Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation

Christophersen, Ingrid E.; Rienstra, Michiel; Roselli, Carolina; Yin, Xiaoyan; Geelhoed, Bastiaan; Barnard, John; Lin, Honghuang; Arking, Dan E.; Smith, Albert V.; Albert, Christine M.; Chaffin, Mark; Tucker, Nathan R.; Li, Molong; Klarin, Derek; Bihlmeyer, Nathan A.; Low, Siew-Kee; Weeke, Peter E.; Mueller-Nurasyid, Martina; Smith, J. Gustav; Brody, Jennifer A.; Niemeijer, Maartje N.; Doerr, Marcus; Trompet, Stella; Huffman, Jennifer E.; Gustafsson, Stefan; Schurmann, Claudia; Kleber, Marcus E.; Lyytikainen, Leo-Pekka; Seppala, Ilkka; Malik, Rainer; Horimoto, Andrea R. V. R.; Perez, Marco; Sinisalo, Juha; Aeschbacher, Stefanie; Theriault, Sebastien; Yao, Jie; Radmanesh, Farid; Weiss, Stefan; Teumer, Alexander; Choi, Seung Hoan; Weng, Lu-Chen; Clauss, Sebastian; Deo, Rajat; Rader, Daniel J.; Shah, Svati H.; Sun, Albert; Hopewell, Jemma C.; Debette, Stephanie; Chauhan, Ganesh; Yang, Qiong; Worrall, Bradford B.; Pare, Guillaume; Kamatani, Yoichiro; Hagemeijer, Yanick P.; Verweij, Niek; Siland, Joylene E.; Kubo, Michiaki; Smith, Jonathan D.; Van Wagoner, David R.; Bis, Joshua C.; Perz, Siegfried; Psaty, Bruce M.; Ridker, Paul M.; Magnani, Jared W.; Harris, Tamara B.; Launer, Lenore J.; Shoemaker, M. Benjamin; Padmanabhan, Sandosh; Haessler, Jeffrey; Bartz, Traci M.; Waldenberger, Melanie; Lichtner, Peter; Arendt, Marina; Krieger, Jose E.; Kahonen, Mika; Risch, Lorenz; Mansur, Alfredo J.; Peters, Annette; Smith, Blair; Lind, Lars; Scott, Stuart A.; Lu, Yingchang; Bottinger, Erwin B.; Hernesniemi, Jussi; Lindgren, Cecilia M.; Wong, Jorge A.; Huang, Jie; Eskola, Markku; Morris, Andrew P.; Ford, Ian; Reiner, Alex P.; Delgado, Graciela; Chen, Lin Y.; Chen, Yii-Der Ida; Sandhu, Roopinder K.; Li, Man; Boerwinkle, Eric; Eisele, Lewin; Lannfelt, Lars; Rost, Natalia; Anderson, Christopher D.; Taylor, Kent D.; Campbell, Archie; Magnusson, Patrik K. E.; Porteous, David J.; Hocking, Lynne J.; Vlachopoulou, Efthymia; Pedersen, Nancy L.; Nikus, Kjell; Orho-Melander, Marju; Hamsten, Anders; Heeringa, Jan; Denny, Joshua C.; Kriebel, Jennifer; Darbar, Dawood; Newton-Cheh, Christopher; Shaffer, Christian M.; Macfarlane, Peter W; Heilmann-Heimbach, Stefanie; Almgren, Peter; Huang, Paul L.; Sotoodehnia, Nona; Soliman, Elsayed Z.; Uitterlinden, Andre G.; Hofman, Albert; Franco, Oscar H.; Voelker, Uwe; Joeckel, Karl-Heinz; Sinner, Moritz F.; Lin, Henry J; Guo, Xiuqing; Dichgans, Martin; Ingelsson, Erik; Kooperberg, Charles; Melander, Olle; Loos, Ruth J. F.; Laurikka, Jari; Conen, David; Rosand, Jonathan; van der Harst, Pim; Lokki, Marja-Liisa; Kathiresan, Sekar; Pereira, Alexandre; Jukema, J. Wouter; Hayward, Caroline; Rotter, Jerome I.; Maerz, Winfried; Lehtimaki, Terho; Stricker, Bruno H. C.; Chung, Mina K.; Felix, Stephan B.; Gudnason, Vilmundur; Alonso, Alvaro; Roden, Dan M.; Kaeaeb, Stefan; Chasman, Daniel I.; Heckbert, Susan R.; Benjamin, Emelia J.; Tanaka, Toshihiro; Lunetta, Kathryn L.; Lubitz, Steven A.; Ellinor, Patrick T.

Published in:
Nature Genetics

DOI:
10.1038/ng.3843

Publication date:
2017

SUPPLEMENTARY MATERIAL
Common and Rare Variant Analyses in >150,000 Individuals Reveal Novel Genetic Loci Associated with Atrial Fibrillation

INDEX

1. **Supplementary figures**

Figure S1. QQ plot for the combined ancestry GWAS meta-analysis

Figure S2. Regional plots from combined ancestry ExWAS meta-analysis

Figure S3. QQ plot for the combined ancestry ExWAS meta-analysis

Figure S4. Manhattan plot African American ancestry GWAS meta-analysis

Figure S5. Manhattan plot European ancestry GWAS meta-analyses

Figure S6. Manhattan plot Asian ancestry GWAS meta-analyses

Figure S7. Manhattan plot of results from incident AF GWAS meta-analysis in Europeans

Figure S8. Manhattan plot of results from prevalent AF GWAS meta-analysis in Europeans

Figure S9. Distinct loci on chromosome 1 and 10, as demonstreated by using approximate joint and conditional association analysis in European ancestry studies with GCTA software.

Figure S10. AF associated loci display pleiotropy across clinical, electrocardiographic, and echocardiographic cardiac phenotypes

Figure S11. AF associated loci are enriched for functional elements

Figure S12. AF associated loci are enriched for eQTLs

2. Supplementary tables

Table S1. Baseline characteristics GWAS

Table S2. Baseline characteristics ExWAS

Table S3. Detailed description of the genes at novel AF loci

Table S4. Results from Asian ancestry SKAT gene based test

Table S5. Single variant association results for the variants that were analyzed in the two significant gene-based tests for *SH3PDX2A* in the Asian ancestry group

Table S6. Results from ancestry-specific GWAS meta-analyses

Table S7. Results from European and Asian ancestry ExWAS meta-analyses

Table S8. Results from European incident AF GWAS meta-analysis

Table S9. Results from European prevalent AF GWAS meta-analysis

Table S10. Comparison of results for common variant loci between the AFGen Consortium combined ancestry analysis and the Biobank Japan study

Table S11. Comparison of results for common variant loci between the AFGen Consortium combined ancestry analysis and the UK Biobank study.

Table S12. Approximate and joint conditional analysis in European ancestry results identify 20 independent genetic loci associated with AF

Table S13. Overlap with AF risk factor GWAS loci

Table S14. Association between novel AF loci and stroke subtypes in the Neuro-CHARGE Stroke Consortium

Table S15. Association between novel AF loci and stroke subtypes in the Metastroke Consortium

Table S16. GO terms enriched in AF associated loci compared to GWAS catalog loci and to 1000 genomes matched loci

Table S17. Summary of top eQTLS within atrial fibrillation associated loci

Table S18. *In silico* eQTL analysis in GTEx database (enclosed excel file)

Table S19. eQTL analysis of in CCAF human atrial tissue samples (enclosed excel file)

Table S20. *In silico* functional evaluation of novel and replicated loci from GWAS and ExWAS combined ancestry analysis

Table S21. Per study overlap of samples between GWAS and exome chip analyses

Table S22. GWAS information per study

Table S23. Quality control and filtering

Table S24. ExWAS information per study

Table S25. Baseline characteristics of African American ancestry replication studies

Table S26. Results from replication in African American ancestry studies

Table S27. Results from DEPICT pathway analysis of GWAS meta-analysis results

Table S28. Top 5 enriched canonical pathways from Ingenuity Pathway Analysis of GWAS meta-analysis results

Table S29. Enriched diseases or functions annotation from Ingenuity canonical pathway analysis of GWAS meta-analysis results

3. Supplementary Methods

Replication of genetic variants specific to African American ancestry GWAS meta-analysis

In silico replication in the Biobank Japan (BBJ)

In silico replication in the UK Biobank

In silico evaluation of novel AF loci

Pathway analysis

1. DEPICT
2. IPA

Detailed description of participating studies

4. Supplementary Results

Ancestry-specific GWAS meta-analyses

GWAS meta-analyses of incident and prevalent AF in Europeans

Replication of genetic variants specific to African American ancestry GWAS meta-analysis

Pathway analysis

1. DEPICT

2. IPA

5. Funding Information

6. Supplementary References

1. SUPPLEMENTARY FIGURES

Figure S1. QQ plot for the combined ancestry GWAS meta-analysis

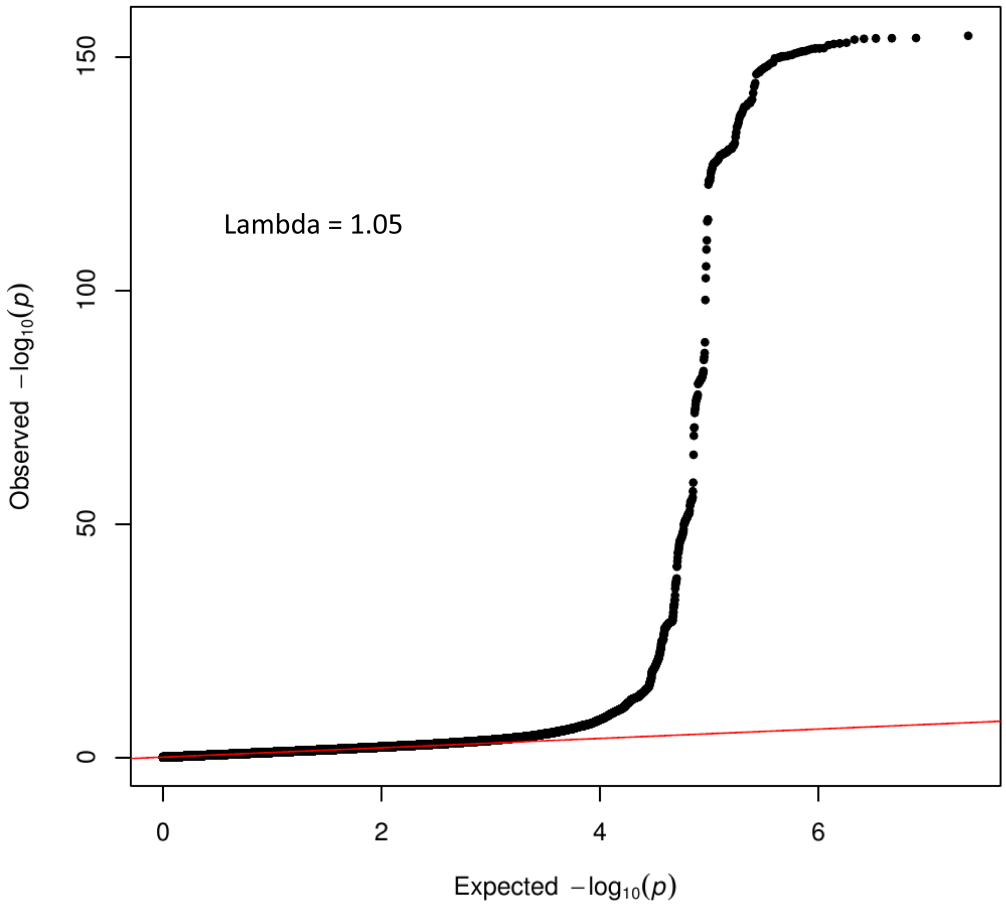


Figure S2. Regional plots from combined ancestry ExWAS meta-analysis

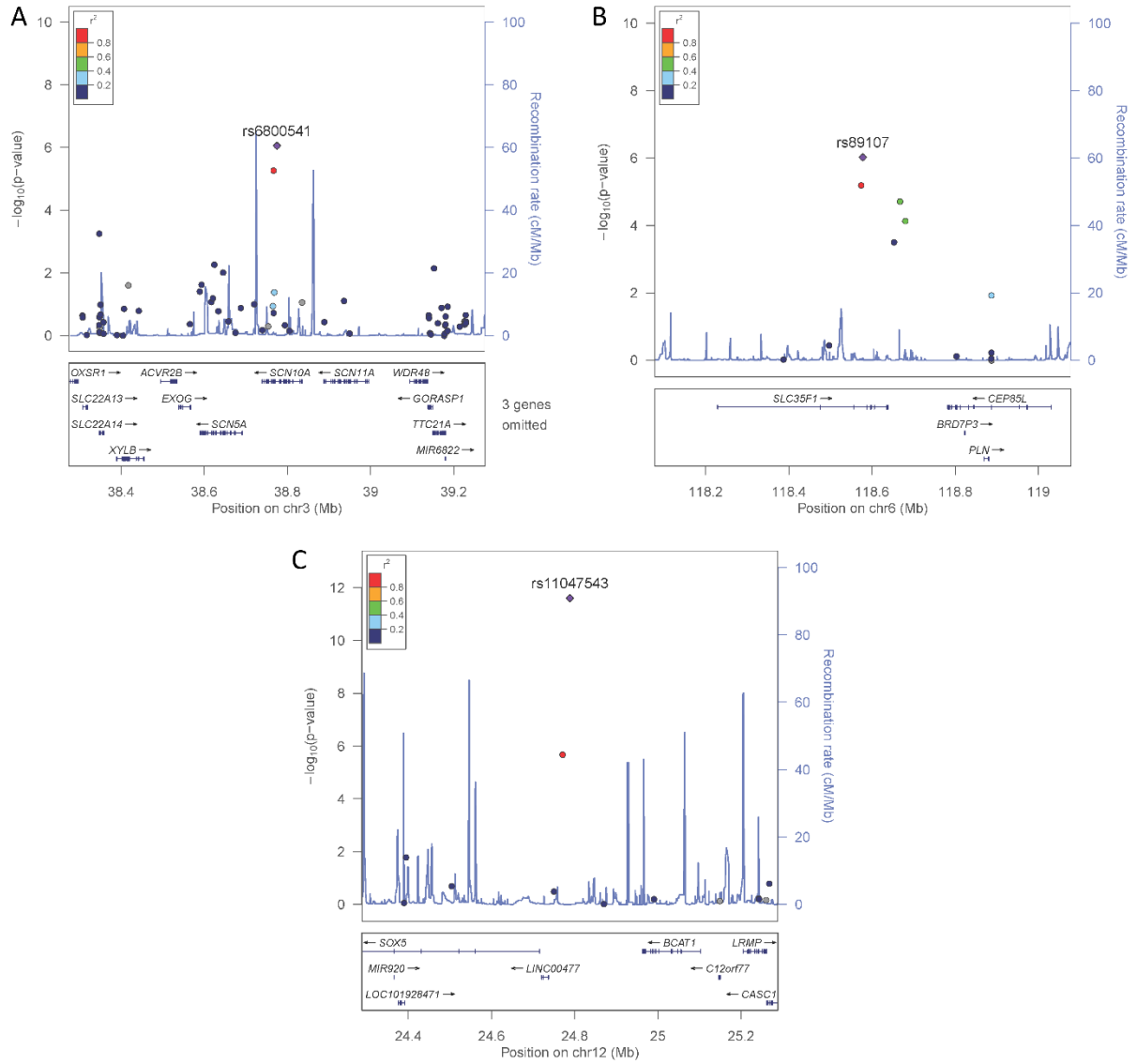


Figure S3. QQ plot for the combined ancestry ExWAS meta-analysis

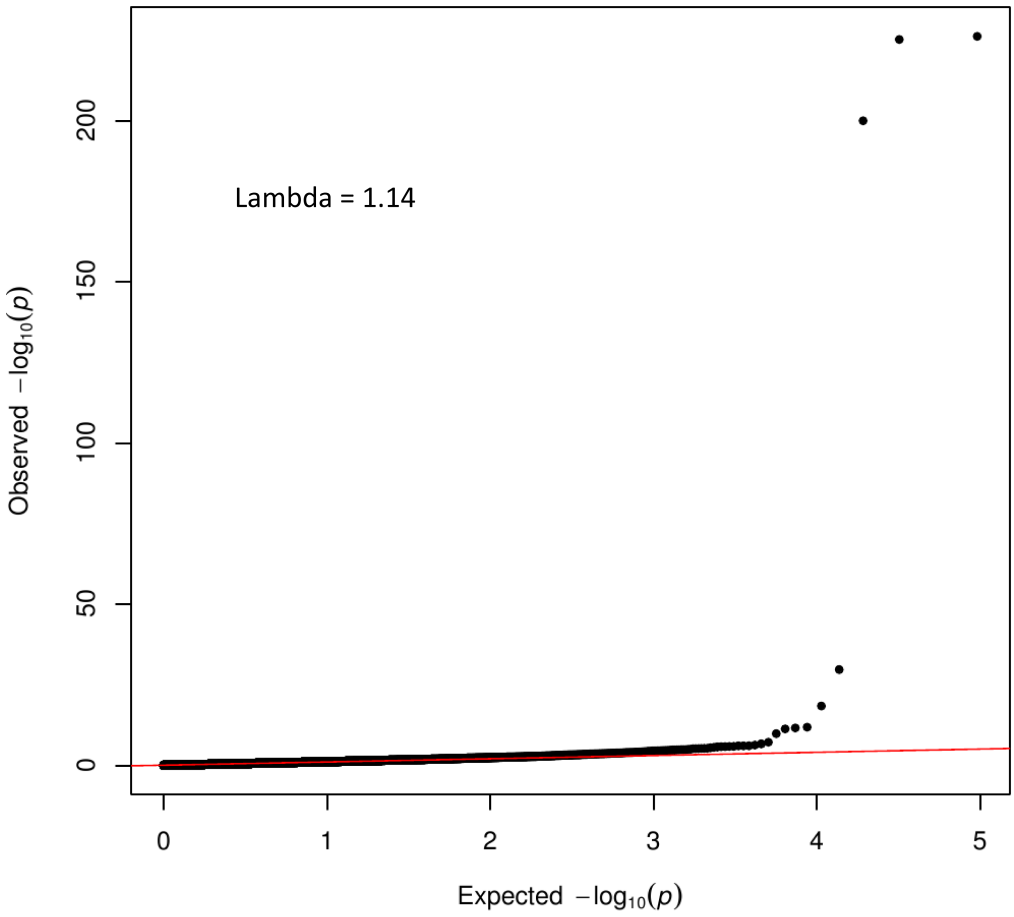


Figure S4. Manhattan plot of results from African American ancestry GWAS meta-analysis

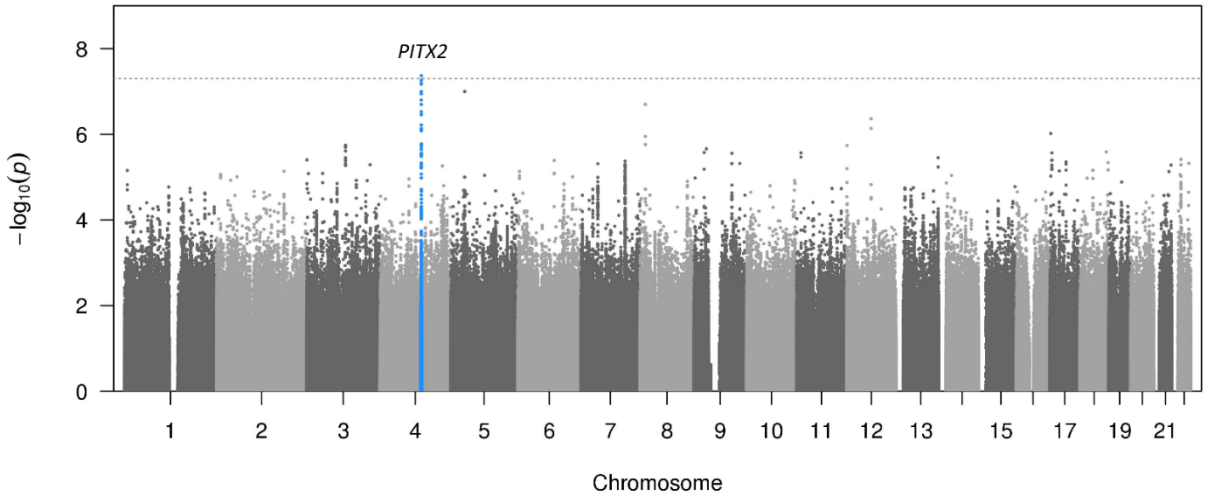


Figure S5. Manhattan plot of results from European ancestry GWAS meta-analysis

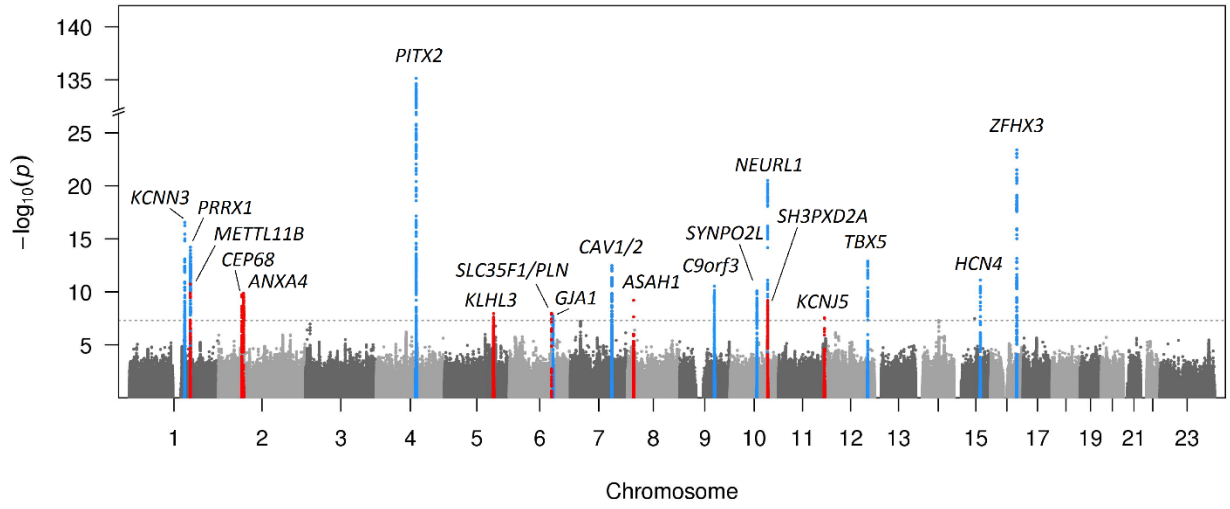


Figure S6. Manhattan plot of results from Asian ancestry GWAS meta-analysis

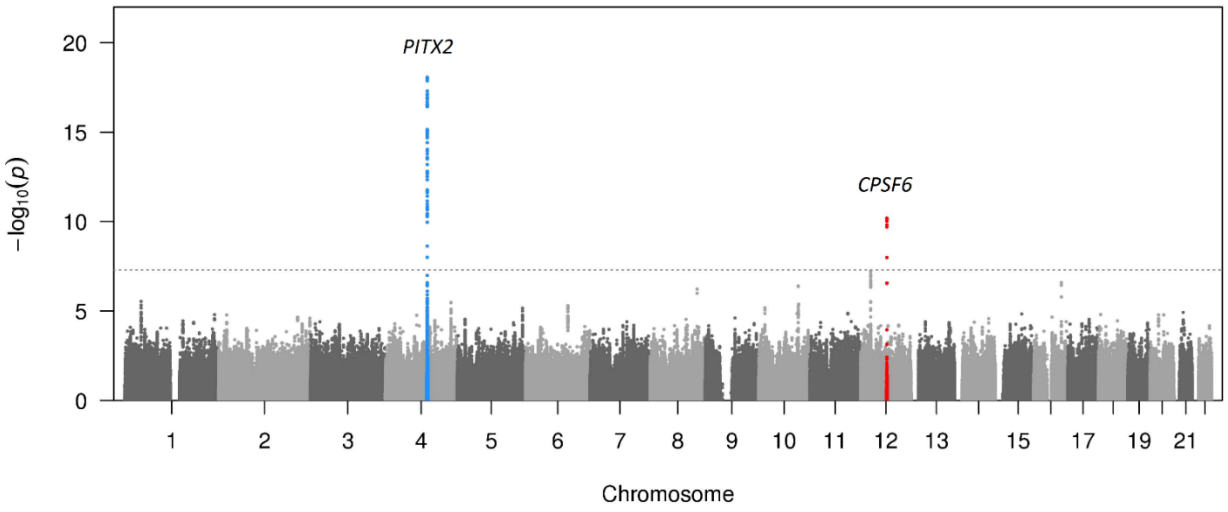


Figure S7. Manhattan plot of results from incident AF GWAS meta-analysis in Europeans

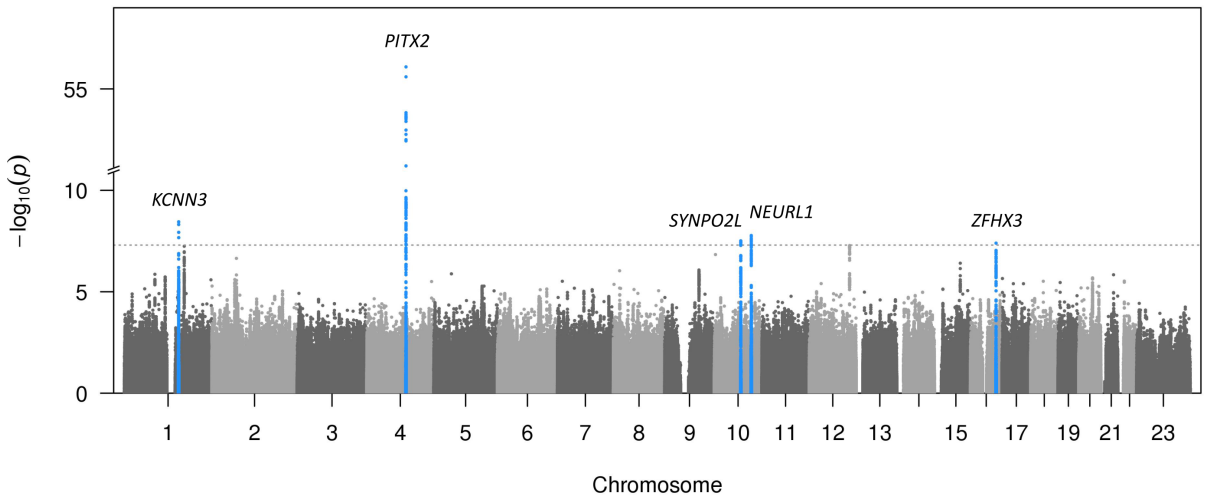


Figure S8. Manhattan plot of results from prevalent AF GWAS meta-analysis in Europeans

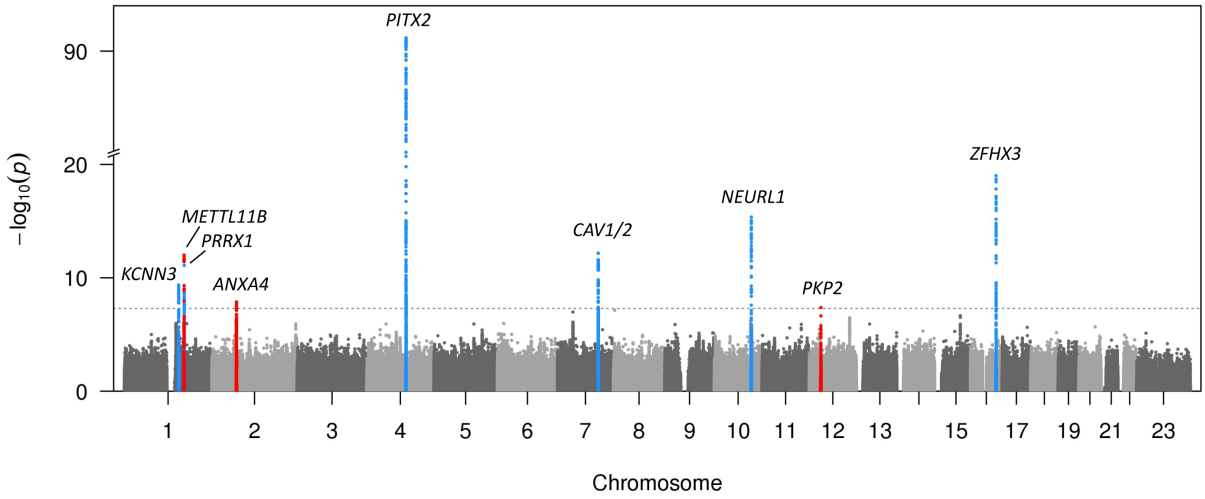
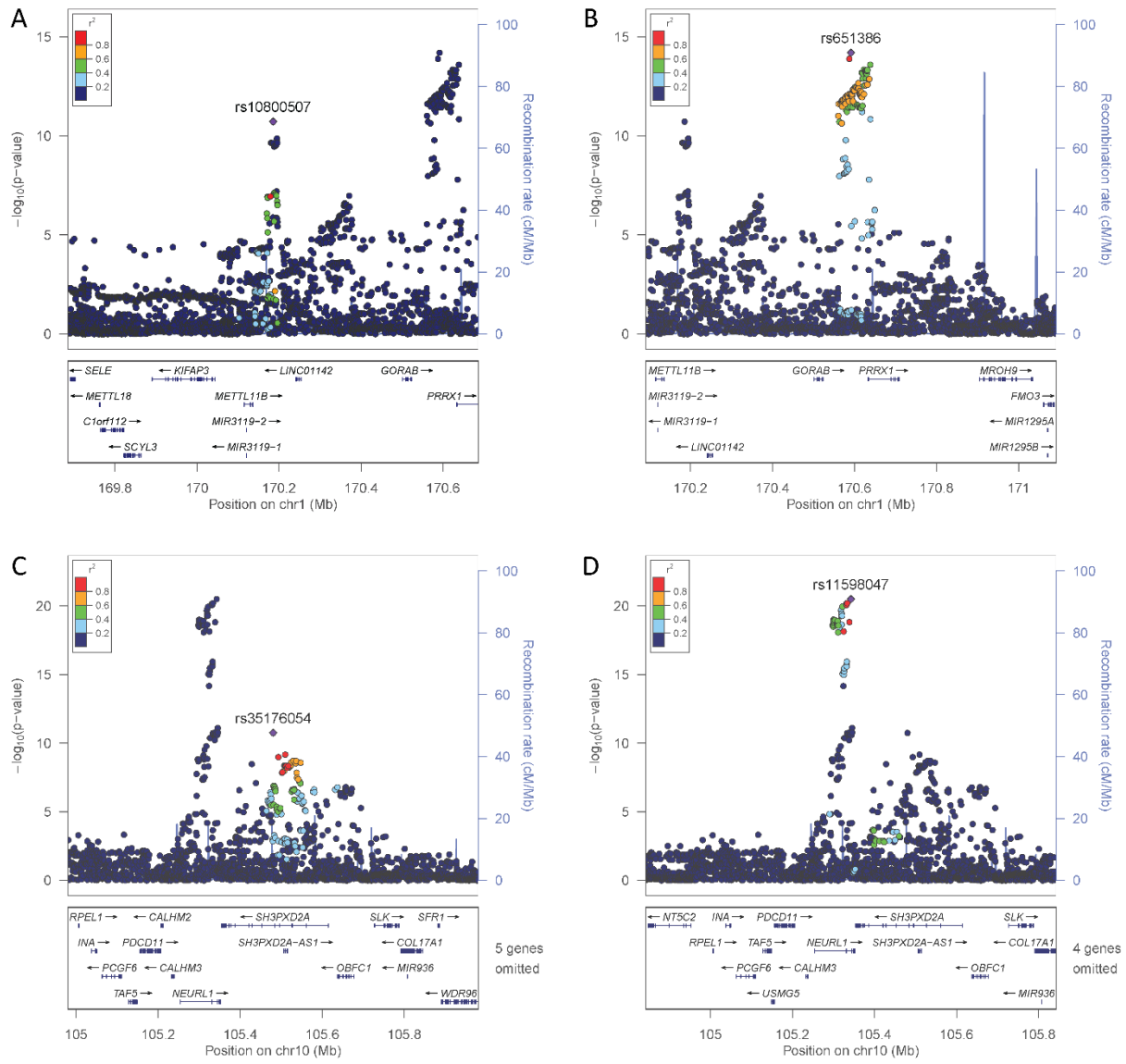


Figure S9. Distinct loci on chromosome 1 and 10, as demonstreated by using approximate joint and conditional association analysis in European ancestry studies with GCTA software.



All conditional analyses were performed using the European ancestry results only, with a European ancestry reference population from the Framingham Heart Study. Regional plots were created using Locus Zoom software[1] with LD information from the European ancestry 1000G reference population. **A-B**, Regional plots of the independent signals at chromosome 1q24; the novel *METTL11B* locus (**A**) and the replicated *PRRX1* locus (**B**). **C-D**, Regional plots of the independent signals at chromosome 10q24; the novel *SH3PXD2A* locus (**C**) and the replicated *NEURL1* locus (**D**).

Figure S10. AF associated loci display pleiotropy across clinical, electrocardiographic, and echocardiographic cardiac phenotypes.

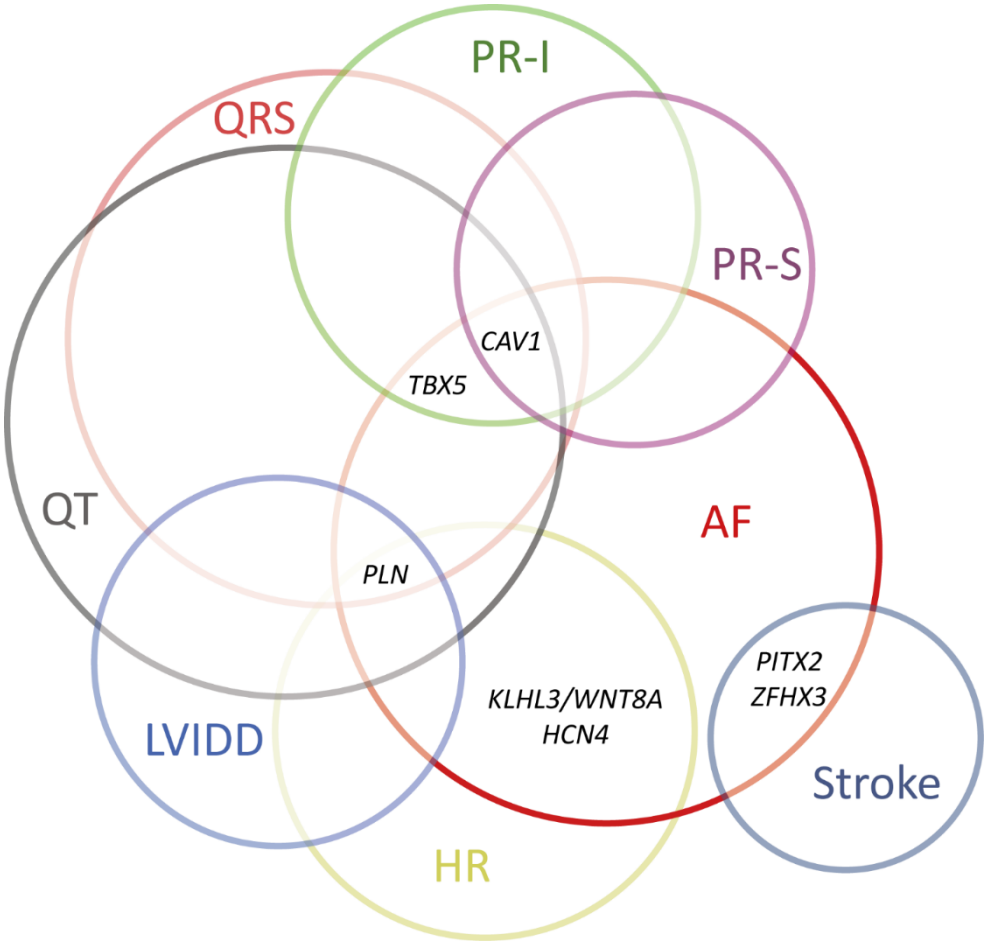
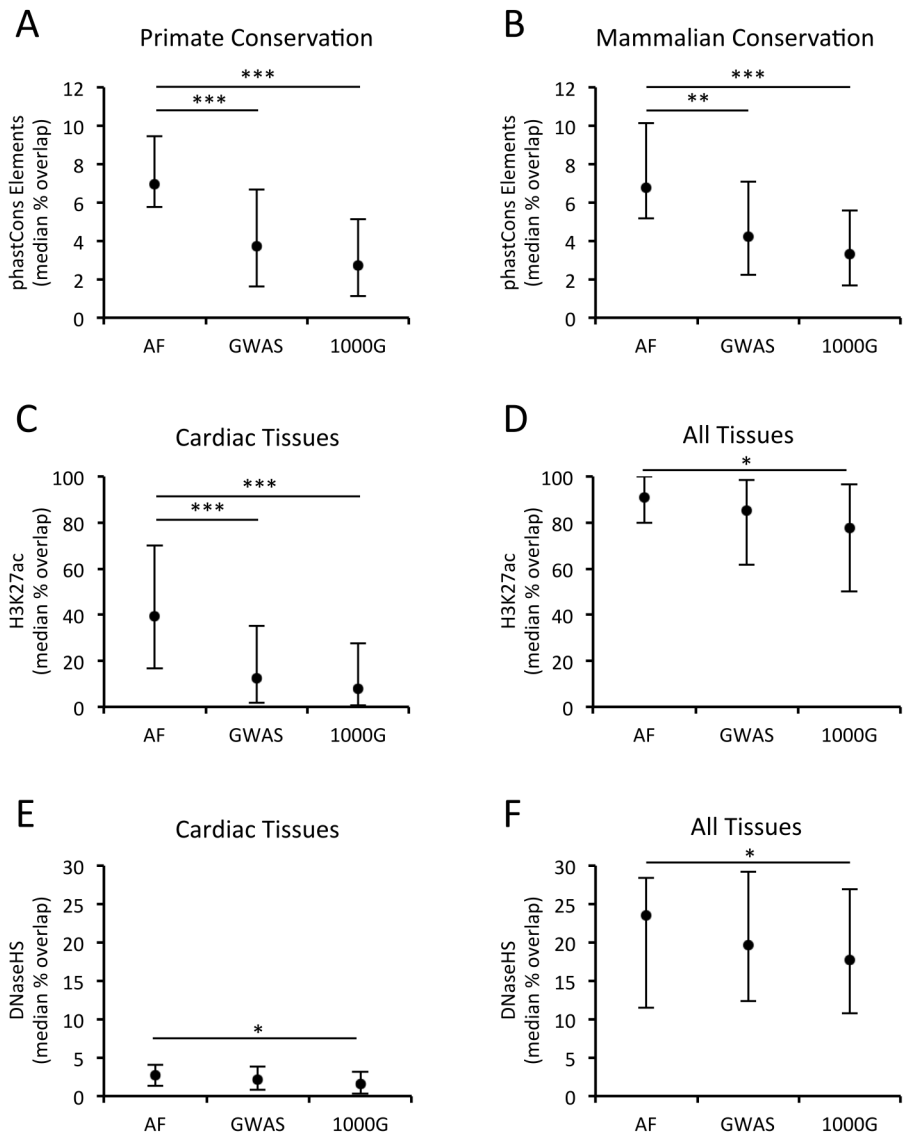


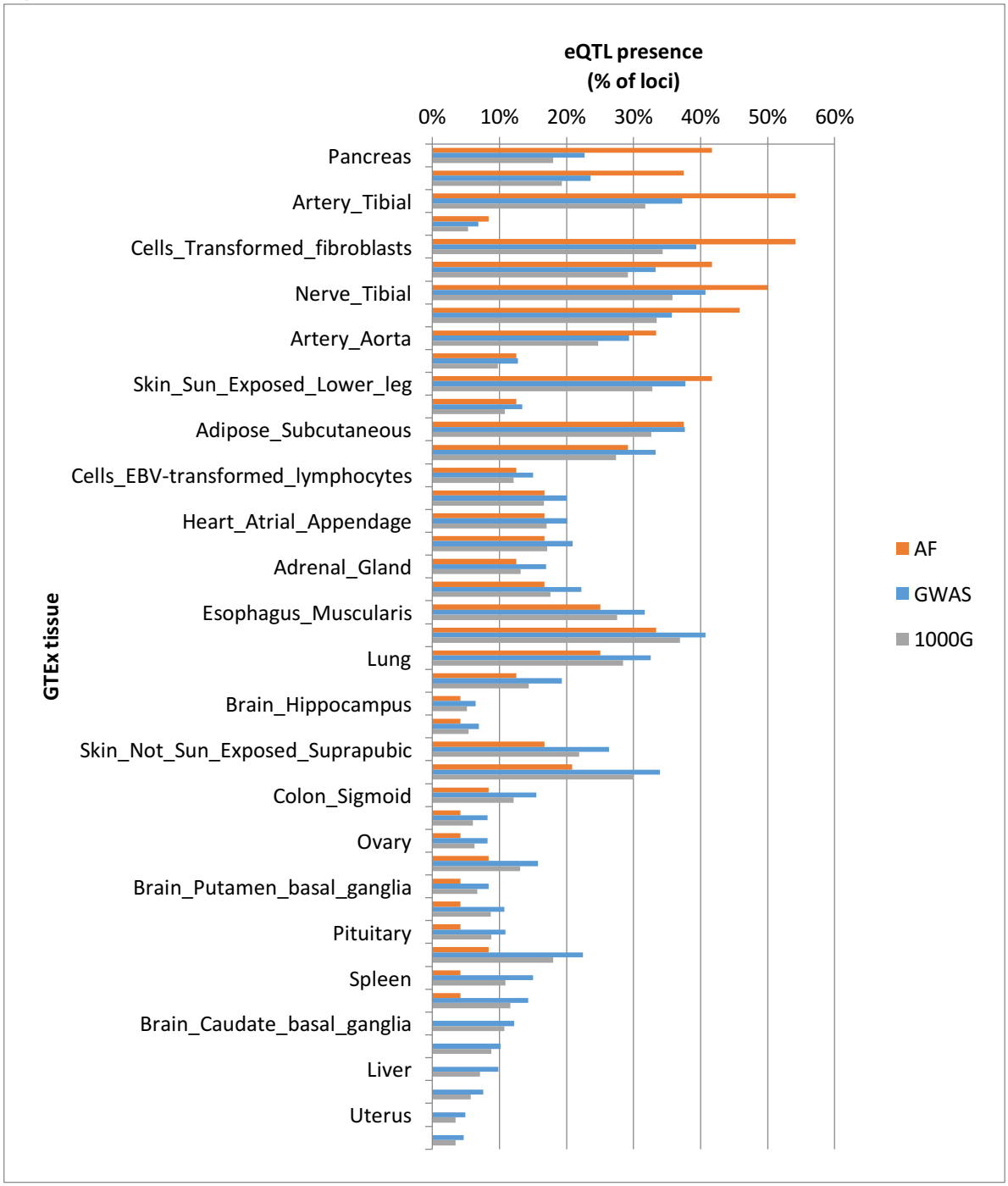
Diagram showing overlap of genetic associations between cardiac phenotypes, identified through interrogation of the NHGRI-EBI GWAS catalog.[2] Gene names are in italic and represent genetic loci identified through GWAS (**Supplementary Table S13**). AF, atrial fibrillation; HR, heart rate; LVIDD, left ventricle internal diastolic diameter on echocardiography; PR-I, PR-interval; PR-S, PR-segment; QRS, QRS-interval; QT, QT-interval.

Figure S11. AF associated loci are enriched for functional elements.



Median overlap of loci by (A) phastCons 46-way primate conserved elements, (B) phastCons 46-way mammalian conserved elements, (C) Roadmap Epigenome cardiac H3K27ac gapped peaks (R atrium, L ventricle, R ventricle, aorta), (D) Any Roadmap Epigenome H3K27ac gapped peak (98 cell types), (E) ENCODE DNaseHS cardiac sites (cardiac fibroblasts, atrial fibroblasts, cardiac myocytes), and (F) ENCODE DNaseHS master sites (125 cell types). *p-value < 0.05. **p-value < 0.01. ***p-value < 0.001 by one-tailed bootstrapping (n=1000). Whiskers, interquartile range; AF, atrial fibrillation associated loci (n=24); GWAS, NHGRI-EBI GWAS catalog associated loci (n=3381); 1000G, 1000 Genomes control loci based on SNPsnap matched variants to AF GWAS hits (n=9093).

Figure S12. AF associated loci are enriched for eQTLs.



The figure shows the proportion of loci with at least one GTEx eQTL for all tissues available in the GTEx database. Only the three top tissues showed significant enrichment of eQTLs for AF loci (pancreas, left ventricle, and tibial artery tissue) AF, atrial fibrillation associated loci (n=24); GWAS, NHGRI-EBI GWAS catalog associated loci (n=3381); 1000G, 1000 Genomes control loci based on SNPsnap matched variants to AF GWAS hits (n=9093).

2. SUPPLEMENTARY TABLES

Table S1. Baseline characteristics GWAS meta-analysis

Enclosed electronic excel file

Table S2. Baseline characteristics for the exome chip meta-analysis

Enclosed electronic excel file

Table S3. Detailed description of the genes at novel AF loci

Chromosomal location, Sentinel variant, *Gene(s)*: Description of the genes at the locus.
GWAS loci
1q24, rs72700118, *METTL11B/KIFAP3*: The most significant variant at 1q24 lies downstream of the closest gene, *METTL11B*, which encodes an N-terminal mono-methyltransferase that regulates DNA-protein interactions.[3] It is an important cell cycle regulator and mediator of DNA repair mechanisms since *METTL11B* knockout mice either die shortly after birth or display various developmental defects.[4] Interestingly, it also has been shown that METTL11B might act as a tumor suppressor protein in breast cancer.[5] *METTL11B* is highly expressed in right atrial and left ventricular tissue in GTEx. Analyses revealed that METTL11B may potentially interact with the atrial specific myosin light chain gene (MYL4) that has been linked to atrial fibrillation.[6,7] The locus also includes the gene *KIFAP3,* for which there also was a significant eQTL in the CCAF human atrial samples (**Supplementary Table S17 and S19**). *KIFAP3* encodes the kinesin associated protein 3, which regulates small G proteins by stimulating GDP/GTP exchange reactions or inhibiting their membrane interactions.[8] The gene is expressed in right atrial and left ventricular human tissue samples in the GTEx database. It is thought that this protein serves as a linker between human chromosome-associated polypeptide (HCAP) and KIF3A/B, a kinesin superfamily protein in the nucleus, and that this motor complex mediates binding to motor proteins enabling mainly anterograde transport of vesicles along microtubules.[9,10] *KIFAP3* variants have previously been associated with increased survival in sporadic amyotrophic lateral sclerosis and a combined phenotype of obesity and endometriosis in GWAS.[11,12] Reduced expression of *KIFAP3* has been demonstrated in clear cell renal carcinomas and was correlated with tumor aggressiveness and poorer patient outcomes,[13] whereas overexpression of the gene has been shown in breast cancer tumors.[14] In addition, *KIFAP3* has been shown to be involved in control of female puberty onset.[15] No relation to cardiac phenotypes have been noted for *KIFAP3* so far.
2p13, rs3771537, ***ANXA4/GMCL1:*** At 2p13, the most significant variant was intronic to *ANXA4*, whereas there were significant eQTLs for *ANXA4, GMCL1, PCYOX1,* and *SNRNP27* in GTEx left ventricle and skeletal muscle tissue (**Supplementary Table S17-S18**). *ANXA4* encodes Annexin 4, which is a Ca^{2+} and phospholipid binding protein that modulates membrane permeability, growth, apoptosis.[16] It has been demonstrated to be overexpressed in various cancers like lung cancer, colorectal cancer or prostate cancer where it enhances tumor invasion and chemotherapy resistance.[17] It has further been shown that *ANXA4* is involved in β-adrenergic signaling since $Anxa4^{-/-}$ mice show increased cellular cAMP levels and enhanced left ventricle contraction force upon adrenergic stimulation, whereas calcium stimulation in the left atrium lead to increased contraction force relative to wildtype mice.[18] Moreover, annexin 4 has been shown to bind to adenylyl cyclase type 5; thus, it has been suggested that annexin 4 directly modulates the β-adrenoceptor cAMP-dependent signal transduction pathway by inhibiting adenylyl cyclase 5.[18] In line with this hypothesis, *ANXA4* has been shown to be upregulated in human failing hearts.[19]

GMCL1, which encodes Germ Cell-Less protein 1, is predominantly expressed in the testis, where it is involved in spermatogenesis.[20,21] It has been demonstrated to regulate chromatin in germ cells by interacting with GAGE12I[22] and might also have a role in oncogenesis since it is expressed in various cancers like B cell lymphoma.[23] A direct link to cardiac physiology or disease; however, is currently missing.

2p14, rs2540949, *CEP68*: The most significant variant at 2p14 was intronic to *CEP68*, which encodes the centrosomal protein 68 that is important for the cell cycle by regulating centrosome cohesion.[24] There were significant eQTLs for *CEP68* in both the CCAF human atrial samples (**Supplementary Table S17 and 19**) and GTEx atrial, left ventricle, and skeletal muscle tissue (**Supplementary Table S17-S18**). At the onset of mitosis CEP68 dissociates from the centrosomes allowing the centrosomes to separate.[25] Variants in *CEP68* has been associated with aspirin-induced asthma[26] and acute urticaria/angioedema induced by non-steroidal anti-inflammatory drugs.[27]

2p31, rs2288327, *TTN/TTN-AS1:* At 2q31 we identified six significant coding variants in the A-band and M-line of titin, which all were predicted to be benign by PolyPhen and SIFT. The *TTN* gene spans 363 exons and the encoded protein stretches through half the length of a sarcomere.[28] Titin ensures sarcomere integrity and elasticity, and binds actin and myosin, which are crucial players in the contractile machinery in striated muscle.[29,30] Truncating mutations in titin have been shown to be the most important cause of dilated cardiomyopathy;[31–35] however, the gene displays considerable variation, making interpretation of mutational findings challenging.[36] Titin has been associated with the QT-interval in previous GWAS,[37,38] but the lead variant in our study (rs2288327) was not in LD with the QT-associated *TTN* variant (rs7561149, r^2=0.004).

5q22, rs337711, *KCNN2:* The variant at 5q22 is located in an intron of the gene *KCNN2* that encodes the small-conductance calcium-activated potassium channel, subfamily N, member 2 or SK2 channel. There was a significant eQTL for *KCNN2* itself in the CCAF human atrial tissue samples (**Supplementary Table S19**). This ion channel is predominantly expressed in the atria[39] and is involved in electrical remodeling resulting in atrial fibrillation.[39,40] In chronic AF, SK2 expression is reduced leading to significant changes in action potential duration (APD), a finding that has been confirmed in knockout mice. Furthermore, SK2 channels have been demonstrated to be involved in ventricular repolarization and also development of ventricular arrhythmias, especially in failing hearts where SK2 channels are upregulated both in patients and animal models.[41–45] Functional analysis revealed that the activation and modulation of SK2 channels is dependent on Ryr2-mediated calcium release[46] and that amiodarone can inhibit SK2 channels in a time- and voltage-independent but calcium-dependent mechanism, partly explaining its antiarrhythmic effects in failing hearts.[47] Additionally, genome-wide association studies have identified *KCNN2* as a susceptibility gene for coronary aneurysms in Kawasaki disease.[48,49] SK2 channels have also been shown to be involved in ischemia-induced neuronal cell death,[50,51] neuronal plasticity and learning,[52–55] drug addiction,[56,57] regulation of sleep duration,[58] and maintenance of the ionic milieu of the inner ear fluid.[59] They may be therapeutic targets for Parkinson's disease, since activation of SK2 channels provides protective effects in human dopaminergic neurons.[60]

5q31, rs2967791, *PKD2L2/KLHL3/WNT8A/FAM13B*: *PKD2L2* encodes the polycystic kidney disease 2-like 2 protein that belongs to the transient receptor potential (TRP) superfamily and is highly expressed

in human brain, kidney, and testis.[61,62] In rodents, it is also expressed in the heart and has been demonstrated to be involved in calcium homeostasis, proliferation, and apoptosis.[61–63]

KLHL3 encodes the gene Kelch Like Family Member 3 that is part of the E3 ubiquitin ligase complex regulating the sodium/chloride cotransporter (NCC), the epithelial sodium channel (ENaC), and the renal outer medullary potassium channel (ROMC) in the kidney.[64,65] It is an important regulator of the electrolyte homeostasis and therefore the blood pressure.[66,67] Genetic variants of *KLHL3* have been described to cause familial hyperkaliemic hypertension. [65,68,69]

WNT8A is a member of the WNT/beta catenin-signaling network that plays an essential role in development and carcinogenesis.[70] *WNT8A* has been demonstrated to regulate body axis extension[71] and neuroectodermal posteriorization.[72] *WNT8A* polymorphisms have been shown to be associated with Hirschsprung's disease and its expression is upregulated in stenotic colon segments in patients.[73] Interestingly, *in vitro* overexpression of WNT8 results in impaired calcium handling[74] and might therefore also be involved in AF pathophysiology.

For the 5q31 locus, we identified an eQTL for the gene *FAM13B* in eQTL enrichment analysis (**Supplemental table S17**). *FAM13B* (syn. *C5ORF5*) consists of 23 exons spanning over 27 kb; the transcript is 5.47 kb and encodes a protein of 915 amino acids.[75] It contains a putative rhoGAP domain at the N-terminus and two bipartite nuclear localization signals and is predominantly expressed in brain and male reproductive tissue[76] (Human Protein Atlas available from www.proteinatlas.org). So far, *FAM13B* has not been reported in a cardiovascular context.

8p22, rs7508, *ASAH1/PCM1*: At 8p22, the lead AF risk variant was associated with decreased expression of *ASAH1* (rs7508; $P = 5.1x10^{-3}$) in CCAF human atrial samples and increased expression of *PCM1* (rs7508; $P = 9.6x10^{-14}$) in both the CCAF samples (**Supplementary Table S17 and S19**) and GTEx left ventricle and skeletal muscle tissue (**Supplementary Table S17-S18**). *ASAH1* encodes the acid ceramidase 1 that is involved in lipid metabolism by degradation of ceramide into sphingosine and free fatty acids within lysosomes.[77,78] Overexpression of ceramidase has been reported in several cancer cell types,[79–81] resulting in increased proliferation[82] and invasiveness,[83,84] predominantly in prostate cancer, which in turn has led to studies showing promising results of ceramidase inhibitors as new cancer therapeutics.[85,86] Ceramidase has also been implicated in Farber's disease (lipogranulomatosis),[87,88] spinal muscular atrophy with myoclonic epilepsy,[89] and Alzheimer's disease.[90] *ASAH1* is highly expressed in the heart.[91] Accumulation of ceramide has been shown to result in oxidative stress, electron transport chain dysfunction, and cardiomyocyte apoptosis in rats.[92,93]

PCM1 encoding pericentriolar material 1, has been demonstrated to be an integral component of centriolar satellites in ciliogenesis.[94] It has also been shown to be involved in neurogenesis ,[95] the centrosomal actin network,[96] hematological neoplasms,[97] and associated with schizophrenia.[98]

10q24, rs35176054, *SH3PXD2A*: The variant at 10q24 is located intronic to the gene *SH3PXD2A* that encodes the SH3 and PX domain-containing protein 2A or Adapter protein TKS5 that plays an essential role in various malignancies. It interacts with Src tyrosine kinase to promote tumor growth and the formation of invadopodia resulting in degradation of extracellular matrix and invasion of cancer cells

into surrounding tissue in breast, ovarian, colon, lung, prostate cancer, melanoma, and glioma.[99–103] Its expression level has been demonstrated to be negatively correlated with tumor size and patient survival in ovarian cancer.[104,105] However, it is also involved in normal embryonic development by regulating neural crest migration[106,107] and in macrophage or microglia physiology.[108,109]

11q24, rs75190942, *KCNJ5:* The genetic variant rs75190942 is located at 11q24 within the gene *KCNJ5*, that encodes the G protein-activated inward rectifier potassium channel 4 (Kir3.4/GIRK4). There was a significant eQTL for *KCNJ5* itself in CCAF human atrial tissue samples (**Supplementary Table S17 and S19**) and in GTEx left ventricle tissue (**Supplementary Table S17-S18**). GIRK4 is known to form heteromeres with Kir3.1/GIRK1/KCNJ3, constituting the I_{KAch} channel complex, which contributes to the regulation of the membrane potential in the sinoatrial node and atria – making it a therapeutic target for AF. This ion channel has been shown to regulate pacemaker activity and recovery of resting heart rate after sympathetic stimulation.[110] GIRK4 inactivation can also rescue arrhythmias that are induced by genetic silencing of funny currents.[111] Furthermore, it determines inducibility, dynamics and termination of atrial fibrillation by regulating action potential duration.[112] Additionally, genetic polymorphisms in *KCNJ5* are associated with early-onset lone atrial fibrillation,[113] whereas mutations in this gene have been shown to cause long QT syndrome.[114] GIRK4 is also expressed in the ventricles and contributes to ventricular repolarization[115] and has been shown to be significantly downregulated in patients with dilated cardiomyopathy.[116] Furthermore, mutations in *KCNJ5* can cause Andersen-Tawil syndrome,[117] primary aldosteronism[118] and has been detected in adrenal tumors.[119] Also, *KCNJ5* is associated with Tourette Syndrome and Attention-Deficit/Hyperactivity Disorder.[120]

ExWAS loci

3p22, rs6800541, *SCN10A:* The variant rs6800541 is located intronic to *SCN10A*, the gene that encodes the sodium channel Nav1.8. It is highly expressed in primary sensory neurons and dorsal root ganglion neurons and has been linked to nociception, painful neuropathy, and multiple sclerosis.[121] Recently, it has been shown that Nav1.8 is also expressed in the heart where it contributes to the late sodium current.[122,123] Genome-wide association studies demonstrated genetic variants in *SCN10A* as risk loci for quantitative ECG traits like PR interval,[124–128] and QRS duration,[126,129,130] as well as for AF[124,126,130,131] and Brugada Syndrome.[132] Also, mutations in *SCN10A* has been shown to be responsible for a large fraction of cases of Brugada Syndrome.[133] Data suggest that *SCN10A* affects cardiac conduction either directly through cardiomyocytes, indirectly through intracardiac neurons, or by modulation of *SCN5A* expression.[134,135]

12p12, rs11047543, *SOX5:* The most significant SNP at 12p12 is located downstream of the *SOX5* gene. SOX5 is a transcription factor that has been shown to be involved in limb development,[136] chondrogenesis,[137] brain development,[138] and lung development.[139] Our current study confirmed previous genome-wide association studies that showed a significant association between *SOX5* and early-onset AF.[124,140] Furthermore, *SOX5* has been demonstrated to be significantly associated with PR-interval,[124] left ventricular mass,[141] resting heart rate,[142] osteoporosis,[143] systemic sclerosis,[144] AIDS,[145] chronic obstructive pulmonary disease,[139] and non-obstructive azoospermia.[146] Additionally, it is involved in the development of lung cancer,[147] hepatocellular carcinoma,[148] follicular lymphoma,[149] and melanoma.[163]

Locus identified in both GWAS and EWAS:
6q22, rs4946333 (GWAS), rs89107 (EWAS), *SLC35F1/PLN:*** At 6q22 we identified a locus including the phospholamban gene (*PLN*), *SLC35F1,* and *CEP85L*. Phospholamban regulates cardiac contractility and relaxation through inhibiting the cardiac muscle sarcoplasmic reticulum calcium ATPase SERCA.[164] Mutations in this gene has been associated with hypertrophic[165,166] and dilated cardiomyopathy.[167,168] *SLC35F1* encodes a member of the solute carrier family 35. *SLC35F1* knockout mice display reduced levels of hemoglobin and lactate dehydrogenase but do not show any further phenotype. Previous GWAS have associated the locus surrounding *SLC35F1/PLN/CEP85L* with resting heart rate,[6,15] QT-interval,[12,14] and left ventricle internal diastolic diameter.[11] One of the variants associated with heart rate by den Hoed et al. also associated with atrial fibrillation in secondary analyses.[6]

Table S4. Results from Asian ancestry SKAT gene based test

Gene	Chr	CMAF	N variants	P-value
Filter: Variants predicted to be damaging				
SH3PXD2A	10q24	0.4	6	4.77×10^{-11}
Filter: Nonsynonymous and splice site variants				
SH3PXD2A	10q24	0.4	11	4.21×10^{-11}

Chr, chromosome; CMAF, cumulative minor allele frequency per gene.

Table S5. Single variant association results for the variants that were analyzed in the two significant gene-based tests for *SH3PDX2A* in the Asian ancestry group.

rsID	Risk/ref allele	Amino acid substitution	RAF, %	OR	95% CI	P-value
rs149867987	A/G	His110Tyr	0.01	16.72	2.23-125.31	0.006
rs200938753*	G/A	Arg761Cys	99.89	1.45	0.74-2.84	0.27
rs202011870*	C/A	Leu396Arg	0.18	4.68	2.97-7.39	3.30E-11
rs201065560*	A/G	Arg1031Cys	0.02	2.03	0.55-7.47	0.29
rs74661743*	G/A	Arg1003Cys	99.93	1.02	0.42-2.47	0.97
rs79061932	G/A	Arg994Cys	99.99	1.13	0.07-18.44	0.93
rs201439736	C/T	Ala886Thr	99.97	1.44	0.46-4.52	0.54
rs201054626*	T/C	Arg302Gln	0.01	4.85	0.83-28.47	0.08
rs143819462	T/C	Arg269Gln	0.01	2.34	0.39-13.93	0.35
rs147297499	T/C	Asp231Asn	0.005	13.31	0.67-264.24	0.09
rs143409187*	T/C	Arg102Gln	0.007	2.85	0.15-55.03	0.49

The gene-based test was significant for the subset of nonsynonymous and splice site variants, which included all listed variants, and the subset of nonsynonymous possibly damaging variants, which included 6 of the listed variants (*). RAF, risk allele frequency; CI, confidence interval; OR, odds ratio.

Table S6. Results from ancestry-specific GWAS meta-analyses

	rsID	Chr	Gene	Location relative to gene	Risk/ref allele	RAF, %	OR	95% CI	P-value
EUR	*15,993 cases, 113,719 referents*								
	Novel associations								
	rs10800507	1q24	*METTL11B/KIFAP3*	Intergenic	C/G	51	1.09	1.06-1.12	$1.87x10^{-11}$
	rs62133983	2p13	***ANXA4***/*GMCL1*	Intronic	G/T	52	1.09	1.06-1.12	1.36x10-10
	rs2723064	2p14	*CEP68*	Intergenic	T/C	61	1.09	1.06-1.12	1.88x10-10
	rs6864727	5q31	***PKD2L2***/*WNT8A/FAM13B*	Intronic	C/T	31	1.08	1.05-1.11	$1.12x10^{-8}$
	rs281868	6q22	***SLC35F1***/*PLN*	Intronic	G/A	50	1.08	1.05-1.10	$1.03x10^{-8}$
	rs7508	8p22	***ASAH1***/*PCM1*	3'UTR	A/G	73	1.10	1.06-1.13	$6.34x10^{-10}$
	rs35176054	10q24	***SH3PXD2A***	Intronic	A/T	13	1.14	1.10-1.18	$1.75x10^{-11}$
	rs75190942	11q24	***KCNJ5***	Intronic	A/C	8	1.18	1.11-1.25	$2.82x10^{-8}$
	rs2921421	15q21	*CGNL1*	Intergenic	G/C	3	1.72	1.42-2.09	$3.29x10^{-8}$
	Previously known associations								
	rs11264280	1q21	*KCNN3*	Intergenic	T/C	32	1.13	1.10-1.16	$2.77x10^{-17}$
	rs651386	1q24	*PRRX1*	Intergenic	A/T	57	1.11	1.08-1.14	$6.23x10^{-15}$
	rs2129977	4q25	*PITX2*	Intergenic	A/G	22	1.45	1.41-1.49	$7.25x10^{-136}$
	rs12664873	6q22	*GJA1*	Intergenic	T/G	69	1.08	1.05-1.12	$1.80x10^{-8}$
	rs11773845	7q31	***CAV1/2***	Intronic	A/C	60	1.10	1.07-1.13	$3.35x10^{-13}$
	rs7026071	9q22	***C9orf3***	Intronic	T/C	41	1.09	1.07-1.12	$2.86x10^{-11}$
	rs10824026	10q22	*SYNPO2L*	Intergenic	A/G	84	1.13	1.09-1.17	$8.29x10^{-11}$
	rs11598047	10q24	***NEURL1***	Intronic	G/A	17	1.18	1.14-1.22	$3.16x10^{-21}$
	rs883079	12q24	***TBX5***	3'UTR	T/C	72	1.11	1.08-1.15	$1.31x10^{-13}$
	rs7183206	15q24	*HCN4*	Intergenic	A/G	15	1.13	1.09-1.18	$7.70x10^{-12}$
	rs2106261	16q22	***ZFHX3***	Intronic	T/C	18	1.19	1.15-1.23	$4.01x10^{-24}$
AA	*641 cases, 4956 referents*								
	rs6843082	4q25	*PITX2*	Intergenic	G/A	30	1.40	1.24-1.58	$4.31x10^{-8}$
AS	*837 cases, 2456 referents*								
	Novel association								
	rs7138621	12q15	*CPSF6*	Intergenic	G/C	95	7.92	4.26-14.73	$6.48x10^{-11}$
	Previously known association								
	rs2723334	4q25	*PITX2*	Intergenic	T/C	70	1.94	1.68-2.25	$8.46x10^{-19}$

The most significant variant at each genetic locus associated with AF is listed. Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants, the closest gene(s) are listed. Chr, chromosome; CI, confidence interval; OR, odds ratio; EUR, European ancestry; AA, African American ancestry; AS, Asian ancestry; RAF, risk allele frequency.

Table S7. Results from European and Asian ancestry ExWAS meta-analysis

	rsID	Chr	Gene	Location relative to gene	Risk/ref allele	RAF, %	OR	95% CI	P-value
EUR	*13,496 cases, 96,273 referents*								
	Novel associations								
	rs6800541	3p22	***SCN10A***	Intronic	T/C	60	1.08	1.05-1.12	8.75×10^{-7}
	rs89107	6q22	***SLC35F1****/PLN*	Intronic	G/A	50	1.09	1.06-1.13	2.71×10^{-7}
	rs11047543	12p12	*SOX5*	Intergenic	G/A	85	1.13	1.08-1.18	4.65×10^{-7}
	Previously known associations								
	rs13376333	1q21	***KCNN3***	Intronic	T/C	31	1.14	1.10-1.17	1.58×10^{-13}
	rs2200733	4q25	*PITX2*	Intergenic	T/C	12	1.60	1.52-1.67	9.95×10^{-90}
	rs3807989	7q31	***CAV1***	Intronic	G/A	59	1.09	1.06-1.13	2.93×10^{-8}
	rs60632610	10q22	***SYNPO2L***	Exonic; nonsyn	C/T	85	1.13	1.08-1.18	2.53×10^{-8}
	rs2106261	16q22	***ZFHX3***	Intronic	A/G	17	1.21	1.16-1.26	3.37×10^{-18}
AS	*8180 cases, 28,612 referents*								
	Novel associations								
	rs55952639	2p14	***CEP68***	Exonic; syn	T/C	76	1.13	1.07-1.18	1.29×10^{-6}
	rs11047543	12p12	*SOX5*	Intergenic	G/A	88	1.18	1.10-1.26	1.16×10^{-6}
	Previously known associations								
	rs17042171	4q25	*PITX2*	Intergenic	A/C	48	1.69	1.62-1.76	4.04×10^{-137}

The most significant variant at each genetic locus associated with AF is listed. Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants the closest gene(s) are listed. Chr, chromosome; CI, confidence interval; OR, odds ratio; EUR, European ancestry; AA, African American ancestry; AS, Asian ancestry; nonsyn, nonsynonymous; syn, synonymous; RAF, risk allele frequency

Table S8. Results from European incident AF GWAS meta-analysis

rsID	Chr	Gene	Location relative to gene	Risk/ref allele	RAF, %	OR	95% CI	P-value
rs11264280	1q21	*KCNN3*	Intergenic	T/C	32	1.12	1.08-1.16	$3.57x10^{-9}$
rs6843082	4q25	*PITX2*	Intergenic	G/A	21	1.38	1.33-1.44	$8.21x10^{-57}$
rs7394190	10q22	*SYNPO2L*	Intergenic	G/A	84	1.15	1.09-1.21	$3.09x10^{-8}$
rs60848348	10q24	***NEURL1***	Intronic	T/C	20	1.13	1.09-1.18	$1.69x10^{-8}$
rs4499262	16q22	***ZFHX3***	Intronic	A/C	17	1.14	1.09-1.19	$4.01x10^{-8}$

The most significant variant at each genetic locus associated with AF is listed. Gene names in bold font indicate that the variant is located within the gene. Chr, chromosome; CI, confidence interval; OR, odds ratio; RAF, risk allele frequency

Table S9. Results from European prevalent AF GWAS meta-analysis

rsID	Chr	Gene	Location relative to gene	Risk/ref allele	RAF, %	OR	95% CI	P-value
Novel associations								
rs72700118	1q24	*METTL11B/KIFAP3*	Intergenic	A/C	11	1.24	1.17-1.31	$9.93x10^{-13}$
rs6546550	2p13	***ANXA4****/GMCL1*	Intronic	C/G	54	1.12	1.08-1.16	$1.36x10^{-8}$
rs1454934	12p11	***PKP2***	Intronic	T/C	16	1.16	1.1-1.22	$4.18x10^{-8}$
Previously known associations								
rs36004974	1q21	***KCNN3***	Intronic	G/A	32	1.14	1.1-1.19	$4.36x10^{-10}$
rs577676	1q24	*PRRX1*	Intergenic	C/T	55	1.15	1.1-1.19	$2.77x10^{-12}$
rs61303432	4q25	*PITX2*	Intergenic	T/C	14	1.71	1.62-1.8	$6.66x10^{-92}$
rs2109514	7q31	*CAV1/2*	Intergenic	A/G	51	1.15	1.11-1.19	$6.73x10^{-13}$
rs11598047	10q24	***NEURL1***	Intronic	G/A	17	1.24	1.18-1.31	$4.34x10^{-16}$
rs2106261	16q22	***ZFHX3***	Intronic	T/C	18	1.25	1.19-1.31	$9.68x10^{-20}$

The most significant variant at each genetic locus associated with AF is listed. Gene names in bold font indicate that the variant is located within the gene. Chr, chromosome; CI, confidence interval; OR, odds ratio; RAF, risk allele frequency

Table S10. Comparison of results for common variant loci between the AFGen Consortium combined ancestry analysis and the Biobank Japan study.

Enclosed electronic excel file

Table S11. Replication of the common variant loci identified in the AFGen Consortium combined ancestry analysis and the UK Biobank study.

Enclosed electronic excel file

Table S12. Approximate and joint conditional analysis in European ancestry GWAS meta-analysis identify 20 independent genetic loci associated with AF

rsID	Chr	Gene	Location relative to gene	P-value
rs11264280	1	*KCNN3*	Intergenic	2.77×10^{-17}
rs10800507	1	*METTL11B*	Intergenic	1.87×10^{-11}
rs651386	1	*PRRX1*	Intergenic	6.23×10^{-15}
rs2723065	2	*CEP68*	Intergenic	1.91×10^{-10}
rs62133983	2	***ANXA4***	Intronic	1.36×10^{-10}
rs2129977*	4	*PITX2*	Intergenic	7.25×10^{-136}
rs6864727	5	***PKD2L2***	Intronic	1.12×10^{-8}
rs281868	6	***SLC35F1***	Intronic	1.03×10^{-8}
rs7773091	6	*GJA1*	Intergenic	2.02×10^{-8}
rs11773845	7	***CAV1***	Intronic	3.35×10^{-13}
rs7508	8	***ASAH1***	3'UTR	6.34×10^{-10}
rs7026071	9	***C9orf3***	Intronic	2.86×10^{-11}
rs11598047	10	***NEURL1***	Intronic	3.16×10^{-21}
rs35176054	10	***SH3PXD2A***	Intronic	1.75×10^{-11}
rs10824026	10	*SYNPO2L*	Intergenic	8.29×10^{-11}
rs75190942	11	***KCNJ5***	Intronic	2.82×10^{-8}
rs883079	12	***TBX5***	3'UTR	1.31×10^{-13}
rs2921421	15	*CGNL1*	Intergenic	3.29×10^{-8}
rs8040533	15	*HCN4*	Intergenic	3.09×10^{-11}
rs2106261	16	***ZFHX3***	Intronic	4.01×10^{-24}

Chr, chromosome; UTR, untranslated region. Bold font indicates that the variant lies within the gene. *The 4q25/*PITX2* region was not analyzed because the complexity of this association signal is not accurately evaluated with the GCTA method (**Online Methods**).

Table S13. Overlap with AF risk factor GWAS loci

rsID	Chr	Closest gene*	rsID GWAS Catalog	LD	GWAS P-Value	HR	PR-S	PR-I	QRS	QT	Echo LVIDD	Stroke
ALL ANCESTRIES												
rs6843082	4	*PITX2(dist=154788);C4orf32(dist=1348486)*	rs6843082	1	$3.41x10^{-155}$							[3]
rs6843082	4	*PITX2(dist=154788);C4orf32(dist=1348486)*	rs12646447	0.51	$1.12x10^{-148}$							[4]
rs6843082	4	*PITX2(dist=154788);C4orf32(dist=1348486)*	rs2200733	0.51	$2.32x10^{-150}$							[5]
rs2967791	5	*KLHL3/WNT8A*	rs7722600	0.15	$1.25x10^{-6}$	[6]						
rs4946333	6	*SLC35F1/PLN*	rs457162	<0.10	0.0686					[7]		
rs4946333	6	*SLC35F1/PLN*	rs11752626	0.43	0.0001					[8]		
rs4946333	6	*SLC35F1/PLN*	rs11970286	0.48	$3.29x10^{-5}$					[9,10]		
rs4946333	6	*SLC35F1/PLN*	rs12210810	<0.10	0.001					[10]		
rs4946333	6	*SLC35F1/PLN*	rs12210733	<0.10	0.001					[7]		
rs4946333	6	*SLC35F1/PLN*	rs89107	0.99	$4.03x10^{-9}$						[11]	
rs4946333	6	*SLC35F1/PLN*	rs3902035	<0.10	0.002					[7]		
rs4946333	6	*SLC35F1/PLN*	rs11756438	0.29	0.0008					[12]		
rs4946333	6	*SLC35F1/PLN*	rs6906287	0.38	$5.84x10^{-5}$				[13]			
rs4946333	6	*SLC35F1/PLN*	rs11153730	0.45	$2.01x10^{-5}$	[6]				[7,14]		
rs4946333	6	*SLC35F1/PLN*	rs281868	1	$2.12x10^{-9}$	[15]						
rs1997572	7	*CAV1*	rs3807989	0.93	$1.47x10^{-14}$		[16]	[9,17,18]	[9]			
rs1997572	7	*CAV1*	rs11773845	0.94	$7.53x10^{-15}$			[19,20]				
rs1997572	7	*CAV1*	rs9920	0.15	0.0005					[7]		
rs883079	12	*TBX5*	rs883079	1	$1.80x10^{-15}$				[21]			
rs883079	12	*TBX5*	rs7312625	0.90	$1.03x10^{-14}$			[22]				
rs883079	12	*TBX5*	rs1895585	0.83	$1.25x10^{-14}$			[19]				
rs883079	12	*TBX5*	rs7135659	0.88	$9.59x10^{-15}$			[20]				
rs883079	12	*TBX5*	rs3825214	0.59	$1.82x10^{-10}$			[9]	[9]	[9]		
rs74022964	15	*HCN4(dist=15659);C15orf60(dist=58235)*	rs4489968	0.77	$4.59x10^{-11}$	[6]						
rs2106261	16	*ZFHX3*	rs879324	0.91	$1.13x10^{-25}$							[3]

Table showing overlap of genetic associations between cardiac phenotypes, identified through interrogation of the NHGRI-EBI GWAS catalog.[2]
Numbers in superscript in the phenotype columns indicate references to the literature. Chr, chromosome; LD, linkage disequilibrium r^2 with lead SNP; HR, heart rate; PR-S, PR-segment; PR-I, PR-interval; LVIDD, Left Ventricle Internal Diastolic Diameter; AF, atrial fibrillation; AFL, atrial flutter.
*For intronic variants, the gene the variant is located within is listed; for intergenic variants, the closest genes upstream and downstream are listed.

Table S14. Association between novel AF loci and stroke subtypes in the Neuro-CHARGE Stroke Consortium

			All stroke		Ischemic stroke		Cardioembolic stroke	
rsID	**Gene***	**Risk/ref allele**	**OR**	**P-value**	**OR**	**P-value**	**OR**	**P-value**
rs72700118	*METTL11B*	A/C	1.01	0.70	1.02	0.61	1.09	0.38
rs3771537	***ANXA4****/GMCL1*	A/C	1.00	0.85	0.99	0.75	0.99	0.88
rs2540949	***CEP68***	A/T	1.04	0.12	1.05	0.09	1.14	0.02
rs2288327	***TTN/TTN-AS1***	G/A	1.05	0.08	1.08	0.02	1.22	0.01
rs337711	***KCNN2***	T/C	0.97	0.16	0.96	0.18	0.97	0.63
rs2967791	***KLHL3****/WNT8A/FAM13B*	T/C	1.03	0.14	1.04	0.10	1.11	0.05
rs4946333	***SLC35F1****/PLN*	G/A	0.97	0.21	0.97	0.18	0.97	0.58
rs7508	***ASAH1***	A/G	1.04	0.12	1.04	0.17	1.11	0.14
rs35176054	***SH3PXD2A***	A/T	1.03	0.38	1.01	0.77	1.07	0.44
rs75190942	***KCNJ5***	A/C	1.01	0.85	1.04	0.45	-	-

OR, odds ratio. *Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants, the closest gene(s) are listed.

Table S15. Association between novel AF loci and stroke subtypes in the Metastroke Consortium

			Ischemic stroke		Cardioembolic stroke		Large vessel disease		Small vessel disease	
rsID	**Gene***	**Risk/ref allele**	**OR**	**P-value**	**OR**	**P-value**	**OR**	**P-value**	**OR**	**P-value**
rs72700118	*METTL11B/ KIFAP3*	A/C	1.07	0.02	1.14	0.02	1.01	0.92	1.04	0.53
rs3771537	***ANXA4****/GMCL1*	A/C	0.99	0.52	1.02	0.57	0.94	0.08	1.00	0.95
rs2540949	***CEP68***	A/T	0.99	0.63	1.03	0.40	1.05	0.18	0.97	0.54
rs2288327	***TTN/TTN-AS1***	G/A	1.02	0.54	1.03	0.61	1.02	0.66	1.07	0.21
rs337711	***KCNN2***	T/C	1.01	0.50	1.08	0.04	1.00	0.90	0.94	0.19
rs2967791	***KLHL3****/WNT8A/FAM13B*	T/C	1.02	0.39	1.05	0.19	1.06	0.15	0.92	0.05
rs4946333	***SLC35F1****/PLN*	G/A	0.98	0.26	0.91	0.01	0.89	0.003	1.01	0.79
rs7508	***ASAH1***	A/G	0.98	0.37	1.00	1.00	1.03	0.45	0.94	0.17
rs35176054	***SH3PXD2A***	A/T	1.01	0.67	1.07	0.25	0.96	0.46	1.10	0.13
rs75190942	***KCNJ5***	A/C	1.02	0.59	1.09	0.31	1.03	0.73	0.98	0.80

OR, odds ratio. *Gene names in bold font indicate that the variant is located within the gene, whereas additional gene names indicate eQTL gene or gene strongly suspected to be causal due to the function of the encoded protein. For intergenic variants, the closest gene(s) are listed.

Table S16. GO Terms Enriched in Atrial Fibrillation Associated Loci Compared to GWAS Catalog Loci and to 1000 Genomes Matched Loci

Gene Ontology Description	P-value	FDR Q-value
Compared to 1000 Genomes Matched Loci		
Small conductance calcium-activated potassium channel activity	9.48×10^{-5}	3.01×10^{-1}
Metal ion transport	1.62×10^{-4}	1.00
Potassium channel activity	2.52×10^{-4}	4.00×10^{-1}
Z disc	2.70×10^{-4}	3.85×10^{-1}
Monovalent inorganic cation transport	3.52×10^{-4}	1.00
Potassium ion transmembrane transport	5.08×10^{-4}	1.00
Cellular potassium ion transport	5.08×10^{-4}	1.00
Potassium ion transmembrane transporter activity	5.70×10^{-4}	6.04×10^{-1}
Regulation of cardiac muscle contraction	6.92×10^{-4}	1.00
Striated muscle tissue development	6.92×10^{-4}	1.00
Potassium ion transport	7.08×10^{-4}	1.00
Cation transport	7.34×10^{-4}	1.00
Regulation of heart rate	9.10×10^{-4}	1.00
Compared to GWAS catalog Loci		
Small conductance calcium-activated potassium channel activity	2.64×10^{-4}	7.43×10^{-1}
Z disc	2.67×10^{-4}	3.34×10^{-1}
Metal ion transport	3.17×10^{-4}	1.00
Potassium channel activity	4.14×10^{-4}	5.83×10^{-1}
Monovalent inorganic cation transport	7.01×10^{-4}	1.00

Table S17. Summary of top eQTLS within atrial fibrillation associated loci

Enclosed electronic excel file

Table S18. *In silico* eQTL analysis in GTEx database

Enclosed electronic excel file

Table S19. eQTL analysis of in CCAF human atrial tissue samples

Enclosed electronic excel file

Table S20. *In silico* functional evaluation of novel and replicated loci from GWAS and ExWAS combined ancestry analysis

Enclosed electronic excel file

Table S21. Per study overlap of samples between GWAS and exome chip analyses

Study	Overlap	
	Cases	**Controls**
BioVU	206	3811
WGHS	934	20,266
FHS - incident	411	1612
FHS - prevalent	181	2123
CHS - incident	922	1979
CHS -prevalent	60	2900
AGES	354	2989
RS	346	2370
CAMP	665	2128
SHIP	99	2710
AFLMU/KORA	349	415
MGH	333	0
ARIC EA	1253	3415
ARIC AA	233	742
MESA	155	2372
GS:SFHS	203	6651
BioMe EA	290	857
BioMe AA	166	2041
BioMe HA	255	2800
BEAT-AF	1520	1516
BBJ	782	0
Total	9717	63,697

Table S22. GWAS information per study

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
AFLMU /KORA	[169]	Illumina HumanCNV370 + Illumina Human550K	BeadStudio	≥98%	$<10^{-5}$	-	-	>1%	P<0.05	1	306,838	SHAPEIT v2.r790 + IMPUTE v.2.1.2	SNPTEST v2.5	7,540,650	1.023
AGES	[170]	Illumina HumanCNV370-Duo BeadChip	BeadStudio	≥97%	$<10^{-6}$	-	-	≥1%	P<0.05	0	329,804	MaCH v.1.0.16 + minimac	ProbABEL, R	I: 7,602,716 P: 6,085,662	I: 1.068 P: 1.006
ANGES	[171]	Illumina Metabochip	GenomeStudio	≥95%	$\geq 10^{-6}$	-	>3.18 SD from the mean removed	-	first 4 PCs	4	121,545	SHAPEIT v.2.r790 + IMPUTE2 v.2.3.0	SNPTEST v2.4.1	5,861,502	P&I: 1.011
ARIC	[172,173]	Affymetrix 6.0	Birdseed	≥95%	$<10^{-5}$	-	-	EA: >0.5% AA: >1%	Analysis committee recommendations	EA: 4 AA: 10	EA: 711,589 AA: 806,416	(1) Pre-phasing with ShapeIt (v1.r532) (2) Imputation with IMPUTE2.1.0	FAST	EA: 9,428,893 AA: 8,978,558	EA: 1.011 AA: 0.991
Beat-AF	[174]	Illumina HumanCoreExome	BeadStudio	≥95%	$>10^{-6}$	-	>3 SD from the mean removed	≥1%	First 10 PCs	10	254,488	SHAPEIT v2.r790 + IMPUTE v.2.3.2	SNPTEST v.2.5	9,309,201	1.022

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
BBJ	[175]	Illumina Human610 Quad and Illumina Human Hap550v3 BeadChip	Beadstudio	≥99%	$>10^{-6}$	-	-	≥1%	First 2 PCs	2	432,042	MaCH + minimac	PLINK v1.07	6,429,092	1.024
BioMe	[176]	Illumina HumanOmni ExpressExome-8 v1.0	zCall (GenomeStudio)	≥90%	$p>10^{-6}$	-	-	≥1%	first 4 PCs	4	768,517	IMPUTE2	SNPTEST v.2.5	EA: 7,022,478 AA: 8,200,353 HA: 8,139,248	EA: 1.008 AA: 1.019 HA: 1.026
BioVU	[177]	Illumina Omni5 + Omni1 + 1M + 660K	GenomeStudio	≥98%	$<10^{-5}$	-	-	≥1%	First 2 PCs	2	4,167,400	IMPUTE2 v2.3.0	PLINK v1.90	660: 3,187,278 omni: 4,373,169	660: 1.003 Omni: 1.01
CCAF	[169]	Hap550 v1&v3 chip + Hap610 v1 chip	BeadStudio	≥95%	FDR>10^{-4}	-	FDR>0.01	≥1%	P<0.05	4	516,461	Shapeit v2.r727 + IMPUTE v.2.3.0	SNPtest v.2.5	8,122,372	1.026
CHS - AA	[178]	HumanOmni1-Quad_v1	GenomeStudio	≥97%	$\geq 10^{-5}$	≤1 in CEPH trios	-	>0.01 %	PCs with P<0.05 and all PCs before the associated PC	3	963,248	IMPUTE version 2.2.2	R	8,152,032	1.001
CHS - EA	[178]	Illumina 370 CNV + ITMAT-Broad-CARe (IBC) Illumina iSELECT chip	BeadStudio	≥97%	$\geq 10^{-5}$	≤2 in CEPH trios	-	>0.01 %	PCs with P<0.05 and all PCs before the associated PC	0	359,592	MaCH + minimac	R	8,278,530	1.045

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
COROGENE	[179]	Illumina Metabochip + CoreExome	GenomeStudio	≥95%	≥10^{-5}	-	-	≥1%	-	0	553,581	IMPUTE v2.2.2	SNPTEST v2.4.1	6,956,681	1.019
FHS	[180,181]	Affymetrix, Gene Chip®, 500K Array Set & 50K Human Gene Focused Panel	BRLMM	≥97%	<10^{-6}		Subject hetero-zygosity >5 SD away from the mean	≥1%	All PCs u-ssociated, p>0.05	0	385,958	Mach1 v1.0.15	R packages kinship, GEE, COXPH	I: 7525764 P: 6556225	I: 1.019 P: 1.04
FINCAVAS	[182]	Illumina Metabochip + CoreExome	GenomeStudio	≥95%	≥10^{-6}	-	>3.23 SD from the mean removed	-	First 4 PCs	4	Metabochip: 120,689 CoreExome: 277,211	SHAPEIT v.2.r790 + IMPUTE2 v.2.3.0	SNPTEST v2.4.1	8,384,365	P&I: 1.04
GS:SFHS	[183]	Illumina Omni Express Plus Exome	BeadStudio	Omni ≥98% Exome ≥99%	<10^{-6}	-	-	Omni <1% Exome <0.01 %	PCs associated after adjustment for sex and age with p<0.05)	1	706,198 (690,759 Autosomes)	ShapeIt2 (pre-phasing), IMPUTE2 (imputation)	ProbABEL	6,563,971	0.997
HNR	[184]	Illumina: Omni Express, Omni1, CoreExomeA and CoreExomeB			<10^{-5}		Subject hetero-zygosity >5 SD away from the mean	MAF ≥0.01 and ≤99.9	First 10 PCs	10	Omni1: 682,618 OmniEx: 646,304 CoreExB: 255,584 CoreExA: 256,445	Impute v.2.3.0	SNPTEST	Excluded due to sample size	
LURIC	[185]	Affymetrix 6.0	Birdseed v.2	≥98%	0.0001	-	-	≥1%	First 3 PCs	3	686,195	IMPUTE v.2	SNPtest v.2.5	7,270,779	1.003

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
MDCS	[186]	Illumina Human Omni Express Exome 1.0	GenomeStudio	≥95%	0.0001	-	-	≥1%	All PCs u-ssociated, p>0.05	0	816,728	IMPUTE v.2	SNPtest v.2.5	I: 8,981,701 P: 5,392,317	I: 0.99 P: 1.00
MESA	[187,188]	Affymetrix 6.0	Birdseed v1.33	≥95%	$<10^{-6}$	-	-	≥1%	First 2 PCs	2	881,666	IMPUTE2	ProbABEL	5,340,434	1.027
MGH AF study	[169]	Affymetrix 6.0	Birdseed	≥97%	$<10^{-6}$	-	-	≥1%	-	0	663,637	IMPUTE v2	PLINK v1.07	6,764,173	1.028
MGH CAMP		Infinium HumanCoreExome-24 BeadChips	zCall (GenomeStudio)	≥95%	$\geq 10^{-6}$	-	-	≥1%	PC1-PC10	10	224,343	IMPUTE2	PLINK v1.08	8,262,143	1.01
MGH Stroke	[3,189]	Affymetrix 6.0 + Illumina 610	Birdseed / GenCall	>95% MAF >5%	$<10^{-6}$	-	>±3 SD from the mean	>5%	-	2	GASROS Affymetrix: 579,083 GASROS Illumi-: 398,434 GOCHA: 521,363	IMPUTE2 v.2.3.0	SNPtest v.2.4.1	Excluded due to sample size	
WTCCC2 Munich	[3,190]	Illumina 660	GenCall	>98%	$>10^{-5}$	-	-	>1%	-	0	495,851	MACH+minimac	SNPTEST	5,891,675	1.019

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	*N* variants for imputation	Imputation software	GWAS Statistical Analysis	*N* variants analyzed	Inflation factor, lambda
PIVUS	[191]	Illumina OmniExpress +Metabochip	GenCall	≥99% (MAF< 5%) or ≥95% (MAF≥ 5%)	$>10^{-6}$	-	>3 SD from the mean	≥1%	First 2 PCs	2	738,879	IMPUTE v.2.2.2	SNPTEST v.2.5	6,045,282	1.006
PREVEND	[192]	Illumina CytoSNP12 v2	GenomeStudio	>95%	$>10^{-6}$	-	-	≥1%	First 5 PCs	5	232,571	IMPUTE1	SNPTEST v.2	5,091,540	1.031
PROSPER	[193]	Illumina Beadchip 660Quad	BeadStudio	≥98%	$<10^{-6}$	-		>1%		4	557,192	IMPUTE v.2.2.2	SNPTEST	7,819,558	1.009
RS	[194]	Illumina Infinium HumanHap550 chip v3.0	BeadStudio	≥98%	$<10^{-6}$	-	>0.336	>1%	First 4 PCs	4	512,849	Mach 1 vs 1.0.151	ProbABEL	RSI: 7,695,631 RS2: 5,543,119 RS3: 5,224,770	P&I RS1: 1.022 RS2: 1.003 RS3: 1.033
SPHFC	[195]	Affymetrix Axion Brazilian Biobank Array	Birdseed v.2	≥97%	$<10^{-6}$	-	-	≥1%	First 3 PCs			IMPUTE v3	PLINK v1.08	7,104,209	1.02
SHIP	[196]	Affymetrix Genome-Wide Human SNP Array 6.0	Birdseed2	≥80%	>0.0001	-	-	≥1%	First 10 PCs	-	905,910	IMPUTE v.2.2.2	QUICKTEST v0.95	5,289,189	0.997
TWINGENE	[197]	Illumina HumanOmni Express	GenCall	≥97%	$>10^{-7}$	-	>5 SD from the mean	≥1%	First 3 PCs	3	644,556	minimac (release 2012-10-03)	SNPTEST v.2.5	7,201,417	0.983

Study	**R**	**Array**	**Calling Algorithm**	**Per variant call rate**	**HWE p-value**	**Mendelian errors**	**Excess heterozygosity**	**MAF**	**Selection criteria for PCs**	**PCs**	***N* variants for imputation**	**Imputation software**	**GWAS Statistical Analysis**	***N* variants analyzed**	**Inflation factor, lambda**
ULSAM	[198]	Illumina Omni2.5+Metabochip	GenCall	≥99% (MAF< 5%) or ≥95% (MAF≥ 5%)	$>10^{-6}$	-	>3 SD from the mean	≥1%	First 2 PCs	2	1,587,454	IMPUTE v.2.2.2	SNPTEST v.2.5	7,297,774	0.996
WGHS	[199]	Illumina HumanHap 300 DuoPlus	BeadStudio v. 3.3	≥90%	$>10^{-6}$	-	-	≥1%	PCs 1,2, & 10	3	332,927	MaCH v.1.0.16 + minimac (release 5/29/2012)	ProbABEL	8,144,887	1.02

Table S23. General principles for quality control and filtering

Pre-imputation:
Per marker quality control: Call rate (exclude markers if <95%) Hardy-Weinberg Equilibrium (exclude markers if marked deviation) Duplicate concordance (exclude markers with high discordance rates) Mendelian inconsistencies (exclude markers with an excess of Mendelian inconsistencies) Genotype completeness (exclude markers with relatively high missingness) Polymorphism check (exclude monomorphic markers which can represent assay failures) **Per individual quality checks typically include:** Principal Component Analysis Exclude samples with high degree of missingness Exclude samples with unusual heterozygosity Exclude monomorphic markers which can represent assay failures **Exclude related individuals for non-family studies**
Imputation:
Cases and controls imputed together **Criteria for imputation:** 1000G release used for imputation: 20110521 Phase 1 Integrated release ALL Gene reference assembly: GRCh37 SNPs oriented to forward/+ strand
Individuals study analysis:
Account for genotype uncertainty of imputed SNPs **Control for population stratification**
Meta-analysis:
Criteria for including variants (GWAS/EWAS) Imputation quality >0.3 MAF ≥0.01 (GWAS), MAF ≥0.005 (EWAS) Variant present in ≥2 studies Effect allele frequency x imputation quality (INFO) x number of cases ≥10 **Criteria for including genes (gene based tests)** Cumulative MAF per gene ≤0.005 **Quality control:** Estimate genomic inflation factor lambda for each study, and adjust if lambda >1 Check distribution of meta-analysis –log10(p-values) using QQ plots

Table S24. Exome chip information per study

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	Total *N* variants analyzed
AFLMU/ MGH AF	[169]	Illumina Infinium HumanExome BeadChip v1.0	CHARGE	-	-	-	Exclude |het|>5 SD	-	p<0.01 in association adjusted for age and sex; derived under exclusion of candidate regions	11	241,465
AGES	[170]	Illumina Exome Chip v1.0	Illumina GenomeStudio2011.1	≥95%	$<10^{-6}$	-	-	-	p<0.05	0	247,501
ARIC	[172]	Illumina HumanExome Beadchip v.1.0	Centrally at CHARGE	0.95	-	-	-	-	First 10 PCs	10	223,577
BBJ	[175]	Infinium OmniExpressExome-8 BeadChip Kit	Illumina GenCall	>0.99	$>10^{-6}$ in control	no trios in samples; QC done using IBS	Yes	Exclude monomorphic in either control or case	Eigenstrates	2	61,024
BEAT-AF	[174]	Illumina HumanCoreExome	BeadStudio	≥95%	$>10^{-6}$	-	> 3 SD from the mean removed	ALL	First 10 PCs	10	495,970
BioMe	[176]	Illumina HumanOmniExpress Exome-8 v1.0	zCall (GenomeStudio)	≥90%	$>10^{-6}$	-	-	≥1%	first 4 PCs	4	241,465
BioVU	[177]	Illumina Infinium HumanExome BeadChip	GenomeStudio	>0.95	$>10^{-6}$	>1 removed	Yes (rate >0.44)	-	first 3 PCs	3	247,039
CHS	[178]	Illumina HumanExome BeadChip v1.0	GenomeStudios	≥97%	None	Any among CEPH trio controls	None	None	5 unless others are associated with the outcome	5	247,870

Study	R	Array	Calling Algorithm	Per variant call rate	HWE p-value	Mendelian errors	Excess heterozygosity	MAF	Selection criteria for PCs	PCs	Total *N* variants analyzed
FHS	[180,181]	Illumina HumanExome BeadChip v1.0	GenomeStudio v. 2011.1 and zCall following CHARGE protocol[200]	-	-	-	-	-	p<0.01 in association adjusted for age and sex	0	247,501
GS:SFHS	[183]	Illumina HumanExome Beadchip v.1-A	GenomeStudio v. 2011.1 CHARGE protocol	0.98	-	-	-	Remove Monomorphic	First 3 PCs	1	247,870
KORA	[201,202]	Illumina Infinium HumanExome BeadChip v1.0	CHARGE	-	-	-	Exclude |het|>5 SD	-	p<0.01 in association adjusted for age and sex; derived under exclusion of candidate regions	11	241,465
LURIC	[185]										Excluded
MESA	[187,200]	Illumina Exome Chip v1.0	GenomeStudio v. 2011.1 and zCall following CHARGE protocol	0.95	$>10^{-6}$	-	-	ALL	Eigenstrates	2	247,039
MGH CAMP		Infinium HumanCoreExome-24 BeadChips	zCall (GenomeStudio)	≥95%	$\geq 10^{-7}$	-	-	≥1%	First 10 PCs	10	247,501
RS	[194]	Illumina Human Exome BeadChip v1.0	zCall following CHARGE	<0.97	-	-	Het excess >0.1 AND Het excess ≤0.9	28,471 monomorphic SNPs were excluded (MAF<1E-9)	First 5	5	247,870
SHIP/SHIP-Trend	[196]	Illumina HumanExome Beadchip v.1.0	SOP v5, zCall v3.3	-	-	-	-	-	First 10 PCs	First 10 PCs	247,039

Study	**R**	**Array**	**Calling Algorithm**	**Per variant call rate**	**HWE p-value**	**Mendelian errors**	**Excess heterozygosity**	**MAF**	**Selection criteria for PCs**	**PCs**	**Total *N* variants analyzed**
WGHS	199,203	Illumina HumanExome Beadchip v.1.1A	GenomeStudio v. 2011.1 and zCall following CHARGE protocol	0.95	-	-	-	-	-	0	247,727
WHI - CT		Illumina Human Exome BeadChip v1.0	GenomeStudio v2010.3	0.95	-	-	-	-	Plink	2	246,670
WHI - OS		Illumina Human Exome BeadChip v1.0	GenomeStudio v2010.3	0.95	-	-	-	-	Plink	2	246,670

Table S25. Baseline characteristics of African American ancestry replication studies

	Cases	Controls	Total
N	447	442	889
Women, %	44	48	46
Age at enrollment, mean (SD)	55 (11)	61 (14)	60 (14)
Age at diagnosis, mean (SD)	58 (14)	-	-
Age range (Q1-Q3)	50-61	52-72	51-69
HTN, %	88	87	88
DM, %	37	41	39
HF, %	24	8	16
MI, %	8	3	6

SD, standard deviation; HTN, hypertension; DM, diabetes mellitus; HF, heart failure; MI, myocardial infarction.

Table S26. Results from replication in African American ancestry studies

rsID	Risk allele	RAF, %	OR	95% CI	P-value
rs115339321	T	97	1.53	0.82-2.18	0.18
rs79433233	A	3	1.36	0.75-2.47	0.31

RAF, risk allele frequency; OR, odds ratio; CI, confidence interval.

Table S27. Results from DEPICT pathway analysis of GWAS meta-analysis results

Original gene set ID	Original gene set description	Nominal P-value
KEGG ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY ARVC	KEGG ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY ARVC	$1.27x10^{-6}$
KEGG_TIGHT_JUNCTION	KEGG TIGHT JUNCTION	$1.75x10^{-6}$
MP:0003157	impaired muscle relaxation	$2.28x10^{-6}$
GO:0016459	myosin complex	$8.31x10^{-6}$
GO:0060429	epithelium development	$1.17x10^{-5}$
MP:0000751	myopathy	$1.25x10^{-5}$
GO:0030855	epithelial cell differentiation	$1.67x10^{-5}$
KEGG HYPERTROPHIC CARDIOMYOPATHY HCM	KEGG HYPERTROPHIC CARDIOMYOPATHY HCM	$3.07x10^{-5}$
REACTOME MUSCLE CONTRACTION	REACTOME MUSCLE CONTRACTION	$4.18x10^{-5}$
GO:0031589	cell-substrate adhesion	$8.50x10^{-5}$

Table S28. Top 5 enriched canonical pathways from Ingenuity Pathway Analysis of GWAS meta-analysis results

Ingenuity Canonical Pathways	P-value	Ratio	Molecules
Coagulation System	0.0088	3/35 (8.6%)	F11, KLKB1, PLAU
Clathrin-mediated Endocytosis Signaling	0.011	7/197 (3.6%)	MET, UBD, FGF17, ACTR2, AAK1, HIP1, PCYOX1
Protein Ubiquitination Pathway	0.013	8/255 (3.1%)	UBD, UBE2G2, USP18, UBE2Q1, BAG1, PSMD5, USP54, PSMD3
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)	0.018	2/17 (11.8%)	FDPS, PMVK
Ephrin Receptor Signaling	0.02	6/174 (3.4%)	ACTR2, SHC1, EFNA3, CREB5, EFNA4, EFNA1

Table S29. Enriched diseases or functions annotation from Ingenuity canonical pathway analysis of GWAS meta-analysis results

Diseases or Functions Annotation	P-value	*N* molecules	Molecules
Arrhythmia of heart ventricle	$3.0x10^{-9}$	12	CASQ2, CSF3, DSG2, HCN4, KCNG2, KCNJ5, PKP2, SCN10A, SCN5A, TBX5, THRA, TTN
Ventricular tachycardia	$1.7x10^{-8}$	10	CASQ2, CSF3, DSG2, HCN4, KCNG2, KCNJ5, PKP2, SCN5A, TBX5, THRA
Tachycardia	$2.5x10^{-8}$	11	CASQ2, CSF3, DSG2, HCN4, KCNG2, KCNJ5, PITX2, PKP2, SCN5A, TBX5, THRA
Arrhythmia	$5.0x10^{-8}$	16	CASQ2, CSF3, DSG2, HCN4, KCNG2, KCNJ5, NR3C1, PITX2,PKP2, PLN, SCN10A, SCN5A, TBX5, THRA, TTN, TUBA8
Ventricular fibrillation	$9.5x10^{-7}$	7	DSG2, KCNG2, KCNJ5, PKP2, SCN5A, THRA, TTN
Cardiomyopathy of heart ventricle	$1.2x10^{-6}$	6	CAV1, DSG2, HCN4, PKP2, SCN5A, TTN
Cardiac fibrillation	$1.6x10^{-6}$	11	DSG2, KCNG2, KCNJ5,NR3C1, PITX2,PKP2, PLN, SCN5A, THRA, TTN, TUBA8
Hypertrophy of cardiac muscle	$5.5x10^{-6}$	10	CAV1, CSF3, FBXO32, IL6R, mir-23, PLAU, RAB1A, SHC1, TBX5, TTN
Arrhythmogenic right ventricular dysplasia	$5.7x10^{-6}$	5	DSG2, HCN4, PKP2, SCN5A, TTN

3. SUPPLEMENTARY METHODS

Replication genotyping of rs115339321 and rs79433233 in African American populations
Custom TaqMan® genotyping probes for rs115339321 and rs79433233 were obtained from Life Technologies. Genotyping was performed on 5 ng of DNA input using the TaqMan® genotyping master mix on a Bio-Rad CFX384 real time PCR instrument. Genotyping was performed in 447 AF cases and 442 referents obtained from four studies (BioVU, Duke Biobank, MGH, and Penn Biobank), with genotype calls being performed by end state fluorescence after 40 cycles.

***In silico* replication in the Biobank Japan (BBJ)**
We performed an *in silico* replication of the most significant variant at each novel AF associated genetic locus in a GWAS of 8,180 cases and 28,612 referents of Asian ancestry. The cases were selected from the Biobank Japan which contains DNA and serum samples collected throughout Japan and AF was defined as persistent or paroxysmal AF diagnosed by a physician. The referents were selected from the Tohoku Medical Megabank organization,[204] the Japan Public Health Centre-based Prospective study, and the Japan Multi-institutional Collaborative Cohort (J-MICC) Study. Samples were genotyped using the Illumina Human OmniExpress BeadChip Kit and Infinium OmniExpressExome BeadChip Kit. Only autosomal variants were included in the GWAS. Variants with call rate <99%, variants that deviated from Hardy-Weinberg equilibrium among control samples ($<1x10^{-6}$), and non-polymorphic variants were excluded.

***In silico* replication in the UK Biobank**
Details of genotyping, imputation, and calculation of principal components of ancestry in the UK biobank interim dataset can be found on the UK biobank website (http://www.ukbiobank.ac.uk/). Briefly, samples were genotyped either by UK BiLEVE Axiom array (UKBL) or UK Biobank Axiom array (UKBB). Both arrays include ~800,000 SNPs and more than 95% of common marker contents are similar. Imputation was phased by modified version of SHAPEIT2 and imputed by IMPUTE2, using a combined panel of UK10K haplotype and 1000G phase 3 as the reference panel. All significant variants detected in the discovery study passed quality control filters in the UK biobank data (imputation quality info ≥ 0.4, variant missing rate < 5%, individual missing rate < 10%, and variant genotype probability > 0.9 in > 90% of the individuals). Variants were then transformed to hard-called genotypes (probability threshold ≥ 0.9, minor allele frequency (MAF) ≥ 0.01, and missing rate per variant <5%). We used logistic regression to test the association between each hard-called variant and risk of AF using an additive genetic model, adjusting for baseline age, sex, array, and the first 15 principal components of ancestry. Quality control, transformation and analyses were performed by QCTOOL and Plink v1.90b. Since we performed an *in silico* replication of 31 variants, we set a conservative significance threshold of 1.6×10^{-3} (0.05 / 31).

***In silico* evaluation of novel AF loci**
All statistically significant variants and genes from GWAS and RVAS analyses were selected for an *in silico* assessment through lookups in the following databases: The Gene Tissue Expression database (GTEx),[205] RegulomeDB,[206] HaploREG,[207] GeneCards (www.genecards.org/), dbSNP.[208] From the GTEx search, we report statistically significant eQTLs in cardiac and skeletal muscle tissues. The NHGRI-EBI GWAS catalog[2] was interrogated with the aim of identifying possible pleiotropy with other cardiovascular phenotypes. At each locus, we defined a region based on LD span ($r^2 > 0.2$) with the lead SNP. We searched the GWAS catalog for all SNPs within these regions and report LD of proxies with the lead SNP when available. LD information was identified using the SNiPA tool[209] (Available at http://www.snipa.org. Accessed 6-24-2016.)

Pathway analyses

Pathway analyses provide a potential route to investigate the collective effects of multiple genetic variants on biological systems. We utilized two different methods for pathway analysis:

1. DEPICT

We ran the analysis DEPICT,[210] which integrates multiple layers of evidence to identify causal genes at GWAS loci. From meta-analysis results, we first performed clumping to identify independent loci using plink.[211] We then performed analysis using DEPICT with the default settings.

2. Ingenuity Pathway Analysis (IPA)

Data were analyzed through the use of QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity). For each of the tested genetic variants, we mapped it back to the reference human genome (NCBI Build 37, 2009) and examined its location relative to RefSeq genes (May 15, 2016). The gene score was defined as the most significant variants that were located within 110kb upstream and 40kb downstream of the gene's most extreme transcript boundaries. Of the 27,011 genes evaluated, 338 reached a score less than $5x10^{-6}$. These genes were then imported into IPA analysis. Fisher's exact test was used to justify the enrichment of each of the canonical pathways.

Detailed Description of participating studies

The meta-analyses described in this manuscript included the following studies described elsewhere: The **Age, Gene/Environment Susceptibility Study (AGES) Reykjavik study**[169]**,** the **Atrial Fibrillation Biobank LMU (AFLMU)** in the context of the **Arrhythmia-Biobank-LMU** (formerly known as **AFNET**) and the **Cooperative Health Research in the Region of Augsburg (KORA)**[169]**,** the **Atherosclerosis Risk in Communities (ARIC) study**[169]**, Cleveland Clinic Lone Atrial Fibrillation GeneBank Study (CCAF)**[169]**,** the **Cardiovascular Health Study (CHS)**[169], **Framingham Heart Study (FHS)**[169], **Massachusetts General Hospital (MGH) AF study**[169], the **Rotterdam Study (RS)**[169], the **Study of Health in Pomerania (SHIP)**[169], **BioVU**[212], the **Women's Genome Health Study (WGHS)**[169], The **PROspective Study of Pravastatin in the Elderly at Risk (PROSPER)**[175], **Biobank Japan (BBJ)**[175], in addition to the studies described here:

ANGES: The Angiography and Genes Study (ANGES) population consists of 1,000 Finnish individuals participating in the ongoing ANGES study. Angiographic, genetic, and covariate data was available for 808 individuals (516 men and 292 women; mean age 62±10). The data was collected between September 2002 and July 2005. All patients underwent coronary angiography at Tampere University Hospital due to clinically suspected coronary artery disease. The study is a cross-sectional study, and after the angiography, patients were treated according to the Finnish Current Care Guidelines. Patients were also interviewed by a study nurse, and a questionnaire was used to collect general information - age, sex, body mass index, alcohol consumption, smoking, medication, as well as traditional risk factors of atherosclerosis and myocardial infarction. The study has been approved by the Ethics Committee of Pirkanmaa Hospital District and written informed consent was obtained from each patient.

BEAT-AF: The Basel Atrial Fibrillation Cohort Study (BEAT-AF) is a prospective observational, multicenter cohort study. Between 2010 and 2014, 1550 patients with documented AF were enrolled across 7 centers in Switzerland. Exclusion criteria were the inability to sign informed consent and the presence of short transient forms of AF. At baseline, patients completed detailed questionnaires about personal,

medical, nutritional and lifestyle factors, current AF symptoms and co-morbidities. Current medications were recorded. A resting 12-lead electrocardiogram (ECG) was recorded and all patients underwent venous blood sampling at the local study center, including DNA from leukocytes. Yearly follow-ups by mailed questionnaires and phone interviews were performed in all patients in order to collect similar information as at baseline and to obtain details about adverse events.
Referents were enrolled from the 'genetic and phenotypic determinants of blood pressure and other cardiovascular risk factors' (GAPP) study, which is an ongoing prospective population-based cohort study among healthy adults in the Principality of Liechtenstein. Between 2010 and 2013, all inhabitants of the Principality of Liechtenstein aged between 25 and 41 years were invited and 2170 agreed to participate in the study. Main exclusion criteria were established cardiovascular disease, chronic kidney disease, diagnosed sleep apnea, a body mass index (BMI) > 35 kg/m2, intake of antidiabetic drugs or any other severe illness. Examinations included detailed assessment of personal, medical, lifestyle and nutritional factors, standardized assessment of weight, height and waist circumference, blood pressure measurement, electrocardiography, bioimpedance analysis, blood, urinary and genetic sampling, spirometry and sleep pulse oximetry with nasal flow measurement. Follow-up examinations are scheduled every 3-5 years. The detailed study design has previously been published.[174]

Bio*Me*: The Mount Sinai Bio*Me* Biobank is an ongoing, prospective, hospital- and outpatient- based population research program operated by The Charles Bronfman Institute for Personalized Medicine (IPM) at Mount Sinai and has enrolled over 33,000 participants since September 2007. Bio*Me* is an Electronic Medical Record (EMR)-linked biobank that integrates research data and clinical care information for consented patients at The Mount Sinai Medical Center, which serves diverse local communities of upper Manhattan with broad health disparities. Bio*Me* populations include 25% of African ancestry (AA), 36% of Hispanic Latino ancestry (HL), 30% of white European ancestry (EA), and 9% of other ancestry. The BioMe disease burden is reflective of health disparities in the local communities. Bio*Me* operations are fully integrated in clinical care processes, including direct recruitment from clinical sites waiting areas and phlebotomy stations by dedicated recruiters independent of clinical care providers, prior to or following a clinician standard of care visit. Recruitment currently occurs at a broad spectrum of over 30 clinical care sites.

Information on AF, age, sex, body mass index (BMI), type 2 diabetes (T2D), hypertension (HYP), heart failure (HFAIL), and myocardial infarct (MI) was derived from participants' EMRs: Age, sex and BMI were derived from the day of enrolment to the BioMe biobank. Prevalent AF cases were defined as Bio*Me* participants with the ICD-9 code 427.31 (atrial fibrillation) and/or 427.32 (atrial flutter) and controls as individuals who have had ECG's but did not have AF or flutter ICD-9 codes. HYP, HFAIL, and MI were defined using the ICD-9 codes 401.*, 428.*, and 410.*, respectively. In addition to the ICD-9 codes, also individuals taking antihypertensive drugs were considered as having HYP. T2D was defined using the eMerge T2D case and control definition algorithms.[213] The algorithms used were developed by a multidisciplinary team of scientists, clinicians and software specialists and have been validated with excellent performance statistics; 100% sensitivity and >98% positive predictive value for cases, and ≥98% sensitivity and ≥98% positive predictive value for controls.

Bio*Me* participants were genotyped with the Illumina HumanOmniExpressExome-8 v1.0 beadchip array and imputed to the 1000 Genomes Project Phase 1 (March12) reference panel using IMPUTE2. Genome-wide association studies (GWAS) were carried out using SNPTEST 2.4.1 after stratifying by self-reported ancestry (AA: 174 AF cases and 2130 controls; EA: 291 AF cases and 860 controls; HL: 277 AF cases and 3081 controls) and adjustment for a) age, sex and the first 4 GWAS PCs (Model1) and b) age, sex, BMI, T2D, HYP, HFAIL, MI, and the first 4 GWAS PCs (Model2). To ensure high quality of the association results, variants with imputation quality < 0.3, Hardy-Weinberg p-value < $1x10^{-5}$ or minor allele frequency < 0.01 were excluded.

BioVU*:* BioVU is the Vanderbilt University Medical Center's biorepository linked to de-identified electronic health records. BioVU operations[212] and ethical oversight[214] 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).[177]

Corogene: The Corogene study was designed as a large cohort to study mainly CAD, but also other related heart diseases such as heart failure and aortic valve disease. We selected the patients from the CAD point of view, and decided to include over 5000 consecutive patients assigned for coronary angiogram. In Finland, coronary angiogram is performed to practically all patients assigned for invasive heart examination. Despite technical developments in diagnostics, coronary angiogram is still the gold standard for evaluating coronaries. The purpose of this study is to follow contemporary trends in coronary heart disease, and related heart disease risk factors, genetics and epigenetics by collecting cohorts referred to heart examination. New cohorts will be collected at 5-year intervals in order to see trends in CAD, its risk factors and epigenetics.

FINCAVAS: The purpose of the Finnish Cardiovascular Study (FINCAVAS) is to construct a risk profile - using genetic, haemodynamic and electrocardiographic (ECG) markers - of individuals at high risk of cardiovascular diseases, events and deaths. All patients scheduled for an exercise stress test at Tampere University Hospital, who gave informed consent to participate, were recruited between October 2001 and December 2007. The total number of participants was 4,567. In addition to repeated measurements of heart rate and blood pressure, digital high-resolution ECG at 500 Hz was recorded continuously during the entire exercise test, including the resting and recovery phases. About 20% of the patients were examined with coronary angiography. Genetic variations known or suspected to alter cardiovascular function or pathophysiology were analyzed to elucidate the effects and interactions of these candidate genes, exercise, and commonly used cardiovascular medications.

GS:SFHS: Generation Scotland: Scottish Family Health Study (GS:SFHS) is a family-based genetic epidemiology study of ~24,000 volunteers from ~7000 families across Scotland with the capacity for follow-up through record linkage and re-contact. Participants completed a demographic, health and lifestyle questionnaire and provided biological samples including DNA, and ~21,500 participants underwent detailed clinical assessment, including anthropometric, cardiovascular, respiratory, cognition and mental health. Genetic analysis (GWAS) is complete on 20,000 participants with full baseline data and CHI linkage, with linkage to SMR, prescriptions and dental records. A full cohort description can be found elsewhere.[183] AF was ascertained as a diagnosis of atrial fibrillation by linkage to one or more inpatient visits with ICD-10 code I48 or ICD-9 427.31 in the Scottish Morbidity Record (SMR1) database before or after recruitment to GS:SFHS.

HNR: The study population of the Heinz Nixdorf Recall (HNR) study has been described in detail elsewhere.[184] Approved by the relevant institutional ethics committees, the study follows strict internal and external quality assurance protocols. Briefly, the study cohort comprises 4,814 men and women

aged 45 – 75 years from the three adjacent Ruhr cities Essen, Bochum and Mülheim/Ruhr. The vast majority of the study population is of central European ancestry. The study area covers a region of approximately 600 km² with almost 1.2 million inhabitants. Subjects were randomly selected from statutory lists of residence and gave informed consent. The baseline examinations were from 2000-2003, the 5-Year follow-Up from 2006-2008 and the 10-Year follow-up from 2011-2015. A standardized digital 12-lead resting surface ECG was sampled at 250 Hz and recorded on a MAC 5000® ECG recorder (GE Healthcare, Freiburg, Germany). ECGs were interpreted automatically using the integrated 12SL-Code® [12SL ECG analysis with age & gender specific criteria. Physician's guide. PN 416791-004 Revision A. GE Medical Systems IT, 2000]. ECG findings were coded and transferred to our database. The ECG-codes #161 and #162 are for atrial fibrillation and atrial flutter, respectively and were combined for the purpose of this analysis.

LURIC: The Ludwigshafen Risk and Cardiovascular Health (LURIC) study is an ongoing prospective study of more than 3,300 individuals of German ancestry in whom cardiovascular and metabolic phenotypes (CAD, MI, dyslipidemia, hypertension, metabolic syndrome and diabetes mellitus) have been defined or ruled out using standardized methodologies in all study participants.[185] Inclusion criteria for LURIC were: German ancestry (limitation of genetic heterogeneity), clinical stability (except for acute coronary syndromes) and availability of a coronary angiogram. Exclusion criteria were: any acute illness other than acute coronary syndromes, any chronic disease where non-cardiac disease predominated and a history of malignancy within the last five years. Genome-wide analyses using the Affymetrix 6.0 have been completed in all participants. A 10-year clinical follow-up for total and cause specific mortality has been completed.

MDCS: The Malmö Diet and Cancer study (MDCS) is a community-based prospective epidemiologic cohort of middle-aged individuals from Southern Sweden.[186] In total, 30,447 subjects attended a baseline exam in 1991-1996, when they filled out a questionnaire and underwent anthropometric and blood pressure measurements. Prevalent or incident cases of atrial fibrillation (AF) were ascertained from nation-wide hospital registers with high validity as described previously.[186] Genome-wide genotyping of single nucleotide variants was performed using the Illumina Human Omni Express Exome BeadChip kit. Genotyping was performed in a nested case-cohort design, including a random subset of 5878 subjects.

MESA: The Multi-Ethnic Study of Atherosclerosis (MESA) is a study of the characteristics of subclinical cardiovascular disease (disease detected non-invasively before it has produced clinical signs and symptoms) and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease. The cohort is a diverse, population-based sample of 6,814 asymptomatic men and women aged 45-84. Approximately 38 percent of the recruited participants are white, 28 percent African American, 22 percent Hispanic, and 12 percent Asian (predominantly of Chinese descent). Participants were recruited during 2000-2002 from 6 field centers across the U.S. (at Wake Forest University; Columbia University; Johns Hopkins University; the University of Minnesota; Northwestern University, and the University of California – Los Angeles). All underwent anthropomorphic measurement and extensive evaluation by questionnaires at baseline, followed by 4 subsequent examinations at intervals of approximately 2-4 years. Age and sex were self-reported. Current AF at baseline was an exclusion criterion. Follow-up phone calls to study participants (every 9-12 months) were used to identify all hospitalizations. Medical records, including discharge diagnoses, were obtained for each hospitalization. Incident AF was defined by International Classification of Disease codes 427.31 or 427.32 (9th revision). In addition, new diagnoses of AF were identified at follow-up by the presence of AF or atrial flutter on a study ECG at Exam 5 (approximately 10 years after baseline).

Further information can be found at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000209.v13.p3.

MGH CAMP: The MGH Cardiology and Metabolic Patient (MGH CAMP) cohort comprises 3857 subjects recruited between 2008 and 2012. Two thirds of the subjects were drawn from patients who had appointments with a physician in the MGH Heart Center, whereas one third were recruited independent of any hospital visit. All subjects had plasma and serum samples collected, as well as blood for genomic DNA. Subjects with known diabetes had vascular reactivity measurements (FMD of brachial artery), while subjects without known diabetes had an oral glucose tolerance test. Exome Core Chip genotyping was performed on all subjects. AF was defined as a self-reported history of fibrillation or flutter at study enrollment, or based on a validated medical record ascertainment algorithm (PPV 88%) that utilizes electrocardiographic and relevant diagnostic, procedure, and medication data.[215]

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.

PIVUS: The participants were randomly sampled from all men and women at age 70 living in Uppsala County in 2001 (www.medsci.uu.se/PIVUS). Of the 2025 individuals invited, 1016 participated. The participants underwent a medical examination including a detailed questionnaire on lifestyle and socioeconomic factors, fasting blood sampling, blood pressure measurement and anthropometric measurements, as previously described.[191] Blood and plasma samples have been frozen until analysis, and blood tests performed include a wide variety of traditional and more recent CVD risk factors, along with DNA extraction. In addition, the individuals have undergone extensive phenotyping including whole body MRI, echocardiography, endothelial function measurements, carotid ultrasound, DXA, and spirometry. The participants have been re-examined at age 75 and 80. AF was defined by 12-lead ECG at the examinations, as well as diagnosis of atrial fibrillation or flutter in the Swedish National Patient Register before or after the baseline examination (inpatient and specialist outpatient care; ICD-9 code, 427.3 and ICD-10 code, I48).

PREVEND: The PREVEND cohort study was founded in 1997, and is an ongoing community-based cohort study including 8592 inhabitants of the city of Groningen, The Netherlands. PREVEND is investigating the natural course of microalbuminaria and its relation to renal and cardiovascular disease. Details of the protocol, AF ascertainment and covariate definitions have been described elsewhere

(www.prevend.org). The baseline screening program consisted of 2 outpatient visits to assess demographic factors, anthropometric measurements, cardiovascular and metabolic risk factors, and health behavior and to collect blood samples and 2 24-h urine samples on 2 consecutive days. Participants were seen at 3-year intervals in the PREVEND outpatient clinic. AF was ascertained if either atrial flutter or AF was present on a 12-lead ECG obtained at one of the three PREVEND follow-up visits, or at an outpatient visit or hospital admission in the two hospitals in the city of Groningen (University Medical Center Groningen and Martini Hospital). Participants without an electrocardiogram (ECG) (n=248), as well as participants with prevalent AF at the baseline screening (n=79) and without GWAS information (n=4632) were excluded, leaving 3633 for analysis.[216]

SPHFC: Participants for the Sao Paolo Heart Failure Cohort (SPHFC) 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 with symptomatic heart failure (stage C) were enrolled. Different heart failure etiologies were included. Patients with prior myocardial infarction (<3 months), unstable angina, hypertrophic cardiomyopathy, valve heart disease candidates to surgical treatment, obstructive pulmonary disease, severe renal or hepatic dysfunction, current history of cancer, severe peripheral arterial disease, cerebrovascular disease and active infection were excluded. AFib status was determined if either atrial flutter or AF was present on a 12-lead ECG at baseline evaluation or prior and could be confirmed by electronic medical record review.

TWINGENE: The Swedish Twin Registry contains data regarding health, health-related behaviors, physical activity, eating habits, and environmental stressors, along with other information from Swedish national registries. TWINGENE includes twins born before 1958 that were contacted to participate at the baseline examination between April 2004 and December 2008.[217] Health and medication data were collected from self-reported questionnaires, while blood sampling and in-person testing, including blood pressure measurement and anthropometrics were completed at a local health care center. Several biomarkers, including lipid profiles, fasting glucose, HbA1C and CRP, have been measured, and aliquoted serum is stored at the Karolinska Institutet Biobank. AF was defined as a diagnosis of atrial fibrillation or flutter in the Swedish National Patient Register before or after the baseline examination (inpatient and specialist outpatient care; ICD-9 code, 427.3 and ICD-10 code, I48).

ULSAM: All men born between 1920 and 1924 in Uppsala, Sweden were invited to participate at age 50 in this longitudinal cohort study that was started in 1970. Participants were reinvestigated at the ages of 60, 70, 77, 82 and 88 years.[198] Blood samples for DNA extraction and main cardiovascular risk factors were available from the investigation at age 70. The participants have undergone extensive phenotyping at repeated time points, including euglycemic clamps, oral glucose tolerance tests, DXA, echocardiography, 24-h ambulatory blood pressure measurement, and a range of biomarkers. AF was defined by 12-lead ECG at the examinations, as well as diagnosis of atrial fibrillation or flutter in the Swedish National Patient Register (inpatient and specialist outpatient care; ICD-9 code, 427.3 and ICD-10 code, I48).

WHI: The Women's Health Initiative (WHI) is one of the largest (n=161,808) studies of women's health ever undertaken in the United States. The WHI studies consisted of randomized CT, which assigned 68,132 women to active or placebo hormone therapy (HT), dietary modification or control, and/or calcium/vitamin D, supplementation or placebo with specific outcomes of common diseases of aging in women, and also an observational study (OS), which collected data on biological and lifestyle factors and health outcomes. A diverse population including 26,045 (17%) women from minority groups were

recruited from 1993-1998 at 40 clinical centers across the U.S. Details of the study design have been previously described.[218,219] For the CT and OS participants enrolled in WHI and who had consented to genetic research, DNA was extracted by the Specimen Processing Laboratory at the Fred Hutchinson Cancer Research Center (FHCRC) using specimens that were collected at the time of enrollment in to the study (between 1993 and 1998).
Baseline AF was determined by an initial questionnaire, which probed for self-reported AF or by presence of AF on the baseline 12-lead electrocardiogram. Women were followed up with a medical history update questionnaire at years 3 to 8, which specifically probed for self-reported AF and hospitalizations.

WTCCC2-Munich: The Wellcome Trust Case Control Consortium 2 Munich (WTCCC2-Munich) study is a hospital-based study on ischemic stroke genetics. Only consecutive European Caucasians recruited from a single dedicated Stroke Unit from South-German origin were selected for this study from the Department of Neurology, Klinikum Großhadern, Ludwig-Maximilians-University, Munich. Age, sex and clinical risk factors were collected. AF was identified by ECG measurement on day of admission. For the German samples controls were Caucasians of German origin participating into the population KORAgen study (www.gsf.de/kora/en/english.html). This survey represents a gender- and age stratified random sample of all German residents of the Augsburg area and consists of individuals 25 to 74 years of age, with about 300 subjects for each 10-year increment. All controls were free of a history of stroke, AF or other cardiovascular diseases.

African American replication studies included:

Penn Medicine Biobank: 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.

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, 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, 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.

4. SUPPLEMENTARY RESULTS

Ancestry-specific GWAS meta-analyses

Separate GWAS in 15,993 cases and 113,719 referents of European ancestry revealed one additional association on chromosome 15q21 (rs2921421, OR 1.72, 95% CI 1.42-2.09, P=3.29x10^{-8}, **Supplementary Table S6**); however, there was only one significant variant at this locus and the variant was imputed with low quality across all studies reducing our confidence in this finding. Additional replication in another European ancestry study is needed to clarify the relevance of rs2921421. In meta-analysis of 837 cases and 2456 referents of Asian ancestry we identified an association on chromosome 12q15 (rs7138621, OR 7.92, 95% CI 4.26-14.73, P=6.48x10^{-11}), which was not significant in *in silico* replication in 8180 cases and 28,612 referents in the Biobank Japan (**Supplementary Table S10**). Separate meta-analyses in individuals of Brazilian and Hispanic descent did not identify additional loci; however, our power was limited in each of these sub-groups.

GWAS meta-analyses of incident and prevalent AF in Europeans

Separate GWAS meta-analyses of incident (7232 cases) and prevalent (8656 cases) AF in Europeans showed similar results to the European ancestry analysis (**Supplementary Tables S8-S9, Supplementary Figs. S7-S8**); however, we did reveal a novel AF locus associated with prevalent AF at chromosome 12p11 (rs1454934, OR 1.16, 95% CI 1.1-1.22, P=4.18x10^{-8}). The most significant variant at this locus was intronic to the gene plakophilin-2 (*PKP2*), which encodes an important component of the desmosome and is known to be associated with arrhythmogenic right ventricular cardiomyopathy[220] and Brugada syndrome.[221,222]

Replication of genetic variants specific to African American ancestry GWAS meta-analysis

The variants rs115339321 (OR 1.53, 95% CI 0.82-2.18, P=0.18) and rs79433233 (OR 1.36, 95% CI 0.75-2.47, P=0.31) were not significantly associated with AF in 447 AF cases and 442 referents of African American ancestry (**Supplementary Table S25-26**). The lack of replication may be caused by the small sample size of the replication study. Further replication in a larger sample of African American ancestry is needed to clarify the role of the variants rs115339321 and rs79433233.

Pathway analyses

1. DEPICT

The most significant pathway identified using the DEPICT software was the arrhythmogenic right ventricular cardiomyopathy (ARVC) pathway (P=1.3x10^{-6}, **Supplementary Table S27**). None of the pathways analyzed reached an FDR <5%.

2. IPA

The most significantly enriched biological pathway was the coagulation system (P=0.0088). In addition, many genes were involved in the clathrin-mediated endocytosis signaling pathway (P=0.011) and the protein ubiquitination pathway (P=0.013). The most significant pathways are listed in **Supplementary Table S28** None of the pathways reached the significance threshold (FDR<5%). In addition, many of the genes investigated were involved in arrhythmia mechanisms (**Supplementary Table S29**).

5. FUNDING INFORMATION

AFLMU/KORA: This work is funded by the European Commission's 7th Framework Programme FP7-HEALTH-2013 No. 602299: EU-CERT-ICD to Dr. Kääb and 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, the BMBF Spitzencluster personalized medicine m4 (01 EX1021E), the LMU Excellence Initiative (42595-6), and the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ, all to Dr. Kääb. The KORA platform is funded by the BMBF and by the State of Bavaria.

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).

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

Biobank Japan: 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.

BioVU: The dataset used in the analyses described were obtained from Vanderbilt University Medical Centers BioVU which is supported by institutional funding and by the Vanderbilt CTSA grant UL1 TR000445 from NCATS/NIH. Genome-wide genotyping was funded by NIH grants RC2GM092618 from NIGMS/OD and U01HG004603 from NHGRI/NIGMS.

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, R01HL068986, 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).

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.

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. 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 1R01 HL092577, 1RO1 HL076784, 1R01 AG028321, Evans Center for Interdisciplinary Biomedical Research ARC on Atrial Fibrillation at Boston University (http://www.bumc.bu.edu/evanscenteribr/the-arcs/the-arcs/ (Benjamin), 6R01-NS 17950 and American Heart Association 09FTF2190028.

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).

HNR: The Heinz Nixdorf Recall Study thanks the Heinz Nixdorf Foundation (Germany) and projects SI 236/8-1 and SI 236/9-1 from the German Research Council for the generous support of this study. We acknowledge the support of the Sarstedt AG & Co. (Nümbrecht, Germany) for laboratory equipment. We are indebted to all study participants and to the dedicated personnel of the study center of the Heinz Nixdorf Recall study. Advisory Board: Meinertz T, Hamburg, Germany (Chair); Bode C, Freiburg, Germany; de Feyter PJ, Rotterdam, Netherlands; Güntert B, Hall i.T., Austria; Gutzwiller F, Bern, Switzerland; Heinen H, Bonn, Germany; Hess O, Bern, Switzerland; Klein B, Essen, Germany; Löwel H, Neuherberg, Germany; Reiser M, Munich, Germany; Schwaiger M, Munich, Germany; Steinmüller C, Bonn, Germany; Theorell T, Stockholm, Sweden; Willich SN, Berlin, Germany.

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

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 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 Family was provided by grants R01-HL-071205, R01-HL-071051, R01-HL-071250, R01-HL-071251, R01-HL-071252, R01-HL-071258, and R01-HL-071259, and by UL1-RR-025005 and UL1RR033176 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. 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 Stroke study: The GOCHA study was supported by the grant R01NS059727 from the NINDS. The GASROS study was supported by the grant U01 NS069208-01 from the NINDS.

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 AF (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.

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).

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 Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG.

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.

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.

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).

WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. A full listing of the WHI investigators can be found at https://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf. We thank the Women's health Initiative investigators, staff, and study participants for their contributions.

WTCCC2-Munich: This work was supported by grants received from the German Federal Ministry of Education and Research (BMBF) in the context of the e:Med program (e:AtheroSysMed), the FP7 European Union project CVgenes@target (261123), the DFG as part of the CRC 1123 (B3), the Corona Foundation, and the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain).

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

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

Role of the Sponsor:
None of the funding agencies had any role in the study design, data collection or analysis, interpretation of the data, writing of the manuscript, or in the decision to submit the manuscript for publication.

6. SUPPLEMENTARY REFERENCES

1. Pruim, R. J. *et al.* LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26,** 2336–2337 (2010).
2. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* **42,** D1001-6 (2014).
3. Traylor, M. *et al.* Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol.* **11,** 951–62 (2012).
4. Dichgans, M. *et al.* Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke.* **45,** 24–36 (2014).
5. Gretarsdottir, S. *et al.* Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann. Neurol.* **64,** 402–9 (2008).
6. den Hoed, M. *et al.* Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. *Nat. Genet.* **45,** 621–31 (2013).
7. Arking, D. E. *et al.* Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. *Nat. Genet.* **46,** 826–36 (2014).
8. Smith, J. G. *et al.* Impact of ancestry and common genetic variants on QT interval in African Americans. *Circ. Cardiovasc. Genet.* **5,** 647–55 (2012).
9. Holm, H. *et al.* Several common variants modulate heart rate, PR interval and QRS duration. *Nat. Genet.* **42,** 117–22 (2010).
10. Pfeufer, A. *et al.* Common variants at ten loci modulate the QT interval duration in the QTSCD Study. *Nat. Genet.* **41,** 407–414 (2009).
11. Vasan, R. S. *et al.* Genetic Variants Associated With Cardiac Structure and Function. *JAMA* **302,** 168 (2009).
12. Newton-Cheh, C. *et al.* Common variants at ten loci influence QT interval duration in the QTGEN Study. *Nat. Genet.* **41,** 399–406 (2009).
13. Ritchie, M. D. *et al.* Genome- and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk. *Circulation* **127,** 1377–85 (2013).
14. Nolte, I. M. *et al.* Common genetic variation near the phospholamban gene is associated with cardiac repolarisation: meta-analysis of three genome-wide association studies. *PLoS One* **4,** e6138 (2009).
15. Eijgelsheim, M. *et al.* Genome-wide association analysis identifies multiple loci related to resting heart rate. *Hum. Mol. Genet.* **19,** 3885–94 (2010).
16. Verweij, N. *et al.* Genetic Determinants of P Wave Duration and PR Segment. *Circ. Cardiovasc. Genet.* **7,** 475–81 (2014).
17. Pfeufer, A. *et al.* Genome-wide association study of PR interval. *Nat. Genet.* **42,** 153–159 (2010).
18. Sano, M. *et al.* Genome-wide association study of electrocardiographic parameters identifies a new association for PR interval and confirms previously reported associations. *Hum. Mol. Genet.* **23,** 6668–76 (2014).
19. Butler, A. A. M. *et al.* Novel loci associated with PR interval in a genome-wide association study of 10 African American cohorts. *Circ. Cardiovasc. Genet.* **5,** 639–646 (2012).
20. Hong, K.-W. *et al.* Identification of three novel genetic variations associated with electrocardiographic traits (QRS duration and PR interval) in East Asians. *Hum. Mol. Genet.* **23,** 6659–67 (2014).
21. Sotoodehnia, N. *et al.* Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nat. Genet.* **42,** 1068–76 (2010).
22. Smith, J. G. *et al.* Genome-wide association studies of the PR interval in African Americans. *PLoS*

Genet. **7,** e1001304 (2011).
23. Petkowski, J. J. *et al.* NRMT2 is an N-terminal monomethylase that primes for its homologue NRMT1. *Biochem. J.* **456,** 453–62 (2013).
24. Bonsignore, L. A. *et al.* NRMT1 knockout mice exhibit phenotypes associated with impaired DNA repair and premature aging. *Mech. Ageing Dev.* **146–148,** 42–52 (2015).
25. Bonsignore, L. A., Butler, J. S., Klinge, C. M. & Schaner Tooley, C. E. Loss of the N-terminal methyltransferase NRMT1 increases sensitivity to DNA damage and promotes mammary oncogenesis. *Oncotarget* **6,** 12248–12263 (2015).
26. Orr, N. *et al.* A mutation in the atrial-specific myosin light chain gene (MYL4) causes familial atrial fibrillation. *Nat. Commun.* **7,** 11303 (2016).
27. Gudbjartsson, D. F. *et al.* Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47,** 435–444 (2015).
28. Shimizu, K. *et al.* SMAP, an Smg GDS-associating protein having arm repeats and phosphorylated by Src tyrosine kinase. *J. Biol. Chem.* **271,** 27013–7 (1996).
29. Shimizu, K., Shirataki, H., Honda, T., Minami, S. & Takai, Y. Complex formation of SMAP/KAP3, a KIF3A/B ATPase motor-associated protein, with a human chromosome-associated polypeptide. *J. Biol. Chem.* **273,** 6591–4 (1998).
30. Hirokawa, N. Stirring up development with the heterotrimeric kinesin KIF3. *Traffic* **1,** 29–34 (2000).
31. Rahmioglu, N. *et al.* Genome-wide enrichment analysis between endometriosis and obesity-related traits reveals novel susceptibility loci. *Hum. Mol. Genet.* **24,** 1185–99 (2015).
32. Landers, J. E. *et al.* Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 9004–9 (2009).
33. Gotoh, M. *et al.* Comprehensive exploration of novel chimeric transcripts in clear cell renal cell carcinomas using whole transcriptome analysis. *Genes. Chromosomes Cancer* **53,** 1018–32 (2014).
34. Telikicherla, D. *et al.* Overexpression of Kinesin Associated Protein 3 (KIFAP3) in Breast Cancer. *J. Proteomics Bioinform.* **5,** 122–126 (2012).
35. Choi, J. *et al.* Kinesin superfamily-associated protein 3 is preferentially expressed in glutamatergic neurons and contributes to the excitatory control of female puberty. *Endocrinology* **149,** 6146–56 (2008).
36. Satoh, A. *et al.* Characterization of human p33/41 (annexin IV), a Ca2+ dependent carbohydrate-binding protein with monoclonal anti-annexin IV antibodies, AS11 and AS17. *Biol. Pharm. Bull.* **20,** 224–9 (1997).
37. Yao, H., Sun, C., Hu, Z. & Wang, W. The role of annexin A4 in cancer. *Front. Biosci. (Landmark Ed.* **21,** 949–57 (2016).
38. Heinick, A. *et al.* Annexin A4 is a novel direct regulator of adenylyl cyclase type 5. *FASEB J.* **29,** fj.14-269837- (2015).
39. Matteo, R. G. & Moravec, C. S. Immunolocalization of annexins IV, V and VI in the failing and non-failing human heart. *Cardiovasc. Res.* **45,** 961–70 (2000).
40. Kimura, T. *et al.* Mouse germ cell-less as an essential component for nuclear integrity. *Mol. Cell. Biol.* **23,** 1304–1315 (2003).
41. Kleiman, S. E. *et al.* Reduced human germ cell-less (HGCL) expression in azoospermic men with severe germinal cell impairment. *J. Androl.* **24,** 670–5
42. Gjerstorff, M. F. *et al.* GAGE cancer-germline antigens are recruited to the nuclear envelope by germ cell-less (GCL). *PLoS One* **7,** e45819 (2012).
43. Fournier, A. *et al.* 1q12 chromosome translocations form aberrant heterochromatic foci

associated with changes in nuclear architecture and gene expression in B cell lymphoma. *EMBO Mol. Med.* **2,** 159–71 (2010).
44. Graser, S., Stierhof, Y.-D. & Nigg, E. A. Cep68 and Cep215 (Cdk5rap2) are required for centrosome cohesion. *J. Cell Sci.* **120,** 4321–31 (2007).
45. Man, X., Megraw, T. L. & Lim, Y. P. Cep68 can be regulated by Nek2 and SCF complex. *Eur. J. Cell Biol.* **94,** 162–72 (2015).
46. Kim, J.-H. *et al.* Genome-wide and follow-up studies identify CEP68 gene variants associated with risk of aspirin-intolerant asthma. *PLoS One* **5,** e13818 (2010).
47. Cornejo-García, J. A. *et al.* Variants of CEP68 gene are associated with acute urticaria/angioedema induced by multiple non-steroidal anti-inflammatory drugs. *PLoS One* **9,** e90966 (2014).
48. Bang, M. L. *et al.* The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circ. Res.* **89,** 1065–72 (2001).
49. Gregorio, C. C. *et al.* The NH2 terminus of titin spans the Z-disc: its interaction with a novel 19-kD ligand (T-cap) is required for sarcomeric integrity. *J. Cell Biol.* **143,** 1013–27 (1998).
50. Linke, W. A. & Granzier, H. A spring tale: new facts on titin elasticity. *Biophys. J.* **75,** 2613–4 (1998).
51. Siu, B. L. *et al.* Familial dilated cardiomyopathy locus maps to chromosome 2q31. *Circulation* **99,** 1022–6 (1999).
52. Roberts, A. M. *et al.* Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci. Transl. Med.* **7,** 270ra6 (2015).
53. Gerull, B. *et al.* Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nat. Genet.* **30,** 201–4 (2002).
54. Herman, D. S. *et al.* Truncations of titin causing dilated cardiomyopathy. *N. Engl. J. Med.* **366,** 619–28 (2012).
55. Akinrinade, O., Alastalo, T.-P. & Koskenvuo, J. W. Relevance of Truncating Titin Mutations in Dilated Cardiomyopathy. *Clin. Genet.* (2016). doi:10.1111/cge.12741
56. Akinrinade, O., Koskenvuo, J. W. & Alastalo, T.-P. Prevalence of Titin Truncating Variants in General Population. *PLoS One* **10,** e0145284 (2015).
57. Marroni, F. *et al.* A genome-wide association scan of RR and QT interval duration in 3 European genetically isolated populations: the EUROSPAN project. *Circ. Cardiovasc. Genet.* **2,** 322–8 (2009).
58. Li, N. *et al.* Ablation of a Ca2+-activated K+ channel (SK2 channel) results in action potential prolongation in atrial myocytes and atrial fibrillation. *J. Physiol.* **587,** 1087–100 (2009).
59. Yu, T. *et al.* Decreased expression of small-conductance Ca2+-activated K+ channels SK1 and SK2 in human chronic atrial fibrillation. *Life Sci.* **90,** 219–227 (2012).
60. Parajuli, N. *et al.* Determinants of ventricular arrhythmias in human explanted hearts with dilated cardiomyopathy. *Eur. J. Clin. Invest.* **45,** 1286–96 (2015).
61. Yu, C.-C. *et al.* Small Conductance Calcium-Activated Potassium Current Is Important in Transmural Repolarization of Failing Human Ventricles. *Circ. Arrhythmia Electrophysiol.* **8,** 667–676 (2015).
62. Terentyev, D. *et al.* Sarcoplasmic reticulum Ca2+ release is both necessary and sufficient for SK channel activation in ventricular myocytes. *AJP Hear. Circ. Physiol.* **306,** H738–H746 (2014).
63. Gui, L. *et al.* Ventricular tachyarrhythmias in rats with acute myocardial infarction involves activation of small-conductance Ca2+-activated K+ channels. *Am. J. Physiol. Heart Circ. Physiol.* **304,** H118-30 (2013).
64. Chang, P.-C. *et al.* Heterogeneous upregulation of apamin-sensitive potassium currents in failing human ventricles. *J. Am. Heart Assoc.* **2,** e004713 (2013).

65. Mu, Y.-H. *et al.* RyR2 modulates a Ca2+-activated K+ current in mouse cardiac myocytes. *PLoS One* **9,** e94905 (2014).
66. Turker, I. *et al.* Amiodarone inhibits apamin-sensitive potassium currents. *PLoS One* **8,** e70450 (2013).
67. Kim, J.-J. *et al.* Identification of KCNN2 as a susceptibility locus for coronary artery aneurysms in Kawasaki disease using genome-wide association analysis. *J. Hum. Genet.* **58,** 521–5 (2013).
68. Lee, J.-K. *et al.* Consortium-Based Genetic Studies of Kawasaki Disease in Korea: Korean Kawasaki Disease Genetics Consortium. *Korean Circ. J.* **45,** 443–8 (2015).
69. Allen, D. *et al.* SK2 channels are neuroprotective for ischemia-induced neuronal cell death. *J. Cereb. Blood Flow Metab.* **31,** 2302–12 (2011).
70. Orfila, J. E. *et al.* Increasing small conductance Ca^{2+} -activated potassium channel activity reverses ischemia-induced impairment of long-term potentiation. *Eur. J. Neurosci.* **40,** 3179–3188 (2014).
71. McKay, B. M. *et al.* Increasing SK2 channel activity impairs associative learning. *J. Neurophysiol.* **108,** 863–70 (2012).
72. Ohtsuki, G., Piochon, C., Adelman, J. P. P. & Hansel, C. SK2 channel modulation contributes to compartment-specific dendritic plasticity in cerebellar Purkinje cells. *Neuron* **75,** 108–120 (2012).
73. Sun, J. *et al.* UBE3A Regulates Synaptic Plasticity and Learning and Memory by Controlling SK2 Channel Endocytosis. *Cell Rep.* **12,** 449–461 (2015).
74. Willis, M. *et al.* Small-conductance calcium-activated potassium type 2 channels (SK2, KCa2.2) in human brain. *Brain Struct. Funct.* (2016). doi:10.1007/s00429-016-1258-1
75. Cadet, J. L. *et al.* Genome-wide DNA hydroxymethylation identifies potassium channels in the nucleus accumbens as discriminators of methamphetamine addiction and abstinence. *Mol. Psychiatry* (2016). doi:10.1038/mp.2016.48
76. Fakira, A. K., Portugal, G. S., Carusillo, B., Melyan, Z. & Morón, J. A. Increased Small Conductance Calcium-Activated Potassium Type 2 Channel-Mediated Negative Feedback on N-methyl-D-aspartate Receptors Impairs Synaptic Plasticity Following Context-Dependent Sensitization to Morphine. *Biol. Psychiatry* **75,** 105–114 (2014).
77. Tatsuki, F. *et al.* Involvement of Ca2+-Dependent Hyperpolarization in Sleep Duration in Mammals. *Neuron* **90,** 70–85 (2016).
78. Kim, S. H. *et al.* Electrogenic transport and K(+) ion channel expression by the human endolymphatic sac epithelium. *Sci. Rep.* **5,** 18110 (2015).
79. Dolga, A. M. *et al.* Subcellular expression and neuroprotective effects of SK channels in human dopaminergic neurons. *Cell Death Dis.* **5,** e999 (2014).
80. Xiao, Y. *et al.* Overexpression of Trpp5 contributes to cell proliferation and apoptosis probably through involving calcium homeostasis. *Mol. Cell. Biochem.* **339,** 155–61 (2010).
81. Guo, L. *et al.* Identification and characterization of a novel polycystin family member, polycystin-L2, in mouse and human: sequence, expression, alternative splicing, and chromosomal localization. *Genomics* **64,** 241–51 (2000).
82. Volk, T., Schwoerer, A. P., Thiessen, S., Schultz, J.-H. & Ehmke, H. A polycystin-2-like large conductance cation channel in rat left ventricular myocytes. *Cardiovasc. Res.* **58,** 76–88 (2003).
83. Ferdaus, M. Z. & McCormick, J. A. The CUL3/KLHL3-WNK-SPAK/OSR1 pathway as a target for antihypertensive therapy. *Am. J. Physiol. Renal Physiol.* **310,** F1389-96 (2016).
84. Schumacher, F.-R. *et al.* Characterisation of the Cullin-3 mutation that causes a severe form of familial hypertension and hyperkalaemia. *EMBO Mol. Med.* **7,** 1285–306 (2015).
85. Shibata, S., Zhang, J., Puthumana, J., Stone, K. L. & Lifton, R. P. Kelch-like 3 and Cullin 3 regulate electrolyte homeostasis via ubiquitination and degradation of WNK4. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 7838–43 (2013).

86. Boyden, L. M. *et al.* Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities. *Nature* **482,** 98–102 (2012).
87. Glover, M. *et al.* Detection of mutations in KLHL3 and CUL3 in families with FHHt (familial hyperkalaemic hypertension or Gordon's syndrome). *Clin. Sci. (Lond).* **126,** 721–6 (2014).
88. Louis-Dit-Picard, H. *et al.* KLHL3 mutations cause familial hyperkalemic hypertension by impairing ion transport in the distal nephron. *Nat. Genet.* **44,** 456–60, S1-3 (2012).
89. Saitoh, T. & Katoh, M. Molecular cloning and characterization of human WNT8A. *Int. J. Oncol.* **19,** 123–7 (2001).
90. Cunningham, T. J., Kumar, S., Yamaguchi, T. P. & Duester, G. Wnt8a and Wnt3a cooperate in the axial stem cell niche to promote mammalian body axis extension. *Dev. Dyn.* **244,** 797–807 (2015).
91. Ma, Y. *et al.* The Chromatin Remodeling Protein Bptf Promotes Posterior Neuroectodermal Fate by Enhancing Smad2-Activated wnt8a Expression. *J. Neurosci.* **35,** 8493–506 (2015).
92. Gao, H. *et al.* Polymorphisms and expression of the WNT8A gene in Hirschsprung's disease. *Int. J. Mol. Med.* **32,** 647–52 (2013).
93. Lozano-Velasco, E. *et al.* Pitx2 impairs calcium handling in a dose-dependent manner by modulating Wnt signalling. *Cardiovasc. Res.* **109,** 55–66 (2016).
94. Lai, F. *et al.* cDNA cloning and genomic structure of three genes localized to human chromosome band 5q31 encoding potential nuclear proteins. *Genomics* **70,** 123–30 (2000).
95. Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347,** 1260419 (2015).
96. GATT, S. ENZYMIC HYDROLYSIS AND SYNTHESIS OF CERAMIDES. *J. Biol. Chem.* **238,** 3131–3 (1963).
97. Gatt, S. Enzymatic hydrolysis of sphingolipids. I. Hydrolysis and synthesis of ceramides by an enzyme from rat brain. *J. Biol. Chem.* **241,** 3724–30 (1966).
98. Seelan, R. S. *et al.* Human acid ceramidase is overexpressed but not mutated in prostate cancer. *Genes. Chromosomes Cancer* **29,** 137–46 (2000).
99. Norris, J. S. *et al.* Combined therapeutic use of AdGFPFasL and small molecule inhibitors of ceramide metabolism in prostate and head and neck cancers: a status report. *Cancer Gene Ther.* **13,** 1045–51 (2006).
100. Musumarra, G., Barresi, V., Condorelli, D. F. & Scirè, S. A bioinformatic approach to the identification of candidate genes for the development of new cancer diagnostics. *Biol. Chem.* **384,** 321–7 (2003).
101. Saad, A. F. *et al.* The functional effects of acid ceramidase overexpression in prostate cancer progression and resistance to chemotherapy. *Cancer Biol. Ther.* **6,** 1455–60 (2007).
102. Selzner, M. *et al.* Induction of apoptotic cell death and prevention of tumor growth by ceramide analogues in metastatic human colon cancer. *Cancer Res.* **61,** 1233–40 (2001).
103. Beckham, T. H. *et al.* Acid ceramidase-mediated production of sphingosine 1-phosphate promotes prostate cancer invasion through upregulation of cathepsin B. *Int. J. Cancer* **131,** 2034–43 (2012).
104. Kus, G., Kabadere, S., Uyar, R. & Kutlu, H. M. Induction of apoptosis in prostate cancer cells by the novel ceramidase inhibitor ceranib-2. *In Vitro Cell. Dev. Biol. Anim.* **51,** 1056–63 (2015).
105. Zeidan, Y. H. *et al.* Molecular targeting of acid ceramidase: implications to cancer therapy. *Curr. Drug Targets* **9,** 653–61 (2008).
106. Frohbergh, M., He, X. & Schuchman, E. H. The molecular medicine of acid ceramidase. *Biol. Chem.* **396,** 759–65 (2015).
107. Koch, J. *et al.* Molecular cloning and characterization of a full-length complementary DNA encoding human acid ceramidase. Identification Of the first molecular lesion causing Farber disease. *J. Biol. Chem.* **271,** 33110–5 (1996).

108. Rubboli, G. *et al.* Spinal muscular atrophy associated with progressive myoclonic epilepsy: A rare condition caused by mutations in ASAH1. *Epilepsia* **56,** 692–8 (2015).
109. Huang, Y. *et al.* Elevation of the level and activity of acid ceramidase in Alzheimer's disease brain. *Eur. J. Neurosci.* **20,** 3489–97 (2004).
110. Li, C.-M. M. *et al.* The human acid ceramidase gene (ASAH): structure, chromosomal location, mutation analysis, and expression. *Genomics* **62,** 223–31 (1999).
111. Baranowski, M., Blachnio, A., Zabielski, P. & Gorski, J. Pioglitazone induces de novo ceramide synthesis in the rat heart. *Prostaglandins Other Lipid Mediat.* **83,** 99–111 (2007).
112. Monette, J. S. *et al.* (R)-α-Lipoic acid treatment restores ceramide balance in aging rat cardiac mitochondria. *Pharmacol. Res.* **63,** 23–29 (2011).
113. Wang, L., Lee, K., Malonis, R., Sanchez, I. & Dynlacht, B. D. Tethering of an E3 ligase by PCM1 regulates the abundance of centrosomal KIAA0586/Talpid3 and promotes ciliogenesis. *Elife* **5,** (2016).
114. Zhang, W. *et al.* MiRNA-128 regulates the proliferation and neurogenesis of neural precursors by targeting PCM1 in the developing cortex. *Elife* **5,** (2016).
115. Farina, F. *et al.* The centrosome is an actin-organizing centre. *Nat. Cell Biol.* **18,** 65–75 (2015).
116. Schwaab, J. *et al.* Limited duration of complete remission on ruxolitinib in myeloid neoplasms with PCM1-JAK2 and BCR-JAK2 fusion genes. *Ann. Hematol.* **94,** 233–8 (2015).
117. Sakamoto, S. *et al.* Four polymorphisms of the pericentriolar material 1 (PCM1) gene are not associated with schizophrenia in a Japanese population. *Psychiatry Research* **216,** 288–289 (2014).
118. Stylli, S. S. *et al.* Expression of the adaptor protein Tks5 in human cancer: prognostic potential. *Oncol. Rep.* **32,** 989–1002 (2014).
119. Burger, K. L. *et al.* Src-dependent Tks5 phosphorylation regulates invadopodia-associated invasion in prostate cancer cells. *Prostate* **74,** 134–48 (2014).
120. Blouw, B. *et al.* The invadopodia scaffold protein Tks5 is required for the growth of human breast cancer cells in vitro and in vivo. *PLoS One* **10,** e0121003 (2015).
121. Oikawa, T. *et al.* Tks5-dependent formation of circumferential podosomes/invadopodia mediates cell-cell fusion. *J. Cell Biol.* **197,** 553–68 (2012).
122. Stylli, S. S., I, S. T. T., Kaye, A. H. & Lock, P. *Prognostic significance of Tks5 expression in gliomas*. *Journal of Clinical Neuroscience* **19,** (2012).
123. Wang, F., Chang, J. T.-H., Kao, C. J. & Huang, R. S. High Expression of miR-532-5p, a Tumor Suppressor, Leads to Better Prognosis in Ovarian Cancer Both In Vivo and In Vitro. *Mol. Cancer Ther.* **15,** 1123–31 (2016).
124. Blouw, B., Seals, D. F., Pass, I., Diaz, B. & Courtneidge, S. A. A role for the podosome/invadopodia scaffold protein Tks5 in tumor growth in vivo. *Eur. J. Cell Biol.* **87,** 555–567 (2008).
125. Murphy, D. A. *et al.* A Src-Tks5 pathway is required for neural crest cell migration during embryonic development. *PLoS One* **6,** e22499 (2011).
126. Cejudo-Martin, P. *et al.* Genetic disruption of the sh3pxd2a gene reveals an essential role in mouse development and the existence of a novel isoform of tks5. *PLoS One* **9,** e107674 (2014).
127. Burger, K. L., Davis, A. L., Isom, S., Mishra, N. & Seals, D. F. The podosome marker protein Tks5 regulates macrophage invasive behavior. *Cytoskeleton (Hoboken).* **68,** 694–711 (2011).
128. Vincent, C., Siddiqui, T. A. & Schlichter, L. C. Podosomes in migrating microglia: components and matrix degradation. *J. Neuroinflammation* **9,** 190 (2012).
129. Mesirca, P. *et al.* The G-protein-gated K+ channel, IKACh, is required for regulation of pacemaker activity and recovery of resting heart rate after sympathetic stimulation. *J. Gen. Physiol.* **142,** 113–26 (2013).
130. Mesirca, P. *et al.* Cardiac arrhythmia induced by genetic silencing of 'funny' (f) channels is

rescued by GIRK4 inactivation. *Nat. Commun.* **5,** 4664 (2014).
131. Bingen, B. O. *et al.* Atrium-Specific Kir3.x determines inducibility, dynamics, and termination of fibrillation by regulating restitution-driven alternans. *Circulation* **128,** 2732–2744 (2013).
132. Jabbari, J. *et al.* Common polymorphisms in KNCJ5 are associated with early-onset lone atrial fibrillation in Caucasians. *Cardiology* **118,** 116–120 (2011).
133. Wang, F. *et al.* The phenotype characteristics of type 13 long QT syndrome with mutation in KCNJ5 (Kir3.4-G387R). *Hear. Rhythm* **10,** 1500–1506 (2013).
134. Liang, B. *et al.* G-protein-coupled inward rectifier potassium current contributes to ventricular repolarization. *Cardiovasc. Res.* **101,** 175–84 (2014).
135. Molina-Navarro, M. M. *et al.* Differential gene expression of cardiac ion channels in human dilated cardiomyopathy. *PLoS One* **8,** e79792 (2013).
136. Kokunai, Y. *et al.* A Kir3.4 mutation causes Andersen-Tawil syndrome by an inhibitory effect on Kir2.1. *Neurology* **82,** 1058–64 (2014).
137. Azizan, E. A. B. & Brown, M. J. Novel genetic determinants of adrenal aldosterone regulation. *Curr. Opin. Endocrinol. Diabetes. Obes.* **23,** 209–17 (2016).
138. Chen, A. X., Nishimoto, K., Nanba, K. & Rainey, W. E. Potassium channels related to primary aldosteronism: Expression similarities and differences between human and rat adrenals. *Mol. Cell. Endocrinol.* **417,** 141–148 (2015).
139. Gomez, L. *et al.* Association of the KCNJ5 gene with Tourette Syndrome and Attention-Deficit/Hyperactivity Disorder. *Genes. Brain. Behav.* **13,** 535–42 (2014).
140. Han, C., Huang, J. & Waxman, S. G. Sodium channel Nav1.8: Emerging links to human disease. *Neurology* **86,** 473–83 (2016).
141. Yang, T. *et al.* Blocking Scn10a channels in heart reduces late sodium current and is antiarrhythmic. *Circ. Res.* **111,** 322–332 (2012).
142. Verkerk, A. O. *et al.* Functional NaV1.8 channels in intracardiac neurons: The link between SCN10A and cardiac electrophysiology. *Circ. Res.* **111,** 333–343 (2012).
143. Chambers, J. C. *et al.* Genetic variation in SCN10A influences cardiac conduction. *Nat. Genet.* **42,** 149–152 (2010).
144. Denny, J. C. *et al.* Identification of genomic predictors of atrioventricular conduction: using electronic medical records as a tool for genome science. *Circulation* **122,** 2016–21 (2010).
145. Savio-Galimberti, E. *et al.* SCN10A/Nav1.8 modulation of peak and late sodium currents in patients with early onset atrial fibrillation. *Cardiovasc. Res.* **104,** 355–63 (2014).
146. Bezzina, C. R. *et al.* Common variants at SCN5A-SCN10A and HEY2 are associated with Brugada syndrome, a rare disease with high risk of sudden cardiac death. *Nat. Genet.* **45,** 1044–1049 (2013).
147. Hu, D. *et al.* Mutations in SCN10A are responsible for a large fraction of cases of Brugada syndrome. *J. Am. Coll. Cardiol.* **64,** 66–79 (2014).
148. Park, D. S. & Fishman, G. I. Nav-igating through a complex landscape: SCN10A and cardiac conduction. *J. Clin. Invest.* **124,** 1460–2 (2014).
149. van den Boogaard, M. *et al.* A common genetic variant within SCN10A modulates cardiac SCN5A expression. *J. Clin. Invest.* **124,** 1844–1852 (2014).
150. Aza-Carmona, M. *et al.* NPPB and ACAN, two novel SHOX2 transcription targets implicated in skeletal development. *PLoS One* **9,** e83104 (2014).
151. Liu, C.-F. & Lefebvre, V. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. *Nucleic Acids Res.* **43,** 8183–203 (2015).
152. Baroti, T. *et al.* Transcription factors Sox5 and Sox6 exert direct and indirect influences on oligodendroglial migration in spinal cord and forebrain. *Glia* **64,** 122–38 (2016).
153. Hersh, C. P. *et al.* SOX5 is a candidate gene for chronic obstructive pulmonary disease

susceptibility and is necessary for lung development. *Am. J. Respir. Crit. Care Med.* **183,** 1482–9 (2011).
154. Olesen, M. S. *et al.* Genetic loci on chromosomes 4q25, 7p31, and 12p12 are associated with onset of lone atrial fibrillation before the age of 40 years. *Can. J. Cardiol.* **28,** 191–5 (2012).
155. Della-Morte, D. *et al.* A follow-up study for left ventricular mass on chromosome 12p11 identifies potential candidate genes. *BMC Med. Genet.* **12,** 100 (2011).
156. Wen, Y. *et al.* Integrative analysis of genome-wide association studies and gene expression profiles identified candidate genes for osteoporosis in Kashin-Beck disease patients. *Osteoporos. Int.* **27,** 1041–6 (2016).
157. Jin, J., Chou, C., Lima, M., Zhou, D. & Zhou, X. Systemic Sclerosis is a Complex Disease Associated Mainly with Immune Regulatory and Inflammatory Genes. *Open Rheumatol. J.* **8,** 29–42 (2014).
158. Le Clerc, S. *et al.* Genomewide association study of a rapid progression cohort identifies new susceptibility alleles for AIDS (ANRS Genomewide Association Study 03). *J. Infect. Dis.* **200,** 1194–201 (2009).
159. Tu, W. *et al.* Genome-Wide Loci Linked to Non-Obstructive Azoospermia Susceptibility May Be Independent of Reduced Sperm Production in Males with Normozoospermia. *Biol. Reprod.* **92,** 41–41 (2015).
160. Al Zeyadi, M. *et al.* Whole genome microarray analysis in non-small cell lung cancer. *Biotechnol. Biotechnol. Equip.* **29,** 111–118 (2015).
161. Wang, D., Han, S., Wang, X., Peng, R. & Li, X. SOX5 promotes epithelial-mesenchymal transition and cell invasion via regulation of Twist1 in hepatocellular carcinoma. *Med. Oncol.* **32,** 461 (2015).
162. Shiseki, M. *et al.* Identification of the SOX5 gene as a novel IGH-involved translocation partner in BCL2-negative follicular lymphoma with t(12;14)(p12.2;q32). *Int. J. Hematol.* **102,** 633–8 (2015).
163. Kordaß, T. *et al.* SOX5 is involved in balanced MITF regulation in human melanoma cells. *BMC Med. Genomics* **9,** 10 (2016).
164. Kranias, E. G. & Hajjar, R. J. Modulation of cardiac contractility by the phospholamban/SERCA2a regulatome. *Circ. Res.* **110,** 1646–60 (2012).
165. Minamisawa, S. *et al.* Mutation of the phospholamban promoter associated with hypertrophic cardiomyopathy. *Biochem. Biophys. Res. Commun.* **304,** 1–4 (2003).
166. Landstrom, A. P., Adekola, B. A., Bos, J. M., Ommen, S. R. & Ackerman, M. J. PLN-encoded phospholamban mutation in a large cohort of hypertrophic cardiomyopathy cases: summary of the literature and implications for genetic testing. *Am. Heart J.* **161,** 165–71 (2011).
167. Schmitt, J. P. *et al.* Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science* **299,** 1410–3 (2003).
168. van Spaendonck-Zwarts, K. Y. *et al.* Genetic analysis in 418 index patients with idiopathic dilated cardiomyopathy: overview of 10 years' experience. *Eur. J. Heart Fail.* **15,** 628–36 (2013).
169. Ellinor, P. T. *et al.* Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat. Genet.* **44,** 670–5 (2012).
170. Harris, T. B. *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am. J. Epidemiol.* **165,** 1076–87 (2007).
171. Raitoharju, E. *et al.* Common variation in the ADAM8 gene affects serum sADAM8 concentrations and the risk of myocardial infarction in two independent cohorts. *Atherosclerosis* **218,** 127–133 (2011).
172. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am. J. Epidemiol.* **129,** 687–702 (1989).
173. Alonso, A. *et al.* Incidence of atrial fibrillation in whites and African-Americans: the Atherosclerosis Risk in Communities (ARIC) study. *Am Hear. J* **158,** 111–117 (2009).

174. Conen, D. *et al.* Genetic and phenotypic determinants of blood pressure and other cardiovascular risk factors (GAPP). *Swiss Med. Wkly.* **143,** w13728 (2013).
175. Sinner, M. F. *et al.* Integrating genetic, transcriptional, and functional analyses to identify 5 novel genes for atrial fibrillation. *Circulation* **130,** 1225–35 (2014).
176. Locke, A. E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518,** 197–206 (2015).
177. Weeke, P. *et al.* Examining rare and low-frequency genetic variants previously associated with lone or familial forms of atrial fibrillation in an electronic medical record system: a cautionary note. *Circ. Cardiovasc. Genet.* **8,** 58–63 (2015).
178. Fried, L. P. *et al.* The cardiovascular health study: Design and rationale. *Ann. Epidemiol.* **1,** 263–276 (1991).
179. Vaara, S. *et al.* Cohort Profile: the Corogene study. *Int. J. Epidemiol.* **41,** 1265–71 (2012).
180. Dawber, T. R., Meadors, G. F. & Moore, F. E. Epidemiological approaches to heart disease: the Framingham Study. *Am. J. Public Health Nations. Health* **41,** 279–81 (1951).
181. Kannel, W. B., Feinleib, M., McNamara, P. M., Garrison, R. J. & Castelli, W. P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am. J. Epidemiol.* **110,** 281–90 (1979).
182. Nieminen, T. *et al.* The Finnish Cardiovascular Study (FINCAVAS): characterising patients with high risk of cardiovascular morbidity and mortality. *BMC Cardiovasc. Disord.* **6,** 9 (2006).
183. Smith, B. H. *et al.* Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *Int. J. Epidemiol.* **42,** 689–700 (2013).
184. Schmermund, A. *et al.* Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: Rationale and design of the Heinz Nixdorf RECALL Study. *Am. Heart J.* **144,** 212–218 (2002).
185. Winkelmann, B. R. *et al.* Rationale and design of the LURIC study--a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. *Pharmacogenomics* **2,** S1-73 (2001).
186. 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).
187. Bild, D. E. *et al.* Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am. J. Epidemiol.* **156,** 871–81 (2002).
188. Rasmussen-Torvik, L. J. *et al.* Fasting glucose GWAS candidate region analysis across ethnic groups in the Multiethnic Study of Atherosclerosis (MESA). *Genet. Epidemiol.* **36,** 384–91 (2012).
189. Genes for Cerebral Hemorrhage on Anticoagulation (GOCHA) Collaborative Group. Exploiting common genetic variation to make anticoagulation safer. *Stroke.* **40,** S64-6 (2009).
190. International Stroke Genetics Consortium (ISGC) *et al.* Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat. Genet.* **44,** 328–33 (2012).
191. Lind, L., Fors, N., Hall, J., Marttala, K. & Stenborg, A. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. *Arterioscler. Thromb. Vasc. Biol.* **25,** 2368–75 (2005).
192. Hillege, H. L. *et al.* Urinary albumin excretion predicts cardiovascular and noncardiovascular mortality in general population. *Circulation* **106,** 1777–82 (2002).
193. Shepherd, J. *et al.* Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet* **360,** 1623–30 (2002).

194. Hofman, A. *et al.* The Rotterdam Study: 2016 objectives and design update. *Eur. J. Epidemiol.* **30,** 661–708 (2015).
195. Gioli-Pereira, L. *et al.* Genetic and ElectroNic medIcal records to predict oUtcomeS in Heart Failure patients (GENIUS-HF) - design and rationale. *BMC Cardiovasc. Disord.* **14,** 32 (2014).
196. Völzke, H. *et al.* Cohort profile: the study of health in Pomerania. *Int. J. Epidemiol.* **40,** 294–307 (2011).
197. Hong, Y., Pedersen, N. L., Brismar, K. & de Faire, U. Genetic and environmental architecture of the features of the insulin-resistance syndrome. *Am. J. Hum. Genet.* **60,** 143–52 (1997).
198. Ingelsson, E., Sundström, J., Arnlöv, J., Zethelius, B. & Lind, L. Insulin resistance and risk of congestive heart failure. *JAMA* **294,** 334–41 (2005).
199. Ridker, P. M. *et al.* Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin. Chem.* **54,** 249–55 (2008).
200. Grove, M. L. *et al.* Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PLoS One* **8,** e68095 (2013).
201. Holle, R., Happich, M., Löwel, H., Wichmann, H. E. & MONICA/KORA Study Group. KORA--a research platform for population based health research. *Gesundheitswes. (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes* **67 Suppl 1,** S19-25 (2005).
202. Wichmann, H.-E., Gieger, C., Illig, T. & MONICA/KORA Study Group. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1,** S26-30 (2005).
203. Paynter, N. P. *et al.* Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann. Intern. Med.* **150,** 65–72 (2009).
204. Kuriyama, S. *et al.* The Tohoku Medical Megabank Project: Design and Mission. *J. Epidemiol.* **26,** 493–511 (2016).
205. The GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45,** 580–5 (2013).
206. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22,** 1790–1797 (2012).
207. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **40,** D930-4 (2012).
208. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29,** 308–11 (2001).
209. Arnold, M., Raffler, J., Pfeufer, A., Suhre, K. & Kastenmüller, G. SNiPA: An interactive, genetic variant-centered annotation browser. *Bioinformatics* **31,** 1334–1336 (2015).
210. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6,** 5890 (2015).
211. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–75 (2007).
212. 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).
213. Wei, W.-Q. *et al.* Impact of data fragmentation across healthcare centers on the accuracy of a high-throughput clinical phenotyping algorithm for specifying subjects with type 2 diabetes mellitus. *J. Am. Med. Inform. Assoc.* **19,** 219–24
214. 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).
215. 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).
216. Vermond, R. A. *et al.* Incidence of Atrial Fibrillation and Relationship With Cardiovascular Events, Heart Failure, and Mortality. *J. Am. Coll. Cardiol.* **66,** 1000–1007 (2015).
217. Magnusson, P. K. E. *et al.* The Swedish Twin Registry: establishment of a biobank and other recent developments. *Twin Res. Hum. Genet.* **16,** 317–29 (2013).
218. Hays, J. *et al.* The women's health initiative recruitment methods and results. *Ann. Epidemiol.* **13,** S18–S77 (2003).
219. Anderson, G. L. *et al.* Implementation of the Women's Health Initiative study design. *Ann. Epidemiol.* **13,** S5–S17 (2003).
220. Gerull, B. *et al.* Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy. *Nat. Genet.* **36,** 1162–1164 (2004).
221. Cerrone, M. *et al.* Missense mutations in plakophilin-2 cause sodium current deficit and associate with a brugada syndrome phenotype. *Circulation* **129,** 1092–1103 (2014).
222. Peters, S. Arrhythmogenic cardiomyopathy and provocable Brugada ECG in a patient caused by missense mutation in plakophilin-2. *Int. J. Cardiol.* **173,** 317–8 (2014).