Supporting Information

Yasuno et al. 10.1073/pnas.1117137108

SI Methods

Ethics. The study protocol was approved by the Yale Human Investigation Committee (HIC protocol #7680). Institutional review board approval for genetic studies, along with written consent from all study participants, was obtained at all participating institutions.

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

Discovery and Replication Cohorts. The discovery case-control samples comprised a genetically and sex-matched Finnish (FI) cohort of 808 cases and 4,393 controls, and combined European (CE) cohort of 1,972 cases and 8,122 controls. The latter cohort consisted of three subcohorts based on the centers that ascertained the case samples: the Netherlands (NL), Germany (DE), and a pan-European (AN: @neurIST) cohort. The replication cohorts included two independent Japanese case-control samples (JP1 and JP2). JP1 consisted of 829 cases and 761 controls; JP2 consisted of 2,282 cases and 905 controls. These cohorts were described in detail elsewhere (1).

Replication Strategy. We used a two-stage design to confirm association signals for loci that showed posterior probability of association (PPA) values between 0.1 and 0.5 in the discovery cohort (1). First, we selected the SNP with the maximum PPA within each region as the candidate for replication genotyping. If this was an imputed SNP and we found a genotyped SNP with a similar PPA as an alternative, we selected the genotyped one. For each region, if a second SNP that was highly correlated with the one selected above was available, we also genotyped it to assure the genotyping quality. In the first stage, we analyzed all of the candidate regions using the larger JP2 cohort (Table 1 and Table S2). For the second stage, we chose the SNPs that showed Bayes factor (BF) > 0.5 in the JP2 cohort with the same risk allele as the discovery cohort, and genotyped them using the JP1 cohort.

Genotyping and Quality Control. For SNPs reported in Table 1, we performed genotyping of the JP1 cohort using either the MassARRAY (Sequenom) assay or the Taqman (Applied Biosystems) platform. JP2 cases were genotyped using the multiplex PCR-based Invader assay (Third Wave Technologies), and JP2 controls were genotyped using the Illumina platform (2). We excluded SNPs if any of the following three conditions were met in either cases or controls: (*i*) fraction of missing genotypes > 0.1; (*ii*) *P* value of the exact test of Hardy–Weinberg equilibrium < 0.001; or (*iii*) minor allele frequency < 0.01.

Statistical Analysis. We tested for association between each SNP and IA by fitting a logistic regression model with an additive effect of allele dosage and sex as a covariate. For each SNP, we obtained a *P* value from the score test (two-sided) and estimated the logarithm of per-allele odds ratio (OR) with SE by maximizing the likelihood. For multilocus analysis (see below), we combined genotypes from JP1 and JP2 incorporating the cohort label into the above model, and analyzed the discovery cohort using the conditional logistic regression as described previously (1).

We performed meta-analysis to combine the cohort-wise results. Our primary analysis was based on the fixed-effects model (JP1 + JP2 for replication, FI + CE + JP1 + JP2 for a combined result). To assess the heterogeneity of ORs between cohorts, we first divided CE into three cohorts (NL, DE, and AN; see above), aiming to analyze data without averaging ORs over the European subcohorts, and then combined six cohorts (i.e., FI + NL + DE + AN + JP1 + JP2) using the random-effects model. We used the restricted maximum-likelihood procedure to estimate the inter-cohort heterogeneity variance (τ^2), from which we calculated Cochran's Q statistic and I^2 statistic, using R-function MiMa (3).

Evaluating the Strength of Association. Besides calculating the test *P* values, we also quantitatively measured the strength of association using the BF and PPA, which provided a probabilistic measure of the strength of the evidence (4). The BF is the ratio of the probabilities of the data under the alternative hypothesis versus the null hypothesis, which can be interpreted as the fold-change of the odds of association before and after observing the data. For computational simplicity, we approximated BF following Wakefield (5). For the prior distribution of the log-OR of every SNP, we assumed a normal distribution with a mean of 0 and SD of $\log(1.5)/\Phi^{-1}(0.975)$, where Φ is the normal distribution function (6). We regarded the association between a SNP and IA as replicated if BF > 10 in the replication cohort (i.e., 10-fold increase in the odds of association after observing replication data) (1). We assumed a uniform prior probability of association of 0.0001 across all of the SNPs (1).

Two-Locus Interaction. We tested for deviation from a linear model in which two SNPs combine to increase the log-odds of disease in an additive fashion by fitting a model with an interaction term between two SNPs in addition to linear terms. The interaction OR and 95% CI were obtained from the maximum-likelihood estimate and the interaction *P* value was obtained from the Wald test.

Cumulative Effect. We evaluated potential clinical implications of the genetic profiles of the IA risk loci following the approach described by Clayton (7). We first fitted a model with additive effects of seven loci [rs9298506, rs1333040, rs12413409, rs9315204, and rs11661542 from the previous report (1), and rs6841581 and rs6538595 from the present study] and then calculated the risk scores for every individual using the estimated log-ORs for the seven SNPs and the individual's genotypes. We depicted the receiver-operating characteristic curve for each ethnic cohort (FI, CE, and JP) by calculating the proportions of cases and controls with risk scores exceeding each possible value. The sibling recurrence risk because of the seven SNPs was estimated by assuming the polygenic model that fits well to our data (7). The fraction of the sibling recurrence risk attributable to the seven loci was calculated by taking the ratio of the logarithm of this value and an epidemiologically estimated value of 4 (8). We also calculated the ratio of the exponential of the mean of the risk scores for control subjects within the top versus bottom 5% or 1% to obtain approximated ORs of disease between these classes.

1. Yasuno K, et al. (2010) Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat Genet* 42:420–425.
2. Kamatani Y, et al. (2009) A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. *Nat Genet* 41: 591–595.
3. Viechtbauer W (2005) Bias and efficiency of meta-analytic variance estimators in the random-effects model. *J Educ Behav Stat* 30:261–293.
4. Stephens M, Balding DJ (2009) Bayesian statistical methods for genetic association studies. *Nat Rev Genet* 10:681–690.
5. Wakefield J (2007) A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet* 81:208–227.
6. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447: 661–678.
7. Clayton DG (2009) Prediction and interaction in complex disease genetics: Experience in type 1 diabetes. *PLoS Genet* 5:e1000540.
8. Schievink WI (1997) Genetics of intracranial aneurysms. *Neurosurgery* 40:651662; discussion 662–663.

Other Supporting Information Files

Table S1 (DOCX)
Table S2 (DOCX)
Table S3 (DOCX)
Table S4 (DOCX)