Curr Cardiovasc Risk Rep (2014) 8:368
DOI 10.1007/s12170-013-0368-z

DIABETES AND INSULIN RESISTANCE (M RUTTER, SECTION EDITOR)

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

Marco Dauriz · James B. Meigs

Published online: 17 January 2014

Abstract 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). Chromosome 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.

Keywords Type 2 diabetes · Cardiovascular disease · Genetic risk · Insulin resistance · Glucose homeostasis

This article is part of the Topical Collection on *Diabetes and Insulin Resistance*

M. Dauriz (✉) · J. B. Meigs (✉)
Massachusetts General Hospital, General Medicine Division, 50 Staniford St. 9th Floor, Boston, MA 02114-2698, USA
e-mail: mdauriz@partners.org
e-mail: jmeigs@partners.org

M. Dauriz · J. B. Meigs
Department of Medicine, Harvard Medical School, Boston, MA, USA

M. Dauriz
Division of Endocrinology and Metabolic Diseases, Department of Medicine, University of Verona Medical School and Hospital Trust of Verona, Verona, Italy

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	Welcome 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

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 forwards 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 harbouring 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 glycemic 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.

Recent type 2 Diabetes Genome-Wide Association Studies

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 ten 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 Fig. 1 and thoroughly detailed in the following sections, there is also compelling evidence that specific T2D loci on chromosome 2q36.3 and 9p21.3 harbour allelic variants in close proximity to each other and genomic regions associated with increased CHD risk.

Table 1 Glossary of unfamiliar terms used in the review

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].

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,

Table 2 List of currently known T2D- and CHD-associated SNPs from most recent GWAS[1,2,3], labelled according to the nearest 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	***CDKN2A/B (9p21.3)***
	rs16927668	T	*PTPRD*	rs3217992	A	***CDKN2A/B (9p21.3)***
	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	*CXCL12*
	rs12242953	G	*VPS26A*	rs12413409	G	*CYP17A1*, *CNNM2*, *NT5C2*
	rs12571751	A	*ZMIZ1*	rs11203042	T	*LIPA*
				rs2246833	T	*LIPA*

Table 2 (continued)

65 T2D associated loci[1]				45+1 CHD associated loci[2,3]		
Chromosome	SNP	Risk allele	Nearest gene	SNP	Risk allele	Nearest gene
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	*COL4A1*, *COL4A2*
				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	*ApoE-ApoC1*
Chr. 20	rs4812829	A	*HNF4A*			
Chr. 21				rs9982601	T	*KCNE2*

Candidate overlapping regions at chromosome 9p21.3 and 2q36.3 are highlighted in bold. Though not genome-wide significant, rs2943634 has been included as putative independent signal historically[3] associated with CAD risk in 2q36.3 locus.

[1] Morris AP *et al.* (DIAGRAMv3 Consortium) *Nat. Genet. 2012* – PMID: 22885922 [14•];

[2] Deloukas P *et al.* (CARDIoGRAMplusC4D Consortium) *Nat. Genet. 2013* – PMID:23202125 [15•];

[3] Samani NJ *et. al.* (WTCC and Cardiogenics Consortium) *NEJM 2007* – PMID: 17634449 [11]

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 glycemic trait-associated variants (fasting insulin, fasting plasma glucose, HOMA-B and HOMA-IR) for any of the SNPs analyzed (Fig. 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

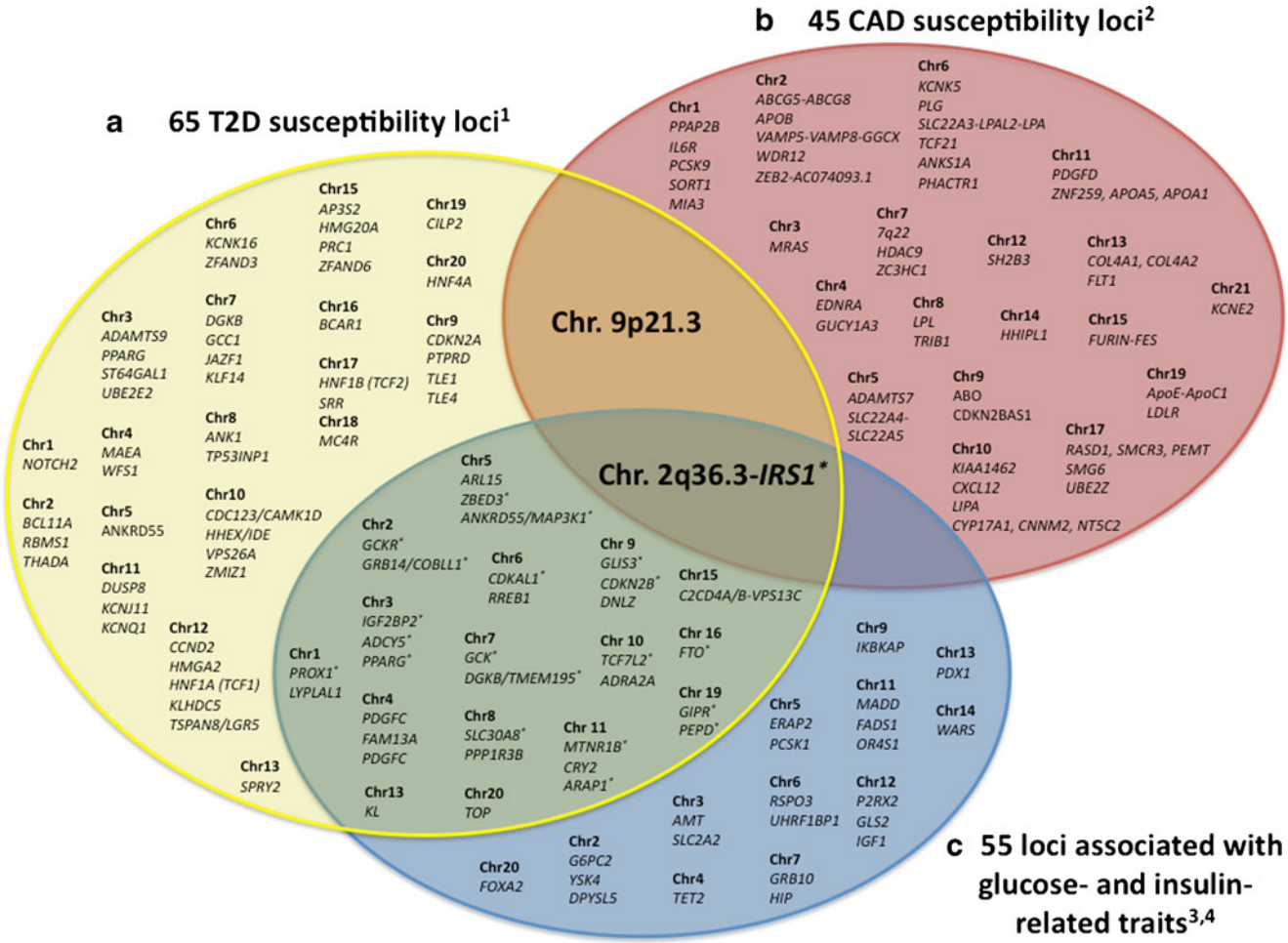

Fig. 1 Overlapping associations among currently known T2D, CHD and glycaemic quantitative trait susceptibility loci from recent GWAS. Loci harbouring one or more common variant(s) associated with the phenotype or trait of interest are listed according to the nearest known gene. The diagram highlights 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

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.

Recent Glycemic 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).

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

		Glycaemic traits $p<5x10^{-8}$			T2D $p<10^{-4}$	Lipids $p<10^{-4}$		Anthropometric measures $p<10^{-4}$	
Nearest gene(s)	Chr.	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		+						

Table 3 (continued)

		Glycaemic traits $p<5x10^{-8}$			T2D $p<10^{-4}$	Lipids $p<10^{-4}$		Anthropometric measures $p<10^{-4}$	
Nearest gene(s)	Chr.	FI	FG	2hGlu		Tg	HDL-C	BMI	WHR
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

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 (Fig. 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] glycemic 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 glycemic traits (FI, FG, post-challenge glucose concentration) to 55 (53 confirmed loci in Scott *et al.* plus two 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 (Fig. 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).

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 glycemic quantitative traits. As pointed out in Fig. 1 and Fig. 2, one of the most promising regions is a large locus spanning ~593 kb located on chromosome 2q36.3 and harbouring 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, 12•, 13, 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 (Fig. 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 significance ($p=1.61x10^{-7}$) 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 2,740 Framingham Heart Study participants. The study examined

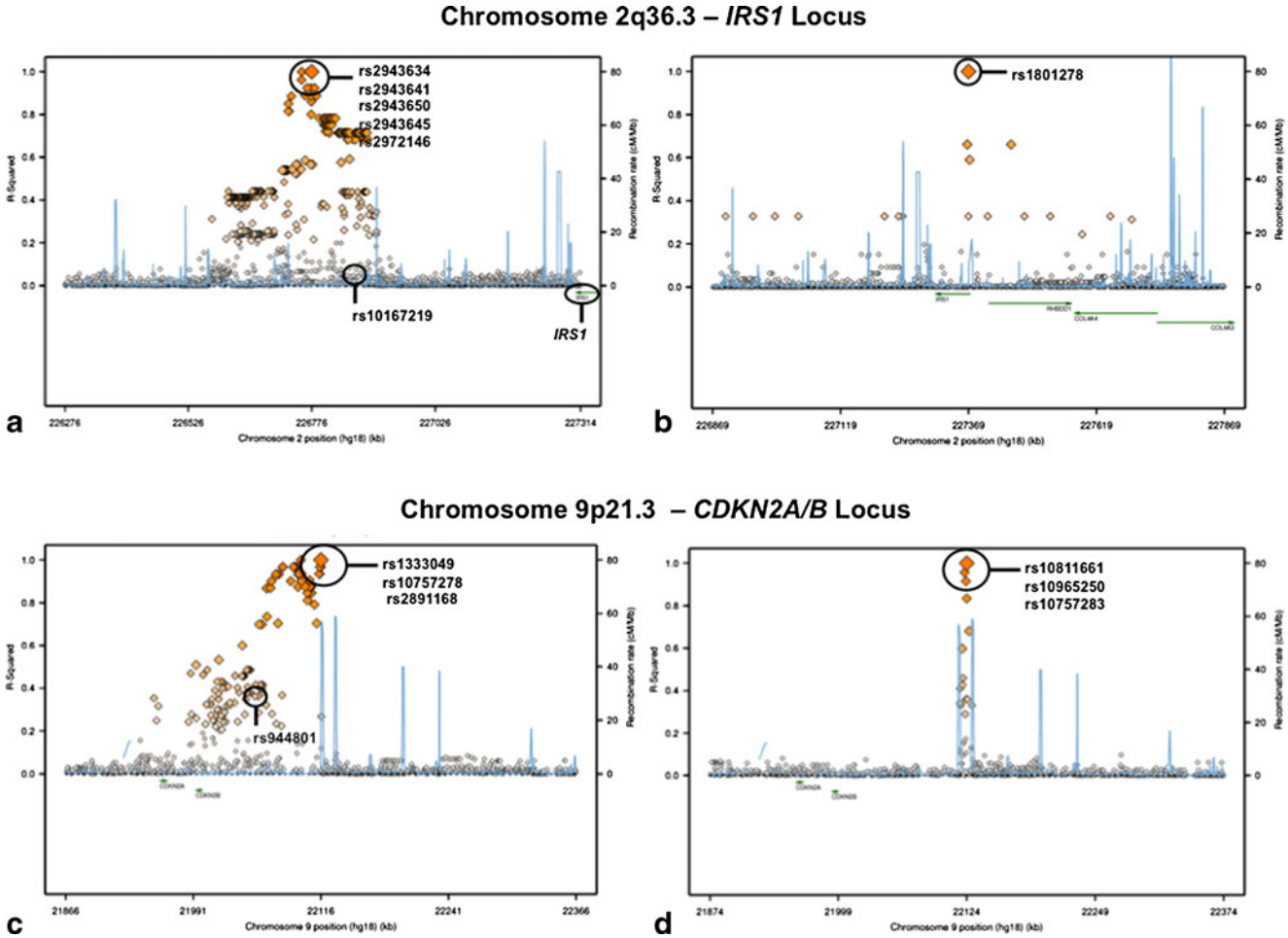

Fig. 2 Linkage disequilibrium patterns among lead SNPs at type 2 diabetes – coronary heart disease loci on chromosomes 2 and 9. The left-hand y-axis of each panel indicates the linkage disequilibrium (LD), represented by the r^2 value, among single nucleotide polymorphisms (SNPs) at the locus, with the 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 ~593 kb 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 ~8.6 kb 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

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 Fig. 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.

The Chromosome 9p21.3 Locus

As shown in Fig. 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 overlap 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 (Fig. 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 so 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 SNP 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 harbouring the CHD-associated SNPs. As shown in Fig. 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 (Fig. 2C). 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.

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 glycemic 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 glycemic 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 potentially 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 harbours 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 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.

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 glycemic 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.

Acknowledgements Dr. Dauriz was supported by research grants of the University of Verona ("CooperInt" Junior Researcher Program 2012 grant and Endocrine Fellowship 2009–2014 grant). Dr. Meigs was supported by NIDKK K24 DK080140 and R01 DK078616. No additional external funding was received for this study. The funders had no role in the design or preparation of the manuscript. The authors have no other conflicts of interest to report.

Compliance with Ethics Guidelines

Conflict of Interest James Meigs serves as a consultant to LipoScience, Inc., and Quest Diagnostics and has received a grant from the NIH.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by the author.

References

Papers of particular interest, published recently, have been highlighted as:

- • Of importance
- •• Of major importance

1. IDF. *IDF Diabetes Atlas 5th ed.* 2011; Available from: http://www.idf.org/diabetesatlas.

2. WHO, *Global Atlas on Cardiovascular Disease Prevention and Control*., ed. P.P. Mendis S, Norrving B editors, Geneva, Switzerland.
3. WHO, *World Health Statistics*, Geneva, Switzerland: World Health Organization.
4. Stranger BE, Stahl EA, Raj T. Progress and promise of genome-wide association studies for human complex trait genetics. Genetics. 2011;187(2):367–83.
5. Broadbent HM 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):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):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;11:60.
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):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):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):140–5.
11. Samani NJ et al. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007;357(5):443–53.
12. • Scott RA 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):991–1005. *Provides the most updated data on the genetic architecture of glycaemic traits.*
13. Rung J et al. Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. Nat Genet. 2009;41(10):1110–5.
14. • Morris AP et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet. 2012;44(9):981–90. *Provides the most updated data on the genetic architecture of T2D.*
15. • Consortium CAD et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet. 2013;45(1): 25–33. *Provides the most updated data on the genetic architecture of CHD.*
16. Cho YS 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):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):638–45.
18. Voight BF et al. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet. 2012;8(8):e1002793.
19. Peden JF, Farrall M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. Hum Mol Genet. 2011;20(R2):R198–205.
20. Reich DE, Lander ES. On the allelic spectrum of human disease. Trends Genet. 2001;17(9):502–10.
21. Manolio TA. Bringing genome-wide association findings into clinical use. Nat Rev Genet. 2013;14(8):549–58.
22. • Manning AK 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):659–69. *Provides the most updated data on the genetic architecture of glycaemic traits.*
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):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):142–8.
25. Kilpelainen TO et al. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. Nat Genet. 2011;43(8):753–60.
26. Teslovich TM et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010;466(7307):707–13.
27. Li G et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell. 2012;148(1–2):84–98.
28. DeFronzo RA. Insulin resistance, lipotoxicity, type 2 diabetes and atherosclerosis: the missing links. The Claude Bernard Lecture 2009. Diabetologia. 2010;53(7):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):149–54.
30. Murabito JM 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):100–12.
31. Helgadottir A et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science. 2007;316(5830):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):372–80.
33. McPherson, R., Chromosome 9p21.3 Locus for CAD: How little we know. J Am Coll Cardiol, 2013.
34. McPherson R, Davies RW. Inflammation and coronary artery disease: insights from genetic studies. Can J Cardiol. 2012;28(6):662–6.
35. Harismendy O et al. 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. Nature. 2011;470(7333):264–8.
36. Almontashiri NA 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):143–7.
37. Solovieff N et al. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet. 2013;14(7):483–95.
38. •• Wray NR et al. Pitfalls of predicting complex traits from SNPs. Nat Rev Genet. 2013;14(7):507–15. *Provides a general sense in designing, interpretation and future projection of studies in the field of human complex traits genetics.*
39. Qi Q et al. Diabetes genetic predisposition score and cardiovascular complications among patients with type 2 diabetes. Diabetes Care. 2013;36(3):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):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):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):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):400–5. *Provides a general sense in designing, interpretation and future projection of studies in the field of human complex traits genetics.* 405e1-3.
44. •• Goldstein DB et al. Sequencing studies in human genetics: design and interpretation. Nat Rev Genet. 2013;14(7):460–70. *Provides a general sense in designing, interpretation and future projection of studies in the field of human complex traits genetics.*
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):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):401–8.
47. Frazer KA et al. Human genetic variation and its contribution to complex traits. Nat Rev Genet. 2009;10(4):241–51.
48. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. PLoS Comput Biol. 2012;8(12):e1002822.
49. Johnson AD et al. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008;24(24):2938–9.

Author Contributions

Dr. Dauriz researched data and wrote the manuscript. Dr. Meigs researched data and wrote and edited the manuscript.