nature genetics

Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants

Myocardial Infarction Genetics Consortium*

We conducted a genome-wide association study testing single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) for association with early-onset myocardial infarction in 2,967 cases and 3,075 controls. We carried out replication in an independent sample with an effective sample size of up to 19,492. SNPs at nine loci reached genome-wide significance: three are newly identified (21q22 near *MRPS6-SLC5A3-KCNE2*, 6p24 in *PHACTR1* and 2q33 in *WDR12*) and six replicated prior observations[1–4] (9p21, 1p13 near *CELSR2-PSRC1-SORT1*, 10q11 near *CXCL12*, 1q41 in *MIA3*, 19p13 near *LDLR* and 1p32 near *PCSK9*). We tested 554 common copy number polymorphisms (>1% allele frequency) and none met the pre-specified threshold for replication ($P < 10^{-3}$). We identified 8,065 rare CNVs but did not detect a greater CNV burden in cases compared to controls, in genes compared to the genome as a whole, or at any individual locus. SNPs at nine loci were reproducibly associated with myocardial infarction, but tests of common and rare CNVs failed to identify additional associations with myocardial infarction risk.

Myocardial infarction is a leading cause of death and disability worldwide[5], with family history being an independent risk factor[6]. The inherited basis for myocardial infarction remains incompletely understood. Whereas the majority of myocardial infarctions occur in individuals >65 y old, 1–5% of younger individuals report a history of myocardial infarction[5,7]. These latter events are associated with substantially greater heritability[8]. Thus, early-onset myocardial infarction is a promising phenotype for genetic mapping.

Genome-wide association studies (GWASs) of common SNPs have been reported for myocardial infarction and coronary artery disease (CAD), with each study finding common SNPs on chromosome 9p21.3 associated with myocardial infarction or CAD[1–3]. In addition to 9p21.3, Samani *et al.* reported six other loci as harboring SNPs associated with CAD[3]. Some of these loci await definitive replication, but even if all were valid, they would explain a small fraction of the risk for myocardial infarction.

Structural variants, another class of human DNA sequence variation, may account for some of the unexplained heritability in myocardial infarction and other common diseases[9]. To our knowledge, no integrated assessment of SNPs and CNVs in the same samples has been reported for myocardial infarction or any other trait. Several technological developments make such systematic surveys now possible, including hybrid oligonucleotide microarrays[10] and analytical methods[11] to simultaneously assess SNPs and CNVs genome-wide in each sample.

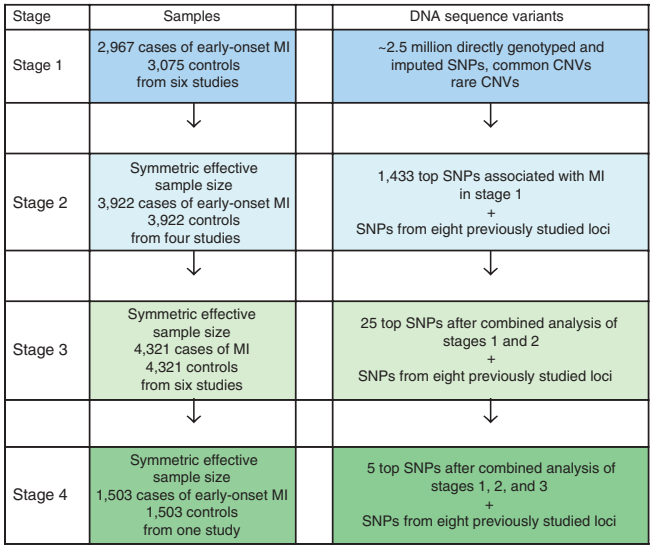

Figure 1 Study design. The GWAS consisted of four stages with an evaluation of common SNPs, common CNPs and rare CNVs in stage 1. The design called for all variants with a $P < 0.001$ to be taken forward to stage 2. As only SNPs met this criterion, 1,441 SNPs were taken forward to stage 2. Thirty-three SNPs were tested in stage 3. Thirteen SNPs were tested in stage 4. Statistical evidence for association was combined across stages 1–4 using meta-analysis.

*The complete list of participants and affiliations is provided at the end of this article. Correspondence should be addressed to S.K. (skathiresan1@partners.org).

Received 13 November 2008; accepted 16 January 2009; published online 8 February 2009; corrected after print 27 May 2009; doi:10.1038/ng.327

Table 1 Participant characteristics of case and control subjects in stage 1 of the GWAS

Study	Italian ATVB Study		Heart Attack Risk in Puget Sound		REGICOR		MGH Premature Coronary Artery Disease Study		FINRISK		Malmö Diet and Cancer Study	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
n	1,693	1,668	505	559	312	317	204	260	167	172	86	99
Ascertainment scheme	Hospital-based	Hospital-based	Community-based	Community-based	Hospital-based	Drawn from community-based cohort	Hospital-based	Hospital-based	Drawn from population-based cohort	Nested case-cohort	Drawn from population-based cohort	Nested case-cohort
Myocardial infarction age criterion	Men or women ≤45	—	Men ≤50 or women ≤60	—	Men ≤50 or women ≤60	—	Men ≤50 or women ≤60	—	Men ≤50 or women ≤60	—	Men ≤50 or women ≤60	—
Country of origin[a]	Italy	Italy	US	US	Spain	Spain	US	US	Finland	Finland	Sweden	Sweden
Mean age (y)[b]	39.4 ± 4.9	39.3 ± 5.0	46.0 ± 6.9	45.2 ± 7.3	45.9 ± 5.8	46.0 ± 5.6	47.0 ± 6.1	53.8 ± 11.1	47.1 ± 6.2	47.1 ± 6.0	48.5 ± 4.4	48.7 ± 4.6
>Female gender (%)	11.4	11.6	51.1	55.5	20.2	21.5	29.9	33.5	33.5	31.4	41.9	42.4
Ever cigarette smoking (%)	87.0	49.3	73.9	41.7	82.8	61.9	74.9	57.3	74.4	58.2	87.2	61.6
Hypertension (%)[c]	32.6	11.9	50.5	30.8	38.0	31.5	33.5	25.3	72.5	68.0	81.4	62.6
Diabetes mellitus (%)[d]	7.8	0.8	14.9	3.0	14.8	6.1	19.2	0.4	17.7	5.9	4.7	1.0
Hypercholesterolemia (%)[e]	60.4	44.4	43.7	26.0	48.9	33.1	79.0	31.3	75.2	48.2	37.2	1.0
Body mass index (kg/m^2)	26.7 ± 4.2	25.0 ± 3.3	29.2 ± 6.8	26.9 ± 5.7	27.5 ± 4.2	27.0 ± 3.9	30.0 ± 7.0	27.9 ± 6.5	29.6 ± 5.0	27.7 ± 4.0	26.9 ± 4.2	25.7 ± 4.3

Values with '±' are means ± s.d. The body-mass index is the weight in kilograms divided by the square of the height in meters.
[a]All cases and controls were of European ancestry. [b]Mean age at myocardial infarction for cases and at age of recruitment for controls. [c]Hypertension was defined as a previous diagnosis of hypertension, on antihypertensive therapy or with recorded systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg. [d]Diabetes mellitus was defined as a previous diagnosis of diabetes or treatment with antidiabetic medications. [e]Hypercholesterolemia was defined as a previous diagnosis of hypercholesterolemia or treatment with lipid-lowering therapy.

We designed a four-stage GWAS of early-onset myocardial infarction with SNPs, common copy number polymorphisms (CNPs) and rare CNVs (**Fig. 1**). Stage 1 consisted of the Myocardial Infarction Genetics Consortium (MIGen), a collection of 2,967 cases of early-onset myocardial infarction (in men ≤50 y old or women ≤60 y old) and 3,075 age- and sex-matched controls free of myocardial infarction from six international sites: Boston and Seattle in the United States, as well as Sweden, Finland, Spain and Italy (**Table 1** and **Supplementary Methods** online). The mean age at the time of myocardial infarction was 41 y among males and 47 y among females. Variants with $P < 0.001$ were advanced through three stages of replication (**Fig. 1**; see Methods for power calculations). Descriptions of the replication studies are provided in **Supplementary Methods** and **Supplementary Tables 1** and **2** online.

After stages 1–4, we observed that SNPs at nine loci were associated with myocardial infarction at a pre-specified threshold for genome-wide significance of $P < 5 \times 10^{-8}$ (corresponding to $P < 0.05$ after adjusting for ~1 million independent tests[12]) (**Tables 2** and **3**). Of these nine, four represent confirmation of associations previously reported by Samani *et al.*[3] (**Table 2**). These four genetic association signals map to 9p21, 1p13 near *CELSR2-PSRC1-SORT1*, 10q11 near *CXCL12* and 1q41 in *MIA3*. In samples fully independent of the two original discovery studies (Wellcome Trust Case Control Consortium and German MI Family Study I), the statistical evidence for these four variants was robust, with the same allele associated in the same direction as the original report (replication *P* ranging from 3×10^{-5} to 1×10^{-30}; **Table 2**).

Three of the loci previously suggested by Samani *et al.* did not replicate (**Table 2**). In samples independent of the two original discovery studies, the statistical evidence for these loci was the following: rs6922269 in *MTHFD1L* (OR = 1.04, 95% CI = 0.99–1.09, $P = 0.08$); rs17228212 in *SMAD3* (OR = 1.01, 95% CI = 0.96–1.05, $P = 0.69$) and rs2943634 on 2q36 (OR = 0.94, 95% CI = 0.90–0.98, $P = 0.01$).

Three previously unreported associations were observed with genome-wide significance (**Table 3**): (i) in an intergenic region between *MRPS6* (mitochondrial ribosomal protein S6), *SLC5A3* (solute carrier family 5 (inositol transporters) member 3) and *KCNE2* (potassium voltage-gated channel, Isk-related family, member 2) on chromosome 21q22 (rs9982601, OR = 1.19, $P = 6 \times 10^{-11}$); (ii) in an intron of *PHACTR1* (phophastase and actin regulator 1) on chromosome 6p24 (rs12526453, OR = 1.13, $P = 1 \times 10^{-9}$); and (iii) in an intron of *WDR12* (WD repeat domain 12) on chromosome 2q33 (rs6725887, OR = 1.17, $P = 1 \times 10^{-8}$).

MRPS6 encodes a subunit of the mitochondrial ribosomal protein 28S[13]. *SLC5A3* is a gene embedded within *MRPS6* and encodes a protein that transports sodium and myo-inositol in response to hypertonic stress[14]. *KCNE2* encodes a subunit of a potassium channel, and mutations in this gene cause inherited arrhythmias[15]. *PHACTR1* is an inhibitor of protein phosphatase 1, an enzyme that dephosphorylates serine and threonine residues on a range of proteins[16]. *WDR12* has been shown to complex with several proteins to enable ribosome biogenesis and cell proliferation[17]. The mechanisms by which gene(s) at these three genomic regions confer increased risk of myocardial infarction remain to be defined.

Of note, the *PHACTR1* locus may lead to myocardial infarction by directly promoting the development of atherosclerosis in the coronary arteries. In an independent GWAS for coronary artery calcification in >10,000 participants from six prospective cohort studies, *PHACTR1* SNPs (along with chromosome 9p21 SNPs) are associated with coronary artery calcification at genome-wide significance (C.J. O'Donnell, National Heart, Lung and Blood Institute, personal communication).

Of the nine loci with convincing association evidence, the remaining two (19p13 near *LDLR* and 1p32 near *PCSK9*) relate to a causal risk factor for myocardial infarction: low-density lipoprotein (LDL) cholesterol. Common, low-frequency and/or rare mutations at *LDLR* and *PCSK9* have previously been shown to influence LDL cholesterol and consequently affect risk for myocardial infarction[18–22]. We confirm that common variants near *LDLR* and *PCSK9* are associated with risk for myocardial infarction. The specific alleles (*LDLR* rs1122608 and *PCSK9* rs11206510) corresponding to higher risk for myocardial

Table 2 Replication evidence for seven previously reported common variants associated with myocardial infarction

Studies (maximum available effective sample size)		Previously reported SNPs with convincing replication evidence				Previously reported SNPs without convincing replication evidence		
	SNP	rs4977574	rs646776	rs17465637	rs1746048	rs6922269	rs17228212	rs2943634
	Chr.	9p21	1p13	1q41	10q11	6q25	15q22	2q36
	Position NCBI35 (bp)	22,088,574	109,530,572	220,890,152	44,095,830	151,294,678	65,245,693	226,776,324
	Nonrisk allele	A	C	A	T	G	C	A
	Risk allele	G	T	C	C	A[a]	T[a]	C[a]
	Risk allele frequency	0.56	0.81	0.72	0.84	0.26[a]	0.73[a]	0.66[a]
	Gene(s) of interest in associated interval	*CDKN2A-CDKN2B*	*CELSR2-PSRC1-SORT1*	*MIA3*	*CXCL12*	*MTHFD1L*	*SMAD3*	None
Stage 1 MIGen (n = 6,046)	OR[b] (95% CI)	1.25 (1.16–1.34)	1.11 (1.02–1.22)	1.17 (1.08–1.27)	1.22 (1.10–1.34)	1.08 (1.00–1.17)	0.98 (0.91–1.06)	0.95 (0.88–1.02)
	P[c]	6.7×10^{-9}	0.040	1.5×10^{-4}	1.6×10^{-4}	0.07	0.68	0.18
Stage 2 PennCATH, MedSTAR (n = 1,750)	OR (95% CI)	1.64 (1.37–1.96)	1.33 (1.09–1.64)	1.09 (0.90–1.33)	1.27 (0.99–1. 64)	1.19 (0.98–1.46)	1.11 (0.91–1.36)	0.99 (0.82–1.20)
	P	5.2×10^{-8}	0.006	0.38	0.06	0.09	0.29	0.92
Stage 3 AMI Gene, Verona, MAHI, IFS, GerMIFS II, INTERHEART (n = 8,642)	OR (95% CI)	1.24 (1.17–1.32)	1.18 (1.08–1.29)	1.11 (1.04–1.19)	1.10 (1.00–1.20)	1.02 (0.95–1.09)	0.97 (0.90–1.04)	0.93 (0.87–1.00)
	P	4.2×10^{-12}	9.2×10^{-5}	0.003	0.046	0.63	0.35	0.048
Stage 4 deCODE (n = 3,006)	OR (95% CI)	1.34 (1.20–1.49)	1.22 (1.06–1.40)	1.11 (0.98 – 1.25)	1.04 (0.87–1.23)	0.98 (0.87–1.11)	1.16 (1.03–1.31)	0.93 (0.83–1.04)
	P	2.4×10^{-7}	0.004	0.10	0.70	0.77	0.02	0.20
Stages 1, 2, 3 + 4 excluding original discovery studies (n = 19,444)	OR (95% CI)	1.28 (1.23–1.33)	1.17 (1.11–1.24)	1.13 (1.08–1.18)	1.14 (1.08–1.21)	1.04 (0.99–1.09)	1.01 (0.96–1.05)	0.94 (0.90–0.98)
	P	1.1×10^{-30}	1.5×10^{-8}	4.9×10^{-7}	3.4×10^{-5}	0.08	0.69	0.01
Stages 1, 2, 3 + 4 including original discovery studies (WTCCC and GerMIFS I)[d] (n = 25,538)	OR (95% CI)	1.29 (1.25–1.34)	1.19 (1.13–1.26)	1.14 (1.10–1.19)	1.17 (1.11–1.24)	1.09 (1.05–1.14)	1.05 (1.01–1.09)	0.95 (0.91–0.98)
	P	2.7×10^{-44}	7.9×10^{-12}	1.4×10^{-9}	7.4×10^{-9}	2.3×10^{-5}	0.02	0.005

[a]Risk allele in two original discovery studies (WTCCC and GerMIFS I) is displayed. For this risk allele, we present the statistical evidence for stages 1–4. [b]Odds ratio based on a fixed-effect–based meta-analysis of odds ratios. [c]P value based on a weighted z-score meta-analysis. [d]For the present study, the phenotype in WTCCC was restricted to myocardial infarction. In the original discovery report by Samani *et al.*, WTCCC included a broader case definition of myocardial infarction or coronary revascularization.

infarction in the present study have recently been correlated with higher LDL cholesterol[4,23,24].

To evaluate the cumulative effect of these nine SNPs on risk for myocardial infarction, we constructed a myocardial infarction genotype score comprised of the nine SNPs, modeling the number of risk alleles carried by each individual in the MIGen GWAS (stage 1). In logistic regression models including age, sex and principal components of ancestry, individuals in the top quintile of myocardial infarction genotype score had greater than twofold increased risk for myocardial infarction compared with bottom quintile (OR = 2.23, 95% CI = 1.89–2.63; $P = 1 \times 10^{-21}$; **Table 4**).

Although this myocardial infarction genotype score confers risk of a magnitude comparable to other established risk factors such as plasma LDL cholesterol (hazard ratio = 1.62, 95% CI = 1.17–2.25 for top versus bottom quintile of LDL cholesterol as previously reported[25]), further studies are required. The specific SNP set will need to include other recent discoveries for myocardial infarction such as the *MRAS* locus[26] as well as additional SNPs related to LDL cholesterol[24]. Nearly all SNPs related to LDL cholesterol affect risk for MI[4]. In addition, the score requires validation in independent studies, preferably those with a prospective cohort design[27]. Finally, gene–gene and gene–environment interactions need to be modeled if such interactions can be reproducibly documented.

Although the GWAS approach has met with some success in myocardial infarction, the confirmed myocardial infarction risk variants, in sum, explain a small fraction of the variance. The current myocardial infarction genotype score explains 2.8% of the variance in risk for early-onset myocardial infarction. Thus, we tested the hypothesis that systematic assessment of CNPs,

Table 3 Newly identified loci and variants associated with myocardial infarction

Studies (maximum available effective sample size)		Newly identified loci			Newly-identified common variants at previously reported loci	
	SNP	rs9982601	rs12526453	rs6725887[a]	rs1122608	rs11206510
	Chr.	21q22	6p24	2q33	19p13	1p32
	Position NCBI35 (bp)	34,520,998	13,035,530	203,454,130	11,024,601	55,268,627
	Nonrisk allele	C	G	T	T	C
	Risk allele	T	C	C	G	T
	Risk allele frequency	0.13	0.65	0.14	0.75	0.81
	Gene(s) of interest in associated interval	*SLC5A3-MRPS6-KCNE2*	*PHACTR1*	*WDR12*	*LDLR*	*PCSK9*
Stage 1	OR[b]	1.20	1.15	1.24	1.18	1.12
MIGen	(95% CI)	(1.07–1.34)	(1.07–1.24)	(1.12–1.38)	(1.09–1.28)	(1.02–1.23)
(n = 6,046)	P[c]	7.8×10^{-4}	4.6×10^{-4}	8.6×10^{-5}	1.7×10^{-4}	0.02
Stage 2	OR	1.34	1.12	1.15	1.19	1.16
WTCCC, GerMIFS I,	(95% CI)	(1.22–1.47)	(1.05–1.21)	(1.04–1.26)	(1.10–1.28)	(1.07–1.26)
PennCATH, MedSTAR	P	1.7×10^{-9}	3.6×10^{-4}	0.003	2.6×10^{-5}	9.1×10^{-4}
(n = 7,844)						
Stage 3	OR	1.09	1.11	1.11	1.13	1.18
AMI Gene, Verona,	(95% CI)	(0.98–1.21)	(1.04–1.19)	(1.02–1.22)	(1.04–1.22)	(1.10–1.28)
MAHI, IFS, GerMIFS II,	P	0.12	0.001	0.02	0.004	2.2×10^{-5}
INTERHEART (n = 8,642)						
Stage 4	OR	1.11	1.10	1.23	1.03	1.11
deCODE	(95% CI)	(0.95–1.30)	(0.97–1.24)	(1.03–1.46)	(0.90–1.18)	(0.89–1.39)
(n = 3,006)	P	0.17	0.13	0.02	0.69	0.37
Stages 1, 2, 3 + 4	OR	1.20	1.12	1.17	1.15	1.15
(n = 25,538)	(95% CI)	(1.14–1.27)	(1.08–1.17)	(1.11–1.23)	(1.10–1.20)	(1.10–1.21)
	P	6.4×10^{-11}	1.3×10^{-9}	1.3×10^{-8}	1.9×10^{-9}	9.6×10^{-9}

[a]For all studies except INTERHEART, where rs4675310 was substituted as a close to perfect proxy to rs6725887 (Hapmap CEU $r^2 = 1.0$). [b]Odds ratio based on a fixed-effect–based meta-analysis of odds ratios. [c]P value based on a weighted z-score meta-analysis.

common and rare, might identify additional loci contributing to myocardial infarction.

We first used the CANARY algorithm[11] to test 554 commonly segregating CNPs (>1% frequency) for association with early-onset myocardial infarction in 2,783 cases and 2,865 controls that passed sample quality control for CNV analysis (see Methods). The estimated genomic control λ for the entire set of CNPs was ~1.23; for 316 CNPs with allele frequency greater than 5%, λ was ~1.05. We did not observe any CNP with evidence for association surpassing our pre-specified threshold for replication of $P < 0.001$. In fact, the strongest association ($P = 0.002$; **Supplementary Table 3** online) did not pass the Bonferroni correction for 554 tests, let alone genome-wide significance for SNPs. A plot of the observed versus expected P value distribution did not show deviation from the null distribution (**Supplementary Fig. 1** online).

To detect rare CNVs, we used Birdseye[11] and restricted analysis to autosomal deletions and duplications that were both rare (<1% frequency in our samples) and large (greater than 100 kb). After stringent quality control filtering (**Supplementary Methods**), the analysis included 5,955 individuals and 8,065 CNVs (39% deletions). The mean number of rare CNVs per individual was 1.35, and the median was 1.

Using the same methods recently described in a successful study of schizophrenia[28], we evaluated case-control differences in rare CNVs across three parameters: the overall burden of rare CNVs genome-wide, the number of genes overlapped by rare CNVs and the total kilobase extent of rare CNVs. Controlling for sample collection site, there were no case-control differences in genome-wide rare CNV rate ($P = 0.39$), the number of genes intersected by rare CNVs ($P = 0.74$) or the total kilobase extent of rare CNVs ($P = 0.77$). Searching for specific loci with increased rates of rare CNVs in cases versus controls, we found only four regions that showed uncorrected P values <0.01; however, the lowest P value after correction for multiple testing was 0.96.

In summary, we screened common SNPs and CNVs (both common and rare) for association with early-onset myocardial infarction in a large sample. Our study suggests four main conclusions. First, there are at least nine regions that harbor common SNPs associated with myocardial infarction at genome-wide significance; three of these are newly described in this study. Second, the magnitude of risk conferred by a common variant bears no relationship to the potential biological value of the specific finding. For example, similarly to the newly identified loci, we find that common variants at *LDLR* and *PCSK9* confer weak effects, and yet study of these two genes has yielded critical insights into atherosclerosis and myocardial infarction. Third, whereas the effects of individual SNPs are modest, the overall effect (in a comparison of extreme quintiles) is higher for a nine-SNP score (~twofold increase in risk). This observation needs to be validated in additional studies. Finally, and in contrast to the positive results for genetic mapping of myocardial infarction via SNP analysis, we were unable to detect common or rare CNVs associated with risk for myocardial infarction.

The remaining inherited risk for myocardial infarction may be due to some combination of common SNPs for which we do not yet have

Table 4 Quintiles of allelic dosage score comprised of nine validated SNPs and risk for early-onset myocardial infarction

Quintile of myocardial infarction genotype score	Odds ratio	95% confidence interval
Quintile 1	1.0 (reference group)	
Quintile 2	1.22	1.04–1.44
Quintile 3	1.43	1.22–1.68
Quintile 4	1.69	1.44–1.99
Quintile 5	2.23	1.89–2.63

P for association of myocardial infarction genotype score with early-onset myocardial infarction: 2×10^{-28}

The nine validated myocardial infarction polymorphisms are as shown in **Table 2** and **Table 3** and include *SLC5A3-MRPS6-KCNE2* rs9982601, *PHACTR1* rs12526453, *WDR12* rs6725887, 9p21.3 rs4977574, *CXCL12* rs1746048, *CELSR2-PSRC1-SORT1* rs646776, *MIA3* rs17465637, *LDLR* rs1122608, and *PCKS9* rs11206510. Risk of early-onset myocardial infarction was assessed in the 2,967 cases and 3,075 controls from stage 1.

sufficient power, CNVs not measured in our analysis, rare point mutations, nonadditive interactions and epigenetic factors, among other possibilities. Approaches to further clarify the genetic architecture of myocardial infarction include larger-scale screens to identify more common SNPs, improved CNV maps and detection methods to enhance statistical power, and sequencing of myocardial infarction loci (and eventually all exons genome-wide) to discover low-frequency and rare variants. In parallel, mechanistic studies in cells, model organisms and humans that are focused on the nine validated loci should improve our understanding of the root causes of myocardial infarction, and consequently, enable better therapies for this disease.

METHODS

Study design and samples. We conducted a genetic association study with four stages as displayed in **Figure 1**. Stage 1 consisted of MIGen, a collection of 2,967 cases of early-onset myocardial infarction (in men ≤50 y old or women ≤60 y old) and 3,075 age- and sex-matched controls free of myocardial infarction from six international sites: Boston and Seattle in the United States as well as Sweden, Finland, Spain and Italy (**Table 1**). At each site, myocardial infarction was diagnosed on the basis of autopsy evidence of fatal myocardial infarction or a combination of chest pain, electrocardiographic evidence of myocardial infarction, or elevation of one or more cardiac biomarkers (creatine kinase or cardiac troponin). The mean age at the time of myocardial infarction was 41 y among male cases and 47 y among female cases.

We took forward SNPs into three stages of replication (stages 2–4; **Fig. 1**). We chose 1,441 SNPs to test in stage 2 on the basis of two criteria: (i) strength of statistical evidence in stage 1 (1,433 SNPs from loci with $P < 10^{-3}$ in stage 1) or (ii) belonging to one of eight reported loci from recent genome-wide association studies for CAD (a common SNP at or near 9p21.3, *CXCL12*, *SMAD3*, *MTHFD1L*, *MIA3*, *CELSR2-PSRC1-SORT1*, 2q36 and *PCSK9*)[3,4].

Stage 2 consisted of comparisons with four recently completed GWAS for myocardial infarction consisting of a symmetric effective sample size of up to 3,922 myocardial infarction cases and 3,922 controls. These studies included the Wellcome Trust Case Control Consortium Coronary Heart Disease study, German MI Family Study I, PennCATH and MedStar (**Supplementary Methods** and **Supplementary Table 1**). In each stage 2 study, the analysis was restricted to the phenotype of myocardial infarction with an age of onset threshold of <66 y for men or women. Although this age cutoff is slightly less restrictive than that used in stage 1, this cutoff is at or below the mean age of first myocardial infarction in the United States (65 y for men and 70 y for women)[5].

We took forward 33 SNPs to stage 3, which consisted of genotyping an additional six studies with a symmetric effective sample size of up to 4,321 myocardial infarction cases and 4,321 controls. These six studies included Acute MI Gene Study/Dortmund Health Study, Verona Heart Study, Mid-America Heart Institute Study, Irish Family Study, German MI Family Study II and INTERHEART (European-ancestry samples) (**Supplementary Methods** and **Supplementary Table 2**). Stage 3 comprised 25 SNPs with the best combined statistical evidence for myocardial infarction from stages 1 and 2 ($P < 10^{-5}$) and the eight previously reported SNPs discussed above. In each stage 3 study, the analysis was restricted to the phenotype of myocardial infarction, and in four of the six studies, an age-of-onset threshold was established at <66 y for men or women.

Thirteen SNPs were taken forward to stage 4, which consisted of association results from deCODE with a symmetric effective sample size of 1,503 cases of early-onset myocardial infarction and 1,503 controls (**Supplementary Table 2**). Stage 4 comprised five SNPs with the best combined statistical evidence from stages 1–3 and the eight previously reported SNPs. In the deCODE study, the analysis was restricted to cases with early-onset myocardial infarction (men <50 y old or women <60 y old). All participants in the 17 studies across stages 1, 2, 3 and 4 gave written informed consent in accordance with the guidelines of local ethical committees.

Genotyping. In stage 1, we studied 727,496 directly genotyped SNPs (Affymetrix 6.0 GeneChip) that passed quality-control filters, as described in the **Supplementary Methods**. In addition, we used these genotyped SNPs and the phased chromosomes from the HapMap CEU sample to impute genotypes for an additional 1,830,248 SNPs with MACH 1.0 software. In previous work, we have shown that imputation is accurate (average concordance rate of 97.9% between imputed and genotyped data for the same SNP) when using MACH 1.0 in samples of European ancestry with the HapMap CEU phased chromosomes as reference[29].

Stage 2 studies were genotyped on either the Affymetrix GeneChip Human Mapping 500K Array Set or Affymetrix 6.0 GeneChip, and imputation of HapMap SNPs was done using either IMPUTE or MACH 1.0 software (**Supplementary Table 1**).

In Stage 3, genotyping was attempted for 33 SNPs in five studies using the iPLEX MassARRAY platform (Sequenom). In the sixth study, German MI Family Study II, SNPs were genotyped using the Affymetrix 6.0 array (**Supplementary Table 2**).

In Stage 4, the deCODE study samples were genotyped on Illumina Infinium HumanHap300 or HumanHap370 chips, and imputation of HapMap SNPs was done using IMPUTE software (**Supplementary Table 2**).

Association of individual SNP genotypes with myocardial infarction. In stage 1, we tested the association of early-onset myocardial infarction with a combined set of ~2.5 million SNPs (directly genotyped and imputed with information content >0.5) using a logistic regression model that accounted for age, sex and study site. The estimated genomic control λ_{1000} was low at 1.01, suggesting little residual confounding due to population stratification. Regardless, association test statistics were corrected using the genomic control method; separate corrections were made for imputed SNPs (with information content >0.5) and genotyped SNPs. We tested imputed genotypes for association after accounting for uncertainty using the "PROPER" option in the SNPTest software package.

In addition, we evaluated an alternate method to account for potential confounding by population stratification within samples of European ancestry. We conducted principal-component analysis as implemented in PLINK software to define axes of ancestry within the six stage 1 studies[30]. The first two principal components separated individuals into clusters that matched study-site labels and revealed the well-known north–south cline in allele frequencies across Europe (**Supplementary Fig. 2** online). Logistic regression analysis with the first two principal components as covariates (instead of study site) led to nearly identical association results (correlation in association statistics was 0.99). In stages 2 and 3, within each study, we examined the association of SNPs with myocardial infarction using logistic regression after adjustment for age and sex. In stage 4, SNPs were tested for association with early-onset myocardial infarction after adjustment for age and sex, with correction of association test statistics using the genomic-control method as previously described[2].

We used two meta-analytic methods to summarize the statistical evidence for each SNP across stages 1–4. We combined odds ratios for a given reference allele on a logarithmic scale weighted by the inverse of their variances using a fixed-effects model. We also combined evidence for association solely on the

basis of *P* values. For each study, we converted the two-sided *P* value to a *z*-statistic and assigned a sign to reflect the direction of the association given the reference allele. Each *z*-score was then weighted with the squared weights summing to 1 and each sample-specific weight being proportional to the square root of the effective number of individuals in the sample. We summed the weighted *z*-statistics across studies and converted the summary *z*-score to a two-sided *P* value.

Statistical analyses were conducted using either PLINK software or in R.

Analyses of myocardial infarction genotype score, common CNVs and rare CNVs. Details for these analyses are provided in **Supplementary Methods**.

Statistical power. Given our inability to identify CNVs associated with myocardial infarction, we estimated our statistical power for such discovery. For common CNPs, we had 78% power to detect a CNP of 25% frequency and effect size of 1.20 at an alpha of 0.001 in 3,000 cases and 3,000 controls. For rare CNVs, we approximated by simulation the statistical power to detect a CNV with a population frequency for the deletion of 1/8,000 (that is, so it would be observed in 1/4,000 live births). We set the relative risk to 20.0 (the effect size seen for several rare variants associated with schizophrenia[28]) and the population disease prevalence to 1/100. We simulated 10,000 datasets for 2,920 cases and 3,035 controls under this model. Using Fisher's exact test to account for small cell sizes, for a type 1 error rate of 0.01 (one-sided test) we had 97% power. The mean case frequency was ~0.5%, and the mean control frequency was ~0.02%. For a similarly rare variant but with a relative risk of 10.0, the average case frequency was ~0.25% (control frequency still 0.02%) and power was lower at 54%.

These simulations suggest that we had good power to detect loci with large effects, although this assumes perfect sensitivity and specificity for detection. For very large deletions, at least, we expect sensitivity to detect such CNVs would be high. However, we may have missed additional loci with CNVs that are less penetrant, rarer or smaller.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

HARPS. The HARPS study was supported by the grants (R01HL056931, P30ES007033) and a contract (N01HD013107) from US National Institutes of Health.

REGICOR. The REGICOR study was partially funded by the Ministerio de Sanidad y Consumo, Instituto de Salud Carlos III (Red HERACLES RD06/0009), the CIBER Epidemiología y Salud Pública, the FIS and AGAUR Generalitat de Catalunya.

Massachusetts General Hospital. The MIGen study was funded by the US National Institutes of Health (NIH) and National Heart, Lung, and Blood Institute's STAMPEED genomics research program through a grant to D.A. S.K. is supported by a Doris Duke Charitable Foundation Clinical Scientist Development Award, a charitable gift from the Fannie E. Rippel Foundation, the Donovan Family Foundation, a career development award from the NIH and the Department of Medicine and Cardiovascular Research Center at Massachusetts General Hospital. J.B.M. is supported by grant K24 DK080140 from the NIH.

Broad Institute. Genotyping was partially funded by The Broad Institute Center for Genotyping and Analysis, which is supported by grant U54 RR020278 from the National Center for Research Resources.

FINRISK. V.S. was supported by the Sigrid Juselius Foundation. L.P. was supported by the Center of Excellence in Complex Disease Genetics of the Academy of Finland, the Nordic Center of Excellence in Disease Genetics and the Finnish Foundation for Cardiovascular Research.

WTCCC Study. The study was funded by the Wellcome Trust. Recruitment of cases for the WTCCC Study was carried out by the British Heart Foundation (BHF) Family Heart Study Research Group and supported by the BHF and the UK Medical Research Council. N.J.S. and S.G.B. hold chairs funded by the BHF.

PennCATH/MedStar. Recruitment of the PennCATH cohort was supported by the Cardiovascular Institute of the University of Pennsylvania. Recruitment of the MedStar cohort was supported by a research grant from GlaxoSmithKline and from the MedStar Research Institute. Genotyping was done at the Center for Applied Genomics at the Children's Hospital of Philadelphia and supported by GlaxoSmithKline through an Alternate Drug Discovery Initiative research alliance award (to M.P.R. and D.J.R.) with the University of Pennsylvania School of Medicine. D.J.R. was supported by a Doris Duke Charitable Foundation Distinguished Clinical Scientist Award.

Verona Heart Study. The study was supported by a grant from the Italian Ministry of University and Research and grants from the Veneto Region and the Cariverona Foundation, Verona, Italy.

Mid-America Heart Institute. T.M. is supported by a career development grant from the NIH.

Irish Family Study. We thank the clinical staff members for their valuable contribution to the collection of families for this study. The research was supported by the Northern Ireland Research and Development Office, a Royal Victoria Hospital Research Fellowship, the Northern Ireland Chest, Heart and Stroke Association, and the Heart Trust Fund (Royal Victoria Hospital).

GerMIFS I and II. The German Study was supported by the Deutsche Forschungsgemeinschaft and the German Federal Ministry of Education and Research in the context of the German National Genome Research Network.

Cardiogenics. Cardiogenics is an EU-funded integrated project (LSHM-CT-2006-037593).

INTERHEART. S.A. holds the Michael G. DeGroote and Heart and Stroke Foundation of Ontario Chair in Population Health and the May Cohen Eli Lilly Endowed Chair in Women's Health Research, McMaster University. We acknowledge the contribution of S. Yusuf who initiated and, together with the Steering Committee, supervised the conduct of the INTERHEART study. We thank members of the Project Office, S. Rangarajan (study coordinator) and K. Hall (laboratory manager), for their assistance in coordinating the genetics component of the INTERHEART project. R.D. is a recipient of a Canada Graduate Scholarship Doctoral Award from the Canadian Institutes for Health Research.

Disclosures. The collection of clinical and sociodemographic data in the Dortmund Health Study was supported by the German Migraine & Headache Society (DMKG) and by unrestricted grants of equal share from Astra Zeneca, Berlin Chemie, Boots Healthcare, Glaxo-Smith-Kline, McNeil Pharma (former Woelm Pharma), MSD Sharp & Dohme and Pfizer to the University of Muenster.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. McPherson, R. *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488–1491 (2007).
2. Helgadottir, A. *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491–1493 (2007).
3. Samani, N.J. *et al.* Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* **357**, 443–453 (2007).
4. Willer, C.J. *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.* **40**, 161–169 (2008).
5. Rosamond, W. *et al.* Heart disease and stroke statistics—2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* **117**, e25–e146 (2008).
6. Lloyd-Jones, D.M. *et al.* Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *J. Am. Med. Assoc.* **291**, 2204–2211 (2004).
7. Centers for Disease Control and Prevention. Prevalence of heart disease—United States, 2005. *MMWR Morb. Mortal. Wkly. Rep.* **56**, 113–118 (2007).
8. Nora, J.J., Lortscher, R.H., Spangler, R.D., Nora, A.H. & Kimberling, W.J. Genetic-epidemiologic study of early-onset ischemic heart disease. *Circulation* **61**, 503–508 (1980).
9. Sebat, J. *et al.* Large-scale copy number polymorphism in the human genome. *Science* **305**, 525–528 (2004).
10. McCarroll, S.A. *et al.* Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat. Genet.* **40**, 1166–1174 (2008).
11. Korn, J.M. *et al.* Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat. Genet.* **40**, 1253–1260 (2008).
12. Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M.J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet. Epidemiol.* **32**, 381–385 (2008).
13. Cavdar Koc, E., Burkhart, W., Blackburn, K., Moseley, A. & Spremulli, L.L. The small subunit of the mammalian mitochondrial ribosome. Identification of the full complement of ribosomal proteins present. *J. Biol. Chem.* **276**, 19363–19374 (2001).
14. Kwon, H.M. *et al.* Cloning of the cDNa for a Na+/myo-inositol cotransporter, a hypertonicity stress protein. *J. Biol. Chem.* **267**, 6297–6301 (1992).
15. Abbott, G.W. *et al.* MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell* **97**, 175–187 (1999).
16. Allen, P.B., Greenfield, A.T., Svenningsson, P., Haspeslagh, D.C. & Greengard, P. Phactrs 1–4: a family of protein phosphatase 1 and actin regulatory proteins. *Proc. Natl. Acad. Sci. USA* **101**, 7187–7192 (2004).

17. Holzel, M. *et al.* Mammalian WDR12 is a novel member of the Pes1-Bop1 complex and is required for ribosome biogenesis and cell proliferation. *J. Cell Biol.* **170**, 367–378 (2005).
18. Hobbs, H.H., Brown, M.S. & Goldstein, J.L. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum. Mutat.* **1**, 445–466 (1992).
19. Linsel-Nitschke, P. *et al.* Lifelong reduction of LDL-cholesterol related to a common variant in the LDL-receptor gene decreases the risk of coronary artery disease—a Mendelian Randomisation study. *PLoS ONE* **3**, e2986 (2008).
20. Abifadel, M. *et al.* Mutations in *PCSK9* cause autosomal dominant hypercholesterolemia. *Nat. Genet.* **34**, 154–156 (2003).
21. Cohen, J.C., Boerwinkle, E., Mosley, T.H. Jr & Hobbs, H.H. Sequence variations in *PCSK9*, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* **354**, 1264–1272 (2006).
22. Kathiresan, S. *et al.* A PCSK9 missense variant associated with a reduced risk of early-onset myocardial infarction. *N. Engl. J. Med.* **358**, 2299–2300 (2008).
23. Kathiresan, S. *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat. Genet.* **40**, 189–197 (2008).
24. Kathiresan, S. *et al.* Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat. Genet.* **41**, 56–65 (2009).
25. Ridker, P.M., Rifai, N., Cook, N.R., Bradwin, G. & Buring, J.E. Non-HDL cholesterol, apolipoproteins A-I and B100, standard lipid measures, lipid ratios, and CRP as risk factors for cardiovascular disease in women. *J. Am. Med. Assoc.* **294**, 326–333 (2005).
26. Erdmann, J. *et al.* Novel susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* advance online publication, doi:10.1038/ng.307 (8 February 2009).
27. Kathiresan, S. *et al.* Polymorphisms associated with cholesterol and risk of cardiovascular events. *N. Engl. J. Med.* **358**, 1240–1249 (2008).
28. International Schizophrenia. Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237–241 (2008).
29. Zeggini, E. *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638–645 (2008).
30. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

The complete list of authors is as follows:
Stage 1 discovery GWA studies:
Manuscript preparation. Sekar Kathiresan[1–4] (leader), Benjamin F Voight[2,3,5], Shaun Purcell[2,3,6], Kiran Musunuru[1–4], Diego Ardissino[7], Pier M Mannucci[8], Sonia Anand[9], James C Engert[10], Nilesh J Samani[11], Heribert Schunkert[12], Jeanette Erdmann[12], Muredach P Reilly[13,14], Daniel J Rader[13,14], Thomas Morgan[15], John A Spertus[16], Monika Stoll[17], Domenico Girelli[18], Pascal P McKeown[19], Chris C Patterson[19], David S Siscovick[20], Christopher J O'Donnell[1,4,21], Roberto Elosua[22], Leena Peltonen[3,23,24], Veikko Salomaa[25], Stephen M Schwartz[20,26], Olle Melander[27], David Altshuler[2–5,28]
Italian Atherosclerosis, Thrombosis and Vascular Biology Study. Diego Ardissino[7], Pier Angelica Merlini[29], Carlo Berzuini[30], Luisa Bernardinelli[30,31], Flora Peyvandi[8], Marco Tubaro[32], Patrizia Celli[33], Maurizio Ferrario[34], Raffaela Fetiveau[34], Nicola Marziliano[34], Giorgio Casari[35], Michele Galli[36], Flavio Ribichini[37], Marco Rossi[38], Francesco Bernardi[39], Pietro Zonzin[40], Alberto Piazza[41], Pier M Mannucci[8]
Heart Attack Risk in Puget Sound. Stephen M Schwartz[20,26], David S Siscovick[20], Jean Yee[20,26], Yechiel Friedlander[42]
Registre Gironi del COR. Roberto Elosua[22], Jaume Marrugat[22,] Gavin Lucas[22], Isaac Subirana[22], Joan Sala[43], Rafael Ramos[44]
Massachusetts General Hospital Premature Coronary Artery Disease Study. Sekar Kathiresan[1–4]
James B Meigs[4,45], Gordon Williams[4,46], David M Nathan[4,47], Calum A MacRae[1,4], Christopher J O'Donnell[1,4,21]
FINRISK. Veikko Salomaa[25], Aki S Havulinna[25], Leena Peltonen[3,23,24]
Malmo Diet and Cancer Study. Olle Melander[27], Goran Berglund[48]
Stage 1 data analysis. Benjamin F Voight[2,3,5] (leader), Sekar Kathiresan[1–4], Joel N Hirschhorn[3,28], Rosanna Asselta[49], Stefano Duga[49], Marta Spreafico[8], Kiran Musunuru[1–4], Mark J Daly[2–4], Shaun Purcell[2,3,6]
Copy number variant analysis. Benjamin F Voight[2,3,5], Shaun Purcell[2,3,6], James Nemesh[3], Joshua M Korn[2,3,5], Steven A McCarroll[2,3,5]
Stage 1 phenotype data assembly. Stephen M Schwartz[20,26] (leader), Jean Yee[20,26], Sekar Kathiresan[1–4], Gavin Lucas[22], Isaac Subirana[22], Roberto Elosua[22]
Stage 1 genome-wide genotyping. Aarti Surti[3], Candace Guiducci[3], Lauren Gianniny[3], Daniel Mirel[3], Melissa Parkin[3], Noel Burtt[3], Stacey B Gabriel[3] (leader)
Replication studies:
Wellcome Trust Case Control Consortium. Nilesh J Samani[11], John R Thompson[50], Peter S Braund[11], Benjamin J Wright[50], Anthony J Balmforth[51], Stephen G Ball[51], Alistair S Hall[51], Wellcome Trust Case Control Consortium[68]
German MI Family Study I. Heribert Schunkert[12], Jeanette Erdmann[12], Patrick Linsel-Nitschke[12], Wolfgang Lieb[12], Andreas Ziegler[52], Inke R König[52], Christian Hengstenberg[53], Marcus Fischer[53], Klaus Stark[53], Anika Grosshennig[12,52], Michael Preuss[12,52], H-Erich Wichmann[54,55], Stefan Schreiber[56]
Cardiogenics. Heribert Schunkert[12], Nilesh J Samani[11], Jeanette Erdmann[12], Willem Ouwehand[57], Christian Hengstenberg[53], Panos Deloukas[23], Michael Scholz[58], Francois Cambien[59], Cardiogenics[68]
PennCATH/MedSTAR. Muredach P Reilly[13,14], Mingyao Li[60], Zhen Chen[60], Robert Wilensky[13,14], William Matthai[14], Atif Qasim[14], Hakon H Hakonarson[61], Joe Devaney[62], Mary-Susan Burnett[62], Augusto D Pichard[62], Kenneth M Kent[62], Lowell Satler[62], Joseph M Lindsay[62], Ron Waksman[62], Christopher W Knouff[69], Dawn M Waterworth[69], Max C Walker[69], Vincent Mooser[69], Stephen E Epstein[62], Daniel J Rader[13,14]
Acute Myocardial Infarction Gene Study/Dortmund Health Study. Thomas Scheffold[63], Klaus Berger[64], Monika Stoll[17], Andreas Huge[17]
Verona Heart Study. Domenico Girelli[18], Nicola Martinelli[18], Oliviero Olivieri[18], Roberto Corrocher[18]
Mid-America Heart Institute. Thomas Morgan[15], John A Spertus[16]
Irish Family Study. Pascal P McKeown[19], Chris C Patterson[19]
German MI Family Study II. Heribert Schunkert[12], Jeanette Erdmann[12], Patrick Linsel-Nitschke[12], Wolfgang Lieb[12], Andreas Ziegler[52], Inke R König[52], Christian Hengstenberg[53], Marcus Fischer[53], Klaus Stark[53], Anika Grosshennig[12,52], Michael Preuss[12,52], H-Erich Wichmann[54,55], Stefan Schreiber[56]
deCODE Study. Hilma Hólm[65], Gudmar Thorleifsson[65], Unnur Thorsteinsdottir[65,66], Kari Stefansson[65,66]
INTERHEART. James C Engert[10], Ron Do[67], Changchun Xie[9], Sonia Anand[9]
MIGen steering committee. Sekar Kathiresan[1–4], Diego Ardissino[7], Pier M Mannucci[8], David Siscovick[20], Christopher J O'Donnell[1,4,21], Nilesh J Samani[11], Olle Melander[27], Roberto Elosua[22], Leena Peltonen[3,23,24], Veikko Salomaa[25], Stephen M Schwartz[20,26], David Altshuler[2–5,28] (leader)

[1]Cardiovascular Research Center and Cardiology Division and [2]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [3]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [4]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA. [5]Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, 02114, USA. [6]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. [7]Division of Cardiology, Azienda Ospedaliero-Universitaria di Parma, 43100 Parma, Italy. [8]Department of Internal Medicine and Medical Specialities, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico, Ospedale Maggiore, Mangiagalli e Regina Elena, University of Milan, 20122 Milan, Italy. [9]Population Health Research Institute, Hamilton Health Sciences and Departments of Medicine and Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, L8L 2X2 Ontario, Canada. [10]Departments of Medicine and Human Genetics, McGill University, Montréal, H3A 1A1 Québec, Canada. [11]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, LE3 9QP, UK. [12]Medizinische Klinik II, Universität zu Lübeck, 23538 Lübeck, Germany. [13]The Institute for Translational Medicine and Therapeutics, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [14]The Cardiovascular Institute, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [15]Department of Pediatrics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA. [16]Mid-America Heart Institute and University of Missouri-Kansas City, Kansas City, Missouri 64111, USA. [17]Leibniz-Institute for Arteriosclerosis Research, University Münster, 48149 Münster, Germany. [18]Department of Clinical and Experimental Medicine, University of Verona, 37134 Verona, Italy. [19]Centre for Public Health, Queen's University Belfast, Institute of Clinical Science, Belfast, BT12 6BJ, Northern Ireland, UK. [20]Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle, Washington 98101, USA. [21]Framingham Heart Study and National, Heart, Lung, and Blood Institute, Framingham, Massachusetts 01702, USA. [22]Cardiovascular Epidemiology and

Genetics, Institut Municipal D'investigacio Medica, and CIBER Epidemiología y Salud Pública, 08003 Barcelona, Spain. [23]Wellcome Trust Sanger Institute, Cambridge CB10 1SA, UK. [24]Institute for Molecular Medicine, University of Helsinki, Helsinki 00029, Finland. [25]Chronic Disease Epidemiology Unit, Department of Health Promotion and Chronic Disease Prevention, National Public Health Institute, Helsinki 00300, Finland. [26]Department of Epidemiology, University of Washington, Seattle, Washington 98195, USA. [27]Department of Clinical Sciences, Hypertension and Cardiovascular Diseases, University Hospital Malmö, Lund University, Malmö 20502, Sweden. [28]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [29]Division of Cardiology, Azienda Ospedaliera Niguarda Ca' Granda, 20162 Milan, Italy. [30]Biostatistics Unit, Medical Research Council, Cambridge, UK and Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK. [31]Department of Applied Health Sciences, University of Pavia, 27100 Pavia, Italy. [32]Division of Cardiology, Ospedale San Filippo Neri, 00135 Rome, Italy. [33]Division of Cardiology, Ospedale San Camillo, 00151 Rome, Italy. [34]Fondazione Istituto di Ricovero e Cura a Carattere Scientifico, Policlinico San Matteo, 27100 Pavia, Italy. [35]Vita-Salute San Raffaele University and San Raffaele Scientific Institute, 20132 Milan, Italy. [36]Division of Cardiology, Ospedale di Livorno, 57100 Livorno, Italy. [37]Division of Cardiology, Ospedale Borgo Trento, University of Verona, 37126 Verona, Italy. [38]Division of Cardiology, Istituto di Ricovero e Cura a Carattere Scientifico, Istituto Clinico Humanitas, 20089 Milan, Italy. [39]Department of Biochemistry and Molecular Biology, University of Ferrara, 44100 Ferrara, Italy. [40]Division of Cardiology, Ospedale di Rovigo, 45100 Rovigo, Italy. [41]Department of Genetics, Biology and Biochemistry, University of Turin, 10126 Turin, Italy. [42]Unit of Epidemiology, Hebrew University-Hadassah School of Public Health, Jerusalem 91120, Israel. [43]Servei de Cardiologia i Unitat Coronària, Hospital de Girona Josep Trueta and Institut de Investigació Biomedica de Girona, 17007 Girona, Spain. [44]Unitat de Recerca i Unitat Docent de Medicina de Familia de Girona, IDIAP Jordi Gol, Institut Català de la Salut, 08007 Barcelona, Spain. [45]General Medicine Division, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, 02114, USA. [46]Cardiovascular Endocrinology Section, Division of Endocrinology, Diabetes, and Hypertension, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [47]Diabetes Center, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [48]Department of Clinical Sciences, Internal Medicine, University Hospital Malmö, Lund University, Malmö 20502, Sweden. [49]Department of Biology and Genetics for Medical Sciences, University of Milan, 20133 Milan, Italy. [50]Department of Health Sciences, University of Leicester, Leicester, LE1 7RH, UK. [51]LIGHT Research Institute, Faculty of Medicine and Health, University of Leeds, Leeds, LS1 3EX, UK. [52]Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, 23538 Lübeck, Germany. [53]Klinik und Poliklinik für Innere Medizin II, Universität Regensburg, 93042 Regensburg, Germany. [54]Institute of Epidemiology, Helmholtz Zentrum München – German Research Center for Environmental Health, 85764 Neuherberg, Germany. [55]Institute of Medical Information Science, Biometry and Epidemiology, Ludwig-Maximilians-Universität Munich, 81377 Munich, Germany. [56]Institut für Klinische Molekularbiologie, Christian-Albrechts Universität, 24105 Kiel, Germany. [57]Department of Haematology, University of Cambridge, Long Road, Cambridge, CB2 2PT, UK. [58]Trium Analysis Online GmbH, 81677 Munich, Germany. [59]INSERM UMR_S 525, Universite Pierre et Marie Curie –Paris 6, Paris 75634, France. [60]Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [61]The Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. [62]Cardiovascular Research Institute, MedStar Research Institute, Washington Hospital Center, Washington, DC 20010, USA. [63]Institute for Heart and Circulation Research of the University of Witten/Herdecke, 44227 Dortmund, Germany. [64]Institute of Epidemiology and Social Medicine, University of Muenster, 48129 Muenster, Germany. [65]deCODE Genetics, 101 Reykjavik, Iceland. [66]Faculty of Medicine, University of Iceland, 101 Reykjavík, Iceland. [67]Department of Human Genetics, McGill University, Montréal, Québec H3A 1B1, Canada. [68]A full list of members is provided in the **Supplementary Note** online. [69]Genetics Division and Drug Discovery, GlaxoSmithKline, King of Prussia, Pennsylvania 19406, USA.

Corrigendum: Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study

Mark S Silverberg, Judy H Cho, John D Rioux, Dermot P B McGovern, Jing Wu, Vito Annese, Jean-Paul Achkar, Philippe Goyette, Regan Scott, Wei Xu, M Michael Barmada, Lambertus Klei, Mark J Daly, Clara Abraham, Theodore M Bayless, Fabrizio Bossa, Anne M Griffiths, Andrew F Ippoliti, Raymond G Lahaie, Anna Latiano, Pierre Paré, Deborah D Proctor, Miguel D Regueiro, A Hillary Steinhart, Stephan R Targan, L Philip Schumm, Emily O Kistner, Annette T Lee, Peter K Gregersen, Jerome I Rotter, Steven R Brant, Kent D Taylor, Kathryn Roeder & Richard H Duerr
Nat. Genet. **41, 216–220 (2009), published online 4 January 2009; corrected after print 28 April 2009**

In the first paragraph of the second column on the third page, rs11209026 A allele was incorrectly listed as rs111209026 A allele. The error has been corrected in the HTML and PDF versions of the article.

Corrigendum: Loss-of-function mutations of an inhibitory upstream ORF in the human hairless transcript cause Marie Unna hereditary hypotrichosis

Yaran Wen, Yang Liu, Yiming Xu, Yiwei Zhao, Rui Hua, Kaibo Wang, Miao Sun, Yuanhong Li, Sen Yang, Xue-Jun Zhang, Roland Kruse, Sven Cichon, Regina C Betz, Markus M Nöthen, Maurice A M van Steensel, Michel van Geel, Peter M Steijlen, Daniel Hohl, Marcel Huber, Giles S Dunnill, Cameron Kennedy, Andrew Messenger, Colin S Munro, Alessandro Terrinoni, Alain Hovnanian, Christine Bodemer, Yves de Prost, Amy S Paller, Alan D Irvine, Rod Sinclair, Jack Green, Dandan Shang, Qing Liu, Yang Luo, Li Jiang, Hong-Duo Chen, Wilson H-Y Lo, W H Irwin McLean, Chun-Di He & Xue Zhang
Nat. Genet. **41, 228–233 (2009), published online 4 January 2009; corrected after print 28 April 2009**

The affiliation of the 24th author, Alessandro Terrinoni, was listed incorrectly. It should read IDI-IRCCS Biochemistry Laboratory c/o Univ. Tor Vergata, 00133 Rome, Italy. The error has been corrected in the HTML and PDF versions of this article.

Addendum: Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing

Qun Pan, Ofer Shai, Leo J Lee, Brendan J Frey & Benjamin J Blencowe
Nat. Genet. **40, 1413–1415 (2008), published online 2 November 2008; addendum published after print 28 April 2009**

The GEO accession number for the mRNA-Seq datasets is GSE13652.

Corrigendum: Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants

Myocardial Infarction Genetics Consortium
Nat. Genet. **41, 334–341 (2009); published online 8 February 2009; corrected after print 27 May 2009**

In the version of this article initially published, the names of four co-authors (Christopher W Knouff, Dawn M Waterworth, Max C Walker, Vincent Mooser) were omitted from the author list. The error has been corrected in the HTML and PDF versions of the article.