DATA SUPPLEMENT

Systematic validation of *RNF213* coding variants in Japanese patients with Moyamoya disease

Yosuke Moteki, MD[1]; Hideaki Onda, MD[1]; Hidetoshi Kasuya, MD[2,5]; Taku Yoneyama, MD[1]; Yoshikazu Okada, MD[1]; Kengo Hirota, MD[2,5]; Maki Mukawa, MD[3]; Tadashi Nariai, MD[3]; Shohei Mitani, MD[4,5]; PhD; Hiroyuki Akagawa, MD[2,5]

[1]Department of Neurosurgery, Neurological Institute, Tokyo Women's Medical University, Tokyo, Japan
[2]Department of Neurosurgery, Medical Center East, Tokyo Women's Medical University, Tokyo, Japan
[3]Department of Neurosurgery, Tokyo Medical and Dental University, Tokyo, Japan
[4]Department of Physiology, Tokyo Women's Medical University, Tokyo, Japan
[5]Tokyo Women's Medical University Institute for Integrated Medical Sciences (TIIMS)

Correspondence should be addressed to Akagawa H.
(akagawa.hiroyuki@twmu.ac.jp)

Supplemental Methods

Confirmation of the rare sequence variants

Twenty-one out of the 30 rare variants detected in the present study were called as true genetic variants by the Variant Quality Score Recalibration (VQSR) [1] in the Exome Aggregation Consortium (ExAC, Cambridge, MA, URL: http://exac.broadinstitute.org, March 2015). All the rest of nine variants were patient-specific and confirmed by direct sequencing: three (p.Q3020L, p.R4062Q and p.E4750K) in our 103 MMD patients (Supplemental Figure III), two (p.M3891V and p.V4765M) by Kamada et al [2] and four (p.L1911I, p.Q3082R, p.W4024R and p.E4917K) by Miyatake et al [3].

The authors would like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at http://exac.broadinstitute.org/about.

Haplotype phasing

In the present study, three individuals were identified to carry multiple rare missense variants (p.Q3020L, p.E4750K, and p.R4927Q together with p.R4810K). Of these, DNA samples of their parents were not available, except for the p.E4750K mutant. In order to determine whether their genotypes were compound heterozygous, we performed direct haplotyping using long range PCR covering the variants of interest, followed by subcloning.

PCR primers were designed to cover the genomic region between the p.R4810K and p.R4927Q loci (1,605bp) for the p.R4927Q mutant as follows:
Ex60-Ex63_Forward (5'-CTCGCAGCCAGTCTCAAAGT-3');
Ex60-Ex63_Reverse (5'-ACACACCAAATGAGCAGCAG-3').
The PCR product was then subcloned into the pMD20-T vector (Takara Bio Inc.) to divide diploid amplicons, and sequenced for haplotype phasing.

Since the genomic distance between the p.Q3020L and p.R4810K loci was too long to be amplified by single PCR（37,752bp）, we combined computational phase inference using LD with proposed direct haplotyping for the p.Q3020L mutant. Within the p.Q3020L to p.R4810K interval, four common SNPs (rs35993981, rs8067292, rs6565681, and rs7223115) forming strong LD (ID'l>0.8) in 89 JPT subjects were selected from the 1000 genomes project database [4]. The p.Q3020L and p.R4810K loci were amplified with neighboring rs35993981 and rs7223115 loci, respectively, using the following PCR primers:
Q3020L-rs35993981_Forward (5'-TGACTACTACAGCCTCATCAAAATG-3');
Q3020L-rs35993981_Reverse (5'-CTGGGAGAGATTTAACAAGGATCTA-3');
rs7223115-R4810K_Forward (5'-TGGAACAGCAAGTACTCTTCACATA-3');
rs7223115-R4810K_Reverse (5'-AGCTTCTAATATGTTTTTGGGGTTC-3').
Direct haplotyping was performed within each amplicon, as described above. We genotyped these four interval SNPs in our 103 MMD patients, and each haplotype phase was then computed using BEAGLE version 3.3.2 [5].

Violin plots of C scores across various subsets of missense changes

Differences in C scores between the candidate variants and other missense changes with specific functional consequences were visualized using violin plots. Five functional categories were set as follows.

i) *Benign*

We extracted ancestral chimpanzee (PanTro4) alleles that were altered and fixed in the human lineage with allele frequencies >99% for the benign missense set. A total of 4,330 of these autosomal missense changes were obtained.

ii) *GWAS*

We downloaded the National Human Genome Research Institute (NHGRI) genome-wide association study (GWAS) catalog (http://www.genome.gov/gwastudies/) [6] on September 15, 2014. A total of 327 autosomal missense SNPs with trait associations were obtained.

iii) *MMD*

We selected 15 candidate variants for MMD susceptibility according to the VT test (C score >10.02, Fisher's exact test $P < 0.05$, Figure 1A).

iv) *Gain-of-function*

We searched the distinct gain-of-function missense variants curated in the Online Mendelian Inheritance in Man (OMIM) database (http://omim.org). A total of 107 autosomal OMIM genes were identified that contained the search words "gain of function" or "gain-of-function" in the ALLELIC VARIANTS field. We extracted a total of 91 autosomal missense variants specified as gain-of-function variants from the variants listed.

v) *Loss-of-function*

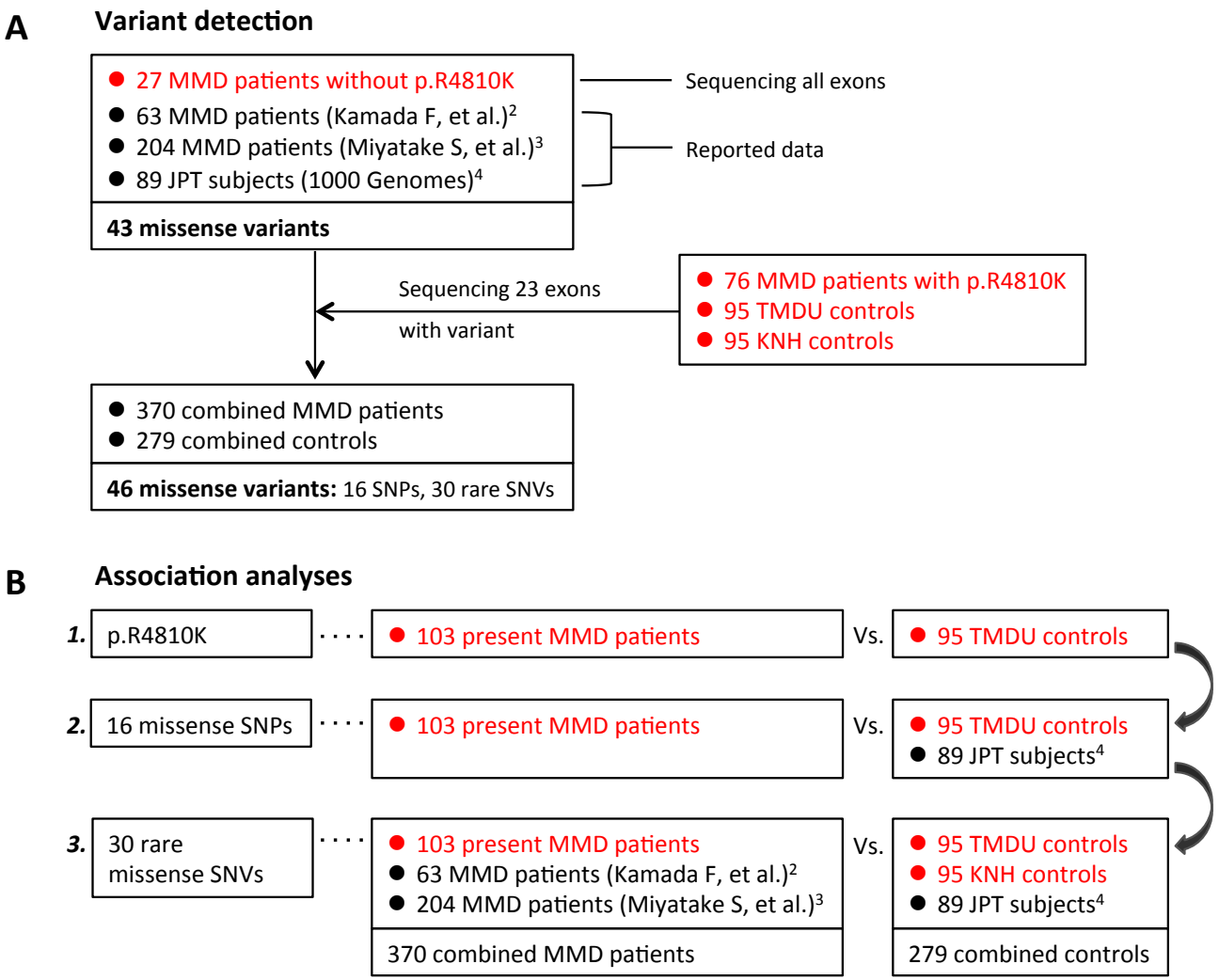
In the same manner as described above, a total of 130 autosomal OMIM genes were identified that contained nonsense variants together with the search words "loss of function" or "loss-of-function" in the ALLELIC VARIANTS field. We extracted a total of 61 autosomal missense variants specified as loss-of-function variants from the variants listed.

The C-scores of these five functional categories were compared using the Kruskal-Wallis test followed by the post hoc Steel-Dwass test.

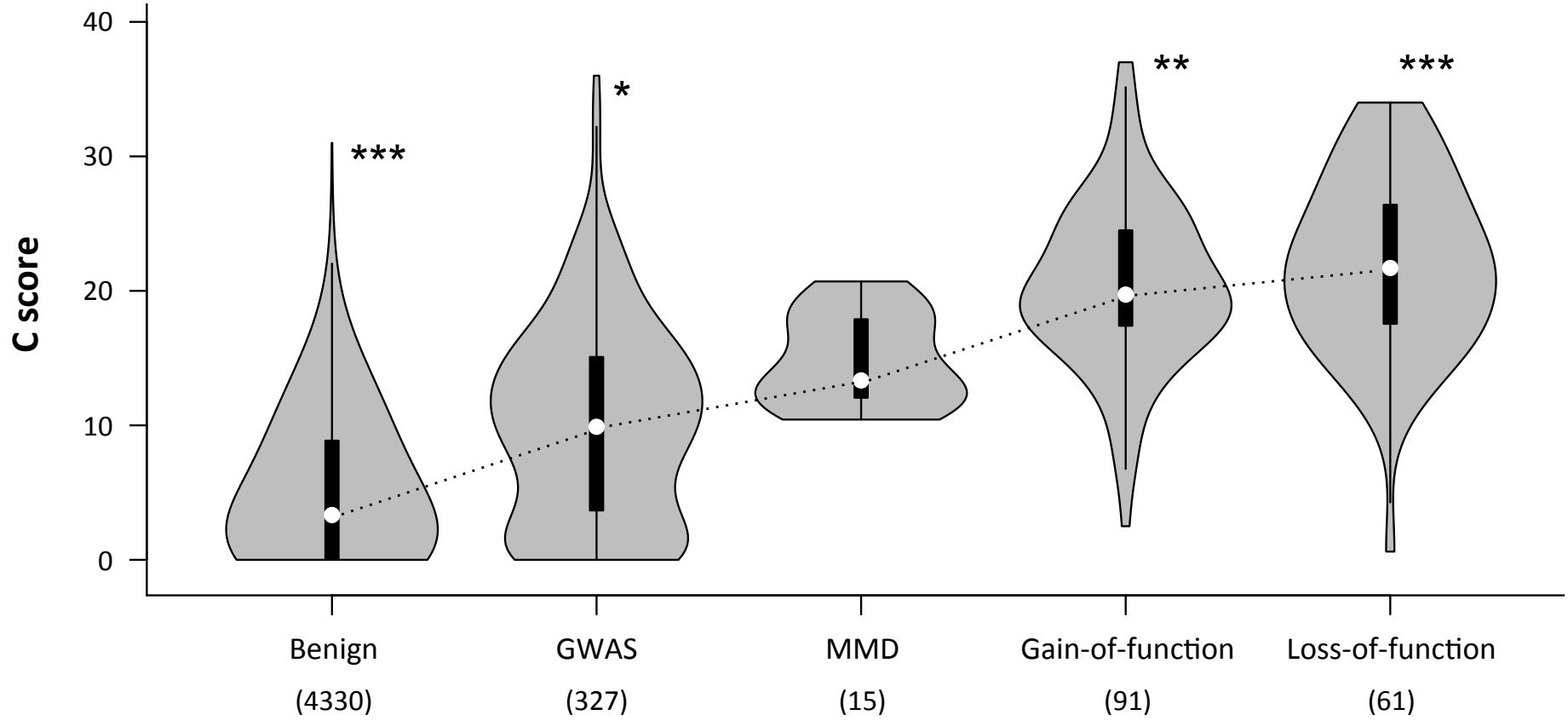
Subordinate comparisons of the GWAS SNPs and the candidate variants for MMD

Among the 327 GWAS SNPs extracted in the previous section, 125 SNPs were reported to be binary trait loci with odds ratios (ORs). In cases where a SNP reported in more than one study, a median OR was employed in the following analysis. These 125 SNPs were subdivided according to first and second quartiles of ORs. We also

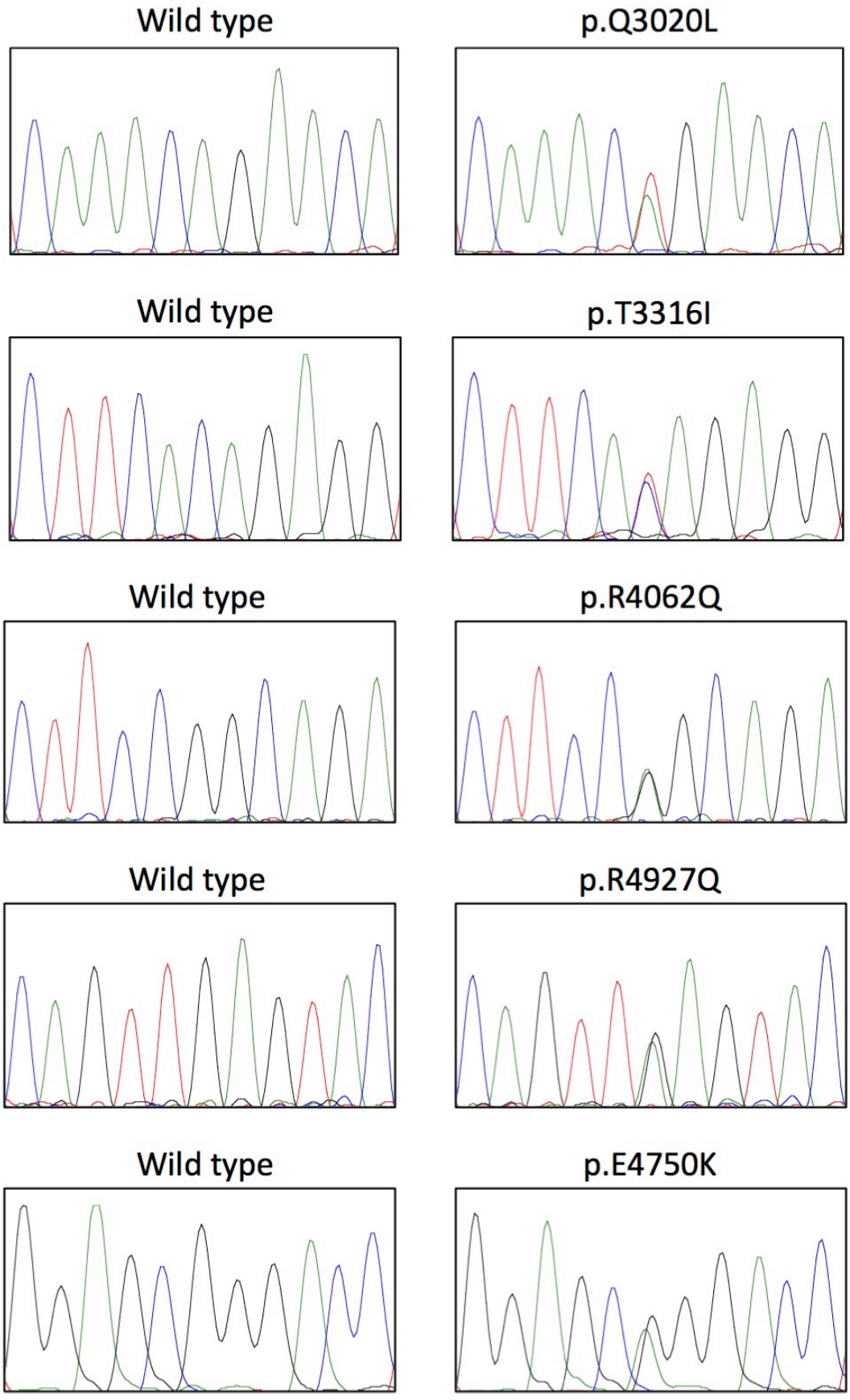
selected nine trait-associated loci of cardiovascular and cardiovascular diseases (e.g. coronary heart disease, stroke) by reason of similar conditions as MMD. From the remaining quantitative trait loci, we selected ten loci associated with serum LDL cholesterol levels. C-scores of the subdivided SNPs were compared with that of the 15 candidate variants for MMD using Steel's test.



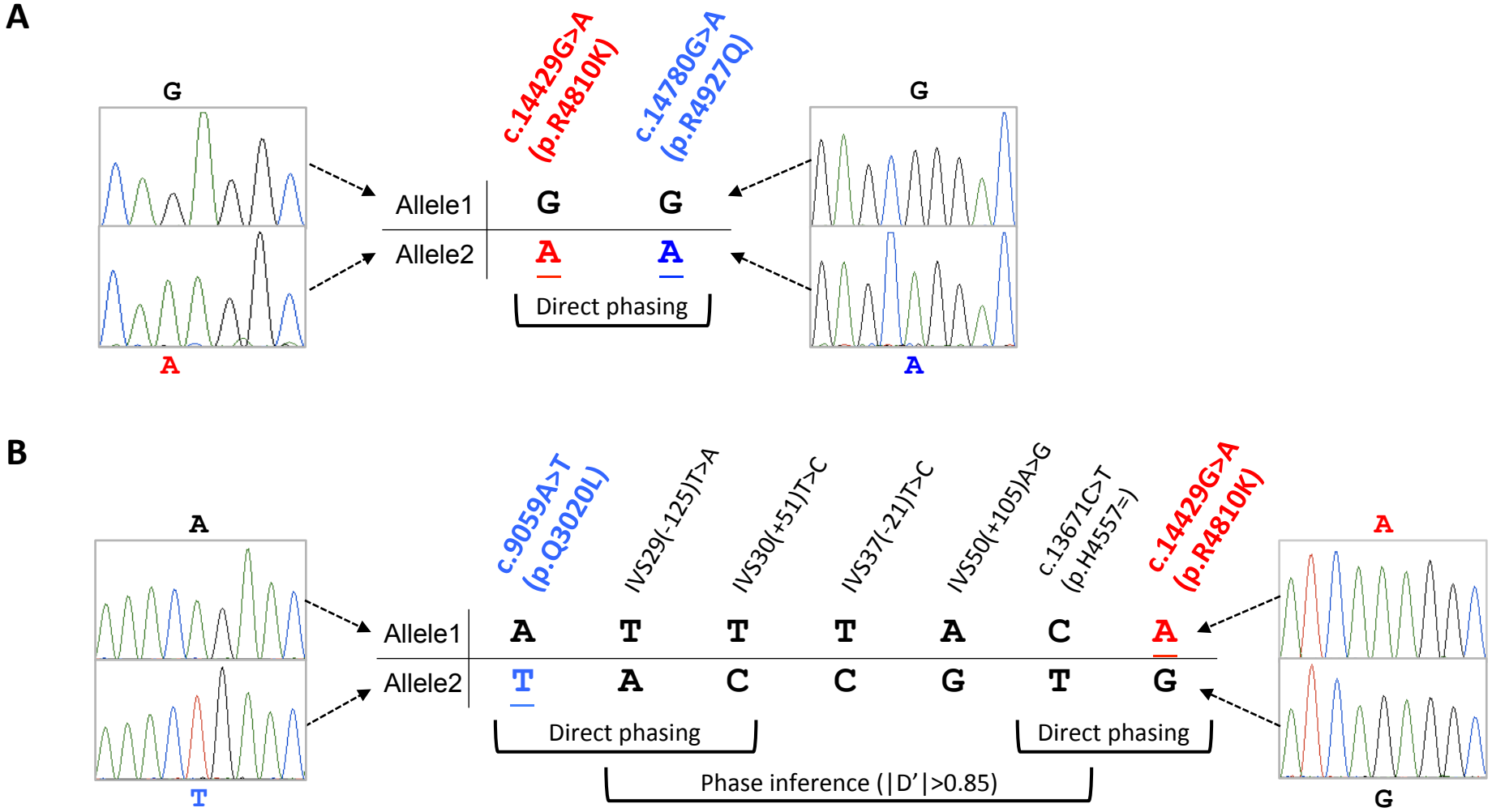
Supplemental Figure I: Flowchart describing the present study. (A) The samples used for variant detection of *RNF213*. (B) Depending on the genotypic frequencies in the patients, analytical samples were increased in stages. Red letters indicate the present DNA samples in hand. SNP: single nucleotide polymorphism, SNV: single nucleotide variant.



Supplemental Figure II: Violin plots of C scores across various functional categories of missense changes. Benign represents ancestral chimpanzee alleles altered in the human lineage; GWAS, listed in the GWAS catalogue [6]; Loss-of-function and Gain-of-function, curated in the OMIM database. Steel-Dwass test *P<0.05, **P<0.005 and ***P<0.0005, significantly different from the MMD candidate variants.



Supplemental Figure III: Chromatograms of five rare missense variants among our 103 MMD patients that ware absent in the controls.



Supplemental Figure IV: Haplotype phasing for patients harboring multiple rare variants. (A) Direct haplotype phasing for patients harboring p.R4810K and p.R4927Q. (B) In patients harboring p.Q3020L and p.R4810K, interval phase inference using tight LD was combined with direct phasing.

Supplemental Table I. Primer sequences for exon resequencing of *RNF213*.

No.	Name	Forward primer (5' -> 3')	Reverse primer (5' -> 3')	Product size
1	exon1	GCTGTGATTTCACTTTCGCA	TGGAAGGGAGGTGAGACATC	295
2	exon2	CGTGGGGAGGATTTCTGTT	GATTTTCCTGGGCAGAGTGA	298
3	exon3	TGGGGTCTTTGCGAGTCTAC	GGAGCATTCTGCCCAAACTA	260
4	exon4	TCGAGCCAAGCTTGATGTAG	AAGCAACGCCAGCACCTT	740
5	exon5	CTCGGCTTGTGGCAGATG	CATGACCTTCCCAGACACCT	539
6	exon6	AAGGGTCGTGAACATTCTGG	ACTAGAAGGGCAGCCTGGAT	695
7	exon7-1	GCATGTTTGCTAGGGTGTGA	TCCCCAAATTCTTCTCCTCC	452
8	exon7-2	AGGAGTCACCGTGTTCTTCC	TCCAGAACCCGAAGAATGTC	573
9	exon8	GCACTCCATGCTACGTTTGA	ACCAACCAGAGACACAACCC	468
10	exon9	TCAGTTGGAGGAGTCTGGCT	CCACAGAGTCAAGGGGTCAT	548
11	exon10	CTCTCACAGGTGTGGTGGTG	TTGCCTGTGTGCTTTCAGTC	489
12	exon11	CTTGCTTCCTGTTGGGACTC	ACTCCGGGTGCACAATGTAG	352
13	exon12	GGACTGTTTCTGTCCCTCCA	GTGAAAGGAAGAATGGCAGC	758
14	exon13	TCGTAGCTGCCATTCTTCCT	GGAGAGACTGCATCGTGCTT	310
15	exon14	AAGGTCAAACAAATGCCCTG	TTGAAATTCCGCACCTATCC	329
16	exon15	GTAGCCACTCCTTGAGCAGC	ACCAACCATCACCATCACCT	522
17	exon16	TTGATGAAGGTTGGGGAGAG	GTTCTCCAGCTTGCCTTCAC	287
18	exon17	GGGCCAGGAGAAGCTTAAAA	GAGCTTCCTCATGTGGCATT	776
19	exon18	TTTGCTCTTTTTGTGGCCTT	ATCTTCTGCCTCCCACCTTT	387
20	exon19	TGCCTGTGTTTGAGGAAGTG	ACAAAAATGCACACGCAAAC	376
21	exon20	CTAGTCCTTCTCCTGGTGGC	TGAGAGCCCTCAACTTTGTTC	298
22	exon21-1	TGTGTGCTCTTGAGTTCGCT	CTCTGCAAGGGTGACCTCTC	503
23	exon21-2	TGCTCAGAGACAGCCACATC	TCTGGCTAAGGAGCAGTGGT	676
24	exon22	AAGAGCTTGGTGTGCCTGTT	GATCAGTGCCCATCTCCCTA	474
25	exon23	GATTACTGCCCAAGACCGAG	GTCACTGCCCTAACGTTGCT	363
26	exon24	GACCACAGGAGGGAGAAGG	GGGGAAATCTGTGAAGAGCA	326
27	exon25	GGCATCTATGTGATCCAGGC	TTCTTGGGAGTACACAGCCC	384
28	exon26-1	CCTGTGTGCTGTTGCAGAGT	GGTGCAGTTGCTTTTGATGA	455
29	exon26-2	AGATTTGTGACCCAGAAGCG	TTTGCATGTAGACAGCCAGG	487
30	exon26-3	CTCCCGCTGATGCTCTTATC	GCTGGTAGTCTTCTCGGTGC	487
31	exon26-4	CTTTGAGGAGGTGGCACTGT	ATACCCATTGTGGGAGACCA	573
32	exon27-1	TGTCCTAGCCCTGGGACTTA	CTTCTGATACTGCGCATCCA	778
33	exon27-2	GTGAAGAGGTTGCACGACAA	GGGGACCACGTCTTTACTGA	233
34	exon28-1	CTCAGTAAGTGCCCTCCAGC	CCTTTCCAGGATTTCAACGA	658
35	exon28-2	GTGCAGACTGGAATTTGGGT	AGTCTCCTAAACGCGCCAT	248
36	exon29-1	CGTGGAGGGTTAAGACACAAA	CTAACCCCATCCATGGTGAC	588
37	exon29-2	CCTCTCTCTTCTGCAATCCG	GTCTGCAGTTGTTCCTCCGT	683
38	exon29-3	CTGTGGGAAAACCAGGCTTA	CGTGGAACCACCTGAAAACT	689
39	exon29-4	TCAGCCTAGATGAAAACGGG	GTACTGCTGCAGGTCCTTCC	698

40	exon29-5	GAAGATTCCCCTCTTCCTGG	ACAGCCTGTGCAATGTCTTG	666
41	exon29-6	GCTCCTCAGACATCCTCGTC	TGGTTGAGTGCGTCGTAGAG	600
42	exon29-7	CCCCAAGGACCAAGAGTACA	CAGGTGTGCCTGAAGGAAAT	691
43	exon29-8	CAGAAGGTGTCTGAGGAGGC	CTACTCCCCACACTTACCGC	367
44	exon30	TCCTTTGGGTTTGGATTCTG	TGACTGTGGCGCTCATTAAC	385
45	exon31	ATGTTTCCCATCACTTTGGG	GCGGGGGAAGAATTGTTATT	311
46	exon32	GAACCAACAGCTCGGAGAAC	AAGAGGTGAGGAGCTTGCCT	800
47	exon33	GCGCAGCAAGATAATGACAA	AGTTCTGCAGATGTGTCCCC	369
48	exon34	TGATCTCTCATCTGGGGACC	ACTGTGTGAGGAAAGTGGGG	445
49	exon35	CCCCACTTTCCTCACACAGT	CACACAGTGAAATGATGCCC	455
50	exon36	GAGGTGGGACAGAAGCACTC	TGTGCTTCTTGGGGAAAAAC	371
51	exon37	CAAAATGTCAAAAGGTGGCA	GGGTCCTGCTGTGTTCTGAT	325
52	exon38	TCTGAGAAGAGCTGGCATCA	GGTGACGTAAATTGGACGCT	397
53	exon39	ACGAGCTCCCATTCAACAAG	GCAAGCTCTACACCTGACCC	314
54	exon40	GTGACCTTAACGTGGGAGGA	CAGAAGCTCTCCATTCCCAG	468
55	exon41	CTGGGAATGGAGAGCTTCTG	CAGATGAAGCAGTGGGTGAG	457
56	exon42	GACACCGAGTCCCAGCTAAG	TCTTCCTTTCGGAAAAGGGT	385
57	exon43	CCCCTGACAAGCAGCATAAT	GCAGAGAAACTGGCCAGAAG	378
58	exon44	TGTGAATGCAAGGAGACAGC	CCTCTGGTGAGGACCTTTGT	600
59	exon45	CGGGTGAGAGGAGTGACTTG	ACTGAGCGGATGACAGGAGT	771
60	exon46	GAGATGCTGCCCAGAGTAGG	TAACAAAACGCTGCGATGA	279
61	exon47	CCAGATCTGAGAGAACCAGGA	CTGCACGGAACAGAGCATT	270
62	exon48	CTGTAGAAGCTTGGGGCATC	GAGGGGATGTAGAAAAGCCC	401
63	exon49	GGGCTTTTCTACATCCCCTC	GAAGCTCTGAGGTGAGTGCC	698
64	exon50	CCTCGAATGATCAACCACCT	GGGCACGTGACTTCCAATAC	372
65	exon51	GCCCTTAGGTAGATCTGGCA	AAATCACCACCATTTGAGGG	794
66	exon52	AAAATTTCCCCCTCAAATGG	CTACAGTGACAAGCTCTGGGG	444
67	exon53	GTCCAGCAGAAGGAAAGCAG	CGGTAGAATGCAGGTGGTTT	329
68	exon54	AACAAATTAGCGTGGGATGC	CCACCCCTCAAACCCTATCT	212
69	exon55	ACACTGGGACAATCTGAGGG	GCTGGAATAAAAACAGCCCA	288
70	exon56	GAACATGGTCCAGGCTTTGT	AAAGGACTCCCTGGGAAGAA	374
71	exon57	CCTCTACCAGGCTCACCATC	CATTCTTTCCAGCAACAGCA	381
72	exon58	TCAAAAGGTCATTAAGTTGGTGG	GCTCAGTCAAAGGCTCTTGC	397
73	exon59	CCACTGGTGGAGTTACTGGG	GAACTCTGCACCGAAAGAGG	383
74	exon60	CTCGCAGCCAGTCTCAAAGT	ATGTTTTTGGGGTTCAAGCA	381
75	exon61	TGCACAAAGCAGGAAAGATG	AACACAGTGCTGGGTGTTGA	300
76	exon62	GCATCAAAAGGGAGCTGAAA	CACTGGGCTTAGGGACTCTG	419
77	exon63	CCACTGCTCCAACTGTGAGA	ACACACCAAATGAGCAGCAG	345
78	exon64	GCACTACGCTGCAGTTTTCC	CACTGAGAAGGCAGAAAGCC	230
79	exon65	CTCGGCTCTCTTACCAGGTG	TGTCTCCCCATCCTTTTCAG	393
80	exon66,67	ACAGGGCAGAACGAAGAGAA	CAGGAAATGGGATTCTGTGG	599
81	exon68	TTACACACGTGAGCCACCAT	AACAGCTCGGCTTTCAAAAA	398

Supplemental Table II. Clinical features of subjects harboring rare missense variants other than p.R4810K.

Pedigree-Individual ID	Ethnicity	Gender	Age at onset	*RNF213* genotype	*RNF213* variant	Phenotype	Onset	Surgical treatment	Familial history of MMD (relationship to the proband)
1-1	Japanese	Female	21	compound heterozygote	Q3020L R4810K	Bilateral MMD	TIA	yes	no
2-1	Japanese	Female	43	heterozygote	T3316I -	Bilateral MMD	infarction	yes	no
3-1	Japanese	Male	8	heterozygote	R4062Q -	Bilateral MMD	TIA	yes	yes (proband)
3-2	Japanese	Male	-	heterozygote	R4062Q -	non MMD	-	-	yes (father of 3-1)
3-3	Japanese	Female	8	heterozygote	R4062Q -	Bilateral MMD	TIA	yes	yes (paternal aunt of 3-1)
*4-1[7]	German	NA	NA	heterozygote	R4062Q -	MMD	NA	NA	NA
5-1	Japanese	Female	1	compound heterozygote	E4750K R4810K	Bilateral MMD	TIA	yes	no
5-2	Japanese	Female	-	heterozygote	E4750K -	non MMD	-	-	yes (mother of 5-1)
5-3	Japanese	Male	-	heterozygote	R4810K -	non MMD	-	-	yes (father of 5-1)
6-1	Japanese	Male	19	heterozygote	R4810K, R4927Q -	Bilateral MMD	TIA	yes	no

*The German MMD patient reported by Liu W et al.[7] NA = not available, TIA = transient ischemic attack.

References

1. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43:491-498.
2. Kamada F, Aoki Y, Narisawa A, Abe Y, Komatsuzaki S, Kikuchi A, Kanno J, Niihori T, Ono M, Ishii N, Owada Y, Fujimura M, Mashimo Y, Suzuki Y, Hata A, Tsuchiya S, Tominaga T, Matsubara Y, Kure S. A genome-wide association study identifies *RNF213* as the first Moyamoya disease gene. *J Hum Genet*. 2010;56:34-40.
3. Miyatake S, Miyake N, Touho H, Nishimura-Tadaki A, Kondo Y, Okada I, Tsurusaki Y, Doi H, Sakai H, Saitsu H, Shimojima K, Yamamoto T, Higurashi M, Kawahara N, Kawauchi H, Nagasaka K, Okamoto N, Mori T, Koyano S, Kuroiwa Y, Taguri M, Morita S, Matsubara Y, Kure S, Matsumoto N. Homozygous c.14576G>A variant of *RNF213* predicts early-onset and severe form of moyamoya disease. *Neurology*. 2012;78:803-810.
4. 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56–65.
5. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies using localized haplotype clustering. *Am J Hum Genet*. 2007;81:1084-1097.
6. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, Parkinson H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 2014;42:D1001-D1006.
7. Liu W, Morito D, Takashima S, Mineharu Y, Kobayashi H, Hitomi T, Hashikata H, Matsuura N, Yamazaki S, Toyoda A, Kikuta K, Takagi Y, Harada KH, Fujiyama A, Herzig R, Krischek B, Zou L, Kim JE, Kitakaze M, Miyamoto S, Nagata K, Hashimoto N, Koizumi A. Identification of *RNF213* as a susceptibility gene for moyamoya disease and its possible role in vascular development. *PLoS One*. 2011;6:e22542.