Exome-wide association study of plasma lipids in >300,000 individuals

SUPPLEMENTARY NOTES

l. Supplementary Table 12: Association Results for 444 independently associated variants with lipid traits
m. Supplementary Table 13: Loci where protein-altering variant is top signal or protein-altering variant explains the GWAS signal
n. Supplementary Table 14: 59 loci where there's a protein-altering variant that is either the top signal, explains the signal or is independent.
o. Supplementary Table 15: Association results for null mutations with $p < 0.001$
p. Supplementary Table 16: HDL-C variants and risk for age-related macular degeneration (AMD)
q. Supplementary Table 17: DNA variants in CETP robustly associate with HDL-C and risk for AMD
r. Supplementary Table 18: Thirty studies from populations of European ancestry contributing to *PCSK9* p.R46L on risk of T2D
s. Supplementary Table 19: Association of LDL-C variants with coronary artery disease (CAD) and type 2 diabetes (T2D)
t. Supplementary Table 20: Definitions of outcomes in UK Biobank PheWAS
u. Supplementary Table 21: sgRNA sequences for functional follow-up experiments

I. Study Descriptions

British 1958 Birth Cohort (1958BC)

The National Child Development Study (NCDS) follows the lives of 17,000 people born in England, Scotland and Wales in a single week of 1958[1]. Also known as the 1958 Birth Cohort Study, it collects information on physical and educational development, economic circumstances, employment, family life, health behaviour, wellbeing, social participation and attitudes. The NCDS is managed by CLS and funded by the Economic and Social Research Council. Since the birth survey in 1958, there have been nine further 'sweeps' of all cohort members at ages 7, 11, 16, 23, 33, 42, 46, 50 and 55. In 2003 (at age 45), 9,000 cohort members also participated in a special bio-medical survey so we could learn more about how development, environments and lifestyles affect people's health.

Anglo–Danish–Dutch Study of Intensive Treatment in People with Screen-Detected Diabetes in Primary Care - Denmark screening cohort (ADDITION)

The Danish ADDITION Study (Anglo–Danish–Dutch Study of Intensive Treatment in People with Screen-Detected Diabetes in Primary Care) is a high-risk screening and intervention study for type 2 diabetes in general practice sampled by Department of General Practice at University of Aarhus, Denmark (ClinicalTrials.gov ID-no: NCT00237548)[2]. The 8,662 participants from the initial screening cohort with available DNA included 1,626 participants with screen-detected and untreated T2D and 7,036 non-diabetic subjects. Patients with T2D were diagnosed by two independent diabetic values at baseline investigation or at one-year follow-up. Phenotypes include anthropometrics, basal fasting biochemistry (e.g. plasma glucose, serum insulin, HbA1C and lipids) and health and lifestyle questionnaires. Here 2238 participants from the Danish screening cohort with information on lipids and exome chip were analyzed.

Age gene/environment susceptibility Reykjavik study (AGES)

The AGES study has been described previously[3]. The study was initiated in 2002 to examine genetic susceptibility and gene/environment interactions related to disease and disability in old age. The AGES study is comprised of 5764 individuals drawn from the Reykjavik Study, a population-based cohort comprised of individuals born between 1907 and 1935 and followed since 1967 by the Icelandic Heart Association. 3219 individuals chosen randomly among 5307 AGES individuals with 'mid-life' data available from the Reykjavik Study were genotyped on a genome-wide association (GWA) array. 2983 individuals randomly selected from the 3219 individuals with GWA were further genotyped for the ExomeChip.

Academic Medical Center Premature Atherosclerosis Study (AMCPAS)

Cases were recruited as part of a prospective cohort study (Academic Medical Centre Amsterdam Premature Atherosclerosis Study (AMC-PAS) with symptomatic CAD before the age of 51 years, defined as MI, coronary revascularization, or evidence of at least 70% stenosis in a major epicardial artery[4].

Atherosclerosis Risk in Communities Study (ARIC-AA, ARIC-EA)

The ARIC study has been described in detail previously[5]. Men and women aged 45-64 years at baseline were recruited from four communities: Forsyth County, North Carolina; Jackson, Mississippi; Minneapolis, Minnesota; and Washington County, Maryland. A total of 15,792 individuals, predominantly White (EA) and African American (AA), participated in the baseline examination in 1987-1989, with three additional triennial follow-up examinations and a fifth exam in 2011-2013.

Anglo-Scandinavian Cardiac Outcome Trial [Scandinavians] (ASCOT-SC)

ASCOT is a randomised control clinical trial investigating the cardiac outcomes of blood pressure lowering and lipid lowering treatments[6]. Of 19,342 hypertensive patients (40–79 years of age with at least three other cardiovascular risk factors) who were randomized to one of two antihypertensive regimens in ASCOT, 10,305 with non-fasting TC concentrations of 6.5 mmol/l or less (measured at the non-fasting screening visit) had been randomly assigned additional atorvastatin 10 mg or placebo. These patients formed the lipid-lowering arm of the study. Only a proportion of United Kingdom, Irish, Sweden, Norway, Finland and Denmark consented to contribute DNA and participate in genetic studies. Blood lipid levels used in this analysis were measured at the (fasting) randomization visit and LDL- C was estimated using the Friedewald equation. The ASCOT-SC sample analysed here is restricted to the patients from Scandinavia, of which a total of 2,468 were genotyped on the Exome-Chip and passed QC for inclusion into this analysis.

Anglo-Scandinavian Cardiac Outcome Trial [UK/Ireland] (ASCOT-UK)

ASCOT is a randomised control clinical trial investigating the cardiac outcomes of blood pressure lowering and lipid lowering treatments[6]. Of 19,342 hypertensive patients (40–79 years of age with at least three other cardiovascular risk factors) who were randomized to one of two antihypertensive regimens in ASCOT, 10,305 with non-fasting TC concentrations of 6.5 mmol/l or less (measured at the non-fasting screening visit) had been randomly assigned additional atorvastatin 10 mg or placebo. These patients formed the lipid-lowering arm of the study. Only a proportion of United Kingdom, Irish, Sweden, Norway, Finland and Denmark consented to contribute DNA and participate in genetic studies. Blood lipid levels used in this analysis were measured at the (fasting) randomization visit and LDL- C was estimated using the Friedewald equation. The ASCOT-UK sample analysed here is restricted to the patients from UK and Ireland, of which a total of 3,246 were genotyped on the Exome-Chip and passed QC for inclusion into this analysis.

Italian Atherosclerosis, Thrombosis, and Vascular Biology Working Group (ATVB-Cases, ATVB-Controls)

ATVB is a nationwide prospective case control study involving 1,693 patients hospitalised for a first ST segment elevation MI before the age of 45 years, and 1,668 healthy subjects matched for age, gender and geographical origin[7].

The BioImage Study (BioImage-African, BioImage-Asian, BioImage-European, BioImage-Hispanic)

The BioImage Study (BioImage Study: A Clinical Study of Burden of Atherosclerotic Disease in an At-Risk Population, NCT00738725), a prospective, observational study aimed at characterizing subclinical atherosclerosis in U.S. adults (55 to 80 years old) at risk for clinical atherosclerotic cardiovascular disease[8]. Between January 2008 and June 2009, the BioImage Study enrolled 7,687 asymptomatic men 55 to 80 years of age and women 60 to 80 years of age who were members of the Humana Health System and residents of the Chicago, Illinois, or Fort Lauderdale, Florida, metropolitan areas. A total of 6,397 individuals with exome chip genotypes and plasma lipids were analyzed.

Vanderbilt University electronic medical record-linked DNA repository (BioVU)

Vanderbilt University Medical Center Biorepository, BioVU, links DNA samples extracted from discarded blood samples from routine clinical testing at Vanderbilt University hospital to de-identified electronic medical records where individual level data can be extracted (e.g. cholesterol levels) and analyzed[9,10]. For the present study, we identified a total of 14156 individuals of European American ancestry with available information on LDL, HDL, total cholesterol, or triglyceride levels.

Bangladesh Risk of Acute Vascular Events study (BRAVE_Cases, BRAVE_Controls)

BRAVE is a retrospective case-control study of first-ever confirmed acute myocardial infarction (MI) in Bangladesh[11]. Patients (male or female; age between 30-80 years) admitted to the emergency rooms of the collaborating hospital in Dhaka, Bangladesh were eligible for inclusion as MI cases on the basis of symptoms, ECG and troponin-I. Controls were hospital based and frequency-matched to cases on age (within 5 year age bands) and sex, and without a self-reported history of cardiovascular disease. Commercial assay kits manufactured by Roche Diagnostics (GmbH, D-68298 Mannheim, Germany) were used to determine total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and triglycerides. All analyses were done on Roche automated clinical chemistry analysers, Hitachi 902, Hitachi Ltd, Tokyo, Japan.

British Genetics of Hypertension Study (BRIGHT)

Participants of the BRIGHT Study are recruited from the Medical Research Council General Practice Framework and other primary care practices in the UK[12]. Each case had a history of hypertension diagnosed prior to 60 years of age with confirmed blood pressure recordings corresponding to seated levels >150/100mmHg (1 reading) or mean of 3 readings >145/95 mmHg. BRIGHT is focused on recruitment of hypertensive individuals with BMI<30. Sample selection for Exome Chip study was based on DNA availability and quantity.

The Coronary Artery Risk Development in Young Adults study (CARDIA-white, CARDIA-black)

The Coronary Artery Risk Development in Young Adults study is a prospective multi-center investigation of the etiology and natural history of cardiovascular disease initiated in 1985-1986. The

study's initial enrollment consisted of 5115 European American and African American men and women between 18 and 30 years old (52% African American and 55% women) recruited from 4 field centers (Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA). The institutional review board at each of the study sites approved the study protocols, and written informed consent was obtained from all participants. Detailed information about the CARDIA study design and methods of data collection have been previously published[13]. Briefly, participants' age, race, and sex were self-reported during the recruitment phase and verified during the baseline clinic visit. Blood pressure was measured at the baseline examination on the right arm using a random-zero sphygmomanometer with the participant seated and following a 5 minute rest. Systolic and diastolic pressures were recorded as Phase I and Phase V Korotkoff sounds. Three measurements were taken at one minute intervals. The average of the second and third measurements was taken as the blood pressure value. The present analysis included 4,151 individuals (2,175 European Americans and 1,976 African Americans) with baseline BP measures and genotype data.

Copenhagen City Heart Study (CCHS)

CCHS is a population-based prospective study initiated in 1976 with follow-up examinations from 1981 to 1983, 1991 to 1994, and 2001 to 2003[14]. Participants were selected on the basis of the national Danish Civil Registration System to reflect the adult Danish population age 20 to ≥80 years. Non-fasting plasma levels of total cholesterol, high-density lipoprotein cholesterol (HDL-C), triglycerides and glucose were measured using colorimetric assays.

Copenhagen General Population Study (CGPS)

The CGPS is a population-based prospective study initiated in 2003 with ongoing enrollment[15,16]. Participants were selected on the basis of the national Danish Civil Registration System to reflect the adult Danish population age 20 to ≥80 years. Data were obtained from a questionnaire, a physical examination, and blood samples including deoxyribonucleic acid extraction. Non-fasting plasma levels of total cholesterol, high-density lipoprotein cholesterol (HDL-C), triglycerides and glucose were measured in fresh samples using colorimetric assays.

Cardiovascular Health Study (CHS-EA, CHS-AA)

The CHS is a population-based cohort study of risk factors for coronary heart disease and stroke in adults ≥65 years conducted across four field centers[17]. The original predominantly Caucasian cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists, and an additional 687 African-Americans were enrolled subsequently for a total sample of 5,888. DNA was extracted from blood samples drawn on all participants at their baseline examination in 1989-90. 750 African-American and 4,021 European- American individuals were genotyped using the Illumina HumanExome BeadChip array. CHS was approved by institutional review committees at each field center and individuals in the present analysis had available DNA and gave informed consent including consent to use of genetic information for the study of cardiovascular disease.

Copenhagen Ischemic Heart Disease Study (CIHDS)

This study comprised 5185 cases with myocardial infarction and other major acute coronary syndromes recruited from Copenhagen University Hospital during the period from 1991 to 2009[15,16]. In addition to a diagnosis of acute coronary syndrome, these cases also had stenosis or atherosclerosis on coronary angiography and/or positive results on exercise electrocardiography. Cases were classified by World Health Organization International Classification of Diseases-Eighth Revision, codes 410 to 414; International Classification of Diseases-Tenth Revision, codes I20 to I25, and through review of all hospital admissions and diagnoses entered in the national Danish Patient Registry and all causes of death entered in the national Danish Causes of Death Registry. Non-fasting plasma levels of total cholesterol, high-density lipoprotein cholesterol (HDL-C), triglycerides and glucose were measured in fresh samples using colorimetric assays.

CROATIA-Korcula (CROATIA-Korcula)

The CROATIA-Korcula study, Croatia, is a family-based, cross-sectional study in the isolated island of Korcula that included 965 examinees aged 18-95[18]. Blood samples were collected in 2007 along with many clinical and biochemical measures and lifestyle and health questionnaires. In the present study 855 genotyped individuals were included in the analysis.

DIABNORD (Exome Chip) (DIABNORD)

The DIABNORD Study is nested within the Västerbotten Health Survey, which is part of the Northern Sweden Health and Disease Study, a population-based prospective cohort study from northern Sweden[19]. Participants with incident type 2 diabetes were identified from the Diabetes Register in Northern Sweden (DiabNorth). A total of 909 Caucasian, non-diabetic participants from the DIABNORD Study had complete genotype and phenotype data necessary for the current analyses. Capillary blood was drawn following an overnight fast. Fasting serum lipid concentrations were measured with a Reflotron bench-top analyzer (Roche Diagnostics Scandinavia AB). Participants were genotyped with Illumina HumanExome Beadchip 12 v1.1. Ethical approval for the DIABNORD Study was obtained from the Regional Ethical Review Board in Umeå, Sweden.

Diabetes register in Vasa (DIREVA)

DIREVA (Diabetes register in Vasa) is a regional project in western Finland. All diabetes patients at all ages in the Vasa region are included. The aim of the registry is to describe the spectrum of diabetes subgroups in western Finland and to link genetic and phenotypic information at diagnosis of diabetes to outcome data and data on response to treatment. The study is coordinated by the Central hospital in Vasa.

The Finnish Diabetes Prevention Study (DPS)

DPS is a prospective randomized controlled trial aimed at preventing the progression from IGT to diabetes[20]. The original DPS was initiated in 1993. A total of 522 middle-aged, overweight subjects with IGT at baseline were randomized into either a lifestyle intervention or a standard-care control group. They were followed for occurrence of diabetes until the year 2000, when the first interim

analysis of the data was carried out as originally planned. At this point, the randomized trial was prematurely terminated due to markedly lower diabetes incidence rate in the lifestyle intervention group as compared to the control group. Since the termination of the randomized phase of the DPS, the original cohorts are no longer offered different treatments. However, all participants are monitored with yearly visits for long-term development of type 2 diabetes and complications.

The Dose Responses to Exercise Training Study (DR's EXTRA)

DR's EXTRA is a 4-year randomized controlled trial on the health effects of aerobic and resistance exercise training and a diet with low saturated fat, high unsaturated fat, and high fiber in a population sample of middle-aged and older men and women[21]. The target population was a representative sample of 3,000 individuals (1,500 men, 1,500 women) who lived in the city of Kuopio in Finland and who were 55-74 years of age in 2002, when they were randomly selected from the national population register. Of these individuals, 2,062 were willing to participate and 1,479 (72%) participated in the baseline examinations in 2005-2006. 1,410 individuals were randomly allocated into one of the six study groups, each of which included about 235 persons.

Duke Catheterization Genetics (Duke-AA-Cases, Duke-AA-Controls, Duke-EA-Cases, Duke-AA-Controls)

The Duke CATHGEN cohort consists of samples collected from individuals undergoing cardiac catheterization at Duke University Medical Center between 2001 and 2010[22]. These samples were matched with the findings of coronary anatomy, fasting chemistry data, as well as development of health habits and cardiovascular disease later in life through follow-up questionnaires. The data on the absence or presence of coronary artery disease were used in this analysis.

The Exeter Family Study of Childhood Health (EFSOCH)

The Exeter Family Study of Childhood Health is a prospective study, set up to test the fetal insulin hypothesis, and to identify genetic polymorphisms that play a role in determining birth weight and early postnatal growth[23]. We recruited 1017 families from a postcode-defined area in central Exeter. Specific inclusion criteria were established to obtain a homogeneous, non-diabetic, UK Caucasian cohort. Detailed anthropometric measurements were taken from both parents at 28 weeksÕ gestation, and from their children at birth, 12 weeks, 1 year and 2 years of age. Insulin and other biochemical analysis were measured in fasting parental samples and an umbilical cord blood sample taken at delivery. Parental and offspring DNA were extracted to allow molecular genetic analysis of candidate genes implicated in fetal growth.

Estonian Genome Center, University of Tartu (EGCUT)

The Estonian cohort is from the population-based biobank of the Estonian Genome Project of University of Tartu (EGCUT)[24]. The whole project is conducted according to the Estonian Gene Research Act and all participants have signed the broad informed consent. The current cohort size is over 51,515, from 18 years of age and up, which reflects closely the age distribution in the adult

Estonian population. Subjects are recruited by the general practitioners and physicians in the hospitals were randomly selected from individuals visiting general practitioners offices or hospitals. Each participant filled out a Computer Assisted Personal interview during 1-2 hours at a doctor's office, including personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, three generations), educational and occupational history and lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, women's health, quality of life). All diseases are defined according to the ICD10 coding. Lipids for current study where directly measured in clinical setting, both phenotype and genotype where available for 1,405 Estonian Biobank participants.

European Prospective Investigation into Cancer and Nutrition - Cardiovascular Disease Study (EPIC-CVD)

EPIC is a multi-centre prospective cohort study of 519,978 participants (366,521 women and 153,457 men, mostly aged 35–70 years) recruited between 1992 and 2000 in 23 centres located in 10 European countries[25]. Participants were invited mainly from population-based registers (Denmark, Germany, certain Italian centres, the Netherlands, Norway, Sweden, UK). Other sampling frameworks included: blood donors (Spain and Turin and Ragusa in Italy); screening clinic attendees (Florence in Italy and Utrecht in the Netherlands); people in health insurance programmes (France); and health conscious individuals (Oxford, UK). About 97% of the participants were of white European ancestry. EPIC-CVD employs a nested case-cohort design, analogous to the EPIC-InterAct study for type-2 diabetes, which established a common set of referents through selection of a random sample of the entire cohort ("subcohort"). Baseline measurements of all serum biomarkers were performed using a Roche MODULAR ANALYTICS EVO analyser by SHL groep in the Netherlands.

EPIC-InterAct (EPIC-InterAct T2D cases)

The InterAct study is a case-cohort study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts, and includes 12,403 incident cases of T2D and a subcohort of 16,154 individuals (including 778 randomly selected incident T2D cases)[26]. Up to 2,192 incident cases of T2D were included in the current analysis.

Family Heart Study (FamHS-EA)

The collection of phenotypes and covariates as well as clinical examination have been previously described (https://dsgweb.wustl.edu/fhscc/)[27]. In brief, the FamHS began in 1992 with the ascertainment of 1,200 families, half randomly sampled and half selected because of an excess of CHD or risk factor abnormalities as compared with age- and sex-specific population rates. The families, with approximately 6,000 European descent subjects, were sampled from four population-based parent studies: the Framingham Heart Study, the Utah Family Tree Study, and two centers for the ARIC study. Informed consent was obtained from all participants, and this project was approved by the Institutional Review Boards of all participating institutions. The participants attended a clinic visit between the years 1994-1996 and a broad range of phenotypes was assessed in the general domains of CHD, atherosclerosis, cardiac and vascular function, inflammation and hemostasis, lipids and lipoproteins, blood pressure, diabetes and insulin resistance, pulmonary function, diet, habitual physical activity, anthropometry, medical history and medication use. Approximately 8 years later,

2,756 subjects belonging to the 510 of the largest and most informative pedigrees were invited for a second clinical exam (2002-2004). The most important CHD risk factors were measured again. Medical history and medication use were updated. A total of 3,794 subjects from the first clinical visit participated in the current study. The subjects were genotyped using the Illumina Infinium HumanExome v1.0 BeadChip.

The Fenland Study (Fenland)

The Fenland Study is an ongoing, population-based cohort study (started in 2005) designed to investigate the association between genetic and lifestyle environmental factors and the risk of obesity, insulin sensitivity, hyperglycemia and related metabolic traits in men and women aged 30 to 55 years[28]. Potential volunteers were recruited from General Practice sampling frames in the Fenland, Ely and Cambridge areas of the Cambridgeshire Primary Care Trust in the UK. Exclusion criteria for the study were: prevalent diabetes, pregnant and lactating women, inability to participate due to terminal illness, psychotic illness, or inability to walk unaided. All participants had measurements done at the MRC Epidemiology Unit Clinical Research Facilities in Ely, Wisbech and Cambridge. Participants attended after an overnight fast for a detailed clinical examination, and blood samples were collected. The Local Research Ethics Committee granted ethical approval for the study and all participants gave written informed consent.

Framingham Heart Study (FHS)

The FHS is a three generational prospective cohort that has been described in detail previously[29]. Individuals were initially recruited in 1948 in Framingham, USA to evaluate cardiovascular disease risk factors. The second generation cohort (5,124 offspring of the original cohort) was recruited between 1971 and 1975. The third generation cohort (4,095 grandchildren of the original cohort) was collected between 2002 and 2005. Fasting lipid levels were measured at exam 1 of the Offspring (1971-1975) and third generation (2002-2005) cohorts, using standard LRC protocols. 8,153 European-American individuals were genotyped using the Illumina HumanExome BeadChip array.

FIN-D2D 2007 (FIN-D2D 2007)

The purpose of the study is to gather information about prevalence of diabetes and cardiovascular diseases and of the risk factors associated with these within the Finnish population[30]. The survey assists in the evaluation of the effects of the national type 2 diabetes prevention plan. The study sample consists of 4,500 people randomly selected from the Finnish population register between the ages of 45 and 74 years and living in one of the three hospital districts chosen for the study: South Ostrobothnia, Central Finland, and Pirkanmaa.

National FINRISK 2007 Study (FINRISK 2007 T2D cases, FINRISK 207 T2D controls)

The Finnish National Public Health Institute performed the FINRISK health study in five areas of Finland during spring 2007 to investigate the people's health behavior and the risk factors of chronic diseases and public health problems[31]. The survey is a continuation of a series of studies begun in

Eastern Finland in 1972 and performed once every five years since then. The purpose of the study is to gather information about the protective and risk factors of the major Finnish public health problems, such as cardiovascular diseases, diseases of the brain and the central nervous system, cancers, diabetes, asthma and allergies, and the prevalence of those factors in the population. The survey also monitors the state of health of the Finnish population.

Finland-United States Investigation of NIDDM Genetics Study (FUSION T2D cases, FUSION T2D controls)

FUSION1 cases included FUSION samples each reporting at least one T2D sibling and Finrisk 2002 T2D cases from a Finnish population-based risk factor survey[32,33]. Controls included 219 subjects from Vantaa, Finland who were NGT at ages 65 and 70 years, NGT spouses of FUSION subjects, and Finrisk 2002 NGT subjects. FUSION1 controls were approximately frequency-matched to the cases by five-year age category, sex, and birth province. FUSION2 includes subjects chosen from the following studies: Dehko 2D (D2D) 2004: a population-based study to screen individuals regarding T2D risk and to prevent T2D development; Finrisk 1987: an early round of the 5-yearly Finrisk national population-based health surveys; Finrisk 2002: a population-based survey of non-communicable diseases in >13,000 individuals aged 25-74 years living in 80 communities of Finland; Action LADA: a study of latent autoimmune diabetes in adults (LADA). Action LADA investigators screened individuals aged 30-69 years with recently-diagnosed diabetes and identified 373 T2D cases who agreed to participate in FUSION; Health 2000: a population-based study of people aged ≥30 years from throughout Finland; Savitaipale Diabetes Study: a study of diabetes in the town of Savitaipale in eastern Finland.

Gene x Lifestyle Interactions and Complex Traits Involved in Elevated Disease Risk (Exome Chip) (GLACIER)

The Gene-Lifestyle interactions And Complex traits Involved in Elevated disease Risk (GLACIER) Study is nested within the Västerbotten Health Survey, which is part of the Northern Sweden Health and Disease Study, a population-based prospective cohort study from northern Sweden[19]. A total of 921 Caucasian, non-diabetic participants from the GLACIER Study had complete genotype and phenotype data necessary for the current analyses. Capillary blood was drawn following an overnight fast. Fasting serum lipid concentrations were measured with a Reflotron bench-top analyzer (Roche Diagnostics Scandinavia AB). Participants were genotyped with Illumina HumanExome Beadchip 12 v1.1. Ethical approval for the GLACIER Study was obtained from the Regional Ethical Review Board in Umeå, Sweden.

Genetics of Diabetes Audit and Reasearch in Tayside Scotland study (GoDARTS_CAD)

A high quality resource, initially funded by the Wellcome Trust and supported by Diabetes UK, has been created with successful recruitment of consented patients with type 2 diabetes and matching controls (non diabetics) throughout Tayside, Scotland[34]. This resource is already available to researchers worldwide and is helping to define genetic factors related to diabetes including susceptibility, complications and response to treatment. This analysis used the coronary artery disease subset of the cohort. First-ever CAD event. Defined as fatal and non-fatal myocardial

infarction, unstable angina or coronary revascularisation. Controls were free of coronary artery disease, stroke and peripheral vascular disease.

Genetics of Diabetes Audit and Research Tayside (GoDARTS-cases, GoDARTS-controls)

A high quality resource, initially funded by the Wellcome Trust and supported by Diabetes UK, has been created with successful recruitment of consented patients with type 2 diabetes and matching controls (non diabetics) throughout Tayside, Scotland[34].

Genetic regulation of arterial pressure in humans in the community (GRAPHIC)

The Genetic Regulation of Arterial Pressure in Humans in the Community (GRAPHIC) Study is a family based population study comprising of 510 nuclear families (two parents aged 40-60 years at recruitment and two adult offspring aged 18-40 years) recruited through primary care in Leicestershire between 2003 and 2005[35]. All subjects are of white European origin. The primary objective of the GRAPHIC study was to investigate the genetic basis of blood pressure variation and all subjects underwent 24-hour ambulatory BP measurements. In addition, all subjects had extensive phenotyping including a full medical history and recording of risk factors and medication, dietary history, physical activity assessment, clinic BP, 12 lead ECG, measurement of height, weight, WHR and skinfold thickness. Available laboratory data include serum and urine electrolytes, plasma lipids measured by NMR and CRP. A total of 1851 subjects were included in this analysis.

Generation Scotland_Scottish Family Health Study (GS-SFHS)

Generation Scotland-SFHS recruited almost 24,000 participants from throughout Scotland. The study enlisted individuals aged 18-65 and their family members. Volunteers were asked to provide information about their lifestyle and diet, their medical history, and samples of blood and urine. Participation was by invitation through local GPs, or families volunteering directly. As the name suggests, the SFHS is based on families so at least one brother or sister of the initial recruit was required and preferably other family members. In the present study 9946 genotyped individuals were included in analysis.

Health2006/Health2008 (Health)

The Health2006 and Health2008 studies are cohort studies of adults aged 18-69 years who live in the greater Copenhagen area[36]. The aim of the studies was to identify lifestyle related risk factors for chronic diseases such as diabetes, heart disease, asthma, musculoskeletal disorders, chronic lung disease and mental disorders. Potential participants were excluded if they emigrated from the study location. Baseline examinations were conducted between 2006 and 2008. Data is collected through two questionnaires pertaining to lifestyle factors and mental health, and through medical exams assessing lung and cardiopulmonary function, and muscle strength. Blood samples were also collected from each participant for genetic and/or biomarker studies. The studies were approved by the Ethical Committee of Copenhagen County and the Danish Data Protection Agency. Here 3616 participants with information on lipids and exome chip were analysed.

Hellenic Isolated Cohorts - MANOLIS cohort (HELIC MANOLIS)

The HELIC (Hellenic Isolated Cohorts; www.helic.org) MANOLIS (Minoan Isolates) collection focuses on the Mylopotamos villages. Recruitment of this population-based sample was primarily carried out at the village medical centres. All individuals were older than 17 years and had to have at least one parent from the Mylopotamos area. The study includes biological sample collection for DNA extraction and lab-based blood measurements, and interview-based questionnaire filling. The phenotypes collected include anthropometric and biometric measurements, clinical evaluation data, biochemical and haematological profiles, self-reported medical history, demographic, socioeconomic and lifestyle information. The study was approved by the Harokopio University Bioethics Committee and informed consent was obtained from every participant. The total sample size in the collection is approximately 1,500 and 825 individuals genotyped on the exome chip and with lipid data available were included in this analysis.

Hellenic Isolated Cohorts - Pomak cohort (HELIC Pomak)

The HELIC (Hellenic Isolated Cohorts; www.helic.org) Pomak collection focuses on the Pomak villages, a set of isolated mountainous villages in the North of Greece. Recruitment of this population-based sample was primarily carried out at the village medical centres. The study includes biological sample collection for DNA extraction and lab-based blood measurements, and interview-based questionnaire filling. The phenotypes collected include anthropometric and biometric measurements, clinical evaluation data, biochemical and haematological profiles, self-reported medical history, demographic, socioeconomic and lifestyle information. The study was approved by the Harokopio University Bioethics Committee and informed consent was obtained from every participant. The total sample size in the collection is approximately 1,700 and 971 individuals genotyped on the exome chip and with lipid data available were included in this analysis.

The Nord-Trondelag Health Study (HUNT-Case)

We included 5,440 individuals with at least one lipid measurement from the second survey of the Nord-Trondelag Health Study (HUNT): 2,662 cases with hospital diagnosed myocardial infarction (primary phenotype) and 2,778 healthy controls without cardiovascular disease matched on sex, birth year (+/- 1 year), and municipality or geographical region to minimize population stratification[37]. HUNT is a population based health study (www.ntnu.edu/hunt) with personal and family medical histories on approximately 120,000 individuals from Nord-Trondelag County, Norway, collected in three surveys (HUNT 1, 2, and 3). Self-reported questionnaires, clinical examination, and non-fasting venous blood samples were collected on 62,816 individuals (66.9% of invited).

The Nord-Trondelag Health Study (HUNT-Control)

We included 5,440 individuals with at least one lipid measurement from the second survey of the Nord-Trondelag Health Study (HUNT): 2,662 cases with hospital diagnosed myocardial infarction (primary phenotype) and 2,778 healthy controls without cardiovascular disease matched on sex, birth year (+/- 1 year), and municipality or geographical region to minimize population stratification[37].

HUNT is a population based health study with personal and family medical histories on approximately 120,000 individuals from Nord-Trondelag County, Norway, collected in three surveys (HUNT 1, 2, and 3)19,35. HUNT 2 was conducted in 1995-97, inviting all residents ³ 20 years of age in Nord-Tr¿ndelag County, Norway. Self-reported questionnaires, clinical examination, and non-fasting venous blood samples were collected on 62,816 individuals (66.9% of invited).

Inter99 (Inter99)

The Inter99 study carried out in 1999-2001 included invitation of 12934 persons aged 30-60 years drawn from an age- and sex-stratified random sample of the population[38]. The baseline participation rate was 52.5%, and the study included 6784 persons. The Inter99 study was a population-based randomized controlled trial (CT00289237, ClinicalTrials.gov) and investigated the effects of lifestyle intervention on CVD. Here 5827 participants with information on lipids and exome chip were analysed.

The Mount Sinai BioMe Biobank (IPM BioMe-African, IPM BioMe-European, IPM BioMe-Hispanic)

The BioMe Biobank is an ongoing, prospective, hospital- and outpatient- based population research program operated by The Charles Bronfman Institute for Personalized Medicine (IPM) at Mount Sinai. BioMe has enrolled over 33,000 participants between September 2007 and December 2015. BioMe is an Electronic Medical Record (EMR)-linked biobank that integrates research data and clinical care information for consented patients at The Mount Sinai Medical Center, which serves diverse local communities of upper Manhattan with broad health disparities. IPM BioMe populations include 25% of African American ancestry (AA), 36% of Hispanic Latino ancestry (HL), 30% of white European ancestry (EA), and 9% of other ancestry. The IPM BioMe disease burden is reflective of health disparities in the local communities. BioMe operations are fully integrated in clinical care processes, including direct recruitment from clinical sites waiting areas and phlebotomy stations by dedicated BioMe recruiters independent of clinical care providers, prior to or following a clinician standard of care visit. Recruitment currently occurs at a broad spectrum of over 30 clinical care sites. Information on anthropometrics, demographics, lipid levels and use of lipid-lowering medication was derived from participants' EMR.

Jackson Heart Study (JHS)

The JHS is a large, population-based observational study evaluating the etiology of cardiovascular, renal, and respiratory diseases among African Americans residing in the three counties (Hinds, Madison, and Rankin) that make up the Jackson, Mississippi metropolitan area[39]. Data and biologic materials have been collected from 5301 participants, including a nested family cohort of 1,498 members of 264 families. The age at enrollment for the unrelated cohort was 35-84 years; the family cohort included related individuals >21 years old. Participants provided extensive medical and social history, had an array of physical and biochemical measurements and diagnostic procedures, and provided genomic DNA during a baseline examination (2000-2004) and two follow-up examinations (2005-2008 and 2009-2012). The study population is characterized by a high prevalence of diabetes, hypertension, obesity, and related disorders. Annual follow-up interviews and cohort surveillance are

ongoing. 2,154 African-American individuals were genotyped using the Illumina HumanExome BeadChip array. Individuals that overlapped with ARIC were randomly split between the two cohorts, except individuals in a known JHS family were kept in JHS.

Kooperative Gesundheitsforschung in der Region Augsburg (KORA)

The KORA Study is a series of population-based epidemiological surveys of persons living in or near the city of Augsburg, Germany[40,41]. All survey participants are residents of German nationality identified through the registration office and between 25 and 75 years old at the time of enrollment. Survey S4 was conducted between 1999 and 2001. KORA F4 is a 7-year follow up of S4. Illumina Exome Chip data is available for participants from KORA F4. Cryptically related persons have been removed as well as population outliers and non-fasting samples. The final study sample consists of 2723 persons.

Lothian Birth Cohort 1921 (LBC 1921)

The LBC1921 cohort consists of 550 relatively healthy individuals, 316 females and 234 males, assessed on cognitive and medical traits at about 79 years of age[42,43]. They were all born in 1921 and most took part in the Scottish Mental Survey of 1932. When tested, the sample had a mean age of 79.1 years (SD = 0.6). They were all Caucasian, community-dwelling, and almost all lived in the Lothian region (Edinburgh city and surrounding area) of Scotland. Genotyping was performed at the Wellcome Trust Clinical Research Facility, Edinburgh using the Illumina HumanExome BeadChip.

Lothian Birth Cohort 1936 (LBC1936)

The LBC1936 consists of 1091 relatively healthy individuals assessed on cognitive and medical traits at about 70 years of age[43,44]. They were all born in 1936 and most took part in the Scottish Mental Survey of 1947. At baseline the sample of 548 men and 543 women had a mean age 69.6 years (SD = 0.8). They were all Caucasian, community-dwelling, and almost all lived in the Lothian region (Edinburgh city and surrounding area) of Scotland. Genotyping was performed at the Wellcome Trust Clinical Research Facility, Edinburgh using the Illumina HumanExome BeadChip.

The London Life Sciences Population Study (LOLIPOP-ExomeChip)

LOLIPOP is an ongoing population based cohort study of 17,606 Indian Asian and 7,766 European men and women aged 35-75 years, recruited from the lists of 58 General Practitioners in West London, United Kingdom[45]. Participants classified as having Indian Asian ancestry reported having all four grandparents born on the Indian subcontinent. Biochemical analysis included total and HDL cholesterol and triglycerides, using standard commercial assays. Aliquots of whole blood, plasma and serum are frozen at –80°C. In the present study, 970 and 1,664 Indian Asian samples were genotyped on the Illumina ExomChip array and the Illumina OmniExpressExome array, respectively.

The London Life Sciences Population Study (LOLIPOP-OmniExpressExome)

LOLIPOP is an ongoing population based cohort study of 17,606 Indian Asian and 7,766 European men and women aged 35-75 years, recruited from the lists of 58 General Practitioners in West London, United Kingdom[45]. Participants classified as having Indian Asian ancestry reported having all four grandparents born on the Indian subcontinent. Biochemical analysis included total and HDL cholesterol and triglycerides, using standard commercial assays. Aliquots of whole blood, plasma and serum are frozen at –80°C. In the present study, 970 and 1,664 Indian Asian samples were genotyped on the Illumina ExomChip array and the Illumina OmniExpressExome array, respectively.

Malmo Diet and Cancer Study (MDC)

The Malmö Diet Cancer Study is an ongoing longitudinal prospective cohort study of the middle-aged population of Malmö designed to screen for dietary habits and genetic markers in order to predict incident cancers in the general population and to screen for cardiovascular risk factors and early atherosclerosis in a sub-sample[46]. 28,000 subjects living in Malmö were during 1992-1996 invited by letter to a clinical examination, food-frequency questionnaire and blood sampling. Individuals were between ages 45-70 at first screening. Self-completed questionnaires, clinical examinations, and extracts from registers are available for the years 1992-2009. A subgroup of the patients totaling less than 5,000 individuals with complete phenotype data was analyzed as one cohort in this study of plasma lipid levels.

Multi-Ethnic Study of Atherosclerosis (MESA-EA, MESA-AA, MESA-Chinese, MESA-Hispanic)

The Multi-Ethnic Study of Atherosclerosis (MESA) is a study of the characteristics of subclinical cardiovascular disease (disease detected non-invasively before it has produced clinical signs and symptoms) and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease[47]. MESA researchers study a diverse, population-based sample of 6,814 asymptomatic men and women aged 45-84. Thirty-eight percent of the recruited participants are White, 28 percent African-American, 22 percent Hispanic, and 12 percent Asian, predominantly of Chinese descent. 2,128 additional individuals from 594 families were recruited through MESA Family by utilizing the existing MESA framework, yielding 3,026 sibpairs divided between African Americans and Hispanic-Americans. Participants were recruited from six field centers across the United States: Wake Forest University, Columbia University, Johns Hopkins University, University of Minnesota, Northwestern University and University of California - Los Angeles. For the present analyses, we included 2490 White, 2496 African American, 2081 Hispanic and 701 Chinese participants.

Metabolic Syndrome in Men Study (METSIM)

The METSIM study aims to investigate the metabolic syndrome, type 2 diabetes, cardiovascular disease, and cardiovascular risk factors[48]. It is an ongoing study of men aged 50 to 70 years, randomly selected from the population registry of the town of Kuopio, in Eastern Finland.

Montreal Heart Institute Biobank (MHI Biobank)

The Montreal Heart Institute (MHI) Biobank is a longitudinal hospital cohort, which was initiated in 2005 with the aim to recruit 30,000 patients of the MHI for clinical and genetic research[49]. Participants are recruited from different departments within the MHI and its affiliated prevention centre. The MHI Biobank collects data by using a detailed questionnaire administered by a research nurse at baseline including demographics, personal and family medical history, diet, tobacco, medication use, as well as depression and hostility questionnaires. Blood, DNA, and plasma are collected at baseline and stored at the Beaulieu-Saucier Pharmacogenomics Centre. The patients health information is confirmed and complemented by the research nurse from the hospital's health record. The cohort's database is updated daily with patient's medical information from the hospital's electronic records including laboratory and lipid measurements and a follow-up study questionnaire is administered every four years. The questionnaire is updated every 4 years. The cohort is comprised of over 20,000 participants as of January 2016 with a median follow up period of 4.2 years. Genotyping was performed on the first 11,556 participants using the Illumina HumanExome chip v1.1. Lipid measurements for the current project were obtained from the hospital records, by selecting the records with sampling dates that were the closest to that of the baseline questionnaire. Following clinical and genetic data cleanup procedures, 6421 patients with both lipid and genotype data were available for analysis.

MOnica Risk, Genetics, Archiving and Monograph project (MORGAM)

MORGAM is a consortium of prospective cohort studies from around Europe[50]. For this project, participants were included from the ATBC study (Finland), Augsburg-KORA (Germany), Brianza (Italy) and PRIME cohorts Belfast (UK), Lille, Strasbourg and Toulouse (all France). A case-cohort design was used comparing incident coronary disease cases with participants randomly selected from within each study. Lipids were measured at baseline using standard approaches.

Northern Finland Birth Cohort 1986 (NFBC1986)

The NFBC1986 study includes 9432 live-born individuals with expected dates of birth between July 1st 1985 and June 30th 1986 in the provinces of Oulu and Lapland, in Finland. The University of Oulu Ethics Committee and the Ethical Committee of Northern Ostrobothnia Hospital District have approved the study. Cohort has been followed up since early pregnancy until adolescence. Growth measurements were obtained from communal child health clinics. Samples were stored at -80 ¼C until analyzed and DNA extracted. Fasting serum total cholesterol and triglycerides were determined using a Hitachi 911 automatic analyser and commercial reagents (Roche, Mannheim, Germany). HDL- and LDL-C were also determined using the same analyzer and methods previously described (Sugiuchi J et al, 1995, ClinChem; Wieland H et al, 1983, J Lipid Res). The intra- and interassay coefficients of variation were 0.7 and 1.5% for total cholesterol, 0.5 and 3.2% for HDL-C, 1.6 and 2.6% for LDL-C, 0.9 and 2.4% for triglycerides.

Oxford BioBank (OBB)

The Oxford Biobank is a collection of 30-50 year old healthy men and women living in Oxfordshire. All participants have undergone a detailed examination at a screening visit, donated DNA and given informed consent to be re-approached. The Oxford Biobank is a resource for medical research to

translate early discoveries to the benefit of patients in the future. In the present study, a total of 4442 individuals were inlcuded in the analysis (http://www.oxfordbiobank.org.uk/).

Ottawa Heart Study (Ottawa-Cases, Ottawa-Controls)

The Ottawa Heart Study is an ongoing, hospital-based study of coronary heart disease at the Ottawa Heart Institute in Ottawa, Canada. All patients at the Institute who undergo coronary artery bypass grafting, coronary artery angiography, or care for acute myocardial infarction are invited to participate in the study. Healthy elderly controls (men > 65y, women > 70y) were recruited via an extensive newspaper and television advertising campaign in the Ottawa community. Controls were carefully interviewed by a physician or nurse to ascertain that they were free of symptoms of possible ischemic arterial disease and had no past history of cardiovascular symptoms, a positive stress test, coronary angiography demonstrating stenosis (>50%) in any artery or clinical cardiovascular events. Individuals with the same ethnic background as the cases (Caucasian) were included in this study (total sample of 1100 cases and 2361 controls). In the present study, a total of 951 cases and 2103 controls were included into analysis.

Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) / Uppsala Longitudinal Study of Adult Men (ULSAM) (PIVUS-ULSAM)

The PIVUS study started in 2001 with the primary aim to investigate the predictive power of different measurements of endothelial function and arterial compliance in a random sample of 1000 subjects aged 70 living in the community of Uppsala[51]. ULSAM is a unique, ongoing, longitudinal, epidemiologic study based on all available men, born between 1920 and 1924, in Uppsala County, Sweden[52]. The men were investigated at the ages of 50, 60, 70, 77, 82 and 88 years. Individuals investigated at age 70 were included in the current analysis. In the present study a total of 2006 samples (944 PIVUS / 1062 ULSAM) were included into analysis.

Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study

The Prevalence, Prediction and Prevention of diabetes (PPP)–Botnia Study is a population-based study in Western Finland carried out from 2004 to 2008 to obtain accurate estimates of prevalence and risk factors for type 2 diabetes, impaired glucose tolerance, impaired fasting glucose and the metabolic syndrome in the adult population and to use this information for prediction and prevention of the disease[53]. The participants were randomly recruited from the national Finnish Population Registry to represent 6 to 7% of the population in the 18- to 75-year age range. Altogether 5,208 individuals participated in the study (54.7% of those invited).

PROCARDIS (Procardis Cases, Procardis Controls)

The PROCARDIS Study is a an ongoing study of coronary heart disease at multiple centers in Europe (the universities of Oxford and Münster; the Karolinska Institute; the Mario Negri Institute; Digilab BioVisioN GmbH; Centre National de Genotypage; Institut de Recerca del Hospital de la Santa Creu I Sant Pau; Università degli Studi di Milano; Clinical Gene Networks AB; CF consulting

S.r.l.; Metabometrix Limited) and AstraZeneca[54]. Families with members with patients with myocardial infarction or symptomatic acute coronary syndrome occurring before the age of 65 as evidenced by typical clinical symptoms, EKG findings, and biomarker elevation were selected. The study aims to identify new susceptibility genes of coronary artery disease through a genome-wide screen.

Pakistan Risk of Myocardial Infarction Study (PROMIS_Cases, PROMIS_Controls)

PROMIS is a retrospective case-control study of first-ever confirmed acute MI in Pakistan[55]. Patients aged 30-80 years who were admitted to the emergency rooms of nine recruitment centres across Pakistan were eligible for inclusion as cases on the basis of symptoms, ECG and troponin levels. Controls were hospital based and frequency-matched to cases on age (within 5 year age bands) and sex, and without a self-reported history of cardiovascular disease. Nonfasting blood samples were drawn from each participant and centrifuged within 45 minutes of venepuncture. Serum samples were stored at -80°C. Total cholesterol, HDL-C, and triglyceride concentrations were measured using enzymatic methods (Roche Diagnostics, USA) at the Center for Non-Communicable Diseases, Pakistan.

Prospective Study of Pravastatin in the Elderly at Risk clinical trial (PROSPER)

PROSPER was a controlled, randomised study involving 2,804 men and 3,000 women aged 70-82, with a history of, or risk factors for cardiovascular disease[56]. Participants were randomised to either 40mg pravastatin per day or matching placebo. A nested case-control design was used for this study, selecting as cases individuals who self-reported a history of coronary disease at baseline or who had a coronary event during follow-up. Controls were participants who were free of cardiovascular disease at baseline and at the end of follow-up, frequency matched to the cases for sex and age (in 5-year bands). Baseline lipid levels were measured using standard assays at the Department of Pathological Biochemistry at the Glasgow Royal Infirmary.

Peking University Health Science Center and the University of Michigan Medical School study of Myocardial Infarction (PUUMA.Capital)

Samples from China were collected by the Joint Institute of the Peking University Health Science Center and the University of Michigan Medical School study of Myocardial Infarction (PUUMA-MI)[57]. PUUMA-MI is a large-scale project designed to study cardiovascular disease and related traits including myocardial infarction (MI) and plasma lipid levels. Fasting plasma lipid levels (including serum total cholesterols, LDL cholesterol, HDL cholesterol and triglycerides) were tested using Roche cobas 8000 modular analyzer series (Indianapolis , IN, USA) in Beijing Shijingshan cohort samples (N=836) and Beckman coulter UniCel DxC 800 Synchron (Brea, CA, USA) in Peking University First Hospital-based samples (N=7,339) after overnight fasting, respectively.

Peking University Health Science Center and the University of Michigan Medical School study of Myocardial Infarction (PUUMA.Case, PUUMA.Control)

Samples from China were collected by the Joint Institute of the Peking University Health Science Center and the University of Michigan Medical School study of Myocardial Infarction (PUUMA-MI)[57]. PUUMA-MI is a large-scale project designed to study cardiovascular disease and related traits including myocardial infarction (MI) and plasma lipid levels. Blood samples were taken in the morning after an overnight fast and collected into vacuum tubes containing EDTA for the measurement of plasma lipids. Clinical chemical analyses were conducted at the central chemistry lab of Peking University Third Hospital. Using Beckman Coulter AU 5800 Auto-Analyzer (Tokyo, Japan), total cholesterol was measured by an enzymatic method (Baiding Biological Engineering Ltd., Beijing, China); triglycerides were measured by an enzymatic (with peroxidase) method (Biosino Bio-Technology Co., Ltd., Beijing, China); and high density lipoprotein cholesterol and low density lipoprotein cholesterol were measured by a liquid selective detergent method (Sekisui Medical Co., Ltd., Tokyo, Japan). The day-to-day coefficients of variation were 0.9%-2.0% for total cholesterol, 1.6% for high density lipoprotein-cholesterol, 1.5% for low density lipoprotein-cholesterol and 0.8% - 2.1% for triglyceride.

SardiNIA study on aging (SardiNIA)

The SardiNIA study is a longitudinal, population-based study that includes 6,921 individuals, representing >60% of the adult population of 4 villages in the Lanusei valley on Sardinia (Italy)[58]. These individuals are clustered in 1,257 multigenerational families, up to 5 generations deep, and have been characterized for hundreds of quantitative traits. All participants gave informed consent to study protocols, which were approved by the Sardinian local research ethic committees: Comitato Etico di Azienda Sanitaria Locale 8, Lanusei (2009/0016600) and Comitato Etico di Azienda Sanitaria Locale 1, Sassari (2171/CE)) and by the NIH Office of Human Subject Research as governed by Italian institutional review board approval.

Steno Diabetes Center (SDC)

Patients with T2D (above 18 years of age) were recruited from the outpatient clinic at Steno Diabetes Center, Gentofte, Denmark[59]. Anthropometrics and basal fasting biochemistry (e.g. plasma glucose, serum insulin, HbA1C and lipids) have been measured. Here 499 T2D cases with information on lipids and exome chip were analysed.

Twins UK (TwinsUK)

The TwinsUK cohort is an adult twin British registry recruited from the general population in the United Kingdom[60]. In the present study, a total of 923 individuals were included into analysis.

Vejle Biobank - T2D cases and Controls (VejleCases, VejleControls)

Vejle Biobank is a sample of clinical-onset T2D patients, and non-diabetic control individuals with matched age and gender distribution, examined at Vejle Hospital during a three year period[59]. Control individuals were non-diabetic by self-report and according to fasting plasma glucose levels. The main objectives of the study were to investigate the development of late diabetic complications and lack of

treatment effect. Anthropometrics, including body fat percentage, basal fasting biochemistry (e.g. plasma glucose, serum insulin, HbA1C and lipids), detailed biochemistry (e.g. measures of kidney function and serum CRP) have been assessed and questionnaires regarding lifestyle, health, diabetic complications and use of anti-diabetic medication filled out. Here 1879 T2D cases and 424 T2D controls with information on lipids and exome chip were analysed.

Women's Genome Health Study (WGHS)

The Women's Genome Health Study (WGHS) is a prospective cohort of initially healthy, female North American health care professionals at least 45 years old at baseline representing participants in the Women's Health Study (WHS) who provided a blood sample at baseline and consent for blood-based analyses[61]. The WHS was a 2x2 trial beginning in 1992-1994 of vitamin E and low dose aspirin in prevention of cancer and cardiovascular disease with about 10 years of follow-up. Since the end of the trial, follow-up has continued in observational mode.

Women's Health Initiative (WHI-EA, WHI-AA)

The WHI is one of the largest (n = 161,808) US studies of women's health. This project was approved by the ethics committee at the Fred Hutchinson Cancer Research Center. The WHI consists of 2 main components: (i) a clinical trial that enrolled 68,132 post-menopausal women aged 50–79 years and randomized them to 1 of 3 placebo-controlled clinical trials of hormone therapy, dietary modification or supplementation with calcium and vitamin D and (ii) an observational study that enrolled 93,676 women of the same age range in a parallel prospective study[62].

West of Scotland Coronary Prevention Study (WOSCOPS)

WOSCOPS was a controlled, randomised study involving 6,595 men aged 45-64, with elevated LDL cholesterol but no history of myocardial infarction[63]. Participants were randomised to either 40mg pravastatin per day or matching placebo. A nested case-control design was used for this study, selecting as cases individuals who had a coronary event during follow-up. Controls were participants who were free of cardiovascular disease at baseline and at the end of follow-up, frequency matched to the cases on age (in 5-year bands). Baseline lipid levels were measured using enzymatic cholesterol and triglyceride assays at the Department of Pathological Biochemistry at the Glasgow Royal Infirmary.

II. Disclosures

L. Adrienne Cupples reports a relationship with Boston Heart Diagnostics. Funding for exome chip genotyping in WGHS was provided by a grant from Amgen to Daniel Chasman. Marit E. Jørgensen holds shares in Novo Nordisk A/S, is PI of an investigator initiated trial sponsored by Astra Zeneca, and has received research grants from Astra Zeneca. Bruce M. Psaty serves on the DSMB for a clinical trial funded by the manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. Peter Sever has received research awards from Pfizer Inc. Neil Poulter has received financial support from several pharmaceutical companies which manufacture either blood pressure lowering or lipid lowering agents, or both, and consultancy fees. Naveed Sattar consulted for Amgen and Sanofi related to PCSK9 inhibitors; and was an investigator on clinical trials of PCSK9 inhibition funded by Amgen

III. Acknowledgments

We acknowledge the funding support for the following contributing studies to the meta-analysis

1958BC: We are grateful for being able to use the British 1958 Birth Cohort DNA collection. Sample collection funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping was funded by the Wellcome Trust. Analysis was supported by BHF grant (Deloukas) RG/14/5/30893. Panos Deloukas's work forms part of the research themes contributing to the translational research portfolio of Barts Cardiovascular Biomedical Research Unit, which is supported and funded by the National Institute for Health Research.

ADDITION: The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen partially funded by an unrestricted donation from the Novo Nordisk Foundation (www.metabol.ku.dk).

AGES: This study has been funded by NIA contract N01-AG-12100 with contributions from NEI, NIDCD and NHLBI, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

AMCPAS: We would like to acknowledge Dr. Mieke Trip for setting up the AMCPAS.

ARIC: The Atherosclerosis Risk in Communities (ARIC) study is carried out as a collaborative study supported by the National Heart, Lung, and Blood Institute (NHLBI) contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C). The authors thank the staff and participants of the ARIC study for their important contributions. Funding support for "Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium" was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419).

BioVU: The data sets used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU, which is supported by institutional funding and by the Vanderbilt CTSA grant UL1 TR000445 from NCATS/NIH. This work was also in part supported by National Institutes of Health grants (U19 HL65962, R01 HL092217, and P50 GM115305).

BRAVE: The BRAVE study genetic epidemiology working group is a collaboration between the Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, UK, the Centre for Control of Chronic Diseases, icddr,b, Dhaka, Bangladesh and the National Institute of Cardiovascular Diseases, Dhaka, Bangladesh. The CHD Exome+ coordinating centre receives core support from the UK Medical Research Council (G0800270), the BHF (SP/09/002), the NIHR and Cambridge Biomedical Research Centre, as well as grants from the

European Research Council (268834) and the European Commission Framework Programme 7 (HEALTH-F2-2012-279233). Exome+ assays were supported by grants from Merck and Pfizer.

BRIGHT: The BRIGHT study is extremely grateful to all the patients who participated in the study and the BRIGHT nursing team. This work forms part of the research programme of the NIHR Cardiovascular Biomedical Research Unit at Barts and The London, Queen Mary University of London, UK.

CARDIA: The CARDIA Study is conducted and supported by the National Heart, Lung, and Blood Institute in collaboration with the University of Alabama at Birmingham (HHSN268201300025C & HHSN268201300026C), Northwestern University (HHSN268201300027C), University of Minnesota (HHSN268201300028C), Kaiser Foundation Research Institute (HHSN268201300029C), and Johns Hopkins University School of Medicine (HHSN268200900041C). CARDIA is also partially supported by the Intramural Research Program of the National Institute on Aging. Exome Chip genotyping was supported from grants R01-HL093029 and U01- HG004729 to M. Fornage. This manuscript has been reviewed and approved by CARDIA for scientific content.

CCHS/CGPS/CIHDS: The CHD Exome+ coordinating centre receives core support from the UK Medical Research Council (G0800270), the BHF (SP/09/002), the NIHR and Cambridge Biomedical Research Centre, as well as grants from the European Research Council (268834) and the European Commission Framework Programme 7 (HEALTH-F2-2012-279233). Exome+ assays were supported by grants from Merck and Pfizer.

CHS: This Cardiovascular Health Study (CHS) research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL068986 , and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CROATIA_Korcula: The CROATIA-Korcula study was funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). We would like to acknowledge the invaluable contributions of the recruitment team in Korcula, the administrative teams in Croatia and Edinburgh and the people of

Korcula. The SNP genotyping for the CROATIA-Korcula cohort was performed in Helmholtz Zentrum MŸnchen, Neuherberg, Germany

DIABNORD: We are grateful to the study participants who dedicated their time and samples to these studies. We also thank the VHS, the Swedish Diabetes Registry and Umeå Medical Biobank staff for biomedical data and DNA extraction. We also thank M Sterner, G Gramsperger and P Storm for their expert technical assistance with genotyping and genotype data preparation.

EFSOCH: The authors thank University of Exeter Medical School. EXTEND data were provided by the Peninsula Research Bank, part of the NIHR Exeter Clinical Research Facility.

EGCUT: EGCUT work was supported through the Estonian Genome Center of University of Tartu by the Targeted Financing from the Estonian Ministry of Science and Education [SF0180142s08]; the Development Fund of the University of Tartu (grant SP1GVARENG); the European Regional Development Fund to the Centre of Excellence in Genomics (EXCEGEN; grant 3.2.0304.11-0312); and through FP7 grant 313010. EGCUT were further supported by the US National Institute of Health [R01DK075787].

EPIC-CVD: We thank all EPIC participants and staff for their contribution to the study, the laboratory teams at the Medical Research Council Epidemiology Unit for sample management and Cambridge Genomic Services for genotyping, Sarah Spackman and Nicola Kerrison for data management, and the teams at the EPIC-CVD and EPIC-InterAct Coordinating Centres for study coordination and administration. Genotyping and lipid assays in EPIC-CVD were principally supported by grants awarded to the University of Cambridge from the EU Framework Programme 7 (HEALTH-F2-2012-279233), the UK Medical Research Council (G0800270) and British Heart Foundation (SP/09/002), and the European Research Council (268834).

EPIC-InterAct: Funding for the InterAct project was provided by the EU FP6 programme (grant number LSHM_CT_2006_037197). We thank all EPIC participants and staff for their contribution to the study. We thank the lab team at the MRC Epidemiology Unit for sample management and Nicola Kerrison for data management.

FamHS: The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes grants DK063491 and NIDDK DK089256.

FHS: Genotyping, quality control and calling of the Illumina HumanExome BeadChip in the Framingham Heart Study (FHS) was supported by funding from the National Heart, Lung and Blood Institute Division of Intramural Research (Daniel Levy and Christopher J. O'Donnell, Principal

Investigators). Support for the centralized genotype calling was provided by Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium through the National Institutes of Health (NIH) American Recovery and Reinvestment Act of 2009 (5RC2HL102419). The NHLBI's Framingham Heart Study is a joint project of the National Institutes of Health and Boston University School of Medicine and was supported by contracts N01-HC-25195 and HHSN268201500001I. A portion of this research was conducted using the Linux Clusters for Genetic Analysis (LinGA) computing resources at Boston University Medical Campus.

FINRISK 2007: The FINRISK 2007 survey was primarily funded by the Finnish National Institute for Health and Welfare (THL). Additional support was obtained from the Academy of Finland and domestic foundations. VS was supported by the Finnish Foundation for Cardiovascular Research

Generation Scotland: Generation Scotland received core funding from the Chief Scientist Office of the Scottish Government Health Directorate CZD/16/6 and the Scottish Funding Council HR03006. Genotyping of the GS:SFHS samples was carried out by staff at the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the UK's Medical Research Council.

GLACIER: We are indebted to the study participants who dedicated their time and samples to these studies. We thank J Hutiainen and Å Ågren (Umeå Medical Biobank) for data organization and K Enquist and T Johansson (Västerbottens County Council) for technical assistance with DNA extraction. We also thank M Sterner, G Gramsperger and P Storm for their expert technical assistance with genotyping and genotype data preparation.

GoDARTS_CAD: We are grateful to all the participants who took part in this study, to the general practitioners, to the Scottish School of Primary Care for their help in recruiting the participants, and to the whole team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. We acknowledge the support of the Health Informatics Centre, University of Dundee for managing and supplying the anonymised data and NHS Tayside, the original data owner. The Wellcome Trust supported the Wellcome Trust UK Type 2 Diabetes Case Control Collection (Go-DARTS) and the Scottish Health Informatics Programme. The Chief Scientist Office supported informatics. Project funded by the UK Medical Research Council (G0601261). Genotyping was funded by the Wellcome Trust. Analysis was supported by BHF grant (Deloukas) RG/14/5/30893.

GRAPHIC: Recruitment for the GRAPHIC Study was funded by the British Heart Foundation. Exome array genotyping was funded by the NIHR and the Wellcome Trust. Nilesh J Samani holds a Chair funded by the British Heart Foundation and is a NIHR Senior Investigator. NM is funded by the NIHR Leicester Cardiovascular Biomedical Research Unit.

Health: The Health2006 was financially supported by grants from the Velux Foundation; The Danish Medical Research Council, Danish Agency for Science, Technology and Innovation; The Aase and Ejner Danielsens Foundation; ALK-Abello A/S, Hørsholm, Denmark, and Research Centre for Prevention and Health, the Capital Region of Denmark. This work was supported by the Timber Merchant Vilhelm Bang's Foundation, the Danish Heart Foundation (Grant number 07-10-R61-A1754-B838-22392F), and the Health Insurance Foundation (Helsefonden) (Grant number 2012B233). The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen partially funded by an unrestricted donation from the Novo Nordisk Foundation (www.metabol.ku.dk).

HELIC MANOLIS and HELIC Pomak: HELIC (Hellenic Isolated Cohorts). This work was funded by the Wellcome Trust (098051) and the European Research Council (ERC-2011-StG 280559-SEPI). The MANOLIS cohort is named in honour of Manolis Giannakakis, 1978-2010. We thank the residents of the Mylopotamos villages, and of the Pomak villages, for taking part. The HELIC study has been supported by many individuals who have contributed to sample collection (including A. Athanasiadis, O. Balafouti, C. Batzaki, G. Daskalakis, E. Emmanouil, C. Giannakaki, M. Giannakopoulou, A. Kaparou, V. Kariakli, S. Koinaki, D. Kokori, M. Konidari, H. Koundouraki, D. Koutoukidis, V. Mamakou, E. Mamalaki, E. Mpamiaki, M. Tsoukana, D. Tzakou, K. Vosdogianni, N. Xenaki, E. Zengini), data entry (T. Antonos, D. Papagrigoriou, B. Spiliopoulou), sample logistics (S. Edkins, E. Gray), genotyping (R. Andrews, H. Blackburn, D. Simpkin, S. Whitehead), research administration (A. Kolb-Kokocinski, S. Smee, D. Walker) and informatics (M. Pollard, J. Randall).

HUNT: The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Centre (Faculty of Medicine, Norwegian University of Science and Technology NTNU), Nord-Trøndelag County Council, Central Norway Health Authority, and the Norwegian Institute of Public Health.

Inter99: The Inter99 was initiated by Torben Jørgensen (PI), Knut Borch-Johnsen (co-PI), Hans Ibsen and Troels F. Thomsen. The steering committee comprises the former two and Charlotta Pisinger. The study was financially supported by research grants from the Danish Research Council, the Danish Centre for Health Technology Assessment, Novo Nordisk Inc., Research Foundation of Copenhagen County, Ministry of Internal Affairs and Health, the Danish Heart Foundation, the Danish Pharmaceutical Association, the Augustinus Foundation, the Ib Henriksen Foundation, the Becket Foundation, and the Danish Diabetes Association.

JHS: We thank the Jackson Heart Study (JHS) participants and staff for their contributions to this work. The JHS is supported by contracts HHSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300050C from the National Heart, Lung, and Blood Institute and the National Institute on Minority Health and Health Disparities.

KORA: The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München - German Research Center for

Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

LBC1921: We thank the Lothian Birth Cohort 1921 (LBC1921) participants and team members who contributed to this study. Phenotype collection was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC), The Royal Society and The Chief Scientist Office of the Scottish Government. Genotyping was funded by the BBSRC. The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged.

LBC1936: We thank the Lothian Birth Cohort 1936 (LBC1936) participants and team members who contributed to this study. Phenotype collection was supported by Research Into Ageing (continues as part of Age UK The Disconnected Mind project). Genotyping was funded by the BBSRC. The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged.

MORGAM: The MORGAM Project received funding during the work from European Union FP 7 projects CHANCES (HEALTH-F3-2010-242244) and BiomarCaRE (278913). This has supported central coordination and part of the activities of the MORGAM Data Centre, at THL in Helsinki, Finland. MORGAM Participating Centres are funded by regional and national governments, research councils, charities, and other local sources. The CHD Exome+ coordinating centre receives core support from the UK Medical Research Council (G0800270), the BHF (SP/09/002), the NIHR and Cambridge Biomedical Research Centre, as well as grants from the European Research Council (268834) and the European Commission Framework Programme 7 (HEALTH-F2-2012-279233). Exome+ assays were supported by grants from Merck and Pfizer.

MESA: Multi-Ethnic Study of Atherosclerosis (MESA) and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079, UL1-TR-000040, and DK063491. MESA Family is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support is provided by grants and contracts R01HL071051, R01HL071205, R01HL071250, R01HL071251, R01HL071258, R01HL071259, by the National Center for Research Resources, Grant UL1RR033176, and the National Center for Advancing Translational Sciences, Grant UL1TR000124. This publication was developed under a STAR research assistance agreement, No. RD831697 (MESA Air), awarded by the U.S Environmental protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication.

NFBC1986: The authors are grateful to the NFBC1986 participants and their families and to the NFBC research staff for data collection. We also thank late Professor Paula Rantakallio (launch of NFBC1966 and 1986), Ms OutiTornwall and Ms MinttuJussila (DNA biobanking). Financial support was received from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), the European Commission (EURO-BLCS, Framework 5 award QLG1-CT-2000-01643), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), NIH/NIMH (5R01MH63706:02), the Medical Research Council, UK (G0500539, G0600705, PrevMetSyn/SALVE) and the Wellcome Trust (project grant GR069224), UK, ENGAGE project and grant agreement HEALTH-F4-2007-201413, and the EU Framework Programme 7 small-scale focused research collaborative project EurHEALTHAgeing 277849. The DNA extractions, sample quality controls, biobank up-keeping and aliquotting were performed in the National Public Health Institute, Biomedicum Helsinki, Finland and supported financially by the Academy of Finland and Biocentrum Helsinki.

OBB: The Oxford Biobank (OBB) is supported by the Oxford Biomedical Research Centre and part of the National NIHR Bioresource.

PROMIS: Field-work, genotyping, and standard clinical chemistry assays in PROMIS were principally supported by grants awarded to the University of Cambridge from the British Heart Foundation, UK Medical Research Council, Wellcome Trust, EU Framework 6–funded Bloodomics Integrated Project, Pfizer, Novartis, and Merck.

PUUMA: The PUUMA project is a collaboration between Peking University Health Science Center and The University of Michigan Medical School. Funding for the project was provided by the University of Michigan Medical School and the Peking University Health Sciences Center Joint Institute for Clinical and Translational Research (Principal Investigators: C.J.W./W.G. and S.K.G./Y.H.)

SardiNIA: This work was supported by Contract NO1-AG-1-2109 from the National Institute of Aging, and in part by a grant from the Italian Ministry of Economy and Finance to the CNR for the Project “FaReBio di Qualità” to F. Cucca.

SDC: The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen partially funded by an unrestricted donation from the Novo Nordisk Foundation (www.metabol.ku.dk).

TwinsUK: TwinsUK study was funded by the Wellcome Trust; European Community’s Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute

for Health Research (NIHR) BioResource Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.

WGHS: The WGHS is supported by HL043851 and HL080467 from the National Heart, Lung, and Blood Institute and CA047988 from the National Cancer Institute with collaborative scientific support and funding for genotyping provided by Amgen.

WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, the US National Institutes of Health and the US Department of Health and Human Services (HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C and HHSN271201100004C). Exome chip data and analysis were supported through the Exome Sequencing Project (NHLBI RC2 HL-102924, RC2 HL-102925 and RC2 HL-102926), the Genetics and Epidemiology of Colorectal Cancer Consortium (NCI CA137088), and the Genomics and Randomized Trials Network (NHGRI U01-HG005152). The authors thank the WHI investigators and staff for their dedication and the study participants for making the program possible.

References and Notes:

1. Strachan, D.P. *et al.* Lifecourse influences on health among British adults: effects of region of residence in childhood and adulthood. *Int J Epidemiol* **36**, 522-31 (2007).
2. Lauritzen, T. *et al.* The ADDITION study: proposed trial of the cost-effectiveness of an intensive multifactorial intervention on morbidity and mortality among people with Type 2 diabetes detected by screening. *Int J Obes Relat Metab Disord* **24 Suppl 3**, S6-11 (2000).
3. Harris, T.B. *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* **165**, 1076-87 (2007).
4. CARDIoGRAMplusC4D Consortium, *et al.* Large scale association analysis of novel genetic loci for coronary artery disease. *Arterioscler Thromb Vasc Biol* **29**, 774-80 (2009).
5. The ARIC investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am J Epidemiol* **129**, 687-702 (1989).
6. Sever, P.S. *et al.* Anglo-Scandinavian Cardiac Outcomes Trial: a brief history, rationale and outline protocol. *J Hum Hypertens* **15 Suppl 1**, S11-2 (2001).
7. Erdmann, J. *et al.* New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat Genet* **41**, 280-2 (2009).
8. Muntendam, P. *et al.* The BioImage Study: novel approaches to risk assessment in the primary prevention of atherosclerotic cardiovascular disease--study design and objectives. *Am Heart J* **160**, 49-57 e1 (2010).
9. Roden, D.M. *et al.* Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther* **84**, 362-9 (2008).
10. Weeke, P. *et al.* Examining rare and low-frequency genetic variants previously associated with lone or familial forms of atrial fibrillation in an electronic medical record system: a cautionary note. *Circ Cardiovasc Genet* **8**, 58-63 (2015).
11. Chowdhury, R. *et al.* The Bangladesh Risk of Acute Vascular Events (BRAVE) Study: objectives and design. *Eur J Epidemiol* **30**, 577-87 (2015).
12. Caulfield, M. *et al.* Genome-wide mapping of human loci for essential hypertension. *Lancet* **361**, 2118-23 (2003).
13. Friedman, G.D. *et al.* CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol* **41**, 1105-16 (1988).
14. Schnohr, P., Lange, P., Scharling, H. & Jensen, J.S. Long-term physical activity in leisure time and mortality from coronary heart disease, stroke, respiratory diseases, and cancer. The Copenhagen City Heart Study. *Eur J Cardiovasc Prev Rehabil* **13**, 173-9 (2006).
15. Frikke-Schmidt, R. *et al.* Association of loss-of-function mutations in the ABCA1 gene with high-density lipoprotein cholesterol levels and risk of ischemic heart disease. *JAMA* **299**, 2524-32 (2008).
16. Varbo, A., Benn, M., Tybjaerg-Hansen, A. & Nordestgaard, B.G. Elevated remnant cholesterol causes both low-grade inflammation and ischemic heart disease, whereas elevated low-density lipoprotein cholesterol causes ischemic heart disease without inflammation. *Circulation* **128**, 1298-309 (2013).
17. Fried, L.P. *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* **1**, 263-76 (1991).
18. Zemunik, T. *et al.* Genome-wide association study of biochemical traits in Korcula Island, Croatia. *Croat Med J* **50**, 23-33 (2009).
19. Hallmans, G. *et al.* Cardiovascular disease and diabetes in the Northern Sweden Health and Disease Study Cohort - evaluation of risk factors and their interactions. *Scand J Public Health Suppl* **61**, 18-24 (2003).
20. Tuomilehto, J. *et al.* Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* **344**, 1343-50 (2001).
21. Kouki, R. *et al.* Diet, fitness and metabolic syndrome--the DR's EXTRA study. *Nutr Metab Cardiovasc Dis* **22**, 553-60 (2012).
22. Kraus, W.E. *et al.* A Guide for a Cardiovascular Genomics Biorepository: the CATHGEN Experience. *J Cardiovasc Transl Res* **8**, 449-57 (2015).
23. Knight, B., Shields, B.M. & Hattersley, A.T. The Exeter Family Study of Childhood Health (EFSOCH): study protocol and methodology. *Paediatr Perinat Epidemiol* **20**, 172-9 (2006).

24. Leitsalu, L. *et al.* Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *Int J Epidemiol* **44**, 1137-47 (2015).
25. Gonzalez, C.A. The European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutr* **9**, 124-6 (2006).
26. InterAct Consortium, *et al.* Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. *Diabetologia* **54**, 2272-82 (2011).
27. Higgins, M. *et al.* NHLBI Family Heart Study: objectives and design. *Am J Epidemiol* **143**, 1219-28 (1996).
28. Rolfe Ede, L. *et al.* Association between birth weight and visceral fat in adults. *Am J Clin Nutr* **92**, 347-52 (2010).
29. Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J. & Castelli, W.P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* **110**, 281-90 (1979).
30. Kotronen, A. *et al.* Non-alcoholic and alcoholic fatty liver disease - two diseases of affluence associated with the metabolic syndrome and type 2 diabetes: the FIN-D2D survey. *BMC Public Health* **10**, 237 (2010).
31. Vartiainen, E. *et al.* Thirty-five-year trends in cardiovascular risk factors in Finland. *Int J Epidemiol* **39**, 504-18 (2010).
32. Valle, T. *et al.* Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. *Diabetes Care* **21**, 949-58 (1998).
33. Scott, L.J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341-5 (2007).
34. Morris, A.D. *et al.* The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. *BMJ* **315**, 524-8 (1997).
35. Tomaszewski, M. *et al.* Genetic architecture of ambulatory blood pressure in the general population: insights from cardiovascular gene-centric array. *Hypertension* **56**, 1069-76 (2010).
36. Thuesen, B.H. *et al.* Cohort Profile: the Health2006 cohort, research centre for prevention and health. *Int J Epidemiol* **43**, 568-75 (2014).
37. Krokstad, S. *et al.* Cohort Profile: the HUNT Study, Norway. *Int J Epidemiol* **42**, 968-77 (2013).
38. Jorgensen, T. *et al.* A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. *Eur J Cardiovasc Prev Rehabil* **10**, 377-86 (2003).
39. Taylor, H.A., Jr. *et al.* Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. *Ethn Dis* **15**, S6-4-17 (2005).
40. Holle, R., Happich, M., Lowel, H., Wichmann, H.E. & Group, M.K.S. KORA--a research platform for population based health research. *Gesundheitswesen* **67 Suppl 1**, S19-25 (2005).
41. Wichmann, H.E., Gieger, C., Illig, T. & Group, M.K.S. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1**, S26-30 (2005).
42. Deary, I.J., Whiteman, M.C., Starr, J.M., Whalley, L.J. & Fox, H.C. The impact of childhood intelligence on later life: following up the Scottish mental surveys of 1932 and 1947. *J Pers Soc Psychol* **86**, 130-47 (2004).
43. Deary, I.J., Gow, A.J., Pattie, A. & Starr, J.M. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. *Int J Epidemiol* **41**, 1576-84 (2012).
44. Deary, I.J. *et al.* The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr* **7**, 28 (2007).
45. Kooner, J.S. *et al.* Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat Genet* **40**, 149-51 (2008).
46. 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).
47. Bild, D.E. *et al.* Multi-Ethnic Study of Atherosclerosis: objectives and design. *Am J Epidemiol* **156**, 871-81 (2002).
48. Stancakova, A. *et al.* Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. *Diabetes* **58**, 1212-21 (2009).
49. Dube, M.P. *et al.* CKM and LILRB5 are associated with serum levels of creatine kinase. *Circ Cardiovasc Genet* **7**, 880-6 (2014).

50. Evans, A. *et al.* MORGAM (an international pooling of cardiovascular cohorts). *Int J Epidemiol* **34**, 21-7 (2005).
51. Lind, L., Fors, N., Hall, J., Marttala, K. & Stenborg, A. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. *Arterioscler Thromb Vasc Biol* **25**, 2368-75 (2005).
52. Hedstrand, H. A study of middle-aged men with particular reference to risk factors for cardiovascular disease. *Ups J Med Sci Suppl* **19**, 1-61 (1975).
53. Isomaa, B. *et al.* A family history of diabetes is associated with reduced physical fitness in the Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study. *Diabetologia* **53**, 1709-13 (2010).
54. Barlera, S., Chiodini, B.D., Franzosi, M.G. & Tognoni, G. [PROCARDIS: A current approach to the study of the genetics of myocardial infarct]. *Ital Heart J Suppl* **2**, 997-1004 (2001).
55. Saleheen, D. *et al.* The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. *Eur J Epidemiol* **24**, 329-38 (2009).
56. Kulbertus, H. & Scheen, A.J. [The PROSPER Study (PROspective study of pravastatin in the elderly at risk)]. *Rev Med Liege* **57**, 809-13 (2002).
57. Tang, C.S. *et al.* Exome-wide association analysis reveals novel coding sequence variants associated with lipid traits in Chinese. *Nat Commun* **6**, 10206 (2015).
58. Pilia, G. *et al.* Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* **2**, e132 (2006).
59. Albrechtsen, A. *et al.* Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. *Diabetologia* **56**, 298-310 (2013).
60. Moayyeri, A., Hammond, C.J., Hart, D.J. & Spector, T.D. The UK Adult Twin Registry (TwinsUK Resource). *Twin Res Hum Genet* **16**, 144-9 (2013).
61. Ridker, P.M. *et al.* Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin Chem* **54**, 249-55 (2008).
62. The Women's Health Initiative Study Group. Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials* **19**, 61-109 (1998).
63. Shepherd, J. *et al.* Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med* **333**, 1301-7 (1995).