

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

Myocardial Infarction Genetics Consortium

Address for Correspondence:

Sekar Kathiresan, M.D.
Center for Human Genetic Research and Cardiovascular Research Center
Massachusetts General Hospital
Harvard Medical School
skathiresan@partners.org

SUPPLEMENTARY METHODS

Stage 1 studies. Stage 1 consisted of cases with early-onset myocardial infarction (MI) and matched controls from six studies. All participants were of European ancestry. The six studies were: i) Italian Atherosclerosis Thrombosis and Vascular Biology (IATVB); ii) Heart Attack Risk in Puget Sound (HARPS); iii) Registre Gironi del COR (Gerona Heart Registry or REGICOR); iv) Massachusetts General Hospital Premature Coronary Artery Disease (MGH PCAD); v) FINRISK; and vi) Malmo Diet and Cancer Study (MDC).

IATVB¹ is a prospective, nationwide, case-control study involving 125 coronary care units in Italy. The cases were patients who were hospitalised for a first MI before the age of 45 years and underwent coronary angiography. Acute MI was defined as resting chest pain lasting more than 30 minutes, accompanied by persistent electrocardiographic changes, and confirmed by an increase in total creatine kinase or in the MB fraction to more than twice the upper normal limits.

The controls were healthy subjects without a history of thromboembolic disease who were unrelated to the patients, but individually matched with them by age, gender and geographical origin. They were enrolled from among the blood donors or staff of the same participating hospitals. Recruitment of cases and controls took place between 1994 and 2007.

HARPS² is a population-based case-control study of cases of early-onset MI and controls matched on age and sex. Eligible case patients were men aged <50 and women <60 diagnosed with a first MI between 1991 to 2002. Cases were ascertained through medical record review at all acute care facilities in King, Pierce, and Snohomish counties, Washington, US. Controls were identified using random digit dialing from these counties (1991-2002) and had no history of cardiovascular disease. Data collected through in-person interviews with each case and control

include information on medical and lifestyle risk factors, and a self-administered food frequency questionnaire.

MGH PCAD³ study is a hospital-based case-control study of cases of early-onset MI and controls matched on age and sex. Eligible case patients were men aged ≤ 50 and women aged ≤ 60 who were hospitalized at MGH with MI between 1999 and 2004. We attempted to recruit all patients admitted to MGH meeting the case definition. Controls were recruited from Boston and its vicinity through a general newspaper advertisement and eligible if they reported no history of cardiovascular disease or cardiac medications. Data collected through in-person interviews and examination for each case and control include risk factors, anthropometry, and blood pressure.

REGICOR⁴ is a population-based case-control study of cases of early-onset MI and controls matched on age and sex. Specifically, eligible case patients were consecutive patients hospitalized with a first MI to the only coronary care unit in the catchment area of Gerona, Spain. Controls were subjects randomly selected from a cross-sectional study of cardiovascular risk factors in the catchment area and were deemed free of MI by history, physical examination, and ECG. Data collected through in-person interviews and examination include risk factors (based on questionnaire and blood samples), anthropometry, and blood pressure.

FINRISK⁵ contributed cases of early-onset MI and controls. Cases were defined as men aged ≤ 50 and women aged ≤ 60 and hospitalized with MI or died of MI, and were obtained from a national population-based survey of risk factors for cardiovascular disease. FINRISK surveys are conducted every 5 years, and from the 23,188 individuals enrolled in FINRISK 1992, 1997, and 2002, we assembled cases and controls. Hospitalization data and mortality data for MI were obtained from the Finnish National Hospital Discharge Register and the Finnish National

Causes-of-Death Register as of December 31, 2004 based on ICD-9 code of 410 and ICD-10 codes of I20 and I21. These registers have excellent validity. Cases included MI prevalent at baseline examination and incident events on follow-up. Controls were randomly selected from participants of the three FINRISK cohorts, who survived until the event date of the case without hospitalization for CHD or for stroke, matched to cases based on age, sex, study area, and FINRISK cohort year. Data collected in both cases and controls at a baseline study examination include risk factors (based on questionnaire and blood samples), anthropometry, and blood pressure measurement.

MDC⁶ contributed cases of early-onset MI and controls matched for age and sex. Cases were defined as men aged ≤ 50 and women aged ≤ 60 and hospitalized with MI, and were obtained from a population-based cohort study of 28,098 men and women living in Malmö, Sweden, between 1991 and 1996. Data on MI were obtained by record linkage to the Swedish National Hospital Discharge Register as of December 31, 2000 based on ICD-9 code 410 and ICD-10 code I21. Cases included MI prevalent at baseline examination and incident events on follow-up. Controls were randomly selected among the 28,098 participants and were free of MI (criteria above) at baseline examination and during follow-up. Data collected in cases and controls at a baseline examination include risk factors (based on questionnaire and blood samples), anthropometry, dietary assessment, and a physical examination.

Stage 2 studies. Stage 2 consisted of cases with early-onset MI and matched controls from four studies. All participants were of European ancestry. The four studies were: i) Wellcome Trust Case Control Consortium MI Study (WTCCC MI), ii) German MI Family Study I (GMIFS I), iii) PennCATH, and iv) MedSTAR.

WTCCC MI⁷ has been recently described. For the present manuscript, we restricted the analysis to cases with MI. Cases recruited based exclusively on coronary revascularization were excluded.

GerMIFS I⁷ has been recently described.

The **Penn-CATH**⁸ cohort is a University of Pennsylvania Medical Center based angiographic study of over 3,800 subjects that has been used previously for replication of novel genes and risk factors for atherosclerotic cardiovascular disease and type 2 diabetes⁹⁻¹². Between July 1998 and March 2003, PennCATH recruited a consecutive cohort of patients undergoing cardiac catheterization at Penn. A total of 3,850 subjects were recruited and all gave written informed consent in a Penn Institutional Review Board approved protocol. Enrollment criteria included any clinical indication for cardiac catheterization and ability to give informed consent. The following data were extracted from the medical record; age, gender, race/ethnicity, past medical (including diabetes, hypertension, dyslipidemia, prior MI and cardiac events), social, family and medication history, cardiovascular risk factors, physical exam including vital signs, weight and height (for BMI). Ethnicity information was self-reported. Data from cardiac catheterization including coronary angiography were recorded. Blood was drawn in a fasting state, DNA (buffy coat) and plasma was isolated, and lipoproteins and glucose were assayed on all samples.

A nested case-control GWAS was recently performed in PennCATH (N=1,401 Caucasians) composed of controls (N=468) who on coronary angiography showed no evidence of coronary artery disease (CAD) and CAD cases (N=933) with one or more coronary vessels with $\geq 50\%$ stenosis equally selected for stable CAD cases without history of MI and CAD cases with a history of MI. Controls were aged over 40 in men and 45 in women. For the purpose of

Stage 2 replication of MIGen findings, PennCATH CAD cases were restricted to those with a history of MI and aged less than 66 (N=415).

The **MedStar** study is a Washington Hospital Center based angiographic study of 1,500 subjects specifically designed for biomarker and genetic association studies of acute and chronic coronary atherosclerosis. MedStar is a cross sectional study of coronary atherosclerosis in a consecutive cohort of selected patients undergoing cardiac catheterization at Washington Hospital between August 2004 and March 2007. All subjects have been enrolled in a Washington Hospital Institutional Review Board approved protocol and all subjects gave written informed consent. Enrollment criteria include any clinical indication for cardiac catheterization and ability to give informed consent. The following data were extracted from the medical record; age, gender, race/ethnicity, past medical, social, family and medication history, cardiovascular risk factors (diabetes, smoking, and hypertension), physical exam including vital signs, weight and height (for BMI), and cardiovascular findings. Ethnicity information was self-reported. Data from cardiac catheterization including coronary angiography were recorded. Coronary angiograms were scored on the day by the interventional cardiologist who performed the procedure and reviewed by a second cardiologist at a later date. Blood was drawn in a 12-hour fasting state (except in those with acute MI), at the time of the initial catheter insertion prior to the administration of any contrast dye for plasma, serum and buffy coat DNA isolation. All demographic, anthropometric, clinical, cardiac catheterization, laboratory, and genetic data are integrated into study database that exclude personal identifiers.

A case-control GWAS similar to PennCATH was performed in MedStar (N=1,322 Caucasians) composed of controls (N=447) who on coronary angiography showed no evidence of CAD and CAD cases (N=874) with one or more coronary vessels with $\geq 50\%$ stenosis divided into stable

CAD cases without history of MI and CAD cases with a history of MI. Controls were aged over 45 in men and women. For the purpose of Stage 2 replication of MIGen findings, MedStar CAD cases were restricted to those with a history of MI aged less than 66 (N=420).

Stage 3 studies. Stage 3 consisted of cases with MI and matched controls from six studies. All participants were of European ancestry except for a subset of individuals in INTERHEART who were of South Asian ancestry. The six studies were: i) Acute Myocardial Infarction (AMI) Gene Study/Dortmund Health Study; ii) Verona Heart Study; iii) Mid-America Heart Institute; iv) Irish Family Study; v) German MI Family Study II; and vi) INTERHEART.

The **AMI Gene Study/Dortmund Health Study**¹³ is a prospective multi-centre registry involving 4 heart centres and 6 cardiologic departments in North-Rhine-Westphalia, Lower Saxony, and Mecklenburg-Western Pomerania. Between 2004 and 2006, we enrolled a cohort of 809 consecutive men younger than 65 and suffering from non-ST-segment elevation MI or ST-segment elevation MI with an onset of symptoms less than 24 hours before admission. All cases underwent cardiac catheterization and interventional or surgical revascularization. After informed consent personal data collection and blood sampling were carried out at admission. Information on history, traditional risk factors and co-morbidities were documented after interventional procedures and clinical stabilization of the patients.

Controls came from the Dortmund Health Study^{13,14}, a population-based survey conducted in the city of Dortmund with the aim to determine the prevalence of headache types, cardiovascular and other chronic diseases and their risk factors in the general population. Sampling for the study was done randomly from the city's population register stratified by five-year age group and gender. History of MI and other cardiovascular conditions was assessed in face-to-face interviews. The study was conducted in 2003-2004. The recruitment protocols and

study procedures were approved by the ethics committees of the University of Witten-Herdecke and the University of Muenster, Germany, respectively.

The **Verona Heart Study**¹⁵ is an ongoing study aimed at identifying new risk factors for CAD and MI in a population of subjects with angiographic documentation of their coronary vessel. The CAD group had angiographically documented severe coronary atherosclerosis, the majority of them being candidates for coronary artery bypass grafting or percutaneous coronary intervention. Control subjects were selected such that they had normal coronary arteries, being submitted to coronary angiography for reasons other than CAD. Controls with history or clinical evidence of atherosclerosis in vascular territories beyond the coronary bed were excluded. Information on MI diagnoses was gathered through medical records showing diagnostic electrocardiogram and enzyme changes, and/or the typical sequelae of MI on ventricular angiography and on echocardiography. The local Ethical Committee approved the study. Informed consent was obtained from all the patients after a full explanation of the study.

The **Mid-America Heart Institute**^{16,17} patients (N = 811) were recruited in successive prospective cohort studies designed to investigate clinical outcomes among survivors of acute coronary syndrome. Age, sex, and BMI-matched control subjects without known CAD (N = 650) consisted of ambulatory outpatients from the same geographic area presenting for routine laboratory testing. All 1,461 subjects submitted for genotyping were of self-reported white/mixed European ancestry. There were no restrictions on age. Subjects have recently been described in further detail^{16,17}.

Recruitment for the **Irish Family Study**¹⁸ took place between August 1999 and October 2004. All subjects were Caucasian whose four grandparents were born in Ireland. Each family had at least one member affected with early-onset MI (disease onset ≤ 55 years for males and ≤ 60

years for females) and at least one unaffected sibling and / or both parents surviving. Unaffected siblings were required to: (i) be older than the affected sibling was at the onset of coronary heart disease; (ii) have no symptoms of angina or possible MI by World Health Organization questionnaire assessments; (iii) have no history of coronary heart disease diagnosed by a doctor; and (iv) have a resting 12 lead ECG record showing no evidence of ischemia or previous MI.

The recruitment for **GerMIFS II** was similar to that for GMIFS I¹⁹. All 1,222 patients had a validated MI with a strong genetic component as documented by an early age of onset (prior to the age of 60 years). Moreover, a positive family history for CAD was documented in 726 (59.4 %) of patients. Patients were identified following their admission for acute treatment of MI or in cardiac rehabilitation clinics. Population-based controls were derived from the MONICA/KORA Augsburg survey S4 1 (n=820) and the PopGen blood donor sample 2 (PopGen-BSP) (n=478).

INTERHEART²⁰⁻²² is a global case/control study of risk factors for acute MI involving 27,098 individuals recruited from 262 centers in 52 countries. Informed written consent to obtain the baseline information and to collect and store the genetic and other biologic specimens was obtained from 21,508 individuals (including all individuals analyzed in this study). To identify incident cases of acute MI, all patients, irrespective of age, admitted to the coronary care unit (or an equivalent cardiology ward) within 24 hours of symptom onset were screened. Cases were eligible if they had characteristic symptoms plus electrocardiogram changes indicative of a new MI (new pathologic Q waves, at least 1 mm ST elevation in any two or more contiguous limb leads or a new left bundle branch block, or new persistent ST-T wave changes diagnostic of a non-Q wave MI) or a plasma level of cardiac troponin level above that considered normal in the hospital/institution where the patient was registered. For each case, at least one control of the

same age (± 5 years) and sex was recruited from the same centre. Controls were defined as individuals who had no previous diagnosis of heart disease or history of exertional chest pain. Eligible controls were classified as i) hospital-based, defined as patients attending the hospital or outpatient clinics for the following reasons: refraction and cataracts, physical check-up, routine pap smear, routine breast exam, elective minor surgery for conditions that were not obviously related to CHD or its risk factors, elective orthopedic surgery (eligibility dependent on ability to complete physical measures) or patients attending the hospital or outpatient clinics for: outpatient fractures, arthritic complaints, plastic surgery, hemorrhoids, hernias, hydroceles, routine colon cancer screening, endoscopy, minor dermatologic disorders; or ii) community-based, defined as visitors or relatives of a patient from a non-cardiac ward, or an unrelated (not first-degree relative) visitor of a cardiac patient. Of the controls in INTERHEART, 58% were hospital-based and 36% of controls were community-based, and results were similar with both types of controls. In the remainder of the controls, 3% were from an undocumented source, and 3% were recruited through the WHO MONICA study in Göteborg, Sweden. Exclusion criteria for controls were identical to those described for cases. Structured questionnaires were administered to all cases and controls to obtain information on demographic factors (including self-reported ethnicity) as well as socioeconomic and health status. For this project, we analyzed individuals with self-reported ethnicity defined as "European".

Stage 4 study. Stage 4 consisted of cases with early-onset MI and controls free of MI from Iceland.

The **deCODE study** patients with MI were recruited through the cardiovascular disease genetics program at deCODE⁸. Individuals who suffered an MI were identified from a registry of over 10,000 individuals who: a) had an MI before the age of 75 in Iceland in the years 1981 to

2002 and satisfy the MONICA criteria, or had MI discharge diagnosis from the major hospitals in Reykjavik in the years 2003 and 2005. MI diagnoses of all individuals in the registry follow strict diagnostic criteria based on signs, symptoms, electrocardiograms, cardiac enzymes and necropsy findings. The patients were contacted through collaborating physicians in the CVD genetics programs at deCODE. For this study, we restricted analysis to cases of MI in men < 50 years of age and women < 60 years of age. The controls used for the study were recruited as a part of various genetic programs at deCODE. The medical history of the controls were unknown unless they had also participated in any of the CVD genetic programs (i.e. MI, stroke, peripheral vascular disease, type II diabetes, obesity, familial combined hyperlipidemia, coronary restenosis, and hypertension). Individuals with known MI, stroke, peripheral vascular or coronary artery disease were excluded as controls.

Informed consent. 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.

MI genotype score. We modeled the cumulative number of MI risk alleles carried by each participant in Stage 1. We constructed a score from the [nine SNPs](#) exceeding $P < 5 \times 10^{-8}$ in [Tables 2 and 3](#). The score was composed of allelic dosage (observed counts of 0, 1, or 2 for genotyped SNPs, or fractional allele counts between 0.0 and 2.0 estimated from the imputation procedure for imputed SNPs), weighted by the effect size of that allele on the MI phenotype (to minimize a potential "winner's curse", the effect size was drawn from the combined Stage 1 + 2 + 3 + 4 evidence), and summed across SNPs. We tested the association of genotype score with MI using logistic regression models after accounting for age, gender, and two principal

components of ancestry. We set the lowest quintile of MI genotype score as the referent group and estimated the increase in odds for MI associated with the remaining quintile groups.

Common and rare CNV analysis. Utilizing a previously defined copy number polymorphism map based on HapMap, we genotyped a set of polymorphic (greater than 1% sample frequency) autosomal deletion and duplication variants using the CANARY algorithm²³. We first conducted quality control filtering at the sample level. We assessed the initial 6,042 samples for quality in copy-number genotyping using three quality metrics reported by the Birdseye method. First, we used estimates generated by the Birdseye Hidden Markov Model²³ to remove any sample which was greater or less than three standard deviations from the average estimate of copy number, genome-wide (the average being approximately two copies). Second, we measured the variability in SNP and copy number polymorphism probe intensities, with each standardized per chromosome. We removed any sample with excessive variability in these estimates on average genome-wide (> 3 standard deviations than the average genome-wide). Next, we removed any sample where more than 2 chromosomes failed any of these three metrics (> 3 standard deviations in estimated copy number or excessive SNP or CNV variability for chromosome). Finally, for samples that had 1 or 2 chromosomes failing these measures, rather than failing the sample, we treated the data as missing. As a result, 5,648 samples were copy number genotyped with CANARY software²³.

We genotyped these samples for the previously defined set of 1,315 copy number polymorphisms characterized on the HapMap sample²⁴. As an initial quality control step, we removed any variant where more than 10% of the copy calls were uncertain (confidence score > 0.1) or missing. In addition, we focused on a set of polymorphisms where at least one allele had

a frequency greater than 1%. This restricted our analysis to 614 copy number polymorphic regions. An additional 59 CNVs were removed for inconsistent genotyping. Thus, we focused on a set of 554 copy number variable regions observed to be polymorphic and well genotyped in a set of 2,783 cases and 2,865 controls that passed copy-number sample quality control.

Association testing was performed using a logistic regression model, where copy number was used as a predictor of early-onset MI. We included two principal components that estimated fine-scale population stratification as covariates in the model. Analyses were conducted using PLINK software.

To detect rare CNVs, we used a Hidden Markov Model, as implemented in the Birdseye package, and focused on rare (less than ~1% sample frequency) and large (greater than 100kb) autosomal deletions and duplications. Using methods recently described²⁵, we evaluated case/control differences in rare CNVs across three parameters: genome-wide CNV rate, number of genes intersected by CNVs, and the total kilobase extent. After filtering for a Birdseye LOD score of 10 or more, and a size of 100kb or greater, we observed 78,013 autosomal CNV calls. Based on visual inspection, we removed a small number of CNVs that were outliers, greater than 5Mb or with a size (kb) to LOD score ratio less than 100. We excluded CNVs in regions of common copy number variation by removing any CNV with more than 50% of its length spanning either i) a genomic interval with 60 or more CNVs in this sample, or ii) a region defined as common based on HapMap data (greater than 3% frequency). We joined nearby, type-consistent CNVs if the combined length of individuals CNVs was at least 80% of the length of those CNVs merged. Finally, we removed outlier individuals with greater than either 10Mb of CNVs or 30 CNVs. This resulted in a final sample of 5,955 individuals and 8,065 CNVs

(39% deletions) for the rare CNV analysis. The mean number of rare CNVs per individual was 1.35 and the median was 1. After filtering, 3,986 individuals (67%) had at least one CNV.

The primary genome-wide burden analysis compared the mean rate of CNVs per person in cases versus controls ("CNV-count"). Following the analysis applied to a recent schizophrenia dataset, we also considered the number of genes (RefSeq gene list, hg18 co-ordinates, +/- 20kb around the largest transcript per gene) intersected by one or more CNVs. Finally, we considered the total kb extent of rare CNVs in cases versus controls. Case/control differences in each of these three metrics were evaluated by permutation as 1-sided tests (predicting a greater burden in cases), controlling for study site.

We also asked whether specific loci showed greater rates in CNVs in cases versus controls, using a similar permutation procedure and calculating empirical p-values corrected for genome-wide multiple testing. Given the absence of compelling evidence from the burden analyses to implicate a particular class of rare CNVs, we performed a single analysis of all 8,065 rare CNVs rather than further stratifying size, type or other characteristics. We tested: i) each unique position of the genome in terms of the number of case and control CNVs spanning that position and ii) a gene-wise statistic, comparing the total number of CNVs intersecting that gene (+/- 20kb) in cases and controls.

All filtering and analysis of CNVs used PLINK software.

SUPPLEMENTARY NOTE

Membership of the Wellcome Trust Case Control Consortium

Management Committee: Paul R Burton¹, David G Clayton², Lon R Cardon³, Nick Craddock⁴, Panos Deloukas⁵, Audrey Duncanson⁶, Dominic P Kwiatkowski^{3,5}, Mark I McCarthy^{3,7}, Willem H Ouwehand^{8,9}, Nilesh J Samani¹⁰, John A Todd², Peter Donnelly (Chair)¹¹

Data and Analysis Committee: Jeffrey C Barrett³, Paul R Burton¹, Dan Davison¹¹, Peter Donnelly¹¹, Doug Easton¹², David M. Evans³, Hin-Tak Leung², Jonathan L Marchini¹¹, Andrew P Morris³, Chris CA Spencer¹¹, Martin D Tobin¹, Lon R Cardon (Co-chair)³, David G Clayton (Co-chair)²

UK Blood Services & University of Cambridge Controls: Antony P Attwood^{5,8}, James P Boorman^{8,9}, Barbara Cant⁸, Ursula Everson¹³, Judith M Hussey¹⁴, Jennifer D Jolley⁸, Alexandra S Knight⁸, Kerstin Koch⁸, Elizabeth Meech¹⁵, Sarah Nutland², Christopher V Prowse¹⁶, Helen E Stevens², Niall C Taylor⁸, Graham R Walters¹⁷, Neil M Walker², Nicholas A Watkins^{8,9}, Thilo Winzer⁸, John A Todd², Willem H Ouwehand^{8,9}

1958 Birth Cohort Controls: Richard W Jones¹⁸, Wendy L McArdle¹⁸, Susan M Ring¹⁸, David P Strachan¹⁹, Marcus Pembrey^{18,20}

Bipolar Disorder (Aberdeen): Gerome Breen²¹, David St Clair²¹; **(Birmingham):** Sian Caesar²², Katherine Gordon-Smith^{22,23}, Lisa Jones²²; **(Cardiff):** Christine Fraser²³, Elaine K Green²³, Detelina Grozeva²³, Marian L Hamshere²³, Peter A Holmans²³, Ian R Jones²³, George Kirov²³, Valentina Moskvina²³, Ivan Nikolov²³, Michael C O'Donovan²³, Michael J Owen²³, Nick Craddock²³; **(London):** David A Collier²⁴, Amanda Elkin²⁴, Anne Farmer²⁴, Richard Williamson²⁴, Peter McGuffin²⁴; **(Newcastle):** Allan H Young²⁵, I Nicol Ferrier²⁵

Coronary Artery Disease (Leeds): Stephen G Ball²⁶, Anthony J Balmforth²⁶, Jennifer H Barrett²⁶, D Timothy Bishop²⁶, Mark M Iles²⁶, Azhar Maqbool²⁶, Nadira Yuldasheva²⁶, Alistair S Hall²⁶; **(Leicester):** Peter S Braund¹⁰, Paul R Burton¹, Richard J Dixon¹⁰, Massimo Mangino¹⁰, Suzanne Stevens¹⁰, Martin D Tobin¹, John R Thompson¹, Nilesh J Samani¹⁰

Crohn's Disease (Cambridge): Francesca Bredin²⁷, Mark Tremelling²⁷, Miles Parkes²⁷; **(Edinburgh):** Hazel Drummond²⁸, Charles W Lees²⁸, Elaine R Nimmo²⁸, Jack Satsangi²⁸; **(London):** Sheila A Fisher²⁹, Alastair Forbes³⁰, Cathryn M Lewis²⁹, Clive M Onnie²⁹, Natalie J Prescott²⁹, Jeremy Sanderson³¹, Christopher G Mathew²⁹; **(Newcastle):** Jamie Barbour³², M Khalid Mohiuddin³², Catherine E Todhunter³², John C Mansfield³²; **(Oxford):** Tariq Ahmad³³, Fraser R Cummings³³, Derek P Jewell³³

Hypertension (Aberdeen): John Webster³⁴; **(Cambridge):** Morris J Brown³⁵, David G Clayton²; **(Evry, France):** G Mark Lathrop³⁶; **(Glasgow):** John Connell³⁷, Anna Dominiczak³⁷; **(Leicester):** Nilesh J Samani¹⁰; **(London):** Carolina A Braga Marcano³⁸, Beverley Burke³⁸, Richard Dobson³⁸, Johannie Gungadoo³⁸, Kate L Lee³⁸, Patricia B Munroe³⁸, Stephen J Newhouse³⁸, Abiodun Onipinla³⁸, Chris Wallace³⁸, Mingzhan Xue³⁸, Mark Caulfield³⁸; **(Oxford):** Martin Farrall³⁹

Rheumatoid Arthritis: Anne Barton⁴⁰, Ian N Bruce⁴⁰, Hannah Donovan⁴⁰, Steve Eyre⁴⁰, Paul D Gilbert⁴⁰, Samantha L Hider⁴⁰, Anne M Hinks⁴⁰, Sally L John⁴⁰, Catherine Potter⁴⁰, Alan J Silman⁴⁰, Deborah PM Symmons⁴⁰, Wendy Thomson⁴⁰, Jane Worthington⁴⁰

Type 1 Diabetes: David G Clayton², David B Dunger^{2,41}, Sarah Nutland², Helen E Stevens², Neil M Walker², Barry Widmer^{2,41}, John A Todd²

Type 2 Diabetes (Exeter): Timothy M Frayling^{42,43}, Rachel M Freathy^{42,43}, Hana Lango^{42,43}, John R B Perry^{42,43}, Beverley M Shields⁴³, Michael N Weedon^{42,43}, Andrew T Hattersley^{42,43}; **(London):** Graham A Hitman⁴⁴; **(Newcastle):** Mark Walker⁴⁵; **(Oxford):** Kate S Elliott^{3,7}, Christopher J Groves⁷, Cecilia M Lindgren^{3,7}, Nigel W Rayner^{3,7}, Nicholas J Timpson^{3,46}, Eleftheria Zeggini^{3,7}, Mark I McCarthy^{3,7}

Tuberculosis (Gambia): Melanie Newport⁴⁷, Giorgio Sirugo⁴⁷; **(Oxford):** Emily Lyons³, Fredrik Vannberg³, Adrian VS Hill³

Ankylosing Spondylitis: Linda A Bradbury⁴⁸, Claire Farrar⁴⁹, Jennifer J Pointon⁴⁸, Paul Wordsworth⁴⁹, Matthew A Brown^{48,49}

AutoImmune Thyroid Disease: Jayne A Franklyn⁵⁰, Joanne M Heward⁵⁰, Matthew J Simmonds⁵⁰, Stephen CL Gough⁵⁰

Breast Cancer: Sheila Seal⁵¹, Michael R Stratton^{51,52}, Nazneen Rahman⁵¹

Multiple Sclerosis: Maria Ban⁵³, An Goris⁵³, Stephen J Sawcer⁵³, Alastair Compston⁵³

Gambian Controls (Gambia): David Conway⁴⁷, Muminatou Jallow⁴⁷, Melanie Newport⁴⁷, Giorgio Sirugo⁴⁷; **(Oxford):** Kirk A Rockett³, Dominic P Kwiatkowski^{3,5}

DNA, Genotyping, Data QC and Informatics (Wellcome Trust Sanger Institute, Hinxton): Suzannah J Bumpstead⁵, Amy Chaney⁵, Kate Downes^{2,5}, Mohammed JR Ghoris⁵, Rhian Gwilliam⁵, Sarah E Hunt⁵, Michael Inouye⁵, Andrew Keniry⁵, Emma King⁵, Ralph McGinnis⁵, Simon Potter⁵, Rathi Ravindrarajah⁵, Pamela Whittaker⁵, Claire Widden⁵, David Withers⁵, Panos Deloukas⁵; **(Cambridge):** Hin-Tak Leung², Sarah Nutland², Helen E Stevens², Neil M Walker², John A Todd²

Statistics (Cambridge): Doug Easton¹², David G Clayton²; **(Leicester):** Paul R Burton¹, Martin D Tobin¹; **(Oxford):** Jeffrey C Barrett³, David M Evans³, Andrew P Morris³, Lon R Cardon³; **(Oxford):** Niall J Cardin¹¹, Dan Davison¹¹, Teresa Ferreira¹¹, Joanne Pereira-Gale¹¹, Ingeleif B Hallgrimsdóttir¹¹, Bryan N Howie¹¹, Jonathan L Marchini¹¹, Chris CA Spencer¹¹, Zhan Su¹¹, Yik Ying Teo^{3,11}, Damjan Vukcevic¹¹, Peter Donnelly¹¹

PIs: David Bentley^{5,54}, Matthew A Brown^{48,49}, Lon R Cardon³, Mark Caulfield³⁸, David G Clayton², Alistair Compston⁵³, Nick Craddock²³, Panos Deloukas⁵, Peter Donnelly¹¹, Martin Farrall³⁹, Stephen CL Gough⁵⁰, Alistair S Hall²⁶, Andrew T Hattersley^{42,43}, Adrian VS Hill³, Dominic P Kwiatkowski^{3,5}, Christopher G Mathew²⁹, Mark I McCarthy^{3,7}, Willem H Ouwehand^{8,9}, Miles Parkes²⁷, Marcus Pembrey^{18,20}, Nazneen Rahman⁵¹, Nilesh J Samani¹⁰, Michael R Stratton^{51,52}, John A Todd², Jane Worthington⁴⁰

¹ Genetic Epidemiology Group, Department of Health Sciences, University of Leicester, Adrian Building, University Road, Leicester, LE1 7RH, UK; ² Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Cambridge, CB2 0XY, UK; ³ Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; ⁴ Department of Psychological Medicine, Henry Wellcome Building, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK; ⁵ The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; ⁶ The Wellcome Trust, Gibbs Building, 215 Euston Road, London NW1 2BE, UK; ⁷ Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Churchill Hospital, Oxford, OX3 7LJ, UK; ⁸ Department of Haematology, University of Cambridge, Long Road, Cambridge, CB2 2PT, UK; ⁹ National Health Service Blood and Transplant, Cambridge Centre, Long Road, Cambridge, CB2 2PT,

UK; ¹⁰ Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, UK; ¹¹ Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK; ¹² Cancer Research UK Genetic Epidemiology Unit, Strangeways Research Laboratory, Worts Causeway, Cambridge CB1 8RN, UK; ¹³ National Health Service Blood and Transplant, Sheffield Centre, Longley Lane, Sheffield S5 7JN, UK; ¹⁴ National Health Service Blood and Transplant, Brentwood Centre, Crescent Drive, Brentwood, CM15 8DP, UK; ¹⁵ The Welsh Blood Service, Ely Valley Road, Talbot Green, Pontyclun, CF72 9WB, UK; ¹⁶ The Scottish National Blood Transfusion Service, Ellen's Glen Road, Edinburgh, EH17 7QT, UK; ¹⁷ National Health Service Blood and Transplant, Southampton Centre, Coxford Road, Southampton, SO16 5AF, UK; ¹⁸ Avon Longitudinal Study of Parents and Children, University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ, UK; ¹⁹ Division of Community Health Services, St George's University of London, Cranmer Terrace, London SW17 0RE, UK; ²⁰ Institute of Child Health, University College London, 30 Guilford St, London WC1N 1EH, UK; ²¹ University of Aberdeen, Institute of Medical Sciences, Foresterhill, Aberdeen, AB25 2ZD, UK; ²² Department of Psychiatry, Division of Neuroscience, Birmingham University, Birmingham, B15 2QZ, UK; ²³ Department of Psychological Medicine, Henry Wellcome Building, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK; ²⁴ SGDP, The Institute of Psychiatry, King's College London, De Crespigny Park Denmark Hill London SE5 8AF, UK; ²⁵ School of Neurology, Neurobiology and Psychiatry, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne, NE1 4LP, UK; ²⁶ LIGHT and LIMM Research Institutes, Faculty of Medicine and Health, University of Leeds, Leeds, LS1 3EX, UK; ²⁷ IBD Research Group, Addenbrooke's Hospital, University of Cambridge, Cambridge, CB2 2QQ, UK; ²⁸ Gastrointestinal Unit, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU UK; ²⁹ Department of Medical & Molecular Genetics, King's College London School of Medicine, 8th Floor Guy's Tower, Guy's Hospital, London, SE1 9RT, UK; ³⁰ Institute for Digestive Diseases, University College London Hospitals Trust, London, NW1 2BU, UK; ³¹ Department of Gastroenterology, Guy's and St Thomas' NHS Foundation Trust, London, SE1 7EH, UK; ³² Department of Gastroenterology & Hepatology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne, NE1 4LP, UK; ³³ Gastroenterology Unit, Radcliffe Infirmary, University of Oxford, Oxford, OX2 6HE, UK; ³⁴ Medicine and Therapeutics, Aberdeen Royal Infirmary, Foresterhill, Aberdeen, Grampian AB9 2ZB, UK; ³⁵ Clinical Pharmacology Unit and the Diabetes and Inflammation Laboratory, University of Cambridge, Addenbrookes Hospital, Hills Road, Cambridge CB2 2QQ, UK; ³⁶ Centre National de Genotypage, 2, Rue Gaston Cremieux, Evry, Paris 91057.; ³⁷ BHF Glasgow Cardiovascular Research Centre, University of Glasgow, 126 University Place, Glasgow, G12 8TA, UK; ³⁸ Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London, Queen Mary's School of Medicine, Charterhouse Square, London EC1M 6BQ, UK; ³⁹ Cardiovascular Medicine, University of Oxford, Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK; ⁴⁰ arc Epidemiology Research Unit, University of Manchester, Stopford Building, Oxford Rd, Manchester, M13 9PT, UK; ⁴¹ Department of Paediatrics, University of Cambridge, Addenbrooke's Hospital, Cambridge, CB2 2QQ, UK; ⁴² Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, Magdalen Road, Exeter EX1 2LU UK; ⁴³ Diabetes Genetics, Institute of Biomedical and Clinical Science, Peninsula Medical School, Barrack Road, Exeter EX2 5DU UK; ⁴⁴ Centre for Diabetes and Metabolic Medicine, Barts and The London, Royal London Hospital, Whitechapel, London, E1 1BB UK; ⁴⁵ Diabetes Research Group, School of

Clinical Medical Sciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK; ⁴⁶ The MRC Centre for Causal Analyses in Translational Epidemiology, Bristol University, Canynge Hall, Whiteladies Rd, Bristol BS2 8PR, UK; ⁴⁷ MRC Laboratories, Fajara, The Gambia; ⁴⁸ Diamantina Institute for Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Woolloongabba, Qld 4102, Australia; ⁴⁹ Botnar Research Centre, University of Oxford, Headington, Oxford OX3 7BN, UK; ⁵⁰ Department of Medicine, Division of Medical Sciences, Institute of Biomedical Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; ⁵¹ Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Sutton, SM2 5NG, UK; ⁵² Cancer Genome Project, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; ⁵³ Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK; ⁵⁴ PRESENT ADDRESS: Illumina Cambridge, Chesterford Research Park, Little Chesterford, Nr Saffron Walden, Essex, CB10 1XL, UK.

List of members of Cardiogenics

Associazione per lo Studio della Trombosi in Cardiologia, 27100 Pavia, Italy: Diego Ardissino

EUROIMMUN AG, 23560 Lübeck, Germany: Christian Krüger, Lars Komorowski, Christian Probst, Ulf Steller

Helmholtz Institut Zentrum München, 85764 Neuherberg, Germany: Thomas Meitinger, Erich Wichmann, Peter Lichtner

Institut National de la Santé et de la Recherche Médicale (INSERM), 75654 Paris, France: Francois Cambien, Laurence Tiret

Johannes Gutenberg-University of Mainz, 55099 Mainz, Germany: Stefan Blankenberg, Karl Lackner, Thomas Münzel

Queen's University of Belfast, Belfast BT71NN, UK: Alun Evans, Frank Kee

Sanquin Blood Supply Foundation, 1066 CX Amsterdam, The Netherlands: Ellen van der Schoot, Jaap Jan Zwaginga

Technische Universität München, 80333 München, Germany: Dietlind Zohlnhoefer

TRIUM Analysis Online GmbH, 81677 München, Germany: Martin Daumer, Michael Scholz

The Chancellor, Master and Scholar of the University of Cambridge, Cambridge CB2 1TS, UK: Willem Ouwehand, Richard Farndale, Nicholas Watkins

University of Leeds, Leeds LS2 9JT Leeds, UK: Alistair Hall, Stephen Ball, Tim Bishop

University of Leicester, Leicester LE1 7RH, UK: Nilesh Samani, Alison Goodall, Paul Burton

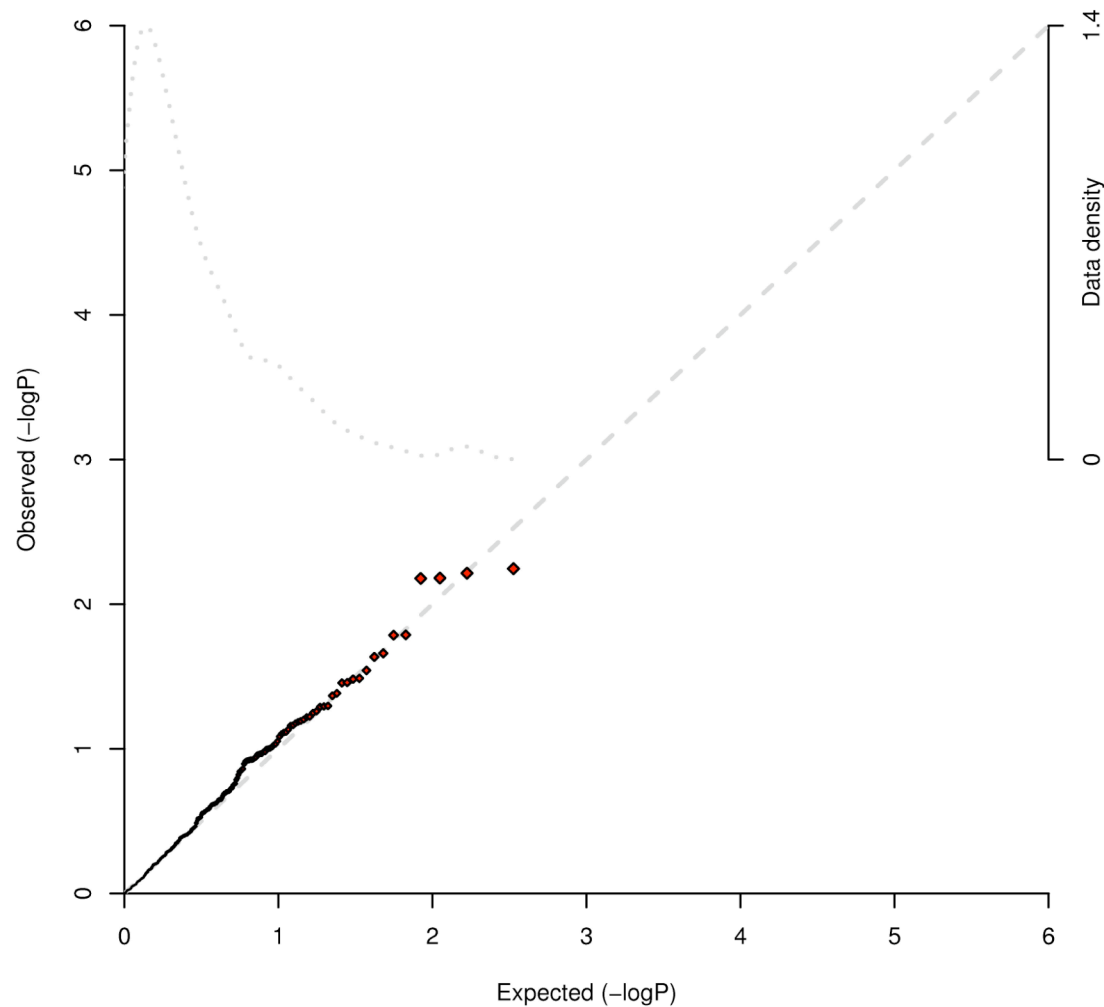
University of Lübeck, 23538 Lübeck, Germany: Heribert Schunkert, Jeanette Erdmann, Andreas Ziegler

University of Regensburg, 93053 Regensburg, Germany: Christian Hengstenberg, Gerd Schmitz

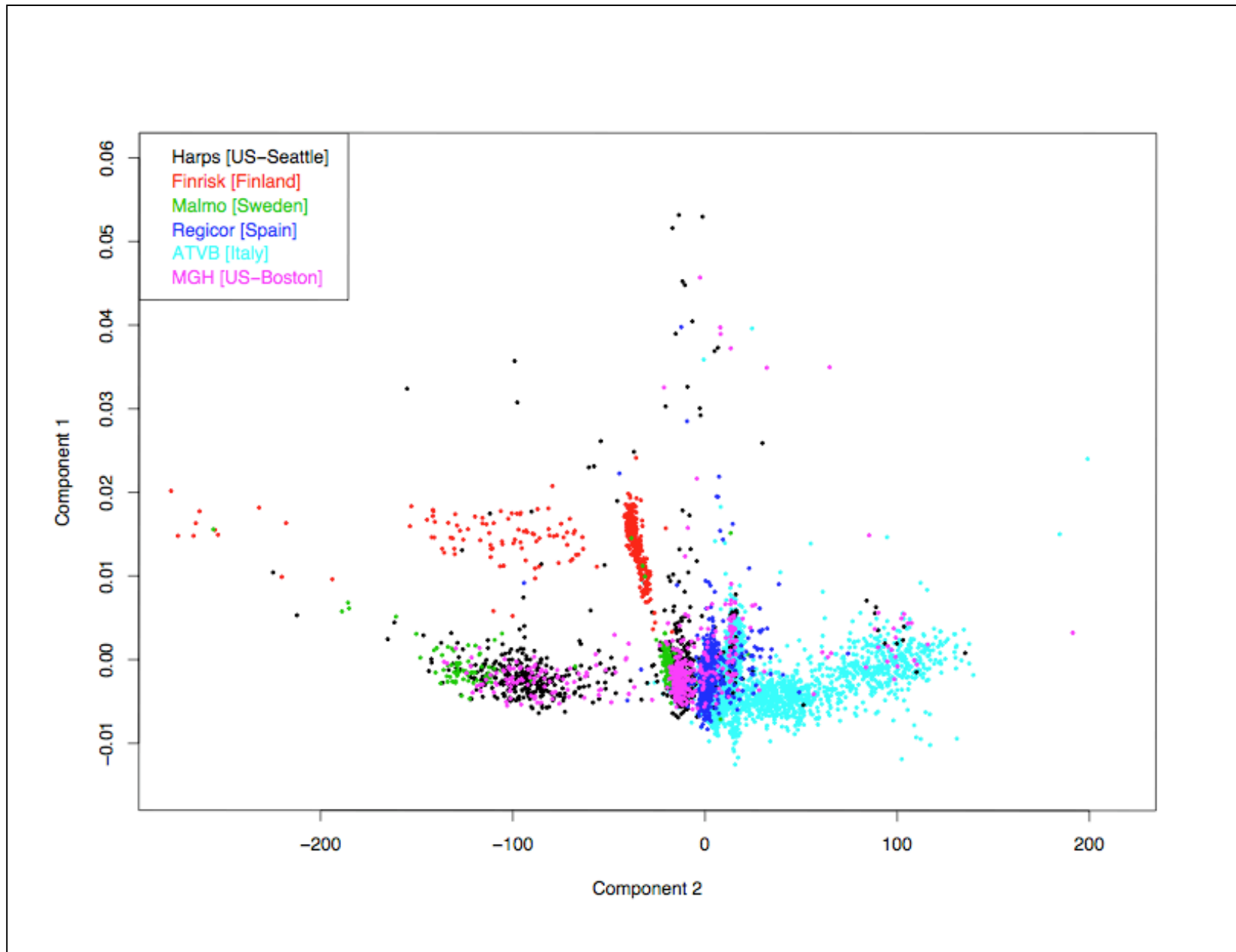
University of Uppsala, 75105 Uppsala, Sweden: Ann-Christin Syvänen, Tomas Axelsson, Ulrika Liljedahl

Wellcome Trust Sanger Institute, Hinxton Cambridge CB10 1SA, UK: Panos Deloukas, Kate Rice

Supplementary Figure 1. Plot of observed versus expected P value distribution for association of 554 common copy number variants with early-onset myocardial infarction. The CANARY algorithm was used to test 554 commonly segregating CNVs (> 1% frequency) for association with early-onset MI in 2,783 cases and 2,865 controls that passed sample quality control for CNV analysis (Methods). The estimated genomic control lambda for the entire set of CNVs was ~1.23; for 316 CNVs with allele frequency greater than 5%, lambda was ~1.05. We did not observe any CNV with evidence for association surpassing our pre-specified threshold for replication of $P < 0.001$. The observed versus expected P value distribution did not show deviation from the null distribution.



Supplementary Figure 2. Principal components of ancestry in the six component Stage 1 studies



Supplementary Table 1. Participant characteristics of case and control subjects in Stage 2 studies

Study	WTCCC MI		German MI Family Study I		PennCATH		MedSTAR	
	cases	Controls	cases	controls	cases	controls	cases	controls
N*	1,414	2,938	875	1,644	415	468	420	447
Effective sample size	3,922 cases and 3,922 controls in Stage 2							
Ascertainment scheme	community-based	community-based	hospital-based men "60 or women " 65	population-based	hospital-based	hospital-based	hospital-based	hospital-based
MI age criterion	<66 years	--	--	--	<66 years	--	<66 years	--
Genotyping platform	Affymetrix 500K	Affymetrix 500K	Affymetrix 500K	Affymetrix 500K	Affymetrix 6.0	Affymetrix 6.0	Affymetrix 6.0	Affymetrix 6.0
Country of origin	U.K.	U.K.	Germany	Germany	U.S.	U.S.	U.S.	U.S.
Mean age (y)[†]	49.3±7.9	44.7±9.3	50.2±7.9	62.5±10.1	50.2±7.5	61.7±9.6	46.8±6.8	59.7±8.9
Female gender (%)	20.2	49.2	32.5	50.5	17.6	51.7	23.8	48.8
Ever cigarette smoking (%)	78.3	NA	70.3	49.3	45.0	34.8	57.5	49.9
Hypertension (%)	41.1	NA	86.5	62.7	57.5	49.2	76.7	60.6
Diabetes mellitus (%)	10.9	NA	12.3	11.0	25.6	11.2	30.2	19.0
Hypercholesterolemia (%)	79.3	NA	76.1	78.0	80.2	62.3	88.4	65.1
Body mass index (kg/m²)	27.6±4.3	NA	27.4±3.6	28.1±4.5	30.1±5.3	28.9±6.4	31.4±6.4	31.3±7.9

Values with '±' are means ± s.d. The body-mass index is the weight in kilograms divided by the square of the height in meters.

* Summary characteristics are provided for maximal number of available subjects. For any given genotype, a subset of these individuals may have been studied.

[†] Mean age at MI for cases and at age of recruitment for controls.

Supplementary Table 2. Participant characteristics of case and control subjects in Stage 3 and Stage 4 validation studies

Stage	Stage 3											Stage 4		
Study	AMI Gene Study /Dortmund Health Study		Verona Heart Study		Mid-America Heart Institute		Irish Family Study		German MI Family Study II		INTERHEART European Ancestry		deCODE	
	cases	controls	cases	controls	cases	controls	cases	controls	cases	controls	cases	controls	cases	controls
N ^a	809	1,132	510	388	811	650	487	629	1,222	1,298	1,886	2,231	728	29,218
Effective sample size	4321 cases, 4321 controls in Stage 3											1503 cases, 1503 controls in Stage 4		
Ascertainment scheme	hospital-based	hospital-based	hospital-based	hospital-based	hospital-based	outpatients	hospital-based	older siblings of cases	hospital-based	hospital-based	hospital-based	hospital-based	population-based	population-based
MI age criterion	men <65	--	men "65 or women " 65	--	no age criterion	--	men "55 or women " 60	--	men "60 or women " 65	--	no age criterion	--	men<50 or women <60	--
Genotyping platform	Sequenom	Sequenom	Sequenom	Sequenom	Sequenom	Sequenom	Sequenom	Sequenom	Affymetrix 6.0	Affymetrix 6.0	Sequenom	Sequenom	Illumina 300/370K	Illumina 300/370K
Country of origin	Germany	Germany	Italy	Italy	U.S.	U.S.	Northern Ireland	Northern Ireland	Germany	Germany	U.S. and Europe	U.S. and Europe	Iceland	Iceland
Mean age (y) ^b	52.2±8.2	52.6±13.7	57.0±9.1	58.9±12.1	61.5±12.7	60.7±12.4	45.9±6.7	55.7±8.0	51.3±7.6	51.2±11.9	61.2±12.1	60.5±12.0	56.7 ± 10.6	48.9 ± 21.6
Female gender (%)	0	53.1	10.6	34.8	32.1	39.0	20.1	55.2	20.3	47.9	28.9	31.3	40.8	61.3
Ever cigarette smoking (%)	65.9	55.9	77.0	43.6	71.0	50.2	82.5	57.6	64.7	55.6	65.3	54.7	94.2	68.9
Hypertension (%) ^c	71.8	62.6	60.9	40.3	60.1	51.1	27.9	28.9	88.1	41.7	44.5	31.4	--	--
Diabetes mellitus (%) ^c	14.6	8.0	18.1	7.3	23.9	11.9	9.4	5.6	15.5	3.5	15.0	6.5	--	--
Hypercholesterolemia (%) ^c	60.6	68.1	74.4	63.0	58.8	50.4	85.8	8.6	80.8	67.3	22.8	15.8	--	--
Body mass index (kg/m ²)	28.0±4.7	27.5±4.9	26.9±3.5	25.4±3.4	29.4	27.9	28.5±4.3	28.2±5.0	27.9±4.0	27.4±4.6	27.3±4.5	26.9±6.9	28.6 ± 5.1	27.0 ± 5.4

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 Summary characteristics are provided for maximal number of available subjects. For any given genotype, a subset of these individuals may have been studied.

^b Mean age at MI for cases and at age of recruitment for controls.

^c Covariate data not routinely ascertained in deCODE study.

Supplementary Table 3. Association of common CNVs with early-onset myocardial infarction

CNPID	Chromosome	Base Pair	N	Alleles	Frequencies	Beta	P Value	Gene Annotation
CNP10955	5	106259294	5646	0,1,2,3,4	0.000354,0.0273,0.96,0.0122,0.000354	-0.4226	0.001502	Gene Desert [+/- 250kb]
CNP12153	14	27739306	5647	1,2,3,4	0.0023,0.981,0.0161,0.00106	0.5933	0.001574	CR597916
CNP10460	3	1760053	5647	0,1,2,3,4	0.000354,0.0446,0.945,0.00921,0.000531	0.3594	0.001728	Gene Desert [+/- 250kb]
CNP10834	4	162557332	5607	1,2,3,4	0.00339,0.979,0.0171,0.000535	0.5677	0.002527	FSTL5
CNP11245	7	87316216	5647	0,1,2,3,4	0.000177,0.0159,0.979,0.00478,0.000177	-0.5482	0.003281	ABCB1,RUNDC3B,SL25A40,ASK/H37,DBF4,ADAM22
CNP10523	3	62689310	5647	1,2,3	0.00106,0.988,0.011	0.7441	0.004402	CADPS
CNP980	6	81343965	5646	0,1,2,3	0.00283,0.0889,0.908,0.000531	0.2527	0.004561	BCKDHB
CNP435	3	82955601	5647	0,1,2,3,4	0.00354,0.106,0.873,0.0158,0.00142	0.2077	0.00492	Gene Desert [+/- 250kb]
CNP11172	7	151090	5647	0,1,2,3,4	0.000177,0.00283,0.939,0.0556,0.00195	0.2983	0.005335	PDGFA,FAM20C
CNP164	1	219408859	5648	0,1,2,3,4	0.0184,0.246,0.72,0.015,0.000885	-0.1446	0.00537	c1orf140
CNP12744	20	59013621	5647	1,2,3,4	0.000177,0.963,0.0326,0.00478	0.3268	0.007434	CDH4
CNP10620	3	168009595	5647	1,2,3,4	0.000531,0.986,0.0136,0.000354	0.5794	0.009685	Gene Desert [+/- 250kb]

References

1. No evidence of association between prothrombotic gene polymorphisms and the development of acute myocardial infarction at a young age. *Circulation* **107**, 1117-22 (2003).
2. Meiner, V. et al. Cholesteryl ester transfer protein (CETP) genetic variation and early onset of non-fatal myocardial infarction. *Ann Hum Genet* **72**, 732-41 (2008).
3. Low, A.F. et al. Aging syndrome genes and premature coronary artery disease. *BMC Med Genet* **6**, 38 (2005).
4. Senti, M., Tomas, M., Marrugat, J. & Elosua, R. Paraoxonase1-192 polymorphism modulates the nonfatal myocardial infarction risk associated with decreased HDLs. *Arterioscler Thromb Vasc Biol* **21**, 415-20 (2001).
5. Vartiainen, E. et al. Cardiovascular risk factor changes in Finland, 1972-1997. *Int J Epidemiol* **29**, 49-56 (2000).
6. Berglund, G., Elmstahl, S., Janzon, L. & Larsson, S.A. The Malmo Diet and Cancer Study. Design and feasibility. *J Intern Med* **233**, 45-51 (1993).
7. Samani, N.J. et al. Genomewide association analysis of coronary artery disease. *N Engl J Med* **357**, 443-53 (2007).
8. Helgadottir, A. et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491-3 (2007).
9. Helgadottir, A. et al. A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nat Genet* **38**, 68-74 (2006).
10. Grant, S.F. et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* **38**, 320-3 (2006).
11. Helgadottir, A. et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet* **40**, 217-24 (2008).
12. Steinthorsdottir, V. et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* **39**, 770-5 (2007).
13. Vennemann, M.M., Hummel, T. & Berger, K. The association between smoking and smell and taste impairment in the general population. *J Neurol* (2008).
14. Happe, S., Vennemann, M., Evers, S. & Berger, K. Treatment wish of individuals with known and unknown restless legs syndrome in the community. *J Neurol* (2008).
15. Martinelli, N. et al. FADS genotypes and desaturase activity estimated by the ratio of arachidonic acid to linoleic acid are associated with inflammation and coronary artery disease. *Am J Clin Nutr* **88**, 941-9 (2008).
16. Morgan, T.M., Krumholz, H.M., Lifton, R.P. & Spertus, J.A. Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *Jama* **297**, 1551-61 (2007).
17. Morgan, T.M. et al. Investigation of 89 candidate gene variants for effects on all-cause mortality following acute coronary syndrome. *BMC Med Genet* **9**, 66 (2008).
18. Meng, W. et al. Chromosome 9p21.3 is associated with early-onset coronary heart disease in the Irish population. *Dis Markers* **25**, 81-5 (2008).

19. Erdmann, J., Grosshennig, A., Braund, P.S., König, I.R., Hengstenberg, C, et al. Novel susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat Genet*, in press.
20. Yusuf, S. et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* **364**, 937-52 (2004).
21. Serre, D. et al. Correction of population stratification in large multi-ethnic association studies. *PLoS ONE* **3**, e1382 (2008).
22. Anand S., X.C., Pare G., Montpetit A., Rangarajan S., et al. Genetic variants associated with myocardial infarction risk factors in over 8,000 individuals from five ethnic groups: the INTERHEART Genetics Study. *Circ Cardiovasc Genet*, in press.
23. Korn, J.M. et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet* **40**, 1253-60 (2008).
24. McCarroll, S.A. et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* **40**, 1166-74 (2008).
25. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237-41 (2008).