Supplementary Information for:

Genome-wide association study of intracranial aneurysm identifies three new risk loci

Katsuhito Yasuno[1,2], Kaya Bilguvar[1,2], Philippe Bijlenga[3], Siew-Kee Low[4], Boris Krischek[5], Georg Auburger[6], Matthias Simon[7], Dietmar Krex[8], Zulfikar Arlier[1,2], Nikhil Nayak[1,2], Ynte M Ruigrok[9], Mika Niemela[10], Atsushi Tajima[11], Mikael von und zu Fraunberg[12], Tamas Doczi[13], Florentina Wirjatijasa[6], Akira Hata[14], Jordi Blasco[15], Agi Oszvald[16], Hidetoshi Kasuya[17], Gulam Zilani[18], Beate Schoch[19], Pankaj Singh[20], Carsten Stüer[21], Roelof Risselada[22], Jürgen Beck[16], Teresa Sola[23], Filomena Ricciardi[6], Arpo Aromaa[24], Thomas Illig[25], Stefan Schreiber[26], Cornelia M van Duijn[27], Leonard H van den Berg[9], Claire Perret[28], Carole Proust[28], Constantin Roder[5], Ali K Ozturk[1,2], Emília Gaál[1,2,10], Daniela Berg[29], Christof Geisen[30], Christoph M Friedrich[31], Paul Summers[18], Alejandro F Frangi[32], Matthew W State [2,33], H Erich Wichmann[25], Monique M B Breteler[27], Cisca Wijmenga[34], Shrikant Mane[35], Leena Peltonen[36], Vivas Elio[23], Miriam CJM Sturkenboom[22], Patricia Lawford[37], James Byrne[18], Juan Macho[15], Erol I Sandalcioglu[19], Bernhard Meyer[21], Andreas Raabe[16], Helmuth Steinmetz[6], Daniel Rüfenacht[38], Juha E Jääskeläinen[12], Juha Hernesniemi[10], Gabriel J E Rinkel[9], Hitoshi Zembutsu[4], Ituro Inoue[11], Aarno Palotie[36], François Cambien[28], Yusuke Nakamura[4], Richard P Lifton[2,39,*], Murat Günel[1,2,*]

1. Departments of Neurosurgery and Neurobiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA.
2. Department of Genetics, Yale Program on Neurogenetics, Yale Center for Human Genetics and Genomics, Yale University School of Medicine, New Haven, Connecticut 06510, USA.
3. Service de Neurochirurgie , Deptartment of Clinical Neurosciences, Geneva Univeristy Hospital, 1211 Geneva 4, Switzerland.
4. Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan.
5. Department of Neurosurgery, University of Tuebingen, Germany.
6. Department of Neurology, Goethe University, Frankfurt am Main, Germany.
7. Department of Neurosurgery, University of Bonn, Bonn, Germany.
8. Klinik und Poliklinik für Neurochirurgie Universitätsklinikum Carl Gustav Carus der Technischen Universität Dresden Fetscherstraße 74 01307 Dresden, Germany.
9. Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands.
10. Department of Neurosurgery, Helsinki University Central Hospital, Helsinki, P.O. Box 266, FI-00029 HUS, Finland.
11. Division of Molecular Life Science, School of Medicine, Tokai University, Shimokasuya 143, Isehara, Kanagawa 259-1193, Japan.
12. Department of Neurosurgery, Kuopio University Hospital, Kuopio FI-70211, Finland.
13. Neurosurgery, University of Pècs Medical School, Pècs, Hungary.
14. Department of Public Health, School of Medicine, Chiba University, Chiba 260-8670, Japan.
15. Department of Vascular Radiology, Hospital Clinic, Barcelona, Spain.
16. Department of Neurosurgery, Goethe University, Frankfurt am Main, Germany.
17. Department of Neurosurgery, Medical Center East, Tokyo Women's University, Tokyo 116-8567, Japan.
18. Nuffield Department of Surgery, John Radcliffe Hospital, University of Oxford, Oxford, UK.
19. Department of Neurosurgery, University Hospital, Essen, Germany.
20. Departments of Medical Physics and Neurosurgery, Royal Hallamshire Hospital, Sheffield, UK.
21. Department of Neurosurgery, Technical University of Munich, Germany.
22. Department of Medical Informatics, Erasmus University Medical Center, 3000CA Rotterdam, The Netherlands.
23. Therapeutic Neuroangiography, Hospital General de Catalunya, San Cugat del Valles, Spain.
24. Department of Health and Functional Capacity, National Public Health Institute, Helsinki, Finland.
25. Institute of Epidemiology, German Research Center for Environmental Health, Helmholtz Zentrum München, Munich, Germany.
26. Institute for Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany.

27. Genetic Epidemiology Unit, Department of Epidemiology and Biostatistics and Department of Clinical Genetics, Erasmus Medical Center, 2040, 3000 CA Rotterdam, The Netherlands.
28. UMR INSERM S937 - University Pierre and Marie Curie, Paris 06, France.
29. Center of Neurology, Department of Neurodegeneration and Hertie Institute for Clinical Brain Research, University of Tuebingen, Germany.
30. Institute of Transfusion Medicine and Immunohaematology, Department of Molecular Haemostasis, DRK Blood Donor Service Baden Wuerttemberg and Hessen, Frankfurt am Main, Germany.
31. Fraunhofer-Institut for Algorithms and Scientific Computing, 53754 Sankt Augustin, Germany.
32. Center for Computational Imaging & Simulation Technologies in Biomedicine, Universitat Pompeu Fabra, Barcelona, Spain.
33. Department of Psychiatry and Child Study Center, Yale University School of Medicine, New Haven, Connecticut 06510, USA.
34. Department of Genetics, University Medical Center Groningen and University of Groningen, 9700 RR Groningen, The Netherlands.
35. Keck Foundation Biotechnology Resource Laboratory, Yale University, 300 George Street, New Haven, Connecticut 06510, USA.
36. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1HH, UK.
37. Academic Unit of Medical Physics, School of Medicine and Biomedical Sciences, University of Sheffield, Sheffield, UK.
38. Neuroradiology - SNI - Clinic Hirslanden, 8032 Zürich, Switzerland.
39. Howard Hughes Medical Institute and Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

* Correspondence should be addressed to R.P.L. (richard.lifton@yale.edu) or M.G. (murat.gunel@yale.edu).

Supplementary Notes

Study subjects

Phenotype description: In all cases, diagnosis of an intracranial aneurysm (IA) was made either with computerized tomography angiogram, magnetic resonance angiogram or cerebral digital subtraction angiogram and confirmed at surgery, when applicable. Rupture of an aneurysm was defined by identification of acute subarachnoid or intracranial hemorrhage (through computerized tomography or magnetic resonance imaging) from a proven aneurysm. Subjects with subarachnoid hemorrhage without saccular IA, non-saccular IA (such as fusiform aneurysms) and those with known genetic syndromes that are believed to predispose to IA (e.g. polycystic kidney disease and Ehlers-Danlos syndrome Type IV) were excluded from the study.

Discovery cohort: Genome-wide case-control data consisted of 11 cohorts (**Supplementary Table 2**). Four of these were reported in our prior IA genome-wide association study (GWAS)[1]: Finnish case-control cohort (FI case-control, 912 cases and 740 controls), Dutch case (NL cases, *n* = 786) and 2 control cohorts (Utrecht neurologically normal subjects[2] (*n* = 450) and the Rotterdam study[3] (*n* = 5,974)), 5 previously-genotyped control cohorts: Finland Health 2000[4] (*n* = 2,138), Northern Finnish Birth Cohort 1966[5] (*n* = 5,302), KORA-gen[6] (*n* = 840), PopGen[7] (*n* = 493) and Illumina's iControlDB (*n* = 3,182, http://www.illumina.com/science/icontroldb.ilmn), and 2 newly ascertained case series: the Germany cohort (DE cases, *n* = 975) which consisted of cases ascertained from Bonn, Dresden, Essen, Frankfurt and Tubingen, and the @neurIST case series (*n* = 641) collected from Germany, Great Britain, Hungary, The Netherlands, Switzerland and Spain.

Replication cohorts: We analyzed 2 independent Japanese case-control cohorts to confirm association signals detected in the discovery cohort. *JP1:* This cohort was described previously[1]. Since the description of this cohort, additional cases (*n* = 334) and controls (*n* = 85) were ascertained. The sample size of this cohort is 829 cases and 761 controls. *JP2:* Case subjects (*n* = 2,282) were obtained from the BioBank Japan at the Institute of Medical Science, the University of Tokyo[8] while control subjects (*n* = 905) from volunteers in the Osaka-Midosuji Rotary Club, Osaka, Japan[9].

Supplementary Table 1: Power estimates for the previous and current studies:

GRR	Previous Study: mean (range)			Current Study: mean (range)		
	Discovery	Replication	Combined	Discovery	Replication	Combined
0.05 ≤ RAF ≤ 0.95 in 3 cohorts						
1.2	0.15 (0.00-0.37)	0.15 (0.00-1.00)	0.05 (0.00-0.23)	0.63 (0.00-0.97)	0.70 (0.00-1.00)	0.56 (0.00-0.96)
1.25	0.39 (0.00-0.76)	0.28 (0.00-1.00)	0.23 (0.00-0.62)	0.83 (0.01-1.00)	0.85 (0.00-1.00)	0.80 (0.00-1.00)
1.3	0.62 (0.00-0.94)	0.42 (0.00-1.00)	0.46 (0.00-0.92)	0.91 (0.06-1.00)	0.91 (0.25-1.00)	0.90 (0.03-1.00)
1.35	0.77 (0.00-1.00)	0.54 (0.00-1.00)	0.66 (0.00-0.99)	0.96 (0.21-1.00)	0.95 (0.38-1.00)	0.95 (0.14-1.00)
0.1 ≤ RAF ≤ 0.9 in 3 cohorts						
1.2	0.17 (0.00-0.37)	0.18 (0.00-1.00)	0.06 (0.00-0.23)	0.70 (0.09-0.97)	0.77 (0.11-1.00)	0.64 (0.01-0.96)
1.25	0.44 (0.01-0.76)	0.32 (0.00-1.00)	0.27 (0.00-0.62)	0.91 (0.29-1.00)	0.91 (0.49-1.00)	0.89 (0.21-1.00)
1.3	0.69 (0.06-0.94)	0.46 (0.00-1.00)	0.54 (0.01-0.92)	0.97 (0.53-1.00)	0.97 (0.65-1.00)	0.97 (0.44-1.00)
1.35	0.84 (0.10-1.00)	0.60 (0.07-1.00)	0.75 (0.04-0.99)	0.99 (0.75-1.00)	0.99 (0.80-1.00)	0.99 (0.70-1.00)
0.2 ≤ RAF ≤0.8 in 3 cohorts						
1.2	0.21 (0.05-0.37)	0.21 (0.00-0.54)	0.09 (0.00-0.23)	0.82 (0.48-0.97)	0.85 (0.58-0.99)	0.78 (0.38-0.96)
1.25	0.54 (0.22-0.76)	0.38 (0.07-0.64)	0.35 (0.08-0.62)	0.98 (0.80-1.00)	0.97 (0.82-1.00)	0.97 (0.77-1.00)
1.3	0.80 (0.37-0.94)	0.54 (0.19-0.81)	0.67 (0.20-0.92)	1.00 (0.97-1.00)	0.99 (0.93-1.00)	1.00 (0.96-1.00)
1.35	0.94 (0.74-1.00)	0.69 (0.35-0.90)	0.87 (0.56-0.99)	1.00 (1.00-1.00)	1.00 (0.99-1.00)	1.00 (0.99-1.00)

A multiplicative model of genotypic relative risk (GRR) of a causal SNP and IA prevalence of 2% were assumed[10]. Sample size of the previous study[1] (case; control): Finnish = (874; 944); Dutch = (706; 5,332); Japanese = (495; 676). Sample size of the current study: Finnish = (808; 4,393); other combined European = (1,972; 8,122); Japanese = (3,111; 1,666). Risk allele frequencies (RAFs) of 3 cohorts were selected from 5% to 95% by increments of 5% but restricted to keep the maximum frequency difference between Finnish and other European cohorts to less than 10%. Similarly, maximum RAF difference between Japanese and one of the European (Finnish or other European) cohorts is kept less than 40%. These maximum differences were determined from the empirical distribution of allele frequencies of genotyped SNPs on chromosome 21 and contained approximately 95% of the SNPs.

For each cohort, we simulated genotype counts of cases and controls using the prevalence, RAF in the population and GRR. We then estimated the logarithm of per-allele odds ratio and the standard error using the logistic regression. To combine multiple cohorts, we used a fixed-effects model for meta-analysis. Bayes factors (BFs) and posterior probabilities of association (PPAs) were calculated as described in Wakefield[11] (see Online Methods). Power to detect association was estimated as the number of times the simulated result meets the criteria given below divided by the number of simulations which was 100.

Discovery = PPA > 0.5 in the discovery cohort (Finnish + Dutch/other European);

Replication = Conditional on Discovery, BF > 10 in the Japanese cohort with the same risk allele as that of the discovery cohort;

Combined = Discovery, PPA > 0.99 in the joint analysis of discovery and replication cohorts and PPA in the joint analysis greater than that in the discovery cohort.

Supplementary Table 2a: Sample quality control process:

Cohort*	Country	Beadchip	Prev†	Cohort Type	Initial Size	Excluded subjects‡						Final Size
						Genotyping And Information Quality	Spurious Relatedness	Cryptic Relatedness	Nearest From Many Subjects	Population Outliers	Not Matched	
FI	Finland	CNV370	Y	case	912	8	0	66	0	3	27	808
case-control				control	740	0	0	85	1	16	38	600
Health2000	Finland	Human610	N	control	2,138	22	0	66	3	31	273	1,743
NFBC1966	Finland	CNV370	N	control	5,302	104	0	1,295	4	24	1,825	2,050
NL cases	The Netherlands	CNV370	Y	case	786	20	14	23	0	21	0	708
Utrecht	The Netherlands	Hap300v1	Y	control	450	7	1	22	1	5	6	408
RS	The Netherlands	Hap550v3	Y	control	5,974	33	1	1,435	1	31	62	4,411
DE cases	Germany	Human610	N	case	975	33	3	132	0	15	3	789
KORA-gen	Germany	Hap550v3	N	control	840	14	0	19	1	6	5	795
PopGen	Germany	Hap550v3	N	control	493	2	0	11	0	3	8	469
@neurIST	see below	Human610	N	case	641	21	6	71	0	43	25	475
iControlDB	unknown	Hap550v1/v3	N	control	3,182	30	8	107	1	478	519	2,039

*Health2000 = Finland Health 2000; NFBC1966 = Northern Finnish Birth Cohort 1966; Utrecht = Utrecht neurologically normal subjects; RS = The Rotterdam study. @neurIST sample consists of subjects from Germany, Great Britain, Hungary, The Netherlands, Switzerland and Spain.
†The column indicates whether the cohort was analyzed in the previous study[1] (Y = yes, N = no).
‡Number of excluded subjects were recorded for each step as described below:
Genotyping and information quality: We excluded from analysis samples that (a) were likely to be contaminated, which were identified by examining distribution of the normalized Theta (or B allele frequency) in the BeadStudio; (b) showed greater than 3% missing genotypes (missing rate > 0.03); (c) belonged to a cluster showing systematically higher missing rates, indicating a batch effect; (d) showed discrepancy between stated and genotypic gender; (e) were outliers with respect to the estimated inbreeding coefficient (> 0.125 or < – 0.03), for which negatively large inbreeding coefficients indicate contamination, while positively large ones could be attributable to incomplete hybridization or to true inbreeding.
Spurious relatedness: Subjects who showed distant (approximately 3rd to 4th degree) relatedness with 5 or more subjects that were associated with high inbreeding coefficients were excluded. While this could be due to true relatedness, it is more likely to be due to long stretches of homozygosity resulting in spurious relationship between subjects.
Cryptic relatedness: We estimated the IBS similarity (the proportion of alleles shared identical-by-state[12]) based on nearly independent SNPs extracted by using the complete-linkage hierarchical clustering[13]. We kept only one subject from duplicates or probable relatives based on the IBS similarity.
Nearest-from-many subjects: We excluded subjects that were indicated to be the nearest neighbor for more than 100 individuals as it suggested contamination.
Population outliers: We excluded outliers based on empirical parameters for each cohort (CE and FI) which were identified using: (a) the mean IBS similarity calculated for each subject by averaging over all possible pairs; (b) the multidimensional scaling (MDS) analysis; (c) the maximum IBS similarity; and (d) the proportion of distant subjects.

Supplementary Table 2b: SNP quality control process:

QC metric	CE				FI			
	Remaining genotyped SNPs (*n*)	Genomic inflation factor (*n* with $P < 5\times10^{-8}$)	Remaining imputed SNPs (*n*)	Genomic inflation factor (*n* with $P < 5\times10^{-8}$)	Remaining genotyped SNPs (*n*)	Genomic inflation factor (*n* with $P < 5\times10^{-8}$)	Remaining imputed SNPs (*n*)	Genomic inflation factor (*n* with $P < 5\times10^{-8}$)
Initial data	294,374	--	3,859,875	--	325,268	--	3,860,940	--
Imputation quality	--	--	2,462,318	1.204 (1,348)	--	--	2,396,529	1.215 (2,820)
Genotyping quality	278,285	1.099 (6)	1,812,624	1.127 (120)	311,023	1.071 (2)	1,808,313	1.114 (145)
Non-redundant SNPs	--	--	1,652,092	--	--	--	1,624,167	--
Allelic $R^2 \geq 0.9$	--	--	1,613,978	1.125 (105)	--	--	1,548,488	1.111 (110)
Differential missingness	235,984	1.097 (6)	669,922	1.092 (30)	296,849	1.069 (1)	1,007,027	1.075 (13)
Remaining SNPs in each cohort		905,906		1.094 (36)		1,303,876		1.074 (14)
Remaining SNPs in both cohorts	831,534 (genotyped = 233,749; imputed = 597,785)							
Final total	831,532 (genotyped = 233,749; imputed = 597,783)							

Initial data:
[*Genotyped SNPs*] Number of SNPs that were genotyped in at least one individual in every sub-cohort after excluding SNPs as described in Preprocessing section in Online Methods. For the sub-cohort definition, see **Supplementary Table 2a**).
[*Imputed SNPs*] All the SNPs analyzed using IMPUTE v1.0.0 software.
Imputation quality: MAF (based on fractional allele dosage score) ≥ 0.01, allelic R^2 (squared correlation between true and imputed genotypes) ≥ 0.3, and average posterior probability for the most likely genotype ≥ 0.9.
Genotyping quality: The following conditions had to be met in each of the cohorts.
[*Genotyped SNPs*] MAF ≥ 0.01, missing rate ≤ 0.005 if $0.01 \leq$ MAF < 0.05 or missing rate ≤ 0.05 if MAF ≥ 0.05, and *P*-value of HWE test $\geq 1 \times 10^{-5}$.
[*Imputed SNPs*] (most likely genotypes with posterior probability ≥ 0.9 were used) MAF ≥ 0.01, missing rate ≤ 0.05 if $0.01 \leq$ MAF < 0.05 or missing rate ≤ 0.1 if MAF ≥ 0.05, and *P*-value of HWE test $\geq 1\times10^{-5}$.
Non-redundant SNPs: At this step, duplicated SNPs between genotyped and imputed ones were identified and the data from the imputation was excluded leading to a non-redundant, non-overlapping data set.
Allelic $R^2 \geq 0.9$: The allelic R^2 is defined as the squared correlation between true and imputed genotypes[14]. We added this filter to keep only the imputed SNPs with high prediction accuracy, uniformly across all the cohorts.
Differential missingness: SNPs were eliminated if the absolute value of the difference of the missing rates between cases and controls > 0.01 and *P*-value of the exact test of missingness $< 1 \times 10^{-6}$. This stringent filter was added as our data consisted of cohorts genotyped using various Illumina beadchips.
Remaining SNPs in each cohort: We merged the genotyped and imputed SNPs in each of the CE and FI cohorts.
Remaining SNPs in both cohorts: We extracted SNPs that passed QC filters in both the CE and FI cohorts. If a SNP was genotyped in one cohort and imputed In the other one, we classified this particular SNP as imputed.
Final total: 2 imputed SNPs, in which association with IA was implicated with PPA > 0.5, were eliminated due to lack of supporting evidence from genotyped SNPs in LD with the imputed SNPs.

Supplementary Table 3: Result of case-control matching:

CE Case Series (*n*=1,972)		The Netherlands		Germany		iControlDB (Study ID)				Control Total (#strata)
		Utrecht	RS	PopGen	KORA-gen	64	65	66	67	
NL cases	708 (35.9%)	323 (79.2%)	3,404 (77.2%)	90 (19.2%)	33 (4.2%)	19 (7.4%)	42 (6.7%)	15 (2.9%)	28 (4.4%)	3,954 (708)
DE cases	789 (40.0%)	56 (13.7%)	600 (13.6%)	341 (72.7%)	579 (72.8%)	82 (32.0%)	174 (27.7%)	179 (34.6%)	217 (34.1%)	2,228 (786)
@neurIST	475 (24.1%)	29 (7.1%)	407 (9.2%)	38 (8.1%)	183 (23.0%)	155 (60.5%)	413 (65.7%)	324 (62.5%)	391 (61.5%)	1,940 (452)
Cohort Total		408	4,411	469	795	256	629	518	636	8,122

FI Case Series (*n*=808)		Helsinki	Kuopio	Health 2000	NFBC 1966	Control Total (#strata)
Helsinki	510 (63.1%)	265 (75.5%)	92 (36.9%)	1,206 (69.2%)	1,295 (63.2%)	2,858 (510)
Kuopio	298 (36.9%)	86 (24.5%)	157 (63.1%)	537 (30.8%)	755 (36.8%)	1,535 (298)
Cohort Total		351	249	1,743	2,050	4,393

Sub-cohorts of CE were defined based on the case series: NL = NL cases (*n* = 708) and matched controls (*n* = 3,954); DE = DE cases (*n* = 789) and matched controls (*n* = 2,228); and AN = @neurIST cases (*n* = 475) and matched controls (*n* = 1,940).

Supplementary Table 4: Summary of the results for 13 SNPs genotyped both in the discovery and replication cohorts:

Locus	SNP	Position	Gene	Risk Allele	Cohort	*P*-value	$\log_{10}$(BF)	PPA	OR (95% CI)	Control RAF	Case RAF
8q11.23	rs1504749	55473264	*5'-SOX17*	C	Discovery	4.6E-07	4.60	0.7984	1.22 (1.13-1.31)	0.21/0.24	0.24/0.27
					Replication	0.095	-0.06		1.09 (0.99-1.20)	0.30	0.33
					Combined	5.4E-07	4.50	0.7604	1.17 (1.10-1.24)		
8q11.23	rs10958409	55489644	*5'-SOX17*	A	Discovery	4.2E-07	4.64	0.8128	1.24 (1.14-1.35)	0.15/0.19	0.18/0.22
					Replication	0.12	-0.11		1.08 (0.98-1.20)	0.28	0.29
					Combined	9.0E-07	4.30	0.6685	1.17 (1.10-1.25)		
8q12.1	rs9298506	55600077	*3'-SOX17*	A	Discovery	1.2E-10	7.94	0.9999	1.33 (1.22-1.45)	0.81/0.76	0.85/0.81
					Replication	0.0012	1.56		1.21 (1.08-1.36)	0.79	0.81
					Combined	1.3E-12	9.85	1.0 - 1.4E-06	1.28 (1.20-1.38)		
9p21.3	rs1333040	22073404	*CDKN2A/B*	T	Discovery	2.5E-16	13.41	1.0 - 3.9E-10	1.32 (1.24-1.41)	0.56/0.45	0.63/0.53
					Replication	1.0E-07	5.18		1.31 (1.19-1.45)	0.66	0.72
					Combined	1.5E-22	19.48	1.0 - 3.3E-16	1.32 (1.25-1.39)		
9p21.3	rs2383207	22105959	*CDKN2A/B*	G	Discovery	9.7E-15	11.89	1.0 - 1.3E-08	1.29 (1.21-1.38)	0.50/0.43	0.56/0.50
					Replication	2.6E-08	5.74		1.32 (1.20-1.45)	0.63	0.70
					Combined	1.5E-21	18.51	1.0 - 3.1E-15	1.30 (1.23-1.37)		
10q24.32	rs12411886	104675289	*CNNM2*	C	Discovery	1.0E-06	4.19	0.6089	1.38 (1.21-1.56)	0.91/0.91	0.94/0.93
					Replication (JP2)	0.0092	0.82		1.20 (1.05-1.37)	0.74	0.77
					Combined	9.5E-08	5.22	0.9437	1.29 (1.17-1.41)		
10q24.32	rs12413409	104709086	*CNNM2*	G	Discovery	7.9E-07	4.29	0.6621	1.38 (1.22-1.57)	0.91/0.91	0.94/0.93
					Replication	0.00014	2.34		1.23 (1.10-1.37)	0.74	0.77
					Combined	1.2E-09	7.00	0.9990	1.29 (1.19-1.40)		
13q13.1	rs9315204	32591837	*STARD13*	T	Discovery	3.3E-07	4.73	0.8443	1.21 (1.13-1.31)	0.21/0.33	0.24/0.39
					Replication	0.0019	1.36		1.18 (1.06-1.31)	0.24	0.27
					Combined	2.5E-09	6.72	0.9981	1.20 (1.13-1.28)		
13q13.1	rs1980781	32598374	*STARD13*	G	Discovery	3.8E-07	4.68	0.8259	1.21 (1.13-1.31)	0.21/0.33	0.24/0.39
					Replication	0.0018	1.39		1.18 (1.07-1.32)	0.24	0.27
					Combined	2.6E-09	6.70	0.9980	1.20 (1.13-1.28)		
13q13.1	rs3742321	32602065	*STARD13*	C	Discovery	2.9E-07	4.79	0.8597	1.22 (1.13-1.31)	0.21/0.33	0.24/0.39
					Replication	0.0049	1.01		1.16 (1.05-1.29)	0.24	0.27

					Combined	6.1E-09	6.35	0.9955	1.20 (1.13-1.27)		
18q11.2	rs4800418	18400738	*RBBP8*	C	Discovery	7.8E-09	6.26	0.9945	1.23 (1.15-1.32)	0.30/0.31	0.35/0.34
					Replication	0.00054	1.84		1.21 (1.09-1.36)	0.21	0.24
					Combined	1.7E-11	8.78	1.0 - 1.7E-05	1.22 (1.15-1.30)		
18q11.2	rs11662668	18433379	*RBBP8*	G	Discovery	2.4E-08	5.80	0.9844	1.20 (1.13-1.29)	0.48/0.44	0.54/0.47
					Replication	1.2E-05	3.29		1.23 (1.12-1.35)	0.60	0.64
					Combined	1.4E-12	9.81	1.0 - 1.5E-06	1.21 (1.15-1.28)		
18q11.2	rs11661542	18477693	*RBBP8*	C	Discovery	5.6E-09	6.39	0.9959	1.21 (1.14-1.30)	0.49/0.44	0.54/0.47
					Replication	4.5E-05	2.79		1.22 (1.11-1.34)	0.61	0.65
					Combined	1.1E-12	9.92	1.0 - 1.2E-06	1.22 (1.15-1.28)		

Gene column indicates a region identifier, which does not imply that the listed gene is the causal one. In the Cohort column, “Replication (JP2)” means that, for the SNP rs12411886, JP1 cohort failed to genotype so that replication cohort includes only the JP2 cohort. Risk allele frequencies (RAFs) for discovery cohort are shown as (RAF of CE)/(RAF of FI).

Supplementary Table 5: Sub-cohort results for 13 SNPs genotyped both in the discovery and replication cohorts:

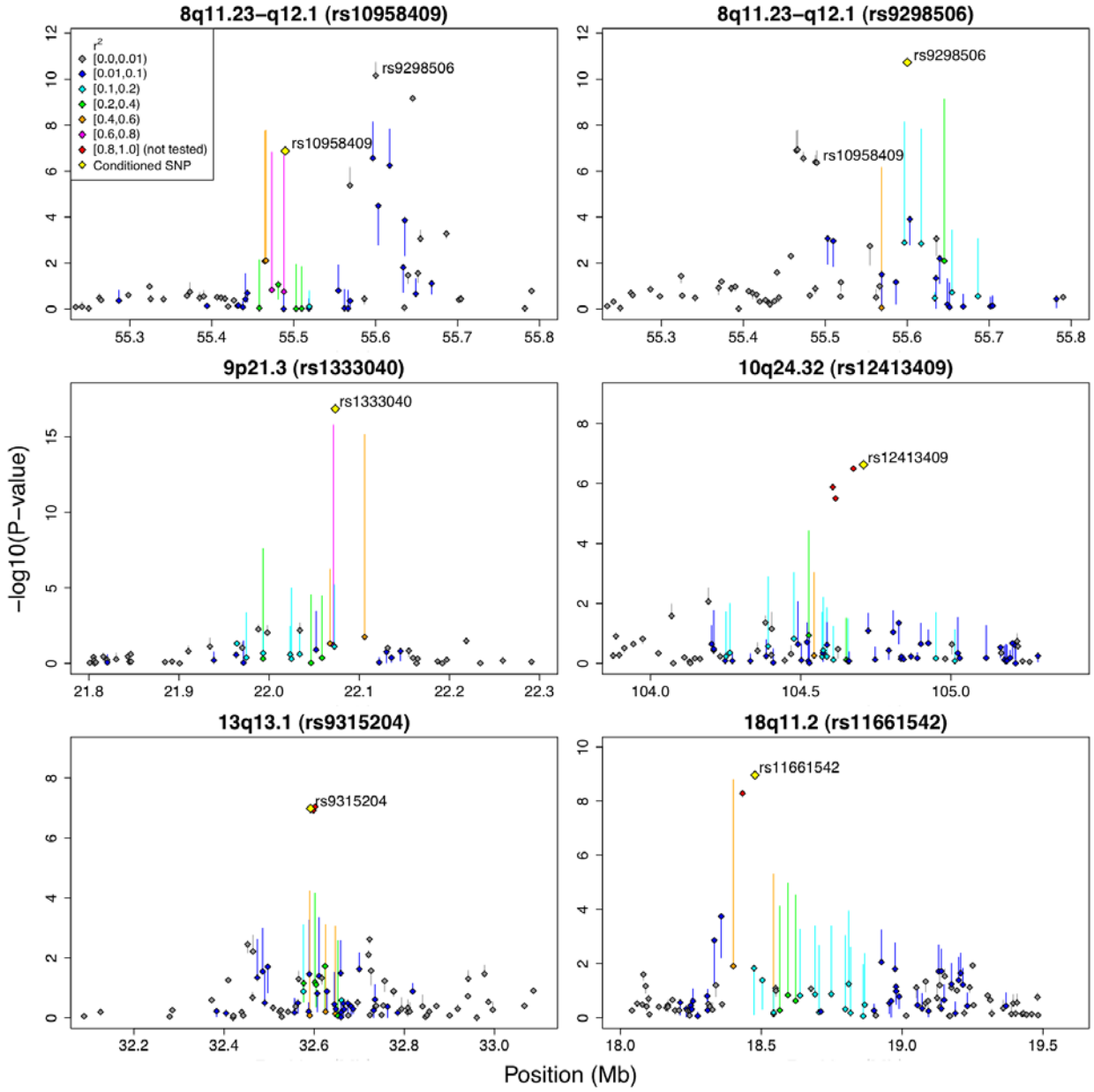
Locus	SNP	Position	Gene	Risk Allele	Cohort	P-value	log_{10}(BF)	PPA	OR (95% CI)
8q11.23	rs1504749	55473264	*5'-SOX17*	C	FI	0.0028	1.23		1.22 (1.07-1.40)
					NL	0.00083	1.67		1.28 (1.11-1.47)
					DE	0.21	-0.16		1.10 (0.95-1.28)
					AN	0.010	0.79		1.29 (1.06-1.58)
					CE	4.5E-05	2.77	0.0551	1.21 (1.10-1.33)
					JP1	0.033	0.41		1.19 (1.01-1.39)
					JP2	0.65	-0.49		1.03 (0.91-1.17)
					FI+NL	8.7E-06	3.43	0.2123	1.25 (1.13-1.37)
					DE+AN	0.011	0.75	0.0006	1.17 (1.04-1.32)
8q11.23	rs10958409	55489644	*5'-SOX17*	A	FI	0.0051	1.03		1.23 (1.06-1.41)
					NL	0.00011	2.34		1.36 (1.16-1.59)
					DE	0.30	-0.22		1.09 (0.93-1.29)
					AN	0.013	0.71		1.32 (1.06-1.65)
					CE	2.3E-05	3.02	0.0955	1.25 (1.12-1.38)
					JP1	0.061	0.23		1.17 (0.99-1.38)
					JP2	0.60	-0.46		1.04 (0.91-1.18)
					FI+NL	3.3E-06	3.81	0.3902	1.28 (1.16-1.43)
					DE+AN	0.021	0.54	0.0003	1.17 (1.02-1.33)
8q12.1	rs9298506	55600077	*3'-SOX17*	A	FI	1.0E-05	3.22		1.39 (1.20-1.61)
					NL	4.0E-05	2.66		1.43 (1.20-1.69)
					DE	0.11	0.05		1.15 (0.97-1.36)
					AN	0.011	0.76		1.34 (1.07-1.69)
					CE	1.6E-06	4.04	0.5233	1.30 (1.17-1.45)
					JP1	0.19	-0.06		1.14 (0.94-1.38)
					JP2	0.0022	1.33		1.25 (1.08-1.44)
					FI+NL	2.2E-09	6.65	0.9978	1.41 (1.26-1.57)
					DE+AN	0.0053	1.02	0.0011	1.22 (1.06-1.39)
9p21.3	rs1333040	22073404	*CDKN2A/B*	T	FI	5.3E-08	5.29		1.39 (1.23-1.56)
					NL	4.8E-05	2.72		1.30 (1.15-1.48)
					DE	5.4E-05	2.66		1.31 (1.15-1.49)
					AN	0.0080	0.88		1.25 (1.06-1.48)
					CE	3.8E-10	7.41	0.9996	1.29 (1.19-1.40)
					JP1	3.4E-05	2.76		1.41 (1.20-1.66)
					JP2	0.00040	1.94		1.26 (1.11-1.43)
					FI+NL	1.9E-11	8.69	1.0 - 2.1E-05	1.35 (1.24-1.47)
					DE+AN	1.6E-06	4.08	0.5465	1.29 (1.16-1.43)
9p21.3	rs2383207	22105959	*CDKN2A/B*	G	FI	4.7E-07	4.48		1.35 (1.20-1.52)
					NL	4.0E-05	2.79		1.30 (1.14-1.47)
					DE	1.4E-05	3.17		1.32 (1.17-1.50)
					AN	0.11	0.06		1.14 (0.97-1.34)
					CE	2.1E-09	6.73	0.9981	1.27 (1.17-1.37)

					JP1	0.00085	1.65		1.30 (1.12-1.52)
					JP2	7.2E-06	3.43		1.33 (1.17-1.50)
					FI+NL	1.1E-10	7.97	0.9999	1.33 (1.22-1.44)
					DE+AN	1.1E-05	3.34	0.1783	1.25 (1.13-1.38)
10q24.32	rs12411886	104675289	*CNNM2*	C	FI	0.043	0.36		1.26 (1.01-1.59)
					NL	0.12	0.10		1.22 (0.95-1.56)
					DE	4.4E-05	2.29		1.67 (1.30-2.14)
					AN	0.022	0.49		1.44 (1.05-1.97)
					CE	5.1E-06	3.45	0.2204	1.43 (1.23-1.67)
					JP1	NA	NA		NA (NA-NA)
					JP2	0.0092	0.82		1.20 (1.05-1.37)
					FI+NL	0.011	0.77	0.0006	1.24 (1.05-1.47)
					DE+AN	4.6E-06	3.34	0.1799	1.58 (1.30-1.92)
10q24.32	rs12413409	104709086	*CNNM2*	G	FI	0.042	0.37		1.27 (1.01-1.59)
					NL	0.096	0.15		1.23 (0.96-1.58)
					DE	4.9E-05	2.27		1.66 (1.30-2.13)
					AN	0.021	0.50		1.44 (1.05-1.97)
					CE	3.9E-06	3.55	0.2613	1.44 (1.23-1.68)
					JP1	0.0073	0.91		1.27 (1.07-1.50)
					JP2	0.0062	0.96		1.21 (1.05-1.38)
					FI+NL	0.0089	0.85	0.0007	1.25 (1.06-1.48)
					DE+AN	4.9E-06	3.32	0.1728	1.58 (1.30-1.91)
13q13.1	rs9315204	32591837	*STARD13*	T	FI	0.00017	2.25		1.27 (1.12-1.44)
					NL	0.0026	1.26		1.25 (1.08-1.45)
					DE	0.11	0.04		1.13 (0.97-1.32)
					AN	0.14	0.02		1.15 (0.95-1.40)
					CE	0.00034	1.99	0.0096	1.19 (1.08-1.30)
					JP1	0.078	0.15		1.16 (0.98-1.37)
					JP2	0.010	0.79		1.20 (1.04-1.37)
					FI+NL	1.6E-06	4.09	0.5532	1.26 (1.15-1.39)
					DE+AN	0.029	0.40	0.0003	1.14 (1.01-1.29)
13q13.1	rs1980781	32598374	*STARD13*	G	FI	0.00024	2.13		1.26 (1.11-1.43)
					NL	0.0024	1.29		1.25 (1.08-1.45)
					DE	0.090	0.09		1.14 (0.98-1.33)
					AN	0.16	-0.03		1.15 (0.95-1.39)
					CE	0.00031	2.03	0.0105	1.19 (1.08-1.30)
					JP1	0.068	0.20		1.17 (0.99-1.39)
					JP2	0.011	0.76		1.19 (1.04-1.37)
					FI+NL	2.0E-06	4.01	0.5035	1.26 (1.14-1.38)
					DE+AN	0.028	0.41	0.0003	1.14 (1.01-1.29)
13q13.1	rs3742321	32602065	*STARD13*	C	FI	0.00018	2.23		1.27 (1.12-1.43)
					NL	0.0018	1.39		1.26 (1.09-1.46)
					DE	0.096	0.08		1.14 (0.98-1.33)
					AN	0.18	-0.05		1.14 (0.94-1.38)

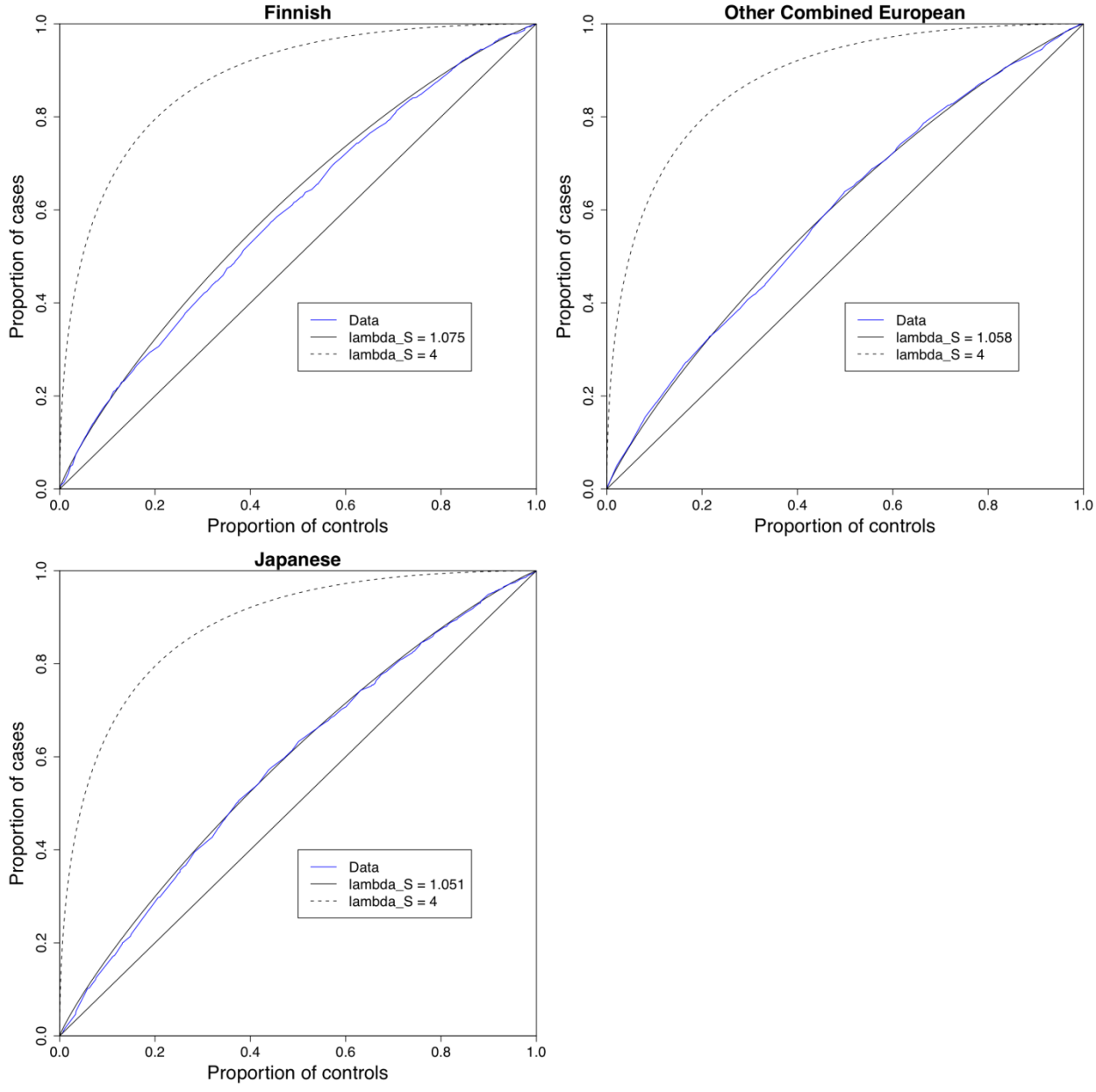
					CE	0.00030	2.04	0.0109	1.19 (1.08-1.30)
					JP1	0.14	-0.01		1.13 (0.96-1.34)
					JP2	0.015	0.65		1.18 (1.03-1.35)
					FI+NL	1.2E-06	4.21	0.6197	1.26 (1.15-1.39)
					DE+AN	0.033	0.36	0.0002	1.14 (1.01-1.29)
18q11.2	rs4800418	18400738	*RBBP8*	C	FI	0.037	0.34		1.14 (1.01-1.30)
					NL	5.0E-05	2.69		1.31 (1.15-1.50)
					DE	0.00015	2.28		1.30 (1.13-1.49)
					AN	0.14	0.0017		1.14 (0.96-1.36)
					CE	2.7E-08	5.70	0.9804	1.27 (1.17-1.38)
					JP1	0.00019	2.11		1.40 (1.17-1.68)
					JP2	0.13	-0.05		1.11 (0.97-1.28)
					FI+NL	1.7E-05	3.16	0.1266	1.22 (1.12-1.34)
					DE+AN	1.0E-04	2.48	0.0294	1.24 (1.11-1.37)
18q11.2	rs11662668	18433379	*RBBP8*	G	FI	0.037	0.31		1.13 (1.01-1.27)
					NL	0.0029	1.22		1.21 (1.07-1.37)
					DE	2.0E-05	3.04		1.32 (1.16-1.50)
					AN	0.061	0.23		1.17 (0.99-1.39)
					CE	8.6E-08	5.24	0.9460	1.24 (1.15-1.34)
					JP1	0.011	0.79		1.21 (1.05-1.41)
					JP2	0.00037	1.97		1.24 (1.10-1.40)
					FI+NL	0.00039	1.93	0.0084	1.17 (1.07-1.27)
					DE+AN	6.4E-06	3.55	0.2602	1.26 (1.14-1.40)
18q11.2	rs11661542	18477693	*RBBP8*	C	FI	0.023	0.47		1.14 (1.02-1.28)
					NL	0.00087	1.66		1.23 (1.09-1.40)
					DE	4.4E-05	2.74		1.30 (1.15-1.48)
					AN	0.044	0.33		1.19 (1.00-1.40)
					CE	3.3E-08	5.63	0.9772	1.25 (1.15-1.35)
					JP1	0.0061	0.97		1.24 (1.06-1.46)
					JP2	0.0023	1.31		1.20 (1.07-1.36)
					FI+NL	8.9E-05	2.51	0.0312	1.18 (1.09-1.29)
					DE+AN	8.6E-06	3.43	0.2121	1.26 (1.14-1.39)

Sub-cohorts of CE (NL, DE and AN) were defined based on the case series (**Supplementary Table 3**): NL = NL cases (*n* = 708) and matched controls (*n* = 3,954); DE = DE cases (*n* = 789) and matched controls (*n* = 2,228); and AN = @neurIST cases (*n* = 475) and matched controls (*n* = 1,940). FI+NL data includes only previously analyzed cases (and their matched controls), while DE+AN data includes newly analyzed cases (and their matched controls).

Supplementary Figure 1: Results of conditional analysis in the discovery cohort: Conditional analysis for each of the associated regions was performed using only the genotyped SNPs in the discovery cohort. Location of each diamond represents the *P*-value after adjusting for the effect of the conditioned SNP, while the other end of the line connected to a particular diamond shows the *P*-value of a single-locus analysis, prior to conditioning. Test statistics were not corrected using genomic control. We used SNPs listed in **Table 2** as the conditioned ones, which are indicated by yellow diamonds. The genotypic squared correlation between the conditioned SNP and a tested SNP is color-coded as indicated in the inset within the top left panel. An interval $[x,y)$ for r^2 indicates $x \leq r^2 < y$. We did not analyze very strongly correlated SNPs with $r^2 \geq 0.8$ because of multiple colinearity. The intervals shown are the same as in **Figure 2**.



Supplementary Figure 2: Receiver-Operating Characteristic curve from 5 associated loci: Prediction using 5 SNPs that are associated with IA (**Table 2**) is shown in blue for 3 ethnically distinct cohorts. These curves were obtained by fitting conditional (Finnish and other combined European) or unconditional logistic regression models as described in Online Methods and by calculating the proportion of cases and controls with risk scores exceeding each possible value. The black curves correspond to the theoretical curves for the polygenic model with the sibling recurrence risk (λ_S) = 1.075 (Finnish), 1.058 (other combined European) and 1.051, respectively[15]. The dished curves correspond to the epidemiologically estimated value, λ_S = 4[16,17].



References

1. Bilguvar, K. et al. Susceptibility loci for intracranial aneurysm in European and Japanese populations. *Nat Genet* **40**, 1472-7 (2008).
2. van Es, M.A. et al. ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. *Lancet Neurol* **6**, 869-77 (2007).
3. Hofman, A. et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol* **22**, 819-29 (2007).
4. Keskitalo, K. et al. Association of serum cotinine level with a cluster of three nicotinic acetylcholine receptor genes (CHRNA3/CHRNA5/CHRNB4) on chromosome 15. *Hum Mol Genet* **18**, 4007-12 (2009).
5. Sabatti, C. et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* **41**, 35-46 (2009).
6. Wichmann, H.E., Gieger, C. & Illig, T. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1**, S26-30 (2005).
7. Krawczak, M. et al. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet* **9**, 55-61 (2006).
8. Nakamura, Y. The BioBank Japan Project. *Clin Adv Hematol Oncol* **5**, 696-7 (2007).
9. Kamatani, Y. et al. A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. *Nat Genet* **41**, 591-5 (2009).
10. Rinkel, G.J., Djibuti, M., Algra, A. & van Gijn, J. Prevalence and risk of rupture of intracranial aneurysms: a systematic review. *Stroke* **29**, 251-6 (1998).
11. Wakefield, J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet* **81**, 208-27 (2007).
12. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
13. Rinaldo, A. et al. Characterization of multilocus linkage disequilibrium. *Genet Epidemiol* **28**, 193-206 (2005).
14. Browning, B. & Browning, S. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet* **84**, 210-23 (2009).
15. Clayton, D. Prediction and interaction in complex disease genetics: experience in type 1 diabetes. *PLoS Genet* **5**, e1000540 (2009).
16. Cannon Albright, L.A. et al. A genealogical assessment of heritable predisposition to aneurysms. *J Neurosurg* **99**, 637-43 (2003).
17. Schievink, W.I. Genetics of intracranial aneurysms. *Neurosurgery* **40**, 651-62; discussion 662-3 (1997).