Genome-wide association study on coronary artery disease in type 1 diabetes suggests beta-defensin 127 as a risk locus

Antikainen, Anni A. V.

2021-02-01

Antikainen , A A V , Sandholm , N , Tregouet , D-A , Charmet , R , McKnight , A J , Ahluwalia , T S , Syreeni , A , Valo , E , Forsblom , C , Gordin , D , Harjutsalo , V , Hadjadj , S , Maxwell , A P , Rossing , P & Groop , P-H 2021 , ' Genome-wide association study on coronary artery disease in type 1 diabetes suggests beta-defensin 127 as a risk locus ' , Cardiovascular Research , vol. 117 , no. 2 , pp. 600-612 . https://doi.org/10.1093/cvr/cvaa045

http://hdl.handle.net/10138/339508
https://doi.org/10.1093/cvr/cvaa045

acceptedVersion

Genome-wide association study on coronary artery disease in type 1 diabetes suggests beta-defensin 127 as a risk locus

Short title: *Coronary Artery Disease GWAS in type 1 diabetes*

Anni A. V. Antikainen (MSc in Tech)*[1,2,3], Niina Sandholm (DSc in Tech)*[1,2,3], David-Alexandre Trëgouet (PhD)[4,5,6], Romain Charmet (PhD)[4,5], Amy Jayne McKnight (PhD)[7], Tarunveer S. Ahluwalia (PhD)[8], Anna Syreeni (MSc)[1,2,3], Erkka Valo (MSc in Tech)[1,2,3], Carol Forsblom (DMSc)[1,2,3], Daniel Gordin (MD, DMSc)[1,2,3,9], Valma Harjutsalo (PhD)[1,2,3,10], Samy Hadjadj (MD, PhD)[11,12], Alexander P. Maxwell (MD, PhD)[7], Peter Rossing (MD, DMSc)[8,13], Per-Henrik Groop (MD, DMSc)[1,2,3,14]

**authors contributed equally*

Underlined affiliation indicate department, where the work was performed

1. Folkhälsan Institute of Genetics, Folkhälsan Research Center, FI-00290, Helsinki, Finland
2. Abdominal Center, Nephrology, University of Helsinki and Helsinki University Hospital, FI-00290, Helsinki, Finland
3. Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, FI-00290, Helsinki, Finland
4. Sorbonne Université, UPMC Univ Paris 06, INSERM UMR_S 1166, Paris, France.
5. ICAN Institute for Cardiometabolism and Nutrition, Paris, France.
6. INSERM UMR_S 1219, Bordeaux Population Health Research Center, Bordeaux University, Bordeaux, France

7. Centre for Public Health, Queens University of Belfast, Northern Ireland, BT7 1NN UK.
8. Steno Diabetes Center Copenhagen, DK 2820 Gentofte, Denmark.
9. Joslin Diabetes Center, Harvard Medical School, Boston MA
10. The Chronic Disease Prevention Unit, National Institute for Health and Welfare, FI-00271, Helsinki, Finland
11. Centre Hospitalier Universitaire de Poitiers, Department of Endocrinology and Diabetology, Poitiers, France; INSERM CIC 1402, Poitiers, France.
12. L'institut du thorax, INSERM, CNRS, UNIV Nantes CHU Nantes, Nantes, France.
13. University of Copenhagen, Copenhagen, Denmark.
14. Department of Diabetes, Central Clinical School, Monash University, Melbourne, Victoria, Australia

All authors contributed substantially to the manuscript and agreed on the content

Author contributions listed in Chapter 6, following the Sources of Funding section as requested by the Editorial Office

Original Article

The manuscript is not submitted elsewhere or under consideration for publication

Corresponding Author:

Per-Henrik Groop, MD DMSc FRCP Edin

University of Helsinki and Helsinki University Hospital

Biomedicum Helsinki

Haartmaninkatu 8, 00290 Helsinki

Finland

Fax: +358 9191 25382

Tel: +358-500-430 436

Email: per-henrik.groop@helsinki.fi

Word count: 8940

ABSTRACT

Aims

Diabetes is a known risk factor for coronary artery disease. There is accumulating evidence that coronary artery disease pathogenesis differs for individuals with type 1 diabetes. However, the genetic background has not been extensively studied. We aimed to discover genetic loci increasing coronary artery disease susceptibility especially in type 1 diabetes, to examine the function of these discoveries and to study the role of the known risk loci in type 1 diabetes.

Methods and Results

We performed the largest genome-wide association study to date for coronary artery disease in type 1 diabetes, comprising 4869 individuals with type 1 diabetes (cases/controls: 941/3928). Two loci reached genome-wide significance, rs1970112 in *CDKN2B-AS1* (OR=1.32, $p=1.50\times10^{-8}$), and rs6055069 on *DEFB127* promoter (OR=4.17, $p=2.35\times10^{-9}$), with consistent results in survival analysis. The *CDKN2B-AS1* variant replicated ($p=0.04$) when adjusted for diabetic kidney disease in three additional type 1 diabetes cohorts (cases/controls: 434/3123). Furthermore, we explored the function of the lead discoveries with a cardio-phenome-wide analysis. Among the eight suggestive loci ($p<1\times10^{-6}$), rs70962766 near *B3GNT2* associated with central blood pressure, rs1344228 near *CNTNAP5* with intima media thickness, and rs2112481 on *GRAMD2B* promoter with serum leucocyte concentration. Finally, we calculated genetic risk scores for individuals with type 1 diabetes with the known susceptibility loci. General population risk variants were modestly but significantly associated with coronary artery disease also in type 1 diabetes ($p=4.21\times10^{-7}$).

Conclusions

While general population coronary artery disease risk loci had limited effect on the risk in type 1 diabetes, for the first time, variants at the *CDKN2B-AS1* locus were robustly associated with coronary artery disease in individuals with type 1 diabetes. The novel finding on β-defensin

DEFB127 promoter provides a link between diabetes, infection susceptibility and coronary artery disease, although pending on future confirmation.

Translational perspective

Genetic association studies enable the discovery of novel genes and genetic pathways associated with the disease. Thus, this study provides an insight into coronary artery disease mechanisms specific to type 1 diabetes. The *DEFB127* discovery may lead to a therapeutic target and improve patient care, if replicated in the future. Furthermore, genetic studies on coronary artery disease in type 1 diabetes are required for accurate personalized treatment plans achieved through genetic data for those with type 1 diabetes.

1. Introduction

A total of 425 million people worldwide have diabetes[1]. One third of them develop diabetic complications including diabetic kidney disease (DKD), blindness, amputations, and cardiovascular disease (CVD), further divided into coronary artery disease (CAD), strokes, and other macrovascular complications. Cardiovascular deaths are the most common cause of mortality in individuals with diabetes[2], and CVD occurs particularly early in life in individuals with type 1 diabetes (T1D). The risk of CVD increases steeply with the severity of DKD[3], and as many as 40% of individuals with T1D and DKD develop CVD by the age of 40 years[4]; nevertheless, even individuals with T1D but without DKD have four times higher age standardized incidence of CAD compared with non-diabetic individuals[3]. Altogether, there is increasing evidence that the incidence, risk factors, and even pathophysiology of CVD and CAD differ in individuals with T1D from those without diabetes, or even from those with type 2 diabetes (T2D)[5].

Classical risk factors for CAD in diabetes include unfavorable glycemic control, high blood pressure, and dyslipidemia[6]. With the improvements in the treatment of risk factors, the incidence of CAD has drastically decreased over the decades[7]. In addition, family history of CAD is an important risk factor both in the general population[8], and among individuals with insulin treated diabetes[9]. Consequently, genome-wide association studies (GWAS) have revealed 163 distinct loci for CAD in the general population[10]. While replication of these loci has been attempted also in individuals with diabetes, most of the results have been inconclusive[11-13]. A previous study suggested that CAD risk associated with the 9p21 locus is magnified in individuals with T2D in the presence of poor glycemic control[14]. Furthermore, candidate gene studies on CAD in diabetes have suggested e.g. haptoglobin 2-2 genotype associated with CVD specifically in diabetes, including T1D[5]. The first genome-wide association study on CAD in individuals with T2D identified variants near the glutamine

synthase *(GLUL)* gene[15], and another GWAS in the ACCORD trial revealed two loci that affected CVD mortality in the intensively treated participants with T2D[16]. A recent GWAS on CAD in 15,666 individuals with diabetes from the UK Biobank found that two of the previously identified CAD loci, the *LPA* and 9p21 regions, were genome-wide significantly associated with CAD also in individuals with diabetes, with the effect sizes similar to the general population[17]. The first GWAS on CAD in individuals with T1D, including 434 CAD cases and 3123 non-CAD patients, suggested novel susceptibility variants for CAD at *CDK18, FAM189A2* and *PKD1* loci, even though these did not reach genome-wide significance in the final analysis[13].

Due to the central role of diabetes and hyperglycemia for the risk of CAD, an intriguing question is, whether novel genetic risk factors can be discovered in individuals with T1D, who tend to have suboptimal glycemic control, who suffer CAD early in life, and who are known to display a differing CAD pathogenesis from individuals without diabetes or even with T2D[5]. In this work, we present the results from the largest GWAS on CAD in T1D to date.

2. Methods

2.1 Patients

2.1.1 FinnDiane study

This study included 5342 Finnish individuals with T1D from the Finnish Diabetic Nephropathy (FinnDiane) study, an ongoing nationwide multicenter study with the aim to identify factors leading to diabetic complications. The participants were diagnosed with T1D with onset age ≤40 years and insulin treatment initiated within two years from diagnosis, if known. The individuals underwent a thorough clinical examination by the attending physician, including blood samples and timed urine collections, at their FinnDiane baseline visit; a subset of patients also participated in similar follow-up visits. Further health related data were obtained from the

patients' medical records. A subset of 740 individuals was included in FinnDiane through collaboration with the Finnish National Institute of Health and Welfare and had only medical records and registry data available.

The study protocol was approved by the ethics committee of the Helsinki and Uusimaa Hospital District (491/E5/2006, 238/13/03/00/2015, and HUS-3313-2018), and the study was performed in accordance with the Declaration of Helsinki. All participants gave informed consent before participation in the study. Summary statistics from the discovery cohort GWAS will be submitted to GRASP data base.

CAD was defined as a hard CAD event by the end of 2015 based on the Finnish Death Registry and Finnish hospital discharge registry using ICD codes I21, I22, and I23 for myocardial infarctions; or procedure codes for coronary bypass surgery or coronary balloon angioplasty (**Supplemental Table 1)**. Controls were individuals without hard CAD events. Only 5% of cases had CAD event before age of 35 years, and thus, controls with age <35 years (N=322) or diabetes duration <15 years (N=151) were excluded from the case-control CAD analysis, including 4869 individuals (941 cases, 3928 controls). Among these, 2590 had normoalbuminuria, while 2113 had DKD, defined as microalbuminuria (N=719), macroalbuminuria (N=544), or end stage renal disease (i.e. dialysis, kidney transplant, or estimated glomerular filtration rate <15 mL/min/1.73m^2; N=850) at the latest available data point.

2.1.2 Replication cohorts

Replication of the lead findings was performed in three studies comprising 3557 T1D individuals with (N_{cases}=434) and without CAD ($N_{controls}$=3123), from France, Denmark, and United Kingdom/Republic of Ireland (UK-ROI), included in a recently published CAD GWAS meta-analysis among T1D individuals[13].

2.2 GWAS Genotyping and data processing

The genome-wide genotyping and imputation procedures have been described earlier[18]. In short, 6152 unique Finnish individuals with diabetes and their relatives were genotyped at the University of Virginia with HumanCoreExome Bead arrays 12-1.0, 12-1.1, and 24-1.0 (Illumina, San Diego, CA, USA). Genotype calling with zCall[19], sample and variant quality control, genotype imputation with minimac3-software[20] and the 1000 Genomes phase 3 reference panel were also performed at the University of Virginia. After quality control 6019 individuals remained; 575 of these did not have T1D and/or data on CAD, resulting in 5342 T1D individuals with CAD and GWAS data available.

2.3 Statistical methods

2.3.1 GWAS analysis

Association analysis was performed with a score test based on estimated allele dosages using Rvtest (version 20160404)[21]. Calendar year of diabetes onset correlates strongly with the incidence of CAD most likely due to the improvements in diabetes management and glycemic control. We have adjusted the analysis for the calendar year of diabetes onset **(Supplemental Figure 1)**, in addition to gender and genotyping batch; and for kinship matrix to account for relatedness and population substructure. The top discoveries were reanalyzed with DKD as an additional adjustment covariate. Results were filtered to those in Hardy Weinberg equilibrium (HWE; $p \geq 5\times10^{-8}$) with minor allele frequency (MAF) ≥0.01 and imputation quality $r^2 \geq 0.6$. Suggestive lead single nucleotide polymorphisms (SNP) were defined as independent SNPs (distance ≥100k base pairs [bp]) with p-value $<10^{-6}$.

2.3.2 Fine mapping of the loci

Fine mapping of causal variants at the lead loci was performed with FINEMAP v1.3[22] based on a stochastic search of most important causal configurations in the GWAS summary

statistics. We used the default parameters and included 1Mbp region up and downstream of each lead variant. Linkage disequilibrium (LD) structure was calculated based on the FinnDiane imputed maximum likelihood genotypic data.

2.3.3 Replication and meta-analysis

For the loci reaching *p*-value 10^{-6}, we looked for *in silico* replication in a recent GWAS for CAD in T1D patients[13]. Variants with MAF≥0.01 and imputation quality r^2≥0.6 were included in the analysis. Meta-analysis was performed for both DKD adjusted and unadjusted models with the metal software (2011-03-25 release)[23] based on effect sizes and standard errors.

2.3.4 Power calculations

Power calculations were conducted with genetic power calculator[24] using relative risks calculated from odds ratios (OR) and assuming a 19.3% disease prevalence as observed in the FinnDiane study. Power to detect an association with $\alpha=5\times10^{-8}$ significance level in the discovery cohort is approximately 80% for variants with ORs of 1.40, 1.63 and 1.90 and minor allele frequencies of 0.50, 0.10 and 0.05 respectively. Furthermore, a variant with minor allele frequency of 0.01 and OR of 3.32 can be discovered with a roughly 75% power. Replication power of the lead findings was evaluated with a 0.01 (0.05/5) significance level by assuming disease prevalence as perceived in the replication cohorts and utilizing discovery cohort lower bound ORs and effect allele frequencies (**Table 1**). Finally, power to replicate CAD loci previously discovered in T1D was evaluated with a 0.003 (0.05/16) significance level (**Supplemental Table 2**), in T2D with a 0.01 (0.05/5) significance level, and in the general population with a 0.00031 (0.05/160) significance level by considering published effect allele frequencies and ORs, more precisely lower ORs when available (**Supplemental Figure 2**).

2.3.5 Survival analysis

The lead genotype ($p<10^{-6}$) effects were modeled with Cox proportional hazard (PH) models with time varying from diabetes onset until the latest follow-up data or a CAD incident. Implemented models incorporated different adjustment covariates: gender, diabetes onset year, age at onset of diabetes and time weighted mean HbA1c level, based on a median of 21 HbA1c measurements (maximum 142) available for 5139 of 5342 patients. Age at diabetes onset and mean HbA1c did not fulfill the PH assumption, and thus, we stratified the HbA1c level (<6.5, 6.5–7.5, 7.5–8.5, 8.5–9.5, 9.5–10.5, 10.5–11.5 or 11.5>) and age at diabetes onset subgroups (<6.5, 6.5–10.5, 10.5–14.5, 14.5–20.5, 20.5–26.5 and 26.5>). Survival modeling, including Kaplan-Meier visualization, were performed in R statistical software with survival and survminer packages.

2.3.6 Phenotypic characterization of lead GWAS loci

Peripheral blood pressures including systolic pressure (SBP, N=4516), diastolic pressure (DBP, N=4514), pulse pressure (PP; SBP – DBP, N=4515) and mean arterial pressure (MAP; DBP - 1/3×PP, N=4515) were measured during the FinnDiane visits or obtained from medical records. In addition, examined variables included cholesterol (N=4597), low-density lipoprotein (LDL, calculated with Friedewald's formula, N=4527), high-density lipoprotein (HDL, N=4568) and log-transformed triglyceride (N=4583) concentrations; two inflammatory markers, leukocyte (N=469) serum concentration and log-transformed C-reactive protein (N=4066) concentration; log-transformed and averaged annual number of antibiotic purchases (N=4744); body mass index (N=4367) and waist to hip ratio (N=3985); and the time weighted mean HbA1c concentration (N=5131). Further arterial stiffness and atherosclerosis measures collected during the FinnDiane visits of the Helsinki district were included. Arterial stiffness was measured with SphygmoCor device (Atcor Medical, Sydney, Australia) by recording peripheral measurements; and by estimating corresponding central blood pressures (N=899)

including central end systolic pressure (CESP) in addition to central SBP, DBP, MAP and PP. Furthermore, subendocardial viability ratio (N=899), time to reflection (N=899), central augmentation index (N=898), brachial pulse wave velocity (PWV, N=391) and central PWV (N=367) were measured, out of which pulse wave velocities only after 2008. The amount of atherosclerotic plaque at common carotid arteries was measured with Esaote ultrasound (Esaote Artlab, Genova, Italy) revealing the diameter (N=463) in addition to intima-media thicknesses of both right (N=457) and left arteries (N=458) **(Supplemental Table 3).**

Associations between the clinical variables and GWAS lead loci ($p<10^{-6}$) were performed with Rvtest[21] by requiring minor allele count ≥5, and adjusting for gender and diabetes onset year. Diabetes duration was restricted to a minimum of five years. Due to inter-correlation between the variables, only the number of SNPs was considered when correcting for multiple testing.

2.3.7 DKD interaction analysis

Interactions between the lead variants and DKD in CAD development were assessed with logistic regression in R statistical software. DKD was defined as a binary: normoalbuminuria and albuminuria/end-stage renal disease. Analyses were adjusted for genotyping batch, gender and diabetes onset calendar year.

2.3.8 Histone modifications and chromatin interactions

Histone modification CHiP-seq data was queried for the lead loci in adult heart left ventricles and iPSC-derived cardiomyocytes in the WashU epigenome browser (http://epigenomegateway.wustl.edu/legacy/) within the "promoter interaction map for cardiovascular disease genetics" public track[25]. Promoter Capture Hi-C (PCHiC) chromatin conformation data in hESC derived cardiomyocytes[26] were queried for lead loci using CHiCP browser (www.chicp.org)[27].

2.3.9 Polygenic risk score

We calculated a genetic risk score for CAD with general population risk variants reported by Erdmann et al.[10], defined as a mean of the SNP dosages, weighted by the corresponding natural logarithm of risk allele OR from an original study (**Supplemental Table 4**). Association between the genetic risk score and CAD was evaluated in R with logistic regression, and model fit was estimated with McFadden pseudo R^2.

2.3.10 Pathway analysis

Genes and genetic pathways were scored with PASCAL software[28]. Gene scoring was performed by counting sum-of-chi-squares statistics from gene regions with 50 kb up- and downstream extensions. Variants with MAF≥0.01 in the 1000 Genomes project and genes with ≤3000 variants were accepted. Genes within the same pathway and 1.0 Mb distance were fused together for pathway scoring. Bonferroni corrected significance thresholds were utilized for significance indication; 5×10^{-6} for genes and 5×10^{-5} for pathways.

3. Results

3.1 Lead findings for CAD

The GWAS included 4869 individuals with T1D, out of which 941 (19%) had suffered a CAD event **(Table 2)**. The mean age at the first CAD event was 52.5 (standard deviation [SD]=10.2) years; 11% of the cases experienced the first CAD event before the age of 40 years, and for 65 (6.9%) cases, death from a cardiovascular cause was the first severe cardiovascular event. GWAS included 8,744,746 SNPs and revealed two loci that were genome-wide significantly associated ($p<5\times10^{-8}$) with CAD **(Figure 1**): a common variant rs1970112 (MAF=0.41) within an intron of the well-known *CDKN2B-AS1* locus on chromosome 9p21 (OR=1.32, 95% confidence interval [CI] 1.2-1.46, $p=1.50\times10^{-8}$), and a low frequency intergenic variant rs6055069 (MAF=0.02) on chromosome 20 within *DEFB127* gene promoter region (OR for

minor T allele 4.17, 95%CI 2.63-6.67, p=2.35×10^{-9}; **Figure 2)**. A total of 10 loci reached a suggestive p-value of <10^{-6} **(Table 1)**. Statistical fine mapping suggested that each lead locus included only one underlying causal variant (**Supplemental Table 5)**; except for the chromosome 11 locus which had a 43% posterior probability to include two causal variants **(Supplemental Figure 3)**.

3.1.1 Survival analyses

Since the risk of CAD increases with diabetes duration[3], we implemented survival models with various adjustment covariates for the two genome-wide significant variants. Both SNPs had significant effect sizes in all models **(Figure 3, Supplemental Table 6)**. Furthermore, the eight suggestive loci had significant genotypic effects in Cox PH models when accounting for diabetes duration and by adjusting for calendar year of diabetes onset, gender, age at diabetes onset and mean HbA1c level **(Supplemental Table 7)**.

3.1.2 Adjustment with DKD status

The *CDKN2B-AS1* variant effect remained significant with a similar OR when adjusted for DKD (OR=1.32, p=6.67×10^{-9}), suggesting that the locus affects CAD risk independently of DKD **(Supplemental Table 8)**. On the contrary, *DEFB127* variant effect size and statistical significance were attenuated when adjusting for DKD (OR=2.70, p=1.09×10^{-6}), which may reflect the high correlation between CAD and DKD in T1D, shared etiology between DKD and CAD, or CAD due to DKD. Adjustment with DKD attenuated the association with CAD for rs70962766 near *B3GNT2* considerably (p=0.0018). However, interaction terms between the variants and DKD in the development of CAD were insignificant for all lead SNPs (**Supplemental Table 8**).

3.1.3 Replication

Replication of the ten loci was attempted by adjusting for DKD as in the previous publication, as well as by reanalyzing without the DKD adjustment from three independent GWAS studies on CAD in T1D, including a total of 434 cases with CAD and 3123 controls without CAD[13]. Due to low allele frequency or limited imputation quality, data were available for five of the lead SNPs in one or more replication studies. Association at the 9p21 reached p=0.09 in replication without DKD adjustment. Despite the unsuccessful replication, meta-analysis across the four cohorts kept the variant genome-wide significant (p=1.19×10^{-8}). When adjusted for DKD, however, the 9p21 replicated (p=0.04), thus improving the combined p-value from the meta-analysis of the three studies and our DKD adjusted statistics to p=1.91×10^{-9}. None of the other loci were replicated (**Table 1, Figure 4**). Of note, with DKD adjustment, *CNTNAP5* showcased the same direction of effect in all four cohorts. Further look-up of the ten lead loci in GWAS data for the general population including 60,801 CAD cases and 123,504 controls was significant only for the *CDKN2B-AS1* locus (rs1970112 p=1.2×10^{-89})[29] (**Supplemental Table 9**).

3.2 Phenotypic characterization of the lead loci

Three of the lead loci represented association with at least one CVD predisposing phenotype (p<0.005, corrected for 10 loci; **Figure 5, Supplemental Table 10)**, thus, elucidating their potential roles in the pathogenesis. rs70962766 near *B3GNT2* was associated with central SBP and CESP, and nominally (p<0.05) with multiple other arterial stiffness measures. rs1344228 near *CNTNAP5* locus was associated with intima media thickness (p<0.005), a strong indicator of vascular disease[30]. In addition, both rs70962766 and rs1344228 were nominally associated with increased HbA1c (p<0.05). The variant rs2112481 near *GRAMD2B* was associated with serum leukocyte level (p<0.005), thus possibly acting by inducing inflammation, and nominally with pulse pressure. The lead variant at *CDKN2B-AS1* was nominally associated

with brachial PWV ($p<0.05$). Arterial stiffness, measured as PWV and characterized by decreased elastic properties of the vessels, precedes hypertension[31].

3.3 Association with gene expression and epigenetic interactions

3.3.1 Expression quantitative trait loci (eQTL)

Associations between the lead loci and gene expression levels at different tissues were inspected from the Genotype-Tissue Expression (GTEx) portal comparing whole genome sequencing genotypes of the study participants to tissue specific RNA-seq gene expression levels. The rs1970112 at *CDKN2B-AS1* was the only locus with significant eQTL associations. The risk allele was associated with increased *CDKN2B* expression in brain cortex (normalized effect size; NES=0.36) in addition to decreased expression in sigmoid colon (NES=-0.29), minor salivary gland (NES=-0.21) and tibial nerve (NES=-0.15) after correcting for the number of tissues ($p_{eQTL}<0.001$). Of note, no *DEFB127* expression was detected in GTEx eQTL tissues, possibly explaining the lack of eQTL associations.

3.3.2 Histone modifications and chromatin interactions

Three of the lead SNPs were located within the promoter region defined as 5kbp upstream of a protein coding gene transcription start site, including rs6055069 4kbp upstream of *DEFB127*, rs2112481 4kb upstream of *GRAMD2B*, and rs574480779 2kbp upstream of *KIF4B*. Each of these three SNPs overlapped a left ventricle histone H3K27ac peak[25], typically found on promoters or enhancers of active genes, thus providing further support for the *DEFB127*, *GRAMD2B* and *KIF4B* as the target genes **(Supplemental Table 10)**. Furthermore, chromatin conformation data on hESC derived cardiomyocytes[26] showed interaction between the rs2112481 containing DNA fragment and *GRAMD2B* and *ALDH7A1* genes, suggesting that the variant affects the transcription of the two genes **(Supplemental Table 10)**.

3.5 CAD loci in the general population

In the general population, GWAS have successfully revealed a total of 163 CAD susceptibility loci[10]. Out of the 156 variants available in our GWAS, only the *CDKN2B-AS1* variant was significant after Bonferroni correction, with a similar effect size as in the general population **(Figure 6, Supplemental Table 4)**. Furthermore, rs3827066 with *PLTP* and *MMP9* as the most likely target genes[10] showcased a stronger effect size for CAD in T1D (OR=1.24 [1.08–1.42]) than in the general population (OR=1.04 [1.03–1.06])[32], with significant heterogeneity in meta-analysis (I^2=84.5%, p=0.011). There were two other variants with significantly higher effect sizes in T1D as well as 11 variants with significantly higher effect sizes in the general population. Of note, we had low power to replicate many of the loci (**Supplemental Figure 2**). However, the genetic risk score based on the general population CAD risk variants was significantly associated with CAD also in T1D (p=2.74×10^{-8}, OR per 0.005 unit increase 1.37 [1.23–1.54]) although with limited model fit (R^2=0.0065); individuals with CAD had slightly higher polygenic risk scores (mean 0.00846, SD 0.00312) than individuals without CAD (mean 0.00781, SD 0.00319), suggesting that individuals with T1D are modestly affected also by the known CAD susceptibility loci **(Supplemental Figure 4)**. Despite a few loci with increased effect sizes in T1D, most of the known CAD risk variants seem to affect those with T1D similarly; effect sizes (β) between the two conditions were correlated (r=0.24, p=3.04×10^{-3}).

3.6 Replication of previous loci

Charmet et al.[13] suggested 21 genetic loci within their first-stage GWAS for CAD in T1D. Sixteen of them were studied in our GWAS with no significant associations despite high statistical power for replication **(Supplemental Table 2)**. Our GWAS did not replicate the T2D CAD risk SNPs rs10911021 near *GLUL* gene[15] (p=0.994, power=0.79), rs9299879 within an intron of *MGMT* gene (p=0.364, power=1.00), or rs57922 close to non-coding RNAs[16] (p=0.279, power=1.00), nor the CAD risk variant rs74617384 at the *LPA* locus[17](p=0.227,

power=0.51). However, rs10811652 at 9p21, the lead SNP in the GWAS on CAD in any diabetes[17], was associated with CAD also in our GWAS (p=1.48×10^{-7}, power=0.41).

3.7 Gene level and Pathway analyses

Mutations increasing CAD risk may accumulate within genes and genetic pathways. Gene scoring supported the *CDKN2B-AS1* association with CAD (p<2.5×10^{-6}, significant after Bonferroni correction). In addition, three genes on chromosome 1 at location 40.94Mb to 41.01Mb (*ZFP69*, *EXO5* and *ZNF684*) reached suggestive significance level, potentially representing the same association signal (p<5×10^{-5} for each, strongest for *EXO5*; **Supplemental Table 12; Supplemental Figure 5**). The *DEFB125-DEFB127* region was associated with CAD, although not significantly after Bonferroni correction (p-values 1.5×10^{-4}, 1.4×10^{-3} and 1.9×10^{-3}, respectively).

Pathway scoring revealed only suggestive pathways (p<5×10^{-3}) including recruitment of nuclear mitotic apparatus protein to mitotic centrosomes, and SET pathway **(Supplemental Table 13)**.

4. Discussion

Genome-wide association studies have revealed multiple loci that account for a great proportion of CAD heritability[10]. While diabetes is a well-known risk factor for CAD, and pathophysiology of CAD may differ in T1D from the general population[5], only a few GWAS have been performed for CAD in individuals with diabetes. We conducted the largest GWAS on CAD in individuals with T1D (N=4869), and identified variants with genome-wide significance in the previously reported 9p21 region within *CDKN2B-AS1,* and on chromosome 20p13 near *DEFB127*. Both loci were also supported by survival analysis from diabetes onset until CAD event.

346 The 9p21 genetic region has previously been associated with CAD in the general population,
347 as well as with T2D *per se*[33], and recently also with CAD in individuals with diabetes[17].
348 However, this is the first time that the locus is associated with CAD in individuals with T1D
349 with genome-wide statistical significance. The effect size was similar for the previously
350 reported lead risk variant at 9p21 in the general population and in T1D. The region includes
351 genes *CDKN2A* and *CDKN2B* encoding for cyclin-dependent kinase inhibitor proteins, which
352 in turn may alter cell proliferation[34]. Similarly to our GWAS, disease risk variants are usually
353 discovered at the non-coding *CDKN2B-AS1* region, which has been suggested to alter histone
354 modifications of other genetic loci by binding to polycomb protein subunits[34] e.g. CBX7 and
355 SUZ12—likely targeting *CDKN2A* and *CDKN2B*, respectively[35,36]. Of note, *CDKN2B-AS1*
356 may act also in trans on *CDKN2B*[34]. Interestingly, Motterle et al. (2012) showed that genetic
357 variation at 9p21 impact *CDKN2A/B* expression in vascular smooth muscle cells as well as
358 their proliferation rate[37], thus potentially inducing vascular injury. In our detailed phenotypic
359 analysis, the locus was nominally associated with higher brachial PVW ($p<0.05$), although not
360 significantly after multiple testing correction ($p<0.005$). Brachial PWV is indicative of arterial
361 stiffness, which often precedes hypertension and eventually arterial diseases. Lastly, *CDKN2B-*
362 *AS1* may play a role in inflammatory pathways in co-operation with YY1 transcription factor[38].

363 The diabetes-specific discovery at 20p13 locus with $p=2.35\times10^{-9}$ for rs6055069 at the
364 discovery stage, is located only 4kb upstream of the *DEFB127* gene encoding β-defensin 127
365 (also known as β-defensin 27). Defensins are small secreted antimicrobial peptides preventing
366 microbial colonization on epithelial surfaces, capable to inhibit the growth of bacteria and
367 fungi. The cationic β-defensins attract negatively charged bacteria, diffuse their hydrophobic
368 part onto the cell membrane of the bacterium and generate pores, eventually leading to cell
369 death. Of note, *DEFB127* has been shown exhibit antimicrobial activity towards E. coli[39].
370 Although *DEFB127* expression is highest at testis and epididymis similarly to many other β-

371 defensins, it is also moderately expressed in the heart, pancreas, kidney, skeletal muscle, liver
372 and lung[40]. Furthermore, β-defensins are believed to contribute to innate and adaptive immune
373 systems[40]. Yang et al. (1999) suggested that β-defensin 2 chemoattracts memory T-cells and
374 immature dendritic cells[41], while Soruri et al. (2007) were unable to replicate this and instead
375 proposed β-defensins 1-4 to chemoattract macrophages[42]. Of note, individuals with diabetes
376 have low grade chronic inflammation[43], which further contributes to atherosclerosis. However,
377 rs6055069 was not significantly associated with inflammatory markers. Interestingly,
378 individuals with diabetes also have more infections[44], and thus, *DEFB127* may affect the risk
379 of CAD through elevated infection susceptibility, even though the variant was not directly
380 associated with the yearly amount of antibiotic purchases. Of note, even autoimmunity driven
381 chronic myocardial inflammation has been suggested to play a role in cardiac dysfunction in
382 T1D[45]. We have previously shown that serum α-defensin (class 1-3) levels are associated with
383 DKD in individuals with T1D[46], but to our knowledge, this is the first report linking genetic
384 variation in the β-defensin region to cardiovascular outcomes.

385 Association at the 9p21 locus replicated when adjusted for DKD in 3557 additional individuals
386 with T1D (rs1970112 p=0.04), suggesting DKD independent mechanisms of action. As the
387 association at chromosome 20p13 did not replicate, the signal remains suggestive, pending
388 further confirmation. However, it should be noted that replication data were available only for
389 322 cases and 1917 controls for the low frequency variant rs6055059, resulting in limited
390 power.

391 We also discovered eight suggestive genetic variants ($p<10^{-6}$) in the initial GWAS analysis,
392 albeit with attenuated associations at the meta-analysis stage, if available in replication cohorts.
393 We further attempted to replicate the lead variants from Nikpay et al. (2015) general population
394 CAD GWAS[29]. Out of the five variants available, including *DEFB127*, only the *CDKN2B-AS1*
395 variant replicated, thus suggesting T1D specific effects.

Three of these suggestive variants showed association with CAD predisposing phenotypes, most importantly rs70962766 (*B3GNT2)* with central blood pressure, rs1344228 (*CNTNAP5*) with atherosclerosis through direct intima-media thickness measures, and rs2112481 (*GRAMD2B*) with leukocyte concentration and therefore inflammation, thus, suggesting potential mechanisms of action. It should be noted that this cardio-phenome-wide analysis for the lead SNPs included a variable number of observations for the studied phenotypes, and only the above-mentioned associations remained significant after correction for ten tested SNPs (but assuming inter-correlation between the CVD related phenotypes).

PCHiC data in cardiomyocytes suggested chromatin interaction between the lead variant rs2112481 4kbp upstream *GRAMD2B,* and the *GRAMD2B* and *ALDH7A1* genes. Of note, an intronic variant in *ALDH7A1* was recently associated with coronary artery calcified atherosclerotic plaque in a GWAS on African Americans with T2D[47]. Furthermore, gene scoring suggested three genes: *ZFP69, EXO5 and ZNF684*, of which *ZFP69* has been previously linked to hyperlipidemia[48].

Among the known CAD risk variants from the general population, only rs1333049 in *CDKN2B-AS1* was significantly associated with CAD in T1D after correction for multiple testing. Nevertheless, the genetic risk score for the known CAD susceptibility loci was modestly but significantly associated with CAD in T1D, suggesting that genetic risk factors discovered in the general population also affect individuals with T1D. Of note, the variant with *PLTP* and *MMP9* as likely target genes showcased a stronger effect size in T1D; *PLTP* has been associated with hypertriglyceridemia especially in obesity and T2D[49].

One limitation of the study is the use of registry data. However, the Finnish administrative registers cover all deaths and hospitalization events, and capture CAD events well[50]. In addition, this study suffered from limited power due to moderate GWAS size. However, this is by far the largest GWAS on CAD in individuals with T1D, a group of people with particularly

high CAD risk. Furthermore, we have integrated our GWAS with multiple different omics data such as eQTL, left ventricle histone modification, and hESC derived cardiomyocyte chromatin conformation data, linking the SNP associations to interesting target genes.

To conclude, this is the first time that variants on chromosome 9p21 in *CDKN2-AS1* locus were genome-wide significantly associated with CAD in T1D. Furthermore, we identified with genome-wide significance a novel locus for CAD in individuals with T1D on the β-defensin 127 promoter region, potentially acting through infection susceptibility, elevated in individuals with diabetes. While this and the suggestive loci require confirmation in further studies, they suggest novel biological mechanisms for cardiovascular complications in diabetes.

5. Funding

This work and the FinnDiane project were supported by Folkhälsan Research Foundation, Wilhelm och Else Stockmann Foundation, Novo Nordisk Foundation [NNFOC0013659], Liv och Hälsa Society, Helsinki University Central Hospital Research Funds (EVO) and Academy of Finland [299200, 275614, and 316664], European Foundation for the Study of Diabetes (EFSD) Young Investigator Research Award funds [to NS], Ida Montini Foundation [to AAVA], an EFSD award supported by EFSD/Sanofi European Diabetes Research Programme in Macrovascular Complications, and Finnish Foundation for Cardiovascular Research. Genotyping of the GWAS data for discovery and replication cohorts was funded by the Juvenile Diabetes Research Foundation (JDRF) within the Diabetic Nephropathy Collaborative Research Initiative [DNCRI; grant 17-2013-7], with GWAS quality control and imputation performed at University of Virginia. Charmet R. was financially supported by a PhD grant from the Region Ile de France (CORDDIM). Recruitment of the UK-ROI cohort was funded by Diabetes UK and the JDRF. Further research involving the UK-ROI cohort was funded by grants from Health Research Board, Science Foundation Ireland, Northern Ireland Public Health Agency (Research & Development Division) and Medical Research Council as part of the USA-Ireland-Northern Ireland research partnership and Department for the Economy NI 15/IA/3152; this funding supported investigators in the Warren 3/UK GoKinD Study Group, which includes, Belfast: A. P. Maxwell, A. J. McKnight, D. A. Savage; Edinburgh: J. Walker; London: S. Thomas, G. C. Viberti; Manchester: A. J. M. Boulton; Newcastle: S. Marshall; Plymouth: A. G. Demaine, B. A. Millward; Swansea: S. C. Bain.

6. Author contributions

A.A.V.A. and N.S. analyzed the data, contributed to the interpretation of the data, drafted the work and wrote the manuscript. N.S. further contributed to acquisition of data and conception

and study design. A.S. and E.V. contributed to acquisition of genetic data. D.G., C.F., V.H. and P.-H.G. contributed to interpretation of data, acquisition of phenotypic data, and to conception and study design. D.-A.T. contributed to data analysis and interpretation. R.C., A.J.M., and T.S.A. contributed to data analysis. S.H., A.P.M., and P.R. contributed to acquisition of genetic and phenotypic data, to conception and study design. A.S., E.V., D.G., C.F., V.H., P.-H.G., D.-A.T., R.C., A.J.M., T.S.A., S.H., A.P.M., and P.R. revised the work critically for important intellectual content. All authors gave final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

7. Acknowledgements

We acknowledge FinnDiane participants and the doctors and nurses participating in the FinnDiane study **(Supplemental Table 13)** and thank the laboratory staff for their skillful work. We acknowledge all participants and participating centers from replication cohorts (Supplemental note: List of participating centers from French cohorts).

Data on CAD in general population have been contributed by CARDIoGRAMplusC4D investigators and have been downloaded from www.CARDIOGRAMPLUSC4D.ORG. The GTEx project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data reported here were obtained from the GTEx Portal on 07/01/2019 from https://gtexportal.org/home/.

8. Conflict of interests

P.-H.G. has received investigator-initiated research grants from Eli Lilly and Roche, is an advisory board member for AbbVie, Astellas AstraZeneca, Boehringer Ingelheim, Cebix, Eli

Lilly, Janssen, Medscape, Merck Sharp & Dohme, Nestle Novartis, Novo Nordisk and Sanofi; and has received lecture fees from AstraZeneca, Boehringer Ingelheim, Eli Lilly, Elo Water, Genzyme, Merck Sharp & Dohme, Medscape, Mundipharma Novo Nordisk, PeerVoice, Sanofi and SCIARC.

S.H. personally or collectively received research grants, honoraria and fees from the following companies in the last 3 years: Astra Zeneca, Bayer, Boehringer Ingelheim, Dinno Santé, Eli Lilly, LVL, Johnson & Johnson, Medtronic, MSD, NovoNordisk, Novartis, Pierre Fabre Santé, Sanofi, Servier, Valbiotis.

Each disclosed relationship is considered modest, and other authors declare no conflicts of interest.

9. References

1. International Diabetes Federation. IDF Diabetes Atlas, 8th edn. *Brussels, Belgium: International Diabetes Federation* 2017;.

2. Pambianco G, Costacou T, Ellis D, Becker DJ, Klein R, Orchard TJ. The 30-year natural history of type 1 diabetes complications: the Pittsburgh Epidemiology of Diabetes Complications Study experience. *Diabetes* 2006;**55**:1463-1469.

3. Harjutsalo V, Thomas MC, Forsblom C, Groop PH, FinnDiane Study Group. Risk of coronary artery disease and stroke according to sex and presence of diabetic nephropathy in type 1 diabetes. *Diabetes Obes Metab* 2018;**20**:2759-2767.

4. Tuomilehto J, Borch-Johnsen K, Molarius A, Forsen T, Rastenyte D, Sarti C, Reunanen A. Incidence of cardiovascular disease in Type 1 (insulin-dependent) diabetic subjects with and without diabetic nephropathy in Finland. *Diabetologia* 1998;**41**:784-790.

5. de Ferranti SD, de Boer IH, Fonseca V, Fox CS, Golden SH, Lavie CJ, Magge SN, Marx N, McGuire DK, Orchard TJ, Zinman B, Eckel RH. Type 1 diabetes mellitus and cardiovascular disease: a scientific statement from the American Heart Association and American Diabetes Association. *Circulation* 2014;**130**:1110-1130.

6. Tolonen N, Forsblom C, Makinen VP, Harjutsalo V, Gordin D, Feodoroff M, Sandholm N, Thorn LM, Waden J, Taskinen MR, Groop PH, on behalf of the FinnDiane Study Group. Different Lipid Variables Predict Incident Coronary Artery Disease in Patients With Type 1 Diabetes With or Without Diabetic Nephropathy: The FinnDiane Study. *Diabetes Care* 2014;**37**:2374-2382.

7. Rawshani A, Rawshani A, Franzen S, Eliasson B, Svensson AM, Miftaraj M, McGuire DK, Sattar N, Rosengren A, Gudbjornsdottir S. Mortality and Cardiovascular Disease in Type 1 and Type 2 Diabetes. *N Engl J Med* 2017;**376**:1407-1418.

8. Myers RH, Kiely DK, Cupples LA, Kannel WB. Parental history is an independent risk factor for coronary artery disease: the Framingham Study. *Am Heart J* 1990;**120**:963-969.

9. Earle K, Walker J, Hill C, Viberti G. Familial clustering of cardiovascular disease in patients with insulin-dependent diabetes and nephropathy. *N Engl J Med* 1992;**326**:673-677.

10. Erdmann J, Kessler T, Munoz Venegas L, Schunkert H. A decade of genome-wide association studies for coronary artery disease: the challenges ahead. *Cardiovasc Res* 2018;**114**:1241-1257.

11. Raffield LM, Cox AJ, Carr JJ, Freedman BI, Hicks PJ, Langefeld CD, Hsu F, Bowden DW. Analysis of a cardiovascular disease genetic risk score in the Diabetes Heart Study. *Acta Diabetol* 2015;**52**:743-751.

12. Qi L, Parast L, Cai T, Powers C, Gervino EV, Hauser TH, Hu FB, Doria A. Genetic susceptibility to coronary heart disease in type 2 diabetes: 3 independent studies. *J Am Coll Cardiol* 2011;**58**:2675-2682.

13. Charmet R, Duffy S, Keshavarzi S, Gyorgy B, Marre M, Rossing P, McKnight AJ, Maxwell AP, Ahluwalia TVS, Paterson AD, Tregouet DA, Hadjadj S. Novel risk genes identified in a genome-wide association study for coronary artery disease in patients with type 1 diabetes. *Cardiovasc Diabetol* 2018;**17**:61-018-0705-0.

14. Doria A, Wojcik J, Xu R, Gervino EV, Hauser TH, Johnstone MT, Nolan D, Hu FB, Warram JH. Interaction between poor glycemic control and 9p21 locus on risk of coronary artery disease in type 2 diabetes. *Jama* 2008;**300**:2389-2397.

15. Qi L, Qi Q, Prudente S, Mendonca C, Andreozzi F, di PN, Sturma M, Novelli V, Mannino GC, Formoso G, Gervino EV, Hauser TH, Muehlschlegel JD, Niewczas MA, Krolewski AS, Biolo G, Pandolfi A, Rimm E, Sesti G, Trischitta V, Hu F, Doria A. Association between a genetic variant related to glutamic acid metabolism and coronary heart disease in individuals with type 2 diabetes. *Jama* 2013;**310**:821-828.

16. Shah HS, Gao H, Morieri ML, Skupien J, Marvel S, Pare G, Mannino GC, Buranasupkajorn P, Mendonca C, Hastings T, Marcovina SM, Sigal RJ, Gerstein HC, Wagner MJ, Motsinger-Reif AA, Buse JB, Kraft P, Mychaleckyj JC, Doria A. Genetic Predictors of Cardiovascular Mortality During Intensive Glycemic Control in Type 2 Diabetes: Findings From the ACCORD Clinical Trial. *Diabetes Care* 2016;**39**:1915-1924.

17. Fall T, Gustafsson S, Orho-Melander M, Ingelsson E. Genome-wide association study of coronary artery disease among individuals with diabetes: the UK Biobank. *Diabetologia* 2018;**61**:2174-2179.

18. Sandholm N, Haukka JK, Toppila I, Valo E, Harjutsalo V, Forsblom C, Groop PH. Confirmation of GLRA3 as a susceptibility locus for albuminuria in Finnish patients with type 1 diabetes. *Sci Rep* 2018;**8**:12408-018-29211-1.

19. Goldstein JI, Crenshaw A, Carey J, Grant GB, Maguire J, Fromer M, O'Dushlaine C, Moran JL, Chambert K, Stevens C, Swedish Schizophrenia Consortium, ARRA Autism Sequencing Consortium, Sklar P, Hultman CM, Purcell S, McCarroll SA, Sullivan PF, Daly MJ, Neale BM. zCall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics* 2012;**28**:2543-2545.

20. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, Schlessinger D, Stambolian D, Loh PR, Iacono WG, Swaroop A, Scott LJ, Cucca F, Kronenberg F, Boehnke M, Abecasis GR, Fuchsberger C. Next-generation genotype imputation service and methods. *Nat Genet* 2016;**48**:1284-1287.

21. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics* 2016;**32**:1423-1426.

22. Benner C, Spencer CC, Havulinna AS, Salomaa V, Ripatti S, Pirinen M. FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* 2016;**32**:1493-1501.

23. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;**26**:2190-2191.

24. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003;**19**:149-150.

562 25. Montefiori LE, Sobreira DR, Sakabe NJ, Aneas I, Joslin AC, Hansen GT, Bozek G, Moskowitz IP,
563 McNally EM, Nobrega MA. A promoter interaction map for cardiovascular disease genetics. *Elife*
564 2018;**7**:e35788.

565 26. Choy MK, Javierre BM, Williams SG, Baross SL, Liu Y, Wingett SW, Akbarov A, Wallace C, Freire-
566 Pritchett P, Rugg-Gunn PJ, Spivakov M, Fraser P, Keavney BD. Promoter interactome of human
567 embryonic stem cell-derived cardiomyocytes connects GWAS regions to cardiac gene networks. *Nat*
568 *Commun* 2018;**9**:2526-018-04931-0.

569 27. Schofield EC, Carver T, Achuthan P, Freire-Pritchett P, Spivakov M, Todd JA, Burren OS. CHiCP: a
570 web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets.
571 *Bioinformatics* 2016;**32**:2511-2513.

572 28. Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. Fast and Rigorous Computation of
573 Gene and Pathway Scores from SNP-Based Summary Statistics. *PLoS Comput Biol* 2016;**12**:e1004714.

574 29. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP,
575 Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI,
576 Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber
577 ME, Lau KW, Lu X, Lu Y, Lyytikainen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati
578 E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade
579 M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier
580 D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, Konig IR, Lannfelt L,
581 Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon
582 FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH,
583 Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger
584 EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T,
585 Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB,
586 Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ,
587 Kooner JS, Kullo IJ, Lehtimaki T, Loos RJF, Melander O, Metspalu A, Marz W, Palmer CN, Perola M,
588 Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM,
589 Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson
590 JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ,
591 Farrall M. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of
592 coronary artery disease. *Nat Genet* 2015;**47**:1121-1130.

593 30. Lorenz MW, Markus HS, Bots ML, Rosvall M, Sitzer M. Prediction of clinical cardiovascular events
594 with carotid intima-media thickness: a systematic review and meta-analysis. *Circulation*
595 2007;**115**:459-467.

596 31. Shirwany NA , Zou MH. Arterial stiffness: a brief review. *Acta Pharmacol Sin* 2010;**31**:1267-1276.

597 32. van der Harst P , Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on
598 the Genetic Architecture of Coronary Artery Disease. *Circ Res* 2018;**122**:433-443.

599 33. Strawbridge RJ , van Zuydam NR. Shared Genetic Contribution of Type 2 Diabetes and
600 Cardiovascular Disease: Implications for Prognosis and Treatment. *Curr Diab Rep* 2018;**18**:59-018-
601 1021-5.

602 34. Congrains A, Kamide K, Ohishi M, Rakugi H. ANRIL: molecular mechanisms and implications in
603 human health. *Int J Mol Sci* 2013;**14**:1278-1292.

35. Kotake Y, Nakagawa T, Kitagawa K, Suzuki S, Liu N, Kitagawa M, Xiong Y. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene* 2011;**30**:1956-1962.

36. Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* 2010;**38**:662-674.

37. Motterle A, Pu X, Wood H, Xiao Q, Gor S, Ng FL, Chan K, Cross F, Shohreh B, Poston RN, Tucker AT, Caulfield MJ, Ye S. Functional analyses of coronary artery disease associated variation on chromosome 9p21 in vascular smooth muscle cells. *Hum Mol Genet* 2012;**21**:4021-4029.

38. Zhou X, Han X, Wittfeldt A, Sun J, Liu C, Wang X, Gan LM, Cao H, Liang Z. Long non-coding RNA ANRIL regulates inflammatory responses as a novel component of NF-kappaB pathway. *RNA Biol* 2016;**13**:98-108.

39. Huang L, Leong SS, Jiang R. Soluble fusion expression and characterization of bioactive human beta-defensin 26 and 27. *Appl Microbiol Biotechnol* 2009;**84**:301-308.

40. Pazgier M, Hoover DM, Yang D, Lu W, Lubkowski J. Human beta-defensins. *Cell Mol Life Sci* 2006;**63**:1294-1313.

41. Yang D, Chertov O, Bykovskaia SN, Chen Q, Buffo MJ, Shogan J, Anderson M, Schroder JM, Wang JM, Howard OM, Oppenheim JJ. Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science* 1999;**286**:525-528.

42. Soruri A, Grigat J, Forssmann U, Riggert J, Zwirner J. beta-Defensins chemoattract macrophages and mast cells but not lymphocytes and dendritic cells: CCR6 is not involved. *Eur J Immunol* 2007;**37**:2474-2486.

43. Saraheimo M, Teppo AM, Forsblom C, Fagerudd J, Groop PH. Diabetic nephropathy is associated with low-grade inflammation in Type 1 diabetic patients. *Diabetologia* 2003;**46**:1402-1407.

44. Simonsen JR, Harjutsalo V, Jarvinen A, Kirveskari J, Forsblom C, Groop PH, Lehto M, FinnDiane Study Group. Bacterial infections in patients with type 1 diabetes: a 14-year follow-up study. *BMJ Open Diabetes Res Care* 2015;**3**:e000067.

45. Lipes MA , Galderisi A. Cardiac autoimmunity as a novel biomarker, mediator, and therapeutic target of heart disease in type 1 diabetes. *Curr Diab Rep* 2015;**15**:30-015-0598-1.

46. Saraheimo M, Forsblom C, Pettersson-Fernholm K, Flyvbjerg A, Groop PH, Frystyk J, FinnDiane Study Group. Increased levels of alpha-defensin (-1, -2 and -3) in type 1 diabetic patients with nephropathy. *Nephrol Dial Transplant* 2008;**23**:914-918.

47. Divers J, Palmer ND, Langefeld CD, Brown WM, Lu L, Hicks PJ, Smith SC, Xu J, Terry JG, Register TC, Wagenknecht LE, Parks JS, Ma L, Chan GC, Buxbaum SG, Correa A, Musani S, Wilson JG, Taylor HA, Bowden DW, Carr JJ, Freedman BI. Genome-wide association study of coronary artery calcified atherosclerotic plaque in African Americans with type 2 diabetes. *BMC Genet* 2017;**18**:105-017-0572-9.

641 48. Chung B, Stadion M, Schulz N, Jain D, Scherneck S, Joost HG, Schurmann A. The diabetes gene
642 Zfp69 modulates hepatic insulin sensitivity in mice. *Diabetologia* 2015;**58**:2403-2413.

643 49. Tzotzas T, Desrumaux C, Lagrost L. Plasma phospholipid transfer protein (PLTP): review of an
644 emerging cardiometabolic risk factor. *Obes Rev* 2009;**10**:403-411.

645 50. Pajunen P, Koukkunen H, Ketonen M, Jerkkola T, Immonen-Raiha P, Karja-Koskenkari P,
646 Mahonen M, Niemela M, Kuulasmaa K, Palomaki P, Mustonen J, Lehtonen A, Arstila M, Vuorenmaa
647 T, Lehto S, Miettinen H, Torppa J, Tuomilehto J, Kesaniemi YA, Pyorala K, Salomaa V. The validity of
648 the Finnish Hospital Discharge Register and Causes of Death Register data on coronary heart disease.
649 *Eur J Cardiovasc Prev Rehabil* 2005;**12**:132-137.

650

651 10. Figure legends

652 **Figure 1: Manhattan plot and QQ-plot of the results showing a genome-wide significant**
653 **association on chromosomes 9 and 20.** A) Red horizontal line on the Manhattan plot
654 illustrates the threshold for genome-wide significance, i.e. p-value$<5\times10^{-8}$. All SNPs within
655 100kb of the genome-wide significant SNPs are highlighted with green. B) QQ-plot shows the
656 expected vs. observed $-\log_{10}(p)$.

657 **Figure 2: LocusZoom plot of the A) chromosome 20 and B) 9p21 regions associated with**
658 **CAD.** Each dot represents a SNPs, with chromosomal position (bp) given on the x-axis, and
659 statistical significance ($-\log_{10}(p$-value)) on the y-axis. Dot color indicates the linkage
660 disequilibrium with the SNP with the smallest *p*-value, marked with purple diamond.

661 **Figure 3: Predicted survivor functions of Cox proportional hazards models and Kaplan-**
662 **Meier survival estimators grouped according to most likely genotypes.** A) Cox PH survivor
663 function of rs6055059 (*DEFB127*) with p-value 8.64×10^{-8}, adjusted for gender and calendar
664 year of diabetes onset (N=4869), B) Kaplan-Meier estimator of rs6055059 (*DEFB127*), C) Cox
665 PH survivor function of rs1970112 (*CDKN2B-AS1*) with p-value 4.13×10^{-8}, adjusted for
666 gender and calendar year of diabetes onset (N=4869), and D) Kaplan-Meier estimator of

667 rs1970112 (*CDKN2B-AS1*)**.** Genotype allele dosages (0-2) represent the alternative allele count
668 (major C for rs6055069, minor C for rs1970112).

669 **Figure 4: Forest plot of odds ratios in FinnDiane and the replication cohorts.** FinnDiane
670 ($N_{cases}/N_{controls}$: 941/3928), France ($N_{cases}/N_{controls}$: 85/1285), Steno ($N_{cases}/N_{controls}$: 237/632) and
671 UK-ROI ($N_{cases}/N_{controls}$: 112/1206). *Upper 95%CI out of bounds.

672 **Figure 5: Cardiophenome-wide analysis of variants discovered in GWAS.** Colored
673 according to normalized CAD risk allele effect size and highlighted with significant non-
674 normalized test p-values. Association tests were conducted for body to mass index (BMI), waist
675 to hip ratio (WHR), systolic blood pressure (SBP), diastolic blood pressure (DBP), mean
676 arterial pressure (MAP), pulse pressure (PP), central augmentation index (CAIX), central
677 systolic blood pressure (CSBP), central diastolic blood pressure (CDBP), central mean arterial
678 pressure (CMAP), central end systolic pressure (CESP), central pulse pressure (CPP),
679 subendocardial viability ratio (SERV), time to reflection (TR), brachial pulse wave velocity
680 (BPWV), central pulse wave velocity (CPWV), left common carotid artery intima media
681 thickness (IMTLEFT), right common carotid artery intima media thickness (IMTRIGHT),
682 mean common carotid artery diameter (DIAMETER), cholesterol concentration (CHOL), low
683 density lipoprotein concentration (LDL), log-transformed triglyceride concentration (TG), high
684 density lipoprotein concentration (HDL), weighted mean HbA1c level (HBA1C), leukocyte
685 concentration (LEUK), log-transformed C-reactive protein concentration (CRP) and log-
686 transformed annual antibiotic purchases (INFECTIONS).

687 **Figure 6: General population CAD risk variant minor allele OR comparison between**
688 **individuals with and without T1D.**

11. Tables

Table 1. **GWAS, replication and meta-analysis results for loci reaching $p<10^{-6}$ in the discovery GWAS.** Chr:Pos_REF/ALT: chromosome and base pair position, reference/alternative (=non-effect/effect) allele; Gene: closest gene(s) underlying or within the promoter region (5kbp upstream), or *within +/- 50kbp, †within +/- 250kbp, or ‡within +/- 500kbp; EAF: Effect (alternative) allele frequency; OR (95% CI): OR and 95% confidence interval; P: p-value; RSQ: imputation r^2 quality estimate; N: N cases/controls in replication; Power: statistical power to observe association with α=0.05/5=0.01 significance level, based on discovery study EAF and 95% lower CI of OR, and 12.2%, 27.3% or 14.4% CAD prevalence in corresponding replication cohort. Dir: effect direction (for effect i.e. alternative allele) in FinnDiane, SDCC, UK-ROI, and French. + indicates predisposition, - protection, ?: data not available.

Discovery stage							Replication			Meta-analysis			Meta-analysis with DKD adjustment*	
SNP	**Chr:Pos_REF/ALT**	**Gene**	**EAF**	**OR (95%CI)**	**P**	**RSQ**	**N**	**Power**	**P**	**Dir**	**OR (95%CI)**	**P**	**OR (95%CI)**	**P**
rs1970112	9:22085598_T/C	*CDKN2B-AS1*	0.41	1.32 (1.2-1.46)	1.50×10^{-8}	0.98	434/3123	0.46	0.09	+ +++	1.27 (1.17-1.38)	1.19×10^{-8}	1.28 (1.18-1.39)	1.91×10^{-9}
rs181176493	2:224157866_G/A	*TRK-TTT15-1(pseudo)*, KCNE4*†	0.01	3.88 (2.33-6.47)	1.98×10^{-7}	0.80	-	-	-	+ ???	3.88 (2.33-6.47)	1.98×10^{-7}	3.62 (2.19-5.98)	4.91×10^{-7}
rs70962766	2:62406999_CTTTTTTTTTTTT/C	*RPSA26(pseudo)*, B3GNT2**	0.99	0.2 (0.11-0.37)	4.68×10^{-7}	0.64	-	-	-	- ???	0.2 (0.11-0.37)	4.68×10^{-7}	0.47 (0.29-0.75)	0.0017
rs138181578	6:45725830_A/AT	*LOC107986519*†, *CLIC5*†	0.99	0.18 (0.09-0.35)	6.78×10^{-7}	0.77	-	-	-	- ???	0.18 (0.09-0.35)	6.78×10^{-7}	0.36 (0.22-0.60)	8.33×10^{-5}
rs10625784	11:106193897_T/TAAA	*LOC643855*, AASDHPPT*†	0.99	0.23 (0.13-0.42)	7.63×10^{-7}	0.83	-	-	-	- ???	0.23 (0.13-0.42)	7.63×10^{-7}	0.44 (0.28-0.70)	0.00045
rs79237700	15:53741612_T/C	*LOC105370826, WDR72**	0.04	1.97 (1.51-2.55)	3.86×10^{-7}	0.75	237/632	0.17	0.79	+ -??	1.82 (1.42-2.33)	2.23×10^{-6}	1.76 (1.37-2.25)	6.88×10^{-6}
rs574480779	5:154390880_TA/T	*KIF4B*	0.03	2.27 (1.64-3.15)	9.36×10^{-7}	0.69	-	-	-	+ ???	2.27 (1.64-3.15)	9.36×10^{-7}	2.14 (1.56-2.95)	2.91×10^{-6}
rs1344228	2:124286924_C/T	*LOC100422580*†, *CNTNAP5*‡	0.11	1.46 (1.26-1.69)	5.64×10^{-7}	1.00	434/3123	0.33	0.71	+ -++	1.33 (1.17-1.51)	8.68×10^{-6}	1.33 (1.17-1.50)	8.30×10^{-6}
rs6055069	20:134284_T/C	*DEFB127*	0.98	0.24 (0.15-0.38)	2.35×10^{-9}	0.62	322/1917	0.95	0.74	- +?-	0.38 (0.25-0.56)	1.36×10^{-6}	0.49 (0.35-0.70)	9.80×10^{-5}
rs2112481	5:125691632_T/G	*GRAMD2B*	0.99	0.10 (0.04-0.23)	6.98×10^{-8}	0.63	434/3123	1	0.35	- +-+	0.67 (0.43-1.01)	0.058	0.73 (0.50-1.06)	0.097

* $N_{FinnDaine}$=926/3777 in meta-analysis with DKD.

Table 2. Clinical characteristics of the subjects. Mean ± SD, or Median ± IQR, or N (%). *Before 30th Dec 2015, including deaths after CVD event; †Death of cardiovascular cause as the first CHD event. DKD: micro- or macroalbuminuria or end stage renal disease.

	Cases	Controls	p-value
N	941	3928	
males, N (%)	530 (56.3)	1992 (50.7)	0.0022
Diabetes onset [calendar year]	1967±9.6	1979±11.1	1.0×10^{-174}
Age at diabetes onset [yr] , Median ± IQR	13.3±11.7	14.3±13.2	9.33×10^{-4}
Age [yr]	52.5±10.2	52.2±10.3	0.48
T1D duration [yr]	37.4±10.4	36.0±10.8	3.12×10^{-4}
Mean HbA1c [%]	8.88±1.34	8.37±1.15	2.98×10^{-22}
Mean HbA1c [mmol/mol]	73.5±14.7	68.0±12.6	2.98×10^{-22}
HbA1c count, Median ± IQR	16±28	23±26	8.00×10^{-22}
Deceased by Dec 31st 2015*	449 (47.7%)	436 (11.1%)	2.62×10^{-150}
Cardiovascular death†	65 (6.9%)	0 (0%)	-
DKD [N cases/controls (%cases)]	689/237 (74%)	1424/2353 (38%)	2.20×10^{-16}

Figure 1

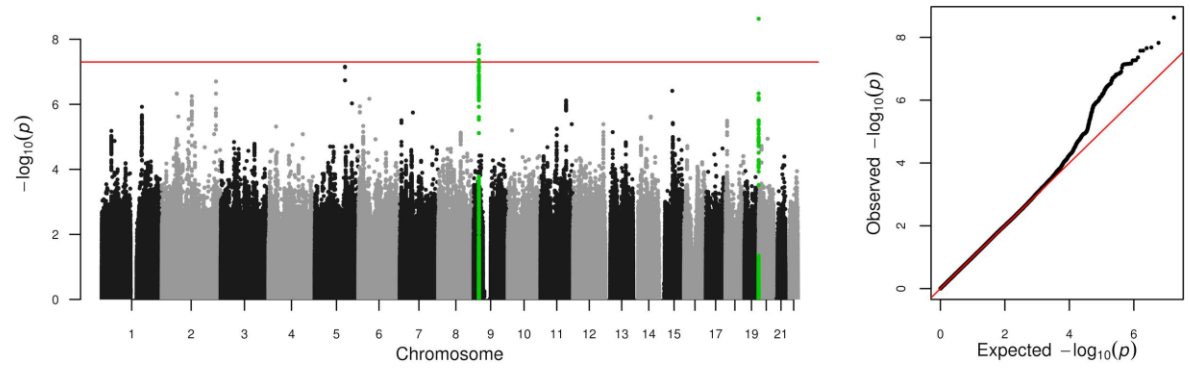

Figure 2

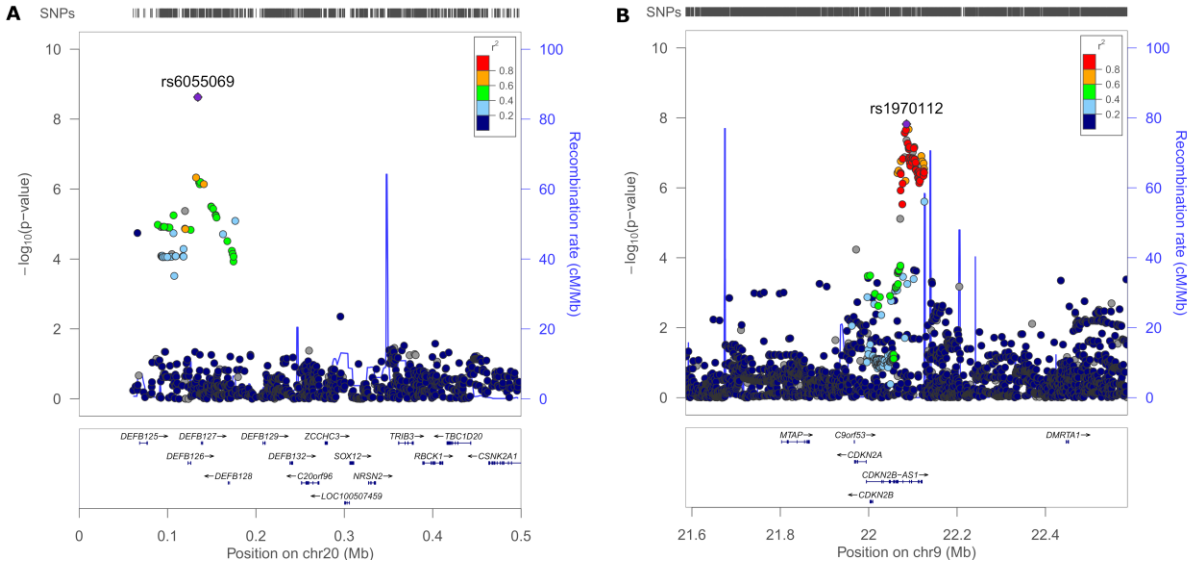

Figure 3

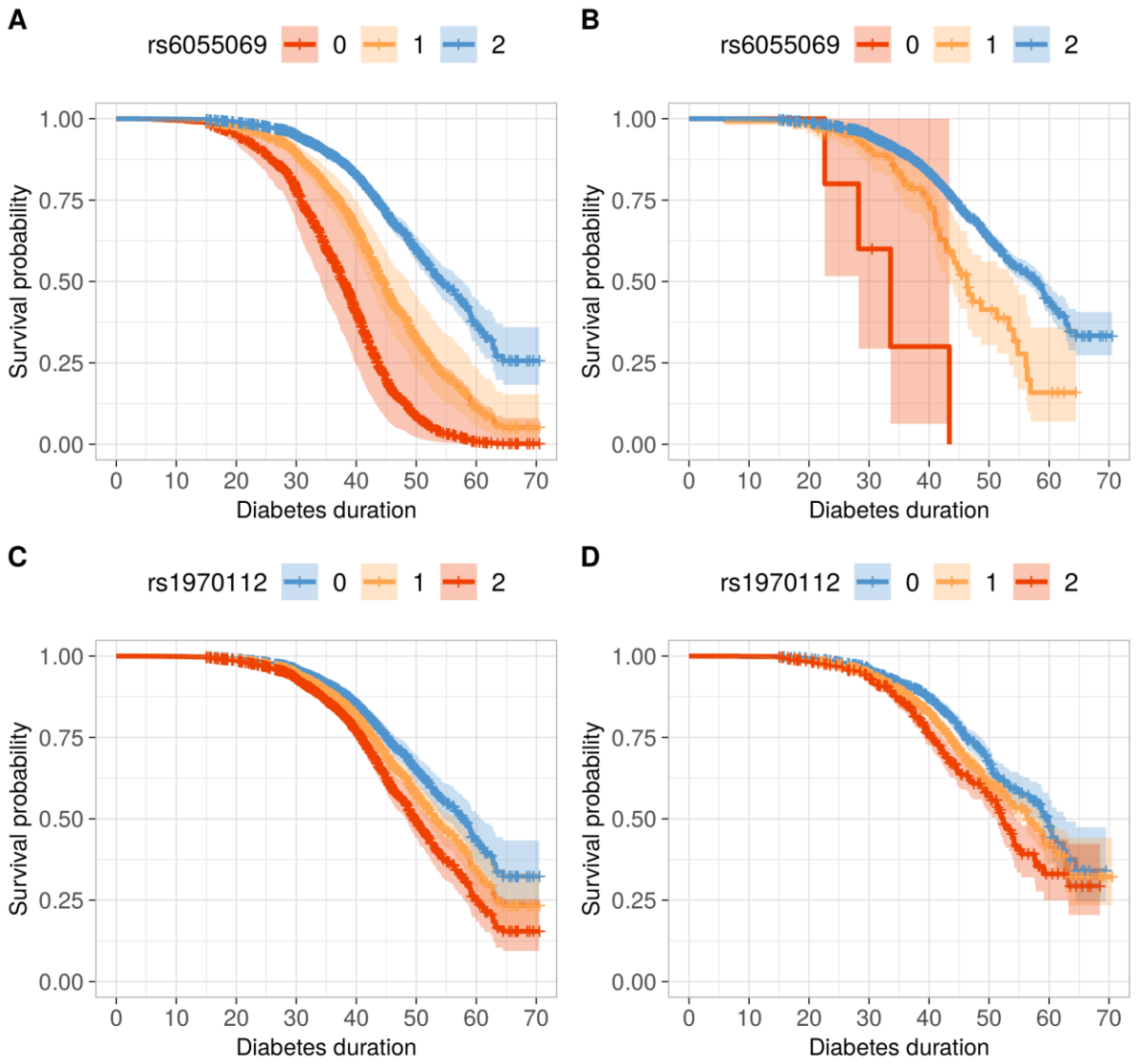
A
rs6055069 0 1 2
Survival probability
1.00
0.75
0.50
0.25
0.00
0 10 20 30 40 50 60 70
Diabetes duration
B
rs6055069 0 1 2
Survival probability
1.00
0.75
0.50
0.25
0.00
0 10 20 30 40 50 60 70
Diabetes duration
C
rs1970112 0 1 2
Survival probability
1.00
0.75
0.50
0.25
0.00
0 10 20 30 40 50 60 70
Diabetes duration
D
rs1970112 0 1 2
Survival probability
1.00
0.75
0.50
0.25
0.00
0 10 20 30 40 50 60 70
Diabetes duration

Figure 4

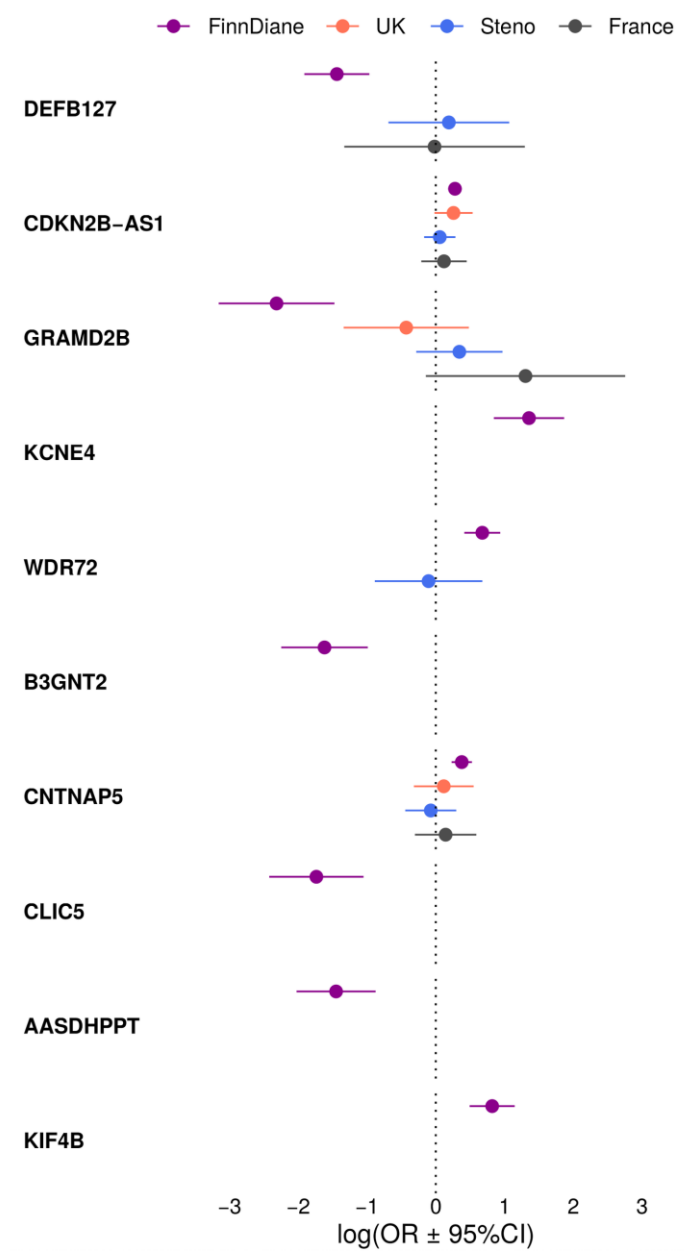
FinnDiane
UK
Steno
France
DEFB127
CDKN2B−AS1
GRAMD2B
KCNE4
WDR72
B3GNT2
CNTNAP5
CLIC5
AASDHPPT
KIF4B
−3
−2
−1
0
1
2
3
log(OR ± 95%CI)

Figure 5

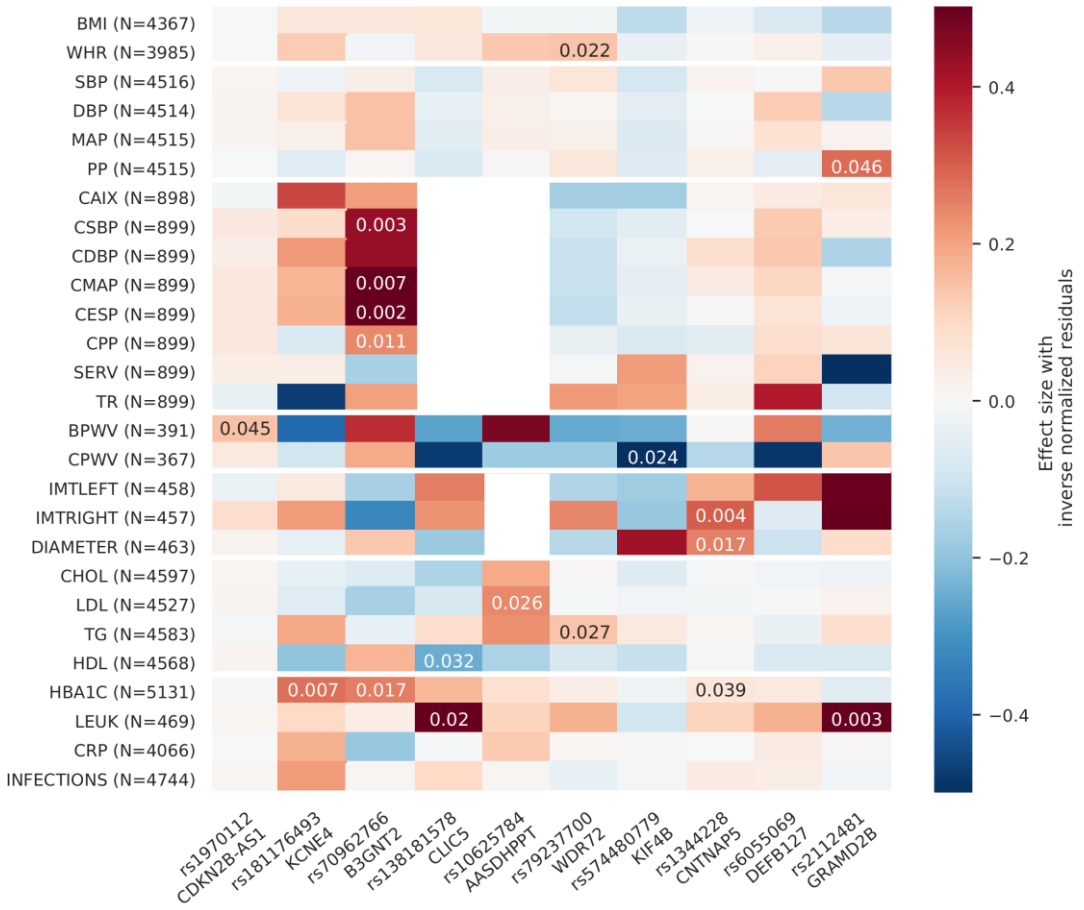

Figure 6