Multi-ethnic genome-wide association study for atrial fibrillation

Roselli, Carolina; Chaffin, Mark D ; Weng, Lu-Chen ; Aeschbacher, Stefanie; Ahlberg, Gustav ; Albert, Christine M.; Almgren, Peter; Alonso, Alvaro; Anderson, Christopher D ; Aragam, Krishna G; Arking, Dan E; Barnard, John; Bartz, Traci M ; Benjamin, Emelia J; Bihlmeyer, Nathan A ; Bis, Joshua C; Bloom, Heather L ; Boerwinkle, Eric; Bottinger, Erwin P; Brody, Jennifer A

DOI:
10.1038/s41588-018-0133-9

Document Version
Peer reviewed version

GWAS for Atrial Fibrillation

SUPPLEMENTARY TEXT AND TABLES

INDEX

Supplementary Tables

Supplementary Tables

Supplementary Table S1. Summary of AF cases and referents by ancestry

Ancestry	AF	Referents
European	55,114	482,295
Japanese	8,180	28,612
African American	1,307	7,660
Brazilian	568	1,096
Hispanic	277	3,081
Total	**65,446**	**522,744**

Supplementary Table S2. Additional novel loci in combined and European ancestry meta-analysis with traditional P-value cutoff (5 x 10^{-8})

Sentinel Variant	Chr	hg19	Risk/Ref Allele	RAF [%]	RR	95% CI	P-value	Nearest Gene(s)	Func	Imp Qual	I^2	P_{HET}
Combined ancestry												
rs12992412	2	148792665	T/A	38	1.04	1.03-1.06	2.30E-08	*MBD5*	intronic	0.99	0	4.78E-01
rs73032363	3	24472866	A/G	70	1.04	1.03-1.06	3.59E-08	*THRB*	intronic	0.98	0	6.89E-01
rs1307274	6	34240576	T/G	8	1.08	1.05-1.11	3.85E-08	*C6orf1,NUDT3*	regulatory reg.	0.92	0	8.76E-01
rs6907805	6	87856003	G/T	51	1.04	1.03-1.06	1.10E-08	*CGA,ZNF292*	intergenic	0.98	0	4.14E-01
rs12208899	6	133474303	A/G	21	1.05	1.03-1.07	1.95E-08	*LINC00326,EYA4*	intergenic	0.98	0	4.22E-01
rs11768850	7	876227	T/C	42	1.04	1.03-1.05	4.96E-08	*SUN1*	intronic	0.98	0	4.14E-01
rs55985730	7	128417044	G/T	6	1.10	1.06-1.14	1.81E-08	*OPN1SW*	upstream	0.87	0	8.05E-01
rs7460121	8	135812416	A/G	10	1.07	1.05-1.10	1.65E-08	*MIR30B*	downstream	0.92	0	7.04E-01
European ancestry												
rs9872035	3	196494702	C/T	45	1.04	1.03-1.06	1.80E-08	*PAK2*	intronic	0.97		
rs210632	6	117880342	A/G	25	1.05	1.03-1.07	2.75E-08	*GOPC*	downstream	1		

Abbreviations, Chr, chromosome, CI, confidence interval, Func, functional consequence (most severe consequence by variant effect predictor), HET, heterogeneity, I^2, I-square, impQual, average imputation quality, P, P-value, RR, relative risk, RAF, risk allele frequency.

Supplementary Table S3. Known loci in combined ancestry meta-analysis

Enclosed electronic excel file

Supplementary Table S4. Gene set enrichment analysis results for combined ancestry meta-analysis

Enclosed electronic excel file

Supplementary Table S5. Novel and known loci in ancestry specific meta-analyses

Enclosed electronic excel file

Supplementary Table S6. Proportion of heritability explained by AF loci

Study	AF-loci h^2_g observed (SE)	AF-loci h^2_g liability scale (SE)	Overall h^2_g liability scale	Proportion explained [%]
25 AF loci (EUR ancestry loci Christophersen et al.)[1]	0.00744 (0.00081)	0.0533 (0.00579)	0.2156	24.74
84 AF loci (EUR ancestry loci from Roselli et al.)	0.01262 (0.00121)	0.0905 (0.00868)	0.2140	42.28

Abbreviations, AF, atrial fibrillation, EUR, European, h^2_g, SNP-heritability, SE, standard error.

Supplementary Table S7. Loci with multiple signals identified by conditional and joint analysis for European ancestry meta-analysis

Enclosed electronic excel file

Supplementary Table S8. Sentinel variants or proxies with missense consequence, for loci from combined ancestry meta-analysis

Sentinel Variant	Sentinel MAF [%]	Proxy	Proxy MAF [%]	LD [r^2]	Location	Gene	Consequence coding sequence	Consequence protein sequence	# In silico predicted as damaging*
Novel loci									
rs187585530	0.5	-	-	-	1:10167425	*UBE4B*	c.1342G>A	p.Gly448Arg	-
rs4484922	31.7	rs4074536	31.7	1.000	1:116310967	*CASQ2*	c.196A>G	p.Thr66Ala	0
rs72926475	13.0	rs34605051	15.7	0.671	2:86693826	*KDM3A*	c.1339T>C	p.Ser447Pro	1
rs2306272	31.8	-	-	-	3:66434643	*LRIG1*	c.1843A>G	p.Met615Val	0
rs3822259	32.1	rs13441	32.4	0.724	4:10099340	*WDR1*	c.553A>G	p.Ile185Val	0
rs10760361	35.3	rs4574	42.5	0.756	9:127177161	*PSMB7*	c.116T>C	p.Val39Ala	0
rs7919685	46.7	rs1935	48.1	0.861	10:64927823	*JMJD1C*	c.7605G>C	p.Glu2535Asp	1
rs7978685	27.9	rs2958149	27.7	0.999	12:57109792	*NACA*	c.2063T>C	p.Leu688Pro	0
rs12298484	32.6	rs11057401	30.7	0.906	12:124427306	*CCDC92*	c.208A>T	p.Ser70Cys	4
Known loci									
rs35504893	24.8	rs3731746	23.2	0.846	2:179430997	*TTN*	c.79862C>T	p.Thr26621Met	2
rs6810325	36.6	rs11718898	36.5	0.996	3:12848822	*CAND2*	c.230T>C	p.Val77Ala	0
rs6790396	39.2	rs6795970	38.0	0.929	3:38766675	*SCN10A*	c.3227T>C	p.Val1076Ala	0
rs60212594	14.7	rs60632610	14.7	0.998	10:75415677	*SYNPO2L*	c.4G>A	p.Gly2Ser	3

Abbreviations, Ala, alanine Arg, arginine, Asp, asparagine, Cys, cysteine, Glu, glutamine, Gly, glycine, Leu, leucine, LD, linkage disequilibrium, MAF, minor allele frequency, Met, methionine, Pro, proline, Ser, serine, Thr, threonine, Val, valine. *Assessed by 5 prediction algorithms: MutationTaster (disease causing automatic or disease causing), SIFT (damaging), LRT (deleterious), Polyphen2 prediction based on HumDiv (probably damaging or possibly damaging), Polyphen2 prediction based on HumVar (probably damaging or possibly damaging).

Supplementary Table S9. Chromatin states for sentinel variants and proxies from Roadmap Epigenomics across all tissues and heart

Enclosed electronic excel file

Supplementary Table S10. Significant cis-eQTLs for sentinel variants from combined ancestry meta-analysis in left atrial tissue from MAGNet

Sentinel Variant	Chr	hg 19	Gene	Effect	P-value	FDR	Effect/Ref Allele
Novel loci							
rs4484922	1	116310818	*CASQ2*	0.72	5.29E-05	4.71E-02	C/G
rs4855075	3	179170494	*GNB4*	1.24	2.61E-12	<1.00E-05	T/C
rs34969716	6	18210109	*TPMT*	0.91	6.73E-07	1.72E-03	A/G
rs12298484	12	124418674	*CCDC92*	0.67	8.40E-06	1.22E-02	T/C
rs10873299	14	77426711	*LINC01629*	1.02	2.20E-20	<1.00E-05	A/G
			RP11-7F17.1	0.77	9.63E-12	<1.00E-05	A/G
rs12908004	15	80676925	*ARNT2*	1.06	4.62E-09	<1.00E-05	G/A
			RP11-210M15.2	1.00	3.51E-08	1.65E-04	G/A
			LINC01314	0.85	3.93E-06	6.74E-03	G/A
rs242557	17	44019712	*MAPT*	0.91	1.76E-12	<1.00E-05	A/G
Known loci							
rs2540949	2	65284231	*CEP68*	0.69	9.40E-07	2.21E-03	A/T
rs60212594	10	75414344	*MYOZ1*	1.49	1.18E-11	<1.00E-05	C/G

Abbreviations: Chr, chromosome, eQTL, expression quantitative trait locus, FDR, false discovery rate, Ref, reference.

Supplementary Table S11. Significant cis-eQTLs for sentinel variants from combined ancestry meta-analysis in GTEx heart tissues

Enclosed electronic excel file

Supplementary Table S12. Probable AF susceptibility genes for loci from combined ancestry meta-analysis

Enclosed electronic excel file

Supplementary Table S13. MetaXcan results based on summary level data from combined ancestry meta-analysis

Enclosed electronic excel file

Supplementary Table S14. Association to diseases and traits in NHGRI-EBI GWAS catalog for sentinel variants or proxies from combined ancestry meta-analysis

Enclosed electronic excel file

Supplementary Table S15. PheWAS results in UK Biobank for sentinel variants from combined ancestry meta-analysis

Enclosed electronic excel file

Supplementary Table S16. 134 loci associated with atrial fibrillation

Enclosed electronic excel file

Supplementary Table S17. Baseline summary for GWAS

Enclosed electronic excel file

Supplementary Table S18. GWAS summary on genotyping, QC, imputation and analysis per study

Enclosed electronic excel file

Supplementary Table S19. Pre-Imputation quality control guidelines

Sample based filters	Check for per sample completeness (e.g. sample call-rate ≥95%) Check for per sample heterozygosity (e.g. heterozygosity should be within mean ± 3s.d.) Inspection of principal component or multi-dimensional scaling plots (exclude genetic ancestry outliers) Removal of related individuals or account for relatedness in the statistical model
Variant based filters	**Exclusion of markers when:** Low genotype completeness (e.g. SNP call rate <98%) Clear deviations from Hardy-Weinberg proportions (e.g. Hardy Weinberg Equilibrium P-value in controls < $1x10^{-6}$) High discordance rates Excess of Mendelian inconsistencies (if parent-offspring pairs available) Rare variants (e.g. minor allele frequency <0.5% or <1%)

Supplementary Table S20. Clinical characteristics of MAGNet samples

Characteristics	**Left atria (n=101)**
Women, n (%)	57 (56.4%)
Age, mean [years] ± SD	59±12
Weight, mean [kg] ± SD	82±22
Height, mean [cm] ± SD	167±17
Hypertension, n (%)	60 (59.4%)
Diabetes, n (%)	20 (19.8%)
History of atrial fibrillation, n (%)	14 (13.9%)

Abbreviations, cm, centimeter, kg, kilogram, SD, standard deviation.

Supplementary Table S21. Phenotype definitions for UK Biobank PheWAS

Phenotype	Definition
Bradyarrhythmia	1) Non-cancer illness code, self-reported (1486) or 2) Diagnoses – main/secondary ICD10 (I44, I44.0-7, I45.0-5, I49.5) or 3) Underlying primary/Contributing secondary cause of death: ICD10 (I44, I44.0-7, I45.0-5, I49.5) or 4) Diagnoses – main/secondary ICD9 (426.0, 426.1, 426.3, 426.4, 426.5, 426.6) or 5) Operation code (1096, 1548, 1549) or 6) Pacemaker (Yes) or 7) Operative procedures – main/secondary OPCS (K60, K60.1-9, K61, K61.1-9)
Coronary Artery Disease	1) Operative procedures – main/secondary OPCS (K40.1-4, K41.1-4, K45.1-5, K49.1-2, K49.8-9, K50.2, K75.1-4, K75.8-9) or 2) Source of first myocardial infarction report (Self-reported only, Hospital admission, Death only)
Type II Diabetes	1) Non-cancer illness code, self-reported (1223) or 2) Diagnoses – main/secondary ICD10 (E11, E11.0-9) or 3) Underlying (primary/secondary) cause of death: ICD10 (E11, E11.0-9)
Heart Failure	1) Non-cancer illness code, self-reported (1076, 1079) or 2) Diagnoses – main/secondary ICD10 (I11.0, I13.0, I13.2, I25.5, I42.0, I42.5, I42.8-9, I50, I50.0, I50.1, I50.9) or 3) Underlying primary/Contributing secondary cause of death: ICD10 (I11.0, I13.0, I13.2, I25.5, I42.0, I42.5, I42.8-9, I50, I50.0, I50.1, I50.9) or 4) Diagnoses – main/secondary ICD9 (425.4, 428.0, 428.1, 428.9) and exclude: 1) Non-cancer illness code, self-reported (1588) or 2) Diagnoses – main/secondary ICD10 (I42.1-2) or 3) Underlying primary/Contributing secondary cause of death: ICD10 (I42.1-2)
Hypercholesterolemia	1) Non-cancer illness code, self-reported (1473) or 2) Diagnoses – main/secondary ICD10 (E78.0-2, E78.4-5) or 3) Underlying primary/Contributing secondary cause of death: ICD10 (E78.0-2, E78.4-5)
Hypertension	1) Non-cancer illness code, self-reported (1065, 1072) or 2) Vascular/heart problems diagnosed by doctor (High blood pressure) or

	3) Diagnoses – main/secondary ICD10 (I10, I11, I11.0, I11.9, I12, I12.0, I12.9, I13, I13.0-2, I13.9, I15, I15.0-2, I15.8-9) or 4) Underlying primary/Contributing secondary cause of death: ICD10 (I10, I11, I11.0, I11.9, I12, I12.0, I12.9, I13, I13.0-2, I13.9, I15, I15.0-2, I15.8-9) or 5) Diagnoses – main/secondary ICD9 (401, 401.0, 401.1, 401.9, 402, 402.0, 402.1, 402.9, 403, 403.0, 403.1, 403.9, 404, 404.0, 404.1, 404.9, 405, 405.0, 405.1, 405.9)
Mitral regurgitation	1) Non-cancer illness code, self-reported (1585) or 2) Diagnoses – main/secondary ICD10 (I05.1-2, I34.0) or 3) Underlying primary/Contributing secondary cause of death: ICD10 (I05.1-2, I34.0) or 4) Diagnoses – main/secondary ICD9 (3942) or 5) Operative procedures – main/secondary OPCS (K34.1)
Peripheral vascular disease	1) Non-cancer illness code, self-reported (1067, 1087, 1088) or 2) Operation code (1102, 1108, 1440) or 3) Diagnoses – main/secondary ICD10 (I70.0, I70.00-01, I70.2, I70.20-21, I70.8, I70.80, I70.9, I70.90, I73.8-9) or 4) Underlying primary/Contributing secondary cause of death: ICD10 (I70.0, I70.00-01, I70.2, I70.20-21, I70.8, I70.80, I70.9, I70.90, I73.8-9) or 5) Diagnoses – main/secondary ICD9 (440.0, 440.2, 443.8, 443.9) or 6) Operative procedures – main/secondary OPCS (X09.3-5, L21.6, L51.3, L51.6, L51.8, L52.1-2, L54.1, L54.4, L54.8, L59.1-8, L60.1-2, L63.1, L63.5, L63.9, L66.7)
Stroke	Source of first stroke report (Self-reported only, Hospital admission, Death only)
Smoking	Smoking status at first visit (Never versus Previous or Current)
BMI	Body mass index at first visit
Height	Standing height at first visit (exclude if <120cm)

Abbreviations: ICD, international classification of diseases and related health problems, OPCS, office of population censuses and surveys: classification of interventions and procedures.

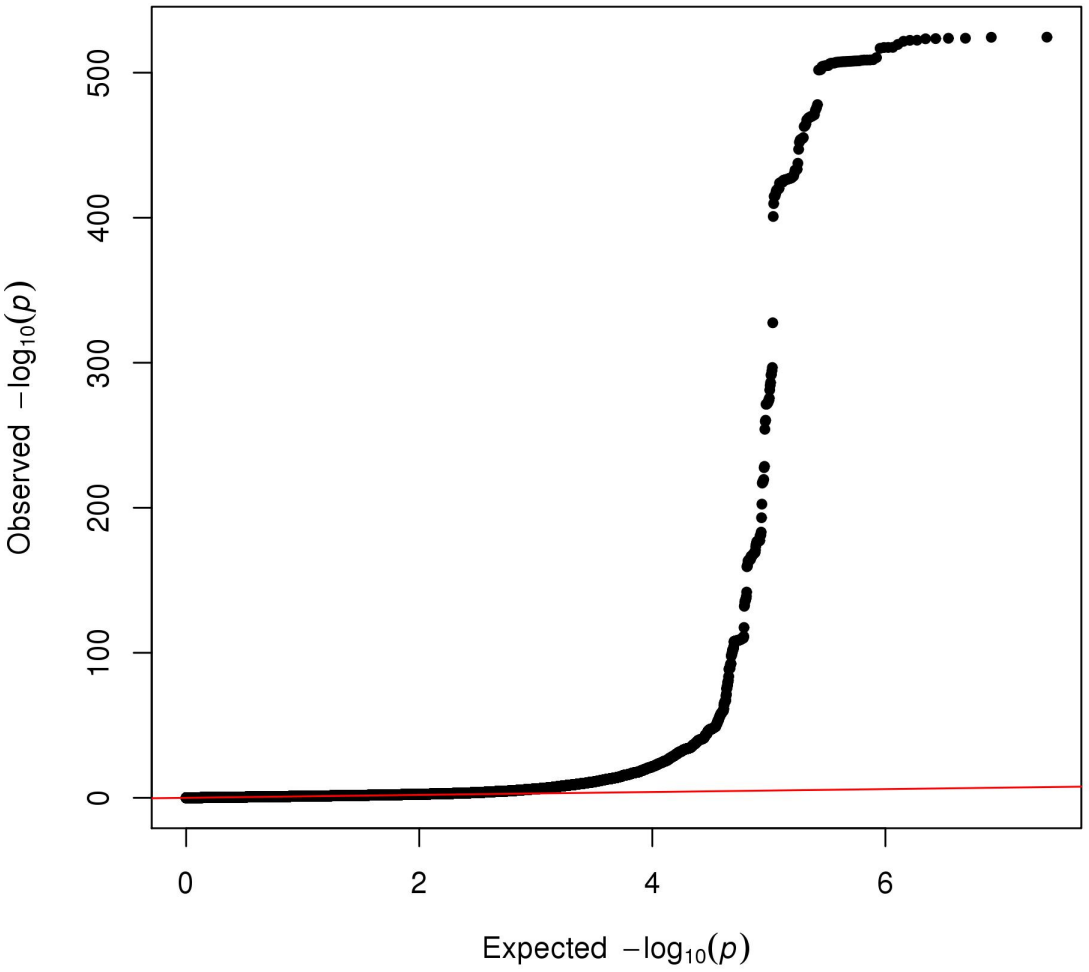
Supplementary Table S22. Summary of cases, referents and sample size for phenotypes in UK Biobank PheWAS

Phenotype	# Cases	# Referents	Sample size
Body-mass index	-	-	393,038
Height	-	-	393,485
Bradyarrhythmia	8,072	386,333	394,405
Coronary artery disease	20,397	374,018	394,415
Heart failure	6,504	387,652	394,156
Hypercholesterolemia	73,382	321,074	394,456
Hypertension	134,901	259,611	394,512
Mitral regurgitation	2,372	392,042	394,414
Peripheral vascular disease	5,117	389,291	394,408
Smoking [ever]	186,224	206,678	392,902
Stroke	9,927	384,489	394,416
Type II Diabetes	19,744	374,670	394,414

Abbreviations: PheWAS, phenome-wide association study.

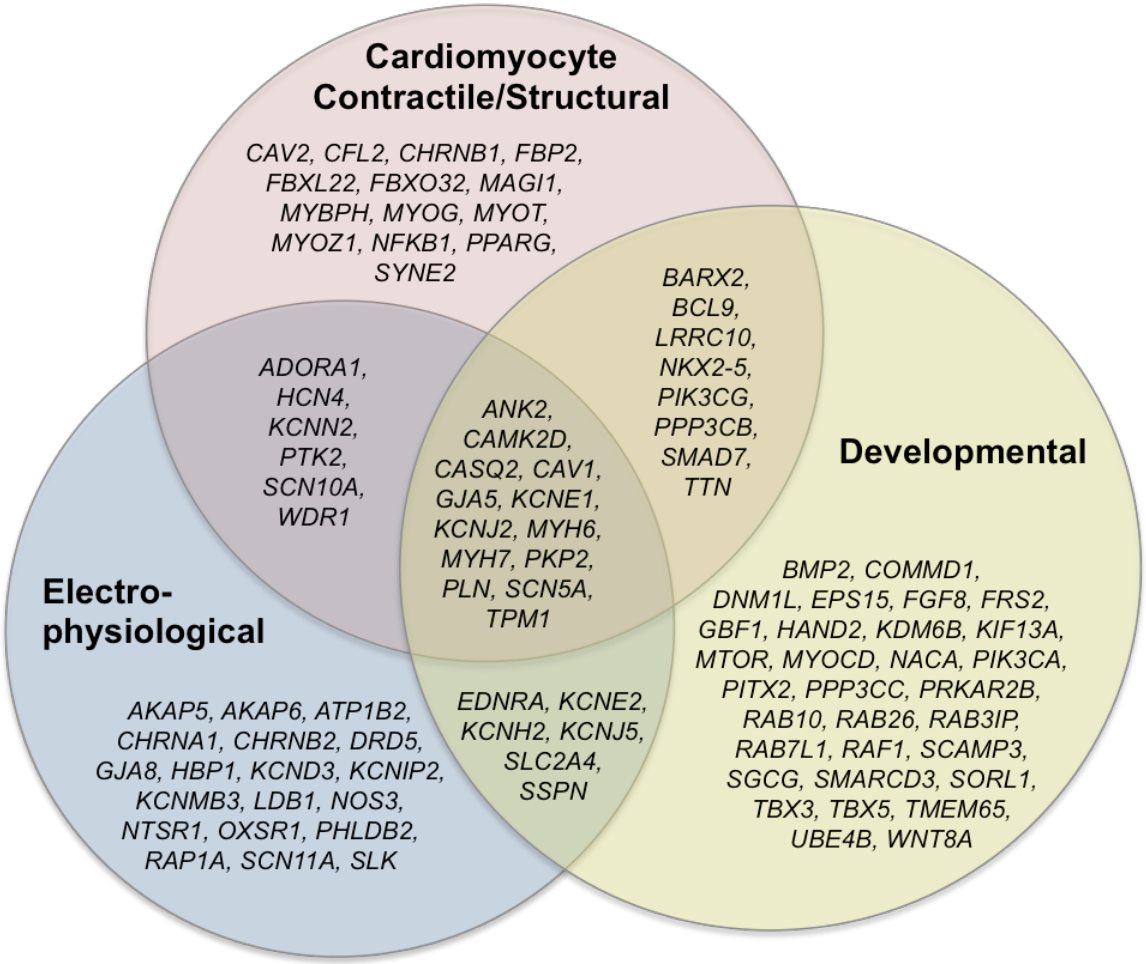
Supplementary Figures

Supplementary Figure S1. Quantile-Quantile plot of combined ancestry meta-analysis

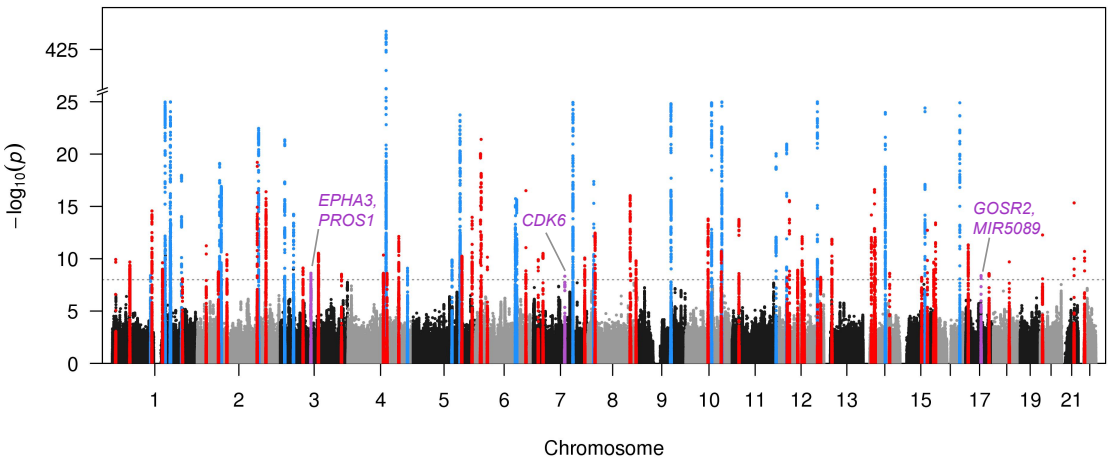


Quantile-Quantile plot of combined ancestry meta-analysis for 12,149,979 included variants and λ_{GC} = 1.0948.

Supplementary Figure S2. Venn diagram for genes near sentinel variants from combined ancestry meta-analysis within enriched gene sets, by functional groups

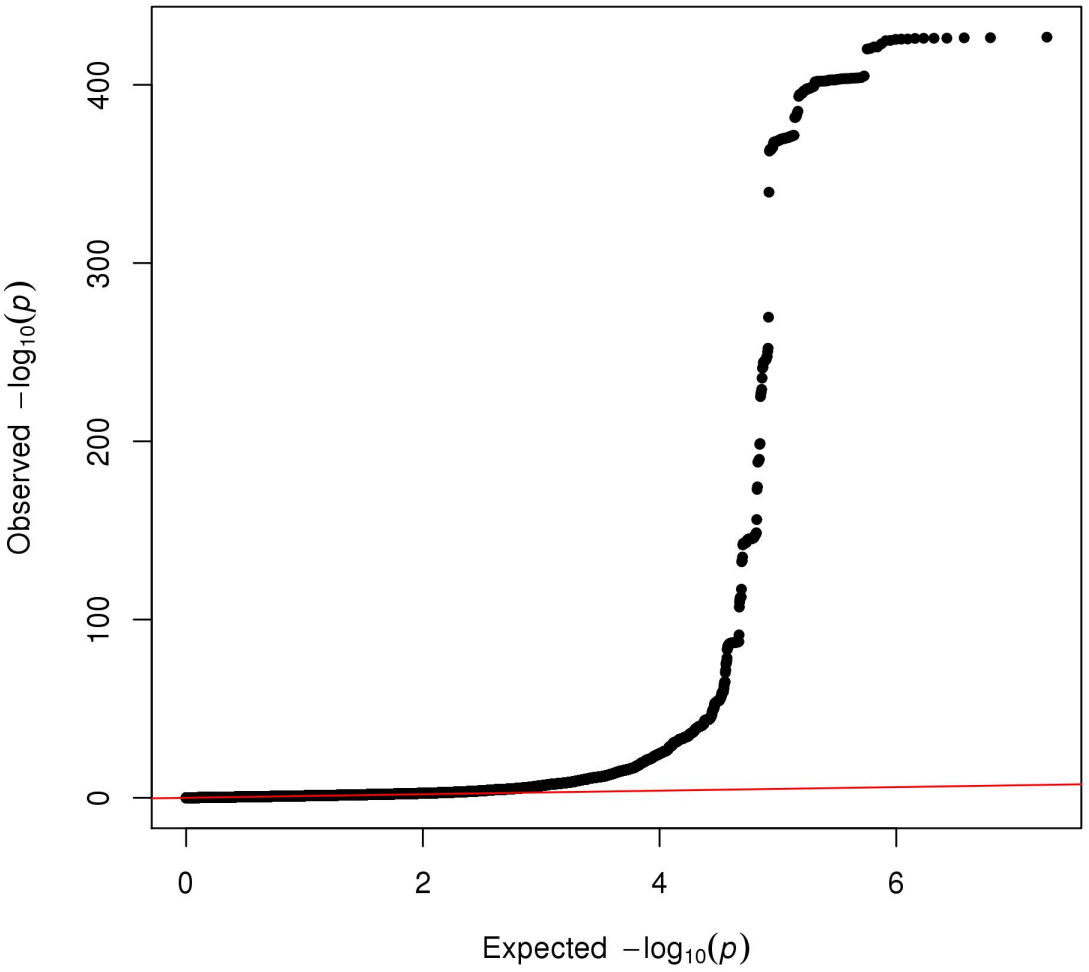


Supplementary Figure S3. Manhattan plot of European ancestry meta-analysis



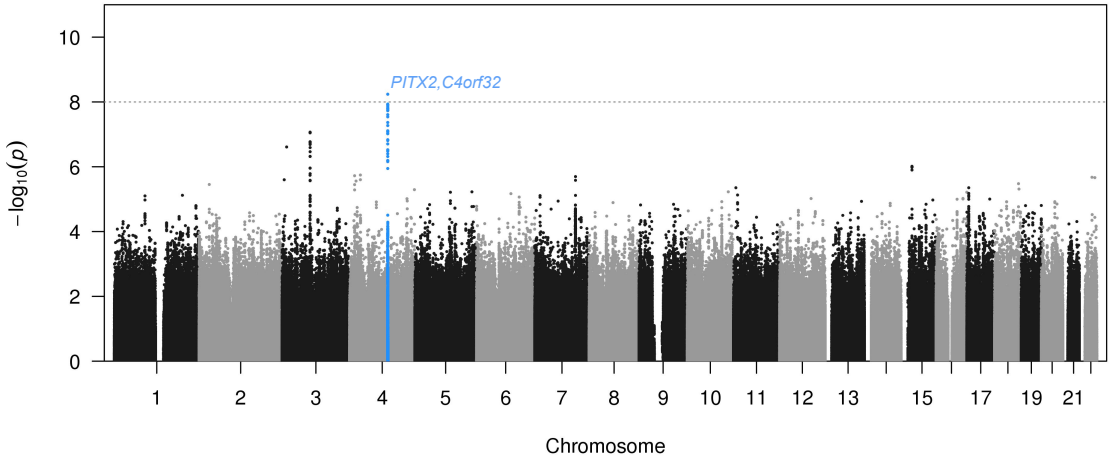
The plot shows novel (red and purple) and known (blue) genetic loci associated with AF at a significance level of $P < 1x10^{-8}$ (dashed line), for the European ancestry meta-analysis. Loci in purple did not reach genome-wide significance in the combined ancestry meta-analysis. Gene labels correspond to the nearest gene(s). The y-axis has a break between –log10(P) of 25 and 400 to emphasize the novel loci.

Supplementary Figure S4. Quantile-Quantile plot of European ancestry meta-analysis



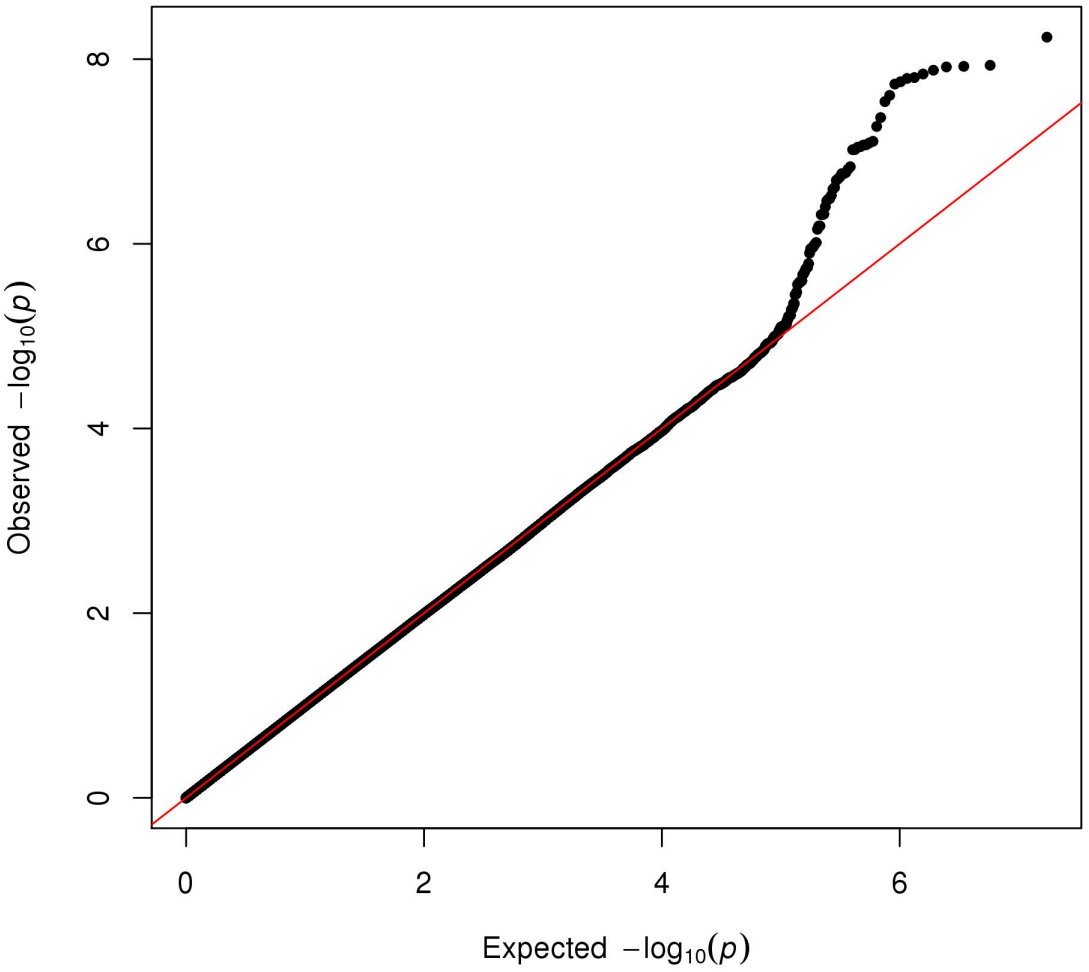
Quantile-Quantile plot of European ancestry meta-analysis for 9,362,422 included variants and λ_{GC} = 1.1194.

Supplementary Figure S5. Manhattan plot of African American meta-analysis



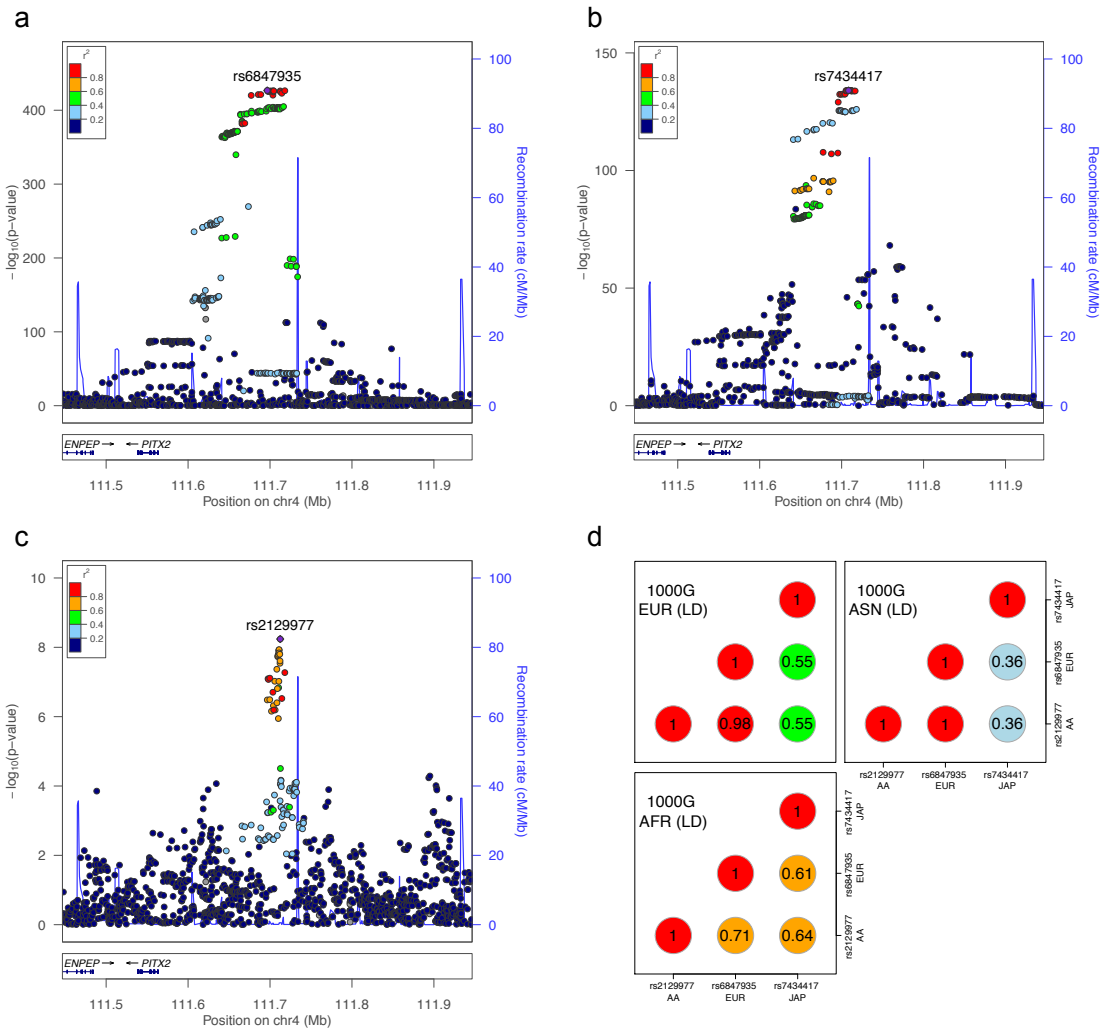
The plot shows known (blue) genetic loci associated with AF at a significance level of $P < 1x10^{-8}$ (dashed line), for the African American ancestry meta-analysis. The gene label corresponds to the nearest genes.

Supplementary Figure S6. Quantile-Quantile plot of African American ancestry meta-analysis



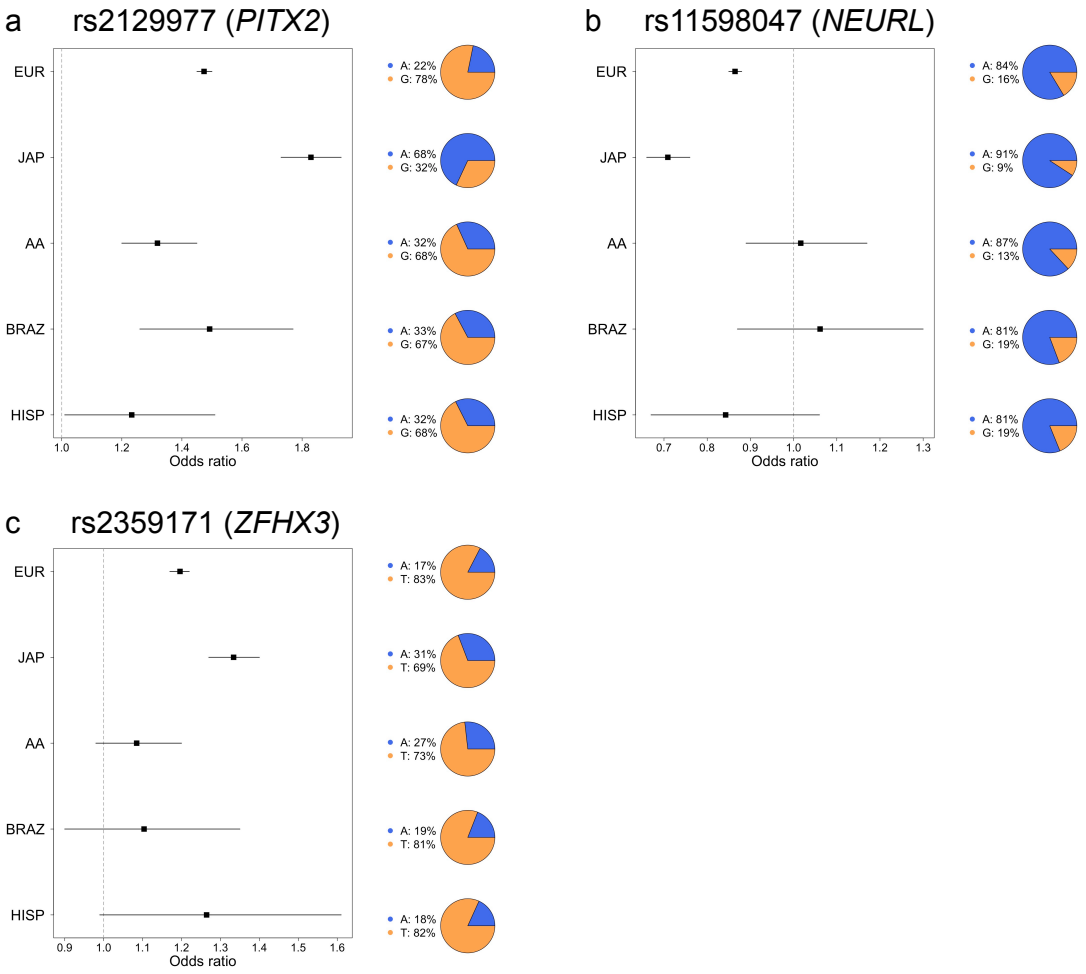
Quantile-Quantile plot of African American ancestry meta-analysis for 8,640,046 included variants and λ_{GC} = 0.997.

Supplementary Figure S7. Regional plots for 4q25 for European, Japanese and African American ancestry



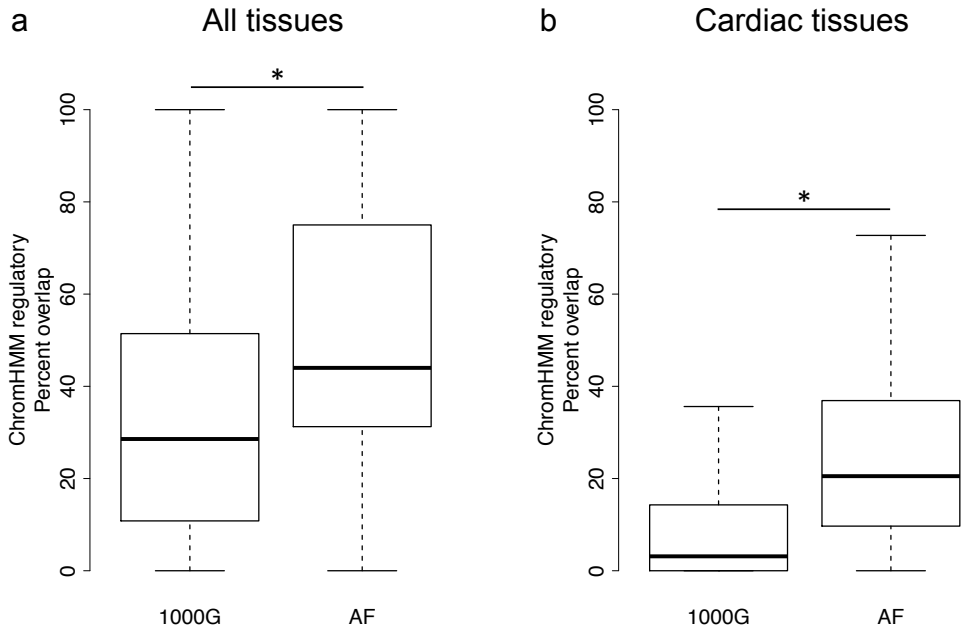
Regional plots of 4q25 created with LocusZoom[2], for European ancestry results (a), Japanese ancestry results (b) and African American ancestry results (c). LD is shown based on 1000 Genomes phase 1 v3 reference, using the populations EUR (a), ASN (b) and AFR (c). Panel d shows the pairwise LD (r^2) for the sentinel variants based on the 1000 Genomes phase 1 v3 reference for each ancestry, calculated with SNiPA[3]. Abbreviations, 1000G, 1000 Genomes, AA, African American, AFR, African, ASN, Asian, EUR, European, JAP, Japanese, LD, linkage disequilibrium.

Supplementary Figure S8. Forest plots of odds ratios, and allele frequency plots, by ancestry for sentinel variants with significant heterogeneity



Forest plots of odds ratios and pie charts of allele frequencies across ancestries for sentinel variants with significant heterogeneity, close to *PITX2* (a), *NEURL* (b), and *ZFHX3* (c). Shown are odds ratios with 95% confidence intervals. Frequencies of the effect allele are depicted in blue, frequencies of the reference allele are depicted in orange. Abbreviations, AA, African American, BRAZ, Brazilian, EUR, European, HISP, Hispanic, JAP, Japanese.

Supplementary Figure S9. Enrichment of atrial fibrillation associated loci across ChromHMM regulatory regions



Percent overlap of loci with regulatory regions (promoter, enhancer, DNase) based on Roadmap Epigenomics Consortium 25-state model across (a) all tissues and (b) cardiac tissues. Each locus includes sentinel variant and proxies with $r^2 > 0.6$. *P < 0.005, one-tailed permutation test (n=1,000). 1000G, 1000 Genomes control loci matched to atrial fibrillation sentinel SNPs via SNPSnap (n=93,000), AF, atrial fibrillation associated loci from combined-ancestry analysis (n=93), sentinel SNP for 1 AF locus could not be matched in SNPSnap and was excluded from this analysis. Whiskers show 1.5 * interquartile range.

Supplementary Notes

Description of participating studies

This manuscript includes the following studies that are described elsewhere: The **Age, Gene/Environment Susceptibility Study (AGES) Reykjavik study**[4], the **Atrial Fibrillation Biobank LMU (AFLMU)** in the context of the **Arrhythmia-Biobank-LMU**[4]**, ANGES**[1]**,** the **Atherosclerosis Risk in Communities (ARIC) study**[4], **BEAT-AF**[1], **Biobank Japan (BBJ)**[5], **Bio*Me***[1], **Cleveland Clinic Lone Atrial Fibrillation GeneBank Study (CCAF)**[4]**,** the **Cardiovascular Health Study (CHS)**[4], **Corogene**[1], **Framingham Heart Study (FHS)**[4], **FINCAVAS**[1], **GS:SFHS**[1], **LURIC**[1], **MDCS**[1], **MESA**[1], **Massachusetts General Hospital (MGH) AF study**[4], **MGH CAMP**[1], **PIVUS**[1], **PREVEND**[1], the **PROspective Study of Pravastatin in the Elderly at Risk (PROSPER)**[6], the **Rotterdam Study (RS)**[4], **SiGN**[7], the **Study of Health in Pomerania (SHIP)**[4], **SPHFC**[1], **TWINGENE**[1], **UK Biobank**[8]**, ULSAM**[1], the **Women's Genome Health Study (WGHS)**[4] and **WTCCC2-Munich**[1]. Additional studies are described as follows:

Australian Familial AF Study: A cohort of probands with familial AF was recruited for genetics studies at the Victor Chang Cardiac Research Institute. Familial AF cases were identified from in-patient and outpatient populations at St Vincent's Hospital and by referral from collaborating physicians throughout Australia. Study subjects underwent clinical evaluation with history, ECG and echocardiogram, and informed consent was obtained from all participants. 151 probands aged <66 years at the time of diagnosis were included in this analysis. The control cohort was comprised of age- and sex-matched individuals (n=151) who had no history of cardiovascular disease.

Danish AF Study: From the Danish Study of Genetic Causes of Atrial Fibrillation (DANFIB) 732 patients with the diagnosis of atrial fibrillation were included. The enrolled individuals were identified after hospitalization at a hospital in the Copenhagen area and/or through a nationwide search using the Danish National Registries. Individuals who accepted participation had blood samples drawn, and clinical information were obtained through questionnaires and subsequently review of available electronic patient health records. A control cohort (n=534) without the diagnosis of atrial fibrillation was recruited from the Copenhagen area. All enrolled participants provided written informed consent.

Duke Biobank: The CATHeterization GENetics (CATHGEN) biorepository collected biospecimens and clinical data on individuals age ≥18 undergoing cardiac catheterization for concern of ischemic heart disease at a single center (Duke University Medical Center) from 2000-2010; a total of N=9334 individuals were collected. Samples were matched at the individual level to clinical data collected at the time of catheterization and stored in the Duke Databank for Cardiovascular Diseases (DDCD). Clinical data included subject demographics, cardiometabolic risk factors, and cardiac history including symptoms, age-of-onset of cardiovascular diseases, coronary anatomy and cardiac function at catheterization, laboratory data, and yearly follow-up for hospitalizations, vital status, and medication use and lifestyle factors. AF cases were defined as individuals who had ever had AF based on any ECG available at Duke University or ICD-9 code for AF used for inpatient or outpatient billing.

EAST - AFNET 4 biomarker substudy (EAST): The samples used here were taken at enrolment into the EAST – AFNET 4 trial from patients consenting to participation in the biomarker substudy. EAST - AFNET 4 is a controlled trial comparing early rhythm control therapy to usual care in patients with recent onset atrial fibrillation (defined as first diagnosis ≤ 12 months or less before randomization) and at least two stroke risk factors as codified in the CHA2DS2-VASc score. Patients are enrolled in 11 European countries. Detailed information is available at www.easttrial.org and in the published design paper[9].

EGCUT: The Estonia biobank is a population-based cohort of the Estonian Genome Center at the University of Tartu (EGCUT), which was established in 2001. Subjects were recruited at random and represent about 5% of the Estonian population.

Genetics in AF (**GENAF):** The Genetics in AF (GENAF) study enrolled individuals with early-onset lone AF before age 50 in Norway between 2009 and 2016[10]. Early-onset was defined as diagnosis of AF before age 50. Lone AF was defined as AF in the absence of clinical or echocardiographic findings of cardiovascular disease, hypertension, and metabolic or pulmonary disease. AF was documented in ECG. All participants underwent clinical examination, including ECG, echocardiography, and blood draw, from which DNA has been extracted. The study conforms to the principles of the Declaration of Helsinki and was approved by the Regional Ethics Committee (REK) in Norway (Protocol reference number: 2009/2224-5). All included patients gave written informed consent.

German MI Family Study (GerMIFS) 6: The study consists of 3,320 individuals from Southern Germany that underwent coronary angiography and cardiac examination. 1,820 were diagnosed to have coronary artery disease with at least 75% stenosis of at least one coronary segment and served as cases. 1,500 individuals were free of coronary artery disease, atrial fibrillation, congestive heart failure, valvular heart disease, and peripheral vascular disease at time of enrollment (control cohort).

Groningen Genetics of Atrial Fibrillation (GGAF): The GGAF cohort (n=2207) is a genotype and phenotype repository of individuals with AF and age- and sex-matched controls from 5 different sources. All studies were approved by the ethical committee, and all individuals provided written informed consent. Individuals with AF (n=1108) were included in 3 registry cohorts at the University Medical Center (www.atrialfibrillationresearch.nl), and Maastricht University Medical Center (AF-Risk n=6). The AF-Risk study (ClinicalTrials.gov Identifier: NCT01510210) is an observational hospital-based cohort (n=500; in GGAF 339) to seek for markers of severity of atrial remodeling and predict outcome of a rhythm control treatment strategy. Patients with a short history of AF were included. Detailed phenotypic information was collected, including non-invasive vascular function measurements, body surface mapping, and detailed information on presence or progression of AF during 5-years follow up is obtained by use of serial ECGs, 24-hour Holter monitoring and recordings from loop recorders. The Young-AF study is an observational hospital-based cohort (n=500; in GGAF 311) to seek describe the phenotypic profile of patients with AF onset at age <60 years and the occurrence of AF progression during 5 years follow up. The phenotypic data that was collected is similar to the AF risk profile study. The Biomarker AF study (ClinicalTrials.gov Identifier: NCT01510197) is an observational hospital-based cohort (n=500; in GGAF 458) to identify a risk profile to guide AF therapy in all-comers with AF. The project is similar in design as the AF risk profile study, with a few modifications. No extra phenotypic information on top of our standard clinical AF protocol was performed, except blood

sampling. Age- and sex-matched individuals without AF (controls) were included from 2 cohorts at the University Medical Center Groningen. The GIPS study is a randomized-controlled trial (n=380; in GGAF 362) to evaluate the effect of metformin treatment on preservation of left ventricular function in patients without diabetes presenting with ST-segment elevation myocardial infarction (STEMI). Mean left ventricular ejection fraction after 4 months, assessed by magnetic resonance imaging was 53.1%, and the use of metformin compared with placebo did not improve left ventricular ejection fraction.[11] The PREVEND cohort study (www.prevend.org) is a community-based cohort study including 8592 inhabitants of the city of Groningen, The Netherlands. PREVEND is one of the AFGen consortium participants, see further for more details on cohort description[12]. In the GGAF cohort we included 742 individuals without AF, not previously included in GWAS.

Genetic Risk Assessment of Defibrillator Events (GRADE): GRADE (The Genetic Risk Assessment of Defibrillator Events) study, designed to identify genetic modifiers of arrhythmic risk. Inclusion criteria were: patients who were ≥18 years of age with a diagnosis of at least moderate systolic left ventricular dysfunction (EF ≤30%), and who had an ICD at the University of Pittsburgh Medical Center, Emory University Medical Center, Massachusetts General Hospital, Ohio State University Medical Center, Mid-Ohio Cardiology or the Pittsburgh Veterans Affairs Medical Center. Subjects were excluded if they had intractable Class IV heart failure, and conditions (other than HF) that were expected to limit survival to less than 6 months. The institutional review boards of participating medical centers approved the study and each patient gave written informed consent prior to participation. This study was conducted according to the guidelines laid down in the Declaration of Helsinki and the trial was registered at www.clinicaltrials.gov (NCT 02045043).

Hopkins: The Johns Hopkins University School of Medicine Atrial Fibrillation Genetics Cohort started collecting data in September 2008 on patients >18 years of age that were referred for catheter ablation of symptomatic drug refractory atrial fibrillation. De-identifiable patient samples and clinical information were collected prior to ablation procedures.

Heart and Vascular Health Study (HVH): The Heart and Vascular Health Study (HVH) is a case-control study of risk factors for development of cardiovascular disease, conducted in the setting of Group Health Cooperative, a large integrated healthcare system in Washington State, USA. In the AF study, plan members assigned a new ICD-9 code of 427.31 or 427.32 in the inpatient or outpatient setting during 2001-2007 were identified. Incident AF was verified by review of medical records with the requirement that the AF be documented by 12-lead electrocardiogram and clinically recognized by a physician, with no previous evidence of AF in the medical record. Control subjects were identified from the Group Health membership, and had no history of atrial fibrillation by medical record review. All subjects had an index date. For cases the index date was the date of AF diagnosis. Control subjects were assigned an index date from the distribution of case index dates. Controls were matched to AF cases on age, sex, race, treated hypertension status, and year of identification. Only individuals age 30 years or older and who self-identified as white were included in this analysis.

Incor Warfarin Study: Participants for the Incor Warfarin Study were prospectively enrolled from the outpatient clinic at the Heart Institute, the University of Sao Paulo Medical School, Sao Paulo, Brazil. Only patients older than 18 years and in current use of warfarin were enrolled. This is a cross-sectional study that included patients treated with warfarin for at least 12

months, prior to enrollment, at the Heart Institute (Incor), University of São Paulo, São Paulo, Brazil. They were enrolled in two different time periods: from September 2011 to March 2012 and from January 2014 to February 2014. Main indications for treatment with warfarin therapy were atrial fibrillation or flutter (59.3%), previous cerebrovascular accident (12.8%), thrombosis or embolus (6.3%) and use of prosthetic heart valves (excluding mechanical valves) requiring chronic anticoagulation (13.5%). Atrial fibrillation status was determined if either atrial flutter or atrial fibrillation was present on a 12-lead ECG at baseline evaluation or prior and could be confirmed by electronic medical record review[13].

Intermountain: The INtermountain Healthcare Biological Samples Collection Project and Investigational REgistry for the On-going Study of Disease Origin, Progression and Treatment (Intermountain INSPIRE Registry) purpose is to collect biological samples, clinical information and laboratory data from Intermountain Healthcare patients. The registry originally collected samples in patients undergoing a coronary angiography as part of the Intermountain Heart Collaborative Study. It has been expanded to collect samples in patients diagnosed with all types of medical conditions, and patients from the general population including those who have not been diagnosed with health related issues. Just over 25,000 individuals have provided samples as part of this registry. The registry enables researchers to develop a comprehensive collection of information that may help in disease management, including determining best medical practices for predicting, preventing and treating medical conditions.

Maastricht AFCT: The purpose of this study was to determine the prevalence of coronary artery disease (CAD) in patients diagnosed with idiopathic paroxysmal AF[14]. Of the patients who underwent cardiac computed tomographic angiography (CTA) in our center between January 2008 and March 2011, we identified a total of 115 consecutive idiopathic paroxysmal AF patients who underwent CTA before electrophysiologic ablation. Patients were compared with age-, sex-, and PROCAM risk score-matched healthy controls in permanent sinus rhythm. All patients were free of hypertension, diabetes, congestive heart failure, previous known coronary artery and peripheral vascular disease, previous stroke, thyroid, pulmonary, and renal disease, and structural abnormalities on echocardiography.

MGH - DOFEGEN: The Genetics of QT Prolongation with Antiarrhythmics (DOFEGEN, ClinicalTrials.gov identifier NCT02439658) is an observational study of subjects admitted to Massachusetts General Hospital for initiation of the anti-arrhythmic medications dofetilide and sotalol, with the purposes of obtaining pharmacogenomic information regarding drug-induced effects on the QT interval. The medications were initiated for either atrial fibrillation or ventricular tachycardia/fibrillation, and subjects underwent repeated ECGs during the hospitalization, which were reviewed by cardiologists for underlying rhythm and validation of computer-measured intervals. The study is presently enrolling, although 66 subjects were included for this investigation.

MGH Stroke Study: The Genetics of Cerebral Hemorrhage on Anticoagulation (GOCHA) study is a multicenter study of the genetics of intracerebral hemorrhage in the USA, based at the Massachusetts General Hospital. The cases are individuals presented with acute primary hemorrhagic stroke, aged more than 55 years. The controls were recruited from ambulatory clinics in the same centers in which cases were enrolled. The Genes Affecting Stroke Risk and Outcome Study (GASROS) is a single-center prospective cohort that enrolled cases with acute ischemic stroke, aged more than 18 years who presented to MGH from 2003 to 2011. Ischemic

stroke was defined as a clinical syndrome of associated with a radiographically proven acute infarction consistent with a vascular pattern and without radiographic evidence of a demyelinating or neoplastic disease or other structural disease. In all subjects, the diagnosis was confirmed by diffusion weighted imaging (DWI) completed within 48 hours after symptom onset. Only patients of self-reported European ancestry were enrolled. Controls were matched to cases on the basis of age, sex and race/ethnicity. In both GOCHA and GASROS, AFib status was determined by reviewing medical records, and/or interview subjects or their families. The diagnosis of AFib was established if the subject either had a pre-existing diagnosis or was diagnosed with AFib in the hospital. The diagnosis was not confirmed by ECG in all cases.

MPP: The Malmö Preventive Project (MPP) was a community-based disease prevention program including 33 346 inhabitants from the city of Malmö in Southern Sweden. Complete birth cohorts between 1921-1949 were invited, and the participation rate was 71%. Participants underwent screening between 1974 to 1992 for cardiovascular risk factors, alcohol abuse, and breast cancer. Between 2002-2006, surviving participants were invited to a reexamination which included blood sampling[15] from which DNA has been extracted. Subjects with prevalent or incident AF were identified from national registers as previously described[16], and cases with DNA were then matched in a 1:1 fashion to controls with DNA from the same cohort by sex, age (±1 year), and date of baseline exam (±1 year). Also, controls required a follow-up exceeding that of the corresponding AF case.

Myocardial Applied Genomics Network (MAGNet) repository: The MAGNet respository (http://www.med.upenn.edu/magnet/) includes samples from normal donors at the time of cardiac transplantation. The study protocol was approved by the Institutional Review Board at the University of Pennsylvania, and all patients provided written informed consent to participate.

Partners HealthCare Biobank (PHB): The Partners HealthCare Biobank is a large research data and sample repository that is embedded within the framework of Partners Personalized Medicine. The Partners Biobank provides banked samples (plasma, serum, DNA and buffy coats), genomic data, and other health information, including data from the Electronic Medical Record (EMR). Using the Research Patient Database Query Tool, a large database of detailed Partners HealthCare electronic health record data, we applied a validated electronic AF ascertainment algorithm[17] to identify unique cases of AF. Briefly, the AF algorithm utilizes diagnostic, procedure, electrocardiographic, and medication data to ascertain the presence of atrial flutter or fibrillation.

Penn Medicine Biobank (Penn): The Penn Medicine BioBank was started in 2009 and aims to recruit patients within the University of Pennsylvania Health System to donate venous blood. All samples are linked to de-identified electronic medical records. Participation is completely voluntary and written and informed consent are obtained prior to sample collection. For this project, all samples were collected within the inpatient and outpatient sections of the cardiovascular division at the University of Pennsylvania. AF cases were limited to adults >18 years of age. AF was ascertained through an ICD-9 diagnosis of atrial fibrillation, atrial flutter or documentation within the medical record.

Texas Cardiac Arrhythmia Institute (TCAI): DECAF trial was conducted at TCAI in 2013 in collaboration with University of Texas at Austin. Four hundred consecutive AF patients

undergoing catheter ablation were enrolled. All participants provided voluntary informed consents. Blood samples were collected before the ablation procedure and labeled with anonymous patient identifier. The researchers at UT Austin responsible for DNA extraction and genetic analysis were blinded about the clinical characteristics and identification of the study participants. AF cases included adults >18 years of age from both sex and all AF types[18].

UCSF: The University of California San Francisco (UCSF) Cardiovascular Research Institute (CVRI) Resource in Arteriosclerosis and Metabolic Disease is an ongoing multi-ethnic study of adults ≥18 years of age which was started in 1989 and now includes 28,000 participants recruited from the UCSF medical system. Within the Resource lies data and biospecimens from nearly 1,000 patients presenting to the electrophysiology laboratory for electrophysiology procedures that were densely phenotyped for electrophysiologic characteristics with biospecimens collected from various intra and extra-cardiac chambers. Phenotyping of all participants was achieved via interview and review of medical records.

UMass: The University Of Massachusetts Medical School (UMMS) Division of Cardiovascular Medicine Research, "Defining time-dependent genetic and transcriptomic responses to cardiac injury among patients with arrhythmias" is an ongoing study of any adult patient ≥18 years of age undergoing an elective electrophysiology study or arrhythmia ablation procedure for a supraventricular or ventricular arrhythmia, including atrial fibrillation (AF). Patients may be in sinus rhythm or atrial fibrillation at the time of the procedure. This study started in 2011 and now includes 580 participants recruited from the UMass Memorial HealthCare Center (UMMHC). De-Identifiable patient samples and clinical information were collected prior to ablation procedures. Using a prospective study design, we quantified plasma expression of 86 miRNAs using high-throughput quantitative reverse transcriptase-polymerase chain reaction (RT-qPCR) in 72 patients undergoing ablation for symptomatic AF and followed through the UMMHC AF Treatment Program. The 72 participants completed a standardized follow-up regimen over 12 months following ablation and were included in the analysis. All participants received per-protocol ECG and cardiac rhythm monitors to assess for AF recurrence. MiRNA levels were normalized and then compared between participants with and without an AF recurrence over the 12-month follow-up period.

Vanderbilt Atrial Fibrillation Registry: The Vanderbilt Atrial Fibrillation Registry is a clinical biorepository for patients with atrial fibrillation. Subjects were prospectively enrolled and longitudinally followed to collect data on disease course and response to AF therapy.

Vanderbilt AF Ablation Registry (VAFAR): The Vanderbilt Atrial Fibrillation Ablation Registry (VAFAR) is a prospective observational registry of subjects undergoing AF ablation (clinicaltrials.gov NCT #02404415). Written informed consent is obtained prior to ablation. DNA is extracted from whole blood collected during the procedure. Baseline clinical data is manually extracted from the medical record and supplemented by patient interview. Subjects are prospectively followed for arrhythmia recurrence post-ablation according to current guidelines.

Vanderbilt BioVU: BioVU is the Vanderbilt University Medical Center's biorepository linked to de-identified electronic health records. BioVU operations[19] and ethical oversight[20] have been described elsewhere. Briefly, DNA is collected from discarded blood samples remaining after routine clinical testing at Vanderbilt outpatient clinics in Nashville, Tennessee and surrounding areas, and is linked to a de-identified version of the patient's electronic health record termed

the “Synthetic Derivative.” AF cases were defined as individuals who were aged >18 years, had an ICD-9 diagnosis for AF or flutter (ICD-9: 427.3, 427.31, and 427.32), or a cardiologist diagnosis of AF as identified by a natural language processing tool from the unstructured free text of the ECG impression. In all instances, patients with a history of a heart transplant were excluded (Current Procedural Terminology: 33935, 3394, and 580; ICD-9: V42.1, 996.83).

Acknowledgments

AFLMU: This work is funded by the European Commission’s Horizon 2020 Framework Programme EU-H2020-PHC-RIA No.633196: CATCH-ME to Dr. Sinner. It is further supported by the DZHK (German Centre for Cardiovascular Research), partner site: Munich Heart Alliance, Munich, Germany to Dr. Kääb, and the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ.

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).

ANGES: The Angiography and Genes Study (ANGES) has been financially supported by the Competitive Research Funding of the Tampere University Hospital (Grant 9M048 and 9N035), the Finnish Cultural Foundation, the Finnish Foundation for Cardiovascular Research, the Emil Aaltonen Foundation, Finland, and the Tampere Tuberculosis Foundation.

ARIC: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), 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. Funding support for “Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium” was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419).

Australian Familial AF Study: This work was supported by the National Health and Medical Research Council of Australia (573732, 1025008), the Estate of the Late R.T. Hall, and the St Vincent’s Clinic Foundation.

BEAT-AF: Swiss National Science Foundation (PP00P3_133681 and PP00P3_159322); Swiss Heart Foundation; University of Basel; University Hospital Basel

BBJ: The BioBank Japan Project was supported by the Ministry of Education, Culture, Sports, Sciences and Technology of the Japanese government.

Bio*Me*: The Mount Sinai Bio*Me* Biobank Program is supported by The Andrea and Charles Bronfman Philanthropies. Analyses of Bio*Me* data was supported in part through the computational resources and staff expertise provided by the Department of Scientific Computing at the Icahn School of Medicine at Mount Sinai.

CCAF: CCAF is funded by National Institutes of Health grants R01 HL090620 and R01 HL111314 to MKC, JB, JS, and DVW, the NIH National Center for Research Resources for Case Western Reserve University and The Cleveland Clinic Clinical and Translational Science Award UL1-RR024989, and the Department of Cardiovascular Medicine philanthropic research fund, Heart and Vascular Institute, Cleveland Clinic

CHS: This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268200960009C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, R01HL130114, and R01HL085251 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Corogene: The Corogene study has been financially supported by the Aarno Koskelo Foundation (partial); the Finnish Foundation for Cardiovascular Research (partial); the Paulo Foundation (partial); Jenny and Antti Wihuri Foundation (partial); EVO funds of Helsinki University Central Hospital; Wellcome Trust, UK (the genetic part).

Danish AF Study: The John and Birthe Meyer Foundation, the Villadsen Family Foundation, the stockbroker Henry Hansen and wife Karla Hansen, born Westergaard, Grant. Danish National Research Foundation Center for Cardiac Arrhythmia, the Research Foundation of the Heart Center Rigshospitalet.

Duke Biobank: This work was supported by grants from the National Heart, Lung and Blood Institute (HL 095987 [Shah] and HL101621 [Kraus].

EAST - AFNET 4 biomarker substudy: Sponsor of the EAST – AFNET 4 biomarker substudy is the Atrial Fibrillation NETwork (AFNET). The EAST – AFNET 4 trial is, supported by the DZHK (German Centre for Cardiovascular Research) and the BMBF (German Ministry of Education and Research), the European Heart Rhythm Association, the Deutsche Herzstiftung, Sanofi, and St Jude Medical. This analysis was partially supported by BHF (British Heart Foundation (FS/13/43/30324 to PK) Leducq Foundation (to PK) and by European Union, grant agreement No 633196 [CATCH ME] (to PK and HO).

EGCUT: Estonian Research Council Grant IUT20-60, EU, H2020 grant 692145, European Union through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012) GENTRANSMED.

FHS: This research was conducted using data and resources from Framingham Heart Study (FHS) 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. HHSN268201500001I; 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. Other support came from 2R01 HL092577, 1R01HL128914.

FINCAVAS: The Finnish Cardiovascular Study (FINCAVAS) has been financially supported by the Competitive Research Funding of the Tampere University Hospital (Grant 9M048 and 9N035), the Finnish Cultural Foundation, the Finnish Foundation for Cardiovascular Research, the Emil Aaltonen Foundation, Finland, and the Tampere Tuberculosis Foundation.

GENAF: Dr. Christophersen is supported by a mobility grant from the Research Council of Norway (240149/F20).

GerMIFS: This work has been funded by German Center for Cardiovascular Research (DZHK), and the Deutsche Forschungsgemeinschaft (DFG) as part of the Sonderforschungsbereich CRC 1123 (B02). This study was also supported by grants from the Fondation Leducq (CADgenomics: Understanding CAD Genes, 12CVD02), the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (e:AtheroSysMed, grant 01ZX1313A-2014), and the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement n° HEALTH-F2-2013-601456 (CVgenes-at-target).

GGAF: The GGAF is supported by funding to the 5 sources that form GGAF. The AF RISK study is supported by the Netherlands Heart Foundation (grant NHS2010B233), and the Center for Translational Molecular Medicine. Both the Young-AF and Biomarker-AF studies are supported by funding from the University Medical Center Groningen. The GIPS-III trial was supported by grant 95103007 from ZonMw, the Netherlands Organization for Health Research and Development. The PREVEND study is supported by the Dutch Kidney Foundation (grant E0.13) and the Netherlands Heart Foundation (grant NHS2010B280). Dr. Rienstra and Prof.Dr. Van Gelder acknowledge support from the Netherlands Cardiovascular Research Initiative: an initiative supported by the Netherlands Heart Foundation, CVON 2014–9: “Reappraisal of Atrial Fibrillation: interaction between hyperCoagulability, Electrical remodelling, and Vascular destabilization in the progression of AF (RACE V)”

GRADE: NIH-NHLBI R01 HL77398 (Genetic Modulators of Sudden Death).

GS:SFHS: Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Dirorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the

Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally" (STRADL) (Reference 104036/Z/14/Z).

Hopkins: The work was made possible by grants K23HL089333 and R01HL116280 to Dr. Nazarian. The Johns Hopkins Atrial Fibrillation Program is funded in part by the Roz and Marvin H Weiner and Family Foundation, the Dr. Francis P. Chiaramonte Foundation, the Marilyn and Christian Poindexter Arrhythmia Research Fund, and the Norbert and Louise Grunwald Cardiac Arrhythmia Research Fund.

HVH: The Heart and Vascular Health Study was supported by grants HL068986, HL085251, HL095080, HL073410, HL091244, and HL088456 from the National Heart, Lung, and Blood Institute, and by the Locke Foundation.

Intermountain: The work was supported by funding from the Dell Loy Hansen Heart Foundation.

LURIC: LURIC was supported by the 7th Framework Program (integrated project AtheroRemo, grant agreement number 201668 and RiskyCAD, grant agreement number 305739) of the European Union, by the INTERREG IV Oberrhein Program (Project A28, Genetic mechanisms of cardiovascular diseases) with support from the European Regional Development Fund (ERDF) and the Wissenschaftsoffensive TMO

Maastricht AFCT: This work was supported by the Maastricht University Medical Center and the Netherlands Cardiovascular Research Initiative: an initiative with support of the Dutch Heart Foundation, CVON 2014-9: Reappraisal of Atrial Fibrillation: interaction between hyperCoagulability, Electrical remodeling, and Vascular destabilization in the progression of AF (RACE V).

MDCS: The MDCS was made possible by grants from the Malmö city council. J. Gustav Smith was supported by The Märta Winkler foundation, Swedish Heart Association, Swedish Heart-Lung Foundation, the European Research Council, the Wallenberg Center for Molecular Medicine at Lund University, the Swedish Research Council, the Crafoord Foundation, governmental funding of clinical research within the Swedish National Health Service, Skåne University Hospital in Lund.

MESA: The Multi-Ethnic Study of Atherosclerosis (MESA) is supported by NIH contracts HHSN2682015000031, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169 and by grants UL1-TR-000040, UL1-TR-001079, and UL1-RR-025005 from NCRR. Funding for MESA SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

MGH AF: This work was supported by grants from the National Institutes of Health to Dr. Ellinor (1RO1HL092577, R01HL128914, K24HL105780). Dr. Christophersen is supported by a mobility grant from the Research Council of Norway [240149/F20]. Dr. Ellinor is also supported by an

Established Investigator Award from the American Heart Association (13EIA14220013) and by the Fondation Leducq (14CVD01). Dr. Lubitz is supported by grants from the NIH (K23HL114724) and by a Doris Duke Charitable Foundation Clinical Scientist Development Award (2014105).

MGH CAMP: The recruitment, collection of samples, and genotyping was supported by Pfizer. Analysis of data was a three-way collaboration between MGH, the Broad Institute, and Pfizer. Dr. Huang is supported by grants from the NIH (NS33335, NS055104).

MGH - DOFEGEN: This work was made possible by grants from the NIH (MR: K23 HL127296-02 and Loan Repayment Program.

MGH Stroke Study: The GOCHA study was supported by the grant R01NS059727 from the NINDS. The GASROS study was supported by the Bugher Foundation of the American Heart Association and the Massachusetts General Hospital Deane Institute for Integrative Study of Atrial Fibrillation and Stroke and the grant U01 NS069208-01 and K23NS064052 (N.S.R.) from the NINDS and by a Doris Duke Charitable Foundation Clinical Scientist Development Award (2014105).

MPP: The re-examination of the MPP study was supported by The Swedish Heart-Lung Foundation, Merck, Sharp & Dohme, Hulda and E Conrad Mossfelts Foundation and Ernhold Lundströms Foundation. J. Gustav Smith was supported by Swedish Heart-Lung Foundation, the Swedish Research Council, the Crafoord Foundation, governmental funding of clinical research within the Swedish National Health Service, Skåne University Hospital in Lund, the European Research Council and the Wallenberg Center for Molecular Medicine at Lund University.

Penn Biobank: This work was supported by grants from the National Institutes of Health to Dr. Deo (K23-DK089118).

PIVUS: These projects were supported by Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellow), Swedish Diabetes Foundation (2013-024), Swedish Research Council (2012-1397 and 2015-02907), and Swedish Heart-Lung Foundation (20140422). Computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. Genotyping was funded by the Wellcome Trust under awards WT064890 and WT086596. Analysis of genetic data was funded by the Wellcome Trust under awards WT098017 and WT090532. Andrew P Morris is a Wellcome Trust Senior Fellow in Basic Biomedical Science, under award WT098017.

PREVEND: The PREVEND study is supported by the Dutch Kidney Foundation (grant E0.13) and the Netherlands Heart Foundation (grant NHS2010B280. Dr. M. Rienstra is supported by grants from the Netherlands Organization for Scientific Research (Veni grant 016.136.055), the EHRA Academic Fellowship program, and from the Netherlands Cardiovascular Research Initiative: an initiative supported by the Netherlands Heart Foundation, CVON 2014–9: “Reappraisal of Atrial Fibrillation: interaction between hyperCoagulability, Electrical remodelling, and Vascular destabilization in the progression of atrial fibrillation (RACE V)”.

PROSPER: The PROSPER study was supported by an investigator initiated grant obtained from Bristol-Myers Squibb. Prof. Dr. J. W. Jukema is an Established Clinical Investigator of the Netherlands Heart Foundation (grant 2001 D 032). Support for genotyping was provided by the

seventh framework program of the European commission (grant 223004) and by the Netherlands Genomics Initiative (Netherlands Consortium for Healthy Aging grant 050-060-810).

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 Organization for Scientific Research (NWO) project nr. 050-060-810.

SHIP: SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthineers, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the Caché Campus program of the InterSystems GmbH.

SPHFC: The Heart Failure Cohort was supported by the Samaritan Hospital of São Paulo and Brazilian Ministry of Health - Programa de Apoio ao Desenvolvimento Institucional do Sistema Único de Saúde (PROADI-SUS), the Sao Paulo Research Agency Fapesp (grant 2013/17368-0).

TWINGENE: These projects were supported by Ministry for Higher Education, GenomEUtwin (EU/QLRT-2001-01254; QLG2-CT-2002-01254), NIH grant DK U01-066134, Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellow), Swedish Diabetes Foundation (2013-024), Swedish Research Council (M-2005-1112, 2009-2298, 2012-1397 and 2015-02907), and Swedish Heart-Lung Foundation (20140422). Computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. We thank the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se) for excellent genotyping.

UCSF: This work was made possible by grants from the NIH (KL2 8K12RR023262) and the American Heart Association Western States Affiliate Beginning Grant-in-Aid to Dr. Marcus.

UK Biobank: This research has been conducted using the UK Biobank resource, application 17488**.**

ULSAM: These projects were supported by Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellow), Swedish Diabetes Foundation (2013-024), Swedish Research Council (2012-1397 and 2015-02907), and Swedish Heart-Lung Foundation (20140422). Computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. Genotyping was funded by the Wellcome Trust under awards WT064890 and WT086596. Analysis of genetic data was

funded by the Wellcome Trust under awards WT098017 and WT090532. We thank the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se) for excellent genotyping.

UMass: This study was supported by Grants 1U01HL105268-01 and KL2RR031981 to Dr. McManus; N01-HC 25195, 6R01-NS 17950, RFA-HL-12-008, and 1R01 HL64753 to Dr. Freedman; R01 HL087201A to Drs. Freedman and Tanriverdi; and R01 HL076784 and 1 R01 AG028321 to Dr. Benjamin from the National Heart, Lung, Blood Institute, National Institutes of Health

Vanderbilt Atrial Fibrillation Registry: This work was supported by National Institutes of Health grants U19 HL65962, HL092217, and an American Heart Association Established Investigator Award (0940116N).

Vanderbilt BioVU: The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU which is supported by institutional funding, the 1S10RR025141-01 instrumentation award, and by the CTSA grant UL1TR000445 from NCATS/NIH.

Vanderbilt AF Ablation Registry (VAFAR): This work was supported by grants from the National Institutes of Health: K23 HL127704 (Shoemaker), and CTSA grant UL1TR000445 from NCATS/NIH.

WGHS: The Women's Genome Health Study (WGHS) is supported by the National Heart, Lung, and Blood Institute (HL043851 and HL080467 [both Buring]) and the National Cancer Institute (CA047988 [Buring] and UM1CA182913 [Buring and Lee]) with collaborative scientific support and funding for genotyping provided by Amgen (Chasman and Ridker). Atrial fibrillation endpoint confirmation was supported by HL-093613 (Albert) and HL116690 (Albert) and a grant from the Harris Family and Watkin's Foundation (Tedrow).

WTCCC2-Munich: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 666881, SVDs@target (to M Dichgans) and No 667375, CoSTREAM (to M Dichgans); the DFG as part of the Munich Cluster for Systems Neurology (EXC 1010 SyNergy) and the CRC 1123 (B3)(to M Dichgans); the Corona Foundation (to M Dichgans); the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain)(to M Dichgans); the e:Med program (e:AtheroSysMed) (to M Dichgans) and the FP7/2007-2103 European Union project CVgenes@target (grant agreement number Health-F2-2013-601456) (to M Dichgans).

Supplementary References

1. Christophersen, I. E. *et al.* Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nat. Genet.* **49,** 946–952 (2017).
2. Pruim, R. J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26,** 2336–7 (2010).
3. Arnold, M., Raffler, J., Pfeufer, A., Suhre, K. & Kastenmüller, G. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* **31,** 1334–6 (2015).
4. Ellinor, P. T. *et al.* Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat. Genet.* **44,** 670–5 (2012).
5. Low, S.-K. *et al.* Identification of six new genetic loci associated with atrial fibrillation in the Japanese population. *Nat. Genet.* **49,** 953–958 (2017).
6. Sinner, M. F. *et al.* Integrating genetic, transcriptional, and functional analyses to identify 5 novel genes for atrial fibrillation. *Circulation* **130,** 1225–35 (2014).
7. Pulit, S. L. *et al.* Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol.* **15,** 174–184 (2016).
8. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. **12,** e1001779 (2015).
9. Kirchhof, P. *et al.* Improving outcomes in patients with atrial fibrillation: Rationale and design of the Early treatment of Atrial fibrillation for Stroke prevention Trial. *Am. Heart J.* **166,** 442–448 (2013).
10. Christophersen, I. E. *et al.* Genetic variation in KCNA5: impact on the atrial-specific potassium current IKur in patients with lone atrial fibrillation. *Eur. Heart J.* **34,** 1517–25 (2013).
11. Lexis, C. P. H. *et al.* Effect of metformin on left ventricular function after acute myocardial infarction in patients without diabetes: the GIPS-III randomized clinical trial. *JAMA* **311,** 1526–35 (2014).
12. Vermond, R. A. *et al.* Incidence of Atrial Fibrillation and Relationship With Cardiovascular Events, Heart Failure, and Mortality: A Community-Based Study From the Netherlands. *J. Am. Coll. Cardiol.* **66,** 1000–7 (2015).
13. Santos, P. C. J. L. *et al.* Development of a pharmacogenetic-based warfarin dosing algorithm and its performance in Brazilian patients: highlighting the importance of population-specific calibration. *Pharmacogenomics* **16,** 865–876 (2015).
14. Weijs, B. *et al.* Patients originally diagnosed with idiopathic atrial fibrillation more often suffer from insidious coronary artery disease compared to healthy sinus rhythm controls. *Hear. Rhythm* **9,** 1923–1929 (2012).
15. Lyssenko, V. *et al.* Clinical Risk Factors, DNA Variants, and the Development of Type 2 Diabetes. *N. Engl. J. Med.* **359,** 2220–2232 (2008).
16. Smith, J. G., Platonov, P. G., Hedblad, B., Engström, G. & Melander, O. Atrial fibrillation in the Malmö Diet and Cancer study: a study of occurrence, risk factors and diagnostic validity. *Eur. J. Epidemiol.* **25,** 95–102 (2010).
17. Khurshid, S., Keaney, J., Ellinor, P. T. & Lubitz, S. A. A Simple and Portable Algorithm for Identifying Atrial Fibrillation in the Electronic Medical Record. *Am. J. Cardiol.* **117,** 221–225 (2016).
18. Mohanty, S. *et al.* Novel association of polymorphic genetic variants with predictors of outcome of catheter ablation in atrial fibrillation: new directions from a prospective study (DECAF). *J. Interv. Card. Electrophysiol.* **45,** 7–17 (2016).

19. Roden, D. M. *et al.* Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin. Pharmacol. Ther.* **84,** 362–9 (2008).
20. Pulley, J., Clayton, E., Bernard, G. R., Roden, D. M. & Masys, D. R. Principles of human subjects protections applied in an opt-out, de-identified biobank. *Clin. Transl. Sci.* **3,** 42–8 (2010).