Supplementary Information
Variants in *ZFHX3* are associated with atrial fibrillation

Emelia J. Benjamin, Kenneth M. Rice, Dan E. Arking, Arne Pfeufer, Charlotte van Noord, Albert V. Smith, Renate B. Schnabel, Joshua C. Bis, Eric Boerwinkle, Moritz F. Sinner, Abbas Dehghan, Steven A. Lubitz, Ralph B. D'Agostino, Sr., Thomas Lumley, Georg B. Ehret, Jan Heeringa, Thor Aspelund, Christopher Newton-Cheh, Martin G. Larson, Kristin D. Marciante, Elsayed Z. Soliman, Fernando Rivadeneira, Thomas J. Wang, Guðný Eiríksdóttir, Daniel Levy, Bruce M. Psaty, Man Li, Alanna M. Chamberlain, Albert Hofman, Ramachandran S. Vasan, Tamara B. Harris, Jerome I. Rotter, W.H. Linda Kao, Sunil K. Agarwal, Bruno H. Ch. Stricker, Ke Wang, Lenore J. Launer, Nicholas L. Smith; Aravinda Chakravarti, André G. Uitterlinden, Philip A Wolf, Nona Sotoodehnia, Anna Köttgen, Cornelia M. van Duijn, Kathryn L. Lunetta, Susan R. Heckbert, Vilmundur Gudnason, Alvaro Alonso, Stefan Kääb, Patrick T. Ellinor, Jacqueline C. Witteman

Supplement Index

Supplementary Information
Variants in *ZFHX3* are associated with atrial fibrillation

Supplementary Table 1
Study-specific details about participants, genotyping, and data cleaning.

Characteristic	Study				
Study Acronym	**AGES**	**ARIC**	**CHS**	**FHS**	**RS**
Study	Age, Gene/ Environment Susceptibility Study	Atherosclerosis Risk in Communities Study	Cardiovascular Health Study	Framingham Heart Study	Rotterdam Study
Study website	http://www.hjarta.is/english/ages	http://www.cscc.unc.edu/aric/	http://www.chs-nhlbi.org/	http://www.framinghamheartstudy.org/about/index.html	http://www.epib.nl/ergo.htm
Array(s)	Illumina HumanCNV370-Duo BeadChip	Affymetrix 6.0	Illumina 370 CNV	Affymetrix GeneChip® Human Mapping 500K Array Set and 50K Human Gene Focused Panel	Illumina Infinium HumanHap550-chip v3.0
Calling algorithm	BeadStudio	Birdseed	BeadStudio	BRLMM	BeadStudio
Data cleaning	▪ Mismatched position ▪ Inconsistent sex, Inconsistent genotypes from other genotyping ▪ Missing haplotype tests	▪ Call rate ▪ Replicate errors ▪ Recorded vs. genotyped sex discrepancy ▪ Discrepant with prior genotyping ▪ First-degree relatives ▪ Outliers identified by IBS clustering and/or EIGENSTRAT	▪ Call rate ▪ Mendelian errors ▪ Replicate errors ▪ Recorded vs. genotyped sex discrepancy ▪ Discrepant with prior genotyping	▪ NCBI & FHS used procedures outlined by Pompanon*	Excluded: ▪ Call rate <97.5% ▪ Excess heterozygosity ▪ Phenotypic & sex mismatch (n=36) ▪ Outliers identified by the IBS clustering analysis with >3 SDs (n=102) ▪ IBS probabilities >97% (n=129)
Individual Participant Exclusions					
Call rate exclusion	<90%	<95%	<95%	<95%	<97.5%; n=209
Excess heterozygosity	NA	NA	NA	ND	>0.336; n=21
SNP Genotyping exclusions					
HWE p-value	$<10^{-6}$	$<10^{-5}$	$<10^{-5}$	$<10^{-6}$	$<10^{-6}$
SNP call rate	<97%	<95% before imputation	<97%	<97%	<90%

Supplementary Information
Variants in *ZFHX3* are associated with atrial fibrillation

Supplementary Table 1
Study-specific details about participants, genotyping, and data cleaning.

Characteristic	Study				
Study Acronym	**AGES**	**ARIC**	**CHS**	**FHS**	**RS**
MAF	<0.01	<0.01	No Dropped SNPs with no heterozygotes	<0.01	<0.01
Imputation software	Mach1 v1.0.16	MACH v1.0.16	BIMBAM v0.99	Mach1 v1.0.15	Mach1 v1.0.15
Imputation Backbone	HapMap CEU	HapMap CEU	HapMap CEU	HapMap CEU	HapMap CEU
NCBI Build	36	35	36	36	36
SNPs used for imputation	308,340	602,642	306,655	378,163	530,683
GWAS Statistical Analysis	ProbABEL, R	ProbABEL, PLINK, R	R, version 2.7	PLINK, R	Mach2QTL GenABEL + PLINK, R

Pompanon *et al.*[*1]
BRLMM denotes the Bayesian Robust Linear Modeling Mahalanobis algorithm[2].
BIMBAM (http://stephenslab.uchicago.edu/software.html)[3]
EIGENSTRAT (http://genepath.med.harvard.edu/~reich/Software.htm)[4]
GenABLE and ProbABEL (http://mga.bionet.nsc.ru/~yurii/ABEL/)[5]
MACH (http://www.sph.umich.edu/csg/abecasis/MaCH/index.html)[6]
PLINK http://pngu.mgh.harvard.edu/purcell/PLINK/[7]

Supplementary Information
Variants in *ZFHX3* are associated with atrial fibrillation

Supplementary Table 2 Study cohort, exclusion criteria and genomic inflation factor (λ) by prevalent vs. incident AF analysis

	Prevalent (n=16,664)				Incident (n=23,854)				
	AGES	**CHS**	**FHS**	**Rotterdam**	**AGES**	**ARIC**	**CHS**	**FHS**	**Rotterdam**
Age, years	≥66	≥65	≥30	≥55	≥66	≥44	≥65	≥30	≥55
Participants before exclusions	3,219	3,372	5,282	5,974	3,219	11,459[c]	3,372	5,282	5,974
Atrial fibrillation	NA	NA	NA	NA	Prevalent	Prevalent	Prevalent	Prevalent	Prevalent
CABG, n	Excluded N=260	Excluded[a]	Excluded[a]	Excluded[a]	Excluded n=260	Excluded[a]	Excluded[a]	No	Yes[a]
Participants after exclusion, n	2,959	3,267	4,464	5,974	2,718	8,086	3,201	4,184	5,665
Genomic inflation factor (λ)[b]	1.062	1.038	1.031	1.020	1.006	0.999	1.045	1.012	1.033

[a]AF that occurred during the same hospital stay as coronary bypass or cardiac valve surgery was not counted as incident AF in ARIC, CHS or RS.
In ARIC individuals with atrial flutter were not counted as atrial fibrillation.
[b]The overall **λ** for the prevalent AF was 1.005, for incident AF was 1.014, and for combined prevalent and incident AF 1.026.
[c]Number of ARIC European ancestry participants with GWAS data is 8,861.

Supplementary Information
Variants in *ZFHX3* are associated with atrial fibrillation

Supplementary Table 3 Characteristics of participants in the five cohorts by prevalent versus incident AF meta-analysis.

Baseline Characteristics	Prevalent AF Analysis				Incident AF Analysis				
	AGES	CHS	FHS	Rotterdam	AGES	ARIC	CHS	FHS	Rotterdam
Participants, n	2,959	3,267	4,464	5,974	2,718	8086	3,201	4,184	5665
Sex, men, n (%)	1,154(39.0)	1,278(39.1)	2,004(44.9)	2,427(40.6)	1,011(37.2)	3814(47.2)	1241(38.8)	1830(43.7)	2282(40.3)
Age[a], years, mean±SD	76.5±5.5	72.3±5.4	65.5±12.7	69.4±9.1	76.3±5.5	57.0±5.7	72.2±5.3	64.7±12.6	69.1±9.0
Age[a], years, minimum-maximum	66-95	65-98	30-100	55-99	66-95	46-70	65-98	30-100	55-99
Hypertension, n (%)	2,260(79.8)	1,711(52.4)	2,263(50.8)	1,997(33.4)	2,145(78.9)	2192 (27.1)	1,677(52.4)	2,062(49.4)	1,866(32.9)
BMI, kg/m^2, mean±SD	27.1±4.4	26.3±4.4	27.7±5.1	26.3±3.7	27.1±4.5	27.0±4.9	26.3±4.4	27.7±5.2	26.3±3.7
Diabetes, n (%)	319(10.8)	392(12.0)	380(8.5)	631(10.6)	289(10.6)	693(8.6)	379(11.8)	334(8.0)	567(10.0)
Myocardial infarction, n (%)	147(5.0)	0	313(7.0)	727(12.2)	130(4.8)	331(4.1)	0	240(5.7)	626(11.1)
Heart failure, n (%)	63(2.1)	0	122(2.7)	194(3.2)	32 (1.2)	314(3.9)	0	55(1.3)	140(2.5)
Atrial fibrillation cases									
Number	241	66	280	309	138	731	763	343	542
Age at AF onset, mean±SD	76.9±6.0	NA	70.6±10.6	NA	80.6±6.0	67.0±6.7	81.2±6.0	77.4±10.5	77.7±7.7

[a]Age was defined as age at DNA collection (baseline examination, AGES, CHS, RS; 1990s, FHS; Visit 2, 1990-1992, ARIC).
BMI, body mass index; NA, age at onset of prevalent AF not available for CHS and RS.

Supplementary Information
Variants in *ZFHX3* are associated with atrial fibrillation

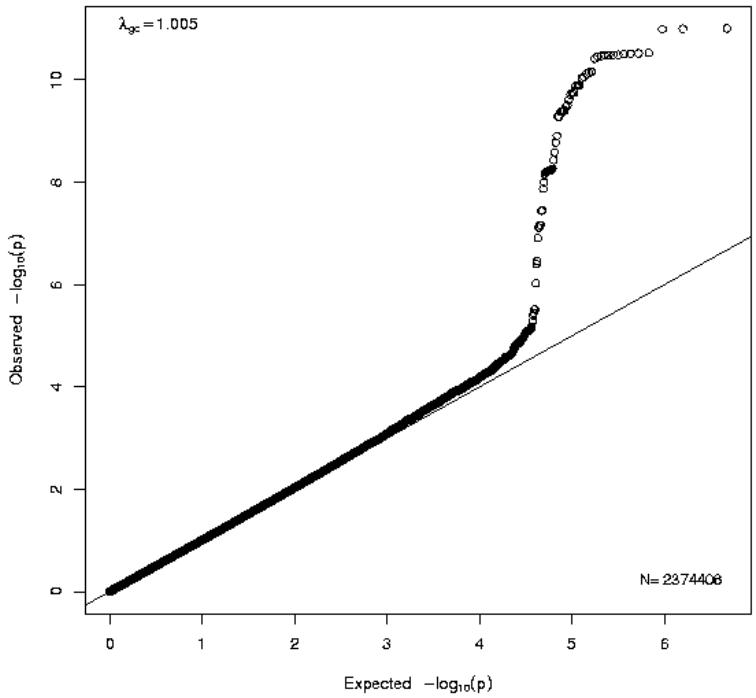
Supplementary Table 4 Summary of cohort-specific results that had $P \leq 4 \times 10^{-7}$ in the meta-analysis of combined prevalent and incident AF

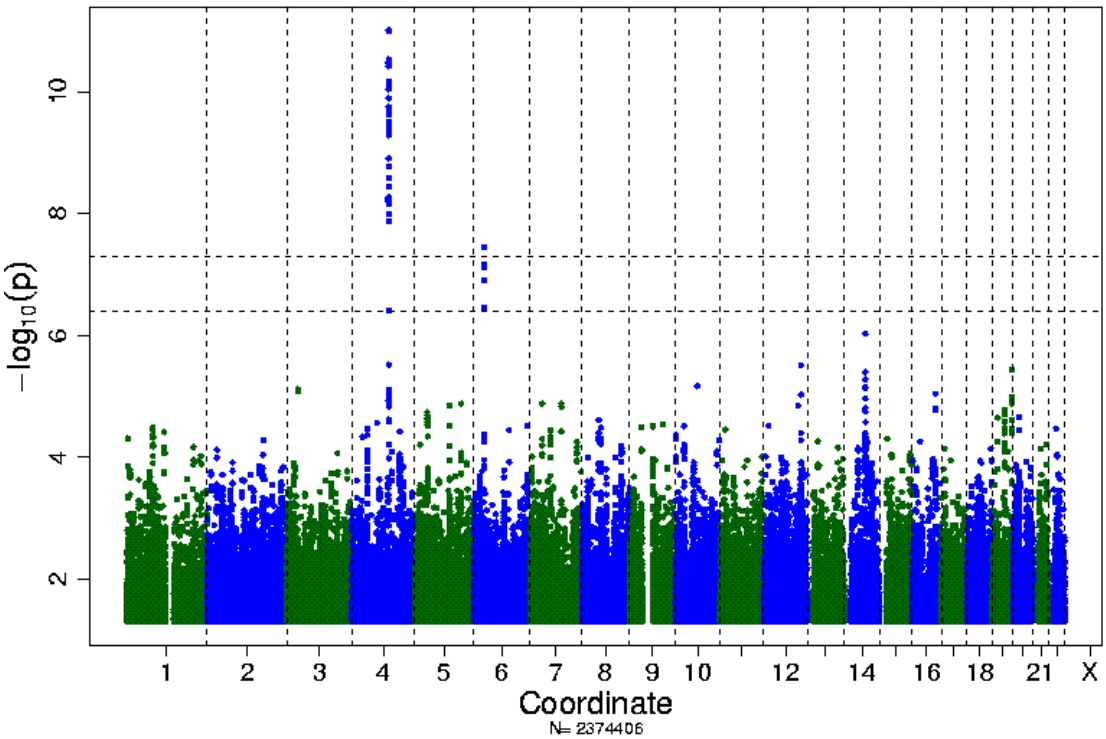
SNP	Chr	Position	Nearby gene	Minor/ major allele	Analysis	AGES	ARIC	CHS	FHS	RS	Meta-analysis	
								β±se RR			β±se RR	*P* value
rs17042171	4	111927736	*PITX2*	A/C	Prevalent	0.54±0.13 1.72	NA	0.90±0.20 2.46	0.50±0.12 1.65	0.20±0.13 1.22	0.47±0.07 1.60	3.1×10^{-11}
					Incident	0.49±0.15 1.63	0.44±0.07 1.55	0.27±0.07 1.31	0.13±0.11 1.14	0.36±0.09 1.43	0.34±0.04 1.40	8.3×10^{-18}
					Combined						0.37±0.03 1.45	6.0×10^{-27}
rs2106261	16	71609121	*ZFHX*3	T/C	Prevalent	0.43±0.11 1.54	NA	0.01±0.29 1.01	0.05±0.13 1.05	0.36±0.10 1.43	0.29±0.04 1.33	9.0×10^{-6}
					Incident	-0.11±0.15 0.90	0.23±0.07 1.26	0.05±0.08 1.05	0.27±0.11 1.31	0.05±0.08 1.05	0.13±0.06 1.14	7.9×10^{-4}
					Combined						0.17±0.03 1.19	2.3×10^{-7}
rs17375901	1	11775103	*MTHFR*	T/C	Prevalent	0.35±0.20 1.42	NA	0.89±0.33 2.44	0.36±0.20 1.43	0.19±0.17 1.21	0.35±0.10 1.42	8.5×10^{-4}
					Incident	0.30±0.24 1.35	0.22±0.10 1.25	0.36±0.12 1.43	-0.04±0.19 0.96	0.33±0.12 1.39	0.26±0.06 1.30	1.2×10^{-5}
					Combined						0.29±0.05 1.34	4.6×10^{-8}

Chr, chromosome, β, regression parameter estimate from the logistic (prevalent AF) or Cox (incident AF) regression analysis; se, standard error; RR, odds ratio (prevalent AF) or hazard ratio (incident AF). Prevalent AF n=16,526; Incident AF n=23,854

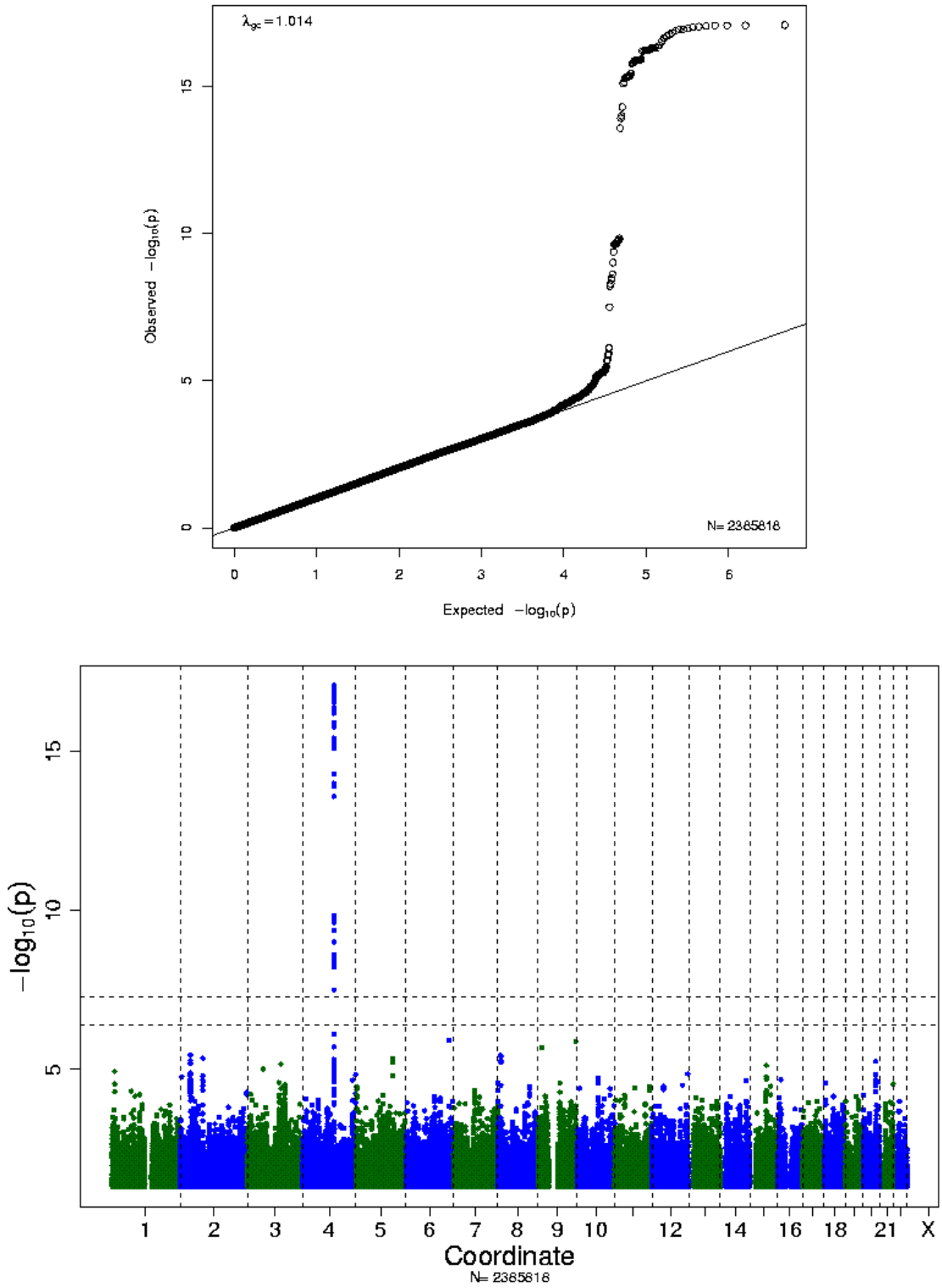
Supplementary Information
Variants in *ZFHX3* are associated with atrial fibrillation

Supplementary Figure 1 Summary of genome-wide scan of prevalent AF.
Top panel is quantile-quantile plot;
Bottom panel is plot of -$\log_{10}$ *P* values by genomic location (Manhattan plot) for the meta-analysis of four prevalent AF studies. Results are displayed by chromosome for SNPs with average minor allele frequency across studies >0.01. SNPs with *P* >0.05 are omitted from the plot. The two horizontal dotted lines are for thresholds of $P=5\times10^{-8}$ and 4×10^{-7}. n= the number of imputed SNPs included.

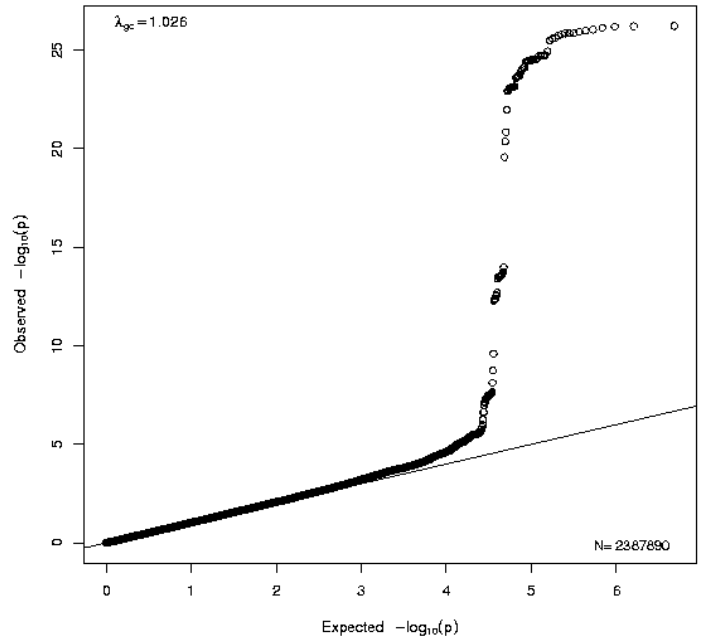


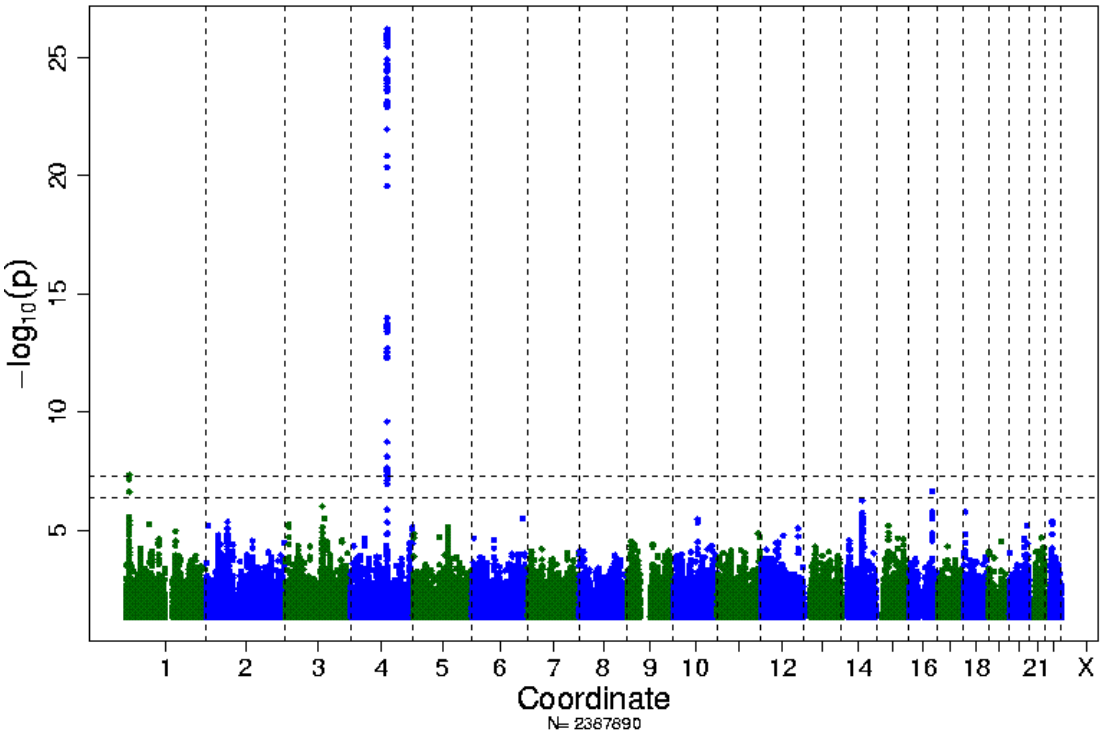


Supplementary Figure 2 Summary of genome-wide scan of incident AF;
Top panel is quantile-quantile plot for incident AF;
Bottom panel is plot of -$\log_{10}$ *P* values by genomic location for the meta-analysis of five incident AF studies. Results are displayed by chromosome for SNPs with average minor allele frequency across studies >0.01. SNPs with *P* >0.05 are omitted from the plot. The two horizontal dotted lines are for thresholds of $P=5\times10^{-8}$ and 4×10^{-7}.



Supplementary Figure 3 Summary of genome-wide scan for combined prevalent and incident AF. Top panel is quantile-quantile plot.
Bottom panel is plot of -$\log_{10}$ *P* values by genomic location for the meta-analysis of combined AF studies (four prevalent and five incident AF). The two horizontal dotted lines are for detection thresholds of $P=5\times10^{-8}$ and $P=4\times10^{-7}$. The total number of imputed SNPs depicted in the plot represents SNPs meeting our filtering criteria, which included average minor allele frequency across the studies >0.01 and at least 6 of 9 studies providing results. The plots omit SNPs with $P>0.05$.





Supplementary Note

AGES: The Age, Gene/Environment Susceptibility Reykjavik Study has been funded by NIH contract N01-AG-12100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

ARIC: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

CHS: The Cardiovascular Health Study research reported in this article was supported by contract numbers N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, and grant numbers U01 HL080295, R01 HL087652, and R01 HLO88456 from the National Heart, Lung, and Blood Institute, with additional contribution from the National Institute of Neurological Disorders and Stroke. DNA handling and genotyping was supported in part by National Center for Research Resources grant M01RR00069 to the Cedars-Sinai General Clinical Research Center Genotyping core and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center. A full list of principal CHS investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm.

FHS: This research was conducted using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine based on analyses by Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No.N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. The Framingham Heart Study research was supported by NIH grants 1R01HL092577-01A1 (PTE, EJB); HL076784, AG028321 (EJB); 6R01-NS 17950 (PAW); T32 HL007575 (SAL); R01 HL093328 and RO1 HL 080124 (RSV); Deane Institute for Integrative Research in Atrial Fibrillation and Stroke (PTE). The Deutsche Forschungsgemeinschaft (German Research Foundation) Research Fellowship SCHN 1149/1-1 supported Gutenberg Heart Study investigator RBS' FHS research.

AFNET / KORA: German National Genome Research Network NGFN 01GS0499 and 01GS0838 (SK), NGFN 01GR0803 (AP), NGFN 01GR0103 (TM), German Federal Ministry of research BMBF 01EZ0874 (AP), German Competence Network on AF (AF-Net) 01 GI 0204/N (SK, HEW), Leducq Foundation 07-CVD 03 (SK), LMU Excellence Initiative (SK), KORA is supported by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria.

RS: The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam; The Netherlands Organization for Scientific Research; The Netherlands Organization for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly; The Netherlands Heart Foundation; the Ministry of Education, Culture and Science; the Ministry of Health Welfare and Sports; the European Commission; and the

Municipality of Rotterdam. Support for genotyping was provided by The Netherlands Organization for Scientific Research (NWO) (175.010.2005.011, 911.03.012) and Research Institute for Diseases in the Elderly (RIDE). This study was supported by The Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810.

Supplementary Information
Variants in *ZFHX3* are associated with atrial fibrillation

Supplementary Methods

The CHARGE AF Consortium[8] included analyses from five community-based cohorts that collected AF cases systematically in longitudinal follow-up and had GWAS data (**Supplementary Table 1** contains study details). All participants included in this analysis were of European descent. In CHARGE cohorts (and AFNET and KORAS4) written informed consent was obtained from each subject, including consent to use DNA for genetic analyses of cardiovascular disease. Consent precluded participant-specific meta-analysis. African American participants from ARIC and CHS studies were not analyzed for the present study.

- **AGES** represents the later follow-up of the midlife Reykjavik (Iceland) Study founded in 1967[9]. AGES was designed to examine the genetic epidemiology of four phenotypes known to alter with advancing age: vascular, neurocognitive, musculoskeletal and body composition. The AGES examinations were conducted between 2002 and 2006 on 5,764 survivors of the Reykjavik Study.
- The **ARIC** study was initiated in 1987 to examine atherosclerosis in middle-aged adults and completed enrollment of 15,792 participants. ARIC was conducted in four US communities (Forsyth County, NC; Jackson, MS; Minneapolis suburbs, MN; and Washington County, MD). Participants were examined about every three years four times and followed for events[10]. Only white ARIC participants were included in the analyses.
- **CHS** is a prospective population-based cohort study of CVD in adults 65 years and older. The four Field Centers (Forsyth County, NC; Sacramento County, CA; Washington County, MD; Pittsburgh, PA) completed enrollment of 5888 participants in 1989-1990 and 1992-1993[11].
- **FHS** is a community-based observational, cohort initiated in 1948 to prospectively investigate CVD and its risk factors. The Original cohort (n=5209) received biennial exams[12]. The Original Cohort children (& spouses), termed the Offspring cohort (n=5214), were recruited in 1971, and were examined every four to eight years[13].
- The community-based **RS** was founded in 1990 to examine the determinants of disease and health in the elderly with a focus on CVD, neurogeriatric, bone and eye diseases. Inhabitants of a Rotterdam suburb (n=10,275) age ≥55 years were invited and 7,983 participants (78%) were examined. The participants were examined up to four times approximately every three years[14].

AF Ascertainment

Studies included initial, paroxysmal, persistent, and permanent atrial fibrillation and atrial flutter (ARIC did not include atrial flutter). Prevalent AF was considered present if AF was observed on baseline electrocardiograms (AGES, CHS, RS), or prior to DNA collection (FHS). Incident AF was defined as AF that first occurred after the collection of DNA (FHS, ARIC), or after the baseline examination (other cohorts). There was no overlap in AF cases between prevalent and incident AF analyses.

- The **AGES** study ascertained AF based on AGES examination Minnesota coded electrocardiograms, and ICD-10 I48 recorded from hospitalizations from 1997 through March, 2008.
- The **ARIC** study determined AF from three sources: electrocardiograms at study visits[15], hospital discharge records and death certificates[16] (first reviewed and confirmed by a cardiologist, latter two reviewed by trained abstractor, ICD-9 code 427.31 or 427.3; ICD-10 I48). Incidence of AF was identified through 2004 as the first occurrence of AF by any of the sources. ARIC excluded 37 atrial flutter without subsequent AF cases.

- In **CHS**, prevalent AF was identified by 12-lead ECG at baseline. Incident AF in up to 16 years follow-up (median 13 years) was identified by the first occurrence of AF on annual CHS study electrocardiograms or the first occurrence of a hospital discharge ICD-9 code for AF. CHS participant hospital discharge diagnosis codes for AF were found to have a sensitivity of 71% for AF[17].
- In **FHS**, all cardiovascular hospital and outside records were routinely obtained and electrocardiograms were recorded at all FHS examinations; AF cases through 2007 were verified by 2 FHS cardiologists[18].
- For the **RS** the ascertainment of AF included review of hospital discharge information, general practitioner diagnoses and RS electrocardiograms. AF was verified by 2 physicians and verified by electrocardiogram review by a cardiologist in the case of a disagreement[19].

AF Replication sample

The replication **German Competence Network for Atrial Fibrillation (AFNET)** is a national registry of patients with prevalent AF onset before 60 years (n=2,145) without structural heart disease. Referent subjects without AF or structural heart disease (n=4,073) were drawn from the population-based **KORA S4** study[20]. Replication genotyping was performed using iPlex single base primer extension with MALDI-TOF mass spectrometry (Sequenom Inc., San Diego, CA). Association testing was performed using logistic regression with an additive genetic model adjusting for age at DNA draw, sex and hypertension status.

Genotyping and Imputation (Supplementary Table 1)

Briefly, the five studies utilized a variety of high-density Illumina (Human CNV370, AGES and CHS; Infinium 550, RS) and Affymetrix (6.0, ARIC; 500K +50K human gene focused, FHS) platforms. Approximately 2.5 million autosomal genotypes were imputed within each study using the Phase II CEU HapMap reference panel (http://hapmap.org) and BIMBAM (CHS; http://stephenslab.uchicago.edu/software.html) or MACH v1.0.15/16 (all others; http://www.sph.umich.edu/csg/abecasis/MaCH/index.html) software. A recent review supports the validity of combining results across statistical and genotyping platforms[21]. For FHS, the imputation model parameters were estimated using an independent subset of all individuals, and then applied to all others. We expressed imputation results as an allelic dosage (fractional value between 0.0 and 2.0). Each cohort performed stringent quality control checks on imputed allele dosages. The imputation engines used for the present project were built on the same essential algorithm. In line with other authors we found no important difference in performance, and no evidence to suggest biases due to choice of imputation software. The excellent concordance between effects seen between cohorts in this and other CHARGE manuscripts provides corroboration.

Statistical methods

Primary GWAS were performed within each cohort separately for prevalent and incident AF using an additive genetic model adjusting for age, sex, and, if relevant, cohort (FHS Original versus Offspring) or site (ARIC, CHS). Prevalent AF was examined with logistic regression; controls for prevalent analyses included all eligible participants without prevalent AF at the time of DNA collection. Incident AF was examined with proportional-hazards regression, using years to AF as the outcome, censoring at death, loss to follow-up or date of last contact. The incident AF analyses included eligible participants without prevalent AF at the time of DNA collection. To account for its pedigree structure data, FHS used generalized estimating equations for logistic analyses and robust variance estimates for proportional-hazards analyses. The variance of the regression parameter was multiplied by λ, and then the regression parameter and adjusted standard error were combined using prospective inverse-variance weighted meta-

analysis. The meta-estimates and *P* values form the principal results. In secondary analyses, heterogeneity of the regression parameters across the nine studies was assessed using Cochran's test.

The person-time of individuals used for the incident AF analyses begins immediately following the time point at which phenotypes were observed for the prevalent AF analysis. Thus, the time spans of the two analyses do not overlap, despite having overlapping referent individuals. Under the martingale property of Cox models, the two analyses are independent. Across our genome-wide data, the correlation of the regression coefficients for the prevalent and incident analyses for the four studies providing both prevalent and incident data (ARIC contributed incident data only) was low, varying from 0.005 to 0.027, empirically confirming the expected independence.

1. Pompanon,F.*, et al.*, *Nat. Rev. Genet* **6**, 847-859 (2005).
2. Rabbee,N.*, et al.*, *Bioinformatics.* **22**, 7-12 (2006).
3. Servin,B.*, et al.*, *PLoS. Genet.* **3**, e114 (2007).
4. Price,A.L.*, et al.*, *Nat. Genet* **38**, 904-909 (2006).
5. Aulchenko,Y.S.*, et al.*, *Bioinformatics.* **23**, 1294-1296 (2007).
6. Li,Y.*, et al.*, *Am J Hum. Genet.* **S79**, 2290 (2008).
7. Purcell,S.*, et al.*, *Am J Hum. Genet.* **81**, 559-575 (2007).
8. Psaty,B.M.*, et al.*, **2**, 73-80 (2009).
9. Harris,T.B.*, et al.*, *Am J Epidemiol.* **165**, 1076-1087 (2007).
10. ARIC investigators, *Am J Epidemiol.* **129**, 687-702 (1989).
11. Fried,L.P.*, et al.*, *Ann. Epidemiol.* **1**, 263-276 (1991).
12. Dawber,T.R.*, et al.*, *Am J Public Health Nations. Health* **41**, 279-281 (1951).
13. Kannel,W.B.*, et al.*, *Am. J. Epidemiol.* **110**, 281-290 (1979).
14. Hofman,A.*, et al.*, *Eur. J Epidemiol.* **22**, 819-829 (2007).
15. Soliman,E.Z.*, et al.*, **40**, 1204-1211 (2009).
16. Alonso,A.*, et al.*, *Am Heart J* (2009).
17. Psaty,B.M.*, et al.*, **96**, 2455-2461 (1997).
18. Benjamin,E.J.*, et al.*, *JAMA* **271**, 840-844 (1994).
19. Heeringa,J.*, et al.*, *Eur. Heart J* **27**, 949-953 (2006).
20. Sinner,M.F.*, et al.*, *Eur. Heart J* **29**, 907-914 (2008).
21. de Bakker,P.I.*, et al.*, *Hum. Mol. Genet* **17**, R122-R128 (2008).