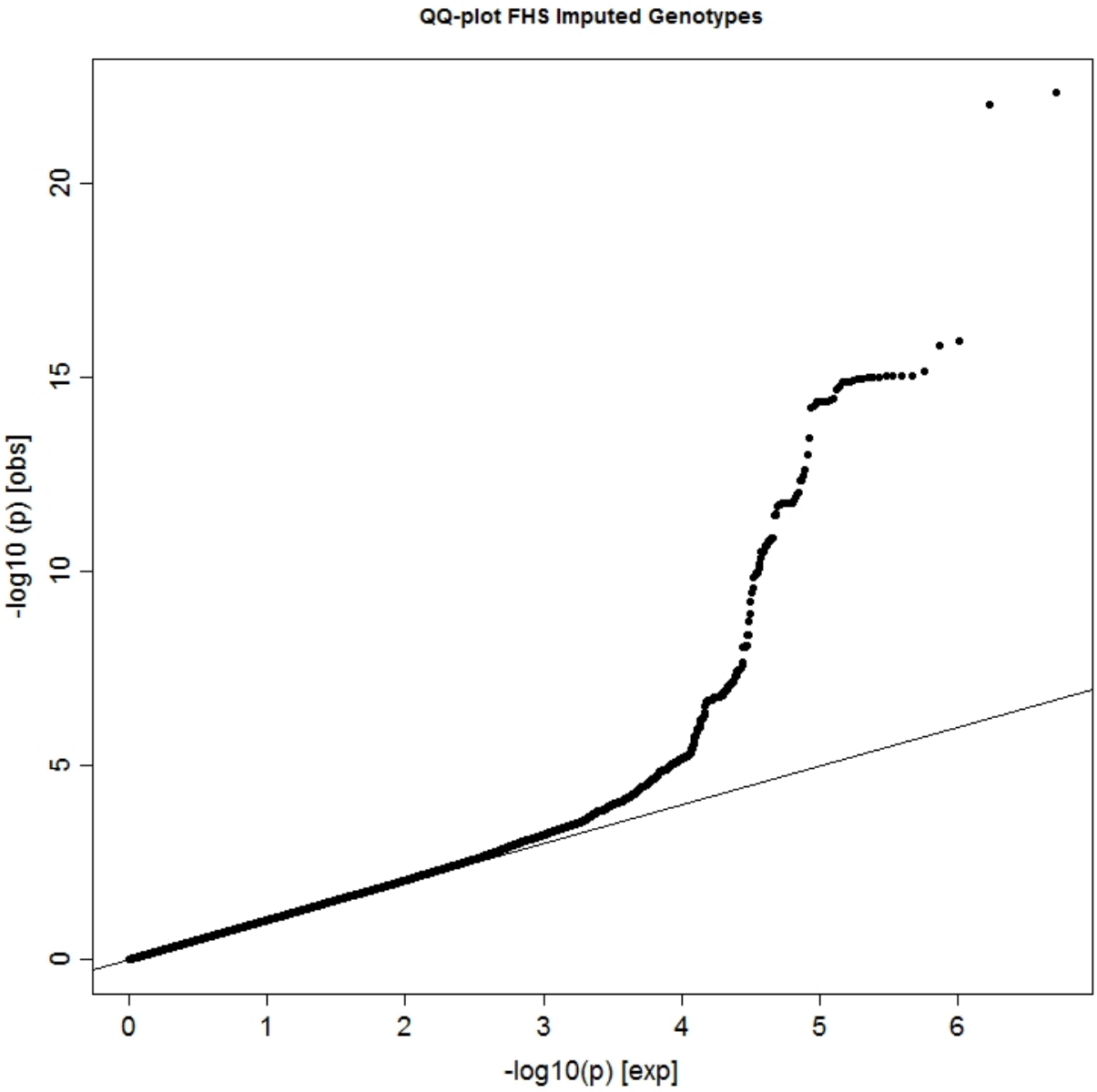
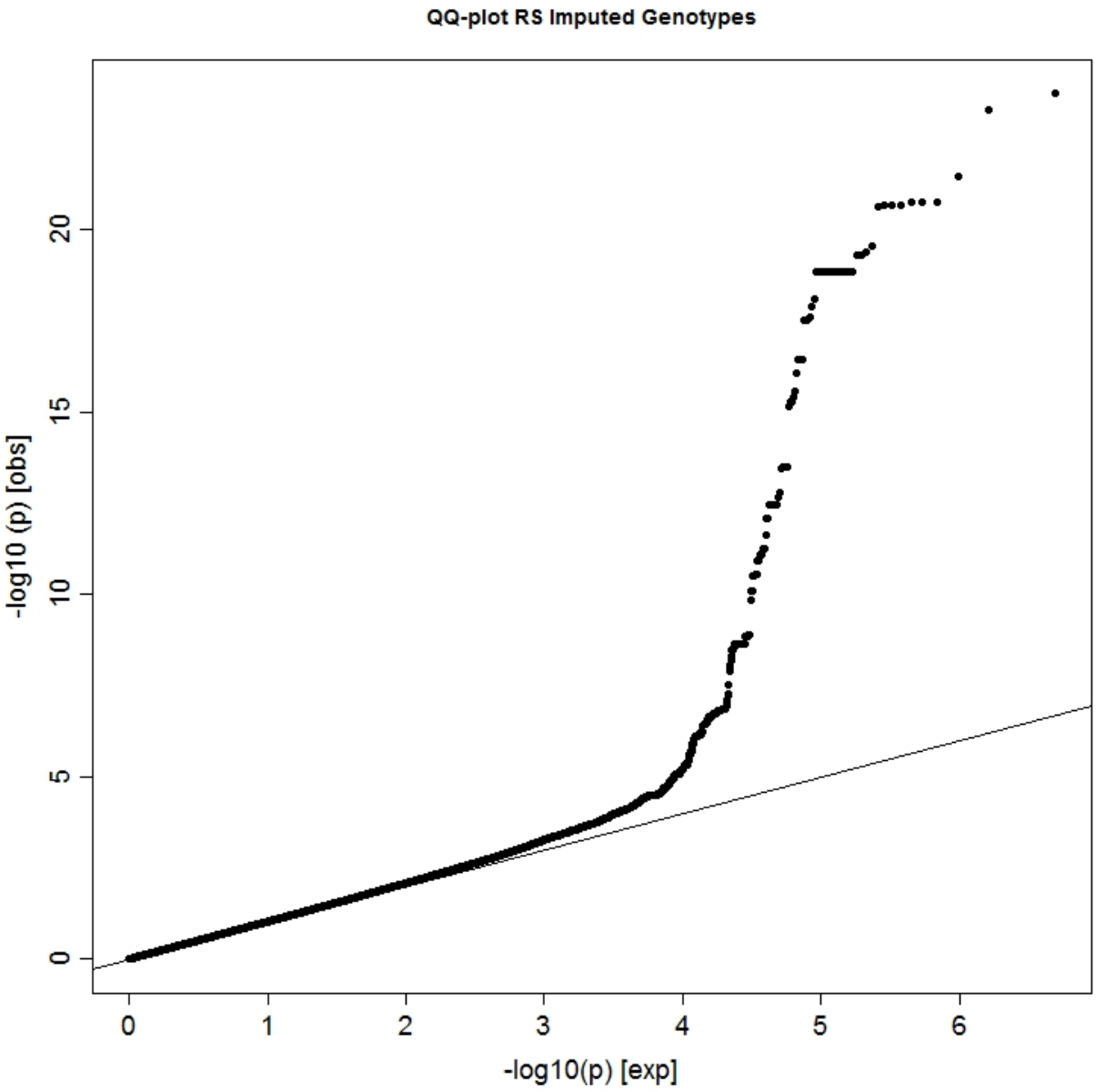
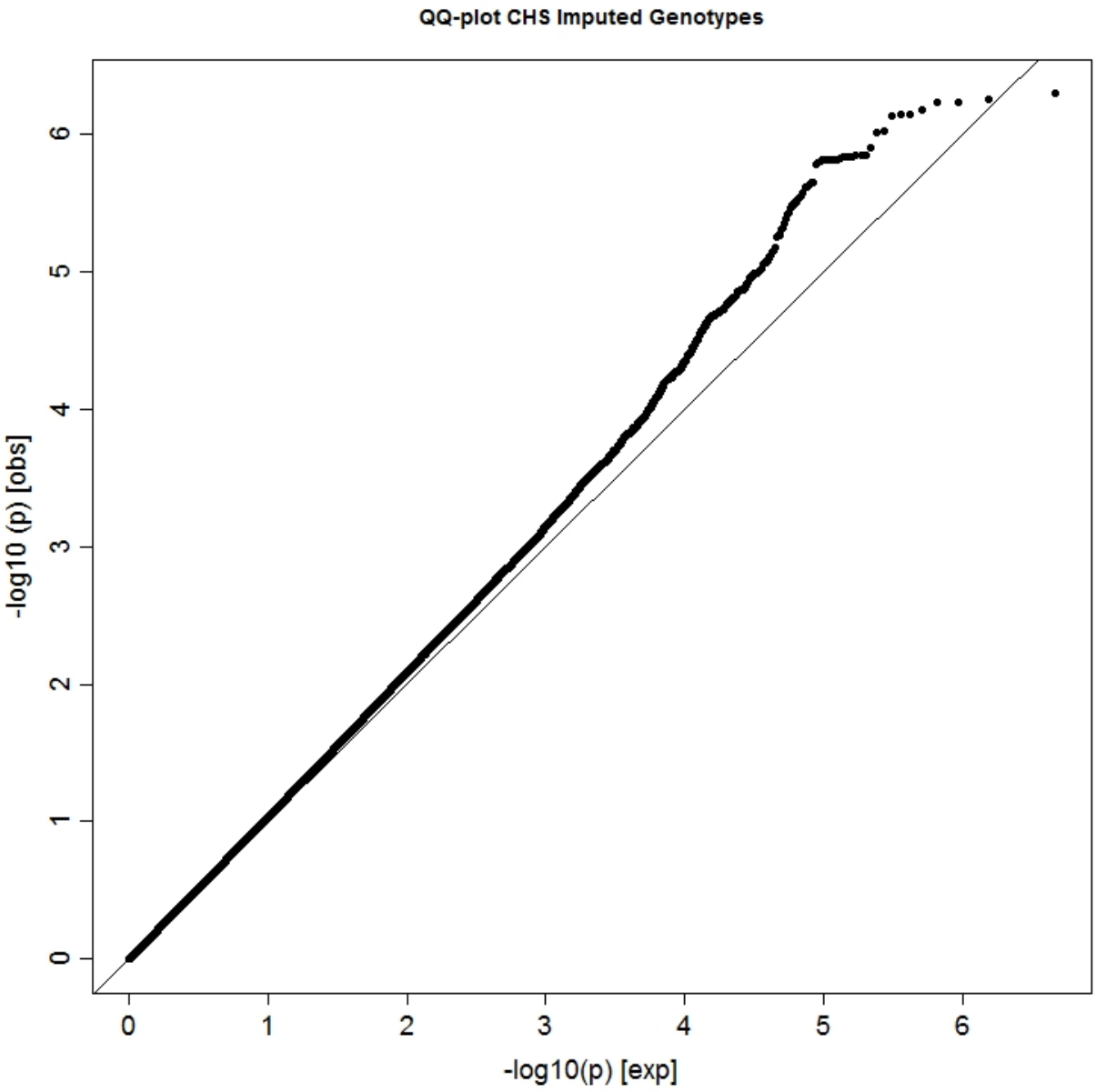
Common variants at ten loci influence myocardial repolarization: the QTGEN consortium

Christopher Newton-Cheh, Mark Eijgelsheim, Kenneth Rice, Paul I. W. de Bakker, Xiaoyan Yin, Karol Estrada, Joshua Bis, Kristin Marciante, Fernando Rivadeneira, Peter A. Noseworthy, Nona Sotoodehnia, Nicholas L. Smith, Jerome I. Rotter, Jan A. Kors, Jacqueline C.M. Witteman, Albert Hofman, Susan R. Heckbert, Christopher J. O'Donnell, André G. Uitterlinden, Bruce M. Psaty, Thomas Lumley, Martin G. Larson, Bruno H.Ch. Stricker

Supplementary Figure 1. Shown are quantile-quantile plots of -log(p-value) of QT interval association tests in Framingham Heart Study (FHS), Rotterdam Study (RS), and Cardiovascular Health Study (CHS).




QQ-plot RS Imputed Genotypes
-log10 (p) [obs]
0
5
10
15
20
0
1
2
3
4
5
6
-log10(p) [exp]


QQ-plot CHS Imputed Genotypes
-log10 (p) [obs]
-log10(p) [exp]
0
1
2
3
4
5
6

Supplementary Table 1. P-value distributions from meta-analysis of 2,543,686 imputed SNPs in 13,685 individuals, binned by $-\log_{10}(P)$. The ratio of observed to expected SNPs exceeding different thresholds is shown.

$-\log_{10}(P)$	<2	<3	<4	<5	<6	<7	<8
observed	27,702	4,260	1,197	568	459	346	304
expected	25,437	2,544	254	25	2.5	0.25	0.03
ratio	1.1	1.7	4.7	22	180	1,360	11,951

Supplementary Table 2. Cohort-specific association results in multivariable regression models including 14 SNPs at 10 loci.

	Framingham Heart Study			Rotterdam Study			Cardiovascular Health Study		
SNP	Beta (SD)	SE	P	Beta (SD)	SE	P	Beta (SD)	SE	P
rs12143842	0.15	0.02	$8x10^{-12}$	0.19	0.03	$9x10^{-14}$	0.17	0.07	$9x10^{-3}$
rs12029454	0.11	0.03	$4x10^{-5}$	0.08	0.03	$7x10^{-3}$	-0.02	0.14	$9x10^{-1}$
rs16857031	0.07	0.03	$4x10^{-3}$	0.17	0.03	$6x10^{-8}$	0.24	0.08	$3x10^{-3}$
rs2074238	-0.47	0.12	$6x10^{-5}$	-0.41	0.06	$3x10^{-10}$	-0.66	0.55	$3x10^{-1}$
rs37062	-0.11	0.02	$2x10^{-8}$	-0.13	0.02	$5x10^{-8}$	-0.13	0.05	$8x10^{-3}$
rs11756438	0.07	0.02	$4x10^{-5}$	0.11	0.02	$6x10^{-8}$	0.12	0.06	$6x10^{-2}$
rs12576239	0.10	0.03	$1x10^{-4}$	0.07	0.03	$3x10^{-2}$	0.12	0.08	$1x10^{-1}$
rs846111	0.13	0.03	$6x10^{-6}$	0.10	0.03	$3x10^{-4}$	0.25	0.08	$4x10^{-3}$
rs4725982	0.06	0.02	$4x10^{-3}$	0.08	0.03	$4x10^{-3}$	0.12	0.06	$5x10^{-2}$
rs8049607	0.10	0.02	$3x10^{-6}$	0.08	0.02	$1x10^{-4}$	-0.01	0.05	$9x10^{-1}$
rs1805128	0.47	0.08	$2x10^{-8}$	0.06	0.06	$3x10^{-1}$	0.79	0.84	$4x10^{-1}$
rs12053903	-0.09	0.02	$3x10^{-6}$	-0.05	0.02	$2x10^{-2}$	-0.12	0.05	$2x10^{-2}$
rs2074518	-0.06	0.02	$2x10^{-4}$	-0.06	0.02	$2x10^{-3}$	-0.09	0.06	$9x10^{-2}$
rs2968864	-0.08	0.02	$3x10^{-4}$	-0.03	0.02	$2x10^{-1}$	-0.12	0.05	$3x10^{-2}$
	Model r^2 = 5.4%			Model r^2 = 6.5%			Model r^2 = 2.3%		

Supplementary Table 3. Coverage statistics for novel loci. For purposes of coverage estimation, we defined the associated interval at each of the five novel loci to be the genomic span surrounding the top association signal(s) bounded by SNPs with r^2 to each independent signal ≥0.20, including all intervening SNPs. Shown are the genomic span of these loci, the mean imputation quality score (ranging from 0 to 1) and the proportion of SNPs in the interval captured at an imputation quality score >0.50 and >0.80. As can be seen, coverage at four of these novel loci is quite good (>90% of all SNPs captured at imputation quality >0.50) but the locus containing *RNF207* is less well covered.

Chr	Gene	Start	Stop	Genomic span (kb)	Mean	>0.50	>0.80
16q	*CNOT1*	57085908	57257853	172	0.89	0.93	0.86
6q	*c6orf204*	118634581	119134543	500	0.95	0.99	0.95
1p	*RNF207*	6201001	6245523	45	0.63	0.71	0.43
16p	*LITAF*	11574706	11631856	57	0.70	0.97	0.27
17q	*LIG3*	30084616	30492053	407	0.92	0.96	0.89

Supplementary Table 4. Quintiles of QT interval score and odds ratios for prolonged QTc. Shown are the odds ratios, 95% confidence intervals and the p-values for each quintile of QT interval score relative to the lowest quintile (q1).

	Rotterdam Study			Framingham Heart Study		
	OR	95% CI	*P*	OR	95% CI	*P*
q1	ref			ref		
q2	1.64	1.01-2.65	0.04	1.24	0.76-2.05	0.39
q3	1.95	1.22-3.11	0.005	1.91	1.21-3.03	0.006
q4	2.26	1.43-3.57	$5x10^{-4}$	1.52	0.94-2.46	0.09
q5	3.08	1.98-4.78	$6x10^{-7}$	2.55	1.64-3.98	$3x10^{-5}$

Supplementary Methods

QT measurement methods. In FHS, paper electrocardiograms recorded on Marquette machines were scanned and digital caliper measurements were made using proprietary software (eResearchTechnology, generations 1 and 2) or using Rigel 1.7.2. (AMPS, LLC, New York, NY, USA, generation 3). The QT interval was measured from the beginning of the QRS to the end of the T wave (or the nadir of the T-U when U waves were present) in two cardiac cycles from lead II, one cycle from lead V2 and one cycle from lead V5.[1] The correlation of a single cycle QT measure between two readers was 0.77 and the coefficient of variation was 2.6%, as previously published.[1] To increase precision, the average of measures from 4 cardiac cycles were used in analyses.

In the Rotterdam Study, electrocardiograms were measured on ACTA electrocardiographs (ESAOTE, Florence, Italy) and digital measurements of the QT interval were made using the Modular ECG Analysis System (MEANS).[2] The operation of the waveform recognition algorithms has been described and validated extensively.[2,3] The MEANS program determines common QRS onset and T offset for all 12 leads together on one representative averaged beat in a 10-second recording by use of template matching techniques. A common QRS onset and T offset are determined over all 12 leads in the representative complex. The presence of U waves or low T waves in individual leads then becomes less relevant. In the Common Standards for Quantitative Electrocardiography (CSE) study, in which different ECG computer programs were compared with a group of expert cardiologists, the measurement performance of MEANS ranked among the best.[3]

In the Cardiovascular Health Study, the electrocardiograms were recorded on MAC PC-DT ECG recorder (Marquette Electronics Inc, Milwaukee, WI, USA) machines and measurements of QT interval made using the Marquette 12SL algorithm, which measures the QT as a global interval measured from the median complex derived from the cardiac cycles occurring in 10 seconds. QT measures were compared to values generated using the Dalhousie Program by the EPICORE reading center,[4] which measures a global interval from a complex obtained with selective averaging of all normally conducted complexes. 12SL values with QT within 40 msec of the Dalhousie Program were accepted for use (97.4%). If the QT interval differed between the two programs by ≥40 msec (only 2.6% of ECGs), then the value closer to the median rate-corrected QT was used.

Genotyping. In FHS, genotyping was performed by Affymetrix (Santa Clara, CA, USA) using the Affymetrix 500K GeneChip array and a custom-designed gene-centric 50K MIP. Affymetrix 500K genotypes were called using the BRLMM algorithm[5] In FHS, the following exclusions were applied to exclude individuals with call rate ≤ 97%, per subject heterozygosity ± 5SD away from mean, or excess Mendelian errors resulting in 8,481 individuals with genotype regardless of phenotype and then to exclude SNPs with HWE $P < 10^{-6}$ (15,586), call rate ≤ 97% (64,511), mishap $P < 10^{-9}$ (45,361), Mendel errors >100 (4,857), minor allele frequency <0.01 (67,269), incompatible strand with HapMap genotypes (release 22, n = 2) and SNPs not present on HapMap (13,394), resulting in a set of 378,163 SNPs to be used in imputation.

In RS, genotyping was performed by the genetic laboratory of the Department of Internal Medicine, Erasmus Medical Center, Rotterdam using the Infinium II HumanHap550K Genotyping BeadChip version 3 (Illumina, San Diego, CA, USA). The Illumina 550K BeadChip array was genotyped in all participants of the original Rotterdam Study cohort with proper quality DNA samples (n = 6,449). Intensity files were analyzed using the BeadStudio Genotyping Module software v.3.1.14. A no-call threshold of 0.15 was applied to a custom-generated cluster file derived from the Illumina-provided cluster file (based on the cluster definitions applied to the HapMap CEPH cohort). In the custom-cluster file 2,308 SNPs with GenCall scores < 0.90 were visually checked by two observers and manually re-clustered or zeroed accordingly. Poorly performing samples with low call rate and 10th percentile GenCall score were excluded prior to calling genotypes. Any samples with a call rate below 97.5% (n = 209), excess autosomal heterozygosity > 0.336 ~FDR < 0.1% (n = 21), mismatch between called and phenotypic sex (n = 36), or if there were outliers identified by the IBS clustering analysis clustering > 3 standard deviations away from the population mean (n = 102) or IBS probabilities > 97% (n = 129) were excluded from the analysis. In total, 5,974 genotyped samples were available after exclusions.

In CHS, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV BeadChip system. Genotypes were called using the Illumina BeadStudio software as above. Samples were excluded from analysis for sex mismatch, discordance with prior genotyping, or call rate < 95%. SNPs were excluded from analysis for HWE $p < 10^{-5}$; SNPs with call

rates < 95% were manually reclustered using the Illumina software. All three studies make use of genotypes that were not specifically generated to examine the genetic basis of QT interval duration.

Associations of poorly imputed SNPs were validated by re-genotyping FHS samples at the Broad Institute using the Sequenom platform (San Diego, CA, USA) and RS samples at the Erasmus Medical Center using Taqman MGB platform using Assays-by-Design (Applied Biosystems, Foster City, CA, USA).

Imputation. In FHS, model parameters were estimated using MACH[6] v 1.0.15 (using flags --rounds 100, --greedy) in 200 unrelated FHS individuals, prioritized for high call rate (≥98.9%), low Mendel error rates and non-outlier status in EIGENSTRAT[7] principal components analysis. With these model parameters, we used MACH[6] to impute allele dosage, defined as the expected number of copies of the minor allele (a fractional value between 0 and 2), of all autosomal SNPs on HapMap CEU based on phased chromosomes of release 22, build 36.

In RS, the following exclusions were applied to identify 512,349 SNPs to be used for imputation: HWE $P < 10^{-6}$, call rate ≤ 98% and minor allele frequency < 0.01. In total 49,117 SNPs were excluded. For setting model parameters 200 random subjects were selected and used for every chromosome to estimate error and recombination rate (with flags --greedy, --rounds 100). Imputation was then performed to impute genotypes oriented to the positive strand of the human genome reference sequence for all autosomal SNPs in HapMap CEU using release 22, build 36. For each SNP in each individual, imputation results are reported as an allele dosage.

In CHS, the following exclusions were applied to identify a final set of 332,946 SNPs: call rate >95%, HWE $P > 10^{-5}$, 2 duplicate errors or Mendelian inconsistencies (for reference CEPH trios). Imputation was performed using BIMBAM[8] v0.95 with reference to HapMap CEU using release 21a, build 35 using one round of imputations and the default expectation-maximization warm-ups and runs. SNPs were excluded for variance on the allele dosage ≤0.01.

References

1. Newton-Cheh, C. et al. QT interval is a heritable quantitative trait with evidence of linkage to chromosome 3 in a genome-wide linkage analysis: The Framingham Heart Study. *Heart Rhythm.* **2**, 277-284 (2005).
2. van Bemmel, J.H., Kors, J.A. & van, H.G. Methodology of the modular ECG analysis system MEANS. *Methods Inf.Med.* **29**, 346-353 (1990).
3. Willems, J.L. et al. A reference data base for multilead electrocardiographic computer measurement programs. *J Am Coll Cardiol* **10**, 1313-21 (1987).
4. Fried, L.P. et al. The Cardiovascular Health Study: design and rationale. *Ann.Epidemiol.* **1**, 263-276 (1991).
5. Rabbee, N. & Speed, T.P. A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics.* **22**, 7-12 (2006).
6. Li, Y. & Abecasis, G.R. Mach 1.0: Rapid haplotype reconstruction and missing genotype inference. *Am J Hum.Genet.* **S79**, 2290 (2006).
7. Price, A.L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat.Genet.* **38**, 904-909 (2006).
8. Servin, B. & Stephens, M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS.Genet* **3**, e114- (2007).