SUPPLEMENTAL MATERIALS:

STUDY DESCRIPTIONS:

Atherosclerosis Risk in Communities Study (ARIC): The Atherosclerosis Risk in Communities (ARIC) study is a community-based prospective cohort of cardiovascular disease and its risk factors. ARIC recruited 15,792 men and women, aged 45-64 years, from 4 communities in the United States (Forsyth Co, NC; Jackson, MS; Minneapolis suburbs, MN; Washington Co, MD) in 1987-1989.[1] Participants were mostly white in the Minnesota and Washington Co field center, white and African-American in Forsyth Co, and exclusively African-American in the Jackson field center. At baseline, participants answered questionnaires on cardiovascular risk factors, had a physical exam, and provided blood samples.

Cardiovascular Health Study (CHS): The CHS is a population-based cohort study of risk factors for CHD and stroke in adults ≥65 years conducted across four field centers.[2] The original cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists; subsequently, an additional predominantly African American cohort of 687 persons was enrolled in 1992-1993 for a total sample of 5,888. DNA was extracted from blood samples drawn on all participants at their baseline examinations. Targeted sequencing was performed at on 1,132 CHS participants of European ancestry who were free of CVD at baseline, consented to genetic testing, and had DNA available for genotyping.

Framingham Heart Study (FHS): The Framingham Heart Study (FHS) is a community-based cohort study composed of three generations of Framingham (MA) residents who are almost exclusively of European descents. The Original cohort (n=5209) was initiated in 1948 whose children and their spouses were enrolled to the Offspring cohort (n=5124) in 1971-1975.[3] The Third Generation cohort (n=4095), mostly the children of the Offspring cohort was enrolled in 2002 to 2005.[4] All participants were examined every 4-8 years. For this study DNA was extracted from either cell lines or blood samples of 1096 participants of Original cohort exam 25 or 26 (n=36, draw date 1997-2001), Offspring cohort exam 6 (n=958, draw date 1995-1996) and Third Generation exam 1 (n=102, draw date 2002-2005).

PHENOTYPE GROUP SELECTION[$]

1. **Atrial fibrillation:** Two hundred (200) subjects from the Massachusetts General Hospital Atrial Fibrillation study with early-onset atrial fibrillation occurring before 66 years of age were selected for sequencing. Participants with evidence of structural heart disease as assessed by echocardiography were excluded.

2. **Blood pressure:** One hundred (100) individuals were selected from the ARIC study, 50 from FHS, and 50 from CHS from both extremes of the standardized residuals of systolic blood pressure and diastolic blood pressure after adjustment for age, age^2, BMI, and study site if applicable. The regression was stratified by sex, and an equal number of individuals from both sexes were chosen for sequencing. Data from the first clinical visit available for an individual was used where there was data from multiple examinations. Systolic blood pressure and

diastolic blood pressure were adjusted if participants were taking antihypertensive medication by adding 10 mm Hg/5 mm Hg. Individuals taking antihypertensive medication for the selection of subjects for the lower tail of the trait; who had a history of heart failure prior to measurement of blood pressure; whose systolic blood pressure was < 60 mm Hg, or diastolic blood pressure was <20 mm Hg; or whose BMI was + 4 standard deviations from the mean were excluded.

3. **Body mass index (BMI):** Two hundred (200) unrelated individuals including 100 participants from the ARIC study, 50 from CHS, and 50 from FHS were sequenced from the high tail of the distribution for BMI based on age- and sex-adjusted residuals. In FHS, subjects were greater than 25 years of age. In the ARIC study and CHS, there were no age restrictions. In all studies, individuals were excluded if BMI < 18.5 kg/m^2.

4. **Bone mineral density:** One hundred (100) individuals were selected for targeted sequencing from CHS and 100 were chosen from FHS who had extremely low femoral neck bone mineral density with approximately twice as many women as men. The selection of participants was based on using a femoral neck bone mineral density T-score (number of standard deviations below young normal values) < -2 and Z-score < -1.5.

5. **C-reactive protein (CRP) level:** One hundred (100) individuals from the ARIC study, 50 from CHS, and 50 from FHS with the highest CRP residuals were chosen from a sex-stratified sample

after adjustment for age, hormone therapy, study site, BMI, and lipid therapy. Participants with residuals greater than four times the standard deviation were excluded.

6. **Carotid intima-media thickness (IMT):** The study sample included 100 subjects from the ARIC study, 50 subjects from CHS, and 50 subjects from FHS, with an equal number of men and women. Participants were selected for sequencing from the high tail of the common carotid IMT distribution.

7. **Echocardiography (left ventricular diastolic dimension):** Fifty (50) unrelated males and 50 unrelated females (total=100) from the highest end of the trait distribution in CHS and FHS were sampled for sequencing after adjustment for age, height, weight, and study site if applicable.

8. **ECG (electrocardiogram) PR interval:** Two hundred (200) subjects from the upper tail of the PR trait distribution based on residuals of a model with PR interval as the dependent variable and age, sex, study center, BMI, and height as the independent variables were selected for sequencing, including 50 men with the highest residuals and 50 women with the highest residuals in the ARIC study, 50 participants from CHS, and 50 participants from FHS after applying exclusions. Individuals with a history of atrial fibrillation at baseline; extreme PR interval (<80 or >320); pacemaker or defibrillator; Wolff-Parkinson-White (WPW) syndrome; third degree AV block; history of heart failure or myocardial infarction; use of digoxin or class I

or class III antiarrhythmic blocking medication; or who were missing covariates used for adjustment were excluded.

9. **ECG (electrocardiogram) QRS interval:** Two hundred (200) subjects were sequenced from the upper tail of the QRS trait distribution including 50 men and 50 women in the ARIC study, 50 individuals from CHS and 50 participants in FHS after applying exclusions. Individuals with atrial fibrillation; history of myocardial infarction or congestive heart failure; a QRS interval > 120; Wolff-Parkinson-White (WPW) syndrome; implantation of a pacemaker, and use of class I and class III antiarrhythmic blocking medication were excluded.

10. **Fasting insulin:** Two hundred (200) subjects were sampled from the high tail of the distribution including 100 participants in the ARIC study, 50 participants in CHS, and 50 participants in FHS. Individuals with known diabetes; who were treated for diabetes; or those with a fasting glucose > 7 mmol/L were excluded. The ARIC study and FHS applied a further exclusion of non-fasting individuals. Participants who were missing hemoglobin A1c values were also excluded in the ARIC study, and subjects with type 1 diabetes were excluded in FHS. Selection was sex balanced.

11. **Hematocrit:** Two hundred (200) individuals were selected from the lower tail of the hematocrit distribution including 100 ARIC study participants, 50 CHS participants, and 50 FHS participants. A 50:50 gender ratio was maintained. The residuals from linear regression of hematocrit as a continuous trait with adjustment for age, sex, and study site for multicenter

cohorts were calculated for each of the three cohorts. Individuals with hematocrits within 3 standard deviations of the sample mean for each cohort were included in the analysis. Individuals with known malignancies; who smoked; or who had renal failure were excluded.

12. **Pulmonary function:** Severe cases of chronic obstructive pulmonary disease (COPD) were selected based on forced expiratory volume in the first second (FEV1) that was less than 65% of the predicted value, and its ratio to forced vital capacity (FEV1/FVC) that was less than the lower limit of normal based on NHANES III prediction equations. A random sample of 200 subjects was selected for sequencing among those who met the severe COPD definition at visits 1 and 2 in the ARIC study, and who had non-missing covariate data.

13. **Retinal venule diameter:** Individuals were selected for sequencing from the highest quartile of the trait distribution adjusted for age and sex from the ARIC study (n=166) and CHS (n=34). All participants had retinal photography and retinal arteriolar and venular caliber measured from computer software using standardized protocols.

14. **Stroke:** Stroke was defined as a focal neurological deficit of presumed vascular cause with sudden onset and lasting for at least 24 hours or until death if the participant died less than 24 hours after the onset of symptoms. Participants with incident ischemic stroke based on clinical and imaging criteria excluding cardioembolic events were eligible for selection. This phenotype, corresponding to both large and small artery atherothrombotic strokes yielded the largest hazard ratio in the CHARGE meta-analysis. From among individuals meeting these criteria, the

individuals with the earliest strokes with onset past the age of 65, and equal numbers of men and women were selected in numbers proportional to the size of the participating cohorts: 80, 70, and 50 individuals from the ARIC study, CHS, and FHS, respectively.

[$]**Note**: Each Phenotype Group often has additional extreme participants that were selected by other Phenotype Groups or Cohort Random Sample.

PLEIOTROPIC REGION TARGET SELECTION

We defined pleiotropy operationally as evidence that a region or locus containing one to many genes displays strong associations ($P < 5 \times 10^{-8}$) with 2 or more traits in multiple genome-wide association studies. Candidate pleiotropic regions were identified on the basis of having two or more cardiovascular disease traits including our target phenotypes with $p < 5 \times 10^{-8}$ in either the NHGRI GWAS Catalog and/or the open access GWAS database[5] (Table S2b). Once these regions were nominated, *key SNPs* were selected in the regions by the following criteria: 1) GWAS trait associations, 2) known gene expression (eQTL) associations, and 3) HapMap CEU LD of $r^2>0.5$ with SNPs identified by GWAS or eQTL approaches. The 5'-most SNP (minus a 100 bp pad) and 3'-most SNP (plus a 100 bp pad) in each locus defined the locus boundaries. The *key SNPs* (n=714 SNPs) were targeted with a ±30 bp pad regardless of genic or intergenic position. A total of 43 gene sub-regions in 7 loci related to cardiovascular disease were annotated (Table 3b) and all exons and untranslated regions (n=494 exons or UTRs) were targeted with ±15 bp pads for sequencing. The full bi-directional promoter region of *COL4A1* and *COL4A2* was additionally selected for targeting. Finally, sub-regions in the 7 loci displaying high sequence conservation

across mammals and/or vertebrates (28-species phastCons scores >500[6]) were selected for targeting (n=705 sub-regions). The union of the selected regions was assessed in all participants from the ARIC study, CHS, and FHS who were selected for targeted sequencing.

Supplemental Methods and Results

Library Preparation and Sequencing

At the Baylor College of Medicine Human Genome Sequencing Center (HGSC), we developed a new SOLiD™ platform-based multiplexed capture sequencing protocol for this project (the complete protocol and primer sequences are accessible from our website https://www.hgsc.bcm.edu/sites/default/files/documents/Preparation_of_SOLiD_Multiplex_Capture_Libraries.pdf). The main features of the new protocol involved utilizing two HGSC-designed adaptors (i.e. TrTA-Mul-P1 and TrTA-Mul-Int) for better ligation efficiency and SOLiD™ barcodes-containing PCR primers (i.e. Mul-pre-BCn) in pre-capture Ligation Mediated PCR (LM-PCR) amplification. Briefly, samples were first subjected to quality control (QC) examination to evaluate DNA integrity and quantity using both agarose gels and PicoGreen assays. Starting with 1*u*g of each sample, pre-capture libraries were generated for each sample and consisted of DNA fragmentation (~120 bp), end repair, 3'-adenylation, adaptor ligation and barcode labeling in pre-capture LM-PCR. After bead purification, PCR products were quantified and their size distribution analyzed using an Agilent Bioanalyzer 2100 DNA 7500 chip. Library methods were supported by the use of Beckman Biomek NX/FX robotic platforms for 96-well processing as well as a custom integrated laboratory information management system.

For multiplexed sequence capture, four color-balanced pre-capture libraries (using barcode sequences BC1-4 and BC5-8) were pooled in equal molar amount (~500ng for each sample) for co-capture with the NimbleGen CHARGE liquid probe set (~2.0 Mb target size). Each pool was hybridized in solution to the probe according to the manufacturer's protocol with minor revisions. Specifically, Human COT1 DNA and adaptor-specific hybridization enhancing oligonucleotides (HEOs) were added into the hybridization reactions to block repetitive genomic sequences and the common adaptor sequences. Post-capture LM-PCR was performed using 12-14 cycles. Capture libraries were quantified using PicoGreen and their size distribution analyzed using the Agilent Bioanalyzer 2100 DNA 7500 chip. The efficiency of the capture was evaluated by performing a qPCR-based quality check on the standard NimbleGen controls. The enrichment of the capture libraries was estimated to range from 7 to 9 fold over the background. Two successfully enriched libraries were further pooled to form an 8-sample pool for multiplexed sequencing on the SOLiD^{TM} platform. Each sequencing pool was subsequently sequenced on one quadrant of a SOLiD^{TM} V4 slide (total 32 samples per SOLiD^{TM} slide) using Life Technologies' Barcode Fragment Sequencing Kits and methods.

Genotype Validation by GWAS data

In order to further validate the genotypes identified by the current project, we performed a cross-validation by the Affymetrix Gene Chip 500K Array Set & 50K Human Gene Focused Panel in 1096 FHS samples. A total of 558 SNPs were shared between the two platforms. After excluding missing genotypes, approximately 98.0% of genotypes were concordant between the two platforms, suggesting high accuracy of the sequenced genotypes.

We also performed proof-of-principle analysis to investigate potential effect of copy number variations (CNVs) in the studied samples. We genotyped FHS samples by the Affymetrix Gene Chip 500K Array, whereas CNVs were called by PennCNV.[7] Only 18 out of 1096 studied samples (1.6%) had copy number segments that overlapped with the 77 targeted regions. We thus concluded that the chance to have CNVs in our samples would be low. We acknowledge that Affymetrix Gene Chip 500K Array is not the optimum approach to detecting CNVs. Future advance in sequencing technologies and variant calling algorithms will be able to integrate CNVs into variant calling.

Association with C-reactive Protein Levels

We performed association analyses for C-reactive protein (CRP) levels, which were measured in each individual cohort according to the manufacturers' protocols. Four genes were selected for sequencing because of their prior association with CRP: *GCKR*, *HNF1A*, *IL6R*, and *LEPR*.[8] We identified 3,376 variants within the 4 CRP regions, including 286 common variants and 3,090 rare variants. Most (279, or 97.6%) of common variants had already been reported in dbSNP or the 1000 Genomes Project, and approximately 40% were present in HapMap CEU samples. The analyses assumed an additive effect of allelic dosage, and were adjusted for age, sex and study site (if applicable) or study cohort (FHS). We found 13 common SNPs were significantly associated with CRP levels after Bonferroni correction (*P*-value cutoff = 0.05/286 ≈.05/28^{-4}), and 9 of them were located in the introns of *HNF1A* (**Supplemental Table 3**).

Selection of Rare Variant Tests

We acknowledge that the power to identify causal rare variants was still low given that they were only present in a small number of individuals. The selection of rare variant tests was based on several considerations: 1) ease of calculation and implementation, including meta-analysis; 2) whether a method can be applied to quantitative, dichotomous and survival outcomes; 3) whether the method allows for covariate adjustment; 4) power to identify variants associated with a trait, relative to other methods; 5) sufficiently-accurate type I error rate when no association is present; and 6) speed of calculation. With these considerations in mind, we decided to take two approaches to testing rare variants, collapsed aggregate statistics or joint analysis of variants using SKAT. For both approaches, statistical power is likely to be improved if non-functional variants can be discarded before testing. For coding regions, nonsynonymous SNPs are usually considered to be functional variants and thus appropriate candidates for rare variant tests. For non-coding regions, however, it is harder to decide which should be included. Although it may be beneficial to restrict to, e.g., predicted and measured transcription-factor binding sites, microRNA targets, splice sites, and DNase hypersensitive sites, this brings a risk of omitting potentially important variants with unknown mechanisms. Given the reliance of these methods on ever-shifting annotation, working groups often customized plans according to their own specific needs. We also investigated the enrichment of rare variants, which spanned a total of 83 gene regions. Seven genes harbored more than 1000 rare variants (**Supplemental Figure S1a**), whereas 5 genes harbored more than 100 nonsynonymous or splicing rare variants (**Supplemental Figure S1b**). These genes were generally longer than others. **Supplemental Figures S2a and S2b** show the genes with the greatest number of rare variants in every 1kb of gene length.

Supplemental Table 1. Summary Statistics of Identified SNPs per Sample

SNP	Total number of SNPs	Proportion of heterozygous SNPs	Transition to transversion (Ti/Tv) ratio
2.5% percentile	1328	53.7%	2.26
25% percentile	1531	61.2%	2.38
Median	1622	65.1%	2.44
75% percentile	1713	68.7%	2.51
97.5% percentile	1892	75.4%	2.65

Supplemental Table 2a: References of Targets for the Phenotype Groups

Target Name	Related References
TMEM18	[9,10]
PRRX1	[11]
CAV2	[11]
CAV1	[11]
ZFHX3	[11]
CACNB2	[12]
CYPI7AI	[12]
PLEKHA7	[12]
ATP2B1	[12]
LEPR	[8]
IL6R	[8]
GCKR	[8]
HNF1A	[8]
2q36.3	[13,14]
IRS1	[13,14]
IGF1	[13,14]
PLN	[15]
EXOG	[16]
SCN10A	[16]
MEF2C	[17]
TFR2	[18]
HK1	[18]
SH2B3	[18]
HTR4	[19]

ADAM19	[19]
WLS	[20]
MEF2C	[20]
JAG1	[20]
NINJ2	[21]
SCN5A	[22]
SOX5	[22]
C12ORF67	[22]
SLC17A4	[23]

Supplemental Table 2b: References of Targets for Pleiotropic Loci

Locus	Related References[†]
6q23.3	Blood cell traits,[18,24] Beta-thalassemia[25]
7q22.3	Platelet aggregation[26], platelet volume,[27] carotid plaque,[23] menopause,[28] eQTLs
7q36.1	Chronic kidney disease,[29] Blood cell traits,[18] Blood pressure,[30] Post-exercise heart rate,[31] eQTLs
8p21.1	vWF levels,[32] FVIII levels,[32] carotid intima-media thickness,[23] eQTLs
11p11.2	Femoral neck bone mineral density,[33] Fasting blood glucose,[13] HDL cholesterol,[34] BMI,[9] Stroke,[21] Arterial stiffness,[35] eQTLs
12q24.12-13	Blood pressure,[12] Hematocrit,[18] Eosinophilia,[36] Type I Diabetes,[37] Asthma,[36] eQTLs, others
13q34	Coronary artery disease,[30] Arterial stiffness,[38] Coronary artery calcium,[39] Weight,[40] Type II Diabetes,[30] eQTLs

[†]GWAS associations were selected from the NHGRI GWAS catalog and/or the open access GWAS database[5] with a focus on cardiovascular disease and related traits. Strong *cis*-eQTL associations were also identified and targeted in these regions from a database of published and unpublished results.

Supplemental Table 3. Most Significant Common Variants in 4 CRP Target Regions from Primary Analysis ($p<1.8x10^{-4}$)

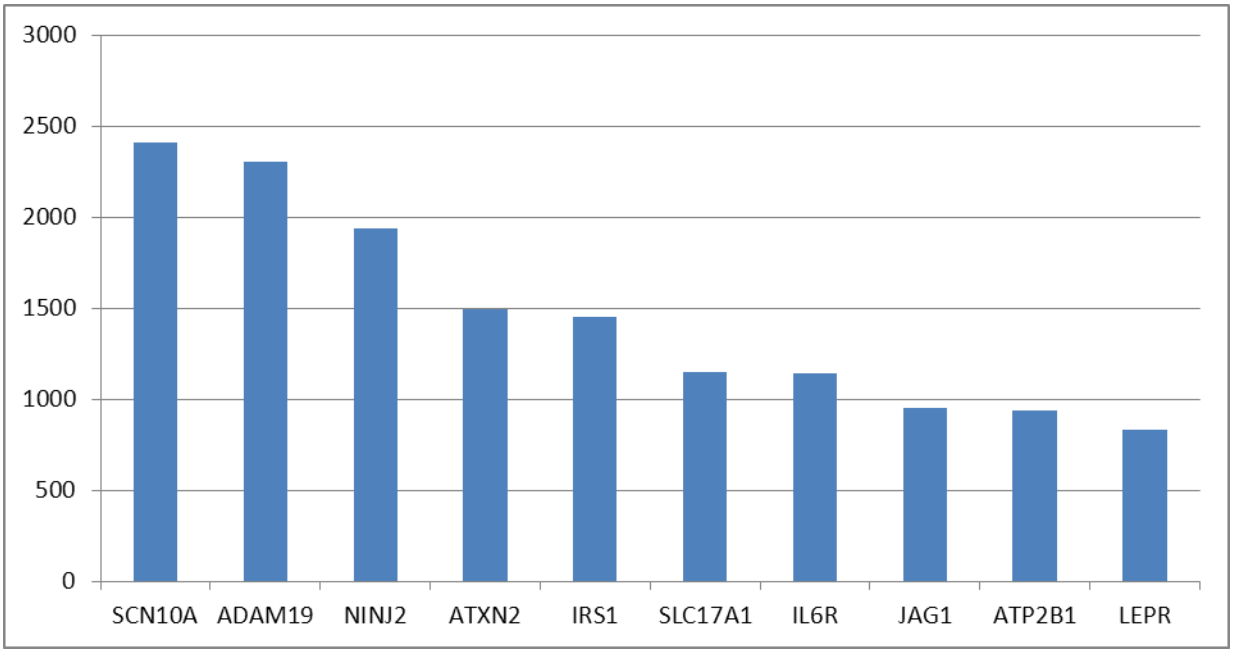
SNP	Alleles[+]	MAF	Function	Gene	Correlation with the GWAS lead SNP (r^2)*	Primary analysis[$]		
						Effect	SE	P-value
rs2244608	A/G	0.32	Intronic	*HNF1A*	0.96	0.155	0.042	$2.0x10^{-7}$
rs1169292	T/C	0.32	Intronic	*HNF1A*	0.89	-0.113	0.034	$8.7x10^{-7}$
rs1169294	A/G	0.32	Intronic	*HNF1A*	0.92	-0.121	0.033	$3.3x10^{-6}$
rs1169306	T/C	0.36	Intronic	*HNF1A*	0.45	-0.098	0.033	$3.6x10^{-6}$
rs2259816	T/G	0.36	Intronic	*HNF1A*	0.45	-0.089	0.032	$3.9x10^{-6}$
rs4129267	T/C	0.39	Intronic	*IL6R*	1	-0.132	0.031	$6.0x10^{-6}$
rs1169300	A/G	0.30	Intronic	*HNF1A*	0.63	-0.108	0.033	$6.8x10^{-6}$
rs2259852	A/G	0.36	Intronic	*HNF1A*	0.45	-0.090	0.032	$1.6x10^{-5}$
rs1169301	T/C	0.30	Intronic	*HNF1A*	0.63	-0.119	0.039	$2.0x10^{-5}$
rs4537545	T/C	0.40	Intronic	*IL6R*	1	-0.134	0.031	$2.6x10^{-5}$
rs2228145	A/C	0.39	Missense	*IL6R*	0.97	0.128	0.031	$3.9x10^{-5}$
rs1169303	A/C	0.49	Intronic	*HNF1A*	0.16	0.086	0.032	$5.1x10^{-5}$
rs7518199	A/C	0.38	Intronic	*IL6R*	0.97	0.127	0.032	$1.2x10^{-4}$

[+]The first one is the tested allele, and the second one is the other allele

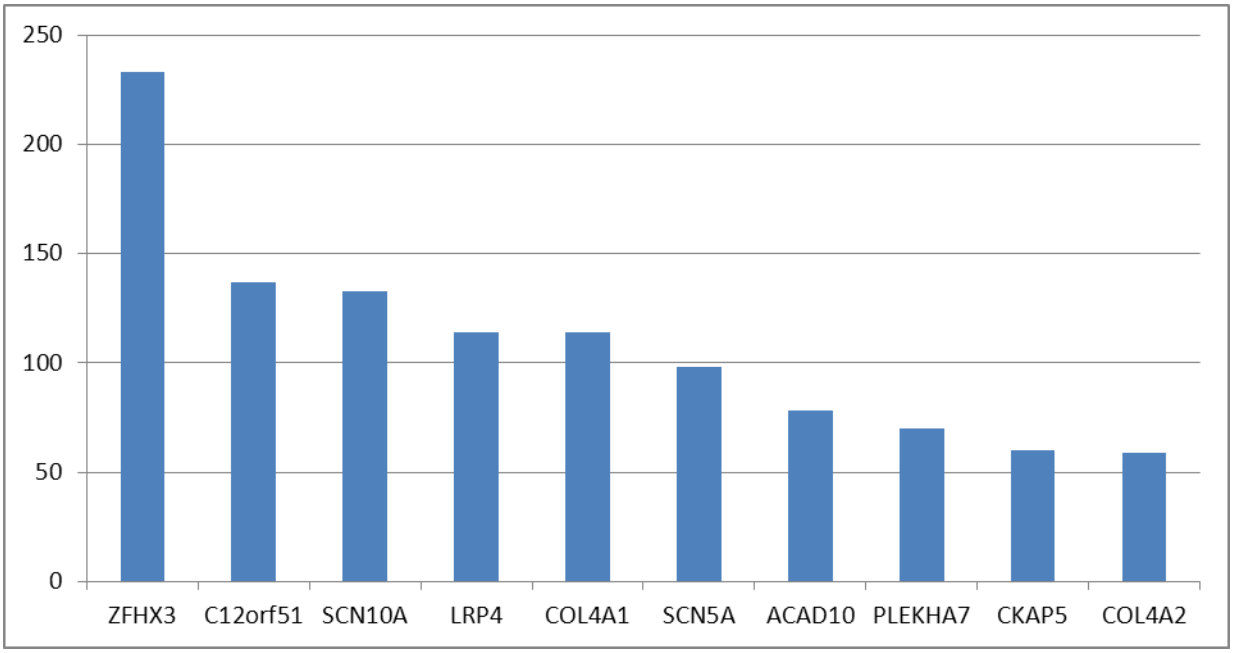
*Correlation with the GWAS lead SNP at each CRP-locus: rs4420065 (*LEPR*), rs4129267 (*IL6R*), rs1260326 (*GKCR*), and rs1183910 (*HNF1A*). It was calculated based on the 1000G Pilot 1 data.

[$]Effect and SE are from the analysis that accounts for sampling weights; *P*-value is from the unweighted analysis.

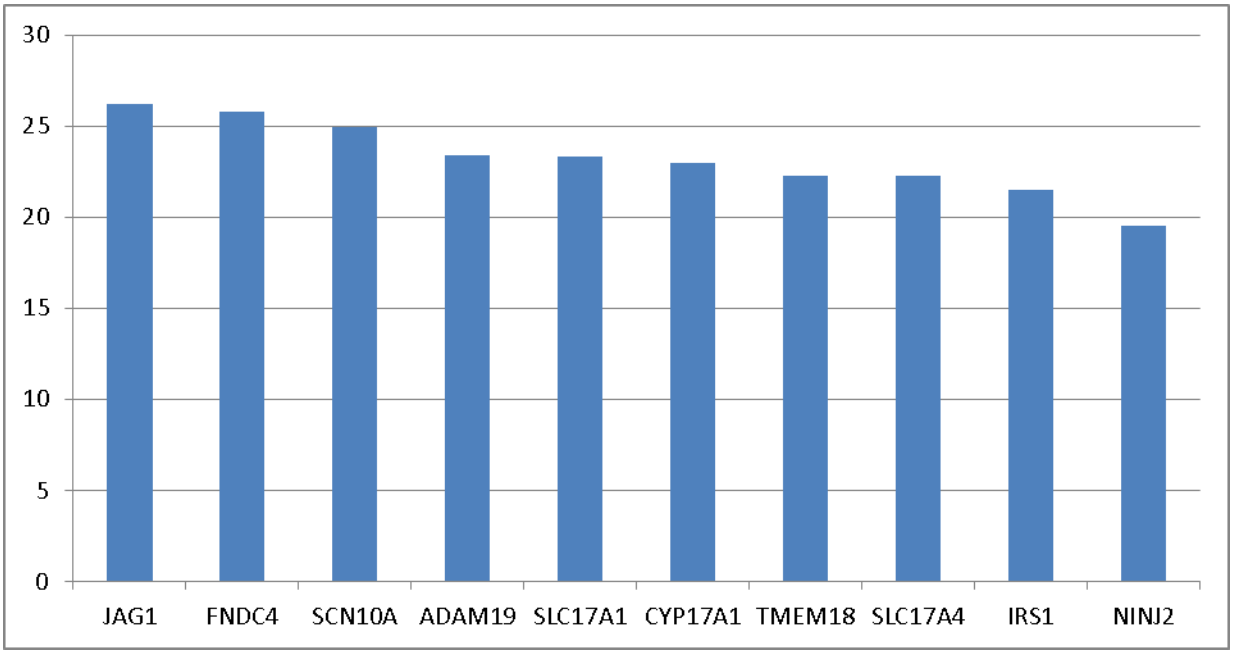
Supplemental Figure S1a. Genes with the greatest number of rare variants



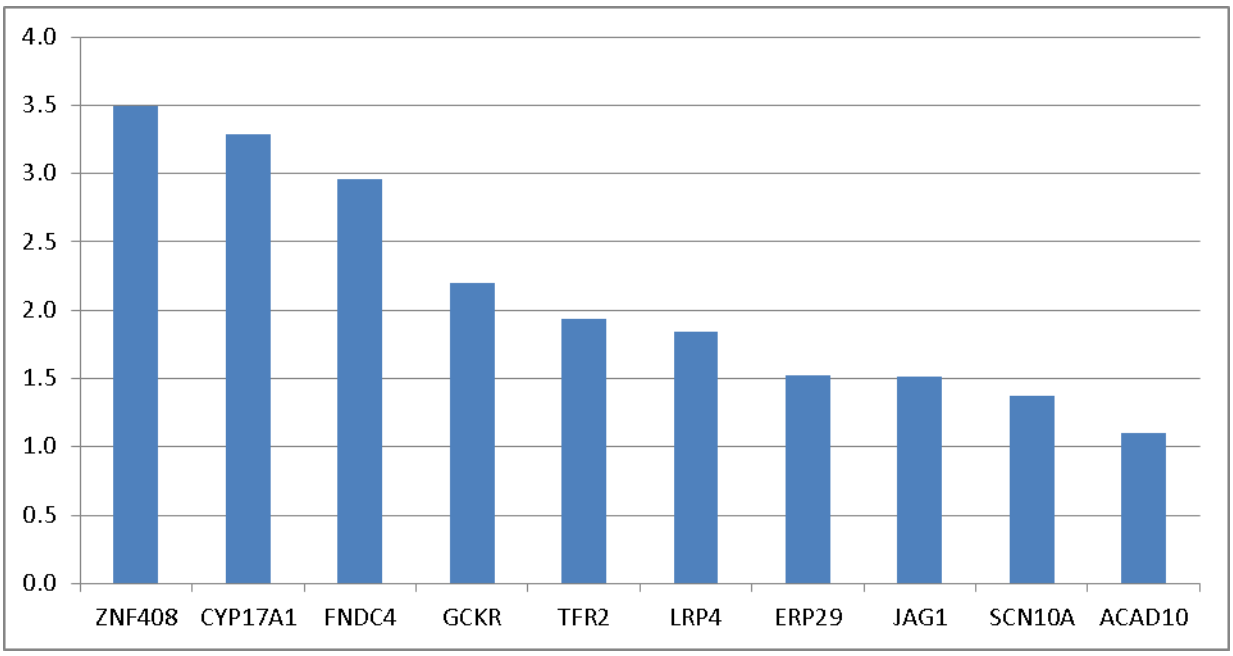
Supplemental Figure S1b. Genes with the greatest number of nonsynonymous or splicing rare variants



Supplemental Figure S2a. Genes with the greatest number of nonsynonymous or splicing rare variants in every 1kb of gene length



Supplemental Figure S2b. Genes with the greatest number of nonsynonymous or splicing rare variants in every 1kb of gene length



References

1. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol*. 1989;129:687-702
2. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol*. 1991;1:263-276
3. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol*. 1979;110:281-290
4. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol*. 2007;165:1328-1335
5. Johnson AD, O'Donnell CJ. An open access database of genome-wide association results. *BMC Med Genet*. 2009;10:6
6. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 2005;15:1034-1050
7. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res*. 2007;17:1665-1674
8. Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C, et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation*. 2011;123:731-738
9. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet*. 2010;42:937-948
10. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet*. 2009;41:25-34
11. Ellinor PT, Lunetta KL, Albert CM, Glazer NL, Ritchie MD, Smith AV, et al. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat Genet*. 2012;44:670-675
12. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, et al. Genome-wide association study of blood pressure and hypertension. *Nat Genet*. 2009;41:677-687
13. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. 2010;42:105-116
14. Soranzo N, Sanna S, Wheeler E, Gieger C, Radke D, Dupuis J, et al. Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes*. 2010;59:3229-3239
15. Vasan RS, Glazer NL, Felix JF, Lieb W, Wild PS, Felix SB, et al. Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. *JAMA*. 2009;302:168-178
16. Sotoodehnia N, Isaacs A, de Bakker PI, Dorr M, Newton-Cheh C, Nolte IM, et al. Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nat Genet*. 2010;42:1068-1076
17. Ikram MK, Sim X, Jensen RA, Cotch MF, Hewitt AW, Ikram MA, et al. Four novel Loci (19q13, 6q24, 12q24, and 5q14) influence the microcirculation in vivo. *PLoS Genet*. 2010;6:e1001184

18. Ganesh SK, Zakai NA, van Rooij FJ, Soranzo N, Smith AV, Nalls MA, et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nat Genet*. 2009;41:1191-1198
19. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet*. 2010;42:45-52
20. Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet*. 2012;44:491-501
21. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. *N Engl J Med*. 2009;360:1718-1728
22. Pfeufer A, van Noord C, Marciante KD, Arking DE, Larson MG, Smith AV, et al. Genome-wide association study of PR interval. *Nat Genet*. 2010;42:153-159
23. Bis JC, Kavousi M, Franceschini N, Isaacs A, Abecasis GR, Schminke U, et al. Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. *Nat Genet*. 2011;43:940-947
24. Soranzo N, Spector TD, Mangino M, Kuhnel B, Rendon A, Teumer A, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat Genet*. 2009;41:1182-1190
25. Nuinoon M, Makarasara W, Mushiroda T, Setianingsih I, Wahidiyat PA, Sripichai O, et al. A genome-wide association identified the common genetic variants influence disease severity in beta0-thalassemia/hemoglobin E. *Hum Genet*. 2010;127:303-314
26. Johnson AD, Yanek LR, Chen MH, Faraday N, Larson MG, Tofler G, et al. Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. *Nat Genet*. 2010;42:608-613
27. Soranzo N, Rendon A, Gieger C, Jones CI, Watkins NA, Menzel S, et al. A novel variant on chromosome 7q22.3 associated with mean platelet volume, counts, and function. *Blood*. 2009;113:3831-3837
28. Stolk L, Zhai G, van Meurs JB, Verbiest MM, Visser JA, Estrada K, et al. Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nat Genet*. 2009;41:645-647
29. Kottgen A, Pattaro C, Boger CA, Fuchsberger C, Olden M, Glazer NL, et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet*. 2010;42:376-384
30. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661-678
31. Vasan RS, Larson MG, Aragam J, Wang TJ, Mitchell GF, Kathiresan S, et al. Genome-wide association of echocardiographic dimensions, brachial artery endothelial function and treadmill exercise responses in the Framingham Heart Study. *BMC Med Genet*. 2007;8 Suppl 1:S2
32. Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation*. 2010;121:1382-1392
33. Rivadeneira F, Styrkarsdottir U, Estrada K, Halldorsson BV, Hsu YH, Richards JB, et al. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet*. 2009;41:1199-1206
34. Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, et al. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet*. 2009;41:35-46
35. Mitchell GF, Verwoert GC, Tarasov KV, Isaacs A, Smith AV, Yasmin, et al. Common genetic variation in the 3'-BCL11B gene desert is associated with carotid-femoral pulse wave velocity

and excess cardiovascular disease risk: the AortaGen Consortium. *Circ Cardiovasc Genet*. 2012;5:81-90

36. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet*. 2009;41:342-347
37. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*. 2009;41:703-707
38. Tarasov KV, Sanna S, Scuteri A, Strait JB, Orru M, Parsa A, et al. COL4A1 is associated with arterial stiffness by genome-wide association scan. *Circ Cardiovasc Genet*. 2009;2:151-158
39. O'Donnell CJ, Kavousi M, Smith AV, Kardia SL, Feitosa MF, Hwang SJ, et al. Genome-wide association study for coronary artery calcification with follow-up in myocardial infarction. *Circulation*. 2011;124:2855-2864
40. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science*. 2007;316:1331-1336