OPEN

Clinical & Translational Immunology (2017) 6, e162; doi:10.1038/cti.2017.51
Official journal of the Australasian Society for Immunology Inc.
www.nature.com/cti

REVIEW

Type 1 diabetes genome-wide association studies: not to be lost in translation

Flemming Pociot[1,2,3]

Genetic studies have identified >60 loci associated with the risk of developing type 1 diabetes (T1D). The vast majority of these are identified by genome-wide association studies (GWAS) using large case–control cohorts of European ancestry. More than 80% of the heritability of T1D can be explained by GWAS data in this population group. However, with few exceptions, their individual contribution to T1D risk is low and understanding their function in disease biology remains a huge challenge. GWAS on its own does not inform us in detail on disease mechanisms, but the combination of GWAS data with other omics-data is beginning to advance our understanding of T1D etiology and pathogenesis. Current knowledge supports the notion that genetic variation in both pancreatic β cells and in immune cells is central in mediating T1D risk. Advances, perspectives and limitations of GWAS are discussed in this review.

Clinical & Translational Immunology (2017) **6,** e162; doi:10.1038/cti.2017.51; published online 1 December 2017

Type 1 diabetes (T1D) is a chronic immune-mediated disease causing attrition and death of the insulin-producing pancreatic β cells, resulting in a life-long requirement for exogenous insulin. The progressive loss of β cells is mainly owing to autoimmune inflammation.[1] Worldwide >20 million people are afflicted with T1D. By 2015, more than half a million children are estimated to be living with T1D and ~86 000 children develop T1D each year (www.diabetesatlas.org). In most countries T1D incidence is increasing by ~3–4% every year, most notably in children and adolescents.[2] Five million people in the USA are expected to have T1D by 2050, including ~600 000 youth.[3,4] Existing treatments do not relieve the disease burden, for example, severe hypoglycemia is common[5,6] and >70% of patients are unable to maintain a healthy HbA1c.[7,8] Life-expectancy is reduced by up to 13 years; even with good HbA1c control, life expectancy is reduced by ~8 years.[9]

Fifteen percent of newly diagnosed T1D patients have a first-degree family member with T1D. The T1D concordance rates are in the range of 30–70% in monozygotic twins and 3–13% in dizygotic twins.[10–12] This non-Mendelian inheritance pattern is characteristic for multi-factorial diseases and results from the contribution of several genes each having only a minor influence on disease development.[13] In addition to genetic predisposition, environmental and epigenetic factors impact the disease susceptibility.[14,15]

In this review, only studies of genetic architecture of T1D are discussed with focus on the translation of genetics into biology. For the last two decades, genome-wide approaches to map the genetic risk have been prevailing. First, as linkage studies using affected sib-pairs and subsequently as genome-wide association studies (GWAS) using a case–control design. Whereas the initial genome-wide linkage studies (GWLS) mainly confirmed established associations (*HLA*, *INS* and *CTLA4*) from previous candidate-gene studies,[16–20] more recent and larger GWLS also provided novel information on the genetic predisposition to T1D.[21–23] Nevertheless, very few of the novel loci identified from these 1st, 2nd and 3rd generation GWLS have been replicated and confirmed in more recent GWAS data sets. This is mainly owing to the inherited limitations in GWLS, which include limited power to narrow down risk variants and to detect risk variants with only minor contribution. An example of a novel region associated with T1D identified by GWLS and subsequently confirmed by GWAS, is the *UBASH3A* region on chromosome 21.[24]

T1D GENETICS–THE PRE-GWAS ERA

Genetic studies have had an essential role in understanding T1D biology.[25] The first reports of genetic association to T1D were for the human leukocyte antigen (HLA) region.[26–28] As this discovery, researchers have tried to understand the underlying mechanisms by which alleles of HLA-encoding genes are responsible for the T1D association. Although much has been learned about the effects of certain HLA alleles on T1D risk, the exact biological mechanism of HLA-conferred susceptibility remains elusive. The extreme polymorphism of the HLA locus makes association analyses complicated. In addition, the strong linkage disequilibrium in the region makes assessment of individual risk variants challenging. The HLA region is the most polymorphic observed in the human genome, with 17 166 unique alleles reported as of July 2017 (http://www.ebi.ac.uk/imgt/hla/stats.html). The genetic risk for T1D in Caucasians is conferred mainly by combinations of HLA-DR and -DQ genes, for example, those encoding DR4-DQ8 (that is, *DRB1*04, DQA1*03–DQB1*03:02*) and

[1]Department of Pediatrics, Herlev and Gentofte Hospital, Herlev, Denmark; [2]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark and [3]Steno Diabetes Center Copenhagen, Gentofte, Denmark
Correspondence: Professor F Pociot, Steno Diabetes Center Copenhagen, Niels Steensensvej 2, Gentofte 2820, Denmark.
E-mail: flemming.pociot@regionh.dk
Received 30 August 2017; revised 15 October 2017; accepted 16 October 2017

Table 1 Type 1 diabetes genome-wide association studies (GWAS) reported in the GWAS Catalogue (www.ebi.ac.uk/gwas)

Study	*Sample description (cases/ controls)*	*Replication sample (cases/ controls)*	*Platform (SNPs passing QC)*	*Significant associations ($P \leqslant 5 \times 10E\text{-}8$)*
WTCCC[36]	1963/2938		Affymetrix (469 557)	10
Todd *et al.*[102]	2000/3000	4000/50 002 997 trios	Affymetrix (NR)[a]	12
Hakonarson *et al.*[103]	561/1 143 467 trios	1333/390 trios	Illumina (543 071)	4
Cooper *et al.*[104]	3561/4646	6225/69 463 064 trios	Affymetrix (335 565)	14
Hakonarson *et al.*[105]	561/1 143 467 trios	946/1 098 364 trios	Illumina (543 071)	1
Barrett *et al.*[87]	7514/9045	4267/4 6704 342 trios	Affymetrix/Illumina (841 622 (imputed))	38
Grant *et al*[106]	563/1 146 483 trios	3303/4673	Illumina (1 000 000)	5
Wallace *et al.*[74]	7514/9045	4840/26 705 766 trios	Affymetrix/Illumina (2 600 000 (imputed))	2
Bradfield *et al.*[107]	9934/16 956	1120 trios	Affymetrix/Illumina (2 540 000 (imputed))	9
Huang *et al.*[108]	16 179[b]		(6 233 112) (imputed)	2
Onengut-Gumuscu.[c39]	6808/128 352 601 ASP 69 trios		Immunochip (138 229)	44[d]

The main type 1 diabetes genome-wide association studies listed with sample sizes for initial discovery and replication, analysis platform and number of significant observations.
[a]Not reported.
[b]This study used 1000 Genomes-based imputation to identify associations from the Wellcome Trust Case Control Consortium phase 1 Data.[36]
[c]This study is a fine mapping study using the Immunochip.[67]
[d]$P \leqslant 3.23 \times 10E-7$ (Immunochip Bonferroni-corrected $P < 0.05$).

DR3-DQ2 (that is, *DRB1*03, DQA1*05–DQB1*02*), and particularly those present in HLA-DQ2/DQ8 heterozygotes are associated with high susceptibility to T1D. In contrast, a particular DQ6 molecule, encoded by *HLA-DQA1*01:02–DQB1*06:02*, is associated with strong protection from the disease, even in the presence of high-risk HLA alleles and/or T1D-associated autoantibodies.[29] Today, >60 loci associated with T1D have been identified. However, the HLA association remains the strongest by far, with reported odds ratios ranging from 0.02 to >11 for specific DR-DQ haplotypes.[30] After HLA, the strongest T1D genetic association comes from polymorphism in the promoter region of the insulin gene (odds ratios = 2.4).[13] Only two other loci, *PTPN22* and *IL2RA*, have consistently reported odds ratios greater than 1.5; most others are in the range of 1.05–1.25 underscoring the importance of the HLA region compared with other loci.[13] Thus, all studies of T1D genetic susceptibility should take HLA into account when interpreting association data for any other candidate loci.[31]

GWAS STUDIES

During the past decade, GWA studies have represented a paradigm shift in strategies for identifying risk genes for complex (multifactorial) human diseases, including T1D. In GWAS a large number (up to millions) of variants are tested in a hypothesis-free context. The first successful GWAS was published in 2005. It investigated patients with age-related macular degeneration and found two SNPs with significantly altered allele frequency compared to healthy controls.[32] Today, >3000 GWAS publications are catalogued by the National Human Genome Research Institute (NHGRI) and the European Bioinformatics Institute (EMBL-EBI).[33] The Catalog is a quality controlled, manually curated and literature-derived collection of all published GWAS assaying at least 100 000 SNPs and all SNP-trait associations with P-values $< 1.0 \times 10^{-5}$ are reported.[34] The GWAS Catalogue reports 64 SNP-trait associations for T1D, but notably the disease-causing variants and genes are still largely unknown. The leading role in these studies belongs to International Consortia, which possess individual DNA samples from various cohorts; among the main leaders have been the T1DGC (International Type 1 Diabetes Genetics Consortium),[35] and the WTCCC (Welcome Trust Case Control Consortium).[36] Noteworthy, GWAS do not necessarily identify the specific gene or genes in a given locus responsible for the observed disease association, and do not typically inform the wider context in which the disease genes operate.[37,38] Thus, GWAS on their own provide limited insights into the molecular mechanisms driving disease. Numerous GWAS have been performed in T1D, Table 1, and identified several genomic regions associated with increased risk to T1D. There is a significant overlap in study populations between the studies, for example, the WTCCC-, T1DGC-, as well as other cohorts are included in several of the GWAS. Notably, all study populations were of European ancestry.

The principle has been to 'name' these GWAS regions after the gene closest to the strongest associated marker (the lead SNP) or, alternatively, after gene(s) with (some) biological significance for the disease pathology. However, almost all T1D GWAS-associated regions contain multiple genes and thorough fine mapping is therefore essential to narrow down the causal variant(s). The most extensive fine mapping has been performed by the Immunochip study,[39] which was designed to make genetic comparisons across autoimmune disorders as informative as possible. The Immunochip genotyping confirmed and narrowed down most GWAS identified risk loci and also identified new T1D-associated regions ($P < 5 \times 10E\text{-}8$).

The heritability obtained from twin and sibling studies ranges from 0.4 to 0.92.[10,12,40–42] That is, assuming the heritability estimates are correct. This may not always be the case as GWAS have been performed in outbred populations with little evidence of familial clustering of the disease. It is thus quite possible that we have overestimated heritability, which often is inferred from twin or family studies. The familial clustering of T1D, in contrast to most other complex diseases, can be explained almost completely by the multiple common variants identified by GWAS. The estimated proportion of heritability explained by currently identified loci is $>80\%$.[43]

As most variants identified through GWAS contribute to only modest effects to disease risk, it is likely that a combination of variants will better capture effects of clinical relevance. The impact of multiple variants on disease prediction and progression has been evaluated in

candidate gene-based studies,[44–46] and more recently using a genome-wide approach.[47,48] Taken together these studies demonstrate that combining information from several loci in a T1D genetic risk score accurately can identify young adults with diabetes.[47]

PREDICTION OF LIKELY TARGETS

To identify the underlying causal disease mechanisms it is important to prioritize among the many GWAS signals. This is often done by integration of other data sets. A common initial step is to overlap genomic features, such as expression quantitative trait loci (eQTLs), transcription factor-binding sites, DNase hypersensitive sites and histone modifications, with SNP position.[49] This is considered a biologically plausible approach and has provided important insight—especially from the combination of genetic data with expression data. SNPs that influence gene expression are called expression quantitative trait loci (eQTL), and it has been demonstrated that complex trait-loci are enriched for eQTLs.[50]

This appears to be the case also for T1D.[51] A potential limitation is that eQTLs often are tissue specific and studies have often used lymfoblastoid cell lines as a proxy for the autoimmune process in T1D. However, even between immune cell populations eQTLs vary.[51,52] Non-synonymous SNPs change amino-acid composition or truncate the protein sequence by introducing a stop-codon. Synonymous SNPs may affect splicing sites resulting in alternative mRNA isoforms. Structured variations as indels may have the same consequences. Although several non-synonymous SNPs have been identified in T1D risk genes, these missense SNPs are as such not enriched in T1D loci and have, with the exceptions of *HLA* and *PTPN22*, not been convincingly demonstrated to be causal. Identification of differential gene expression profiles in T1D cases and control subjects or in T1D model systems may suggest disease mechanisms/pathways for follow-up studies.[37,53,54] Finally, it was recently demonstrated that GWAS loci, including T1D loci, were enriched for SNPs mapping to regulatory element.[55,56] Thus, several approaches are used for prioritization of risk variants/genes for follow-up studies. In addition, there are several computational tools available for prioritizing SNPs for further downstream analysis (for example, see https://omictools.com/snp-prioritization-category).

Table 2 Candidate genes affecting β-cell functions

Gene (Chromosome)	*Variant(s)*	*Function/pathway affected*	*Reference*
INS (11p15.5)	INS VNTR class I rs7111341 rs11564705[a]	β-cell expression level	70,71
IFIH1 (2q24.2)	rs1990760 rs3747517	MDA5 signalling	73
GLIS3 (9p24.2)	rs7020673	β-cell development β-cell apoptosis GLUT2 expression	109,110
PTPN2 (18p11.21)	rs1893217 rs2542151[a]	Inflammation and virus-induced β-cell apoptosis	111–113
CTSH (15q25.1)	rs3825932 rs11856301[a]	Cytokine-induced apoptosis Insulin transcription	99
BACH2 (6q15)	rs11755527	Cytokine-induced apoptosis	114
TYK2 (19p13.2)	rs2304256	Inflammation and virus-induced β-cell apoptosis	75
CLEC16A (16p13.13)	rs12444268 rs12708716 rs11865121[a]	Autophagy/mitophagy Insulin secretion	115,116

Based on [85] is listed.
[a]T1D candidate genes where experimental studies support their functional significance in β cells. Adapted from[58,62] variant(s) shows the lead SNP identified in GWAS and in cases where the lead SNP is not the most likely one to be functional then the potentially functionally SNP.

TARGET GENES SUBSEQUENTLY FOUND BY FUNCTIONAL STUDIES TO INFLUENCE PATHOPHYSIOLOGY

Determining the mechanisms of action of T1D risk variants is challenging, owing to interaction effects, cell type-specific gene expression, the local tissue milieu, the temporal course of gene expression and complicating environmental factors. Great efforts over the last few years have highlighted potentially functionally target genes that influence T1D pathophysiology. These have mainly been studied in β cells or immune cells. In T1D, the pancreatic β cells in the islets of Langerhans are selectively destroyed by the immune system resulting in absolute insulin deficiency. At least 40% of the genes in the T1D susceptibility loci are expressed in human islets and β cells, where they according to recent studies modulate the β-cell response to the immune system.[37,53,57-62] At least half of these are regulated by cytokines *in vitro*,[37,53,57] an often used model system for T1D pathogenesis. To explore causality of these gene variants in β-cells, functional studies in experimental models, for example, knockdown and overexpression studies, is necessary. Furthermore, studies on knock-in and knock-out mice will, in many cases, aid in the understanding of how the candidate genes affect the disease pathogenesis and contribute to the risk of T1D. This field is in its early stages, but recent studies (reviewed in[58–62]) have identified candidate genes that affect β-cell function or survival in T1D settings, Table 2.

Functional studies of immune cells support the potential impact of T1D risk variants on gene regulation in immune cells.[51,63,64] HLA has a pivotal role as antigen-presenting molecules and as such in the autoimmune process. For non-HLA T1D SNPs strongest evidence comes from their role in regulating gene expression. A recent review found close to 100 eQTLs in different immune cells for T1D risk SNPs.[63] Noteworthy, eQTLs vary between different immune cells, for example, between $CD4^+$ and $CD8^+$ T cells, underlying the importance of studying fractionated immune cell populations. Several studies point to delineation of regulatory or effector pathways as autoreactive $CD4^+$ T cells as key immunological mechanisms affected by T1D genetic risk.[65,66]

NEW INSIGHTS ARISING FROM GWAS

The wealth of data generated by GWAS has also informed our biological understanding of disease processes. This comes from, for example, identification of pleiotropic risk loci, by assessing both disease susceptibility and protection, and by studying the potential role of non-coding RNAs.

Overlapping etiological factors in autoimmune diseases have been recognized for a long time due shared clinical and immunological features. Also, T1D share genetic susceptibility loci with a number of other IMD (immune-mediated diseases). This has been recognized for the HLA region for a long time and more recently GWAS studies have added a considerable number of pleiotropic susceptibility loci, that is, SNPs that confer susceptibility to more than one IMD (reviewed in[67]). Interestingly, T1D loci primarily show concordant overlap, that is, same SNP allele confer risk, with other seropositive autoimmune diseases (for example, autoimmune thyroid disease, rheumatoid arthritis, celiac disease), whereas discordant association, that is, same

Table 3 Examples of pleiotropic non-HLA loci in type 1 diabetes and other immune-mediated diseases

IMD	*Loci*
Alopecia areata	1p13.2, 11q13.1, 12q13.2, 12q24.12
Autoimmune thyroid disease	1p13.2, 2q33.2, 6q15
Celiac disease	2q33.2, 3p21.31, 4q27, 6q15, 6q23.3, 6q25.3, 12q24.12, 15q25.1, 18p11.21, 18q22.2
Crohn's disease	1p13.2, 1q32.1, 11q13.1, 16p11.2, 18p11.21, 19p13.2, 19q13.33
Inflammatory bowel disease	1q32.1, 2q24.2, 6q23.3, 18p11.21, 19q13.33
Juvenile rheumatoid arthritis	1p13.2, 12q24.12, 19p13.2
Multiple sclerosis	6q15, 6q25.3, 11q13.1, 12p13.31, 12q14.1, 16p13.13, 18q22.2, 19p13.2
Narcolepsy	15q25.1
Primary biliary cirrhosis	2q32.3, 6q23.3, 11q13.1, 12q24.12, 16p13.13, 17q12, 17q21.31, 19p13.2
Primary sclerosing cholangitis	12q24.12, 14q24.1
Psoriasis	2q24.2, 19p13.2
Rheumatoid arthritis	1p13.2, 2q11.2, 2q32.3, 2q33.2, 4p15.2, 6q15, 6q23.3, 10p15.1, 12q24.12, 17q12, 19p13.2, 21q22.3
Systemic lupus erythematosus	1p13.2, 1q32.1, 2q32.3, 6q23.3
Ulcerative colitis	1q32.1, 2q24.2, 4q27, 6q23.3, 18p11.21
Vitiligo	1p13.2, 2q24.2, 10p15.1, 12q24.12, 21q22.3

Abbreviations: AA, alopecia areata, ATD, autoimmune thyroid disease, CEL, celiac disease, CRO, Crohn's disease, IBD, inflammatory bowel disease, JRA, juvenile rheumatoid arthritis, MS, multiple sclerosis, NAR, narcolepsy, PBC, primary biliary cirrhosis, PSC, primary sclerosing cholangitis, PSO, psoriasis, RA, rheumatoid arthritis, SLE, Systemic lupus erythematosus, T1D, type 1 diabetes, UC, ulcerative colitis, VIT, vitiligo. Approximately half of the type 1 diabetes peak SNPs show association with another disease. Adapted from[117]

SNP allele shows risk in opposite directions, is more common to seronegative IMD, Table 3. The study of these pleiotropic risk variant holds great potentials for unraveling the functional significance of these variants.

It has been acknowledged for many years that whereas specific HLA haplotypes confer the strongest genetic risk for T1D other HLA haplotypes provide protection against T1D. The DR2-DQ6 (*DQB1*06:02*) haplotype is strongly protective against T1D[29] and in the rare T1D patients that are *DQB1*06:02* positive, the autoimmune process appears to be unique.[68] Although the exact underlying mechanism(s) for the observed HLA susceptibility is (are) not fully understood, the basic function of HLA molecules, that is, peptide presentation, is compliant with the observation of both predisposing and protective HLA molecules in T1D risk. The second T1D susceptibility gene identified in the pre-GWAS era, the insulin gene (*INS*), also has both predisposing and protective variants.[69] In this case, the effect is ascribed to tissue specificity and expression level.[70,71]

Interestingly, fine mapping studies of GWAS-identified candidate genes have demonstrated that in some cases rare variants with opposite effect exist. For example, GWAS-identified non-synonymous variants (rs1990760 and rs3747517) in the *IFIH1* locus on 2q24.2 primarily mediate susceptibility through increased type I interferon production.[72] In addition to these common SNPs, Nejentsev *et al.*[73] reported the presence of four rare SNPs associated with T1D (rs35667974, rs35337543, rs35732034, rs35744605). These rare alleles are associated with protection from T1D in contrast to the common SNPs.[73] The location of T1D-associated SNPs within the *IFIH1* sequence suggests that these point mutations are either negatively or positively affecting the *IFIH1*-encoded protein MDA5 signaling via a mechanism involving the C-terminal end of the gene.[73] Tyrosine Kinase 2 (*TYK2*) is located on 19p13.2 and harbors a non-synonymous SNP (rs2304256) that causes a missense mutation in TYK2 associated with a lower risk of T1D.[74] This is most likely caused by changes in TYK2 expression that dampens the type I interferon response in virus-infected β cells.[75] However, reduced TYK2 expression caused by rare promoter mutations predispose to virus-induced diabetes in rodents[76] and in the Japanese population.[77] Thus, interpretation of association signals from GWAS should be cautious, even when fine mapping has narrowed it down to a single candidate gene.

Therefore, to fully understand disease pathogenesis from GWAS, it is important to analyze the data in the context of complementary datasets, such as transcriptomics, metabolomics, proteomics, under conditions relevant for the disease. It has been advocated that analysis at the pathway, network or protein complex level is the next step in the process of GWAS data mining.[78] Present evidence suggests genetic risk variants for T1D are organized in pathways, physically interact with one another, and are enriched for protein–protein interaction network modules.[37,79,80]

To analyze high-throughput data different data filtering approaches are often taken. Intrinsic data filtering uses information from the dataset itself, such as filtering genetic variants based on linkage to other variants of interest in that data set. Extrinsic data filtering is based on information outside of the data set, such as the inclusion of genomic annotations from separate studies such as ENCODE. At present, both intrinsic and extrinsic data filtering are essential for efficient characterization of T1D genetics. Often multistage and meta-dimensional analysis approaches are used to explore relationships between data sets. In brief, multistage analysis sequentially examines relationships between each data set, and also between each data set and the trait, for example, as the analysis of eQTLs, which includes analysis of genetic variants and gene expression levels. Meta-dimensional analysis takes advantage of simultaneous combination of multiple data types into a single search space to construct a final model. This form of analysis may use several types of data modelling and integration strategies toward the final analysis.[81] Clearly, there is a need for further developing systems biology approaches to provide new insight on T1D biology. We recently used such an approach to construct protein interaction networks from all genes located in non-HLA loci associated with T1D combined with tissue-specic transcriptomic data to identify significantly regulated network.[37,82–85]. This indicates that protein networks can add biological context to candidate genes identified through GWAS.

Interestingly, >90% of disease-associated SNPs map within the non-coding regions of the genome such as promoters, enhancers, intergenic regions and ncRNA genes,[82] suggesting a regulatory role. This supports the concept that changes in regulation of gene

expression are a key factor in T1D development. Our understanding of the modes of action of most ncRNAs, excluding miRNAs, is still very rudimentary.[83] It is a major challenge to develop tools and models that will capture the function of ncRNAs, especially the longer ones, where function is executed by structured elements rather than defined by linear sequences.[84] Nevertheless, it is foreseen that a more in-depth understanding of ncRNAs may in near future open new strategies for diagnostics, classification and personalized therapeutic regiments in T1D and other immune-mediated diseases.[85]

LIMITATIONS OF PUBLISHED GWAS, OTHER CHALLENGES AND PRIORITIES IN THE FIELD

Although GWAS have a range of potentials as discussed above there is a number of limitations and challenges as well. Missing heritability owing to the undiscovered effect of rare variants and epistasis are two key concerns.[86] GWAS studies predominantly test for association of common SNPs, that is, with minor allele frequency > 5%. Rare variants do not seem to account for a major part of T1D susceptibility, though it has been demonstrated that rare variant in, for example, *IFIH1* affect T1D risk.[73] Epistasis is difficult to test properly, but most studies suggest that this is not a major player in T1D predisposition.[20,87,88] However, it is possible that many GWAS SNPs having low or moderate risk in themselves, that is, *P*-values just above $5 \times 10E\text{-}8$, interact to confer a significant combined effect.

Epigenetics is the collectively heritable changes in phenotype due to processes that arise independent of primary DNA sequence. Epigenetic mechanisms include DNA methylation, histone modifications and RNA interference. All of them are associated with regulation and determination of the cellular transcriptome, thereby pivotal to cell function.[89] Because epigenetic modifications are inherited across generations, but are not assayed by genotyping chips or by whole-genome sequencing,[90] it is difficult to exclude the possibility that epigenetic alterations could account for a proportion of T1D risk normally attributed to genetics.

Interestingly, accumulating evidence suggests that ncRNAs are involved in T1D pathogenesis, see above, genetic association between T1D and histone deacetylases exists,[20] and histone deacetylases inhibitors promote β-cell development, proliferation, differentiation and function.[91] Finally, increased DNA methylation variability in T1D across different immune effector cell types has been reported.[14] Nevertheless, at this time, estimates of epigenetic missing heredity are not widespread for complex diseases. However, this is a major research area to be explored to complete our understanding of T1D genetic risk.

GWAS in T1D have been focused on populations of European descent. However, the degree to which observation gained from these studies is transferable to other populations has not been extensively explored. Nevertheless, this has facilitated the success of GWAS owing to more homogeneous populations being studied.[92–94] It is essential to perform genetic studies in non-European populations to bring medical advances from genetic studies to populations worldwide. Also characterizing risk variants beyond what can be achieved with populations of European descent alone will be crucial. The motivation for genetic studies in diverse populations is obvious as differences in disease allele frequency and LD patterns, in phenotypic prevalence, and in effect size as well as differences in rare variants exist. Several observations suggest that no single population is sufficient for fully uncovering the variants underlying T1D in all populations.

Common statistical analysis of GWAS data tests single-nucleotide polymorphisms (SNPs) individually.[95] However, single SNP analysis may be underpowered, especially for low-frequency SNPs.[96] Individual genes, gene networks and pathways are all genetic entities that are likely to have multiple SNPs that function simultaneously to affect diseases and traits. Several novel methods for analyzing GWAS data are being developed.[97]

A final restraint in GWAS is the classical case–control design. The power of GWAS to identify a true association between a SNP and trait is dependent on the phenotypic variance within the population explained by the SNP. However, to date most GWAS of T1D have identified risk loci by comparing cases versus controls independent of heterogenic disease sub-phenotypes, persisting autoantibody positivity, disease progression, for example, loss of β-cell function/mass, complication status and so on. Studies are now emerging looking at genetic risk for autoimmunity development before clinical onset,[98] for decline in β-cell function after diagnosis,[99] for persistent autoimmunity[100,101] among other T1D-related sub-phenotypes. Intriguingly, a number of new associations are reported from these studies emphasizing that we still have a lot to learn about T1D genetics.

CONCLUSIONS

This review summarizes some of the advances and challenges in GWAS-identification of T1D-associated risk variants. GWAS have led to the identification of >60 loci associated with T1D. Importantly, these loci seem to explain most of the heritability of T1D, which is in contrast to most other complex disorders. Thus, it is timely to dissect and translate this genetic predisposition to a deeper understanding of disease biology. Current genetic understanding is being leveraged by complementary studies in epigenetics, transcriptomics, proteomics, metabolomics and lipidomics of both β cells and immune cells. Novel approaches as next-generation-sequencing of both DNA and especially RNA with the ability to identify allelic imbalance, to quantify gene expression in a transcript-specific manner, and to capture unexpected alternative splicing, truncation and post-transcriptional modification events hold great promises. The use of inducible pluripotent stem cells holds great potential for functional studies of, for example, β-cells in the context of relevant genetic risk profiles. A challenge not easily accomplished by the use of human donor pancreatic islets. It is expected that this genetic insight will transform the landscape of common complex diseases such as T1D and lead to novel treatment, preventive strategies and enable precision medicine based on genetic profiling.

CONFLICT OF INTEREST

The author declares no conflict of interest.

1 Mathis D, Vence L, Benoist C. beta-Cell death during progression to diabetes. *Nature* 2001; **414**: 792–798.
2 You WP, Henneberg M. Type 1 diabetes prevalence increasing globally and regionally: the role of natural selection and life expectancy at birth. *BMJ Open Diabetes Res Care* 2016; **4**: e000161.
3 Dabelea D, Mayer-Davis EJ, Saydah S, Imperatore G, Linder B, Divers J *et al.* Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *JAMA* 2014; **311**: 1778–1786.
4 Imperatore G, Boyle JP, Thompson TJ, Case D, Dabelea D, Hamman RF *et al.* Projections of type 1 and type 2 diabetes burden in the U.S. population aged <20 years through 2050: dynamic modeling of incidence, mortality, and population growth. *Diabetes Care* 2012; **35**: 2515–2520.
5 Workgroup on Hypoglycemia ADA. Defining and reporting hypoglycemia in diabetes: a report from the American Diabetes Association Workgroup on Hypoglycemia. *Diabetes Care* 2005; **28**: 1245–1249.
6 Gubitosi-Klug RA, Braffett BH, White NH, Sherwin RS, Service FJ, Lachin JM *et al.* Risk of severe hypoglycemia in type 1 diabetes over 30 years of follow-up in the DCCT/EDIC study. *Diabetes Care* 2017; **40**: 1010–1016.

7 Nathan DM, Cleary PA, Backlund JY, Genuth SM, Lachin JM, Orchard TJ *et al.* Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med* 2005; **353**: 2643–2653.
8 The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *N Engl J Med* 1993; **329**: 977–986.
9 Writing Group for the DERG, Orchard TJ, Nathan DM, Zinman B, Cleary P, Brillon D *et al.* Association between 7 years of intensive treatment of type 1 diabetes and long-term mortality. *JAMA* 2015; **313**: 45–53.
10 Hyttinen V, Kaprio J, Kinnunen L, Koskenvuo M, Tuomilehto J. Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study. *Diabetes* 2003; **52**: 1052–1055.
11 Olmos P, A'Hern R, Heaton DA, Millward BA, Risley D, Pyke DA *et al.* The significance of the concordance rate for type 1 (insulin-dependent) diabetes in identical twins. *Diabetologia* 1988; **31**: 747–750.
12 Kyvik KO, Green A, Beck-Nielsen H. Concordance rates of insulin dependent diabetes mellitus: a population based study of young Danish twins. *BMJ* 1995; **311**: 913–917.
13 Pociot F, Akolkar B, Concannon P, Erlich HA, Julier C, Morahan G *et al.* Genetics of type 1 diabetes: what's next? *Diabetes* 2010; **59**: 1561–1571.
14 Paul DS, Teschendorff AE, Dang MA, Lowe R, Hawa MI, Ecker S *et al.* Increased DNA methylation variability in type 1 diabetes across three immune effector cell types. *Nat Commun* 2016; **7**: 13555.
15 Rewers M, Ludvigsson J. Environmental risk factors for type 1 diabetes. *Lancet* 2016; **387**: 2340–2348.
16 Davies JL, Kawaguchi Y, Bennett ST, Copeman JB, Cordell HJ, Pritchard LE *et al.* A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 1994; **371**: 130–136.
17 Hashimoto L, Habita C, Beressi JP, Delepine M, Besse C, Cambon-Thomsen A *et al.* Genetic mapping of a susceptibility locus for insulin-dependent diabetes mellitus on chromosome 11q. *Nature* 1994; **371**: 161–164.
18 Mein CA, Esposito L, Dunn MG, Johnson GC, Timms AE, Goy JV *et al.* A search for type 1 diabetes susceptibility genes in families from the United Kingdom. *Nat Genet* 1998; **19**: 297–300.
19 Cox NJ, Wapelhorst B, Morrison VA, Johnson L, Pinchuk L, Spielman RS *et al.* Seven regions of the genome show evidence of linkage to type 1 diabetes in a consensus analysis of 767 multiplex families. *Am J Hum Genet* 2001; **69**: 820–830.
20 Nerup J, Pociot F. European Consortium for IS. A genomewide scan for type 1-diabetes susceptibility in Scandinavian families: identification of new loci with evidence of interactions. *Am J Hum Genet* 2001; **69**: 1301–1313.
21 Concannon P, Chen WM, Julier C, Morahan G, Akolkar B, Erlich HA *et al.* Genome-wide scan for linkage to type 1 diabetes in 2,496 multiplex families from the Type 1 Diabetes Genetics Consortium. *Diabetes* 2009; **58**: 1018–1022.
22 Concannon P, Erlich HA, Julier C, Morahan G, Nerup J, Pociot F *et al.* Type 1 diabetes: evidence for susceptibility loci from four genome-wide linkage scans in 1,435 multiplex families. *Diabetes* 2005; **54**: 2995–3001.
23 Morahan G, Mehta M, James I, Chen WM, Akolkar B, Erlich HA *et al.* Tests for genetic interactions in type 1 diabetes: linkage and stratification analyses of 4,422 affected sib-pairs. *Diabetes* 2011; **60**: 1030–1040.
24 Concannon P, Onengut-Gumuscu S, Todd JA, Smyth DJ, Pociot F, Bergholdt R *et al.* A human type 1 diabetes susceptibility locus maps to chromosome 21q22.3. *Diabetes* 2008; **57**: 2858–2861.
25 Polychronakos C, Li Q. Understanding type 1 diabetes through genetics: advances and prospects. *Nat Rev Genet* 2011; **12**: 781–792.
26 Cudworth AG, Woodrow JC. Letter: HL-A antigens and diabetes mellitus. *Lancet* 1974; **2**: 1153.
27 Nerup J, Platz P, Andersen OO, Christy M, Lyngsoe J, Poulsen JE *et al.* HL-A antigens and diabetes mellitus. *Lancet* 1974; **2**: 864–866.
28 Singal DP, Blajchman MA. Histocompatibility (HL-A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. *Diabetes* 1973; **22**: 429–432.
29 Pugliese A, Gianani R, Moromisato R, Awdeh ZL, Alper CA, Erlich HA *et al.* HLA-DQB1*0602 is associated with dominant protection from diabetes even among islet cell antibody-positive first-degree relatives of patients with IDDM. *Diabetes* 1995; **44**: 608–613.
30 Erlich H, Valdes AM, Noble J, Carlson JA, Varney M, Concannon P *et al.* HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. *Diabetes* 2008; **57**: 1084–1092.
31 Noble JA, Erlich HA. Genetics of type 1 diabetes. *Cold Spring Harb Perspect Med* 2012; **2**: a007732.
32 Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C *et al.* Complement factor H polymorphism in age-related macular degeneration. *Science* 2005; **308**: 385–389.
33 MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* 2017; **45**: D896–D901.
34 Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009; **106**: 9362–9367.
35 Rich SS, Akolkar B, Concannon P, Erlich H, Hilner JE, Julier C *et al.* Overview of the type I diabetes genetics consortium. *Genes Immun* 2009; **10**: S1–S4.
36 Wellcome Trust. Case Control C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; **447**: 661–678.
37 Bergholdt R, Brorsson C, Palleja A, Berchtold LA, Floyel T, Bang-Berthelsen CH *et al.* Identification of novel type 1 diabetes candidate genes by integrating genome-wide association data, protein-protein interactions, and human pancreatic islet gene expression. *Diabetes* 2012; **61**: 954–962.
38 Zhong H, Yang X, Kaplan LM, Molony C, Schadt EE. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *Am J Hum Genet* 2010; **86**: 581–591.
39 Onengut-Gumuscu S, Chen WM, Burren O, Cooper NJ, Quinlan AR, Mychaleckyj JC *et al.* Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat Genet* 2015; **47**: 381–386.
40 Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J *et al.* Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. *Diabetologia* 1992; **35**: 1060–1067.
41 Reich T, James JW, Morris CA. The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. *Ann Hum Genet* 1972; **36**: 163–184.
42 Lam HV, Nguyen DT, Nguyen CD. Sibling method increases risk assessment estimates for type 1 diabetes. *PLoS ONE* 2017; **12**: e0176341.
43 Wray NR, Yang J, Goddard ME, Visscher PM. The genetic interpretation of area under the ROC curve in genomic profiling. *PLoS Genet* 2010; **6**: e1000864.
44 Andersen ML, Rasmussen MA, Porksen S, Svensson J, Vikre-Jorgensen J, Thomsen J *et al.* Complex multi-block analysis identifies new immunologic and genetic disease progression patterns associated with the residual beta-cell function 1 year after diagnosis of type 1 diabetes. *PLoS ONE* 2013; **8**: e64632.
45 Steck AK, Wong R, Wagner B, Johnson K, Liu E, Romanos J *et al.* Effects of non-HLA gene polymorphisms on development of islet autoimmunity and type 1 diabetes in a population with high-risk HLA-DR,DQ genotypes. *Diabetes* 2012; **61**: 753–758.
46 Winkler C, Krumsiek J, Lempainen J, Achenbach P, Grallert H, Giannopoulou E *et al.* A strategy for combining minor genetic susceptibility genes to improve prediction of disease in type 1 diabetes. *Genes Immun* 2012; **13**: 549–555.
47 Oram RA, Patel K, Hill A, Shields B, McDonald TJ, Jones A *et al.* A type 1 diabetes genetic risk score can aid discrimination between Type 1 and Type 2 diabetes in young adults. *Diabetes Care* 2016; **39**: 337–344.
48 Patel KA, Oram RA, Flanagan SE, De Franco E, Colclough K, Shepherd M *et al.* Type 1 diabetes genetic risk score: a novel tool to discriminate monogenic and type 1 diabetes. *Diabetes* 2016; **65**: 2094–2099.
49 Hou L, Zhao H. A review of post-GWAS prioritization approaches. *Front Genet* 2013; **4**: 280.
50 Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* 2010; **6**: e1000888.
51 Ram R, Mehta M, Nguyen QT, Larma I, Boehm BO, Pociot F *et al.* Systematic evaluation of genes and genetic variants associated with type 1 diabetes susceptibility. *J Immunol* 2016; **196**: 3043–3053.
52 Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A *et al.* Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet* 2012; **44**: 502–510.
53 Eizirik DL, Sammeth M, Bouckenooghe T, Bottu G, Sisino G, Igoillo-Esteve M *et al.* The human pancreatic islet transcriptome: expression of candidate genes for type 1 diabetes and the impact of pro-inflammatory cytokines. *PLoS Genet* 2012; **8**: e1002552.
54 Pociot F, Kaur S, Nielsen LB. Effects of the genome on immune regulation in type 1 diabetes. *Pediatr Diabetes* 2016; **17**: 37–42.
55 Hoffman MM, Ernst J, Wilder SP, Kundaje A, Harris RS, Libbrecht M *et al.* Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res* 2013; **41**: 827–841.
56 Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome. *Genome Res* 2012; **22**: 1748–1759.
57 Storling J, Brorsson CA. Candidate genes expressed in human islets and their role in the pathogenesis of type 1 diabetes. *Curr Diabetes Rep* 2013; **13**: 633–641.
58 Floyel T, Kaur S, Pociot F. Genes affecting beta-cell function in type 1 diabetes. *Curr Diabetes Rep* 2015; **15**: 97.
59 Lopes M, Kutlu B, Miani M, Bang-Berthelsen CH, Storling J, Pociot F *et al.* Temporal profiling of cytokine-induced genes in pancreatic beta-cells by meta-analysis and network inference. *Genomics* 2014; **103**: 264–275.
60 Santin I, Dos Santos RS, Eizirik DL. Pancreatic beta cell survival and signaling pathways: effects of type 1 diabetes-associated genetic variants. *Methods Mol Biol* 2016; **1433**: 21–54.
61 Santin I, Eizirik DL. Candidate genes for type 1 diabetes modulate pancreatic islet inflammation and beta-cell apoptosis. *Diabetes Obes Metab* 2013; **15**: 71–81.
62 Storling J, Pociot F. Type 1 diabetes candidate genes linked to pancreatic islet cell inflammation and beta-cell apoptosis. *Genes* 2017; **8**: 2.
63 Ram R, Morahan G. Effects of type 1 diabetes risk alleles on immune cell gene expression. *Genes* 2017; **8**: 6.
64 Bakay M, Pandey R, Hakonarson H. Genes involved in type 1 diabetes: an update. *Genes* 2013; **4**: 499–521.
65 Arif S, Tree TI, Astill TP, Tremble JM, Bishop AJ, Dayan CM *et al.* Autoreactive T cell responses show proinflammatory polarization in diabetes but a regulatory phenotype in health. *J Clin Invest* 2004; **113**: 451–463.
66 Lawson JM, Tremble J, Dayan C, Beyan H, Leslie RD, Peakman M *et al.* Increased resistance to CD4+CD25hi regulatory T cell-mediated suppression in patients with type 1 diabetes. *Clin Exp Immunol* 2008; **154**: 353–359.

67 Parkes M, Cortes A, van Heel DA, Brown MA. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nat Rev Genet* 2013; **14**: 661–673.
68 Pugliese A, Kawasaki E, Zeller M, Yu L, Babu S, Solimena M *et al.* Sequence analysis of the diabetes-protective human leukocyte antigen-DQB1*0602 allele in unaffected, islet cell antibody-positive first degree relatives and in rare patients with type 1 diabetes. *J Clin Endocrinol Metab* 1999; **84**: 1722–1728.
69 Bennett ST, Wilson AJ, Cucca F, Nerup J, Pociot F, McKinney PA *et al.* IDDM2-VNTR-encoded susceptibility to type 1 diabetes: dominant protection and parental transmission of alleles of the insulin gene-linked minisatellite locus. *J Autoimmun* 1996; **9**: 415–421.
70 Pugliese A, Zeller M, Fernandez A Jr., Zalcberg LJ, Bartlett RJ, Ricordi C *et al.* The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. *Nat Genet* 1997; **15**: 293–297.
71 Vafiadis P, Bennett ST, Todd JA, Nadeau J, Grabs R, Goodyer CG *et al.* Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. *Nat Genet* 1997; **15**: 289–292.
72 Gorman JA, Hundhausen C, Errett JS, Stone AE, Allenspach EJ, Ge Y *et al.* The A946T variant of the RNA sensor IFIH1 mediates an interferon program that limits viral infection but increases the risk for autoimmunity. *Nat Immunol* 2017; **18**: 744–752.
73 Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science* 2009; **324**: 387–389.
74 Wallace C, Smyth DJ, Maisuria-Armer M, Walker NM, Todd JA, Clayton DG. The imprinted DLK1-MEG3 gene region on chromosome 14q32.2 alters susceptibility to type 1 diabetes. *Nat Genet* 2010; **42**: 68–71.
75 Marroqui L, Dos Santos RS, Floyel T, Grieco FA, Santin I, Op de Beeck A *et al.* TYK2, a candidate gene for type 1 diabetes modulates apoptosis and the innate immune response in human pancreatic beta-cells. *Diabetes* 2015; **64**: 3808–3817.
76 Izumi K, Mine K, Inoue Y, Teshima M, Ogawa S, Kai Y *et al.* Reduced Tyk2 gene expression in beta-cells due to natural mutation determines susceptibility to virus-induced diabetes. *Nat Commun* 2015; **6**: 6748.
77 Nagafuchi S, Kamada-Hibio Y, Hirakawa K, Tsutsu N, Minami M, Okada A *et al.* TYK2 promoter variant and diabetes mellitus in the Japanese. *EBioMedicine* 2015; **2**: 744–749.
78 Gustafsson M, Nestor CE, Zhang H, Barabasi AL, Baranzini S, Brunak S *et al.* Modules, networks and systems medicine for understanding disease and aiding diagnosis. *Genome Med* 2014; **6**: 82.
79 Berchtold LA, Storling ZM, Ortis F, Lage K, Bang-Berthelsen C, Bergholdt R *et al.* Huntingtin-interacting protein 14 is a type 1 diabetes candidate protein regulating insulin secretion and beta-cell apoptosis. *Proc Natl Acad Sci USA* 2011; **108**: E681–E688.
80 Bergholdt R, Storling ZM, Lage K, Karlberg EO, Olason PI, Aalund M *et al.* Integrative analysis for finding genes and networks involved in diabetes and other complex diseases. *Genome Biol* 2007; **8**: R253.
81 Vaitsiakhovich T, Drichel D, Herold C, Lacour A, Becker T. METAINTER: meta-analysis of multiple regression models in genome-wide association studies. *Bioinformatics* 2015; **31**: 151–157.
82 Ricano-Ponce I, Wijmenga C. Mapping of immune-mediated disease genes. *Annu Rev Genomics Hum Genet* 2013; **14**: 325–353.
83 Ulitsky I. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nat Rev Genet* 2016; **17**: 601–614.
84 Seemann SE, Mirza AH, Hansen C, Bang-Berthelsen CH, Garde C, Christensen-Dalsgaard M *et al.* The identification and functional annotation of RNA structures conserved in vertebrates. *Genome Res* 2017; **27**: 1371–1383.
85 Mirza AH, Kaur S, Pociot F. Long non-coding RNAs as novel players in beta cell function and type 1 diabetes. *Hum Genomics* 2017; **11**: 17.
86 Zuk O, Hechter E, Sunyaev SR, Lander ES. The mystery of missing heritability: genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* 2012; **109**: 1193–1198.
87 Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA *et al.* Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 2009; **41**: 703–707.
88 Clayton DG. Prediction and interaction in complex disease genetics: experience in type 1 diabetes. *PLoS Genet* 2009; **5**: e1000540.
89 Xie Z, Chang C, Zhou Z. Molecular mechanisms in autoimmune type 1 diabetes: a critical review. *Clin Rev Allergy Immunol* 2014; **47**: 174–192.
90 Trerotola M, Relli V, Simeone P, Alberti S. Epigenetic inheritance and the missing heritability. *Hum Genomics* 2015; **9**: 17.
91 Christensen DP, Dahllof M, Lundh M, Rasmussen DN, Nielsen MD, Billestrup N *et al.* Histone deacetylase (HDAC) inhibition as a novel treatment for diabetes mellitus. *Mol Med* 2011; **17**: 378–390.
92 Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet* 2003; **361**: 598–604.
93 Jakobsson M, Scholz SW, Scheet P, Gibbs JR, VanLiere JM, Fung HC *et al.* Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* 2008; **451**: 998–1003.
94 Tishkoff SA, Kidd KK. Implications of biogeography of human populations for 'race' and medicine. *Nat Genet* 2004; **36**: S21–S27.
95 Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet* 2012; **90**: 7–24.
96 Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* 2014; **95**: 5–23.
97 Barnett I, Mukherjee R, Lin X. Thegeneralized higher criticism for testing snp-set effects in genetic association studies. *J Am Stat Assoc* 2017; **112**: 64–76.
98 Torn C, Hadley D, Lee HS, Hagopian W, Lernmark A, Simell O *et al.* Role of type 1 diabetes-associated SNPs on risk of autoantibody positivity in the TEDDY study. *Diabetes* 2015; **64**: 1818–1829.
99 Floyel T, Brorsson C, Nielsen LB, Miani M, Bang-Berthelsen CH, Friedrichsen M *et al.* CTSH regulates beta-cell function and disease progression in newly diagnosed type 1 diabetes patients. *Proc Natl Acad Sci USA* 2014; **111**: 10305–10310.
100 Brorsson CA, Onengut S, Chen WM, Wenzlau J, Yu L, Baker P *et al.* Novel association between immune-mediated susceptibility loci and persistent autoantibody positivity in type 1 diabetes. *Diabetes* 2015; **64**: 3017–3027.
101 Brorsson CA, Pociot F. Type 1 Diabetes Genetics C Shared genetic basis for type 1 diabetes, islet autoantibodies, and autoantibodies associated with other immune-mediated diseases in families with type 1 diabetes. *Diabetes Care* 2015; **38** (Suppl. 2): S8–S13.
102 Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V *et al.* Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007; **39**: 857–864.
103 Hakonarson H, Grant SF, Bradfield JP, Marchand L, Kim CE, Glessner JT *et al.* A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. *Nature* 2007; **448**: 591–594.
104 Cooper JD, Smyth DJ, Smiles AM, Plagnol V, Walker NM, Allen JE *et al.* Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat Genet* 2008; **40**: 1399–1401.
105 Hakonarson H, Qu HQ, Bradfield JP, Marchand L, Kim CE, Glessner JT *et al.* A novel susceptibility locus for type 1 diabetes on Chr12q13 identified by a genome-wide association study. *Diabetes* 2008; **57**: 1143–1146.
106 Grant SF, Qu HQ, Bradfield JP, Marchand L, Kim CE, Glessner JT *et al.* Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes. *Diabetes* 2009; **58**: 290–295.
107 Bradfield JP, Qu HQ, Wang K, Zhang H, Sleiman PM, Kim CE *et al.* A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet* 2011; **7**: e1002293.
108 Huang J, Ellinghaus D, Franke A, Howie B, Li Y. 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase 1 Data. *Eur J Hum Genet* 2012; **20**: 801–805.
109 Nogueira TC, Paula FM, Villate O, Colli ML, Moura RF, Cunha DA *et al.* GLIS3, a susceptibility gene for type 1 and type 2 diabetes, modulates pancreatic beta cell apoptosis via regulation of a splice variant of the BH3-only protein Bim. *PLoS Genet* 2013; **9**: e1003532.
110 Yang Y, Chang BH, Chan L. Sustained expression of the transcription factor GLIS3 is required for normal beta cell function in adults. *EMBO Mol Med* 2013; **5**: 92–104.
111 Colli ML, Moore F, Gurzov EN, Ortis F, Eizirik DL. MDA5 and PTPN2, two candidate genes for type 1 diabetes, modify pancreatic beta-cell responses to the viral by-product double-stranded RNA. *Hum Mol Genet* 2010; **19**: 135–146.
112 Moore F, Colli ML, Cnop M, Esteve MI, Cardozo AK, Cunha DA *et al.* PTPN2, a candidate gene for type 1 diabetes, modulates interferon-gamma-induced pancreatic beta-cell apoptosis. *Diabetes* 2009; **58**: 1283–1291.
113 Santin I, Moore F, Colli ML, Gurzov EN, Marselli L, Marchetti P *et al.* PTPN2, a candidate gene for type 1 diabetes, modulates pancreatic beta-cell apoptosis via regulation of the BH3-only protein Bim. *Diabetes* 2011; **60**: 3279–3288.
114 Marroqui L, Santin I, Dos Santos RS, Marselli L, Marchetti P, Eizirik DL. BACH2, a candidate risk gene for type 1 diabetes, regulates apoptosis in pancreatic beta-cells via JNK1 modulation and crosstalk with the candidate gene PTPN2. *Diabetes* 2014; **63**: 2516–2527.
115 Soleimanpour SA, Ferrari AM, Raum JC, Groff DN, Yang J, Kaufman BA *et al.* Diabetes susceptibility genes Pdx1 and Clec16a function in a pathway regulating mitophagy in beta-cells. *Diabetes* 2015; **64**: 3475–3484.
116 Soleimanpour SA, Gupta A, Bakay M, Ferrari AM, Groff DN, Fadista J *et al.* The diabetes susceptibility gene Clec16a regulates mitophagy. *Cell* 2014; **157**: 1577–1590.
117 Pociot F, Lernmark A. Genetic risk factors for type 1 diabetes. *Lancet* 2016; **387**: 2331–2339.