Technische Universität München

TUM School of Computation, Information and Technology

Genetic investigation of coronary artery disease by Mendelian randomization and transcriptome-wide association studies

Ling Li

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Burkhard Rost

Prüfer*innen der Dissertation:1. Prof. Dr. Julien Gagneur

2. Prof. Dr. Heribert Schunkert

3. apl. Prof. Dr. Bertram Müller-Myhsok

Die Dissertation wurde am 08.08.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 30.07.2023 angenommen.

Acknowledgments

Foremost, I would like to thank the members of my dissertation committee including: Prof. Dr. Burkhard Rost**,** Prof. Dr. Julien Gagneur, Prof. Dr. Bertram Muller-Myhsok, and Prof. Dr. Heribert Schunkert, not only for their time and extreme patience, but for their intellectual contributions to my development as a scientist.

I want to express my sincere gratitude to my advisor Prof. Heribert Schunkert. It's my honor to join his group. He is an inspiring mentor, very open-minded to new ideas, and willing to give support to young researchers. During my PhD study, he has closely supervised my work and given full support and guidance at all levels of my PhD research. He gave me great opportunities to train in research presentation and scientific writing, which are critical for my academic career. I appreciate all his knowledgeable input, kindness, enthusiasm, and encouragement throughout my PhD training.

To my supervisor Prof. Julien Gagneur, I thank him for providing the opportunity to pursue my PhD research at the Deutsches Herzzentrum München and Technische Universität München, where I can apply advanced bioinformatic techniques in translational research. He is a leading expert in bioinformatics and I have enhanced my technical skills under his guidance. I acknowledge the collaboration between the labs of Prof. Gagneur and Prof. Schunkert which inspired me to perform the first CAD TWAS allowing the great integration of 'wet' and 'dry' lab-based research. I thank his valuable advice for my PhD work.

I would like to express my special thanks to Dr. Lingyao Zeng and Dr. Zhifen Chen. I thank Lingyao for introducing me to Prof. Schunkert's lab and giving me lots of help at beginning of my PhD study when I faced challenges as an international student. Zhifen is a passionate researcher who always has lots of new ideas. It's very enjoyable to discuss science

topics with her. I appreciate her valuable advice and experiment validation for my TWAS findings. I also want to thank Zhifen for helping revise my dissertation.

To my lab mates, I thank them for the enjoyable lab life and support to my research work. I'd like to give special thanks to Dr. Moritz von Scheidt and Dr. Thorsten Keßler. Moritz has been continuously helping me to obtain data from external collaborators. Thorsten shared creative ideas for my PhD work. I also want to thank Shiyue Li who helped conduct validation experiments.

Last but not least, I want to thank my family and friends for their constant support through the ups and downs during my PhD study. This dissertation is dedicated to the memory of my father.

Abstract

Coronary artery disease (CAD) is the leading cause of death globally, which is in part caused by genetic variants. The majority of risk loci of CAD identified by genome-wide association study (GWAS) are located in non-coding regions that hamper their function interpretation. So, follow-up studies are needed.

Polygenic risk score (PRS) is a genetic estimate of an individual's liability to a trait or disease, calculated from a set of independent risk variants based on large-scale GWASs. The PRS has been widely applied to study the genetic association between complex traits. Mendelian randomization (MR) uses genetic variants as instrumental variables to infer whether risk factors (exposures) causally affect a health outcome, which is broadly used in observational epidemiology. The combination of PRS and MR ('PRS+MR' strategy) could improve detection rates for causal relationships which can be particularly useful when evaluating associations between genetic liability for a given trait and hundreds of diverse health outcomes. Transcriptome-wide association studies (TWAS) have been recently proposed as an invaluable tool for annotating GWAS risk loci by investigating the potential gene expression regulatory mechanisms underlying variant-trait associations. The signals identified by TWAS reflect associations between genetically regulated expression (GReX) and complex traits. In this dissertation, we applied these two bioinformatic approaches to study the genetic association between intelligence and CAD and to filter risk genes for CAD in a tissue-specific fashion.

The first work is to study the genetic association between intelligence and CAD using the 'PRS+MR' strategy. In this study, we first estimated a genetic intelligence score (gIQ) in samples from ten CAD case-control studies (n=34,083) from CARDIoGRAMplusC4D cohorts and UK Biobank (UKB) (n=427,306) based on 242 variants independently associated with

intelligence. Meta-analysis using a fixed-effect size model indicated the increase of gIQ by one standard deviation (SD) led to a relative decrease of CAD risk by 5% (95% confidence interval [CI], 0.93 to 0.98; *P*=4.93e-5). From UKB, we observed significant inverse correlations between gIQ and lifestyle factors of CAD including smoking, body mass index (BMI), type 2 diabetes (T2D), hypertension, and a positive correlation with high-density lipoprotein cholesterol (HDL-C). We also observed positive correlations for gIQ with measured intelligence and educational attainment. The association between gIQ and CAD risk was largely attenuated after the adjustment of lifestyle factors, suggesting mediatory roles of these lifestyle factors in the pathway of linking high gIQ and low risk of CAD. The effects of gIQ on lifestyle factors were also largely attenuated after the adjustment of measured intelligence and educational attainment, suggesting their mediatory roles in the pathways between gIQ and lifestyle factors. Finally, the associations between intelligence and CAD as well as its lifestyle factors were confirmed through the two-sample MR method. In conclusion, using genetic approaches, we depicted a pathway from gIQ to CAD risk. The higher gIQ is associated with the higher measured intelligence and longer educational attainment, both of which appear to reduce the prevalence of unfavourable factors of CAD including BMI, smoking, T2D, and hypertension, and increase HDL-C, which subsequently reduce the incidence of CAD.

The second work applied the TWAS methodology for filtering risk genes of CAD as well as for identification of the tissues in which differential expression affect rise. In this study, we first trained expression prediction models for nine risk tissues of CAD from the two largest reference panels, the Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET) and the Genotype-Tissue Expression (GTEx). We next applied these prediction models to imputing individual-level GReX from genotype data of 11 cohorts. By performing association analysis between GReX and CAD risk, we identified 114 transcriptome-wide significant genes (*P*<3.85e-6). Of these, 96 resided within GWAS loci of CAD (known), and

18 independent of GWAS loci (novel). Stepwise analyses bridged the known genes with a series of pathophysiological pathways related to CAD, including lipid metabolism, inflammation, angiogenesis, high blood pressure, etc. The *in-silico* analyses showed that the novel genes were associated with lipid metabolism in both genotype data of human samples and expression data of an atherosclerosis mouse model. Of these novel genes, *KPTN* and *RGSI9*, which were rarely studied before, gave significant signals in liver tissue by TWAS analysis. The CRISPR/Cas9-based gene knockdown of the two genes in the human hepatocyte cell line resulted in reduced secretion of APOB100 and lipids in the cell culture medium. In conclusion, our CAD TWAS work i) prioritized candidate genes in a tissue-specific fashion, and ii) identified 18 novel genes to be associated with CAD by their connection to lipid metabolism.

In summary, the two studies of my dissertation using different genetic approaches, on the one hand depicted the genetic pathway from high intelligence to low risk of CAD, and on the other hand pinpointed risk genes for CAD in a tissue-specific fashion. These studies expand our knowledge scope of the genetic etiology of CAD from different perspectives.

Abbreviations

CAD	Coronary artery disease
GWAS	Genome-wide association study
PheWAS	Phenome-wide association study
TWAS	Transcriptome-wide association study
MR	Mendelian randomization
PRS	Polygenic risk score
GReX	Genetically regulated expression
NGS	Next-generation sequencing
1KG	The 1000 Genome project
HRC	The Haplotype Reference Consortium
WGS	Whole-genome sequencing
WES	Whole-exome sequencing
RNAseq	RNA sequencing
eQTL	Expression quantitative trait locus
GTEx	The Genotype-Tissue Expression project
STARNET	The Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task
UKB	UK Biobank
LD	Linkage disequilibrium
SNP	Single nucleotide polymorphism
gIQ	Genetic intelligence score
AOR	Atherosclerotic aortic root

COR	Coronary
MAM	Internal mammary artery
TIB	Tibial artery
LIV	Liver
SF	Subcutaneous fat
VAF	Visceral abdominal fat
SKLM	Skeletal muscle
BLD	Blood
LDL-C	Low-density lipoprotein cholesterol
HDL-C	High-density lipoprotein cholesterol
TG	Triglycerides
TC	Total cholesterol
APOA	Apolipoprotein A
APOB	Apolipoprotein B

Table of Contents

I Introduction

1.1 Coronary artery disease (CAD)

1.1.1 Epidemiology of CAD

Epidemiologic and experimental cardiovascular research has improved the understanding of coronary artery disease (CAD) pathophysiology. Preventive and therapeutic strategies based on this knowledge decreased age-adjusted CAD mortality by over 50% in the past 30 years[1]. However, CAD remains the leading cause of death worldwide. Globally, CAD affects ~126 million individuals globally (1,655 persons per 100,000) representing 1.72% of the world population, and caused ~9 million deaths in 2017[2]. With the aging of the population, prevalence is expected to exceed 1845 persons per 100,000 by 2030. According to the European Commission and European Society of Cardiology (ESC), CAD causes 12.5% of all death and costs over €100 billion yearly in Europe. The enormous health care and economic burdens urge better preventive and therapeutic approaches to tackle the disease.

1.1.2 Biological mechanisms underlying CAD

CAD is clinically presented as the accumulation of atherosclerotic plaques within the wall of coronary arteries that provide nutrients and oxygen to the heart. CAD can be caused by dyslipidemia, dysfunction of endothelial cells (EC), vascular smooth muscle cells (VSMC) or fibroblasts, immune-inflammatory reactions, hyperinsulinism, and abnormal glucose metabolism[3]. Although the underlying mechanisms of CAD are intensively investigated the pathophysiology of CAD has not been fully elucidated yet. Generally, EC damage is thought to be the initial step of CAD, which may start early in life. These cells show dysfunction in regulating vascular tone via nitric oxide signaling. The increased amount of lipoprotein particles enhances activation and adhesion of monocytes and promotes the cholesterol-loading

of macrophages, or ‘foam cells’. VSMC and fibroblasts underlying the EC layer proliferate, migrate and remodel the vessels. In addition, inflammation and immune response trigger the vicious cycle of EC damage and deposition of fat and calcium inside the artery wall, all of which accelerate atherosclerotic plaque formation, narrow the vessels, and eventually obstruct the blood flow[4].

1.1.3 Risk factors for CAD

Various risk factors affect the multifactorial etiology of CAD. Risk factors of CAD could be classified into non- and modifiable types[5–7]. Non-modifiable risk factors include increased age, male gender, ethnicity, and family history of CAD. Modifiable risk factors include dyslipidemia, smoking, body-mass index (BMI), hypertension, diabetes mellitus, obesity, poor diet, lack of physical activity, and stressful lifestyle. Dyslipidemia could be caused by dysregulated levels of lipoprotein (a) (LPA), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglycerides (TG), and other lipoproteins.

1.2 Genome-wide association study (GWAS)

Genome-wide association study (GWAS) is a methodology for studying associations between genetic variants and phenotypes across the entire genome by testing allele frequency differences of genetic variants among individuals with and without the trait. Importantly, ancestry background needs to be similar between affected and control individuals[8]. Genetic variants in GWAS include mainly single-nucleotide polymorphisms (SNPs) and a small portion of structured variants, e.g., insertion/deletion polymorphism (InDel). The classic workflow of conducting a GWAS test is shown in Figure 1[9]. First, researchers need to recruit appropriate participants, from whom they collect phenotype data and tissue samples for DNA extraction and genotyping. Second, genetic data will be obtained using microarray genotyping technology, or next-generation sequencing (NGS) technology like whole-genome sequencing

(WGS) and whole-exome sequencing (WES). In addition, currently, targeted sequencing is becoming a rapid and cost-effective way of detecting known and novel variants in selected genomic regions. Third, researchers have to conduct quality control (QC) on both variant and sample levels after calling genotype from the array or NGS platforms. The QC is an essential step to remove factors that might cause bias on association test, such as removing low-quality variants, removing variants and individuals with low calling rate, detecting the population strata, and removing samples of genetically related. Fourth, the imputation step predicts undetected variants from detected variants based on genetics reference panels. The 1000 Genome (1KG) from the International Genome Sample Resource[10] and the Haplotype Reference Consortium (HRC)[11] are the two most popular genetic reference panels. Fifth, the imputed genotype data will be used for testing associations between genetic variants and phenotypes. According to the International HapMap project and other studies, the human genome has approximately one million independent common genetic variants on average[12], which makes the Bonferroni correcting threshold of $P<5e\text{-}8$ (false discovery rate 0.05/1e6) the most popular significance threshold of GWAS. PLINK[13] and CGTA[14] (Genome-wide Complex Trait Analysis) are the two most popular tools of GWAS analysis.

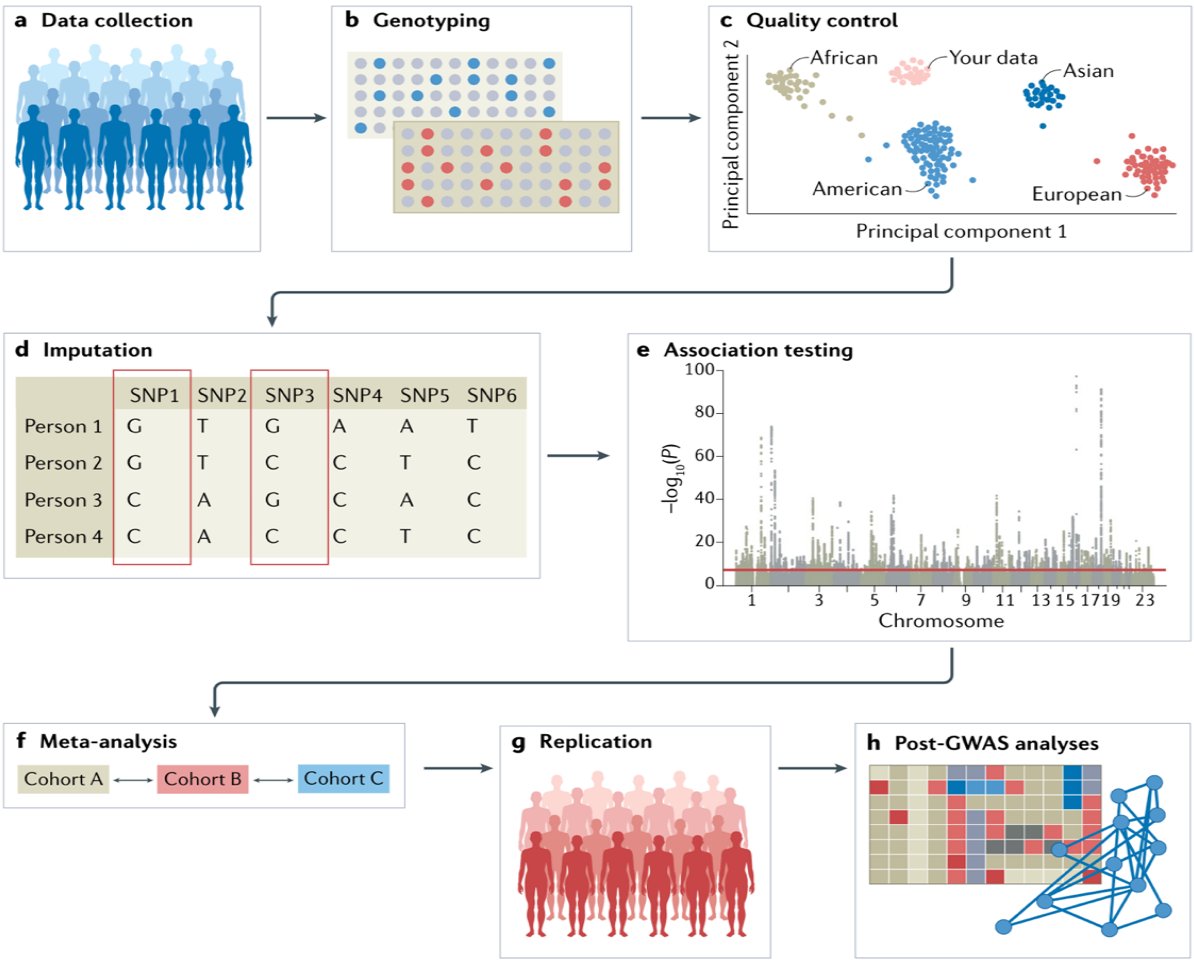


Figure 1 Overview of steps for conducting GWAS. The figure was adopted from Emil Uffelmann[9]

GWAS tests associations of genetic variants with traits on the basis of the single-SNP model at a genomic scale. A group of significant variants in the same locus are frequently observed in association tests because of the linkage disequilibrium (LD) relationship among common variants, which challenges the identification of the causal variants. Several attempts could help to prioritize the candidate causal variants at a locus. The simplest method is to pick the lead SNP (the one with the lowest association *P*-value) in a genomic region (e.g., a 1-Mb window centered on the locus). The method is based on the hypothesis of a single causal variant at a genomic region, and the top SNP captures the maximum amount of variation by its LD relations with other significant SNPs in the vicinity. Yet, the hypothesis may have several limitations. First, even if the hypothesis is true, the genotyping and imputation may not capture

all the variants at a locus. Second, the hypothesis might not be true given that multiple causal variants could exist at a locus. In this case, a single SNP does not fully represent potential causal variants, especially the ones not LD-linked with the lead[15]. To overcome these limitations, a tool, conditional and joint multiple-SNP analysis of GWAS summary statistics (COJO), was developed to identify additional variants associated with complex diseases in a genomic locus[15]. COJO, embedded in CGTA, requires two inputs, the GWAS summary statistics, and the reference genome of the same ethnic background, to estimate LD-correlations between SNPs, which largely facilitate the studies without individual-level genotype data[14, 15]. COJO starts with conditional analysis of the lead-associated SNPs, followed by a stepwise selection of regional SNPs by conditioning the effects of lead SNPs. COJO performs a stepwise model to select independently associated SNPs (*e.g.*, $P<5e-8$) based on conditional *P* values and estimates the joint effects of all selected SNPs after the model has been optimized. Compared with the method of regional lead SNP picking, COJO may identify multiple independent candidate causal variants in the same regions and can capture larger phenotype variations in the genomic regions.

1.3 Post-GWAS era

1.3.1 Fine-mapping

A decade of GWASs have uncovered thousands of genomic variants associated with complex human traits and diseases. In the post-GWAS era, efforts were directed to delineate the biological mechanisms underlying the associations between risk variants and traits/diseases. The biggest challenge originates from the non-coding feature of the majority of the GWAS loci, which hampers their functional interpretation. To address this, many different approaches have been developed including expression quantitative trait loci (eQTL) analysis, colocalization analysis, and tissue of action (TOA) analysis. Many collective resources, like

NIH Roadmap consortium, ENCODE, FANTOM, and the IHEC consortium provide functional annotation to genetic variants and regions. These indispensable context-specific resources can be integrated into fine-mapping methods to pinpoint regulatory mechanisms of GWAS variants on disease risk.

A direct way of understanding the effects of variants showing association by GWAS is to test the effects of these variants on gene expression of cells or tissues, namely eQTL analysis. An eQTL explains a fraction of gene expression regulated by genetic variants at this locus[16]. The genetics-of-gene-expression panels, such as the Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET)[17] and the Genotype-Tissue Expression (GTEx)[18], have largely facilitated eQTL studies. These panels usually collected genotype and tissue or cell-type transcriptomic data from hundred participants. Standard eQTL analysis firstly maps GWAS variants to a gene region and then tests the association between individual respective genotypes and expression levels in a tissue- or cell-type-specific manner. Based on the distance between variants and genes in genome architecture, eQTLs have been classified into *cis* and *trans*. Conventionally, variants residing within ±1Mb of a gene's transcription starting site (TSS) or gene-body regions are classified as *cis* variants; otherwise, variants are classified as *trans* variants. Although most studies to date focus on *cis*-eQTLs, *trans*-eQTL studies are catching up[17, 19]. Besides eQTL, splicing QTL (sQTL), protein QTL (pQTL), and methylation QTL (mQTL) are increasingly investigated driven by the need of detangling the complexity of variants' function and availability of corresponding omics datasets[20–22].

A QTL variant is sometimes associated with multiple genes and a gene often has multiple QTL variants, which challenges the identification of causal variants and genes. Based on QTL analysis, scientists have introduced another strategy, colocalization, to prioritize variants and genes by integrating GWAS and QTL signals[23–25]. Colocalization seeks to

identify shared causal variants between a molecular profile (e.g., gene expression or protein level) and a disease trait in a genomic locus, therefore prioritizing candidate causal genes from GWAS loci. COLOC is a popular tool for colocalization between pairs of genetic association studies using the Bayesian model (Figure 2)[26]. COLOC takes summary data of eQTL and GWAS as input to compute the probability of five hypotheses: H0 corresponding to no eQTL and no GWAS association in the region; H1 corresponding to association with eQTL but no GWAS association, or vice-versa (H2); H3 corresponding to eQTL and GWAS association but independent signals; H4 corresponding to shared eQTL and GWAS associations[23, 27]. MOLOC, a multi-trait extension of COLOC, was designed to compare association signals for multiple traits which could be used to pinpoint the regulatory mechanism of GWAS variants[26]. In the case of colocalization among eQTL, mQTL, and GWAS signals, the eQTL signal helps pinpoint the responsible genes in GWAS loci, while the mQTL signal helps identify the epigenetic mechanisms that impact gene expression, and which in turn affect disease risk[28].

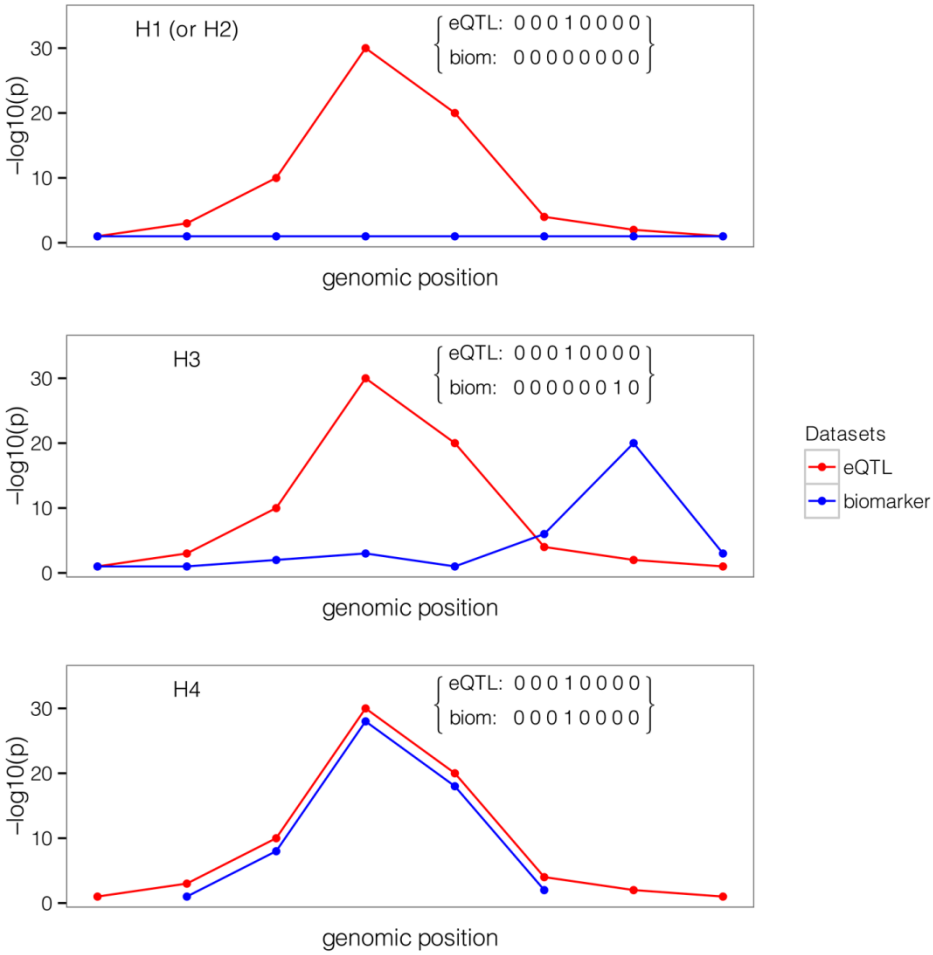


Figure 2 Example of one configuration under different hypotheses. A configuration is represented by a binary vector for each trait of (0,1) at a length of n=9, the number of shared variants in a region. The two traits shown in the figure are eQTL and GWAS signal. The value of 1 means the variant of the position is causally related to the trait, 0 vice versa. The top plot, corresponding to hypotheses 1 and 2, means that only one trait has a causal variant in the locus. The middle plot, corresponding to hypothesis 3, means that two traits have different causal variants in the locus. The bottom plot, corresponding to hypothesis 4, means that both traits share the same causal variant in the locus. The figure was adapted from Giambartolomei[26].

The tissue of action (TOA) score exerts a new strategy to partition the effects of variants on disease in the context of tissues or cell types[29]. For each GWAS signal, the TOA method systematically partitions posterior probabilities from the Bayesian fine-mapping method by

integrating tissue-specific functional annotation and expression data. The TOA score reflects the probability of genetic variant functioning in a specific tissue or cell type. As complex diseases are usually multiple tissues or cell types involved, the TOA score provides valuable guidance for in vivo/vitro experiment validation in the post-GWAS era.

1.3.2 Phenome-wide association study (PheWAS)

Phenome-wide association study (PheWAS), an alternative/complementary approach to the GWAS methodology, is to estimate associations of single genetic variants with a wide range of phenotypes. A fundamental difference between GWAS and PheWAS designs is the direction of inference: PheWAS studies pleiotropy of SNPs, a genotype-to-phenotype strategy; GWAS studies causality of SNPs, a phenotype-to-genotype strategy[30].

The PheWAS approach was originally developed due to the widespread availability of both anonymized human clinical electronic health record (EHR) data and matched genotype data. One of the main components of EHR is the International Classifications of Diseases version 9 (ICD9) codes, which includes information of 17,000 phenotypes binned into different hierarchy codes. In most cases, PheWASs divide cohorts into “cases” if participants have ICD9 codes related to a specific trait/disease, or “controls” if participants do not. Usually, GWAS SNPs are logical starting points for PheWAS because of the availability of association data. Staring with a known disease-associated SNP, a PheWAS study performs association tests for the specific SNP with a series of related traits. The significance level of PheWAS analysis is estimated via a Bonferroni correction as $\alpha=0.05/N$ (N the number of feasible models tested).

One of the major advantages of PheWAS is its potential to identify genomic variants with pleiotropic properties. Investigating cross-phenotype associations of SNPs makes PheWAS an important tool for understanding genetic associations among diseases/traits and gene-disease associations, as well as elucidating mechanisms of GWAS risk loci[31–34]. For

instance, a PheWAS work revealed that *CCDC92*, a novel GWAS risk locus of CAD, likely affected CAD through insulin resistance pathways[31]. PheWAS has been proposed as a possible aid in drug development through elucidating mechanisms of action, identifying alternative indications, or predicting adverse drug events[35].

PheWAS design has challenges. PheWAS is a hypothesis-generating approach, which is challenged by multiple comparison testing. The *N* used for calculating the significance threshold can be varied greatly in different PheWAS designs. Suppose a SNP is tested for associations with all traits listed in the ICD9, and the Bonferroni significance is 2.94e-6=0.05/17,000. If we had multiple SNPs tested at the same time, the threshold would be even higher. If we only tested associations of SNPs or SNP sets with several phenotypes, the threshold would be lower. This phenomenon could lead to false positive or false negative associations. Since PheWAS heavily relies on the annotation of traits/diseases, the depth and breadth of annotation data influence the PheWAS results. Like GWAS, association regression in PheWAS faces the problem of covering all possible covariates, which makes further validation of PheWAS findings necessary.

1.3.3 Mendelian randomization (MR)

As mentioned in section 1.1.3, many risk factors are associated with CAD, such as LDL-C, HDL-C, TG, inflammatory cytokines, larger BMI, Diabetes, and blood pressure. Finding the causal nature of risk factors on the disease process is critically important in epidemiology because these modifiable risk factors represent promising targets for primary intervention and drug development[36]. Mendelian randomization (MR) uses genetic variants as instrumental variables (IVs) to infer whether risk factors (exposures) causally affect a health outcome which is CAD in our case (Figure 3)[37]. In addition, the associations between exposures and

outcomes reflected by MR analyses are genetic associations that are independent of confounding factors.

In MR studies variants must fulfil three assumptions if they shall be used as instrumental varbles: i) the genetic variants are associated with the exposure (i.e. the risk factor - relevance assumption); ii) the genetic variants are associated with the outcome through the studied exposure only (exclusion restriction assumption); iii) the variants are independent of other cofounders which could affect the outcome (independence assumption)[37, 38]. Mendel's Law of Independent Assortment that the alleles segregate randomly when passed from parents to offspring, forms the foundation of MR. In addition, the germline genotypes are determined at conception, they precede the investigated outcomes, and therefore observed associations cannot be explained by reverse causation[39]. However, the wide existing pleiotropy of genetic variants and the LD association among genetic variants raise scientists' caution when performing MR analysis.

There are two types of MR designs: one-sample MR and two-sample MR (Figure 3)[39]. For one-sample MR analysis, the SNP-exposure associations and SNP-outcome associations are estimated from the same sample, set. In two-sample MR, the SNP-exposure associations are estimated in one sample set and the SNP-outcome association is estimated in a second sample set. Compared with the one-sample MR study, the two-sample MR study doesn't require exposure and outcome to be measured in all data sets, which allows researchers to use GWAS summary statistics from large consortia. This also improves transparency and reproducibility.

The classic steps of performing two-sample MR analysis are: i) identify genetic variants from a well-conducted GWAS; ii) obtain SNP-exposure associations from data set 1; iii) obtain SNP-outcome associations from data set 2; iv) harmonize SNP effects on exposure and

outcome; v) generate MR estimates; vi) perform sensitivity analyses. The inverse variance weighted (IVW) method is the most efficient estimate of the causal effect when all genetic variants are valid instruments[40]. IVV causal estimates could be biased when genetic variants exhibit horizontal pleiotropy because it violates the assumption of exclusion restriction. As mentioned in section 1.3.2, pleiotropy of genetic variants is inevitable in most MR studies with the increased knowledge of the genetic basis of complex diseases. For sensitivity analysis, scientists developed other robust methods, like MR Egger[41] and weighted median[42], that can provide valid causal inferences under weaker assumptions than the standard IVW mythology. Egger regression can identify the presence of 'directional' pleiotropy and provide a less biased causal estimate. The weighted median estimator provides a consistent estimate of the causal effect even when up to 50% of the information contributing to the analysis comes from genetic variants that are invalid IVs[42].

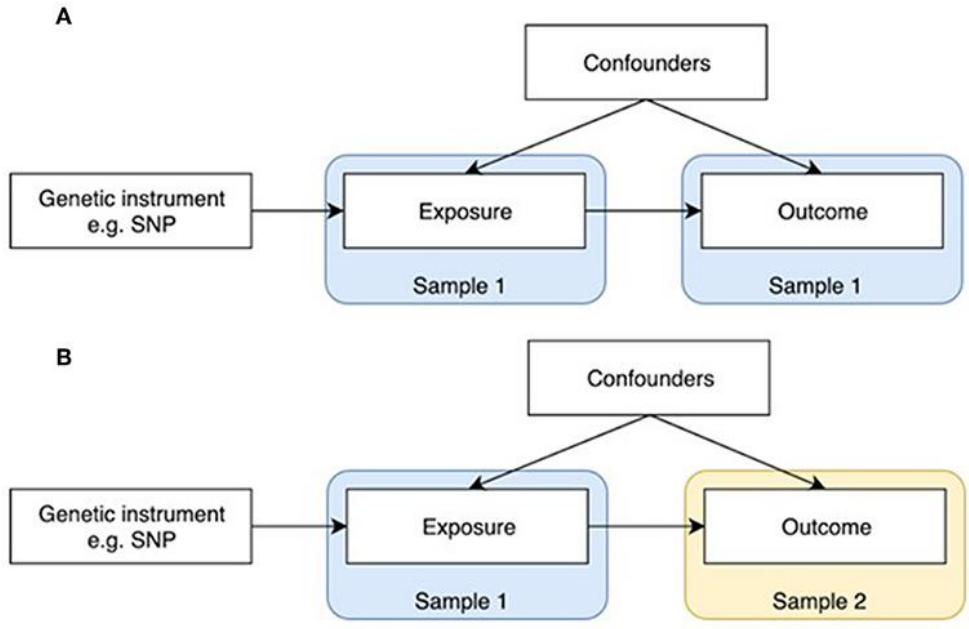


Figure 3 One-sample and two-sample Mendelian randomization designs. (A) One-sample Mendelian randomization (MR) measures both exposure and outcome in the same population.

(B) In two-sample MR, associations of genetic variants with exposures and outcomes are measured in non-overlapping populations. The Figure was adapted from Zheng[39].

The MR studies have identified many risk biomarkers, traits, and diseases relevant to the pathogenesis of CAD (Figure 4)[43]. Risk factors displaying confirmation by MR may be considered causal and include LDL-C, TG, LPA, IL6, non-fasting glucose, diabetes, obesity, adiponectin, blood pressure, and telomere length. Vice versa, MR studies challenged the roles of many biomarkers including HDL-C, CRP, lipoprotein-associated phospholipase A2 (LP-PLA_2), homocysteine, fibrinogen, bilirubin, and uric acid. Drugs targeting causal risk factors defined by MR are attractive treatment targets for CAD. For example, statins are well-known to reduce cardiovascular events and mortality in CAD patients because of their significant lipid-lowering functions[44]. We have shown in our MR studies that genetically modulated educational attainment may have implications for a series of unfavorable risk factor profiles, such as BMI and hypertension, and, subsequently, affects the prevalence of CAD[45]. These findings strengthen the importance of campaigns for enabling adequate schooling for preventing CAD.

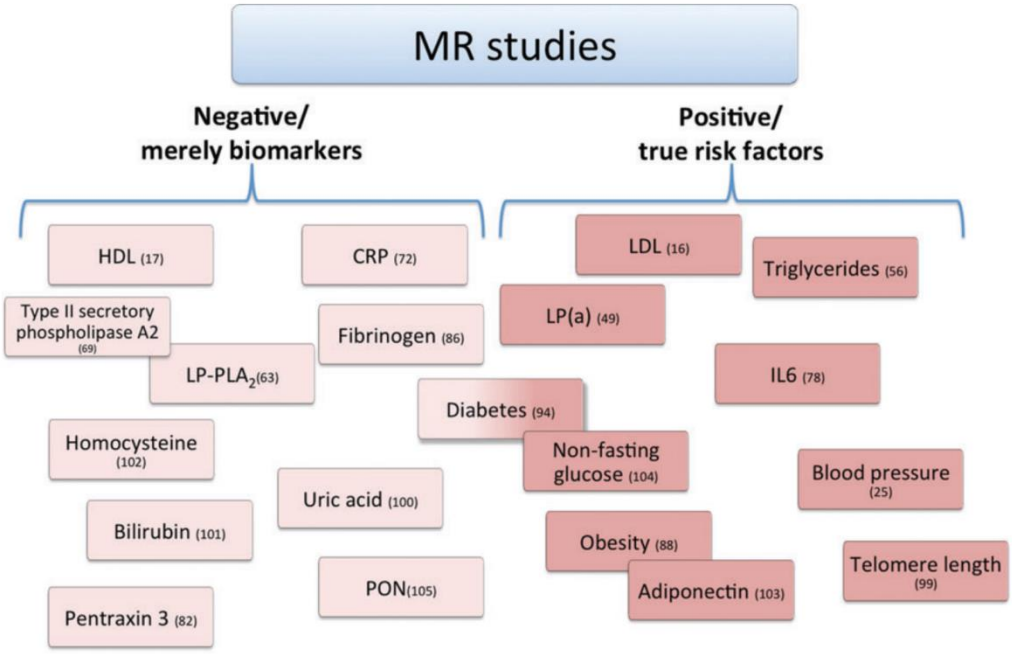


Figure 4. Risk factors of coronary artery diseases identified by MR analyses. The figure was adapted from Jansen[43].

1.4 Genetics-of-gene-expression panels

In the post-GWAS era, studying GWAS variants' effects on gene expression is a critical step to elucidating the genetic basis of common diseases or traits, which makes large-scale examination of genotype and transcript data in the context of disease-relevant tissues or cell types indispensable. The fast development of high-throughput technologies, such as the genotyping array, whole-genome sequencing (WGS), whole-exome sequencing (WES), RNA sequencing (RNA-seq), and single-cell RNA sequencing (scRNA-seq) accelerates the efficiency and affordability of generating such massive genotype and gene expression data.

The Genotype-Tissue Expression project (GTEx) is one of the most popular comprehensive public resources to study tissue-specific gene expression and regulation[18]. Samples of the GTEx project were collected from 54 non-diseased tissue sites of up to 1000 individuals for molecular assays including genotyping, WGS, WES, and RNA-seq. Among all

GTEx donors, 84.6% are white, 67.1% are male, 32.1% are between 50-59 years old, and 36% are between 60-70 years old. Furthermore, the GTEx consortium has also profiled scRNA data for eight tissues including breast, esophagus mucosa, esophagus muscularis, heart, lung, skeletal muscle, prostate, and skin. Scientists can download the processed data, such as expression and eQTL data, from the project portal, or apply for the secured data, such as genotyping data and phenotype data, from the dbGAP platform. The GTEx portal also provides a user-friendly interface and search engine at https://gtexportal.org/home/.

Different from GTEx which collected samples from healthy tissues, the Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET) represents the unique genetics-of-gene-expression panel for CAD study[17]. Tissue samples of STARNET were obtained from ~600 CAD patients during open thorax surgery, including atherosclerotic aortic root (AOR), atherosclerotic-lesion-free internal mammary artery (MAM), blood (BLD), subcutaneous fat (SF), visceral abdominal fat (VAF), skeletal muscle (SKLM), and liver (LIV). The STARNET patients are mainly Caucasians (31% females), 32% had diabetes, 75% had hypertension, 67% had hyperlipidemia, and 33% had an MI before age 60. By New York Heart Association criteria, 45% were class I, 42% class II, 9% class III, and 1% class IV CAD. DNA was genotyped with the OmniExpress Exome array (Illumina, ~900k SNPs), and RNA was sequenced using the HighSeq2000 platform. Researchers could apply the data from the dbGAP database. A study comparing tissue-specific eQTLs between STARNET and GTEs showed that STARNET had more eQTLs coinciding with CAD-associated risk SNPs suggesting that risk SNPs were more active in the focus of infection[17]. STARNET also provides a browser for exploring co-expression models inside and between tissues, which is at http://starnet.mssm.edu/[46].

1.5 UK Biobank (UKB)

UK Biobank (UKB) is the largest publicly available resource with deep phenotyping and genomic data of around 500,000 UK residents aged 40-69 years old initially recruited from 2006 to 2010[47]. The study continuously collects extensive phenotypic and genotypic data from all participants, including data from questionnaires, physical measures, sample assays, accelerometry, multimodal imaging, genome-wide genotyping, and longitudinal follow-up for a wide range of health-related outcomes.

The genome-wide genotyping data of ~500,000 participants released in 2016 was performed using the UK Biobank Axiom Array. Approximately 850,000 variants were directly measured, with>90million variants imputed using the HRC and UK10K + 1KG reference panels. UKB provides different levels of genotyping data, such as before or after imputation, to meet the needs of most applicants. Starting from 2016, UKB gradually released WES data of participants. By the fourth season of 2021, the WES data of ~500,000 participants were available for all participants[48]. The WES data measures the regions of the genome (about 2%) that encode proteins and is particularly suitable for identifying disease-causing rare variants (see section 1.6). By 2021, the database also released telomere data for the ~500,000 participants and metabolic data for ~127,000 participants. UKB's imaging study includes measures such as white matter hyper intensities derived from the brain scans, visceral fat derived from the abdominal scans, and left ventricular ejection fraction derived from the cardiac scans. The imaging data of ~50,000 participants were released in 2020, extra 50,000 participants' data could be expected in the near future. All data release information can be found at https://www.ukbiobank.ac.uk/enable-your-research/about-our-data. UKB also developed a Research Analysis Platform (RAP) powered by the DNAnexus (https://www.dnanexus.com/), providing a secure, scalable, and cloud-based environment for researchers to use UKB biomedical resources.

The greatest sample size and widest scope of phenotype data make UKB the most valuable resource for researchers of multiple fields. According to their statistics, the number of annual publications using data of UKB increased from only one in 2012 to 690 in 2021. The UKB resource was applied to multiple investigations of CAD revealing novel risk loci for CAD[49], identifying risk loci associated with aortic valve area from genotyping and imaging data[50], and studying the interaction between polygenetic risk scores and monogenic mutations[51], etc.

1.6 Rare mutation association study in complex disease

There are two controversial hypotheses about the contribution of genetic variants to individual susceptibility to common complex diseases: the common diseases common variant (CDCV) hypothesis, and the common disease rare variant (CDRV) hypothesis. The CDCV assumes that common variants (e.g., MAF>0.01) are major contributors to complex diseases. The CDRV supposes that rare genetic variants (e.g. MAF<0.01) play a more important role compared to common variants[52]. Many studies have shown that rare variants usually have a larger effect size on disease development than common variants which empowers the unique role of rare variants in investing genetic pathogenesis of complex diseases[53]. Same to GWAS, the association test for rare variants is a kind of hypothesis-free approach in which researchers do not start with a certain functional hypothesis. The enrichment of rare variants provides hypothesis-free evidence for gene causality. The rare missense, loss-of-function (LoF), and gain-of-function (GoF) variants mimic the impact of gene knocking out or enhancing experiments, which are ideal for functional analysis to elucidate new disease mechanisms. Rare variants showing association for a disease could provide favorable targets for drug development as their alleles mimic the effect of a modulated drug target. In addition, rare variants with larger effect size on disease development could be biomarkers for personalized medicine.

The workflow of the rare variant association study is shown in Figure 5[54]. The first step is to set up a work target that determines the variant calling platforms, genotyping, or sequencing. Compared to the genotyping array, the NGS technology has the advantages of massive scaling and high sensitivity in identifying rare variants. The next step is to call genotype data and perform quality filtering on the raw data. For data called from genotyping array, the quality control (QC) step is similar to GWAS, except that rare variants would be kept. When data is called from NGS platforms, extra QC parameters are required, such as sequencing depth, coverage, heterozygosity/contamination rate, etc. The third step is to assay function of rare variants using different in-silico methods, such as LRT score[55], MutationTaster[56], PolyPhen-2 HumDiv, PolyPhen-2 HumVar[57], and SIFT[58]. There is a variety of annotation sources to predict the consequence of genetic variants. sdbNSFP (v4.1) is the largest database designed to facilitate the functional annotation step by providing deleteriousness prediction and functional annotation for all potential non-synonymous and splice-site single nucleotide variants (a total of 84,013,093) in the human genome[59]. ClinVar is a freely accessible, public archive of reports of the associations of genetic variations with human phenotypes, with supporting evidence[60]. Tools like Variant Effect Predictor (VEP)[61], snpEff[62], and ANNOVAR[63] assemble these tools and resources into a set providing a single-line command to annotate the effect of genetic variants.

The next step is to perform an association test between functional variants and diseases/traits. The models designed for association analysis can be generally classified into two types: single-variant level and gene- or region-based aggression tests. In single-variant tests, the association is typically evaluated by linear regression for continuous traits and by logistic regression for binary traits. Single variant tests for rare variants are only suitable when the sample size is large enough; otherwise, they are less powerful than the tests for common variants. Different from single-variant tests, gene- or region-based aggression tests increase the

statistical power by evaluating the cumulative effects of multiple genetic variants in a gene or region. Based on varying assumptions about the underlying genetic models, methods for aggression test can be broadly categorized into five classes: burden test, adaption burden test, variance-component test, combined burden and variance-component (VC) test, and exponential-combination (EC) test[54]. Fisher's exact test for categorical phenotypes is the simplest combined test. The Sequence Kernel Association Test (SKAT), a widely used score-based VC test, is robust to groupings including variants with both positive and negative effects[64].

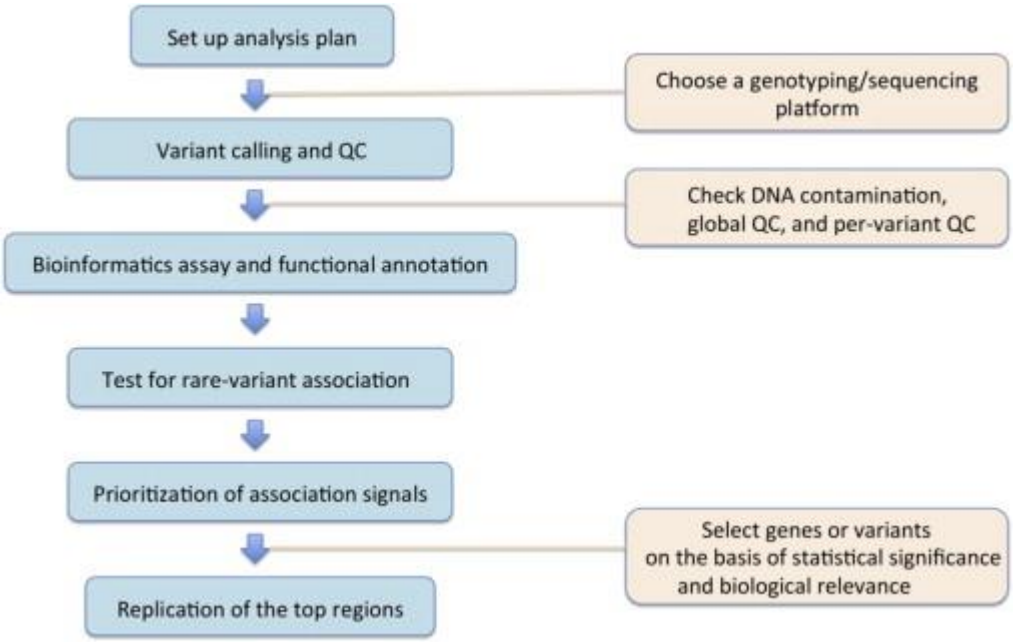


Figure 5 Workflow of rare variant association study. The figure was adapted from Lee[54].

The association test of rare variants informs the posterior probability of the disease relevance of a gene, which could guide the decision-making for biological validation. By performing a meta-analysis of exome-chip studies of European descent involving 42,335 patients and 78,240 controls, the CARDIoGRAMplusC4D Consortium identified associations of variants in *ANGPTL4*, *LPL*, and *SVEP1* with CAD[65]. Following the genetic study, an *in-*

vivo investigation was carried out to elucidate the mechanism of the novel gene *SVEP1* in CAD development[66].

1.7 Transcriptome-wide association study

Transcriptome-wide association studies (TWAS) have been recently proposed as an invaluable tool for annotating GWAS risk loci by investigating the potential gene regulatory mechanisms underlying variant-trait associations. The expression level of a gene can be decomposed into three components: a genetically determined component, a component altered by the trait itself, and a component determined by the remaining factors (including environment)[67]. Based on the assumption, TWAS reflects associations between genetically regulated gene expression (GReX) and diseases or traits.

The TWAS methodologies are broadly separated into two categories: the individual-level predictor and the GWAS summary-based predictor. The first step of both methodologies is to train prediction models from reference genetics-of-gene-expression panels which scale genotype and expression data simultaneously (Figure 6). Then the individual-level predictors would apply prediction models to impute expression profiles from individual genotypes of GWAS cohorts and perform association test between predicted expression and traits (section 1.7.1). However, the GWAS summary-based predictors take GWAS summary data as input to examine associations between intermediate gene expression levels and phenotypes (section 1.7.2).

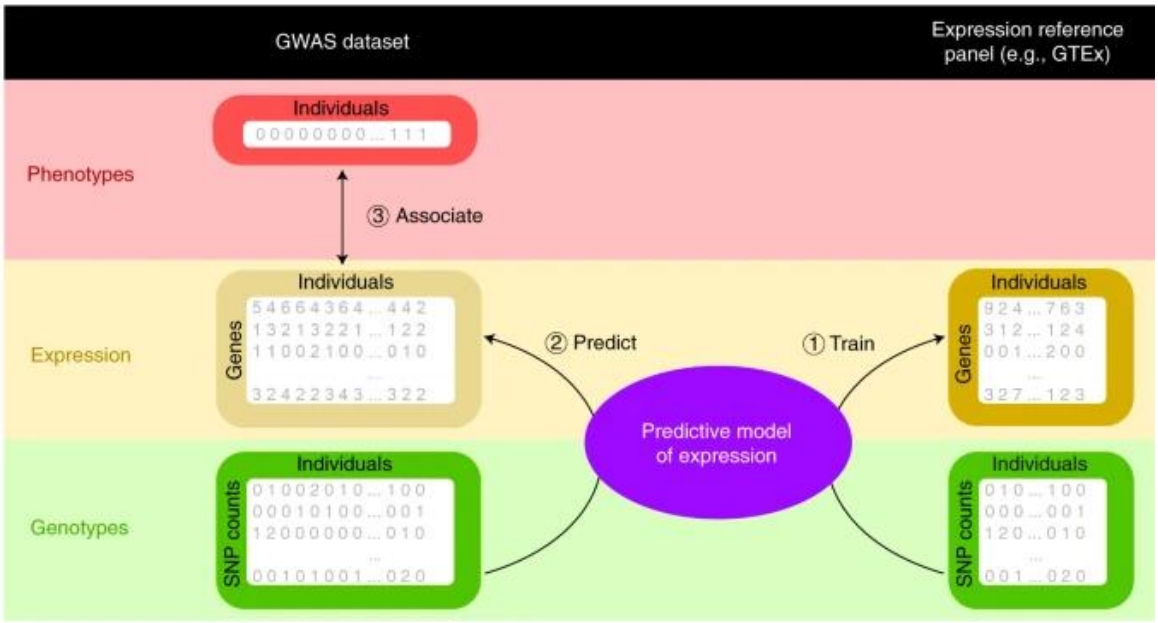


Figure 6 An overview of the TWAS. Briefly, TWAS involves i) training tissue-specific prediction models from references; ii) predicting genetically regulated expression (GReX) from genotype data using prediction models; iii) associating GReX with phenotypes. The figure was adapted from Wainberg[68].

1.7.1 Individual-level predictors

PrediXcan was the first individual-level predictor which aggregates impacts of variant set on gene expression[67]. Studies have suggested that aggregating variants by integrating gene expression or other omics may better explain underlying biological mechanisms and increase the power of association studies beyond GWAS[7]. Based on PrediXcan, Gao developed EpiXcan[69] which outperformed PrediXcan in prediction performance by integrating epigenome annotation from Roadmap Epigenomics Mapping Consortium (REMC)[70]. It firstly calculates SNP priors by using a hierarchical Bayesian model (qtlBHM) that jointly analyzes REMC epigenome annotations and eQTL statistics. Then, priors are transformed with an adaptive mapping function to penalty factors, which are then utilized by the weighted elastic net (WENet), a model selection technique that combines LASSO and Ridge regression and

seeks to identify which predictor variables to include in a regression model. The WENet model analyzes SNP priors, genotypes, and gene expression traits to estimate genetically regulated expression across different. The simulated expression value for each gene is

$$y = X \times \beta + \varepsilon$$

Here, *X* denotes the genotype matrix of cis-SNPs included in the prediction model, β denotes the coefficient vector of the cis-SNPs, '×' denotes matrix-vector product, and ε denotes residual value. The criterion of the WENet model can be written as:

$$C_{WENet}(\theta, \lambda, \alpha) = \sum_{i=1}^{n}[y_i - X_i\beta]^2 + \lambda\alpha|\beta|_w + \lambda(1 - \alpha)\beta^T W\beta,$$

$$|\beta|_w = \sum_{j=1}^{m} w_j|\beta_j|$$

In the above equations, *n* is the number of samples used for training model; X_i, $1 \le i \le n$, is the *i*-th row-vector of matrix *X* containing genotypes with dosages from 0 to 2; *m* is the number of cis-SNPs included in the model; w is the weight matrix that stores the penalty factors for SNPs. The α parameter is set to 0.5 and λ is estimated via cross-validation (CV).

We could apply the EpiXcan pipeline to train prediction models in a tissue-specific fashion from genetics-of-gene-expression panels, such as GTEx and STARNET (see section 1.4), and use these models to predict gene expression from individual-level genotype data. Finally, we can perform association analyses on predicted expression using regular models, such as logistic regression for categorical traits, and linear regression for quantitative traits.

1.7.2 Summary-based predictors

In many cases, the individual-level genotype data are not available, but the GWAS summary data are accessible via general collective resources, such as the GWAS catalog, or trait-specific

consortia, such as CARDIoGRAMplusC4D for CAD and GLGG for lipid traits[71]. In this case, the GWAS summary-based predictors become valuable. The tool Summary PrediXcan (S-PrediXcan) was derived from PrediXcan by the same research group[27]. S-PrediXcan uses prediction models trained by PrediXcan, takes summary data as input, and estimates the Z-score (Wald statistic) of the association between predicted gene expression and a phenotype. The Z-score for gene g is estimated as:

$$Z_g \approx \sum_i^m w_{ij} \frac{\widehat{\sigma_i}}{\widehat{\sigma_g}} \frac{\widehat{\beta_i}}{se(\widehat{\beta_i})}$$

In this formula, *m* is the number of cis-SNPs included in a gene's prediction model, w_{ij} is the weight of *i*-th SNP in the expression prediction model, $\widehat{\sigma_i}$ is the estimated variance of *i*-th SNP, $\widehat{\sigma_g}$ is the estimated variance of gene g, $\widehat{\beta_i}$ is the estimated effect size of *i*-th SNP from GWAS summary data, and $se(\widehat{\beta_i})$ is the standard error of *i*-th variant in GWAS summary data. The weights of cis-SNPs on gene expression derived from EpiXcan can also be applied.

FUSION is another summary-based TWAS tool[72]. Different from the elastic-net(ENet) model used by PrediXcan or EpiXcan, FUSION uses the Bayesian sparse linear mixed model (BSLMM)[73] in estimating weights for cis-SNPs. The FUSION package also provides other weights-estimating models including BLUP, ENet, and top SNPs. A comparison study has shown that FUSION and S-PrediXcan are consistent and complementary to each other[74]. Different from most TWAS tools focusing on cis-SNPs, Bayesian genome-wide TWAS (BGW-TWAS) leverages both cis- and trans-eQTL information for a TWAS[75]. With the accumulation of knowledge about functional genomics, the strategy of integrating multi-omics is *en vogue*. MOSTWAS is the recently published multi-omics strategy for TWAS analyses[76].

1.7.3 Opportunities and challenges of TWAS

TWAS, on the one hand, is a promising approach to prioritizing causal genes at GWAS loci in a tissue-specific fashion, on the other hand, the method can identify novel risk genes residing outside of GWAS risk loci. For instance, Wu prioritized 48 risk genes for breast cancer by performing TWAS analyses on 229,000 participants, of which 14 genes were independent of any reported GWAS loci[77]. Gene silence experiments on novel genes identified 11 novels that had effects on cell proliferation and/or colony-forming efficiency. In another work, the gene expression imputations across multiple brain regions in over one million participants identified 67 non-MHC (major histocompatibility complex) risk genes for schizophrenia, of which 14 did not fall within previous GWAS loci[78].

Despite the great contribution of providing insights into complex diseases, TWAS has certain limitations. TWAS combines eQTL reference panels with large-scale genotype data to test associations between genes and diseases. The pleiotropy features of eQTLs, and LD associations among eQTLs often result in multiple gene hits at the same locus[68]. In some cases, multiple genes significant at the same locus are due to co-expression with causal genes or co-regulated by the same set of eQTLs or those in high LD. Therefore, further validation is needed to check the causality. Another big challenge of TWAS is how to select appropriate tissues for testing association with diseases. Most candidate causal genes drop out after switching to tissue with has no clear mechanistic relationship to the trait due to the lack of sufficient expression or sufficiently heritable expression[68]. For TWAS analysis, the eQTL reference panels are critical in building prediction models. The differences in sample collections, expression scaling techniques, disease status of panel samples, etc., could create consistent outcomes of TWAS but might also be complementary to each other. Therefore,

combing TWAS results from different reference panels could provide richer insights into diseases[17].

1.8 Polygenic risk score (PRS)

1.8.1 Definition and analysis pipeline of PRS

A polygenic risk score (PRS) is a genetic estimate of an individual's liability to a trait or disease, calculated from a set of independent variants usually based on large-scale GWAS data[79]. For each individual, PRS is a sum of the number of risk alleles at each variant (0, 1, 2), which could be unweighted, or weighted by its effect size estimated derived from GWAS[80].

The PRS analysis process is shown in Figure 7[79]. PRS can be characterized by the use of base and target data. The risk loci and betas or odds ratio (OR) weights are retrieved from base data and then applied to calculate PRS for target data. The QC for both base and target data sets are similar to standards of classic GWAS. In addition, the "QC checklist" specific to PRS analysis has to be outlined: i) using base GWAS data with heritability $h^2>0.05$; ii) specifying effect and non-effect alleles from base GWAS data; iii) using target data with sample size of ≥100 individuals (or effective sample sizes>100 for case/control data) if PRS would be applied for association test[79]. There are two main options for approximating PRS. The classic method is called the clumping + thresholding ('C+T') method, which clumps SNPs passing the GWAS significance threshold (see section 1.2) so that SNPs retained are largely independent of each other. Another method is beta shrinkage which all SNPs are included, accounting for the LD between them. Many tools have been developed for calculating PRS, like PLINK[81], PRSice-2[82], LDpred[83], lassosum[84], etc.

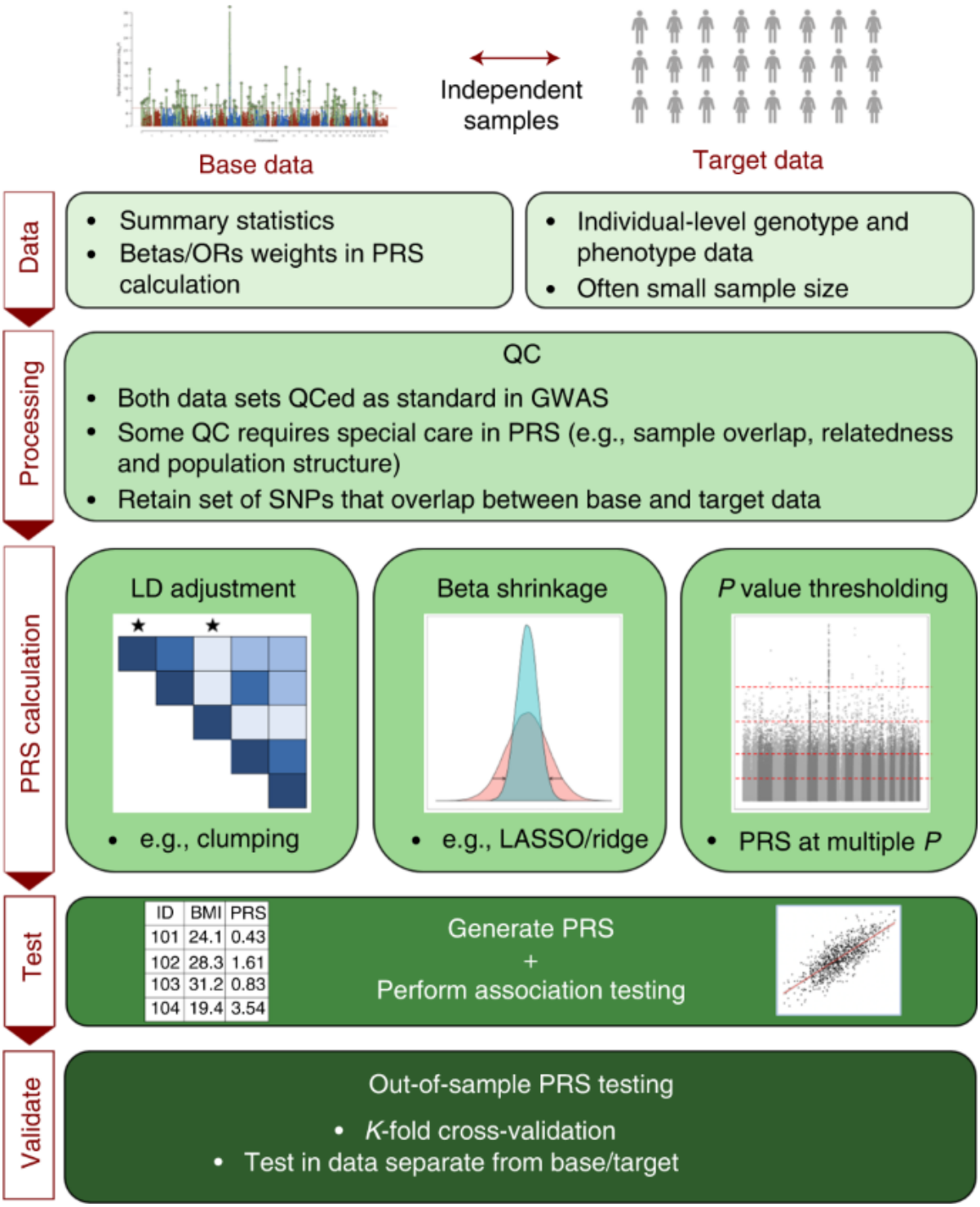


Figure 7 The PRS analysis process. The figure was adapted from Choi[79].

1.8.2 Application of PRS

In basic science, PRS has been used to evaluate associations between phenotypes and to elucidate risk factors that may play a mediating role along the causal pathway to disease[85].

The MR method has the advantage of featuring causal relationships among traits (see section 1.3.3). Therefore, the combination of PRS and MR analysis can improve detection rates for causal relationships which can be particularly useful when evaluating associations between genetic liability for a given trait and hundreds of diverse health outcomes[85]. We named this 'PRS+MR' strategy. Zeng et al. detected a negative association between PRS of education attainment and CAD risk, then they used the MR method to deduct the causal pathway of how genetic education attainment reduces CAD risk through its impacts on a series of risk factors[45]. This strategy was also applied to one of my PhD works which studied the genetic association between intelligence score and CAD risk (see section 3.1)[86].

In the clinical setting, PRS can be implemented in three key stages (Figure 8)[87]. First, PRS contributes to risk stratification in an apparently healthy population by screening for high-risk groups followed by intensified strategies for disease prevention and early intervention. Second, PRS could be used in clinical diagnosis when people are in the early phase of diseases without significant clinical diagnostic signs. Third, it is possible that in the future, PRS could contribute to personalized medicine and outcome prediction.

There is a long way to go before PRS could be massively applied in clinical settings. PRS itself is not strong enough in the prediction of disease incidence. In CAD, PRS didn't outperform in predicting subsequent CAD events as compared to other clinical risk predictors. But when both were combined, it was more accurate than either PRS or clinical risk predictors alone[87]. Studies also showed that the combination of PRS in subjects affected by monogenic variants with low to moderate penetrance could increase the accuracy of risk prediction. The average probability of CAD by the age of 75 years old subjects was found to increase from 13% in noncarriers of familial hypercholesterolemia (FH) variants to 41% in carriers of FH variants[88]. Likewise, in carriers of FH variants, a substantial gradient of risk was observed

depending on the PRS. Similar trends were also observed in breast cancer[88] and prostate cancer[89].

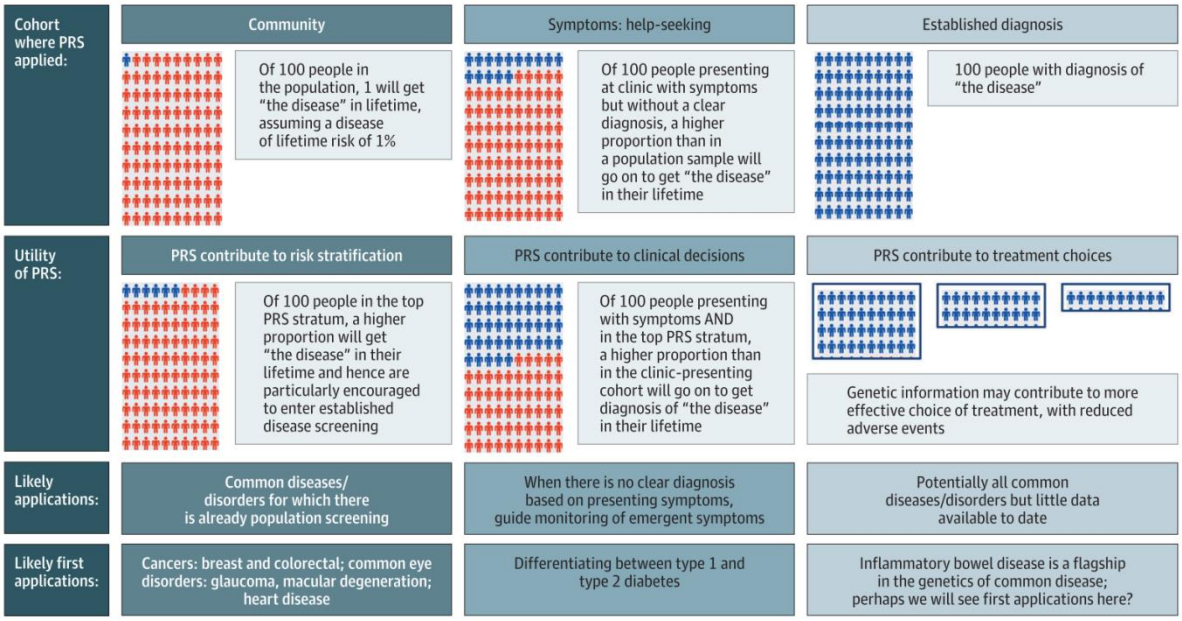


Figure 8 An overview of the population cohorts where polygenic risk scores could be applied. The figure was adapted from Wray[87].

1.8.3 Further discussion about PRS

Ethnicity

By far, most GWAS studies were conducted on individuals of European decedent, and genomic research of the non-European population is significantly underrepresented. More and more scientists put their attention to trans-ethnic GWASs because of genetic diversity among ethnicities[49, 90]. Along with increased identification of ethnicity-specific and trans-ethnic loci, a question about PRS study was raised: which is better, ethnicity-specific or trans-ethnic PRS? The latest work about CAD GWAS identified 8 novel Japanese-specific risk loci as well as 35 novel trans-ethnic loci[90]. This work also suggested that PRS derived from trans-ethnic

loci outperformed PRS derived either from Japanese or European risk loci. However, further work is needed to confirm this phenomenon.

Multi-PRS strategy

In classic studies, the PRS is estimated based on risk loci associated with the target trait. Inspired by the pleiotropy feature of genetic variants and genetic correlations between complex traits, the multiple polygenic risk score (multiPRS) approach improved prediction performance by using the joint predictive power of multiple polygenic scores in one regression model[91, 92]. For instance, the formula of multiPRS for CAD could be represented as

$$multiPRS_{CAD} = \beta_1 PRS_{CAD} + \beta_2 PRS_{Diabetes} + \beta_3 PRS_{Hypertension} + \beta_3 PRS_{LDL-C} + \cdots covrs$$

, in which $multiPRS_{CAD}$ is a weighted combination of CAD and other risk traits' PRS. The mulitPRS approach may be useful in investigating developmental, multivariate and gene-environment interplay issues, stratifying individuals according to the risk of conditions, and eventually, improving performance in personalized medicine.

Polygenic resilience score

Polygenic resilience is a reverse concept of PRS. It studies genetic variants that promote resistance to disease by reducing the penetrance of risk loci, wherein resilience and risk loci operate orthogonally to each other[93]. Currently, there is only one study about polygenic resilience score conducted by Hess et al[93]. Based on schizophrenia PRS, they stratified samples into different risk groups. The controls from high PRS groups were defined as high resistance controls. From the high-risk group, they retrieved genetic resilience variants by performing GWAS analysis. The works suggested that positive correlation between polygenic resilience score and PRS in the whole cohort, but no correlation in case samples. The

correlation coefficient reached the highest in controls of the high PRS group which verified the hypothesis of the orthogonal relation between genetic risk and resilience. Studying polygenic resilience factors could give us extra insights into why participants in the high PRS group are resilient to developing disease and might promote a more sophisticated PRS model with better performance than the current one.

1.9 Genetics of CAD

1.9.1 GWAS of CAD

Decades of GWASs have generated wealthy knowledge of genetic factors contributing to CAD etiology. Due to the endeavor of large CAD consortia, as well as national and international collaboration, 321 genome-wide significant loci have been associated with CAD (Figure 8)[94]. The sample size and diversity of participants are still increasing in multiple biobanks, such as CARDIoGRAMplusC4D, UKB, Japanese biobank, Million Hearts GWAS, Million Veteran Program, and All of Us Research Program. A combination of these biobanks would increase the possibility of identifying novel risk loci and studying ethnicity-specific risk lock. Current knowledge of CAD GWAS brings us to the post-GWAS era to i) elucidate the disease-associated mechanisms underlying CAD loci, ii) prioritize potential causal genes and novel drug targets for the disease, and iii) harness CAD genetic variations as the age-independent biomarkers for risk stratification, disease prevention, and personalized medicine.

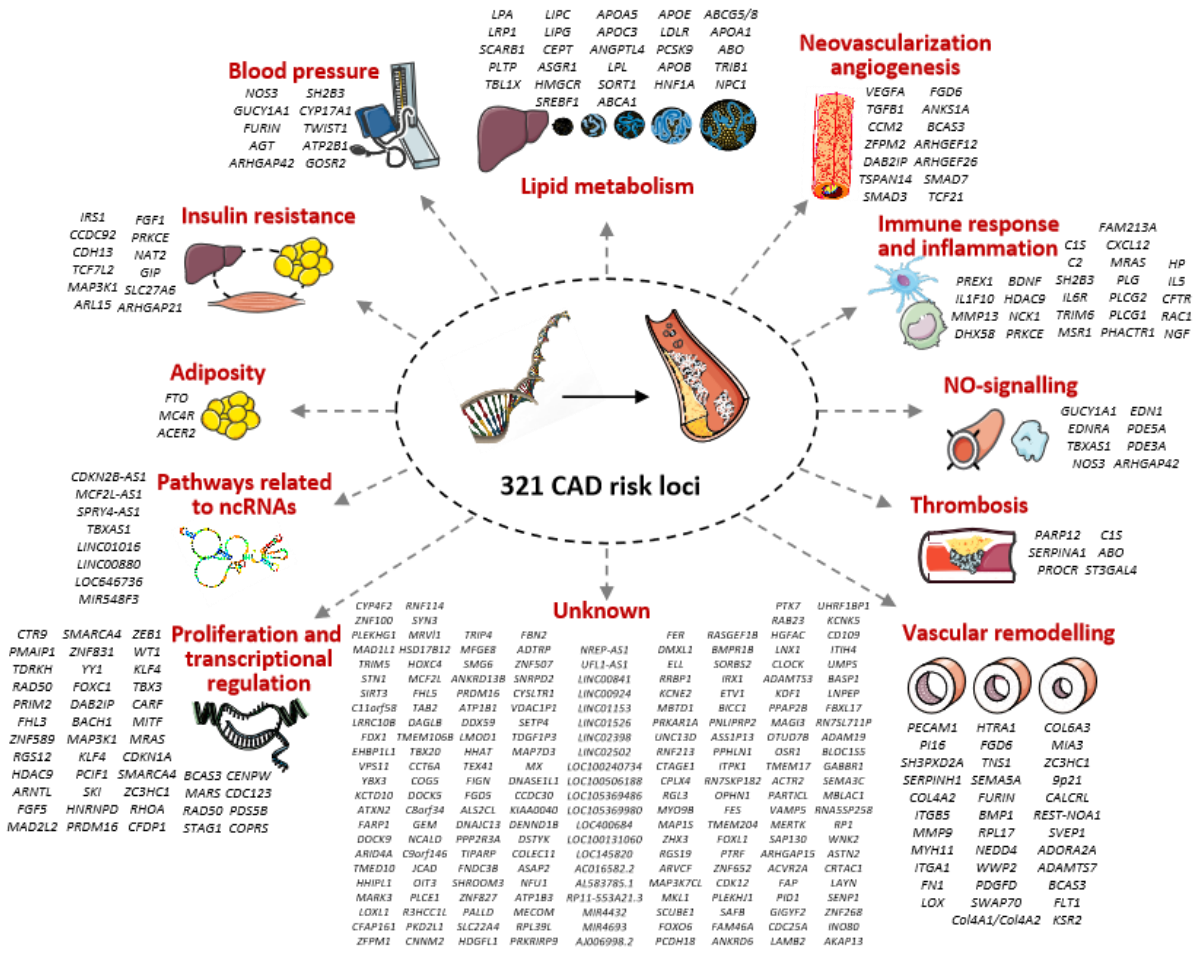


Figure 9 Genes mapped to 321 CAD risk loci and related pathophysiological pathways of atherosclerosis. The figure was adapted from Chen[94].

1.9.2 Prioritization of CAD causal genes by in-silico methods

The *in-silico* methods of prioritizing risk genes have been discussed in section 1.3. Strategies of pinpointing causal genes for CAD include but are not limited to i) vicinity of genes to risk loci; ii) eQTL mapping in CAD-relevant tissues according to reference panels like GTEx and STARNET; iii) monogenic mutations on candidate genes associated with CAD or its risk factors; iv) fine-mapping with functional annotation; v) colocalization between GWAS signal and genes' eQTL or other quantitative signals; vi) phenome-wide screening for pleiotropy of risk loci; vii) cross-mapping with consortia of mouse models like Hybrid Mouse Diversity Panel (HMDP)[95], the International Mouse Phenotyping Consortium (IMPC)[96] or Mouse

Genome Informatics database (MGI)[97]; viii) integrative genomic analysis (IGA) to combine results from multiple in-silico methods, which has emerged as a powerful strategy for identifying causal genes[98]. Moreover, the fast accumulation of single-cell data and metabolomics data would greatly benefit causal gene selections and pathogenic mechanism study[99, 100]. However, TWAS of CAD, which could systematically identify tissue-specific risk genes for CAD has not been performed yet. This forms the rationale of my PhD research on CAD TWAS.

II Methods

2.1 Preprocessing of 11 GWAS cohorts

Genotype data of ten CARDIoGRAMplusC4D[86, 101–109] cohorts and UKB[47] were used in both projects building the PhD thesis (Table 1).

Table 1 Overview of eleven individual-level genotype cohorts.

Study	Array	CAD case		Control	
		N_case_female(%N_cases)	N_case	N_control_female(%N_controls)	N_control
GerMIFSI	Affymetrix Mapping 500K Array Set	207(33.3)	622	795(51.3)	1551
GerMIFSII	Affymetrix Genome-Wide Human SNP Array 6.0	244(20.5)	1192	604(48.1)	1256
GerMIFSIII	Affymetrix Genome-Wide Human SNP Array 5.0/6.0	212(20.1)	1055	696(48.3)	1441
GerMIFSIV	Affymetrix Genome-Wide Human SNP Array 6.0	336(35.2)	954	697(61.4)	1136
GerMIFSV	Illumina HumanOmniExpress/Omniuni_2.5	593(24.3)	2437	827(52.5)	1574
GerMIFSVI	Illumina PsychChip_v1-1	492(30)	1639	609(51.3)	1186
GerMIFSVII	Infinium Global Screening Array-24	1031(33.7)	3062	1886(54.5)	3462
WTCCC	Affymetrix Genome-Wide Human SNP Array 6.0	395(20.8)	1900	1481(50.9)	2911

Cardiogenics	Illumina Human660W-Quad	49(12.8)	382	239(59.2)	404
MIGen	Affymetrix Mapping 500K Array Set	646(22.3)	2901	733(24.3)	3018
UKB	UK BiLEVE Axiom array n~=50,000 UK Biobank Axiom array n~=450,000	4465 (22.0)	20310	231869(57.0)	406996

2.1.1 Ten CARDIoGRAMplusC4D cohorts

Individua-level genotyping data of 17,687 cases and 17,854 controls were collected from ten CARDIoGRAMplusC4D cohorts including the German Myocardial Infarction Family Studies (GerMIFS) I-VII[86, 101–106], Wellcome Trust Case Control Consortium (WTCCC)[107], LURIC[108] study and Myocardial Infarction Genetics Consortium (MIGen)[109]. All samples were of European descent, mostly from Germany or UK. The subjects building the MIGen cohort were from several European countries and the United States.

All studies went through the following quality filtering steps (as mentioned in section 2) using plink (v1.9)[13] before imputation: individual-level calling rate≥0.98, SNP-level calling rate>0.95, MAF>0.01, sex consistency between the reported and the genotype-derived, deviation from HWE *P*>1e-5, Identity By Descent (IBD)<0.125 (individuals were distant to each other more than the third generation), heterozygosity rate within mean ± 3*SD. After quality filtering of genotype level, all studies except GerMIFS VII were imputed according to the reference of 1000 Genomes Phase I integrated variant (v3)[10] using SHAPEIT2[110] and IMPUTE2[111]. For GerMIFS VII, the reference data was HRC[11]. All SNPs were mapped to NCBI GRCh37/hg19. After genotype imputation, we conducted quality filtering again using the following parameters: SNP-level calling rate>0.98, MAF>0.01, and deviation from HWE *P*>1e-5.

2.1.2 Genotype data of UKB

The UKB is globally the largest biobank with deep phenotypic and genomic data (see section 1.5)[47]. We obtained imputed genotype data from 3rd release of UKB. The data were imputed using a combination of two reference panels. The first panel is the HRC reference panel[11]. The second is a merged reference panel of the UK10K haplotype panel and the 1000 Genomes references. If variants were imputed from both reference panels, the HRC imputation result

was retained. A series of QC were done on the imputed data: MAF>0.00001, imputation info>0.4, SNP-level calling rate>0.95, individual-level calling rate≥0.98, sex consistency, kinship coefficient<0.088, and deviation from HWE *P*>1e-5.

All phenotypes were defined by either self-reported, hospital episode, and/or death registry data. The definition of CAD encompassed individuals with fatal or normal myocardial infarction (MI), percutaneous transluminal coronary angioplasty (PTCA), or coronary artery bypass grafting (GABAD). MI was defined as hospital admission or cause of death due to ICD10 I21-I24, I25.2, ICD9 410-412, and self-reported 1075. PTCA, CABG, and triple heart bypass were defined as hospital admission or cause of death due to OPCS-4 K40-K46, K49, K50.1, K75, and self-reported 1070, 1095, 1523. Finally, we got 20,310 hard CAD cases and randomly selected 25,000 non-CAD samples as controls.

2.2 Methods used in the Intelligence-CAD project

In this work, we firstly constructed a genetic intelligence score (gIQ) based on 242 SNPs independently associated with intelligence[112] for participants from ten CARDIoGRAMplusC4D[86, 101–109] cohorts and UKB[47]. The 'C+T' method (see section 1.8.1) of PRS estimating was applied to estimate gIQ. In short, gIQ is a sum of the weighted dosage of effect alleles of 242 independent intelligence variants. As to missing variants in the genotype data, we replaced them with reference allele frequencies.

We tested the association between gIQ and CAD risk in eleven cohorts using logistic regression models. For all cohorts except UKB, two principal components were added to the regression model to adjust the bias of population stratification. As the genotype data of UKB was scaled from two array platforms and has a relatively more complicated ethnic background, the top five principal components and array platform were added to the regression model.

In UKB, we also studied associations between gIQ and serial risk factors of CAD including smoking, hypertension, BMI, T2D, HDL-C, and LDL-C. Logistic regression models were applied for binary traits, and linear regression models were applied for quantitative traits. All regression models were adjusted by the top five principal components and array platforms. Next, the significant risk factors were applied as adjustments of regression models between gIQ and CAD to study their mediatory roles.

Because of genetic overlaps between intelligence and educational attainment, the effects of intelligence on CAD and its risk factors might be false positive. We defined the direct effect of intelligence as the effect of intelligence that was not mediated by educational attainment. Lee et al. reported 1271 independent SNPs associated with educational attainment through a meta-analysis in 1.1 million persons[113]. Seven SNPs associated with both intelligence and educational attainment were excluded during recalculating gIQ. Then same regression analysis was performed to study the direct effects of intelligence.

To further study causal pathways from intelligence/education to CAD risk, a multivariable two-sample MR analysis was carried out. The GWAS summary data of CAD and its risk factors, educational attainment were acquired from CARDIoGRAMplusC4D (CAD)[104], GIANT (BMI)[114], TAG (smoking)[115], GLGC (HDL-C, LDL-C)[71], SSGAC (educational attainment)[113], and DIAGRAM (T2D)[116]. Three MR methods were used including IVW, MR-egger, and weighted median. Lastly, MR sensitivity analyses were performed for intelligence and educational attainment respectively. SNPs moderately associated with CAD and risk factors ($P<0.001$) were removed from intelligence SNP and education SNPs respectively.

2.3 Methods used in the TWAS project

As introduced in section 1.7.1, the first step was to train prediction models for risk tissues of CAD. We adopted the existing expression prediction models trained using the EpiXcan pipeline by Zhang et al., including models of AOR, MAM, LIV, SF, VAF, BLD, and SKLM based on the STARNET panel[17], and models of AOR, LIV, BLD, SF, VAF, and SKLM based on the GTEx panel[18]. We trained prediction models for another two risk tissues of CAD, arterial wall coronary (COR) and tibial artery (TIB) datasets which are only available in the GTEx panel (V7), using the EpiXcan pipeline. All individuals used for prediction models were restricted to European descent. First, genetic variants were filtered out if they matched one of the QC parameters: calling rate<0.95, MAF<0.01, and HWE<1e-6. Second, for expression data, we did sample-level quantile normalization and gene-level inverse quantile normalization using preprocess codes of the PredicDB pipeline. Third, we calculated SNP priors using qtlBHM that jointly analyzed epigenome annotations of aorta derived from REMC[70]. Lastly, the SNP priors, genotype data, and expression data were jointly applied to 10-fold cross-validated WENet to train predicting models by deploying the EpiXcan pipeline[69]. The predictors were filtered by cross-validated prediction R^2>0.01.

Next, we applied prediction models of nine risk tissues to impute GReX from individual genotype data of ten CARDIoGRAMplusC4D cohorts[86, 101–109] and UKB[47], which, in total, having of 37,997 cases and 42,854 controls. For each tissue, we performed association tests between GReX and CAD using the logistic regression model in 11 genotype cohorts and performed meta-analysis for all cohorts to get summarized TWAS statistics. In total, we identified 114 genes representing 193 gene-tissue pairs thresholding Bonferroni-corrected significance.

Then, we did a series of *in-silico* analyses to check the plausibility of the TWAS gene list. i) We compared the genomic position between TWAS genes and GWAS loci using MAGMA[117]. ii) We compared TWAS genes with gene lists resulting from colocalization analyses. In this part, the tool coloc[23, 27] integrated eQTL data from either GTEx or STARNET and GWAS summary data from CARDIoGRAMplusC4D[65], providing significant gene-tissue pairs which posterior probability of hypothesis 4, PP4≥0.55 (section 1.3.1). iii) To check the biological function and pathogenicity of TWAS genes, we did pathway enrichment analysis using ClueGO[118] and disease enrichment analysis based on DisGeNET[119]. iv) We also performed gene-based rare variant association analysis using WES data from UKB on 200,632 participants. The damaging rare variants residing in TWAS genes were defined by one of five *in-silico* methods (LRT score, MutationTaster, PolyPhen-2 HumDiv, PolyPhen-2 HumVar, and SIFT) by the annotation resource dbNSFP 4.1a[59] using the VEP tool[61]. For gene-based tests, we used Fisher's exact test for binary traits and the linear regression model for quantitative traits.

Finally, our attention was focused on 18 genes that resided outside known GWAS loci, which we called "novel". To study the susceptibility of novel genes, we did *in-silico* analyses and *in-vitro* validation. We studied the genetic association between novel genes and a series of lipids traits in human genotype data from UKB. We also studied their association with lipid traits in atherosclerosis mouse models from the Hybrid Mouse Diversity Panel (HMDP). Both human genotype data and mouse expression data suggested connections between novel genes and lipid traits. Of these novel genes, *KPTN* and *RGSI9,* which were rarely studied before, gave significant signals in liver tissue by TWAS analysis. So, we finally carried out knockdown experiments in human hepatocytes using dual CRISPR strategy for *KPTN* and *RGS19*. Cells for measurement of secretion of triglycerides, cholesterol, and APOB100 were cultured for 16

hours in serum-free medium. Medium triglycerides and cholesterol were enriched five times by vacuum centrifuge and measured with colorimetric kits, triglycerides (cobas) and CHOL2 (cobas), respectively.

III Discussion

3.1 Genetic association between Intelligence and CAD

This work applied 'PRS+MR' strategy which firstly use PRS to filter genetic association between traits, then apply MR method to study casual effects on associated traits (see section 1.8.2).

We firstly tested association of gIQ with CAD and its risk factors. In this part, we observed that one standard deviation (SD) increase of gIQ was related to a 5% decrease of CAD risk (odds ratio [OR] of 0.95; 95% confidence interval [CI] 0.93 to 0.98; *P*=4.93e–5), which was validated in UKB (OR=0.97; 95% CI 0.96 to 0.99; *P*=6.4e–4). In UKB data, we also observed significant inverse correlations between gIQ and risk factors of CAD including BMI (OR=0.899; 95% CI 0.886 to 0.911; *P*=5.4e-49), smoking (OR=0.981; 95% CI 0.975 to 0.987; *P*=8.3e-10), T2D (OR=0.966; 95% CI 0.951 to 0.980; *P*=4.1e-6), hypertension (OR=0.987; 95% CI 0.981 to 0.993; *P*=3.8e-5), and a positive correlation with HDL-C (OR=1.007; 95%CI 1.006 to 1.008; *P*=1.3e-29). The associations of gIQ with CAD and its risk factors were largely attenuated after the adjustment of measured intelligence and educational attainment. Same phenomena happened between gIQ and CAD risk with the adjustment of risk factors significantly associated with gIQ. These findings suggested intermediate role of risk factors, measured intelligence, and educational attainment between gIQ and CAD.

Next, we applied two-sample MR analyses to depict casual pathway from genetic intelligence to CAD risk. One SD increase of intelligence resulted in decrease of CAD risk by

25% (OR=0.75; 95% CI 0.69 to 0.81; *P*<1e-10), decrease of BMI by 0.1 kg/m2 (95% CI -0.16 to -0.14; *P*=1.02e-3), decrease of T2D risk by 15% (OR=85; 95% CI 0.77 to 0.95). One SD increase in the education years resulted in decrease of risk of CAD by 38% (OR=0.62; 95% CI 0.58 to 0.66; *P*<1e-10), decrease of BMI by 0.32 kg/m2 (95%CI -0.37 to -0.27; *P*<1e–10), increase of HDL-C by 0.19 mmol/L(95% CI 0.14 to 0.25; *P*<1e-10), decrease of the risk of smoking by 43% (OR=0.57; 95%CI 0.501 to 0.642; *P*<1e-10), and decrease of T2D risk by 47% (OR= 0.53; 95%CI 0.49 to 0.57; *P*<1e-10). The effects of educational attainment on CAD and its risk factors displayed the same direction as intelligence but were stronger in magnitude.

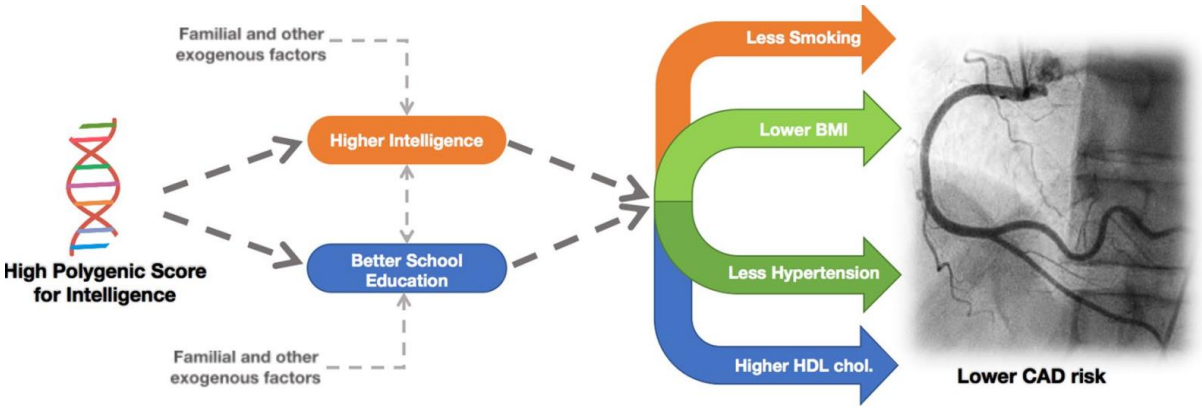


Figure 10 Pathways from Pathway from higher gIQ to lower risk of CAD.

In conclusion, using genetic approaches, we depicted a pathway from gIQ to CAD risk The higher gIQ is associated with the higher measured intelligence and longer educational attainment, both of which appear to reduce the prevalence of risk factors of CAD including BMI, smoking, T2D, and hypertension, and increase HLD-C, which in concert subsequently reduce the prevalence of CAD (Figure 10). Moreover, the effects of educational attainment on risk factors and CAD appear to be stronger than the effects of intelligence. Thus, repetitive campaigns throughout schooling may be worthwhile for preventive reasons as they may ameliorate the association between gIQ and unhealthy lifestyle.

Author contributions

Heribert Schunkert and **Ling Li** designed the study and wrote the manuscript. Shichao Pang, Lingyao Zeng, and Ulrich Güldener provided technical support and gave conceptual advice.

3.2 TWAS for CAD

Our work is the first to systematically identify susceptibility genes of CAD in tissue-specific fashion using TWAS method. The main findings of this work were summarized in Figure 11. We first trained gene expression prediction models for nine CAD-relevant tissues using EpiXcan based on two largest genetics-of-gene-expression panels, STARNET and GTEx. We next explored these prediction models on individual level genotype data of 37,997 CAD cases and 42,854 controls. Our TWAS identified 114 CAD risk genes to be differentially expressed by genetic means. Of these, 96 genes were resided within ±1Mb region of previously identified by GWAS loci (a sort of positive control) and 18 were novel. For 114 TWAS genes, we performed stepwise analyses to prove their plausibility, biological function, and pathogenicity to CAD, including analyses for colocalization, damaging mutation, pathway enrichment, phenome-wide associations with human data and expression-traits correlations using mouse data. Finally, we focused on two novel genes, *RSG19* and *KPTN*, and conducted CRISPR/Cas9-based knockdown experiments for them in human hepatocytes. We observed reduced secretion of APOB100 and lipids in the cell culture medium, i.e., a functional explanation for the association findings.

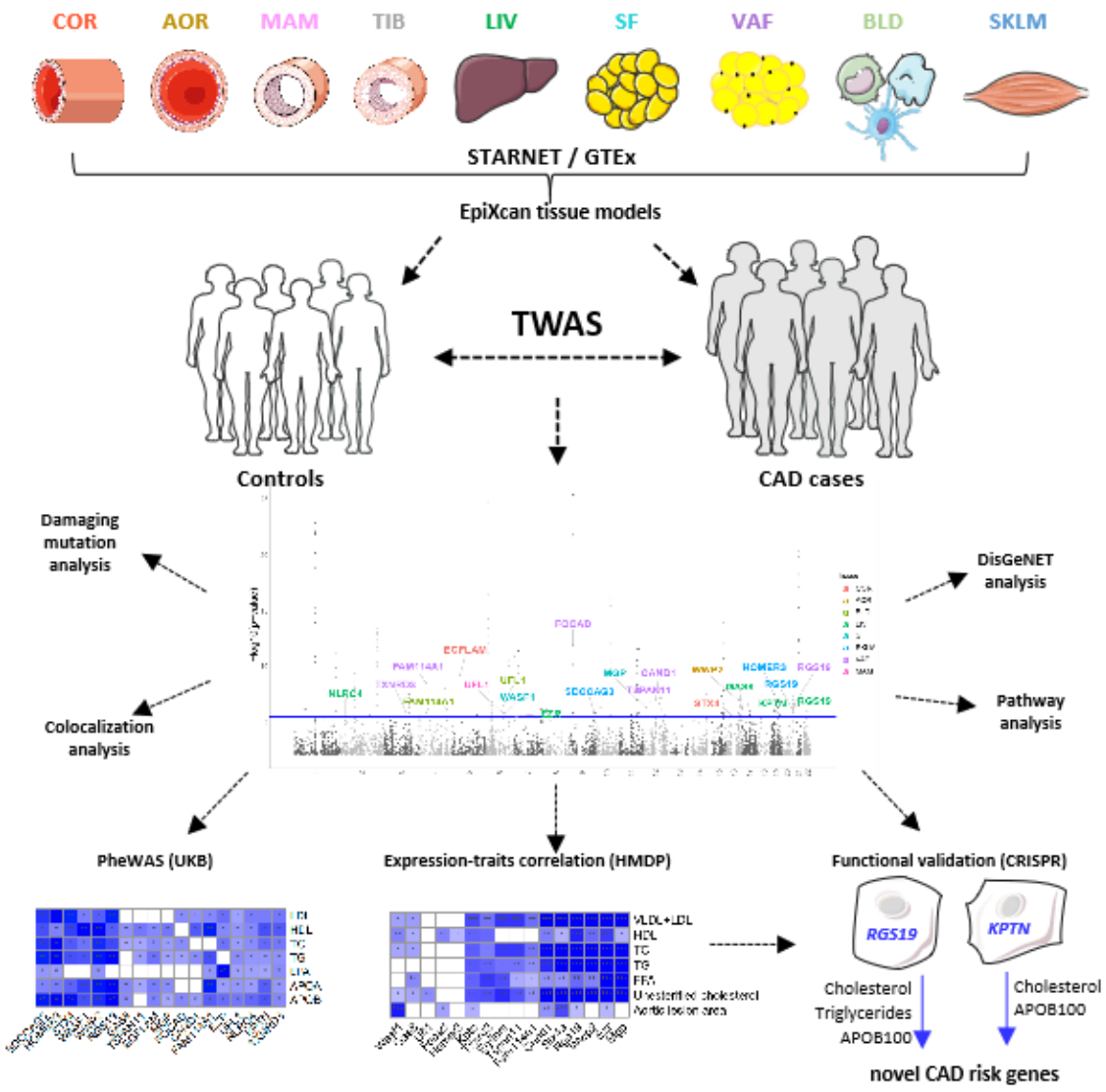


Figure 11 Schematic illustration of CAD TWAS.

This work discovered 18 novel genes to be associated with CAD, and functionally evaluated 96 genes within CAD GWAS loci, for example by indicating their tissue(s) of action. The downstream analyses of these genes revealed their intermediate cardiometabolic phenotypes bridging gene variants with their effects on CAD. Our result provides a substantial step towards prioritization of genes at respective GWAS loci as well as their tissues of actions. In this respect, 46 genes identified by this TWAS are known for effects in pathophysiological pathways related to CAD, including lipid metabolism, inflammation, angiogenesis,

transcriptional regulation, cell proliferation, NO signaling, and high blood pressure, to name a few giving credibility to the association findings.

Our *in-silico* analyses on novel genes suggested they were associated with lipid traits in both human genotype data and mouse expression data. The hypothesis was furtherly confirmed by the latest lipid GWAS paper conducted by GLGLC consortium (Table 2)[120]. Two novel genes, *KPTN* and *RGS19*, were firstly confirmed to be associated with lipid metabolism from both *in-silico* and *in-vitro* data. So, we believe that our study on novel genes may provide novel insights into molecular etiology of CAD.

Table 2 8 novel genes were within 500kb of lipid risk loci

Gene	Lipid risk loci
HOMER3	EUR_TC_rs10423802;EUR_HDL_rs60570301;EUR_logTG_rs1966500;HIS_nonHDL_rs2238675;HIS_TC_rs150641967;rs376645231;EAS_logTG_rs58542926;EAS_nonHDL_rs58542926;EUR_LDL_rs58542926;EUR_logTG_rs58542926;EUR_nonHDL_rs58542926;EUR_TC_rs58542926;HIS_LDL_rs58542926;SAS_LDL_rs58542926;SAS_nonHDL_rs58542926;SAS_TC_rs58542926;SAS_logTG_rs8107974;EAS_TC_rs10401969
KPTN	EUR_HDL_rs12609461;EUR_logTG_rs10408163;EUR_nonHDL_rs12461923;EUR_HDL_rs3112494;EUR_logTG_rs62129968
NLRC4	EUR_TC_rs62142080
RGS19	EUR_TC_rs35201382;EUR_LDL_rs6090040;EUR_nonHDL_20_62692060_C_A;EUR_HDL_rs8126001;EUR_logTG_rs8126001;EUR_TC_rs73147887
SDCCAG3	EUR_nonHDL_rs3780190;EUR_TC_rs3780190;EUR_LDL_rs13301660
STX4	EUR_TC_rs1870293;EUR_LDL_rs35468353;EUR_nonHDL_rs73530203;EUR_logTG_rs7196161;EUR_HDL_rs41440449
TXNRD3	EUR_LDL_rs9862203;EUR_TC_rs9862203

WWP2	EUR_TC_rs76116020;EUR_HDL_rs141904578;EUR_logTG_rs6499240;EUR_LDL_rs181501802;EUR_TC_rs181501802;EUR_nonHDL_rs62049432;AFR_LDL_rs200535533;rs374818812

*EUR, European; EAS, east Asian; SAS, South Asian; AFR, African.

We must admit that our study has limitation embedded in TWAS mythology. Since TWAS are strongly dependent on the reference panel linking genetic signatures with gene expression, it had to be expected that STARNET- and GTEx-based predictive models display differences in gene-CAD associations. The difference may be due to different sample sizes used for training predictions models, different disease states (subjects with and without CAD), intravital or postmortem sample collection, as well as different transcript abundance and genotype coverage leading to differences in expression associated SNPs in our reference panels. Even so, a fair consistency of TWAS results between STARNET- and GTEx-based models gave us rationality of combing results derived from both panels to increase the power of capturing risk genes. Second, although TWAS facilitates candidate risk gene prioritization, LD-link between SNPs, co-regulation or co-expression in cis at a given locus limits the precise determination of the culprit gene. Indeed, at 12 loci we observed signals for three or more TWAS genes. This strengthens the importance of integration of other genetic analyses to improve risk gene prioritization. In our work, a series of stepwise analyses were performed on TWAS gene list to furtherly study their disease-causing mechanism, like damaging variant association, pathway enrichments, genetic association with other phenotypes and expression-traits association statistics. Lastly, all findings by *in-silico* methods have to be furtherly validated and functionally explained by in-vitro or in-vivo methods. Our prove-of-concept experiment on *KPTN* and *RGS19* in human hepatocytes validated our hypothesis about their associations with lipid traits. But extra efforts are necessary to clearly depict the molecular mechanisms.

In summary, our TWAS study based on two genetics-of-gene-expression panels created a set of gene-centered and tissue-annotated associations for CAD, providing insightful guidance for further biological investigation and therapeutic development.

Author contributions

Heribert Schunkert**, Ling Li**, and Zhifen Chen designed the study and wrote the manuscript. Ling Li ran the bioinformatic analyses. Zhifen Chen, Shuangyue Li, and Andrea Steiner performed the wet lab experiments. Julien Gagneur, Moritz von Scheidt, Ulrich Güldene1, Simon Koplev, Angela Ma, Ke Hao, Calvin Pan, Aldons J. Lusis, Shichao Pang, Thorsten Kessler, Raili Ermel, Katyayani Sukhavasi, Arno Ruusalepp, Jeanette Erdman, Jason C. Kovacic, Johan L.M. Björkegren provided research data, technical support and gave conceptual advice.

IV Conclusion and outlook

I conducted two original research on post-GWAS studies during my graduate training, including a PRS study and the CAD TWAS. In the PRS study, I used genetic methods to firstly verify the epidemiology phenomenon of reverse association between intelligence and CAD risk and depicted a pathway between them. In the TWAS study, we pinpointed risk genes of CAD as well as their action tissues which empowered our understanding of the molecular mechanisms of CAD. The positive experimental validations of two novel genes might lay the foundation of the therapeutic development of CAD. The two projects inspired me to explore potential transcriptome changes due to the polygenic risk in a tissue-specific fashion.

Emerging wave of TWAS studies focuses on investigating the role of spicing variants and transcripts in CAD. The data sets of tissue splicing variants (sQTL) and transcripts are

available in several public resources, such as STARNET and GTEx. However, the lack of cellular level omics data is a disadvantage of the current TWAS analysis. The CAD risk linked to the cell-type specific function of a gene might be obscured due to the mix cellular profile in a tissue. The booming single-cell technology will soon tackle the current disadvantage and TWAS analysis will be performed in a cell type- or subtype-specific manner. Moreover, the concept of TWAS analysis will be applied to several other omics datasets, such as epigenomics, proteomics, and metabolomics, when tissue- or cell-specific data will be available in scale. The different layers of biological data could be integrated for a multi-omics TWAS analysis to uncover the interaction among the layers of cellular profiles. The increasing complexity in the data integration will undoubtedly challenge the statistic algorisms. In this scenario, machine learning (ML) will embark to unlock the myth of complex common diseases.

Finally, I'd like close my thesis using poetry from a Chinese famous poet Qu Yuan: Long, long had been my road and far, far was the journey; I would go up and down to seek my heart's desire (路漫漫其修远兮，我将上下而且求索).

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton A, Benjamin EJ, Benziger CP, Bonny A, Brauer M, Brodmann M, Cahill TJ, Carapetis JR, Catapano AL, Chugh S, Cooper LT, Coresh J, Criqui MH, DeCleene NK, Eagle KA, Emmons-Bell S, Feigin VL, Fernández-Sola J, Fowkes FGR, Gakidou E, Grundy SM, He FJ, Howard G, Hu F, Inker L, Karthikeyan G, Kassebaum NJ, Koroshetz WJ, Lavie C, Lloyd-Jones D, Lu HS, Mirijello A, Misganaw AT, Mokdad AH, Moran AE, Muntner P, Narula J, Neal B, Ntsekhe M, Oliveira GMM, Otto CM, Owolabi MO, Pratt M, Rajagopalan S, Reitsma MB, Ribeiro ALP, Rigotti NA, Rodgers A, Sable CA, Shakil SS, Sliwa K, Stark BA, Sundström J, Timpel P, Tleyjeh II, Valgimigli M, Vos T, Whelton PK, Yacoub M, Zuhlke LJ, Abbasi-Kangevari M, Abdi A, Abedi A, Aboyans V, Abrha WA, Abu-Gharbieh E, Abushouk AI, Acharya D, Adair T, Adebayo OM, Ademi Z, Advani SM, Afshari K, Afshin A, Agarwal G, Agasthi P, Ahmad S, Ahmadi S, Ahmed MB, Aji B, Akalu Y, Akande-Sholabi W, Aklilu A, Akunna CJ, Alahdab F, Al-Eyadhy A, Alhabib KF, Alif SM, Alipour V, Aljunid SM, Alla F, Almasi-Hashiani A, Almustanyir S, Al-Raddadi RM, Amegah AK, Amini S, Aminorroaya A, Amu H, Amugsi DA, Ancuceanu R, Anderlini D, Andrei T, Andrei CL, Ansari-Moghaddam A, Anteneh ZA, Antonazzo IC, Antony B, Anwer R, Appiah LT, Arabloo J, Ärnlöv J, Artanti KD, Ataro Z, Ausloos M, Avila-Burgos L, Awan AT, Awoke MA, Ayele HT, Ayza MA, Azari S, Darshan BB, Baheiraei N, Baig AA, Bakhtiari A, Banach M, Banik PC, Baptista EA, Barboza MA, Barua L, Basu S, Bedi N, Béjot Y, Bennett DA, Bensenor IM, Berman AE, Bezabih YM, Bhagavathula AS, Bhaskar S, Bhattacharyya K, Bijani A, Bikbov B, Birhanu MM, Boloor A, Brant LC, Brenner H, Briko NI, Butt ZA, dos Santos FLC, Cahill LE, Cahuana-Hurtado L, Cámera

LA, Campos-Nonato IR, Cantu-Brito C, Car J, Carrero JJ, Carvalho F, Castañeda-Orjuela CA, Catalá-López F, Cerin E, Charan J, Chattu VK, Chen S, Chin KL, Choi JYJ, Chu DT, Chung SC, Cirillo M, Coffey S, Conti S, Costa VM, Cundiff DK, Dadras O, Dagnew B, Dai X, Damasceno AAM, Dandona L, Dandona R, Davletov K, de la Cruz-Góngora V, de la Hoz FP, de Neve JW, Denova-Gutiérrez E, Molla MD, Derseh BT, Desai R, Deuschl G, Dharmaratne SD, Dhimal M, Dhungana RR, Dianatinasab M, Diaz D, Djalalinia S, Dokova K, Douiri A, Duncan BB, Duraes AR, Eagan AW, Ebtehaj S, Eftekhari A, Eftekharzadeh S, Ekholuenetale M, El Nahas N, Elgendy IY, Elhadi M, El-Jaafary SI, Esteghamati S, Etisso AE, Eyawo O, Fadhil I, Faraon EJA, Faris PS, Farwati M, Farzadfar F, Fernandes E, Prendes CF, Ferrara P, Filip I, Fischer F, Flood D, Fukumoto T, Gad MM, Gaidhane S, Ganji M, Garg J, Gebre AK, Gebregiorgis BG, Gebregzabiher KZ, Gebremeskel GG, Getacher L, Obsa AG, Ghajar A, Ghashghaee A, Ghith N, Giampaoli S, Gilani SA, Gill PS, Gillum RF, Glushkova E V., Gnedovskaya E V., Golechha M, Gonfa KB, Goudarzian AH, Goulart AC, Guadamuz JS, Guha A, Guo Y, Gupta R, Hachinski V, Hafezi-Nejad N, Haile TG, Hamadeh RR, Hamidi S, Hankey GJ, Hargono A, Hartono RK, Hashemian M, Hashi A, Hassan S, Hassen HY, Havmoeller RJ, Hay SI, Hayat K, Heidari G, Herteliu C, Holla R, Hosseini M, Hosseinzadeh M, Hostiuc M, Hostiuc S, Househ M, Huang J, Humayun A, Iavicoli I, Ibeneme CU, Ibitoye SE, Ilesanmi OS, Ilic IM, Ilic MD, Iqbal U, Irvani SSN, Islam SMS, Islam RM, Iso H, Iwagami M, Jain V, Javaheri T, Jayapal SK, Jayaram S, Jayawardena R, Jeemon P, Jha RP, Jonas JB, Jonnagaddala J, Joukar F, Jozwiak JJ, Jürisson M, Kabir A, Kahlon T, Kalani R, Kalhor R, Kamath A, Kamel I, Kandel H, Kandel A, Karch A, Kasa AS, Katoto PDMC, Kayode GA, Khader YS, Khammarnia M, Khan MS, Khan MN, Khan M, Khan EA, Khatab K, Kibria GMA, Kim YJ, Kim GR, Kimokoti RW, Kisa S, Kisa A, Kivimäki M, Kolte D, Koolivand A, Korshunov VA,

Laxminarayana SLK, Koyanagi A, Krishan K, Krishnamoorthy V, Defo BK, Bicer BK, Kulkarni V, Kumar GA, Kumar N, Kurmi OP, Kusuma D, Kwan GF, la Vecchia C, Lacey B, Lallukka T, Lan Q, Lasrado S, Lassi ZS, Lauriola P, Lawrence WR, Laxmaiah A, LeGrand KE, Li MC, Li B, Li S, Lim SS, Lim LL, Lin H, Lin Z, Lin RT, Liu X, Lopez AD, Lorkowski S, Lotufo PA, Lugo A, Nirmal KM, Madotto F, Mahmoudi M, Majeed A, Malekzadeh R, Malik AA, Mamun AA, Manafi N, Mansournia MA, Mantovani LG, Martini S, Mathur MR, Mazzaglia G, Mehata S, Mehndiratta MM, Meier T, Menezes RG, Meretoja A, Mestrovic T, Miazgowski B, Miazgowski T, Michalek IM, Miller TR, Mirrakhimov EM, Mirzaei H, Moazen B, Moghadaszadeh M, Mohammad Y, Mohammad DK, Mohammed S, Mohammed MA, Mokhayeri Y, Molokhia M, Montasir AA, Moradi G, Moradzadeh R, Moraga P, Morawska L, Velásquez IM, Morze J, Mubarik S, Muruet W, Musa KI, Nagarajan AJ, Nalini M, Nangia V, Naqvi AA, Swamy SN, Nascimento BR, Nayak VC, Nazari J, Nazarzadeh M, Negoi RI, Kandel SN, Nguyen HLT, Nixon MR, Norrving B, Noubiap JJ, Nouthe BE, Nowak C, Odukoya OO, Ogbo FA, Olagunju AT, Orru H, Ortiz A, Ostroff SM, Padubidri JR, Palladino R, Pana A, Panda-Jonas S, Parekh U, Park EC, Parvizi M, Kan FP, Patel UK, Pathak M, Paudel R, Pepito VCF, Perianayagam A, Perico N, Pham HQ, Pilgrim T, Piradov MA, Pishgar F, Podder V, Polibin R V., Pourshams A, Pribadi DRA, Rabiee N, Rabiee M, Radfar A, Rafiei A, Rahim F, Rahimi-Movaghar V, Rahman MHU, Rahman MA, Rahmani AM, Rakovac I, Ram P, Ramalingam S, Rana J, Ranasinghe P, Rao SJ, Rathi P, Rawal L, Rawasia WF, Rawassizadeh R, Remuzzi G, Renzaho AMN, Rezapour A, Riahi SM, Roberts-Thomson RL, Roever L, Rohloff P, Romoli M, Roshandel G, Rwegerera GM, Saadatagah S, Saber-Ayad MM, Sabour S, Sacco S, Sadeghi M, Moghaddam SS, Safari S, Sahebkar A, Salehi S, Salimzadeh H, Samaei M, Samy AM, Santos IS, Santric-Milicevic MM, Sarrafzadegan N, Sarveazad A, Sathish T,

Sawhney M, Saylan M, Schmidt MI, Schutte AE, Senthilkumaran S, Sepanlou SG, Sha F, Shahabi S, Shahid I, Shaikh MA, Shamali M, Shamsizadeh M, Shawon MSR, Sheikh A, Shigematsu M, Shin MJ, Shin J Il, Shiri R, Shiue I, Shuval K, Siabani S, Siddiqi TJ, Silva DAS, Singh JA, Singh A, Skryabin VY, Skryabina AA, Soheili A, Spurlock EE, Stockfelt L, Stortecky S, Stranges S, Abdulkader RS, Tadbiri H, Tadesse EG, Tadesse DB, Tajdini M, Tariqujjaman M, Teklehaimanot BF, Temsah MH, Tesema AK, Thakur B, Thankappan KR, Thapar R, Thrift AG, Timalsina B, Tonelli M, Touvier M, Tovani-Palone MR, Tripathi A, Tripathy JP, Truelsen TC, Tsegay GM, Tsegaye GW, Tsilimparis N, Tusa BS, Tyrovolas S, Umapathi KK, Unim B, Unnikrishnan B, Usman MS, Vaduganathan M, Valdez PR, Vasankari TJ, Velazquez DZ, Venketasubramanian N, Vu GT, Vujcic IS, Waheed Y, Wang Y, Wang F, Wei J, Weintraub RG, Weldemariam AH, Westerman R, Winkler AS, Wiysonge CS, Wolfe CDA, Wubishet BL, Xu G, Yadollahpour A, Yamagishi K, Yan LL, Yandrapalli S, Yano Y, Yatsuya H, Yeheyis TY, Yeshaw Y, Yilgwan CS, Yonemoto N, Yu C, Yusefzadeh H, Zachariah G, Zaman S Bin, Zaman MS, Zamanian M, Zand R, Zandifar A, Zarghi A, Zastrozhin MS, Zastrozhina A, Zhang ZJ, Zhang Y, Zhang W, Zhong C, Zou Z, Zuniga YMH, Murray CJL, Fuster V (2020) Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. J Am Coll Cardiol 76:2982–3021. doi: 10.1016/J.JACC.2020.11.010

2. Khan MA, Hashim MJ, Mustafa H, Baniyas MY, Suwaidi SKBM Al, AlKatheeri R, Alblooshi FMK, Almatrooshi MEAH, Alzaabi MEH, Darmaki RS Al, Lootah SNAH (2020) Global Epidemiology of Ischemic Heart Disease: Results from the Global Burden of Disease Study. Cureus 12. doi: 10.7759/CUREUS.9349

3. Libby P, Buring JE, Badimon L, Hansson GK, Deanfield J, Bittencourt MS, Tokgözoğlu

L, Lewis EF (2019) Atherosclerosis. Nat Rev Dis Prim 5. doi: 10.1038/S41572-019-0106-Z

4. Libby P (2021) The changing landscape of atherosclerosis. Nature 592:524–533. doi: 10.1038/S41586-021-03392-8

5. JC B, TE G, E K (2021) Risk Factors For Coronary Artery Disease. Risk Factors Coron Artery Dis 1–219. doi: 10.3109/9781420014570

6. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P (2007) Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. BMJ 335:136–141. doi: 10.1136/BMJ.39261.471806.55

7. Saçlı H, Kara İ, Kırali MK (2018) Focus on Coronary Atherosclerosis. Atheroscler - Yesterday, Today Tomorrow. doi: 10.5772/INTECHOPEN.77301

8. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, Martin HC, Lappalainen T, Posthuma D (2021) Genome-wide association studies. Nat Rev Methods Prim 1:59. doi: 10.1038/s43586-021-00056-9

9. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, Martin HC, Lappalainen T, Posthuma D (2021) Genome-wide association studies. Nat Rev Methods Prim 2021 11 1:1–21. doi: 10.1038/s43586-021-00056-9

10. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, Gabriel SB, Gibbs RA, Green ED, Hurles ME, Knoppers BM, Korbel JO, Lander ES, Lee C, Lehrach H, Mardis ER, Marth GT, McVean GA, Nickerson DA, Schmidt JP, Sherry ST, Wang J, Wilson RK, Boerwinkle E, Doddapaneni H, Han Y, Korchina V, Kovar C, Lee S, Muzny D, Reid JG, Zhu Y,

Chang Y, Feng Q, Fang X, Guo X, Jian M, Jiang H, Jin X, Lan T, Li G, Li J, Li Y, Liu S, Liu X, Lu Y, Ma X, Tang M, Wang B, Wang G, Wu H, Wu R, Xu X, Yin Y, Zhang D, Zhang W, Zhao J, Zhao M, Zheng X, Gupta N, Gharani N, Toji LH, Gerry NP, Resch AM, Barker J, Clarke L, Gil L, Hunt SE, Kelman G, Kulesha E, Leinonen R, McLaren WM, Radhakrishnan R, Roa A, Smirnov D, Smith RE, Streeter I, Thormann A, Toneva I, Vaughan B, Zheng-Bradley X, Grocock R, Humphray S, James T, Kingsbury Z, Sudbrak R, Albrecht MW, Amstislavskiy VS, Borodina TA, Lienhard M, Mertes F, Sultan M, Timmermann B, Yaspo ML, Fulton L, Ananiev V, Belaia Z, Beloslyudtsev D, Bouk N, Chen C, Church D, Cohen R, Cook C, Garner J, Hefferon T, Kimelman M, Liu C, Lopez J, Meric P, O'Sullivan C, Ostapchuk Y, Phan L, Ponomarov S, Schneider V, Shekhtman E, Sirotkin K, Slotta D, Zhang H, Balasubramaniam S, Burton J, Danecek P, Keane TM, Kolb-Kokocinski A, McCarthy S, Stalker J, Quail M, Davies CJ, Gollub J, Webster T, Wong B, Zhan Y, Campbell CL, Kong Y, Marcketta A, Yu F, Antunes L, Bainbridge M, Sabo A, Huang Z, Coin LJM, Fang L, Li Q, Li Z, Lin H, Liu B, Luo R, Shao H, Xie Y, Ye C, Yu C, Zhang F, Zheng H, Zhu H, Alkan C, Dal E, Kahveci F, Garrison EP, Kural D, Lee WP, Leong WF, Stromberg M, Ward AN, Wu J, Zhang M, Daly MJ, DePristo MA, Handsaker RE, Banks E, Bhatia G, Del Angel G, Genovese G, Li H, Kashin S, McCarroll SA, Nemesh JC, Poplin RE, Yoon SC, Lihm J, Makarov V, Gottipati S, Keinan A, Rodriguez-Flores JL, Rausch T, Fritz MH, Stütz AM, Beal K, Datta A, Herrero J, Ritchie GRS, Zerbino D, Sabeti PC, Shlyakhter I, Schaffner SF, Vitti J, Cooper DN, Ball E V., Stenson PD, Barnes B, Bauer M, Cheetham RK, Cox A, Eberle M, Kahn S, Murray L, Peden J, Shaw R, Kenny EE, Batzer MA, Konkel MK, Walker JA, MacArthur DG, Lek M, Herwig R, Ding L, Koboldt DC, Larson D, Ye K, Gravel S, Swaroop A, Chew E, Lappalainen T, Erlich Y, Gymrek M, Willems TF, Simpson JT, Shriver MD, Rosenfeld JA, Bustamante CD, Montgomery SB, De La Vega FM, Byrnes

JK, Carroll AW, DeGorter MK, Lacroute P, Maples BK, Martin AR, Moreno-Estrada A, Shringarpure SS, Zakharia F, Halperin E, Baran Y, Cerveira E, Hwang J, Malhotra A, Plewczynski D, Radew K, Romanovitch M, Zhang C, Hyland FCL, Craig DW, Christoforides A, Homer N, Izatt T, Kurdoglu AA, Sinari SA, Squire K, Xiao C, Sebat J, Antaki D, Gujral M, Noor A, Ye K, Burchard EG, Hernandez RD, Gignoux CR, Haussler D, Katzman SJ, Kent WJ, Howie B, Ruiz-Linares A, Dermitzakis ET, Devine SE, Kang HM, Kidd JM, Blackwell T, Caron S, Chen W, Emery S, Fritsche L, Fuchsberger C, Jun G, Li B, Lyons R, Scheller C, Sidore C, Song S, Sliwerska E, Taliun D, Tan A, Welch R, Wing MK, Zhan X, Awadalla P, Hodgkinson A, Li Y, Shi X, Quitadamo A, Lunter G, Marchini JL, Myers S, Churchhouse C, Delaneau O, Gupta-Hinch A, Kretzschmar W, Iqbal Z, Mathieson I, Menelaou A, Rimmer A, Xifara DK, Oleksyk TK, Fu Y, Liu X, Xiong M, Jorde L, Witherspoon D, Xing J, Browning BL, Browning SR, Hormozdiari F, Sudmant PH, Khurana E, Tyler-Smith C, Albers CA, Ayub Q, Chen Y, Colonna V, Jostins L, Walter K, Xue Y, Gerstein MB, Abyzov A, Balasubramanian S, Chen J, Clarke D, Fu Y, Harmanci AO, Jin M, Lee D, Liu J, Mu XJ, Zhang J, Zhang Y, Hartl C, Shakir K, Degenhardt J, Meiers S, Raeder B, Casale FP, Stegle O, Lameijer EW, Hall I, Bafna V, Michaelson J, Gardner EJ, Mills RE, Dayama G, Chen K, Fan X, Chong Z, Chen T, Chaisson MJ, Huddleston J, Malig M, Nelson BJ, Parrish NF, Blackburne B, Lindsay SJ, Ning Z, Zhang Y, Lam H, Sisu C, Challis D, Evani US, Lu J, Nagaswamy U, Yu J, Li W, Habegger L, Yu H, Cunningham F, Dunham I, Lage K, Jespersen JB, Horn H, Kim D, Desalle R, Narechania A, Sayres MAW, Mendez FL, Poznik GD, Underhill PA, Mittelman D, Banerjee R, Cerezo M, Fitzgerald TW, Louzada S, Massaia A, Yang F, Kalra D, Hale W, Dan X, Barnes KC, Beiswanger C, Cai H, Cao H, Henn B, Jones D, Kaye JS, Kent A, Kerasidou A, Mathias R, Ossorio PN, Parker M, Rotimi CN, Royal CD, Sandoval K, Su Y, Tian Z, Tishkoff S, Via M,

Wang Y, Yang H, Yang L, Zhu J, Bodmer W, Bedoya G, Cai Z, Gao Y, Chu J, Peltonen L, Garcia-Montero A, Orfao A, Dutil J, Martinez-Cruzado JC, Mathias RA, Hennis A, Watson H, McKenzie C, Qadri F, LaRocque R, Deng X, Asogun D, Folarin O, Happi C, Omoniwa O, Stremlau M, Tariyal R, Jallow M, Joof FS, Corrah T, Rockett K, Kwiatkowski D, Kooner J, Hien TT, Dunstan SJ, ThuyHang N, Fonnie R, Garry R, Kanneh L, Moses L, Schieffelin J, Grant DS, Gallo C, Poletti G, Saleheen D, Rasheed A, Brooks LD, Felsenfeld AL, McEwen JE, Vaydylevich Y, Duncanson A, Dunn M, Schloss JA (2015) A global reference for human genetic variation. Nature 526:68–74

11. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, Luo Y, Sidore C, Kwong A, Timpson N, Koskinen S, Vrieze S, Scott LJ, Zhang H, Mahajan A, Veldink J, Peters U, Pato C, Van Duijn CM, Gillies CE, Gandin I, Mezzavilla M, Gilly A, Cocca M, Traglia M, Angius A, Barrett JC, Boomsma D, Branham K, Breen G, Brummett CM, Busonero F, Campbell H, Chan A, Chen S, Chew E, Collins FS, Corbin LJ, Smith GD, Dedoussis G, Dorr M, Farmaki AE, Ferrucci L, Forer L, Fraser RM, Gabriel S, Levy S, Groop L, Harrison T, Hattersley A, Holmen OL, Hveem K, Kretzler M, Lee JC, McGue M, Meitinger T, Melzer D, Min JL, Mohlke KL, Vincent JB, Nauck M, Nickerson D, Palotie A, Pato M, Pirastu N, McInnis M, Richards JB, Sala C, Salomaa V, Schlessinger D, Schoenherr S, Slagboom PE, Small K, Spector T, Stambolian D, Tuke M, Tuomilehto J, Van Den Berg LH, Van Rheenen W, Volker U, Wijmenga C, Toniolo D, Zeggini E, Gasparini P, Sampson MG, Wilson JF, Frayling T, De Bakker PIW, Swertz MA, McCarroll S, Kooperberg C, Dekker A, Altshuler D, Willer C, Iacono W, Ripatti S, Soranzo N, Walter K, Swaroop A, Cucca F, Anderson CA, Myers RM, Boehnke M, McCarthy MI, Durbin R, Abecasis G, Marchini J (2016) A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 48:1279–1283. doi: 10.1038/ng.3643

12. (2005) A haplotype map of the human genome. Nature 437:1299–1320. doi: 10.1038/NATURE04226

13. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, Sham PC (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. doi: 10.1086/519795

14. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: A tool for genome-wide complex trait analysis. Am J Hum Genet 88:76–82. doi: 10.1016/j.ajhg.2010.11.011

15. Yang J, Ferreira T, Morris AP, Medland SE, F Madden PA, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, Frayling TM, McCarthy MI, Hirschhorn JN, Goddard ME, Visscher PM, Genet Author manuscript N (2013) Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits Genetic Investigation of ANthropometric Traits (GIANT) Consortium 4 , DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium NIH Public Access Author Manuscript. Nat Genet 44:369–372. doi: 10.1038/ng.2213

16. Nica AC, Dermitzakis ET (2013) Expression quantitative trait loci: present and future. Philos Trans R Soc B Biol Sci 368. doi: 10.1098/RSTB.2012.0362

17. Franzén O, Ermel R, Cohain A, Akers NK, Di Narzo A, Talukdar HA, Foroughi-Asl H, Giambartolomei C, Fullard JF, Sukhavasi K, Köks S, Gan LM, Giannarelli C, Kovacic JC, Betsholtz C, Losic B, Michoel T, Hao K, Roussos P, Skogsberg J, Ruusalepp A, Schadt EE, Björkegren JLM (2016) Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. Science (80-) 353:827–830. doi: 10.1126/science.aad6970

18. GTEx Consortium (2013) The Genotype-Tissue Expression (GTEx) project. Nat Genet 45:580–5. doi: 10.1038/ng.2653

19. Võsa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, Kirsten H, Saha A, Kreuzhuber R, Yazar S, Brugge H, Oelen R, de Vries DH, van der Wijst MGP, Kasela S, Pervjakova N, Alves I, Favé M-J, Agbessi M, Christiansen MW, Jansen R, Seppälä I, Tong L, Teumer A, Schramm K, Hemani G, Verlouw J, Yaghootkar H, Sönmez Flitman R, Brown A, Kukushkina V, Kalnapenkis A, Rüeger S, Porcu E, Kronberg J, Kettunen J, Lee B, Zhang F, Qi T, Hernandez JA, Arindrarto W, Beutner F, Dmitrieva J, Elansary M, Fairfax BP, Georges M, Heijmans BT, Hewitt AW, Kähönen M, Kim Y, Knight JC, Kovacs P, Krohn K, Li S, Loeffler M, Marigorta UM, Mei H, Momozawa Y, Müller-Nurasyid M, Nauck M, Nivard MG, Penninx BWJH, Pritchard JK, Raitakari OT, Rotzschke O, Slagboom EP, Stehouwer CDA, Stumvoll M, Sullivan P, 't Hoen PAC, Thiery J, Tönjes A, van Dongen J, van Iterson M, Veldink JH, Völker U, Warmerdam R, Wijmenga C, Swertz M, Andiappan A, Montgomery GW, Ripatti S, Perola M, Kutalik Z, Dermitzakis E, Bergmann S, Frayling T, van Meurs J, Prokisch H, Ahsan H, Pierce BL, Lehtimäki T, Boomsma DI, Psaty BM, Gharib SA, Awadalla P, Milani L, Ouwehand WH, Downes K, Stegle O, Battle A, Visscher PM, Yang J, Scholz M, Powell J, Gibson G, Esko T, Franke L (2021) Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nat Genet 2021 539 53:1300–1310. doi: 10.1038/s41588-021-00913-z

20. Nurnberg ST, Guerraty MA, Wirka RC, Rao HS, Pjanic M, Norton S, Serrano F, Perisic L, Elwyn S, Pluta J, Zhao W, Testa S, Park Y, Nguyen T, Ko Y-A, Wang T, Hedin U, Sinha S, Barash Y, Brown CD, Quertermous T, Rader DJ (2020) Genomic profiling of human vascular cells identifies TWIST1 as a causal gene for common vascular diseases.

PLoS Genet 16. doi: 10.1371/JOURNAL.PGEN.1008538

21. Yao C, Chen G, Song C, Keefe J, Mendelson M, Huan T, Sun BB, Laser A, Maranville JC, Wu H, Ho JE, Courchesne P, Lyass A, Larson MG, Gieger C, Graumann J, Johnson AD, Danesh J, Runz H, Hwang S-J, Liu C, Butterworth AS, Suhre K, Levy D (2018) Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease. Nat Commun 2018 91 9:1–11. doi: 10.1038/s41467-018-05512-x

22. Ciuculete DM, Boström AE, Voisin S, Philipps H, Titova OE, Bandstein M, Nikontovic L, Williams MJ, Mwinyi J, Schiöth HB (2017) A methylome-wide mQTL analysis reveals associations of methylation sites with GAD1 and HDAC3 SNPs and a general psychiatric risk score. Transl Psychiatry 2017 71 7:e1002–e1002. doi: 10.1038/tp.2016.275

23. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V (2014) Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLoS Genet 10:e1004383. doi: 10.1371/journal.pgen.1004383

24. Hormozdiari F, van de Bunt M, Segrè AV, Li X, Joo JWJ, Bilow M, Sul JH, Sankararaman S, Pasaniuc B, Eskin E (2016) Colocalization of GWAS and eQTL Signals Detects Target Genes. Am J Hum Genet 99:1245. doi: 10.1016/J.AJHG.2016.10.003

25. Wu Y, Broadaway KA, Raulerson CK, Scott LJ, Pan C, Ko A, He A, Tilford C, Fuchsberger C, Locke AE, Stringham HM, Jackson AU, Narisu N, Kuusisto J, Pajukanta P, Collins FS, Boehnke M, Laakso M, Lusis AJ, Civelek M, Mohlke KL (2019) Colocalization of GWAS and eQTL signals at loci with multiple signals identifies

additional candidate genes for body fat distribution. Hum Mol Genet 28:4161–4172. doi: 10.1093/HMG/DDZ263

26. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V (2014) Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLoS Genet 10:e1004383. doi: 10.1371/journal.pgen.1004383

27. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, Torstenson ES, Shah KP, Garcia T, Edwards TL, Stahl EA, Huckins LM, Consortium G, Nicolae DL, Cox NJ, Im HK Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. doi: 10.1038/s41467-018-03621-1

28. Giambartolomei C, Liu JZ, Zhang W, Hauberg M, Shi H, Boocock J, Pickrell J, Jaffe AE, Pasaniuc B, Roussos P (2018) A Bayesian framework for multiple trait colocalization from summary association statistics. Bioinformatics 34:2538–2545. doi: 10.1093/BIOINFORMATICS/BTY147

29. Torres JM, Abdalla M, Payne A, Fernandez-Tajes J, Thurner M, Nylander V, Gloyn AL, Mahajan A, McCarthy MI (2020) A Multi-omic Integrative Scheme Characterizes Tissues of Action at Loci Associated with Type 2 Diabetes. Am J Hum Genet 107:1011–1028. doi: 10.1016/j.ajhg.2020.10.009

30. Hebbring SJ (2014) The challenges, advantages and future of phenome-wide association studies. Immunology 141:157–165. doi: 10.1111/imm.12195

31. Klarin D, Zhu QM, Emdin CA, Chaffin M, Horner S, McMillan BJ, Leed A, Weale ME, Spencer CCA, Aguet F, Segrè A V., Ardlie KG, Khera A V., Kaushik VK, Natarajan P,

Kathiresan S (2017) Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. Nat Genet 49:1392–1397. doi: 10.1038/ng.3914

32. Klarin D, Lynch J, Aragam K, Chaffin M, Assimes TL, Huang J, Lee KM, Shao Q, Huffman JE, Natarajan P, Arya S, Small A, Sun Y V., Vujkovic M, Freiberg MS, Wang L, Chen J, Saleheen D, Lee JS, Miller DR, Reaven P, Alba PR, Patterson O V., DuVall SL, Boden WE, Beckman JA, Gaziano JM, Concato J, Rader DJ, Cho K, Chang KM, Wilson PWF, O'Donnell CJ, Kathiresan S, Tsao PS, Damrauer SM (2019) Genome-wide association study of peripheral artery disease in the Million Veteran Program. Nat Med 2019 258 25:1274–1279. doi: 10.1038/s41591-019-0492-5

33. Verma A, Lucas A, Verma SS, Zhang Y, Josyula N, Khan A, Hartzel DN, Lavage DR, Leader J, Ritchie MD, Pendergrass SA (2018) PheWAS and Beyond: The Landscape of Associations with Medical Diagnoses and Clinical Measures across 38,662 Individuals from Geisinger. Am J Hum Genet 102:592–608. doi: 10.1016/J.AJHG.2018.02.017

34. Bush WS, Oetjens MT, Crawford DC (2016) Unravelling the human genome-phenome relationship using phenome-wide association studies. Nat. Rev. Genet. 17:129–145

35. Diogo D, Tian C, Franklin CS, Alanne-Kinnunen M, March M, Spencer CCA, Vangjeli C, Weale ME, Mattsson H, Kilpeläinen E, Sleiman PMA, Reilly DF, McElwee J, Maranville JC, Chatterjee AK, Bhandari A, Nguyen KDH, Estrada K, Reeve MP, Hutz J, Bing N, John S, MacArthur DG, Salomaa V, Ripatti S, Hakonarson H, Daly MJ, Palotie A, Hinds DA, Donnelly P, Fox CS, Day-Williams AG, Plenge RM, Runz H (2018) Phenome-wide association studies across large population cohorts support drug target validation. Nat Commun 2018 91 9:1–13. doi: 10.1038/s41467-018-06540-3

36. Østergaard SD, Mukherjee S, Sharp SJ, Proitsi P, Lotta LA, Day F, Perry JRB, Boehme

KL, Walter S, Kauwe JS, Gibbons LE, Larson EB, Powell JF, Langenberg C, Crane PK, Wareham NJ, Scott RA, Albert MS, Albin RL, Apostolova LG, Arnold SE, Asthana S, Atwood CS, Baldwin CT, Barber RC, Barmada MM, Barnes LL, Beach TG, Becker JT, Beecham GW, Beekly D, Bigio EH, Bird TD, Blacker D, Boeve BF, Bowen JD, Boxer A, Burke JR, Buxbaum JD, Cairns NJ, Cantwell LB, Cao C, Carlson CS, Carlsson CM, Carney RM, Carrasquillo MM, Carroll SL, Chui HC, Clark DG, Corneveaux J, Cribbs DH, Crocco EA, Cruchaga C, De Jager PL, DeCarli C, Demirci FY, Dick M, Dickson DW, Duara R, Ertekin-Taner N, Evans D, Faber KM, Fallon KB, Farlow MR, Farrer LA, Ferris S, Foroud TM, Frosch MP, Galasko DR, Gearing M, Geschwind DH, Ghetti B, Gilbert JR, Glass JD, Goate AM, Graff-Radford NR, Green RC, Growdon JH, Haines JL, Hakonarson H, Hamilton RL, Hamilton-Nelson KL, Hardy J, Harrell LE, Head E, Honig LS, Huebinger RM, Huentelman MJ, Hulette CM, Hyman BT, Jarvik GP, Jicha GA, Jin LW, Jun G, Kamboh MI, Karydas A, Kaye JA, Kim R, Kowall NW, Kramer JH, Kukull WA, Kunkle BW, LaFerla FM, Lah JJ, Leverenz JB, Levey AI, Li G, Lieberman AP, Lin CF, Lopez OL, Lunetta KL, Lyketsos CG, Mack WJ, Marson DC, Martin ER, Martiniuk F, Mash DC, Masliah E, Mayeux R, McCormick WC, McCurry SM, McDavid AN, McKee AC, Mesulam M, Miller BL, Miller CA, Miller JW, Montine TJ, Morris JC, Murrell JR, Myers AJ, Naj AC, Olichney JM, Pankratz VS, Parisi JE, Partch A, Paulson HL, Pericak-Vance MA, Perry W, Peskind E, Petersen RC, Pierce A, Poon WW, Potter H, Quinn JF, Raj A, Raskind M, Reiman EM, Reisberg B, Reitz C, Ringman JM, Roberson ED, Rogaeva E, Rosen HJ, Rosenberg RN, Sager MA, Sano M, Schellenberg GD, Schneider JA, Schneider LS, Seeley WW, Smith AG, Sonnen JA, Spina S, St George-Hyslop P, Stern RA, Tanzi RE, Thornton-Wells TA, Trojanowski JQ, Troncoso JC, Tsuang DW, Valladares O, VanDeerlin VM, Van Eldik LJ, Vardarajan BN, Vinters H V., Vonsattel JP, Wang LS, Weintraub S, Welsh-Bohmer KA,

Williamson J, Wishnek S, Woltjer RL, Wright CB, Younkin SG, Yu CE, Yu L, Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Chapman J, Russo G, Hamshere M, Singh Pahwa J, Escott-Price V, Badarinarayan N, Dowzell K, Williams A, Jones N, Thomas C, Stretton A, Morgan A, Taylor S, Lovestone S, Powell J, Proitsi P, Lupton MK, Brayne C, Rubinsztein DC, Gill M, Lawlor B, Lynch A, Morgan K, Brown K, Passmore P, Craig D, McGuinness B, Todd S, Johnston J, Holmes C, Mann D, David Smith A, Love S, Kehoe PG, Hardy J, Mead S, Fox N, Rossor M, Collinge J, Maier W, Jessen F, Heun R, Schürmann B, Ramirez A, Becker T, Herold C, Lacour A, Drichel D, van den Bussche H, Heuser I, Kornhuber J, Wiltfang J, Dichgans M, Frölich L, Hampel H, Hüll M, Rujescu D, Goate A, Kauwe JSK, Cruchaga C, Nowotny P, Morris JC, Mayo K, Livingston G, Bass NJ, Gurling H, McQuillin A, Gwilliam R, Deloukas P, Al-Chalabi A, Shaw CE, Singleton AB, Guerreiro R, Mühleisen TW, Nöthen MM, Moebus S, Jöckel KH, Klopp N, Wichmann HE, Carrasquillo MM, Shane Pankratz V, Younkin SG, Holmans P, Donovan MO, Owen MJ, Williams J, Langenberg C, Scott RA, Sharp SJ, Forouhi NG, Kerrison ND, Sims M, Lucarelli DME, Barroso I, Deloukas P, McCarthy MI, Arriola L, Balkau B, Barricarte A, Boeing H, Franks PW, Gonzalez C, Grioni S, Kaaks R, Key TJ, Navarro C, Nilsson PM, Overvad K, Palli D, Panico S, Ramón Quirós J, Rolandsson O, Sacerdote C, Sánchez MJ, Slimani N, Tjonneland A, Tumino R, van der Daphne AL, van der Schouw YT, Riboli E, Wareham NJ (2015) Associations between Potentially Modifiable Risk Factors and Alzheimer Disease: A Mendelian Randomization Study. PLoS Med 12. doi: 10.1371/JOURNAL.PMED.1001841

37. Boef AGC, Dekkers OM, Le Cessie S (2015) Mendelian randomization studies: a review of the approaches used and the quality of reporting. Int J Epidemiol 44:496–511. doi: 10.1093/IJE/DYV071

38. Lawlor DA (2016) Commentary: Two-sample Mendelian randomization: opportunities and challenges. Int J Epidemiol 45:908. doi: 10.1093/IJE/DYW127

39. Zheng J, Frysz M, Kemp JP, Evans DM, Davey Smith G, Tobias JH (2019) Use of Mendelian Randomization to Examine Causal Inference in Osteoporosis. Front Endocrinol (Lausanne) 10. doi: 10.3389/FENDO.2019.00807

40. de Bakker PIW, Ferreira MAR, Jia X, Neale BM, Raychaudhuri S, Voight BF (2008) Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet 17. doi: 10.1093/hmg/ddn288

41. Burgess S, Thompson SG (2017) Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol 32:377–389. doi: 10.1007/s10654-017-0255-x

42. Bowden J, Davey Smith G, Haycock PC, Burgess S (2016) Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol 40:304–314. doi: 10.1002/gepi.21965

43. Jansen H, Samani NJ, Schunkert H (2014) Mendelian randomization studies in coronary artery disease. Eur Heart J 35:1917–1924. doi: 10.1093/EURHEARTJ/EHU208

44. Nakagawa Y (2013) Role of Statins in Coronary Artery Disease. Chonnam Med J 49:1. doi: 10.4068/CMJ.2013.49.1.1

45. Zeng L, Ntalla I, Kessler T, Kastrati A, Erdmann J, Danesh J, Watkins H, Samani NJ, Deloukas P, Schunkert H (2019) Genetically modulated educational attainment and coronary disease risk. Eur Heart J 40:2413–2420. doi: 10.1093/EURHEARTJ/EHZ328

46. Koplev S, Seldin M, Sukhavasi K, Ermel R, Pang S, Zeng L, Bankier S, Di Narzo A, Cheng H, Meda V, Ma A, Talukdar H, Cohain A, Amadori L, Argmann C, Houten SM,

Franzén O, Mocci G, Meelu OA, Ishikawa K, Whatling C, Jain A, Jain RK, Gan L-M, Giannarelli C, Roussos P, Hao K, Schunkert H, Michoel T, Ruusalepp A, Schadt EE, Kovacic JC, Lusis AJ, Björkegren JLM (2022) A mechanistic framework for cardiometabolic and coronary artery diseases. Springer US

47. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J (2018) The UK Biobank resource with deep phenotyping and genomic data. Nature 562:203–209. doi: 10.1038/s41586-018-0579-z

48. Backman JD, Li AH, Marcketta A, Sun D, Mbatchou J, Kessler MD, Benner C, Liu D, Locke AE, Balasubramanian S, Yadav A, Banerjee N, Gillies C, Damask A, Liu S, Bai X, Hawes A, Maxwell E, Gurski L, Watanabe K, Kosmicki JA, Rajagopal V, Mighty J, Jones M, Mitnaul L, Stahl E, Coppola G, Jorgenson E, Habegger L, Salerno WJ, Shuldiner AR, Lotta LA, Overton JD, Cantor MN, Reid JG, Yancopoulos G, Kang HM, Marchini J, Baras A, Abecasis GR, Ferreira MA (2021) Exome sequencing and analysis of 454,787 UK Biobank participants. Nat 2021 599:1–7. doi: 10.1038/s41586-021-04103-z

49. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN (2021) Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants

50. Córdova-Palomera A, Tcheandjieu C, Fries JA, Varma P, Chen VS, Fiterau M, Xiao K, Tejeda H, Keavney BD, Cordell HJ, Tanigawa Y, Venkataraman G, Rivas MA, Ré C, Ashley E, Priest JR (2020) Cardiac Imaging of Aortic Valve Area from 34 287 UK Biobank Participants Reveals Novel Genetic Associations and Shared Genetic Comorbidity with Multiple Disease Phenotypes. Circ Genomic Precis Med 13:609–617.

doi: 10.1161/CIRCGEN.120.003014

51. Khera A V., Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT, Kathiresan S (2018) Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 2018 509 50:1219–1224. doi: 10.1038/s41588-018-0183-z

52. Schork NJ, Murray SS, Frazer KA, Topol EJ (2009) Common vs. Rare Allele Hypotheses for Complex Diseases. Curr Opin Genet Dev 19:212. doi: 10.1016/J.GDE.2009.04.010

53. Garcia CK (2018) Insights from human genetic studies of lung and organ fibrosis. J Clin Invest 128:36–44. doi: 10.1172/JCI93556

54. Lee S, Abecasis GR, Boehnke M, Lin X (2014) Rare-variant association analysis: Study designs and statistical tests. Am. J. Hum. Genet. 95:5–23

55. Chun S, Fay JC (2009) Identification of deleterious mutations within three human genomes. Genome Res 19:1553–1561. doi: 10.1101/GR.092619.109

56. Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, Liu X (2015) Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum Mol Genet 24:2125–2137. doi: 10.1093/hmg/ddu733

57. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods 2010 74 7:248–249. doi: 10.1038/nmeth0410-248

58. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009 47 4:1073–1081. doi: 10.1038/nprot.2009.86

59. Liu X, Li C, Mou C, Dong Y, Tu Y (2020) dbNSFP v4: a comprehensive database of transcript-specific functional predictions and annotations for human nonsynonymous and splice-site SNVs. Genome Med 12:1–8. doi: 10.1186/S13073-020-00803-9/FIGURES/4

60. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, Hoffman D, Jang W, Kaur K, Liu C, Lyoshin V, Maddipatla Z, Maiti R, Mitchell J, O'leary N, Riley GR, Shi W, Zhou G, Schneider V, Maglott D, Holmes JB, Kattman BL (2019) ClinVar: improvements to accessing data. Nucleic Acids Res 48:835–844. doi: 10.1093/nar/gkz972

61. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F (2016) The Ensembl Variant Effect Predictor. Genome Biol 17. doi: 10.1186/s13059-016-0974-4

62. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6:80. doi: 10.4161/FLY.19695

63. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38. doi: 10.1093/NAR/GKQ603

64. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X (2011) Rare-Variant Association Testing for Sequencing Data with the Sequence Kernel Association Test. Am J Hum Genet 89:82. doi: 10.1016/J.AJHG.2011.05.029

65. Investigators MIG and CardiEC (2016) Coding Variation in ANGPTL4, LPL, and

SVEP1 and the Risk of Coronary Disease. N Engl J Med 374:1134. doi: 10.1056/NEJMOA1507652

66. Winkler MJ, Müller P, Sharifi AM, Wobst J, Winter H, Mokry M, Ma L, van der Laan SW, Pang S, Miritsch B, Hinterdobler J, Werner J, Stiller B, Güldener U, Webb TR, Asselbergs FW, Björkegren JLM, Maegdefessel L, Schunkert H, Sager HB, Kessler T (2020) Functional investigation of the coronary artery disease gene SVEP1. Basic Res Cardiol 115:1–15. doi: 10.1007/S00395-020-00828-6/FIGURES/7

67. Gamazon ER, Wheeler HE, Shah KP, Mozaffari S V, Aquino-Michaels K, Carroll RJ, Eyler AE, Denny JC, Nicolae DL, Cox NJ, Im HK, Consortium Gte (2015) A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 47:1091–1098. doi: 10.1038/ng.3367

68. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K, Björkegren JLM, Im HK, Pasaniuc B, Rivas MA, Kundaje A (2019) Opportunities and challenges for transcriptome-wide association studies. Nat Genet 51:592–599. doi: 10.1038/s41588-019-0385-z

69. Zhang W, Voloudakis G, Rajagopal VM, Readhead B, Dudley JT, Schadt EE, Björkegren JLM, Kim Y, Fullard JF, Hoffman GE, Roussos P (2019) Integrative transcriptome imputation reveals tissue-specific and shared biological mechanisms mediating susceptibility to complex traits. Nat Commun 10:1–13. doi: 10.1038/s41467-019-11874-7

70. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, Quon G, Sandstrom RS, Eaton ML, Wu YC, Pfenning AR, Wang X, Claussnitzer M, Liu Y, Coarfa C, Harris RA, Shoresh N, Epstein

CB, Gjoneska E, Leung D, Xie W, Hawkins RD, Lister R, Hong C, Gascard P, Mungall AJ, Moore R, Chuah E, Tam A, Canfield TK, Hansen RS, Kaul R, Sabo PJ, Bansal MS, Carles A, Dixon JR, Farh KH, Feizi S, Karlic R, Kim AR, Kulkarni A, Li D, Lowdon R, Elliott G, Mercer TR, Neph SJ, Onuchic V, Polak P, Rajagopal N, Ray P, Sallari RC, Siebenthall KT, Sinnott-Armstrong NA, Stevens M, Thurman RE, Wu J, Zhang B, Zhou X, Beaudet AE, Boyer LA, De Jager PL, Farnham PJ, Fisher SJ, Haussler D, Jones SJM, Li W, Marra MA, McManus MT, Sunyaev S, Thomson JA, Tlsty TD, Tsai LH, Wang W, Waterland RA, Zhang MQ, Chadwick LH, Bernstein BE, Costello JF, Ecker JR, Hirst M, Meissner A, Milosavljevic A, Ren B, Stamatoyannopoulos JA, Wang T, Kellis M (2015) Integrative analysis of 111 reference human epigenomes. Nature 518:317–329. doi: 10.1038/nature14248

71. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang HY, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkilä K, Hyppönen E, Isaacs A, Jackson AU, Johansson Å, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikäinen LP, Magnusson PKE, Mangino M, Mihailov E, Montasser ME, Müller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen AK, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van Den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney ASF, Döring A, Elliott P, Epstein SE, Eyjolfsson GI, Gigante B, Goodarzi MO, Grallert H, Gravito ML, Groves CJ, Hallmans G, Hartikainen AL, Hayward C, Hernandez D, Hicks AA, Holm H, Hung YJ, Illig T, Jones MR, Kaleebu P,

Kastelein JJP, Khaw KT, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin SY, Lindström J, Loos RJF, Mach F, McArdle WL, Meisinger C, Mitchell BD, Müller G, Nagaraja R, Narisu N, Nieminen TVM, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stancáková A, Stirrups K, Swift AJ, Tiret L, Uitterlinden AG, Van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen YDI, Collins FS, Cooper RS, Danesh J, Dedoussis G, De Faire U, Feranil AB, Ferrières J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllensten U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Järvelin MR, Jula A, Kähönen M, Kaprio J, Kesäniemi A, Kivimaki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, März W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njølstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PEH, Sheu WHH, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, Van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BHR, Ordovas JM, Boerwinkle E, Palmer CNA, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Kathiresan S, Mohlke KL, Ingelsson E, Abecasis GR (2013) Discovery and refinement of loci associated with lipid levels. Nat Genet 2013 4511

45:1274–1283. doi: 10.1038/ng.2797

72. Mancuso N, Gayther S, Gusev A, Zheng W, Penney KL, Kote-Jarai Z, Eeles R, Freedman M, Haiman C, Pasaniuc B, Henderson BE, Benlloch S, Schumacher FR, Olama AA Al, Muir K, Berndt SI, Conti D V., Wiklund F, Chanock S, Stevens VL, Tangen CM, Batra J, Clements J, Gronberg H, Pashayan N, Schleutker J, Albanes D, Weinstein S, Wolk A, West C, Mucci L, Cancel-Tassin G, Koutros S, Sorensen KD, Maehle L, Neal DE, Hamdy FC, Donovan JL, Travis RC, Hamilton RJ, Ingles SA, Rosenstein B, Lu YJ, Giles GG, Kibel AS, Vega A, Kogevinas M, Park JY, Stanford JL, Cybulski C, Nordestgaard BG, Brenner H, Maier C, Kim J, John EM, Teixeira MR, Neuhausen SL, De Ruyck K, Razack A, Newcomb LF, Lessel D, Kaneva R, Usmani N, Claessens F, Townsend PA, Dominguez MG, Roobol MJ, Menegaux F, Khaw KT, Cannon-Albright L, Pandha H, Thibodeau SN, Hunter DJ, Kraft P (2018) Large-scale transcriptome-wide association study identifies new prostate cancer risk regions. Nat Commun 9:1–11. doi: 10.1038/s41467-018-06302-1

73. Zhou X, Carbonetto P, Stephens M (2013) Polygenic modeling with bayesian sparse linear mixed models. PLoS Genet 9. doi: 10.1371/JOURNAL.PGEN.1003264

74. Fryett JJ, Inshaw J, Morris AP, Cordell HJ (2018) Comparison of methods for transcriptome imputation through application to two common complex diseases. Eur J Hum Genet 2018 2611 26:1658–1667. doi: 10.1038/s41431-018-0176-5

75. Luningham JM, Chen J, Tang S, De Jager PL, Bennett DA, Buchman AS, Yang J (2020) Bayesian Genome-wide TWAS Method to Leverage both cis- and trans-eQTL Information through Summary Statistics. Am J Hum Genet 107:714–726. doi: 10.1016/J.AJHG.2020.08.022

76. Bhattacharya A, Li Y, Love MI (2021) MOSTWAS: Multi-Omic Strategies for

Transcriptome-Wide Association Studies. PLOS Genet 17:e1009398. doi: 10.1371/JOURNAL.PGEN.1009398

77. Wu L, Shi W, Long J, Guo X, Michailidou K, Beesley J, Bolla MK, Shu XO, Lu Y, Cai Q, Al-Ejeh F, Rozali E, Wang Q, Dennis J, Li B, Zeng C, Feng H, Gusev A, Barfield RT, Andrulis IL, Anton-Culver H, Arndt V, Aronson KJ, Auer PL, Barrdahl M, Baynes C, Beckmann MW, Benitez J, Bermisheva M, Blomqvist C, Bogdanova N V., Bojesen SE, Brauch H, Brenner H, Brinton L, Broberg P, Brucker SY, Burwinkel B, Caldés T, Canzian F, Carter BD, Castelao JE, Chang-Claude J, Chen X, Cheng TYD, Christiansen H, Clarke CL, Collée M, Cornelissen S, Couch FJ, Cox D, Cox A, Cross SS, Cunningham JM, Czene K, Daly MB, Devilee P, Doheny KF, Dörk T, Dos-Santos-Silva I, Dumont M, Dwek M, Eccles DM, Eilber U, Eliassen AH, Engel C, Eriksson M, Fachal L, Fasching PA, Figueroa J, Flesch-Janys D, Fletcher O, Flyger H, Fritschi L, Gabrielson M, Gago-Dominguez M, Gapstur SM, García-Closas M, Gaudet MM, Ghoussaini M, Giles GG, Goldberg MS, Goldgar DE, González-Neira A, Guénel P, Hahnen E, Haiman CA, Håkansson N, Hall P, Hallberg E, Hamann U, Harrington P, Hein A, Hicks B, Hillemanns P, Hollestelle A, Hoover RN, Hopper JL, Huang G, Humphreys K, Hunter DJ, Jakubowska A, Janni W, John EM, Johnson N, Jones K, Jones ME, Jung A, Kaaks R, Kerin MJ, Khusnutdinova E, Kosma VM, Kristensen VN, Lambrechts D, Le Marchand L, Li J, Lindström S, Lissowska J, Lo WY, Loibl S, Lubinski J, Luccarini C, Lux MP, MacInnis RJ, Maishman T, Kostovska IM, Mannermaa A, Manson JAE, Margolin S, Mavroudis D, Meijers-Heijboer H, Meindl A, Menon U, Meyer J, Mulligan AM, Neuhausen SL, Nevanlinna H, Neven P, Nielsen SF, Nordestgaard BG, Olopade OI, Olson JE, Olsson H, Peterlongo P, Peto J, Plaseska-Karanfilska D, Prentice R, Presneau N, Pylkäs K, Rack B, Radice P, Rahman N, Rennert G, Rennert HS, Rhenius V, Romero A, Romm J, Rudolph A, Saloustros E, Sandler DP, Sawyer EJ, Schmidt MK,

Schmutzler RK, Schneeweiss A, Scott RJ, Scott CG, Seal S, Shah M, Shrubsole MJ, Smeets A, Southey MC, Spinelli JJ, Stone J, Surowy H, Swerdlow AJ, Tamimi RM, Tapper W, Taylor JA, Terry MB, Tessier DC, Thomas A, Thöne K, Tollenaar RAEM, Torres D, Truong T, Untch M, Vachon C, Van Den Berg D, Vincent D, Waisfisz Q, Weinberg CR, Wendt C, Whittemore AS, Wildiers H, Willett WC, Winqvist R, Wolk A, Xia L, Yang XR, Ziogas A, Ziv E, Dunning AM, Pharoah PDP, Simard J, Milne RL, Edwards SL, Kraft P, Easton DF, Chenevix-Trench G, Zheng W (2018) A transcriptome-wide association study of 229,000 women identifies new candidate susceptibility genes for breast cancer. Nat Genet 50:968–978. doi: 10.1038/s41588-018-0132-x

78. Huckins LM, Dobbyn A, Ruderfer DM, Hoffman G, Wang W, Pardiñas AF, Rajagopal VM, Als TD, T. Nguyen H, Girdhar K, Boocock J, Roussos P, Fromer M, Kramer R, Domenici E, Gamazon ER, Purcell S, Johnson JS, Shah HR, Klein LL, Dang KK, Logsdon BA, Mahajan MC, Mangravite LM, Toyoshiba H, Gur RE, Hahn CG, Schadt E, Lewis DA, Haroutunian V, Peters MA, Lipska BK, Buxbaum JD, Hirai K, Perumal TM, Essioux L, Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, Lee P, Bulik-Sullivan B, Collier DA, Huang H, Pers TH, Agartz I, Agerbo E, Albus M, Alexander M, Amin F, Bacanu SA, Begemann M, Belliveau RA, Bene J, Bergen SE, Bevilacqua E, Bigdeli TB, Black DW, Bruggeman R, Buccola NG, Buckner RL, Byerley W, Cahn W, Cai G, Campion D, Cantor RM, Carr VJ, Carrera N, Catts S V., Chambert KD, Chan RCK, Chen RYL, Chen EYH, Cheng W, Cheung EFC, Chong SA, Cloninger CR, Cohen D, Cohen N, Cormican P, Craddock N, Crowley JJ, Curtis D, Davidson M, Davis KL, Degenhardt F, Del Favero J, Demontis D, Dikeos D, Dinan T, Djurovic S, Donohoe G, Drapeau E, Duan J, Dudbridge F, Durmishi N, Eichhammer P, Eriksson J, Escott-Price V, Fanous AH, Farrell MS, Frank J, Franke L, Freedman R,

Freimer NB, Friedl M, Friedman JI, Fromer M, Genovese G, Georgieva L, Giegling I, Giusti-Rodríguez P, Godard S, Goldstein JI, Golimbet V, Gopal S, Gratten J, de Haan L, Hammer C, Hamshere ML, Hansen M, Hansen T, Hartmann AM, Henskens FA, Herms S, Hirschhorn JN, Hoffmann P, Hofman A, Hollegaard M V., Hougaard DM, Ikeda M, Joa I, Julia A, Kahn RS, Kalaydjieva L, Karachanak-Yankova S, Karjalainen J, Kavanagh D, Keller MC, Kennedy JL, Khrunin A, Kim Y, Klovins J, Knowles JA, Konte B, Kucinskas V, Kucinskiene ZA, Kuzelova-Ptackova H, Kahler AK, Laurent C, Keong JLC, Lee SH, Legge SE, Lerer B, Li M, Li T, Liang KY, Lieberman J, Limborska S, Loughland CM, Lubinski J, Lonnqvist J, Macek M, Magnusson PKE, Maher BS, Maier W, Mallet J, Marsal S, Mattheisen M, Mattingsdal M, McCarley RW, McDonald C, McIntosh AM, Meier S, Meijer CJ, Melegh B, Melle I, Mesholam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mokrab Y, Morris DW, Mors O, Murphy KC, Murray RM, Myin-Germeys I, Muller-Myhsok B, Nelis M, Nenadic I, Nertney DA, Nestadt G, Nicodemus KK, Nikitina-Zake L, Nisenbaum L, Nordin A, O'Callaghan E, O'Dushlaine C, O'Neill FA, Oh SY, Olincy A, Olsen L, Van Os J, Pantelis C, Papadimitriou GN, Papiol S, Parkhomenko E, Pato MT, Paunio T, Pejovic-Milovancevic M, Perkins DO, Pietiläinen O, Pimm J, Pocklington AJ, Powell J, Price A, Pulver AE, Purcell SM, Quested D, Rasmussen HB, Reichenberg A, Reimers MA, Richards AL, Roffman JL, Salomaa V, Sanders AR, Schall U, Schubert CR, Schulze TG, Schwab SG, Scolnick EM, Scott RJ, Seidman LJ, Shi J, Sigurdsson E, Silagadze T, Silverman JM, Sim K, Slominsky P, Smoller JW, So HC, Spencer CCA, Stahl EA, Stefansson H, Steinberg S, Stogmann E, Straub RE, Strengman E, Strohmaier J, Stroup TS, Subramaniam M, Suvisaari J, Svrakic DM, Szatkiewicz JP, Soderman E, Thirumalai S, Toncheva D, Tosato S, Veijola J, Waddington J, Walsh D, Wang D, Wang Q, Webb BT, Weiser M, Wildenauer DB, Williams NM, Williams S, Witt SH, Wolen AR, Wong

EHM, Wormley BK, Xi HS, Zai CC, Zheng X, Zimprich F, Wray NR, Stefansson K, Visscher PM, Adolfsson R, Andreassen OA, Blackwood DHR, Bramon E, Børglum AD, Cichon S, Darvasi A, Ehrenreich H, Esko T, Gejman P V., Gill M, Gurling H, Hultman CM, Iwata N, Jablensky A V., Jonsson EG, Kendler KS, Kirov G, Knight J, Lencz T, Levinson DF, Li QS, Liu J, Malhotra AK, McCarroll SA, McQuillin A, Moran JL, Mortensen PB, Mowry BJ, Nothen MM, Ophoff RA, Owen MJ, Palotie A, Pato CN, Petryshen TL, Posthuma D, Rietschel M, Riley BP, Rujescu D, Sham PC, Sklar P, Clair DS, Weinberger DR, Wendland JR, Werge T, Daly MJ, Sullivan PF, O'Donovan MC, Rajagopal VM, Grove J, Mortensen PB, Pedersen CB, Pedersen MG, Nordentoft M, Bybjerg-Grauholm J, Bækvad-Hansen M, Hansen CS, Sullivan P, Devlin B, Sieberts SK, Cox NJ, Im HK (2019) Gene expression imputation across multiple brain regions provides insights into schizophrenia risk. Nat Genet 51:659–674. doi: 10.1038/s41588-019-0364-4

79. Choi SW, Mak TSH, O'Reilly PF (2020) Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc 2020 159 15:2759–2772. doi: 10.1038/s41596-020-0353-1

80. Lewis CM, Vassos E (2020) Polygenic risk scores: From research tools to clinical instruments. Genome Med 12:1–11. doi: 10.1186/S13073-020-00742-5/TABLES/2

81. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience 4. doi: 10.1186/S13742-015-0047-8

82. Choi SW, O'Reilly PF (2019) PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience 8. doi: 10.1093/GIGASCIENCE/GIZ082

83. Vilhjálmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S, Genovese G, Loh PR, Bhatia G, Do R, Hayeck T, Won HH, Neale BM, Corvin A, Walters JTR, Farh

KH, Holmans PA, Lee P, Bulik-Sullivan B, Collier DA, Huang H, Pers TH, Agartz I, Agerbo E, Albus M, Alexander M, Amin F, Bacanu SA, Begemann M, Belliveau RA, Bene J, Bergen SE, Bevilacqua E, Bigdeli TB, Black DW, Bruggeman R, Buccola NG, Buckner RL, Byerley W, Cahn W, Cai G, Campion D, Cantor RM, Carr VJ, Carrera N, Catts S V., Chambert KD, Chan RCK, Chen RYL, Chen EYH, Cheng W, Cheung EFC, Chong SA, Cloninger CR, Cohen D, Cohen N, Cormican P, Craddock N, Crowley JJ, Curtis D, Davidson M, Davis KL, Degenhardt F, Del Favero J, Delisi LE, Demontis D, Dikeos D, Dinan T, Djurovic S, Donohoe G, Drapeau E, Duan J, Dudbridge F, Durmishi N, Eichhammer P, Eriksson J, Escott-Price V, Essioux L, Fanous AH, Farrell MS, Frank J, Franke L, Freedman R, Freimer NB, Friedl M, Friedman JI, Fromer M, Georgieva L, Gershon ES, Giegling I, Giusti-Rodrguez P, Godard S, Goldstein JI, Golimbet V, Gopal S, Gratten J, Grove J, De Haan L, Hammer C, Hamshere ML, Hansen M, Hansen T, Haroutunian V, Hartmann AM, Henskens FA, Herms S, Hirschhorn JN, Hoffmann P, Hofman A, Hollegaard M V., Hougaard DM, Ikeda M, Joa I, Julia A, Kahn RS, Kalaydjieva L, Karachanak-Yankova S, Karjalainen J, Kavanagh D, Keller MC, Kelly BJ, Kennedy JL, Khrunin A, Kim Y, Klovins J, Knowles JA, Konte B, Kucinskas V, Kucinskiene ZA, Kuzelova-Ptackova H, Kahler AK, Laurent C, Keong JLC, Lee SH, Legge SE, Lerer B, Li M, Li T, Liang KY, Lieberman J, Limborska S, Loughland CM, Lubinski J, Lnnqvist J, Macek M, Magnusson PKE, Maher BS, Maier W, Mallet J, Marsal S, Mattheisen M, Mattingsdal M, McCarley RW, McDonald C, McIntosh AM, Meier S, Meijer CJ, Melegh B, Melle I, Mesholam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mokrab Y, Morris DW, Mors O, Mortensen PB, Murphy KC, Murray RM, Myin-Germeys I, Mller-Myhsok B, Nelis M, Nenadic I, Nertney DA, Nestadt G, Nicodemus KK, Nikitina-Zake L, Nisenbaum L, Nordin A, O'Callaghan E, O'Dushlaine C, O'Neill FA, Oh SY, Olincy A, Olsen L, Van Os J, Pantelis C,

Papadimitriou GN, Papiol S, Parkhomenko E, Pato MT, Paunio T, Pejovic-Milovancevic M, Perkins DO, Pietilinen O, Pimm J, Pocklington AJ, Powell J, Price A, Pulver AE, Purcell SM, Quested D, Rasmussen HB, Reichenberg A, Reimers MA, Richards AL, Roffman JL, Roussos P, Ruderfer DM, Salomaa V, Sanders AR, Schall U, Schubert CR, Schulze TG, Schwab SG, Scolnick EM, Scott RJ, Seidman LJ, Shi J, Sigurdsson E, Silagadze T, Silverman JM, Sim K, Slominsky P, Smoller JW, So HC, Spencer CCA, Stahl EA, Stefansson H, Steinberg S, Stogmann E, Straub RE, Strengman E, Strohmaier J, Stroup TS, Subramaniam M, Suvisaari J, Svrakic DM, Szatkiewicz JP, Sderman E, Thirumalai S, Toncheva D, Tooney PA, Tosato S, Veijola J, Waddington J, Walsh D, Wang D, Wang Q, Webb BT, Weiser M, Wildenauer DB, Williams NM, Williams S, Witt SH, Wolen AR, Wong EHM, Wormley BK, Wu JQ, Xi HS, Zai CC, Zheng X, Zimprich F, Wray NR, Stefansson K, Adolfsson R, Andreassen OA, Visscher PM, Blackwood DHR, Bramon E, Buxbaum JD, Børglum AD, Cichon S, Darvasi A, Domenici E, Ehrenreich H, Esko T, Gejman P V., Gill M, Gurling H, Hultman CM, Iwata N, Jablensky A V., Jonsson EG, Kendler KS, Kirov G, Knight J, Lencz T, Levinson DF, Li QS, Liu J, Malhotra AK, McCarroll SA, McQuillin A, Moran JL, Mowry BJ, Nthen MM, Ophoff RA, Owen MJ, Palotie A, Pato CN, Petryshen TL, Posthuma D, Rietschel M, Riley BP, Rujescu D, Sham PC, Sklar P, St. Clair D, Weinberger DR, Wendland JR, Werge T, Daly MJ, Sullivan PF, O'Donovan MC, Kraft P, Hunter DJ, Adank M, Ahsan H, Aittomäki K, Baglietto L, Berndt S, Blomquist C, Canzian F, Chang-Claude J, Chanock SJ, Crisponi L, Czene K, Dahmen N, Dos Santos Silva I, Easton D, Eliassen AH, Figueroa J, Fletcher O, Garcia-Closas M, Gaudet MM, Gibson L, Haiman CA, Hall P, Hazra A, Hein R, Henderson BE, Hopper JL, Irwanto A, Johansson M, Kaaks R, Kibriya MG, Lichtner P, Lund E, Makalic E, Meindl A, Meijers-Heijboer H, Müller-Myhsok B, Muranen TA, Nevanlinna H, Peeters PH, Peto J, Prentice

RL, Rahman N, Sánchez MJ, Schmidt DF, Schmutzler RK, Southey MC, Tamimi R, Travis R, Turnbull C, Uitterlinden AG, Van Der Luijt RB, Waisfisz Q, Wang Z, Whittemore AS, Yang R, Zheng W, Kathiresan S, Pato M, Pato C, Stahl E, Zaitlen N, Pasaniuc B, Kenny EE, Schierup MH, De Jager P, Patsopoulos NA, McCarroll S, Daly M, Purcell S, Chasman D, Neale B, Goddard M, Patterson N, Price AL (2015) Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. Am J Hum Genet 97:576–592. doi: 10.1016/J.AJHG.2015.09.001

84. Mak TSH, Porsch RM, Choi SW, Zhou X, Sham PC (2017) Polygenic scores via penalized regression on summary statistics. Genet Epidemiol 41:469–480. doi: 10.1002/GEPI.22050

85. Richardson TG, Harrison S, Hemani G, Smith GD (2019) An atlas of polygenic risk score associations to highlight putative causal relationships across the human phenome. Elife 8. doi: 10.7554/ELIFE.43657

86. Li L, Pang S, Zeng L, Güldener U, Schunkert H (2021) Genetically determined intelligence and coronary artery disease risk. Clin Res Cardiol 110:211–219. doi: 10.1007/s00392-020-01721-x

87. Wray NR, Lin T, Austin J, McGrath JJ, Hickie IB, Murray GK, Visscher PM (2021) From Basic Science to Clinical Application of Polygenic Risk Scores: A Primer. JAMA Psychiatry 78:101–109. doi: 10.1001/JAMAPSYCHIATRY.2020.3049

88. Fahed AC, Wang M, Homburger JR, Patel AP, Bick AG, Neben CL, Lai C, Brockman D, Philippakis A, Ellinor PT, Cassa CA, Lebo M, Ng K, Lander ES, Zhou AY, Kathiresan S, Khera A V. (2020) Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. Nat Commun 2020 111 11:1–9. doi: 10.1038/s41467-020-17374-3

89. Darst BF, Sheng X, Eeles RA, Kote-Jarai Z, Conti D V., Haiman CA (2021) Combined Effect of a Polygenic Risk Score and Rare Genetic Variants on Prostate Cancer Risk. Eur Urol 80:134–138. doi: 10.1016/j.eururo.2021.04.013

90. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, Matsunaga H, Ieki H, Ozaki K, Onouchi Y, Takahashi A, Nomura S, Morita H, Akazawa H, Kim C, Seo J sun, Higasa K, Iwasaki M, Yamaji T, Sawada N, Tsugane S, Koyama T, Ikezaki H, Takashima N, Tanaka K, Arisawa K, Kuriki K, Naito M, Wakai K, Suna S, Sakata Y, Sato H, Hori M, Sakata Y, Matsuda K, Murakami Y, Aburatani H, Kubo M, Matsuda F, Kamatani Y, Komuro I (2020) Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. Nat Genet 52:1169–1177. doi: 10.1038/s41588-020-0705-3

91. Krapohl E, Patel H, Newhouse S, Curtis CJ, Von Stumm S, Dale PS, Zabaneh D, Breen G, O'Reilly PF, Plomin R (2017) Multi-polygenic score approach to trait prediction. Mol Psychiatry 2018 235 23:1368–1374. doi: 10.1038/mp.2017.163

92. Tremblay J, Haloui M, Attaoua R, Tahir R, Hishmih C, Harvey F, Marois-Blanchet FC, Long C, Simon P, Santucci L, Hizel C, Chalmers J, Marre M, Harrap S, Cífková R, Krajčoviechová A, Matthews DR, Williams B, Poulter N, Zoungas S, Colagiuri S, Mancia G, Grobbee DE, Rodgers A, Liu L, Agbessi M, Bruat V, Favé MJ, Harwood MP, Awadalla P, Woodward M, Hussin JG, Hamet P (2021) Polygenic risk scores predict diabetes complications and their response to intensive blood pressure and glucose control. Diabetologia 64:2012–2025. doi: 10.1007/S00125-021-05491-7/FIGURES/5

93. Hess JL, Tylee DS, Mattheisen M, Adolfsson R, Agartz I, Agerbo E, Albus M, Alexander M, Amin F, Andreassen OA, Arranz MJ, Bacanu SA, Bakker S, Band G,

Barroso I, Begemann M, Bellenguez C, Belliveau RA, Bender S, Bene J, Bergen SE, Bevilacqua E, Bigdeli TB, Black DW, Blackburn H, Blackwell JM, Blackwood DHR, Bramon E, Brown MA, Bruggeman R, Buccola NG, Buckner RL, Bulik-Sullivan B, Bumpstead SJ, Buxbaum JD, Byerley W, Cahn W, Cai G, Campion D, Cantor RM, Carr VJ, Carrera N, Casas JP, Catts S V., Chambert KD, Chan RYL, Chan RCK, Chen EYH, Cheng W, Cheung EFC, Chong SA, Cichon S, Cloninger CR, Cohen D, Cohen N, Collier DA, Cormican P, Corvin A, Craddock N, Crespo-Facorro B, Crowley JJ, Daly MJ, Darvasi A, Davidson M, Davis KL, Degenhardt F, Del Favero J, Deloukas P, Demontis D, Dikeos D, Dinan T, Djurovic S, Domenici E, Donnelly P, Donohoe G, Drapeau E, Dronov S, Duan J, Dudbridge F, Duncanson A, Durmishi N, Edkins S, Ehrenreich H, Eichhammer P, Eriksson J, Escott-Price V, Esko T, Essioux L, Fanous AH, Farh KH, Farrell MS, Frank J, Franke L, Freedman R, Freeman C, Freimer NB, Friedl M, Friedman JI, Fromer M, Gejman P V., Genovese G, Georgieva L, Giannoulatou E, Giegling I, Gill M, Gillman M, Giusti-Rodríguez P, Godard S, Goldstein JI, Golimbet V, Gopal S, Gratten J, Gray E, Gurling H, Gwilliam R, de Haan L, Hall J, Hammer C, Hammond N, Hamshere ML, Hansen M, Hansen T, Haroutunian V, Hartmann AM, Hellenthal G, Henskens FA, Herms S, Hirschhorn JN, Hoffmann P, Hofman A, Hollegaard M V., Huang H, Hultman CM, Hunt SE, Ikeda M, Iwata N, Iyegbe C, Jablensky A V., Jankowski J, Jayakumar A, Joa I, Jönsson EG, Julià A, Kähler AK, Kahn RS, Kalaydjieva L, Karachanak-Yankova S, Karjalainen J, Kavanagh D, Keller MC, Kendler KS, Kennedy JL, Khrunin A, Kim Y, Kirov G, Klovins J, Knight J, Knowles JA, Konte B, Kucinskas V, Kucinskiene ZA, Kuzelova-Ptackova H, Langford C, Laurent C, Lawrie S, Lee SH, Lee P, Lee J, Legge SE, Lencz T, Lerer B, Levinson DF, Lewis CM, Li T, Li QS, Li M, Liang KY, Liddle J, Lieberman J, Limborska S, Lin K, Linszen DH, Liu J, Lönnqvist J, Loughland CM, Lubinski J, Macek M, Magnusson

PKE, Maher BS, Maier W, Malhotra AK, Mallet J, Markus HS, Marsal S, Mata I, Mathew CG, Mattingsdal M, McCann OT, McCarley RW, McCarroll SA, McCarthy MI, McDonald C, McIntosh AM, McQuillin A, Meier S, Meijer CJ, Melegh B, Melle I, Mesholam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mokrab Y, Moran JL, Morris DW, Mowry BJ, Müller-Myhsok B, Murphy KC, Murray RM, Myin-Germeys I, Neale BM, Nelis M, Nenadic I, Nertney DA, Nestadt G, Nicodemus KK, Nikitina-Zake L, Nisenbaum L, Nordin A, Nöthen MM, O'Callaghan E, O'Donovan MC, O'Dushlaine C, O'Neill FA, Oh SY, Olincy A, Olsen L, Ophoff RA, Van Os J, Owen MJ, Palmer CNA, Palotie A, Pantelis C, Papadimitriou GN, Papiol S, Parkhomenko E, Pato MT, Pato CN, Paunio T, Pearson R, Cairns MJ, DeLisi LE, Gershon ES, Kelly BJ, Lam M, Norgren N, Paciga SA, Tooney PA, Wu JQ, Pejovic-Milovancevic M, Perkins DO, Pers TH, Petryshen TL, Pietiläinen O, Pimm J, Pirinen M, Plomin R, Pocklington AJ, Posthuma D, Potter SC, Powell J, Price A, Pulver AE, Purcell SM, Quested D, Rasmussen HB, Rautanen A, Ravindrarajah R, Reichenberg A, Reimers MA, Richards AL, Ricketts M, Rietschel M, Riley BP, Ripke S, Roffman JL, Roussos P, Ruderfer DM, Rujescu D, Salomaa V, Sanders AR, Sawcer SJ, Schall U, Schubert CR, Schulze TG, Schwab SG, Scolnick EM, Scott RJ, Seidman LJ, Sham PC, Shi J, Sigurdsson E, Silagadze T, Silverman JM, Sim K, Sklar P, Slominsky P, Smoller JW, So HC, Söderman E, Spencer CCA, Clair DS, Stahl EA, Stogmann E, Strange A, Straub RE, Strengman E, Strohmaier J, Stroup TS, Su Z, Subramaniam M, Sullivan PF, Suvisaari J, Svrakic DM, Szatkiewicz JP, Tashakkori-Ghanbaria A, Thirumalai S, Toncheva D, Tosato S, Trembath RC, Veijola J, Visscher PM, Viswanathan AC, Vukcevic D, Waddington J, Waller M, Walsh D, Walshe M, Walters JTR, Wang Q, Wang D, Webb BT, Weinberger DR, Weisbrod M, Weiser M, Wendland JR, Weston P, Whittaker P, Widaa S, Wiersma D, Wildenauer DB, Williams S, Williams NM, Witt

SH, Wolen AR, Wong EHM, Wood NW, Wormley BK, Wray NR, Xi HS, Zai CC, Zheng X, Zimprich F, Børglum AD, Als TD, Grove J, Werge T, Mortensen PB, Mors O, Nordentoft M, Hougaard DM, Byberg-Grauholm J, Bækvad-Hansen M, Greenwood TA, Tsuang MT, Curtis D, Steinberg S, Sigurdsson E, Stefánsson H, Stefánsson K, Edenberg HJ, Holmans P, Faraone S V., Glatt SJ (2019) A polygenic resilience score moderates the genetic risk for schizophrenia. Mol Psychiatry 2019 263 26:800–815. doi: 10.1038/s41380-019-0463-8

94. Chen Z, Schunkert H (2021) Genetics of coronary artery disease in the post-GWAS era. J Intern Med. doi: 10.1111/JOIM.13362

95. Lusis AJ, Seldin MM, Allayee H, Bennett BJ, Civelek M, Davis RC, Eskin E, Farber CR, Hui S, Mehrabian M, Norheim F, Pan C, Parks B, Rau CD, Smith DJ, Vallim T, Wang Y, Wang J (2016) The Hybrid Mouse Diversity Panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. J Lipid Res 57:925–942. doi: 10.1194/jlr.R066944

96. Brown SDM, Moore MW (2012) Towards an encyclopaedia of mammalian gene function: the International Mouse Phenotyping Consortium. Dis Model Mech 5:289–292. doi: 10.1242/DMM.009878

97. Shaw DR (2016) Searching the mouse genome informatics (MGI) resources for information on mouse biology from genotype to phenotype. Curr Protoc Bioinforma 2016:1.7.1-1.7.16. doi: 10.1002/cpbi.18

98. Hao K, Ermel R, Sukhavasi K, Cheng H, Ma L, Li L, Amadori L, Koplev S, Franzén O, d'Escamard V, Chandel N, Wolhuter K, Bryce NS, Venkata VRM, Miller CL, Ruusalepp A, Schunkert H, Björkegren JLM, Kovacic JC (2021) Integrative Prioritization of Causal Genes for Coronary Artery Disease. Circ Genomic Precis Med.

doi: 10.1161/CIRCGEN.121.003365

99. Kott KA, Vernon ST, Hansen T, de Dreu M, Das SK, Powell J, de St Groth BF, Di Bartolo BA, McGuire HM, Figtree GA (2020) Single-cell immune profiling in coronary artery disease: The role of state-of-the-art immunophenotyping with mass cytometry in the diagnosis of atherosclerosis. J Am Heart Assoc 9. doi: 10.1161/JAHA.120.017759

100. McGarrah RW, Crown SB, Zhang GF, Shah SH, Newgard CB (2018) Cardiovascular Metabolomics. Circ Res 122:1238–1258. doi: 10.1161/CIRCRESAHA.117.311002

101. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, König IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Brænne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H (2007) Genomewide association analysis of coronary artery disease. N Engl J Med 357:443–453. doi: 10.1056/NEJMoa072366

102. Erdmann J, Großhennig A, Braund PS, König IR, Hengstenberg C, Hall AS, Linsel-Nitschke P, Kathiresan S, Wright B, Trégouët DA, Cambien F, Bruse P, Aherrahrou Z, Wagner AK, Stark K, Schwartz SM, Salomaa V, Elosua R, Melander O, Voight BF, O'Donnell CJ, Peltonen L, Siscovick DS, Altshuler D, Merlini PA, Peyvandi F, Bernardinelli L, Ardissino D, Schillert A, Blankenberg S, Zeller T, Wild P, Schwarz DF, Tiret L, Perret C, Schreiber S, Mokhtari NE El, Schäfer A, März W, Renner W, Bugert P, Klüter H, Schrezenmeir J, Rubin D, Ball SG, Balmforth AJ, Wichmann HE, Meitinger T, Fischer M, Meisinger C, Baumert J, Peters A, Ouwehand WH, Deloukas P, Thompson JR, Ziegler A, Samani NJ, Schunkert H (2009) New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet 41:280–282. doi:

10.1038/ng.307

103. Erdmann J, Willenborg C, Nahrstaedt J, Preuss M, Konig IR, Baumert J, Linsel-Nitschke P, Gieger C, Tennstedt S, Belcredi P, Aherrahrou Z, Klopp N, Loley C, Stark K, Hengstenberg C, Bruse P, Freyer J, Wagner AK, Medack A, Lieb W, Grosshennig A, Sager HB, Reinhardt A, Schafer A, Schreiber S, El Mokhtari NE, Raaz-Schrauder D, Illig T, Garlichs CD, Ekici AB, Reis A, Schrezenmeir J, Rubin D, Ziegler A, Wichmann H-E, Doering A, Meisinger C, Meitinger T, Peters A, Schunkert H (2011) Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. Eur Heart J 32:158–168. doi: 10.1093/eurheartj/ehq405

104. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, CHopewell J, Webb TR, Zeng L, Dehghan A, Alver M, MArmasu S, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikäinen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith A V., Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, De Andrade M, De Vries PS, Van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, König IR, Lannfelt L, Lieb W, Lind L, MLindgren C, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FUR, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, EStirrups K, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples L, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, SHall A, Hamsten A, Harris TB, LHazen S,

Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimäki T, Loos RJF, Melander O, Metspalu A, März W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M (2015) A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 47:1121–1130. doi: 10.1038/ng.3396

105. Stitziel NO, Won HH, Morrison AC, Peloso GM, Do R, Lange LA, Fontanillas P, Gupta N, Duga S, Goel A, Farrall M, Saleheen D, Ferrario P, König I, Asselta R, Merlini PA, Marziliano N, Notarangelo MF, Schick U, Auer P, Assimes TL, Reilly M, Wilensky R, Rader DJ, Kees Hovingh G, Meitinger T, Kessler T, Kastrati A, Laugwitz KL, Siscovick D, Rotter JI, Hazen SL, Tracy R, Cresci S, Spertus J, Jackson R, Schwartz SM, Natarajan P, Crosby J, Muzny D, Ballantyne C, Rich SS, O'Donnell CJ, Abecasis G, Sunyaev S, Nickerson DA, Buring JE, Ridker PM, Chasman DI, Austin E, Ye Z, Kullo IJ, Weeke PE, Shaffer CM, Bastarache LA, Denny JC, Roden DM, Palmer C, Deloukas P, Lin DY, Tang ZZ, Erdmann J, Schunkert H, Danesh J, Marrugat J, Elosua R, Ardissino D, McPherson R, Watkins H, Reiner AP, Wilson JG, Altshuler D, Gibbs RA, Lander ES, Boerwinkle E, Gabriel S, Kathiresan S (2014) Inactivating mutations in NPC1L1 and protection from coronary heart disease. N Engl J Med 371:2072–2082. doi: 10.1056/NEJMoa1405386

106. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, Zeng L, Ntalla I, Lai FY, Hopewell JC, Giannakopoulou O, Jiang T, Hamby SE, Di Angelantonio E, Assimes TL, Bottinger EP, Chambers JC, Clarke R, Palmer CNA, Cubbon RM, Ellinor

P, Ermel R, Evangelou E, Franks PW, Grace C, Gu D, Hingorani AD, Howson JMM, Ingelsson E, Kastrati A, Kessler T, Kyriakou T, Lehtimäki T, Lu X, Lu Y, März W, McPherson R, Metspalu A, Pujades-Rodriguez M, Ruusalepp A, Schadt EE, Schmidt AF, Sweeting MJ, Zalloua PA, Alghalayini K, Keavney BD, Kooner JS, Loos RJF, Patel RS, Rutter MK, Tomaszewski M, Tzoulaki I, Zeggini E, Erdmann J, Dedoussis G, Björkegren JLM, Schunkert H, Farrall M, Danesh J, Samani NJ, Watkins H, Deloukas P (2017) Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet 49:1385–1391. doi: 10.1038/ng.3913

107. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, Todd JA, Donnelly P, Barrett JC, Davison D, Easton D, Evans D, Leung HT, Marchini JL, Morris AP, Spencer CCA, Tobin MD, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse C V., Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Breen G, St. Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Hamshere ML, Holmans PA, Jones IR, Kirov G, Moskvina V, Nikolov I, O'Donovan MC, Owen MJ, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, Young AH, Ferrier IN, Ball SG, Balmforth AJ, Barrett JH, Bishop DT, Iles MM, Maqbool A, Yuldasheva N, Hall AS, Braund PS, Dixon RJ, Mangino M, Stevens S, Thompson JR, Bredin F, Tremelling M, Parkes M, Drummond H, Lees CW, Nimmo ER, Satsangi J, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Mathew CG, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, Ahmad T, Cummings FR, Jewell DP, Webster J, Brown MJ, Lathrop GM, Connell J, Dominiczak A, Braga Marcano CA, Burke B, Dobson R, Gungadoo J, Lee KL, Munroe PB, Newhouse SJ, Onipinla A, Wallace C, Xue M, Caulfield M, Farrall M, Barton A, Bruce

IN, Donovan H, Eyre S, Gilbert PD, Hider SL, Hinks AM, John SL, Potter C, Silman AJ, Symmons DPM, Thomson W, Worthington J, Dunger DB, Widmer B, Frayling TM, Freathy RM, Lango H, Perry JRB, Shields BM, Weedon MN, Hattersley AT, Hitman GA, Walker M, Elliott KS, Groves CJ, Lindgren CM, Rayner NW, Timpson NJ, Zeggini E, Newport M, Sirugo G, Lyons E, Vannberg F, Hill AVS, Bradbury LA, Farrar C, Pointon JJ, Wordsworth P, Brown MA, Franklyn JA, Heward JM, Simmonds MJ, Gough SCL, Seal S, Stratton MR, Rahman N, Ban SM, Goris A, Sawcer SJ, Compston A, Conway D, Jallow M, Rockett KA, Bumpstead SJ, Chaney A, Downes K, Ghori MJR, Gwilliam R, Hunt SE, Inouye M, Keniry A, King E, McGinnis R, Potter S, Ravindrarajah R, Whittaker P, Widden C, Withers D, Cardin NJ, Ferreira T, Pereira-Gale J, Hallgrimsdóttir IB, Howie BN, Spencer CCA, Su Z, Teo YY, Vukcevic D, Bentley D, Compston A (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678. doi: 10.1038/nature05911

108. Winkelmann BR, März W, Boehm BO, Zotz R, Hager J, Hellstern P, Senges J (2001) Rationale and design of the LURIC study - A resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. Pharmacogenomics 2. doi: 10.1517/14622416.2.1.s1

109. Anderson CD, Rosand J, Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles G V, Clark NR, Ma'ayan A, Kim JY, Kim JY, Song KS, Lee YH, Seo JS, Jelinek J, Goldschmidt-Clermont PJ, Issa JP, Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW, Krizsan-Agbas D, Pedchenko T, Smith PG, Li C, Bazzano LA, Rao DC, Hixson JE, He J, Gu D, Gu CC, Shimmin LC, Jaquish CE, Schwander K, Liu DP, Huang J, Lu F, Cao J, Chong S, Lu X, Kelly TN, Sharma P, Kumar J, Garg G, Kumar A,

Patowary A, Karthikeyan G, Ramakrishnan L, Brahmachari V, Sengupta S, Pedrinelli R, Ballo P, Fiorentini C, Denti S, Galderisi M, Ganau A, Germanò G, Innelli P, Paini A, Perlini S, Salvetti M, Zacà V, Besingi W, Johansson Å, Cao TH, Quinn PA, Sandhu JK, Voors AA, Lang CC, Parry HM, Mohan M, Jones DJL, Ng LL, Heart N, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups KK, König IR, Cazier J-B, Johansson Å, Hall AS, Lee J-YJ-Y, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikäinen L-P, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R, Doney ASF, Mokhtari NEE El, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han B-G, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki M-L, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Müller-Nurasyid M, Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schäfer A, Sivananthan M, Song C, Stewart AFR, Tan S-T, Thorgeirsson G, Schoot CE van der, Wagner PJ, Wells G a., Wild PS, Yang T-P, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Brambilla P, Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM, Ferrières J, Gauguier D, Go AS, Goodall AH, Gudnason VV, Hazen SL, Holm H, Iribarren C, Jang Y, Kähönen M, Kee F, Kim H-S, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R, Lee J-YJ-Y, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Trégouët

D-A, Virtamo J, Wallentin L, Wareham NJN, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS, Pastinen TM, Syvänen A-C, Hovingh GK, Dedoussis G, Franks PW, Lehtimäki T, Metspalu A, Zalloua PA, Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell CJ, Reilly MP, März W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdottir U, Danesh J, Palmer CNA, Roberts R, Watkins H, Schunkert H, Samani NJ, Willer CJ, Schmidt EM, Al E, Facp AFR, Rodriguez-lopez L, Vargas-ayala G, Mc SH, Serna DC, Lozano-nuevo JJ, Rubio-guerra AF, Studies TIC for BPG-WA, Franks WT, Zhou DH, Wylie BJ, Money BG, Graesser DT, Frericks HL, Sahota G, Rienstra CM, Rakyan VK, Down TA, Balding DJ, Beck S, Hager J, Kamatani Y, Cazier J-B, Youhanna S, Ghassibe-Sabbagh M, Platt DE, Abchee AB, Romanos J, Khazen G, Othman R, Badro DA, Haber M, Salloum AK, Douaihy B, Shasha N, Kabbani S, Sbeite H, Chammas E, Bayeh H el, Rousseau F, Zelenika D, Gut I, Lathrop M, Farrall M, Gauguier D, Zalloua PA, Myocardial Infarction Genetics C, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand SS, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario MM, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Yee J, Friedlander Y, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Meigs JB, Williams G, Nathan DM, MacRae CA, Havulinna AS, Berglund G, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Daly MJ, Nemesh J, Korn JM, McCarroll SA, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall AS, Wellcome Trust Case Control C, Linsel-Nitschke P, Lieb W, Ziegler A, Konig IR,

Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann H-EE, Schreiber S, Ouwehand WH, Deloukas P, Scholz M, Cambien F, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney JM, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser V, Epstein SE, Scheffold T, Berger K, Huge A, Martinelli N, Olivieri O, Corrocher R, McKeown PP, Erdmann E, Konig IR, Holm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Do R, Xie C, Siscovick DS, Liang L, Willis-Owen SAG, Laprise C, Wong KCC, Davies GA, Hudson TJ, Binia A, Hopkin JM, Yang I V., Grundberg E, Busche S, Hudson M, Rönnblom L, Pastinen TM, Schwartz DA, Lathrop GM, Moffatt MF, Cookson WOCM, Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM, Mayer B, Erdmann J, Schunkert H, Myocardial Infarction Genetics C, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand SS, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario MM, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Yee J, Friedlander Y, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Meigs JB, Williams G, Nathan DM, MacRae CA, Havulinna AS, Berglund G, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Daly MJ, Nemesh J, Korn JM, McCarroll SA, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall AS, Wellcome

Trust Case Control C, Linsel-Nitschke P, Lieb W, Ziegler A, Konig IR, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann H-EE, Schreiber S, Ouwehand WH, Deloukas P, Scholz M, Cambien F, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney JM, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser V, Epstein SE, Scheffold T, Berger K, Huge A, Martinelli N, Olivieri O, Corrocher R, McKeown PP, Erdmann E, Konig IR, Holm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Do R, Xie C, Siscovick DS, The CARDIoGRAMplusC4D Consortium, Studies TIC for BPG-WA, The CARDIoGRAMplusC4D Consortium, Preuss M, König IR, Thompson JR, Erdmann J, Absher D, Assimes TL, Blankenberg SS, Boerwinkle E, Chen L, Cupples LA, Hall AS, Halperin E, Hengstenberg C, Holm H, Laaksonen R, Li M, Marz W, McPherson R, Musunuru K, Nelson CP, Burnett MS, Epstein SE, O'Donnell CJ, Quertermous T, Rader DJ, Roberts R, Schillert A, Stefansson K, Stewart AFR, Thorleifsson G, Voight BF, Wells G a., Ziegler A, Kathiresan S, Reilly MP, Samani NJ, Schunkert H, Avenell A, Broom J, Brown TJ, Poobalan A, Aucott L, Stearns SC, Smith WCS, Jung RT, Campbell MK, Grant AM, Rakyan VK, Down TA, Balding DJ, Beck S, Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AFR, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NEE El, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason VV, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD,

Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJP, Khaw K-T, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups KK, Stoll M, Tang WHW, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJN, Wells G a., Wichmann H-EE, Wild PS, Willenborg C, Witteman JCM, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg SS, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J, Samani NJ, Sharma P, Garg G, Kumar A, Mohammad F, Ramesh S, Slatkin M, Myocardial T, Genetics I, Investigators C, Udali S, Guarini P, Moruzzi S, Choi S, Friso S, Schübeler D, Setten J Van, Isgum I, Smolonska J, Ripke S, Jong PA De, Oudkerk M, Koning H De, Lammers JJ, Zanen P, Groen HJM, Boezen HM, Postma DS, Wijmenga C, Viergever MA, Th WP, Bakker PIW De (2015) Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 478:103–109. doi: 10.1038/nrg3000.Epigenome-Wide

110. O'Connell J, Gurdasani D, Delaneau O, Pirastu N, Ulivi S, Cocca M, Traglia M, Huang J, Huffman JE, Rudan I, McQuillan R, Fraser RM, Campbell H, Polasek O, Asiki G,

Ekoru K, Hayward C, Wright AF, Vitart V, Navarro P, Zagury JF, Wilson JF, Toniolo D, Gasparini P, Soranzo N, Sandhu MS, Marchini J (2014) A General Approach for Haplotype Phasing across the Full Spectrum of Relatedness. PLoS Genet 10:1004234. doi: 10.1371/journal.pgen.1004234

111. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5:1000529. doi: 10.1371/journal.pgen.1000529

112. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, De Leeuw CA, Nagel M, Awasthi S, Barr PB, Coleman JRI, Grasby KL, Hammerschlag AR, Kaminski JA, Karlsson R, Krapohl E, Lam M, Nygaard M, Reynolds CA, Trampush JW, Young H, Zabaneh D, Hägg S, Hansell NK, Karlsson IK, Linnarsson S, Montgomery GW, Muñoz-Manchado AB, Quinlan EB, Schumann G, Skene NG, Webb BT, White T, Arking DE, Avramopoulos D, Bilder RM, Bitsios P, Burdick KE, Cannon TD, Chiba-Falek O, Christoforou A, Cirulli ET, Congdon E, Corvin A, Davies G, Deary IJ, Derosse P, Dickinson D, Djurovic S, Donohoe G, Conley ED, Eriksson JG, Espeseth T, Freimer NA, Giakoumaki S, Giegling I, Gill M, Glahn DC, Hariri AR, Hatzimanolis A, Keller MC, Knowles E, Koltai D, Konte B, Lahti J, Le Hellard S, Lencz T, Liewald DC, London E, Lundervold AJ, Malhotra AK, Melle I, Morris D, Need AC, Ollier W, Palotie A, Payton A, Pendleton N, Poldrack RA, Räikkönen K, Reinvang I, Roussos P, Rujescu D, Sabb FW, Scult MA, Smeland OB, Smyrnis N, Starr JM, Steen VM, Stefanis NC, Straub RE, Sundet K, Tiemeier H, Voineskos AN, Weinberger DR, Widen E, Yu J, Abecasis G, Andreassen OA, Breen G, Christiansen L, Debrabant B, Dick DM, Heinz A, Hjerling-Leffler J, Ikram MA, Kendler KS, Martin NG, Medland SE, Pedersen NL, Plomin R, Polderman TJC, Ripke S, Van Der Sluis S, Sullivan PF, Vrieze SI, Wright

MJ, Posthuma D (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet 50:912–919. doi: 10.1038/s41588-018-0152-6

113. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, Nguyen-Viet TA, Bowers P, Sidorenko J, Karlsson Linnér R, Fontana MA, Kundu T, Lee C, Li H, Li R, Royer R, Timshel PN, Walters RK, Willoughby EA, Yengo L, Agee M, Alipanahi B, Auton A, Bell RK, Bryc K, Elson SL, Fontanillas P, Hinds DA, McCreight JC, Huber KE, Litterman NK, McIntyre MH, Mountain JL, Noblin ES, Northover CAM, Pitts SJ, Sathirapongsasuti JF, Sazonova O V., Shelton JF, Shringarpure S, Tian C, Vacic V, Wilson CH, Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, Turley P, Chen GB, Emilsson V, Meddens SFW, Oskarsson S, Pickrell JK, Thom K, Timshel P, Vlaming R de, Abdellaoui A, Ahluwalia TS, Bacelis J, Baumbach C, Bjornsdottir G, Brandsma JH, Concas MP, Derringer J, Furlotte NA, Galesloot TE, Girotto G, Gupta R, Hall LM, Harris SE, Hofer E, Horikoshi M, Huffman JE, Kaasik K, Kalafati IP, Karlsson R, Kong A, Lahti J, van der Lee SJ, Leeuw C de, Lind PA, Lindgren KO, Liu T, Mangino M, Marten J, Mihailov E, Miller MB, van der Most PJ, Oldmeadow C, Payton A, Pervjakova N, Peyrot WJ, Qian Y, Raitakari O, Rueedi R, Salvi E, Schmidt B, Schraut KE, Shi J, Smith A V., Poot RA, St Pourcain B, Teumer A, Thorleifsson G, Verweij N, Vuckovic D, Wellmann J, Westra HJ, Yang J, Zhao W, Zhu Z, Alizadeh BZ, Amin N, Bakshi A, Baumeister SE, Biino G, Bønnelykke K, Boyle PA, Campbell H, Cappuccio FP, Davies G, De Neve JE, Deloukas P, Demuth I, Ding J, Eibich P, Eisele L, Eklund N, Evans DM, Faul JD, Feitosa MF, Forstner AJ, Gandin I, Gunnarsson B, Halldórsson B V., Harris TB, Heath AC, Hocking LJ, Holliday EG, Homuth G, Horan MA, Hottenga JJ, de Jager PL, Joshi PK, Jugessur A, Kaakinen MA, Kähönen M, Kanoni S, Keltigangas-Järvinen L, Kiemeney LALM, Kolcic I, Koskinen

S, Kraja AT, Kroh M, Kutalik Z, Latvala A, Launer LJ, Lebreton MP, Levinson DF, Lichtenstein P, Lichtner P, Liewald DCM, Loukola LLCSA, Madden PA, Mägi R, Mäki-Opas T, Marioni RE, Marques-Vidal P, Meddens GA, McMahon G, Meisinger C, Meitinger T, Milaneschi Y, Milani L, Montgomery GW, Myhre R, Nelson CP, Nyholt DR, Ollier WER, Palotie A, Paternoster L, Pedersen NL, Petrovic KE, Porteous DJ, Räikkönen K, Ring SM, Robino A, Rostapshova O, Rudan I, Rustichini A, Salomaa V, Sanders AR, Sarin AP, Schmidt H, Scott RJ, Smith BH, Smith JA, Staessen JA, Steinhagen-Thiessen E, Strauch K, Terracciano A, Tobin MD, Ulivi S, Vaccargiu S, Quaye L, van Rooij FJA, Venturini C, Vinkhuyzen AAE, Völker U, Völzke H, Vonk JM, Vozzi D, Waage J, Ware EB, Willemsen G, Attia JR, Bennett DA, Berger K, Bertram L, Bisgaard H, Boomsma DI, Borecki IB, Bültmann U, Chabris CF, Cucca F, Cusi D, Deary IJ, Dedoussis G V., van Duijn CM, Eriksson JG, Franke B, Franke L, Gasparini P, Gejman P V., Gieger C, Grabe HJ, Gratten J, Groenen PJF, Gudnason V, van der Harst P, Hayward C, Hinds DA, Hoffmann W, Hyppönen E, Iacono WG, Jacobsson B, Järvelin MR, Jöckel KH, Kaprio J, Kardia SLR, Lehtimäki T, Lehrer SF, Magnusson PKE, Martin NG, McGue M, Metspalu A, Pendleton N, Penninx BWJH, Perola M, Pirastu N, Pirastu M, Polasek O, Posthuma D, Power C, Province MA, Samani NJ, Schlessinger D, Schmidt R, Sørensen TIA, Spector TD, Stefansson K, Thorsteinsdottir U, Thurik AR, Timpson NJ, Tiemeier H, Tung JY, Uitterlinden AG, Vitart V, Vollenweider P, Weir DR, Wilson JF, Wright AF, Conley DC, Krueger RF, Smith GD, Hofman A, Laibson DI, Medland SE, Meyer MN, Yang J, Johannesson M, Visscher PM, Esko T, Koellinger PD, Cesarini D, Benjamin DJ, Alver M, Bao Y, Clark DW, Day FR, Furlotte NA, Joshi PK, Kemper KE, Kleinman A, Langenberg C, Mägi R, Trampush JW, Verma SS, Wu Y, Lam M, Zhao JH, Zheng Z, Boardman JD, Campbell H, Freese J, Harris KM, Hayward C, Herd P, Kumari M, Lencz T, Luan J,

Malhotra AK, Metspalu A, Milani L, Ong KK, Perry JRB, Porteous DJ, Ritchie MD, Smart MC, Smith BH, Tung JY, Wareham NJ, Wilson JF, Beauchamp JP, Conley DC, Esko T, Lehrer SF, Magnusson PKE, Oskarsson S, Pers TH, Robinson MR, Thom K, Watson C, Chabris CF, Meyer MN, Laibson DI, Yang J, Johannesson M, Koellinger PD, Turley P, Visscher PM, Benjamin DJ, Cesarini D (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 50:1112–1121. doi: 10.1038/s41588-018-0147-3

114. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, Yengo L, Ferreira T, Marouli E, Ji Y, Yang J, Jones S, Beaumont R, Croteau-Chonka DC, Winkler TW, Hattersley AT, Loos RJF, Hirschhorn JN, Visscher PM, Frayling TM, Yaghootkar H, Lindgren CM (2019) Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet 28:166–174. doi: 10.1093/hmg/ddy327

115. Furberg H, Kim Y, Dackor J, Boerwinkle E, Franceschini N, Ardissino D, Bernardinelli L, Mannucci PM, Mauri F, Merlini PA, Absher D, Assimes TL, Fortmann SP, Iribarren C, Knowles JW, Quertermous T, Ferrucci L, Tanaka T, Bis JC, Furberg CD, Haritunians T, McKnight B, Psaty BM, Taylor KD, Thacker EL, Almgren P, Groop L, Ladenvall C, Boehnke M, Jackson AU, Mohlke KL, Stringham HM, Tuomilehto J, Benjamin EJ, Hwang SJ, Levy D, Preis SR, Vasan RS, Duan J, Gejman P V., Levinson DF, Sanders AR, Shi J, Lips EH, McKay JD, Agudo A, Barzan L, Bencko V, Benhamou S, Castellsagué X, Canova C, Conway DI, Fabianova E, Foretova L, Janout V, Healy CM, Holcátová I, Kjaerheim K, Lagiou P, Lissowska J, Lowry R, MacFarlane T V., Mates D, Richiardi L, Rudnai P, Szeszenia-Dabrowska N, Zaridze D, Znaor A, Lathrop M, Brennan P, Bandinelli S, Frayling TM, Guralnik JM, Milaneschi Y, Perry JRB, Altshuler

D, Elosua R, Kathiresan S, Lucas G, Melander O, O'Donnell CJ, Salomaa V, Schwartz SM, Voight BF, Penninx BW, Smit JH, Vogelzangs N, Boomsma DI, De Geus EJC, Vink JM, Willemsen G, Chanock SJ, Gu F, Hankinson SE, Hunter DJ, Hofman A, Tiemeier H, Uitterlinden AG, Van Duijn CM, Walter S, Chasman DI, Everett BM, Paré G, Ridker PM, Li MD, Maes HH, Audrain-Mcgovern J, Posthuma D, Thornton LM, Lerman C, Kaprio J, Rose JE, Ioannidis JPA, Kraft P, Lin DY, Sullivan PF (2010) Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet 42:441–447. doi: 10.1038/ng.571

116. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, Pervjakova N, Pers TH, Johnson AD, Eicher JD, Jackson AU, Ferreira T, Lee Y, Ma C, Steinthorsdottir V, Thorleifsson G, Qi L, Van Zuydam NR, Mahajan A, Chen H, Almgren P, Voight BF, Grallert H, Müller-Nurasyid M, Ried JS, Rayner NW, Robertson N, Karssen LC, van Leeuwen EM, Willems SM, Fuchsberger C, Kwan P, Teslovich TM, Chanda P, Li M, Lu Y, Dina C, Thuillier D, Yengo L, Jiang L, Sparso T, Kestler HA, Chheda H, Eisele L, Gustafsson S, Frånberg M, Strawbridge RJ, Benediktsson R, Hreidarsson AB, Kong A, Sigurðsson G, Kerrison ND, Luan J, Liang L, Meitinger T, Roden M, Thorand B, Esko T, Mihailov E, Fox C, Liu CT, Rybin D, Isomaa B, Lyssenko V, Tuomi T, Couper DJ, Pankow JS, Grarup N, Have CT, Jørgensen ME, Jørgensen T, Linneberg A, Cornelis MC, van Dam RM, Hunter DJ, Kraft P, Sun Q, Edkins S, Owen KR, Perry JRB, Wood AR, Zeggini E, Tajes-Fernandes J, Abecasis GR, Bonnycastle LL, Chines PS, Stringham HM, Koistinen HA, Kinnunen L, Sennblad B, Mühleisen TW, Nöthen MM, Pechlivanis S, Baldassarre D, Gertow K, Humphries SE, Tremoli E, Klopp N, Meyer J, Steinbach G, Wennauer R, Eriksson JG, Männistö S, Peltonen L, Tikkanen E, Charpentier G, Eury E, Lobbens S, Gigante B, Leander K, McLeod O, Bottinger EP, Gottesman O, Ruderfer D, Blüher M, Kovacs P, Tonjes A, Maruthur NM, Scapoli C, Erbel R, Jöckel KH,

Moebus S, de Faire U, Hamsten A, Stumvoll M, Deloukas P, Donnelly PJ, Frayling TM, Hattersley AT, Ripatti S, Salomaa V, Pedersen NL, Boehm BO, Bergman RN, Collins FS, Mohlke KL, Tuomilehto J, Hansen T, Pedersen O, Barroso I, Lannfelt L, Ingelsson E, Lind L, Lindgren CM, Cauchi S, Froguel P, Loos RJF, Balkau B, Boeing H, Franks PW, Barricarte Gurrea A, Palli D, van der Schouw YT, Altshuler D, Groop LC, Langenberg C, Wareham NJ, Sijbrands E, van Duijn CM, Florez JC, Meigs JB, Boerwinkle E, Gieger C, Strauch K, Metspalu A, Morris AD, Palmer CNA, Hu FB, Thorsteinsdottir U, Stefansson K, Dupuis J, Morris AP, Boehnke M, McCarthy MI, Prokopenko I (2017) An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes 66:2888–2902. doi: 10.2337/db16-1253

117. De Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015) MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLoS Comput Biol 11:1004219. doi: 10.1371/journal.pcbi.1004219

118. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J (2009) ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25:1091–1093. doi: 10.1093/bioinformatics/btp101

119. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, Furlong LI (2020) The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Res 48:D845–D855. doi: 10.1093/nar/gkz1021

120. Graham SE, Clarke SL, Wu K-HH, Kanoni S, Zajac GJM, Ramdas S, Surakka I, Ntalla I, Vedantam S, Winkler TW, Locke AE, Marouli E, Hwang MY, Han S, Narita A, Choudhury A, Bentley AR, Ekoru K, Verma A, Trivedi B, Martin HC, Hunt KA, Hui Q, Klarin D, Zhu X, Thorleifsson G, Helgadottir A, Gudbjartsson DF, Holm H, Olafsson

I, Akiyama M, Sakaue S, Terao C, Kanai M, Zhou W, Brumpton BM, Rasheed H, Ruotsalainen SE, Havulinna AS, Veturi Y, Feng Q, Rosenthal EA, Lingren T, Pacheco JA, Pendergrass SA, Haessler J, Giulianini F, Bradford Y, Miller JE, Campbell A, Lin K, Millwood IY, Hindy G, Rasheed A, Faul JD, Zhao W, Weir DR, Turman C, Huang H, Graff M, Mahajan A, Brown MR, Zhang W, Yu K, Schmidt EM, Pandit A, Gustafsson S, Yin X, Luan J, Zhao J-H, Matsuda F, Jang H-M, Yoon K, Medina-Gomez C, Pitsillides A, Hottenga JJ, Willemsen G, Wood AR, Ji Y, Gao Z, Haworth S, Mitchell RE, Chai JF, Aadahl M, Yao J, Manichaikul A, Warren HR, Ramirez J, Bork-Jensen J, Kårhus LL, Goel A, Sabater-Lleal M, Noordam R, Sidore C, Fiorillo E, McDaid AF, Marques-Vidal P, Wielscher M, Trompet S, Sattar N, Møllehave LT, Thuesen BH, Munz M, Zeng L, Huang J, Yang B, Poveda A, Kurbasic A, Lamina C, Forer L, Scholz M, Galesloot TE, Bradfield JP, Daw EW, Zmuda JM, Mitchell JS, Fuchsberger C, Christensen H, Brody JA, Feitosa MF, Wojczynski MK, Preuss M, Mangino M, Christofidou P, Verweij N, Benjamins JW, Engmann J, Kember RL, Slieker RC, Lo KS, Zilhao NR, Le P, Kleber ME, Delgado GE, Huo S, Ikeda DD, Iha H, Yang J, Liu J, Leonard HL, Marten J, Schmidt B, Arendt M, Smyth LJ, Cañadas-Garre M, Wang C, Nakatochi M, Wong A, Hutri-Kähönen N, Sim X, Xia R, Huerta-Chagoya A, Fernandez-Lopez JC, Lyssenko V, Ahmed M, Jackson AU, Irvin MR, Oldmeadow C, Kim H-N, Ryu S, Timmers PRHJ, Arbeeva L, Dorajoo R, Lange LA, Chai X, Prasad G, Lorés-Motta L, Pauper M, Long J, Li X, Theusch E, Takeuchi F, Spracklen CN, Loukola A, Bollepalli S, Warner SC, Wang YX, Wei WB, Nutile T, Ruggiero D, Sung YJ, Hung Y-J, Chen S, Liu F, Yang J, Kentistou KA, Gorski M, Brumat M, Meidtner K, Bielak LF, Smith JA, Hebbar P, Farmaki A-E, Hofer E, Lin M, Xue C, Zhang J, Concas MP, Vaccargiu S, van der Most PJ, Pitkänen N, Cade BE, Lee J, van der Laan SW, Chitrala KN, Weiss S, Zimmermann ME, Lee JY, Choi HS, Nethander M, Freitag-Wolf S,

Southam L, Rayner NW, Wang CA, Lin S-Y, Wang J-S, Couture C, Lyytikäinen L-P, Nikus K, Cuellar-Partida G, Vestergaard H, Hildalgo B, Giannakopoulou O, Cai Q, Obura MO, van Setten J, Li X, Schwander K, Terzikhan N, Shin JH, Jackson RD, Reiner AP, Martin LW, Chen Z, Li L, Highland HM, Young KL, Kawaguchi T, Thiery J, Bis JC, Nadkarni GN, Launer LJ, Li H, Nalls MA, Raitakari OT, Ichihara S, Wild SH, Nelson CP, Campbell H, Jäger S, Nabika T, Al-Mulla F, Niinikoski H, Braund PS, Kolcic I, Kovacs P, Giardoglou T, Katsuya T, Bhatti KF, de Kleijn D, de Borst GJ, Kim EK, Adams HHH, Ikram MA, Zhu X, Asselbergs FW, Kraaijeveld AO, Beulens JWJ, Shu X-O, Rallidis LS, Pedersen O, Hansen T, Mitchell P, Hewitt AW, Kähönen M, Pérusse L, Bouchard C, Tönjes A, Chen Y-DI, Pennell CE, Mori TA, Lieb W, Franke A, Ohlsson C, Mellström D, Cho YS, Lee H, Yuan J-M, Koh W-P, Rhee SY, Woo J-T, Heid IM, Stark KJ, Völzke H, Homuth G, Evans MK, Zonderman AB, Polasek O, Pasterkamp G, Hoefer IE, Redline S, Pahkala K, Oldehinkel AJ, Snieder H, Biino G, Schmidt R, Schmidt H, Chen YE, Bandinelli S, Dedoussis G, Thanaraj TA, Kardia SLR, Kato N, Schulze MB, Girotto G, Jung B, Böger CA, Joshi PK, Bennett DA, De Jager PL, Lu X, Mamakou V, Brown M, Caulfield MJ, Munroe PB, Guo X, Ciullo M, Jonas JB, Samani NJ, Kaprio J, Pajukanta P, Adair LS, Bechayda SA, de Silva HJ, Wickremasinghe AR, Krauss RM, Wu J-Y, Zheng W, den Hollander AI, Bharadwaj D, Correa A, Wilson JG, Lind L, Heng C-K, Nelson AE, Golightly YM, Wilson JF, Penninx B, Kim H-L, Attia J, Scott RJ, Rao DC, Arnett DK, Walker M, Koistinen HA, Chandak GR, Yajnik CS, Mercader JM, Tusié-Luna T, Aguilar-Salinas CA, Villalpando CG, Orozco L, Fornage M, Tai ES, van Dam RM, Lehtimäki T, Chaturvedi N, Yokota M, Liu J, Reilly DF, McKnight AJ, Kee F, Jöckel K-H, McCarthy MI, Palmer CNA, Vitart V, Hayward C, Simonsick E, van Duijn CM, Lu F, Qu J, Hishigaki H, Lin X, März W, Parra EJ, Cruz M, Gudnason V, Tardif J-C, Lettre G, 't Hart LM, Elders PJM,

Damrauer SM, Kumari M, Kivimaki M, van der Harst P, Spector TD, Loos RJF, Province MA, Psaty BM, Brandslund I, Pramstaller PP, Christensen K, Ripatti S, Widén E, Hakonarson H, Grant SFA, Kiemeney LALM, de Graaf J, Loeffler M, Kronenberg F, Gu D, Erdmann J, Schunkert H, Franks PW, Linneberg A, Jukema JW, Khera A V., Männikkö M, Jarvelin M-R, Kutalik Z, Cucca F, Mook-Kanamori DO, van Dijk KW, Watkins H, Strachan DP, Grarup N, Sever P, Poulter N, Rotter JI, Dantoft TM, Karpe F, Neville MJ, Timpson NJ, Cheng C-Y, Wong T-Y, Khor CC, Sabanayagam C, Peters A, Gieger C, Hattersley AT, Pedersen NL, Magnusson PKE, Boomsma DI, de Geus EJC, Cupples LA, van Meurs JBJ, Ghanbari M, Gordon-Larsen P, Huang W, Kim YJ, Tabara Y, Wareham NJ, Langenberg C, Zeggini E, Kuusisto J, Laakso M, Ingelsson E, Abecasis G, Chambers JC, Kooner JS, de Vries PS, Morrison AC, North KE, Daviglus M, Kraft P, Martin NG, Whitfield JB, Abbas S, Saleheen D, Walters RG, Holmes M V., Black C, Smith BH, Justice AE, Baras A, Buring JE, Ridker PM, Chasman DI, Kooperberg C, Wei W-Q, Jarvik GP, Namjou B, Hayes MG, Ritchie MD, Jousilahti P, Salomaa V, Hveem K, Åsvold BO, Kubo M, Kamatani Y, Okada Y, Murakami Y, Thorsteinsdottir U, Stefansson K, Ho Y-L, Lynch JA, Rader DJ, Tsao PS, Chang K-M, Cho K, O'Donnell CJ, Gaziano JM, Wilson P, Rotimi CN, Hazelhurst S, Ramsay M, Trembath RC, van Heel DA, Tamiya G, Yamamoto M, Kim B-J, Mohlke KL, Frayling TM, Hirschhorn JN, Kathiresan S, Boehnke M, Natarajan P, Peloso GM, Brown CD, Morris AP, Assimes TL, Deloukas P, Sun Y V., Willer CJ (2021) The power of genetic diversity in genome-wide association studies of lipids. Nature 600:675–679. doi: 10.1038/s41586-021-04064-3

121. Ray A (2022) Machine learning in postgenomic biology and personalized medicine. 1–26. doi: 10.1002/widm.1451

122. Xu Y, Vuckovic D, Ritchie SC, Danesh J, Akbari P, Jiang T, Grealey J, Butterworth AS, Ouwehand WH, Roberts DJ, Angelantonio E Di, Soranzo N, Inouye M (2022) Machine learning optimized polygenic scores for blood cell traits identify sex-specific trajectories and genetic correlations with disease. Cell Genomics 2:100086. doi: 10.1016/j.xgen.2021.100086

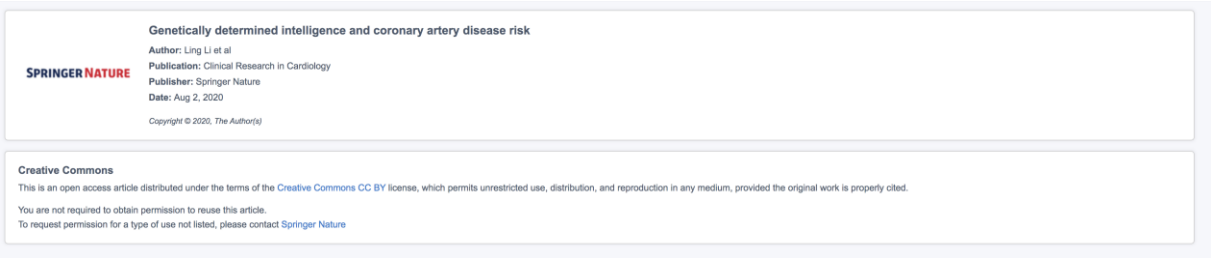
123. Miao Z, Humphreys BD, McMahon AP, Kim J (2021) Multi-omics integration in the age of million single-cell data. Nat Rev Nephrol 17:710–724. doi: 10.1038/s41581-021-00463-x

Publications during PhD

1. **Li L,** Pang SC, Zeng LY, et al. Genetically determined intelligence and coronary artery disease risk[J]. Clinical Research in Cardiology, 2021, 110(2): 211-219.
2. **Li L,** Chen ZF, von Scheidt M, et al. Transcriptome-wide association study of coronary artery disease identifies novel susceptibility genes[J]. Basic research in cardiology, 2022, 117(1): 1-20.
3. Yang CH, Starnecker F, Pang SC, et al., **Li L**. Polygenic risk for coronary artery disease in the Scottish and English population[J]. BMC cardiovascular disorders, 2021, 21(1): 1-9.
4. Zeng L, Moser S, Mirza-Schreiber N, et al., **Li L**. Cis-epistasis at the LPA locus and risk of cardiovascular diseases[J]. Cardiovascular research, 2022, 118(4): 1088-1102.
5. Hao K, Ermel R, Sukhavasi K, et al., **Li L**. Integrative prioritization of causal genes for coronary artery disease[J]. Circulation: Genomic and Precision Medicine, 2022, 15(1): e003365.
6. Aragam KG, Jiang T,Goel A, et al., **Li L**. Discovery and systematic characterization of risk variants and genes for 2 coronary artery disease in over a million participants. (under review).

Approval from each publisher

Publication 1

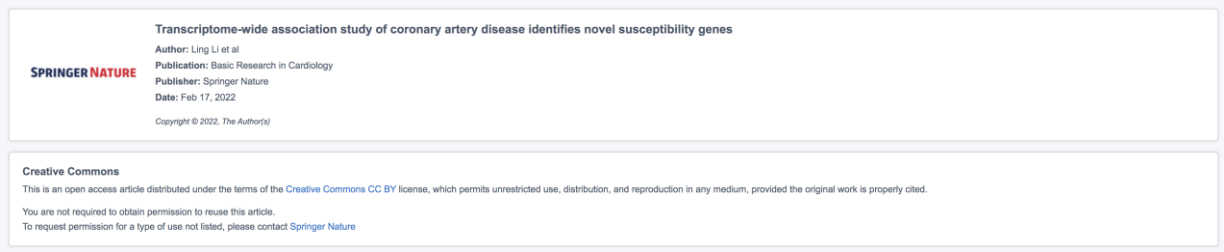


SPRINGER NATURE

Genetically determined intelligence and coronary artery disease risk

Author: Ling Li et al
Publication: Clinical Research in Cardiology
Publisher: Springer Nature
Date: Aug 2, 2020

Publication 2



SPRINGER NATURE

Transcriptome-wide association study of coronary artery disease identifies novel susceptibility genes

Author: Ling Li et al
Publication: Basic Research in Cardiology
Publisher: Springer Nature
Date: Feb 17, 2022

Appendix I Genetically determined intelligence and coronary artery disease risk

Clinical Research in Cardiology
https://doi.org/10.1007/s00392-020-01721-x

ORIGINAL PAPER

Genetically determined intelligence and coronary artery disease risk

Ling Li[1] · Shichao Pang[1] · Lingyao Zeng[1] · Ulrich Güldener[1] · Heribert Schunkert[1,2]

Received: 7 May 2020 / Accepted: 23 July 2020

Abstract

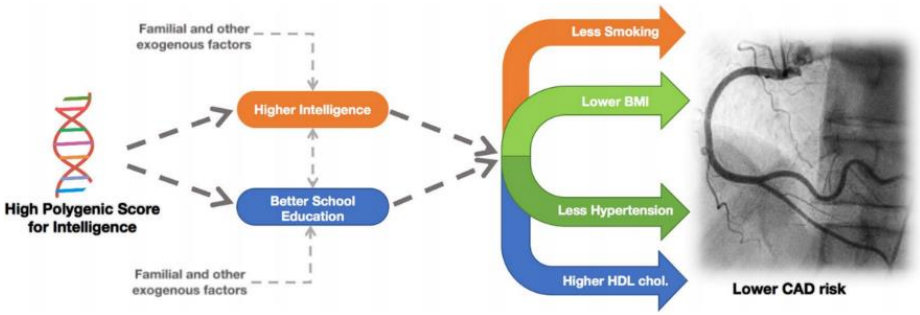
Background Epidemiological studies have shown inverse association between intelligence and coronary artery disease (CAD) risk, but the underlying mechanisms remain unclear.

Methods Based on 242 SNPs independently associated with intelligence, we calculated the genetic intelligence score (gIQ) for participants from 10 CAD case–control studies ($n = 34{,}083$) and UK Biobank ($n = 427{,}306$). From UK Biobank, we extracted phenotypes including body mass index (BMI), type 2 diabetes (T2D), smoking, hypertension, HDL cholesterol, LDL cholesterol, measured intelligence score, and education attainment. To estimate the effects of gIQ on CAD and its related risk factors, regression analyses was applied. Next, we studied the mediatory roles of measured intelligence and educational attainment. Lastly, Mendelian randomization was performed to validate the findings.

Results In CAD case–control studies, one standard deviation (SD) increase of gIQ was related to a 5% decrease of CAD risk (odds ratio [OR] of 0.95; 95% confidence interval [CI] 0.93 to 0.98; $P = 4.93e{-}5$), which was validated in UK Biobank (OR = 0.97; 95% CI 0.96 to 0.99; $P = 6.4e{-}4$). In UK Biobank, we also found significant inverse correlations between gIQ and risk factors of CAD including smoking, BMI, T2D, hypertension, and a positive correlation with HDL cholesterol. The association signals between gIQ and CAD as well as its risk factors got largely attenuated after the adjustment of measured intelligence and educational attainment. The causal role of intelligence in mediating CAD risk was confirmed by Mendelian randomization analyses.

Conclusion Genetic components of intelligence affect measured intelligence and educational attainment, which subsequently affect the prevalence of CAD via a series of unfavorable risk factor profiles.

Graphic abstract



Keywords Coronary artery disease · Intelligence · Educational attainment · Genetic association · Genetic risk score · Smoking · Obesity

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00392-020-01721-x) contains supplementary material, which is available to authorized users.

Extended author information available on the last page of the article

Published online: 02 August 2020

Appendix I Genetically determined intelligence and coronary artery disease risk

Introduction

Epidemiological studies have shown an inverse association between intelligence score and risk of coronary artery disease (CAD) [1, 2]. Higher intelligence is also inversely associated with risk factors of CAD, like smoking and obesity [3–5]. Moreover, there is evidence for association between higher intelligence and longer educational attainment [6, 7] which may be an important mediator in reducing CAD risk [8]. However, the mechanisms linking higher intelligence with a decreased risk of CAD remain unclear.

Genome wide association studies (GWAS) have identified large numbers of genetic variants, typically single nucleotide polymorphisms (SNPs), associated with a wide range of complex traits providing opportunities of exploring the relationships between traits. Polygenic risk scores defined as sum of trait-associated SNPs weighted by effect size derived from large-scale GWAS measure the liability of individuals developing such traits [9, 10]. Thereby polygenic risk scores become an important genetic tool for studying association between traits [8, 11]. Two-sample Mendelian randomization (MR) is another genetic method of accessing causal relationships among traits which requires summary statistics of GWAS instead of full individual level genotype data and phenotypic measurements [12].

Savage et al. performed genome-wide association meta-analysis in 269,867 individuals and identified 242 SNPs independently associated with intelligence [13]. We used the statistics of these intelligence SNPs to perform both regression analysis of the individual-level polygenic score and two-sample MR analysis to study the association between intelligence and CAD risk, and to explore potential pathways from a higher genetic intelligence score to lower CAD risk.

Methods

Cohorts description of individual-level genotype data

Individual level genotype data were collected from ten case–control studies of CAD as discovery set [14–21]. All participants were of European descent, mostly from the Germany and UK. The replication set was from UK Biobank [22] which includes genotypes of 487,409 individuals derived from two different genotyping array platforms.

The data of UK Biobank were also applied to characterize interplay between intelligence and risk factors of CAD including body mass index (BMI), type 2 diabetes (T2D), HDL cholesterol, LDL cholesterol, hypertension, and smoking behavior. These traits were either self-reported or extracted from hospital episodes or death registries as reported by UK Biobank [22]. Intelligence scores were measured in UK Biobank through a 13-item verbal-numeric reasoning test designed to assess the ability of solving problems that require logic and reasoning ability, independent of acquired knowledge (field ID 20016). The total range of intelligence as measured by this score was from 0 to 13 arbitrary unit. Details of corresponding studies, data preprocessing and traits definition of data from UK Biobank are shown at Supplementary Notes and Table S1.

Intelligence associated variants

Savage et al. performed GWAS meta-analysis of 14 independent epidemiological cohorts of European descent and reported 242 independent SNPs with genome-wide significant association ($P < 5e{-}8$) to intelligence scores [13]. We estimated effect size for each SNP from GWAS summary statistic table using method by Zhu et al. [23]. Details are shown at Supplementary Notes and Table S2.

Statistics

The summary statistics of 242 independent SNPs of intelligence were applied to calculate the individual-level weighted genetic score of intelligence for each study. Firstly, each variant was given a value from 0 to 2 according to the presence of the intelligence allele in the imputed genotype data of each participant, which was then multiplied with the effect size of the variant on intelligence. For variants with missing genotypes in the imputed data, the reference allele frequency was applied. Then we summed these values of 242 variants for each participant as the polygenic score of intelligence, namely the genetic intelligence score (gIQ). Afterwards, the continuous gIQ was standardized into z-scores with mean of 0 and standard deviation (SD) of 1. By logistic regression analyses, we estimated effects of gIQ on CAD risk for each study separately. To control the bias due to population stratification or different genotyping platforms, the first two principle components for 10 CAD studies were added as adjustments of the regression model. In UK Biobank, because of more complex population structure, we employed top five principle components and array platforms for this data set. Lastly, the fixed-effect size meta-analysis was performed to estimate the combined effects across all CAD studies. Based on gIQ, all individuals were evenly separated into low, medium and high groups to study the distribution of cases and controls along with increasing gIQ.

Albeit the gIQ reflects intelligence at first place, the SNPs utilized in this score may be pleiotropic and thus affects other traits [24–26]. Seven of 242 intelligence SNPs were reported to be associated with educational attainment through a large scale GWAS cohort which detected 1271 education-associated SNPs [27]. We thus re-evaluated the association between gIQ and CAD risks after exclusion of 7 SNPs overlapping with educational attainment to estimate the direct effects of intelligence.

In UK Biobank, we estimated effects of gIQ on measured intelligence, educational attainment, risk factors of CAD including BMI, T2D, smoking, HDL cholesterol, LDL cholesterol and hypertension. Definitions of these traits are shown at Supplementary Notes and Table S3. Logistic regression was applied to binary traits like T2D, smoking, hypertension; and linear regression was for continuous traits like measured intelligence, educational attainment, BMI, HDL cholesterol, and LDL cholesterol. Top five principle components and array platforms were used as adjustment of regression models. We also studied phenotypical association of measured intelligence with educational attainment and CAD incidence in UK Biobank. Additionally, to avoid the genetic influence of education derived from genetic overlaps between intelligence and education, we re-estimated the effects of intelligence on CAD and its risk factors by eliminating seven overlapping SNPs [27].

Two-sample Mendelian randomization analysis

Mendelian randomization (MR) is a method using genetic variants as instruments to study causal relationships between exposures and outcomes [28]. We introduced the multivariable two-sample MR analysis to investigate the direct casual effects of intelligence and educational attainment on CAD and its risk factors. This approach taking GWAS summary statistics as input measures effects of one standard deviation (SD) change in intelligence or educational attainment. As bias can be introduced in two-sample MR when using genetic consortia that have partially overlapping sets of participants, we selected consortia without overlaps. The GWAS summary statistics of CAD and its risk factors, educational attainment were acquired from CARDIoGRAMplusC4D (CAD) [17], GIANT (BMI) [29], TAG (smoking) [30], GLGC (HDL cholesterol, LDL cholesterol) [31], SSGAC (educational attainment) [27], and DIAGRAM (T2D) [32]. Elaborate description of these five studies can be found at Supplementary Notes and Table S4.

To address the influence of genetic overlaps between education and intelligence, we eliminated seven SNPs that are both associated with intelligence and educational attainment in MR analysis. Three MR methods including inverse-variance-weighted average (IVW) [33], MR-egger [34] and weighted median [35] were applied. Relationships significant ($P < 0.05$) in at least two of three methods were identified to be reliable and shown by IVW results. Lastly, sensitivity analysis of effects of intelligence and educational attainment on CAD were performed by excluding SNPs that were moderately associated with risk factors of CAD ($P < 0.001$) from intelligence SNPs and education SNPs respectively. Details are shown at Supplementary Notes.

Results

Effect of gIQ on the risk of CAD

Ten case–control studies of CAD with 16,144 CAD cases and 17,939 controls were included in this study. Majority of participants were from the Germany and UK. Individual-level genotype data and elaborate phenotype data from UK Biobank were used as validating set containing 20,310 CAD cases which were defined by either self-reported, or hospital episode and death registry data, and 406,996 controls. (Supplementary Notes and Table S1). For each cohort, we generated gIQ based on 242 SNPs reported to be genome-wide significantly associated with intelligence [13].

The score in participants of the 10 CAD studies was normally distributed (Fig. S1). Meta-analysis using fixed-effect size model indicated relative decrease of CAD risk by 5% (95% confidence interval [CI], 0.93 to 0.98; $P = 4.93e{-}5$) along with per 1-SD increase in gIQ (Fig. 1). When individuals were equally grouped into a low, medium and high group of gIQ, risk of CAD steadily decreased with an odds ratio (OR) of the high group vs low group being 0.89 (95% CI 0.84 to 0.93; $P = 6.2e{-}6$, Fig. 2).

Data from the UK Biobank confirmed the inverse association between gIQ and CAD risk with an OR = 0.97 (95% CI 0.96 to 0.99; $P = 6.4e{-}4$, Fig. 1). The risk of high gIQ group was 7% lower than the low gIQ group ($P = 0.0005$) in UK Biobank. As expected, the association between gIQ and CAD risk was abolished after adjustment for measured intelligence defining measured intelligence as an intermediary trait between gIQ and CAD risk (Fig. 3).

Bidirectional association between intelligence and education

In UK Biobank, we found that 1-SD increase of gIQ increased measured intelligence score by 0.29 unit ($P < 1e{-}10$) and prolonged years spent in school by 0.45 year. In addition, one more year spent in school increased the measured intelligence score by 0.16 unit ($P < 1e{-}10$). Vice versa, one unit increase in measured intelligence prolonged years spent in school by 0.98 year ($P < 1e{-}10$). Both the measured intelligence and educational attainment had inverse effects on CAD risk. See results in Table S5.

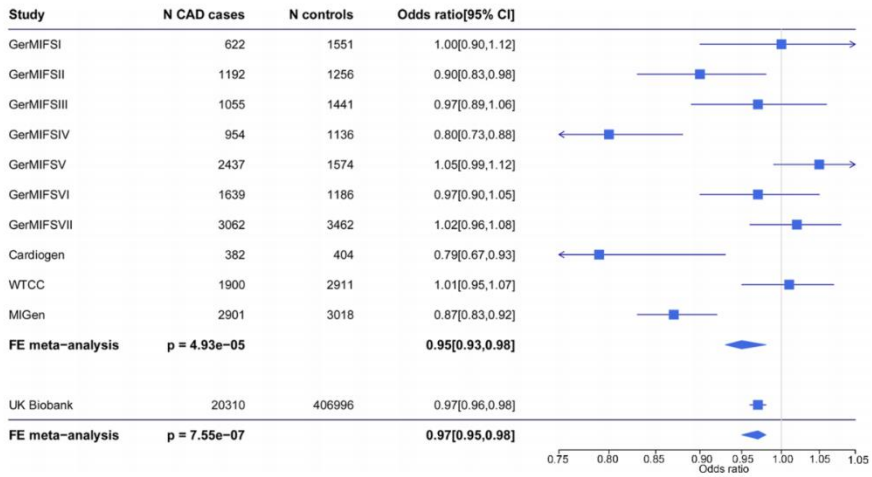


Fig. 1 Association of gIQ and CAD risk. The genetic intelligence score was calculated in 10 case-controls studies of CAD and UK Biobank respectively. Logistic regression was performed to evaluate the association between gIQ and CAD risks in each study. Fixed-effect size meta-analysis was performed to combine all studies. Forest plot shows regression result in each study and the overall effect size. The gIQ was inversely associated with CAD risk

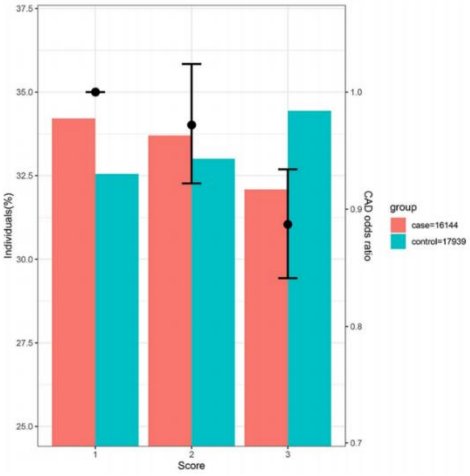


Fig. 2 Distribution of cases and controls according to gIQ. Individuals in 10 CAD studies were evenly grouped into a low (score = 1), medium (score = 2) and high (score = 3) group according to their gIQ. The OR is incidence of CAD relative to low group. Risk of CAD decreases along the increases of gIQ

Effects of gIQ on risk factors of CAD

We next asked, in UK Biobank data, whether the association between gIQ and CAD risk was mediated by traditional risk factors of CAD, and whether such effects were dependent of measured intelligence and educational attainment. We found strong associations of gIQ with BMI, smoking, T2D, HDL cholesterol, and hypertension (Fig. 3). The effects of gIQ on CAD risk factors were largely attenuated after adjustment for measured intelligence or educational attainment (Fig. 3 and Table S6), suggesting that measured intelligence and educational attainment mainly mediated associations between gIQ and these risk factors. The analyses after removal of seven SNPs overlapping between intelligence and educational attainment obtained quantitatively and qualitatively similar effects of gIQ on CAD and its risk factors (Fig. S2).

We also studied the mediatory roles of these risk factors on the association between gIQ and CAD risk by applying them as adjustments to the regression model. Adjusting for individual risk factor or risk factors combined markedly attenuated association signal between gIQ and CAD risk (Fig. S3), indicating these risk factors were involved in mediating the association between gIQ and CAD risk.

Mendelian randomization validation

To substantiate our observations, we performed multivariable two-sample MR analysis taking intelligence or

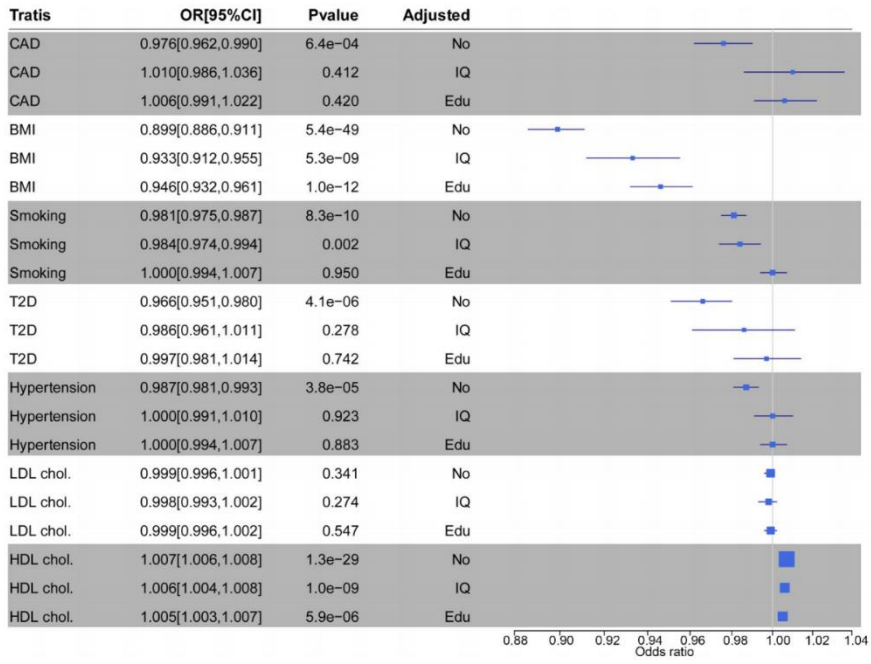


Fig. 3 Associations of gIQ with CAD and it risk factors including BMI, smoking, T2D, HDL cholesterol, LDL cholesterol, and hypertension in UK Biobank. The OR for BMI is shown as logarithm of the linear regression coefficient. 'Adjusted' indicates the regression model between gIQ and trait after adjustment for measured intelligence (IQ), or length of school years completed (Edu), or neither of the two (No). The gIQ had inverse effects on BMI, T2D, smoking, and hypertension and a positive effect on HDL cholesterol. The association signals were largely attenuated by measured intelligence and educational attainment

educational attainment as exposures, CAD and its risk factors as outcomes. The estimates of the direct effects on outcomes for intelligence and education were generally in a consistent direction (Fig. 4). 1-SD increase of intelligence resulted in decrease of CAD risk by 25% (OR = 0.75; 95% CI 0.69 to 0.81; $P < 1e{-}10$), decrease of BMI by 0.1 kg/m^2 (95% CI − 0.16 to − 0.14; $P = 1.02e{-}3$), decrease of T2D risk by 15% (OR = 85; 95% CI 0.77 to 0.95). A SD increase in the education years resulted in decrease of risk of CAD by 38% (OR = 0.62; 95% CI 0.58 to 0.66; $P < 1e{-}10$), decrease of BMI by 0.32 kg/m^2 (95%CI − 0.37 to − 0.27; $P < 1e{-}10$), increase of HDL cholesterol by 0.19 mmol/L (95% CI 0.14 to 0.25; $P < 1e{-}10$), decrease of the risk of smoking by 43% (OR = 0.57; 95%CI 0.501 to 0.642; $P < 1e{-}10$), and decrease of T2D risk by 47% (OR = 0.53; 95%CI 0.49 to 0.57; $P < 1e{-}10$). The effects of educational attainment on CAD and its risk factors displayed the same direction as intelligence but were stronger in magnitude. See details at supplementary notes and Table S7.

Lastly, MR sensitivity analysis were performed for intelligence and educational attainment respectively. For intelligence, SNPs moderately associated ($P < 0.001$) with CAD ($n = 5$), BMI ($n = 45$), and HDL cholesterol ($n = 5$) were removed from intelligence SNPs. The sensitivity analysis showed 1-SD increase in intelligence to decrease the risk of CAD by 22% (OR for IVW method of 0.78; 95% CI 0.72 to 0.84; $P = 5.6e{-}10$). Same as intelligence, SNPs moderately associated ($P < 0.001$) with CAD ($n = 13$), BMI ($n = 155$), HDL cholesterol ($n = 6$), LDL cholesterol ($n = 5$), and smoking ($n = 2$) were removed from education SNPs. The sensitivity analysis showed 1-SD increase in education years to decrease the risk of CAD by 34% (OR for IVW method of 0.66; 95%CI 0.62 to 0.70; $P < 1e{-}10$). Results are shown at Table S8.

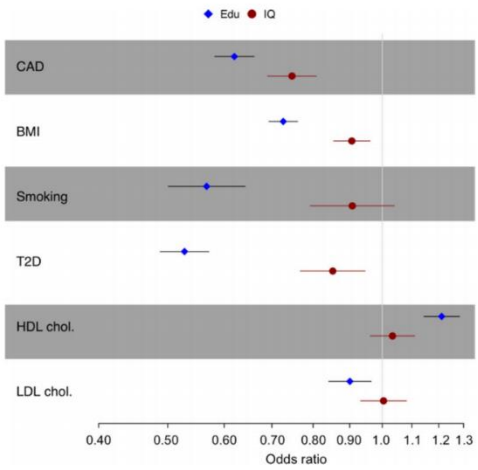


Fig. 4 The result of MR analyses. Error bars indicate 95% confidence intervals around the estimated effects calculated using multivariable two-sample MR. The effects on outcomes for intelligence and educational attainment were generally in consistent directions. But the effects of educational attainment are quantitatively stronger than intelligence

Discussion

Epidemiological studies have revealed that increased intelligence correlates with reduced CAD risk [1, 2]. Consistently, our study shows that 1-SD increase of gIQ based on accumulated effects of genetic variants associated with intelligence, results in 5% decrease in the risk of CAD. The CAD risk in the high group of gIQ is relatively lower by 11% than in the low group. The observation was replicated in the UK Biobank. Interestingly, the inverse association got largely attenuated after adjustment for measured intelligence and educational attainment supporting the hypothesis that these traits play a role in modulating CAD risk.

Our study also shows the inverse effects of gIQ on health-related outcomes including BMI, smoking, T2D, hypertension, and a positive effect on HDL cholesterol, which are well-known for their influences on CAD risk [36–40]. Same as for CAD, these association signals appear to be largely mediated by measured intelligence and educational attainment. It can be concluded that these risk factors mediate the association between gIQ and CAD risk individually and collectively.

Our study confirms that intelligence and educational attainment are genetically and phenotypically associated with each other [6, 7]. Like in the present study, a recent study by our group states that educational attainment is inversely associated with CAD risk which appears to be mediated by risk factors such as BMI and smoking [8, 26]. Interestingly, our current study indicates that the effects of educational attainment on CAD and its risk factors are quantitatively stronger than respective effects of intelligence. All these findings indicate that improving educational attainment can have potential benefits in improving decision-making regarding health-relevant lifestyle factors and reducing risk of CAD and other health-related outcomes.

Polygenic risk score and two-sample MR are two genetic approaches of investigating association between traits. Compared with the traditional epidemiologic approach, the genetic approach is unlikely to be confounded by lifestyle or environmental factors as genotypes are stable over lifetime [11]. The utilization of genetic methods is limited, however, by false discovery because of horizontal pleiotropy, a phenomenon explained by the fact that variants may affect multiple traits through different pathways [9, 10]. The complex interplay of intelligence and educational attainment caused by their genetic roots limits a precise causal relationship between intelligence and CAD as well as its risk factors. In our study, we aimed to exclude genetic overlaps between intelligence and education to highlight putative causal effects of intelligence on CAD and its risk factors. Indeed, this notion was furtherly confirmed by MR analysis and the MR sensitivity analysis after excluding SNPs marginally associated with risk factors of CAD from intelligence (or education).

There are some limitations in our study. First, the intelligence SNPs utilized in this study were identified from a large GWAS meta-analysis based 14 independent epidemiological cohorts of European ancestry [13]. To avoid bias due to difference in population genetics, we restricted our analysis to cohorts from Germany, UK, and others of European ancestry. Second, there might be other health-related or socioeconomic factors that interplay with intelligence and CAD risk [26]. Specially, environmental exposures can be important confounders of association between intelligence and CAD risk. Third, the measured intelligence obtained in UK Biobank through a 13-item verbal-numeric reasoning test does not equal to real intelligence whose full scopes are unspecifiable. Moreover, educational attainment defined as years spent in schools in this study has a wide spectrum in various countries. Last, the two-sample MR analyses are likely to be biased if two studies contains overlapping participants or cohorts which are quite common in large-scale GWAS meta-analysis [41]. We tried best to choose studies that are of European ancestry and have minimal overlaps to avoid such bias in two-sample MR analysis.

In conclusion, using genetic approaches, we depicted a pathway from gIQ to CAD risk (Fig. 5). The higher gIQ is associated with the higher measured intelligence and longer educational attainment, both of which appear to reduce the prevalence of risk factors of CAD including BMI, smoking,

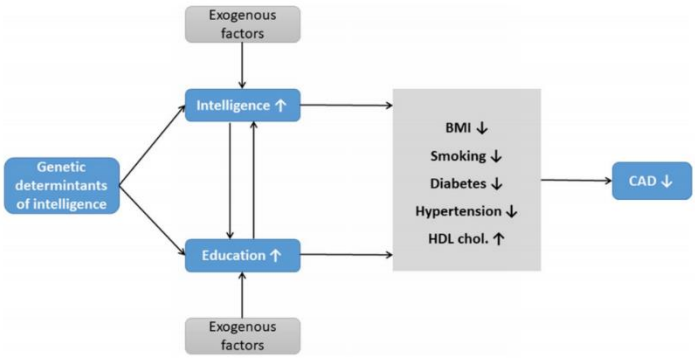


Fig. 5 Pathway from higher gIQ to lower risk of CAD. Our study shows inverse effects of genetic determinants of intelligence on CAD and its risk factors including BMI, smoking, hypertension, and T2D and positive effects on HDL cholesterol. These association signals are mediated by measured intelligence and educational attainment, which two are bidirectionally associated with each other

T2D and hypertension, and increase HDL cholesterol, which in concert subsequently reduce the prevalence of CAD. Moreover, the effects of educational attainment on risk factors and CAD appear to be stronger than the effects of intelligence. Thus, repetitive campaigns throughout schooling may be worthwhile for preventive reasons as they may ameliorate the association between gIQ and unhealthy lifestyle.

Acknowledgements Open Access funding provided by Projekt DEAL.

Funding We gratefully acknowledge the support of the German Federal Ministry of Education and Research (BMBF) within the framework of ERA-NET on Cardiovascular Disease (Druggable-MI-genes: 01KL1802), within the scheme of target validation (Block-CAD: 16GW0198K), the German Centre of Cardiovascular Research (DZHK) Munich Heart Alliance, within the framework of the e:Med research and funding concept (AbCD-Net: 01ZX1706C). As a Co-applicant for the British Heart Foundation (BHF)/German Centre of Cardiovascular Research (DZHK)-collaboration, we kindly acknowledge their support and funding. Further support has been received from the German Research Foundation (DFG) as part of the Sonderforschungsbereich SFB 1123 (B02) and the Sonderforschungsbereich SFB TRR 267 (B05). Furthermore, Bavarian State Ministry of Health and Care founded this work within its framework of DigiMed Bayern (Grant No: DMB-1805-0001). The work has also been supported by the German Federal Ministry of Economics and Energy in its scheme of ModulMax (Grant No: ZF4590201BA8). We got authorization for applicating data from UK Biobank (project ID: 25214).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Silventoinen K, Modig-Wennerstad K, Tynelius P, Rasmussen F (2007) Association between intelligence and coronary heart disease mortality: a population-based cohort study of 682 361 Swedish men. Eur J Cardiovasc Prev Rehabil 14:555–560. https://doi.org/10.1097/HJR.0b013e328014672e
2. Batty GD, Deary IJ, Benzeval M, Der G (2010) Does IQ predict cardiovascular disease mortality as strongly as established risk factors? Comparison of effect estimates using the West of Scotland Twenty-07 cohort study. Eur J Prev Cardiol 17:24–27. https://doi.org/10.1097/HJR.0b013e328321311b
3. Weiser M, Zarka S, Werbeloff N et al (2010) Cognitive test scores in male adolescent cigarette smokers compared to non-smokers: A population-based study. Addiction 105:358–363. https://doi.org/10.1111/j.1360-0443.2009.02740.x
4. Kanazawa S (2014) Intelligence and obesity. Curr Opin Endocrinol Diabetes Obes 21:339–344. https://doi.org/10.1097/MED.0000000000000091
5. Tabriz AA, Sohrabi MR, Parsay S et al (2015) Relation of intelligence quotient and body mass index in preschool children: a community-based cross-sectional study. Nutr Diabetes 5:e176–e176. https://doi.org/10.1038/nutd.2015.27
6. Deary IJ, Strand S, Smith P, Fernandes C (2007) Intelligence and educational achievement. Intelligence 35:13–21. https://doi.org/10.1016/j.intell.2006.02.001
7. Calvin CM, Fernandes C, Smith P et al (2010) Sex, intelligence and educational achievement in a national cohort of over 175,000 11-year-old schoolchildren in England. Intelligence 38:424–432. https://doi.org/10.1016/j.intell.2010.04.005
8. Zeng L, Ntalla I, Kessler T et al (2019) Genetically modulated educational attainment and coronary disease risk. Eur Heart J 40:2413–2420. https://doi.org/10.1093/eurheartj/ehz328
9. Maier RM, Visscher PM, Robinson MR, Wray NR (2017) Embracing polygenicity: a review of methods and tools for psychiatric genetics research. Psychol Med 48:1055-1067. https://doi.org/10.1017/S0033291717002318
10. Richardson TG, Harrison S, Hemani G, Smith GD (2019) An atlas of polygenic risk score associations to highlight putative causal

Appendix I Genetically determined intelligence and coronary artery disease risk

relationships across the human phenome. Elife 8:e43657. https://doi.org/10.7554/eLife.43657
11. Nelson CP, Hamby SE, Saleheen D et al (2015) Genetically determined height and coronary artery disease. N Engl J Med 372:1608–1618. https://doi.org/10.1056/NEJMoa1404881
12. Hartwig FP, Davies NM, Hemani G, Smith GD (2016) Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique Editorial. Int J Epidemiol 45:1717–1726. https://doi.org/10.1093/ije/dyx028
13. Savage JE, Jansen PR, Stringer S et al (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet 50:912–919. https://doi.org/10.1038/s41588-018-0152-6
14. Samani NJ, Erdmann J, Hall AS et al (2007) Genomewide association analysis of coronary artery disease. N Engl J Med 357:443–453. https://doi.org/10.1056/NEJMoa072366
15. Erdmann J, Großhennig A, Braund PS et al (2009) New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet 41:280–282. https://doi.org/10.1038/ng.307
16. Erdmann J, Willenborg C, Nahrstaedt J et al (2011) Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. Eur Heart J 32:158–168. https://doi.org/10.1093/eurheartj/ehq405
17. Nikpay M, Goel A, Won HH et al (2015) A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 47:1121–1130. https://doi.org/10.1038/ng.3396
18. Stitziel NO, Won HH, Morrison AC et al (2014) Inactivating mutations in NPC1L1 and protection from coronary heart disease. N Engl J Med 371:2072–2082. https://doi.org/10.1056/NEJMoa1405386
19. Nelson CP, Goel A, Butterworth AS et al (2017) Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet 49:1385–1391. https://doi.org/10.1038/ng.3913
20. Winkelmann BR, März W, Boehm BO et al (2001) Rationale and design of the LURIC study—a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. Pharmacogenomics 2:S1-S73. https://doi.org/10.1517/14622416.2.1.s1
21. Kathiresan S, Voight BF, Purcell S et al (2009) Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 41:334–341. https://doi.org/10.1038/ng.327
22. Bycroft C, Freeman C, Petkova D et al (2018) The UK Biobank resource with deep phenotyping and genomic data. Nature 562:203–209. https://doi.org/10.1038/s41586-018-0579-z
23. Zhu Z, Zhang F, Hu H et al (2016) Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48:481–487. https://doi.org/10.1038/ng.3538
24. Gratten J, Visscher PM (2016) Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. Genome Med 8:78. https://doi.org/10.1186/s13073-016-0332-x
25. Davies NM, Hill WD, Anderson EL et al (2019) Multivariable two-sample mendelian randomization estimates of the effects of intelligence and education on health. Elife 8:e43990. https://doi.org/10.7554/eLife.43990
26. Schunkert H, Erdmann J, Samani NJ et al (2020) Genetics of educational attainment and coronary risk in Mendelian randomization studies. Eur Heart J 41:894–895. https://doi.org/10.1093/eurheartj/ehz946
27. Lee JJ, Wedow R, Okbay A et al (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 50:1112–1121. https://doi.org/10.1038/s41588-018-0147-3
28. Smith GD, Hemani G (2014) Mendelian randomization: geneticanchorsfor causal inference in epidemiological studies. Hum Mol Genet 23:R89–R98. https://doi.org/10.1093/hmg/ddu328
29. Pulit SL, Stoneman C, Morris AP et al (2019) Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet 28:166–174. https://doi.org/10.1093/hmg/ddy327
30. Furberg H, Kim Y, Dackor J et al (2010) Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet 42:441–447. https://doi.org/10.1038/ng.571
31. Willer CJ, Schmidt EM, Sengupta S et al (2013) Discovery and refinement of loci associated with lipid levels. Nat Genet 45:1274–1285. https://doi.org/10.1038/ng.2797
32. Scott RA, Scott LJ, Mägi R et al (2017) An expanded genome-wide association study of type 2 diabetes in Europeans. Diabetes 66:2888–2902. https://doi.org/10.2337/db16-1253
33. de Bakker PIW, Ferreira MAR, Jia X et al (2008) Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet 17:R122–R128. https://doi.org/10.1093/hmg/ddn288
34. Burgess S, Thompson SG (2017) Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol 32:377–389. https://doi.org/10.1007/s10654-017-0255-x
35. Bowden J, Davey Smith G, Burgess S (2015) Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 44:512–525. https://doi.org/10.1093/ije/dyv080
36. Formentini FS, Zaina Nagano FE, Lopes Neto FDN et al (2019) Coronary artery disease and body mass index: what is the relationship? Clin Nutr ESPEN 34:87–93. https://doi.org/10.1016/j.clnesp.2019.08.008
37. Ambrose JA, Barua RS (2004) The pathophysiology of cigarette smoking and cardiovascular disease: an update. J Am Coll Cardiol 43:1731–1737. https://doi.org/10.1016/j.jacc.2003.12.047
38. Ali KM, Wonnerth A, Huber K, Wojta J (2012) Cardiovascular disease risk reduction by raising HDL cholesterol—current therapies and future opportunities. Br J Pharmacol 167:1177–1194. https://doi.org/10.1111/j.1476-5381.2012.02081.x
39. Kjeldsen SE (2018) Hypertension and cardiovascular risk: general aspects. Pharmacol Res 129:95–99. https://doi.org/10.1016/j.phrs.2017.11.003
40. Bulik-Sullivan B, Finucane HK, Anttila V et al (2015) An atlas of genetic correlations across human diseases and traits. Nat Genet 47:1236–1241. https://doi.org/10.1038/ng.3406
41. Burgess S, Davies NM, Thompson SG (2016) Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol 40:597–608. https://doi.org/10.1002/gepi.21998

Appendix I Genetically determined intelligence and coronary artery disease risk

Clinical Research in Cardiology

Affiliations

Ling Li[1] · Shichao Pang[1] · Lingyao Zeng[1] · Ulrich Güldener[1] · Heribert Schunkert[1,2]

✉ Heribert Schunkert
schunkert@dhm.mhn.de

[1] Department of Cardiology, Deutsches Herzzentrum München, Technische Universität München, Lazarettstr. 36, 80636 Munich, Germany

[2] Deutsches Zentrum für Herz- und Kreislaufforschung (DZHK), Munich Heart Alliance, Munich, Germany

Appendix II Transcriptome-wide association study of coronary artery disease identifies novel susceptibility genes

Basic Research in Cardiology (2022) 117:6
https://doi.org/10.1007/s00395-022-00917-8

ORIGINAL CONTRIBUTION

Transcriptome-wide association study of coronary artery disease identifies novel susceptibility genes

Ling Li[1,2,3] · Zhifen Chen[1,3] · Moritz von Scheidt[1,3] · Shuangyue Li[1,3] · Andrea Steiner[1,3] · Ulrich Güldener[1,3] · Simon Koplev[4] · Angela Ma[4] · Ke Hao[4] · Calvin Pan[5] · Aldons J. Lusis[5,6,7] · Shichao Pang[1,3] · Thorsten Kessler[1,3,7] · Raili Ermel[8] · Katyayani Sukhavasi[8] · Arno Ruusalepp[8,9] · Julien Gagneur[2] · Jeanette Erdmann[10,11] · Jason C. Kovacic[12,13,14] · Johan L. M. Björkegren[4,9,15] · Heribert Schunkert[1,3]

Received: 2 November 2021 / Revised: 18 January 2022 / Accepted: 1 February 2022 / Published online: 17 February 2022

Abstract
The majority of risk loci identified by genome-wide association studies (GWAS) are in non-coding regions, hampering their functional interpretation. Instead, transcriptome-wide association studies (TWAS) identify gene-trait associations, which can be used to prioritize candidate genes in disease-relevant tissue(s). Here, we aimed to systematically identify susceptibility genes for coronary artery disease (CAD) by TWAS. We trained prediction models of nine CAD-relevant tissues using EpiXcan based on two genetics-of-gene-expression panels, the Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET) and the Genotype-Tissue Expression (GTEx). Based on these prediction models, we imputed gene expression of respective tissues from individual-level genotype data on 37,997 CAD cases and 42,854 controls for the subsequent gene-trait association analysis. Transcriptome-wide significant association (i.e. $P < 3.85e{-}6$) was observed for 114 genes. Of these, 96 resided within previously identified GWAS risk loci and 18 were novel. Stepwise analyses were performed to study their plausibility, biological function, and pathogenicity in CAD, including analyses for colocalization, damaging mutations, pathway enrichment, phenome-wide associations with human data and expression-traits correlations using mouse data. Finally, CRISPR/Cas9-based gene knockdown of two newly identified TWAS genes, *RGS19* and *KPTN*, in a human hepatocyte cell line resulted in reduced secretion of APOB100 and lipids in the cell culture medium. Our CAD TWAS work (i) prioritized candidate causal genes at known GWAS loci, (ii) identified 18 novel genes to be associated with CAD, and iii) suggested potential tissues and pathways of action for these TWAS CAD genes.

Keywords Coronary artery disease · Transcriptome-wide association study · Genome-wide association study · Genetically regulated expression

Introduction

Coronary artery disease (CAD), a leading cause of premature death worldwide, is influenced by interactions of lifestyle, environmental, and genetic risk factors [43]. Genome-wide association studies (GWAS) have identified over 200 risk loci for CAD [11, 17, 35]. Most of them are located in non-coding regions which hampers their functional interpretation. Expression quantitative traits loci (eQTLs) to some extent explain the genomic effects of GWAS signals [19, 61, 64]. By leveraging effects of multiple *cis*-eQTL variants on gene expression, transcriptome-wide association studies (TWAS) search primarily for gene-trait associations. The approach first builds on prediction models of gene expression from reference panels that correlate genotype patterns with transcript levels in tissues which are relevant for the phenotype. Prediction models are then used to impute tissue-specific gene expression based on genotypes with a given trait in individuals of GWAS cohorts [21]. Since TWAS signals reflect association between transcriptome-wide genetically regulated expression (GReX) and traits or diseases, the approach can be used to prioritize candidate genes across disease-relevant tissues. Thereby, TWAS may point to causal genes at risk loci identified by GWAS and

Ling Li and Zhifen Chen have contributed equally to this work.

✉ Heribert Schunkert
schunkert@dhm.mhn.de

Extended author information available on the last page of the article

thus provide further insights into biological mechanisms [62, 70]. Moreover, TWAS increase the sensitivity to identify susceptibility genes missed by traditional GWAS analyses. Here we performed a TWAS to identify novel susceptibility genes for CAD comprising more than 80,000 individuals with genotype data along with validation and exploratory analyses for the associated genes.

Materials and methods

Prediction models of nine tissues based on two reference panels

The starting point of this investigation was two large human biobanks with individual-level data on genome-wide genotypes as well as mRNA expression levels in multiple tissues with relevance for CAD. These include atherosclerotic aortic wall (AOR), atherosclerotic-lesion-free internal mammary artery (MAM), liver (LIV), blood (BLD), subcutaneous fat (SF), visceral abdominal fat (VAF), and skeletal muscle (SKLM) in the Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET) [20], and AOR, LIV, BLD, SF, VAF, and SKLM in the Genotype-Tissue Expression (GTEx) [1] (Supplementary Table 1). Arterial wall coronary (COR) and tibial artery (TIB) datasets were only available in the GTEx. The pipeline used for training prediction models was EpiXcan which was built on the basis of PrediXcan but with improved prediction performance by integrating epigenomic annotation data into model-training process [21, 70]. The samples used for training models were restricted to European ancestry. We adopted the existing expression prediction models established by Zhang except COR and TIB tissues which were not covered yet [70].

We established predictive models for COR and TIB tissues using the same parameters as other tissues [70]. In brief, we first filtered the genotype and expression data of COR and TIB from GTEx v7 [1]. For genotype data, variants with call rate < 0.95, minor allele frequency (MAF) < 0.01, and Hardy Weinberg equilibrium (HWE) < 1e−6 were removed. For expression, we used quality-controlled data and performed sample-level quantile normalization and gene-level inverse quantile normalization using preprocess codes of PredicDB pipeline [21]. We then calculated SNP priors using hierarchical Bayesian model (qtlBHM) [40] that jointly analyzed epigenome annotations of aorta derived from Roadmap Epigenomics Mapping Consortium (REMC) [5], and eQTL statistics. The SNP priors (Supplementary Table 2), genotype data and expression data were jointly applied to tenfold cross-validated weighted elastic-net to train prediction models [70].

Both STARNET- and GTEx-based models were filtered by cross-validated prediction $R^2 > 0.01$ [28, 68]. The summary statistics of sample sizes used for training models and the transcript numbers of genes covered by each predicting model are shown in Supplementary Table 1.

GWAS cohorts

For the discovery cohort, individual-level genotyping data were collected from ten CAD GWAS, a subset of CARDIoGRAMplusC4D, including the German Myocardial Infarction Family Studies (GerMIFS) I–VII [16, 18, 38, 47, 48, 52, 56], Wellcome Trust Case Control Consortium (WTCCC) [7], LURIC [65], and Myocardial Infarction Genetics Consortium (MIGen) [2]. We used a part of individual-level data from UK Biobank (UKB) as the replication cohort [8], by extracting 20,310 CAD cases according to hospital episodes or death registries as reported and randomly selecting 25,000 non-CAD participants as controls. The detailed information about selection criteria of case and control were described at elsewhere [38]. In total, genotypes of 37,997 cases and 42,854 controls were included in our transcriptome-wide association studies (TWAS) of CAD (Supplementary Table 3). The preprocessing steps of genotyping data are as previously described [38].

TWAS analysis

The transcriptome-wide association analysis was performed using prediction models of nine tissues for imputing individual-level GReX from CAD cases and controls of 11 GWAS cohorts and by association of these tissue-specific GReX with CAD risk in each cohort. To test the replicability of TWAS results, we used ten GWAS cohorts as discovery set and UKB as the replication set to test replicability within STARNET- and GTEx-based models, respectively. We compared the consistency of TWAS results between STARNET- and GTEx-based models of the six overlapping tissues using all genotype data. Then, we merged TWAS genes resulted from two reference-based panels as the final list. Finally, we annotated the TWAS genes list by over 200 CAD loci identified by GWASs [17, 35] using MAGMA [37]. Gene resided in the ± 1 Mb regions around known GWAS loci were marked as the known, otherwise genes were marked as the novel.

Colocalization of the eQTL and GWAS signals

Colocalization analysis was performed using COLOC, a Bayesian statistical methodology that takes GWAS and eQTL data as inputs, and tests the posterior probabilities of hypothesis #4 (PP4) that there are shared casual variants for each locus [23]. The summary statistics of GWAS meta-analysis were obtained from CARDIoGRAMplusC4D Consortium [47], and the eQTL data of nine tissues from

STARNET [20] and GTEx [1]. The significance threshold is PP4 > 0.55.

Co-expression network for protein coding and lncRNA genes

We used RNA-seq data of STARNET [20] to calculate expression correlations between long non-coding RNA (lncRNA) genes and protein-coding genes in seven tissues. Co-expression pairs with absolute Pearson correlation coefficient larger than 0.4 were considered to be significant. The co-expression network was displayed by cytoscape [34].

Gene set enrichment analyses

Pathway enrichment analysis was carried out using ClueGO (v2.5.2) [6], a plugin of cytoscape [34], based on collated gene sets from public databases including Gene Ontology (GO) [26], KEGG [30], Reactome [12], and WikiPathways [55]. Gene sets with false discovery rate (FDR) by right-sided hypergeometric test less than 0.05 were considered to be significant.

Furthermore, we also studied the diseases or traits associated with risk genes by performing disease enrichment analysis based on DisGeNET [50], the largest publicly available datasets of genes and variants association of human diseases. FDR < 0.05 was used for thresholding.

Rare damaging variants association analysis

To investigate association of damaging variants in TWAS genes with CAD, we used whole-exome sequencing (WES) data of 200,632 participants from UKB [27]. The WES data were processed following the Functional Equivalence (FE) protocol. We performed quality control on the WES data by filtering variants with calling rate < 0.9 and variants with HWE < 1e−6. For the relevant traits, besides CAD, we considered i) three lifestyle factors including body mass index (BMI), diabetes, hypertension; ii) seven categories of blood lipids including low-density lipoproteins cholesterol (LDL-C), high density lipoproteins cholesterol (HDL-C), apolipoprotein A (APOA), apolipoprotein B (APOB), Lipoprotein(a) (LPA), total cholesterol (TC) and triglycerides (TG); iii) four inflammation related factors including C-reactive protein (CRP), lymphocyte count (Lymphocyte), monocyte count (Monocyte) and neutrophil count (Neutrophil). In total, 15 traits were studied.

We defined damaging variants as (i) MAF < 0.01; (ii) annotated into following one of the three classes: loss-of-function (LoF) variants (stop-gained, splice site disrupting, or frameshift variants), pathogenic variants in ClinVar [36], or missense variants predicted to be damaging by one of five computer prediction algorithms (LRT score, MutationTaster, PolyPhen-2 HumDiv, PolyPhen-2 HumVar, and SIFT). The Ensembl Variant Effect Predictor (VEP) [45] and its plugin loftee [31], and annotation databases dbNSFP 4.1a [14] and ClinVar (GRCh38) [36] were used for annotating damaging mutations.

For each analysis, samples were classified into carriers or noncarriers of the gene's damaging mutations. For binary traits, we used Fisher's exact test to check if there was incidence difference of mutation carrying between case and controls. For the quantitative traits, we used linear regression model with adjustments of sex, first five principal components, and lipid medication status to investigate the associations between mutation carrying status and traits. We used nominal significance threshold ($P < 0.05$), given that coding variants are rather rare, and the case–control sample sizes were not balanced which might increase false negative rate.

Association of variants in novel genes with lipid traits

For 18 novel risk genes, we performed association analysis for variants located in respective loci (± 1 Mb) with lipid traits using genotyping data of UKB. The lipid traits include levels of LDL-C, HDL-C, APOA, APOB, LPA, TC, and TG. The variants were filtered by MAF > 0.01, and imputation info score > 0.4. The association test was performed using PLINK2 [10] with adjustment of sex, age, first five principal components, and lipid medication status. The numbers of independent SNPs were estimated using Genetic type 1 error calculator (GEC) tool [39].

Expression-trait association study using mouse data

The hybrid mouse diversity panel (HMDP) is a set of 105 well-characterized inbred mouse strains on a 50% C57BL/6J genetic background [42]. To specifically study atherosclerosis in the HMDP, transgene implementation of human APOE-Leiden and cholesteryl ester transfer protein was performed, promoting distinct atherosclerotic lesion formation [4]. A Western diet containing 1% cholesterol was fed for 16 weeks. Subsequently, gene expression was quantified in aorta and liver of these mice and lesion size was assessed in the proximal aorta using oil red O staining. Fourteen other related traits were measured too, including liver fibrosed area, body weight, TC, VLDL-C (very low-density lipoprotein cholesterol) + LDL-C, HDL-C, TGs, unesterified cholesterol, free fatty acid (FFA), Il-1b, Il-6, Tnfa, Mcp-1, and M-csf. From HMDP, we extracted significant association pairs between TWAS genes and 15 risk traits by applying significance $P < 0.05$.

Experimental validation of *KPTN* and *RGS19* in human cells

To knock down *KPTN* and *RGS19*, two sgRNAs targeting shared exons of all transcription isoforms were delivered by lentivirus into a Cas9-expression huh7, a human hepatoma cell line. Exon 4 of *KPTN* and exon 5 of *RGS19* were targeted by a dual CRISPR strategy to create a 40 bp and 130 bp frame shift deletion, respectively. SgRNAs were carried by Lenti-Guide-Puro vector (addgene, #52963) and infected cells were treated with 10 ug/ml puromycin treatment for 3 days to eliminate the negative cell. Positive targeted cells were expanded in culture and passaged for assays. Cells for measurement of secretive triglycerides, cholesterol, and APOB100 were cultured for 16 h in serum-free medium. Medium triglycerides and cholesterol were enriched for five times by vacuum centrifuge and measured with colorimetric kits, triglyceride (cobas), and CHOL2 (cobas), respectively. The amount of medium APOB100 was measured with an ELISA kit (MABTECH).

RNA isolation and sequencing

Total RNA from huh7 cells was isolated using RNEasy Plus Mini Kit (Qiagen) (control cells, $n=3$; knockout cells, $n=3$). Quantity and quality of the isolated RNAs were measured by Fragment Analyzer (Agilent). RNA sequencing (RNA-seq) was performed by BGI TECH SOLUTIONS (HONGKONG) using strand specific library preparation with mRNA enrichment, paired-end sequencing with 100 bp read length on the DNBSEQ platform and 20 M clean read pairs per sample. Clean reads were mapped onto the GRCh38.p12. Expression quantifications, differential expression, and gene set enrichment were performed according to BGI RNA-seq pipeline.

Results

Transcriptome-wide significant genes for CAD

The study design is shown in Fig. 1. Expression prediction models of nine tissues were derived from two reference panels, STARNET [20] and GTEx [1], using EpiXcan pipeline [70] (Materials and methods). We applied these models to impute transcriptome-wide GReX of nine tissues from individual-level genotype data of 11 GWAS cohorts (Supplementary materials; Supplementary Fig. 1–2; Supplementary Tables 1–3) [2, 7, 8, 16, 18, 38, 47, 52, 56, 65]. We next associated the GReX with CAD risk in each cohort (Supplementary materials). The results revealed replicability of TWAS genes when taking ten CARDIoGRAMplusC4D cohorts as discovery and UKB as replication set within the STARNET- and GTEx-based prediction models, respectively (Supplementary Fig. 2I–II; Supplementary Fig. 3). The results also showed consistency and complementarity of TWAS findings in six shared tissues between two reference-based prediction models (Supplementary Fig. 2III–IV; Supplementary Figs. 4–5). Therefore, we combined the results based on the two reference models for the final list of TWAS genes (Supplementary Fig. 2 V).

From STARNET-based models 129 gene-tissues pairs and from GTEx-based models 106 gene-tissue pairs were significantly associated with CAD (Bonferroni-corrected significance based on 12,995 genes, $P<3.85e{-}6$). Since 42 pairs overlapped between the two panels (Supplementary Fig. 5), the total number of gene-tissue pairs was 193. Given that some genes displayed association in multiple tissues, the final list of significant TWAS genes for CAD was 114 genes (Fig. 2; Supplementary Fig. 6; Supplementary Table 4). Of these, 95 gene-tissue association pairs were confirmed using another commonly used fine-mapping tool (COLOC) [23] with posterior probabilities of shared causal variants in each locus larger than 0.55 (PP4 > 0.55; Materials and methods; Supplementary Table 5; Supplementary Fig. 7).

Of the 114 TWAS genes, 46 genes displayed genetically mediated differential expression in AOR, 28 in MAM, 25 in LIV, 23 in VAF, 22 in SKLM, 18 in SF, 16 in BLD, 10 in TIB, and 5 in COR (Fig. 3a). Most genes revealed significant associations in only a single tissue; 38 were significant in more than one, almost all having consistent directions of association between predicted expression levels and CAD across tissues (Fig. 3b).

Among the 114 genes, 102 were protein-coding and 12 were lnRNA genes (Supplementary Table 4). The STARNET data showed that most lncRNAs were positively co-expressed with a surrounding gene in affected tissues (Supplementary Fig. 8). *LINC00310* was the only exception, which displayed complex co-expression patterns with other genes.

Respective genes were found in 63 genomic regions, thus several regions represented multiple genes with significant associations. Six regions had multiple TWAS genes with shared GWAS and eQTL signals in respective tissues, like 1p13.3 and 2p33.2 (Supplementary Figs. 9–10; Supplementary Table 5). On the other hand, in 39 regions expression of only a single gene was found to be significantly associated, which makes these genes likely candidates for mediating causal effects, particularly, if these genes reside within GWAS risk loci for CAD (these genes are indicated in Supplementary Table 6).

Most TWAS genes ($n=96$) could be positionally annotated to the 1 Mb region around one of the over 200 GWAS loci that are currently known to be genome-wide significantly associated with CAD [11, 17, 35]. Therefore, we marked these as known genes (Supplementary

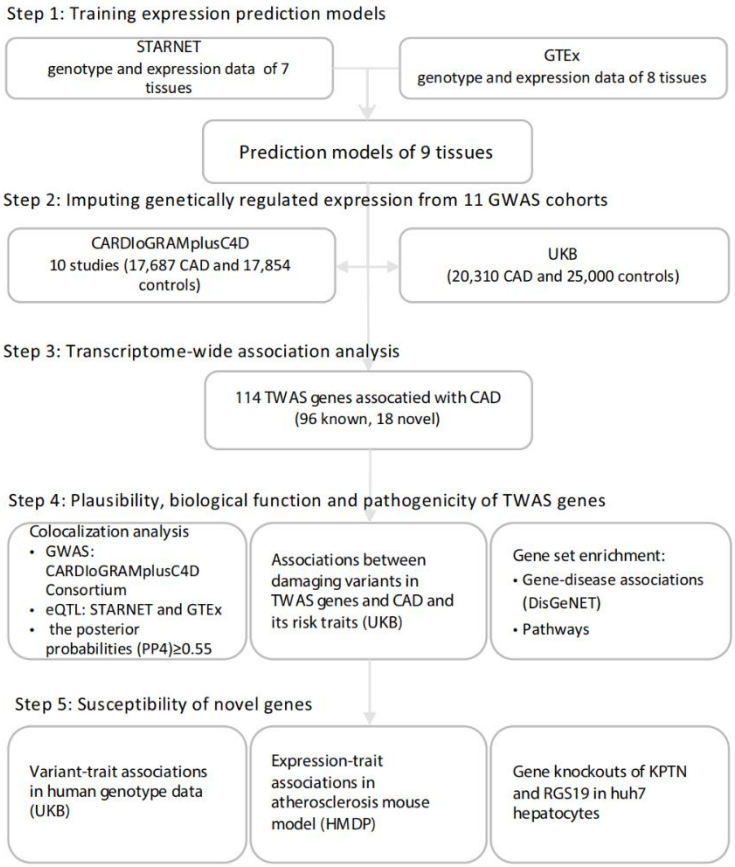


Fig. 1 The study design. Step 1, we trained prediction models using EpiXcan from two eQTL panels, the Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task (STARNET) and the Genotype-Tissue Expression (GTEx) for nine tissues. Step 2, the prediction models were applied to impute genetically regulated expression (GReX) from individual-level genotype data of ten CARDIoGRAMplusC4D sets and UK Biobank (UKB). Step 3, we associated transcriptome-wide GReX with risk of coronary artery disease (CAD) (Supplementary Results) and identified 114 transcriptome-wide significant genes (TWAS genes). Of these, 96 resided within genome-wide significant (GWAS) loci and 18 outside of known GWAS loci (novel genes). Step 4, we tested the plausibility of novel TWAS genes by conducting colocalization analysis and studying effects of damaging mutations, as well as gene set enrichment analyses. Step 5, we explored potential mechanisms of novel genes by testing association with risk traits of CAD in human genotype data of UKB, and association between expressions and risk traits of CAD in atherosclerosis mouse models from the Hybrid Mouse Diversity Panel (HMDP). Lastly, we carried out CRISPR/Cas9-based knockdown experiment for two novel genes *RGS19* and *KPTN* in human hepatocyte cell lines to experimentally validate related functions

Table 6). On the other hand, 18 genes resided outside of these regions and were labeled as novel genes (Table 1). Most novel genes were tissue-specific, except *RGS19*, *FAM114A1* and *UFL1* which displayed evidence for differential expression in multiple tissues.

Pathways and diseases enriched by TWAS genes

We carried out two types of gene set enrichment tests to further study the biological relevance of genes giving signals in this TWAS. First, we studied disease-gene sets from the DisGeNET platform which is one of the largest publicly

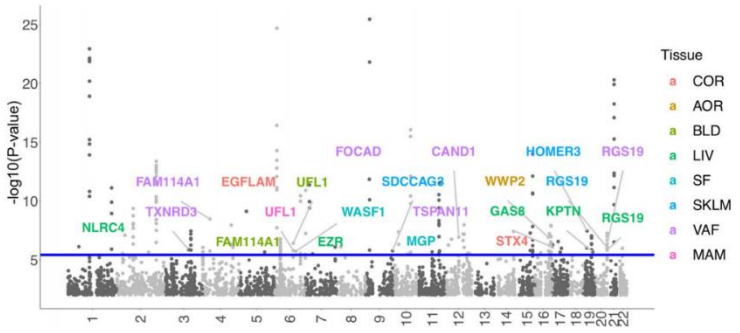


Fig. 2 Manhattan plot of CAD TWAS results. The association results from STARNET- and GTEx-based models were integrated by lowest *P* values. The blue line marks $P=3.85e{-}6$, i.e. transcriptome-wide significance. Each point corresponds to an association test between gene-tissue pair. 18 novel TWAS genes were highlighted. Supplementary Fig. 6 identifies all genes identified by their genetically-modulated association signals. The color code identifies the tissue in which the genes were differentially expressed by genetic means: *AOR* aorta, *COR* coronary artery, *MAM* mammary artery, *BLD* blood, *LIV* liver, *SF* subcutaneous fat, *VAF* visceral abdominal fat, *SKLM* skeletal muscle

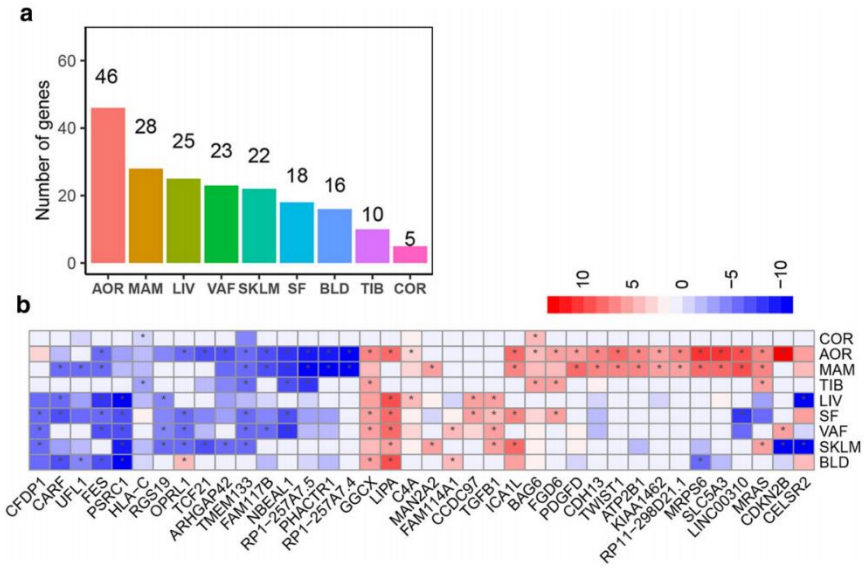


Fig. 3 Tissue distribution of 114 TWAS genes of CAD. **a** Number of transcriptome-wide significant genes across tissues. **b** Heatmap plot of 38 genes identified in more than one tissues. The color codes indicate direction of effects. Cells marked with * represent significant gene-tissue pairs ($P<3.85e{-}6$). *AOR* aorta, *COR* coronary artery, *MAM* mammary artery, *BLD* blood, *LIV* liver, *SF* subcutaneous fat, *VAF* visceral abdominal fat, *SKLM* skeletal muscle, *TIB* tibial artery

available collections of genes and variants associated with human diseases [50]. The results showed that genes discovered by TWAS were primarily enriched for CAD, coronary atherosclerosis, and hypercholesterolemia (Supplementary Table 7), adding to the plausibility of our TWAS findings.

Table 1 18 TWAS genes residing outside of published GWAS loci

Cytoband	Gene	Tissue	Z score	SE	P value	From[a]
2p22.3	*NLRC4*	LIV	−3.383	0.044	3.04E−06	STARNET
3q21.3	*TXNRD3*	VAF	2.566	0.059	1.36E−06	STARNET
4p14	*FAM114A1*	VAF	4.026	0.050	3.44E−09	GTEx
4p14	*FAM114A1*	BLD	4.845	0.037	1.80E−06	GTEx
5p13.2	*EGFLAM*	COR	5.596	0.047	7.70E−10	GTEx
6q16.1	*UFL1*	MAM	−5.246	0.038	1.62E−06	STARNET
6q16.1	*UFL1*	BLD	−4.687	0.038	8.70E−05	STARNET
6q16.1	*UFL1*	BLD	−4.955	0.042	3.96E−07	GTEx
6q21	*WASF1*	SF	4.320	0.059	1.91E−06	STARNET
6q25.3	*EZR*	LIV	−3.187	0.025	3.53E−06	STARNET
9p21.3	*FOCAD*	VAF	8.348	0.068	1.44E−12	GTEx
9q34.3	*SDCCAG3*	SKLM	−3.015	0.061	1.74E−06	STARNET
12p11.21	*TSPAN11*	VAF	2.285	0.065	1.79E−07	STARNET
12p12.3	*MGP*	SF	−3.412	0.040	5.67E−07	GTEx
12q14.3	*CAND1*	VAF	−2.355	0.030	1.19E−07	GTEx
16p11.2	*STX4*	COR	3.347	0.056	2.59E−06	GTEx
16q22.1	*WWP2*	AOR	4.491	0.029	5.67E−06	STARNET
16q22.1	*WWP2*	AOR	6.570	0.031	1.19E−07	GTEx
16q24.3	*GAS8*	LIV	0.189	0.041	8.32E−07	GTEx
19p13.11	*HOMER3*	SKLM	4.647	0.030	3.52E−08	GTEx
19q13.32	*KPTN*	LIV	−3.076	0.076	2.17E−06	STARNET
20q13.33	*RGS19*	LIV	−4.913	0.028	1.52E−06	GTEx
20q13.33	*RGS19*	VAF	−4.545	0.030	4.63E−07	GTEx
20q13.33	*RGS19*	SKLM	−5.026	0.024	1.42E−06	STARNET
20q13.33	*RGS19*	SKLM	−5.298	0.018	9.29E−07	GTEx

TWAS transcriptome-wide association study, *STARNET* the Stockholm-Tartu Atherosclerosis Reverse Network Engineering panel, *GTEx* the Genotype-Tissue Expression panel, *AOR* aorta, *COR* coronary artery, *MAM* mammary artery, *BLD* blood, *LIV* liver, *SF* subcutaneous fat, *VAF* visceral abdominal fat, *SKLM* skeletal muscle

[a]Association statistics from either STARNET- or GTEx-based models

In line with these results, gene set enrichment analysis based on GO [26], KEGG [30], Reactome [12], and WikiPathways [55] databases showed that the TWAS genes were highly enriched for pathways involved in cholesterol metabolism and regulation of lipoprotein levels. To a lesser extent, risk genes were enriched in regulation of blood pressure as well as development and morphogenesis of the heart and the aortic valve (Supplementary Table 8).

Effects of damaging variants in TWAS genes

We next searched in exome-sequencing data of 200,643 participants from UKB for rare damaging variants in TWAS genes (either loss-of-function mutations or mutations predicted to be adverse by one of five in-silico methods, allele frequency < 0.01) (Materials and methods). In 97 genes we detected such variants. Expectedly these damaging mutations were very rare which limits the power of gene-based burden tests to observe association with risk of CAD or one of 14 CAD-related cardiometabolic traits we tested (15 traits in total). Nevertheless, associations of eight genes with risk traits reached Bonferroni-corrected significance ($P < 3.44e{-}5$; 0.05/97genes × 15traits) (Fig. 4; Supplementary Tables 9–10). Mutations of lipoprotein lipase (*LPL*), a critical regulator of lipid metabolism [29, 60], were evidently associated with lipid traits, including levels of HDL-C (beta = −0.106; $P = 4.54e{-}68$), APOA (beta = −0.062; $P = 6.25e{-}47$), APOB (beta = 0.025; $P = 1.38e{-}12$), and TG (beta = 0.241; $P = 1.47e{-}68$). *ABCG5*, encoding a sterol transfer protein [69], was associated with LDL-C (beta = 0.12; $P = 3.66e{-}10$), TC (beta = 0.16; $P = 8.63e{-}10$). *PCSK9*, a drug target for cholesterol lowering [13], was associated with LDL-C (beta = −0.01; $P = 4.29e{-}7$) and APOB (beta = −0.03; $P = 4.4e{-}10$). A mutation of *SARS* was associated with APOB (beta = −0.02; $P = 5.92e{-}7$), *MAT2A* with lymphocyte counts (beta = 1.34; $P = 3.41E{-}28$), and *JCAD* (odds ratio [OR] = 1.31; 95% confidence interval [CI] 1.18–1.46; $P = 5.77e{-}7$) as well as *ARHGAP42* (OR = 2.08; 95% CI 1.65–2.59; $P = 2.22e{-}9$) were associated with risk of diabetes. We also observed nominally

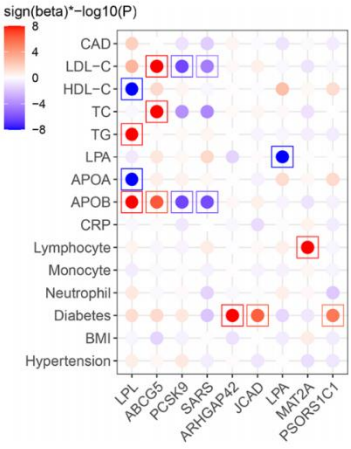


Fig. 4 Effects of damaging variants in TWAS genes on CAD and its risk traits. Sign(beta)*−log10(p) displays direction and significance of gene-trait associations. When the Sign(beta)*−log10(P)>8, they were trimmed to 8. The gene-trait association pairs reached Bonferroni-significance $P<3.44e{-}5$ were highlighted in box. *CAD* coronary artery disease, *LDL-C* low-density lipoproteins cholesterol, *VLDL-C* very low-density lipoprotein cholesterol, *HDL-C* high density lipoproteins cholesterol, *APOA* apolipoprotein A, *APOB* apolipoprotein B, *TC* total cholesterol, *TG* triglycerides, *CRP* C-reactive protein, *BMI* body mass index

significant associations of several genes with CAD: *LPL* [29, 60] (OR = 1.168; CI 1.034–1.036; $P=0.016$), *NOS3* [15] (OR = 1.143; 95% CI 1.109–1.279; $P=0.02$), and *ADAMTS7* [32] (OR = 1.062; 95% CI 1.011–1.115; $P=0.016$) (Supplementary Tables 9–10).

Novel genes associated with risk factors in human and mouse data

We next associated single nucleotide polymorphisms (SNPs) in the regions of ± 1 Mb around the 18 novel TWAS genes to study their associations with a series of lipid traits including LDL-C, HDL-C, APOA, APOB, LPA, TC, and TG in UKB (Materials and methods). There were 883 independent SNPs estimated by GEC. Bonferroni-corrected significance $P<8.09e{-}6$ (0.05/883 × 7 lipid traits) was observed for numerous respective lead variants, of which *RGS19*, *SDCCAG3*, *EZR*, *HOMER3*, and *WWP2* reached genome-wide significant association ($P<5e{-}8$) with multiple lipid traits (Fig. 5a; Supplementary Table 11).

Next, we extracted expression-trait association statistics of TWAS genes from HMDP, which brings together genotypes and expression data from atherosclerosis mouse models [42]. Based on the expression data from mouse aorta and liver tissues, 55 TWAS genes were significantly associated with aortic lesion area and 14 further cardiovascular traits ($P<0.05$; Supplementary Table 12). Expression levels of seven novel genes, i.e. *Rgs19*, *Kptn*, *Ezr*, *Stx4a*, *Cand1*, *Focad*, and *Wasf1*, were associated with aortic lesion

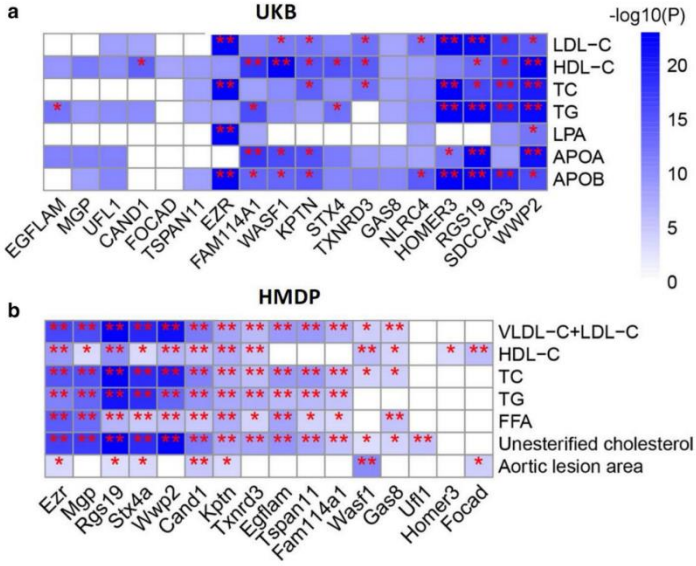


Fig. 5 Novel risk genes were associated with lipid traits. **a** Data from UK Biobank (UKB) indicated that lead variants inside the boundary of risk genes were associated with lipid traits with Bonferroni-corrected significance levels (*$P<8.09e{-}6$), or by genome-wide significance (**$P<5e{-}8$). **b** Expression levels of novel genes were likewise associated with lipid traits and aortic lesion area in an atherosclerosis mouse model from the hybrid mouse diversity panel (HMDP). *$P<0.05$; **FDR < 0.05. *LDL-C* low-density lipoproteins cholesterol, *VLDL-C* very low-density lipoprotein cholesterol, *HDL-C* high density lipoproteins cholesterol, *APOA* apolipoprotein A, *APOB* apolipoprotein B, *TC* total cholesterol, *TG* triglycerides, *FFA* free fatty acid

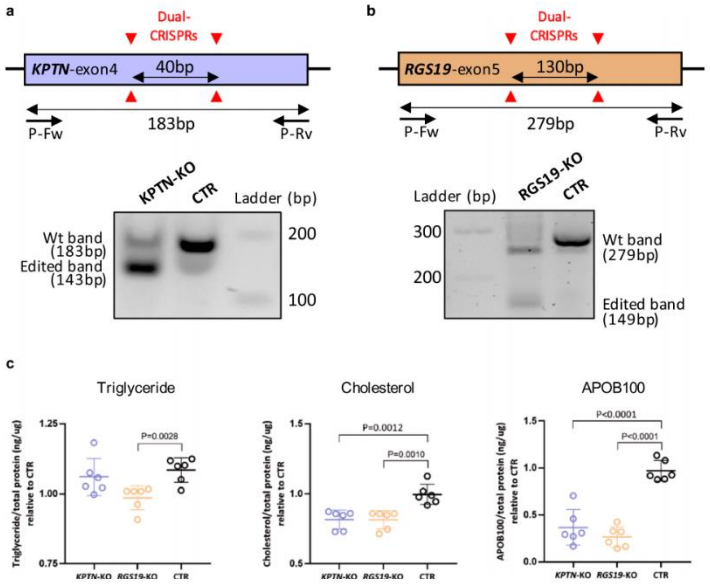


Fig. 6 Targeting of *KPTN* and *RGS19* reduced lipids and APOB secretion of human liver cells. **a** Two sgRNAs were used to target the exon4 of *KPTN* (shared exon among isoforms) in a Cas9-expressing huh7 liver cell line. The dual CRISPR strategy created a 40 bp frame shift deletion in the gene and profound reduction of *KPTN* at both mRNA and protein levels (Supplementary Fig. 11c, d). The primers (P-Fw and P-Rv) used for analyzing the CRISPR editing as indicated. **b** The same strategy was used for *RGS19* targeting, which resulted in a 130 bp frame shift deletion in the gene, and reduction of mRNA and protein (Supplementary Fig. 11c, d). **c** Reduced triglyceride and cholesterol levels in knockout (KO) cell lines were detected by colorimetric method and APOB100 secretion was measured by human APOB100 Elisa ($n=6$). Triglyceride, cholesterol, and APOB100 levels were normalized to total protein and compared between the KO and control (CTR) cell lines

area (Fig. 5b), a commonly used measure for atherosclerotic plaque formation in mice. Additionally, we found the novel genes were associated with at least one lipid trait in the mouse model.

Knockdown of *RGS19* and *KPTN* in human liver cells

Potential functional implications of all novel genes, based on the literature, are summarized in Supplementary Table 13. We additionally aimed to validate two exemplary novel TWAS genes by in vitro studies. Based on the above in-silico annotations we focused these studies on novel genes identified in liver with potential effects on lipids, the top risk factor for CAD (Fig. 5). Among the five genes identified in liver including *NLRC4*, *EZR*, *GAS8*, *KPTN*, and *RGS19*, the last two were, not only the least studied but also associated with nearly a full spectrum of lipid traits in human or mouse data (Fig. 5). In addition, both *KPTN* and *RGS19* are indeed expressed in hepatocyte (Supplementary Fig. 11a, b). Finally, both *KPTN* and *RGS19* are located within lipid loci identified recently in more than one million individuals [24]. Therefore, we decided to test the influence of *KPTN* and *RGS19* on lipid metabolism of liver cells.

We generated gene knockout (KO) huh7 cell lines by a dual CRISPR strategy (Materials and methods), which substantially reduced expression of the respective genes (Supplementary Fig. 11c, d). We measured secretion levels of TG, cholesterol and APOB in gene-targeted versus control cells. Notably, under normal circumstances, human hepatocytes synthesize cholesterol, assemble TG and APOB100, and secrete these particles in form of VLDL-C [58]. Compared to control huh7 cells, we found reduced APOB and cholesterol levels in culture medium of *KPTN*-KO cells (Fig. 6a, c). In culture medium of *RGS19*-KO cells we also detected reduced levels of APOB100, cholesterol, and TG (Fig. 6b, c), in line with strong associations of this gene with an array of lipid traits in both human genotyping and mouse expression data sets (Fig. 5).

We further corroborated our experimental results by performing RNA sequencing (RNA-seq) on *KPTN*-KO and

RGS19-KO hepatocytes. In comparison to control cells, dysregulated genes in *KPTN*-KO and *RGS19*-KO hepatocytes ($P < 0.05$; Supplementary Tables 14–15) were indeed enriched for lipid metabolism (Supplementary Fig. 12). For *KPTN*-KO hepatocytes, the top four significantly enriched pathways plausibly contribute to CAD risk. Pathways ranked 1 and 3, 'regulation of cholesterol esterification' and 'LDL particle remodeling', strongly suggested that *KPTN* can affect CAD risk via cholesterol metabolism (Supplementary Fig. 12a, b). For *RGS19*-KO hepatocytes, the dysregulated genes were enriched for both cholesterol and triglycerides metabolisms (Supplementary Fig. 12c, d) and eight of the top ten significant enriched pathways were related to lipid metabolism, consistent with the reduced secretion of cholesterol and triglyceride of *RGS19*-KO cells (Fig. 6c).

Discussion

In a stepwise approach, we first generated models which allow to predict gene expression based on genotypes in nine tissues. Next, we applied these models to individual-level genotype data on more than 80,000 CAD cases and controls to perform a transcriptome-wide association analysis. We identified 114 genes with differential expression by genetic means in CAD patients. Many signals were highly plausible as they resided within loci displaying genome-wide significant association with CAD by traditional GWAS. By in-silico analyses, these genes were markedly enriched in established pathways for the disease. Moreover, damaging variants in these genes showed association with CAD risk or its underlying traits in whole exome sequencing data from UKB. Importantly, we also identified 18 genes without prior evidence for their involvement in CAD by GWAS, many of which were found to be associated with lipid metabolism in human and mouse data.

Only a minority of genes residing within published CAD GWAS loci have been validated experimentally for their underlying causal role in atherosclerosis. Our data provide a substantial step towards prioritization of genes at respective GWAS loci [17, 35], because the TWAS association finding is based on expression levels of specific genes in defined tissues. In this respect, 46 genes identified by this TWAS are known for effects in pathophysiological pathways related to CAD, including lipid metabolism, inflammation, angiogenesis, transcriptional regulation, cell proliferation, NO signaling, and high blood pressure, to name a few (Supplementary Table 6), giving credibility to the association findings.

Interestingly, our TWAS uncovered eight novel gene-CAD associations in fat tissue, including *MGP* and *WASF1* in SF, and *CAND1*, *FAM114A1*, *FOCAD*, *RGS19*, *TSPAN11*, and *TXNRD3* in VAF, representing half of the novel genes. All these genes also showed significant association with multiple lipid traits in a mouse atherosclerosis model (Fig. 5b). Given many CAD patients that are overweight or obese, it will be of great interest to identify how these genes modify cardiometabolic traits leading to cardiovascular disorders. In this respect our TWAS could provide a list of candidate genes and related targetable cardiometabolic traits. In addition, it is of surprise to unveil 22 genes linking SKLM to CAD risk, and eight were unique to this tissue, including *HOMER3*, *SDCCAG3*, *MTAP*, *NME9*, *PSMA4*, *SLC2A12*, *UNC119B*, and *VAMP5*, the first two being novel. *SDCCAG3* or *ENTR1* encodes endosome associated trafficking regulator 1 and involves in recycling of *GLUT1* (glucose transporter type 1), supplying the major energy source for muscle contraction. SKLM-based metabolism may have a protective role in CAD as suggested by the many cardioprotective effects of sports [44, 54]. Gene targets enhancing SKLM function in this respect might be effective in CAD prevention, a field relatively unexplored thus far. Here, for the first time, quantitative traits regulated genes in SKLM were associated with CAD by TWAS, providing novel evidence for genes that could modulate CAD risk by their functions in SKLM.

Many novel TWAS genes revealed association with lipid traits in both genotype-trait data of human biobank and expression-trait data of atherosclerosis mouse model. For example, *KPTN* and *RGS19*, both novel genes displaying significant TWAS results for CAD—based on their genetically-modulated expression profiles of liver tissue—also showed significant association with various lipid traits as well as aortic lesion area in the atherosclerosis mouse model. Moreover, both gene loci harbor SNPs which are significantly associated with several lipids including LDL-C, HDL-C, TC, and/or TG in human genotype data. Based on these observations, we functionally validated the roles of these two novel genes by studying lipid levels in human liver cells, i.e. the tissue that displayed evidence for differential expression by TWAS. Indeed, we observed that knockout of the two genes lowered secretion of APOB and lipids into culture medium. *KPTN*, kaptin (actin binding protein), a member of the *KPTN*, *ITFG2*, *C12orf66*, and *SZT2* (KICSTOR) protein complex, is a lysosome-associated negative regulator of the mechanistic target of rapamycin complex 1 (mTORC1) signaling [67]. By investigating dysregulated genes of *KPTN*-KO hepatocytes, we found many genes of mTORC1 pathway to be upregulated (Supplementary Fig. 12b), including a subunit of mTORC1, namely, *MLST8* (MTOR associated protein, LST8 Homolog). Interestingly, many lysosome genes were also significantly upregulated including *LAMP1* (lysosomal associated membrane protein 1), *ACP2* (acid phosphatase 2, lysosomal), *AP1B1* (adaptor related protein complex 1 subunit beta 1), *ATP6V0C* (ATPase H+ transporting V0 subunit c), *CTNS* (cystinosin, lysosomal cystine transporter), *CTSA*

(cathepsin A), and *CLTB* (clathrin light chain B). Lysosomes promote lipid catabolism and transport, and maintain cellular lipid homeostasis [57]. Activated mTORC1 located on lysosome membrane [67], acts as a sensor of lysosomal lipids [57], such as cholesterol and phosphatidic acid, which exert as building block for cellular and subcellular membrane system. In fact, cholesterol mediates mTORC1 activation at the lysosome [9]. The interaction of mTORC1 and lysosome may promote lipid-sensing and lipid-trafficking to support the function of other subcellular organelles [46, 57]. These results hint the enhanced cellular usage of cholesterol via mTOR-lysosome axis in *KPTN*-KO hepatocytes. In addition, several lipoprotein genes were downregulated as well, including *APOA1*, *APOA2*, *APOA4*, and *APOB*. Both processes might contribute to the reduced cholesterol secretion and the association with CAD.

RGS19 belongs to the *RGS* (regulators of G-protein signaling) family, who are regulators for G-protein-coupled receptors (GPCRs) [49]. *RGS19* inhibits GPCR signal transduction by increasing the GTPase activity of G-protein alpha subunits, thereby transforming them into an inactive GDP-bound form [53, 59]. The targeting GPCR of *RGS19* has not been observed before, and how *RGS19* regulates lipid metabolism remains unclear. The *RGS19* locus was first reported to be associated with TC and TG in 2017 [33, 41]. We observed significant association of this gene with CAD and functionally validated its role in TG and cholesterol secretion. A potential mechanism could be related to PPARα pathway that regulates the expression of genes involving hepatic lipogenesis and lipid storage [63, 66]. PPARα also regulates cholesterol, bile acid homeostasis, and sphingolipid metabolism in the liver [22]. Many genes in PPARα pathway were significantly downregulated in *RGS19*-KO hepatocytes, including *FABP1* (also known as liver fatty acid binding protein), *PLTP* (phospholipid transfer protein), *APOA1*, *APOA2*, and *APOC3* (Supplementary Fig. 12d). *RGS19* is a regulator for G-protein-coupled receptors (GPCR). Interestingly, we found six dysregulated GPCRs in *RGS19*-KO hepatocytes, including, *ADGRL2*, *CELSR1*, *ADGRV1*, *OXER1*, *LGR5*, and *LGR4* (Supplementary Fig. 12d). Furthermore, one of them, *OXER1,* an activator PPARα [51], was also downregulated in *RGS19*-KO cells. All in all, one hypothesis could be that *RGS19* associated GPCR signaling affects the PPARα pathway, and thereby lipid metabolism and CAD risk. Previous and current studies concordantly suggest from different angles that *RGS19* has a role in lipid metabolism and our data further indicate that this function might meditate its effects on CAD risk.

There are certain limitations in our study. First, we observed that about 15% of gene-tissue pairs displayed some degree of heterogeneity in the association findings with CAD risk across the cohorts (Supplementary Table 4). While this number is relatively low and likely result from a play of chance when association findings are being compared across individuals with relatively small case–control samples, it might also indicate some degree of population specific effects within European ancestries from UK, Germany, France, and Italy. Second, since TWAS are strongly dependent on the reference panel linking genetic signatures with gene expression, it had to be expected that STARNET- and GTEx-based predictive models display some differences in gene-CAD associations. STARNET-based TWAS identified 129 gene-tissue pairs, whereas GTEx-based TWAS identified 106 gene-tissue pairs. Yet, 42 gene-tissue pairs were shared between the two analyses, and effect sizes for the shared genes were highly concordant ($\rho = 0.97$). An average of 62% overlapping genes was observed in the matched tissues of two reference-based models, and the resulting size of expression-CAD associations was linearly consistent with an average $\rho = 0.72$. The relatively small differences may be due to different sample sizes used for training predictive models [70], different disease states (subjects with and without CAD), intravital (STARNET) or post mortem (GTEx) sample collection, as well as different transcript abundance and genotype coverage leading to differences in expression associated SNPs in our reference panels [20, 25]. Given a fair consistency between the two data sources, we combined results derived from both panels to increase the power for capturing risk genes. Third, although TWAS facilitates candidate risk gene prioritization, co-regulation or co-expression *in cis* at a given locus limits the precise determination of the culprit gene [62]. Indeed, at 12 loci we observed signals for three or more TWAS genes. For instance, in LIV tissue TWAS identified five genes at 1p13.3, *ATXN7L2*, *CELSR2*, *PSMA5*, *PSRC1*, *SARS*, and *SORT1* which were co-regulated by same risk variant set, confusing prioritization of the causal gene. While *CELSR2*, *PSRC1*, and *SORT1* were previously shown to act on lipid metabolism [3], we found that damaging mutations in *SARS* were also associated with serum levels of HDL-C and APOA. Thus, a combined effects of some or all genes at this locus may contribute to the association signal. In addition, all lncRNA genes identified by our study displayed co-expression with their neighboring coding genes, which makes it difficult to determine their casual effects. Nevertheless, in combining TWAS data with other genetic analyses, e.g. effects of damaging mutations, genetic association with other phenotypes and expression-traits association statistics, we aimed to improve risk gene prioritization, and to provide deeper insights of possible disease-causing mechanisms. For instance, *LPL* is well-known for its protective role against CAD by lowering lipids [29, 60], and our analyses showed that damaging *LPL* mutations were associated with higher lipid levels. Last, as with all statistical methods, there are certain limitations and assumptions associated with TWAS. Further evolution and improvement of these methods, as well as functional

validation experiments, will assuredly improve the accuracy of these studies.

In summary, our TWAS study based on two genetics-of-gene-expression panels identified 114 gene expression-CAD associations, of which 18 were novel. The extended analyses with multiple datasets supported the reliability of the CAD TWAS signals in prioritizing candidate risk genes and identifying novel associations in a tissue-specific manner. Functional validation of two novel genes, *RGS19* and *KPTN,* lend support to our TWAS findings and provide strong evidence for their role in lipid metabolism. Thus, our study created a set of gene-centered and tissue-annotated associations for CAD, providing insightful guidance for further biological investigation and therapeutic development.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00395-022-00917-8.

Acknowledgements We'd like to thank Dr. Wen Zhang, Pro. Panos Roussos for sharing the prediction models of seven tissues trained from either STARNET or GTEx panels, as well as the EpiXcan pipeline.

Author contributions HS, LL, ZC designed the study and wrote the manuscript. LL ran analyses. ZC, SL and AS performed experiment. MS, UG, SK, AM, KH, CP, AJL, SP, TK, RE, KS, AR, JG, JE, JCK, JLMB provided research data, technical support and gave conceptual advice.

Funding Open Access funding enabled and organized by Projekt DEAL. The work was funded by the German Federal Ministry of Education and Research (BMBF) within the framework of ERA-NET on Cardiovascular Disease (Druggable-MI-genes: 01KL1802), within the scheme of target validation (BlockCAD: 16GW0198K), COMMITMENT (01ZX1904A), and within the scheme of e:Med research and funding concept (AbCD-Net: 01ZX1706C). As a Co-applicant of the British Heart Foundation (BHF)/German Centre of Cardiovascular Research (DZHK)-collaboration (DZHK-BHF: 81X2600522) and the Leducq Foundation for Cardiovascular Research (PlaqOmics: 18CVD02), we gratefully acknowledge their funding. Additional support has been received from the German Research Foundation (DFG) as part of the Sonderforschungsbereich SFB 1123 (B02) and the Sonderforschungsbereich SFB TRR 267 (B05). Further, we kindly acknowledge the support of the Bavarian State Ministry of Health and Care who funded this work with DigiMed Bayern (grant No: DMB-1805–0001) within its Masterplan "Bayern Digital II" and of the German Federal Ministry of Economics and Energy in its scheme of ModulMax (grant No: ZF4590201BA8). Finally, we acknowledge the support of DZHK with Postdoc Start-up Grant ("Förderkennzeichen").

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Tools and data EpiXcan pipeline: https://bitbucket.org/roussoslab/epixcan/src/master/, and predictive models based on STARNET and GTEx databases: http://predictdb.org/.

PrediXcan pipeline: https://github.com/hakyim/PrediXcan.
qtlBHM: https://github.com/rajanil/qtlBHM.

STARNET database: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001203.v1.p1. Project ID: 13585.
GTEx database: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000424.v8.p2. Project ID: 20848.
UK Biobank: https://www.ukbiobank.ac.uk/. Project ID: 25214.
MAGMA: https://ctg.cncr.nl/software/magma.
R package for colocalization analysis, coloc: https://cran.r-project.org/web/packages/coloc/vignettes/vignette.html.
DisGeNET: https://www.disgenet.org/.
CARDIoGRAMplusC4D Consortium: http://www.cardiogramplusc4d.org/.

References

1. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, Mohammadi P, Park YS, Parsana P, Segrè AV, Strober BJ, Zappala Z, Cummings BB, Gelfand ET, Hadley K, Huang KH, Lek M, Li X, Nedzel JL, Nguyen DY, Noble MS, Sullivan TJ, Tukiainen T, MacArthur DG, Getz G, Addington A, Guan P, Koester S, Little AR, Lockhart NC, Moore HM, Rao A, Struewing JP, Volpi S, Brigham LE, Hasz R, Hunter M, Johns C, Johnson M, Kopen G, Leinweber WF, Lonsdale JT, McDonald A, Mestichelli B, Myer K, Roe B, Salvatore M, Shad S, Thomas JA, Walters G, Washington M, Wheeler J, Bridge J, Foster BA, Gillard BM, Karasik E, Kumar R, Miklos M, Moser MT, Jewell SD, Montroy RG, Rohrer DC, Valley D, Mash DC, Davis DA, Sobin L, Barcus ME, Branton PA, Abell NS, Balliu B, Delaneau O, Frésard L, Gamazon ER, Garrido-Martín D, Gewirtz ADH, Gliner G, Gloudemans MJ, Han B, He AZ, Hormozdiari F, Li X, Liu B, Kang EY, McDowell IC, Ongen H, Palowitch JJ, Peterson CB, Quon G, Ripke S, Saha A, Shabalin AA, Shimko TC, Sul JH, Teran NA, Tsang EK, Zhang H, Zhou YH, Bustamante CD, Cox NJ, Guigó R, Kellis M, McCarthy MI, Conrad DF, Eskin E, Li G, Nobel AB, Sabatti C, Stranger BE, Wen X, Wright FA, Ardlie KG, Dermitzakis ET, Lappalainen T, Battle A, Brown CD, Engelhardt BE, Montgomery SB, Handsaker RE, Kashin S, Karczewski KJ, Nguyen DT, Trowbridge CA, Barshir R, Basha O, Bogu GK, Chen LS, Chiang C, Damani FN, Ferreira PG, Hall IM, Howald C, Im HK, Kim Y, Kim-Hellmuth S, Mangul S, Monlong J, Muñoz-Aguirre M, Ndungu AW, Nicolae DL, Oliva M, Panousis N, Papasaikas P, Payne AJ, Quan J, Reverter F, Sammeth M, Scott AJ, Sodaei R, Stephens M, Urbut S, Van De Bunt M, Wang G, Xi HS, Yeger-Lotem E, Zaugg JB, Akey JM, Bates D, Chan J, Claussnitzer M, Demanelis K, Diegel M, Doherty JA, Feinberg AP, Fernando MS, Halow J, Hansen KD, Haugen E, Hickey PF, Hou L, Jasmine F, Jian R, Jiang L, Johnson A, Kaul R, Kibriya MG, Lee K, Li JB, Li Q, Lin J, Lin S, Linder S, Linke C, Liu Y, Maurano MT, Molinie B, Nelson J, Neri FJ, Park Y, Pierce BL, Rinaldi NJ, Rizzardi LF, Sandstrom R, Skol A, Smith KS, Snyder MP, Stamatoyannopoulos J, Tang H, Wang L, Wang M, Van Wittenberghe N, Wu F, Zhang R, Nierras

CR, Carithers LJ, Vaught JB, Gould SE, Lockart NC, Martin C, Addington AM, Koester SE, Undale AH, Smith AM, Tabor DE, Roche NV, McLean JA, Vatanian N, Robinson KL, Valentino KM, Qi L, Hunter S, Hariharan P, Singh S, Um KS, Matose T, Tomaszewski MM, Barker LK, Mosavel M, Siminoff LA, Traino HM, Flicek P, Juettemann T, Ruffier M, Sheppard D, Taylor K, Trevanion SJ, Zerbino DR, Craft B, Goldman M, Haeussler M, Kent WJ, Lee CM, Paten B, Rosenbloom KR, Vivian J, Zhu J (2017) Genetic effects on gene expression across human tissues. Nature 550:204–213. https://doi.org/10.1038/nature24277

2. Anderson CD, Rosand J, Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles G V, Clark NR, Ma'ayan A, Kim JY, Kim JY, Song KS, Lee YH, Seo JS, Jelinek J, Goldschmidt-Clermont PJ, Issa JP, Kim M, Long TI, Arakawa K, Wang R, Yu MC, Laird PW, Krizsan-Agbas D, Pedchenko T, Smith PG, Li C, Bazzano LA, Rao DC, Hixson JE, He J, Gu D, Gu CC, Shimmin LC, Jaquish CE, Schwander K, Liu DP, Huang J, Lu F, Cao J, Chong S, Lu X, Kelly TN, Sharma P, Kumar J, Garg G, Kumar A, Patowary A, Karthikeyan G, Ramakrishnan L, Brahmachari V, Sengupta S, Pedrinelli R, Ballo P, Fiorentini C, Denti S, Galderisi M, Ganau A, Germanò G, Innelli P, Paini A, Perlini S, Salvetti M, Zacà V, Besingi W, Johansson Å, Cao TH, Quinn PA, Sandhu JK, Voors AA, Lang CC, Parry HM, Mohan M, Jones DJL, Ng LL, Heart N, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, Goldstein BA, Stirrups KK, König IR, Cazier J-B, Johansson Å, Hall AS, Lee J-YJ-Y, Willer CJ, Chambers JC, Esko T, Folkersen L, Goel A, Grundberg E, Havulinna AS, Ho WK, Hopewell JC, Eriksson N, Kleber ME, Kristiansson K, Lundmark P, Lyytikäinen L-P, Rafelt S, Shungin D, Strawbridge RJ, Thorleifsson G, Tikkanen E, Van Zuydam N, Voight BF, Waite LL, Zhang W, Ziegler A, Absher D, Altshuler D, Balmforth AJ, Barroso I, Braund PS, Burgdorf C, Claudi-Boehm S, Cox D, Dimitriou M, Do R, Doney ASF, Mokhtari NEE, Eriksson P, Fischer K, Fontanillas P, Franco-Cereceda A, Gigante B, Groop L, Gustafsson S, Hager J, Hallmans G, Han B-G, Hunt SE, Kang HM, Illig T, Kessler T, Knowles JW, Kolovou G, Kuusisto J, Langenberg C, Langford C, Leander K, Lokki M-L, Lundmark A, McCarthy MI, Meisinger C, Melander O, Mihailov E, Maouche S, Morris AD, Müller-Nurasyid M, Nikus K, Peden JF, Rayner NW, Rasheed A, Rosinger S, Rubin D, Rumpf MP, Schäfer A, Sivananthan M, Song C, Stewart AFR, Tan S-T, Thorgeirsson G, Schoot CE van der, Wagner PJ, Wells GA, Wild PS, Yang T-P, Amouyel P, Arveiler D, Basart H, Boehnke M, Boerwinkle E, Brambilla P, Cambien F, Cupples AL, de Faire U, Dehghan A, Diemert P, Epstein SE, Evans A, Ferrario MM, Ferrières J, Gauguier D, Go AS, Goodall AH, Gudnason VV, Hazen SL, Holm H, Iribarren C, Jang Y, Kähönen M, Kee F, Kim H-S, Klopp N, Koenig W, Kratzer W, Kuulasmaa K, Laakso M, Laaksonen R, Lee J-YJ-Y, Lind L, Ouwehand WH, Parish S, Park JE, Pedersen NL, Peters A, Quertermous T, Rader DJ, Salomaa V, Schadt E, Shah SH, Sinisalo J, Stark K, Stefansson K, Trégouët D-A, Virtamo J, Wallentin L, Wareham NJN, Zimmermann ME, Nieminen MS, Hengstenberg C, Sandhu MS, Pastinen TM, Syvänen A-C, Hovingh GK, Dedoussis G, Franks PW, Lehtimäki T, Metspalu A, Zalloua PA, Siegbahn A, Schreiber S, Ripatti S, Blankenberg SS, Perola M, Clarke R, Boehm BO, O'Donnell CJ, Reilly MP, März W, Collins R, Kathiresan S, Hamsten A, Kooner JS, Thorsteinsdottir U, Danesh J, Palmer CNA, Roberts R, Watkins H, Schunkert H, Samani NJ, Willer CJ, Schmidt EM, Al E, Facp AFR, Rodriguez-lopez L, Vargas-ayala G, Mc SH, Serna DC, Lozano-nuevo JJ, Rubio-guerra AF, Studies TIC for BPG-WA, Franks WT, Zhou DH, Wylie BJ, Money BG, Graesser DT, Frericks HL, Sahota G, Rienstra CM, Rakyan VK, Down TA, Balding DJ, Beck S, Hager J, Kamatani Y, Cazier J-B, Youhanna S, Ghassibe-Sabbagh M, Platt DE, Abchee AB, Romanos J, Khazen G, Othman R, Badro DA, Haber M, Salloum AK, Douaihy B, Shasha N, Kabbani S, Sbeite H, Chammas E, Bayeh H, Rousseau F, Zelenika D, Gut I, Lathrop M, Farrall M, Gauguier D, Zalloua PA, Myocardial Infarction Genetics C, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand SS, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario MM, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Yee J, Friedlander Y, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Meigs JB, Williams G, Nathan DM, MacRae CA, Havulinna AS, Berglund G, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Daly MJ, Nemesh J, Korn JM, McCarroll SA, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall AS, Wellcome Trust Case Control C, Linsel-Nitschke P, Lieb W, Ziegler A, Konig IR, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann H-EE, Schreiber S, Ouwehand WH, Deloukas P, Scholz M, Cambien F, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney JM, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser V, Epstein SE, Scheffold T, Berger K, Huge A, Martinelli N, Olivieri O, Corrocher R, McKeown PP, Erdmann E, Konig IR, Holm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Do R, Xie C, Siscovick DS, Liang L, Willis-Owen SAG, Laprise C, Wong KCC, Davies GA, Hudson TJ, Binia A, Hopkin JM, Yang IV, Grundberg E, Busche S, Hudson M, Rönnblom L, Pastinen TM, Schwartz DA, Lathrop GM, Moffatt MF, Cookson WOCM, Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM, Mayer B, Erdmann J, Schunkert H, Myocardial Infarction Genetics C, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, Anand SS, Engert JC, Samani NJ, Schunkert H, Erdmann J, Reilly MP, Rader DJ, Morgan T, Spertus JA, Stoll M, Girelli D, McKeown PP, Patterson CC, Siscovick DS, O'Donnell CJ, Elosua R, Peltonen L, Salomaa V, Schwartz SM, Melander O, Altshuler D, Merlini PA, Berzuini C, Bernardinelli L, Peyvandi F, Tubaro M, Celli P, Ferrario MM, Fetiveau R, Marziliano N, Casari G, Galli M, Ribichini F, Rossi M, Bernardi F, Zonzin P, Piazza A, Yee J, Friedlander Y, Marrugat J, Lucas G, Subirana I, Sala J, Ramos R, Meigs JB, Williams G, Nathan DM, MacRae CA, Havulinna AS, Berglund G, Hirschhorn JN, Asselta R, Duga S, Spreafico M, Daly MJ, Nemesh J, Korn JM, McCarroll SA, Surti A, Guiducci C, Gianniny L, Mirel D, Parkin M, Burtt N, Gabriel SB, Thompson JR, Braund PS, Wright BJ, Balmforth AJ, Ball SG, Hall AS, Wellcome Trust Case Control C, Linsel-Nitschke P, Lieb W, Ziegler A, Konig IR, Hengstenberg C, Fischer M, Stark K, Grosshennig A, Preuss M, Wichmann H-EE, Schreiber S, Ouwehand WH, Deloukas P, Scholz M, Cambien F, Li M, Chen Z, Wilensky R, Matthai W, Qasim A, Hakonarson HH, Devaney JM, Burnett MS, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Knouff CW, Waterworth DM, Walker MC, Mooser V, Epstein SE, Scheffold T, Berger K, Huge A, Martinelli N, Olivieri O, Corrocher R, McKeown PP, Erdmann E, Konig IR, Holm H, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Do R, Xie C, Siscovick DS, The CARDIoGRAMplusC4D Consortium, Studies TIC for BPG-WA, The CARDIoGRAMplusC4D Consortium, Preuss M, König IR, Thompson JR, Erdmann J, Absher D, Assimes TL, Blankenberg SS, Boerwinkle E, Chen L, Cupples LA, Hall AS, Halperin E,

Appendix II Transcriptome-wide association study of coronary artery disease identifies novel susceptibility genes

Hengstenberg C, Holm H, Laaksonen R, Li M, Marz W, McPherson R, Musunuru K, Nelson CP, Burnett MS, Epstein SE, O'Donnell CJ, Quertermous T, Rader DJ, Roberts R, Schillert A, Stefansson K, Stewart AFR, Thorleifsson G, Voight BF, Wells GA, Ziegler A, Kathiresan S, Reilly MP, Samani NJ, Schunkert H, Avenell A, Broom J, Brown TJ, Poobalan A, Aucott L, Stearns SC, Smith WCS, Jung RT, Campbell MK, Grant AM, Rakyan VK, Down TA, Balding DJ, Beck S, Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AFR, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NEE El, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason VV, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJP, Khaw K-T, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Mühleisen TW, Muhlestein JB, Münzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nöthen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schäfer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz SM, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups KK, Stoll M, Tang WHW, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJN, Wells GA, Wichmann H-EE, Wild PS, Willenborg C, Witteman JCM, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, März W, Hengstenberg C, Blankenberg SS, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J, Samani NJ, Sharma P, Garg G, Kumar A, Mohammad F, Ramesh S, Slatkin M, Myocardial T, Genetics I, Investigators C, Udali S, Guarini P, Moruzzi S, Choi S, Friso S, Schübeler D, Setten J Van, Isgum I, Smolonska J, Ripke S, Jong PA De, Oudkerk M, Koning H De, Lammers JJ, Zanen P, Groen HJM, Boezen HM, Postma DS, Wijmenga C,Viergever MA, Th WP, Bakker PIW De (2015) Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 478:103–109. https://doi.org/10.1038/nrg3000.Epigenome-Wide

3. Arvind P, Nair J, Jambunathan S, Kakkar VV, Shanker J (2014) CELSR2-PSRC1-SORT1 gene expression and association with coronary artery disease and plasma lipid levels in an Asian Indian cohort. J Cardiol 64:339–346. https://doi.org/10.1016/j.jjcc.2014.02.012
4. Bennett BJ, Davis RC, Civelek M, Orozco L, Wu J, Qi H, Pan C, Packard RRS, Eskin E, Yan M, Kirchgessner T, Wang Z, Li X, Gregory JC, Hazen SL, Gargalovic PS, Lusis AJ (2015) Genetic architecture of atherosclerosis in mice: a systems genetics analysis of common inbred strains. PLoS Genet 11:1005711. https://doi.org/10.1371/journal.pgen.1005711
5. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, Farnham PJ, Hirst M, Lander ES, Mikkelsen TS, Thomson JA (2010) The NIH roadmap epigenomics mapping consortium. Nat Biotechnol 28:1045–1048. https://doi.org/10.1038/nbt1010-1045
6. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J (2009) ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25:1091–1093. https://doi.org/10.1093/bioinformatics/btp101
7. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, Todd JA, Donnelly P, Barrett JC, Davison D, Easton D, Evans D, Leung HT, Marchini JL, Morris AP, Spencer CCA, Tobin MD, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse CV, Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Breen G, St. Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Hamshere ML, Holmans PA, Jones IR, Kirov G, Moskvina V, Nikolov I, O'Donovan MC, Owen MJ, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, Young AH, Ferrier IN, Ball SG, Balmforth AJ, Barrett JH, Bishop DT, Iles MM, Maqbool A, Yuldasheva N, Hall AS, Braund PS, Dixon RJ, Mangino M, Stevens S, Thompson JR, Bredin F, Tremelling M, Parkes M, Drummond H, Lees CW, Nimmo ER, Satsangi J, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Mathew CG, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, Ahmad T, Cummings FR, Jewell DP, Webster J, Brown MJ, Lathrop GM, Connell J, Dominiczak A, Braga Marcano CA, Burke B, Dobson R, Gungadoo J, Lee KL, Munroe PB, Newhouse SJ, Onipinla A, Wallace C, Xue M, Caulfield M, Farrall M, Barton A, Bruce IN, Donovan H, Eyre S, Gilbert PD, Hider SL, Hinks AM, John SL, Potter C, Silman AJ, Symmons DPM, Thomson W, Worthington J, Dunger DB, Widmer B, Frayling TM, Freathy RM, Lango H, Perry JRB, Shields BM, Weedon MN, Hattersley AT, Hitman GA, Walker M, Elliott KS, Groves CJ, Lindgren CM, Rayner NW, Timpson NJ, Zeggini E, Newport M, Sirugo G, Lyons E, Vannberg F, Hill AVS, Bradbury LA, Farrar C, Pointon JJ, Wordsworth P, Brown MA, Franklyn JA, Heward JM, Simmonds MJ, Gough SCL, Seal S, Stratton MR, Rahman N, Ban SM, Goris A, Sawcer SJ, Compston A, Conway D, Jallow M, Rockett KA, Bumpstead SJ, Chaney A, Downes K, Ghori MJR, Gwilliam R, Hunt SE, Inouye M, Keniry A, King E, McGinnis R, Potter S, Ravindrarajah R, Whittaker P, Widden C, Withers D, Cardin NJ, Ferreira T, Pereira-Gale J, Hallgrimsdóttir IB, Howie BN, Spencer CCA, Su Z, Teo YY, Vukcevic D, Bentley D, Compston A (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678.https://doi.org/10.1038/nature05911
8. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J (2018) The UK Biobank resource with deep phenotyping and genomic data. Nature 562:203–209. https://doi.org/10.1038/s41586-018-0579-z
9. Castellano BM, Thelen AM, Moldavski O, Feltes M, Van Der Welle REN, Mydock-McGrane L, Jiang X, Van Eijkeren RJ, Davis OB, Louie SM, Perera RM, Covey DF, Nomura DK, Ory DS, Zoncu R (2017) Lysosomal cholesterol activates mTORC1 via an SLC38A9-Niemann-Pick C1 signaling complex. Science 355:1306–1311. https://doi.org/10.1126/SCIENCE.AAG1417
10. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4:7. https://doi.org/10.1186/s13742-015-0047-8

11. Chen Z, Schunkert H (2021) Genetics of coronary artery disease in the post-GWAS era. J Int Med. https://doi.org/10.1111/JOIM.13362
12. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, Caudy M, Garapati P, Gopinath G, Jassal B, Jupe S, Kalatskaya I, MayMahajan S, May B, Ndegwa N, Schmidt E, Shamovsky V, Yung C, Birney E, Hermjakob H, D'Eustachio P, Stein L (2011) Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res 39:D691–D697. https://doi.org/10.1093/nar/gkq1018
13. Dadu RT (2014) Ballantyne CM (2014) Lipid lowering with PCSK9 inhibitors. Nat Rev Cardiol 1110(11):563–575. https://doi.org/10.1038/nrcardio.2014.84
14. Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, Liu X (2015) Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum Mol Genet 24:2125–2137. https://doi.org/10.1093/hmg/ddu733
15. Emdin CA, Khera AV, Klarin D, Natarajan P, Zekavat SM, Nomura A, Haas M, Aragam K, Ardissino D, Wilson JG, Schunkert H, McPherson R, Watkins H, Elosua R, Bown MJ, Samani NJ, Baber U, Erdmann J, Gormley P, Palotie A, Stitziel NO, Gupta N, Danesh J, Saleheen D, Gabriel S, Kathiresan S (2018) Phenotypic consequences of a genetic predisposition to enhanced nitric oxide signaling. Circulation 137:222–232. https://doi.org/10.1161/CIRCULATIONAHA.117.028021
16. Erdmann J, Großhennig A, Braund PS, König IR, Hengstenberg C, Hall AS, Linsel-Nitschke P, Kathiresan S, Wright B, Trégouët DA, Cambien F, Bruse P, Aherrahrou Z, Wagner AK, Stark K, Schwartz SM, Salomaa V, Elosua R, Melander O, Voight BF, O'Donnell CJ, Peltonen L, Siscovick DS, Altshuler D, Merlini PA, Peyvandi F, Bernardinelli L, Ardissino D, Schillert A, Blankenberg S, Zeller T, Wild P, Schwarz DF, Tiret L, Perret C, Schreiber S, El MNE, Schäfer A, März W, Renner W, Bugert P, Klüter H, Schrezenmeir J, Rubin D, Ball SG, Balmforth AJ, Wichmann HE, Meitinger T, Fischer M, Meisinger C, Baumert J, Peters A, Ouwehand WH, Deloukas P, Thompson JR, Ziegler A, Samani NJ, Schunkert H (2009) New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet 41:280–282. https://doi.org/10.1038/ng.307
17. Erdmann J, Kessler T, Munoz Venegas L, Schunkert H (2018) A decade of genome-wide association studies for coronary artery disease: the challenges ahead. Cardiovasc Res 114:1241–1257. https://doi.org/10.1093/cvr/cvy084
18. Erdmann J, Willenborg C, Nahrstaedt J, Preuss M, Konig IR, Baumert J, Linsel-Nitschke P, Gieger C, Tennstedt S, Belcredi P, Aherrahrou Z, Klopp N, Loley C, Stark K, Hengstenberg C, Bruse P, Freyer J, Wagner AK, Medack A, Lieb W, Grosshennig A, Sager HB, Reinhardt A, Schafer A, Schreiber S, El Mokhtari NE, Raaz-Schrauder D, Illig T, Garlichs CD, Ekici AB, Reis A, Schrezenmeir J, Rubin D, Ziegler A, Wichmann H-E, Doering A, Meisinger C, Meitinger T, Peters A, Schunkert H (2011) Genome-wide association study identifies a new locus for coronary artery disease on chromosome 10p11.23. Eur Heart J 32:158–168. https://doi.org/10.1093/eurheartj/ehq405
19. Foroughi Asl H, Talukdar HA, Kindt ASD, Jain RK, Ermel R, Ruusalepp A, Nguyen KDH, Dobrin R, Reilly DF, Schunkert H, Samani NJ, Braenne I, Erdmann J, Melander O, Qi J, Ivert T, Skogsberg J, Schadt EE, Michoel T, Björkegren JLM (2015) Expression quantitative trait loci acting across multiple tissues are enriched in inherited risk for coronary artery disease. Circ Cardiovasc Genet 8:305–315. https://doi.org/10.1161/CIRCGENETICS.114.000640
20. Franzén O, Ermel R, Cohain A, Akers NK, Di Narzo A, Talukdar HA, Foroughi-Asl H, Giambartolomei C, Fullard JF, Sukhavasi K, Köks S, Gan LM, Giannarelli C, Kovacic JC, Betsholtz C, Losic B, Michoel T, Hao K, Roussos P, Skogsberg J, Ruusalepp A, Schadt EE, Björkegren JLM (2016) Cardiometabolic risk loci share downstream cis- and trans-gene regulation across tissues and diseases. Science (80-) 353:827–830. https://doi.org/10.1126/science.aad6970
21. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, Eyler AE, Denny JC, Nicolae DL, Cox NJ, Im HK, Consortium Gte (2015) A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 47:1091–1098. https://doi.org/10.1038/ng.3367
22. Ghonem NS, Assis DN, Boyer JL (2015) Fibrates and cholestasis. Hepatology 62:635–643. https://doi.org/10.1002/HEP.27744
23. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, Plagnol V (2014) Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet 10:e1004383. https://doi.org/10.1371/journal.pgen.1004383
24. Graham SE, Clarke SL, Wu K-HH, Kanoni S, Zajac GJM, Ramdas S, Surakka I, Ntalla I, Vedantam S, Winkler TW, Locke AE, Marouli E, Hwang MY, Han S, Narita A, Choudhury A, Bentley AR, Ekoru K, Verma A, Trivedi B, Martin HC, Hunt KA, Hui Q, Klarin D, Zhu X, Thorleifsson G, Helgadottir A, Gudbjartsson DF, Holm H, Olafsson I, Akiyama M, Sakaue S, Terao C, Kanai M, Zhou W, Brumpton BM, Rasheed H, Ruotsalainen SE, Havulinna AS, Veturi Y, Feng Q, Rosenthal EA, Lingren T, Pacheco JA, Pendergrass SA, Haessler J, Giulianini F, Bradford Y, Miller JE, Campbell A, Lin K, Millwood IY, Hindy G, Rasheed A, Faul JD, Zhao W, Weir DR, Turman C, Huang H, Graff M, Mahajan A, Brown MR, Zhang W, Yu K, Schmidt EM, Pandit A, Gustafsson S, Yin X, Luan J, Zhao J-H, Matsuda F, Jang H-M, Yoon K, Medina-Gomez C, Pitsillides A, Hottenga JJ, Willemsen G, Wood AR, Ji Y, Gao Z, Haworth S, Mitchell RE, Chai JF, Aadahl M, Yao J, Manichaikul A, Warren HR, Ramirez J, Bork-Jensen J, Kårhus LL, Goel A, Sabater-Lleal M, Noordam R, Sidore C, Fiorillo E, McDaid AF, Marques-Vidal P, Wielscher M, Trompet S, Sattar N, Møllehave LT, Thuesen BH, Munz M, Zeng L, Huang J, Yang B, Poveda A, Kurbasic A, Lamina C, Forer L, Scholz M, Galesloot TE, Bradfield JP, Daw EW, Zmuda JM, Mitchell JS, Fuchsberger C, Christensen H, Brody JA, Feitosa MF, Wojczynski MK, Preuss M, Mangino M, Christofidou P, Verweij N, Benjamins JW, Engmann J, Kember RL, Slieker RC, Lo KS, Zilhao NR, Le P, Kleber ME, Delgado GE, Huo S, Ikeda DD, Iha H, Yang J, Liu J, Leonard HL, Marten J, Schmidt B, Arendt M, Smyth LJ, Cañadas-Garre M, Wang C, Nakatochi M, Wong A, Hutri-Kähönen N, Sim X, Xia R, Huerta-Chagoya A, Fernandez-Lopez JC, Lyssenko V, Ahmed M, Jackson AU, Irvin MR, Oldmeadow C, Kim H-N, Ryu S, Timmers PRHJ, Arbeeva L, Dorajoo R, Lange LA, Chai X, Prasad G, Lorés-Motta L, Pauper M, Long J, Li X, Theusch E, Takeuchi F, Spracklen CN, Loukola A, Bollepalli S, Warner SC, Wang YX, Wei WB, Nutile T, Ruggiero D, Sung YJ, Hung Y-J, Chen S, Liu F, Yang J, Kentistou KA, Gorski M, Brumat M, Meidtner K, Bielak LF, Smith JA, Hebbar P, Farmaki A-E, Hofer E, Lin M, Xue C, Zhang J, Concas MP, Vaccargiu S, van der Most PJ, Pitkänen N, Cade BE, Lee J, van der Laan SW, Chitrala KN, Weiss S, Zimmermann ME, Lee JY, Choi HS, Nethander M, Freitag-Wolf S, Southam L, Rayner NW, Wang CA, Lin S-Y, Wang J-S, Couture C, Lyytikäinen L-P, Nikus K, Cuellar-Partida G, Vestergaard H, Hildalgo B, Giannakopoulou O, Cai Q, Obura MO, van Setten J, Li X, Schwander K, Terzikhan N, Shin JH, Jackson RD, Reiner AP, Martin LW, Chen Z, Li L, Highland HM, Young KL, Kawaguchi T, Thiery J, Bis JC, Nadkarni GN, Launer LJ, Li H, Nalls MA, Raitakari OT, Ichihara S, Wild SH, Nelson CP, Campbell H, Jäger S, Nabika T, Al-Mulla F, Niinikoski H, Braund PS, Kolcic I, Kovacs P, Giardoglou T, Katsuya T, Bhatti KF, de Kleijn D, de Borst GJ, Kim EK, Adams HHH, Ikram MA, Zhu X, Asselbergs FW, Kraaijeveld AO, Beulens JWJ, Shu X-O, Rallidis

LS, Pedersen O, Hansen T, Mitchell P, Hewitt AW, Kähönen M, Pérusse L, Bouchard C, Tönjes A, Chen Y-DI, Pennell CE, Mori TA, Lieb W, Franke A, Ohlsson C, Mellström D, Cho YS, Lee H, Yuan J-M, Koh W-P, Rhee SY, Woo J-T, Heid IM, Stark KJ, Völzke H, Homuth G, Evans MK, Zonderman AB, Polasek O, Pasterkamp G, Hoefer IE, Redline S, Pahkala K, Oldehinkel AJ, Snieder H, Biino G, Schmidt R, Schmidt H, Chen YE, Bandinelli S, Dedoussis G, Thanaraj TA, Kardia SLR, Kato N, Schulze MB, Girotto G, Jung B, Böger CA, Joshi PK, Bennett DA, De Jager PL, Lu X, Mamakou V, Brown M, Caulfield MJ, Munroe PB, Guo X, Ciullo M, Jonas JB, Samani NJ, Kaprio J, Pajukanta P, Adair LS, Bechayda SA, de Silva HJ, Wickremasinghe AR, Krauss RM, Wu J-Y, Zheng W, den Hollander AI, Bharadwaj D, Correa A, Wilson JG, Lind L, Heng C-K, Nelson AE, Golightly YM, Wilson JF, Penninx B, Kim H-L, Attia J, Scott RJ, Rao DC, Arnett DK, Walker M, Koistinen HA, Chandak GR, Yajnik CS, Mercader JM, Tusié-Luna T, Aguilar-Salinas CA, Villalpando CG, Orozco L, Fornage M, Tai ES, van Dam RM, Lehtimäki T, Chaturvedi N, Yokota M, Liu J, Reilly DF, McKnight AJ, Kee F, Jöckel K-H, McCarthy MI, Palmer CNA, Vitart V, Hayward C, Simonsick E, van Duijn CM, Lu F, Qu J, Hishigaki H, Lin X, März W, Parra EJ, Cruz M, Gudnason V, Tardif J-C, Lettre G, 't Hart LM, Elders PJM, Damrauer SM, Kumari M, Kivimaki M, van der Harst P, Spector TD, Loos RJF, Province MA, Psaty BM, Brandslund I, Pramstaller PP, Christensen K, Ripatti S, Widén E, Hakonarson H, Grant SFA, Kiemeney LALM, de Graaf J, Loeffler M, Kronenberg F, Gu D, Erdmann J, Schunkert H, Franks PW, Linneberg A, Jukema JW, Khera AV, Männikkö M, Jarvelin M-R, Kutalik Z, Cucca F, Mook-Kanamori DO, van Dijk KW, Watkins H, Strachan DP, Grarup N, Sever P, Poulter N, Rotter JI, Dantoft TM, Karpe F, Neville MJ, Timpson NJ, Cheng C-Y, Wong T-Y, Khor CC, Sabanayagam C, Peters A, Gieger C, Hattersley AT, Pedersen NL, Magnusson PKE, Boomsma DI, de Geus EJC, Cupples LA, van Meurs JBJ, Ghanbari M, Gordon-Larsen P, Huang W, Kim YJ, Tabara Y, Wareham NJ, Langenberg C, Zeggini E, Kuusisto J, Laakso M, Ingelsson E, Abecasis G, Chambers JC, Kooner JS, de Vries PS, Morrison AC, North KE, Daviglus M, Kraft P, Martin NG, Whitfield JB, Abbas S, Saleheen D, Walters RG, Holmes MV, Black C, Smith BH, Justice AE, Baras A, Buring JE, Ridker PM, Chasman DI, Kooperberg C, Wei W-Q, Jarvik GP, Namjou B, Hayes MG, Ritchie MD, Jousilahti P, Salomaa V, Hveem K, Åsvold BO, Kubo M, Kamatani Y, Okada Y, Murakami Y, Thorsteinsdottir U, Stefansson K, Ho Y-L, Lynch JA, Rader DJ, Tsao PS, Chang K-M, Cho K, O'Donnell CJ, Gaziano JM, Wilson P, Rotimi CN, Hazelhurst S, Ramsay M, Trembath RC, van Heel DA, Tamiya G, Yamamoto M, Kim B-J, Mohlke KL, Frayling TM, Hirschhorn JN, Kathiresan S, Boehnke M, Natarajan P, Peloso GM, Brown CD, Morris AP, Assimes TL, Deloukas P, Sun YV, Willer CJ (2021) The power of genetic diversity in genome-wide association studies of lipids. Nature 600:675–679. https://doi.org/10.1038/s41586-021-04064-3

25. GTEx Consortium (2013) The genotype-tissue expression (GTEx) project. Nat Genet 45:580–585. https://doi.org/10.1038/ng.2653
26. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la Cruz N, Tonellato P, Jaiswal P, Seigfried T, White R (2004) The gene oncology (GO) database and informatics resource. Nucleic Acids Res 32:D258–D261. https://doi.org/10.1093/nar/gkh036
27. Van Hout CV, Tachmazidou I, Backman JD, Hoffman JD, Liu D, Pandey AK, Gonzaga-Jauregui C, Khalid S, Ye B, Banerjee N, Li AH, O'Dushlaine C, Marcketta A, Staples J, Schurmann C, Hawes A, Maxwell E, Barnard L, Lopez A, Penn J, Habegger L, Blumenfeld AL, Bai X, O'Keeffe S, Yadav A, Praveen K, Jones M, Salerno WJ, Chung WK, Surakka I, Willer CJ, Hveem K, Leader JB, Carey DJ, Ledbetter DH, Cardon L, Yancopoulos GD, Economides A, Coppola G, Shuldiner AR, Balasubramanian S, Cantor M, Nelson MR, Whittaker J, Reid JG, Marchini J, Overton JD, Scott RA, Abecasis GR, Yerges-Armstrong L, Baras A (2020) Exome sequencing and characterization of 49,960 individuals in the UK Biobank. Nature 586:749–756. https://doi.org/10.1038/s41586-020-2853-0
28. Huckins LM, Dobbyn A, Ruderfer DM, Hoffman G, Wang W, Pardiñas AF, Rajagopal VM, Als TD, Nguyen H, Girdhar K, Boocock J, Roussos P, Fromer M, Kramer R, Domenici E, Gamazon ER, Purcell S, Johnson JS, Shah HR, Klein LL, Dang KK, Logsdon BA, Mahajan MC, Mangravite LM, Toyoshiba H, Gur RE, Hahn CG, Schadt E, Lewis DA, Haroutunian V, Peters MA, Lipska BK, Buxbaum JD, Hirai K, Perumal TM, Essioux L, Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, Lee P, Bulik-Sullivan B, Collier DA, Huang H, Pers TH, Agartz I, Agerbo E, Albus M, Alexander M, Amin F, Bacanu SA, Begemann M, Belliveau RA, Bene J, Bergen SE, Bevilacqua E, Bigdeli TB, Black DW, Bruggeman R, Buccola NG, Buckner RL, Byerley W, Cahn W, Cai G, Campion D, Cantor RM, Carr VJ, Carrera N, Catts SV, Chambert KD, Chan RCK, Chen RYL, Chen EYH, Cheng W, Cheung EFC, Chong SA, Cloninger CR, Cohen D, Cohen N, Cormican P, Craddock N, Crowley JJ, Curtis D, Davidson M, Davis KL, Degenhardt F, Del Favero J, Demontis D, Dikeos D, Dinan T, Djurovic S, Donohoe G, Drapeau E, Duan J, Dudbridge F, Durmishi N, Eichhammer P, Eriksson J, Escott-Price V, Fanous AH, Farrell MS, Frank J, Franke L, Freedman R, Freimer NB, Friedl M, Friedman JI, Fromer M, Genovese G, Georgieva L, Giegling I, Giusti-Rodríguez P, Godard S, Goldstein JI, Golimbet V, Gopal S, Gratten J, de Haan L, Hammer C, Hamshere ML, Hansen M, Hansen T, Hartmann AM, Henskens FA, Herms S, Hirschhorn JN, Hoffmann P, Hofman A, Hollegaard MV, Hougaard DM, Ikeda M, Joa I, Julia A, Kahn RS, Kalaydjieva L, Karachanak-Yankova S, Karjalainen J, Kavanagh D, Keller MC, Kennedy JL, Khrunin A, Kim Y, Klovins J, Knowles JA, Konte B, Kucinskas V, Kucinskiene ZA, Kuzelova-Ptackova H, Kahler AK, Laurent C, Keong JLC, Lee SH, Legge SE, Lerer B, Li M, Li T, Liang KY, Lieberman J, Limborska S, Loughland CM, Lubinski J, Lonnqvist J, Macek M, Magnusson PKE, Maher BS, Maier W, Mallet J, Marsal S, Mattheisen M, Mattingsdal M, McCarley RW, McDonald C, McIntosh AM, Meier S, Meijer CJ, Melegh B, Melle I, Mesholam-Gately RI, Metspalu A, Michie PT, Milani L, Milanova V, Mokrab Y, Morris DW, Mors O, Murphy KC, Murray RM, Myin-Germeys I, Muller-Myhsok B, Nelis M, Nenadic I, Nertney DA, Nestadt G, Nicodemus KK, Nikitina-Zake L, Nisenbaum L, Nordin A, O'Callaghan E, O'Dushlaine C, O'Neill FA, Oh SY, Olincy A, Olsen L, Van Os J, Pantelis C, Papadimitriou GN, Papiol S, Parkhomenko E, Pato MT, Paunio T, Pejovic-Milovancevic M, Perkins DO, Pietiläinen O, Pimm J, Pocklington AJ, Powell J, Price A, Pulver AE, Purcell SM, Quested D, Rasmussen HB, Reichenberg A, Reimers MA, Richards AL, Roffman JL, Salomaa V, Sanders AR, Schall U, Schubert CR, Schulze TG, Schwab SG, Scolnick EM, Scott RJ, Seidman LJ, Shi J, Sigurdsson E, Silagadze T, Silverman JM, Sim K, Slominsky P, Smoller JW, So HC, Spencer CCA, Stahl EA, Stefansson H, Steinberg S, Stogmann E, Straub RE, Strengman E, Strohmaier J, Stroup TS, Subramaniam M, Suvisaari J, Svrakic DM, Szatkiewicz JP, Soderman E, Thirumalai S, Toncheva D, Tosato S, Veijola J, Waddington J, Walsh D, Wang D, Wang Q, Webb BT, Weiser M, Wildenauer DB, Williams NM, Williams S, Witt SH,

Wolen AR, Wong EHM, Wormley BK, Xi HS, Zai CC, Zheng X, Zimprich F, Wray NR, Stefansson K, Visscher PM, Adolfsson R, Andreassen OA, Blackwood DHR, Bramon E, Børglum AD, Cichon S, Darvasi A, Ehrenreich H, Esko T, Gejman PV, Gill M, Gurling H, Hultman CM, Iwata N, Jablensky AV, Jonsson EG, Kendler KS, Kirov G, Knight J, Lencz T, Levinson DF, Li QS, Liu J, Malhotra AK, McCarroll SA, McQuillin A, Moran JL, Mortensen PB, Mowry BJ, Nothen MM, Ophoff RA, Owen MJ, Palotie A, Pato CN, Petryshen TL, Posthuma D, Rietschel M, Riley BP, Rujescu D, Sham PC, Sklar P, Clair DS, Weinberger DR, Wendland JR, Werge T, Daly MJ, Sullivan PF, O'Donovan MC, Rajagopal VM, Grove J, Mortensen PB, Pedersen CB, Pedersen MG, Nordentoft M, Bybjerg-Grauholm J, Bækvad-Hansen M, Hansen CS, Sullivan P, Devlin B, Sieberts SK, Cox NJ, Im HK (2019) Gene expression imputation across multiple brain regions provides insights into schizophrenia risk. Nat Genet 51:659–674. https://doi.org/10.1038/s41588-019-0364-4

29. Investigators MIG and CardiEC (2016) Coding variation in ANGPTL4, LPL, and SVEP1 and the risk of coronary disease. N Engl J Med 374:1134. https://doi.org/10.1056/NEJMOA1507652
30. Kanehisa M (2000) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30. https://doi.org/10.1093/nar/28.1.27
31. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, Aguilar Salinas CA, Ahmad T, Albert CM, Ardissino D, Atzmon G, Barnard J, Beaugerie L, Benjamin EJ, Boehnke M, Bonnycastle LL, Bottinger EP, Bowden DW, Bown MJ, Chambers JC, Chan JC, Chasman D, Cho J, Chung MK, Cohen B, Correa A, Dabelea D, Daly MJ, Darbar D, Duggirala R, Dupuis J, Ellinor PT, Elosua R, Erdmann J, Esko T, Färkkilä M, Florez J, Franke A, Getz G, Glaser B, Glatt SJ, Goldstein D, Gonzalez C, Groop L, Haiman C, Hanis C, Harms M, Hiltunen M, Holi MM, Hultman CM, Kallela M, Kaprio J, Kathiresan S, Kim BJ, Kim YJ, Kirov G, Kooner J, Koskinen S, Krumholz HM, Kugathasan S, Kwak SH, Laakso M, Lehtimäki T, Loos RJF, Lubitz SA, Ma RCW, MacArthur DG, Marrugat J, Mattila KM, McCarroll S, McCarthy MI, McGovern D, McPherson R, Meigs JB, Melander O, Metspalu A, Neale BM, Nilsson PM, O'Donovan MC, Ongur D, Orozco L, Owen MJ, Palmer CNA, Palotie A, Park KS, Pato C, Pulver AE, Rahman N, Remes AM, Rioux JD, Ripatti S, Roden DM, Saleheen D, Salomaa V, Samani NJ, Scharf J, Schunkert H, Shoemaker MB, Sklar P, Soininen H, Sokol H, Spector T, Sullivan PF, Suvisaari J, Tai ES, Teo YY, Tiinamaija T, Tsuang M, Turner D, Tusie-Luna T, Vartiainen E, Ware JS, Watkins H, Weersma RK, Wessman M, Wilson JG, Xavier RJ, Neale BM, Daly MJ, MacArthur DG (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581:434–443. https://doi.org/10.1038/s41586-020-2308-7
32. Kessler T, Zhang L, Liu Z, Yin X, Huang Y, Wang Y, Fu Y, Mayr M, Ge Q, Xu Q, Zhu Y, Wang X, Schmidt K, De Wit C, Erdmann J, Schunkert H, Aherrahrou Z, Kong W (2015) ADAMTS-7 inhibits re-endothelialization of injured arteries and promotes vascular remodeling through cleavage of thrombospondin-1. Circulation 131:1191–1201. https://doi.org/10.1161/CIRCULATIONAHA.114.014072
33. Klimentidis YC, Arora A, Newell M, Zhou J, Ordovas JM, Renquist BJ, Wood AC (2020) Phenotypic and genetic characterization of lower LDL cholesterol and increased type 2 diabetes risk in the UK biobank. Diabetes 69:2194–2205. https://doi.org/10.2337/db19-1134
34. Kohl M, Wiese S, Warscheid B (2011) Cytoscape: software for visualization and analysis of biological networks. Methods Mol Biol 696:291–303. https://doi.org/10.1007/978-1-60761-987-1_18
35. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, Matsunaga H, Ieki H, Ozaki K, Onouchi Y, Takahashi A, Nomura S, Morita H, Akazawa H, Kim C, Seo J, Higasa K, Iwasaki M, Yamaji T, Sawada N, Tsugane S, Koyama T, Ikezaki H, Takashima N, Tanaka K, Arisawa K, Kuriki K, Naito M, Wakai K, Suna S, Sakata Y, Sato H, Hori M, Sakata Y, Matsuda K, Murakami Y, Aburatani H, Kubo M, Matsuda F, Kamatani Y, Komuro I (2020) Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. Nat Genet 52:1169–1177. https://doi.org/10.1038/s41588-020-0705-3
36. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, Hoffman D, Jang W, Kaur K, Liu C, Lyoshin V, Maddipatla Z, Maiti R, Mitchell J, O'leary N, Riley GR, Shi W, Zhou G, Schneider V, Maglott D, Holmes JB, Kattman BL (2019) ClinVar: improvements to accessing data. Nucleic Acids Res 48:835–844.https://doi.org/10.1093/nar/gkz972
37. De Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015) MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 11:1004219. https://doi.org/10.1371/journal.pcbi.1004219
38. Li L, Pang S, Zeng L, Güldener U, Schunkert H (2021) Genetically determined intelligence and coronary artery disease risk. Clin Res Cardiol 110:211–219. https://doi.org/10.1007/s00392-020-01721-x
39. Li MX, Yeung JMY, Cherny SS, Sham PC (2012) Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. Hum Genet 131:747–756. https://doi.org/10.1007/S00439-011-1118-2
40. Li YI, Van De Geijn B, Raj A, Knowles DA, Petti AA, Golan D, Gilad Y, Pritchard JK (2016) RNA splicing is a primary link between genetic variation and disease. Science (80-) 352:600–604. https://doi.org/10.1126/science.aad9417
41. Liu DJ, Peloso GM, Yu H, Butterworth AS, Wang X, Mahajan A, Saleheen D, Emdin C, Alam D, Alves AC, Amouyel P, Angelantonio E DI, Arveiler D, Assimes TL, Auer PL, Baber U, Ballantyne CM, Bang LE, Benn M, Bis JC, Boehnke M, Boerwinkle E, Bork-Jensen J, Bottinger EP, Brandslund I, Brown M, Busonero F, Caulfield MJ, Chambers JC, Chasman DI, Chen YE, Chen YDI, Chowdhury R, Christensen C, Chu AY, Connell JM, Cucca F, Cupples LA, Damrauer SM, Davies G, Deary IJ, Dedoussis G, Denny JC, Dominiczak A, Dubé MP, Ebeling T, Eiriksdottir G, Esko T, Farmaki AE, Feitosa MF, Ferrario M, Ferrieres J, Ford I, Fornage M, Franks PW, Frayling TM, Frikke-Schmidt R, Fritsche LG, Frossard P, Fuster V, Ganesh SK, Gao W, Garcia ME, Gieger C, Giulianini F, Goodarzi MO, Grallert H, Grarup N, Groop L, Grove ML, Gudnason V, Hansen T, Harris TB, Hayward C, Hirschhorn JN, Holmen OL, Huffman J, Huo Y, Hveem K, Jabeen S, Jackson AU, Jakobsdottir J, Jarvelin MR, Jensen GB, Jørgensen ME, Jukema JW, Justesen JM, Kamstrup PR, Kanoni S, Karpe F, Kee F, Khera AV, Klarin D, Koistinen HA, Kooner JS, Kooperberg C, Kuulasmaa K, Kuusisto J, Laakso M, Lakka T, Langenberg C, Langsted A, Launer LJ, Lauritzen T, Mliewald DC,

Lin LA, Linneberg A, Loos RJF, Lu Y, Lu X, Mägi R, Malarstig A, Manichaikul A, Manning AK, Mäntyselkä P, Marouli E, Masca NGD, Maschio A, Meigs JB, Melander O, Metspalu A, Morris AP, Morrison AC, Mulas A, Müller-Nurasyid M, Munroe PB, Neville MJ, Nielsen SF, Nielsen JB, Nordestgaard BG, Ordovas JM, Mehran R, O'Donnell CJ, Orho-Melander M, Molony CM, Muntendam P, Padmanabhan S, Palmer CNA, Pasko D, Patel AP, Pedersen O, Perola M, Peters A, Pisinger C, Pistis G, Polasek O, Poulter N, Psaty BM, Rader DJ, Rasheed A, Rauramaa R, Reilly DF, Reiner AP, Renström F, Rich SS, Ridker PM, Rioux JD, Robertson NR, Roden DM, Rotter JI, Rudan I, Salomaa V, Samani NJ, Sanna S, Sattar N, Schmidt EM, Scott RA, Sever P, Sevilla RS, Shaffer CM, Sim X, Sivapalaratnam S, Small KS, Smith AV, Smith BH, Somayajula S, Southam L, Spector TD, Speliotes EK, Starr JM, Stirrups KE, Stitziel N, Strauch K, Stringham HM, Surendran P, Tada H, Tall AR, Tang H, Tardif JC, Taylor KD, Trompet S, Tsao PS, Tuomilehto J, Tybjaerg-Hansen A, Zuydam NRV, Varbo A, Varga TV, Virtamo J, Waldenberger M, Wang N, Wareham NJ, Warren HR, Weeke PE, Weinstock J, Wessel J, Wilson JG, Wilson PWF, Xu M, Yaghootkar H, Young R, Zeggini E, Zhang H, Zheng NS, Zhang W, Zhang Y, Zhou W, Zhou Y, Zoledziewska M, Howson JMM, Danesh J, McCarthy MI, Cowan CA, Abecasis G, Deloukas P, Musunuru K, Willer CJ, Kathiresan S (2017) Exome-wide association study of plasma lipids in > 300,000 individuals. Nat Genet 49:1758–1766.https://doi.org/10.1038/NG.3977

42. Lusis AJ, Seldin MM, Allayee H, Bennett BJ, Civelek M, Davis RC, Eskin E, Farber CR, Hui S, Mehrabian M, Norheim F, Pan C, Parks B, Rau CD, Smith DJ, Vallim T, Wang Y, Wang J (2016) The Hybrid Mouse Diversity Panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. J Lipid Res 57:925–942. https://doi.org/10.1194/jlr.R066944
43. Malakar AK, Choudhury D, Halder B, Paul P, Uddin A, Chakraborty S (2019) A review on coronary artery disease, its risk factors, and therapeutics. J Cell Physiol 234:16812–16823. https://doi.org/10.1002/jcp.28350
44. McGough IJ, Steinberg F, Gallon M, Yatsu A, Ohbayashi N, Heesom KJ, Fukuda M, Cullen PJ (2014) Identification of molecular heterogeneity in SNX27-retromermediated endosome-to-plasma-membrane recycling. J Cell Sci 127:4940–4953. https://doi.org/10.1242/jcs.156299
45. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F (2016) The ensembl variant effect predictor. Genome Biol. https://doi.org/10.1186/s13059-016-0974-4
46. Meng Y, Heybrock S, Neculai D, Saftig P (2020) Cholesterol handling in lysosomes and beyond. Trends Cell Biol 30:452–466. https://doi.org/10.1016/J.TCB.2020.02.007
47. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, Zeng L, Ntalla I, Lai FY, Hopewell JC, Giannakopoulou O, Jiang T, Hamby SE, Di Angelantonio E, Assimes TL, Bottinger EP, Chambers JC, Clarke R, Palmer CNA, Cubbon RM, Ellinor P, Ermel R, Evangelou E, Franks PW, Grace C, Gu D, Hingorani AD, Howson JMM, Ingelsson E, Kastrati A, Kessler T, Kyriakou T, Lehtimäki T, Lu X, Lu Y, März W, McPherson R, Metspalu A, Pujades-Rodriguez M, Ruusalepp A, Schadt EE, Schmidt AF, Sweeting MJ, Zalloua PA, Alghalayini K, Keavney BD, Kooner JS, Loos RJF, Patel RS, Rutter MK, Tomaszewski M, Tzoulaki I, Zeggini E, Erdmann J, Dedoussis G, Björkegren JLM, Schunkert H, Farrall M, Danesh J, Samani NJ, Watkins H, Deloukas P (2017) Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet 49:1385–1391. https://doi.org/10.1038/ng.3913
48. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, CHopewell J, Webb TR, Zeng L, Dehghan A, Alver M, Marmasu S, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikäinen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, De Andrade M, De Vries PS, Van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, König IR, Lannfelt L, Lieb W, Lind L, MLindgren C, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FUR, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, EStirrups K, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples L, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, SHall A, Hamsten A, Harris TB, LHazen S, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimäki T, Loos RJF, Melander O, Metspalu A, März W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M (2015) A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 47:1121–1130.https://doi.org/10.1038/ng.3396
49. Stewart A, Fisher RA (2015) Introduction: G protein-coupled receptors and RGS proteins. Prog Mol Biol Transl Sci 133:1–11. https://doi.org/10.1016/BS.PMBTS.2015.03.002
50. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, Furlong LI (2020) The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Res 48:D845–D855. https://doi.org/10.1093/nar/gkz1021
51. Powell WS, Rokach J (2015) Biosynthesis, biological effects, and receptors of hydroxyeicosatetraenoic acids (HETEs) and oxoeicosatetraenoic acids (oxo-ETEs) derived from arachidonic acid. Biochim Biophys Acta 1851:340–355. https://doi.org/10.1016/J.BBALIP.2014.10.008
52. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann H-E, Barrett JH, König IR, Stevens SE, Szymczak S, Tregouet D-A, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Brænne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H (2007) Genomewide association analysis of coronary artery disease. N Engl J Med 357:443–453. https://doi.org/10.1056/NEJMoa072366
53. Sangphech N, Osborne BA, Palaga T (2014) Notch signaling regulates the phosphorylation of Akt and survival of lipopolysaccharide-activated macrophages via regulator of G protein signaling 19 (*RGS19*). Immunobiology 219:653–660. https://doi.org/10.1016/j.imbio.2014.03.020
54. Sixt S, Beer S, Blüher M, Korff N, Peschel T, Sonnabend M, Teupser D, Thiery J, Adams V, Schuler G, Niebauer J (2010) Long- but not short-term multifactorial intervention with focus on exercise training improves coronary endothelial dysfunction in diabetes mellitus type 2 and coronary artery disease. Eur Heart J 31:112–119. https://doi.org/10.1093/eurheartj/ehp398
55. Slenter DN, Kutmon M, Hanspers K, Riutta A, Windsor J, Nunes N, Mélius J, Cirillo E, Coort SL, Digles D, Ehrhart F, Giesbertz P, Kalafati M, Martens M, Miller R, Nishida K, Rieswijk L,

Waagmeester A, Eijssen LMT, Evelo CT, Pico AR, Willighagen EL (2018) WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Res 46:D661–D667. https://doi.org/10.1093/nar/gkx1064

56. Stitziel NO, Won HH, Morrison AC, Peloso GM, Do R, Lange LA, Fontanillas P, Gupta N, Duga S, Goel A, Farrall M, Saleheen D, Ferrario P, König I, Asselta R, Merlini PA, Marziliano N, Notarangelo MF, Schick U, Auer P, Assimes TL, Reilly M, Wilensky R, Rader DJ, Kees Hovingh G, Meitinger T, Kessler T, Kastrati A, Laugwitz KL, Siscovick D, Rotter JI, Hazen SL, Tracy R, Cresci S, Spertus J, Jackson R, Schwartz SM, Natarajan P, Crosby J, Muzny D, Ballantyne C, Rich SS, O'Donnell CJ, Abecasis G, Sunyaev S, Nickerson DA, Buring JE, Ridker PM, Chasman DI, Austin E, Ye Z, Kullo IJ, Weeke PE, Shaffer CM, Bastarache LA, Denny JC, Roden DM, Palmer C, Deloukas P, Lin DY, Tang ZZ, Erdmann J, Schunkert H, Danesh J, Marrugat J, Elosua R, Ardissino D, McPherson R, Watkins H, Reiner AP, Wilson JG, Altshuler D, Gibbs RA, Lander ES, Boerwinkle E, Gabriel S, Kathiresan S (2014) Inactivating mutations in NPC1L1 and protection from coronary heart disease. N Engl J Med 371:2072–2082. https://doi.org/10.1056/NEJMoa1405386
57. Thelen AM, Zoncu R (2017) Emerging roles for the lysosome in lipid metabolism. Trends Cell Biol 27:833–850. https://doi.org/10.1016/J.TCB.2017.07.006
58. Tiwari S, Siddiqi SA (2012) Intracellular trafficking and secretion of VLDL. Arterioscler Thromb Vasc Biol 32:1079–1086. https://doi.org/10.1161/ATVBAHA.111.241471
59. Tso PH, Yung LY, Wang Y, Wong YH (2011) RGS19 stimulates cell proliferation by deregulating cell cycle control and enhancing Akt signaling. Cancer Lett 309:199–208. https://doi.org/10.1016/j.canlet.2011.06.002
60. Tsutsumi K (2003) Lipoprotein lipase and atherosclerosis. Curr Vasc Pharmacol 1:11–17. https://doi.org/10.2174/1570161033386673
61. Vilne B, Schunkert H (2018) Integrating genes affecting coronary artery disease in functional networks by multi-OMICs approach. Front Cardiovasc Med 5:89. https://doi.org/10.3389/fcvm.2018.00089
62. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K, Björkegren JLM, Im HK, Pasaniuc B, Rivas MA, Kundaje A (2019) Opportunities and challenges for transcriptome-wide association studies. Nat Genet 51:592–599. https://doi.org/10.1038/s41588-019-0385-z
63. Wang Y, Nakajima T, Gonzalez FJ, Tanaka N (2020) PPARs as metabolic regulators in the liver: lessons from liver-specific PPAR-null mice. Int J Mol Sci. https://doi.org/10.3390/IJMS21062061
64. Wild PS, Zeller T, Schillert A, Szymczak S, Sinning CR, Deiseroth A, Schnabel RB, Lubos E, Keller T, Eleftheriadis MS, Bickel C, Rupprecht HJ, Wilde S, Rossmann H, Diemert P, Cupples LA, Perret C, Erdmann J, Stark K, Kleber ME, Epstein SE, Voight BF, Kuulasmaa K, Li M, Schäfer AS, Klopp N, Braund PS, Sager HB, Demissie S, Proust C, König IR, Wichmann H-E, Reinhard W, Hoffmann MM, Virtamo J, Burnett MS, Siscovick D, Wiklund PG, Qu L, El Mokthari NE, Thompson JR, Peters A, Smith AV, Yon E, Baumert J, Hengstenberg C, März W, Amouyel P, Devaney J, Schwartz SM, Saarela O, Mehta NN, Rubin D, Silander K, Hall AS, Ferrieres J, Harris TB, Melander O, Kee F, Hakonarson H, Schrezenmeir J, Gudnason V, Elosua R, Arveiler D, Evans A, Rader DJ, Illig T, Schreiber S, Bis JC, Altshuler D, Kavousi M, Witteman JCM, Uitterlinden AG, Hofman A, Folsom AR, Barbalic M, Boerwinkle E, Kathiresan S, Reilly MP, O'Donnell CJ, Samani NJ, Schunkert H, Cambien F, Lackner KJ, Tiret L, Salomaa V, Munzel T, Ziegler A, Blankenberg S (2011) A genome-wide association study identifies *LIPA* as a susceptibility gene for coronary artery disease. Circ Cardiovasc Genet 4:403–412. https://doi.org/10.1161/CIRCGENETICS.110.958728
65. Winkelmann BR, März W, Boehm BO, Zotz R, Hager J, Hellstern P, Senges J (2001) Rationale and design of the LURIC study—a resource for functional genomics, pharmacogenomics and long-term prognosis of cardiovascular disease. Pharmacogenomics. https://doi.org/10.1517/14622416.2.1.s1
66. Wolfrum C, Borrmann CM, Börchers T, Spener F (2001) Fatty acids and hypolipidemic drugs regulate peroxisome proliferator-activated receptors alpha - and gamma-mediated gene expression via liver fatty acid binding protein: a signaling path to the nucleus. Proc Natl Acad Sci USA 98:2323–2328. https://doi.org/10.1073/PNAS.051619898
67. Wolfson RL, Chantranupong L, Wyant GA, Gu X, Orozco JM, Shen K, Condon KJ, Petri S, Kedir J, Scaria SM, Abu-Remaileh M, Frankel WN, Sabatini DM (2017) KICSTOR recruits GATOR1 to the lysosome and is necessary for nutrients to regulate mTORC1. Nature 543:438–442. https://doi.org/10.1038/nature21423
68. Wu L, Shi W, Long J, Guo X, Michailidou K, Beesley J, Bolla MK, Shu XO, Lu Y, Cai Q, Al-Ejeh F, Rozali E, Wang Q, Dennis J, Li B, Zeng C, Feng H, Gusev A, Barfield RT, Andrulis IL, Anton-Culver H, Arndt V, Aronson KJ, Auer PL, Barrdahl M, Baynes C, Beckmann MW, Benitez J, Bermisheva M, Blomqvist C, Bogdanova NV, Bojesen SE, Brauch H, Brenner H, Brinton L, Broberg P, Brucker SY, Burwinkel B, Caldés T, Canzian F, Carter BD, Castelao JE, Chang-Claude J, Chen X, Cheng TYD, Christiansen H, Clarke CL, Collée M, Cornelissen S, Couch FJ, Cox D, Cox A, Cross SS, Cunningham JM, Czene K, Daly MB, Devilee P, Doheny KF, Dörk T, Dos-Santos-Silva I, Dumont M, Dwek M, Eccles DM, Eilber U, Eliassen AH, Engel C, Eriksson M, Fachal L, Fasching PA, Figueroa J, Flesch-Janys D, Fletcher O, Flyger H, Fritschi L, Gabrielson M, Gago-Dominguez M, Gapstur SM, García-Closas M, Gaudet MM, Ghoussaini M, Giles GG, Goldberg MS, Goldgar DE, González-Neira A, Guénel P, Hahnen E, Haiman CA, Håkansson N, Hall P, Hallberg E, Hamann U, Harrington P, Hein A, Hicks B, Hillemanns P, Hollestelle A, Hoover RN, Hopper JL, Huang G, Humphreys K, Hunter DJ, Jakubowska A, Janni W, John EM, Johnson N, Jones K, Jones ME, Jung A, Kaaks R, Kerin MJ, Khusnutdinova E, Kosma VM, Kristensen VN, Lambrechts D, Le Marchand L, Li J, Lindström S, Lissowska J, Lo WY, Loibl S, Lubinski J, Luccarini C, Lux MP, MacInnis RJ, Maishman T, Kostovska IM, Mannermaa A, Manson JAE, Margolin S, Mavroudis D, Meijers-Heijboer H, Meindl A, Menon U, Meyer J, Mulligan AM, Neuhausen SL, Nevanlinna H, Neven P, Nielsen SF, Nordestgaard BG, Olopade OI, Olson JE, Olsson H, Peterlongo P, Peto J, Plaseska-Karanfilska D, Prentice R, Presneau N, Pylkäs K, Rack B, Radice P, Rahman N, Rennert G, Rennert HS, Rhenius V, Romero A, Romm J, Rudolph A, Saloustros E, Sandler DP, Sawyer EJ, Schmidt MK, Schmutzler RK, Schneeweiss A, Scott RJ, Scott CG, Seal S, Shah M, Shrubsole MJ, Smeets A, Southey MC, Spinelli JJ, Stone J, Surowy H, Swerdlow AJ, Tamimi RM, Tapper W, Taylor JA, Terry MB, Tessier DC, Thomas A, Thöne K, Tollenaar RAEM, Torres D, Truong T, Untch M, Vachon C, Van Den Berg D, Vincent D, Waisfisz Q, Weinberg CR, Wendt C, Whittemore AS, Wildiers H, Willett WC, Winqvist R, Wolk A, Xia L, Yang XR, Ziogas A, Ziv E, Dunning AM, Pharoah PDP, Simard J, Milne RL, Edwards SL, Kraft P, Easton DF, Chenevix-Trench G, Zheng W (2018) A transcriptome-wide association study of 229,000 women identifies new candidate susceptibility genes for breast cancer. Nat Genet 50:968–978. https://doi.org/10.1038/s41588-018-0132-x

69. Zein AA, Kaur R, Hussein TOK, Graf GA, Lee J-Y (2019) ABCG5/G8: a structural view to pathophysiology of the hepato-biliary cholesterol secretion. Biochem Soc Trans 47:1259. https://doi.org/10.1042/BST20190130
70. Zhang W, Voloudakis G, Rajagopal VM, Readhead B, Dudley JT, Schadt EE, Björkegren JLM, Kim Y, Fullard JF, Hoffman GE, Roussos P (2019) Integrative transcriptome imputation reveals tissue-specific and shared biological mechanisms mediating susceptibility to complex traits. Nat Commun 10:1–13. https://doi.org/10.1038/s41467-019-11874-7

Authors and Affiliations

Ling Li[1,2,3] · Zhifen Chen[1,3] · Moritz von Scheidt[1,3] · Shuangyue Li[1,3] · Andrea Steiner[1,3] · Ulrich Güldener[1,3] · Simon Koplev[4] · Angela Ma[4] · Ke Hao[4] · Calvin Pan[5] · Aldons J. Lusis[5,6,7] · Shichao Pang[1,3] · Thorsten Kessler[1,3,7] · Raili Ermel[8] · Katyayani Sukhavasi[8] · Arno Ruusalepp[8,9] · Julien Gagneur[2] · Jeanette Erdmann[10,11] · Jason C. Kovacic[12,13,14] · Johan L. M. Björkegren[4,9,15] · Heribert Schunkert[1,3]

1 Department of Cardiology, German Heart Center Munich, Technical University Munich, Lazarettstraße 36, 80636 Munich, Germany

2 Fakultät für Informatik, Technische Universität München, Munich, Germany

3 Deutsches Zentrum für Herz- und Kreislaufforschung (DZHK), Partner Site Munich Heart Alliance, Munich, Germany

4 Department of Genetics and Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029-6574, USA

5 Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

6 Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

7 Department of Microbiology, Immunology and Molecular Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

8 Department of Cardiac Surgery, The Heart Clinic, Tartu University Hospital, Tartu, Estonia

9 Clinical Gene Networks AB, Stockholm, Sweden

10 DZHK (German Research Centre for Cardiovascular Research), Partner Site Hamburg/Lübeck/Kiel, Lübeck, Germany

11 Institute for Cardiogenetics, University of Lübeck, Lübeck, Germany

12 Victor Chang Cardiac Research Institute, Darlinghurst, Australia

13 St Vincent's Clinical School, University of New South Wales, Sydney, Australia

14 Icahn School of Medicine at Mount Sinai, Cardiovascular Research Institute, New York, NY 10029-6574, USA

15 Department of Medicine, Huddinge, Karolinska Institutet, Karolinska Universitetssjukhuset, Stockholm, Sweden