SUPPLEMENTAL MATERIAL

Supplemental Methods

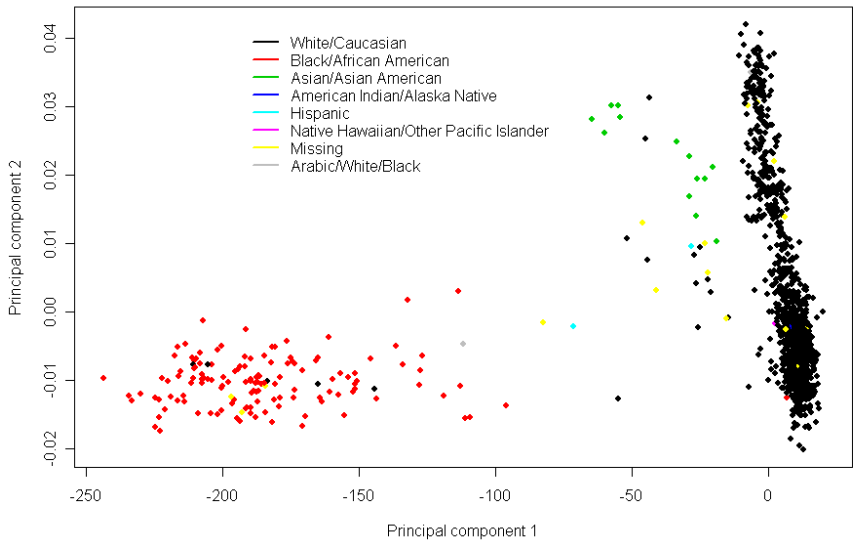
IBC Genotyping and Quality Control

DNA was isolated from peripheral blood mononuclear cells in all subjects and DNA quality was assessed utilizing optical absorbance and minigels. DNA was amplified and hybridized with IBC arrays using the Illumina platform according to manufacturer specifications. Details regarding chip design have been published[1] and a list of genes and SNPs on the array is available at http://bmic.upenn.edu/cvdsnp. Because this was a new platform, we utilized a set of 5051 samples from multiple different IBC studies to derive a customized cluster file for genotype calling using BeadStudio™ software. We checked for consistency in genotyping within each of 6 duplicate sample pairs, and the genotype concordance rate was > 99.99%. The overall genotype call rate across all samples was 99.6%.

We then performed quality control measures to exclude unreliable samples. First, samples were excluded if the genotype call rate was < 97.5% or if the sample showed excess or deficient heterozygosity (inbreeding coefficient $|F| > 0.1$). Next, we assessed for cryptic relatedness or erroneous duplicates among participants using identical by descent estimation. We also compared all IBC genotype in the Stage 1 Discovery population with similar data from populations of different ancestries and used multi-dimensional scaling (MDS) analysis to determine subjects that shared similar genetic ancestry (Supplemental Figure 1). We used this information to exclude those subjects with genetically inferred non-Caucasian ancestry from the Stage 1 analysis.

We performed quality control measures to exclude unreliable SNPs. For the analysis reported here, we eliminated SNPs with genotype call rate < 95%, with minor allele frequency (MAF) < 1% in either cases or controls or if there was significant departure from Hardy-Weinberg equilibrium ($p < 10^{-}$

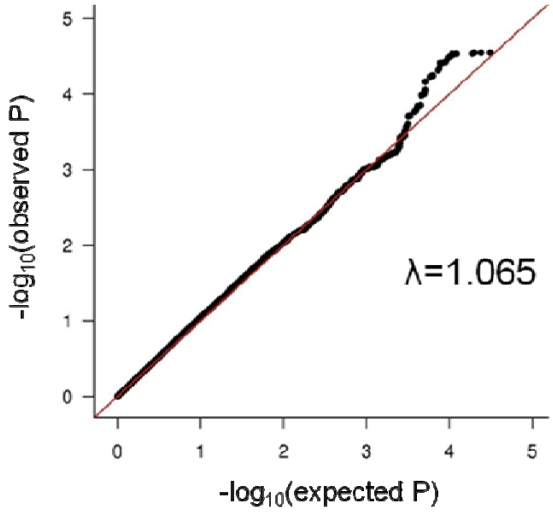
[6]). This resulted in 31,682 autosomal SNPs in our final analysis in the Stage 1 Discovery Population. See Supplemental Table 5 for Hardy-Weinberg equilibrium p-values for reported SNPs.



Supplemental Figure 1. MDS analysis of genetic race using all IBC genotypes. Self-identified race is indicated by color. We limited our IBC analysis to genetically inferred Caucasian Race.

Analysis of IBC array Data

Our primary analysis compared genotype frequencies in cases of all-cause heart failure with frequencies in controls adjusting for age, gender, study site and the first two principal components from the MDS analysis to avoid confounding by residual differences in ancestry within our Stage 1discovery cohort of genetically-inferred Caucasians. All analyses were conducted using PLINK software[2]. After association analysis was conducted, we compared the observed p-value distribution with that expected under the null distribution that assumes no association, no population stratification, and genotyping bias, and calculated the genomic control inflation factor λ, defined as the median of the observed 1-df chi-squared statistics divided by 0.455 (Supplemental Figure 2) [3]. Secondary Analyses were conducted in an analogous fashion with additional adjustment for history of hypertension and history of diabetes mellitus.



Supplemental Figure 2. QQ- plot for IBC Analysis in the Caucasian Discovery Cohort.

Power Calculations:

We estimated power using genetic power calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/). Based on a Stage 1 sample size of 1,590 cases and 577 controls and $\alpha = 5 \times 10^{-5}$, we estimate 80% power detect a OR of 1.36 or greater for a risk allele (and 80% power to detect an OR of 0.74 or less for a protective allele) for a variant with minor allele frequency of 0.40. Based on a combined Stage 1 and Stage 2 samples size of 1899 cases and 2891 controls and $\alpha = 5 \times 10^{-5}$, we estimate 80% power detect an OR of 1.20 or greater for a risk allele (and 80% power to detect an OR of 0.84 or less for a protective allele) for a variant with minor allele frequency of 0.40.

Imputation of Genotypes at Associated Loci

To better understand our two replicated SNPs associated with heart failure in Caucasians, we increased genotype coverage at these loci by imputing IBC genotypes with HapMap CEU data. We

conducted imputations using a computationally efficient hidden Markov model based algorithm as implemented in software MACH.[4]

MACH combines our genotyped data with phased chromosomes from the HapMap CEU samples and then infers the unknown genotypes in the study sample probabilistically by searching for similar stretches of flanking haplotype in the HapMap CEU reference sample. In the imputation procedure, we used the genotype data from the 31,682 IBC SNPs that passed quality control. We only analyzed SNPs that passed the following imputation QC criteria: $R^2 > 0.3$, information content > 0.5, and MAF > 0.01 in both cases and controls. Since there is uncertainty involved in the imputation, to take into account of imputation ambiguity, we analyzed case-control associations for imputed SNPs using software SNPTEST[5]. We then constructed association plots and LD plots at rs1739843 and rs6787362 using SNAP[6].

Comparison of Top Framingham heart failure SNPs with IBC proxies

We selected SNPs associated with incident heart failure from Larson et al[7] and utilized SNAP to identify proxies on the IBC array within 500 kb of these SNPs (Supplemental Table 6). Proxy SNPs at *C9orf39* and at *RYR2* were present on the IBC array within 500 kB of the Framingham hits, but with variable LD based on HapMap data. We then compared association P-values for these proxies identified in our primary analysis (which included adjustment for age, sex, site, and principal components from MDS). Proxies at *C9orf39* showed no association with heart failure, but several SNPs at *RYR2* showed modest associations in our data.

Supplemental Table 1. Top Associations with All-Cause Heart Failure in the Stage 1 Caucasian Discovery Population using the IBC array.*

Gene	Chr	Top SNP	Coordinate	Variant Type	Primary Analysis		Secondary Analysis		Minor Allele	Major Allele	Minor Allele Frequencies		Number of associated IBC SNPs
					P-value	OR (95%CI)	P-value	OR (95%CI)			Cases	Controls	
HSPB7	**1**	**rs1739843**	**16215841**	**Intronic**	**2.83E-05**	**0.75 (0.65, 0.85)**	7.66E-06	0.71 (0.62, 0.83)	A	G	0.367	0.436	1
FRMD4B	**3**	**rs6787362**	**69310069**	**Intronic**	**4.51E-05**	**0.65 (0.53, 0.80)**	3.38E-04	0.67 (0.53, 0.83)	G	A	0.098	0.140	1
PCSK6	**15**	**rs7174882**	**99836368**	**Intronic**	**5.70E-05**	**0.76 (0.66, 0.87)**	5.74E-04	0.78 (0.67, 0.89)	C	A	0.449	0.511	2
PDE4D	**5**	**rs16877169**	**58813004**	**Intronic**	**5.90E-05**	**0.63 (0.51, 0.79)**	4.74E-04	0.65 (0.52, 0.83)	A	C	0.075	0.119	4
ARHGAP26	5	rs17099475	142138783	Intronic	1.41E-04	1.64 (1.27, 2.11)	2.28E-05	1.77 (1.36, 2.31)	G	A	0.109	0.071	1
GAS7	17	rs16959235	9862480	Intronic	1.60E-04	0.65 (0.52, 0.81)	2.96E-04	0.65 (0.52, 0.81)	A	G	0.079	0.115	1
SORCS1	10	rs12360161	108592969	Intronic	1.71E-04	0.55 (0.40, 0.75)	4.62E-04	0.55 (0.40, 0.75)	C	A	0.035	0.062	1
CASP8	2	rs6435072	201824041	Intronic	1.97E-04	0.67 (0.55, 0.83)	1.18E-03	0.67 (0.55, 0.83)	A	C	0.102	0.141	2
RNF216	7	rs12668000	5796948	Intergenic	2.65E-04	0.76 (0.66, 0.88)	7.53E-04	0.76 (0.66, 0.88)	G	C	0.311	0.371	1
FAM174A	5	rs12514622	100022279	Intergenic	4.48E-04	0.78 (0.67, 0.89)	7.72E-04	0.78 (0.67, 0.89)	A	G	0.298	0.354	1
AL359649.8	13	rs1183183	110934706	Intergenic	4.94E-04	0.75 (0.64, 0.88)	3.12E-04	0.75 (0.64, 0.88)	A	C	0.197	0.248	1
MMP3	11	rs520540	102214635	Coding S	5.00E-04	0.79 (0.69, 0.90)	2.57E-03	0.79 (0.69, 0.90)	A	G	0.485	0.541	2
NOS1AP	1	rs10918594	160297312	Intergenic	5.94E-04	0.78 (0.67, 0.90)	1.12E-02	0.78 (0.67, 0.90)	C	G	0.304	0.366	2
LIF	22	rs715605	28970308	Intronic	6.01E-04	0.68 (0.55, 0.85)	5.20E-05	0.68 (0.55, 0.85)	G	A	0.086	0.123	1
POLR2A	17	rs2269459	7353762	Intronic	6.58E-04	0.77 (0.66, 0.89)	2.95E-03	0.77 (0.66, 0.89)	A	G	0.247	0.297	1
CACNA1C	12	rs16929388	2405678	Intronic	7.49E-04	1.72 (1.25, 2.36)	4.70E-03	1.61 (1.16, 2.24)	G	A	0.075	0.044	1
ITGB4	17	rs3862481	71232322	Intronic	7.59E-04	0.74 (0.61, 0.88)	2.82E-04	0.77 (0.66, 0.90)	A	G	0.144	0.185	1
NTRK2	9	rs1443439	86712230	Intronic	8.80E-04	0.53 (0.37, 0.78)	2.80E-03	0.73 (0.62, 0.87)	A	C	0.025	0.043	1
PRKAG2	7	rs2727537	151003979	Intronic	9.08E-04	0.79 (0.68, 0.91)	1.39E-03	0.80 (0.69, 0.93)	G	A	0.344	0.400	1
MMP13	11	rs562736	102332123	Upstream	9.15E-04	1.27 (1.10, 1.46)	3.26E-03	1.25 (1.08, 1.45)	A	T	0.455	0.402	2
ADRA1A	8	rs12541572	26764903	Intronic Reg	9.28E-04	1.45 (1.16, 1.81)	6.25E-04	1.50 (1.19, 1.89)	A	G	0.140	0.100	1
F13A1	6	rs1742932	6249530	Intronic	9.74E-04	1.30 (1.11, 1.51)	1.54E-04	1.37 (1.16, 1.61)	G	A	0.306	0.253	1

*For each gene data for the most strongly associated SNP is summarized. OR are for the minor allele. The primary analysis is adjusted for age, sex, study site, and MDS. The secondary analysis is additionally adjusted for history of diabetes and hypertension. SNPs selected for Stage 2 are in bold.

Supplemental Table 2. Top Associations with Heart Failure Subtypes in the Stage 1 Caucasian Discovery Population using the IBC array.*

Gene	Chr	Top SNP	Coordinate	Variant Type	Primary Analysis		Secondary Analysis		Minor Allele	Major Allele	Minor Allele Frequencies		Number of associated IBC SNPs
					P-value	OR (95%CI)	P-value	OR (95%CI)			Cases	Controls	
Ischemic Heart Failure													
PCSK6	**15**	**rs7174882**	**99836368**	**Intronic**	**1.56E-05**	**0.70 (0.60, 0.83)**	5.70E-05	0.69 (0.57, 0.83)	C	A	0.432	0.511	2
F2RL1	5	rs2243073	76151489	Intronic	1.69E-04	1.73 (1.30, 2.30)	3.83E-04	1.79 (1.30, 2.46)	A	T	0.111	0.072	1
AL445929.5	13	rs9544230	75717997	Intergenic	1.83E-04	1.43 (1.18, 1.72)	1.66E-04	1.50 (1.22, 1.85)	A	G	0.290	0.227	1
PDE4D	5	rs16877169	58813004	Intronic	2.16E-04	0.60 (0.46, 0.79)	1.55E-03	0.61 (0.44, 0.83)	A	C	0.070	0.119	1
TRBV2	7	rs11764352	141644584	Intronic	3.63E-04	0.72 (0.60, 0.86)	2.26E-03	0.73 (0.60, 0.89)	A	T	0.258	0.321	1
EVC	4	rs7674034	5793954	Intronic	4.21E-04	0.73 (0.61, 0.87)	4.75E-04	0.70 (0.58, 0.85)	G	A	0.299	0.360	1
F7	13	rs6041	112820708	Intronic	7.41E-04	0.65 (0.51, 0.83)	1.23E-02	0.70 (0.53, 0.93)	A	G	0.108	0.140	3
STAT3	17	rs2306580	37745206	Intronic	9.83E-04	1.73 (1.25, 2.40)	3.28E-03	1.73 (1.20, 2.50)	G	C	0.091	0.053	1
Nonischemic Heart Failure													
LIPC	15	rs4581654	56600858	Intronic	4.85E-05	1.50 (1.23, 1.82)	2.78E-04	1.46 (1.19, 1.78)	A	C	0.264	0.199	3
HSPB7	1	rs1739843	16215841	Intronic	1.46E-04	0.73 (0.62, 0.85)	1.50E-04	0.72 (0.61, 0.85)	A	G	0.358	0.436	1
LIF	22	rs715605	28970308	Intronic	1.75E-04	0.59 (0.45, 0.78)	2.21E-04	0.58 (0.44, 0.78)	G	A	0.076	0.123	1
ABO	9	rs8176694	135127467	Intronic	2.02E-04	1.49 (1.21, 1.85)	7.37E-04	1.46 (1.17, 1.82)	G	A	0.219	0.163	2
GPX2	14	rs2737844	64478262	Intronic	3.09E-04	0.72 (0.61, 0.86)	1.64E-03	0.75 (0.62, 0.89)	A	G	0.255	0.320	2
TNFSF4	1	rs1234313	171432870	Intronic	5.56E-04	1.37 (1.14, 1.64)	1.67E-03	1.35 (1.12, 1.63)	A	G	0.335	0.266	1
ADRA1A	8	rs12541572	26764903	Intronic Reg	6.10E-04	1.55 (1.21, 1.98)	5.58E-04	1.58 (1.22, 2.05)	A	G	0.148	0.100	1
IGF1R	15	rs1546713	97289389	Intronic	6.61E-04	0.74 (0.62, 0.88)	1.47E-03	0.74 (0.62, 0.89)	A	G	0.289	0.352	1
RYR2	1	rs2618717	235798955	Intronic	8.30E-04	1.36 (1.13, 1.62)	5.62E-04	1.39 (1.15, 1.67)	G	A	0.320	0.261	1
CALU	7	rs4731513	128167231	Intronic	8.90E-04	0.75 (0.63, 0.88)	1.77E-03	0.75 (0.63, 0.90)	G	A	0.330	0.393	1
ATP6V1B1	2	rs2239484	71026916	Intronic	9.13E-04	1.58 (1.21, 2.07)	3.81E-04	1.67 (1.26, 2.21)	G	A	0.130	0.086	1
ANGPT2	8	rs11989215	6383317	Intronic	9.42E-04	1.33 (1.12, 1.58)	2.31E-04	1.40 (1.17, 1.68)	G	A	0.368	0.300	1
POLR2A	17	rs2269459	7353762	Intronic	9.60E-04	0.74 (0.61, 0.88)	2.20E-03	0.74 (0.61, 0.90)	A	G	0.236	0.297	1

*For each gene data for the most strongly associated SNP is summarized. Odds Ratios are for the minor allele. The primary analysis is adjusted for age, sex, study site, and MDS. The secondary analysis is additionally adjusted for history of diabetes and hypertension. SNPs selected for Stage 2 are in bold.

Supplemental Table 3. Associations in African Americans

Gene	Chr	SNP	SNP type	Minor/Major allele	HapMap YRI MAF	Control MAF	Case MAF	OR	P-value
HSPB7	1	rs1739843	Intronic	A/G	0.23	0.281	0.300	1.09 (0.92, 1.29)	0.34
FRMD4B	3	rs6787362	Intronic	G/A	0.093	0.099	0.083	0.83 (0.64, 1.10)	0.19
PCSK6	15	rs7174882	Intronic	C/A	0.422	0.452	0.474	1.09 (0.93, 1.28)	0.29
PDE4D	5	rs16877169	Intronic	A/C	0.441	0.339	0.304	0.84 (0.72, 0.99)	0.039

*N= 635 cases, 714 controls. OR is for the minor allele, adjusted for age, sex, and study site.

Supplemental Table 4. Analysis of subgroups and intermediate phenotypes in African Americans

Gene	SNP	Subgroup	OR	P	$OR_{DM, HTN}$	$P_{DM, HTN}$
PDE4D	rs16877169	All	0.84 (0.72, 0.99)	0.039	0.84 (0.69, 1.02)	.07
		Ischemic	0.92 (0.74, 1.13)	0.43	0.91 (0.72, 1.15)	.43
		Nonischemic	0.78 (0.63, 0.95)	.015	0.77 (0.63, 1.00)	.054

n = 635 cases (300 ischemic, 324 nonischemic, 11 other) and 714 controls.
OR is for the minor allele, adjusted for age, sex, and study site. $OR_{DM, HTN}$ is additionally adjusted for history of diabetes and hypertension

Supplemental Table 5.

Gene	SNP	HWE P-values		
		Stage 1 Caucasians	Stage 2 Caucasians	African Americans
HSPB7	**rs1739843**	**0.47**	**0.57**	**0.92**
FRMD4B	**rs6787362**	**0.66**	**0.77**	**0.30**
PCSK6	**rs7174882**	**0.11**	**0.91**	**0.37**
PDE4D	**rs16877169**	**0.14**	**0.54**	**0.19**
ABO	rs8176694	0.73	-	-
ADRA1A	rs12541572	0.70	-	-
AL359649.8	rs1183183	0.08	-	-
AL445929.5	rs9544230	0.31	-	-
ANGPT2	rs11989215	0.96	-	-
ARHGAP26	rs17099475	0.63	-	-
ATP6V1B1	rs2239484	0.91	-	-
CACNA1C	rs16929388	0.49	-	-
CALU	rs4731513	0.06	-	-
CASP8	rs6435072	0.08	-	-
EVC	rs7674034	0.38	-	-
F13A1	rs1742932	0.47	-	-
F2RL1	rs2243073	0.05	-	-
F7	rs6041	0.42	-	-
FAM174A	rs12514622	0.16	-	-
GAS7	rs16959235	0.42	-	-
GPX2	rs2737844	0.79	-	-
IGF1R	rs1546713	0.49	-	-
ITGB4	rs3862481	0.57	-	-
LIF	rs715605	1.0	-	-
LIPC	rs4581654	0.61	-	-
MMP13	rs562736	0.32	-	-
MMP3	rs520540	0.79	-	-
NOS1AP	rs10918594	0.84	-	-
NTRK2	rs1443439	1.0	-	-
POLR2A	rs2269459	0.91	-	-
POLR2A	rs2269459	0.91	-	-
PRKAG2	rs2727537	0.61	-	-
RNF216	rs12668000	0.53	-	-
RYR2	rs2618717	0.79	-	-
SORCS1	rs12360161	1.0	-	-
STAT3	rs2306580	0.63	-	-
TNFSF4	rs1234313	0.54	-	-
TRBV2	rs11764352	0.79	-	-

Supplemental Table 6. Comparison with Published Framingham 100K Heart Failure GWAS Findings

SNP	Chr	Location	Nearest Gene	Framingham Association Method	Framingham P-value	IBC proxies (< 500 kB)	Distance (kB)	R^2	IBC P-values		
									All Cause Heart Failure	Ischemic Heart Failure	Nonischemic Heart Failure
rs740363	10	118565596	*KIAA1598*	FBAT	8.82E-06	rs992528	413	0.016	-	-	-
rs10511633	9	17151527	*C9orf39*	FBAT	2.59E-05	rs1442526	234	0.712	0.45	0.67	0.38
						rs2499053	326	0.34	0.43	0.90	0.12
						rs2499054	326	0.274	0.86	0.76	0.50
						rs10963034	162	0.036	0.45	0.82	0.33
						rs10963045	181	0.036	0.45	0.93	0.35
rs10515869	5	163444804	-	GEE	4.72E-05	none	-	-	-	-	-
rs1176486	10	132315529	-	GEE	1.49E-04	none	-	-	-	-	-
rs9313999	5	163444569	-	GEE	1.55E-04	none	-	-	-	-	-
rs939698	1	235857303	*RYR2*	GEE	3.60E-04	rs663152	2	1	0.161	0.095	0.48
						rs1759123	3	0.194	0.015	0.014	0.046
						rs2618717	58	0.007	0.016	0.560	0.0008
						rs6663810	372	0.001	0.060	0.006	0.47

References

(1) Keating BJ, Tischfield S, Murray SS, Bhangale T, Price TS, Glessner JT, Galver L, Barrett JC, Grant SF, Farlow DN, Chandrupatla HR, Hansen M, Ajmal S, Papanicolaou GJ, Guo Y, Li M, Derohannessian S, de Bakker PI, Bailey SD, Montpetit A, Edmondson AC, Taylor K, Gai X, Wang SS, Fornage M, Shaikh T, Groop L, Boehnke M, Hall AS, Hattersley AT, Frackelton E, Patterson N, Chiang CW, Kim CE, Fabsitz RR, Ouwehand W, Price AL, Munroe P, Caulfield M, Drake T, Boerwinkle E, Reich D, Whitehead AS, Cappola TP, Samani NJ, Lusis AJ, Schadt E, Wilson JG, Koenig W, McCarthy MI, Kathiresan S, Gabriel SB, Hakonarson H, Anand SS, Reilly M, Engert JC, Nickerson DA, Rader DJ, Hirschhorn JN, Fitzgerald GA. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS ONE* 2008;3(10):e3583.

(2) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007 September;81(3):559-75.

(3) Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999 December;55(4):997-1004.

(4) Li Y, Abecasis GR. Mach 1.0: Rapid Haplotype Reconstruction and Missing Genotype Inference. *Am J Hum Genet* 2006 March;S79(3):2290.

(5) Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007 July;39(7):906-13.

(6) Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008 December 15;24(24):2938-9.

(7) Larson MG, Atwood LD, Benjamin EJ, Cupples LA, D'Agostino RB, Sr., Fox CS, Govindaraju DR, Guo CY, Heard-Costa NL, Hwang SJ, Murabito JM, Newton-Cheh C, O'Donnell CJ, Seshadri S, Vasan RS, Wang TJ, Wolf PA, Levy D. Framingham Heart Study 100K project: genome-wide associations for cardiovascular disease outcomes. *BMC Med Genet* 2007;8 Suppl 1:S5.