Online Supplement

Genome-wide association analysis identifies multiple loci related with resting heart rate

Eijgelsheim et al.

Online Appendix Content:

I. Supplemental Methods

II. Supplemental Tables 1 – 3

III. Supplemental Figures 1 – 2

IV. Supplemental References

V. Full text funding statement and Acknowledgements

VI. Author contribution list

I. Supplemental Methods

Cohort descriptions

For the description of the CHARGE consortium and participating cohorts we refer to the CHARGE design paper(1). In short, the AGES – Reykjavik Study represents a sample from the population-based Reykjavik Study, originally comprising 31,795 individuals. Between 2002 and 2006 the AGES – Reykjavik Study 5,764 survivors of the original cohort were re-examined(2). The ARIC study is a population-based prospective cohort study, sponsored by the National Heart, Lung, and Blood Institute, and included 15,792 individuals aged 45 to 64 years at baseline from 4 US communities(3). CHS is a National Heart, Lung, and Blood Institute sponsored population-based cohort study in adults over 65 years of age conducted at 4 field centers. Originally CHS 5,201 predominantly white individuals were enrolled in 1989-1990(4). The FHS started in 1948 with the recruitment of an original cohort of 5,209 men and women aged 28 to 62 years of age(5). In 1971, children and spouses of children of the original cohort were enrolled(6). The third generation cohort enrollment started in 2002 and comprises 4,095 children of offspring cohort participants.(7) The RS is a prospective population-based cohort study comprising 7,983 subjects aged 55 years or older (RS-I)(8). In 2000-2001, an additional 3,011 individuals aged 55 years or older (RS-II) were recruited (9).

In addition to cohorts in the CHARGE consortium, additional cohorts participated and are discussed below. EUROSPAN consists of five isolated population cohorts. For the current effort we used ECG measurements that were available in three of the participating studies, namely ERF, MICROS and ORCADES. The studied populations originate from Rucphen, (The Netherlands), South Tyrol (Italy) and the Orkney Islands (Scotland, UK), respectively. The Erasmus Rucphen Study (ERF) is derived from a recent genetic isolate in the southwest Netherlands. This population was founded in the middle of the 18th century by approximately 150 individuals and was isolated until the last few decades. Twenty couples living in the region in the 19th century were chosen. These couples parented a minimum of 6 children, each of whom was baptized between 1880 and 1900 in the community church. All

living descendants of these pairs (as well as their spouses), ascertained on the basis of municipal and baptismal records, were traced and invited to participate. The MICROS study is part of the genomic health care program 'GenNova' and was carried out in three villages of the Val Venosta on the populations of Stelvio, Vallelunga and Martello. This study was an extensive survey carried out in South Tyrol (Italy) in the period 2001-2003. An extensive description of the study is available elsewhere(10). Briefly, study participants were volunteers from three isolated villages located in the Italian Alps, bordering with Austria and Switzerland. Due to geographical, historical and political reasons, the entire region experienced prolonged isolation from surrounding populations. The Orkney Complex Disease Study (ORCADES) is an ongoing family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with the high levels of endogamy historically. Data for participants aged 18-88 years, from a subgroup of ten islands, were used for this analysis. The KORA Study is a series of population-based epidemiological surveys of persons living in or near the city of Augsburg, Germany. All survey participants are residents of German nationality identified through the registration office and between 25 and 74 years old at the time of enrollment. Survey S3 was conducted between 1994 and 1995 and survey S4 between 1999 and 2001. KORA F3, a follow up examination at 10 years of follow up after S3, occurred in 2004 and 2005(11). The Netherlands Study of Depression and Anxiety (NESDA) is an ongoing 8-year longitudinal cohort study to examine (predictors of) the long-term course of depression and anxiety disorders. The rationales, methods and recruitment strategy have been described elsewhere(12). Subjects were recruited in various settings (community, general practices, mental health organizations) and include subjects with and without depression and anxiety disorders.The SardiNIA study is a population based study, that recruited and phenotyped 6,148 individuals, male and female, ages 14–102 years, from a cluster of four towns in the Lanusei Valley of Sardinia(13). For the GWA scans a total of 4,305 related individuals were examined(14). Genotyped individuals had four Sardinian grandparents and were selected without regard to their phenotypes. The Study of Health in

Pomerania (SHIP) is a longitudinal population-based cohort study in West Pomerania, a region in the northeast of Germany with a total population of 212,157 inhabitants(15). A two-stage cluster sampling method adopted from the WHO MONICA Project Augsburg, Germany yielded 12 five-year age strata for both genders, each including 292 individuals. For the baseline cohort, a sample of 6,267 eligible subjects aged 20 to 79 years was drawn from population registries where all German citizens are registered. The final study population compromised 4,310 subjects (response proportions 69%). The Twins UK Registry comprises unselected, mostly female volunteers ascertained from the general population through national media campaigns in the UK(16). Means and ranges of quantitative phenotypes in Twins UK were similar to an age-matched singleton sample from the general population(17). Zygosity was determined by standardized questionnaire and confirmed by DNA fingerprinting. Written informed consent was obtained from all participants before they entered the studies, which were approved by the local research ethics committee.

Phenotype measurement

In ARIC, RR interval was measured automatically from 12-lead ECGs performed at baseline. The study ECGs were recorded using MAC PC ECG machines (Marquette Electronics, Milwaukee, Wisconsin) in all four clinical centers. All ECGs were visually inspected for technical errors and inadequate quality. ECGs were initially processed in a central laboratory at the EPICORE Center (University of Alberta, Edmonton, Alberta, Canada) and during later phases of the study at the EPICARE Center (Wake Forest University, Winston-Salem, North Carolina). Initial ECG processing was done by the Dalhousie ECG program, and processing was later repeated with the 2001 version of the GE Marquette 12 SL program (GE Marquette, Milwaukee, Wisconsin).

In AGES, RR interval duration was automatically measured from 12-lead electrocardiograms using the Marquette 12 SL analysis program (General Electric Marquette Medical Division, Milwaukee, Wisconsin, USA).

In CHS, electrocardiograms were recorded using MAC PC ECG machines (Marquette Electronics, Milwaukee, Wisconsin) in all clinical centers. ECGs were initially processed in a central laboratory at the EPICORE Center (University of Alberta, Edmonton, Alberta, Canada) and during later phases of the study, at the EPICARE Center (Wake Forest University, Winston-Salem, North Carolina). All ECGs were visually inspected for technical errors and inadequate quality. RR interval was calculated from heart rate using the baseline ECG for eligible subjects. Initial ECG processing was done by the Dalhousie ECG program, and processing was later repeated with the 2001 version of the GE Marquette 12 SL program (GE Marquette, Milwaukee, Wisconsin).

In FHS, paper electrocardiograms 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 RR interval between QRS complexes from consecutive beats in sinus rhythmwas measured, after excluding all premature atrial or ventricular beats. The RR interval trait examined was the average RR interval from up to 4 cardiac cycles.

In KORA F3 and S4 12-lead resting electrocardiograms were recorded with digital recording systems (KORA F3: Mortara Portrait, Mortara Inc., Milwaukee, USA; KORA S4: Hörmann Bioset 9000, Hörmann Medizinelektronik, Germany). The Mortara Portrait determines RR interval by the proprietary XL-ECG algorithm which has not been published but has shown to be in good accordance with other published electrocardiogram measurement algorithms(18); RR interval in KORA S4 (Hörmann Bioset) was determined using the Hannover EKG analysis software (HES-Version 3.22-12) by computerized analysis of all leads and all cycles of a 10 second interval as described earlier.(19,20) In the Hannover algorithm RR intervals are taken as the intervals between the reference points detected in adjacent QRS complexes. The mean RR interval was computed, after exclusion of RR intervals that immediately precede and follow any premature ventricular complex. Only ECGs classified as "appropriate technical quality" according to visual inspection were used. The ECG examinations in both studies were performed according to a standard protocol, after ten minutes resting in supine position.

In MICROS 12-lead resting ECGs were recorded using a digital recording system (Mortara Portrait, Mortara, Milwaukee, WI, USA). The Mortara Portrait determines RR interval by the proprietary XL-ECG algorithm. In brief, computerized analysis of an averaged cycle computed from all leads and all cycles of the 10 second recording after exclusion of ectopic beats was performed. RR intervals were determined as the intervals between the reference points detected in adjacent QRS complexes. The median RR interval was computed, after exclusion of RR intervals that immediately precede and follow any premature ventricular complex.

In NESDA, all respondents received an average physiological recording of 100 minutes, performed with the ‘Vrije Universiteit Ambulatory Monitoring System”(21). The VU-AMS is a light-weight ambulatory device that records the electrocardiogram (ECG) from three electrodes in lead II configuration and changes in thorax impedance (dZ) from four electrodes placed at chest and back of the subjects. From the R-waves in the ECG signal the inter beat intervals (IBI) were computed online and the IBI time series was visually checked

for ectopic beats or missed R-waves offline. For the GWA mean heart rate was computed across valid IBIs across ten minutes of quiet supine rest at the start of the recording period.

In ORCADES digital 10 second ECGs were taken after 10 minutes supine rest, using a PC link with RR interval calculated using CardioView software (NUMED cardiac diagnostics, Sheffield, UK).

In the RS-I and RS-II, electrocardiograms were recorded on ACTA electrocardiographs (ESAOTE, Florence, Italy) and digital measurements of the RR intervals were made using the Modular ECG Analysis System (MEANS)(22). The MEANS program locates the QRS complexes and determines a stable reference point in each complex. The QRS detector of MEANS operates on multiple simultaneously recorded leads, which are transformed to a detection function that brings out the QRS complexes among the other parts of the signal. RR intervals are taken as the intervals between the reference points in adjacent QRS complexes. The median RR interval was computed, after exclusion of RR intervals that immediately precede and follow any premature ventricular complex. SHIP and ERF both use the same ECG analysis methods as the Rotterdam Study.

In SardiNIA the RR interval was calculated from heart rate that was measured from the electrocardiogram (measured on Cardiette 600 machines) during physical examination.

In TwinsUK, ECG data were available on 3,043 individuals before exlcusions. Two thousand seven hundred twenty six had automated measurements of the RR interval by the Cardiofax ECG-9020K (Nihon Kohden UK Ltd., Middlesex, UK) and 317 were scored manually using a high-resolution digitizing board (GTCO CalComp Peripherals, USA). The dataset for analyses consequently included 2,727 individuals, of which 1980 were DZ twins (i.e. 990 pairs) and 747 singletons. These singletons included 474 MZ twins of which the mean RR interval of both twins was used to optimize information.

Genotyping and imputation

A summary of the following information is given in **eTable 1**.

In AGES, blood was drawn at entry into the AGES Reykjavik study (2002-2006). Using the Illumina 370CNV BeadChip array, 3,219 participants were genotyped. Samples were excluded from the dataset based on sample failure, genotype mismatch with reference panel and genotypic-phenotypic sex mismatch. Standard Illumina genotyping protocols were followed. Prior to genotype imputation SNPs were excluded using filters for call rate (<97%), Hardy-Weinberg disequilibirum ($P<1x10^{-6}$), mishap ($P< 1x10^{-9}$), MAF (<0.01) and mismatched position between Illumina, dbSNP and/or HapMap, resulting in 325,094 SNPs available for usage in imputation. Imputation was performed using MACH v1.0.16(23) using flags: --rounds 100, --greedy in 200 unrelated individuals to generate model parameters. These parameters were used to impute autosomal HapMap CEU SNPs (release 22, build 36) allele dosages for each SNP resulting in 2,532,729 SNPs for analysis.

In ARIC, Affymetrix 6.0 array genotypes were obtained in 8,861 self-identified whites: 734 individuals were excluded for the following reasons: 1) discordant with previous genotype data, 2) genotypic sex did not match phenotypic sex, 3) suspected first degree relative of an included individual based on genome-wide genotype data, 4) genetic outlier (as assessed by average Identity by State (IBS) using PLINK and >8 standard deviations along any of first 10 principal components in EIGENSTRAT after 5 iterations. To be included for imputation, SNPs were required to have minor allele frequency ≥0.01, call rate ≥95% and HWE p $>1x10^{-5}$, resulting in 602,642 autosomal SNPs. Imputation was performed using MACH v1.0.16(23) with reference to HapMap CEU human genome release 21a build 35.

In CHS, genomic DNA was extracted from blood samples drawn on all participants at their baseline examination in 1989-90. Genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV BeadChip system in 2007-08. Genotypes were called using the Illumina BeadStudio software. Filtering of both individuals and SNPs was performed to ensure robustness for genetic analysis. 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, or zeroed accordingly. A final SNP set of 306,655 SNPs was identified. Imputation was performed using BIMBAM (i) v0.99 with reference to HapMap CEU using release 22 human genome build 36, 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.

In the EUROSPAN cohorts genotyping was performed using the Illumina 300 HumanHap Bead Chip. ERF included an additional 200 subjects previously genotyped on Affymetrix 250k. In ERF and MICROS only SNPs with a call rate>98%, MAF 1% and HWE $P>10^{-6}$ were used for imputation using MACH v1.0.16(23) and HapMap CEU release 22 build 36 as reference. The remaining 292,917 SNPs were used for imputation based on HapMap CEU release 22 build 36 using MACH v1.0.16(23). In ORCADES, subjects were excluded if they fulfilled any of the following criteria: genotypic call rate <97%, mismatch between reported and genotypic sex, unexpectedly low genomic sharing with first degree relatives, excess autosomal heterozygosity, or outliers identified by IBS clustering analysis. SNPs were excluded on the basis of minor allele frequency (< 1%), HWE ($P<10^{-5}$), call rate (<97%). MACH v1.0.15(23) was used to impute SNPs from HapMap CEU release 22 build 36 .

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. 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), strand incompatible 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. Imputation model parameters were estimated using MACH v1.0.15(23) (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 principal components analysis. With these model parameters, we used MACH to impute allele dosage, defined as the expected number of copies (a fractional value between 0 and 2), of all autosomal SNPs on HapMap CEU based on phased chromosomes of release 22, build 36.

In KORA, genomic DNA was hybridized in accordance with the manufacturer's standard recommendations. Genotypes from Affymetrix arrays were determined using the BRLMM (F3) or Birdseed (S4) clustering algorithm. The minimum call rate per SNP and person for genotypes carried to the next step were set to 93%. Genotypes from each study were imputed for the entire set of polymorphic HapMap SNPs (phased haplotypes: HapMap release 21 for imputation of KORA F3 and HapMap release 22 for imputation of KORA S4). MACH(23) software was used to compute allele dosages for all SNPs from the HapMap reference files.

In NESDA, GWA genotyping of a subsample of the NESDA participants was performed as part of the GAIN initiative to detect genetic variation associated with major depression.(24) Genotyping was performed on Affymetrix 500K GeneChip arrays with subsequent imputation using IMPUTE and the HapMap (release 22, build 36) reference population. A detailed description of genotyping and quality control criteria is given in Sullivan et al.(25).

All RS participants with available DNA were genotyped using Illumina Infinium II HumanHap BeadChips at the Department of Internal Medicine, Erasmus Medical Center following manufacturer's protocols. RS-I participants (n=6,449) were genotyped with 550K (V.3) single and duo chips, while RS-II participants (n=2,516) were genotyped with 550K (V.3) duo and 610K Quad chips. Genotype calling was performed in RS-I using BeadStudio software (version 0.3.10.14) and GenomeStudio in RS-II. Participants with call rate < 97.5%, excess autosomal heterozygosity, sex mismatch, or outlying identity-by-state clustering estimates were excluded. After quality control 5,974 RS-I participants and 2,157 RS-II participants remained in the GWAS datasets.

In SardiNIA, Genotyping was carried out on 3,329 individuals using Affymetrix 10K and on 1,412 individuals using the Affymetrix 500K, of whom 436 individuals were genotyped with both arrays. Genotype data was properly assessed by stringent quality control criteria as described in detail elsewhere.(14) In short, markers were removed for low call rate (<90%), MAF <0.05, HWE $P<10^{-3}$ (10K) or $P<10^{-6}$ (500K), Mendelian inconsistencies or unlikely genotypes. Additional genotypes from the entire set of polymorphic HapMap SNPs were imputed using the MACH software(23) This round of imputation was performed only in 1,412 individuals genotyped with the Affymetrix Mapping 500K Array Set. For the remaining 2,893 individuals genotyped with the Affymetrix Mapping 10K Array, mostly offspring and siblings of the 1,412 individuals that were genotyped with the Affymetrix Mapping 500K Array Set, we took advantage of the relatedness among individuals to impute missing genotypes as described elsewhere(14,26). Our inference approach allowed us to account for uncertainty in genotype assignment, estimating, instead of the most likely genotype, an expected genotype score, representing the expected number of copies of a reference allele (a fractional number between 0 and 2).

In SHIP, samples were genotyped using the Affymetrix Human SNP Array 6.0. Hybridization of genomic DNA was done in accordance with the manufacturer's standard recommendations. The genetic data analysis workflow was created using InforSense software. Genetic data were stored using the Caché (InterSystems) database. Genotypes were determined using the Birdseed2 clustering algorithm. For quality control purposes, several control samples where added. On the chip level, only subjects with a genotyping rate on QC probesets (QC callrate) of at least 86% were included. The overall genotyping efficiency of the GWA was 98.6 %. Imputation of genotypes in SHIP was performed with IMPUTE v0.5.0 based on HapMap (release 22, build 36).

Samples from the TwinsUK cohort were genotyped with the Infinium assay (Illumina, San Diego, USA) using four different SNP arrays, the Hap300 Duo, Hap300, Hap550 and Hap610(27). SNP calling was performed using the Illuminus software(28). SNPs were excluded if they violated Hardy–Weinberg equilibrium (HWE) ($p<1.0\times10^{-4}$); had

genotype call rates <95%; or had a minor allele frequency (MAF) of less than 0.01. Individuals were excluded if the sample call rate was less than 95%, autosomal heterozygosity was outside the expected range, genotype concordance was over 97% with another sample and the sample was of lesser call rate, non-Caucasian ancestry either self-identified or identified by cluster analysis in STRUCTURE(29) or multidimensional scaling by comparison to the three HapMap phase 2 reference populations (CEU, YRI, CHB+JPT), or unexplained relatedness (estimated proportion of allele shared identical by descent >0.05) to >120 other samples(30). This resulted in GWAS data being available for 5,295 twins from the TwinsUK cohort.

Genotype-phenotype association method In AGES and ARIC, RR interval residuals were tested for association with ProbABEL, which uses dosage values (0-2, continuous) as a predictor in a linear regression framework(31). In CHS, RR interval was linearly regressed on individual allele dosage (0-2), adjusting for covariates of age, sex, clinic, and BMI. Analyses were performed using R software. In FHS, linear mixed effects regression model accounting for within pedigree correlation was applied for association testing of the RR interval residuals with imputed allele dosage under an additive genetic model using the linear mixed effects model of the kinship package in R to account for relatedness. In KORA F3, KORA S4, RS-I and RS-II, RR interval residuals were tested for association with genotypes using MACH2QTL, which uses dosage value as a predictor in a linear regression framework(23). All three EUROSPAN cohorts performed the association analysis using ProbABEL(31). In ERF, a mixed-model approach was implemented to account for the family structure. Adjustments were made for sex, age and body mass index in an additive genetic model. Standard errors of effect estimates and association p-values were adjusted using genomic control(32) to correct the distribution of p-values that could be skewed in inbred populations. In SardiNIA, genotype scores were included in the family-based association test for evaluating the additive effect of each marker, as described elsewhere(26) and implemented in Merlin(33). Due to computational constraints, we divided large pedigrees into sub-units with

"bit-complexity" of 19 or less (typically, 20-25 individuals) before all the analyses. In SHIP and NESDA, RR intervals were tested for association using SNPTEST v1.1.5 with a linear regression model based on allele dosage and adjustment for sex, age and BMI. In TwinsUk, RR interval was adjusted for age, sex, and body mass index using regression models. In addition a covariate for the method of measurement (automatically vs. manually scored) of the ECG was incorporated. RR interval residuals were used for further analyses. Association between RR interval residuals and autosomal SNPs was tested in SNPTEST v1.1.4. As the TwinsUK cohort data consisted partly of dizygotic twins, the variances of the regression coefficients were corrected for the sibship relations using the Huber-White method for robust variance estimation in R(34,35).

II. Supplemental Tables

Supplemental Table 1. Summary methods per cohort

	Array	Calling algorithm	SNP call rate filter	HWE p-value filter	MAF filter	Imputation software	NCBI Build	Statistical Analysis	Number of SNPs	Lambda
AGES	Illumina 370CNV	BeadStudio	<97%	$<10^{-6}$	<1%	Mach v1.0.16	Build 36	ProbABEL	2,532,729	1.012
ARIC	Affymetrix 6.0	Birdseed	<95%	$<10^{-5}$	<1%	Mach v1.0.16	Build 35	ProbABEL	2,557,232	1.017
CHS	Illumina 370CNV	BeadStudio	<97%	$<10^{-5}$	<1%	BimBam v0.99	Build 36	R	2,333,043	1.043
ERF	Illumina 300, Affymetrix 250k	Beadstudio	<98%	$<10^{-6}$	<1%	MACH v1.0.16	Build 36	ProbABEL	2,543,887	1.027
FHS	Affymetrix 500k + gene-centric 50k MIP	BRLMM	<97%	$<10^{-6}$	<1%	Mach v1.0.15	Build 36	Kinship package in R	2,540,128	1.016
KORA F3	Affymetrix 500k	BRLMM	<93%	$<10^{-5}$	<5%	Mach v1.0.10	Build 35	Mach2QTL	2,557,252	0.997
KORA S4	Affymetrix 6.0	Birdseed	<93%	$<10^{-5}$	<5%	Mach v1.0.16	Build 36	Mach2QTL	2,543,887	0.995
MICROS	Illumina 300	Beadstudio	<98%	$<10^{-6}$	<1%	MACH v1.0.16	Build 36	ProbABEL	2,543,887	0.979
NESDA	Affymetrix 500k	PERLEGEN	<95%	none	<1%	IMPUTE v0.5.0	Build 36	SNPTEST	2,081,096	1.012
ORCADES	Illumina 300	Beadstudio	<97%	$<10^{-5}$	<1%	MACH v1.0.15	Build 36	ProbABEL	2,543,887	1.055
RS-I	Illumina 550k	BeadStudio	<98%	$<10^{-6}$	<1%	Mach v1.0.15	Build 36	Mach2QTL	2,543,887	1.016
RS-II	Illumina 550k	GenomeStudio	<98%	$<10^{-6}$	<1%	Mach v1.0.16	Build 36	Mach2QTL	2,543,887	1.015
SardiNIA	Affymetrix 10k, Affymetrix 500k	BRLMM	<90%	$<10^{-3}$ (10k), $<10^{-6}$ (500k)	<5%	MACH	Build 35	MERLIN	2,252,228	1.027
SHIP	Affymetrix 6.0	Birdseed	none	none	none	IMPUTE v0.5.0	Build 36	SNPTEST	2,748,910	1.007
TwinsUK	Illumina 300k Duo, Illumina 300, Illumina 550k, Illumina 610k	Beadstudio	<95%	$<10^{-4}$	<1%	IMPUTE v0.5.0	Build 36	SNPTEST	2,237,157	1.003

Supplemental table 2: Complete list of all genome-wide significant SNP from the RRGEN meta analysis

See Excel-file: RRGEN-summary-genome-wide-SNPs.xls

Supplemental Table 3. rs-numbers and annotation of quality controlled genotyped index SNPs in multi-variant polygenic model that in aggregate resulted in the maximal explained variance.

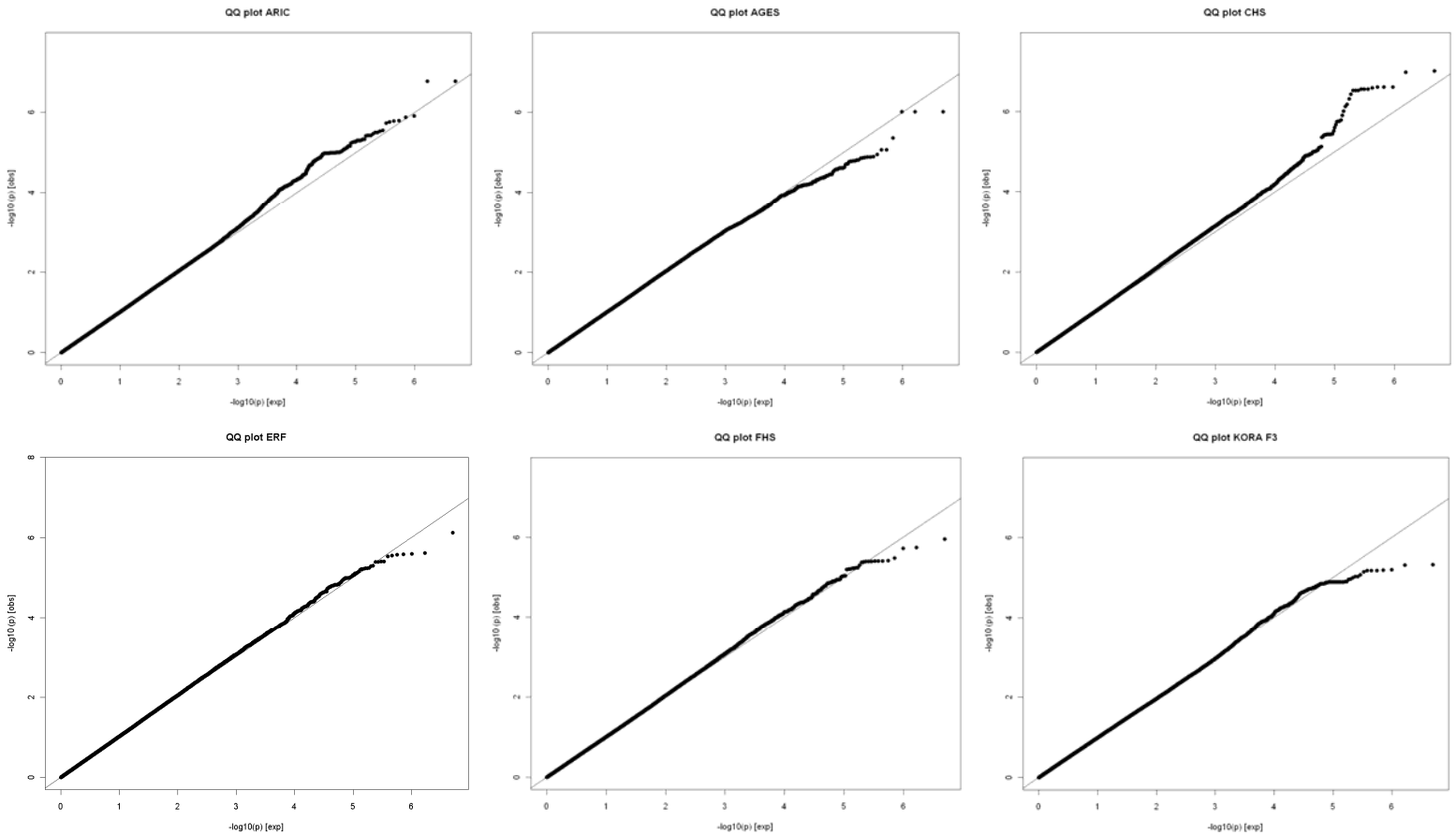
SNP	CHR	POS	Nearest Gene annotation
rs272564	1p34	44,784,860	*c1orf164*
rs11579530	1q32	206,197,041	Intergenic, near ***CD34***
rs16845015	2q21	141,430,306	*LRP1B*
rs1873164	2q31	287,738,258	Upstream *CCDC141*
rs4894803	3q26	173,282,950	*FNDC3B*
rs7612445	3q26	180,655,287	Upstream *GNB4*
rs281868	6q22	118,684,736	***SLC35F1***
rs2269579	6q22	122,198,112	Intergenic, ***GJA1***
rs314370	7q22	100,291,144	***SLC12A9***
rs17099385	10q25	121,680,302	Intergenic, *SEC23IP*
rs174546	11q12	61,326,406	3' UTR ***FADS1 / FADS3***
rs2238018	12p13	2,048,922	*CACNA1C*
rs4246224	12p12	24,675,406	Intergenic, ***SOX5 / BCAT1***
rs2200155	12p11	33,626,202	Intergenic, *SYT10*
rs2887596	12q14	76,678,888	Intergenic, *NAV3*
rs4981691	14q11	28,738,258	Intergenic, *PRKD1*
rs365990	14q12	22,931,651	Non synonymous coding ***MYH6***
rs223116	14q12	23,046,850	Intergenic, ***MYH7*** */ NGDN / ZFHX2*
rs2883661	16p12	19,899,685	Intergenic, *GPRC5B*
rs1364215	16q22	63,455,370	Intergenic, *CDH11*

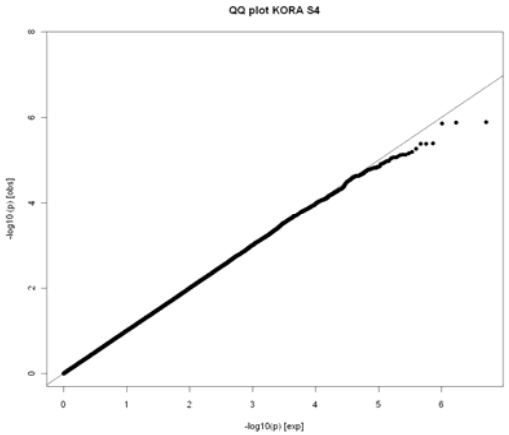
CHR=chromosome, POS=position

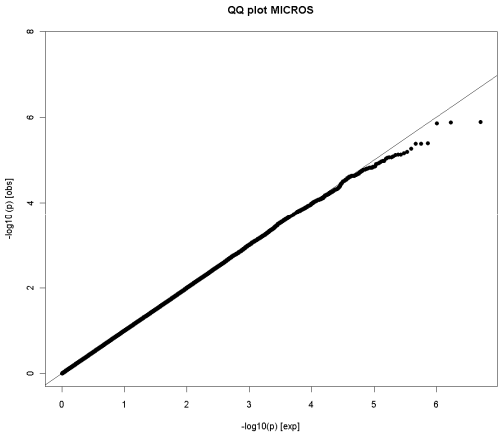
Genome-wide loci from main analysis are printed in bold

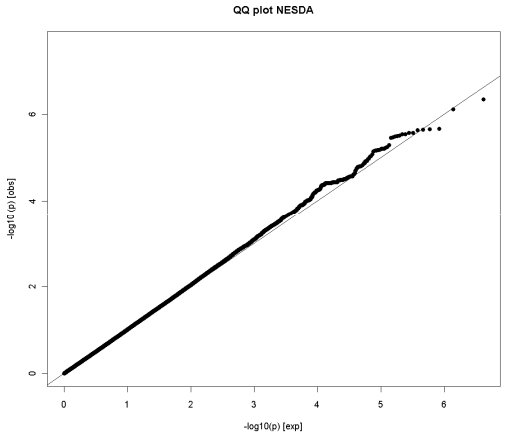
III. Supplemental Figures

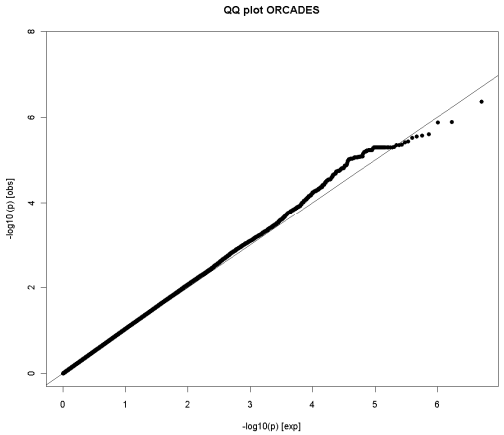
Supplemental Figure 1. Quantile – Quantile plots of observed/expected $-\log_{10}(P)$ per cohort and meta-analysis.

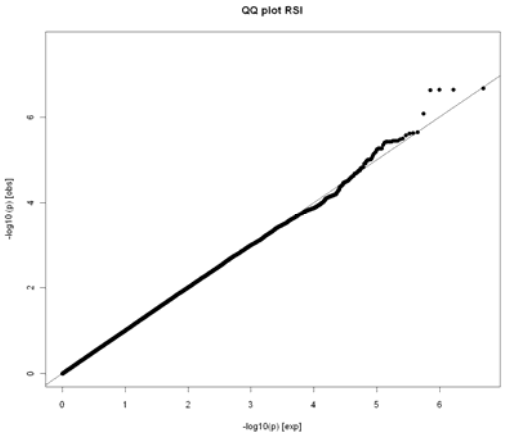


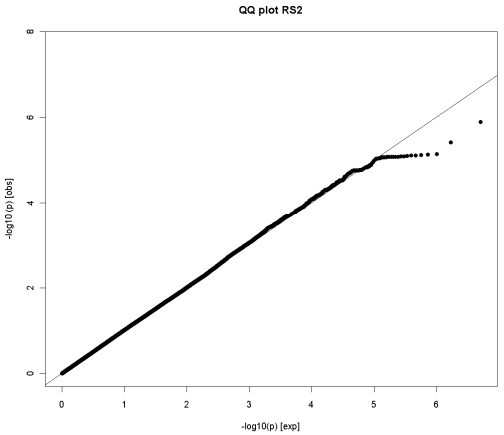

QQ plot KORA S4
-log10 (p) [obs]
-log10(p) [exp]

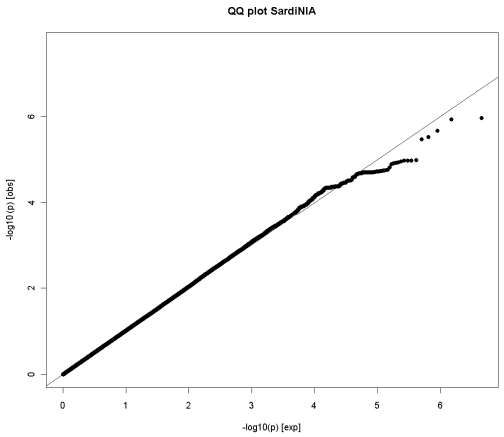

QQ plot MICROS
-log10 (p) [obs]
-log10(p) [exp]

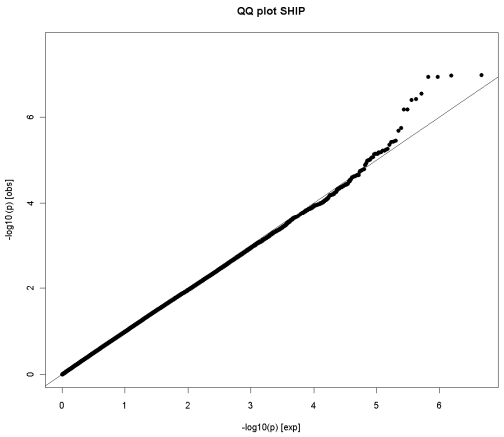

QQ plot NESDA
-log10 (p) [obs]
-log10(p) [exp]

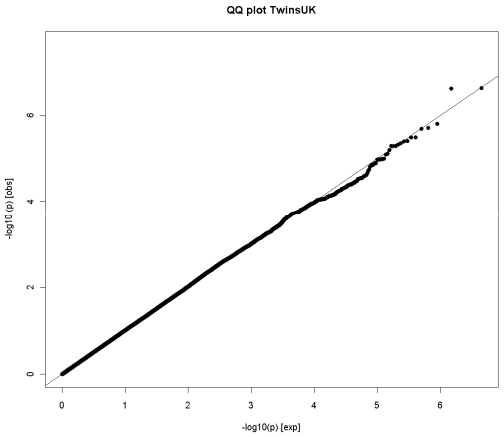

QQ plot ORCADES
-log10 (p) [obs]
-log10(p) [exp]

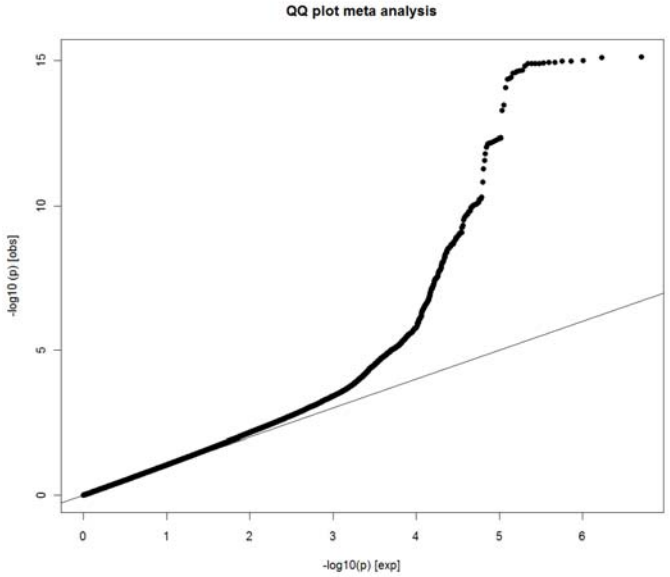

QQ plot RSI
-log10 (p) [obs]
-log10(p) [exp]


QQ plot RS2
-log10 (p) [obs]
-log10(p) [exp]

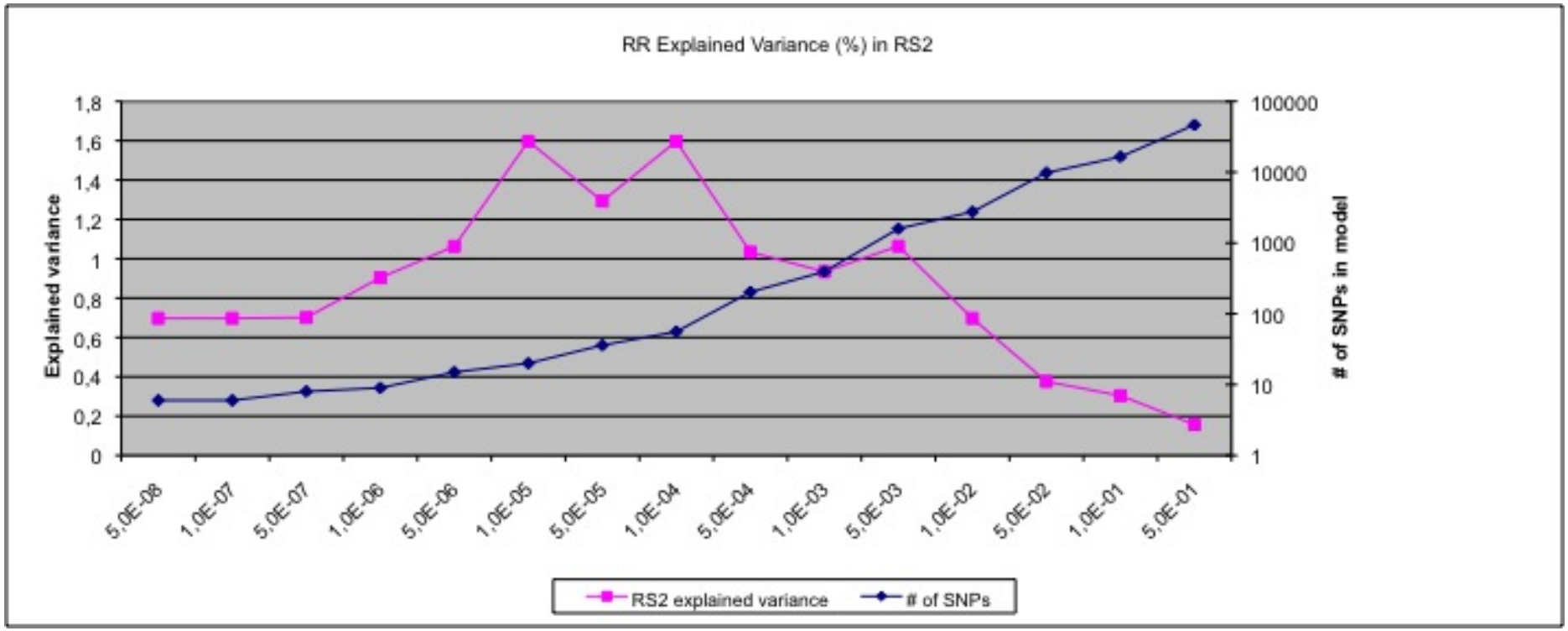

QQ plot SardiNIA
-log10(p) [obs]
-log10(p) [exp]


QQ plot SHIP
-log10(p) [obs]
-log10(p) [exp]


QQ plot TwinsUK
-log10(p) [obs]
-log10(p) [exp]


QQ plot meta analysis
-log10(p) [obs]
-log10(p) [exp]

Supplemental Figure 2. Graphical presentation of the multiple variant modeling approach estimating explained variance in the RS-II sample.



X-axis: P-value thresholds below which all independent signals meeting this threshold were included in the regression model to estimate the explained variance in the RS-II sample.

Left Y-axis: explained variance in the RS-II sample for any given score at the different P-value thresholds.

Right Y-axis: number of independent SNPs included in the model for a given P-value threshold.

IV. Supplemental References

1. Psaty, B.M., O'Donnell, C.J., Gudnason, V., Lunetta, K.L., Folsom, A.R., Rotter, J.I., Uitterlinden, A.G., Harris, T.B., Witteman, J.C.M., Boerwinkle, E. *et al.* (2009) Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of Prospective Meta-Analyses of Genome-Wide Association Studies From 5 Cohorts, *Circ. Cardiovasc. Genet.,* **2,** 73-80.
2. Harris, T.B., Launer, L.J., Eiriksdottir, G., Kjartansson, O., Jonsson, P.V., Sigurdsson, G., Thorgeirsson, G., Aspelund, T., Garcia, M.E., Cotch, M.F. *et al.* (2007) Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am.J. Epidemiol.,* **165,** 1076-1087.
3. (1989) The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am. J. Epidemiol.,* **129,** 687-702.
4. Fried, L.P., Borhani, N.O., Enright, P., Furberg, C.D., Gardin, J.M., Kronmal, R.A., Kuller, L.H., Manolio, T.A., Mittelmark, M.B., Newman, A. *et al.* (1991) The Cardiovascular Health Study: design and rationale. *Ann. Epidemiol.,* **1,** 263-276.
5. Dawber, T.R., Meadors, G.F. and Moore, F.E., Jr. (1951) Epidemiological approaches to heart disease: the Framingham Study. *Am. J. Public Health Nations Health,* **41,** 279-281.
6. Feinleib, M., Kannel, W.B., Garrison, R.J., McNamara, P.M. and Castelli, W.P. (1975) The Framingham Offspring Study. Design and preliminary data. *Preventive medicine,* **4,** 518-525.
7. Splansky, G.L., Corey, D., Yang, Q., Atwood, L.D., Cupples, L.A., Benjamin, E.J., D'Agostino, R.B., Sr., Fox, C.S., Larson, M.G., Murabito, J.M. *et al.* (2007) The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am. J. Epidemiol.,* **165,** 1328-1335.
8. Hofman, A., Grobbee, D.E., de Jong, P.T. and van den Ouweland, F.A. (1991) Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur. J. Epidemiol.,* **7,** 403-422.
9. Hofman, A., Breteler, M.M., van Duijn, C.M., Krestin, G.P., Pols, H.A., Stricker, B.H., Tiemeier, H., Uitterlinden, A.G., Vingerling, J.R. and Witteman, J.C. (2007) The Rotterdam Study: objectives and design update. *Eur. J. Epidemiol.,* **22,** 819-829.
10. Pattaro, C., Marroni, F., Riegler, A., Mascalzoni, D., Pichler, I., Volpato, C.B., Dal Cero, U., De Grandi, A., Egger, C., Eisendle, A. *et al.* (2007) The genetic study of three population microisolates in South Tyrol (MICROS): study design and epidemiological perspectives. *BMC Med. Genet.,* **8,** 29.
11. Wichmann, H.E., Gieger, C. and Illig, T. (2005) KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen,* **67 Suppl 1,** S26-30.
12. Penninx, B.W., Beekman, A.T., Smit, J.H., Zitman, F.G., Nolen, W.A., Spinhoven, P., Cuijpers, P., De Jong, P.J., Van Marwijk, H.W., Assendelft, W.J. *et al.* (2008) The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Iin. J. Meth. Psych. Res.,* **17,** 121-140.
13. Pilia, G., Chen, W.M., Scuteri, A., Orru, M., Albai, G., Dei, M., Lai, S., Usala, G., Lai, M., Loi, P. *et al.* (2006) Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet.,* **2,** e132.
14. Scuteri, A., Sanna, S., Chen, W.M., Uda, M., Albai, G., Strait, J., Najjar, S., Nagaraja, R., Orru, M., Usala, G. *et al.* (2007) Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.,* **3,** e115.
15. John, U., Greiner, B., Hensel, E., Ludemann, J., Piek, M., Sauer, S., Adam, C., Born, G., Alte, D., Greiser, E. *et al.* (2001) Study of Health In Pomerania (SHIP): a health examination survey in an east German region: objectives and design. *Soz. Praventivmed.,* **46,** 186-194.

16. Spector, T.D. and Williams, F.M. (2006) The UK Adult Twin Registry (TwinsUK). *Twin Res. Hum. Genet.,* **9,** 899-906.
17. Andrew, T., Hart, D.J., Snieder, H., de Lange, M., Spector, T.D. and MacGregor, A.J. (2001) Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin. Res.,* **4,** 464-477.
18. Michelucci, A., Mortara, D., Lazzeri, C., Barletta, G., Capalbo, A., Badia, T., Del Bene, R., Bano, C., Gensini, G.F. and Franchi, F. (2002) Simultaneous assessment of electrocardiographic parameters for risk stratification: validation in healthy subjects. *Ital. Heart. J.,* **3,** 308-317.
19. Pfeufer, A., Jalilzadeh, S., Perz, S., Mueller, J.C., Hinterseer, M., Illig, T., Akyol, M., Huth, C., Schopfer-Wendels, A., Kuch, B. *et al.* (2005) Common variants in myocardial ion channel genes modify the QT interval in the general population: results from the KORA study. *Circ. Res.,* **96,** 693-701.
20. Zywietz, C., Borovsky, D., Gotsch, G. and Joseph, G. (1990) Methodology of ECG interpretation in the Hannover program. *Methods Inf Med,* **29,** 375-385.
21. Licht, C.M., de Geus, E.J., Zitman, F.G., Hoogendijk, W.J., van Dyck, R. and Penninx, B.W. (2008) Association between major depressive disorder and heart rate variability in the Netherlands Study of Depression and Anxiety (NESDA). *Arch. Gen. Psychiatry.,* **65,** 1358-1367.
22. van Bemmel, J.H., Kors, J.A. and van Herpen, G. (1990) Methodology of the modular ECG analysis system MEANS. *Methods Inf. Med.,* **29,** 346-353.
23. Li, Y. and Abecasis, G.R. (2006) Mach 1.0: Rapid Haplotype Reconstruction and Missing Genotype Inference. *Am. J. Hum. Genet.,* **S79,** 2290.
24. Boomsma, D.I., Willemsen, G., Sullivan, P.F., Heutink, P., Meijer, P., Sondervan, D., Kluft, C., Smit, G., Nolen, W.A., Zitman, F.G. *et al.* (2008) Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. *Eur. J. Hum. Genet.,* **16,** 335-342.
25. Sullivan, P.F., de Geus, E.J., Willemsen, G., James, M.R., Smit, J.H., Zandbelt, T., Arolt, V., Baune, B.T., Blackwood, D., Cichon, S. *et al.* (2009) Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol. Psychiatry,* **14,** 359-375.
26. Sanna, S., Jackson, A.U., Nagaraja, R., Willer, C.J., Chen, W.M., Bonnycastle, L.L., Shen, H., Timpson, N., Lettre, G., Usala, G. *et al.* (2008) Common variants in the GDF5-UQCC region are associated with variation in human height. *Nat. Genet.,* **40,** 198-203.
27. Richards, J.B., Rivadeneira, F., Inouye, M., Pastinen, T.M., Soranzo, N., Wilson, S.G., Andrew, T., Falchi, M., Gwilliam, R., Ahmadi, K.R. *et al.* (2008) Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet,* **371,** 1505-1512.
28. Teo, Y.Y., Inouye, M., Small, K.S., Gwilliam, R., Deloukas, P., Kwiatkowski, D.P. and Clark, T.G. (2007) A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics,* **23,** 2741-2746.
29. Pritchard, J.K., Stephens, M. and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics,* **155,** 945-959.
30. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.,* **81,** 559-575.
31. Aulchenko, Y.S., Ripke, S., Isaacs, A. and van Duijn, C.M. (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics,* **23,** 1294-1296.
32. Devlin, B., Roeder, K. and Wasserman, L. (2001) Genomic control, a new approach to genetic-based association studies. *Theor. Popul. Biol.,* **60,** 155-166.
33. Abecasis, G.R., Cherny, S.S., Cookson, W.O. and Cardon, L.R. (2002) Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.,* **30,** 97-101.

34. Huber, P.J. (1967) The behavior of maximum likelihood estimates under non-standard conditions. *Proc. Fifth Berkeley Symposium Math. Stat.,* **1,** 221-223.
35. White, H. (1982) Maximum Likelihood Estimation of Misspecified Models. . *Econometrica* **50,** 1-26.

V. Funding and Acknowledgements

Funding The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute [contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367, R01HL086694 and U10HL054512]; National Human Genome Research Institute [contract U01HG004402]; and National Institutes of Health [contract HHSN268200625226C]. Infrastructure was partly supported by a component of the National Institutes of Health and National Institutes of Health Roadmap for Medical Research [grant UL1RR025005].

The Age, Gene/Environment Susceptibility Reykjavik Study has been funded by National Institutes of Health [contract N01-AG-12100], the National Institute of Aging Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

The CHS research reported in this article was supported by the National Heart, Lung, and Blood Institute [contract numbers N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, grant numbers U01 HL080295, R01 HL087652, and R01 HL088456], with additional contribution from the National Institute of Neurological Disorders and Stroke. A full list of principal CHS investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center for Research Resources [grant M01RR00069] to the Cedars-Sinai General Clinical Research Center Genotyping core and National Institute of Diabetes and Digestive and Kidney Diseases [grant DK063491] to the Southern California Diabetes Endocrinology Research Center.

The ERF study was supported by grants from the Netherlands Organisation for Scientific Research [Pionier, 047.016.009, 047.017.043], Erasmus MC and the Centre for Medical Systems Biology (CMSB; National Genomics Initiative).

The FHS research was conducted in part using data and resources from the Framingham Heart Study of the National Heart, Lung, and Blood Institute of the National

Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study [contract number N01-HC-25195] and its contract with Affymetrix, Inc for genotyping services [contract number 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. Electrocardiographic measurements in the third generation of FHS were supported by grants from the National Institutes of Health and the Doris Duke Charitable Foundation to CNC.

The KORA Augsburg studies were financed by the Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany and supported by grants from the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Part of this work was financed by the German National Genome Research Network (NGFN-2 and NGFN-plus), the German National Competence network on atrial fibrillation (AFNET) and the Bioinformatics for the Functional Analysis of Mammalian Genomes program (BFAM) by grants to SK [NGFN 01GS0499, 01GS0838 and AF-Net 01GI0204/N], AP [GFN 01GR0803, 01EZ0874], HEW [NGFN 01GI0204] and to TM [NGFN 01GR0103]. SK is also supported by a grant from Fondation Leducq. Furthermore, this research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ.

The MICROS study in South Tyrol was supported by the Ministry of Health and Department of Educational Assistance, University and Research of the Autonomous Province of Bolzano and the South Tyrolean Sparkasse Foundation.

NESDA was supported by the Geestkracht program of ZonMW [grant 10-000-1002]; matching funds from universities and mental health care institutes involved in NESDA (GGZ Buitenamstel-Geestgronden, Rivierduinen, University Medical Center Groningen, GGZ

Lentis, GGZ Friesland, GGZ Drenthe). Genotyping was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health, and analysis was supported by grants from GAIN and the NIMH (MH081802). Genotype data were obtained from dbGaP (http://www.ncbi.nlm.nih.gov/dbgap, accession number phs000020.v1.p1).

ORCADES was supported by the Scottish Executive Health Department and the Royal Society and is now a component of the EU Framework 6 project EUROSPAN [Contract LSHG-CT-2006-018947]. PN acknowledges the Medical Research Council for funding.

The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments [grant 175.010.2005.011, 911-03-012]. This study is funded by the Research Institute for Diseases in the Elderly [014-93-015; RIDE2] and the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) [050-060-810]. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The Rotterdam Study used resources from the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fur Forschung und Technology [grant #01 AK 803 A-H, # 01 IG 07015]. This work was also supported by the Netherlands Heart Foundation (NHF) [2007B221 and 2009R014 to ME].

The SardiNIA team was supported by the National Institute on Aging [contract NO1-AG-1-2109] from and in part by the Intramural Research Program of the National Institutes of Health , National Institute on Aging.

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

01ZZ9603, 01ZZ0103 and 01ZZ0403], the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg - West Pomerania. Generation of genome-wide data has been supported by the Federal Ministry of Education and Research [grant 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.

Twins UK was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013) [grant HEALTH-F2-2008-201865-GEFOS] and (FP7/2007-2013) [ENGAGE grant HEALTH-F4-2007-201413] and the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254). The study also receives support from the Dept. of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. The project also received support from a Biotechnology and Biological Sciences Research Council (BBSRC) project [grant G20234]. Analyses were performed on the Genetic Cluster Computer, which is financed by an NWO Medium Investment [grant 480-05-003] and by the Faculty of Psychology and Education of the VU University, Amsterdam, The Netherlands.

Acknowledgements The authors thank the staff and participants of the ARIC study for their important contributions. We thank T. Aspelund and G. Eiriksdottir for their contribution to collecting, analyzing and preparing the AGES Reykjavik Study data. We thank P. Arp, M. Jhamai, Dr M. Moorhouse, M. Verkerk, and S. Bervoets for their help in creating the Rotterdam Study GWAS database. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We would like to thank Dr. T. A. Knoch, L.V. de Zeeuw, A. Abuseiris, and R. de Graaf as well as their institutions, the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid. We would like to thank the Sardinian volunteers who generously supported the SardiNIA study and made it possible. We also acknowledge the support of the administration of Lanusei, Ilbono, Arzana and Elini (Sardinia, Italy). The authors are grateful to S. Funke for the opportunity to use his Server Cluster for SNP Imputation, and to the contributions of A. Teumer, A. Hoffmann, and A. Petersmann in generating the SNP data for the SHIP study. For the ERF Study we are grateful to all patients and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection. We would like to acknowledge the invaluable contributions of L. Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney. For the MICROS study, we thank the primary care practitioners R. Stocker, S. Waldner, T. Pizzecco, J. Plangger, U. Marcadent and the personnel of the Hospital of Silandro (Department of Laboratory Medicine) for their participation and collaboration in the research project. For TwinsUK we thank the staff from the Genotyping Facilities at the Wellcome Trust Sanger Institute for sample preparation, quality control and genotyping led by Leena Peltonen and Panos Deloukas; Le Centre National de Génotypage, France, led by Mark Lathrop, for genotyping; Duke University, North Carolina, USA, led by David Goldstein, for genotyping; and the Finnish Institute of Molecular Medicine, Finnish Genome Center,

University of Helsinki, led by Aarno Palotie, for genotyping. Genotyping was also performed by CIDR as part of a National Eye Institute NIH project grant (PI: Terri Young).

The authors acknowledge the essential role of the Cohorts for Heart and Aging Research in Genome Epidemiology Consortium in development and support of this manuscript. CHARGE members include the Netherlands's Rotterdam Study (RS), Framingham Heart Study (FHS), Cardiovascular Health Study (CHS), the NHLBI's Atherosclerosis Risk in Communities (ARIC) Study, and the NIA's Iceland Age, Gene/Environment Susceptibility (AGES) Study.

VI. Author contribution list

RRGEN Study concept and design: E.B., P.I.W.d.B., C.N.C., C.v.D., C.O.D., M.E., S.H., A.C.M., M.M., A.Pf., N.S., B.H.C.S.

Cohort Study concept and design: E.B., H.C., C.N.C., C.v.D., C.O.D., S.B.F., E.J.C.d.G., V.G., T.H., A.H., H.K.K., E.G.L., M.G.L., L.L., B.A.O., B.W.P., A.Pe., P.P.P., B.M.P., D.S.S., T.D.S., M.U., H.V., H.E.W., J.F.W., A.F.W.

Acquisition of data: A.A., E.B., D.C., C.N.C., C.v.D., C.O.D., M.D., F.E., K.E., A.F., C.F., C.G., V.G., A.A.H., G.H., S.J.H., A.I., Y.J., S.K., E.G.L., M.G.L., C.M.L., T.M., S.S.N., B.A.O., A.Pe., A.Pf., S.P., P.P.P., B.M.P., F.R., J.R., I.M.R., T.D.S., B.H.C.S., A.G.U., M.U., S.H.W., H.E.W., J.F.W., J.C.W.

Statistical analysis and interpretation of data: D.E.A., Y.S.A., P.I.W.d.B., J.B., C.N.C., C.O.D., M.D., M.E., F.E., K.E., C.F., E.J.C.d.G., C.G., A.A.H., S.J.H., G.H., J.J.H., A.I., S.K., A.C.M., F.M., M.M., I.M.N., P.Na., P.No., A.Pf., K.M.R., F.R., H.S., N.S., S.S., A.V.S.

Drafting of the manuscript: P.I.W.d.B., C.N.C, M.E., A.C.M., N.S.

Critical revision of the manuscript: A.A., J.B., E.B., D.C., H.C., C.N.C., C.v.D., C.O.D., M.D., M.E., A.F., C.G., A.I., A.M., K.E., C.F., S.B.F., E.J.C.d.G., A.H., S.H., A.A.H., G.H., A.I., S.K., H.K.K., E.G.L., M.G.L., F.M., T.M., M.M., I.M.N., P.Na., P.No., S.P., A.Pe., A.Pf., P.P.P., B.M.P., K.M.R., F.R., J.R., I.M.R., S.S., D.S.S., H.S., A.V.S., B.H.C.S., K.V.T., A.G.U., H.V., H.E.W., J.F.W., A.F.W., J.C.W.

Obtained funding: E.B., H.C., C.N.C., C.v.D., S.B.F., E.J.C.d.G., V.G., T.H., A.H., Y.J., S.K., H.K.K., L.L., B.A.O., B.W.P., A.Pf., P.P.P, B.M.P., J.R., H.S., N.S., T.D.S., B.H.C.S., A.G.U., M.U., H.V., H.E.W., J.F.W., A.F.W., J.C.W.

Final approval of manuscript: all authors.