

## ONLINE METHODS

In each center, the local institutional review board reviewed and approved all study procedures; written informed consent was obtained from each participant, including consent to use DNA for genetic analyses of cardiovascular disease.

**Discovery cohorts.** For a detailed description of the study cohorts used in the discovery phase, please see the **Supplementary Note**.

**Replication cohorts.** Independent subjects with atrial fibrillation were identified from the German Competence Network for Atrial Fibrillation (AFNET) Study, and controls without atrial fibrillation were obtained from the KORA S4 study for the replication phase of the current study. Independent subjects with atrial fibrillation and controls without atrial fibrillation were identified from the HVH study for replication; included in the replication sample were atrial fibrillation cases of  $\geq 66$  years of age or with clinically recognized structural heart disease at atrial fibrillation diagnosis and referent subjects without atrial fibrillation, who were frequency matched to atrial fibrillation cases on the basis of age, sex, hypertension and year of identification. Independent subjects with early-onset atrial fibrillation were identified from the Massachusetts General Hospital (MGH) Atrial Fibrillation Study; referent subjects without atrial fibrillation were drawn from the local hospital catchment.

The Health Aging and Body Composition (Health ABC) Study is a National Institute on Aging-sponsored ongoing cohort study of the factors that contribute to incident disability and the decline in function of healthier older persons, with a particular emphasis on changes in body composition in old age. Health ABC enrolled well-functioning, community-dwelling black ( $n = 1,281$ ) and white ( $n = 1,794$ ) men and women aged 70–79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all-black Medicare-eligible residents in the Pittsburgh and Memphis, Tennessee, metropolitan areas. The key components of Health ABC include a baseline exam, annual follow-up clinical exams and phone contacts every 6 months to identify major health events and document functional status between clinic visits.

The Malmö Study consists of cases with prevalent or incident atrial fibrillation from two population-based cohorts from Malmö, Sweden, (Malmö Diet and Cancer and the reexamination of the Malmö Preventive Project identified from national registers as previously described) that were matched 1:1 to controls from the same cohort by sex, age ( $\pm 1$  year), date of baseline exam ( $\pm 1$  year) and requirement for a follow-up exam exceeding that for the corresponding case.

The Ottawa Heart Atrial Fibrillation study consists of individuals with lone atrial fibrillation or atrial fibrillation and hypertension, recruited from the Arrhythmia Clinic at the University of Ottawa Heart Institute (UOHI). Enrollment requires at least one episode of electrocardiographically documented atrial fibrillation, characterized by erratic atrial activity without distinct P waves and irregularly irregular QRS intervals. Exclusion criteria consist of a history of coronary artery disease, a left ventricular ejection fraction of  $< 50\%$  or significant valvular disease on echocardiography. Control subjects were drawn from the control arm of the Ottawa Heart Genomics Study, an ongoing case-control study for coronary artery disease at the UOHI. Male control subjects were  $\geq 65$  years of age, and female controls were  $\geq 70$  years of age. Control subjects with a documented history of atrial fibrillation were excluded from this study. All cases and controls were of western European ancestry.

**Atrial fibrillation GWAS in Japanese.** We used 843 atrial fibrillation cases who participated in the BioBank Japan project between 2003 and 2006. Control subjects consisted of 2,444 Japanese individuals registered in BioBank Japan as subjects with 11 diseases (hepatic cirrhosis, osteoporosis, colorectal cancer, breast cancer, prostate cancer, lung cancer, uterine myoma, amyotrophic lateral sclerosis, drug eruption, gallbladder and bile duct cancer and pancreatic cancer) and 906 healthy volunteers recruited from the Osaka Midosuji Rotary Club.

Illumina Human610-Quad and HumanHap550v3 Genotyping BeadChips were used for case and control groups, respectively. We applied quality control criteria (call rate of  $\geq 0.99$  in both cases and controls and Hardy-Weinberg equilibrium test  $P$  of  $\geq 1.0 \times 10^{-6}$  in the control population); 430,963 SNPs on all chromosomes passed the quality control filters. All cluster plots were examined by visual inspection by trained personnel to exclude SNPs with ambiguous calls.

**Genotyping.** Detailed information on the genotyping platforms and exclusions in each cohort for the GWAS meta-analysis are provided (**Supplementary Table 1**). Replication genotyping was performed using TaqMan assays (Applied Biosystems) in the Ottawa sample or Sequenom iPLEX single-base primer extensions with MALDI-TOF mass spectrometry (Sequenom) for AFNET/KORA S4, HVH, Malmö and MGH samples.

**Statistical analysis.** For the meta-analysis, over 2.5 million HapMap SNPs were imputed within each study using the HapMap CEU population. MACH v1.0.1x was used by AFNET, AGES, Rotterdam (RS-1), Vanderbilt, MGH, FHS, ARIC, the Cleveland Clinic (CC) and WGH; BIMBAM was used by HVH and CHS; and IMPUTE v0.5 was used by SHIP. In studies for which population structure was associated with the atrial fibrillation phenotype (FHS, MGH and Cleveland Clinic), analyses were adjusted for the principal components of genotype associated with phenotype<sup>28</sup>. The primary analysis in each center used logistic or proportional hazards regression, as appropriate, adjusting for age at DNA draw and sex. ARIC and CHS also adjusted for study site. Each SNP was modeled using an additive genetic effect. The ratio of observed-to-expected variance in the imputed SNP genotype counts<sup>29</sup>, the MACH Rsq statistic, which is a variation on this metric, or a measure of the observed statistical information associated with the imputed genotype that was computed by IMPUTE, was used as a quality control metric for imputed SNPs. All three metrics range from 0 to 1, with 1 indicating high imputation quality and 0 indicating no imputation information. For each SNP, all studies with quality scores greater than 0.10 were included in meta-analyses. For each SNP, a fixed-effects model was used for meta-analysis of the genotype logistic regression parameters (log odds ratios), using inverse-variance weights as implemented in the meta-analysis utility METAL. Before meta-analysis, genomic control was applied to each study having a genomic control inflation factor ( $\lambda$ ) of  $> 1.0$ , by multiplying the standard error of the SNP regression parameter by the square root of the study-specific  $\lambda$  value. A total of 2,609,549 SNPs with average minor allele frequencies of  $\geq 0.01$  across participating studies were included in meta-analyses. We preset a threshold of  $P < 5 \times 10^{-8}$  corresponding to Bonferroni adjustment for 1 million independent tests as our criterion for genome-wide significance<sup>30</sup>. Our preset criterion for replication was that the meta-analysis of the discovery and replication studies would have a smaller  $P$  value than the discovery meta-analysis.

**Prediction of SNP function and eQTL analyses.** The proxy of each of the three previously published and seven newly identified top SNPs was obtained from SNAP Proxy Search. The HapMap (release 22) CEU population was used as the reference panel, and the  $r^2$  threshold was 0.8. We limited the maximum physical distance to 500 kb. The seven newly identified top SNPs, along with their proxies, were then used for SNP function and eQTL analysis. eQTL analysis was performed by searching against the Genotype-Tissue Expression eQTL browser, which compiled data sets collected from multiple studies. Functional annotation of these SNPs was obtained from the dbSNP database. Nonsynonymous SNPs were selected and submitted to PolyPhen-2 and SIFT for functional effect prediction.

28. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).

29. de Bakker, P.I. *et al.* Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **17**, R122–R128 (2008).

30. International HapMap Consortium. The International HapMap Project. *Nature* **426**, 789–796 (2003).