LETTERS

Meta-analysis identifies six new susceptibility loci for atrial fibrillation

Patrick T Ellinor[1–4,75], Kathryn L Lunetta[5,6,75], Christine M Albert[4,7,8,75], Nicole L Glazer[9,75], Marylyn D Ritchie[10,11,75], Albert V Smith[12,13,75], Dan E Arking[14,75], Martina Müller-Nurasyid[15–17,75], Bouwe P Krijthe[18,19,75], Steven A Lubitz[1,7,75], Joshua C Bis[9,75], Mina K Chung[20,21,75], Marcus Dörr[22,75], Kouichi Ozaki[23,75], Jason D Roberts[24], J Gustav Smith[25,26], Arne Pfeufer[27,28], Moritz F Sinner[1,6,15], Kurt Lohman[29], Jingzhong Ding[30], Nicholas L Smith[9,31–33], Jonathan D Smith[20,34], Michiel Rienstra[1,35], Kenneth M Rice[36], David R Van Wagoner[20,21], Jared W Magnani[6,37], Reza Wakili[15], Sebastian Clauss[15], Jerome I Rotter[38], Gerhard Steinbeck[15], Lenore J Launer[39], Robert W Davies[40], Matthew Borkovich[24], Tamara B Harris[39], Honghuang Lin[37], Uwe Völker[41], Henry Völzke[42], David J Milan[1,2], Albert Hofman[18,19], Eric Boerwinkle[43], Lin Y Chen[44], Elsayed Z Soliman[45], Benjamin F Voight[26], Guo Li[9], Aravinda Chakravarti[14], Michiaki Kubo[46], Usha B Tedrow[4,7,8], Lynda M Rose[7], Paul M Ridker[4,7,8], David Conen[47], Tatsuhiko Tsunoda[48], Tetsushi Furukawa[49], Nona Sotoodehnia[9,50], Siyan Xu[5,51], Naoyuki Kamatani[52], Daniel Levy[6], Yusuke Nakamura[53], Babar Parvez[54], Saagar Mahida[1], Karen L Furie[4,55], Jonathan Rosand[3,4,55], Raafia Muhammad[54], Bruce M Psaty[9,32,33,56], Thomas Meitinger[27,28], Siegfried Perz[57], H-Erich Wichmann[58–60], Jacqueline C M Witteman[18,19], W H Linda Kao[61,62], Sekar Kathiresan[26,63], Dan M Roden[54,64], Andre G Uitterlinden[18,19,65], Fernando Rivadeneira[18,19,65], Barbara McKnight[36], Marketa Sjögren[66], Anne B Newman[67], Yongmei Liu[68], Michael H Gollob[24], Olle Melander[66], Toshihiro Tanaka[23,76], Bruno H Ch Stricker[17,19,65,69,70,76], Stephan B Felix[22,76], Alvaro Alonso[71,76], Dawood Darbar[54,76], John Barnard[72,76], Daniel I Chasman[4,7,76], Susan R Heckbert[9,32,33,76], Emelia J Benjamin[6,37,51,73,76], Vilmundur Gudnason[12,13,76] & Stefan Kääb[15,74,76]

Atrial fibrillation is a highly prevalent arrhythmia and a major risk factor for stroke, heart failure and death[1]. We conducted a genome-wide association study (GWAS) in individuals of European ancestry, including 6,707 with and 52,426 without atrial fibrillation. Six new atrial fibrillation susceptibility loci were identified and replicated in an additional sample of individuals of European ancestry, including 5,381 subjects with and 10,030 subjects without atrial fibrillation ($P < 5 \times 10^{-8}$). Four of the loci identified in Europeans were further replicated *in silico* in a GWAS of Japanese individuals, including 843 individuals with and 3,350 individuals without atrial fibrillation. The identified loci implicate candidate genes that encode transcription factors related to cardiopulmonary development, cardiac-expressed ion channels and cell signaling molecules.

Genome-wide association studies in individuals of European descent have identified three genomic regions associated with atrial fibrillation on chromosomes 4q25 (*PITX2*)[2–5], 16q22 (*ZFHX3*)[3,5] and 1q21 (*KCNN3*)[2]. Studies of electrocardiographic traits have also identified a number of loci associated with atrial fibrillation[6,7]. However, despite these findings, much of the heritability of atrial fibrillation remains unexplained, justifying the search for additional genetic variants underlying atrial fibrillation risk[8]. Large-scale meta-analysis of GWAS results is a powerful method to identify additional genetic variation underlying traits and conditions. We therefore conducted a meta-analysis of multiple well-phenotyped GWAS samples of European ancestry to identify additional atrial fibrillation susceptibility loci.

Six prospective cohort and 10 prevalent study samples contributed to the discovery analysis, which was adjusted for age and sex (**Table 1**, Online Methods and **Supplementary Note**). Atrial fibrillation status was systematically ascertained in each sample (Online Methods and **Supplementary Note**). After application of quality control SNP exclusion criteria in each study (**Supplementary Table 1**), meta-analysis was performed, applying genomic control to each study. The genomic control inflation factor for the meta-analysis was 1.042 for the full set of SNPs and 1.040 after omitting all SNPs within 500 kb of the association signals that reached genome-wide significance.

Table 1 Subject characteristics

Cohort	Cohort type	Participants (n)	AF (n)	Males (%)	Age at DNA collection (mean ± s.d.)	Age at DNA collection (range)	Age of AF onset (mean ± s.d.)	Hypertension n (%)	Body mass index, in kg/m² (mean ± s.d.)	Diabetes n (%)	Myocardial infarction n (%)	Heart failure n (%)
Incident AF	Overall N											
ARIC	Cohort	8,890	802	4,181 (47.0)	54.3 ± 5.7	44–66	67.8 ± 6.8	2,376 (26.7)	27.0 ± 4.8	763 (8.6)	354 (4.0)	325 (3.7)
AGES	Cohort	2,959	158	1,154 (39.0)	76.5 ± 5.5	66–95	75.0 ± 8.8	2,595 (87.7)	27.1 ± 4.4	319 (10.8)	189 (6.39)	55 (1.9)
CHS	Cohort	3,204	764	1,242 (38.8)	72.2 ± 5.3	65–98	81.2 ± 6.0	1,678 (52.4)	26.3 ± 4.4	377 (11.8)	0	0
FHS	Cohort	4,062	310	1,771 (43.6)	64.7 ± 12.6	31–101	77.8 ± 10.6	2,001 (49.3)	27.7 ± 5.2	328 (8.07)	231 (5.7)	55 (1.4)
RS-I	Cohort	5,665	542	2,282 (40.3)	69.1 ± 9.0	55–99	77.6 ± 7.7	1,866 (32.9)	26.3 ± 3.7	567 (10.0)	632 (11.2)	156 (2.8)
WGHS	Cohort	20,836	648	0	54.1 ± 7.0	43–89	68.0 ± 8.2	5,022 (24.1)	25.9 ± 4.9	503 (2.4)	0	17 (0.1)
Prevalent AF	N per group											
AFNET	Cases	468	–	236 (50.4)	51.8 ± 7.2	29–74	51.3 ± 7.6	252 (53.8)	28.0 ± 4.9	36 (7.7)	6 (1.8)	14 (4.8)
KORA	Referents	438	–	219 (50.0)	56.2 ± 7.1	45–69	–	185 (42.2)	27.7 ± 4.5	37 (8.4)	6 (1.3)	13 (2.9)
AGES[a]	Cases	241	–	88 (55.7)	78.5 ± 5.9	67–95	80.9 ± 6.2	143 (90.5)	27.7 ± 4.4	20 (12.7)	8 (5.1)	5 (3.2)
	Referents	2,718	–	70 (36.1)	76.1 ± 5.4	66–94	80.4 ± 5.4	2,002 (78.2)	27.0 ± 4.5	269 (10.5)	122 (4.8)	27 (1.1)
CC	Cases	496	–	375 (75.6)	58.8 ± 10.7	20–84	51.7 ± 12.0	269 (54.2)	30.2 ± 6.2	28 (5.6)	0 (0)	0 (0)
	Referents	2,971	–	1,124 (37.8)	28.5 ± 22.2	0–87	–	–	–	–	–	–
HVH	Cases	95	–	28 (29.5)	59.5 ± 6.5	40–68	57.4 ± 6.4	50 (52.6)	34.1 ± 9.9	14 (14.7)	0	0
	Referents	193	–	106 (54.9)	59.5 ± 6.0	40–69	–	153 (79.3)	31.4 ± 7.2	31 (16.1)	9 (4.7)	7 (3.6)
CHS[a]	Cases	67	–	38 (56.7)	76.3 ± 5.8	66–90	–	35 (52.2)	26.6 ± 4.3	14 (20.9)	0	0
	Referents	3,204	–	1,242 (38.8)	72.2 ± 5.3	65–98	–	1,678 (52.4)	26.3 ± 4.4	377 (11.8)	0	0
FHS[a]	Cases	253	–	151 (59.7)	76.9 ± 9.9	45–97	70.9 ± 10.8	180 (71.1)	27.4 ± 4.8	41 (16.21)	60 (23.7)	57 (0.23)
	Referents	4,151	–	1,807 (43.5)	64.7 ± 12.6	31–101	–	2,036 (49.1)	27.7 ± 5.2	329 (7.9)	235 (5.7)	55 (1.32)
MGH	Cases	366	–	295 (80.6)	53.4 ± 10.5	21–77	46.1 ± 11.7	85.8 (22.7)	27.8 ± 5.0	12 (3.2)	4 (1.1)	10 (2.8)
MIGEN	Referents	911	–	485 (53.2)	47.9 ± 8.8	18–83	–	–	–	–	–	–
RS-I[a]	Cases	309	–	145 (46.9)	76.2 ± 8.7	56–98	–	131 (42.4)	25.9 ± 3.6	64 (20.7)	69 (22.3)	54 (17.5)
	Referents	5,665	–	2,282 (40.3)	69.1 ± 9.0	55–99	–	1,866 (32.9)	26.3 ± 3.7	567 (10.0)	632 (11.2)	156 (2.8)
SHIP	Cases	107	–	69 (64.5)	65.1 ± 11.5	21–81	–	59 (55.1)	29.6 ± 5.1	23 (21.5)	14 (13.1)	44 (41.1)
	Referents	1,816	–	906 (49.9)	50.7 ± 14.9	21–81	–	437 (24.1)	27.2 ± 4.5	131 (7.2)	54 (3.0)	157 (8.6)
Vanderbilt	Cases	1,081	–	738 (68.3)	59.5 ± 12.6	16–87	51.5 ± 14.7	625 (57.8)	30.7 ± 6.9	197 (18.2)	95 (8.8)	191 (17.6)
	Referents	880	–	551 (62.6)	50.0 ± 17.4	18–91	–	463 (52.6)	28.0 ± 5.8	180 (20.5)	299 (34.0)	120 (13.6)

AF, atrial fibrillation.
[a]Studies included community-based prospectively ascertained prevalent atrial fibrillation cases. The other prevalent atrial fibrillation studies were of a case-control design.

The quantile-quantile plot of the expected versus observed *P*-value distributions for association of the 2,609,549 SNPs analyzed is shown (**Supplementary Fig. 1**). We identified ten loci that exceeded our preset threshold for genome-wide significance ($P < 5 \times 10^{-8}$) (**Fig. 1**). The three loci most significantly associated with atrial fibrillation were at previously identified atrial fibrillation susceptibility loci on chromosomes 4q25 near *PITX2* (rs6817105; $P = 1.8 \times 10^{-74}$)[4], 16q22 in *ZFHX3* (rs2106261; $P = 3.2 \times 10^{-16}$)[3,5] and 1q21 in *KCNN3* (rs6666258; $P = 2.0 \times 10^{-14}$)[2] (**Table 2**).

Seven new genomic loci were associated with atrial fibrillation with $P < 5 \times 10^{-8}$ in the discovery stage (**Table 2**). The most significantly associated SNP in each of the seven new loci was genotyped and tested for association with atrial fibrillation in an additional 3,132 to 5,289 independent individuals with atrial fibrillation and 8,159 to 11,148 referent individuals derived from six studies of individuals of European ancestry (**Supplementary Table 2**). Six of the loci associated with atrial fibrillation in the discovery stage met our criteria for independent replication. Study-specific replication results are detailed (**Supplementary Table 3**). The results from meta-analysis of the discovery and replication results are shown (**Table 2**), as are regional plots (**Fig. 2**). Recognizing that the genes in closest physical proximity to the associated SNPs are not always the causative genes, we report below the genetic associations in order of statistical significance and describe the nearest gene.

The most significant new association in the discovery stage was on chromosome 1q24 (rs3903239; overall $P = 8.4 \times 10^{-14}$) in *PRRX1*, which encodes a homeodomain transcription factor highly expressed in the developing heart, particularly in connective tissue. Biological interaction between *PRRX1* and a related homeobox transcription factor gene, *PRRX2*, results in abnormalities of great vessel development in a mouse knockout model[9]. In a separate *PRRX1* knockout, fetal pulmonary vasculature development was impaired[10].

A second locus was identified on chromosome 7q31 (rs3807989; overall $P = 3.6 \times 10^{-12}$) in *CAV1*, which encodes caveolin-1, a cellular membrane protein involved in signal transduction. *CAV1* is selectively expressed in the atria[11], and its knockout has been associated with dilated cardiomyopathy[12]. The CAV1 protein colocalizes with and negatively regulates the activity of KCNH2 (ref. 13), a potassium
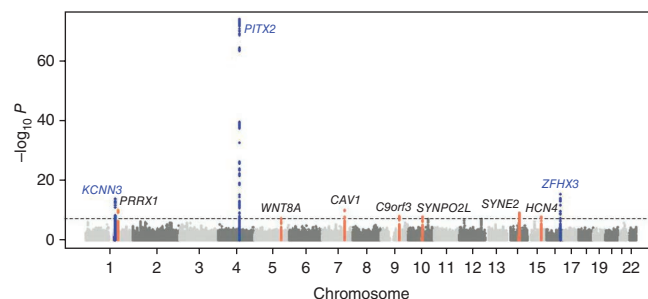

Figure 1 Manhattan plot of meta-analysis results for genome-wide association with atrial fibrillation. The $-\log_{10}$ (*P* value) is plotted against the physical position of each SNP on each chromosome. The threshold for genome-wide significance, $P < 5 \times 10^{-8}$, is indicated by the dashed line. The three previously reported loci for atrial fibrillation are indicated in blue, and the seven new loci that exceeded the genome-wide significance threshold are indicated in orange.

Table 2 Summary of GWAS meta-analysis results with $P < 5 \times 10^{-8}$

						Discovery			Replication		Overall	
SNP	Locus	Closest gene	SNP location relative to closest gene	Minor/ major allele	MAF (%)	RR (95% CI)	Meta P value	I^2 (%), P value	RR (95% CI)	Meta P value	RR (95% CI)	Meta P value
rs6666258	1q21	KCNN3-PMVK	Intronic	C/G	29.9	1.18 (1.13–1.23)	2.0×10^{-14}	42.3, 0.04	–	–	–	–
rs3903239	1q24	PRRX1	46 kb upstream	G/A	44.7	1.14 (1.10–1.18)	9.1×10^{-11}	53.2, 6.3×10^{-3}	1.13 (1.06–1.20)	2.0×10^{-4}	**1.14 (1.10–1.17)**	**8.4×10^{-14}**
rs6817105	4q25	PITX2	150 kb upstream	C/T	13.1	1.64 (1.55–1.73)	1.8×10^{-74}	80.8, 1.4×10^{-10}	–	–	–	–
rs2040862	5q31	WNT8A	Intronic	T/C	17.8	1.15 (1.09–1.21)	3.2×10^{-8}	10, 0.34	1.04 (0.96–1.12)	3.6×10^{-1}	1.12 (1.07–1.17)	2.5×10^{-7}
rs3807989	7q31	CAV1	Intronic	A/G	40.4	0.88 (0.84–0.91)	9.6×10^{-11}	10, 0.34	0.93 (0.88–0.97)	2.7×10^{-3}	**0.90 (0.87–0.92)**	**3.6×10^{-12}**
rs10821415	9q22	C9orf3	Intronic	A/C	42.4	1.13 (1.08–1.18)	7.9×10^{-9}	49.5, 0.015	1.09 (1.04–1.15)	7.2×10^{-4}	**1.11 (1.08–1.15)**	**4.2×10^{-11}**
rs10824026	10q22	SYNPO2L	5 kb upstream	G/A	15.8	0.85 (0.81–0.90)	1.7×10^{-8}	37.9, 0.06	0.91 (0.83–0.99)	3.5×10^{-2}	**0.87 (0.83–0.91)**	**4.0×10^{-9}**
rs1152591	14q23	SYNE2	Intronic	A/G	47.6	1.13 (1.09–1.18)	6.2×10^{-10}	25.7, 0.16	1.12 (1.06–1.19)	1.9×10^{-4}	**1.13 (1.09–1.17)**	**5.8×10^{-13}**
rs7164883	15q24	HCN4	Intronic	G/A	16.0	1.16 (1.10–1.22)	1.3×10^{-8}	0, 0.85	1.24 (1.16–1.32)	1.3×10^{-10}	**1.19 (1.14–1.24)**	**2.8×10^{-17}**
rs2106261	16q22	ZFHX3	Intronic	T/C	17.6	1.24 (1.17–1.30)	3.2×10^{-16}	58.8, 1.6×10^{-3}	–	–	–	–

MAF, minor allele frequency; RR, relative risk. I^2 represents the proportion of variability in the effect size due to between-study variability. We did not attempt replication of the previously published genetic loci associated with atrial fibrillation on chromosomes 1q21 (KCNN3)[2], 4q25 (PITX2)[4] and 16q22 (ZFHX3)[3,5].

channel involved in cardiac repolarization; the corresponding KCNH2 gene was found to be associated with atrial fibrillation in a candidate gene association study, although not in our present analysis[14]. The top SNP at the CAV1 locus identified in the current study, rs3807989, was previously identified in a GWAS of the PR and QRS intervals and related to atrial fibrillation[6,7]. The relationship between other previously reported PR-associated loci and atrial fibrillation are reported (**Supplementary Table 4**). Of note, significant associations with atrial fibrillation were observed for SNPs related to the PR interval in SOX5, TBX5, SCN5A and SCN10A[6,7].

The third locus was on chromosome 14q23 (rs1152591; overall $P = 5.8 \times 10^{-13}$) and was located in an intron of SYNE2, which encodes numerous nesprin-2 isoforms, some of which are highly expressed in the heart and skeletal muscle. Nesprin-2 localizes throughout the sarcomere and is involved in maintaining nuclear structural integrity by anchoring the nucleus to the cytoskeleton. In a candidate gene approach, mutations in SYNE2 were found to segregate in some families with Emery-Dreifuss muscular dystrophy[15], which is characterized by skeletal muscle atrophy, cardiomyopathy and cardiac conduction defects.

The fourth locus was on chromosome 9q22 (rs10821415; overall $P = 4.2 \times 10^{11}$) and was located in an ORF on chromosome 9. Genes at this locus include FBP1 and FBP2, which are important for gluconeogenesis. Autosomal recessive FBP1 deficiency has been described, but cardiovascular features did not seem to be prominent[16]. Variants at 9q22 have been implicated in the regulation of height, pulmonary function, angiogenesis and attention deficit hyperactivity disorder, although rs10821415 is not in substantial linkage disequilibrium (LD) with any of these SNPs ($r^2 < 0.30$).

A fifth signal was located at 15q24 (rs7164883; overall $P = 2.8 \times 10^{-17}$) in the first intron of HCN4. The HCN4 protein is the predominant cardiac hyperpolarization-activated cyclic nucleotide–gated channel and is highly expressed in the sinoatrial node. HCN4 activity underlies the funny current (I_f) that governs cardiac pacemaking, and mutations in HCN4 have been associated with various forms of sinus nodal dysfunction[17,18].

A sixth locus on chromosome 10q22 (rs10824026; overall $P = 4.0 \times 10^{-9}$) was located 5 kb upstream of SYNPO2L and 20 kb upstream of MYOZ1. The proteins encoded by SYNPO2L and MYOZ1 are both expressed in skeletal and cardiac muscle, localize to the Z-disc and interact with numerous other proteins. However, the precise role of either gene in cardiovascular physiology is unknown[19,20]. A mouse knockout of MYOZ1 showed increased calcineurin activity and cardiac hypertrophy in response to pressure overload. However, candidate gene approaches have not supported a prominent role for MYOZ1 mutations in causing familial dilated cardiomyopathy[21]. Of note, the SYNPO2L locus is located within a previously reported atrial fibrillation susceptibility locus identified in a family with autosomal dominant atrial fibrillation[22].

One other locus was identified in the meta-analysis of atrial fibrillation in the WNT8A gene (rs2040862; $P = 3.2 \times 10^{-8}$); however, this association failed to replicate in additional independent cohorts with atrial fibrillation (replication $P = 0.36$; combined $P = 2.5 \times 10^{-7}$).

There was evidence of significant heterogeneity in the discovery meta-analysis at the previously published atrial fibrillation susceptibility signals at 4q25 near PITX2 and at 16q22 in ZFHX3 (**Table 2**). Effect heterogeneity at the PITX2 locus has already been observed[4,23,24].

We then sought to determine whether the top SNPs or their proxies at each locus were associated with alterations in gene expression in an expression quantitative trait locus (eQTL) database, the Genotype-Tissue Expression eQTL browser. The top SNP at the SYNPO2L locus, rs10824026, is in strong LD with a SNP, rs12570126 ($r^2 = 0.932$), that was found to correlate with the expression of MYOZ1 and SYNPO2L ($P = 1.5 \times 10^{-6}$; data derived from lymphoblastoid cell lines in 270 individuals from the HapMap Consortium)[25]. Furthermore, the top SNP at the SYNPO2L locus is in LD with a nonsynonymous SNP in SYNPO2L, rs3812629 ($r^2 = 0.8$; encoding the SYNPO2L P707L variant with respect to transcript Q9H987-1), which is predicted to be damaging by both the PolyPhen-2 and the SIFT algorithms. None of the other identified atrial fibrillation risk SNPs were associated with variations in gene expression in the Genotype-Tissue Expression eQTL browser.
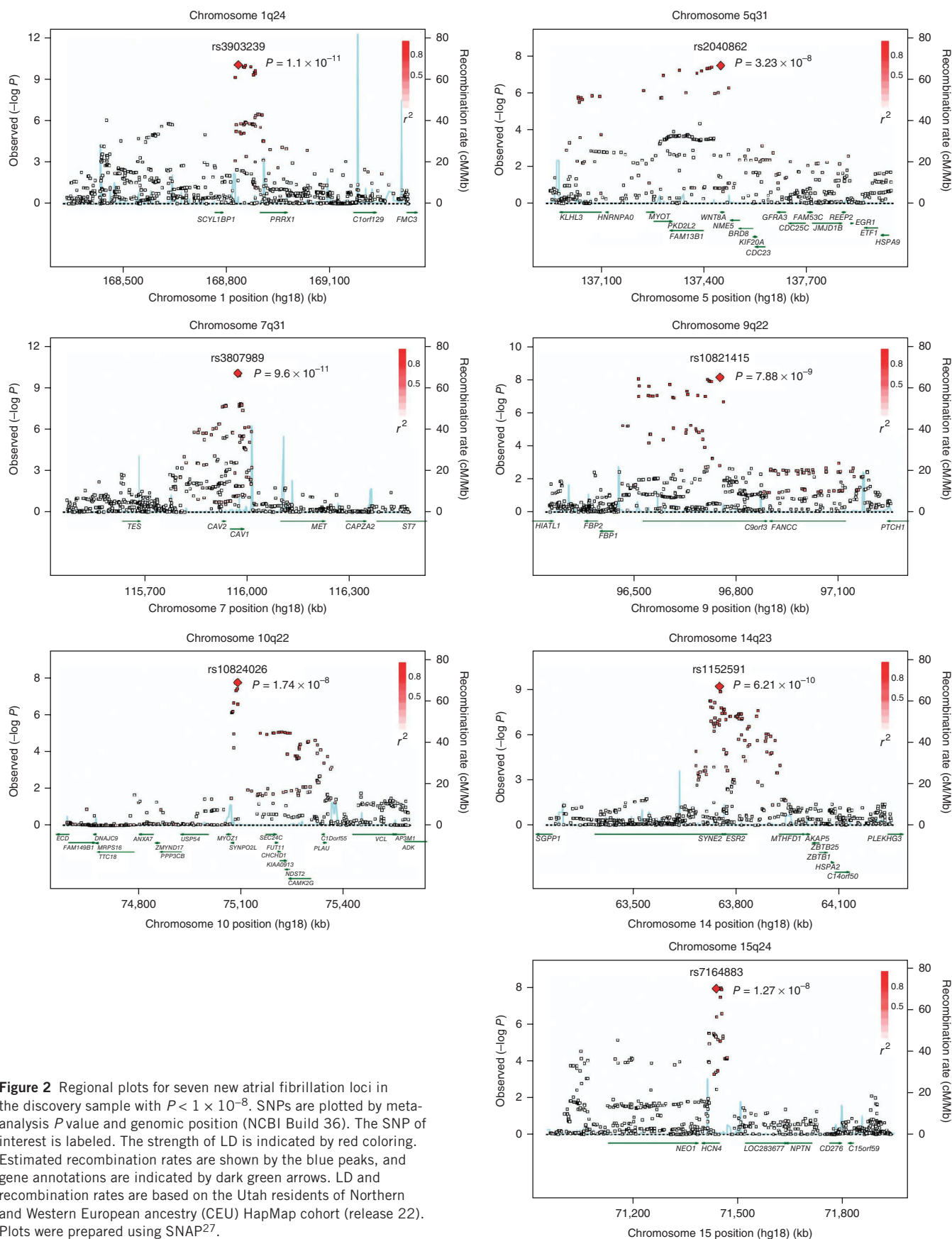
Figure 2 Regional plots for seven new atrial fibrillation loci in the discovery sample with $P < 1 \times 10^{-8}$. SNPs are plotted by meta-analysis P value and genomic position (NCBI Build 36). The SNP of interest is labeled. The strength of LD is indicated by red coloring. Estimated recombination rates are shown by the blue peaks, and gene annotations are indicated by dark green arrows. LD and recombination rates are based on the Utah residents of Northern and Western European ancestry (CEU) HapMap cohort (release 22). Plots were prepared using SNAP[27].

We next examined the generalizability of our findings by examining our results in a separate GWAS of individuals of Japanese ancestry, including 843 subjects with and 3,350 subjects without atrial fibrillation in the Japan BioBank study (**Supplementary Fig. 2a,b**). Within the Japanese GWAS, only the chromosome 4q25 (*PITX2*) locus exceeded the preset threshold for genome-wide significance (rs2634073; odds ratio (OR) = 1.84, 95% confidence interval (CI) = 1.59–2.13; $P = 3.7 \times 10^{-17}$; **Supplementary Figs. 2a** and **3** and **Supplementary Table 5**). At the previously published locus at 16q22 (*ZFHX3*)[3,5], rs12932445 was associated with atrial fibrillation in participants of Japanese ancestry ($P = 6.8 \times 10^{-4}$). The relationship between atrial fibrillation and variants at the *KCNN3-PMVK* locus on chromosome 1q21 failed to replicate ($P = 0.17$); however, a regional plot of this locus revealed a distinct signal at the rs7514452 SNP—approximately 375 kb away from rs6666258, the top SNP in the European ancestry sample—that was modestly associated with atrial fibrillation in the Japanese sample ($P = 4.9 \times 10^{-5}$; $R^2 = 0.002$). At *PRRX1* and *CAV1*, the top SNPs in the European samples were also associated with atrial fibrillation in Japanese individuals. For the loci near *PRRX1* and *C9orf3*, alternate SNPs in the Japanese cohort were more significantly associated with atrial fibrillation than the top SNP at each locus in Europeans (**Supplementary Fig. 3** and **Supplementary Table 5**).

Our study was subject to a number of limitations. To maximize both the power and the generalizability of our study, we included all available individuals with atrial fibrillation; thus, some individuals had comorbidities, such as systolic dysfunction and hypertension. However, none of the identified risk variants for atrial fibrillation were strongly associated with systolic dysfunction in the EchoGen Consortium[26], a meta-analysis of echocardiographic data from 5 community-based cohorts consisting of over 12,000 individuals of European descent (P value $< 1 \times 10^{-5}$). Further, when our replication results were adjusted for hypertension status, the identified variants remained significantly associated with atrial fibrillation (**Supplementary Table 3**). Ultimately, the development of a comprehensive risk score incorporating clinical, biochemical and genetic marker data will be necessary to clarify the incremental benefit of our findings in clinical care. Our eQTL analyses were limited to data available within the Genotype-Tissue Expression eQTL browser; future eQTL analyses in cardiac tissue may be helpful in identifying a relationship between the SNPs associated with atrial fibrillation risk and gene expression. Finally, we acknowledge that the identified variants may not be causal but may represent causal elements in the same or different molecular pathways; future statistical, bioinformatic and biological analyses investigating potential genetic interactions are warranted. Fine mapping and deep resequencing will be necessary to uncover the genetic architecture accompanying the identified common atrial fibrillation susceptibility signals.

In summary, our GWAS meta-analysis for atrial fibrillation has identified six new susceptibility loci in or near plausible candidate genes involved in pacemaking activity, signal transduction and cardiopulmonary development. Our results show that atrial fibrillation has multiple genetic associations and identifies new targets for biological investigation.

URLs. BIMBAM, http://quartus.uchicago.edu/~yguan/bimbam/index.html; Genotype-Tissue Expression eQTL browser, http://www.ncbi.nlm.nih.gov/gtex/GTEX2/gtex.cgi; iControlDB database, http://www.illumina.com/science/icontroldb.ilmn; IMPUTE, https://mathgen.stats.ox.ac.uk/impute/impute.html; MACH, http://www.sph.umich.edu/csg/abecasis/MACH/; METAL, http://www.sph.umich.edu/csg/abecasis/metal/index.html; PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/; SIFT, http://sift.jcvi.org/; SNAP, http://www.broadinstitute.org/mpg/snap/ldsearch.php.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

1. Fuster, V. *et al.* ACC/AHA/ESC 2006 Guidelines for the Management of Patients with Atrial Fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients With Atrial Fibrillation): developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *Circulation* **114**, e257–e354 (2006).
2. Ellinor, P.T. *et al.* Common variants in *KCNN3* are associated with lone atrial fibrillation. *Nat. Genet.* **42**, 240–244 (2010).
3. Benjamin, E.J. *et al.* Variants in *ZFHX3* are associated with atrial fibrillation in individuals of European ancestry. *Nat. Genet.* **41**, 879–881 (2009).
4. Gudbjartsson, D.F. *et al.* Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* **448**, 353–357 (2007).
5. Gudbjartsson, D.F. *et al.* A sequence variant in *ZFHX3* on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat. Genet.* **41**, 876–878 (2009).
6. Pfeufer, A. *et al.* Genome-wide association study of PR interval. *Nat. Genet.* **42**, 153–159 (2010).
7. Holm, H. *et al.* Several common variants modulate heart rate, PR interval and QRS duration. *Nat. Genet.* **42**, 117–122 (2010).
8. Lubitz, S.A. *et al.* Association between familial atrial fibrillation and risk of new-onset atrial fibrillation. *J. Am. Med. Assoc.* **304**, 2263–2269 (2010).
9. Bergwerff, M. *et al.* Loss of function of the *Prx1* and *Prx2* homeobox genes alters architecture of the great elastic arteries and ductus arteriosus. *Virchows Arch.* **436**, 12–19 (2000).
10. Ihida-Stansbury, K. *et al.* Paired-related homeobox gene *Prx1* is required for pulmonary vascular development. *Circ. Res.* **94**, 1507–1514 (2004).
11. Volonte, D., McTiernan, C.F., Drab, M., Kasper, M. & Galbiati, F. Caveolin-1 and caveolin-3 form heterooligomeric complexes in atrial cardiac myocytes that are

required for doxorubicin-induced apoptosis. *Am. J. Physiol. Heart Circ. Physiol.* **294**, H392–H401 (2008).

12. Zhao, Y.Y. *et al.* Defects in caveolin-1 cause dilated cardiomyopathy and pulmonary hypertension in knockout mice. *Proc. Natl. Acad. Sci. USA* **99**, 11375–11380 (2002).

13. Lin, J. *et al.* The regulation of the cardiac potassium channel (HERG) by caveolin-1. *Biochem. Cell Biol.* **86**, 405–415 (2008).

14. Sinner, M.F. *et al.* The non-synonymous coding I_{Kr}-channel variant *KCNH2*-K897T is associated with atrial fibrillation: results from a systematic candidate gene-based analysis of *KCNH2* (*HERG*). *Eur. Heart J.* **29**, 907–914 (2008).

15. Zhang, Q. *et al.* Nesprin-1 and -2 are involved in the pathogenesis of Emery Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. *Hum. Mol. Genet.* **16**, 2816–2833 (2007).

16. Matsuura, T. *et al.* Two newly identified genomic mutations in a Japanese female patient with fructose-1,6-bisphosphatase (FBPase) deficiency. *Mol. Genet. Metab.* **76**, 207–210 (2002).

17. Schulze-Bahr, E. *et al.* Pacemaker channel dysfunction in a patient with sinus node disease. *J. Clin. Invest.* **111**, 1537–1545 (2003).

18. Milanesi, R., Baruscotti, M., Gnecchi-Ruscone, T. & DiFrancesco, D. Familial sinus bradycardia associated with a mutation in the cardiac pacemaker channel. *N. Engl. J. Med.* **354**, 151–157 (2006).

19. Frey, N. & Olson, E.N. Calsarcin-3, a novel skeletal muscle–specific member of the calsarcin family, interacts with multiple Z-disc proteins. *J. Biol. Chem.* **277**, 13998–14004 (2002).

20. Beqqali, A. *et al.* CHAP is a newly identified Z-disc protein essential for heart and skeletal muscle function. *J. Cell Sci.* **123**, 1141–1150 (2010).

21. Arola, A.M. *et al.* Mutations in *PDLIM3* and *MYOZ1* encoding myocyte Z line proteins are infrequently found in idiopathic dilated cardiomyopathy. *Mol. Genet. Metab.* **90**, 435–440 (2007).

22. Brugada, R. *et al.* Identification of a genetic locus for familial atrial fibrillation. *N. Engl. J. Med.* **336**, 905–911 (1997).

23. Lubitz, S.A. *et al.* Independent susceptibility markers for atrial fibrillation on chromosome 4q25. *Circulation* **122**, 976–984 (2010).

24. Kääb, S. *et al.* Large scale replication and meta-analysis of variants on chromosome 4q25 associated with atrial fibrillation. *Eur. Heart J.* **30**, 813–819 (2009).

25. Stranger, B.E. *et al.* Population genomics of human gene expression. *Nat. Genet.* **39**, 1217–1224 (2007).

26. Vasan, R.S. *et al.* Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. *J. Am. Med. Assoc.* **302**, 168–178 (2009).

27. Johnson, A.D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**, 2938–2939 (2008).

[1]Cardiovascular Research Center, Massachusetts General Hospital, Charlestown, Massachusetts, USA. [2]Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, Massachusetts, USA. [3]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. [4]Harvard Medical School, Boston, Massachusetts, USA. [5]Department of Biostatistics, School of Public Health, Boston University, Boston, Massachusetts, USA. [6]National Heart, Lung, and Blood Institute's and Boston University's Framingham Heart Study, Framingham, Massachusetts, USA. [7]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [8]Division of Cardiology, Brigham and Women's Hospital, Boston, Massachusetts, USA. [9]Department of Medicine, Cardiovascular Health Research Unit, University of Washington, Seattle, Washington, USA. [10]Center for Systems Genomics, Pennsylvania State University, University Park, Pennsylvania, USA. [11]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania, USA. [12]Icelandic Heart Association Research Institute, Kopavogur, Iceland. [13]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [14]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [15]Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-University, Munich, Germany. [16]Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany. [17]Institute of Genetic Epidemiology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [18]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. [19]Netherlands Consortium for Healthy Aging (NCHA), The Netherlands. [20]Department of Cardiovascular Medicine, Heart and Vascular Institute, Cleveland Clinic, Cleveland, Ohio, USA. [21]Department of Molecular Cardiology, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA. [22]Department of Internal Medicine B, Ernst Moritz Arndt University Greifswald, Greifswald, Germany. [23]Laboratory for Cardiovascular Diseases, RIKEN Center for Genomic Medicine, Kanagawa, Japan. [24]Arrhythmia Research Laboratory, University of Ottawa Heart Institute, Ottawa, Ontario, Canada. [25]Department of Cardiology, Lund University, Lund, Sweden. [26]Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [27]Institute of Human Genetics, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [28]Institute of Human Genetics, Technische Universität (TU) München, Klinikum Rechts der Isar, Munich, Germany. [29]Department of Biostatistical Sciences, Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA. [30]Sticht Center on Aging, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. [31]Seattle Epidemiologic Research and Information Center of the Department of Veterans Affairs Office of Research and Development, Seattle, Washington, USA. [32]Group Health Research Institute, Group Health Cooperative, Seattle, Washington, USA. [33]Department of Epidemiology, University of Washington, Seattle, Washington, USA. [34]Department of Cell Biology, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA. [35]Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [36]Department of Biostatistics, University of Washington, Seattle, Washington, USA. [37]Cardiology Section, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA. [38]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA. [39]Laboratory of Epidemiology, Demography, and Biometry, Intramural Research Program, National Institute on Aging, US National Institutes of Health, Bethesda, Maryland, USA. [40]Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, Ottawa, Ontario, Canada. [41]Interfaculty Institute for Genetics and Functional Genomics, Ernst Moritz Arndt University Greifswald, Greifswald, Germany. [42]Institute for Community Medicine, Ernst Moritz Arndt University Greifswald, Greifswald, Germany. [43]Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, USA. [44]Cardiac Arrhythmia Center, Cardiovascular Division, Department of Medicine, University of Minnesota Medical School, Minneapolis, Minnesota, USA. [45]Epidemiological Cardiology Research Center (EPICARE), Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA. [46]Laboratory for Genotyping Development, RIKEN Center for Genomic Medicine, Kanagawa, Japan. [47]Department of Medicine, University Hospital, Basel, Switzerland. [48]Laboratory for Medical Informatics, RIKEN Center for Genomic Medicine, Kanagawa, Japan. [49]Department of Bio-informational Pharmacology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan. [50]Division of Cardiology, University of Washington, Seattle, Washington, USA. [51]Department of Epidemiology, School of Public Health, Boston University, Boston, Massachusetts, USA. [52]Laboratory for Statistical Analysis, RIKEN Center for Genomic Medicine, Kanagawa, Japan. [53]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan. [54]Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. [55]Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts, USA. [56]Department of Health Services, University of Washington, Seattle, Washington, USA. [57]Institute for Biological and Medical Imaging, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [58]Institute of Epidemiology I, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. [59]Chair of Epidemiology, Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University, Munich, Germany. [60]Klinikum Grosshadern, Munich, Germany. [61]Department of Epidemiology, Johns Hopkins University, Baltimore, Maryland, USA. [62]The Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University, Baltimore, Maryland, USA. [63]Cardiology Division, Massachusetts General Hospital, Boston, Massachusetts, USA. [64]Department of Pharmacology, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. [65]Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. [66]Department of Clinical Sciences, Lund University, Malmo, Sweden. [67]Department of Epidemiology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. [68]Department of Epidemiology & Prevention, Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. [69]Inspectorate for Health Care, The Hague, The Netherlands. [70]Department of Medical Informatics, Erasmus Medical Center, Rotterdam, The Netherlands. [71]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA. [72]Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA. [73]Preventive Medicine Section, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA. [74]Munich Heart Alliance, Munich, Germany. [75]These authors contributed equally to this work. [76]These authors jointly directed this work. Correspondence should be addressed to P.T.E. (pellinor@partners.org) or S. Kääb (skaab@med.uni-muenchen.de).

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 Cooperative Health Research in the Region Augsburg (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 Heart and Vascular Health (HVH) study for replication; included in the replication sample were atrial fibrillation cases of ≥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 that were matched 1:1 to controls from the same cohort by sex, age (±1 year), date of baseline exam (±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 ≥65 years of age, and female controls were ≥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 ≥0.99 in both cases and controls and Hardy-Weinberg equilibrium test *P* of ≥1.0×10^{-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 WGHS; BIMBAM was used by HVH and CHS; and IMPUTE v0.5 was used by Study of Health in Pomerania (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[28]. 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[29], 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 (λ) of >1.0, by multiplying the standard error of the SNP regression parameter by the square root of the study-specific λ value. A total of 2,609,549 SNPs with average minor allele frequencies of ≥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[30]. 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).