. Dev Cell 7: 117–123.
37. Alvarez Y, Cederlund ML, Cottell DC, Bill BR, Ekker SC, et al. (2007) Genetic determinants of hyaloid and retinal vasculature in zebrafish. BMC Dev Biol 7: 114.
38. Takagi Y, Kikuta K, Nozaki K, Fujimoto M, Hayashi J, et al. (2007) Expression of hypoxia-inducing factor-1 alpha and endoglin in intimal hyperplasia of the middle cerebral artery of patients with Moyamoya disease. Neurosurgery 60: 338–345.
39. El Ramahi KM, Al Rayes HM (2000) Systemic lupus erythematosus associated with moyamoya syndrome. Lupus 9: 632–636.
40. Ogawa K, Nagahiro S, Arakaki R, Ishimaru N, Kobayashi M, et al. (2003) Anti-alpha-fodrin autoantibodies in Moyamoya disease. Stroke 34: e244–246.
41. Czartoski T, Hallam D, Lacy JM, Chun MR, Becker K (2005) Postinfectious vasculopathy with evolution to moyamoya syndrome. J Neurol Neurosurg Psychiatry 76: 256–259.
42. Sharfstein SR, Ahmed S, Islam MQ, Najjar MI, Ratushny V (2007) Case of moyamoya disease in a patient with advanced acquired immunodeficiency syndrome. J Stroke Cerebrovasc Dis 16: 268–272.
43. Somarajan A, Ashalatha R, Syam K (2005) Moya Moya disease: an unusual clinical presentation. J Assoc Physicians India 53: 49–51.
44. Tanigawara T, Yamada H, Sakai N, Andoh T, Deguchi K, et al. (1997) Studies on cytomegalovirus and Epstein-Barr virus infection in moyamoya disease. Clin Neurol Neurosurg 99(Suppl): S225–228.
45. Ueno M, Oka A, Koeda T, Okamoto R, Takeshita K (2002) Unilateral occlusion of the middle cerebral artery after varicella-zoster virus infection. Brain Dev 24: 106–108.
46. Bitzer M, Topka H (1995) Progressive cerebral occlusive disease after radiation therapy. Stroke 26: 131–136.
47. Bauersachs S, Ulbrich SE, Gross K, Schmidt SE, Meyer HH, et al. (2006) Embryo-induced transcriptome changes in bovine endometrium reveal species-specific and common molecular markers of uterine receptivity. Reproduction 132: 319–331.
48. Zhang S, Kim CC, Batra S, McKerrow JH, Loke P (2010) Delineation of diverse macrophage activation programs in response to intracellular parasites and cytokines. PLoS Negl Trop Dis 4: e648.
49. Phillips JA, 3rd, Poling JS, Phillips CA, Stanton KC, Austin ED, et al. (2008) Synergistic heterozygosity for TGFbeta1 SNPs and BMPR2 mutations modulates the age at diagnosis and penetrance of familial pulmonary arterial hypertension. Genet Med 10: 359–365.

Expansion of Intronic GGCCTG Hexanucleotide Repeat in *NOP56* Causes SCA36, a Type of Spinocerebellar Ataxia Accompanied by Motor Neuron Involvement

Hatasu Kobayashi,[1,4] Koji Abe,[2,4] Tohru Matsuura,[2,4] Yoshio Ikeda,[2] Toshiaki Hitomi,[1] Yuji Akechi,[2] Toshiyuki Habu,[3] Wanyang Liu,[1] Hiroko Okuda,[1] and Akio Koizumi[1,*]

Autosomal-dominant spinocerebellar ataxias (SCAs) are a heterogeneous group of neurodegenerative disorders. In this study, we performed genetic analysis of a unique form of SCA (SCA36) that is accompanied by motor neuron involvement. Genome-wide linkage analysis and subsequent fine mapping for three unrelated Japanese families in a cohort of SCA cases, in whom molecular diagnosis had never been performed, mapped the disease locus to the region of a 1.8 Mb stretch (LOD score of 4.60) on 20p13 (D20S906–D20S193) harboring 37 genes with definitive open reading frames. We sequenced 33 of these and observed a large expansion of an intronic GGCCTG hexanucleotide repeat in *NOP56* and an unregistered missense variant (Phe265Leu) in *C20orf194*, but we found no mutations in *PDYN* and *TGM6*. The expansion showed complete segregation with the SCA phenotype in family studies, whereas Phe265Leu in *C20orf194* did not. Screening of the expansions in the SCA cohort cases revealed four additional occurrences, but none were revealed in the cohort of 27 Alzheimer disease cases, 154 amyotrophic lateral sclerosis cases, or 300 controls. In total, nine unrelated cases were found in 251 cohort SCA patients (3.6%). A founder haplotype was confirmed in these cases. RNA foci formation was detected in lymphoblastoid cells from affected subjects by fluorescence in situ hybridization. Double staining and gel-shift assay showed that (GGCCUG)n binds the RNA-binding protein SRSF2 but that $(CUG)_6$ does not. In addition, transcription of *MIR1292*, a neighboring miRNA, was significantly decreased in lymphoblastoid cells of SCA patients. Our finding suggests that SCA36 is caused by hexanucleotide repeat expansions through RNA gain of function.

Autosomal-dominant spinocerebellar ataxias (SCAs) are a heterogeneous group of neurodegenerative disorders characterized by loss of balance, progressive gait, and limb ataxia.[1–3] We recently encountered two unrelated patients with intriguing clinical symptoms from a community in the Chugoku region in western mainland Japan.[4] These patients both showed complicated clinical features, with ataxia as the first symptom, followed by characteristic late-onset involvement of the motor neuron system that caused symptoms similar to those of amyotrophic lateral sclerosis (ALS [MIM 105400]).[4] Some SCAs (SCA1 [MIM 164400], SCA2 [MIM 183090], SCA3 [MIM 607047], and SCA6 [MIM 183086]) are known to slightly affect motor neurons; however, their involvement is minimal and the patients usually do not develop skeletal muscle and tongue atrophies.[4] Of particular interest is that RNA foci have been recently demonstrated in hereditary disorders caused by microsatellite repeat expansions or insertions in the noncoding regions of their gene.[5–7] The unique clinical features in these families have seldom been described in previous reports; therefore, we undertook a genetic analysis.

A similar form of SCA was observed in five Japanese cases from a cohort of 251 patients with SCA, in whom molecular diagnosis had not been performed, who were followed by the Department of Neurology, Okayama University Hospital. These five cases originated from a city of 450,000 people in the Chugoku region. Thus, we suspected the presence of a founder mutation common to these five cases, prompting us to recruit these five families (pedigrees 1–5) (Figure 1, Table 1). This study was approved by the Ethics Committee of Kyoto University and the Okayama University institutional review board. Written informed consent was obtained from all subjects. An index of cases per family was investigated in some depth: IV-4 in pedigree 1, II-1 in pedigree 2, III-1 in pedigree 3, II-1 in pedigree 4, and II-1 in pedigree 5. The mean age at onset of cerebellar ataxia was 52.8 ± 4.3 years, and the disease was transmitted by an autosomal-dominant mode of inheritance. All affected individuals started their ataxic symptoms, such as gait and truncal instability, ataxic dysarthria, and uncoordinated limbs, in their late forties to fifties. MRI revealed relatively confined and mild cerebellar atrophy (Figure 2A). Unlike individuals with previously known SCAs, all affected individuals with longer disease duration showed obvious signs of motor neuron involvement (Table 1). Characteristically, all affected individuals exhibited tongue atrophy with fasciculation, although its degree of severity varied (Figure 2B). Despite severe tongue atrophy in some cases, their swallowing function was relatively preserved, and they were allowed oral intake even at a later point after onset. In addition to tongue atrophy, skeletal muscle atrophy and fasciculation in the limbs and trunk appeared in advanced cases.[4] Tendon reflexes were generally mildly to severely hyperreactive in most

[1]Department of Health and Environmental Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [2]Department of Neurology, Graduate School of Medicine, Dentistry and Pharmaceutical Science, Okayama University, Okayama, Japan; [3]Radiation Biology Center, Kyoto University, Kyoto, Japan
[4]These authors contributed equally to this work
*Correspondence: koizumi.akio.5v@kyoto-u.ac.jp
DOI 10.1016/j.ajhg.2011.05.015.

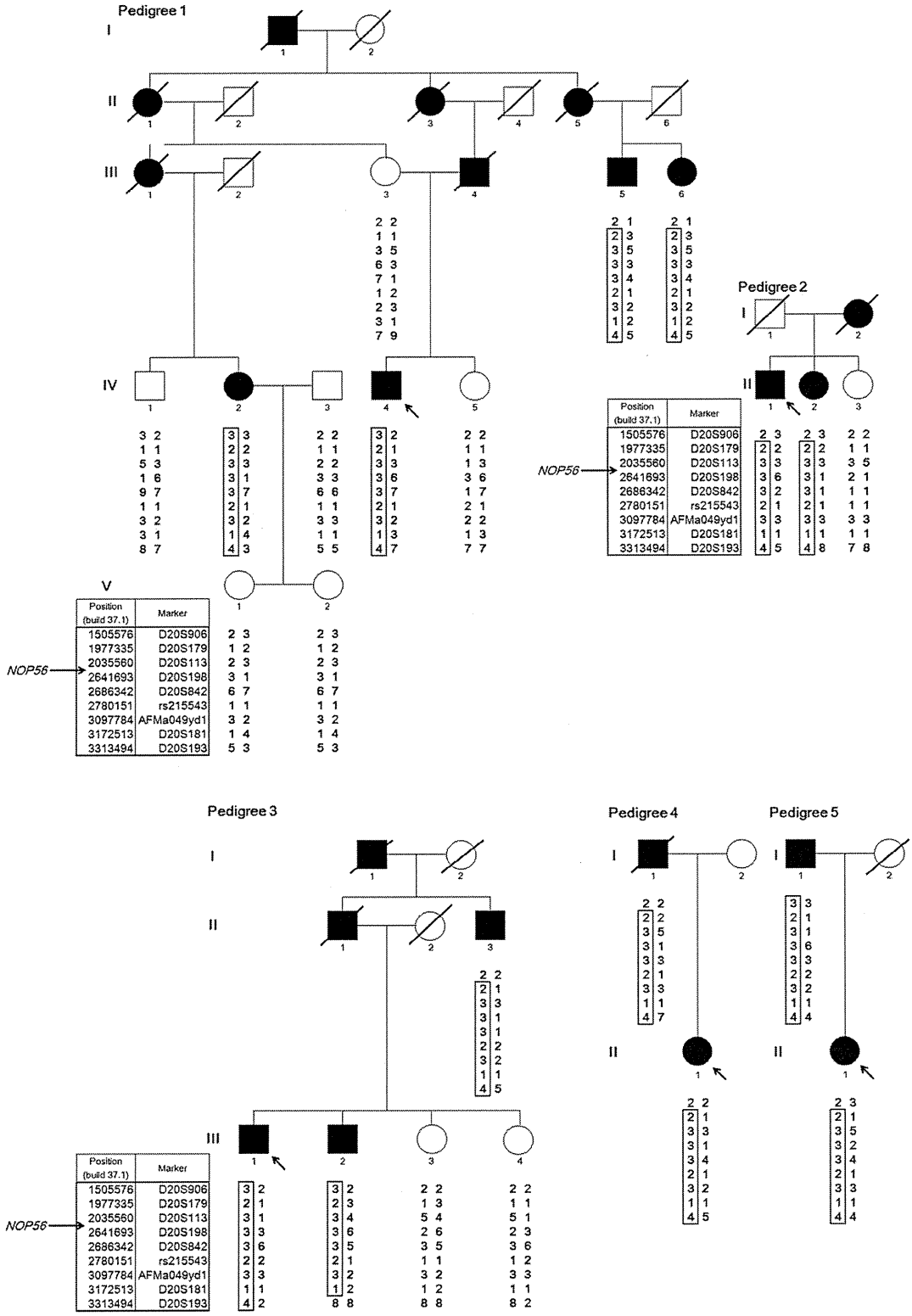


Figure 1. Pedigree Charts of the Five SCA Families
Haplotypes are shown for nine markers from D20S906 (1,505,576 bp) to D20S193 (3,313,494 bp), spanning 1.8 Mb on chromosome 20p13. *NOP56* is located at 2,633,254–2,639,039 bp (NCBI build 37.1). Filled and unfilled symbols indicate affected and unaffected individuals, respectively. Squares and circles represent males and females, respectively. A slash indicates a deceased individual. The putative founder haplotypes among patients are shown in boxes constructed by GENHUNTER.[8] Arrows indicate the index case. The pedigrees were slightly modified for privacy protection.

Table 1. Clinical Characteristics of Affected Subjects

Pedigree No.	Patient ID	Gender	Onset Age (yr)	Current Age (yr)	Ataxia	Motor Neuron Involvement			Genotype of GGCCTG Repeats
						Skeletal Muscle Atrophy	Skeletal Muscle Fasciculation	Tongue Atrophy/ Fasciculation	
1	III-5	M	50	70	+++	N.D.	N.D.	N.D.	g.263397_263402[6]+(1800)
	III-6	F	52	68	++	+	+	+	g.263397_263402[6]+(2300)
	IV-2	F	57	63	+	-	-	+	g.263397_263402[6]+(2300)
	IV-4	M	50	59	+	-	-	+	g.263397_263402[6]+(2300)
2	II-1	M	55	77	+++	++	+	+	g.263397_263402[6]+(2200)
	II-2	F	53	70	++	N.D.	N.D.	N.D.	g.263397_263402[6]+(2200)
3	II-3	M	58	77	++	++	+	+	g.263397_263402[3]+(2300)
	III-1	M	56	62	+	-	-	±	g.263397_263402[8]+(2200)
	III-2	M	51	61	++	+	+	+	g.263397_263402[6]+(1800)
4	I-1	M	57	died in 2001 at 83	++	N.D.	N.D.	N.D.	g.263397_263402[5]+(1800)
	II-1	F	48	61	++	+	±	++	g.263397_263402[6]+(2000)
5	I-1	M	57	86	++	+++	+	+	g.263397_263402[5]+(2000)
	II-1	F	47	58	++	+	+	+	g.263397_263402[8]+(1700)
	SCA#1	M	52	69	+++	+++	+++	+++	g.263397_263402[5]+(2200)
	SCA#2	F	43	53	+++	-	-	+	g.263397_263402[6]+(1800)
	SCA#3	M	55	60	++	-	-	++	g.263397_263402[8]+(1700)
	SCA#4	M	57	81	+++	+	+	+++	g.263397_263402[5]+(2200)
Mean			52.8						
SD			4.3						

N.D., not determined.

affected individuals, none of whom displayed severe lower limb spasticity or extensor plantar response. Electrophysiological studies were performed in an affected individual. Nerve conduction studies revealed normal findings in all of the cases that were examined; however, an electromyogram showed neurogenic changes only in cases with skeletal muscle atrophy, indicating that lower motor neuropathy existed in this particular disease. Progression of motor neuron involvement in this SCA was typically limited to the tongue and main proximal skeletal muscles in both upper and lower extremities, which is clearly different from typical ALS, which usually involves most skeletal muscles over the course of a few years, leading to fatal results within several years.

We conducted genome-wide linkage analysis for nine affected subjects and eight unaffected subjects in three informative families (pedigrees 1–3; Figure 1). For genotyping, we used an ABI Prism Linkage Mapping Set (Version 2; Applied Biosystems, Foster City, CA, USA) with 382 markers, 10 cM apart, for 22 autosomes. Fine-mapping markers (approximately 1 cM apart) were designed according to information from the uniSTS reference physical map in the NCBI database. A parametric linkage analysis was carried out in GENEHUNTER[8] with the assumption of an autosomal-dominant model. The disease allele frequency was set at 0.000001, and a phenocopy frequency of 0.000001 was assumed. Population allele frequencies were assigned equal portions of individual alleles. We performed multipoint analyses for autosomes and obtained LOD scores. We considered LOD scores above 3.0 to be significant.[8] Genome-wide linkage analysis revealed a single locus on chromosome 20p13 with a LOD score of 3.20. Fine mapping increased the LOD score to 4.60 (Figure 3). Haplotype analysis revealed two recombination events in pedigree 3, delimiting a1.8 Mb region (D20S906–D20S193) (Figure 1). We further tested whether the five cases shared the haplotype. As shown in Figure 1, pedigrees 4 and 5 were confirmed to have the same haplotype as pedigrees 1, 2, and 3, indicating that the 1.8 Mb region is very likely to be derived from a common ancestor.

The 1.8 Mb region harbors 44 genes (NCBI, build 37.1). We eliminated two pseudogenes and five genes (*LOC441938*, *LOC100289473*, *LOC100288797*, *LOC100289507*, and *LOC100289538*) from the candidates. Evidence view showed that the first, fourth, and fifth genes were not found in the contig in this region, whereas the second and third

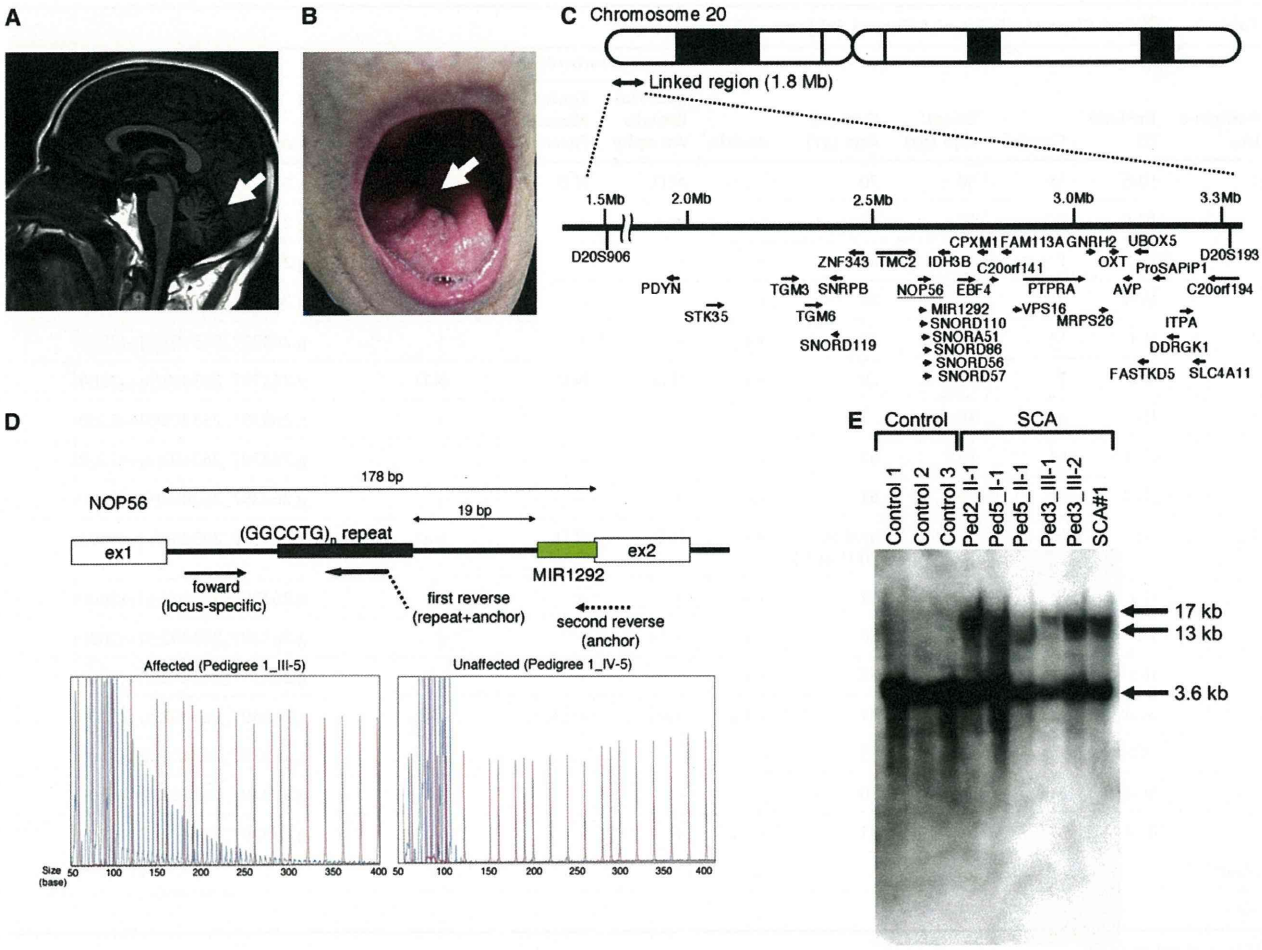


Figure 2. Motor Neuron Involvement and (GGCCTG)n Expansion in the First Intron of *NOP56*

(A) MRI of an affected subject (SCA#3) showed mild cerebellar atrophy (arrow) but no other cerebral or brainstem pathology.

(B) Tongue atrophy (arrow) was observed in SCA#1.

(C) Physical map of the 1.8-Mb linkage region from D20S906 (1,505,576 bp) to D20S193 (3,313,494 bp), with 33 candidate genes shown, as well as the direction of transcription (arrows).

(D) The upper portion of the panel shows the scheme of primer binding for repeat-primer PCR analysis. In the lower portion, sequence traces of the PCR reactions are shown. Red lines indicate the size markers. The vertical axis indicates arbitrary intensity levels. A typical saw-tooth pattern is observed in an affected pedigree.

(E) Southern blotting of LCLs from SCA cases and three controls. Genomic DNA (10 μg) was extracted from Epstein-Barr virus (EBV)-immortalized LCLs derived from six affected subjects (Ped2_II-1, Ped3_III-1, Ped3_III-2, Ped5_I-1, Ped5_II-1, and SCA#1) and digested with 2 U of *AvrII* overnight (New England Biolabs, Beverly, MA, USA). A probe covering exon 4 of *NOP56* (452 bp) was subjected to PCR amplification from human genomic DNA with the use of primers (Table S3) and labeled with ^{32}P-dCTP.

genes are not assigned to orthologous loci in the mouse genome. Sequence similarities among paralog genes defied direct sequencing of four genes: *SIRPD* [NM 178460.2], *SIRPB1* [NM 603889], *SIRPG* [NM 605466], and *SIRPA* [NM 602461]. Thus, we sequenced 33 of 37 genes (*PDYN* [MIM 131340], *STK35* [MIM 609370], *TGM3* [MIM 600238], *TGM6* [NM_198994.2], *SNRPB* [MIM 182282], *SNORD119* [NR_003684.1], *ZNF343* [NM_024325.4], *TMC2* [MIM 606707], *NOP56* [NM_006392.2], *MIR1292* [NR_031699.1], *SNORD110* [NR_003078.1], *SNORA51* [NR_002981.1], *SNORD86* [NR_004399.1], *SNORD56* [NR_002739.1], *SNORD57* [NR_002738.1], *IDH3B* [MIM 604526], *EBF4* [MIM 609935], *CPXM1* [NM_019609.4], *C20orf141* [NM_080739.2], *FAM113A* [NM_022760.3], *VPS16* [MIM 608550], *PTPRA* [MIM 176884], *GNRH2* [MIM 602352], *MRPS26* [MIM 611988], *OXT* [MIM 167050], *AVP* [MIM 192340], *UBOX5* [NM_014948.2], *FASTKD5* [NM_021826.4], *ProSAPiP1* [MIM 610484], *DDRGK1* [NM_023935.1], *ITPA* [MIM 147520], *SLC4A11* [MIM 610206], and *C20orf194* [NM_001009984.1]) (Figure 2C). All noncoding and coding exons, as well as the 100 bp up- and downstream of the splice junctions of these genes, were sequenced in two index cases (IV-4 in pedigree1 and III-1 in pedigree 3) and in three additional cases (II-1 in pedigree 2, II-1 in pedigree 4, and II-1 in pedigree 5) with the use of specific primers (Table S1 available online). Eight unregistered variants were found among the two index cases. Among these, there was a coding variant, c.795C>G

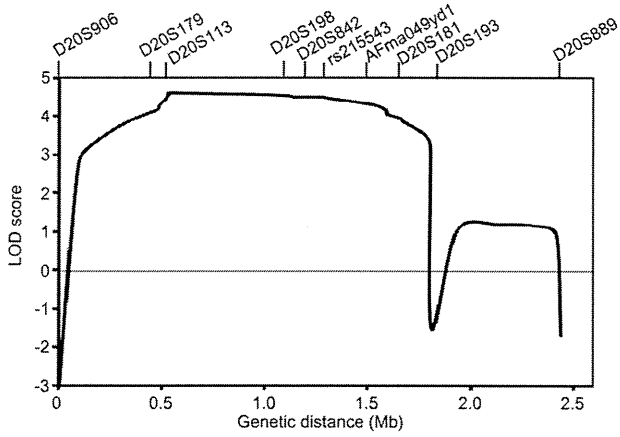


Figure 3. Multipoint Linkage Analysis with Ten Markers on Chromosome 20p13

(p.Phe265Leu), in *C20orf194*, whereas the other seven included one synonymous variant, c.1695T>A (p.Leu565-Leu), in *ZNF343* and six non-splice-site intronic variants (Table S2). We tested segregation by sequencing exon 11 of *C20orf194* in IV-2 and III-5 in pedigree 1. Neither IV-2 nor III-5 had this variant. We thus eliminated *C20orf194* as a candidate. Missense mutations in *PDYN* and *TGM6*, which have been recently reported as causes of SCA, mapped to 20p12.3-p13,[9,10] but none were detected in the five index cases studied here (Table S2).

Possible expansions of repetitive sequences in these 33 genes were investigated when intragenic repeats were indicated in the database (UCSC Genome Bioinformatics). Expansions of the hexanucleotide repeat GGCCTG (rs68063608) were found in intron 1 of *NOP56* (Figure 2D) in all five index cases through the use of a repeat-primed PCR method.[11–13] An outline of the repeat-primed PCR experiment is described in Figure 2D. In brief, the fluorescent-dye-conjugated forward primer corresponded to the region upstream of the repeat of interest. The first reverse primer consisted of four units of the repeat (GGCCTG) and a 5′ tail used as an anchor. The second reverse primer was an "anchor" primer. These primers are described in Table S3. Complete segregation of the expanded hexanucleotide was confirmed in all pedigrees, and the maximum repeat size in nine unaffected members was eight (data not shown).

In addition to the SCA cases in five pedigrees, four unrelated cases (SCA#1–SCA#4) were found to have a (GGCCTG)n allele through screening of the cohort SCA patients (Table 1). Neurological examination was reevaluated in these four cases, revealing both ataxia and motor neuron dysfunction with tongue atrophy and fasciculation (Table 1). In total, nine unrelated cases were found in the 251 cohort patients with SCA (3.6%). For confirmation of the repeat expansions, Southern blot analysis was conducted in six affected subjects (Ped2_II-1, Ped3_III-1, Ped3_III-2, Ped5_I-1, Ped5_II-1, and SCA#1). The data showed >10 kb of repeat expansions in the lymphoblastoid cell lines (LCLs) obtained from the SCA patients (Figure 2E). Furthermore, the numbers of GGCCTG repeat expansion were estimated by Southern blotting in 11 other cases. The expansion analysis revealed approximately 1500 to 2500 repeats in 17 cases (Table 1). There was no negative association between age at onset and the number of GGCCTG repeats ($n = 17$, $r = 0.42$, $p = 0.09$; Figure S1) and no obvious anticipation in the current pedigrees.

To investigate the disease specificity and disease spectrum of the hexanucleotide repeat expansions, we tested the repeat expansions in an Alzheimer disease (MIM 104300) cohort and an ALS cohort followed by the Department of Neurology, Okayama University Hospital. We also recruited Japanese controls, who were confirmed to be free from brain lesions through MRI and magnetic resonance angiography, which was performed as described previously.[14] Screening of the 27 Alzheimer disease cases and 154 ALS cases failed to detect additional cases with repeat expansions. The GGCCTG repeat sizes ranged from 3 to 8 in 300 Japanese controls (5.9 ± 0.8 repeats), suggesting that the >10 kb repeat expansions were mutations.

Expression of *Nop56*, an essential component of the splicing machinery,[15] was examined by RT-PCR with the use of primers for wild-type mouse *Nop56* cDNA (Table S3). Expression of *Nop56* mRNA was detected in various tissues, including CNS tissue, and a very weak signal was detected in spinal cord tissue (Figure 4A). Immunohistochemistry using an anti-mouse Nop56 antibody (Santa Cruz Biotechnology, Santa Cruz, CA, USA) detected the Nop56 protein in Purkinje cells of the cerebellum as well as motor neurons of the hypoglossal nucleus and the spinal cord anterior horn (Figure 4B), suggesting that these cells may be responsible for tongue and muscle atrophy in the trunk and limbs, respectively. Immunoblotting also confirmed the presence of Nop56 in neural tissues (Figure 4C), where Nop56 is localized in both the nucleus and cytoplasm.

Alterations of *NOP56* RNA expression and protein levels in LCLs from patients were examined by real-time RT-PCR and immunoblotting. The primers for quantitative PCR of human *NOP56* cDNA are described in Table S3. Immunoblotting was performed with the use of an anti-human NOP56 antibody (Santa Cruz Biotechnology, Santa Cruz, CA, USA). We found no decrease in *NOP56* RNA expression or protein levels in LCLs from these patients (Figure 5A). To investigate abnormal splicing variants of *NOP56*, we performed RT-PCR using the primers covering the region from the 5′ UTR to exon 4 around the repeat expansion (Table S3); however, no splicing variant was observed in LCLs from the cases (Figure 5B). We also performed immunocytochemistry for NOP56 and coilin, a marker of the Cajal body, where NOP56 functions.[16] NOP56 and coilin distributions were not altered in LCLs of the SCA patients (Figure 5C), suggesting that qualitative or quantitative changes in the Cajal body did not occur. These results indicated that haploinsufficiency could not explain the observed phenotype.