UNIVERSITA' DEGLI STUDI DI VERONA

DIPARTIMENTO DI MEDICINA

SCUOLA di DOTTORATO di
SCIENZE DELLA VITA E DELLA SALUTE

DOTTORATO di RICERCA in
SCIENZE MEDICHE, CLINICHE E SPERIMENTALI

XXVIII ciclo/2013

Tesi di Dottorato

DETERMINANTI GENETICI DI DIABETE MELLITO DI TIPO 2 E FENOTIPI CARDIOMETABOLICI ASSOCIATI

S.S.D. MED13

Coordinatore:
Prof. ENZO BONORA, MD, PhD

Tutor:
Prof. ENZO BONORA, MD, PhD

Dottorando:
Dr. MARCO DAURIZ, MD

UNIVERSITA' DEGLI STUDI DI VERONA

DEPARTMENT of MEDICINE

GRADUATE SCHOOL of
GRADUATE SCHOOL FOR HEALTH AND LIFE SCIENCES

DOCTORAL PROGRAM in
MEDICAL, CLINICAL AND EXPERIMENTAL SCIENCES

XXVIII cicle/2013

Doctoral Thesis

GENETIC DETERMINANTS OF TYPE 2 DIABETES AND ASSOCIATED CARDIOMETABOLIC DISORDERS

S.S.D. MED13

Coordinator:
Prof. ENZO BONORA, MD, PhD

Tutor:
Prof. ENZO BONORA, MD, PhD

Doctoral candidate:
Dr. MARCO DAURIZ, MD

Genetic determinants of type 2 diabetes and associated cardiometabolic disorders –

Determinanti genetici di diabete mellito di tipo 2 e fenotipi cardiometabolici associati

Marco Dauriz

Doctoral Thesis - Tesi di Dottorato

Verona, 11 May 2016

ISBN 978-88-69250-12-5

EXAMINATION COMMITTEE

Prof. Enzo Bonora
University of Verona (Doctoral Program Coordinator)

Prof. Cristina Giannattasio
University of Milano Bicocca (Component)

Prof. Marcello Govoni
University of Ferrara (Component)

Prof. Giovanni Targher
University of Verona (Component)

ACKNOWLEDGMENTS

The work presented in this thesis was conducted in part at the Division of Endocrinology and Metabolic Diseases, Department of Medicine, University of Verona Medical School and Hospital Trust of Verona, Verona, Italy, and in part at the General Medicine Division, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA.

The relentless supervision and the overall guidance provided by Prof. James B. Meigs (Harvard Medical School), Prof. Riccardo C. Bonadonna (University of Parma), and by Prof. Enzo Bonora (University of Verona) are gratefully acknowledged.

The individual contribution and financial disclosure statement of the research groups and collaborators at each participating Institution involved in the projects herein presented are listed in the original publication at the respective publisher's website.

The work of Dr. M. Dauriz was supported by the financial contribution provided by the University of Verona (2012 *CooperInt* Junior Researcher Program 2012; 2014-2015 Research Fellowship) and by the Autonomous Province of Trento (2009-2014 Fellowship Program in Endocrinology, Diabetes and Metabolism).

TYPE 2 DIABETES and ASSOCIATED CARDIOMETABOLIC DISORDERS

Marco Dauriz

Certainty in science went out of fashion early last century following Heisenberg's postulates.
We now see ranges and confidence intervals around any given mode, mean, or median.
Truth, in short, is a dependent variable.

R. David G. Leslie & Eric S. Kilpatrick
Diabetes Care, 2009 Jan;32(1):e11

To my family

TABLE OF CONTENTS

List of Publications

PUBLICATIONS AND MANUSCRIPTS BASED ON THE STUDIES DESCRIBED IN THIS THESIS

Chapter 2

Marco Dauriz, James B. Meigs

Current Insights into the Joint Genetic Basis of Type 2 Diabetes and Coronary Heart Disease

Curr Cardiovasc Risk Rep (2014) 8(1):368 – PMID: 24729826

Chapter 3

Marco Dauriz, Bianca C. Porneala, Xiuqing Guo, Lawrence F. Bielak, Patricia A. Peyser, Nefertiti H. Durant, Mercedes R. Carnethon, Riccardo C. Bonadonna, Enzo Bonora, Donald W Bowden, Jose C. Florez, Myriam Fornage, Marie-France Hivert, David R. Jacobs Jr, Edmond K. Kabagambe, Cora E. Lewis, Joanne M Murabito, Laura J. Rasmussen-Torvik, Stephen S. Rich, Jason L. Vassy, Jie Yao, Jeffrey J Carr, Sharon L. Kardia, David Siscovick, Christopher J. O'Donnell, Jerome I. Rotter, Josée Dupuis, James B. Meigs

Association of a 62 Variants Type 2 Diabetes Genetic Risk Score With Markers of Subclinical Atherosclerosis: A Transethnic, Multicenter Study

Circ Cardiovasc Genet. 2015 Jun;8(3):507-15 – PMID: 25805414

Chapter 4

Maddalena Trombetta[*], *Marco Dauriz*[*], Sara Bonetti[*], Daniela Travia, Maria Linda Boselli, Lorenza Santi, Enzo Bonora, Riccardo C Bonadonna

[*] Equal contribution.

Is common genetic variation at IRS1, ENPP1 and TRIB3 loci associated with cardio metabolic phenotypes in type 2 diabetes? An exploratory analysis of the Verona Newly Diagnosed Type 2 Diabetes Study

Nutr Metab Cardiovasc Dis. 2016 Mar;26(3):232-8

Chapter 5

Marco Dauriz, Bianca C Porneala, Josée Dupuis, Joanne M Murabito, Adrianne L Cupples, Jose C Florez, James B Meigs

A Genetic Risk Score of 96 Variants linked with Type 2 Diabetes and Cardiometabolic Risk Traits is Associated with Cardiovascular Mortality in 29-years Follow-up of the Framingham Heart Study

In preparation

OTHER RELEVANT PUBLICATIONS

Vassy JL, Hivert MF, Porneala B, *Dauriz M*, Florez JC, Dupuis J, Siscovick DS, Fornage M, Rasmussen-Torvik LJ, Bouchard C, Meigs JB.

Polygenic type 2 diabetes prediction at the limit of common variant detection

Diabetes. 2014 Jun;63(6):2172-82

Walford GA, Porneala BC, *Dauriz M*, Vassy JL, Cheng S, Rhee EP, Wang TJ, Meigs JB, Gerszten RE, Florez JC

Metabolite traits and genetic risk provide complementary information for the prediction of future type 2 diabetes

Diabetes Care. 2014 Sep;37(9):2508-14

Cornes BK, Brody JA, Nikpoor N, Morrison AC, Dang HC, Ahn BS, Wang S, *Dauriz M*, Barzilay JI, Dupuis J, Florez JC, Coresh J, Gibbs RA, Kao WH, Liu CT, McKnight B, Muzny D, Pankow JS, Reid JG, White CC, Johnson AD, Wong TY, Psaty BM, Boerwinkle E, Rotter JI, Siscovick DS, Sladek R, Meigs JB

Association of levels of fasting glucose and insulin with rare variants at the chromosome 11p11.2-MADD locus: Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium Targeted Sequencing Study

Circ Cardiovasc Genet. 2014 Jun;7(3):374-82 – PMID: 24951664

Wessel J, Chu AY, Willems SM, Wang S, Yaghootkar H, Brody JA, *Dauriz M*, Hivert MF, Raghavan S, Lipovich L, Hidalgo B, Fox K, Huffman JE, An P, Lu Y, Rasmussen-Torvik LJ, Grarup N, Ehm MG, Li L, Baldridge AS, Stančáková A, Abrol R, Besse C, Boland A, Bork-Jensen J, Fornage M, Freitag DF, Garcia ME, Guo X, Hara K, Isaacs A, Jakobsdottir J, Lange LA, Layton JC, Li M, Hua Zhao J, Meidtner K, Morrison AC, Nalls MA, Peters MJ, Sabater-Lleal M, Schurmann C, Silveira A, Smith AV, Southam L, Stoiber MH, Strawbridge RJ, Taylor KD, Varga TV, Allin KH, Amin N, Aponte JL, Aung T, Barbieri C, Bihlmeyer NA, Boehnke M, Bombieri C, Bowden DW, Burns SM, Chen Y, Chen YD, Cheng CY, Correa A, Czajkowski J, Dehghan A, Ehret GB, Eiriksdottir G, Escher SA, Farmaki AE, Frånberg M, Gambaro G, Giulianini F, Goddard WA 3rd, Goel A, Gottesman O, Grove ML, Gustafsson S, Hai Y, Hallmans G, Heo J, Hoffmann P, Ikram MK, Jensen RA, Jørgensen ME, Jørgensen T, Karaleftheri M, Khor CC, Kirkpatrick A, Kraja AT, Kuusisto J, Lange EM, Lee IT, Lee WJ, Leong A, Liao J, Liu C, Liu Y, Lindgren CM, Linneberg A, Malerba G, Mamakou V, Marouli E, Maruthur NM, Matchan A, McKean-Cowdin R, McLeod O, Metcalf GA, Mohlke KL, Muzny DM, Ntalla I, Palmer ND, Pasko D, Peter A, Rayner NW, Renström F, Rice K, Sala CF, Sennblad B, Serafetinidis I, Smith JA, Soranzo N, Speliotes EK, Stahl EA, Stirrups K, Tentolouris N, Thanopoulou A, Torres M, Traglia M, Tsafantakis E, Javad S, Yanek LR, Zengini E, Becker DM, Bis JC, Brown JB, Cupples LA, Hansen T, Ingelsson E, Karter AJ, Lorenzo C, Mathias RA, Norris JM, Peloso GM, Sheu WH, Toniolo D, Vaidya D, Varma R, Wagenknecht LE, Boeing H, Bottinger EP, Dedoussis G, Deloukas P, Ferrannini E, Franco OH, Franks PW, Gibbs RA, Gudnason V, Hamsten A, Harris TB, Hattersley AT, Hayward C, Hofman A, Jansson JH, Langenberg C, Launer LJ, Levy D, Oostra BA, O'Donnell CJ, O'Rahilly S, Padmanabhan S, Pankow JS, Polasek O, Province MA, Rich SS, Ridker PM, Rudan I, Schulze MB, Smith BH, Uitterlinden AG, Walker M, Watkins H, Wong TY, Zeggini E; EPIC-InterAct Consortium, Laakso M, Borecki IB,

Chasman DI, Pedersen O, Psaty BM, Tai ES, van Duijn CM, Wareham NJ, Waterworth DM, Boerwinkle E, Kao WH, Florez JC, Loos RJ, Wilson JG, Frayling TM, Siscovick DS, Dupuis J, Rotter JI, Meigs JB, Scott RA, Goodarzi MO

Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility

Nat Commun. 2015 Jan 29;6:5897 – PMID: 25631608

Willems SM, Cornes BK, Brody JA, Morrison AC, Lipovich L, *Dauriz M*, Chen Y, Liu CT, Rybin DV, Gibbs RA, Muzny D, Pankow JS, Psaty BM, Boerwinkle E, Rotter JI, Siscovick DS, Vasan RS, Kaplan RC, Isaacs A, Dupuis J, van Duijn CM, Meigs JB

Association of the IGF1 gene with fasting insulin levels.

Eur J Hum Genet. 2016 Feb 10. doi: 10.1038/ejhg.2016.4 – PMID: 26860063

Dauriz M, Meigs JB

The power of numbers.

Invited Commentary - Diabetologia. 2016 Apr 26 – PMID: 27115413

Buzzetti R, Prudente S, Copetti M, *Dauriz M*, Simona Zampetti, Garofalo M, Penno G, Trischitta V

Genetic prediction of common forms of diabetes mellitus and related chronic complications. Still much work to do.

Under submission

Chapter 1

Preface / Prefazione

Preface

Since the earliest steps in this Doctoral Program I envisioned this journey as consisting of three steps: (1) to become comfortable with the state-of-the-art of genetic epidemiology and get organizational skills to correctly and productively frame ideas, manage datasets and biobanks and to deal with international research consortia; (2) to learn how to compellingly write research proposals and scientific reports, discuss and share results, critically review and support the work of other research groups involved in the field of diabetes, obesity and cardiovascular disease from a wide variety of perspectives (from wet lab to clinical epidemiology); (3) to finalize for publication original research papers on the topics outlined above.

To this end, in 2013-2014 I had the opportunity to get a formal training in scientific methodology and writing, epidemiology, medical and population genetics at Harvard University in Boston. This experience led to the publication of several papers in high-impact scientific journals and continuing collaboration with the Framingham Heart Study group, MAGIC and CHARGE Consortia is still fruitfully ongoing. Since my arrival back in Italy in July 2014 I have continued to work on the projects I started in Boston, while disseminating scientific knowledge by actively attending to several national and international meetings.

The unprecedented experience accrued while working with the Framingham Investigators and other international leaders in the field of (genetic) epidemiology led to create the overarching framework of my current activity, which aims at bringing together the genetic determinants of diabetes, obesity, cardiovascular disease and intermediate traits and to understand their relationship with the pathophysiology of the glucose-insulin system.

T2D is a complex disease characterized by a high prevalence and incidence worldwide, and recognizes genetic and non-genetic (environmental) risk factors as underlying determinants. CVD are currently one of the leading causes of death and are also often clinically associated to T2D. Recent large-scale genome-wide association studies (GWAS) have identified common genetic risk variants associated with a higher propensity of developing T2D, CVD and intermediate cardiometabolic phenotypes.

The goal of the research project herein presented was three-fold: (1) to critically revise the available literature about the genetic determinants of type 2 diabetes (T2D), coronary heart disease (CHD) and intermediate phenotypes (sub-diabetic hyperglycemia, measures of subclinical atherosclerosis (SCA) and associated risk conditions), aimed at searching for potential overlapping areas of shared genetic background; (2) to verify whether the genetic determinants of T2D, and particularly those associated with insulin resistance, are also associated with measures of SCA; (3) to verify whether a genetic risk score comprised of the genetic determinants of T2D, myocardial infarction, stroke, atrial fibrillation, sudden cardiac death, coronary heart disease, is associated with an excess risk of all-cause mortality and/or CVD death.

In detail, the present research exercise aimed at exploring the common genetic background of T2D, CVD and sub-diabetic forms of hyperglycemia by means of three exemplifying studies herein outlined. The first study verified whether the genetic risk for T2D, as represented by the aggregate burden of T2D risk loci (either as a whole or by distinct functional sub-groups, representative of loci with prior evidence of association with defective beta-cell function and/or increased insulin resistance), is associated with SCA traits in multi-ethnic cohorts.

The second study verified the hypothesis that the common genetic variability at loci gatekeepers of the insulin signaling transduction pathway are associated with

insulin resistance, beta-cell dysfunction, pathologic electrocardiogram, and/or increased SCA in patients affect by newly-diagnosed T2D.
The third study verified whether the composite of the genetic determinants of T2D and intermediate CVD risk traits is associated with a higher mortality in the Framingham Offspring Study.

Prefazione

Il presente progetto di ricerca si compone di tre parti: (1) revisione della letteratura relativa ai determinanti genetici di diabete mellito tipo 2, malattie coronariche e fenotipi intermedi (forme sub-diabetiche di iperglicemia, forme subcliniche di aterosclerosi e fattori di rischio associati) alla ricerca di possibili aree di sovrapposizione; (2) verificare se i determinanti di rischio genetico per diabete tipo 2, ed in particolare quelli maggiormente associati a insulino-resistenza, sono anche associati a misure di aterosclerosi subclinica; (3) verificare se uno score di rischio genetico costituito dai determinanti genetici di diabete tipo 2, infarto miocardico, stroke, fibrillazione atriale, morte cardiaca improvvisa, malattie coronariche è associato a mortalità per tutte le cause e/o mortalità per malattie cardiovascolari.

Il diabete mellito di tipo 2 (T2D) è una malattia complessa ad alta prevalenza e incidenza che riconosce fattori genetici e non-genetici quali determinanti causali. Le malattie cardiovascolari (CVD) sono una delle maggiori cause di morte e sono spesso associate a T2D. Studi di associazione genome-wide hanno identificato varianti genetiche comuni associate a T2D, CVD e fenotipi cardiometabolici intermedi.

Questo percorso di ricerca si è proposto di individuare le basi genetiche comuni a T2D, CVD e forme sub-diabetiche di iperglicemia attraverso tre studi esemplificativi. Nel primo studio è stato verificato se il rischio genetico per T2D sia associato, in aggregato e/o in sottogruppi funzionali distinti (disfunzione beta-cellulare o insulino-resistenza), a tratti di aterosclerosi subclinica (ATS) in coorti multi-etniche. Il secondo studio ha testato l'ipotesi che la variabilità genetica comune dei loci principalmente coinvolti nella trasduzione del segnale insulinico siano associati a insulino-resistenza, funzione beta-cellulare, anomalie elettrocardiografiche e/o aterosclerosi subclinica in soggetti affetti da T2D neo-diagnosticato. Nel terzo studio è stato indagato se il rischio genetico per T2D e tratti di rischio cardiometabolico sia associato ad aumentata mortalità nel Framingham Offspring Study.

Chapter 2

Current Insights into the Joint Genetic Basis of Type 2 Diabetes and Coronary Heart Disease

Curr Cardiovasc Risk Rep (2014) 8(1):368

Authors and affiliations are listed in Chapter 6.1

2.1 ABSTRACT

English - The large-scale genome-wide association studies conducted so far identified numerous allelic variants associated with type 2 diabetes (T2D), coronary heart disease (CHD) and related cardiometabolic traits. Many T2D- and some CHD-risk loci are also linked with metabolic traits that are hallmarks of insulin resistance (lipid profile, abdominal adiposity). 9p21.3 and 2q36.3, are the most consistently replicated loci appearing to share genetic risk for both T2D and CHD. Although many glucose- or insulin-related trait variants are also linked with T2D risk, none of them is associated with CHD. Hence, while T2D and CHD are strongly clinically linked together, further ongoing analyses are needed to clarify the existence of a shared underlying genetic signature of these complex traits. The present review summarizes an updated picture of T2D-CHD genetics as of 2013, aiming to provide a platform for targeted studies dissecting the contribution of genetics to the phenotypic heterogeneity of T2D and CHD.

Italian - Gli studi su larga scala del genoma hanno sinora identificato numerose varianti alleliche associate a diabete mellito tipo 2 (DMT2), malattie coronariche (CHD) e a fenotipi intermedi di rischio cardiometabolico. Molti loci associati a DMT2 ed alcuni associati a CHD sono anche associati a tratti metabolici caratteristici dell'insulino-resistenza (profilo lipidico, adiposità addominale). I loci 9p21.3 e 2q36.3 sono stati più volte identificati quali determinanti genetici di rischio all'intersezione tra DMT2 e CHD. Benchè tuttavia molte varianti genetiche associate a glicemia, insulinemia e fenotipi ad essi correlati siano anche associate a un aumentato rischio per DMT2, nessuna è risultata associata a CHD. Pertanto, benchè sia evidente che DMT2 e CHD sono fortemente associati clinicamente, ulteriori studi sono necessari per chiarire se questi complessi fenotipi riconoscano una base genetica comune. In questo riassunto viene fornito

un quadro aggiornato delle conoscenze relativa ai determinanti genetici di DMT2 e CHD e si propone di fornire una base per esplorare l'eterogeneità fenotipica che caratterizza DMT2 e CHD.

Abbreviations

CARDIoGRAM	Coronary ARtery DIsease Genome wide Replication and Meta-analysis Consortium
C4D	Coronary Artery Disease Genetics Consortium
DIAGRAMv3	DIAbetes Genetics Replication and Meta-analysis
MAGIC	Meta-Analyses of Glucose- and Insulin-related traits Consortium
WTCC	Wellcome Trust Case Control Consortium
GWAS	Genome Wide Association Study
T2D	Type 2 Diabetes
CHD	Coronary Heart Disease
MI	Myocardial Infarction
LD	Linkage Disequilibrium
SNP	Single Nucleotide Polymorphism
MAF	Minor Allele Frequency
BMI	Body Mass Index
WHR	Waist-to-Hip Ratio
HOMA-B	Homeostatic Model Assessment of Beta-Cell Function
HOMA-IR	Homeostatic Model Assessment of Insulin Resistance
HUVEC	HUman Vascular Endothelial Cells

2.2 INTRODUCTION

The global epidemic of type 2 diabetes (T2D) and associated cardiovascular diseases is increasing tremendously despite great efforts in prevention and treatment [1]. Cardiovascular diseases, especially coronary heart disease (CHD), represent the leading cause of death worldwide [2] and alarming projections for upcoming years require new and more effective strategies [3].

Better understanding of mechanisms underlying disease etiology and disease pathogenesis is the *sine qua non* to move forward and is a major goal of recent genetic studies on T2D and CHD [4]. Both T2D and CHD constitute the paradigm of common complex traits and have been an exciting and highly productive arena in the field of genetics: the last decade witnessed an impressive growth of available information about the genetic architecture of T2D and CHD. Interestingly, the growing amount of available information has revealed many apparently overlapping genetic signals that share association with T2D and CHD, especially in and near chromosome 9p21.3 [5-9] and 2q36.3 [10, 11], and at several other loci harboring variants associated with fasting glucose or insulin and other cardiometabolic traits (for instance, levels of lipids and anthropometric measures) that increase risk for CHD and/or T2D [12, 13].

The present review will outline and discuss the results from large-scale association analyses for T2D [14], CHD [15] and glycaemic traits [12] published in the last year (2012-2103), and integrate the evidence on chromosomal regions at 9p21.3 and 2q36.3 loci to provide a plausible, though not exhaustive, explanation at the genetic level of the common soil underlying CHD, T2D and associated metabolic traits.

2.2.1 Recent Type 2 Diabetes Genome-Wide Association Studies (GWAS)

In 2012, the DIAbetes Genetics Replication and Meta-analysis Consortium published the largest to date association analysis for T2D (DIAGRAMv3) [14]. The study, combined with the 2011 genome-wide association study (GWAS, see the glossary in **Table 1**) of Cho *et al.* [16] in roughly 55,000 East Asians, brought to 65 the number of independent T2D susceptibility loci (**Table 2**), thus further extending an effort begun a few years ago [17] to unveil the common allelic architecture of T2D. The strategy took advantage of the experience accumulated in the field of GWAS and the availability of the Metabochip custom array [18] for cost-effective follow-up genotyping. The case-control, two-stage DIAGRAMv3 meta-analysis was conducted in nearly 150,000 subjects (34,840 T2D cases and 114,981 controls) mostly of European ancestry from 38 independent cohorts. The study found 10 new T2D variants of modest effect size in or near *ZMIZ1*, *ANK1*, *KLHDC5*, *TLE1*, *ANKRD55*, *CILP2*, *MC4R*, *BCAR1*, *HMG20A* and *GRB14*. Linkage disequilibrium (LD) analysis and previous reports showed that the lead SNP at many of these loci was also associated with T2D-related metabolic traits that overlap CHD risk factors such as body-mass index (BMI), waist circumference, and insulin resistance (*MC4R*), triglyceride concentration (*MC4R*, *CILP2*), waist-to-hip ratio (WHR) (*GRB14*), HDL-cholesterol (*GRB14*, *CILP2*) and total-cholesterol (*CILP2*). Interestingly, as clearly shown in **Figure 1** and thoroughly detailed in the following sections, there is also compelling evidence that specific T2D loci on chromosome 2q36.3 and 9p21.3 harbor allelic variants in close proximity to each other and marking genomic regions associated with increased CHD risk.

2.2.2 Recent Coronary Heart Disease GWAS

As detailed in **Table 2** the number of loci currently known to be associated with coronary heart disease at genome-wide significance level have reached 45, thanks to the joint effort undertaken by the CARDIoGRAM-C4D Consortium on a sample of nearly 200,000 individuals (63,746 CHD cases and 130,681 controls in Stage1 + Stage2) [15]. This study, published in early 2013, confirmed previous findings [11, 19], discovered 15 new genome-wide significant loci and tested them by a thorough association analysis with traditional CHD risk factors. Twelve loci (*APOB*, *ABCG5-ABCG8*, *PCSK9*, *SORT1*, *ABO*, *LDLR*, *APOE* and *LPA*) showed genome-wide significance for association with at least one lipid trait in the expected direction. The CHD-raising allele was also associated with abnormal lipid levels, the strongest association being with LDL-cholesterol; *CYP17A1-NT5C2*, *SH2B3*, *GUCY1A3*, *FES* and *ZC3HC1* were associated with blood pressure; *CYP17A1-CNNM2-NT5C2* and *RAI1-PEMT-RASD1* loci were associated with BMI and WHR. Notably, there was no overlap with specific T2D or glycaemic trait-associated variants (fasting insulin, fasting plasma glucose, HOMA-B and HOMA-IR) for any of the SNPs analyzed (**Figure 1**).

Taken together, the overall spectrum of 65 T2D and 45 CHD genome-wide associated common variants explain only a small fraction (~10% each) of disease heritability, thus leaving a large unfilled space under the umbrella of the common variant/common disease hypothesis [20]. Indeed, a great proportion of common genetic variance is predicted to occur in non-coding regions at the level of structural variation, such as deletions, insertions, inversions and copy number variants, which might be imperfectly tagged or under-represented in current GWAS arrays [21]. Large scale sequencing studies currently underway may help to fill in some of the unfilled space under the umbrella of the genetic basis of T2D

and CHD by identifying less common or regulatory variants underlying these diseases.

2.2.3 Recent Glycaemic Quantitative Traits GWAS

Valuable details concerning quantitative risk factors were added to the overall picture in 2012 by two large-scale association analyses from the Meta-Analyses of Glucose and Insulin-related traits (MAGIC) Consortium [12, 22] that further enlightened our understanding of the genetic determinants of overlapping risk factors for T2D and CHD (see **Table 3**).

The joint meta-analysis by Manning *et al.* [22] in nearly 100,000 non-diabetic subjects of European ancestry investigated the genetic variability of insulin resistance by testing on a genome-wide basis the interaction of body mass index with fasting glucose and insulin. Based on previous experience from MAGIC [23] a new computational approach accounting for potential interactions between BMI and genetic variants was applied, enabling the discovery of 13 previously unknown SNPs associated with fasting insulin (FI) or fasting glucose (FG) at genome-wide significance. Among the FI-loci, the lead SNP in or near *IRS1*, *COBLL1-GRB14*, *PDGFC* or *LYPLA1* was also associated with an increased risk for T2D (**Figure 1, Table 3**), the strongest signal being for the chr2q36.3-*IRS1* locus (rs2943634). Notably, as detailed in **Table 3**, the risk allele of most of the FI-SNPs identified were also associated with metabolic phenotypes related to insulin resistance and CHD risk (for instance, detrimental lipid profile, higher WHR). None of the FG-loci showed association with any insulin resistance-cardiometabolic trait, and only *ARAP1* was associated with T2D (**Table 3**).

These results are complementary to the GWAS conducted by Scott *et al.* [12], which identified 41 previously undiscovered [23, 24] glycaemic associations in up

to 133,010 non-diabetic individuals of European descent by combining previous discovery MAGIC data with newly Metabochip-genotyped samples. Scott *et al.* and Manning *et al.* jointly raised the number of non-overlapping loci influencing glycaemic traits (FI, FG, post-challenge glucose concentration) to 55 (53 confirmed loci in Scott *et al.* plus 2 additional and potentially independent signals from Manning *et al.*, associated with FG and lying, respectively, in or near *OR4S1* and *DPSYL5* genes); 34 of them are also at least nominally associated with increased T2D risk (**Figure 1**), and most of the FI-raising loci showed directionally consistent associations with abdominal obesity and/or higher triglycerides-to-HDL cholesterol ratio (Tg/HDL) (**Table 3**).

2.3 THE CHROMOSOME 2q36.3-*IRS1* LOCUS

The evidence described above suggests that loci associated with signatures of insulin resistance are fairly good candidates mechanistically linking the overlap between T2D, CHD and glycaemic quantitative traits. As pointed out in **Figure 1** and **Figure 2**, one of the most promising regions is a large locus spanning ~593 kb located on chromosome 2q36.3 and harboring the *IRS1* gene, a key mediator along the insulin signaling pathway. Over the past few years many large-scale association studies from different research groups including Manning *et al.* and Scott *et al.* led to the identification of a cluster of SNPs (rs2943634, rs2043640, rs2943641, rs2943650, rs2972146, rs2943645) in high LD with each other ($0.75<r^2<1.00$; 1000 Genomes Pilot 1 CEU population) and associated with T2D, CHD, increased FI, higher Tg/HDL and/or low subcutaneous-to-visceral fat ratio [11-14, 22, 25, 26]. A recent basic science report by Li *et al.* [27] also clarified that these variants are located in two major sites ~600 kb and ~1 Mb downstream from the *IRS1* gene promoter and might physically regulate *IRS1* gene expression

by looping interactions, explaining how putative regulatory regions far from *IRS1* might regulate insulin sensitivity. The variant rs2943634 deserves a special mention (**Figure 2A**) as the only one SNP discovered so far in 2q36.3 region directly associated with increased CHD risk –though at slightly below genome-wide significant ($p=1.61x10^{-7}$) by the WTCC and Cardiogenics Consortium GWAS effort in 2007 [11].

That said, since insulin resistance and its associated traits have also been proposed as common pathophysiological background underlying CHD risk and the diabetic atherogenic context [28], Lim *et al.* [29] early in 2013 further investigated whether the genetic variation at 2q36.3 locus might also affect CHD risk *via* subclinical atherosclerosis in a sample of 2740 Framingham Heart Study participants. The study examined the cluster of SNPs described above along with 195 additional genotyped or imputed SNPs in 2q36.3 locus, testing them for association with subclinical atherosclerosis traits, but failed to find any correlation, despite an adequate sample size and detailed phenotypic characterization. The only significant association between rs10167219 (r^2 with rs2943634 = 0.07) and ankle-brachial index (ABI) was not confirmed after a validation step in a larger ABI meta-analysis [30].

On the other hand, Bacci *et al.* [10] found that functional candidate variants of insulin signaling genes, including *IRS1* G972R (rs1801278) (regional plot shown in **Figure 2B**), *ENPP1* K121Q (rs1044498) and *TRIB3* Q84R (rs229549), summed in a genetic risk score (GRS), jointly nominally predicted a composite endpoint of incident cardiovascular events in a sample of 733 type 2 diabetic patients, The GRS was also associated with decreased insulin sensitivity, and functional analysis in human vascular endothelial cells (HUVEC) showed that the GRS was inversely related with insulin-stimulated nitric oxide synthase activity.

Hence, depending on the outcome measured, whether atherosclerotic plaque formation or coronary heart disease events, current insights on 2q36.3 locus are

still far from conclusive with much remaining to be understood at a mechanistic level.

2.4 THE CHROMOSOME 9p21.3 LOCUS

As shown in **Figure 1** and **Table 2**, only two of the 65 T2D genome-wide associated loci [14, 16] but none of the common variants at these loci clearly overlaps any of the 45 CHD loci [15]. The example provided by 9p21.3 locus, a large genomic region spanning ~53 kb, is paradigmatic in this sense, owing to its unique haplotype structure (**Figure 2B**). Notably, this locus is associated with both CHD and T2D in European ancestry individuals [5, 7-9] and also in Chinese Han individuals as shown in 2011 by Cheng *et al.* [6].

The 9p21.3 locus has been extensively studied over the past years and has been historically primarily linked with an increased risk of CHD and myocardial infarction [11, 31], as confirmed by the recent GWAS conducted by CARDIoGRAM-C4D Consortium [15]. The numerous CHD-associated SNPs identified thus far in this interval are characterized by high LD with each other, thus representing a distinct region robustly associated with CHD. In 2013 additional insights in the haplotype structure of this CHD-risk interval have become available. In a case-control study conducted in nearly 3,700 non-diabetic white subjects, Fan *et al.* [32] successfully showed that atherosclerotic plaque formation is determined by a set of allelic variants physically distinct from the haplotype that predicts MI, namely, vulnerable plaque rupture and thrombosis. The 9p21.3 locus does not house protein-coding genes; the closest, *CDKN2ABS1*, *CDKN2A/B*, and *ABO,* are 120 kb from the principal index SNPs at the locus. As well highlighted by a recent editorial by McPherson [33], a mechanistic explanation to unambiguously clarify the contribution of 9p21.3 CHD-associated

SNPs to atherosclerosis and MI is still missing. Long range regulatory interactions with distant coding regions, tissue-specific effects of 9p21.3 CHD susceptibility SNPs and interactions with inflammation have been hypothesized, [34], but current results are conflicting and a clear mechanistic model for the genetic effects at this locus remain to be identified [35, 36].

With respect to T2D risk, as found by Morris *et al.* [14], chromosome 9p21.3 also encompasses variants strongly associated with T2D (**Table 2**) and spatially arranged in a very tight genomic region adjacent but distinct from that harboring the CHD-associated SNPs. As shown in **Figure 2C-D**, it is well ascertained that the haplotype structure of 9p21.3 locus stands on two main regions or "blocks" [8]: one large segment spans roughly 44 kb and hosts the CHD LD region (lead-SNP: rs1333049); on the other side of a recombination peak lies a 4kb T2D-associated block (lead-SNP: rs10811661). The LD between the respective lead-SNPs of T2D and CHD blocks [15] is very low (r^2<0.009; 1000 Genomes Pilot 1 CEU population). The two regions have a low chance of mixing together during recombination, thus suggesting a distinct pattern of inheritance.

However, the DIAGRAMv3 GWAS identified an additional lead SNP at a putative independent secondary T2D signal (rs944801; r^2=0.01 with rs10811661) [14] within the CHD-haplotype block (**Figure 3C**). This T2D-associated SNP is in modest LD (r^2=0.35) with the CHD lead-SNP (rs1333049), thus indicating a potential region close to the CDKN2A/B genes jointly affecting CHD and T2D.

Functional studies to parse in depth the contribution, if any, to both T2D and CHD of this and other variants within 9p21.3 locus is a challenging task that is worth pursuing further.

2.5 SUMMARY

Large-scale GWAS have been a powerful tool to uncover common genetic signatures strongly associated with common complex diseases like T2D, CHD and associated cardiometabolic traits. Here we reviewed the most recent findings in this field, highlighting the hitherto confirmed overlapping associations among T2D, CHD and glycaemic trait susceptibility loci.

The papers in the last year by the DIAGRAM, Cardiogram-C4D and MAGIC consortia showed that a few GWAS-discovered loci overlap both T2D and CHD risk, and for quantitative traits, a larger fraction of glycaemic trait raising alleles are also associated with T2D risk and CHD quantitative risk factors. In particular, FI-raising alleles show a directionally consistent link with increased T2D risk and adverse lipid and anthropometric measures. These results suggest that many FI-associated loci represent insulin resistance loci that potential provide a genetic underpinning for joint T2D-CHD risk. The 2q36.3-*IRS1* locus in particular has emerged as a crossroad for signals associated with T2D-CHD risk. However, a firm and comprehensive functional explanation of the role played by 2q36.3-*IRS1* remains to be shown, especially towards CHD risk. For instance, 2q36.3 locus harbors variants that, taken together, seem to play heterogeneous genetic effects on atherosclerotic plaque formation/rupture [10, 29]. Interestingly, compelling evidence exists for the association between cardiovascular events and the candidate functional variant *IRS1* G972R (rs1801278) [10]. Unfortunately, this variant lacks of GWAS confirmation despite being quite common (MAF 5.4%), probably because no proxy for rs1801278 mapping in or near other known variants in 2q36.3 locus is presently available in any available SNP data set. Thus, absence of evidence for a clear role of this variant is due to absence of evidence, not evidence of no role. Genotyping of this variant in large, independent samples is needed for firm confirmation of this coding variant's role in CHD risk.

The role of the 9p21.3 locus on T2D-CHD risk needs further elucidation, as well. It has a peculiar haplotype structure organized in two contiguous but distinct blocks conferring risk, respectively, for T2D and CHD/MI. However there appears to be a variant, rs944801, that may be an independent secondary T2D signal amidst the CHD-haplotype block. Targeted confirmatory association and functional studies are needed to further investigate joint risk of T2D-CHD in this haplotype block.

2.6 IMPLICATIONS AND FUTURE DIRECTIONS

A number of possible confounding elements may explain why association results should be taken with, perhaps, a grain of salt [21, 37, 38]. First, as pointed out by Wray *et al*. [38], GWAS are capable, by design, to explain only a small fraction (currently 10%, on average) of disease heritability and are intrinsically underpowered to uncover the "missing inheritance" carried by rare and low-frequency variants; second, the nature of the association is essentially statistical and in most cases doesn't tell much about the functional effect, if any, of the SNPs identified, thus limiting the predictive power of the loci discovered so far [21]; third, most of the GWAS SNPs lie in non-coding DNA regions and might work as regulatory or chromatin-modulating variants with unknown distant *cis*/*trans* effect on gene expression [37]; and finally, possible limitations including imperfect tagging due to insufficiently dense SNP arrays cannot be excluded.

Another possibility is that diabetic and non-diabetic individuals might have distinct mechanisms of CHD risk. For instance, an increased burden of T2D-associated GWAS risk variants is associated with cardiovascular disease risk in individuals with T2D [39], but CHD risk at chromosome 9p21.3 is only raised in

T2D among those with elevated HbA1c levels [40], and the recently discovered variant on chromosome 1q25 associated with glutamic acid metabolism and CHD risk in T2D has not been observed in large scale non-diabetic CHD GWAS [41]. Further dissection of the joint genetic association of T2D and CHD versus the interaction of T2D on genetic risk for CHD will require additional careful untangling in large scale association studies and follow-up functional and physiological studies.

Future research might also focus on pleiotropy analyses of variants with less stringent evidence for genome-wide significance. For instance, as detailed in **Table 3**, the link between glycaemic trait raising alleles with lipids and BMI is physiologically consistent and statistically convincing for "true" associations, though in most cases not strong enough to reach $p<5x10^{8}$. Whether these loci that appear to be associated with more than one trait are true pleiotropic loci or more a function of the known trait correlations (that is, greater adiposity is a well-known correlate of insulin resistance) remains to be elucidated. In addition, studies that leverage extended genealogy [42] to catch more of the "missing heritability" and improve polygenic risk prediction [43] combined with targeted re-sequencing and fine-mapping studies of confirmed loci like 2q36.3 and 9p21.3 may also help to untangle the joint association of T2D-CHD [44].

Furthermore, increasing the prior probability to find "true" associations would be of paramount help. To this end it might be wise to focus on studies of carefully selected, deeply phenotyped population samples with *a priori* stronger genetic background like early-onset diabetes [45] or cohorts free of confounding factors like long standing (sub)diabetic hyperglycaemia [46]. The availability of detailed assessments of beta-cell function and insulin sensitivity (instead of surrogate markers) as well as the accessibility of tissue- and cell-repositories within these population samples will also provide the unique opportunity to mechanistically

unravel the genetic signature of T2D and/or CHD.

Greater understanding of the genetic associations underlying T2D-CHD risk in the setting of a global pandemic of T2D and CHD is a timely challenge for improved population health and the sustainability of healthcare systems. The tremendous abundance of discoveries made by large-scale association studies published in 2012-2013 now needs further translation into mechanistic insights and improved clinical practice. However, this promise for discoveries achieved in the field of diabetes and cardiometabolic disease genetics is becoming ever closer.

2.7 FIGURES and TABLES

Table 1 - Glossary of Unfamiliar Terms

SNP: A Single Nucleotide Polymorphism (SNP) is a single base-pair change in the DNA sequence and is a class of common human genetic variations [47]. A genetic variant is usually considered as "common" if its Minor Allele Frequency (MAF) is over 5%, i.e. the less frequently inherited allele on one of the two DNA strands has a prevalence over 5% in the population of interest.
Linkage Disequilibrium: The difference between the expected and the observed frequencies of two SNPs under the assumption of independence is a common way to determine and measure the structure of haplotypes in genetic linkage analysis. This probability is called linkage disequilibrium (LD) and is expressed as a correlation coefficient (r^2) between pairs of SNPs, with r^2 ranging between 0 and 1) The higher the r^2, the higher the probability that two SNPs are non-randomly inherited together during recombination.
GWAS: Genome-Wide Association Study. GWAS have become global scientific efforts begun over 10 years ago to analyze DNA sequence variations and to identify their possible association with common diseases by a hypothesis-free approach. For a general overview of basic principles, experimental design and overall computational strategy underlying GWAS we recommend the recent publication of Bush *et al.* [48].
Lead SNP: is the representative variant in a genomic region (or "locus") most significantly associated with the disease or trait in a GWAS. As a general agreement, the lead SNP "tags" (or is a "tag-SNP") an LD region and is named as being associated with the nearest known gene at the locus, if any. Further mapping and function studies are also required to determine if the lead SNP at a locus is actually associated with the named gene or has any molecular functional significance related to the disease or trait being studied.
Statistical significance in GWAS: Since currently available GWAS genotyping platforms allow to test millions of SNPs together against one or more traits of interest across thousands of individuals, the agreement of what has to be considered statistically significant (i.e. accepted as true association rather than happened by chance) takes into account the nominal Pearson's statistical significance threshold (0.05) and the number of apparently independent association tests in the human genome. It is estimated that in individuals of European ancestry there are about 1 million uncorrelated ("independent") common SNPs, hence, the resulting threshold is $5x10^{-8}$ or 0.05 divided by 1 million.
Pleiotropy: Describes a single genetic variant or multiple variants at the same locus that affect one or more phenotypic traits. If such genetic variation acts as possible underlying cause for an observed cross-phenotype association, then pleiotropy occurs [37].

GWAS[1,2,3], labelled according to the nearest the index SNP.

	65 T2D associated loci[1]			45+1 CHD associated loci[2,3]		
Chromosome	**SNP**	**Risk allele**	**Nearest gene**	**SNP**	**Risk allele**	**Nearest gene**
Chr. 1	rs10923931	T	*NOTCH2*	rs4845625	T	*IL6R*
	rs2075423	G	*PROX1*	rs11206510	T	*PCSK9*
				rs602633	C	*SORT1*
				rs17114036	A	*PPAP2B*
				rs17464857	T	*MIA3*
Chr. 2	rs243021	A	*BCL11A*	rs6544713	T	*ABCG5-ABCG8*
	rs780094	C	*GCKR*	rs515135	G	*APOB*
	rs13389219	C	*GRB14*	rs1561198	A	*VAMP5-VAMP8-GGCX*
	rs2943640	C	***IRS1 (2q36.3)***	rs6725887	C	*WDR12*
	rs7593730	C	*RBMS1*	rs2252641	G	*ZEB2-AC074093.1*
	rs11899863	C	*THADA*	rs2943634[3]	C	***2q36.3*** *(1.61x10^{-7})*
Chr. 3	rs6795735	C	*ADAMTS9*	rs9818870	T	*MRAS*
	rs11717195	T	*ADCY5*			
	rs4402960	T	*IGF2BP2*			
	rs1801282	C	*PPARG*			
	rs12497268	G	*PSMD6*			
	rs17301514	A	*ST64GAL1*			
	rs7612463	C	*UBE2E2*			
Chr. 4	rs6819243	T	*MAEA*	rs1878406	T	*EDNRA*
	rs1801214	T	*WFS1*	rs7692387	G	*GUCY1A3*
Chr. 5	rs459193	G	*ANKRD55*	rs7173743	T	*ADAMTS7*
	rs6878122	G	*ZBED3*	rs273909	C	*SLC22A4-SLC22A5*
Chr. 6	rs10440833	A	*CDKAL1*	rs10947789	T	*KCNK5*
	rs3734621	C	*KCNK16*	rs4252120	T	*PLG*
	rs4299828	A	*ZFAND3*	rs2048327	G	*SLC22A3-LPAL2-LPA*
				rs12190287	C	*TCF21*
				rs12205331	C	*ANKS1A*
				rs9369640	C	*PHACTR1*
Chr. 7	rs17168486	T	*DGKB*	rs12539895	A	*7q22*
	rs17867832	T	*GCC1*	rs2023938	G	*HDAC9*
	rs4607517	A	*GCK*	rs11556924	C	*ZC3HC1*
	rs849134	A	*JAZF1*			
	rs13233731	G	*KLF14*			
Chr. 8	rs516946	C	*ANK1*	rs264	G	*LPL*
	rs3802177	G	*SLC30A8*	rs2954029	A	*TRIB1*
	rs7845219	T	*TP53INP1*			
Chr. 9	rs10965250	G	***CDKN2A/B (9p21.3)***	rs579459	C	*ABO*
	rs10758593	A	*GLIS3*	rs1333049	C	***CDKN2BAS1 (9p21.3)***
	rs16927668	T	*PTPRD*	rs3217992	A	
	rs2796441	G	*TLE1*			
	rs13292136	C	*TLE4*			
Chr. 10	rs12779790	G	*CDC123/CAMK1D*	rs2505083	C	*KIAA1462*
	rs5015480	C	*HHEX/IDE*	rs501120	A	*CXCL12*
	rs7903146	T	*TCF7L2*	rs2047009	C	
	rs12242953	G	*VPS26A*	rs12413409	G	*CYP17A1, CNNM2, NT5C2*
	rs12571751	A	*ZMIZ1*	rs11203042	T	*LIPA*
				rs2246833	T	
Chr. 11	rs1552224	A	*ARAP1 (CENTD2)*	rs974819	A	*PDGFD*
	rs2334499	T	*DUSP8*	rs9326246	C	*ZNF259, APOA5, APOA1*
	rs5215	C	*KCNJ11*			
	rs163184	G	*KCNQ1*			
	rs10830963	G	*MTNR1B*			
Chr. 12	rs11063069	G	*CCND2*	rs3184504	T	*SH2B3*
	rs1531343	C	*HMGA2*			
	rs12427353	G	*HNF1A (TCF1)*			
	rs10842994	C	*KLHDC5*			
	rs4760790	A	*TSPAN8/LGR5*			
Chr. 13	rs1359790	G	*SPRY2*	rs9515203	T	*COL4A1, COL4A2*
				rs4773144	G	
				rs9319428	A	*FLT1*
Chr. 14				rs2895811	C	*HHIPL1*
Chr. 15	rs2028299	C	*AP3S2*	rs17514846	A	*FURIN-FES*
	rs4502156	T	*C2CD4A*			
	rs7177055	A	*HMG20A*			
	rs12899811	G	*PRC1*			
	rs11634397	G	*ZFAND6*			
Chr. 16	rs7202877	T	*BCAR1*			
	rs9936385	C	*FTO*			
Chr. 17	rs4430796	G	*HNF1B (TCF2)*	rs12936587	G	*RASD1, SMCR3, PEMT*
	rs2447090	A	*SRR*	rs2281727	C	*SMG6*
				rs15563	C	*UBE2Z*
Chr. 18	rs11873305	A	*MC4R*			
Chr. 19	rs10401969	C	*CILP2*	rs1122608	G	*LDLR*
	rs8108269	G	*GIPR*	rs445925	C	*ApoE-ApoC1*
	rs8182584	T	*PEPD*	rs2075650	G	
Chr. 20	rs4812829	A	*HNF4A*			
Chr. 21				rs9982601	T	*KCNE2*

Table 3 – Associations of 55 confirmed glycaemic loci with T2D and/or cardiometabolic traits.

Nearest gene(s)	Chr	Glycaemic traits $p<5x10^{-8}$			T2D $p<10^{-4}$	Lipids $p<10^{-4}$		Anthropometric measures $p<10^{-4}$	
		FI	FG	2hGlu		Tg	HDL-C	BMI	WHR
GRB14	2	+			+	+	-		+
IRS1	2	+			+	+	-		
PPARG	3	+			+		-		
ANKRD55-MAP3K	5	+			+	+			-
ARL15	5	+			+			-	
FTO	16	+			+		-	+	
PEPD	19	+			+		-		
LYPLAL1	1	+							+
YSK4	2	+							
TET2	4	+							
PDGFC	4	+					-		
FAM13A	4	+					-		
UHRF1BP1	6	+							
RSPO3	6	+							+
HIP1	7	+						+	
IGF1	12	+							
PPP1R3B	8	+	+	+			-		
GCKR	2	+	+	+	+	+			
TCF7L2	10	+	+	+	+				
IGF2BP2	3		+	+	+				
ADCY5	3		+	+	+				
GCK	7		+	+	+				
VPS13C-C2CD4A/B	15		+	+	+				
GIPR	19		+	+	+			-	
PROX1	1		+		+				
ZBED3	5		+		+				
CDKAL1	6		+		+				
DGKB-TMEM195	7		+		+				
SLC30A8	8		+		+				
CDKN2B	9		+		+				
GLIS3	9		+		+				
MTNR1B	11		+		+				
ARAP1	11		+		+				
KL	13		+		+				
TOP1	20		+		+				
DPSYL5[*]	2		+			-			
G6PC2	2		+						
AMT	3		+						
SLC2A2	3		+						
PCSK1	5		+						
RREB1	6		+						
GRB10	7		+						
IKBKAP	9		+						
DNLZ	9		+						
ADRA2A	10		+						
CRY2	11		+						
OR4S1[*]	11		+						
MADD	11		+						
FADS1	11		+			-	+		
GLS2	12		+						
P2RX2	12		+						
PDX1	13		+						
WARS	14		+						
FOXA2	20		+						
ERAP2	5			+					

Adapted from Scott RA *et al. Nat. Genet*. 2012 – PMID: 22885924 and from [*]Manning AK *et al. Nat. Genet.* 2012 – PMID: 22581228. The 55 loci harboring one or more allelic variants associated with glycaemic traits are shown according to the nearest known gene(s). +/-, effect direction of the glycemic trait raising allele; T2D, type 2 diabetes; Tg, triglycerides; HDL-C, HDL-cholesterol; BMI, body mass index (Kg/m^2); WHR, waist-to-hip ratio; FI, fasting insulin; FG, fasting glucose; 2hGlu, 2-hour post-challenge plasma glucose concentration.

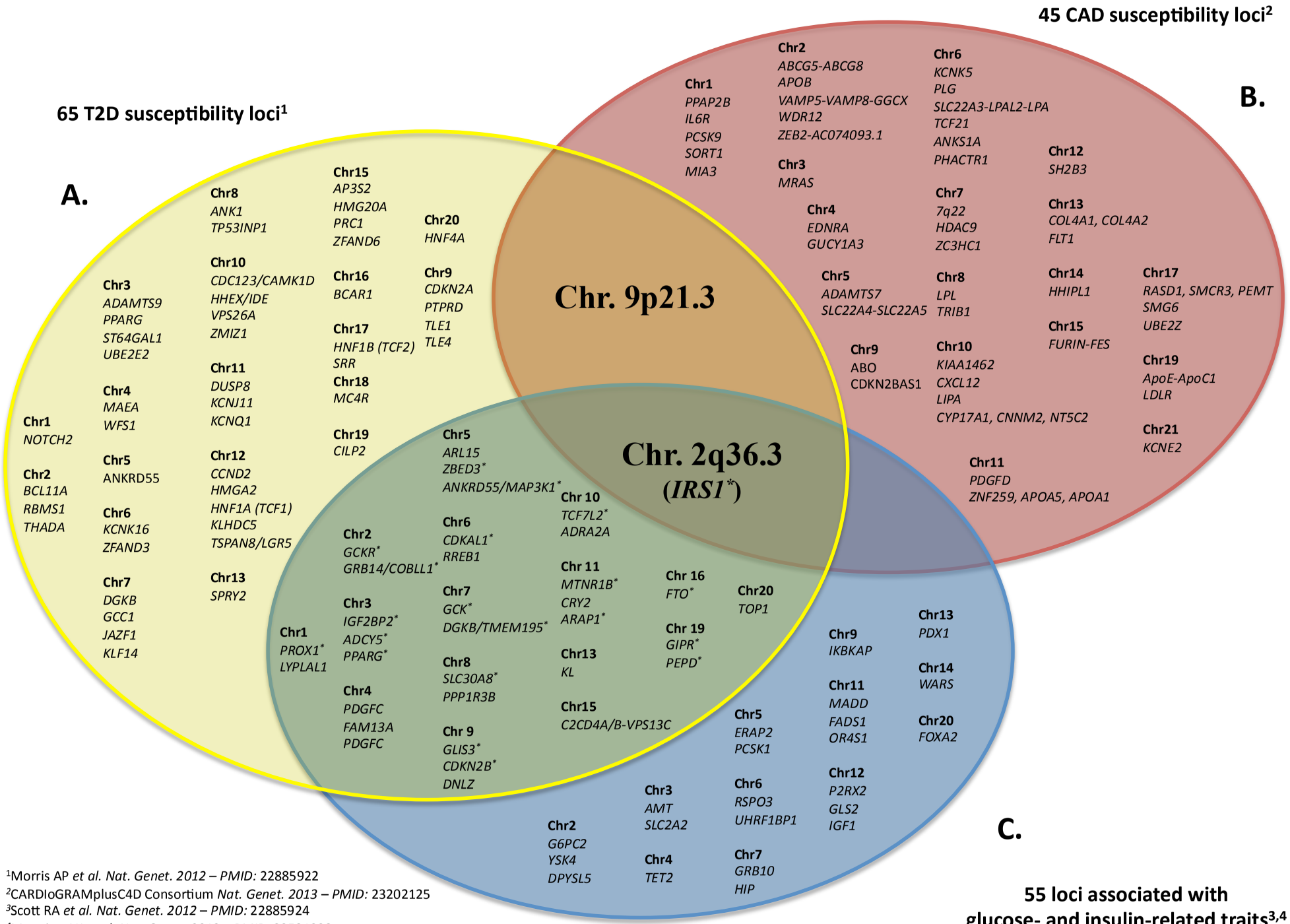
Figure 1. Overlapping associations among currently known T2D, CHD and glycaemic quantitative trait susceptibility loci from recent GWAS.

Loci harboring one or more common variant(s) associated with the phenotype or trait of interest are listed according to the nearest known gene. The diagram highlights the overlapping associations among A, B and C sets. Set **A**, coloured in yellow, comprises 65 confirmed type 2 diabetes (T2D) susceptibility loci, from [1]Morris A.P. *et al.* (DIAGRAMv3 Consortium) *Nat. Genet.* 2012 (PMID: 22885922) [14]. Set **B** (red) shows the 45 confirmed coronary heart disease (CHD) susceptibility loci from [2]Deloukas P. *et al.* (CARDIoGRAMplusC4D Consortium) *Nat. Genet.* 2013 (PMID: 23202125) [15]. Set **C** (blue) shows the 55 confirmed loci associated with glucose- and insulin-related traits (fasting glucose, fasting insulin, 2 hour post-challenge glucose), from [3]Scott R.A. *et al. Nat. Genet.* 2012 (PMID: 22885924) [12] and [4]Manning A.K. *et al. Nat. Genet.* 2012 (PMID: 22581228) [22].

The intersection between set A and set C comprises 34 loci associated with both T2D (at $p<0.05$ or lower) and glycaemic quantitative traits ($p<5x10^{-8}$); loci reaching genome wide significance for association with both T2D and quantitative traits are marked by an asterisk (*).

Chromosome 2q36.3- ***IRS1*** is a starred locus also linked with detrimental levels of other cardiometabolic traits (for instance, higher triglycerides-to-HDL cholesterol ratio or low subcutaneous-to-visceral fat ratio) and harboring a variant (rs2943634) strongly associated with increased CHD risk ($p=1.61x10^{-7}$, Samani NJ *et al. NEJM* 2007-PMID: 17634449) [11]. The chromosome, 2q36.3-*IRS1* locus, lying at the convergence of A, B and C sets, is a joint T2D_CHD locus.

Chromosome 9p21.3 is a locus at the intersection of A and C sets characterized by two contiguous but distinct haplotype blocks harboring variants associated with T2D or CHD and separated by a recombination peak. A potential overlap of a T2D SNP lying in the CHD block at 9p21.3 makes this locus a promising candidate for a shared genetic risk for both T2D and CHD.



[1]Morris AP *et al. Nat. Genet. 2012 – PMID:* 22885922
[2]CARDIoGRAMplusC4D Consortium *Nat. Genet. 2013 – PMID:* 23202125
[3]Scott RA *et al. Nat. Genet. 2012 – PMID:* 22885924
[4]Manning AK *et al. Nat. Genet. 2012 – PMID:* 22581228

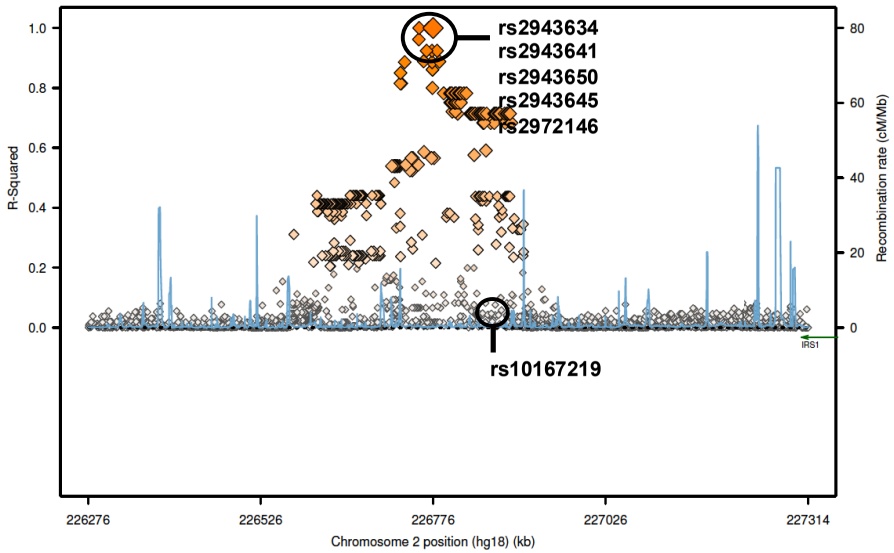
Figure 2: Linkage Disequilibrium Patterns Among Lead SNPs at Type 2 diabetes – Coronary Heart Disease Loci on Chromosomes 2 and 9

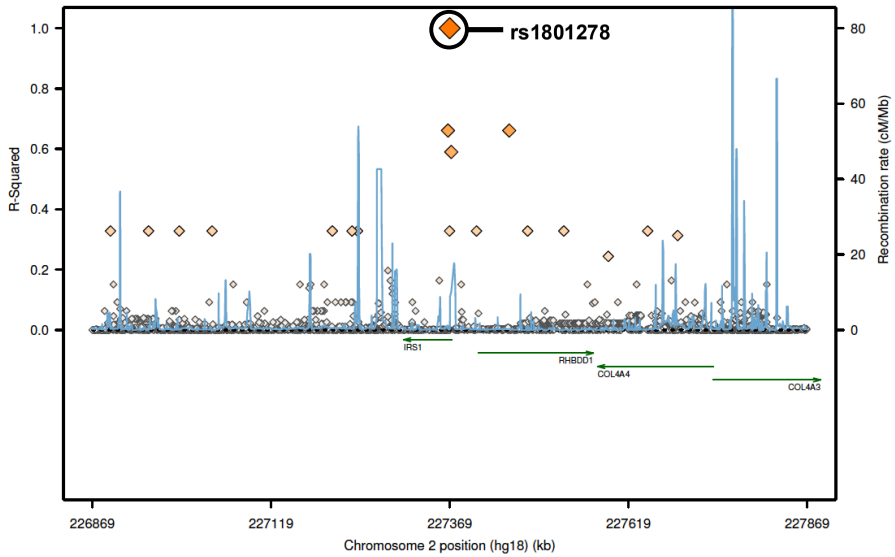
The left-hand y-axis of each panel indicates the linkage disequilibrium (LD), represented by the r^2 value, among single nucleotide polymorphisms (SNPs) at the locus, with the brightness of each point proportional to the r^2 value for that SNP. The right-hand y-axis indicates the recombination rate, plotted as the blue line, with high values indicating frequent recombination at that spot on the chromosomal position, plotted as the x-axis in each panel. LD data come from sequence-based SNP genotype data from the low-coverage sequencing pilot (Pilot 1) of the 1000 Genomes Project. This data set uses phased genotypes for 179 individuals from the HapMap CEU, YRI and JPT+CHB panels. Inter-SNP distances are measured in hg18 coordinates. Data were plotted using SNAP (Johnson A.D. *et al.*, *Bioinformatics* 2008-PMID: 18974171 [49]).

The top panel illustrates the chromosome 2q36.3 (left hand panel) and the *IRS1* (right hand panel) locus. At 2q36.3, the SNP rs2943634 is in high LD with SNPs rs2943641, rs2943650, rs2943645 and rs2972146 (associated with coronary heart disease (CHD), type 2 diabetes (T2D), fasting insulin, waist circumference and triglyceride/HDL cholesterol ratio, all r^2 >0.75), but in low LD with rs10167219 (ankle brachial index, r^2=0.05). At *IRS1*, SNP rs1801278 (CHD, insulin resistance) is ~593kb from rs2943634 and not in LD with any 2q36.3 SNP. Note the low LD and scarcity of SNPs in and around *IRS1*, indicating relatively high conservation (low variation) of base pairs around this important gene.

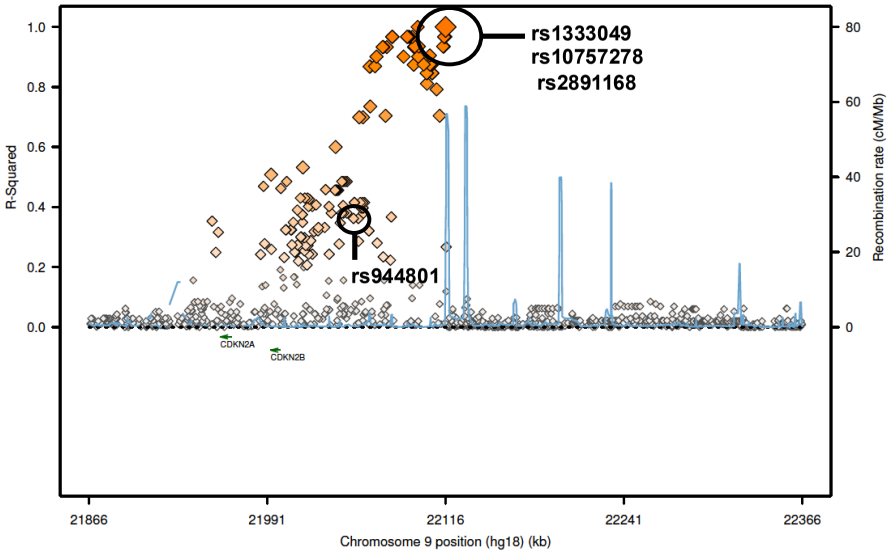
The bottom panel illustrates the chromosome 9p21.3-*CDKN2A/B* locus, with the region of SNPs associated with CHD (left hand panel) separated from the region of SNPs associated with T2D (right hand panel) by a large recombination peak (blue line). The lead SNP for CHD (rs1333049) is only ~8.6kb from but essentially uncorrelated with the lead SNP for T2D (rs10811661, r^2=0.009). However, a potential additional SNP associated with T2D, rs944801, lies in the CHD region and is modestly correlated with rs1333049 (r^2=0.35), indicating a potentially joint T2D – CHD genetic region upstream from the *CDKN2A/B* genes.

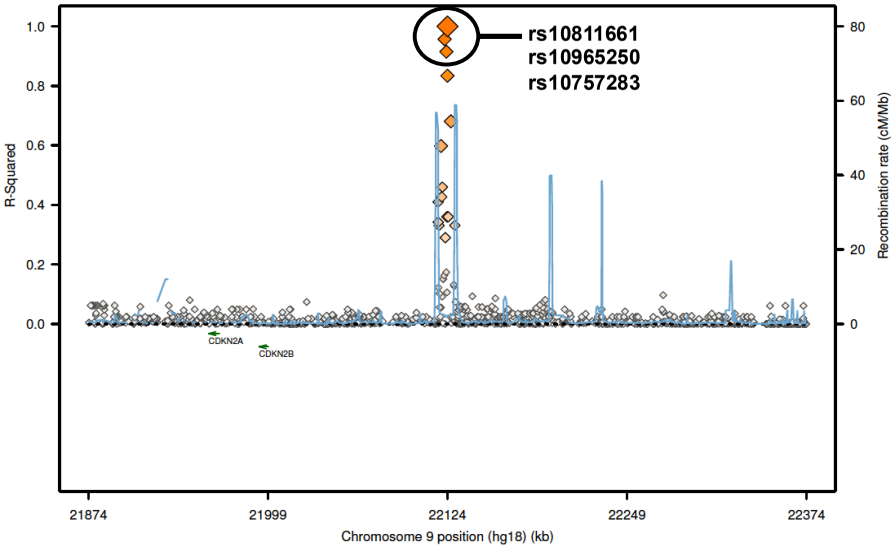
Chromosome 2q36.3 – *IRS1* Locus





Chromosome 9p21.6 – *CDKN2A/B* Locus





2.8 REFERENCES

1. **International Diabetes Federation.** IDF Diabetes Atlas 5th ed. 2011; Available from: http://www.idf.org/diabetesatlas.
2. **World Health Organization**, Global Atlas on Cardiovascular Disease Prevention and Control., ed. P.P. Mendis S, Norrving B editors, Geneva, Switzerland.
3. **World Health Organization**, World Health Statistics, Geneva, Switzerland: World Health Organization.
4. **Stranger, B.E.**, Stahl E.A., and Raj T., Progress and promise of genome-wide association studies for human complex trait genetics. Genetics, 2011. 187(2): p. 367-83.
5. **Broadbent, H.M.**, et al., Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. Hum Mol Genet, 2008. 17(6): p. 806-14.
6. **Cheng, X.**, et al., The same chromosome 9p21.3 locus is associated with type 2 diabetes and coronary artery disease in a Chinese Han population. Diabetes, 2011. 60(2): p. 680-4.
7. **Gori, F.**, et al., Common genetic variants on chromosome 9p21 are associated with myocardial infarction and type 2 diabetes in an Italian population. BMC Med Genet, 2010. 11: p. 60.
8. **Silander, K.**, et al., Worldwide patterns of haplotype diversity at 9p21.3, a locus associated with type 2 diabetes and coronary heart disease. Genome Med, 2009. 1(5): p. 51.
9. **Zeggini, E.**, et al., Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science, 2007. 316(5829): p. 1336-41.
10. **Bacci, S.**, et al., Joint effect of insulin signaling genes on cardiovascular events and on whole body and endothelial insulin resistance. Atherosclerosis, 2013. 226(1): p. 140-5.
11. **Samani, N.J.**, et al., Genomewide association analysis of coronary artery

disease. N Engl J Med, 2007. 357(5): p. 443-53.

12. **Scott, R.A.**, et al., Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet, 2012. 44(9): p. 991-1005.
13. **Rung, J.**, et al., Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. Nat Genet, 2009. 41(10): p. 1110-5.
14. **Morris, A.P.**, et al., Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet, 2012. 44(9): p. 981-90.
15. **Coronary Artery Disease Consortium**, et al., Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet, 2013. 45(1): p. 25-33.
16. **Cho, Y.S.**, et al., Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. Nat Genet, 2012. 44(1): p. 67-72.
17. **Zeggini, E.**, et al., Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet, 2008. 40(5): p. 638-45.
18. **Voight, B.F.**, et al., The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet, 2012. 8(8): p. e1002793.
19. **Peden, J.F.** and M. Farrall, Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. Hum Mol Genet, 2011. 20(R2): p. R198-205.
20. **Reich, D.E.** and E.S. Lander, On the allelic spectrum of human disease. Trends Genet, 2001. 17(9): p. 502-10.
21. **Manolio, T.A.**, Bringing genome-wide association findings into clinical use. Nat Rev Genet, 2013. 14(8): p. 549-58.
22. **Manning, A.K.**, et al., A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet, 2012. 44(6): p. 659-69.

23. **Dupuis, J.**, et al., New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet, 2010. 42(2): p. 105-16.
24. **Saxena, R.**, et al., Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat Genet, 2010. 42(2): p. 142-8.
25. **Kilpelainen, T.O.**, et al., Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. Nat Genet, 2011. 43(8): p. 753-60.
26. **Teslovich, T.M.**, et al., Biological, clinical and population relevance of 95 loci for blood lipids. Nature, 2010. 466(7307): p. 707-13.
27. **Li, G.**, et al., Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell, 2012. 148(1-2): p. 84-98.
28. **DeFronzo, R.A.**, Insulin resistance, lipotoxicity, type 2 diabetes and atherosclerosis: the missing links. The Claude Bernard Lecture 2009. Diabetologia, 2010. 53(7): p. 1270-87.
29. **Lim, S.**, et al., Common variants in and near IRS1 and subclinical cardiovascular disease in the Framingham Heart Study. Atherosclerosis, 2013. 229(1): p. 149-54.
30. **Murabito, J.M.**, et al., Association between chromosome 9p21 variants and the ankle-brachial index identified by a meta-analysis of 21 genome-wide association studies. Circ Cardiovasc Genet, 2012. 5(1): p. 100-12.
31. **Helgadottir, A.**, et al., A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science, 2007. 316(5830): p. 1491-3.
32. **Fan, M.**, et al., Two chromosome 9p21 haplotype blocks distinguish between coronary artery disease and myocardial infarction risk. Circ Cardiovasc Genet, 2013. 6(4): p. 372-80.
33. **McPherson, R.**, Chromosome 9p21.3 Locus for CAD: How little we know. J Am Coll Cardiol, 2013.
34. **McPherson, R.** and R.W. Davies, Inflammation and coronary artery disease: insights from genetic studies. Can J Cardiol, 2012. 28(6): p. 662-6.
35. **Harismendy, O.**, et al., 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. Nature, 2011. 470(7333): p. 264-8.

36. **Almontashiri, N.A.**, et al., Interferon-gamma activates expression of p15 and p16 regardless of 9p21.3 coronary artery disease risk genotype. J Am Coll Cardiol, 2013. 61(2): p. 143-7.
37. **Solovieff, N.**, et al., Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet, 2013. 14(7): p. 483-95.
38. **Wray, N.R.**, et al., Pitfalls of predicting complex traits from SNPs. Nat Rev Genet, 2013. 14(7): p. 507-15.
39. **Qi, Q.**, et al., Diabetes genetic predisposition score and cardiovascular complications among patients with type 2 diabetes. Diabetes Care, 2013. 36(3): p. 737-9.
40. **Doria, A.**, et al., Interaction between poor glycemic control and 9p21 locus on risk of coronary artery disease in type 2 diabetes. JAMA, 2008. 300(20): p. 2389-97.
41. **Qi, L.**, et al., Association between a genetic variant related to glutamic acid metabolism and coronary heart disease in individuals with type 2 diabetes. JAMA, 2013. 310(8): p. 821-8.
42. **Zaitlen, N.**, et al., Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS Genet, 2013. 9(5): p. e1003520.
43. **Chatterjee, N.**, et al., Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. Nat Genet, 2013. 45(4): p. 400-5, 405e1-3.
44. **Goldstein, D.B.**, et al., Sequencing studies in human genetics: design and interpretation. Nat Rev Genet, 2013. 14(7): p. 460-70.
45. **Morini, E.**, et al., IRS1 G972R polymorphism and type 2 diabetes: a paradigm for the difficult ascertainment of the contribution to disease susceptibility of 'low-frequency-low-risk' variants. Diabetologia, 2009. 52(9): p. 1852-7.
46. **Trombetta, M.**, et al., PPARG2 Pro12Ala and ADAMTS9 rs4607103 as "insulin resistance loci" and "insulin secretion loci" in Italian individuals. The GENFIEV study and the Verona Newly Diagnosed Type 2 Diabetes Study (VNDS) 4. Acta Diabetol, 2013. 50(3): p. 401-8.

47. **Frazer, K.A.**, et al., Human genetic variation and its contribution to complex traits. Nat Rev Genet, 2009. 10(4): p. 241-51.
48. **Bush, W.S.** and J.H. Moore, Chapter 11: Genome-wide association studies. PLoS Comput Biol, 2012. 8(12): p. e1002822.
49. **Johnson, A.D.**, et al., SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics, 2008. 24(24): p. 2938-9.

Chapter 3

Association of a 62 Variants Type 2 Diabetes Genetic Risk Score With Markers of Subclinical Atherosclerosis: A Transethnic, Multicenter Study

Circ Cardiovasc Genet. 2015 Jun;8(3):507-15

Authors and affiliations are listed in Chapter 6.1

3.1 ABSTRACT

English

Background - Type 2 diabetes (T2D) and cardiovascular disease (CVD) share risk factors and subclinical atherosclerosis (SCA) predicts events in those with and without diabetes. T2D genetic risk may predict both T2D and SCA. The hypothesis tested in this research project was that greater T2D genetic risk is associated with higher extent of SCA.

Methods and Results - In a cross-sectional analysis including up to 9,210 European Americans, 3,773 African Americans, 1,446 Hispanic Americans and 773 Chinese Americans without known CVD and enrolled in the FHS, CARDIA, MESA and GENOA studies, a 62 T2D-loci genetic risk score (GRS_{62}) was tested for association with measures of SCA, including coronary artery (CACS) or abdominal aortic calcium score (AACS), common (CCA-IMT) and internal (ICA-IMT) carotid artery intima-media thickness, and ankle-brachial index (ABI). Ancestry-stratified linear regression models were used, with random effects accounting for family relatedness when appropriate, applying a genetic-only (adjusted for sex) and a full SCA risk factors adjusted model (significance = $p < 0.01 = 0.05/5$, number of traits analyzed). An inverse association with CACS in MESA Europeans (fully adjusted p=0.004) and with CCA-IMT in FHS (p=0.009) was not confirmed in other study cohorts, either separately or in meta-analysis. Secondary analyses showed no consistent associations with β-cell and insulin resistance sub-GRS in FHS and CARDIA.

Conclusions - SCA appears unlikely to have a major genetic component linked to a burden of 62 T2D loci identified by large genome-wide association studies. A

shared T2D-CVD genetic basis, if any, might become apparent from the functional annotation of both T2D and CVD risk loci.

Italian

Premesse - Diabete mellito tipo 2 (T2D) e malattie cardiovascolari (CVD) sono fortemente associati clinicamente e condividono fattori di rischio che esercitano il loro effetto già a livello subclinico. Ampi studi di associazione "genome-wide" hanno individuato numerosi loci di rischio per T2D, molti dei quali associati in particolare a indici di funzione beta-cellulare o resistenza insulinica (IR). Esiste inoltre una certa sovrapposizione tra loci associati a iperglicemia subdiabetica e iperinsulinemia e fenotipi intermedi di IR che sono anche fattori di rischio per CVD. Non è tuttavia chiaro se la predisposizione genetica a sviluppare T2D agisca quale comune denominatore per lo sviluppo di T2D e aterosclerosi subclinica (SCA).

Obiettivo - In questo studio è stata testata l'ipotesi che un numero incrementale di varianti genetiche di rischio per T2D confermate dai più recenti studi di associazione genome-wide sia associato con aumentati valori di alcuni indici di SCA.

Disegno sperimentale - In un'analisi trasversale di 9,210 Europei Americani, 3,773 Africani Americani, 1,446 Ispanici Americani e 773 Cinesi Americani con anamnesi negativa per CVD ed arruolati negli studi FHS (Framingham Heart Study), CARDIA (Coronary Artery Risk Development in Young Adults), MESA (Multi-Ethnic Study of Atherosclerosis) e GENOA (Genetic Epidemiology Network of Atherosclerosis), è stata verificata l'associazione tra misure di SCA (calcium score coronarico e in aorta addominale, spessore intima-media in carotide comune ed interna, indice caviglia-braccio) ed uno score di rischio

genetico (GRS) composto da 62 polimorfismi (tag-SNP) noti per essere associati ad aumentato rischio di T2D. Il livello di significatività statistica è stato posto a $p<0.01$ ($p=0.05/5$, il numero di indici SCA analizzati) per ciascuno dei modelli statistici utilizzati (modello base corretto per sesso; modello completo corretto per un ampio numero di fattori di rischio cardiovascolari).

Risultati - Il GRS per T2D non è risultato significativamente associato con SCA. Un'associazione negativa tra GRS e calcium score coronarico in MESA (Europei Americani) e con spessore intima-media in carotide comune in FHS ($p=0.004$ e 0.009, rispettivamente, modello completo) non ha trovato conferma nelle altre coorti.

Conclusioni - Il rischio genetico per T2D rappresentato da 62 tag-SNP non è associato con SCA in differenti gruppi etnici, nonostante una dettagliata fenotipizzazione, un ampio campione ed un disegno sperimentale multicentrico. E' possibile che le basi genetiche comuni del rischio di T2D e CVD possano essere chiarite quando saranno individuate le varianti funzionali all'interno dei loci marcati dai 62 tag-SNP.

Abbreviations

FHS	The Framingham Heart Study
CARDIA	The Coronary Artery Risk Development in Young Adults
MESA	The Multi-Ethnic Study of Atherosclerosis
GENOA	The Genetic Epidemiology Network of Arteriopathy
T2D	Type 2 Diabetes
IGT	Impaired Glucose Tolerance
CVD	Cardiovascular Disease
SCA	Sub-Clinical Atherosclerosis
CACS	Coronary Artery Calcium Score
AACS	Abdominal Aortic Calcium Score
CCA-IMT	Common Carotid Artery Intima-Media Thickness
ICA-IMT	Internal Carotid Artery Intima-Media Thickness
ABI	Ankle-Brachial Index
GWAS	Genome-Wide Association Study
SNP	Single Nucleotide Polymorphism
GRS	Genetic Risk Score
IDF	International Diabetes Federation
FG	Fasting Glucose
FI	Fasting Insulin
Tg	Triglycerides
HDL-C	High Density Lipoprotein Cholesterol
BMI	Body Mass Index
WHR	Waist-to-Hip Ratio

3.2 INTRODUCTION

3.2.1 Background

Type 2 diabetes (T2D) and cardiovascular disease (CVD) are clinically associated in adults [1, 2] and are an increasing public health and economic scourge in the US [3, 4] and worldwide[5-7] (**Figure 1-3**). Better prevention strategies require comprehension of the constellation of risk factors and mediators underlying both T2D and CVD [8, 9].

T2D and CVD share a common metabolic milieu (**Figure 4**) that triggers metabolic and vascular dysfunction starting at subclinical disease stages (**Figure 5**) or even at birth [10] due to genetic and non-genetic risk factors. In particular, many recently identified common genetic variants increasing risk for T2D also are associated with CVD risk factors [11] (**Figure 6**) and so might also confer risk for subclinical atherosclerosis (SCA) [12].

Recently, a set of 36 T2D single nucleotide polymorphisms (SNPs) previously identified in large genome-wide association studies (GWAS) as affecting T2D risk was found to be associated with an increased risk of cardiovascular complications in type 2 diabetic patients [2, 13]. It has been shown that an additive genetic risk score (GRS_{62}) comprised of 62 validated T2D-associated SNPs [14-17] is a validated predictor of incident T2D in European and African Americans[18, 19].

3.2.2 Objective

The present work sought to test the hypothesis that the T2D genetic burden, as represented by the polygenic T2D GRS_{62}, is also positively associated in cross-sectional analyses with variation in measures of SCA, including coronary artery

(CACS) or abdominal aortic calcium score (AACS), internal (ICA-IMT) or common carotid artery intima-media thickness (CCA-IMT), and ankle-brachial index (ABI).

To maximize the sample size a multicenter transethnic association study was conducted in large population samples from four studies currently ongoing across the US (**Figure 7**): the Framingham Heart Study (FHS), the Coronary Artery Risk Development in Young Adults (CARDIA) [20], the Multi-Ethnic Study of Atherosclerosis (MESA) [21, 22] and the Genetic Epidemiology Network of Arteriopathy (GENOA) [23, 24].

3.3 METHODS

3.3.1 Study populations

The Offspring Cohort of the Framingham Heart Study

Analyses were conducted for each measured SCA trait on a range of 1,111 up to 2,822 participants of European ancestry from the Offspring Cohort of the FHS [13, 25]. Outcomes of interest and related clinical characteristics included in the present analyses were obtained at Offspring examination cycles 6 (for analyses of ICA-IMT and CCA-IMT) and 7 (for CACS and AACS). ABI was measured between the two examinations and the covariates used in the analysis of ABI were from the examination with the closest date to the ABI evaluation. Details regarding clinical assessment of participants and technical information about imaging tests and indices calculations have been published previously [13, 25, 26].

The CARDIA Study

Analyses were conducted for the available SCA traits on 816 African Americans and 1,635 European Americans enrolled in the CARDIA Study [20]. Analyses were limited to participants whose genotype information and clinical and anthropometric characteristics were available for all predictors of interest. Data on SCA from follow-up visit at years 20 (ICA-IMT and CCA-IMT) and 25 (CACS) were used.

The MESA Study

The MESA Study was designed to prospectively evaluate the development and progression of atherosclerotic disease. The complete design and protocols of MESA have been published previously [22]. The selection included individuals from the resident list of adults from the urban areas of the recruiting centers with emphasis on ethnic diversity. The present study included up to 2,526 participants of European ancestry, 1,611 African Americans, 773 Asian Americans and 1,446 Hispanic Americans from examination year 1 (2000-2001).

The GENOA Study

The longitudinal Genetic Epidemiology Network of Arteriopathy (GENOA) Study is one of four networks in the NHLBI Family-Blood Pressure Program and aims to elucidate the genetics of target organ complications of hypertension [24]. GENOA recruited European and African American sibships with at least 2 individuals with clinically diagnosed essential hypertension before age 60 years. European Americans were recruited from the Rochester, MN Field Center and African Americans were recruited from the Jackson, MS Field Center. Current analyses were conducted on CACS measures and genotypes available for 969 European Americans and 535 African Americans.

In all study cohorts, participants with a personal history of CVD defined as myocardial infarction, stroke, coronary angioplasty and/or amputation not due to injury, when applicable, were excluded from the analyses.

3.3.2. Assessment of subclinical atherosclerosis (SCA)

SCA measures were determined in a similar fashion in all studies by means of carotid ultrasonography intima-media thickness, subcategorized for common and internal carotid (CCA-IMT and ICA-IMT), computed tomography scan for CACS and AACS, and ABI for peripheral artery disease, as detailed elsewhere [13, 27-30]. FHS and MESA study participants had measurements available for all five SCA traits. CARDIA participants had ICA-IMT, CCA-IMT [31] and CACS [32] measurements available for analyses. Evaluation and interpretation of CACS measures in MESA were conducted as published elsewhere [32, 33]. In GENOA, CACS was measured in European Americans with an Imatron C-150 electron beam CT (EBCT) scanner (Imatron Inc., South San Francisco, CA) using a previously described protocol [34]. In GENOA African Americans, CACS was measured with standard scanning protocols developed as part of the NHLBI's MESA and CARDIA studies [32].

3.3.3 Genotyping

Genotyping in FHS was conducted using the Affymetrix GeneChip Human Mapping 500K Array supplemented with the Affymetrix 50K array, while CARDIA and MESA Study participants were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 (Santa Clara, California) [35]. GENOA Study participants were genotyped on the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, 2007) or the Illumina Human 1M-Duo BeadChip (Illumina,

2010) at the Mayo Clinic, Rochester, MN. Quality control measures and imputation strategy for missing genotypes were extensively detailed in previous reports [19, 35-37] for both FHS and CARDIA samples, while complete information on genotyping and imputation quality of MESA and GENOA samples are available in the **Appendix**.

3.4 STATISTICAL ANALYSIS

The GRS in FHS and CARDIA European Americans was calculated as the sum of the number of risk alleles (0, 1, or 2) at each locus, weighted by its published effect size (natural log-transformed) from the DIAbetes Genetics Replication And Meta-analysis (DIAGRAMv3) [14]. For CARDIA African Americans and for each MESA and GENOA ethnic group an unweighted GRS was used, calculated by summing the risk alleles across the loci. ICA-IMT, CCA-IMT, AACS, CACS, fasting insulin, triglycerides and HDL-cholesterol were log-transformed to reduce skewness. Descriptive data were expressed as mean ± standard error, if not otherwise indicated. Multivariable linear regression models were used for CARDIA and MESA cohorts. Multivariable linear regression models with random effect to account for family relatedness were used to test the association of an additive 62 T2D SNPs GRS (**Appendix Table 1**) with measures of SCA in FHS and GENOA.

For each SCA trait models adjusted only for sex (genetic-only model) and for a comprehensive set of SCA risk factors (full model), were used as shown in **Appendix Table 2**. Principal components were also included in GENOA and MESA models to control for population stratification in each ethnic group. The fully-adjusted model included comprehensive CVD risk factors: sex, age, waist

circumference, body mass index (BMI), triglycerides, HDL-cholesterol, LDL-cholesterol, fasting insulin, fasting glucose, systolic blood pressure (SBP), hypertension/diabetes and/or lipid medication, physical activity, smoking status, family history of T2D and/or CVD. SBP was excluded in the analysis for ankle-brachial index (ABI) since ABI is calculated from SBPs at ankle and arm.

Subsidiary analyses were conducted by using two subsets of the 62 T2D SNPs comprised of 20 tag-SNPs thought to be associated with beta-cell function (GRS_{β}) or 10 associated with insulin resistance (GRS_{IR}) [19] in the FHS and CARDIA cohorts to further elucidate possible mechanistic pathways, testing the hypothesis that genetic risk for IR in particular would be associated with SCA.

Post-hoc power calculations using QUANTO 1.2 software showed that for a sample size of 1,835 individuals, there was 80% power to detect association of GRS_{62} explaining 0.64% of the variance in SCA traits with type 1 error rate set at $p < 0.01$ ($p = 0.05$ divided by the number of traits (5) analyzed). There was 80% power to detect association of GRS_{β} and GRS_{IR} each explaining 0.73% or more of the variance in SCA traits with type 1 error set at $p < 0.005$ (0.05/[5 traits x 2 GRS]).

In order to replicate the primary FHS analyses in European Americans and to verify whether they might be extended to different ancestral groups, association analyses of the GRS_{62} with CACS, ICA-IMT and CCA-IMT in CARDIA, MESA and GENOA cohorts were then conducted separately within each ethnicity. Then, association results from each cohort were meta-analyzed using a fixed effect approach, separately for European and African Americans, with a two-sided $p<0.01$ as threshold for significance.

All statistical analyses were carried out with SAS 9.2 (SAS Institute Inc., Cary, NC, USA) and R 2.9.2 [38].

3.5 RESULTS

Up to 7,952 European Americans, 2,124 Africans Americans, 773 Asian Americans and 1,446 Hispanic Americans from four cohort studies were analyzed. Clinical and anthropometric features and measures of SCA traits are shown for each study cohort in **Appendix Table 3** and **Appendix Table 4**.

Overall, study participants were of a wide age and BMI range. Prevalence of diabetes and abdominal obesity was much higher in African Americans than in other ethnicities. Participant characteristics within each ethnic group were comparable across all cohorts with the proportion of males and females being equally distributed, except in GENOA African Americans where women comprised 74.2% of the participants. The T2D GRS_{62} was approximately normally distributed with a range from 48.3 to 83.3 in European Americans and from 46.8 to 83.2 in African Americans over all cohorts. African Americans in CARDIA had higher mean GRS than European Americans, while the opposite was the case in MESA and GENOA cohorts. In MESA Asian and Hispanic Americans the T2D GRS_{62} spanned, respectively, from 48.1 to 73.6 and from 48.5 to 79.6 (**Appendix Figure 1** and **Appendix Figure 2**). GRS_{β} and GRS_{IR} were normally distributed and ranged from 12.2 to 31.9 and from 3.3 to 18.0 in FHS and CARDIA European Americans, respectively, while in CARDIA African Americans the GRS_{β} ranged from 13.2 to 28.4 and the GRS_{IR} ranged from 5.0 and 16.9 (**Appendix Figure 2**).

The primary analyses in FHS showed a significant inverse association between GRS_{62} and CCA-IMT ($p = 0.009$, fully adjusted model), which was not replicated in CARDIA or MESA (**Table 1**) European Americans. In the MESA European Americans, there was evidence for a significant inverse association between GRS_{62} and CACS ($p = 0.004$, fully adjusted model), which was not replicated in other cohorts (**Table 1**). Lack of evidence for a significant association between

GRS_{62} and SCA was confirmed by meta-analyses of up to 12,983 individuals from four cohorts for the available SCA traits, i.e. CACS, CCA-IMT and ICA-IMT (**Figure 8**).

Supplemental analyses showed that ICA-IMT was negatively associated with GRS_{β} in FHS in the fully adjusted model ($p = 0.007$, **Appendix Table 5**), but the finding was not replicated in CARDIA European Americans. The GRS_{IR} was not significantly associated with any of the SCA traits in any of the models in either the FHS cohort or in either CARDIA ethnic group (**Appendix Table 6**).

3.6 DISCUSSION

The primary finding of the present exercise was that there is no evidence of a statistically significant association between the genetic burden for T2D, based on a 62 T2D SNP GRS, and a wide set of subclinical atherosclerosis traits in a large US adult population sample. Results were consistent for all four ancestral groups studied. An inverse association of the GRS_{62} with CCA-IMT in FHS was not confirmed in two other cohorts either in replication comparisons or meta-analysis. A 10 SNP GRS and a 20 SNP GRS representing variants associated, respectively, with beta-cell function or insulin resistance also was not statistically significantly associated with SCA. Starting with the plausible hypothesis that T2D genetic risk would also be associated with SCA, a rigorous approach was applied including conservative correction for multiple trait tests, replication studies in separate cohorts (thereby reducing type 1 error) and meta-analysis of a sufficiently large sample size (increasing power). It can be therefore conclusively stated that this particular formulation of T2D genetic risk is not associated with higher indices of SCA in the study participating in the present study.

These results can be compared to other recent studies. As recently highlighted [11], the genetic signatures of T2D, coronary heart disease and glycemic quantitative traits have some elements of commonality, though limited only to 2q36.3 and 9p21.3 chromosomal regions. Notably, a major proportion of fasting insulin-associated loci showed a directionally consistent association with T2D risk and CVD quantitative risk factors (i.e. adverse lipid profile and abdominal adiposity) but none of the glycemic quantitative traits appears to be directly associated with CVD-risk. While Qi et al. [2] showed that a GRS comprised of 36 T2D genetic variants was associated with an increased risk of CVD complications (a composite endpoint including fatal or nonfatal coronary heart disease and stroke) in European Americans affected by T2D, other attempts failed to identify an etiological link between SCA traits and T2D genetic variation at candidate 2q36.3-*IRS1* locus in the Framingham Heart Study [13]. Additionally, while Doria et al.[39] showed that the effect of genotype at 9p21.3 locus on CVD events is raised only in persons with T2D who have poor glycemic control, Rivera et al. [40] have recently shown that, compared to non-T2D individuals, a higher number of variants at 9p21.3 locus is associated with the severity of coronary artery disease comorbid with T2D. These lines of evidence suggest that T2D and non-T2D subjects might have different mechanisms leading to CVD events and that within T2D cases hyperglycemia might act as permissive environment leading to the full expression of CVD-risk genetics. These data, together with the null results of the present study, both with our full T2D GRS_{62} and with the two sub-scores (GRS_{IR} and GRS_{β}), collectively suggest that in the general population T2D and CVD are not genetically linked together through SCA, the association of T2D genetics being so far observed only with CVD events but not with early subclinical disease.

Strengths of the present exercise include a validated T2D GRS aligned to the current level of evidence, a detailed phenotypic characterization, a comprehensive selection of covariates, confounders and mediators, as well as a careful control of

type 1 and type 2 errors by means of a large sample size from the general population and a multicenter replication strategy in different ethnicities, as discussed above. Additionally, given the strong age-calcification relationship across young adulthood, mid-life and older ages [41, 42], the wide range of age among different cohorts allowed to capture the whole spectrum from early- to late-onset calcification.

However, the approach proposed here might have been weakened by multiple interactions among different SNPs within the GRS: several of the component genes in the score may be indeed associated with SCA, but the component score might not be significantly associated if the effect was diluted by the other variants. Furthermore, the 62 genome-wide significant SNPs used explained only a fraction (around 10%) of the total T2D phenotypic variance in other studies [43] and did not represent actual functional allelic variants that have yet to be discovered. It should be also acknowledged that, if it is assumed that T2D contributes to CVD, excluding individuals with a personal CVD history might have resulted in a population enriched for protective factors especially among those with higher T2D GRS, which might also explain the borderline negative association of the GRS_{62} with CCA IMT in FHS and CACS in MESA European Americans. However, in preliminary sensitivity analyses conducted in FHS and CARDIA the T2D GRS_{62} allele distribution resulted to be comparable between people with positive personal CVD history and the population actually used in the association analyses (data not shown), so that, for consistency with the main focus of the present research on subclinical atherosclerosis individuals with clinical evidence of incident or prevalent CVD were excluded, as defined above in the Methods section. Lastly, while the GRS could be confidently used to depict the T2D genetic risk for European and Mexican Americans, and therefore reasonably allow to claim robustness of the obtained results, the GRS was not best tailored to fit T2D genetics in African or Asian Americans, given their different haplotype

structure and allele frequencies [37] and given that most of GWAS hits have been discovered and confirmed in people of European ancestry.

Results of this research project have several implications and point to future directions. The present study provided compelling evidence that the genetic burden of T2D risk as represented by this GRS_{62} formulation was not associated with SCA. This suggests that either T2D is not working as reasonable foundation of clinical CVD through SCA, at least at the genetic level, or that more complex formulation of T2D genetic risk might be associated with SCA, or that no large common genetic soil [12] underlies both T2D and CVD.

However, T2D and SCA are strongly linked clinically [1, 2] and the prevalence of CVD events and the burden of CVD risk factors are higher in people with diabetes. Therefore, although the last decade has shown an increasing expansion in the understanding of the genetic signature of complex traits, new approaches incorporating functional, structural and/or regulatory annotation [44, 45] into disease prediction is needed to untangle the missing link, if any, between T2D and CVD at the genetic level. Furthermore, given that screening for SCA in asymptomatic individuals at intermediate CVD risk improves CVD risk stratification [46] and that current polygenic scores slightly but not remarkably outperform clinical models [18], functional interrogation of T2D and CVD genetic architecture is necessary to further optimize polygenic risk prediction of either T2D or CVD or both.

3.7 CONCLUSIONS

In conclusion, the present exercise places an additional step in the wider framework of CVD prediction and it is expected that it will serve as placeholder for future mechanistic investigations. Given the global burden of T2D and CVD

in the era of precision medicine and prospective patient-oriented healthcare, it is timely to acquire additional knowledge about the genetic determinants of T2D, CVD and intermediate traits to improve risk prediction and the ability to discover newly targeted therapeutic molecules.

3.8 FIGURES and TABLES

Table 1 – Association between a 62 T2D SNPs genotype risk score (GRS) and prevalent subclinical atherosclerosis measures in linear regression models of FHS, CARDIA, GENOA[§§] and MESA cohorts.

	FHS		**CARDIA**				**GENOA**			
	European Americans		**European Americans**		**African Americans**		**European Americans**		**African Americans[§§]**	
Basic Model	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*
CACS	-0.012±0.026	0.66	0.011±0.010	0.27	-0.007±0.016	0.64	0.003±0.017	0.88	0.025±0.029	0.40
AACS	-0.029±0.032	0.36	-	-	-	-	-	-	-	-
ICA-IMT	-0.008±0.004	0.03	0.000±0.001	0.88	0.000±0.001	0.67	-	-	-	-
CCA-IMT	-0.002±0.002	0.21	0.000±0.000	0.94	0.001±0.001	0.08	-	-	-	-
ABI	0.000±0.001	0.72	-	-	-	-	-	-	-	-
Full Model										
CACS	-0.012±0.028	0.67	0.006±0.010	0.56	-0.018±0.015	0.24	0.001±0.015	0.95	0.041±0.026	0.12
AACS	-0.017±0.033	0.61	-	-	-	-	-	-	-	-
ICA-IMT	-0.009±0.004	0.02	0.000±0.001	0.63	0.000±0.001	0.72	-	-	-	-
CCA-IMT	-0.004±0.002	**0.009**	0.000±0.000	0.54	0.000±0.001	0.47	-	-	-	-
ABI	0.001±0.001	0.13	-	-	-	-	-	-	-	-

Table 1 – *continue*

	MESA							
	European Americans		**Asian Americans**		**African Americans**		**Hispanic Americans**	
Basic Model	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*
CACS	-0.026±0.010	0.01	-0.027±0.012	0.14	0.019±0.014	0.17	-0.004±0.014	0.76
AACS	-0.042±0.023	0.07	-0.022±0.048	0.65	0.003±0.039	0.93	0.029±0.031	0.33
ICA-IMT	-0.001±0.002	0.75	0.001±0.004	0.85	-0.001±0.003	0.65	0.006±0.003	0.02
CCA-IMT	-0.001±0.001	0.33	-0.001±0.002	0.65	-0.000±0.001	0.79	-8.98E-06±0.001	0.99
ABI	3.27E-05±0.001	0.95	0.000±0.001	0.63	-0.000±0.001	0.64	0.000±0.001	0.61
Full Model								
CACS	-0.029±0.009	**0.004**	-0.027±0.019	0.16	0.016±0.014	0.28	0.002±0.015	0.88
AACS	-0.012±0.019	0.53	-0.033±0.039	0.40	-0.027±0.035	0.44	0.035±0.028	0.20
ICA-IMT	-0.001±0.002	0.63	-0.000±0.004	0.98	-0.002±0.003	0.52	0.006±0.003	0.02
CCA-IMT	-0.001±0.001	0.30	-0.001±0.002	0.77	-0.000±0.001	0.87	9.49E-05±0.001	0.93
ABI	0.000±0.001	0.44	0.000±0.001	0.67	-0.001±0.001	0.36	0.000±0.001	0.85

CACS, coronary artery calcium score; AACS, abdominal aorta calcium score; ICA, internal carotid artery; CCA, common carotid artery; IMT, intima-media thickness; ABI, ankle-brachial index. Basic Model: SCA trait = GRS + sex + k; Full Model: fully adjusted model (see Methods and Supplementary Table S7 for details). §§African Americans in GENOA had a genetic risk score limited to 55 of 62 T2D SNPs. Data expressed as mean ± standard error. Sample sizes (*N* [min-max]): FHS N_{Eur} [1,111-2,822]; CARDIA: N_{Eur} [1,267-1,635], N_{Afr} [562-816]; GENOA: N_{Eur}=969, N_{Afr}=535; MESA: N_{Eur} [760-2,526], N_{Asi} [247-773]; N_{Afr} [343-1,611], N_{His} [496-1,446].

Figure 1 – Global projections for the diabetes epidemic 2010-2030 - *IDF Diabetes Atlas 2011*

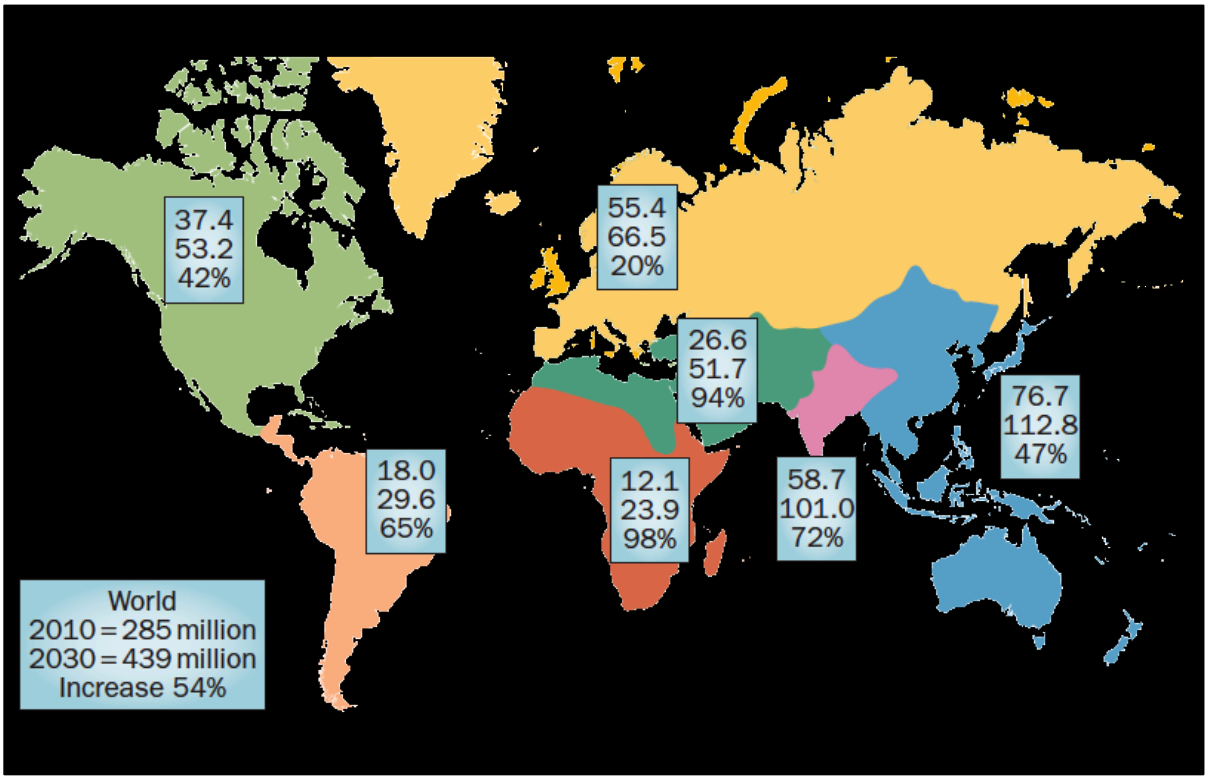


Figure 2 – Worldwide prospective prevalence (2010-2030) of diabetes and sub-threshold hyperglycemia (IGT, Impaired Glucose Tolerance) – *IDF 2011*

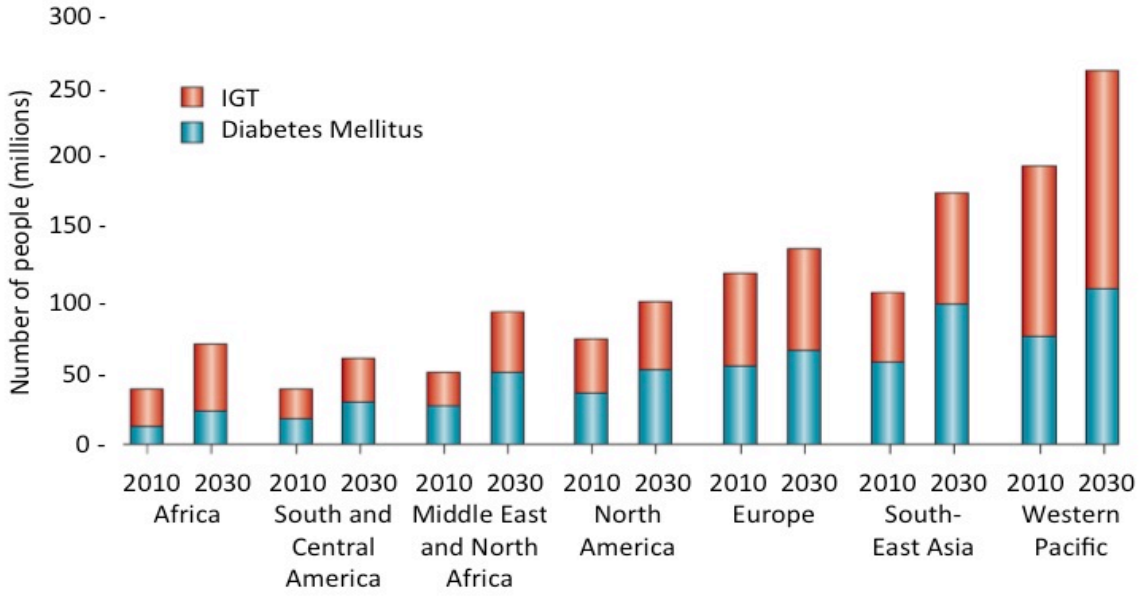


Figure 3 – Percent of medical condition-specific expenditures associated with diabetes in U.S.A. – ***American Diabetes Association 2013***

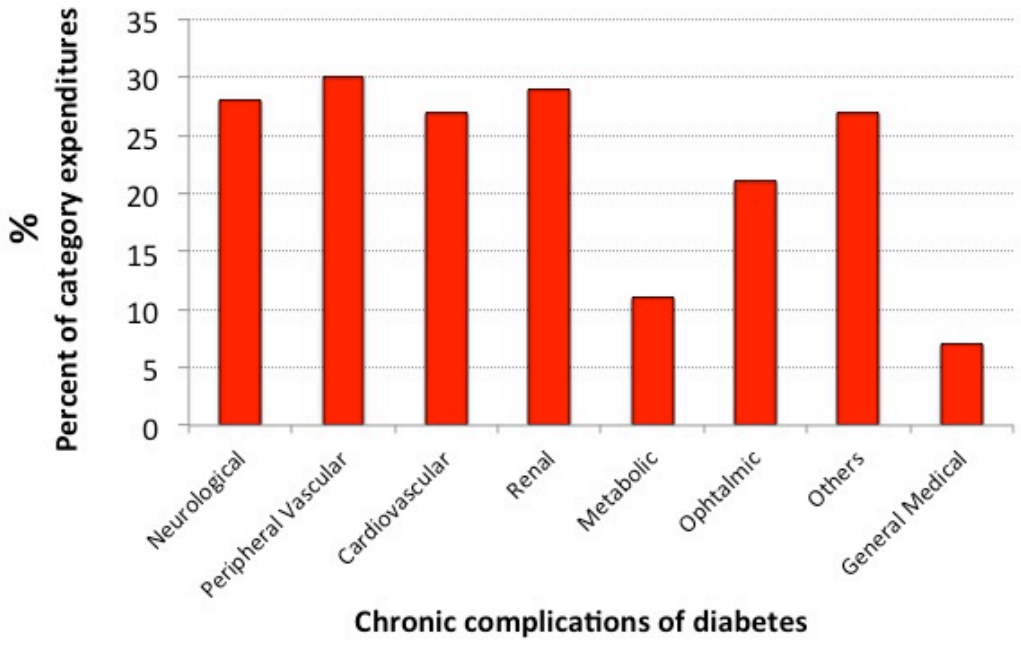
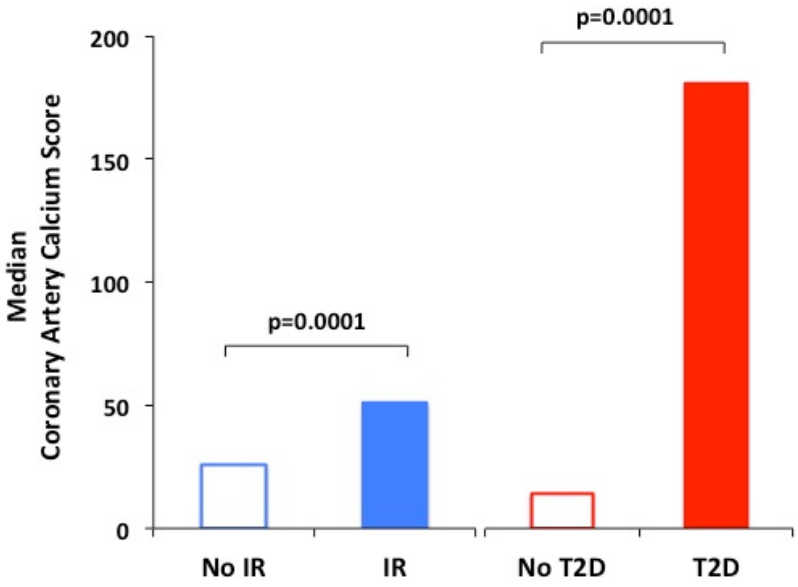


Figure 4 – CVD risk factors and the common T2D-CVD metabolic inflammatory milieu

TRADITIONAL CVD RISK FACTORS	"NON-TRADITIONAL" CVD RISK FACTORS
Hyperglycemia	Endothelial dysfunction
Hypertension	Impaired fibrinolysis
Dyslipidemia	Inflammation
Obesity	Microalbuminuria
Cigarette smoking	Increased homocysteine levels
Physical inactivity	Vascular wall abnormalities

Adapted from Lorber D. ***et al. Diabetes Metab. Syndr. Obes. 2014***

Figure 5 – Atherosclerosis is higher in type 2 diabetes but also in intermediate, insulin resistant (IR) phenotypes: The Framingham Offspring Study.



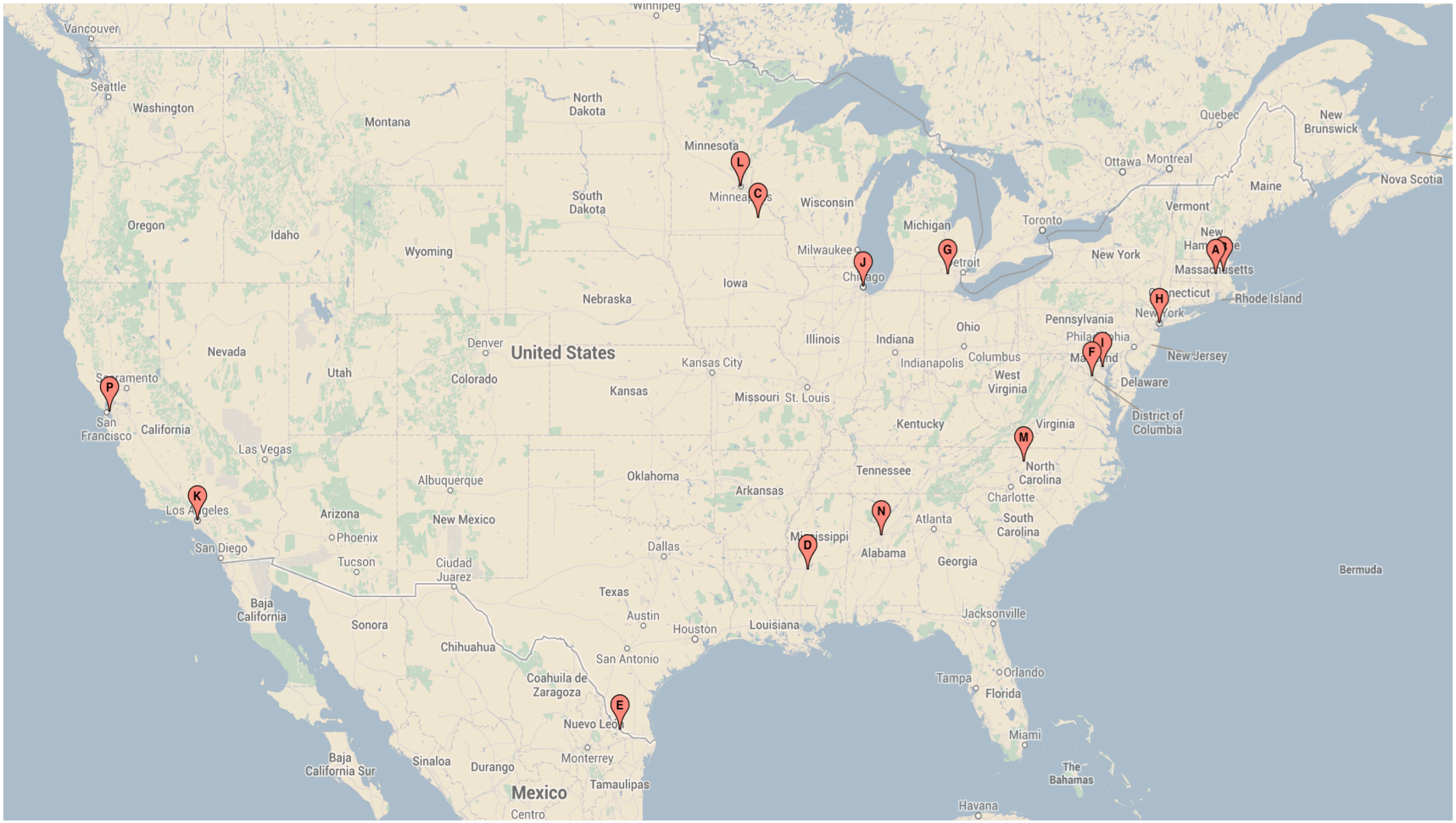
Modified from *Meigs et al. Diabetes Care 2002*

Figure 6 – Association of some glycaemic traits loci with type 2 diabetes and/or insulin resistance-associated metrics overlapping coronary heart disease risk factors.

Locus	Chr	T2D P<-E04	FG P<E05	FI P<E05	Tg/HDL P<-E04	BMI P<-E04	WHR P<-E04
GRB14	2	■		■	■		■
IRS1	2	■		■	■		
PPARG	3	■		■	■		
ANKRD55	5	■		■	■		
ARL15	5	■		■			
FTO	16	■		■	■	■	
PEPD	19	■		■	■		
PPP1R3B	8	■	■	■	■		
GCKR	2	■	■	■	■		
TCF7L2	10	■	■	■			
LYPLAL1	1			■			■
PDGFC	4			■	■		
FAM13A	4			■	■		
RSPO3	6			■			■
HIP1	7			■		■	

Modified from *Scott et al. Nat. Genet. 2012*

Figure 7 - Study participating centers.



A. FHS - NHLBI's Framingham Heart Study - Framingham, MA
B. FHS - Boston University, Harvard Medical School, Massachusetts General Hospital - Boston, MA
C. GENOA – Rochester, MN
D. GENOA – Jackson, MS
E. GENOA – Rio Grande City – Rio Grande City, TX
F. NHLBI, NIH – Bethesda, MD
G. GENOA – University of Michigan - Ann Arbor, MI
H. MESA – Columbia University, New York City, NY
I. MESA – Johns Hopkins University, Baltimore, MD
J. MESA & CARDIA – Northwestern University, Chicago, IL
K. MESA – UCLA – Los Angeles, CA
L. MESA – University of Minnesota, Twin Cities – Minneapolis, MN
M. MESA – Wake Forest University – Winston-Salem, NC
N. CARDIA – University of Alabama – Birmingham, AL
O. CARDIA – University of Minnesota – Minneapolis, MN
P. CARDIA – Kaiser Permanente – Oakland, CA

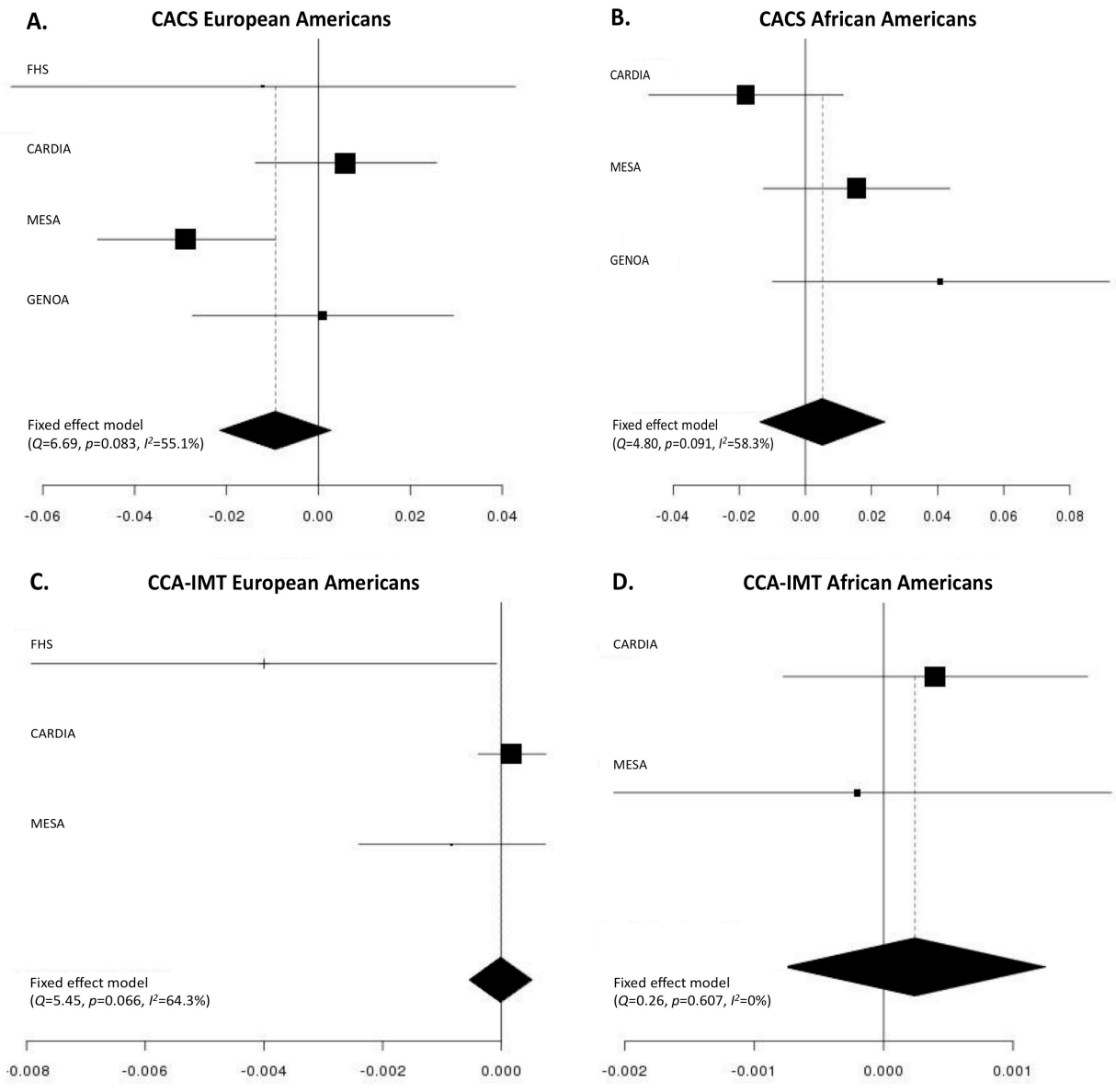
Table 1 – Association between a 62 T2D SNPs genotype risk score (GRS) and prevalent subclinical atherosclerosis measures in linear regression models of FHS, CARDIA, GENOA[§§] and MESA cohorts.

	FHS		CARDIA				GENOA			
	European Americans		European Americans		African Americans		European Americans		African Americans[§§]	
Basic Model	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*
CACS	-0.012±0.026	0.66	0.011±0.010	0.27	-0.007±0.016	0.64	0.003±0.017	0.88	0.025±0.029	0.40
AACS	-0.029±0.032	0.36	-	-	-	-	-	-	-	-
ICA-IMT	-0.008±0.004	0.03	0.000±0.001	0.88	0.000±0.001	0.67	-	-	-	-
CCA-IMT	-0.002±0.002	0.21	0.000±0.000	0.94	0.001±0.001	0.08	-	-	-	-
ABI	0.000±0.001	0.72	-	-	-	-	-	-	-	-
Full Model										
CACS	-0.012±0.028	0.67	0.006±0.010	0.56	-0.018±0.015	0.24	0.001±0.015	0.95	0.041±0.026	0.12
AACS	-0.017±0.033	0.61	-	-	-	-	-	-	-	-
ICA-IMT	-0.009±0.004	0.02	0.000±0.001	0.63	0.000±0.001	0.72	-	-	-	-
CCA-IMT	-0.004±0.002	**0.009**	0.000±0.000	0.54	0.000±0.001	0.47	-	-	-	-
ABI	0.001±0.001	0.13	-	-	-	-	-	-	-	-

	MESA							
	European Americans		Asian Americans		African Americans		Hispanic Americans	
Basic Model	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*
CACS	-0.026±0.010	0.01	-0.027±0.012	0.14	0.019±0.014	0.17	-0.004±0.014	0.76
AACS	-0.042±0.023	0.07	-0.022±0.048	0.65	0.003±0.039	0.93	0.029±0.031	0.33
ICA-IMT	-0.001±0.002	0.75	0.001±0.004	0.85	-0.001±0.003	0.65	0.006±0.003	0.02
CCA-IMT	-0.001±0.001	0.33	-0.001±0.002	0.65	-0.000±0.001	0.79	-8.98E-06±0.001	0.99
ABI	3.27E-05±0.001	0.95	0.000±0.001	0.63	-0.000±0.001	0.64	0.000±0.001	0.61
Full Model								
CACS	-0.029±0.009	**0.004**	-0.027±0.019	0.16	0.016±0.014	0.28	0.002±0.015	0.88
AACS	-0.012±0.019	0.53	-0.033±0.039	0.40	-0.027±0.035	0.44	0.035±0.028	0.20
ICA-IMT	-0.001±0.002	0.63	-0.000±0.004	0.98	-0.002±0.003	0.52	0.006±0.003	0.02
CCA-IMT	-0.001±0.001	0.30	-0.001±0.002	0.77	-0.000±0.001	0.87	9.49E-05±0.001	0.93
ABI	0.000±0.001	0.44	0.000±0.001	0.67	-0.001±0.001	0.36	0.000±0.001	0.85

CACS, coronary artery calcium score; AACS, abdominal aorta calcium score; ICA, internal carotid artery; CCA, common carotid artery; IMT, intima-media thickness; ABI, ankle-brachial index. Basic Model: SCA trait = GRS + sex + k; Full Model: fully adjusted model (see Methods and Supplementary Table S7 for details). §§African Americans in GENOA had a genetic risk score limited to 55 of 62 T2D SNPs. Data expressed as mean ± standard error. Sample sizes (*N* [min-max]): FHS N_{Eur} [1,111-2,822]; CARDIA: N_{Eur} [1,267-1,635], N_{Afr} [562-816]; GENOA: N_{Eur}=969, N_{Afr}=535; MESA: N_{Eur} [760-2,526], N_{Asi} [247-773]; N_{Afr} [343-1,611], N_{His} [496-1,446].

Figure 6 – Meta-analysis of GRS_{62} association testing with CACS, CCA-IMT, ICA-IMT measures across all study cohorts stratified by European and African Americans.



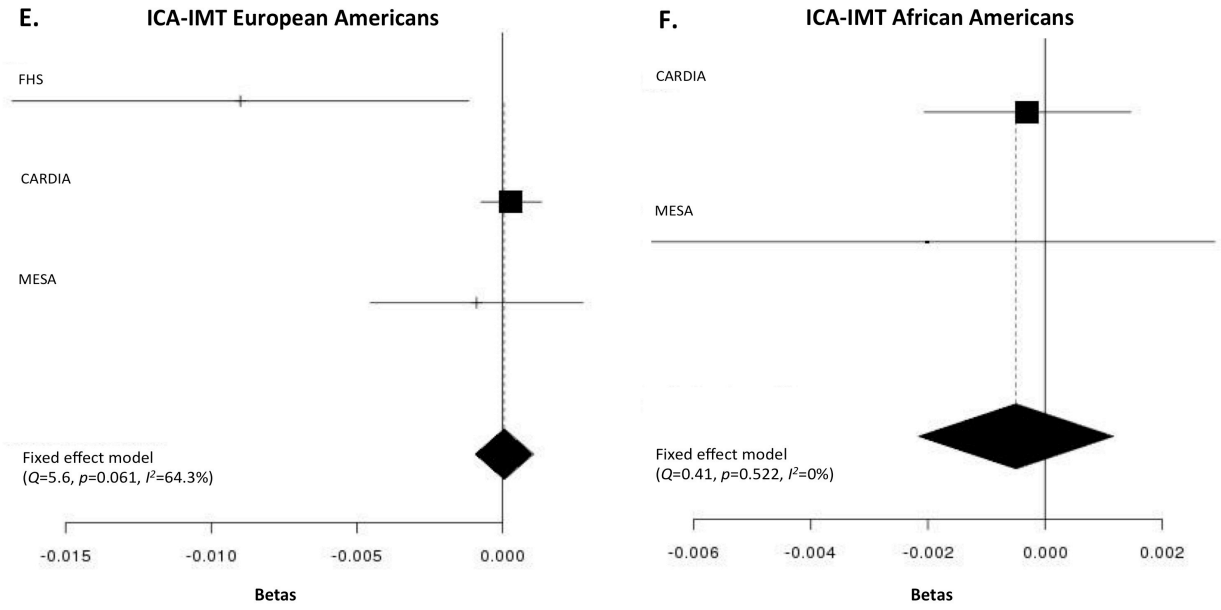


Figure 6 Legend
CACS (panel-A, N_{meta}=4,780; panel-B, N_{meta}=1,835), CCA-IMT (panel-C, N_{meta}=6,220; panel-D, N_{meta}=2,190), ICA-IMT (panel-E, N_{meta}=5,842; panel-F, N_{meta}=2,109); GRS_{62}, genetic risk score comprised of 62 single nucleotide polymorphisms associated with type 2 diabetes; CACS, coronary artery calcium score; CCA-IMT and ICA-IMT.

3.9 REFERENCES

1. **Sattar N.** Revisiting the links between glycaemia, diabetes and cardiovascular disease. *Diabetologia*. 2013;56:686-95.
2. **Qi Q**, Meigs JB, Rexrode KM, Hu FB and Qi L. Diabetes genetic predisposition score and cardiovascular complications among patients with type 2 diabetes. *Diabetes Care*. 2013;36:737-9.
3. **American Diabetes Association**. Economic Costs of Diabetes in the U.S. in 2012. *Diabetes Care*. 2013.
4. **Dabelea D**, Saydah S, Imperatore G, Linder B, Divers J, Bell R, *et al.* Prevalence of Type 1 and Type 2 Diabetes Among Children and Adolescents From 2001 to 2009. *JAMA : the journal of the American Medical Association*. 2014;311:1778.
5. **Whiting DR**, Guariguata L, Weil C and Shaw J. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Res Clin Pract*. 2011;94:311-21.
6. **International Diabetes Federation**. IDF Diabetes Atlas 5th ed.
7. **Morrish NJ**, Wang SL, Stevens LK, Fuller JH and Keen H. Mortality and causes of death in the WHO Multinational Study of Vascular Disease in Diabetes. *Diabetologia*. 2001;44 Suppl 2:S14-21.
8. **Wang CC** and Reusch JE. Diabetes and cardiovascular disease: changing the focus from glycemic control to improving long-term survival. *Am J Cardiol*. 2012;110:58B-68B.
9. **Ingelsson E**, Sullivan LM, Murabito JM, Fox CS, Benjamin EJ, Polak JF, *et al.* Prevalence and prognostic impact of subclinical cardiovascular disease in individuals with the metabolic syndrome and diabetes. *Diabetes*. 2007;56:1718-26.
10. **Zhang H**, Dellsperger KC and Zhang C. The link between metabolic abnormalities and endothelial dysfunction in type 2 diabetes: an update. *Basic Res Cardiol*. 2012;107:237.
11. **Dauriz M** and Meigs JB. Current Insights into the Joint Genetic Basis of Type 2 Diabetes and Coronary Heart Disease. *Curr Cardiovasc Risk Rep*. 2014;8:368.

12. **Stern MP**. Diabetes and cardiovascular disease. The "common soil" hypothesis. *Diabetes*. 1995;44:369-74.
13. **Lim S**, Hong J, Liu CT, Hivert MF, White CC, Murabito JM, *et al.* Common variants in and near IRS1 and subclinical cardiovascular disease in the Framingham Heart Study. *Atherosclerosis*. 2013;229(1):149-54.
14. **Morris AP**, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*. 2012;44:981-90.
15. **Voight BF**, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet*. 2010;42:579-89.
16. **Wellcome Trust Case Control Consortium**. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661-78.
17. **Zeggini E**, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet*. 2008;40:638-45.
18. **Hivert MF**, Vassy JL and Meigs JB. Susceptibility to type 2 diabetes mellitus-from genes to prevention. *Nature reviews Endocrinology*. 2014;10:198-205.
19. **Vassy JL**, Hivert MF, Porneala B, Dauriz M, Florez JC, Dupuis J, *et al.* Polygenic type 2 diabetes prediction at the limit of common variant detection. *Diabetes*. 2014.
20. **Vassy JL**, Durant NH, Kabagambe EK, Carnethon MR, Rasmussen-Torvik LJ, Fornage M, *et al.* A genotype risk score predicts type 2 diabetes from young adulthood: the CARDIA study. *Diabetologia*. 2012;55:2604-12.
21. **Pletcher MJ**, Sibley CT, Pignone M, Vittinghoff E and Greenland P. Interpretation of the coronary artery calcium score in combination with conventional cardiovascular risk factors: the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation*. 2013;128:1076-84.
22. **Bild DE**, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, *et al.* Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol*.

2002;156:871-81.

23. **Daniels PR**, Kardia SL, Hanis CL, Brown CA, Hutchinson R, Boerwinkle E, *et al.* Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med*. 2004;116:676-81.

24. **FBPP Investigators**. Multi-center genetic study of hypertension: The Family Blood Pressure Program (FBPP). *Hypertension*. 2002;39:3-9.

25. **Rosito GA**, Massaro JM, Hoffmann U, Ruberg FL, Mahabadi AA, Vasan RS, *et al.* Pericardial fat, visceral abdominal fat, cardiovascular disease risk factors, and vascular calcification in a community-based sample: the Framingham Heart Study. *Circulation*. 2008;117:605-13.

26. **O'Donnell CJ**, Cupples LA, D'Agostino RB, Fox CS, Hoffmann U, Hwang SJ, *et al.* Genome-wide association study for subclinical atherosclerosis in major arterial territories in the NHLBI's Framingham Heart Study. *BMC Med Genet*. 2007;8 Suppl 1:S4.

27. **Hoffmann U**, Siebert U, Bull-Stewart A, Achenbach S, Ferencik M, Moselewski F, *et al.* Evidence for lower variability of coronary artery calcium mineral mass measurements by multi-detector computed tomography in a community-based cohort--consequences for progression studies. *Eur J Radiol*. 2006;57:396-402.

28. **Murabito JM**, Guo CY, Fox CS and D'Agostino RB. Heritability of the ankle-brachial index: the Framingham Offspring study. *Am J Epidemiol*. 2006;164:963-8.

29. **O'Leary DH**, Polak JF, Kronmal RA, Manolio TA, Burke GL and Wolfson SK, Jr. Carotid-artery intima and media thickness as a risk factor for myocardial infarction and stroke in older adults. Cardiovascular Health Study Collaborative Research Group. *N Engl J Med*. 1999;340:14-22.

30. **Wong ND**, Lopez VA, Allison M, Detrano RC, Blumenthal RS, Folsom AR, *et al.* Abdominal aortic calcium and multi-site atherosclerosis: the Multiethnic Study of Atherosclerosis. *Atherosclerosis*. 2011;214:436-41.

31. **Polak JF**, Person SD, Wei GS, Godreau A, Jacobs DR, Jr., Harrington A, *et al.* Segment-specific associations of carotid intima-media thickness with

cardiovascular risk factors: the Coronary Artery Risk Development in Young Adults (CARDIA) study. *Stroke; a journal of cerebral circulation*. 2010;41:9-15.

32. **Carr JJ**, Nelson JC, Wong ND, McNitt-Gray M, Arad Y, Jacobs DR, Jr., et al. Calcified coronary artery plaque measurement with cardiac CT in population-based studies: standardized protocol of Multi-Ethnic Study of Atherosclerosis (MESA) and Coronary Artery Risk Development in Young Adults (CARDIA) study. *Radiology*. 2005;234:35-43.

33. **McClelland RL**, Chung H, Detrano R, Post W and Kronmal RA. Distribution of coronary artery calcium by race, gender, and age: results from the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation*. 2006;113:30-7.

34. **Bielak LF**, Sheedy PF, 2nd and Peyser PA. Coronary artery calcification measured at electron-beam CT: agreement in dual scan runs and change over time. *Radiology*. 2001;218:224-9.

35. **Lemaitre RN**, Tanaka T, Tang W, Manichaikul A, Foy M, Kabagambe EK, *et al.* Genetic loci associated with plasma phospholipid n-3 fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium. *PLoS Genet*. 2011;7:e1002193.

36. **Lettre G**, Palmer CD, Young T, Ejebe KG, Allayee H, Benjamin EJ, *et al.* Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: the NHLBI CARe Project. *PLoS Genet*. 2011;7:e1001300.

37. **Gabriel SB**, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, *et al.* The structure of haplotype blocks in the human genome. *Science*. 2002;296:2225-9.

38. **R Core Team**. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. http://wwwR-projectorg. Vienna, Austria, 2013.

39. **Doria A**, Wojcik J, Xu R, Gervino EV, Hauser TH, Johnstone MT, *et al.* Interaction between poor glycemic control and 9p21 locus on risk of coronary artery disease in type 2 diabetes. *JAMA : the journal of the American Medical Association*. 2008;300:2389-97.

40. **Rivera NV**, Carreras-Torres R, Roncarati R, Viviani-Anselmi C, De Micco F, Mezzelani A, *et al.* Assessment of the 9p21.3 locus in severity of coronary artery disease in the presence and absence of type 2 diabetes. *BMC Med Genet*. 2013;14:11.

41. **Berry JD**, Liu K, Folsom AR, Lewis CE, Carr JJ, Polak JF, *et al.* Prevalence and progression of subclinical atherosclerosis in younger adults with low short-term but high lifetime estimated risk for cardiovascular disease: the coronary artery risk development in young adults study and multi-ethnic study of atherosclerosis. *Circulation*. 2009;119:382-9.

42. **Kronmal RA**, McClelland RL, Detrano R, Shea S, Lima JA, Cushman M, *et al.* Risk factors for the progression of coronary artery calcification in asymptomatic subjects: results from the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation*. 2007;115:2722-30.

43. **Wray NR**, Yang J, Hayes BJ, Price AL, Goddard ME and Visscher PM. Pitfalls of predicting complex traits from SNPs. *Nat Rev Genet*. 2013;14:507-15.

44. **ENCODE Project Consortium**, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C and Snyder M. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57-74.

45. **Rosenbloom KR**, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, *et al.* ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic acids research*. 2013;41:D56-63.

46. **Peters SA**, den Ruijter HM, Bots ML and Moons KG. Improvements in risk stratification for the occurrence of cardiovascular disease by imaging subclinical atherosclerosis: a systematic review. *Heart*. 2012;98:177-84.

Chapter 4

Is common genetic variation at IRS1, ENPP1 and TRIB3 loci associated with cardiometabolic phenotypes in type 2 diabetes? An exploratory analysis of the Verona Newly Diagnosed Type 2 Diabetes Study

Nutr Metab Cardiovasc Dis. 2016 Mar;26(3):232-8

Authors and affiliations are listed in Chapter 6.1

4.1 ABSTRACT

English

Background and Aims - Insulin resistance is a hallmark of type 2 diabetes (T2DM), it is often accompanied by defective beta-cell function (BF) and is involved in the pathophysiology of cardiovascular disease (CVD). Commonalities among these traits may recognize a genetic background, possibly involving the genetic variation of insulin signaling pathway genes. We conducted an exploratory analysis by testing whether common genetic variability at *IRS1*, *ENPP1* and *TRIB3* loci is associated with cardiovascular risk traits and metabolic phenotypes in T2DM.

Methods and Results - In 597 drug-naïve, GADA-negative, newly-diagnosed T2DM patients we performed: 1) genotyping of 10 independent single-nucleotide polymorphisms covering ~90% of common variability at *IRS1*, *ENPP1* and *TRIB3* loci; 2) carotid artery ultrasound; 3) standard ECG (n=450); 4) euglycaemic insulin clamp to assess insulin sensitivity; 5) 75g-OGTT to estimate BF (derivative and proportional control) by mathematical modeling. False discovery rate of multiple comparisons was set at 0.20. After adjustment for age, sex and smoking status, rs4675095-*T* (*IRS1*) and rs4897549-*A* (*ENPP1*) were significantly associated with carotid atherosclerosis severity, whilst rs7265169-*A (TRIB3)* was associated with ECG abnormalities. Rs858340-*G* (*ENPP1*) was significantly associated with decreased insulin sensitivity, independently of age, sex and body-mass-index. No consistent relationships were found with BF.

Conclusions - Some associations were found between intermediate phenotypes of CVD and common genetic variation of gatekeepers along the insulin signaling

pathway. These results need be replicated to support the concept that in T2DM the CVD genetic risk clock may start ticking long before hyperglycemia appears.

Italian

Premesse e Scopo - L'insulino-resistenza (IR) è una caratteristica peculiare del diabete tipo 2 (DMT2), è coinvolta nella fisiopatologia delle malattie cardiovascolari (CVD) e spesso si accompagna ad una compromessa funzione beta-cellulare (BF). In questo contesto, è possibile che la variabilità genetica comune di alcuni geni coinvolti nella cascata del segnale insulinico possa spiegare, almeno in parte, la relazione esistente tra DMT2, CVD e BF. In questo studio abbiamo pertanto cercato di verificare se la variabilità genetica comune dei loci *IRS1*, *ENPP1* e *TRIB3* è associata a tratti di rischio cardiovascolare e fenotipi metabolici in soggetti affetti da DMT2 arruolati nel Verona Newly Diagnosed Type 2 Diabetes Study (VNDS).

Metodi e Risultati - In 597 soggetti con DMT2 neodiagnosticato, privi di trattamento farmacologico e con negatività degli anticorpi anti-GAD sono stati effettuati: 1) genotipizzazione di 10 polimorfismi indipendenti e selezionati per catturare il 90% della variabilità genetica comune dei loci *IRS1*, *ENPP1* e *TRIB3*, noti quali principali regolatori della cascata del segnale insulinico; 2) eco-Doppler carotideo; 3) ECG standard (n=450); 4) clamp euglicemico iperinsulinemico, gold-standard per la determinazione della sensibilità insulinica; 5) stima della BF con modello matematico nelle sue componenti derivativa e proporzionale. Dopo correzione per età, sesso e abitudine tabagica i polimorfismi rs4675095-*T* (*IRS1*)

and rs4897549-*A* (*ENPP1*) sono risultati significativamente associati a più severa aterosclerosi carotidea, mentre rs7265169-*A (TRIB3)* era associato ad anormalità ischemiche dell'ECG. Dopo correzione per età, sesso e indice di massa corporea, rs858340-*G* (*ENPP1*) era significativamente associati a maggiore IR; non è stata rilevata nessuna relazione significativa tra BF ed i polimorfismi in studio.

Conclusioni - La variabilità genetica comune dei principali geni regolatori della cascata del segnale insulinico potrebbe spiegare, almeno in parte, l'associazione tra IR e CVD nel DMT2. Benchè siano necessari studi di replicazione in più ampie coorti di soggetti, questi risultati suggeriscono che nel DMT2 il rischio genetico per CVD verosimilmente agisce ben prima che il fenotipo clinico dell'iperglicemia si manifesti ed esercita i propri effetti sul fenotipo cardio-metabolico sin dalle fasi più precoci della malattia diabetica.

4.2 INTRODUCTION

Insulin resistance (IR) and its associated traits may act as a common soil for both type 2 diabetes mellitus (T2DM) and cardiovascular diseases (CVD) [1, 2]. The genetic annotation of hitherto confirmed allelic variants associated with the risk of T2DM and CVD is hoped to unveil mechanistic explanations at the genetic level of the strong clinical link existing between these highly prevalent complex traits [3].

Some non-synonymous allelic variants harbored in or near insulin signaling pathway genes, namely *ENPP1* (ectonucleotide pyrophosphatase/phosphodiesterase 1), *IRS1* (insulin receptor substrate 1) and *TRIB3* (tribble homolog 3), have raised considerable interest [4-6] given their association with defective insulin action and impaired endothelial cell function.

The very same variants have been recently confirmed as jointly increasing the CVD risk [5], possibly by affecting systemic and endothelial insulin sensitivity [4, 5]. *In vitro* [7, 8] and *in vivo* studies [4, 5, 9] showed that *ENPP1* K121Q (rs1044498), *IRS1* G972R (rs1801278) and *TRIB3* Q84R (rs2295490) are associated with an increased risk of T2DM and CVD through their effect on the endothelial nitric oxide synthase and the fibrinolysis system. However, other studies on the same loci found no significant association with T2DM or intermediate CVD risk traits [10, 11].

Thus, some evidence supports the association of non-synonymous genetic variants of *IRS1*, *TRIB3* and *ENPP1* with overall CVD risk, but their clinical applicability to unambiguously identify high-risk subjects remains a still unanswered question.

IRS1, *ENPP1*, *TRIB3* loci have been associated with CVD and T2DM risk based upon findings related to non-synonymous variants entailing changes of protein function. Genes harboring mutations resulting in rare monogenic disorders often have been found to be involved in the genetic risk of common complex disorders

through common variants, a most striking example being the role played by some MODY genes on T2DM pathogenesis [12]. We hypothesized that something analogous might occur also at *IRS1*, *ENPP1*, *TRIB3* loci. Indeed, their actual role may be wider than the one suggested by their non-synonymous variants and may be due also to common variability. Therefore, we conducted an exploratory analysis in the Verona Newly Diagnosed Type 2 Diabetes Study (VNDS) of 13 common polymorphisms, selected to comprehensively capture the genetic variation at three candidate loci (*ENPP1*, *IRS1* and *TRIB3*) previously reported as being associated with the insulin signaling cascade. We investigated whether they are individually associated with one or more of four (4) outcome traits, including ECG ischemic abnormalities and subclinical atherosclerosis readouts at common carotid artery as representative of the "cardiovascular" domain, and insulin resistance (IR) and beta cell function (BF), as representative of the "metabolic" domain.

4.3 METHODS

4.3.1 Study population

The VNDS is an ongoing study aimed at building a biobank of patients with newly-diagnosed T2DM. A detailed description of the overall experimental approach have been previously published [13] and is available in the online **Supplementary Material**. The research was approved by the Human Investigation Committee of the Verona City Hospital. Each subject signed a written informed consent upon recruitment.

4.3.2 Genotyping

Thirteen independent tag-SNPs of *IRS1* (rs4675095 and rs1801278), *ENPP1* (rs6939185, rs858340, rs1044498, rs9493119, rs4897549) and *TRIB3* (rs6076472, rs6139007, rs7265169, rs2295490, rs12626158, rs6115830) loci were selected to capture at least 90% of the common genetic variability (**Supplementary Fig. S1**) by GEVALT (GEnotype Visualization and ALgorithmic Tool) software. A peripheral blood sample was collected from each patient and DNA was extracted by standard salting-out method. Genotypes were assessed by Veracode technique (Illumina Inc, CA), applying the GoldenGate Genotyping Assay according to manufacturer's instructions. The selected SNPs were in low linkage disequilibrium (LD), with r^2 between the SNPs at each locus comprised between 0.0 and 0.16. Further information are provided in **Supplementary Figure S1** and **Supplementary Table S3**. Ten SNPs were successfully genotyped (rs4675095, rs1801278, rs6939185, rs858340, rs9493119, rs4897549, rs6076472, rs6139007, rs7265169, rs6115830) in all study participants. Genotyping of rs1044498, rs2295490 and rs12626158 failed due to technical issues. None of them was in LD with other genotyped SNPs at the same loci, as provided in **Supplementary Table S3**, so that it was not possible to replace them with other highly correlated SNPs in the association analyses.

4.3.3 Cardiometabolic Phenotyping

A standard 12-lead electrocardiogram (ECG) was performed in 450 study participants (CardioDirect 12 unit; Metasoft 3.9 software). Presence of ischemic abnormalities was recorded according to Minnesota code and categorized as suggestive for "definite", "probable" or "possible" coronary heart disease (CHD) [14].

High-resolution B-Mode echo-color Doppler of common carotid artery was performed by a single operator in 597 subjects with a 10-MHz linear probe with axial resolution of 0.01 mm (Esaote Wall Track System, Esaote S.p.A., Genova, Italy). Common carotid intima-media thickness (CC-IMT) was estimated by scanning the posterior wall of common carotid artery at 1 cm from carotid bifurcation. Patients were classified in three categories: no carotid atherosclerosis; impaired CC-IMT and/or stenosis <40%, stenosis >40%. The cutoff of 40% was adopted based on previous experience [15].

Blood pressure was measured at the upper left arm in all subjects and classified according to 2013 ESH/ESC guidelines.

Metabolic tests were carried out on two separate days in random order. On one day a frequently sampled, prolonged (240 or 300 min) OGTT (75 g) was performed and beta-cell function (BF) was reconstructed by mathematical modeling, as previously described [16]. By this method, BF is described by two parameters:

1. Derivative (or dynamic) control (DC): the response of the beta cell to the rate of increase of plasma glucose;
2. Proportional (or static) control (PC): the response of the beta cell to glucose concentration per se.

On a separate day, a euglycaemic insulin clamp was performed to assess insulin sensitivity [13].

Beta cell modeling and clamp derived insulin sensitivity can be considered reference methods to assess insulin secretion and action in vivo in man. However, to facilitate the comparison with widely used OGTT-derived surrogate indexes, we also report the insulinogenic index, the corrected insulin response and the Matsuda Index of insulin sensitivity (see below) in the **Supplemental Material**.

4.3.4 Analytical Methods

Plasma glucose was measured in duplicate with a Beckman Glucose Analyzer II (Beckman Instruments, Fullerton, CA, USA) or with an YSI 2300 Stat Plus Glucose&Lactate Analyzer (YSI Inc., Yellow Springs, OH, USA), at bedside. Serum C-peptide and insulin were measured by chemiluminescence [13]. GAD-antibodies were measured by immunoradiometry (CentAK, Medipan, Germany); glycosylated haemoglobin and serum lipids by standard in-house methods.

4.3.5 Calculations

The amount of glucose metabolized during the last 60 min of the clamp (*M*-value, reference insulin sensitivity; units: μmol/min/m^2 BSA) was computed with standard formulae [16].

Mathematical modeling of glucose and C-peptide time series following the 75-g oral glucose challenge were performed and DC and PC of BF were computed, as previously described [13]. Modeling details are found in the **Supplemental Material**.

The following OGTT-derived indices were also computed and are reported in the **Supplemental Material**:

1. Insulinogenic Index: $(Insulin_{30'} - Insulin_{0'})/(Glucose_{30'} - Glucose_{0'})$; units: mU/mmol;
2. Corrected Insulin Response at time 120' of the OGTT ($CIR_{120'}$): $Insulin_{120'}/[Glucose_{120'}\cdot(Glucose_{120'} - 3.89)]$; units: $mU\cdot L/mmol^2$;
3. Matsuda Index of insulin sensitivity: $10{,}000/[(Glucose_{0'}\cdot Insulin_{0'})\cdot(\text{Mean OGTT glucose concentration})\cdot(\text{mean OGTT insulin concentration})]^{1/2}$.

4.4 STATISTICAL ANALYSIS

Data are presented as median and interquartile range, unless otherwise indicated. Hardy-Weinberg equilibrium was tested by chi-square test. For each SNP, the allele with the lowest occurrence in our population was considered as the effect allele in the analyses according to an additive model of inheritance. Statistical models were unadjusted (M1) or adjusted (M2) for relevant covariates (age, sex and smoking status for cardiovascular readouts, or age, sex and BMI for metabolic readouts). Logistic regression models were applied to test the association between genotypes and cardiovascular phenotypes (outcome traits: carotid artery atherosclerosis, ECG ischemic abnormalities). Only a genetic additive model was tested. Generalized Linear Models, as implemented in the SPSS software, with or without repeated measures as appropriate, were carried out to test the associations between genetic variability and metabolic traits (log-transformed or square-root transformed, if needed, unless the latter displayed strong deviation from the Gaussian distribution which could not be corrected by transformation. In the latter case, which applied only to DC of BF, non-parametric (Kruskal-Wallis) tests were applied with no correction for covariates).

Since the present study is an exploratory analysis, the control of the family wise error rate (FEW) of 50 multiple comparisons (10 SNPs by 5 outcome variables: ECG, carotid atherosclerosis, clamp-assessed insulin sensitivity and DC and PC of BF), according to Bonferroni's correction, was deemed too conservative. Thus, we applied the two stage step-up linear procedure of Benjamini-Krieger-Yekutieli (BKY) [17], a recent evolution of the Benjamini-Hochberg's method, to control the false discovery rate (FDR) (see **Supplemental Material** for details). Selections of FDRs ranging from 0.25 to 0.05 can be found in previous papers, the lower figure being conceptually analogous to Bonferroni's correction in strongly favouring protection against false positive results vs. the risk of declaring

false negative findings. We selected an FDR of 0.20, the highest acceptable FDR according to Benjamini and Yekutieli [18] in line with the exploratory nature of this work. All statistical tests were performed by the SPSS 22.0 software.

4.5 RESULTS

In this study we report the data collected in 597 VNDS patients: anthropometric, clinical and metabolic features are summarized in **Table I**. Ten independent SNPs were successfully genotyped and were all in Hardy-Weinberg equilibrium (**Table II**). Rs1044498, rs2295490 and rs12626158 were not included in the statistical analyses due to poor quality genotyping.

At a FDR set at 0.20, the BKY procedure [17] rejected 5 null hypotheses, i.e. accepted 5 results as statistically significant. These are presented herein below. Other findings with nominal (i.e. $p<0.05$) statistical significance, but not rejected by BKY, are not presented.

Rs4897549-*A* (*ENPP1*) and rs4675095-*T* (*IRS1*) were significantly associated with a greater severity of carotid atherosclerosis (p=0.01 and p=0.009, respectively), independently of age, sex and smoking status (**Table III, panel A and B**)). In a secondary analysis, rs4897549-*A* and rs4675095-*T* were jointly tested for independent association with the presence of carotid atherosclerosis (p=0.009 and p=0.014, respectively, **Table III, panel D**). Their associations with carotid atherosclerosis were confirmed to be independent with odd ratios almost superimposable to the ones found in the previous analysis. Rs7265169 (*TRIB3*) was associated with ECG ischemic abnormalities (**Table III, panel C**).

The *G* major allele of rs858340 in *ENPP1* was significantly (p=0.008) associated with impaired insulin sensitivity (**Figure 1**). Furthermore**,** rs6939185 (*ENPP1*) was associated with altered derivative control of BF (*p*=0.024, by Kruskal-Wallis). However, the derivative control (units: pmol per square meter of BSA; median [I.Q. range]) was 488 [105 - 1054] in rs6939185 GG carriers, 377 [0 - 821] in AG carriers and 565 [125 - 1187] in AA carriers. Thus, the relationship appeared biologically inconsistent and will not be discussed further.

4.6 DISCUSSION

In this study we explored the possible associations of common genetic variation in or near three loci harboring insulin signaling pathway genes with cardiometabolic phenotypes in a well-characterized sample of patients with newly-diagnosed T2DM. Since diabetes itself is equipotent to previous stroke or MI in determining mortality risk, we selected for our analysis the two metabolic phenotypes, i.e. insulin sensitivity and beta cell function (BF), which are at the core of the pathophysiology of T2DM, and two cardiovascular intermediate phenotypes (ECG and carotid artery ultrasound scan) which are in widespread clinical use.

On the genetic side, given the role of insulin action in many cells involved in atherogenesis, we assessed the common genetic variability of three known gatekeepers of insulin signaling, *ENPP1*, *IRS1* and *TRIB3,* previously associated with cardiovascular risk and/or intermediate phenotypes through their non-synonymic variants.

We report that rs4897549 (*ENPP1*) and rs4675095 (*IRS1*) were independently associated with increased carotid atherosclerosis (**Table III, panel A and B and**

D), while rs7265169 (*TRIB3*) was associated with ECG ischemic abnormalities (**Table III, panel C**). Interestingly, previous *in vitro* studies suggested that alterations in insulin pathway genes promote endothelial dysfunction, hence predisposing to CVD [19, 20]. Additionally, while previous studies showed that the non-synonymic *ENPP1* K121Q, *IRS1* G972R and *TRIB3* Q84R variants were associated with cardiovascular events in human subjects [19-21], our findings support the hypothesis that the variance of subclinical CVD risk traits may be explained to some extent also by common genetic variation in or near these loci as well.

Rs858340 at *ENPP1* was associated with insulin resistance (**Figure 1**), thus extending previous observations on the detrimental effect of *ENPP1* genetic variability on insulin action [22]. Unfortunately, failure of rs1044498 (*ENPP1*) genotyping – for which a large body of literature exists [5, 9] – did not allow to directly test this specific variant against the outcome traits. Owing to the selection criteria used to pick up the SNPs for the current study it was impossible to replace rs1044498 with other nearby, highly-correlated SNPs (using a r^2 threshold of 0.6, at the minimum). However, since the genetic variants selected for our analyses captured a large proportion of the genetic variability at each locus, our results further strengthen the message that multiple IR-associated variants are harbored within the LD block structure of *ENPP1* locus.

None of the variants at *IRS1* and *TRIB3* was associated with insulin resistance. This is in line with some studies on *IRS1* G972R (rs1801278) conducted in T2DM and non-T2DM humans [23, 24]. However, a more extensive documentation in humans and rodents [25-27], although mostly based on surrogate insulin-sensitivity indices, supports the negative effect of some *IRS1* and *TRIB3* variants on insulin sensitivity *in vitro* and *in vivo*.

Inspired by the case of some MODY genes, which harbour common genetic variants affecting T2DM risk [12], and by the primary role played by insulin secretion in the pathogenesis of T2DM, we tested the genetic variation of *ENPP1*, *TRIB3* and *IRS1* for association with BF. One polymorphism in *ENPP1* (rs6939185) showed a significant association with decreased derivative control of BF. However, its relationship to BF appeared to be biologically inconsistent. Indeed, our FDR threshold (0.20) is compatible with one out of 5 accepted findings being spurious. Thus, common genetic variability of *ENPP1* was associated with both carotid atherosclerosis (through rs4897549) and insulin resistance (through rs858340).

No significant relationships between common genetic variability of *IRS1*, or *ENPP1* or *TRIB3* and beta cell function could be claimed, in spite of some nominal statistical significance (data not shown). Previous *in vitro* and *in vivo* studies showed that *ENPP1* and *TRIB3* affect beta-cell survival [28, 29], thereby being potentially able to affect beta-cell mass. Indeed, the gain-of-function variants *ENPP1* K121Q (rs1044498) and *TRIB3* Q84R (rs2295490) increase beta cell apoptosis [20, 28]. However, no associations between changes in beta cell function and common genetic variability of these loci can be declared on the basis of our findings.

Taken together, our exploratory analysis suggests that common genetic variability of *ENPP1* may play a pivotal role in cardiometabolic phenotypes, in that it may be implicated in carotid atherosclerosis and insulin resistance. Genetic variability at *IRS1* and *TRIB3* may play independent roles in carotid atherosclerosis and ischemia-related alterations in ECG, respectively.

The most important strength of our study is the use of state-of-art methods to assess insulin sensitivity and BF, instead of surrogate markers, in a large sample of patients with newly-diagnosed T2DM. Furthermore, study participants were 60

years old on average (enough to convey the effect, if any, of low-penetrance genetic determinants) and presumably not yet influenced by long-standing hyperglycemia or pharmacological glucose-lowering treatment, which could modify phenotype-genotype interaction.

However, our study has some limitations: i- it is not a population based study, although the VNDS cohort is fairly representative of Italian patients with T2DM [14]; ii- the cohort includes a large number of men, possibly reflecting a gender-related referral bias; iii- the relatively low number of patients may have limited the statistical power; iv- the absence of a replication cohort cautions against the generalizability of our findings. Therefore, we have consulted the publicly available MAGIC Consortium database looking for association results relative to surrogate indexes of insulin sensitivity (HOMA-IR, IR-Homeostatic Model Assessment and ISI, Insulin Sensitivity Index) and BF (HOMA-B, BF-Homeostatic Model Assessment and $CIR_{30'}$, Corrected Insulin Response) [30, 31]. As a result, among the four relevant SNPs, rs4675095-*A* (*IRS1*) was negatively associated with HOMA-IR and HOMA-B (at p=1.17xE-04 and p=4.2x10E-3, respectively), while there was no robust association with ISI (p=0.08) or $CIR_{30'}$ (p=0.94); rs858340-*T* (*ENPP1*) was nominally associated with lower HOMA-IR (p=0.031) but not with ISI (p=0.079) (**Supplementary Table S6-S9**). Among the other SNPs considered, it is worth to mention the positive association of rs6139007-*T* (*TRIB3*) with HOMA-IR (p=0.029) and the negative association of rs9493119-*A* (*ENPP1*) with $CIR_{30'}$ (0.003). Hence, none of the results from our analysis was robustly replicated in the MAGIC database, which, once again, highlights both the intrinsic limitations of surrogate indexes and the relatively limited statistical power of our database.

4.7 CONCLUSIONS

In summary, in this exploratory analysis, *IRS1*, *ENNP1* and *TRIB3*, known to be associated with T2DM and harboring genes playing a prominent role in mediating insulin signaling, may modulate a number of cardiometabolic phenotypes in patients of Italian ancestry with newly-diagnosed T2DM. Although replication studies in separate deeply-phenotyped cohorts are needed to corroborate our results, our findings suggest that *ENPP1* may be a genetic locus potentially implicated in the association both with insulin resistance in diabetes and with cardiovascular disease. Other gatekeepers of the insulin signaling pathway, specifically *IRS1* and *TRIB3*, might play additive roles in the pathogenesis of cardiovascular complications in patients with T2DM. Finally, since our study was conducted in patients with newly diagnosed T2DM, our findings are compatible with the hypothesis that the genetic clock of CVD in T2DM may start ticking long before the onset of overt diabetic hyperglycemia.

Acknowledgements

Replication data on glycemic traits have been contributed by MAGIC investigators and have been downloaded from www.magicinvestigators.org. Monica Zardini and Federica Moschetta (Department of Medicine, University of Verona) provided superb technical help. We thank Massimo Delledonne, Alberto Ferrarini and Chiara Zusi for genotyping assistance at the Center of Functional Genomics (University of Verona).

4.8 FIGURES and TABLES

Table I. Clinical and metabolic features of the VNDS population.

Variable		ALL
N (M/F)		597 (415/182)
Age (years)		59 [52-66]
BMI ($Kg \cdot m^{-2}$)		29.3 [26.5-32.8]
Waist (cm)		100 [93-108]
Smokers (%)*		47.9 [44.8-50.9]
$HbA_{1c\ DCCT}$ (%)		6.7 [6.2-7.5]
$HbA_{1c\ IFCC}$ (mmol/mol)		49.7 [44.3-58.5]
Common carotid artery atherosclerosis (*N=597*)		
absent (%)*		34.7 [31.8-37.6]
impaired IMT or stenosis <40%		59.8 [56.8-62.8]
stenosis ≥40%		5.5 [4.1-6.9]
ECG ischemic abnormalities[§] (*N=489*)		
absent (%)*		70.6 [67.8-73.4]
possible		20.9 [18.4-23.4]
probable		2.2 [1.3-3.1]
definite		6.3 [4.8-7.8]
Insulin Sensitivity (*N=597*)		
M-clamp (µmol/min/m^2 BSA)		607 [380-865]
Matsuda Index		3.0 [2.1-4.7]
Beta-cell Function (*N=595*)		
Derivative Control ($pmol/m^2$ BSA)·$(mmol \cdot L^{-1} \cdot min^{-1})^{-1}$		444 [68-938]
Proportional Control ($pmol \cdot min^{-1} \cdot m^{-2}$ BSA)	**$ISR_{5.5}$**	151 [110-191]
	ISR_8	206 [149-282]
	ISR_{11}	326 [228-473]
	ISR_{15}	510 [331-764]
	ISR_{20}	750 [461-1143]

Data expressed as median and interquartile range [IQR]. *Percentage is given as point estimate at 95% confidence. §ECG abnormalities are classified according to the Minnesota code. BMI, Body Mass Index; SBP, Systolic Blood Pressure; DBP, Diastolic Blood Pressure; HbA_{1c} DCCT, Diabetes Control and Complication Trial-Aligned Hemoglobin A_{1c}; HbA_{1c} IFCC, International Federation of Clinical Chemistry-Aligned Hemoglobin A_{1c}; IMT, Intima-Media Thickness. ISR_n, Insulin Secretion Rate at any given (*n*) mM of plasma glucose.

Table II. Properties of the 10 genetic variants included in the VNDS association analyses.

Locus	Chr	SNP	MAF*	Alleles	H-W Equlibrium
				Minor[§]/Other	*Y/N, (p)*
IRS1	2	rs4675095	0.09	**A**/T	Y, (0.66)
IRS1	2	rs1801278	0.07	**A**/G	Y, (0.57)
ENPP 1	6	rs858340	0.27	**A**/G	Y, (0.40)
ENPP 1	6	rs6939185	0.37	**A**/G	Y, (0.84)
ENPP 1	6	rs9493119	0.06	**G**/A	Y, (0.051)
ENPP 1	6	rs4897549	0.28	**A**/G	Y, (0.51)
TRIB3	20	rs6139007	0.28	**G**/A	Y, (0.11)
TRIB3	20	rs7265169	0.09	**A**/C	Y, (0.87)
TRIB3	20	rs6115830	0.47	**A**/G	Y, (0.17)
TRIB3	20	rs6076472	0.31	**C**/A	Y, (0.15)

* MAF, Minor Allele Frequency in the VNDS study population. [§] Minor allele (considered as effect allele, in bold) defined according to the MAF in our population. H-W, Hardy Weinberg equilibrium; *Y*, yes; *N*, no.

Table III. **Effects of rs4897549-*A* (*ENPP1*), rs4675095-*T* (*IRS1*) and rs7265169-*A* (*TRIB3*) alleles on measured CVD risk traits in the VNDS study participants.**

A - Odds ratio of carotid atherosclerosis in carriers of rs4897549-*A* allele (*ENPP1*). **B**- Odds ratio of carotid atherosclerosis in carriers of rs4675095-*T* allele (*IRS1*). **C**- Odds ratio of abnormal ECG in carriers of rs7265169-*A* allele (*TRIB3*). **D**- Simultaneous, independent association of rs4897549-*A* (*ENPP1*) and rs4675095-*T* (*IRS1*) alleles with carotid atherosclerosis. ECG abnormalities were classified as suggestive for "probable" or "definite" coronary heart disease (CHD) according to the Minnesota code [14]. All analyses were adjusted for age, sex and smoking status.

A.	**rs4897549-*A* (*ENPP1*)**		
Phenotype	**adjusted OR**	**95% C.I.**	***P*-value**
Impaired CC-IMT and/or stenosis <40%	1.26	0.92 – 1.72	0.15
Stenosis > 40%	2.64	1.40 – 4.96	0.003
			0.01 ($P_{overall}$)

B.	**rs4675095-*T* (*IRS1*)**		
Phenotype	**adjusted OR**	**95% C.I.**	***P*-value**
Impaired CC-IMT and/or stenosis <40%	2.05	1.27 – 3.32	0.003
Stenosis > 40%	2.69	0.86 – 8.47	0.09
			0.009 ($P_{overall}$)

C.	**rs7265169-*A* (*TRIB3*)**		
Phenotype	**adjusted OR**	**95% C.I.**	***P*-value**
"Probable" CHD	1.79	1.05 – 3.04	0.033
"Definite" CHD	2.9	1.29 – 6.56	0.01
			0.014 ($P_{overall}$)

D.	**rs4897549-*A* (*ENPP1*)**			**rs4675095-T (*IRS1*)**		
Phenotype	**adj OR**	**95% C.I.**	***P*-value**	**adj OR**	**95% C.I.**	***P*-value**
Impaired CC-IMT and/or stenosis <40%	1.23	0.89 – 1.69	0.21	2.08	1.28 – 3.39	0.003
Stenosis > 40%	2.55	1.36 – 4.81	0.004	2.59	0.81 – 8.27	0.11
			0.014 ($P_{overall}$)			0.009 ($P_{overall}$)

Figure 1 – Association between clamp-assessed insulin sensitivity and rs858340-*G* (*ENPP1*).

The G allele of rs858340 (*ENPP1*) is significantly associated with impaired insulin sensitivity in patients with newly diagnosed type 2 diabetes (p=0.008, after correction for age, sex and BMI). Data are presented as median and interquartile range. BSA, body surface area.

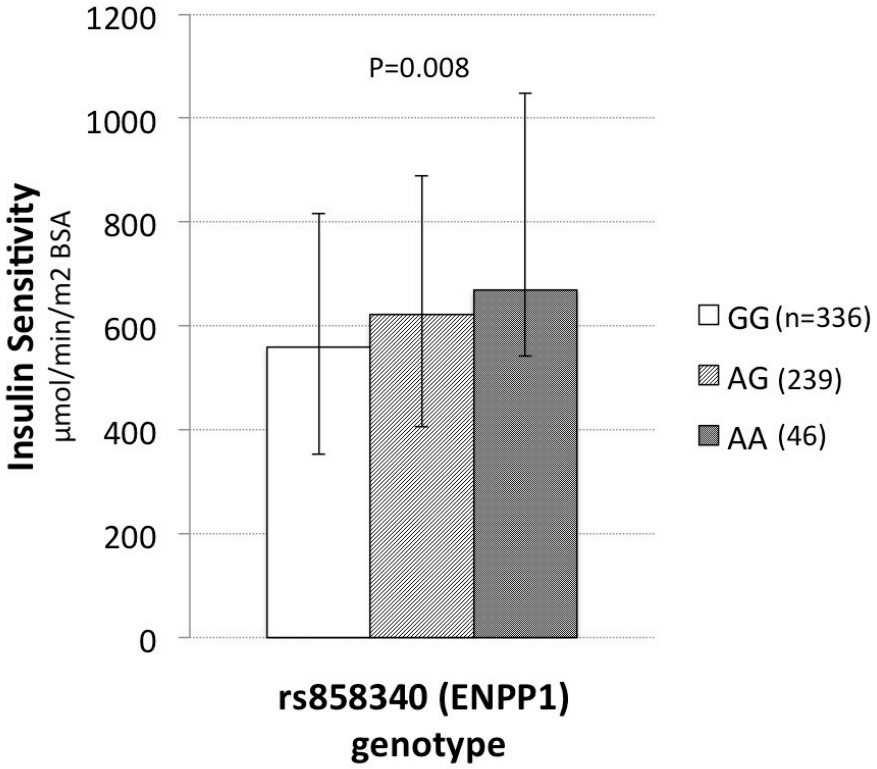
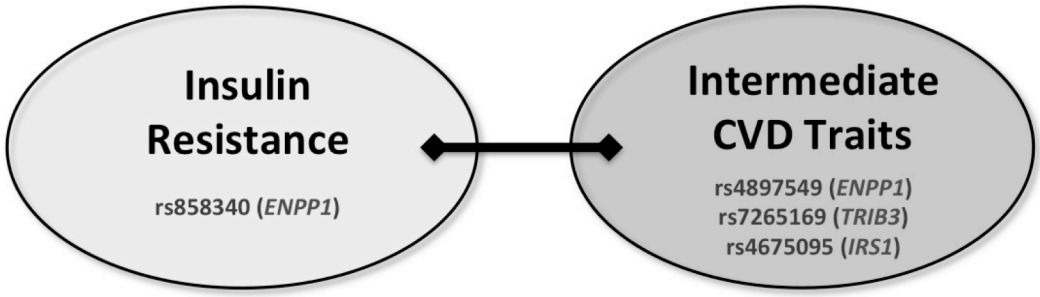

1200
1000
800
600
400
200
0
P=0.008
Insulin Sensitivity
µmol/min/m2 BSA
GG (n=336)
AG (239)
AA (46)
rs858340 (ENPP1)
genotype

Figure 2 – Relationships of intermediate cardiovascular disease traits and insulin resistance with genetic variants at insulin signalling pathway loci.

The figure highlights the role of insulin resistance (IR, light grey set) as a well-established predictor of cardiovascular disease (CVD, dark grey set). In this context, our study suggests that the common genetic variability of *ENPP1* is associated with IR and intermediate CVD phenotypes, while *TRIB3* is associated with ECG ischemic abnormalities and *IRS1* with carotid atherosclerosis.



4.9 REFERENCES

1. **DeFronzo RA**, Ferrannini E: Insulin resistance. A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease. *Diabetes Care* 1991, 14(3):173-194.
2. **Stern MP**: Diabetes and cardiovascular disease. The "common soil" hypothesis. *Diabetes* 1995, 44(4):369-374.
3. **Dauriz M**, Meigs JB: Current Insights into the Joint Genetic Basis of Type 2 Diabetes and Coronary Heart Disease. *Current cardiovascular risk reports* 2014, 8(1):368.
4. **Bacci S**, Prudente S, Copetti M, Spoto B, Rizza S, Baratta R, Di Pietro N, Morini E, Di Paola R, Testa A *et al*: Joint effect of insulin signaling genes on cardiovascular events and on whole body and endothelial insulin resistance. *Atherosclerosis* 2013, 226(1):140-145.
5. **Prudente S**, Morini E, Trischitta V: Insulin signaling regulating genes: effect on T2DM and cardiovascular risk. *Nat Rev Endocrinol* 2009, 5(12):682-693.
6. **Rung J**, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proenca C, Bacot F, Balkau B, Belisle A, Borch-Johnsen K *et al*: Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nat Genet* 2009, 41(10):1110-1115.
7. **Andreozzi F**, Formoso G, Prudente S, Hribal ML, Pandolfi A, Bellacchio E, Di Silvestre S, Trischitta V, Consoli A, Sesti G: TRIB3 R84 variant is associated with impaired insulin-mediated nitric oxide production in human endothelial cells. *Arterioscler Thromb Vasc Biol* 2008, 28(7):1355-1360.
8. **Formoso G**, Di Tomo P, Andreozzi F, Succurro E, Di Silvestre S, Prudente S, Perticone F, Trischitta V, Sesti G, Pandolfi A *et al*: The TRIB3 R84 variant is associated with increased carotid intima-media thickness in vivo and with enhanced MAPK signalling in human endothelial cells. *Cardiovasc Res* 2011, 89(1):184-192.
9. **Tang ST**, Shen XR, Tang HQ, Wang CJ, Wei W, Zhang Q, Wang Y: Association of the ENPP1 K121Q polymorphism with susceptibility to type 2 diabetes in

different populations: evidence based on 40 studies. *Endocr J* 2014, 61(11):1093-1103.

10. **Lim S**, Hong J, Liu CT, Hivert MF, White CC, Murabito JM, O'Donnell CJ, Dupuis J, Florez JC, Meigs JB: Common variants in and near IRS1 and subclinical cardiovascular disease in the Framingham Heart Study. *Atherosclerosis* 2013, 229(1):149-154.
11. **Lyon HN**, Florez JC, Bersaglieri T, Saxena R, Winckler W, Almgren P, Lindblad U, Tuomi T, Gaudet D, Zhu X *et al*: Common variants in the ENPP1 gene are not reproducibly associated with diabetes or obesity. *Diabetes* 2006, 55(11):3180-3184.
12. **Vaxillaire M**, Bonnefond A, Froguel P: The lessons of early-onset monogenic diabetes for the understanding of diabetes pathogenesis. *Best practice & research Clinical endocrinology & metabolism* 2012, 26(2):171-187.
13. **Bonetti S**, Trombetta M, Boselli ML, Turrini F, Malerba G, Trabetti E, Pignatti PF, Bonora E, Bonadonna RC: Variants of GCKR affect both beta-cell and kidney function in patients with newly diagnosed type 2 diabetes: the Verona newly diagnosed type 2 diabetes study 2. *Diabetes Care* 2011, 34(5):1205-1210.
14. **Bonora E**, Targher G, Formentini G, Calcaterra F, Lombardi S, Marini F, Zenari L, Saggiani F, Poli M, Perbellini S *et al*: The Metabolic Syndrome is an independent predictor of cardiovascular disease in Type 2 diabetic subjects. Prospective data from the Verona Diabetes Complications Study. *Diabet Med* 2004, 21(1):52-58.
15. **Willeit J**, Kiechl S, Oberhollenzer F, Rungger G, Egger G, Bonora E, Mitterer M, Muggeo M: Distinct risk profiles of early and advanced atherosclerosis: prospective results from the Bruneck Study. *Arterioscler Thromb Vasc Biol* 2000, 20(2):529-537.
16. **Bonadonna RC**, Heise T, Arbet-Engels C, Kapitza C, Avogaro A, Grimsby J, Zhi J, Grippo JF, Balena R: Piragliatin (RO4389620), a novel glucokinase activator, lowers plasma glucose both in the postabsorptive state and after a glucose challenge in patients with type 2 diabetes mellitus: a mechanistic study. *J Clin Endocrinol Metab* 2010, 95(11):5028-5036.

17. **Benjamini Y**, Krieger AM, Yekutieli D: Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* 2006, 93(3):491-507.
18. **Benjamini Y**, Yekutieli D: Quantitative trait loci analysis using the false discovery rate. *Genetics* 2005, 171(2):783-789.
19. **Marini MA**, Frontoni S, Mineo D, Bracaglia D, Cardellini M, De Nicolais P, Baroni A, D'Alfonso R, Perna M, Lauro D *et al*: The Arg972 variant in insulin receptor substrate-1 is associated with an atherogenic profile in offspring of type 2 diabetic patients. *J Clin Endocrinol Metab* 2003, 88(7):3368-3371.
20. **Prudente S**, Sesti G, Pandolfi A, Andreozzi F, Consoli A, Trischitta V: The mammalian tribbles homolog TRIB3, glucose homeostasis, and cardiovascular diseases. *Endocrine reviews* 2012, 33(4):526-546.
21. **Gong HP**, Wang ZH, Jiang H, Fang NN, Li JS, Shang YY, Zhang Y, Zhong M, Zhang W: TRIB3 functional Q84R polymorphism is a risk factor for metabolic syndrome and carotid atherosclerosis. *Diabetes Care* 2009, 32(7):1311-1313.
22. **Grarup N**, Urhammer SA, Ek J, Albrechtsen A, Glumer C, Borch-Johnsen K, Jorgensen T, Hansen T, Pedersen O: Studies of the relationship between the ENPP1 K121Q polymorphism and type 2 diabetes, insulin resistance and obesity in 7,333 Danish white subjects. *Diabetologia* 2006, 49(9):2097-2104.
23. **Koch M**, Rett K, Volk A, Maerker E, Haist K, Deninger M, Renn W, Haring HU: Amino acid polymorphism Gly 972 Arg in IRS-1 is not associated to lower clamp-derived insulin sensitivity in young healthy first degree relatives of patients with type 2 diabetes. *Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association* 1999, 107(5):318-322.
24. **Almind K**, Bjorbaek C, Vestergaard H, Hansen T, Echwald S, Pedersen O: Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin-dependent diabetes mellitus. *Lancet* 1993, 342(8875):828-832.
25. **Baroni MG**, Arca M, Sentinelli F, Buzzetti R, Capici F, Lovari S, Vitale M, Romeo S, Di Mario U: The G972R variant of the insulin receptor substrate-1 (IRS-1) gene, body fat distribution and insulin-resistance. *Diabetologia* 2001, 44(3):367-372.

26. **Hribal ML**, Tornei F, Pujol A, Menghini R, Barcaroli D, Lauro D, Amoruso R, Lauro R, Bosch F, Sesti G *et al*: Transgenic mice overexpressing human G972R IRS-1 show impaired insulin action and insulin secretion. *Journal of cellular and molecular medicine* 2008, 12(5B):2096-2106.

27. **Prudente S**, Hribal ML, Flex E, Turchi F, Morini E, De Cosmo S, Bacci S, Tassi V, Cardellini M, Lauro R *et al*: The functional Q84R polymorphism of mammalian Tribbles homolog TRB3 is associated with insulin resistance and related cardiovascular risk in Caucasians from Italy. *Diabetes* 2005, 54(9):2807-2811.

28. **Di Paola R**, Caporarello N, Marucci A, Dimatteo C, Iadicicco C, Del Guerra S, Prudente S, Sudano D, Miele C, Parrino C *et al*: ENPP1 affects insulin action and secretion: evidences from in vitro studies. *PloS one* 2011, 6(5):e19462.

29. **Prudente S**, Scarpelli D, Chandalia M, Zhang YY, Morini E, Del Guerra S, Perticone F, Li R, Powers C, Andreozzi F *et al*: The TRIB3 Q84R polymorphism and risk of early-onset type 2 diabetes. *J Clin Endocrinol Metab* 2009, 94(1):190-196.

30. **Prokopenko I**, Poon W, Magi R, Prasad BR, Salehi SA, Almgren P, Osmark P, Bouatia-Naji N, Wierup N, Fall T *et al*: A central role for GRB10 in regulation of islet function in man. *PLoS Genet* 2014, 10(4):e1004235.

31. **Dupuis J**, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL *et al*: New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010, 42(2):105-116.

Chapter 5

A Genetic Risk Score of 96 Variants linked with Type 2 Diabetes and Cardiometabolic Risk Traits is Associated with Cardiovascular Mortality in 29-years Follow-up of the Framingham Heart Study

Presented at the American Diabetes Association 75th Scientific Meeting

June 5-9, 2015 - Boston, MA, USA

Authors and affiliations are listed in Chapter 7.1

5.1 ABSTRACT

English

Background and Aims - Cardiovascular diseases (CVD) are a major cause of death and are often associated with type 2 diabetes (T2D). Genome-wide studies (GWAS) identified loci associated with T2D, CVD and traits leading to early death. We investigated whether these loci in aggregate carry a higher risk of all-cause and CVD mortality in the FHS.

Methods - We computed an unweighted genetic risk score (GRS) of 96 variants selected by effect-size within respective GWAS to represent the top 25% of GWAS variants for the following traits: T2D, coronary artery disease, myocardial infarction (MI), stroke, sudden cardiac death, heart rate, long QT-interval, heavy smoking and 15-years all-cause mortality. We used pooled logistic regressions with genetic-only (GRS adjusted for sex) and full CVD risk factors adjusted models (sex, age, smoking, prevalent non-fatal CVD) to test the association of 96-GRS with all-cause and MI/stroke mortality in 3,426 FHS participants across 29 years follow-up ($p<0.025$ ($p=0.05/2$) for significance).

Results - Prevalence of non-fatal CVD, T2D and smoking was 7.5, 6.1 and 26.4% at baseline and 18.5, 15.9 and 13.2%, respectively, at the beginning of the last period considered. Cumulative incidence of fatal MI/stroke and all-cause mortality was 5.1 and 22.5%, respectively. The 96-GRS was associated with MI/stroke mortality in both genetic-only (OR[95%CI]: 1.04[1.0-1.1], p=0.006) and fully adjusted model (1.04[1-1.1], p=0.009). Association with all-cause mortality did not reach our statistical significance criteria (1.01[1-1.03], p=0.029, genetic-only; 1.02[1-1.03], p=0.034, fully adjusted).

Conclusions - An aggregate burden of 96 GWAS variants with the largest effect size on cardiometabolic traits is predictor of MI/stroke death in longitudinal analysis of a large population of European ancestry. Further studies need to specify the impact of cardiometabolic disease genetics on current mortality prediction models.

Italian

Premesse e Scopo - Il diabete tipo 2 (T2D) è una malattia complessa ad alta prevalenza e incidenza che riconosce fattori genetici e non-genetici quali determinanti causali. Le malattie cardiovascolari (CVD) sono una delle maggiori cause di morte e sono spesso associate a T2D. Studi di associazione genome-wide hanno identificato varianti genetiche comuni associate a T2D, CVD e fenotipi cardiometabolici intermedi. Questo studio verifica l'ipotesi se il rischio genetico per T2D e tratti di rischio cardiometabolico si associno ad aumentata mortalità nello studio Framingham.

Metodi e Risultati - Popolazione: 3,426 soggetti arruolati nello studio FHS e seguiti con follow-up massimo di 29 anni. E' stato calcolato un GRS_{96} composto da 96 tag-SNPs selezionati, in base al rispettivo effect-size, per essere rappresentativi del quartile più alto per ciascuno dei seguenti tratti all'interno dei rispettivi GWAS: T2D, malattie coronariche, infarto miocardico (MI), stroke, morte cardiaca improvvisa, frequenza cardiaca, QT-lungo, abitudine tabagica, mortalità per tutte le cause a 15 anni. Analisi: regressione logistica "pooled". Modelli: M1=GRS+sex; M2: M1+età, fumo, storia personale di CVD non fatale.

Endpoints: (1) mortalità per tutte le cause; (2) mortalità per MI/stroke. Significatività: $p<0.025$ (=0.05/2).

Conclusioni - La prevalenza di CVD non fatali, T2D e tabagismo era pari a 7.5, 6.1 e 26.4% al baseline (Pool I) e 18.5, 15.9 e 13.2%, rispettivamente, all'inizio dell'ultimo periodo considerato (Pool IV). L'incidenza cumulativa di MI/stroke fatali e mortalità per tutte le cause era 5.1 e 22.5%, rispettivamente. Il GRS_{96} era associato a mortalità per MI/stroke sia nel modello M1 (OR [95%CI]: 1.04 [1-1.1], p=0.006), sia nel modello M2 (1.04 [1-1.1], p=0.009). L'associazione con mortalità per tutte le cause non raggiungeva la significatività statistica (1.01 [1.0-1.03], p=0.029, modello M1; 1.02 [1.0-1.03], p=0.034, modello M2).

5.2 INTRODUCTION

Gauging how various diseases and injuries are affecting the living is a way to measure the effectiveness of a country's health system and to better (re)direct human and economic resources to effectively improve the public health.

Type 2 diabetes (T2D) and cardiovascular diseases (CVD) are well known to be clinically associated in adults [1, 2] and are becoming an increasing public health and economic scourge in US [3] and worldwide [4-6]. Moreover, cardiometabolic diseases are presently the major cause of death, according to a WHO 2013 report [7], and multiple large genome-wide association studies (GWAS) have thus far identified numerous loci associated with cardiometabolic diseases and conditions that often lead to early death. It can be reasonably hypothesized that if index SNPs at these loci are marking genes or regions with important functional significance for health, then the aggregate burden of these variants should be associated with an increased risk of mortality.

This project sought to investigate in the FHS SHARe Study sample a "mortality" genetic risk score comprised of up to 96 candidate single nucleotide polymorphisms (SNPs) identified from published GWAS. The principle behind the selection of relevant SNPs was built upon a thorough literature review of currently available evidence (as of late 2014) from GWAS on mortality associated traits or phenotypes, including a spectrum of cardiometabolic diseases spanning from cardiovascular disease risk factors, particularly T2D, to overt cardiovascular diseases [8-17].

We hypothesized that a genetic risk score summed from up to 96 mortality phenotype-associated SNPs identified from published large scale GWAS would be associated with all-cause and cardiovascular disease mortality in the FHS mortality follow-up data.

This project was initiated with the expectation to provide orthogonal information to usual all-cause and CVD mortality risk prediction rules, which are commonly based on clinical measures and information. This approach would be of valuable practical utility to link the available genetic knowledge on cardiometabolic risk and all-cause mortality with the clinical tools currently in use to stratify the cardiometabolic risk and to manage disease care and prevention at the level of single individual and on larger population scale.

5.3 METHODS

Multiple GWAS have identified loci associated with diseases and conditions that often lead to early death. Our literature review identified about 96 SNPs potentially associated with mortality. To test this hypothesis we have used SHARe genotypes for 96 index SNPs from 96 loci reported in large-scale cardiometabolic disease / cardiovascular disease GWAS to be associated with myocardial infarction or coronary heart disease (CHD) death, stroke or stroke death, ECG-measured high heart rate or prolonged QRS interval or sudden cardiac death, T2D, heavy cigarette smoking behavior, or all-cause mortality. SNPs were selected by an exhaustive literature review of current GWAS and a subsequent assembly of a master list of about 96 top potentially mortality-related SNPs. The SNPs are shown in **Table 1**.

For the 96 mortality SNPs, we have generated a genetic burden risk score by summing the presence of 0, 1 or 2 risk loci across the 96 SNPs of interest, where a higher score indicates a greater burden of potentially mortality-associated loci. The genetic risk score was un-weighted (simple allele counting). The genetic risk score was tested for association with mortality in the FHS SHARe mortality-

linked follow-up data in linear additive genetic models. The primary analysis was focused on all-cause mortality and secondary analysis was focused on cardiovascular disease mortality. Other considerations explored in the analysis included testing mortality from participant entry to the study until death versus from the date of genotyping until death, to try and account for survival biases inherent in mid-life, mid-cohort study collection of genetic information. Models were designed to control versus stratify by prevalent cardiovascular disease and T2D to account for confounding by diseases linked both to the genetic exposure and elevated risk for early mortality.

Statistical analysis

We used pooled logistic regressions with genetic-only (Model 1: GRS adjusted for sex) and full CVD risk factors adjusted models (Model 2: GRS adjusted for sex, age, smoking, prevalent CVD) to test the association of the 96-GRS with all-cause and MI/stroke mortality in 3,426 FHS participants over a 29-years follow-up period. Threshold for significance was declared at $p<0.025$ ($p=0.05/2$, the number of endpoints).

In summary:

- *Study design*: longitudinal, population-based.
- *Study participants*: 3,426 subjects of European Ancestry enrolled in the Framingham Heart Study
- *Exclusion criteria*: none.
- *Exposure*: We computed an un-weighted genetic risk score (GRS) of 96 variants selected by effect-size within respective GWAS to represent the top 25% of GWAS variants for the traits listed in **Table 1**.
- *Endpoints*: (1) All-cause Mortality; (2) Myocardial infarction/Stroke Mortality

Composition of the genetic risk score*		
Phenotype	**SNP#**	**PMID**
Type 2 Diabetes	36	22885922
Coronary Artery Disease	31	23202125
Smoking/cigarettes per day	7	20418890
Myocardial Infarction	7	22397355
Stroke	5	23041239
Sudden Cardiac Death	4	23583979
Heart Rate	3	23593153
Long QT-Interval	3	19305408

*More details are provided in Table 2

5.4 RESULTS

Highlights (full descriptives are available in Table 2)

Genetic Risk Score distribution at study entry

	Pool 1 (1983-1991)
GRS_{96} (whole population)	88.2 (5.6)
GRS_{96} (in people dead by stroke or MI)	89.9 (6.0)
GRS_{96} (in people dead by all- causes)	88.8 (6.5)

Cumulative incidence of deaths over 29-years of follow-up	
Stroke + MI (% of all population)	5.1 %
All-cause (% of all population)	22.5 %

The 96-GRS was significantly associated with Myocardial Infarction/stroke mortality

	OR [95% C.I.]	***P*-value**
Model 1	1.04 [1.0-1.1]	0.006
Model 2	1.04 [1.0-1.1]	0.009

Association with all-cause mortality did NOT reach statistical significance ($p<0.025$)

	OR [95% C.I.]	***P*-value**
Model 1	1.01[1-1.03]	0.029
Model 2	1.02[1-1.03]	0.034

-Model 1: Mortality outcome (either ALL-cause or MI/stroke)= 96-GRS+Sex
-Model 2: Mortality outcome = 96-GRS + sex, age, smoking, prevalent CVD (defined as defined as history of myocardial infarction, angina pectoris, and/or coronary insufficiency)

5.5 DISCUSSION

Type 2 diabetes, prediabetic hyperglycemia and related CVD complications are becoming a dramatically increasing burden for the healthcare system in the U.S. [18]. The genetic knowledge in the field of these complex cardiometabolic diseases is expected to substantially contribute in the refinement of future healthcare strategies and to improve the clinical prediction tools currently available to assess diabetes risk development and related CVD complications and/or mortality. Large-scale genetic studies (GWAS) have successfully outlined the common variants genetic architecture of T2D and many other cardiometabolic diseases in people of European ancestry [17]. Some genetic insights also exist on the common allelic variation of genetic loci found to be associated with mortality-related phenotypes or all-cause mortality [8-17].

Further population-based studies are still needed to specify how the allelic spectrum of cardiometabolic disease and mortality phenotype-associated SNPs can contribute to better define mortality prediction for population and personalized prevention strategies. Study of these questions in FHS SHARe permits analyses and interpretations in a large population sample followed-up over a time frame long enough to detect possible associations, if any, between genetics and mortality.

The growing number of cardiometabolic disease risk loci (especially associated with type 2 diabetes) and mortality phenotype-associated SNPs discovered since the beginning of GWAS era might be a valuable tool requiring further testing to verify the ability to predict mortality. The availability of the FHS SHARe mortality-linked follow-up data affords the unique opportunity to test hypotheses about the ultimate population impact of modern genetic discoveries and variation on the cardiometabolic risk and all-cause mortality in the US population.

In this exercise we successfully tested the association of MI/stroke mortality with a genetic risk score summed from up to 96 mortality phenotype-associated SNPs identified from published large scale GWAS for association with mortality in the FHS SHARe mortality follow-up data.

Although our results are far from being conclusive, as they need to be replicated in independent cohorts, the encouraging results obtained in this pilot study with a restricted set of around 100 SNPs encourage further studies exploring the impact of the entire spectrum of cardiometabolic disease genetic variants on current mortality prediction models, in order to properly represent the most updated GWAS landscape on cardiometabolic risk traits or phenotypes.

As stated above, T2D and CVD are common diseases that often clinically occur together. They carry a high risk of death (either all-cause or due to MI/stroke) and share common risk factors. Both T2D and CVD have a strong genetic background and it is therefore possible that they lead to early (CVD) death through common genetic pathways. Recent evidences have shown that the genetic predisposition to T2D, as modeled by a composite genetic risk score, is associated with increased all-cause mortality risk (Leong A et al., 2016) and with non-fatal and fatal CVD (Borglykke A et al., 2012). It is however currently unknown whether the higher all-cause and MI/stroke death rates carried by the genetic burden of T2D risk occur independently of the individual genetic predisposition to CVD. It might also be possible that the genetic risk for T2D represents a permissive (genetic) environment that paves the way to the overt manifestation of CVD genetic risk. All these questions, therefore, require specific study designs; as such, starting from the pilot analyses of the project herein presented, future studies are advocated to untangle, at a genetic level, the common soil underlying the strong clinical link existing between T2D and CVD.

5.6 TABLES

Table 1 – 96 SNPs included in the "mortality" genetic risk score

#	Nearest gene	SNP	Chr	Risk allele	Disease/Trait	PMID #
1.	*TCF7L2*	rs7903146	10	T	T2D	22885922
2.	*CDKN2A/B*	rs10811661	9	T	T2D	22885922
3.	*CDKAL1*	rs7756992	6	G	T2D	22885922
4.	*SLC30A8*	rs3802177	8	G	T2D	22885922
5.	*THADA*	rs10203174	2	C	T2D	22885922
6.	*FTO*	rs9936385	16	C	T2D	22885922
7.	*IGF2BP2*	rs4402960	3	T	T2D	22885922
8.	*PPARG*	rs1801282	3	C	T2D	22885922
9.	*HMGA2*	rs2261181	12	T	T2D	22885922
10.	*HHEX/IDE*	rs1111875	10	C	T2D	22885922
11.	*ADCY5*	rs11717195	3	T	T2D	22885922
12.	*JAZF1*	rs849135	7	G	T2D	22885922
13.	*ARAP1 (CENTD2)*	rs1552224	11	A	T2D	22885922
14.	*DGKB*	rs17168486	7	T	T2D	22885922
15.	*HNF1B (TCF2)*	rs11651052	17	A	T2D	22885922
16.	*MTNR1B*	rs10830963	11	G	T2D	22885922
17.	*ZBED3*	rs6878122	5	G	T2D	22885922
18.	*IRS1*	rs2943640	2	C	T2D	22885922
19.	*WFS1*	rs4458523	4	G	T2D	22885922
20.	*KHLDC5*	rs10842994	12	C	T2D	22885922
21.	*ANK1*	rs516946	8	C	T2D	22885922
22.	*KCNQ1*	rs163184	11	G	T2D	22885922
23.	*ADAMTS9*	rs6795735	3	C	T2D	22885922
24.	*ZMIZ1*	rs12571751	10	A	T2D	22885922
25.	*CILP2*	rs10401969	19	C	T2D	22885922
26.	*BCAR1*	rs7202877	16	T	T2D	22885922
27.	*UBE2E2*	rs1496653	3	A	T2D	22885922
28.	*HNF1A (TCF1)*	rs12427353	12	G	T2D	22885922
29.	*ANKRD55*	rs459193	5	G	T2D	22885922
30.	*CCND2*	rs11063069	12	G	T2D	22885922
31.	*MC4R*	rs12970134	18	A	T2D	22885922
32.	*HMG20A*	rs7177055	15	A	T2D	22885922
33.	*PRC1*	rs12899811	15	G	T2D	22885922

34.	*SPRY2*	rs1359790	13	G	T2D	22885922
35.	*NOTCH2*	rs10923931	1	T	T2D	22885922
36.	*GCK*	rs4607517	7	A	T2D	22885922
37.	*ZNF365*	rs2077316	10	C	Sudden Cardiac Death (SCD)	23593153
38.	*BAZ2B*	rs4665058	2	A	Sudden Cardiac Death (SCD)	21738491
39.	*RAB3GAP1*	rs6730157	2	G	Sudden Cardiac Death (SCD)	23593153
40.	*HDAC9*	rs2107595	7p21.1	A	STROKE (ischemic)	23041239
41.	*ZFHX3*	rs879324	16q22.3	A	STROKE (ischemic)	23041239
42.	*Intergenic*	rs13407662	2p16.2	T	STROKE (ischemic)	23041239
43.	*NINJ2*	rs12425791	12p13.33	A	STROKE (ischemic)	19369658
44.	*PITX2*	rs6843082	4q25	G	STROKE (ischemic)	23041239
45.	*BDNF*	rs6265	11p14.1	C	Smoking initiation	20418890
46.	*DBH*	rs3025343	9q34.2	G	Smoking cessation	20418890
47.	*CDKN2BAS1*	rs3217992	9p21	A	Coronary artery disease	23202125
48.	*CDKN2BAS1*	rs1333049	9p21	C	Coronary artery disease	23202125
49.	*KCNE2*	rs9982601	21q22.11	T	Coronary artery disease	23202125
50.	*SORT1*	rs602633	1	C	Coronary artery disease	23202125
51.	*WDR12*	rs6725887	2	C	Coronary artery disease	23202125
52.	*ApoE-ApoC1*	rs2075650	19	G	Coronary artery disease	23202125
53.	*PPAP2B*	rs17114036	1p32.2	A	Coronary artery disease	23202125
54.	*LDLR*	rs1122608	19p13.2	G	Coronary artery disease	23202125
55.	*PHACTR1*	rs9369640	6p24.1	C	Coronary artery disease	23202125
56.	*ZC3HC1*	rs11556924	7	C	Coronary artery disease	23202125
57.	*COL4A1, COL4A2*	rs9515203	13	T	Coronary artery disease	23202125
58.	*ADAMTS7*	rs7173743	5	T	Coronary artery disease	23202125
59.	*ABO*	rs579459	9	C	Coronary artery disease	23202125
60.	*PDGFD*	rs974819	11	A	Coronary artery disease	23202125
61.	*SH2B3*	rs3184504	12	T	Coronary artery disease	23202125
62.	*TCF21*	rs12190287	6	C	Coronary artery disease	23202125
63.	*HHIPL1*	rs2895811	14q32	C	Coronary artery disease	23202125
64.	*KIAA1462*	rs2505083	10	C	Coronary artery disease	23202125
65.	*PCSK9*	rs11206510	1	T	Coronary artery disease	23202125
66.	*RASD1, SMCR3, PEMT*	rs12936587	17p11.2	G	Coronary artery disease	23202125
67.	*SLC22A3-LPAL2-LPA*	rs2048327	6	G	Coronary artery disease	23202125
68.	*SLC22A4-SLC22A5*	rs273909	5	C	Coronary artery disease	23202125
69.	*ZNF259, APOA5, APOA1*	rs9326246	11	C	Coronary artery disease	23202125

70.	*APOB*	rs515135	2	G	Coronary artery disease	23202125
71.	*CXCL12*	rs501120	10q11	A	Coronary artery disease	23202125
72.	*MRAS*	rs9818870	3	T	Coronary artery disease	23202125
73.	*ABCG5-ABCG8*	rs6544713	2	T	Coronary artery disease	23202125
74.	*GUCY1A3*	rs7692387	4	G	Coronary artery disease	23202125
75.	*KCNK5*	rs10947789	6	T	Coronary artery disease	23202125
76.	*PLG*	rs4252120	6	T	Coronary artery disease	23202125
77.	*MIA3*	rs17464857	1q41	T	Coronary artery disease	23202125
78.	*CHRNA3*	rs1051730	15q25.1	G	Cigarettes per day (CPD)	20418890
79.	*LOC100188947*	rs1329650	10q23.32	T	Cigarettes per day (CPD)	20418890
80.	*CYP2A6,EGLN2*	rs3733829	19q13.2	G	Cigarettes per day (CPD)	20418890
81.	*CYP2A6,RAB4D*	rs4105144	19q13.2	C	Cigarettes per day (CPD)	20418888
82.	*CHRNB3,CHRNA6*	rs6474412	8p11.21	T	Cigarettes per day (CPD)	20418888
83.	*WRN*	rs6997892	8	G	15-year all-cause mortality	22397355
84.	*IGF1R*	rs2684766	15	T	15-year all-cause mortality	22397355
85.	*TRIM32*	rs10817931	9	A	15-year all-cause mortality	22397355
86.	*MAT2B*	rs1421783	5	C	15-year all-cause mortality	22397355
87.	*IGF1R*	rs11630259	15	T	15-year all-cause mortality	22397355
88.	*APOE*	rs7412	19	C	15-year all-cause mortality	22397355
89.	*APOE*	rs429358	19	C	15-year all-cause mortality	22397355
90.	*NOS1AP*	rs12143842	1q	T	QT-interval	19305408
91.	*CNOT1 ,GINS3, NDRG4, SLC38A7, GOT2*	rs37062	16q	G	QT-interval	19305408
92.	*SLC35F1, PLN, ASF1A*	rs11756438	6q	A	QT-interval	19305408
93.	*CCDC141*	rs17362588	2	A	Heart Rate	23583979
94.	*GJA1*	rs1015451	6	C	Heart Rate	23583979
95.	*CD46*	rs11118555	1	A	Heart Rate	23583979
96.	*MYH6*	rs365990	14	G	Heart Rate	23583979

Table 1 – Baseline anthropometric and clinical characteristics of Framingham Offspring Cohort participants in 4 pooled study examination.

	Pool 1 (1983-1991)	Pool 2 (1991-1998)	Pool 3 (1998-2005)	Pool 4 (2005-2011)
N	3,426	3,337	3,209	2,723
N (%) "Healthy" (i.e., non-CVD and subjects not taking meds for DM/hypertension/lipids)	2,716 (79.3)	2,314 (69.2)	1,709 (53.3)	936 (34.4)
N (males, %) All	1,617 (47.2)	1,511 (45.2)	1,480 (46.1)	1,227 (45.1)
Age, years	48.0 (9.9)	54.6 (9.8)	61.3 (9.6)	66.8 (9.2)
BMI, Kg/m^{-2}	26.3 (4.7)	27.4 (4.9)	28.2 (5.3)	28.2 (5.3)
Systolic blood pressure, mmHg	123.3 (16.9)	125.8 (18.6)	127.0 (18.7)	128.5 (17.3)
Lipids, mg/dL				
Total Cholesterol	210.2 (41.0)	205.0 (37.1)	200.3 (36.7)	185.9 (37.3)
LDL-C	133.4 (36.2)	126.6 (33.1)	119.8 (32.8)	105.2 (31.2)
HDL-C	51.3 (14.8)	50.11 (15.1)	53.8 (17.0)	57.4 (18.1)
Triglycerides	120.5 (115.2)	147.6 (115.7)	137.2 (89.6)	118.1 (69.5)
Lipid lowering medication, N (%)	34 (1.0)	232 (6.9)	676 (21.1)	1,1,51 (42.3)
Diabetics, N (%)	209 (6.1)	377 (11.3)	421 (13.1)	433 (15.9)
Diabetes medication, N (% of diabetics)	54 (25.8)	104 (27.6)	221 (52.5)	253 (58.4)
Fasting plasma glucose, mg/dL	94.0 (20.4)	100.5 (27.6)	104.1 (26.5)	106.6 (23.6)
Hypertension, N (%)	1,012 (29.5)	1,064 (31.8)	1,468 (45.8)	1,594 (58.6)
Antihypertensive medication, N (% of hypertensive subjects)	519 (51.3)	578 (54.3)	1,090 (74.3)	1,354 (84.9)
Smoking status, N (%)				
Current smoker	903 (26.4)	643 (19.2)	507 (15.8)	359 (13.2)
Previous smoker	1,387 (40.5)	1,583 (47.4)	1,627 (50.7)	1,417 (52.0)
Never smoker	1,136 (33.2)	1,117 (33.4)	1,075 (33.5)	947 (34.8)
Personal CVD* history, N (%)	257 (7.5)	367 (11.0)	379 (11.8)	505 (18.5)
GRS_{96} (whole population)	88.2 (5.6)	88.2 (5.6)	88.1 (5.6)	88.1 (5.6)
GRS_{96} (in people dead by stroke or MI)	89.9 (6.0)	90.8 (6.0)	89.0 (5.6)	89.4 (6.1)
GRS_{96} (in people dead by all- causes)	88.8 (6.5)	90.1 (6.1)	88.2 (5.7)	88.4 (5.5)
CVD events (fatal + non-fatal) **, N (%)	278 (8.1)	386 (11.5)	459 (14.3)	560 (20.6)
Deaths, N (%)				
Stroke + myocardial infarction (% of all population)	21 (0.6)	18 (0.5)	80 (2.5)	55 (2.0)
Stroke + myocardial infarction in Healthy, N (% of healthy)	6 (0.2)	10 (0.4)	17 (1.0)	4 (0.4)
Cancer (% of all deaths)	34 (39.5)	55 (48.3)	107 (33.3)	95 (38.2)
Non-cancer deaths (% of all population)	52 (1.5)	60 (1.8)	214 (6.7)	154 (5.7)
All-cause (% of all population)	86 (2.5)	115 (3.5)	321 (10.0)	249 (9.1)

*CVD, cardiovascular disease, defined as history of myocardial infarction, angina pectoris, coronary insufficiency. AF, atrial fibrillation. LVH-ECG, left ventricular hypertrophy as evaluated by electrocardiogram. GRS, genetic risk score.

**CVD events (fatal + non-fatal), composite endpoint comprising fatal myocardial infarction, fatal stroke, non-fatal MI, non-fatal stroke, coronary artery disease (angina pectoris, coronary insufficiency).

Data expressed as mean ± standard deviation (SD) unless otherwise indicated.

5.7 REFERENCES

1. **Qi, Q.**, et al., Diabetes genetic predisposition score and cardiovascular complications among patients with type 2 diabetes. Diabetes Care, 2013. **36**(3): p. 737-9.
2. **Sattar, N.**, Revisiting the links between glycaemia, diabetes and cardiovascular disease. Diabetologia, 2013. **56**(4): p. 686-95.
3. **American Diabetes Association**, Economic costs of diabetes in the U.S. in 2012. Diabetes Care, 2013. **36**(4): p. 1033-46.
4. **International Diabetes Federation**, IDF Diabetes Atlas 5th ed. 2011; Available from: http://www.idf.org/diabetesatlas.
5. **Morrish, N.J.**, et al., Mortality and causes of death in the WHO Multinational Study of Vascular Disease in Diabetes. Diabetologia, 2001. **44 Suppl 2**: p. S14-21.
6. **Whiting, D.R.**, et al., IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. Diabetes Res Clin Pract, 2011. **94**(3): p. 311-21.
7. **WHO, World Health Statistics 2013**, Geneva, Switzerland: World Health Organization.
8. **Arking, D.E.**, et al., Identification of a sudden cardiac death susceptibility locus at 2q24.2 through genome-wide association in European ancestry individuals. PLoS Genet, 2011. **7**(6): p. e1002158.
9. **Coronary Artery Disease Consortium**, et al., Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet, 2013. **45**(1): p. 25-33.
10. **Huertas-Vazquez, A.**, et al., Novel loci associated with increased risk of sudden cardiac death in the context of coronary artery disease. PLoS One, 2013. **8**(4): p. e59905.
11. **Ikram, M.A.**, et al., Genomewide association studies of stroke. N Engl J Med, 2009. **360**(17): p. 1718-28.
12. **Newton-Cheh, C.**, et al., Common variants at ten loci influence QT interval duration in the QTGEN Study. Nat Genet, 2009. **41**(4): p. 399-406.

13. **Thorgeirsson, T.E.**, et al., Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. Nat Genet, 2010. **42**(5): p. 448-53.
14. **Tobacco and C. Genetics**, Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet, 2010. **42**(5): p. 441-7.
15. **Traylor, M.**, et al., Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol, 2012. **11**(11): p. 951-62.
16. **Walter, S.**, et al., Genetic, physiological, and lifestyle predictors of mortality in the general population. Am J Public Health, 2012. **102**(4): p. e3-10.
17. **Morris, A.P.**, et al., Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet, 2012. **44**(9): p. 981-90.
18. **American Diabetes Association**, Economic Costs of Diabetes in the U.S. in 2012. Diabetes Care, 2013.

Chapter 6

Supplemental material

6.1

Authors and affiliations

Chapter 2 - Current Insights into the Joint Genetic Basis of Type 2 Diabetes and Coronary Heart Disease

Marco Dauriz, MD[1,2,3], James B. Meigs, MD, MPH[1,2]

1. Massachusetts General Hospital, General Medicine Division, 50 Staniford St., 9th Floor, Boston, MA 02114-2698, USA
2. Department of Medicine, Harvard Medical School, Boston, MA, USA
3. Division of Endocrinology and Metabolic Diseases, Department of Medicine, University of Verona Medical School and Hospital Trust of Verona, Verona, Italy

Chapter 3 - Association of a 62 Variants Type 2 Diabetes Genetic Risk Score With Markers of Subclinical Atherosclerosis: A Transethnic, Multicenter Study

Marco Dauriz, MD[1-4], Bianca C. Porneala, MS[1, 2], Xiuqing Guo, PhD[5], Lawrence F. Bielak, DDS, MPH[6], Patricia A. Peyser, PhD[6], Nefertiti H. Durant, MD[7], Mercedes R. Carnethon, PhD[8], Riccardo C. Bonadonna, MD[9], Enzo Bonora, MD, PhD[3], Donald W. Bowden, PhD[10, 11], Jose C. Florez, MD, PhD[2, 12, 13], Myriam Fornage, PhD[14], Marie-France Hivert, MD[15-17], David R. Jacobs, Jr., PhD[18], Edmond K. Kabagambe, PhD[19], Cora E. Lewis, MD[20], Joanne M. Murabito, MD[21, 22], Laura J. Rasmussen-Torvik, PhD[8], Stephen S. Rich, PhD[4], Jason L. Vassy, MD[2, 23, 24], Jie Yao, MD, MS[5], Jeffrey J. Carr, MD[25], Sharon L.R. Kardia, PhD[6], David Siscovick, MD[26], Christopher J. O'Donnell, MD[2, 22, 27], Jerome I. Rotter, MD[5], Josée Dupuis, PhD[22, 28], James B. Meigs, MD, MPH[1, 2]

1. Massachusetts General Hospital, General Medicine Division, Boston, MA, USA
2. Department of Medicine, Harvard Medical School, Boston, MA, USA
3. Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, University of Verona Medical School and Hospital Trust of Verona, Verona, Italy
4. Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

5. Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA
6. Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA
7. Division of Pediatrics and Adolescent Medicine, Department of Pediatrics, University of Alabama Birmingham School of Medicine, Birmingham, AL, USA
8. Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
9. Division of Endocrinology, Department of Clinical and Experimental Medicine, University of Parma School of Medicine and AOI of Parma, Parma, Italy
10. Centers for Diabetes Research and Human Genomics, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA
11. Department of Biochemistry and Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA
12. Diabetes Unit, Department of Medicine, Massachusetts General, Boston, MA, USA
13. Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA
14. Institute of Molecular Medicine and Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX
15. Harvard Pilgrim Health Care Institute, Department of Population Medicine, Harvard Medical School, Boston, MA, USA
16. Diabetes Unit, Massachusetts General Hospital, Boston, MA USA
17. Division of Endocrinology and Metabolism, Department of Medicine, Université de Sherbrooke, Sherbrooke, Québec, Canada
18. Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA
19. Department of Epidemiology, University of Alabama Birmingham School of Public Health, Birmingham, AL, USA
20. Division of Epidemiology, Department of Medicine, Vanderbilt University, Nashville, TN, USA
21. Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA

22. National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA
23. Section of General Internal Medicine, VA Boston Healthcare System, Boston, MA, USA
24. Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, Boston, MA, USA
25. Wake Forest University Health Sciences, Winston-Salem, NC, USA
26. Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology University of Washington, Seattle, WA, USA
27. Cardiology Division, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
28. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

Chapter 4 - Is common genetic variation at IRS1, ENPP1 and TRIB3 loci associated with cardiometabolic phenotypes in type 2 diabetes? An exploratory analysis of the Verona Newly Diagnosed Type 2 Diabetes Study (VNDS) 5

Maddalena Trombetta M[1,2*], Marco Dauriz M[1,2*], Sara Bonetti[1*], Daniela Travia[2], Maria Linda Boselli[1], Lorenza Santi[1], Enzo Bonora[1,2], Riccardo C Bonadonna[3,4]
[*] Equal contribution.

1. Department of Medicine, University of Verona, Verona, Italy
2. Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Hospital Trust of Verona, Verona, Italy
3. Division of Endocrinology, Department of Clinical and Experimental Medicine, University of Parma School of Medicine
4. Azienda Ospedaliera Universitaria, Parma, Italy

Chapter 5 - A genetic risk score of 96 variants linked with type 2 diabetes and cardiometabolic risk traits is associated with cardiovascular mortality in 29-years follow-up of the Framingham Heart Study

Marco Dauriz[1], Bianca C Porneala[2,3], Josée Dupuis[3], Joanne M Murabito[4,5], Adrianne L Cupples[5], Jose C Florez[2], James B Meigs[2]

1. Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, University of Verona Medical School and Hospital Trust of Verona, Verona, Italy
2. Massachusetts General Hospital, General Medicine Division, Boston, MA, USA
3. Department of Medicine, Harvard Medical School, Boston, MA, USA
4. Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA
5. National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA

6.2

Supplemental material Chapter 3

Association of a 62 Variants Type 2 Diabetes Genetic Risk Score With Markers of Subclinical Atherosclerosis: A Transethnic, Multicenter Study

Experimental design

A multivariable linear regression model with random effects to account for family relatedness was applied, where appropriate, to test the association of subclinical atherosclerosis (SCA) measures with an additive genetic risk score (GRS_{62}) comprised of 62 single nucleotide polymorphisms (SNPs) known to be linked with type 2 diabetes (T2D) risk (**Appendix Table 1**) [1]. Many of them are associated with either beta-cell function or insulin resistance (IR) physiology. Therefore, as described in [2], we used prior genetic and physiologic evidence [1, 3-6] to define a sub-GRS comprised of 20 T2D SNPs mainly associated with beta-cell function (GRS_{β}) and a sub-GRS comprised of 10 T2D SNPs associated with peripheral insulin resistance (GRS_{IR}), with each locus weighted in European Americans by the same effect size as in the GRS_{62}.

For each SCA trait a genetic-only model (adjusted for sex) and a full atherosclerosis risk factors adjusted model (**Appendix Table 2**) were applied. Clinical and anthropometric characteristics of study cohorts are shown in **Appendix Table 3** and **Appendix Table 4**. The GRS_{β} and GRS_{IR} were tested only in FHS and CARDIA study samples (**Appendix Table 5-6**).

Many of the 62 tag-SNPs associated with T2D (**Appendix Table 1**) are also known to be associated with SCA risk factors/confounders. Genome.gov (http://www.genome.gov/), a catalog of published GWAS, and PheGenI (http://www.ncbi.nlm.nih.gov/gap/phegeni), a phenotype-oriented resource housed at the National Center for Biotechnology Information, were interrogated. Risk factors listed in the catalogs as being associated with one or more of the known 62 T2D loci were included, among others, in the full model (BMI, waist circumference, systolic blood pressure, fasting insulin, fasting glucose, triglycerides, HDL-cholesterol and LDL-cholesterol). Therefore, the basic model could be described as “purely” genetic, as it tested the association of a T2D GRS

alone with SCA traits, after adjustment for sex, while the full model accounted for the overall spectrum of confounders, mediators and/or risk factors for SCA available in the dataset, by adjusting for a comprehensive list of atherosclerosis risk factors.

Genotyping

MESA: Caucasian, Hispanic, and Chinese American participants were genotyped on the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) at the Affymetrix Research Services Lab. 6,880 samples passed initial genotyping QC. African American samples were genotyped at the Broad Institute of Harvard and MIT as part of the CARe project [7]. Affymetrix performed wet lab hybridization assay, and plate-based genotype calling using Birdseed v2. Sample QC was based on call rates and contrast QC (cQC) statistics. Broad performed similar QC for CARe sample. Additional sample and SNP QC were carried out at University of Virginia, including sample call rate, sample cQC, and sample heterozygosity by ethnicity at the sample level; outlier plates checking by call rate, median cQC or heterozygosity at plate level. Four samples were removed due to low call rate (<95%). Cryptic sample duplicates or unresolved cryptic duplicates were dropped. Unresolved gender mismatches were also dropped. At the SNP level, we excluded monomorphic SNPs across all samples; SNPs with missing rate was > 5% or observed heterozygosity > 53% were also excluded. Additional genotypes were imputed to the 1000 Genomes Phase I integrated variant set (NCBI build 37 / hg19) separately in each ethnic group using the program IMPUTEv2. We used data freezes from 23 Nov 2010 (low-coverage whole-genome) and 21st May 2011 (high-coverage exome), phased haplotypes released March 2012 (v3), and phased haplotypes for 1,092 individuals and over 39 million variants. All imputed and genotyped SNPs were

aligned to the '+' strand of the human genome reference sequence (NCBI Build 37). The Affymetrix annotation file "GenomeWideSNP_6.na31.annot.csv" was used for all matching of probe set IDs with RS IDs.

GENOA: GENOA Study participants were genotyped on the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, 2007) or the Illumina Human 1M-Duo BeadChip (Illumina, 2010) at the Mayo Clinic, Rochester, MN. African American sibships for the GENOA study were identified using hypertensive participants from the Atherosclerosis Risk in Communities Study (ARIC) as probands. Genotypes were obtained for 92 additional GENOA participants who were also in the ARIC Study and who could not be genotyped on either platform using the GENOA blood sample. Genotyping for the ARIC study was done at the Broad Institute on the Affymetrix 6.0 platform. For all genotyping platforms used, samples and SNPs with a call rate <95% were removed. Samples demonstrating sex mismatch, duplicate samples, and samples with low identity-by-state with all other samples were also removed. Imputation was performed with the single-step approach implemented in Markov Chain Haplotyper (MaCH) 1.0.16 [8]. The reference panel was composed of the HapMap phased haplotypes (release 22). Imputation was performed separately for participants genotyped on the Affymetrix 6.0 as part of the GENOA study, participants genotyped on the Illumina Human 1M-Duo BeadChip, and participants genotyped on the Affymetrix 6.0 as part of the ARIC Study. Since only a small number of directly genotyped SNPs overlap on the Affymetrix and Illumina platforms, imputed dosages were used for all.

In GENOA African Americans the GRS was limited to 55 SNPs due to poor imputation quality for 7 SNPs. FHS and CARDIA genotyping and imputation strategy have been previously detailed elsewhere[2, 7, 9, 10].

Appendix Table 1 - 62 independent loci and relative tag-SNPs associated with Type 2 Diabetes from DIAGRAMv3 [1]

Chr	SNP-risk allele	Locus	OR	Chr	SNP-risk allele	Locus	OR
1	rs2075423-G	*PROX1 or PPP2R5A*	1.07	9	rs10758593-A	*GLIS3*	1.06
1	rs10923931-T	*NOTCH2*	1.08	9	rs17791513-A	*TLE4*	1.12
2	rs10203174-C	*THADA*	1.14	9	rs2796441-G	*TLE1*	1.07
2	rs243088-T	*BCL11A*	1.07	9	rs16927668-T	*PTPRD*	1.04
2	rs13389219-C	*GRB14*	1.07	10	rs11257655-T	*CDC123/CAMK1D*	1.07
2	rs2943640-C	*IRS1*	1.10	10	rs7903146-T	*TCF7L2*	1.39
2	rs7569522-A	*RBMS1*	1.05	10	rs1111875-C	*HHEX/IDE*	1.11
2	rs780094-C	*GCKR*	1.06	10	rs12571751-A	*ZMIZ1 or PPIF*	1.08
3	rs11717195-T	*ADCY5*	1.11	10	rs12242953-G	*VPS26A*	1.07
3	rs1496653-A	*UBE2E2*	1.09	11	rs10830963-G	*MTNR1B*	1.10
3	rs4402960-T	*IGF2BP2*	1.13	11	rs1552224-A	*ARAP1 (CENTD2)*	1.11
3	rs1801282-C	*PPARG*	1.13	11	rs163184-G	*KCNQ1*	1.09
3	rs6795735-C	*ADAMTS9*	1.08	11	rs5215-C	*KCNJ11*	1.07
3	rs12497268-G	*PSMD6*	1.03	11	rs2334499-T	*DUSP8 or HCCA2 (YY1AP1)*	1.04
3	rs17301514-A	*ST64GAL1*	1.05	12	rs7955901-C	*TSPAN8/LGR5*	1.07
4	rs6819243-T	*MAEA*	1.07	12	rs12427353-G	*HNF1A (TCF1)*	1.08
4	rs4458523-G	*WFS1*	1.10	12	rs2261181-T	*HMGA2*	1.13
5	rs6878122-G	*ZBED3 or PDE8B*	1.10	12	rs10842994-C	*KLHDC5 or PPFIBP1*	1.10
5	rs459193-G	*ANKRD55*	1.08	13	rs1359790-G	*SPRY2*	1.08
6	rs7756992-G	*CDKAL1*	1.17	15	rs4502156-T	*C2CD4A or VPS13C*	1.06
6	rs3734621-C	*KCNK16*	1.07	15	rs11634397-G	*ZFAND6*	1.05
6	rs4299828-A	*ZFAND3*	1.04	15	rs12899811-G	*PRC1*	1.08
7	rs17168486-T	*DGKB*	1.11	15	rs2007084-G	*AP3S2*	1.02
7	rs10278336-A	*GCK*	1.07	15	rs7177055-A	*HMG20A*	1.08
7	rs849135-G	*JAZF1*	1.11	16	rs9936385-C	*FTO*	1.13
7	rs17867832-T	*GCC1 or PAX-4*	1.09	16	rs7202877-T	*BCAR1*	1.12
7	rs13233731-G	*KLF14*	1.05	17	rs2447090-A	*SRR*	1.04
8	rs3802177-G	*SLC30A8*	1.14	18	rs12970134-A	*MC4R*	1.08
8	rs7845219-T	*TP53INP1*	1.06	19	rs10401969-C	*CILP2*	1.13
8	rs516946-C	*ANK1*	1.09	19	rs8182584-T	*PEPD*	1.04
9	rs10811661-T	*CDKN2A/B*	1.18	20	rs4812829-A	*HNF4A*	1.06

Appendix Table 2 – Outline of models applied in the association analysis of Genetic Risk Scores (GRS) with subclinical atherosclerosis traits, plus covariates.

BASIC MODEL	GRS, sex	
FULL MODEL	GRS, sex	Age BMI Waist circumference Systolic blood pressure (SBP)* Fasting insulin Fasting glucose Triglycerides HDL-Cholesterol LDL-Cholesterol Family history of T2D Family history of CVD Smoking status Physical activity Diabetes medication Hypertension medication Lipid-lowering medication

Appendix Table 3 – Subclinical atherosclerosis measures, anthropometric and clinical characteristics in FHS and CARDIA cohorts.

	FHS		CARDIA			
	Exam 6	*Exam 7*	*Exam year 20*		*Exam year 25*	
Ethnicity	**European Americans**		**African Americans**	**European Americans**	**African Americans**	**European Americans**
N (male %)	2459 (44.8%)	1111 (44.8%)	816 (38.6%)	1635 (45.9 %)	811 (38.8%)	1621 (45.9 %)
Age (yr)	57.9 ± 9.6	58.9 ± 8.9	44.4 ± 3.8	45.5 ± 3.3	49.4 ± 3.8	50.6 ± 3.3
BMI (kg/m^2)	27.6 ± 4.9	28.1 ± 4.9	31.7 ± 7.6	27.9 ± 6.7	32.2 ± 7.8	28.2 ± 6.2
Waist circumference (cm)	96.5 ± 12.7	96.5 ± 12.7	94.7 ± 15.7	89.8 ± 15.1	97.1 ± 15.8	91.6 ± 15.6
Systolic blood pressure (mmHg)	127.1 ± 18.3	124.9 ± 17.7	119.1 ± 15.5	112.2 ± 12.5	122.2 ± 14.5	114.5 ± 13.7
Fasting glucose (mg/dL)	101.3 ± 22.9	99.9 ± 18.2	102.3 ± 30.2	97.8 ± 21.2	102.1 ± 34.6	96.7 ± 20.4
Fasting insulin (pmol/L)	-	14.3 ± 8.6	17.1 ± 12.2	13.5 ± 9.1	13.39 ± 14.1	9.6 ± 7.3
Triglycerides (mg/dL)	136.1 ± 88.2	132.6 ± 86.5	96.4 ± 58.4	116.9 ± 82.4	101.4 ± 67.1	120.0 ± 86.3
HDL-cholesterol (mg/dL)	51.9 ± 16.1	53.9 ± 15.9	54.1 ± 16.3	54.4 ± 17.2	57.7 ± 17.3	58.7 ± 18.5
LDL-cholesterol (mg/dL)	127.4 ± 32.9	121.4 ± 31.3	110.4 ± 33.6	110.3 ± 30.5	109.2 ± 33.9	113.4 ± 30.9
Parental history of diabetes (%)	19.8	19.8	17.9	9.5	17.6	9.4
Parental history of CVD (%)	43.2	41.9	39.8	41.3	39.9	41.2
Diabetes (%)	7.1	6.3	10.9	3.4	13.3	6.5
Smokers (never/former/current - %)	35.8/48.9/15.2[*]	39.5/50.7/9.8[*]	59.9/40.1[**]	46.1/53.9[**]	62.2/37.9[**]	50.7/49.3[**]
Physical activity	-	-	287.5±285.4[§]	370.3±260.8[§]	264.4±257.5[§]	388.0±280.9[§]
Genetic Risk Score	66.7 ± 5.3	66.7 ± 5.2	69.2 ± 4.5	66.4 ± 5.2	69.2 ± 4.5	66.4 ± 5.2
Comorbidity status						
Diabetes medication (%)	3.3	2.9	7.9	3.4	10.7	4.6
Hypertension medication (%)	23.4	24.8	23.1	10.2	41.8	31.8
Lipid-lowering medication (%)	9.8	13.9	-	-	-	-
Subclinical atherosclerosis traits						
AACS (Agatston unit)	-	1458.6 ± 2332.3	-	-	-	-
CACS (Agatston unit)	-	229.8 ± 550.8	-	-	31.7 ± 154.6 (n=586)	49.2 ± 252.1 (n=1267)
CCA-IMT (mm)	0.5 ± 0.4 (n=2340)	-	0.7 ± 0.1 (n=617)	0.7 ± 0.1 (n=1379)	-	-
ICA-IMT (mm)	0.8 ± 1.9 (n=2035)	-	0.6 ± 0.2 (n=562)	0.6 ± 0.2 (n=1332)	-	-
ABI	1.1 ± 0.1 (n=2822)		-	-	-	-

Data expressed as mean±standard deviation, if not otherwise indicated. AACS, abdominal aorta calcium score; CACS, coronary artery calcium score; CCA, common carotid artery; ICA, internal carotid artery; IMT, intima-media thickness; ABI, ankle-brachial index. [*]Smoking status categorized as never/former/current in FHS. [**]Smoking status categorized as never/ever in CARDIA. [§]Physical activity is expressed as Total Intensity Score, according to the CARDIA Physical Activity History Questionnaire (Pereira MA *et al.*; PMID: 9243481).

Appendix Table 4 – Subclinical atherosclerosis measures, anthropometric and clinical characteristics in MESA and GENOA cohorts.

	MESA				GENOA	
Ethnicity	**European Americans**	**Asian Americans**	**African Americans**	**Hispanic Americans**	**European Americans**	**African Americans[§§]**
N (male %)	2526 (47.7%)	773 (49.2%)	1611 (46.1%)	1446 (48.3%)	969 (40.9%)	535 (25.8%)
Age (yr)	62.7 ± 10.2	62.4 ± 10.4	62.3 ± 10.1	61.4 ± 10.3	58.9 ± 9.5	68.5 ± 7.7
BMI (kg/m^2)	27.7 ± 5.1	23.9 ± 3.3	30.2 ± 5.9	29.5 ± 5.2	30.7 ± 6.3	32.7 ± 7.2
Waist circumference (cm)	97.9 ± 14.5	87.1 ± 9.8	101.3 ± 14.7	100.7 ± 13.1	100.3 ± 16.2	101.1 ± 15.4
Systolic blood pressure (mmHg)	123.5 ± 20.5	124.6 ± 21.7	131.8 ± 21.8	126.8 ± 21.9	131.4 ± 16.8	137.5 ± 21.0
Fasting glucose (mg/dL)	91.3 ± 21.6	99.2 ± 28.6	100.3 ± 32.7	103.9 ± 39.4	104.6 ± 24.5	111.6 ± 37.7
Fasting insulin (pmol/L)	9.1 ± 5.6	9.6 ± 12.5	11.5 ± 27.5	11.8 ± 15.7	54.2 ± 40.3	80.6 ± 87.5
Triglycerides (mg/dL)	133 ± 90.1	143.1 ± 85.7	105.2 ± 70.5	158.4 ± 101.8	159.0 ± 96.9	101.0 ± 63.3
HDL-cholesterol (mg/dL)	52.4 ± 15.8	49.3 ± 12.4	52.3 ± 15.2	47.5 ± 13.1	52.4 ± 15.6	57.1 ± 16.5
LDL-cholesterol (mg/dL)	117.1 ± 30.3	115.1 ± 28.8	116.7 ± 33.3	119.9 ± 32.9	122.7 ± 32.1	114.6 ± 35.3
Parental history of diabetes (%)	-	-	-	-	29.4	40.9
Parental history of CVD (%)	44.6/33.2/2.8[†]	14.5/23.3/1.2[†]	31.9/31/7[†]	31.2/23.8/3.2[†]	57.5[††]	56.1[††]
Diabetes (%)	5.9	13.5	17.4	17.8	13.5	35.5
Smokers (never/former/current - %)	33.1/66.9	69.6/30.4	26.6/73.4	40.7/59.3	52.5/37.3/10.2	60.4/31.4/8.2
Regular physical activity (daily hours)	12.8 ± 4.9	9.9 ± 4.4	14.4 ± 7.1	11.6 ± 5.9	3.7 ± 2.5[§]	1.0 ± 1.8[§]
Genetic Risk Score	63.9 ± 4.7	61.5 ± 4.2	56.5 ± 4.7	62.7 ± 4.8	64.2 ± 4.9	57.0 ± 3.9
Comorbidity status						
Diabetes medication (%)	4.6	9.2	13.6	15.8	8.9	32.0
Hypertension medication (%)	33.3	29.1	50.3	32.9	68.4	80.8
Lipid-lowering medication (%)	18.3	14.1	15.8	13.3	27.0	40.9
Subclinical atherosclerosis traits						
AACS (Agatston unit)	1668.4 ± 2581.4 (n=760)	1044.7 ± 2015.4 (n=247)	887.2 ± 1737.7 (n=343)	1044.6 ± 1898.4 (n=496)	-	-
CACS (Agatston unit)	338.6 ± 577.2 (n=1433)	205.8 ± 374.3 (n=392)	294.0 ± 582.8 (n=714)	281.4 ± 567.2 (n=659)	201.6 ± 467.2	236.3 ± 583.0
CCA-IMT (mm)	0.9 ± 0.2 (n=2501)	0.8 ± 0.2 (n=770)	0.9 ± 0.2 (n=1573)	0.9 ± 0.2 (n=1431)	-	-
ICA-IMT (mm)	1.1 ± 0.6 (n=2475)	0.9 ± 0.5 (n=766)	1.1 ± 0.6 (n=1547)	1.0 ± 0.6 (n=1399)	-	-
ABI	1.1 ± 0.1 (n=2494)	1.1 ± 0.1 (n=768)	1.1 ± 0.1 (n=1432)	1.3 ± 0.1 (n=1430)	-	-

Data expressed as mean±standard error, if not otherwise indicated. AACS, abdominal aorta calcium score; CACS, coronary artery calcium score; CCA, common carotid artery; ICA, internal carotid artery; IMT, intima-media thickness; ABI, ankle-brachial index. [†]CVD is categorized in MESA as myocardial infarction/stroke/amputation not due to injury. [††]Expressed as parental history of coronary heart disease in GENOA. [*]Smoking status categorized as never/ever in MESA. [**]Smoking status categorized as never/former/current in GENOA. [§]Physical activity categorized as moderate or heavy. [§§]African Americans in GENOA had an available genetic risk score limited to 55 of 62 T2D SNPs.

Appendix Table 5 – Association between prevalent subclinical atherosclerosis measures and a T2D genotype risk score (GRS) comprised of 20 tag SNPs mostly linked with beta-cell function (GRS_{β}) in linear regression models of FHS and CARDIA cohorts.

	FHS		CARDIA			
	European Americans		African Americans		European Americans	
Basic Model	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*
CACS	-0.023±0.05	0.64	-0.011±0.03	0.71	0.005±0.02	0.76
AACS	-0.072±0.06	0.23	-	-	-	-
ICA-IMT	-0.016±0.01	0.01	-0.001±0.00	0.55	0.001±0.00	0.31
CCA-IMT	-0.002±0.00	0.41	-0.001±0.00	0.60	0.000±0.00	0.69
ABI	-9.33E+08±0.00	0.94	-	-	-	-
Full Model						
CACS	-0.021±0.05	0.69	-0.027±0.03	0.36	0.001±0.02	0.95
AACS	-0.012±0.06	0.85	-	-	-	-
ICA-IMT	-0.018±0.01	**0.007**	-0.001±0.00	0.68	0.001±0.00	0.15
CCA-IMT	-0.004±0.00	0.15	-0.001±0.00	0.49	0.001±0.00	0.21
ABI	0.001±0.00	0.46	-	-	-	-

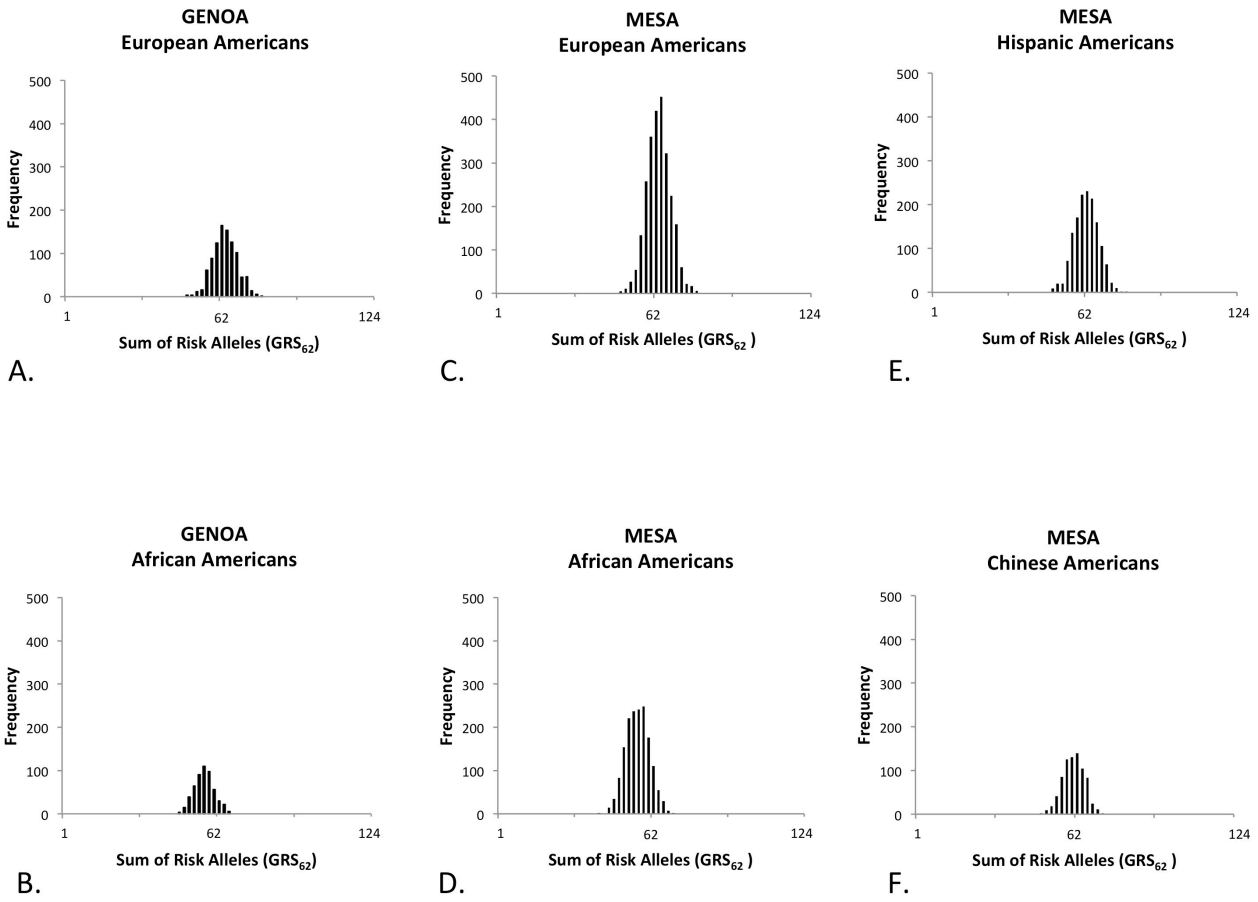
AACS, abdominal aorta calcium score; CACS, coronary artery calcium score; CCA, common carotid artery; ICA, internal carotid artery; IMT, intima-media thickness; ABI, ankle-brachial index. Data expressed as mean±standard error.

Appendix Table 6 – Association between prevalent subclinical atherosclerosis measures and a T2D genotype risk score (GRS) comprised of 10 tag SNPs mostly linked with insulin resistance (GRS_{IR}) in linear regression models of FHS and CARDIA cohorts.

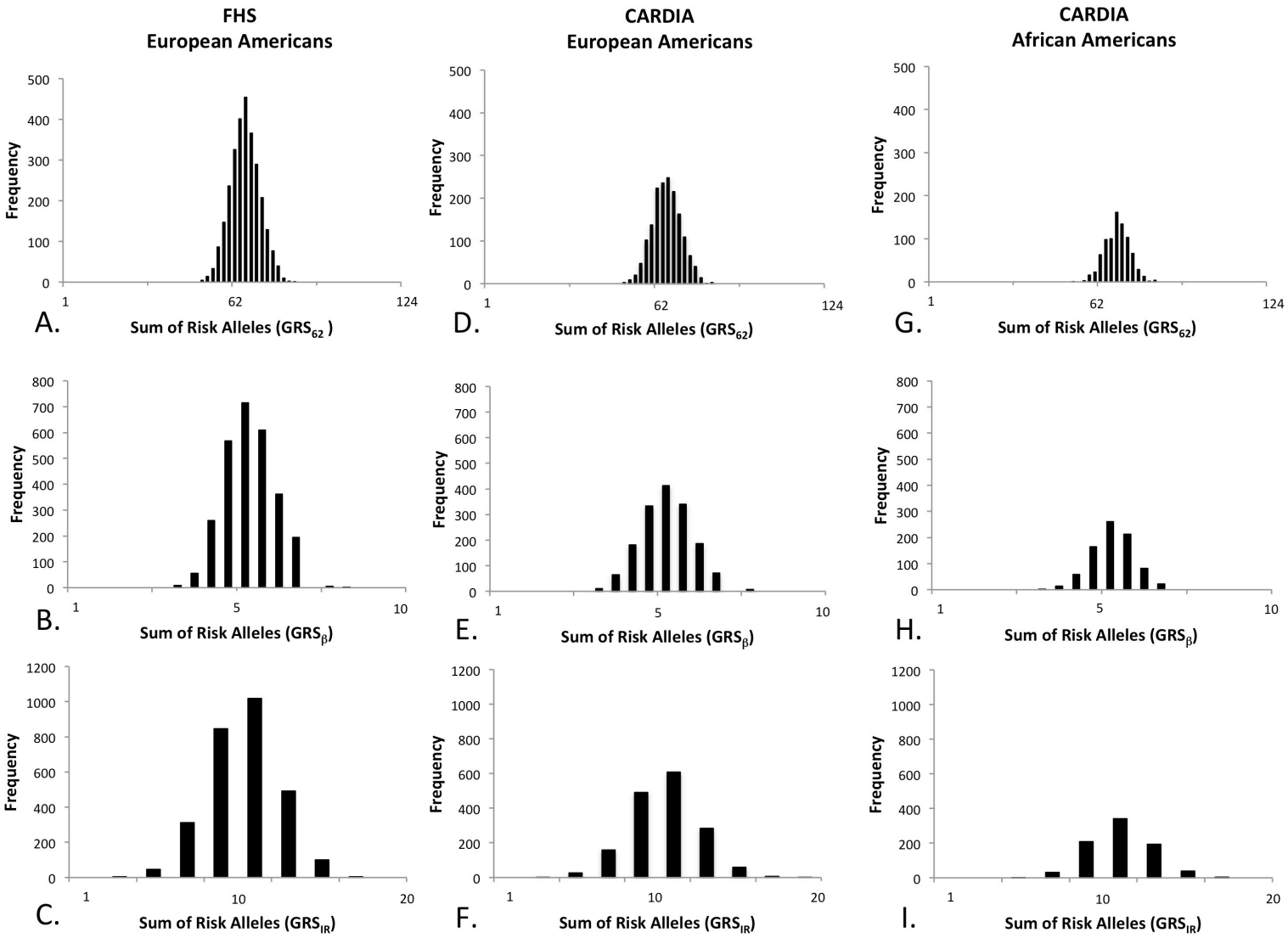
	FHS		CARDIA			
	European Americans		African Americans		European Americans	
Basic Model	*Beta±SE*	*P*	*Beta±SE*	*P*	*Beta±SE*	*P*
CACS	-0.004±0.07	0.95	-0.011±0.04	0.98	0.041±0.03	0.11
AACS	0.112±0.08	0.17	-	-	-	-
ICA-IMT	-0.003±0.01	0.77	0.000±0.00	0.93	-0.001±0.00	0.71
CCA-IMT	-0.004±0.00	0.29	0.001±0.00	0.46	-0.000±0.00	0.77
ABI	-7.74-06±0.00	0.99	-	-	-	-
Full Model						
CACS	0.036±0.08	0.65	0.005±0.04	0.89	0.025±0.03	0.32
AACS	0.056±0.09	0.57	-	-	-	-
ICA-IMT	-0.005±0.01	0.65	0.001±0.00	0.84	-0.000±0.00	0.83
CCA-IMT	-0.009±0.00	0.01	0.001±0.00	0.58	-0.001±0.00	0.63
ABI	0.001±0.00	0.56	-	-	-	-

AACS, abdominal aorta calcium score; CACS, coronary artery calcium score; CCA, common carotid artery; ICA, internal carotid artery; IMT, intima-media thickness; ABI, ankle-brachial index. Data expressed as mean±standard error.

Appendix Figure 1 – Distribution of the total sum of risk alleles comprised in the T2D GRS_{62} in GENOA (panel A and B) and MESA cohorts (panels C to F), stratified by ethnicity.



Appendix Figure 2 – Distribution of the total sum of risk alleles comprised in the T2D GRS_{62}, GRS_{β} and GRS_{IR} in FHS (panel A, B, C, respectively) and in CARDIA cohorts (panel D to I), stratified by ethnicity.



SUPPLEMENTARY REFERENCES

1. **Morris, A.P.**, et al., Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*, 2012. **44**(9): p. 981-90.
2. **Vassy, J.L.**, et al., Polygenic type 2 diabetes prediction at the limit of common variant detection. Diabetes, 2014.
3. **Dimas, A.S.**, et al., Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes*, 2013.
4. **Scott, R.A.**, et al., Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet*, 2012. **44**(9): p. 991-1005.
5. **Manning, A.K.**, et al., A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet,* 2012. **44**(6): p. 659-69.
6. **Lindgren, C.M.**, et al., Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet,* 2009. **5**(6): p. e1000508.
7. **Lettre, G.**, et al., Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: the NHLBI CARe Project. *PLoS Genet,* 2011. **7**(2): p. e1001300.
8. **Li YA, G.R.**, Mach 1.0: Rapid haplotype reconstruction and missing genotype inference. *Am J Hum Genet* 2006. **S79**: p. 2290.
9. **Lemaitre, R.N.**, et al., Genetic loci associated with plasma phospholipid n-3 fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium. *PLoS Genet*, 2011. **7**(7): p. e1002193.
10. **Gabriel, S.B.**, et al., The structure of haplotype blocks in the human genome. *Science*, 2002. **296**(5576): p. 2225-9

7.3

Supplemental material Chapter 4

Is common genetic variation at IRS1, ENPP1 and TRIB3 loci associated with cardiometabolic phenotypes in type 2 diabetes? An exploratory analysis of the Verona Newly Diagnosed Type 2 Diabetes Study

Experimental Design

The Verona Newly Diagnosed Type 2 Diabetes Study (VNDS) is an ongoing study aiming at building a biobank of patients with newly diagnosed Type 2 diabetes. As of Jan 1 2002, all patients referred to the Division of Endocrinology and Metabolic Diseases of University of Verona School of Medicine, whose diabetes has been diagnosed in the last six months, are asked to participate in this research. The clinical evidence on which the diagnosis of Type 2 diabetes has been made is reviewed and the diagnosis confirmed [1]. Patients are drug-naïve or, if already treated with antidiabetic drugs, undergo a treatment washout of at least one week before metabolic tests are performed. Among the exclusion criteria are age>75 years, non-Italian ancestry, insulin treatment, presence of anti-GAD antibodies, malignancies, and any condition severely impairing liver and/or kidney function. In this study, we report the data collected in 509 patients, whose characteristics are summarized in **Table 1**.

All subjects consumed a weight-maintaining diet containing 200-250 g of carbohydrate/day for at least three days before studies. Body weight was stable in all subjects for at least 1 month before studies. No subject participated in any heavy exercise. Each subject gave informed written consent before participating in the research, which was approved by the Human Investigation Committee of the Verona City Hospital. Measurements of standard clinical phenotypes were collected in all patients. Metabolic tests were carried out on two separate days in random order. On both days, patients were admitted to the Metabolic Clinic Research Center at 07:30 after an overnight fast. All studies were carried out in a quiet, temperature controlled (22° C) room.

On one day an OGTT (75 g) was performed to assess beta cell function. For ethical reasons, the OGTT was not performed in patients presenting with FPG greater than 15 mmol/l. During the entire test patients were sitting in a comfortable cardiac chair. One teflon (21 g) venous catheter was inserted into an antecubital vein for blood sampling and kept patent with heparinized normal saline solution. After a 30' rest to establish baseline and after collecting a 20 cc blood sample for leukocyte DNA extraction, at time = 0' subjects ingested 75 g of glucose in 300 ml of water over 5 min. Blood samples to measure glucose, C-peptide and insulin concentrations were collected at times -10', 0', +15', +30', +45', +60', +90', +120', +150', +180', +210' and +240'. In some patients further blood samples were collected at +270' and +300'. Urines were collected to measure glycosuria.

On a separate day, a euglycemic insulin clamp was performed to assess insulin sensitivity [2]. During the entire test patients were lying in bed. One teflon catheter was introduced into an antecubital vein for the infusion of test substances. Another teflon catheter was placed retrogradely into a wrist vein for sampling arterialized venous blood, according to the "hot box" technique. After a 30' rest in bed to establish baseline, indirect calorimetry (at least 40') was

performed as previously described, for a companion study [1]. At the end of calorimetric measures, baseline blood samples were collected and a standard euglycemic insulin (intravenous prime: 4.8 $nmol \cdot m^{-2}$ BSA; continuous infusion: 240 $pmol \cdot min^{-1} \cdot m^{-2}$ BSA) clamp was performed [1]. Plasma glucose was allowed to decline until it reached 5.5 mmol/l, after which glucose clamping started with a glucose concentration goal of 5 mmol/l. The duration of the glucose clamp was at least of 120', but it was prolonged, if and as needed, to ensure at least 60' of insulin clamp at euglycemia in each patient. Timed blood samples were collected to measure hormone and substrate levels. In the last 45' of the clamp indirect calorimetry was repeated to assess substrate oxidation and energy production rates for a companion study. Urines were collected to measure urea excretion rate.
In both metabolic tests, all blood samples were collected in pre-chilled tubes and readily spun at 1,500 g. Plasma and serum specimens were stored at –80° C.

Mathematical Modelling of Beta Cell Function

The analysis of the glucose and C-peptide curves during the OGTT follows the general strategy described in previous publications [3, 4] with some modifications and builds upon previous works from other laboratories [5, 6]. The kinetics of C-peptide is described with a two-compartment model, in which the two pools (1 and 2) exchange with each other and the irreversible loss of the hormone is from pool 1, the same where C-peptide concentration is measured. C-peptide kinetic parameters are computed according to the equations by Van Cauter et al. [7].
Herein are the equations describing the model of glucose induced insulin secretion during an OGTT:

$$dcp_1(t)/dt = ISR(t) + cp_2 \cdot k_{12} - (k_{01} + k_{21}) \cdot cp_1 \quad (Eq.1)$$

where ISR = insulin secretion rate, cp_1 = C-peptide mass in the sampling (accessible) compartment, cp_2 = C-peptide mass in the remote compartment, k_{12}

and k_{21} = rate constants of the exchange between the two C-peptide compartments, and k_{01} = rate constant of the irreversible loss of C-peptide from the accessible compartment. Note that the values of the volume of distribution of C-peptide pool 1 (accessible compartment), k_{12}, k_{21}, and k_{01} are computed according to the equations by Van Cauter et al. [7].

$$ISR(t) = BSR + DSR(t) + PSR(t) \quad (Eq.2)$$

where BSR = basal insulin secretion rate, DSR = insulin secretion rate due to the derivative (or dynamic) component, and PSR = insulin secretion rate due the proportional (or static) component.

$$BSR = CP_{ss} \cdot V_1 \cdot k_{01} \quad (Eq.\ 3)$$

where CP_{ss} is basal C-peptide concentration and V_1 is the volume of the accessible compartment of C-peptide.

From the modeling viewpoint, DSR(t) and PSR(t) are the components which in intravenous glucose tolerance tests or hyperglycemic clamps describe classical first phase insulin secretion and second phase insulin secretion, respectively. Furthermore, from a physiological viewpoint, the sum of BSR and PSR(t) describes the relationship linking glucose concentration and insulin secretion rate, in the absence of the derivative component (DSR).

DSR(t) and PSR(t) are mathematically defined as follows:

$$DSR(t) = X1(t) \cdot t^{-1} \quad (Eq.\ 4)$$

$$dX1(t)/dt = s1 \cdot [dG(t)/dt]/[\log(1.1+t)] - X1(t) \cdot t^{-1} \quad \text{if } dG(t)/dt > 0 \quad (Eq.\ 5)$$

$$dX1(t)/dt = - X1(t) \cdot t^{-1} \quad \text{if } dG(t)/dt \leq 0 \quad (Eq.\ 6)$$

, where *s1* = glucose sensitivity of derivative control of insulin secretion, G = plasma glucose concentration, X1 = C-peptide (insulin) mass made available for the derivative component of insulin secretion, t = time constant of the derivative component of insulin secretion, and the term log(1.1 + t) accomodates the time-associated decline of *s1* documented in humans during a hyperglycemic stimulus [8].

$$PSR(t) = X2(t) \cdot d^{-1} \quad (Eq.\ 7)$$

$$dX2(t) / dt = s2 \cdot [G(t) - q] - X2(t) \cdot d^{-1} \quad (Eq.\ 8)$$

where *s2* = glucose sensitivity of the proportional component of insulin secretion, X2 = C-peptide (insulin) mass made available for the proportional component of insulin secretion, d = time constant of the proportional component of insulin secretion, q = glucose threshold above which the beta-cell responds with the proportional component of insulin secretion to plasma glucose concentration.

This model was implemented in the SAAM 1.2 software (SAAM Institute, Seattle, WA) [9] to estimate its unknown parameters. Numerical values of the unknown parameters were estimated by using nonlinear least squares. Weights were chosen optimally, i.e., equal to the inverse of the variance of the measurement errors, which were assumed to be additive, uncorrelated, with zero mean, and a coefficient of variation (CV) of 6-8%. The unknown parameters of the model are: CP_{ss}, *s1*, t, *s2*, d, and q. They were estimated with good precision, as shown by their CVs (**Supplementary Table S1**).

A good fit of the model to data was obtained as shown by the table of the weighted residuals (**Supplementary Table S2**).

There are two main physiological outputs of the model:

1. derivative control (units: [pmol·m^{-2} BSA] · [mmol·l^{-1}·min^{-1}] $^{-1}$): it is the amount of insulin secreted in response to a rate of glucose increase of 1 mmol/l per min which lasts for 1 minute;

2. stimulus-response curve linking glucose concentration (x axis) to insulin secretion rate (y axis): as explained above, it is the sum of BSR and PSR. With the purpose of avoiding artefactual increases in the power of statistical analyses, we used the stimulus-response curve at the pre-determined glucose concentrations of 5.5, 8.0, 11.0, 15.0 and 20.0 mmol/l.

Statistical analysis

In this paper there are 5 outcome variables of interest (carotid atherosclerosis, ECG, insulin sensitivity, derivative control and proportional controls of beta cell function) and 10 SNPs of interest, giving rise to 50 comparisons and to the issue of correcting for false positive findings. We applied the two stage step-up linear procedure of Benjamini-Krieger-Yekutieli (BKY) [10] to control the false discovery rate (FDR) in our findings. The BKY procedure is a refinement of an earlier attempt [11] to improve the classical Benjamini-Hochberg method to control FDR [12]. It is based on the idea that in many sets of multiple comparisons the number of true null hypotheses is less than the number of hypotheses tested, and that this bit of information is relevant to compute FDR in multiple comparisons.

Briefly, the BKY procedure tries to exploit the information content of the distribution of the *p* values of multiple comparisons. In **Supplementary Figure**

S2, the *p* values of the 50 comparisons performed in this study (*y*-axis) are plotted against their rank (*x*-axis), from the smallest to the largest. If all 50 null hypotheses tested were true, the *p* values should follow the straight line of **Supplementary Figure S2**. The observed *p* values at some point deviate from the expected distribution. The BKY procedure uses the data to estimate the number of true null hypotheses, which by definition in our case can be only ≤ 50. With an iterative process of linear regression analyses, in our case the number of true null hypotheses was found to be 41, not 50. With a FDR set at 0.20, which is the highest acceptable FDR according to Benjamini and Yekutieli [13], this led us to reject 5 null hypotheses, i.e. to accept 5 results as statistically significant at a FDR of 0.20.

Figure S1 - Genomic position and LD values of the *ENPP1, IRS1* and *TRIB3* genotyped variants, as selected by GEVALT (GEnotype Visualization and ALgorithmic Tool) software in the VNDS study sample.

The upper portion of each figure shows the gene and the genomic position of the genotyped polymorphisms. The lower portion of the figure shows the LD value, calculated as r^2, among the SNPs. The dotted lines connect each SNP name and position with the corresponding cell in the LD matrix. Increasing level of LD is shown by darker grayscale. Each number enclosed in the grey diamonds below each locus should be divided by 100 to obtain the actual r^2 value (i.e. 6 means r^2 = 6/100 = 0.06 etc …).

Figure S1

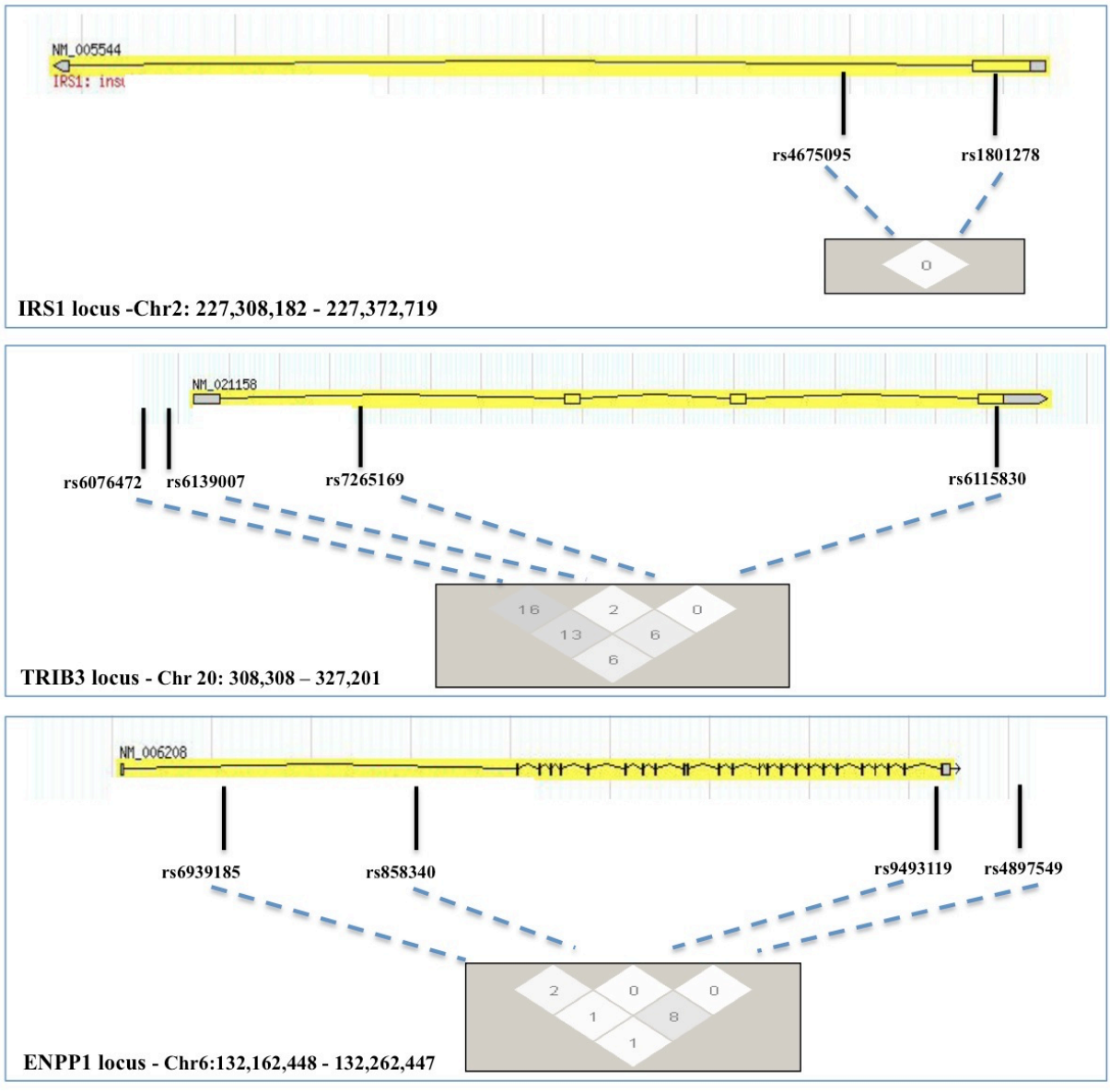
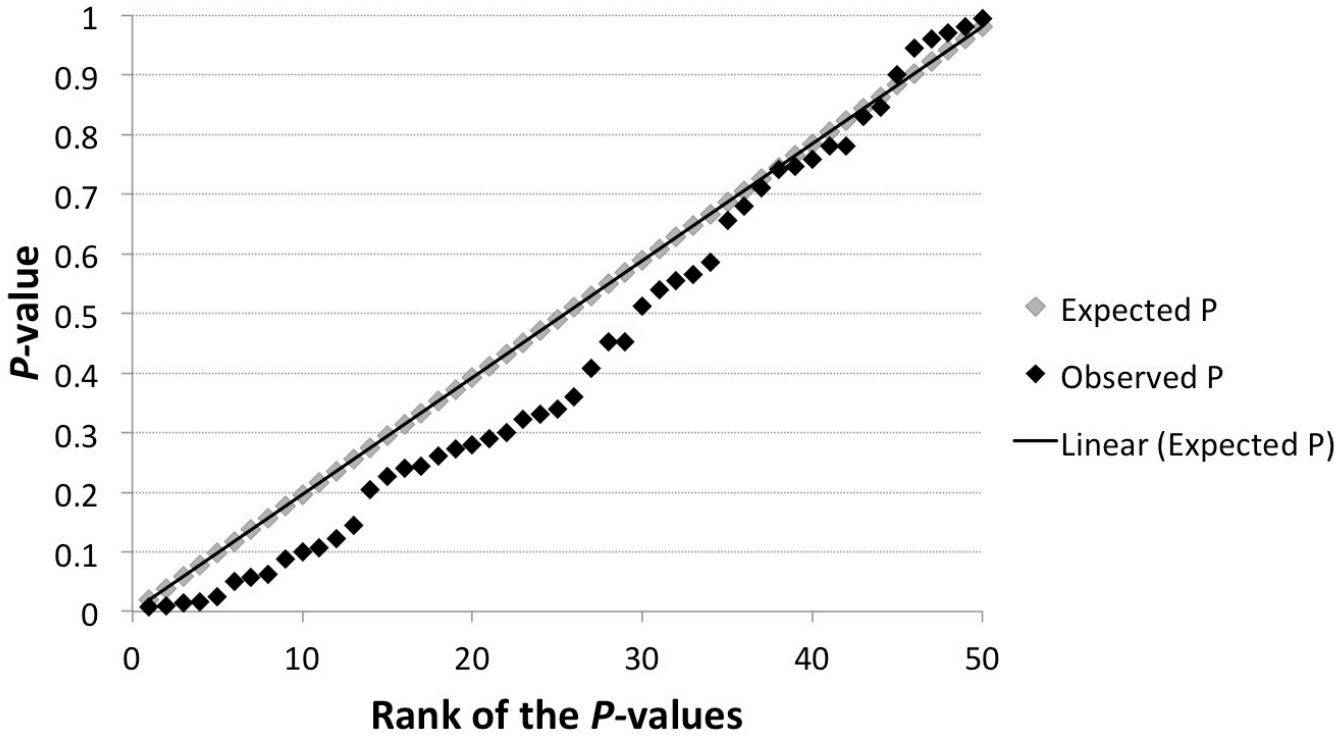

A.
NM_005544
IRS1: ins
rs4675095
rs1801278
0
IRS1 locus -Chr2: 227,308,182 - 227,372,719
B.
NM_021158
rs6076472
rs6139007
rs7265169
rs6115830
16
2
0
13
6
6
TRIB3 locus - Chr 20: 308,308 – 327,201
C.
NM_006208
rs6939185
rs858340
rs9493119
rs4897549
2
0
0
1
8
1
ENPP1 locus - Chr6:132,162,448 - 132,262,447

Figure S2 - *P* values of the 50 comparisons of the main outcome variables (*y*-axis) *vs.* the rank of the same *p* values from the lowest to the highest (*x*-axis).

The straight line indicates the expected distribution of the *p* values if all the null hypotheses were true, and, therefore, to be accepted, i.e. if no "true" associations between SNPs and cardiometabolic phenotypes existed.


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
P-value
0
10
20
30
40
50
Rank of the P-values
Expected P
Observed P
Linear (Expected P)

SUPPLEMENTARY REFERENCES

1. **Monauni T**, Zenti MG, Cretti A, Daniels MC, Targher G, Caruso B, Caputo M, McClain D, Del Prato S, Giaccari A *et al*: Effects of glucosamine infusion on insulin secretion and insulin action in humans. *Diabetes* 2000, 49(6):926-935.
2. **Bonadonna RC**, del Prato S, Bonora E, Gulli G, Solini A, DeFronzo RA: Effects of physiological hyperinsulinemia on the intracellular metabolic partition of plasma glucose. *Am J Physiol* 1993, 265(6 Pt 1):E943-953.
3. **Cali AM**, Bonadonna RC, Trombetta M, Weiss R, Caprio S: Metabolic abnormalities underlying the different prediabetic phenotypes in obese adolescents. *J Clin Endocrinol Metab* 2008, 93(5):1767-1773.
4. **Weiss R**, Caprio S, Trombetta M, Taksali SE, Tamborlane WV, Bonadonna R: Beta-cell function across the spectrum of glucose tolerance in obese youth. *Diabetes* 2005, 54(6):1735-1743.
5. **Cobelli C**, Toffolo GM, Dalla Man C, Campioni M, Denti P, Caumo A, Butler P, Rizza R: Assessment of beta-cell function in humans, simultaneously with insulin sensitivity and hepatic extraction, from intravenous and oral glucose tests. *Am J Physiol Endocrinol Metab* 2007, 293(1):E1-E15.
6. **Mari A**, Camastra S, Toschi E, Giancaterini A, Gastaldelli A, Mingrone G, Ferrannini E: A model for glucose control of insulin secretion during 24 h of free living. *Diabetes* 2001, 50 Suppl 1:S164-168.
7. **Van Cauter E**, Mestrez F, Sturis J, Polonsky KS: Estimation of insulin secretion rates from C-peptide levels. Comparison of individual and standard kinetic parameters for C-peptide clearance. *Diabetes* 1992, 41(3):368-377.
8. **Toschi E**, Camastra S, Sironi AM, Masoni A, Gastaldelli A, Mari A, Ferrannini E, Natali A: Effect of acute hyperglycemia on insulin secretion in humans. *Diabetes* 2002, 51 Suppl 1:S130-133.
9. **Foster DM**, Boston RC, Jacquez JA, Zech L: A resource facility for kinetic analysis: modeling using the SAAM computer programs. *Health Phys* 1989, 57 Suppl 1:457-466.

10. **Benjamini Y**, Krieger AM, Yekutieli D: Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* 2006, 93(3):491-507.
11. **Benjamini Y**, Hochberg Y: On the adaptive control of the false discovery fate in multiple testing with independent statistics. *J Educ Behav Stat* 2000, 25(1):60-83.
12. **Benjamini Y**, Hochberg Y: Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* 1995, 57(1):289-300.
13. **Benjamini Y**, Yekutieli D: Quantitative trait loci analysis using the false discovery rate. *Genetics* 2005, 171(2):783-789.

Table S1. Coefficients of variation of the beta cell model parameters. CP_{ss}= basal C-peptide concentration; *s1* = parameter regulating glucose sensitivity of derivative control of insulin secretion, t = time constant of derivative control of insulin secretion, *s2* = glucose sensitivity of proportional control of insulin secretion, δ = time constant of proportional control of insulin secretion, θ: glycemic threshold of proportional control of insulin secretion.

Model Parameter	Coefficients of Variation (%)	
	Median	I.Q. Range
CP_{ss}	10.8	6.9-18.3
s1	40.7	24.2-83.4
τ	60.3	58.2-61.2
s2	16.2	11.9-22.4
δ	33.0	21.4-66.7
θ	13.4	8.5-22.0

Table S2. Weighted residuals of the model fit to the C-peptide data of the OGTT. Data are presented as means±SD. The weighted residuals are a quantitative point-by-point assessment of the goodness-of-fit of the model to the data: a theoretically perfect fit should generate weighted residuals with mean 0 and SD of 1.

C-Peptide weighted residuals

Time	15’	30’	45’	60’	90’	120’	150’	180’	210’	240’
Mean	-0.397	+0.118	+0.199	+0.267	+0.115	-0.026	+0.142	+0.015	+0.026	+0.107
SD	1.03	1.14	1.197	1.263	1.302	1.213	1.319	1.268	1.277	1.164

Table S3. Pairwise Linkage Disequilibrium among the 13 SNPs of *IRS1*, *ENNP1* and *TRIB3* loci in 1000Genomes Pilot1, panel CEU.*

SNP	Proxy	Distance	RSquared	DPrime	Chr	Coordinate_HG18	Locus
rs6076472	rs12626158	9540	0.104	0.68	20	318150	TRIB3
rs6076472	rs7265169	4137	0.098	0.757	20	312747	TRIB3
rs6076472	rs6139007	179	0.082	1	20	308789	TRIB3
rs6076472	rs6115830	16616	0.069	0.524	20	325226	TRIB3
rs6076472	rs2295490	8295	0.005	0.095	20	316905	TRIB3
rs6139007	rs12626158	9361	0.134	0.608	20	318150	TRIB3
rs6139007	rs6115830	16437	0.098	0.55	20	325226	TRIB3
rs6139007	rs2295490	8116	0.042	1	20	316905	TRIB3
rs6139007	rs7265169	3958	0.014	1	20	312747	TRIB3
rs7265169	rs2295490	4158	0.027	0.287	20	316905	TRIB3
rs7265169	rs6115830	12479	0.014	0.572	20	325226	TRIB3
rs7265169	rs12626158	5403	0.011	0.535	20	318150	TRIB3
rs2295490	rs12626158	1245	0.055	0.426	20	318150	TRIB3
rs2295490	rs6115830	8321	0	0.061	20	325226	TRIB3
rs12626158	rs6115830	7076	0.292	0.569	20	325226	TRIB3
rs6939185	rs9493119	72231	0.033	0.522	6	132253111	ENPP1
rs6939185	rs858340	20007	0.02	0.198	6	132200887	ENPP1
rs6939185	rs1044498	33181	0.005	0.149	6	132214061	ENPP1
rs6939185	rs4897549	80303	0	0.012	6	132261183	ENPP1
rs858340	rs9493119	52224	0.051	0.459	6	132253111	ENPP1
rs858340	rs4897549	60296	0.051	0.247	6	132261183	ENPP1
rs858340	rs1044498	13174	0.003	0.259	6	132214061	ENPP1
rs1044498	rs9493119	39050	0.278	0.726	6	132253111	ENPP1
rs1044498	rs4897549	47122	0.119	0.463	6	132261183	ENPP1
rs9493119	rs4897549	8072	0.019	0.254	6	132261183	ENPP1
rs4675095	rs1801278	n/a	n/a	n/a	n/a	n/a	IRS1
rs1801278	rs4675095	n/a	n/a	n/a	n/a	n/a	IRS1

* As provided by the SNAP software v. 2.2 housed at http://www.broadinstitute.org, Broad Institute (Boston, MA, USA).
n/a, not available.

Supplementary Table S4. Clinical and metabolic features of the VNDS study population.

Variable	All
N (M/F)	597 (415/182)
Age (years)	59 [52-66]
BMI (Kg/m^2)	29.3 [26.5-32.8]
Waist circumference (cm)	100 [93-108]
Smokers (%)	47.9
Fasting plasma glucose (mmol/l)	7.0 [6.1-7.9]
2hr plasma glucose (mmol/l)	13.0 [10.5-16.1]
$HbA1c_{DCCT}$ (%)	6.7 [6.2-7.5]
$HbA1c_{IFCC}$ (mmol/mol)	49.7 [44.3-58.5]
Triglycerides (mmol/l)	1.4 [1.0-2.0]
HDL-cholesterol (mmol/l)	1.1 [1.0-1.3]
Total cholesterol (mmol/l)	4.9 [4.3-5.6]
SBP (mmHg)	138 [124-150]
DBP (mmHg)	82 [80-90]
Insulin Sensitivity (µmol/min/m^2 BSA)	607 [380-865]
Insulinogenic Index (mU/mmol)	4.1 [2.2-7.3]
$CIR_{120'}$ (mU·L/$mmol^2$)	0.5 [0.2-1.3]

Data are presented as median [I.Q. range]; BMI, Body Mass Index; SBP, Systolic Blood Pressure; DBP, Diastolic Blood Pressure; $HbA1c_{DCCT}$, Diabetes Control and Complication Trial-Aligned Hemoglobin A1c; $HbA1c_{IFCC}$, International Federation of Clinical Chemistry-Aligned Hemoglobin A1c; $CIR_{120'}$, Corrected Insulin $Response_{120'}$

Supplementary Table S5. Association of *IRS1, ENPP1* and *TRIB3* genotypes with insulin sensitivity as assessed by euglycemic hyperinsulinemic clamp and Matsuda Index.

SNP (*locus*)	Effect allele†	Genotype			*P*-value
rs858340 (*ENPP1*)	**A**	**AA**	**AG**	**GG**	
Matsuda Index		3.1±0.3	3.4±0.1	3.3±0.1	0.82
M clamp*		820±69.1	664±23.3	623±19.9	**0.008**
rs6939185 (*ENPP1*)	**A**	**AA**	**AG**	**GG**	
Matsuda Index		3.7±0.3	3.3±0.1	3.2±0.1	0.01
M clamp		727±44.1	653±20.8	638±25.6	0.05
rs9493119 (*ENPP1*)	**G**	**GG**	**AG**	**AA**	
Matsuda Index		2.7±0.6	3.9±0.4	3.3±0.1	0.21
M clamp		623±19.9	664±23.3	820±69.1	0.24
rs4897549 (*ENPP1*)	**A**	**AA**	**AG**	**GG**	
Matsuda Index		3.2±0.3	3.4±0.2	3.4±0.1	0.45
M clamp		662±60.7	679±24.8	634±20.1	0.45
rs6115830 (*TRIB3*)	**A**	**AA**	**AG**	**GG**	
Matsuda Index		3.3±0.2	3.4±0.1	3.2±0.2	0.64
M clamp		667±27.9	655±22.3	639±30.8	0.23
rs6139007 (*TRIB3*)	**G**	**GG**	**AG**	**AA**	
Matsuda Index		2.8±0.3	3.5±0.2	3.3±0.1	0.68
M clamp		668±53.3	677±23.9	638±21.0	0.14
rs6076472 (*TRIB3*)	**C**	**CC**	**AC**	**AA**	
Matsuda Index		3.4±0.3	3.3±0.1	3.3±0.1	0.76
M clamp		667±53.2	651±23.7	663±21.3	0.10
rs7265169 (*TRIB3*)	**A**	**AA**	**AC**	**CC**	
Matsuda Index		3.07±0.62	3.21±0.24	3.36±0.09	0.74
M clamp		384±112.1	594±32.7	668.8±17.0	0.24
rs1801278 (*IRS1*)	**A**	**AA**	**AG**	**GG**	
Matsuda Index		3.9±2.6	3.4±0.2	3.3±0.1	0.85
M clamp		829±356.3	655±40.7	653±16.1	0.74
rs4675095 (*IRS1*)	**A**	**AA**	**AT**	**TT**	
Matsuda Index		2.3±0.3	3.7±0.3	3.3±0.1	0.23
M clamp		672±43.0	674±39.7	651±16.2	0.59

Generalized Linear Model, additive genetic model (outcome trait: insulin sensitivity, expressed as Matsuda Index or M-clamp; covariates: age, sex, BMI). † The minor allele, defined according to the MAF in our population, was considered as effect allele. * M clamp unit: µmol/min/m^2 Body Surface Area. Matsuda Index and M-clamp were expressed as mean±SE.

Supplementary Table S6. MAGIC lookup: Ln(HOMA-IR)

SNP	Effect allele[†]	Other allele	MAF	Effect	SE	*P*-value
rs4675095 (*IRS1*)	A	T	0.067	-3.40E-02	8.80E-03	0.0001171
rs1801278 (*IRS1*)	n/a	n/a	n/a	n/a	n/a	n/a
rs858340 (*ENPP1*)	T	C	0.246	-9.50E-03	4.40E-03	0.03142
rs6939185 (*ENPP1*)	A	G	0.396	-3.40E-03	4.10E-03	0.4157
rs9493119 (*ENPP1*)	A	G	0.062	9.00E-04	1.10E-02	0.9295
rs4897549 (*ENPP1*)	T	C	0.178	-3.60E-03	4.50E-03	0.4319
rs6139007 (*TRIB3*)	T	C	0.195	1.30E-02	6.10E-03	0.02906
rs7265169 (*TRIB3*)	A	C	0.106	-9.50E-03	1.40E-02	0.4873
rs6115830 (*TRIB3*)	T	C	0.379	1.50E-03	4.30E-03	0.72
rs6076472 (*TRIB3*)	T	G	0.274	-1.60E-03	6.90E-03	0.8129

[†]According to the MAGIC database. n/a, not available; Ln, natural logarithm. HOMA-IR, IR-Homeostatic Model Assessment; MAF, Minor Allele Frequency; SE, Standard Error.

Supplementary Table S7. MAGIC lookup: Ln(HOMA-B)

SNP	Effect allele[†]	Other allele	MAF	Effect	SE	*P*-value
rs4675095 (*IRS1*)	A	T	0.067	-2.10E-02	7.50E-03	0.004202
rs1801278 (*IRS1*)	n/a	n/a	n/a	n/a	n/a	n/a
rs858340 (*ENPP1*)	T	C	0.246	-4.20E-03	3.60E-03	0.2396
rs6939185 (*ENPP1*)	A	G	0.396	-1.90E-03	3.40E-03	0.569
rs9493119 (*ENPP1*)	A	G	0.062	-4.90E-03	9.40E-03	0.6041
rs4897549 (*ENPP1*)	T	C	0.178	2.20E-03	3.70E-03	0.5512
rs6139007 (*TRIB3*)	T	C	0.195	9.00E-03	5.00E-03	0.07133
rs7265169 (*TRIB3*)	A	C	0.106	-9.50E-03	1.40E-02	0.4873
rs6115830 (*TRIB3*)	T	C	0.379	8.00E-04	3.60E-03	0.8275
rs6076472 (*TRIB3*)	T	G	0.274	4.00E-04	6.40E-03	0.9553

[†]According to the MAGIC database. n/a, not available; Ln, natural logarithm. HOMA-B, Beta-cell function - Homeostatic Model Assessment; MAF, Minor Allele Frequency; SE, Standard Error.

Supplementary Table S8. MAGIC lookup: Insulin Sensitivity Index (ISI)[*]

SNP	Effect allele[†]	Other allele	MAF	Effect	SE	*P*-value
rs4675095 (*IRS1*)	A	T	0.067	9.70E-02	5.50E-02	0.076463
rs1801278 (*IRS1*)	n/a	n/a	n/a	n/a	n/a	n/a
rs858340 (*ENPP1*)	C	T	0.246	7.00E-03	2.70E-02	0.792024
rs6939185 (*ENPP1*)	G	A	0.396	-2.40E-02	2.40E-02	0.327156
rs9493119 (*ENPP1*)	A	G	0.062	-8.90E-03	6.70E-02	0.894582
rs4897549 (*ENPP1*)	C	T	0.178	-1.80E-02	2.40E-02	0.451828
rs6139007 (*TRIB3*)	T	C	0.195	-9.90E-04	3.80E-02	0.979483
rs7265169 (*TRIB3*)	n/a	n/a	n/a	n/a	n/a	n/a
rs6115830 (*TRIB3*)	C	T	0.379	-1.70E-02	2.50E-02	0.500486
rs6076472 (*TRIB3*)	T	G	0.274	2.00E-02	2.90E-02	0.503135

[†]According to the MAGIC database. n/a, not available; Ln, natural logarithm. MAF, Minor Allele Frequency; SE, Standard Error. [*] Calculated as Matsuda Index, as follows: $10{,}000/[(Glucose_{0'}\cdot Insulin_{0'})\cdot(\text{Mean OGTT glucose concentration})\cdot(\text{mean OGTT insulin concentration})]^{1/2}$

Supplementary Table S9. MAGIC lookup: $CIR_{30'}$

SNP	Effect allele†	Other allele	MAF	Effect	SE	*P*-value
rs4675095 (*IRS1*)	A	T	0.067	-3.70E-03	5.00E-02	0.940667
rs1801278 (*IRS1*)	n/a	n/a	n/a	n/a	n/a	n/a
rs858340 (*ENPP1*)	C	T	0.246	4.10E-02	2.30E-02	0.079624
rs6939185 (*ENPP1*)	G	A	0.396	3.50E-02	2.10E-02	0.095807
rs9493119 (*ENPP1*)	A	G	0.062	-1.60E-01	5.50E-02	0.002835
rs4897549 (*ENPP1*)	C	T	0.178	1.80E-03	2.20E-02	0.935155
rs6139007 (*TRIB3*)	T	C	0.195	-3.00E-02	3.20E-02	0.351976
rs7265169 (*TRIB3*)	C	A	0.106	-6.90E-04	3.50E-02	0.984125
rs6115830 (*TRIB3*)	C	T	0.379	-1.80E-02	2.10E-02	0.399347
rs6076472 (*TRIB3*)	T	G	0.274	-3.50E-02	2.60E-02	0.172351

†According to the MAGIC database. n/a, not available; Ln, natural logarithm. $CIR_{30'}$, Corrected Insulin Response at 30' after a 75g-OGTT; MAF, Minor Allele Frequency; SE, Standard Error.