Elsevier Editorial System(tm) for The Lancet
Manuscript Draft

Manuscript Number: THELANCET-D-18-01364R1

Title: GENOMIC INSIGHTS INTO THE CAUSES OF TYPE 2 DIABETES

Article Type: Invited Review

Corresponding Author: XX Claudia Langenberg,

Corresponding Author's Institution:

First Author: Claudia Langenberg

Order of Authors: Claudia Langenberg; Luca A Lotta

Abstract: Genome-wide association studies of type 2 diabetes have implicated up to ~250 genomic regions in disease predisposition, with evidence for causal variants and genes emerging for several of these regions. Understanding of the underlying mechanisms, including the interplay between beta-cell failure, insulin sensitivity, appetite regulation, and adipose storage has been facilitated by the integration of multi-dimensional data on diabetes-related intermediate phenotypes, detailed genomic annotations, functional experiments and now multi-"omic" molecular features. Studies in diverse ethnicities and examples from population isolates have highlighted the value and need for a broader genomic approach to this global disease. Ongoing trans-ethnic discovery efforts and large-scale biobanks in diverse populations and ancestries may help to address some of the existing "Eurocentric" bias. Despite rapid progress in the discovery of the highly-polygenic architecture of type 2 diabetes, dominated by common alleles with small, cumulative effects on disease risk, current knowledge has shown little clinical utility for disease prediction or prevention, and only small contributions to subtype classification or stratified approaches to treatment. Successful development of academia-industry partnerships for exome or genome sequencing in large Biobanks can deliver economies of scale, with implications for the future of genomics-focused research.

GENOMIC INSIGHTS INTO THE CAUSES OF TYPE 2 DIABETES

Claudia Langenberg and Luca A. Lotta

MRC Epidemiology Unit, University of Cambridge, United Kingdom.

Abstract

Genome-wide association studies of type 2 diabetes have implicated up to ~250 genomic regions in disease predisposition, with evidence for causal variants and genes emerging for several of these regions. Understanding of the underlying mechanisms, including the interplay between beta-cell failure, insulin sensitivity, appetite regulation, and adipose storage has been facilitated by the integration of multi-dimensional data on diabetes-related intermediate phenotypes, detailed genomic annotations, functional experiments and now multi-"omic" molecular features. Studies in diverse ethnicities and examples from population isolates have highlighted the value and need for a broader genomic approach to this global disease. Ongoing trans-ethnic discovery efforts and large-scale biobanks in diverse populations and ancestries may help to address some of the existing "Eurocentric" bias. Despite rapid progress in the discovery of the highly-polygenic architecture of type 2 diabetes, dominated by common alleles with small, cumulative effects on disease risk, current knowledge has shown little clinical utility for disease prediction or prevention, and only small contributions to subtype classification or stratified approaches to treatment. Successful development of academia-industry partnerships for exome or genome sequencing in large Biobanks can deliver economies of scale, with implications for the future of genomics-focused research.

1. A global view of type 2 diabetes genomics

The genetic basis of type 2 diabetes

In the last decade, hypothesis-free genome wide association studies (GWAS) have been the single most important contributor to identifying genetic determinants of type 2 diabetes (T2D), leading to the discovery of ~100 associated genomic regions or loci.[1-4] The last year has seen a game-change, a leap forward from smaller, cumulative advances to the description of now up to ~250 genome-wide significant loci,[5-8] including a large meta-analysis currently only available in pre-publication format.[8] Several developments have enabled such rapid progress.

A tripling in effective sample size was achieved through the integration of large-scale, accessible resources such as the UK Biobank (http://www.ukbiobank.ac.uk), teamed with the openness of diabetes researchers worldwide to share results of previous studies (GWAS results: http://diagram-consortium.org; exome sequencing and genotyping results: http://www.type2diabetesgenetics.org). This has led to the inclusion of up to 74,000 T2D cases and 820,000 controls of European-descent,[7, 8] with an accompanying increase in power.

The breadth and depth of genetic variation ascertainment has also dramatically improved. Dense and accurate genotype imputation using the sequenced haplotypes of the 1000 Genomes[5] and now the Haplotype Reference Consortium[7-9] has enabled the interrogation of over 10 million genetic variants, with a ~100-fold increase compared with early GWAS efforts.[1] Direct genotyping or sequencing of common to rare alleles of the exonic areas of the genome has enabled a better ascertainment of coding variation.[7, 10] As a consequence, T2D susceptibility variants now range from 0.02-50% in minor allele frequency (MAF), and from 1.04 to 8.05 in per-allele odds ratio.[5-8] While most signals are led by common variants with ever smaller effects, new risk alleles include several that are low-frequency or rare.[5-8] This is consistent with the model of heritability of T2D derived from whole-genome sequencing experiments, characterised by a prominent contribution to heritability of common variation, a small contribution of rare variation and evidence of low selective pressure on predisposition alleles.[10] Future meta-analyses of multiple large biobanks are likely to expand the catalogue of susceptibility variants to some ultra-rare alleles with extreme effect sizes (odds ratios >10), in addition to finding even more common regulatory alleles with very small effects (odds ratios <1.01; **Figure 1**). Direct genotyping and sequencing will be critical to replicate and identify associations for risk variants that are in the rare allele frequency spectrum (MAF below 0.5%),[11] specifically if discovered based on imputed genotypes given the difficulties in accurately imputing these variants even with expanded population reference panels.[9]

Clinical translation of these genomic associations critically depends on our understanding of underlying mechanisms. Establishing causal variants and variant-gene links has been a challenge in genetic studies. Denser imputation,[5, 8] extended genotyping in coding or metabolic-trait associated regions,[3, 6, 7] direct sequencing,[10, 12] larger sample size,[6-8] integration of extensive genomic and regulatory annotations,[13, 14] and progress in analytical fine-mapping approaches[15, 16] have all made this fundamental task easier. For a given association signal, fine-mapping has been able to considerably narrow the size of the genomic region that likely contains the causal variant as well as the list of plausible causal variants in that genomic region.[5, 7, 8, 14] Identified missense or nonsense variants with evidence of causal association[7, 8] are now amenable to more direct *in vitro* and *in vivo* experimental follow-up. These represent critical advances in the translation of robust associations into biological understanding.

Trans-ethnic discovery efforts

The global burden of T2D with large differences in risk within and between populations warrants a global genomic approach to study its predisposition. Despite the progress in the discovery of its genetic basis, T2D is no exception when it comes to the continued underrepresentation of ethnic diversity in genetic research and discovery efforts.[17] Genetic studies in multiple ethnicities are valuable for several reasons: (a) susceptibility variants may be present at appreciable allele frequency only in non-European populations, as elegantly shown by the discovery of associations at the *HNF1A* (via exome sequencing)[18] *SLC16A11* (via GWAS),[19] and *IGF2* (via exome genotyping)[20] loci in populations of Latin American origin; (b) for shared susceptibility loci, trans-ethnic studies increase statistical power for new discoveries;[4, 6] (c) the diverse linkage disequilibrium patterns across ancestries increase the resolution of fine-mapping analyses for the identification of causal variants;[2, 4] (d) exposure to diverse environments may reveal the effect of susceptibility variants which are masked in other settings (i.e. gene-environment interactions).

Important evidence has been emerging from GWAS of South Asian, East Asian, Latin American and African American populations[19, 21-25] and trans-ethnic discovery efforts;[4, 6] however, studies for non-European ancestries remain small in comparison. Ongoing trans-ethnic discovery efforts under the umbrella of the DIAMANTE consortium include over 170,000 T2D cases with around 45% of the effective sample size accounted for by non-European ancestries, currently including African (7%), East Asian (23%), Hispanic or Latino (6%) and South Asian (9%) ancestry participants (personal communication Prof Andrew Morris, University of Liverpool). While this represents a substantial advance compared to previous studies, research in the field remains heavily "Eurocentric".

Genetically isolated populations

Recent studies have clearly demonstrated the value of studying genetically isolated populations. In these settings deleterious variants with large phenotypic effect may raise by chance to higher allele frequencies, due to a phenomenon known as allelic drift.[26] This makes it easier to identify disease associations for such variants in isolated compared to admixed populations, in which these variants may not be present or may be very rare. Such findings can provide insight into the aetiology of T2D that is generalizable outside of the context of the particular population in which they were discovered.

In the Inuit population from Greenland, homozygote carriers of a loss-of-function variant in *TBC1D4* (p.Arg684*) were found to have a ~10 fold higher risk of diabetes, and ~1 standard deviation higher glucose and insulin at 2 hours following an oral glucose challenge.[27] The variant segregates at high frequency in Inuits (MAF 17%), while being rare or monomorphic in other populations. Risk allele carriers have lower levels of GLUT4 in skeletal muscle[27] (**Figure 2**). This finding highlights the causal role of insufficient GLUT4-mediated glucose uptake in muscle for postprandial hyperglycaemia and T2D risk (**Figure 2**). This corroborates evidence from the first report of a mutation significantly impairing GLUT4 translocation, identified in a child with acanthosis nigricans and extreme postprandial hyperinsulinemia carrying a heterozygous premature stop mutation (p.Arg363*) in *TBC1D4*.[28] In the same population, exome sequencing has revealed the association of loss-of-function variant in *ADCY3* with obesity and diabetes,[29] a finding supported publicly available trans-ethnic T2D exome sequencing studies (http://www.type2diabetesgenetics.org),[29] as well as studies of consanguineous families with severe obesity.[30]

Other examples also highlight convergence of evidence from studies of monogenic disease and population isolates. An autosomal dominant missense mutation in *AKT2* was originally identified as the cause of hyperinsulinemia and diabetes in a family with severe insulin resistance and partial lipodystrophy.[31] A recent exome-array and sequencing study detected a low-frequency *AKT2* coding variant (p.Pro50Thr, MAF 1.1% in Finnish participants, but not detected in other populations) to be associated with higher fasting insulin levels,[32] lower uptake of glucose in insulin-sensitive tissues

(shown in a separate recall-by-genotype study)[33] and higher T2D risk, further highlighting the role of *AKT2* in insulin sensitivity.

In Samoans, a founder population with a high prevalence of obesity and T2D, a common (26% allele frequency in Samoans but extremely rare in other populations) "thrifty" missense variant in *CREBRF* (rs373863828, p.Arg457Gln) was associated with substantially higher body mass index (1.4 kg/m^2 per allele).[34] Interestingly, the adiposity-raising allele was associated with lower fasting glucose and protection from T2D. In an adipocyte model, overexpression of Arg457Gln selectively decreased energy use and increased fat storage, in line with its effect on body fat, potentially highlighting the metabolic benefits of a greater capacity of fat storage.[35]

2. Insights into pathways to diabetes through genomic discovery

Genetic studies of diabetes-related intermediate phenotypes in non-diabetic individuals have emerged as a way to gain mechanistic insights into T2D susceptibility that is complementary to disease-focused discovery GWAS.

Glycaemic control and susceptibility to diabetes

Genetic studies of glucose and insulin related measures have been defined by the struggle between sample size and degree of refinement of phenotype ascertainment. Large efforts have focused on widely-available, simple measures as fasting glucose or insulin, and glycated haemoglobin.[36-42] In contrast, efforts based on "gold-standard" measures involving frequently-sampled oral glucose tolerance tests or continuous intravenous measurements that are difficult to obtain at scale have had limited sample sizes.[43-48] The former approach has been successful in the discovery of loci influencing glycaemic traits in non-diabetic individuals, and helped to identify insulin secretory effects as a major driver of associations for several of the common diabetes susceptibility loci (**Figure 3**).[37, 40, 41, 49] These studies also revealed considerable aetiologic heterogeneity in pathways to T2D, highlighting the multifactorial nature of T2D predisposition.[37, 40, 41, 49] Gold-standard based studies have provided in-depth physiologic characterisations of diabetes susceptibility variants,[43, 45] but also shown promise for discovery by identifying loci (e.g. *GRB10*,[44] *BCL2*,[46] *FAM19A2*,[46] *NAT2*[47]) that have eluded discovery in much larger meta-analyses of more widely-available, simpler measures or indices. Both approaches have been instrumental in understanding the underlying mechanisms of common T2D predisposition and the genetic influences on circulating glucose levels, insulin secretion and resistance.

Overlaying diabetes susceptibility variants, glycaemic traits and pancreatic islet regulatory and functional data[13, 14, 50-55] has provided the foundation for an improved understanding of mechanisms linking beta-cell glucose sensing and insulin secretion with T2D risk. High-throughput functional screens of gene silencing in human beta-cell lines are now available and can empower systematic characterisation of the functional impact of novel likely-causal genes on insulin secretion.[55]

A future challenge in this field will be for intermediate trait studies to stay apace with the rapid increases in sample size of diabetes association analyses. This is complicated by the practical difficulties of obtaining fasting samples, let alone more invasive "gold standard" intravenous measurements, in large biobanks.

Excess overall fat and "central" role of peripheral fat

Excess fat is the hallmark of overeating and lack of physical exercise and has been a major focus of genetic research. Large-scale studies of body mass index and related measures have linked genes highly expressed in the central nervous system with general obesity in different ancestries.[56-58] This has provided complementary evidence to original discoveries that have revealed the fundamental role of appetite regulation in monogenic obesity,[59] in particular the leptin-melanocortin axis.

Observational epidemiology has clearly shown that, for a given level of overall adiposity, the distribution of fat in the body is associated with susceptibility to insulin resistance, diabetes and its complications.[60-65] Genetic discovery approaches integrating multiple insulin-resistance related phenotypes have identified variants associated with insulin resistance, increased risk of diabetes and coronary disease, but lower fat mass in peripheral body compartments, in particular legs and subcutaneous regions.[35] These insulin resistance loci are enriched with genes harbouring mutations in Mendelian forms of lipodystrophy and are associated with increased odds of severe lipodystrophic insulin resistance.[35] In the context of other evidence about the protective role of fat deposition in peripheral compartments,[45, 66-72] these results suggest that reduced ability to safely store excess energy

in the peripheral regions of the body leads to ectopic fat storage and higher cardio-metabolic risk in the general population, similar to clinical manifestations of severe forms of lipodystrophy (**Figure 4**).

Molecular mechanisms underlying this specific aetiology include the impaired ability to generate new adipocytes and the regulation of gene expression in these cells, which is supported by experimental evidence around *PPARG*,[73] *KLF14*,[70, 74] *IRS1*, *CCDC92*, *DNAH10* and *L3MBTL3*.[35] Impaired intravascular lipoprotein lipase (LPL)-mediated lipolysis, the mechanism that regulates lipid buffering from the circulation to peripheral tissues, is also implicated[35] (**Figure 4**). Further studies are necessary to understand how fat deposition in specific body compartments influences metabolic disease risk in the general population and move beyond the very simple notion of "apple" and "pear" body shapes.

Prioritising causal pathways by integrating multi-omic data with clinical outcomes

As the sample size of genetic studies has rapidly grown, so has the ability to measure detailed molecular features in biological samples using high-throughput technology. Global patterns of methylation and other epi-genetic features (epigenomics),[75] gene expression (transcriptomics),[76] proteins (proteomics)[77, 78] or metabolites (metabolomics)[79-81] can now be measured at epidemiological scale enabling genetic mapping in genome-wide studies.

Genetic studies of "-omics" molecular features can help advance the understanding of the causes of diabetes (and other complex) diseases in multiple ways, by (a) characterising the phenotypic consequences of diabetes susceptibility variants identified by GWAS; (b) helping to identify causal variants and genes at known susceptibility loci; (c) enabling the estimation of causal associations between molecular traits and disease risk using the principles of Mendelian randomisation.

So far, no studies have systematically followed-up associations of T2D susceptibility variants with circulating metabolomic or proteomic profiles. Studies overlaying regulatory annotations and gene expression in pancreatic beta-cells with T2D GWAS results show the value of transcriptomic analyses for the identification of diabetes susceptibility genes and mechanistic understanding.[13] The increasing availability of similar data on a variety of cell types (including other relevant metabolic tissues, such as skeletal muscle, adipose, liver) will make it possible to systematically assess the relevance of different tissues, cell types and tissue-specific mechanisms in T2D pathophysiology.

"Mendelian Randomization" studies using genetics to assess causal associations between molecular traits and T2D risk have traditionally focused on specific pathways and biomarkers rather than "-omics" profiles. This partly reflects challenges in applying causal inference frameworks in the context of correlated and co-regulated molecular exposures such as blood metabolites or proteins.[82] For example, studies of the branched chain amino acid pathway have provided human genetic evidence of multi-directional causal relationships between their metabolism, insulin resistance and risk of diabetes,[83] building upon observations dating back to the 1960s.[84-86] This and other metabolomics genetic studies[79-81] illustrate how variation at key metabolic regulators (e.g. enzymes) affects a large set of biologically and phenotypically related measures within a pathway. Hence, inference from associations with diabetes may be restricted to a specific locus of regulatory importance rather than levels of one or more metabolites. Studies of the epigenetic patterns associated with obesity and T2D have illustrated that these are mostly consequences rather than the cause of disease processes.[87]

New methods integrating genomic and other "-omics" data have been developed to facilitate more sophisticated studies in this field.[88-91] While individual multi-omic studies have already been conducted in sample sizes of several thousands, these deeply phenotyped epidemiological studies are still comparatively small in the context of genomic research and the relevance of identified loci to metabolic diseases remains largely unexplored. However, they can serve as models for what might be achieved when technology and cost developments enable high-throughput multi-omic phenotyping at

253 the scale of large national Biobanks with hundreds of thousands of participants systematically
254 followed up for a broad range of diseases.
255
256
257

3. Genomics and therapeutics in T2D

The Chief Medical Officer's Report "Generation Genome" published in 2017 highlighted key areas in which genomics can inform disease therapeutics,[92] providing a helpful framework to describe recent developments of genomics-informed therapeutics in T2D (**Table 1**).

New targets

The identification of novel drug targets is a major stated objective of T2D genetic studies. The finding that drug targets supported by human genetics evidence are more likely to succeed in the selective drug development process[93-95] and recent success in the development of new lipid-lowering drugs following genetic findings from different approaches and sources at the *PCSK9*, *LPA*, *APOC3* and *ANGPTL3* loci[96, 97] have exponentially increased interest in this approach.

Three diabetes-susceptibility genes identified in early GWAS studies[1] encode drug targets for existing glucose lowering therapy (*PPARG*, thiazolidinediones; *KCNJ11*/*ABCC8*, sulfonylureas), suggesting that other genes identified through hypothesis-free approaches may become new drug targets. However, human genetics has played a peripheral role in the development of the most recently approved classes of glucose-lowering drugs, including dipeptidyl peptidase-4 inhibitors,[98] GLP1R agonists[98] and sodium-glucose co-transporter-2 inhibitors. [99]

With over a hundred genetic loci now robustly associated with diabetes, why is there still not a clear "PCSK9-like" example? Generic and therapeutic-area specific obstacles may play a role. First, several of the causal genes implicated by GWAS have only recently emerged, but it takes several years for new drugs to enter clinical development. Specifically targeting pancreatic islets, adipocytes, skeletal myocytes or brain cells, as opposed to circulating proteins is challenging. Safety concerns have been a limitation in the development of appetite suppressants modifying genetically-validated targets, as illustrated by the failure of first generation agonists of the melanocortin 4 receptor due to on-target side effects.[100] Also, many loci for T2D act via impaired insulin secretion. The existence of different classes of approved and widely-used insulin secretagogues may limit the interest of pharmaceutical companies in new drug development in this area, given the focus on developing commercially-differentiated products.[94]

Protective loss-of-function variants are particularly interesting for drug development purposes because they provide insights into the likely consequences of inhibiting a gene product and, if carriers are healthy, provide initial evidence of the likely safety of pharmacological inhibition.[93] Sequencing of the early T2D GWAS gene *SLC30A8*, encoding a pancreatic islet zinc transporter (ZnT8), has identified rare loss-of-function variants associated with protection against T2D (odds ratio for carriers, 0.34).[101] More recently, an exome array genotyping study in populations of Latino descent identified a protective variant (odds ratio per allele, 0.80) in *IGF2* associated with incorrect splicing of isoform 2 of the gene, suggesting that selectively inhibiting this isoform in relevant tissues may be therapeutically exploited.[20] However, efforts to inhibit ZnT8 or insulin-like growth factor 2 have yet to reach clinical development. The potential therapeutic implications of recently reported protective associations of a loss-of-function variant in *GPR151* against obesity, diabetes and coronary artery disease also deserve consideration.[102]

Genetically-tailored treatment

Diabetes medicine holds some of the most elegant examples of tailoring treatment to the specific underlying genetic or molecular defect, but all of these relate to monogenic forms of the disease.[103-105] No such examples exist for common susceptibility loci, but this area is certainly understudied.

Response to glucose-lowering treatment shows a measurable degree of heritability and evidence of polygenicity,[106] indirectly suggesting that combinations of multiple alleles might be able to identify

patients who would be more or less responsive to certain drugs. Opportunities in this field arise from the definition and characterization of polygenic scores combining common variants that capture a particular aetiology. Variants with large effects, such as those from population isolates or rare variants with large effects from GWAS in admixed populations could also provide suitable basis for pharmacogenetics applications. By studying the functional consequences of all theoretically possible missense variants in *PPARG*, Majithia and colleagues elegantly showed that diabetes-associated mutations in the gene display heterogeneous *in vitro* response to thiazolidinediones,[107] which could provide the basis for tailored therapy or dosing in carriers of these specific alleles, as illustrated in initial case reports.[108]

Drug dosing or response

Efforts to identify interactions between genetic background and T2D treatment have been the subject of a recent systematic review,[109] reporting that research in the field is mostly based on observational studies rather than randomized controlled trials and candidate gene rather than hypothesis-free approaches, with a few notable exceptions discussed below. In a pharmacogenetic clinical trial, Srinivasan et al. found that *TCF7L2* variants associated with T2D influences the acute response to both glipizide and metformin in people with risk factors for T2D or treatment-naïve T2D patients.[110] Two genome-wide association studies have identified common genetic variants at the *ATM*[111] and *SLC2A2*[112] loci associated with response to metformin. The difference in the effect of metformin for these variants was estimated at around ~0.15-0.17% of HbA1c per allele, roughly corresponding to a daily dose of ~250 mg of metformin.[111, 112] A genome-wide discovery embedded into a clinical trial found novel associations for common and rare variants in *PRPF31*, *CPA6*, and *STAT3* with metformin response.[113] While these findings are important for the understanding of genetic susceptibility to drug response, the low price of metformin and pragmatic focus on reaching the HbA1c therapeutic target or the maximum tolerated dose of this drug are barriers to clinical use of these genetic tests.

Drug repurposing

If "pharmacomimetic" genetic variants can be used to find new targets, they could theoretically be used to find new indications for existing drugs. While there are not yet any established examples of genetically-directed repurposing of approved drugs in diabetes, Imamura et al. used a systematic bioinformatics approach to identify new T2D drug targets, revealing potential repurposing opportunities for drugs targeting the gene products of *GSK3B* and *JUN*.[23] Recent findings around lipoprotein lipase may offer an example of genetically-driven extension of the target population for drugs that are in active development. In late 2016, we reported a gain-of-function variant (rs328, p.Ser447*) in *LPL* associated with insulin sensitivity and protection from diabetes and an independent a loss-of-function variant (rs1801177, p.Asp36Asn) associated with higher diabetes risk.[35] This followed directionally consistent findings for triglyceride levels and heart disease[114], leading to the hypothesis that the several agents targeting the LPL pathway that are in development for the treatment of hypertriglyceridemia[97, 115-118] could also be valuable as insulin sensitizing agents. The association with diabetes of rs328 has since been replicated[119] and the variant or its proxies have emerged in recent GWAS of diabetes,[7, 8] with consistent findings published for a loss-of-function in the natural LPL-inhibitor *ANGPTL4* (rs116843064, p.Glu40Lys),[119] lending powerful support to this hypothesis.

Drug safety

Genetic variants have been used to understand both desired and undesired secondary effects of pharmacological modulation. Similar to statins and genetic variants at their target *HMGCR*,[120, 121] also cholesterol lowering alleles at *NPC1L1* (encoding the target of ezetimibe)[122] and *PCSK9* (PCSK9 inhibitors)[122-124] are associated with a lower risk of coronary heart disease, but higher diabetes risk.

While this suggested that also non-statin cholesterol lowering agents could be associated with higher diabetes risk, recent randomized controlled trials of ezetimibe[125] and PCSK9-inhibitors[126] did not identify large or statistically-significant diabetogenic effects.

Cardiovascular associations of genetic variants that mimic diabetes medications are particularly insightful, given the regulatory requirement that glucose-lowering drugs should not be associated with a higher risk of cardiovascular disease[127] and the ongoing paradigm shift from glycemic control to prevention of complications in diabetes management.[128] The concomitant publication of randomized controlled trials of a glucagon-like peptide receptor 1 (GLP1R) agonists[129, 130] and of genetic studies of a putative gain-of-function variant of *GLP1R* (rs10305492, p.Ala316Thr),[131] both showing cardio-protective associations for GLP1R activation, illustrates this concept. A similar approach has shown cardio-protective associations for a functional variant in *ABCC8* (rs757110, p.Ala1369Ser),[132] providing genetic insights into the cardiovascular effects of sulfonylureas for which clinical trial evidence is inconclusive.[133]

4. Clinical relevance and future outlook

Genomics in the clinic

In addition to diabetes therapeutics, genomics has been proposed to help disease prediction and diagnosis of common subtypes. The polygenic architecture of T2D dominated by many common variants with small effects and tagging several different aetiologies has critical implications for both applications.

The level of prediction achieved by common alleles contrasts with patients' often more deterministic understanding of what "genetic risk" constitutes, which is informed by examples of highly penetrant causes of monogenic diseases. Accurate prediction in combination with preventive lifestyle interventions can be used for targeted primary prevention and avoidance of complications arising from metabolic dysregulation that is present for years before diagnosis.[134] However, existing diabetes prediction models perform well in the general population and can be used to non-invasively identify individuals at high risk.[135] Earlier prospective studies using up to 65 variants have shown that polygenic T2D risk scores improve prediction performance only modestly when considered over and above risk factors that can assessed non-invasively, such as age, sex, body mass index or family history.[136, 137] New approaches based on machine learning and complex statistical modelling have been proposed as improved methods for genetic prediction.[138] Critically, the availability of good inexpensive predictors from a patient's anamnesis or examination (e.g. family history, BMI) together with clinically established tests (i.e. blood glucose, HbA1c) that inform both future risk prediction and diagnosis greatly limits the scope for introduction of genetic information for T2D prediction in the clinic. Evidence exists that established genetic variants have the weakest relative effect and add the least to prediction in people with highest levels of traditional risk factors and hence at highest absolute risk, the exact subgroup of the population in which preventive interventions and the cost-effectiveness of screening would be greatest.

Diabetes is a multifactorial disease and it has been proposed that genetics could help classify common disease subtypes. Apart from gestational diabetes and rare, specific Mendelian forms for which genetics already helps to guide diagnosis and treatment, diabetes is currently crudely classified into two broad types (>90% T2D) based on clinical presentation and rapid requirement of insulin.[139] Thomas et al. have provided evidence for the presentation of T1D up to the sixth decade of life and that a polygenic score specifically associated with type 1 but not type 2 diabetes can help to rule out T1D in late onset cases.[140] But due to the overwhelming predominance of T2D at older ages, the score's positive predictive value is too low to confidently identify late onset T1D patients, in whom initial management may not be optimal if misdiagnosed as T2D. Interesting studies aiming to identify T2D subgroups using data-driven agnostic approaches have recently emerged. For instance, using six diabetes-related parameters (glutamate decarboxylase antibodies, age at diagnosis, BMI, HbA1c, and indices of insulin secretion and resistance), Ahlqvist et al. have provided evidence of five subtypes of T2D that differ in disease trajectories and risk of complications.[141] Using electronic medical records, Li et al. previously reported evidence of three subtypes,[142] suggesting that data-driven aetiologic classifications are influenced by context and data availability. Investigations using "-omics" measurements may further extend these initial attempts at a more refined disease categorisation. While the existance of different aetiologic subtypes in T2D is widely accepted, a robust and definitive classification is missing. In contrast with approaches aimed at classifying T2D in subgroups and categories, it has been proposed that a more nuanced approach to aetiologic classification would better suit the highly polygenic and multifactorial background of this disease.[143]

Genomic medicine and academia-industry partnerships

In the UK, transformative sequencing projects are currently underway. The 100,000 Genome Project delivers the benefits of genomic medicine to NHS cancer and rare disease patients now and has so far sequenced over 55,000 whole genomes (https://www.genomicsengland.co.uk/the-100000-

genomes-project/). This project has shown the need for economies of scale in the delivery of genomic medicine, and NHS England is recommissioning and modernising NHS Genomic Laboratory Services to develop a first-class genomic service. UK Biobank (http://www.ukbiobank.ac.uk/) has transformed opportunities for population research in this country and internationally and is entering a new phase with the announcement to exome-sequence all 500,000 participants by 2019. Such an accelerated timeline was only possible through major funding by a consortium of five pharmaceutical companies, brought together by Regeneron Pharmaceuticals, following from their first initiative to sequence 50,000 participants in collaboration with GSK. In the USA, a partnership between Regeneron and the Geisinger Health System for the exome sequencing of over 50,000 people has already shown the value of integrating genetic and electronic health record data at scale.[144, 145] These are only two examples of commercial partners having access to participant and patient data, including information collected as part of routine clinical care in electronic health records. The financial benefits of such an effort are something to be considered in the light of the scientific opportunities that this investment and the generated sequence information will provide to biomedical researchers worldwide when they gain access, as will be the case for UK Biobank.

At the same time, large-scale Biobanks collecting genetic, physiological, longitudinal electronic health records and other health data have been established in many countries around the word, focusing on clinical hospital populations (e.g. BioVU: https://www.vumc.org/dbmi/biovu), insurance or care provider populations (e.g. Million Veteran Program: https://www.research.va.gov/mvp/), or national population cohorts (e.g. China Kadoorie Biobank, German National Cohort etc). The NIH funded 1 million persons "All of Us" Research Program was originally launched as a "new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes".[146] It specifically aims to actively recruit ethnic minority populations to help address the existing bias by including sufficient numbers of some of the many currently underrepresented groups.

Engaging patients in research

Genomic sequencing puts the patient at the centre of drug discovery and validation. Close collaboration between academia, the pharmaceutical and other industries can catalyse the development of novel therapies for T2D based on genomic insight. To make full use of these opportunities for patients in this country and elsewhere, the research community needs to engage in a dialogue with patients and the public about genomic medicine and research and their implications for uses and misuses of genomic data.

Notwithstanding the potential that such partnerships have to improve health research and outcomes, research commissioned by the Wellcome Trust (https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf) and work delivered by Genomics England as part of their "Genomics Conversation" (https://www.genomicsengland.co.uk/a-year-of-conversations-about-genomics) have highlighted that patients' and the public have concerns about data safety, i.e. non-legitimate uses, and commercial access to health data, including pharmaceutical and insurance companies.

The Chief Medical Officer's Report "Generation Genome" considered the ethical, social and legal implications of genomic medicine in this country[92] and highlighted the need for highest levels of data security for storage of identifiable data. However, it is not possible to give an absolute guarantee of data security and the potential harms arising from criminal data breaches need to be put in proportion with harms arising from restricting legitimate research uses of health data. For people living in the UK, universal free access to the National Health Service means that there is less reason to fear discrimination with regards health care insurance on the grounds of genetic testing. While there is currently no explicit legislation, the existing voluntary agreement with insurance providers also means that an estimated 95% of insurance customers would not need to disclose genetic test results for example for life assurance, critical illness cover, or income protection, as disclosure is tied to the policy value. The House of Commons Select Committee on Science and Technology recently recommended to extend the existing voluntary agreement, but closely monitor patient's views and the experiences in other countries with a legal prohibition (https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/349/34908.htm#_idTextAnchor041).

Conclusions

Advances in genomic research have facilitated rapid progress in clarifying the genetic basis of T2D and characterising causal variants and variant-gene links. Future opportunities lie in larger-scale sequencing, discovery across diverse ancestries, studies in genetically isolated populations and in massive-scale biobanks. Successful development of academia-industry partnerships can deliver economies of scale, with implications for the future of genomics-focused research.

Review Methods

We searched Pubmed from inception to March 1st 2018 using the following search strategy: (Diabetes Mellitus, Type 2[MeSH] OR NIDDM OR Maturity-Onset Diabetes OR Diabetes Mellitus, Noninsulin-Dependent OR Diabetes Mellitus, Adult-Onset OR Adult-Onset Diabetes Mellitus OR Diabetes Mellitus, Adult Onset OR Diabetes Mellitus, Ketosis-Resistant OR Diabetes Mellitus, Ketosis Resistant OR Ketosis-Resistant Diabetes Mellitus OR Diabetes Mellitus, Maturity-Onset OR Diabetes Mellitus, Maturity Onset OR Diabetes Mellitus, Non Insulin Dependent OR Diabetes Mellitus, Non-Insulin-Dependent OR Non-Insulin-Dependent Diabetes Mellitus OR Diabetes Mellitus, Noninsulin Dependent OR Diabetes Mellitus, Slow-Onset OR Diabetes Mellitus, Slow Onset OR Slow-Onset Diabetes Mellitus OR Diabetes Mellitus, Stable OR Stable Diabetes Mellitus OR Diabetes Mellitus, Type II OR Maturity-Onset Diabetes Mellitus OR Maturity Onset Diabetes Mellitus OR MODY OR Type 2 Diabetes Mellitus OR Noninsulin-Dependent Diabetes Mellitus OR T2D OR T2DM OR Type 2 Diabetes[tiab] OR Type 2 diabetes mellitus OR diabetes[ti]) AND (Genome-Wide Association Study[MeSH] OR Association Studies, Genome-Wide OR Association Study, Genome-Wide OR Genome-Wide Association Studies OR Studies, Genome-Wide Association OR Study, Genome-Wide Association OR Genome Wide Association Scan OR Genome Wide Association Studies OR GWA Study OR GWA Studies OR Studies, GWA OR Study, GWA OR Whole Genome Association Analysis OR Whole Genome Association Study OR Genome Wide Association Analysis OR Genome Wide Association Study). To look for new studies published in pre-publication (non peer-reviewed) form, we searched BioRxiv using the advanced search function (https://www.biorxiv.org/search): articles posted in the "genetics" or "genomics" collections, with the key word "diabetes" in the title or abstract, posted between the 1st of January 2017 and the 9th of April 2018. These literature searches were integrated with reference files of the authors and their colleagues, reference lists of original articles, reviews, and meta-analyses. Given exhaustive reviews on early genetic association studies by McCarthy[1] and Morris[2], amongst others, we focused on recent developments and articles providing insights into clinical translation of genetic findings.

Tables

Table 1. Contribution of genetic findings to T2D therapeutics in key areas.

Area of contribution	Rationale	Considerations and examples in T2D research
New drug target identification	In retrospective analyses, drugs with human genetics support are more likely to successfully transition through the drug development pipeline. Rapid development of new lipid-lowering drugs with genetic validation illustrates potential. Loss-of-function variants provide insights into likely efficacy and safety of inhibition, while gain-of-function on stimulation of target.	Genes encoding the targets of glucose lowering agents have been found by early GWAS, but new classes of diabetes drugs have not been developed as a result of human genetics findings. Most loci identified by GWAS have not lead to new drug development. Protective loss-of-function variants in *SLC30A8* and *IGF2* provide interesting examples that still await new drug development.
Mutation specific treatment	Pharmacological interventions may be particularly effective in patients with particular underlying aetiology or genetic predisposition.	Prominent examples are from Mendelian genetics, lack of examples for common forms of diabetes with polygenic aetiologic contribution. Elegant exemplar from systematic study of all possible missense variants of *PPARG*. Opportunities in specific areas have not been fully exploited.
Drug dosing or response	Drugs may require dose-adjustment according to genetic background.	Common variants at the *ATM* and *SLC2A2* loci have been robustly associated with response to metformin, but genetic testing is not used in the clinic. Several studies including a recent trial have proposed an effect of *TCF7L2* variants on response to glucose-lowering drugs.
Drug repurposing	As with new target identification, genetic variants that "mimic" existing therapeutic agents may provide the basis for repurposing.	No established example of repurposing from other therapeutic areas to diabetes. Recent GWAS have explored repurposing opportunities using bioinformatics approaches. Genetic findings around the *LPL* pathway may provide basis for extension of future indications and target population for emerging drugs targeting this pathway.
Drug safety	Genetic variants can inform on desired and undesired secondary effects of pharmacological modulation of the encoded drug target. In diabetes, it is critical to study the cardiovascular safety of existing and new agents.	The example of low-density lipoprotein cholesterol lowering genetic variants in genes encoding targets of cholesterol lowering therapy (i.e. *HMGCR*, *NPC1L1* and *PCSK9*) and diabetes risk illustrate power and challenges of genetic approaches, given the partial consistency between genetic and clinical trial results. *GLP1R* and *ABCC8* variants have been used to gain insights into cardiovascular safety of existing glucose-lowering drugs.

Abbreviations: GWAS, genome-wide association studies.

Figure Legends

Figure 1. Illustrative representation of genome-wide studies in type 2 diabetes and their power to detect certain types of susceptibility alleles for a given sample size. Susceptibility alleles above the solid black lines are detectable with a given approach. The graph is informed by the results of actual historical[2] and current[8] GWAS studies as well as whole-genome and exome sequencing studies that provided an empirical model for the genetic architecture of type 2 diabetes.[9] Exemplar genetic susceptibility loci are reported in the figure. Abbreviations: GWAS, genome-wide association studies; OR, odds ratio.

Figure 2. Aetiologic model for the role of TBC1D4 in GLUT4 translocation and insulin-mediated glucose uptake in the skeletal muscle.

Figure 3. Models for normal and impaired insulin secretion. Genetic variants affecting these processes result in impaired insulin secretion and higher diabetes risk (right panel), e.g. variants at *KCNJ11* and *ABCC8* identified in genome-wide association studies.

Figure 4. Aetiologic model for the contribution of peripheral adipose storage capacity to metabolic and cardiovascular disease and role of adipogenesis and intravascular lipolysis in this process. Some of the images have been samples and modified from SMART, Servier Medical Art, (URL: https://smart.servier.com/) under Creative Commons Attribution 3.0 Unported License (https://creativecommons.org/licenses/by/3.0/).

References

1. McCarthy MI. Genomics, type 2 diabetes, and obesity. *The New England journal of medicine* 2010; **363**(24): 2339-50.
2. Morris AP. Progress in defining the genetic contribution to type 2 diabetes susceptibility. *Current opinion in genetics & development* 2018; **50**: 41-51.
3. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics* 2012; **44**(9): 981-90.
4. DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature genetics* 2014; **46**(3): 234-44.
5. Scott RA, Scott LJ, Magi R, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* 2017; **66**(11): 2888-902.
6. Zhao W, Rasheed A, Tikkanen E, et al. Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nature genetics* 2017; **49**(10): 1450-7.
7. Mahajan A, Wessel J, Willems SM, et al. Refining The Accuracy Of Validated Target Identification Through Coding Variant Fine-Mapping In Type 2 Diabetes. *Nature genetics* 2018;**50**(4): 559-71.
8. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping of an expanded set of type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *bioRxiv* 2018.
9. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics* 2016; **48**(10): 1279-83.
10. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes. *Nature* 2016; **536**(7614): 41-7.
11. 1000 Genomes Project Consortium, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**(7422): 56-65.
12. Steinthorsdottir V, Thorleifsson G, Sulem P, et al. Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nature genetics* 2014; **46**(3): 294-8.
13. Pasquali L, Gaulton KJ, Rodriguez-Segui SA, et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nature genetics* 2014; **46**(2): 136-43.
14. Gaulton KJ, Ferreira T, Lee Y, et al. Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nature genetics* 2015; **47**(12): 1415-25.
15. Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature genetics* 2012; **44**(4): 369-75, S1-3.
16. Wellcome Trust Case Control Consortium, Maller JB, McVean G, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nature genetics* 2012; **44**(12): 1294-301.
17. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature* 2016; **538**(7624): 161-4.
18. SIGMA Type 2 Diabetes Consortium, Estrada K, Aukrust I, et al. Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. *JAMA : the journal of the American Medical Association* 2014; **311**(22): 2305-14.
19. SIGMA Type 2 Diabetes Consortium, Williams AL, Jacobs SB, et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* 2014; **506**(7486): 97-101.
20. Mercader JM, Liao RG, Bell AD, et al. A Loss-of-Function Splice Acceptor Variant in IGF2 Is Protective for Type 2 Diabetes. *Diabetes* 2017; **66**(11): 2903-14.
21. Cho YS, Chen CH, Hu C, et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nature genetics* 2011; **44**(1): 67-72.

22. Kooner JS, Saleheen D, Sim X, et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nature genetics* 2011; **43**(10): 984-9.
23. Imamura M, Takahashi A, Yamauchi T, et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. *Nature communications* 2016; **7**: 10531.
24. Parra EJ, Below JE, Krithika S, et al. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. *Diabetologia* 2011; **54**(8): 2038-46.
25. Ng MC, Shriner D, Chen BH, et al. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS genetics* 2014; **10**(8): e1004517.
26. Andersen MK, Pedersen CE, Moltke I, Hansen T, Albrechtsen A, Grarup N. Genetics of Type 2 Diabetes: the Power of Isolated Populations. *Current diabetes reports* 2016; **16**(7): 65.
27. Moltke I, Grarup N, Jorgensen ME, et al. A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. *Nature* 2014; **512**(7513): 190-3.
28. Dash S, Sano H, Rochford JJ, et al. A truncation mutation in TBC1D4 in a family with acanthosis nigricans and postprandial hyperinsulinemia. *Proceedings of the National Academy of Sciences of the United States of America* 2009; **106**(23): 9350-5.
29. Grarup N, Moltke I, Andersen MK, et al. Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. *Nature genetics* 2018; **50**(2): 172-4.
30. Saeed S, Bonnefond A, Tamanini F, et al. Loss-of-function mutations in ADCY3 cause monogenic severe obesity. *Nature genetics* 2018; **50**(2): 175-9.
31. George S, Rochford JJ, Wolfrum C, et al. A family with severe insulin resistance and diabetes due to a mutation in AKT2. *Science* 2004; **304**(5675): 1325-8.
32. Manning A, Highland HM, Gasser J, et al. A Low-Frequency Inactivating AKT2 Variant Enriched in the Finnish Population Is Associated With Fasting Insulin Levels and Type 2 Diabetes Risk. *Diabetes* 2017; **66**(7): 2019-32.
33. Latva-Rasku A, Honka MJ, Stancakova A, et al. A Partial Loss-of-Function Variant in AKT2 Is Associated With Reduced Insulin-Mediated Glucose Uptake in Multiple Insulin-Sensitive Tissues: A Genotype-Based Callback Positron Emission Tomography Study. *Diabetes* 2018; **67**(2): 334-42.
34. Minster RL, Hawley NL, Su CT, et al. A thrifty variant in CREBRF strongly influences body mass index in Samoans. *Nature genetics* 2016; **48**(9): 1049-54.
35. Lotta LA, Gulati P, Day FR, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nature genetics* 2017; **49**(1): 17-26.
36. Prokopenko I, Langenberg C, Florez JC, et al. Variants in MTNR1B influence fasting glucose levels. *Nature genetics* 2009; **41**(1): 77-81.
37. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature genetics* 2010; **42**(2): 105-16.
38. Saxena R, Hivert MF, Langenberg C, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nature genetics* 2010; **42**(2): 142-8.
39. Soranzo N, Sanna S, Wheeler E, et al. Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes* 2010; **59**(12): 3229-39.
40. Scott RA, Lagou V, Welch RP, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nature genetics* 2012; **44**(9): 991-1005.
41. Manning AK, Hivert MF, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nature genetics* 2012; **44**(6): 659-69.
42. Wheeler E, Leong A, Liu CT, et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS medicine* 2017; **14**(9): e1002383.

43. Ingelsson E, Langenberg C, Hivert MF, et al. Detailed physiologic characterization reveals diverse mechanisms for novel genetic Loci regulating glucose and insulin metabolism in humans. *Diabetes* 2010; **59**(5): 1266-75.
44. Prokopenko I, Poon W, Magi R, et al. A central role for GRB10 in regulation of islet function in man. *PLoS genetics* 2014; **10**(4): e1004235.
45. Scott RA, Fall T, Pasko D, et al. Common genetic variants highlight the role of insulin resistance and body fat distribution in type 2 diabetes, independent of obesity. *Diabetes* 2014; **63**(12): 4378-87.
46. Walford GA, Gustafsson S, Rybin D, et al. Genome-Wide Association Study of the Modified Stumvoll Insulin Sensitivity Index Identifies BCL2 and FAM19A2 as Novel Insulin Sensitivity Loci. *Diabetes* 2016; **65**(10): 3200-11.
47. Knowles JW, Xie W, Zhang Z, et al. Identification and validation of N-acetyltransferase 2 as an insulin sensitivity gene. *The Journal of clinical investigation* 2015; **125**(4): 1739-51.
48. Wood AR, Jonsson A, Jackson AU, et al. A Genome-Wide Association Study of IVGTT-Based Measures of First-Phase Insulin Secretion Refines the Underlying Physiology of Type 2 Diabetes Variants. *Diabetes* 2017; **66**(8): 2296-309.
49. Dimas AS, Lagou V, Barker A, et al. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* 2014; **63**(6): 2158-71.
50. Lyssenko V, Lupi R, Marchetti P, et al. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *The Journal of clinical investigation* 2007; **117**(8): 2155-63.
51. Moran I, Akerman I, van de Bunt M, et al. Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell metabolism* 2012; **16**(4): 435-48.
52. van de Bunt M, Manning Fox JE, Dai X, et al. Transcript Expression Data from Human Islets Links Regulatory Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic Traits to Their Downstream Effectors. *PLoS genetics* 2015; **11**(12): e1005694.
53. van de Bunt M, Lako M, Barrett A, et al. Insights into islet development and biology through characterization of a human iPSC-derived endocrine pancreas model. *Islets* 2016; **8**(3): 83-95.
54. van de Bunt M, Gaulton KJ, Parts L, et al. The miRNA profile of human pancreatic islets and beta-cells and relationship to type 2 diabetes pathogenesis. *PloS one* 2013; **8**(1): e55272.
55. Thomsen SK, Ceroni A, van de Bunt M, et al. Systematic Functional Characterization of Candidate Causal Genes for Type 2 Diabetes Risk Variants. *Diabetes* 2016; **65**(12): 3805-11.
56. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**(7538): 197-206.
57. Akiyama M, Okada Y, Kanai M, et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nature genetics* 2017; **49**(10): 1458-67.
58. Turcot V, Lu Y, Highland HM, et al. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nature genetics* 2018; **50**(1): 26-41.
59. O'Rahilly S. Harveian Oration 2016: Some observations on the causes and consequences of obesity. *Clinical medicine* 2016; **16**(6): 551-64.
60. Yusuf S, Hawken S, Ounpuu S, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. *Lancet* 2005; **366**(9497): 1640-9.
61. InterAct Consortium, Langenberg C, Sharp SJ, et al. Long-term risk of incident type 2 diabetes and measures of overall and regional obesity: the EPIC-InterAct case-cohort study. *PLoS medicine* 2012; **9**(6): e1001230.
62. Stefan N, Häring HU, Hu FB, Schulze MB. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. *Lancet Diabetes Endocrinol* 2013; **1**(2): 10.
63. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015; **518**(7538): 187-96.

64. Emdin CA, Khera AV, Natarajan P, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. *JAMA : the journal of the American Medical Association* 2017; **317**(6): 626-34.
65. Dale CE, Fatemifar G, Palmer TM, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. *Circulation* 2017; **135**(24): 2373-88.
66. Danforth E, Jr. Failure of adipocyte differentiation causes type II diabetes mellitus? *Nature genetics* 2000; **26**(1): 13.
67. Virtue S, Vidal-Puig A. Adipose tissue expandability, lipotoxicity and the Metabolic Syndrome--an allostatic perspective. *Biochimica et biophysica acta* 2010; **1801**(3): 338-49.
68. Pinnick KE, Nicholson G, Manolopoulos KN, et al. Distinct developmental profile of lower-body adipose tissue defines resistance against obesity-associated metabolic complications. *Diabetes* 2014; **63**(11): 3785-97.
69. Kilpelainen TO, Zillikens MC, Stancakova A, et al. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nature genetics* 2011; **43**(8): 753-60.
70. Small KS, Hedman AK, Grundberg E, et al. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nature genetics* 2011; **43**(6): 561-4.
71. Yaghootkar H, Scott RA, White CC, et al. Genetic evidence for a normal-weight "metabolically obese" phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes. *Diabetes* 2014; **63**(12): 4369-77.
72. Yaghootkar H, Lotta LA, Tyrrell J, et al. Genetic Evidence for a Link Between Favorable Adiposity and Lower Risk of Type 2 Diabetes, Hypertension, and Heart Disease. *Diabetes* 2016; **65**(8): 2448-60.
73. Tontonoz P, Hu E, Spiegelman BM. Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. *Cell* 1994; **79**(7): 1147-56.
74. Small KS, Todorcevic M, Civelek M, et al. Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. *Nature genetics* 2018; **50**(4): 572-80.
75. Portela A, Esteller M. Epigenetic modifications and human disease. *Nature biotechnology* 2010; **28**(10): 1057-68.
76. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; **348**(6235): 648-60.
77. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nature communications* 2017; **8**: 14357.
78. Sun BB, Maranville JC, Peters JE, et al. Consequences Of Natural Perturbations In The Human Plasma Proteome. *bioRxiv* 2017.
79. Suhre K, Shin SY, Petersen AK, et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 2011; **477**(7362): 54-60.
80. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. *Nature genetics* 2014; **46**(6): 543-50.
81. Long T, Hicks M, Yu HC, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nature genetics* 2017; **49**(4): 568-78.
82. Wittemans LBL, Lotta, LA, Langenberg, C. Prioritising Risk Factors for Type 2 Diabetes: Causal Inference Through Genetic Approaches. *Current diabetes reports* 2018. *In press.*
83. Lotta LA, Scott RA, Sharp SJ, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. *PLoS medicine* 2016; **13**(11): e1002179.
84. Felig P, Marliss E, Cahill GF, Jr. Plasma amino acid levels and insulin secretion in obesity. *The New England journal of medicine* 1969; **281**(15): 811-6.

85. Newgard CB, An J, Bain JR, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell metabolism* 2009; **9**(4): 311-26.
86. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nature medicine* 2011; **17**(4): 448-53.
87. Wahl S, Drong A, Lehne B, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 2017; **541**(7635): 81-6.
88. Gamazon ER, Wheeler HE, Shah KP, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nature genetics* 2015; **47**(9): 1091-8.
89. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nature genetics* 2016; **48**(3): 245-52.
90. Hore V, Vinuela A, Buil A, et al. Tensor decomposition for multiple-tissue gene expression experiments. *Nature genetics* 2016; **48**(9): 1094-100.
91. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature genetics* 2016; **48**(5): 481-7.
92. Annual report of the Chief Medical Officer 2016: Generation Genome. URL: https://www.gov.uk/government/publications/chief-medical-officer-annual-report-2016-generation-genome.
93. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. *Nature reviews Drug discovery* 2013; **12**(8): 581-94.
94. Cook D, Brown D, Alexander R, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nature reviews Drug discovery* 2014; **13**(6): 419-31.
95. Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nature genetics* 2015; **47**(8): 856-60.
96. Khera AV, Kathiresan S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nature reviews Genetics* 2017; **18**(6): 331-44.
97. Dewey FE, Gusarova V, Dunbar RL, et al. Genetic and Pharmacologic Inactivation of ANGPTL3 and Cardiovascular Disease. *The New England journal of medicine* 2017; **377**(3): 211-21.
98. Drucker DJ, Nauck MA. The incretin system: glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors in type 2 diabetes. *Lancet* 2006; **368**(9548): 1696-705.
99. Gallo LA, Wright EM, Vallon V. Probing SGLT2 as a therapeutic target for diabetes: basic physiology and consequences. *Diabetes & vascular disease research* 2015; **12**(2): 78-89.
100. Greenfield JR, Miller JW, Keogh JM, et al. Modulation of blood pressure by central melanocortinergic pathways. *The New England journal of medicine* 2009; **360**(1): 44-52.
101. Flannick J, Thorleifsson G, Beer NL, et al. Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nature genetics* 2014; **46**(4): 357-63.
102. Emdin CA, Khera AV, Chaffin M, et al. Analysis of predicted loss-of-function variants in UK Biobank identifies variants protective for disease. *Nature communications* 2018; **9**: 1613.
103. Pearson ER, Flechtner I, Njolstad PR, et al. Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations. *The New England journal of medicine* 2006; **355**(5): 467-77.
104. McCarthy MI, Hattersley AT. Learning from molecular genetics: novel insights arising from the definition of genes for monogenic and type 2 diabetes. *Diabetes* 2008; **57**(11): 2889-98.
105. Oral EA, Simha V, Ruiz E, et al. Leptin-replacement therapy for lipodystrophy. *The New England journal of medicine* 2002; **346**(8): 570-8.
106. Zhou K, Donnelly L, Yang J, et al. Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. *Lancet Diabetes Endocrinol* 2014; **2**(6): 481-7.
107. Majithia AR, Tsuda B, Agostini M, et al. Prospective functional classification of all possible missense variants in PPARG. *Nature genetics* 2016; **48**(12): 1570-5.
108. Agostini M, Schoenmakers E, Beig J, et al. A Pharmacogenetic Approach to the Treatment of Patients with PPARG Mutations. *Diabetes* 2018. *In press.* doi: 10.2337/db17-1236.

109. Maruthur NM, Gribble MO, Bennett WL, et al. The pharmacogenetics of type 2 diabetes: a systematic review. *Diabetes care* 2014; **37**(3): 876-86.
110. Srinivasan S, Kaur V, Chamarthi B, et al. TCF7L2 Genetic Variation Augments Incretin Resistance and Influences Response to a Sulfonylurea and Metformin: The Study to Understand the Genetics of the Acute Response to Metformin and Glipizide in Humans (SUGAR-MGH). *Diabetes care* 2018; **41**(3): 554-61.
111. GoDARTS and UKPDS Diabetes Pharmacogenetics Study Group, Wellcome Trust Case Control Consortium, et al. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nature genetics* 2011; **43**(2): 117-20.
112. Zhou K, Yee SW, Seiser EL, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nature genetics* 2016; **48**(9): 1055-9.
113. Rotroff DM, Yee SW, Zhou K, et al. Genetic Variants in CPA6 and PRPF31 are Associated with Variation in Response to Metformin in Individuals with Type 2 Diabetes. *Diabetes* 2018. *In press.* doi: 10.2337/db17-1164.
114. Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators, et al. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *The New England journal of medicine* 2016; **374**(12): 1134-44.
115. Larsson M, Caraballo R, Ericsson M, et al. Identification of a small molecule that stabilizes lipoprotein lipase in vitro and lowers triglycerides in vivo. *Biochemical and biophysical research communications* 2014; **450**(2): 1063-9.
116. Geldenhuys WJ, Aring D, Sadana P. A novel Lipoprotein lipase (LPL) agonist rescues the enzyme from inhibition by angiopoietin-like 4 (ANGPTL4). *Bioorganic & medicinal chemistry letters* 2014; **24**(9): 2163-7.
117. Gaudet D, Alexander VJ, Baker BF, et al. Antisense Inhibition of Apolipoprotein C-III in Patients with Hypertriglyceridemia. *The New England journal of medicine* 2015; **373**(5): 438-47.
118. Graham MJ, Lee RG, Brandt TA, et al. Cardiovascular and Metabolic Effects of ANGPTL3 Antisense Oligonucleotides. *The New England journal of medicine* 2017; **377**(3): 222-32.
119. Liu DJ, Peloso GM, Yu H, et al. Exome-wide association study of plasma lipids in >300,000 individuals. *Nature genetics* 2017; **49**(12): 1758-66.
120. Sattar N, Preiss D, Murray HM, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet* 2010; **375**(9716): 735-42.
121. Swerdlow DI, Preiss D, Kuchenbaecker KB, et al. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2015; **385**(9965): 351-61.
122. Lotta LA, Sharp SJ, Burgess S, et al. Association Between Low-Density Lipoprotein Cholesterol-Lowering Genetic Variants and Risk of Type 2 Diabetes: A Meta-analysis. *JAMA : the journal of the American Medical Association* 2016; **316**(13): 1383-91.
123. Schmidt AF, Swerdlow DI, Holmes MV, et al. PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol* 2017; **5**(2): 97-105.
124. Ference BA, Robinson JG, Brook RD, et al. Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. *The New England journal of medicine* 2016; **375**(22): 2144-53.
125. Blazing MA, IMPROVE-IT Investigators. Incidence of New-onset Diabetes Mellitus in the IMPROVE-IT Trial. *European Society of Cardiology Congress 2015*.
126. de Carvalho LSF, Campos AM, Sposito AC. Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Inhibitors and Incident Type 2 Diabetes: A Systematic Review and Meta-analysis With Over 96,000 Patient-Years. *Diabetes care* 2018; **41**(2): 364-7.
127. FDA. Guidance for Industry. Diabetes Mellitus — Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes 2008. URL: https://www.fda.gov/downloads/Drugs/Guidances/ucm071627.pdf.
128. Lipska KJ, Krumholz HM. Is Hemoglobin A1c the Right Outcome for Studies of Diabetes? *JAMA : the journal of the American Medical Association* 2017; **317**(10): 1017-8.

129. Marso SP, Bain SC, Consoli A, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *The New England journal of medicine* 2016; **375**(19): 1834-44.
130. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *The New England journal of medicine* 2016; **375**(4): 311-22.
131. Scott RA, Freitag DF, Li L, et al. A genomic approach to therapeutic target validation identifies a glucose-lowering GLP1R variant protective for coronary heart disease. *Science translational medicine* 2016; **8**(341): 341ra76.
132. Emdin CA, Klarin D, Natarajan P, et al. Genetic Variation at the Sulfonylurea Receptor, Type 2 Diabetes, and Coronary Heart Disease. *Diabetes* 2017; **66**(8): 2310-5.
133. Palmer SC, Mavridis D, Nicolucci A, et al. Comparison of Clinical Outcomes and Adverse Events Associated With Glucose-Lowering Drugs in Patients With Type 2 Diabetes: A Meta-analysis. *JAMA : the journal of the American Medical Association* 2016; **316**(3): 313-24.
134. Tabak AG, Jokela M, Akbaraly TN, Brunner EJ, Kivimaki M, Witte DR. Trajectories of glycaemia, insulin sensitivity, and insulin secretion before diagnosis of type 2 diabetes: an analysis from the Whitehall II study. *Lancet* 2009; **373**(9682): 2215-21.
135. Kengne AP, Beulens JW, Peelen LM, et al. Non-invasive risk scores for prediction of type 2 diabetes (EPIC-InterAct): a validation of existing models. *Lancet Diabetes Endocrinol* 2014; **2**(1): 19-29.
136. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *The New England journal of medicine* 2008; **359**(21): 2208-19.
137. Talmud PJ, Cooper JA, Morris RW, et al. Sixty-five common genetic variants and prediction of type 2 diabetes. *Diabetes* 2015; **64**(5): 1830-40.
138. Shigemizu D, Abe T, Morizono T, et al. The construction of risk prediction models using GWAS data and its application to a type 2 diabetes prospective cohort. *PloS one* 2014; **9**(3): e92549.
139. Farmer A, Fox R. Diagnosis, classification, and treatment of diabetes. *Bmj* 2011; **342**: d3319.
140. Thomas NJ, Jones SE, Weedon MN, Shields BM, Oram RA, Hattersley AT. Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *Lancet Diabetes Endocrinol* 2018; **6**(2): 122-9.
141. Ahlqvist E, Storm P, Karajamaki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* 2018.
142. Li L, Cheng WY, Glicksberg BS, et al. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. *Science translational medicine* 2015; **7**(311): 311ra174.
143. McCarthy MI. Painting a new picture of personalised medicine for diabetes. *Diabetologia* 2017; **60**(5): 793-9.
144. Dewey FE, Murray MF, Overton JD, et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science* 2016; **354**(6319).
145. Abul-Husn NS, Manickam K, Jones LK, et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science* 2016; **354**(6319).
146. Collins FS, Varmus H. A new initiative on precision medicine. *The New England journal of medicine* 2015; **372**(9): 793-5.

GENOMIC INSIGHTS INTO THE CAUSES OF TYPE 2 DIABETES

Claudia Langenberg and Luca A. Lotta

MRC Epidemiology Unit, University of Cambridge, United Kingdom.

Abstract

Genome-wide association studies of type 2 diabetes have implicated up to ~250 genomic regions in ~~disease~~ disease predisposition, ~~at several of which~~with ~~there is now strong~~ evidence ~~is now emerging~~ for ~~the underlying~~ causal ~~alleles~~ variants and genes emerging for several of these regions. Understanding of the underlying mechanisms, including the interplay between beta-cell failure, insulin sensitivity, appetite regulation, and adipose storage has been facilitated by t~~T~~he integration of multi-dimensional data on ~~glycaemic and anthropometric~~diabetes-related intermediate phenotypes, detailed genomic annotations, functional ~~work results~~experiments and now multi-"omic" molecular features ~~has furthered the understanding of mechanisms underpinning these associations, including the interplay between beta-cell failure, insulin sensitivity, appetite regulation and adipose storage~~. Studies in diverse ethnicities and examples from population isolates have highlighted the value and need for a broader genomic approach to this global disease. Ongoing trans-ethnic discovery efforts and ~~emerging~~ large-scale b~~B~~iobanks in diverse populations and ancestries may help to address some of the existing "Eurocentric" bias. ~~While studies in diverse ethnicities and selected examples from population isolates have stressed the value and need for a broader genomic approach to this global disease, research in the field is still heavily "Eurocentric".~~ Despite rapid progress in the discovery of the highly-polygenic architecture of type 2 diabetes, dominated by common alleles with small, cumulative effects on disease risk, current knowledge has shown ~~no~~ little clinical utility for disease prediction or prevention, and only small contributions to subtype classification or stratified approaches to treatment. Successful development of academia-industry partnerships for exome or genome sequencing in large Biobanks can deliver economies of scale, with implications for the future of genomics-focused research.

~~Development of academia industry partnerships can help to deliver economies of scale and provide opportunities for genomics informed drug development and validation.~~

Formatted: Numbering: Continuous

Formatted: Not Highlight

1. A global view of type 2 diabetes genomics

The genetic basis of type 2 diabetes

In the last decade, hypothesis-free genome wide association studies (GWAS) have been the single most important contributor to identifying genetic determinants of type 2 diabetes (T2D), leading to the discovery of ~100~~20~~ associated genomic regions or loci.[1-4 ~~1-7~~] "The last year has seen a game-change, a leap forward from smaller, cumulative advances to the description of now up to ~250 genome-wide significant loci,[5-8] including a large meta-analysis currently only available in pre-publication format.[8 ~~8~~] Several developments have enabled such rapid progress.

A tripling in effective sample size was achieved through the integration of large-scale, accessible resources such as the UK Biobank (http://www.ukbiobank.ac.uk), teamed with the openness of diabetes researchers worldwide to share results of previous studies (GWAS results: http://diagram-consortium.org; exome sequencing and genotyping results: http://www.type2diabetesgenetics.org~~/~~). This has led to the inclusion of up to~~more than~~ 74,000 T2D cases and 820,000 controls of European-descent,[7, 8~~8~~] with an accompanying increase in power.

The breadth and depth of genetic variation ascertainment has also dramatically improved. Dense and accurate genotype imputation using the sequenced haplotypes of the 1000 Genomes[5] and now the Haplotype Reference Consortium[7-9~~8~~] has enabled the interrogation of ~~over 25~~over 10 million genetic variants, with a ~100-fold increase compared with early GWAS efforts.[1] Direct genotyping or sequencing of common to rare alleles of the exonic areas of the genome has enabled a better ascertainment of coding variation.[7, 10~~7~~] As a consequence, T2D susceptibility variants now range from 0.02-50% in minor allele frequency (MAF), and from 1.04 to 8.05 in per-allele odds ratio.[5-8] While most signals are led by common variants with ever smaller effects, new risk alleles include ~~15~~ several that are low-frequency or rare~~, including two examples of very rare variants with odds ratio ~8~~.[~~6~~8 5-8] This is consistent with the model of heritability of T2D derived from ~~sequencing experiments~~whole-genome sequencing experiments, characterised by a prominent contribution to heritability of common variation, a small contribution of rare variation and evidence of low selective pressure on predisposition alleles.[9~~10~~] Future meta-analyses of multiple large biobanks are likely to expand the catalogue of susceptibility variants to some ultra-rare alleles with extreme effect sizes (odds ratios >10), in addition to finding even more common regulatory alleles with very small effects (odds ratios <1.01; **Figure 1**). Direct genotyping and sequencing will be critical to replicate and identify associations for ~~imputed~~ risk variants that are in the rare allele frequency spectrum (MAF below 0.5%),[11] specifically if discovered based on imputed genotypes~~...~~,[~~11~~] given the [~~11~~]difficulties ~~of~~high ~~quality ation of~~ in accurately imputing these variants even with expanded population reference panels.[9]

Clinical translation of these genomic associations critically depends on our understanding of underlying mechanisms. Establishing causal variants and variant-gene links has been a ~~critical~~ challenge in genetic studies. Denser imputation,[5, 8~~5, 8~~] extended genotyping in coding or metabolic-trait associated regions,[3, 6, 7 ~~3, 6, 7~~] direct sequencing,[10, 12~~10~~] larger sample size,[6-8~~8~~] integration of extensive genomic and regulatory annotations,[13, 14~~11, 12~~] and progress in analytical fine-mapping approaches[15, 16~~13, 14~~] have all made this fundamental task easier. For a given association signal, fine-mapping has been able to considerably narrow the size of the genomic region that likely contains the causal variant as well as the list of plausible causal variants in that ~~window~~genomic region.[5, 7, 8, 14~~5, 8~~] ~~[7]8 Of 380 signals amenable to fine mapping in the latest analysis, 18 could be attributed to a single causal variant with a posterior probability of association >99%, with over 50 signals at 44 loci exceeding 80% for the strongest variant.[8] While most of these likely causal alleles map to regulatory regions in pancreatic islets, several overlap enhancers or promoters in other diabetes-related tissues as adipose, skeletal muscle and liver. Also, there are now a number of likely causal~~ Identified missense or nonsense variants with evidence of causal association[7, 8] are now amenable to more direct *in vitro*

Formatted: Not Highlight
Field Code Changed
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Field Code Changed
Field Code Changed
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Tab stops: 0.75 cm, Left
Formatted: Font: 12 pt, No underline, Font color: Auto, English (United States)
Formatted: Font: 11 pt
Formatted: Font: (Default) Times New Roman, 11 pt, Underline
Field Code Changed
Formatted: Font: (Default) Times New Roman
Field Code Changed
Formatted: Font: (Default) Times New Roman
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: 12 pt, Font color: Auto, English (United States)
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman

and *in vivo* experimental follow-up.[7, 87, 8] These represent critical advances in the translation of robust associations into biological understanding.

Field Code Changed

Field Code Changed

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Trans-ethnic discovery efforts

~~A~~The global ~~health issue like~~burden of T2D with large differences in risk within and between populations warrants a global genomic approach to study its predisposition. Despite the progress in the discovery of its genetic basis, T2D is no exception when it comes to the continued underrepresentation of ethnic diversity in genetic research and discovery efforts.[17~~15~~] Genetic studies in multiple ethnicities are valuable for several reasons: (a) susceptibility variants may be present at appreciable allele frequency only in non-European~~certain~~ populations, as elegantly shown by the discovery of associations at the *HNF1A* (via exome sequencing)[18] *SLC16A11* (via GWAS),[19~~16~~] ~~*HNF1A*[17]~~ and *IGF2* (via exome genotyping)[20~~18~~] loci in populations of Latin American origin; (b) for shared susceptibility loci, trans-ethnic studies increase statistical power for new discoveries;[4, 6~~4, 6~~] (c) the diverse linkage disequilibrium patterns across ancestries increase the resolution of fine-mapping analyses for the identification of causal variants;[2, 4~~2, 4~~] (d) exposure to diverse environments may reveal the effect of susceptibility variants which are masked in other settings (i.e. gene-environment interactions).

Important evidence has been emerging from GWAS of South Asian, East Asian, Latin American and African American populations[19, 21-25 ~~16, 19-23~~] and trans-ethnic discovery efforts;[~~4, 6~~4, 6] however, studies for non-European ancestries remain small in comparison. ~~Collaborative~~ Ongoing trans-ethnic discovery efforts ~~are now underway~~ under the umbrella of the DIAMANTE consortium ~~with~~ include over 170,000 T2D cases ~~across 120 studies and~~with around 45% of the effective sample size accounted for by non-European ancestries, currently including African (7%), East Asian (23%), Hispanic or Latino (6%) and South Asian (9%) ancestry participants (personal communication Prof Andrew Morris, University of Liverpool). While this represents a substantial advance compared to previous studies, research in the field remains heavily "Eurocentric". ~~This will represent a substantial advance in employing genetic diversity across populations for T2D genetic discovery and will improve our ability to generalise findings with important implications for drug discovery and clinical translation for non-European descent populations, many of which are at high risk of T2D.~~

Genetically isolated populations

Recent studies have clearly demonstrated the value of studying genetically isolated populations. In these settings~~, extended linkage disequilibrium and higher allelic frequencies of deleterious variants due to greater allelic drift[24] allowed unique insights into the aetiology of T2D.~~ deleterious variants with large phenotypic effect may raise by chance to higher allele frequencies, due to a phenomenon known as allelic drift.[26] This makes it easier to identify disease associations for such variants in isolated compared to admixed populations, in which these variants may not be present or may be very rare. Such findings can provide insight into the aetiology of T2D that is generalizable outside of the context of the particular population in which they were discovered.

In the Inuit population from Greenland, homozygote carriers of a loss-of-function variant in *TBC1D4* (p.Arg684*) were found to have a ~10 fold higher risk of diabetes, and ~1 standard deviation higher glucose and insulin at 2 hours following an oral glucose challenge.[27~~25~~] The variant segregates at high frequency in Inuits (MAF 17%), while being rare or monomorphic in other populations~~, and influences risk in a fashion consistent with recessive inheritance~~.[~~25~~] ~~The mechanism of action of this variant is also particularly insightful.~~ Risk allele carriers ~~had~~ have lower levels of GLUT4 in skeletal muscle,[27] ~~which is critical in the insulin-mediated control of post-prandial glycaemia~~(**Figure 2**). This finding highlights the causal role of insufficient GLUT4-mediated glucose uptake in muscle for postprandial hyperglycaemia and T2D risk (**Figure 2**). This corroborates evidence from the first report of a mutation significantly impairing GLUT4 translocation, identified in a child with acanthosis nigricans and extreme postprandial hyperinsulinemia carrying a heterozygous premature stop mutation (p.Arg363*) in *TBC1D4*.[28~~26~~]

Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Field Code Changed
Formatted: Font: (Default) Times New Roman
Formatted: Font: Not Italic
Field Code Changed
Field Code Changed
Formatted: Font: Not Italic
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: Bold
Field Code Changed
Formatted: Tab stops: 0.75 cm, Left

~~—~~In the same population, exome sequencing has revealed the association of ~~More recently, an analysis of rare loss-of function variants from exome sequencing in nine Greenlandic family trios identified a~~ loss-of-function variant in *ADCY3* with obesity and diabetes,[29] a finding supported~~that decreases gene expression via a splicing defect.[27] The variant was associated with greater BMI, abdominal fat, insulin resistance, dyslipidaemia and risk of T2D, with evidence for a recessive inheritance.[27] While this variant was not present in non-Greenlandic populations, the authors were able to identify heterozygous carriers of seven other predicted loss-of function variants in *ADCY3* in~~ publicly available trans-ethnic T2D exome sequencing studies (~~URL:~~ http://www.type2diabetesgenetics.org),[29] as well as studies of consanguineous ~~in~~ families with severe obesity,[30] ~~and demonstrate their enrichment in T2D cases versus controls. A parallel study identified the gene by sequencing patients with severe obesity.[28] This demonstrates the success of prioritizing genomic regions for systematic follow-up in other populations based on strong evidence from population isolates or extreme phenotypes in family studies.~~

Other examples also highlight convergence of evidence from studies of monogenic disease and population isolates. An autosomal dominant ~~heterozygous~~ missense mutation in *AKT2* was originally identified as the cause ~~fasting and postprandial~~of ~~-~~hyperinsulinemia and diabetes in a family with severe insulin resistance and partial lipodystrophy.[31~~29~~] A recent exome-array and sequencing study detected a low-frequency *AKT2* coding variant (p.Pro50Thr, MAF 1.1% in Finnish participants, but not detected in other populations) to be~~, which was~~ associated with higher fasting insulin levels,[32~~30~~] lower uptake of glucose in insulin-sensitive tissues (shown in a separate recall-by-genotype study)[33~~31~~] and higher T2D risk, further highlighting the role of *AKT2* in insulin sensitivity.

~~Another interesting example comes from a~~In ~~study of~~ Samoans, a founder population with a high prevalence of obesity and T2D, ~~in whom~~ a common ~~population-specific~~ (26% allele frequency in Samoans but extremely rare in other populations) "thrifty" missense variant in *CREBRF* (rs373863828, p.Arg457Gln) ~~showed an association~~was associated with substantially higher body mass index (1.4 kg/m^2 per allele).[34~~32~~] Interestingly, the adiposity-raising allele was associated with lower fasting glucose and protection from T2D. In an adipocyte model, overexpression of Arg457Gln selectively decreased energy use and increased fat storage, in line with its effect on ~~overall~~ body fat~~, greater abdominal (waist) and gluterofemoral (hip) fat~~, potentially highlighting the metabolic benefits of a greater capacity of fat storage.[35~~33~~]

Field Code Changed

Formatted: Underline

Formatted: Default Paragraph Font, Font: (Default) +Body, 12 pt, Font color: Black, English (United States)

Field Code Changed

Field Code Changed

Formatted: No underline, Not Highlight

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

2. Insights into pathways to diabetes through genomic discovery

Genetic studies of diabetes-related intermediate phenotypes~~continuous metabolic traits~~ in non-diabetic individuals have emerged as a way to gain mechanistic insights into T2D susceptibility that is complementary to ~~disease~~ disease-focused discovery GWAS.

Glycaemic ~~regulation~~ control and susceptibility to diabetes

~~In the absence of intermediate metabolic traits associations, the interpretation and mechanistic follow up of diabetes susceptibility loci would be in some cases impossible.~~ Genetic studies of glucose and insulin related ~~glycaemic and insulinaemic~~ measures have been defined by the struggle between sample size and degree of refinement of phenotype ascertainment. Large efforts have focused on widely-available, ~~and~~ simple measures as fasting glucose or insulin, and glycated haemoglobin ~~or insulin,~~ [36-42] ~~[34-40]~~ ~~while smaller ones one~~ In contrast, efforts based on "gold-standard" ~~assays~~ measures involving frequently-sampled oral glucose tolerance tests or continuous intravenous measurements that are difficult to obtain at scale have had limited sample sizes.[43-48] ~~[41-47]~~ The former approach has been ~~very~~ successful in the discovery of loci influencing glycaemic traits ~~and insulinaemic levels~~ in non-diabetic individuals, and helped to identify insulin secretory ~~and has provided large scale datasets for the initial characterise~~effects as ~~ation of diabetes associated variants~~a major driver of associations for several of the common diabetes susceptibility loci (**Figure 3**).[37, 40, 41, 49] These studies also revealed considerable aetiologic heterogeneity in pathways to T2D, highlighting the multifactorial nature of T2D predisposition.[37, 40, 41, 49] ~~. The latter has been used primarily for~~Gold-standard based studies have provided ~~–~~in-depth physiologic characterisations of diabetes susceptibility variants ~~(REFS)~~,[43, 45] but ~~has~~ also shown promise for ~~as a~~ discovery ~~approach for the identification of novel~~by identifying loci (e.g. *GRB10*,[44] *BCL2*,[46] *FAM19A2*,[46] *NAT2*[47]) that have eluded discovery in much larger meta-analyses of more widely-available, ~~but~~ simpler measures or indices ~~(REFS)~~. Both approaches have been instrumental in understanding the underlying mechanisms of common T2D predisposition and the genetic influences on circulating glucose levels, insulin secretion and resistance.[35, 36, 38, 39, 44]

~~Studies of intermediate traits highlighted the that dominant role of impaired insulin secretion is as a major driver of associations for several of the common diabetes susceptibility loci (**Figure 3**).~~[37, 40, 41, 49] ~~These studies , but also revealed considerable aetiologic heterogeneity.~~[37, 40, 41, 49] ~~[35, 38, 39, 48]~~ Overlaying diabetes susceptibility variants, glycaemic traits and pancreatic islet regulatory and functional data[13, 14, 50-55] ~~[11, 12, 49-54]~~ has provided the foundation for an improved understanding of mechanisms linking beta-cell glucose sensing and insulin secretion with T2D risk. High-throughput functional screens of gene silencing in human beta-cell lines are now available ~~that~~and can empower systematic characterisation of the functional impact of novel likely-causal genes on insulin secretion.[55] ~~Studies of "gold standard" measures have highlighted novel and specific genetic associations.~~[45, 46]

~~In the absence of intermediate metabolic traits associations, the interpretation and mechanistic follow up of diabetes susceptibility loci would be in some cases impossible.~~ A future challenge in this field will be for intermediate trait studies to stay apace with the rapid ~~inflation~~ increases in sample size of diabetes association analyses. This ~~, a task that~~ is complicated by the practical difficulties~~y~~ of obtaining fasting samples, let alone more invasive "gold standard" intravenous measurements, in large biobanks.

Excess overall fat and "central" role of peripheral fat

Excess fat is the hallmark of overeating and lack of physical exercise and has been a major focus of genetic research. Large-scale studies of body mass index and related measures have linked genes highly expressed in the central nervous system with general obesity in different ancestries.[56-58] ~~[55-57]~~

Formatted: No underline, Font color: Auto, Not Superscript/ Subscript
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: Italic, No underline, Font color: Auto
Field Code Changed
Formatted: Font: Italic, No underline, Font color: Auto
Field Code Changed
Formatted: Font: Italic, No underline, Font color: Auto
Field Code Changed
Formatted: Font: Italic, No underline, Font color: Auto
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman, No underline, Font color:
Formatted: Font: Bold, No underline, Font color: Auto
Field Code Changed
Formatted: No underline, Font color: Auto
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: No underline, Font color: Auto
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman, No underline, Font color:
Field Code Changed
Field Code Changed

This has provided complementary evidence to original discoveries that have revealed the fundamental role of appetite regulation in monogenic obesity,[59~~58~~] in particular the leptin-melanocortin axis.

Observational epidemiology has clearly shown that, for a given level of overall adiposity, the distribution of fat in the body is associated with susceptibility to insulin resistance, diabetes and its complications.[60-65] Genetic discovery approaches integrating multiple insulin~~-~~ resistance related phenotypes [59~~-63~~] ~~However, the complex relationship between regional body fat deposition and metabolic risk go beyond the simple notion of "apple" and "pear" shaped body. This concept is reinforced by the aetiologic and phenotypic heterogeneity of partial and generalized lipodystrophies.[64, 65] Using integrative genomic approaches, we and others have identified 53 genomic regions associated with~~Integrative genomic approaches to the identification of insulin resistance genetic determinants have identified ~~numerous genetic~~ variants associated with~~,~~ insulin resistance, increased~~higher~~ risk of diabetes and coronary disease, but lower fat mass in ~~the~~ peripheral body compartments, in particular legs and subcutaneous regions.[35~~33~~] These insulin resistance loci ~~were~~ are enriched with genes harbouring mutations in Mendelian forms of lipodystrophy and ~~were~~ are associated with increased odds of severe lipodystrophic insulin resistance.[35 ~~33~~] ~~In~~ In the context of other evidence about the protective role of fat deposition in peripheral compartments,[45, 66-72 ~~44, 66-71~~] ~~these~~ these results suggest that reduced ~~mean that being unable~~ability to safely store excess energy in the peripheral regions of the body ~~may~~ leads to ectopic fat storage and~~,~~ higher ~~circulating lipids and~~ cardio-metabolic risk in the general population, similar to clinical manifestations of ~~as in~~ severe forms of lipodystrophy (**Figure ~~2~~4**).

Molecular mechanisms underlying this specific aetiology ~~have only started to emerge. These~~ include the impaired ability to generate new adipocytes and the regulation of gene expression in these cells, ~~as suggested~~ which is supported by ~~initial~~ experimental evidence around *PPARG*,[73] *KLF14*,[70, 74~~66~~] *IRS1*, *CCDC92*, *DNAH10* and *L3MBTL3*.[35~~69~~ ~~33, 70~~ ~~;~~ ~~as well as an i~~Impaired intravascular lipoprotein lipase (LPL)-mediated lipolysis,[~~33~~] ~~the~~ the mechanism that regulates lipid buffering from the circulation to peripheral tissues, is also implicated[35] (**Figure ~~2~~4**). Further studies ~~are necessary to understand whether and how these mechanisms play into the well-established relationship between centripetal fat distribution and metabolic risk and~~are necessary to understand how fat deposition in specific body compartments influences metabolic disease risk ~~of disease~~ in the general population and move beyond the very simple notion of "apple" and "pear" body shape~~sd body~~.

Prioritising causal pathways by integrating multi-omic data with clinical outcomes

As the sample size of genetic studies has rapidly grown, so has ~~our~~ the ability to measure detailed molecular features in biological samples using high-throughput technology. Global patterns of methylation and other epi-genetic features (epigenomics),[75~~72~~] gene expression (transcriptomics),[76~~73~~] proteins (proteomics)[77, 78~~74, 75~~] or metabolites (metabolomics)[79-81~~76-78~~] can now be measured at epidemiological scale ~~and show enough heritability to~~ enabling ~~e~~genetic mapping in genome-wide studies.

Genetic studies of "-omics" molecular features can help advance the understanding of the causes of diabetes (and other complex) diseases in multiple ways, by (a) characterising the phenotypic consequences of diabetes susceptibility variants identified by GWAS; (b) helping to identify causal variants and genes at known susceptibility loci; (c) enabling the estimation of causal associations between molecular traits and disease risk using the principles of Mendelian randomisation.

So far, no studies have systematically followed-up associations of T2D susceptibility variants with circulating metabolomic or proteomic profiles. Studies overlaying regulatory annotations and gene expression in pancreatic beta-cells with T2D GWAS results show the value of ~~"-omics"~~transcriptomic ~~technology~~ analyses for the identification of diabetes susceptibility genes and mechanistic understanding.[82~~13~~] The increasing availability of similar data on a variety of cell types

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: Italic, No underline, Font color: Auto
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman, No underline, Font color:
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman, No underline, Font color:
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

(including other relevant metabolic tissues, such as skeletal muscle, adipose, liver) will make it possible to systematically assess the relevance of different tissues, cell types and tissue-specific mechanisms in T2D pathophysiology.

"Mendelian Randomization" studies using genetics to assess causal associations between molecular traits and T2D risk have traditionally focused on specific pathways and biomarkers rather than "-omics" profiles. This ~~is~~ partly ~~the~~ reflect~~ions~~ ~~of~~ challenges in applying causal inference frameworks in the context of correlated and co-regulated molecular exposures such as blood metabolites or proteins ~~{REF}~~.[82] For example, studies of the branched chain amino acid pathway have provided human genetic evidence of multi-directional causal relationships between their metabolism, insulin resistance and risk of diabetes,[83] building upon observations dating back to the 1960s.[84-86] This and other metabolomics genetic studies[79-81] illustrate how variation at key metabolic regulators (e.g. enzymes) affects a large set of biologically and phenotypically related measures within a pathway. Hence, inference from associations with diabetes may be restricted to a specific locus of regulatory importance rather than levels of one or more metabolites. Studies of the epigenetic patterns associated with obesity and ~~secondarily with~~ T2D have illustrated ~~how, often,~~ that these ~~these complex molecular patterns~~ are mostly ~~a the~~ consequences rather than the cause of disease processes.[87]

N~~Several n~~ew methods ~~that~~ integrating~~e~~ genomic ~~etic~~ and other "~~-~~"~~-~~omics"~~s~~" data have been developed to facilitate ~~these~~more sophisticated studies in this field.[88-91] ~~In principle, these approaches can be used to systematically study the causal influences of molecular traits on diabetes risk. In practice, they present a number of challenges that complicate aetiologic inference.~~[79-83]

While individual multi-omic studies have already been conducted in sample sizes of several thousands, these deeply phenotyped epidemiological studies are still comparatively small in the context of genomic research and the~~ir~~ relevance of identified loci to metabolic diseases remains largely unexplored. However, they can serve as models for what might be achieved when technology and cost developments enable high-throughput multi-omic phenotyping at the scale of large national Biobanks with hundreds of thousands of participants systematically followed up for a broad range of diseases.

~~Building upon the first wave of discovery, large-scale meta-analyses of GWAS of metabolite profiles, proteins and epigenetic markers are underway[88] which will enable the systematic study of causal relationships between genetic differences in circulating metabolitemolecular patterns and risk of diabetes and other health outcomes. These investigations may also explain the association of disease at dozens of T2D associated genetic loci where the underlying molecular mechanisms remain elusive. Ultimately, the hope is that these multi-layered genetic investigations will reveal novel associations and make sense of known ones, but the distance between expectations and achievement in this field remains wide.~~

Formatted: No underline, Font color: Auto, Highlight
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: No underline, Font color: Auto, Not Highlight
Formatted: No underline, Font color: Auto, Not Highlight
Field Code Changed
Field Code Changed

3. Genomics and therapeutics in T2D

The Chief Medical Officer's Report "Generation Genome" published in 2017 highlighted key areas in which genomics can inform disease therapeutics,[89–92] providing a helpful framework to describe recent developments of genomics-informed therapeutics in T2D (**Table 1**).

New targets

The identification of novel drug targets is a major stated objective of T2D genetic studies. The finding that drug targets supported by human genetics evidence are more likely to succeed in the selective drug development process[93-95 90-92] and recent success in the development of lipid-loweringnew lipid-lowering drugs following genetic findings from different approaches and sources at the *PCSK9*, *LPA*, *APOC3* and *ANGPTL3* loci[96, 97 93, 94] have exponentially increased interest in this fieldapproach.

Three diabetes-susceptibility -candidate genes identified in early GWAS studies[1] encode drug targets for existing glucose lowering therapy (*PPARG*, thiazolidinediones; *KCNJ11*/*ABCC8*, sulfonylureas). . Mutations in these genes cause severe monogenic diseases characterized by loss of glycaemic control[95-97] and common variants at these loci were identified in the first wave of diabetes GWAS.[1] This suggestsuggesting that other genes identified through hypothesis-free approaches may provide usefulbecome new drug targetss, even if common lead variants have small effects.
However, human genetics has played a peripheral role in the development of the most recently approved classes of glucose-lowering drugs, including dipeptidyl peptidase-4 inhibitors,[98,98] GLP1R agonists[98,98] and sodium-glucose co-transporter-2 inhibitors.[99,99]

With ~250over a hundred genetic loci now robustly associated with diabetes,[8] , why is there still not a clear "PCSK9-like" example? Generic and therapeutic-area specific obstacles may play a role. First, several of the causal genes implicated by GWAS have only recently emerged, but it takes several years for new drugs to enter clinical development. Specifically targeting pancreatic islets, adipocytes, skeletal myocytes or brain cells, as opposed to circulating proteins is challenging. Safety concerns have been a limitation in the development of appetite suppressants modifying genetically-validated targets, as illustrated by the failure of first generation agonists of the melanocortin 4 receptor due to on-target side effects.[100,100] Also, many loci for T2D act via impaired insulin secretion.[48] The existence of different classes of approved and widely-used insulin secretagogues may limit the interest of pharmaceutical companies in new drug development in this area, given the focus on developing commercially-differentiated products.[94,90]

Protective loss-of-function variants are particularly interesting for drug development purposes because they provide insights into the likely consequences of inhibiting a gene product and, if carriers are healthy, provide someinitial implicit evidence of the likely safety profile of pharmacological inhibition of the existence of a therapeutic window.[93 92] Sequencing of the early T2D GWAS gene *SLC30A8*, encoding a pancreatic islet zinc transporter (ZnT8), has identified rare An interesting example has been the discovery by sequencing of loss-of-function variants in *SLC30A8*, encoding a pancreatic islet zinc transporter (ZnT8), associated with protection against T2D (odds ratio for carriers, 0.34).[101,101] More recently, an exome array genotyping study study in populations of Latino descent identified a protective variant (odds ratio per allele, 0.80) in *IGF2* associated with incorrect splicing of isoform 2 of the gene, suggesting that selectively inhibiting this isoform in relevant tissues may be therapeutically exploited.[20,18] However, efforts to inhibit ZnT8 or insulin-like growth factor 2 have yet to reach clinical development. The potential therapeutic implications of newly discoveredrecently reported protective associations of a loss-of-function variant in *GPR151* against obesity, diabetes and coronary artery disease of a loss-of-function variant in *GPR151* also deserve consideration.[102]

Genetically-tailored treatment

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman, No underline, Font color:
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: No underline, Font color: Auto, Not Highlight
Formatted: No underline, Font color: Auto, Not Highlight
Field Code Changed
Formatted: Font: Italic, No underline, Font color: Auto
Field Code Changed
Field Code Changed
Formatted: Font: Italic, No underline, Font color: Auto
Field Code Changed

Diabetes medicine holds some of the most elegant examples of tailoring treatment to the specific underlying genetic or molecular defect, but all of these relate to monogenic forms of the disease.[103-105] [102 104] No such examples exist for common susceptibility loci, but this area is certainly understudied.

Response to glucose-lowering treatment shows a measurable degree of heritability and evidence of polygenicity,[106] indirectly suggesting that combinations of multiple alleles might be able to identify patients who would be more or less responsive to certain drugs. Opportunities in this field arise from (a) the definition and characterization of polygenic scores combining common variants that capture a particular aetiology, Variants with large effects, such as those ; (b) variants with large effect size from population isolates; (c), or rare variants with very large effects from GWAS in admixed populations. could also provide suitable basis for pharmacogenetics applications. By studying the functional consequences of all theoretically possible missense variants in *PPARG*, Majithia and colleagues elegantly showed that diabetes-associated mutations in the gene display heterogeneous *in vitro* response to thiazolidinediones,[107106, 107] which could provide the basis for tailored therapy or dosing in carriers of these specific alleles, as illustrated in initial case reports.[108]

Drug dosing or response

The several eEfforts to identify interactions between treatment and genetic background and T2Din treatment T2D have been the subject of a recent systematic review,[109108] which revealed the reliance onhowreporting that research in the field is mostly based on observational studies rather than randomized controlled trials and candidate gene rather than hypothesis-free approaches, with a few notable exceptions discussed below. In a pharmacogenetic clinical trial, Srinivasan et al. found that *TCF7L2* variants associated with T2D influences the acute response to both glipizide and metformin in people with risk factors for T2D or treatment-naïve T2D patients.[110] Two Two genome-wide association studies have identified common genetic variants at the *ATM*[111] and [109] , and *SLC2A2*[112110] loci associated with response to metformin. The difference in the effect of metformin effect size ofor these variants on HbA1c was estimated at around ~0.15-0.17% of HbA1c per allele, roughly corresponding to a daily dose of ~250 mg of metformin.[111, 112109, 110110] A new genome-wide studydiscovery embedded into a clinical trial found novel associations for common and rare variants in *PRPF31*, *CPA6*, and *STAT3* with metformin response.[113] HoweverWhile these findings are important for the understanding of genetic susceptibility to drug response, tgenetic testing for these variants has yet to enter the clinic. The very low price of metformin coupled with the and widely adopted pragmatic approach offocus on reaching the HbA1c therapeutic target or the maximum tolerated dose of this drug are barriers to clinical use of thisese genetic tests. genetic testing in this setting. entering the clinic.[112]

Drug repurposing

If "pharmacomimetic" genetic variants can be used to find new targets, they could theoretically be used to find new indications for existing drugs. While there , but there are not yet any established currently no clear examples of genetically-directed repurposing of approved drugs in diabetes. , Imamura et al. used a systematic bioinformatics approach to identify new T2D drug targets, revealing potential repurposing opportunities for drugs targeting the gene products of *GSK3B* and *JUN*.[23] Recent findings around lipoprotein lipase may offer an example of genetically-driven extension of the target population for drugs that are in active development. In late 2016, we reported a gain-of-function variant (rs328, p.Ser447*) in *LPL* associated with insulin sensitivity and protection from diabetes and an independent a loss-of-function variant (rs1801177, p.Asp36Asn) associated with higher diabetes risk. [3335] This followed directionally consistent findings for triglyceride levels and heart disease[114] [111] , supporting leading to the hypothesis that the several agents targeting the LPL pathway that are in development for the treatment of hypertriglyceridemia[94, 112-11597, 115-118] could also be valuable as insulin sensitizing agents. The association with diabetes of rs328 has since been replicated[116119] and

Field Code Changed
Field Code Changed
Formatted: No underline, Font color: Auto, Not Highlight
Formatted: No underline, Font color: Auto, Not Highlight
Formatted: No underline, Font color: Auto, Not Highlight
Formatted: No underline, Font color: Auto, Not Highlight
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: Italic, No underline, Font color: Auto
Field Code Changed
Field Code Changed
Formatted: Font: Not Italic, No underline, Font color: Auto
Field Code Changed
Formatted: Font: (Default) Times New Roman, No underline, Font color: Auto
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: No underline, Font color: Auto, Not Highlight
Formatted: No underline, Font color: Auto, Not Highlight
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

the variant or its proxies have emerged in recent GWAS of diabetes,[7, ~~8~~ ~~7, 8~~] with consistent findings published for a loss-of-function in the natural LPL-inhibitor *ANGPTL4* (rs116843064, p.Glu40Lys),[~~116~~119] lending powerful support to this hypothesis.

Drug safety

Genetic variants have ~~also~~ been used to understand both desired and undesired secondary effects of pharmacological modulation. ~~For instance, statins are associated with cardiovascular protection, but with a modest increase in the risk of new-onset diabetes, a pattern observed also for cholesterol lowering variants in or near~~ Similar to statins and genetic variants at their target *HMGCR*,[~~117, 118~~120, 121] a~~A~~lso ~~genetic studies of non-statin cholesterol lowering targets~~ cholesterol lowering alleles at ~~including~~ *NPC1L1* (encoding the target of ezetimibe)[~~119~~122] and *PCSK9* (PCSK9 inhibitors)[~~119 121~~122-124] ~~have shown associations~~ are associated with a lower risk of coronary heart disease, but higher diabetes risk~~ for cholesterol lowering alleles~~. While t~~T~~his ~~has,~~ suggested~~ing~~ ~~modest diabetogenic effects for~~that also~~that also~~ non-statin cholesterol lowering agents could be associated with higher diabetes risk, ~~might have modest diabetogenic effects. It is important to note that~~However, recent randomized controlled trials of ezetimibe[125~~122~~] and PCSK9-inhibitors[126~~123~~] did not ~~find associations with statistically significant higher diabetes risk, ruling out large~~identify large or statistically-significant diabetogenic effects. ~~A meta-analysis of trials of PCSK9 inhibitors has revealed statistically significant albeit very modest increases in glycemic markers, but no significant association with diabetes risk.~~[~~124~~]

Cardiovascular associations of genetic variants that mimic diabetes medications are particularly insightful, given the regulatory requirement that glucose-lowering drugs should not be associated with a higher risk of cardiovascular disease[125~~127~~] and the ongoing paradigm shift from glycemic control to prevention of complications in diabetes management.[~~126~~128] The concomitant publication of randomized controlled trials of a glucagon-like peptide receptor 1 (GLP1R) agonists[~~127, 128~~129, 130] and of genetic studies of a putative gain-of-function variant of *GLP1R* (rs10305492, p.Ala316Thr),[~~129~~131] both showing cardio-protective associations for GLP1R activation, illustrates this concept. A similar approach has shown cardio-protective associations for a functional variant in *ABCC8* (rs757110, p.Ala1369Ser),[~~130~~ ~~132~~] providing genetic insights into the cardiovascular effects of sulfonylureas for which clinical trial~~s~~ evidence is inconclusive.[~~131~~133]

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: (Default) Times New Roman, No underline, Font color:
Formatted: Font: (Default) Times New Roman, No underline, Font color:
Formatted: Font: (Default) Times New Roman, No underline, Font color:
Field Code Changed

4. Clinical relevance and future outlook

Genomics in the clinic

In addition to diabetes therapeutics, genomics has been proposed to help disease prediction and diagnosis of common subtypes. The polygenic architecture of T2D dominated by many common variants with small effects and tagging several different aetiologies has critical implications for both applications.

The level of prediction achieved by common alleles contrasts with patients' often more deterministic understanding of what "genetic risk" constitutes, which is informed by examples of highly penetrant causes of monogenic diseases. Accurate prediction in combination with preventive lifestyle interventions can be used for targeted primary prevention and avoidance of complications arising from metabolic dysregulation that is present for years before diagnosis.[~~132~~134] However, existing diabetes prediction models perform well in the general population and can be used to non-invasively identify individuals at high risk.[~~133~~135] Earlier prospective studies using up to 65 variants have shown that polygenic T2D risk scores improve prediction performance only ~~very marginally~~modestly when considered over and above risk factors that can assessed non-invasively, such as age, sex, body mass index or family history.[~~134, 135~~136, 137] New approaches based on machine learning and complex statistical modelling ~~such as those~~have been ~~-~~proposed ~~by Shigemizu et~~to ~~-~~as improved methods for genetic prediction.~~have the potential to improve on these results. al.~~[138] ~~could perhaps further improve on these results. In the context of a clinically established "minimally" invasive test,~~ Critically, the ~~existence~~availability of good inexpensive predictors from a patient's anamnesis or examination (e.g. family history, BMI) together with clinically established ~~inexpensive~~ tests (i.e. ~~–~~blood glucose, HbA1c) that inform both ~~that can inform both diagnosis and~~ future risk prediction and diagnosis~~prediction, i.e. blood glucose levels or HbA1c,~~ greatly limits~~s~~ the scope for introduction ~~practical use~~ of ~~genetically~~ genetic information for~~in~~ T~~predicted risk~~2D prediction in the clinic ~~currently has no meaningful clinical contribution to make~~. Evidence exists that established genetic variants have the weakest relative effect and add the least to prediction in people with highest levels of traditional risk factors and hence at highest absolute risk, the exact subgroup of the population in which preventive interventions and the cost-effectiveness of screening would be greatest.~~i.e. blood glucose levels or HbA1c, genetically predicted risk currently has no meaningful clinical contribution to make. Evidence exists that established genetic variants have the weakest relative effect and add the least to prediction in people with highest levels of traditional risk factors and hence at highest absolute risk; the exact subgroup of the population in which preventive interventions and the cost-effectiveness of screening would be greatest.~~

Diabetes is a multifactorial disease and it has been proposed that genetics could help classify common disease subtypes. Apart from gestational diabetes and rare, specific Mendelian forms for which genetics already helps to guide diagnosis and treatment, diabetes is currently crudely classified into two broad types (>~~–~~90% T2D) based on clinical presentation and rapid requirement of insulin.[139~~136~~] Thomas et al. have provided evidence for the presentation of T1D up to the sixth decade of life and that a polygenic score specifically associated with type 1 but not type 2 diabetes can help to rule out T1D in late onset cases.[~~137~~140] But due to the overwhelming predominance of T2D at older ages, the score's positive predictive value is too low to confidently identify late onset T1D patients, in whom initial management may not be optimal if misdiagnosed as T2D.

~~Recent~~ Interesting studies~~efforts to classify~~aiming to identify T2D subgroups ~~on the basis of electronic health records, physiological measures or genetic data~~ using data-driven agnostic approaches ~~are interesting, but their clinical relevance is yet to be demonstrated~~have recently emerged. ~~and perhaps a more nuanced approach to aetiologic classification would better suit the highly polygenic and multifactorial background of this disease.~~[~~138~~]For instance, using six diabetes-related ~~variables~~parameters (glutamate decarboxylase antibodies, age at diagnosis, BMI, HbA1c, and ~~homeostatic model assessment 2 estimates~~indices of insulin secretion and ~~β cell function and insulin~~ resistance), Ahlqvist et al. have provided evidence of five subtypes of T2D that differ in disease

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

trajectories and risk of complications.[141] Using electronic medical records, Li et al. ~~had~~ previously ~~found~~reported evidence of three subtypes,[142] suggesting that data-driven aetiologic c~~a~~lassifications are influenced by context and data availability. Investigations using "-omics" measurements may further extend these initial attempts at a more refined disease categorisation. While the existance of different aetiologic subtypes in T2D is widely accepted, a robust and definitive classification is missing. In contrast with approaches aimed at classifying T2D in subgroups and categories, it has been proposed that a more nuanced approach to aetiologic classification would better suit the highly polygenic and multifactorial background of this disease.[143]

~~Required at diagosis~~

Genomic medicine and academia-industry partnerships

In the UK, transformative sequencing projects are currently underway. The 100,000 Genome Project delivers the benefits of genomic medicine to NHS cancer and rare disease patients now and has so far sequenced over ~~50~~55,000 whole genomes (https://www.genomicsengland.co.uk/the-100000-genomes-project/). This project has shown the need for economies of scale in the delivery of genomic medicine, and NHS England is recommissioning and modernising NHS Genomic Laboratory Services to develop a first-class genomic service. UK Biobank (http://www.ukbiobank.ac.uk/) has transformed opportunities for population research in this country and internationally and is entering a new phase with the announcement to exome-sequence all 500,000 participants by 2019. Such an accelerated timeline was only possible through major funding by a consortium of five pharmaceutical companies, brought together by Regeneron Pharmaceuticals, following from their first initiative to sequence 50,000 participants in collaboration with GSK. In the USA, a partnership between Regeneron and the Geisinger Health System for the exome sequencing of over 50,000 people has already shown the value of integrating genetic ~~data~~ and electronic health record data~~s~~ at ~~a~~ ~~large~~ scale.[144, 145] ~~This~~ These ~~is~~ are only ~~one~~ two examples of commercial partners having access to participant and patient data, including information collected as part of routine clinical care in electronic health records. The financial benefits of such an effort are something to be considered in the light of the scientific opportunities that this investment and the generated sequence information will provide to biomedical researchers worldwide when they gain access, as will be the case for UK Biobank ~~researchers~~.

~~.~~

At the same time, large-scale Biobanks collecting genetic, physiological, longitudinal electronic health records and other health data have been ~~are being~~ established in many countries around the word, focusing on clinical hospital populations (e.g. BioVU: https://www.vumc.org/dbmi/biovu), insurance or care provider populations (e.g. Million Veteran Program: https://www.research.va.gov/mvp/), or national population cohorts (e.g. China Kadoorie Biobank, German National Cohort etc). The NIH funded 1 million persons "All of Us" Research Program was~~,~~ originally launched as a "new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes".~~,~~[146] It specifically aims to actively recruit ethnic minority populations to help address the existing bias by including ~~. Whether or not it will be successful in accruing~~ sufficient numbers ~~required for genetic research~~ of some of the many currently underrepresented groups ~~remains to be established~~.

Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: Not Italic, No underline, Font color: Auto
Formatted: Font: Not Italic
Formatted: Justified, Tab stops: 0.75 cm, Left
Formatted: Font color: Auto
Formatted: Font color: Auto
Field Code Changed
Formatted: Justified, Tab stops: 0.75 cm, Left
Formatted: Font color: Auto
Formatted: Font color: Auto
Field Code Changed

Engaging patients in research

Genomic sequencing puts the patient at the centre of drug discovery and validation. Close collaboration between academia, the pharmaceutical and other industries can catalyse the development of novel therapies for T2D based on genomic insight. To make full use of these opportunities for patients in this country and elsewhere, the research community needs to engage in a dialogue with patients and the public about genomic medicine and research and their implications for uses and misuses of genomic data.

Notwithstanding the potential that such partnerships have to improve health research and outcomes, research commissioned by the Wellcome Trust (https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf) and work delivered by Genomics England as part of their "Genomics Conversation" (https://www.genomicsengland.co.uk/a-year-of-conversations-about-genomics) have highlighted that patients' and the public have concerns about data safety, i.e. non-legitimate uses, and commercial access to health data, including pharmaceutical and insurance companies.

The Chief Medical Officer's Report "Generation Genome" considered the ethical, social and legal implications of genomic medicine in this country[92~~89~~] and highlighted the need for highest levels of data security for storage of identifiable data. However, it is not possible to give an absolute guarantee of data security and the potential harms arising from criminal data breaches need to be put in proportion with harms arising from restricting legitimate research uses of health data. For people living in the UK, universal free access to the National Health ~~System~~ Service means that there is less reason to fear discrimination with regards health care insurance on the grounds of genetic testing. While there is currently no explicit legislation, the existing voluntary agreement with insurance providers also means that an estimated 95% of insurance customers would not need to disclose genetic test results for example for life assurance, critical illness cover, or income protection, as disclosure is tied to the policy value. The House of Commons Select Committee on Science and Technology recently recommended to extend~~hat~~ the existing voluntary agreement, but ~~is extended~~closely monitor ~~but that~~ patient's views and the experiences in other countries with a legal prohibition ~~are closely monitored~~ (https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/349/34908.htm#_idTextAnchor041).

Conclusions

Advances in genomic research have facilitated rapid progress in clarifying the genetic basis of T2D and characterising causal variants and variant-gene links. Future opportunities lie in larger-scale sequencing, discovery across diverse ancestries, studies in genetically isolated populations and in massive-scale biobanks. Successful development of academia-industry partnerships ~~has brought~~ can deliver economies of scale, with implications for the future of genomics-focused research.~~informed drug development and population research in this field.~~

Review Methods

We searched Pubmed from inception to March 1st 2018 using the following search strategy: (Diabetes Mellitus, Type 2[MeSH] OR NIDDM OR Maturity-Onset Diabetes OR Diabetes Mellitus, Noninsulin-Dependent OR Diabetes Mellitus, Adult-Onset OR Adult-Onset Diabetes Mellitus OR Diabetes Mellitus, Adult Onset OR Diabetes Mellitus, Ketosis-Resistant OR Diabetes Mellitus, Ketosis Resistant OR Ketosis-Resistant Diabetes Mellitus OR Diabetes Mellitus, Maturity-Onset OR Diabetes Mellitus, Maturity Onset OR Diabetes Mellitus, Non Insulin Dependent OR Diabetes Mellitus, Non-Insulin-Dependent OR Non-Insulin-Dependent Diabetes Mellitus OR Diabetes Mellitus, Noninsulin Dependent OR Diabetes Mellitus, Slow-Onset OR Diabetes Mellitus, Slow Onset OR Slow-Onset Diabetes Mellitus OR Diabetes Mellitus, Stable OR Stable Diabetes Mellitus OR Diabetes Mellitus, Type II OR Maturity-Onset Diabetes Mellitus OR Maturity Onset Diabetes Mellitus OR MODY OR Type 2 Diabetes Mellitus OR Noninsulin-Dependent Diabetes Mellitus OR T2D OR T2DM OR Type 2 Diabetes[tiab] OR Type 2 diabetes mellitus OR diabetes[ti]) AND (Genome-Wide Association Study[MeSH] OR Association Studies, Genome-Wide OR Association Study, Genome-Wide OR Genome-Wide Association Studies OR Studies, Genome-Wide Association OR Study, Genome-Wide Association OR Genome Wide Association Scan OR Genome Wide Association Studies OR GWA Study OR GWA Studies OR Studies, GWA OR Study, GWA OR Whole Genome Association Analysis OR Whole Genome Association Study OR Genome Wide Association Analysis OR Genome Wide Association Study). To look for new studies published in pre-publication (non peer-reviewed) form, we searched BioRxiv using the advanced search function (https://www.biorxiv.org/search): articles posted in the "genetics" or "genomics" collections, with the key word "diabetes" in the title or abstract, posted between the 1st of January 2017 and the 9th of April 2018. Thesesis literature searches ~~was~~ were integrated with reference files of the authors and their colleagues, reference lists of original articles, reviews, and meta-analyses. Given exhaustive reviews on early genetic association studies by McCarthy[1] and Morris[2], amongst others, we focused on recent developments and articles providing insights into clinical translation of genetic findings.

Formatted: Default Paragraph Font, Font: (Default) +Body, 12 pt, English (United States), Not Expanded by / Condensed by

Formatted: No underline, Font color: Auto, Superscript

Formatted: No underline, Font color: Auto, Superscript

Tables

Table 1. Contribution of genetic findings to T2D therapeutics in key areas.

Area of contribution	Rationale	Considerations and examples in T2D research
New drug target identification	In retrospective analyses, drugs with human genetics support are more likely to successfully transition through the drug development pipeline. Rapid development of new lipid-lowering drugs with genetic validation illustrates potential. Loss-of-function variants provide insights into likely efficacy and safety of inhibition, while gain-of-function on stimulation of target.	Genes encoding the targets of glucose lowering agents have been found by early GWAS, but new classes of diabetes drugs have not been developed as a result of human genetics findings. Most loci identified by GWAS have not lead to new drug development. Protective loss-of-function variants in *SLC30A8* and *IGF2* provide interesting examples that still await new drug development.
Mutation specific treatment	Pharmacological interventions may be particularly effective in patients with particular underlying aetiology or genetic predisposition.	Prominent examples are from Mendelian genetics, lack of examples for common forms of diabetes with polygenic aetiologic contribution. Elegant exemplar from systematic study of all possible missense variants of *PPARG*. Opportunities in specific areas have not been fully exploited.
Drug dosing or response	Drugs may require dose-adjustment according to genetic background.	Common variants at the *ATM* and *SLC2A2* loci have been robustly associated with response to metformin, but genetic testing is not used in the clinic. Several studies including a recent trial have proposed an effect of *TCF7L2* variants on response to glucose-lowering drugs.[108, 10]
Drug repurposing	As with new target identification, genetic variants that "mimic" existing therapeutic agents may provide the basis for repurposing.	No established example of repurposing from other therapeutic areas to diabetes. Recent GWAS have explored repurposing opportunities using bioinformatics approaches. Genetic findings around the *LPL* pathway may provide basis for extension of future indications and target population for emerging ~~lipid~~ ~~lowering~~ drugs targeting this pathway.
Drug safety	Genetic variants can inform on desired and undesired secondary effects of pharmacological modulation of the encoded drug target. In diabetes, it is critical to study the cardiovascular safety of existing and new agents.	The example of low-density lipoprotein cholesterol lowering genetic variants in genes encoding targets of cholesterol lowering therapy (i.e. *HMGCR*, *NPC1L1* and *PCSK9*) and diabetes risk illustrate power and challenges of genetic approaches, given the partial consistency between genetic and clinical trial results. *GLP1R* and *ABCC8* variants have been used to gain insights into cardiovascular safety of existing glucose-lowering drugs.

Abbreviations: GWAS, genome-wide association studies.

Formatted: Font: Italic, No underline, Font color: Auto

Field Code Changed

Formatted: No underline, Font color: Auto

Field Code Changed

Figure Legends

Figure 1. Illustrative representation of genome-wide studies in type 2 diabetes and their power to detect certain types of susceptibility alleles for a given sample size. Susceptibility alleles above the solid black lines are detectable with a given approach. The graph is informed by the results of actual historical[2] and current[8] GWAS studies as well as whole-genome and exome sequencing studies that provided an empirical model for the genetic architecture of type 2 diabetes.[9] Exemplar genetic susceptibility loci are reported in the figure. Abbreviations: GWAS, genome-wide association studies; OR, odds ratio.

Figure 2. Aetiologic model for the role of TBC1D4 in GLUT4 translocation and insulin-mediated glucose uptake in the skeletal muscle.

Figure 3. Models for normal and impaired insulin secretion. Genetic variants affecting these processes result in impaired insulin secretion and higher diabetes risk (right panel), e.g. variants at *KCNJ11* and *ABCC8* identified in genome-wide association studies.

Figure ~~2~~4. Aetiologic model for the contribution of peripheral adipose storage capacity to metabolic and cardiovascular disease and role of adipogenesis and intravascular lipolysis in this process. Some of the images have been samples and modified from SMART, Servier Medical Art, (URL: https://smart.servier.com/) under Creative Commons Attribution 3.0 Unported License (https://creativecommons.org/licenses/by/3.0/).

Formatted: Font: (Default) Times New Roman

Formatted: Space Before: 0 pt, After: 0 pt

Formatted: Space Before: 0 pt, After: 0 pt

Formatted: Font: Not Bold, No underline, Font color: Auto

Formatted: Font: Not Bold, No underline, Font color: Auto

Formatted: Font: Not Bold, Italic, No underline, Font color: Auto

Formatted: Font: Not Bold, No underline, Font color: Auto

Formatted: Font: Not Bold, Italic, No underline, Font color: Auto

Formatted: Font: Not Bold, No underline, Font color: Auto

Formatted: Font: Not Bold

References

1. McCarthy MI. Genomics, type 2 diabetes, and obesity. *The New England journal of medicine* 2010; **363**(24): 2339-50.
2. Morris AP. Progress in defining the genetic contribution to type 2 diabetes susceptibility. *Current opinion in genetics & development* 2018; **50**: 41-51.
3. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics* 2012; **44**(9): 981-90.
4. DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature genetics* 2014; **46**(3): 234-44.
5. Scott RA, Scott LJ, Magi R, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* 2017; **66**(11): 2888-902.
6. Zhao W, Rasheed A, Tikkanen E, et al. Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nature genetics* 2017; **49**(10): 1450-7.
7. Mahajan A, Wessel J, Willems SM, et al. Refining The Accuracy Of Validated Target Identification Through Coding Variant Fine-Mapping In Type 2 Diabetes. *bioRxiv* 2017.
8. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping of an expanded set of type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *bioRxiv* 2018.
9. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes. *Nature* 2016; **536**(7614): 41-7.
10. Steinthorsdottir V, Thorleifsson G, Sulem P, et al. Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nature genetics* 2014; **46**(3): 294-8.
11. Pasquali L, Gaulton KJ, Rodriguez-Segui SA, et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nature genetics* 2014; **46**(2): 136-43.
12. Gaulton KJ, Ferreira T, Lee Y, et al. Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nature genetics* 2015; **47**(12): 1415-25.
13. Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature genetics* 2012; **44**(4): 369-75, S1-3.
14. Wellcome Trust Case Control Consortium, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nature genetics* 2012; **44**(12): 1294-301.
15. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature* 2016; **538**(7624): 161-4.
16. SIGMA Type 2 Diabetes Consortium, et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* 2014; **506**(7486): 97-101.
17. SIGMA Type 2 Diabetes Consortium, et al. Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. *JAMA : the journal of the American Medical Association* 2014; **311**(22): 2305-14.
18. Mercader JM, Liao RG, Bell AD, et al. A Loss-of-Function Splice Acceptor Variant in IGF2 Is Protective for Type 2 Diabetes. *Diabetes* 2017; **66**(11): 2903-14.
19. Cho YS, Chen CH, Hu C, et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nature genetics* 2011; **44**(1): 67-72.
20. Kooner JS, Saleheen D, Sim X, et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nature genetics* 2011; **43**(10): 984-9.
21. Imamura M, Takahashi A, Yamauchi T, et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. *Nature communications* 2016; **7**: 10531.

22. Parra EJ, Below JE, Krithika S, et al. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. *Diabetologia* 2011; **54**(8): 2038-46.
23. Ng MC, Shriner D, Chen BH, et al. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS genetics* 2014; **10**(8): e1004517.
24. Andersen MK, Pedersen CE, Moltke I, Hansen T, Albrechtsen A, Grarup N. Genetics of Type 2 Diabetes: the Power of Isolated Populations. *Current diabetes reports* 2016; **16**(7): 65.
25. Moltke I, Grarup N, Jorgensen ME, et al. A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. *Nature* 2014; **512**(7513): 190-3.
26. Dash S, Sano H, Rochford JJ, et al. A truncation mutation in TBC1D4 in a family with acanthosis nigricans and postprandial hyperinsulinemia. *Proceedings of the National Academy of Sciences of the United States of America* 2009; **106**(23): 9350-5.
27. Grarup N, Moltke I, Andersen MK, et al. Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. *Nature genetics* 2018; **50**(2): 172-4.
28. Saeed S, Bonnefond A, Tamanini F, et al. Loss-of-function mutations in ADCY3 cause monogenic severe obesity. *Nature genetics* 2018; **50**(2): 175-9.
29. George S, Rochford JJ, Wolfrum C, et al. A family with severe insulin resistance and diabetes due to a mutation in AKT2. *Science* 2004; **304**(5675): 1325-8.
30. Manning A, Highland HM, Gasser J, et al. A Low-Frequency Inactivating AKT2 Variant Enriched in the Finnish Population Is Associated With Fasting Insulin Levels and Type 2 Diabetes Risk. *Diabetes* 2017; **66**(7): 2019-32.
31. Latva-Rasku A, Honka MJ, Stancakova A, et al. A Partial Loss-of-Function Variant in AKT2 Is Associated With Reduced Insulin-Mediated Glucose Uptake in Multiple Insulin-Sensitive Tissues: A Genotype-Based Callback Positron Emission Tomography Study. *Diabetes* 2018; **67**(2): 334-42.
32. Minster RL, Hawley NL, Su CT, et al. A thrifty variant in CREBRF strongly influences body mass index in Samoans. *Nature genetics* 2016; **48**(9): 1049-54.
33. Lotta LA, Gulati P, Day FR, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nature genetics* 2017; **49**(1): 17-26.
34. Prokopenko I, Langenberg C, Florez JC, et al. Variants in MTNR1B influence fasting glucose levels. *Nature genetics* 2009; **41**(1): 77-81.
35. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature genetics* 2010; **42**(2): 105-16.
36. Saxena R, Hivert MF, Langenberg C, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nature genetics* 2010; **42**(2): 142-8.
37. Soranzo N, Sanna S, Wheeler E, et al. Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes* 2010; **59**(12): 3229-39.
38. Scott RA, Lagou V, Welch RP, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nature genetics* 2012; **44**(9): 991-1005.
39. Manning AK, Hivert MF, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nature genetics* 2012; **44**(6): 659-69.
40. Wheeler E, Leong A, Liu CT, et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS medicine* 2017; **14**(9): e1002383.
41. Ingelsson E, Langenberg C, Hivert MF, et al. Detailed physiologic characterization reveals diverse mechanisms for novel genetic Loci regulating glucose and insulin metabolism in humans. *Diabetes* 2010; **59**(5): 1266-75.

42. Strawbridge RJ, Dupuis J, Prokopenko I, et al. Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. *Diabetes* 2011; **60**(10): 2624-34.
43. Prokopenko I, Poon W, Magi R, et al. A central role for GRB10 in regulation of islet function in man. *PLoS genetics* 2014; **10**(4): e1004235.
44. Scott RA, Fall T, Pasko D, et al. Common genetic variants highlight the role of insulin resistance and body fat distribution in type 2 diabetes, independent of obesity. *Diabetes* 2014; **63**(12): 4378-87.
45. Walford GA, Gustafsson S, Rybin D, et al. Genome-Wide Association Study of the Modified Stumvoll Insulin Sensitivity Index Identifies BCL2 and FAM19A2 as Novel Insulin Sensitivity Loci. *Diabetes* 2016; **65**(10): 3200-11.
46. Knowles JW, Xie W, Zhang Z, et al. Identification and validation of N-acetyltransferase 2 as an insulin sensitivity gene. *The Journal of clinical investigation* 2016; **126**(1): 403.
47. Wood AR, Jonsson A, Jackson AU, et al. A Genome-Wide Association Study of IVGTT-Based Measures of First-Phase Insulin Secretion Refines the Underlying Physiology of Type 2 Diabetes Variants. *Diabetes* 2017; **66**(8): 2296-309.
48. Dimas AS, Lagou V, Barker A, et al. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* 2014; **63**(6): 2158-71.
49. Lyssenko V, Lupi R, Marchetti P, et al. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *The Journal of clinical investigation* 2007; **117**(8): 2155-63.
50. Moran I, Akerman I, van de Bunt M, et al. Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell metabolism* 2012; **16**(4): 435-48.
51. van de Bunt M, Manning Fox JE, Dai X, et al. Transcript Expression Data from Human Islets Links Regulatory Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic Traits to Their Downstream Effectors. *PLoS genetics* 2015; **11**(12): e1005694.
52. van de Bunt M, Lako M, Barrett A, et al. Insights into islet development and biology through characterization of a human iPSC-derived endocrine pancreas model. *Islets* 2016; **8**(3): 83-95.
53. van de Bunt M, Gaulton KJ, Parts L, et al. The miRNA profile of human pancreatic islets and beta-cells and relationship to type 2 diabetes pathogenesis. *PloS one* 2013; **8**(1): e55272.
54. Thomsen SK, Ceroni A, van de Bunt M, et al. Systematic Functional Characterization of Candidate Causal Genes for Type 2 Diabetes Risk Variants. *Diabetes* 2016; **65**(12): 3805-11.
55. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**(7538): 197-206.
56. Akiyama M, Okada Y, Kanai M, et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nature genetics* 2017; **49**(10): 1458-67.
57. Turcot V, Lu Y, Highland HM, et al. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nature genetics* 2018; **50**(1): 26-41.
58. O'Rahilly S. Harveian Oration 2016: Some observations on the causes and consequences of obesity. *Clinical medicine* 2016; **16**(6): 551-64.
59. Yusuf S, Hawken S, Ounpuu S, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. *Lancet* 2005; **366**(9497): 1640-9.
60. InterAct Consortium, Langenberg C, Sharp SJ, et al. Long-term risk of incident type 2 diabetes and measures of overall and regional obesity: the EPIC-InterAct case-cohort study. *PLoS medicine* 2012; **9**(6): e1001230.
61. Stefan N, Häring HU, Hu FB, Schulze MB. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. *Lancet Diabetes Endocrinol* 2013; **1**(2): 10.
62. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015; **518**(7538): 187-96.

63. Emdin CA, Khera AV, Natarajan P, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. *JAMA : the journal of the American Medical Association* 2017; **317**(6): 626-34.
64. Garg A. Acquired and inherited lipodystrophies. *The New England journal of medicine* 2004; **350**(12): 1220-34.
65. Semple RK, Savage DB, Cochran EK, Gorden P, O'Rahilly S. Genetic syndromes of severe insulin resistance. *Endocrine reviews* 2011; **32**(4): 498-514.
66. Danforth E, Jr. Failure of adipocyte differentiation causes type II diabetes mellitus? *Nature genetics* 2000; **26**(1): 13.
67. Virtue S, Vidal-Puig A. Adipose tissue expandability, lipotoxicity and the Metabolic Syndrome--an allostatic perspective. *Biochimica et biophysica acta* 2010; **1801**(3): 338-49.
68. Pinnick KE, Nicholson G, Manolopoulos KN, et al. Distinct developmental profile of lower-body adipose tissue defines resistance against obesity-associated metabolic complications. *Diabetes* 2014; **63**(11): 3785-97.
69. Kilpelainen TO, Zillikens MC, Stancakova A, et al. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nature genetics* 2011; **43**(8): 753-60.
70. Small KS, Hedman AK, Grundberg E, et al. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nature genetics* 2011; **43**(6): 561-4.
71. Yaghootkar H, Scott RA, White CC, et al. Genetic evidence for a normal-weight "metabolically obese" phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes. *Diabetes* 2014; **63**(12): 4369-77.
72. Portela A, Esteller M. Epigenetic modifications and human disease. *Nature biotechnology* 2010; **28**(10): 1057-68.
73. GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; **348**(6235): 648-60.
74. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nature communications* 2017; **8**: 14357.
75. Sun BB, Maranville JC, Peters JE, et al. Consequences Of Natural Perturbations In The Human Plasma Proteome. *bioRxiv* 2017.
76. Suhre K, Shin SY, Petersen AK, et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 2011; **477**(7362): 54-60.
77. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. *Nature genetics* 2014; **46**(6): 543-50.
78. Long T, Hicks M, Yu HC, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nature genetics* 2017; **49**(4): 568-78.
79. Gamazon ER, Wheeler HE, Shah KP, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nature genetics* 2015; **47**(9): 1091-8.
80. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nature genetics* 2016; **48**(3): 245-52.
81. Hore V, Vinuela A, Buil A, et al. Tensor decomposition for multiple-tissue gene expression experiments. *Nature genetics* 2016; **48**(9): 1094-100.
82. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature genetics* 2016; **48**(5): 481-7.
83. Barbeira AN, Dickinson SP, Torres JM, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *bioRxiv* 2017.
84. Lotta LA, Scott RA, Sharp SJ, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. *PLoS medicine* 2016; **13**(11): e1002179.
85. Felig P, Marliss E, Cahill GF, Jr. Plasma amino acid levels and insulin secretion in obesity. *The New England journal of medicine* 1969; **281**(15): 811-6.

86. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nature medicine* 2011; **17**(4): 448-53.
87. Newgard CB, An J, Bain JR, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell metabolism* 2009; **9**(4): 311-26.
88. Langenberg C. Integrating Omics with Clinical Outcomes. *2017 Genomics of Common Diseases Meeting* 2017. September 2017, Cambridge, United Kingdom.
89. Annual report of the Chief Medical Officer 2016: Generation Genome. 2017. URL: https://www.gov.uk/government/publications/chief-medical-officer-annual-report-2016-generation-genome. Accessed March 5th 2018.
90. Cook D, Brown D, Alexander R, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nature reviews Drug discovery* 2014; **13**(6): 419-31.
91. Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nature genetics* 2015; **47**(8): 856-60.
92. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. *Nature reviews Drug discovery* 2013; **12**(8): 581-94.
93. Khera AV, Kathiresan S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nature reviews Genetics* 2017; **18**(6): 331-44.
94. Dewey FE, Gusarova V, Dunbar RL, et al. Genetic and Pharmacologic Inactivation of ANGPTL3 and Cardiovascular Disease. *The New England journal of medicine* 2017; **377**(3): 211-21.
95. Barroso I, Gurnell M, Crowley VE, et al. Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature* 1999; **402**(6764): 880-3.
96. Bowman P, Flanagan SE, Edghill EL, et al. Heterozygous ABCC8 mutations are a cause of MODY. *Diabetologia* 2012; **55**(1): 123-7.
97. Gloyn AL, Pearson ER, Antcliff JF, et al. Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. *The New England journal of medicine* 2004; **350**(18): 1838-49.
98. Drucker DJ, Nauck MA. The incretin system: glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors in type 2 diabetes. *Lancet* 2006; **368**(9548): 1696-705.
99. Gallo LA, Wright EM, Vallon V. Probing SGLT2 as a therapeutic target for diabetes: basic physiology and consequences. *Diabetes & vascular disease research* 2015; **12**(2): 78-89.
100. Greenfield JR, Miller JW, Keogh JM, et al. Modulation of blood pressure by central melanocortinergic pathways. *The New England journal of medicine* 2009; **360**(1): 44-52.
101. Flannick J, Thorleifsson G, Beer NL, et al. Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nature genetics* 2014; **46**(4): 357-63.
102. Pearson ER, Flechtner I, Njolstad PR, et al. Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations. *The New England journal of medicine* 2006; **355**(5): 467-77.
103. McCarthy MI, Hattersley AT. Learning from molecular genetics: novel insights arising from the definition of genes for monogenic and type 2 diabetes. *Diabetes* 2008; **57**(11): 2889-98.
104. Oral EA, Simha V, Ruiz E, et al. Leptin-replacement therapy for lipodystrophy. *The New England journal of medicine* 2002; **346**(8): 570-8.
105. Zhou K, Donnelly L, Yang J, et al. Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. *Lancet Diabetes Endocrinol* 2014; **2**(6): 481-7.
106. Majithia AR, Tsuda B, Agostini M, et al. Prospective functional classification of all possible missense variants in PPARG. *Nature genetics* 2016; **48**(12): 1570-5.
107. Majithia AR, Flannick J, Shahinian P, et al. Rare variants in PPARG with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. *Proceedings of the National Academy of Sciences of the United States of America* 2014; **111**(36): 13127-32.

108. Maruthur NM, Gribble MO, Bennett WL, et al. The pharmacogenetics of type 2 diabetes: a systematic review. *Diabetes care* 2014; **37**(3): 876-86.
109. GoDARTS and UKPDS Diabetes Pharmacogenetics Study Group, et al. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nature genetics* 2011; **43**(2): 117-20.
110. Zhou K, Yee SW, Seiser EL, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nature genetics* 2016; **48**(9): 1055-9.
111. Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *The New England journal of medicine* 2016.
112. Larsson M, Caraballo R, Ericsson M, et al. Identification of a small molecule that stabilizes lipoprotein lipase in vitro and lowers triglycerides in vivo. *Biochemical and biophysical research communications* 2014; **450**(2): 1063-9.
113. Geldenhuys WJ, Aring D, Sadana P. A novel Lipoprotein lipase (LPL) agonist rescues the enzyme from inhibition by angiopoietin-like 4 (ANGPTL4). *Bioorganic & medicinal chemistry letters* 2014; **24**(9): 2163-7.
114. Gaudet D, Alexander VJ, Baker BF, et al. Antisense Inhibition of Apolipoprotein C-III in Patients with Hypertriglyceridemia. *The New England journal of medicine* 2015; **373**(5): 438-47.
115. Graham MJ, Lee RG, Brandt TA, et al. Cardiovascular and Metabolic Effects of ANGPTL3 Antisense Oligonucleotides. *The New England journal of medicine* 2017; **377**(3): 222-32.
116. Liu DJ, Peloso GM, Yu H, et al. Exome-wide association study of plasma lipids in >300,000 individuals. *Nature genetics* 2017.
117. Sattar N, Preiss D, Murray HM, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet* 2010; **375**(9716): 735-42.
118. Swerdlow DI, Preiss D, Kuchenbaecker KB, et al. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2015; **385**(9965): 351-61.
119. Lotta LA, Sharp SJ, Burgess S, et al. Association Between Low-Density Lipoprotein Cholesterol-Lowering Genetic Variants and Risk of Type 2 Diabetes: A Meta-analysis. *JAMA : the journal of the American Medical Association* 2016; **316**(13): 1383-91.
120. Schmidt AF, Swerdlow DI, Holmes MV, et al. PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol* 2017; **5**(2): 97-105.
121. Ference BA, Robinson JG, Brook RD, et al. Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. *The New England journal of medicine* 2016; **375**(22): 2144-53.
122. Blazing MAobotI Ii. New-onset Diabetes Mellitus in the IMPROVE-IT Trial. *European Society of Cardiology Congress 2015* 2015.
123. Sabatine MS, Leiter LA, Wiviott SD, et al. Cardiovascular safety and efficacy of the PCSK9 inhibitor evolocumab in patients with and without diabetes and the effect of evolocumab on glycaemia and risk of new-onset diabetes: a prespecified analysis of the FOURIER randomised controlled trial. *Lancet Diabetes Endocrinol* 2017.
124. de Carvalho LSF, Campos AM, Sposito AC. Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Inhibitors and Incident Type 2 Diabetes: A Systematic Review and Meta-analysis With Over 96,000 Patient-Years. *Diabetes care* 2018; **41**(2): 364-7.
125. FDA. Guidance for Industry. Diabetes Mellitus — Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes 2008.
126. Lipska KJ, Krumholz HM. Is Hemoglobin A1c the Right Outcome for Studies of Diabetes? *JAMA : the journal of the American Medical Association* 2017; **317**(10): 1017-8.
127. Marso SP, Bain SC, Consoli A, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *The New England journal of medicine* 2016; **375**(19): 1834-44.
128. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *The New England journal of medicine* 2016; **375**(4): 311-22.

129. Scott RA, Freitag DF, Li L, et al. A genomic approach to therapeutic target validation identifies a glucose-lowering GLP1R variant protective for coronary heart disease. *Science translational medicine* 2016; **8**(341): 341ra76.
130. Emdin CA, Klarin D, Natarajan P, et al. Genetic Variation at the Sulfonylurea Receptor, Type 2 Diabetes, and Coronary Heart Disease. *Diabetes* 2017; **66**(8): 2310-5.
131. Palmer SC, Mavridis D, Nicolucci A, et al. Comparison of Clinical Outcomes and Adverse Events Associated With Glucose-Lowering Drugs in Patients With Type 2 Diabetes: A Meta-analysis. *JAMA : the journal of the American Medical Association* 2016; **316**(3): 313-24.
132. Tabak AG, Jokela M, Akbaraly TN, Brunner EJ, Kivimaki M, Witte DR. Trajectories of glycaemia, insulin sensitivity, and insulin secretion before diagnosis of type 2 diabetes: an analysis from the Whitehall II study. *Lancet* 2009; **373**(9682): 2215-21.
133. Kengne AP, Beulens JW, Peelen LM, et al. Non-invasive risk scores for prediction of type 2 diabetes (EPIC-InterAct): a validation of existing models. *Lancet Diabetes Endocrinol* 2014; **2**(1): 19-29.
134. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *The New England journal of medicine* 2008; **359**(21): 2208-19.
135. Talmud PJ, Cooper JA, Morris RW, et al. Sixty-five common genetic variants and prediction of type 2 diabetes. *Diabetes* 2015; **64**(5): 1830-40.
136. Farmer A, Fox R. Diagnosis, classification, and treatment of diabetes. *Bmj* 2011; **342**: d3319.
137. Thomas NJ, Jones SE, Weedon MN, Shields BM, Oram RA, Hattersley AT. Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *Lancet Diabetes Endocrinol* 2018; **6**(2): 122-9.
138. McCarthy MI. Painting a new picture of personalised medicine for diabetes. *Diabetologia* 2017; **60**(5): 793-9.

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman, No underline, Font color:

1. McCarthy MI. Genomics, type 2 diabetes, and obesity. *The New England journal of medicine* 2010; **363**(24): 2339-50.
2. Morris AP. Progress in defining the genetic contribution to type 2 diabetes susceptibility. *Current opinion in genetics & development* 2018; **50**: 41-51.
3. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics* 2012; **44**(9): 981-90.
4. Replication DIG, Meta-analysis C, Asian Genetic Epidemiology Network Type 2 Diabetes C, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature genetics* 2014; **46**(3): 234-44.
5. Scott RA, Scott LJ, Magi R, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. *Diabetes* 2017; **66**(11): 2888-902.
6. Zhao W, Rasheed A, Tikkanen E, et al. Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nature genetics* 2017; **49**(10): 1450-7.
7. Mahajan A, Wessel J, Willems SM, et al. Refining The Accuracy Of Validated Target Identification Through Coding Variant Fine-Mapping In Type 2 Diabetes. *bioRxiv* 2017.
8. Mahajan A, Taliun D, Thurner M, et al. Fine-mapping of an expanded set of type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *bioRxiv* 2018.
9. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics* 2016; **48**(10): 1279-83.
10. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes. *Nature* 2016; **536**(7614): 41-7.

11. Genomes Project C, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**(7422): 56-65.
12. Steinthorsdottir V, Thorleifsson G, Sulem P, et al. Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nature genetics* 2014; **46**(3): 294-8.
13. Pasquali L, Gaulton KJ, Rodriguez-Segui SA, et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nature genetics* 2014; **46**(2): 136-43.
14. Gaulton KJ, Ferreira T, Lee Y, et al. Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nature genetics* 2015; **47**(12): 1415-25.
15. Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature genetics* 2012; **44**(4): 369-75, S1-3.
16. Wellcome Trust Case Control C, Maller JB, McVean G, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nature genetics* 2012; **44**(12): 1294-301.
17. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature* 2016; **538**(7624): 161-4.
18. Consortium STD, Estrada K, Aukrust I, et al. Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. *JAMA : the journal of the American Medical Association* 2014; **311**(22): 2305-14.
19. Consortium STD, Williams AL, Jacobs SB, et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* 2014; **506**(7486): 97-101.
20. Mercader JM, Liao RG, Bell AD, et al. A Loss-of-Function Splice Acceptor Variant in IGF2 Is Protective for Type 2 Diabetes. *Diabetes* 2017; **66**(11): 2903-14.
21. Cho YS, Chen CH, Hu C, et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nature genetics* 2011; **44**(1): 67-72.
22. Kooner JS, Saleheen D, Sim X, et al. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nature genetics* 2011; **43**(10): 984-9.
23. Imamura M, Takahashi A, Yamauchi T, et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. *Nature communications* 2016; **7**: 10531.
24. Parra EJ, Below JE, Krithika S, et al. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. *Diabetologia* 2011; **54**(8): 2038-46.
25. Ng MC, Shriner D, Chen BH, et al. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS genetics* 2014; **10**(8): e1004517.
26. Andersen MK, Pedersen CE, Moltke I, Hansen T, Albrechtsen A, Grarup N. Genetics of Type 2 Diabetes: the Power of Isolated Populations. *Current diabetes reports* 2016; **16**(7): 65.
27. Moltke I, Grarup N, Jorgensen ME, et al. A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. *Nature* 2014; **512**(7513): 190-3.
28. Dash S, Sano H, Rochford JJ, et al. A truncation mutation in TBC1D4 in a family with acanthosis nigricans and postprandial hyperinsulinemia. *Proceedings of the National Academy of Sciences of the United States of America* 2009; **106**(23): 9350-5.
29. Grarup N, Moltke I, Andersen MK, et al. Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. *Nature genetics* 2018; **50**(2): 172-4.
30. Saeed S, Bonnefond A, Tamanini F, et al. Loss-of-function mutations in ADCY3 cause monogenic severe obesity. *Nature genetics* 2018; **50**(2): 175-9.
31. George S, Rochford JJ, Wolfrum C, et al. A family with severe insulin resistance and diabetes due to a mutation in AKT2. *Science* 2004; **304**(5675): 1325-8.
32. Manning A, Highland HM, Gasser J, et al. A Low-Frequency Inactivating AKT2 Variant Enriched in the Finnish Population Is Associated With Fasting Insulin Levels and Type 2 Diabetes Risk. *Diabetes* 2017; **66**(7): 2019-32.

33. Latva-Rasku A, Honka MJ, Stancakova A, et al. A Partial Loss-of-Function Variant in AKT2 Is Associated With Reduced Insulin-Mediated Glucose Uptake in Multiple Insulin-Sensitive Tissues: A Genotype-Based Callback Positron Emission Tomography Study. *Diabetes* 2018; **67**(2): 334-42.
34. Minster RL, Hawley NL, Su CT, et al. A thrifty variant in CREBRF strongly influences body mass index in Samoans. *Nature genetics* 2016; **48**(9): 1049-54.
35. Lotta LA, Gulati P, Day FR, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nature genetics* 2017; **49**(1): 17-26.
36. Prokopenko I, Langenberg C, Florez JC, et al. Variants in MTNR1B influence fasting glucose levels. *Nature genetics* 2009; **41**(1): 77-81.
37. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature genetics* 2010; **42**(2): 105-16.
38. Saxena R, Hivert MF, Langenberg C, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nature genetics* 2010; **42**(2): 142-8.
39. Soranzo N, Sanna S, Wheeler E, et al. Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes* 2010; **59**(12): 3229-39.
40. Scott RA, Lagou V, Welch RP, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nature genetics* 2012; **44**(9): 991-1005.
41. Manning AK, Hivert MF, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nature genetics* 2012; **44**(6): 659-69.
42. Wheeler E, Leong A, Liu CT, et al. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS medicine* 2017; **14**(9): e1002383.
43. Ingelsson E, Langenberg C, Hivert MF, et al. Detailed physiologic characterization reveals diverse mechanisms for novel genetic Loci regulating glucose and insulin metabolism in humans. *Diabetes* 2010; **59**(5): 1266-75.
44. Prokopenko I, Poon W, Magi R, et al. A central role for GRB10 in regulation of islet function in man. *PLoS genetics* 2014; **10**(4): e1004235.
45. Scott RA, Fall T, Pasko D, et al. Common genetic variants highlight the role of insulin resistance and body fat distribution in type 2 diabetes, independent of obesity. *Diabetes* 2014; **63**(12): 4378-87.
46. Walford GA, Gustafsson S, Rybin D, et al. Genome-Wide Association Study of the Modified Stumvoll Insulin Sensitivity Index Identifies BCL2 and FAM19A2 as Novel Insulin Sensitivity Loci. *Diabetes* 2016; **65**(10): 3200-11.
47. Knowles JW, Xie W, Zhang Z, et al. Identification and validation of N-acetyltransferase 2 as an insulin sensitivity gene. *The Journal of clinical investigation* 2015; **125**(4): 1739-51.
48. Wood AR, Jonsson A, Jackson AU, et al. A Genome-Wide Association Study of IVGTT-Based Measures of First-Phase Insulin Secretion Refines the Underlying Physiology of Type 2 Diabetes Variants. *Diabetes* 2017; **66**(8): 2296-309.
49. Dimas AS, Lagou V, Barker A, et al. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* 2014; **63**(6): 2158-71.
50. Lyssenko V, Lupi R, Marchetti P, et al. Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *The Journal of clinical investigation* 2007; **117**(8): 2155-63.
51. Moran I, Akerman I, van de Bunt M, et al. Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell metabolism* 2012; **16**(4): 435-48.
52. van de Bunt M, Manning Fox JE, Dai X, et al. Transcript Expression Data from Human Islets Links Regulatory Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic Traits to Their Downstream Effectors. *PLoS genetics* 2015; **11**(12): e1005694.

53. van de Bunt M, Lako M, Barrett A, et al. Insights into islet development and biology through characterization of a human iPSC-derived endocrine pancreas model. *Islets* 2016; **8**(3): 83-95.
54. van de Bunt M, Gaulton KJ, Parts L, et al. The miRNA profile of human pancreatic islets and beta-cells and relationship to type 2 diabetes pathogenesis. *PloS one* 2013; **8**(1): e55272.
55. Thomsen SK, Ceroni A, van de Bunt M, et al. Systematic Functional Characterization of Candidate Causal Genes for Type 2 Diabetes Risk Variants. *Diabetes* 2016; **65**(12): 3805-11.
56. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**(7538): 197-206.
57. Akiyama M, Okada Y, Kanai M, et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nature genetics* 2017; **49**(10): 1458-67.
58. Turcot V, Lu Y, Highland HM, et al. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nature genetics* 2018; **50**(1): 26-41.
59. O'Rahilly S. Harveian Oration 2016: Some observations on the causes and consequences of obesity. *Clinical medicine* 2016; **16**(6): 551-64.
60. Yusuf S, Hawken S, Ounpuu S, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. *Lancet* 2005; **366**(9497): 1640-9.
61. InterAct C, Langenberg C, Sharp SJ, et al. Long-term risk of incident type 2 diabetes and measures of overall and regional obesity: the EPIC-InterAct case-cohort study. *PLoS medicine* 2012; **9**(6): e1001230.
62. Stefan N, Häring HU, Hu FB, Schulze MB. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. *Lancet Diabetes Endocrinol* 2013; **1**(2): 10.
63. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015; **518**(7538): 187-96.
64. Emdin CA, Khera AV, Natarajan P, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. *JAMA : the journal of the American Medical Association* 2017; **317**(6): 626-34.
65. Dale CE, Fatemifar G, Palmer TM, et al. Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus: A Mendelian Randomization Analysis. *Circulation* 2017; **135**(24): 2373-88.
66. Danforth E, Jr. Failure of adipocyte differentiation causes type II diabetes mellitus? *Nature genetics* 2000; **26**(1): 13.
67. Virtue S, Vidal-Puig A. Adipose tissue expandability, lipotoxicity and the Metabolic Syndrome--an allostatic perspective. *Biochimica et biophysica acta* 2010; **1801**(3): 338-49.
68. Pinnick KE, Nicholson G, Manolopoulos KN, et al. Distinct developmental profile of lower-body adipose tissue defines resistance against obesity-associated metabolic complications. *Diabetes* 2014; **63**(11): 3785-97.
69. Kilpelainen TO, Zillikens MC, Stancakova A, et al. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nature genetics* 2011; **43**(8): 753-60.
70. Small KS, Hedman AK, Grundberg E, et al. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nature genetics* 2011; **43**(6): 561-4.
71. Yaghootkar H, Scott RA, White CC, et al. Genetic evidence for a normal-weight "metabolically obese" phenotype linking insulin resistance, hypertension, coronary artery disease, and type 2 diabetes. *Diabetes* 2014; **63**(12): 4369-77.
72. Yaghootkar H, Lotta LA, Tyrrell J, et al. Genetic Evidence for a Link Between Favorable Adiposity and Lower Risk of Type 2 Diabetes, Hypertension, and Heart Disease. *Diabetes* 2016; **65**(8): 2448-60.
73. Tontonoz P, Hu E, Spiegelman BM. Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. *Cell* 1994; **79**(7): 1147-56.

74. Small KS, Todorcevic M, Civelek M, et al. Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. *Nature genetics* 2018; **50**(4): 572-80.
75. Portela A, Esteller M. Epigenetic modifications and human disease. *Nature biotechnology* 2010; **28**(10): 1057-68.
76. Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; **348**(6235): 648-60.
77. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nature communications* 2017; **8**: 14357.
78. Sun BB, Maranville JC, Peters JE, et al. Consequences Of Natural Perturbations In The Human Plasma Proteome. *bioRxiv* 2017.
79. Suhre K, Shin SY, Petersen AK, et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 2011; **477**(7362): 54-60.
80. Shin SY, Fauman EB, Petersen AK, et al. An atlas of genetic influences on human blood metabolites. *Nature genetics* 2014; **46**(6): 543-50.
81. Long T, Hicks M, Yu HC, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nature genetics* 2017; **49**(4): 568-78.
82. Wittemans LBL, Lotta, L.A., Langenberg, C. Prioritising Risk Factors for Type 2 Diabetes: Causal Inference Through Genetic Approaches. *Current diabetes reports* 2018.
83. Lotta LA, Scott RA, Sharp SJ, et al. Genetic Predisposition to an Impaired Metabolism of the Branched-Chain Amino Acids and Risk of Type 2 Diabetes: A Mendelian Randomisation Analysis. *PLoS medicine* 2016; **13**(11): e1002179.
84. Felig P, Marliss E, Cahill GF, Jr. Plasma amino acid levels and insulin secretion in obesity. *The New England journal of medicine* 1969; **281**(15): 811-6.
85. Newgard CB, An J, Bain JR, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell metabolism* 2009; **9**(4): 311-26.
86. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nature medicine* 2011; **17**(4): 448-53.
87. Wahl S, Drong A, Lehne B, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 2017; **541**(7635): 81-6.
88. Gamazon ER, Wheeler HE, Shah KP, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nature genetics* 2015; **47**(9): 1091-8.
89. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nature genetics* 2016; **48**(3): 245-52.
90. Hore V, Vinuela A, Buil A, et al. Tensor decomposition for multiple-tissue gene expression experiments. *Nature genetics* 2016; **48**(9): 1094-100.
91. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature genetics* 2016; **48**(5): 481-7.
92. Annual report of the Chief Medical Officer 2016: Generation Genome. 2017.
93. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. *Nature reviews Drug discovery* 2013; **12**(8): 581-94.
94. Cook D, Brown D, Alexander R, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nature reviews Drug discovery* 2014; **13**(6): 419-31.
95. Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nature genetics* 2015; **47**(8): 856-60.
96. Khera AV, Kathiresan S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nature reviews Genetics* 2017; **18**(6): 331-44.
97. Dewey FE, Gusarova V, Dunbar RL, et al. Genetic and Pharmacologic Inactivation of ANGPTL3 and Cardiovascular Disease. *The New England journal of medicine* 2017; **377**(3): 211-21.

98. Drucker DJ, Nauck MA. The incretin system: glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors in type 2 diabetes. *Lancet* 2006; **368**(9548): 1696-705.
99. Gallo LA, Wright EM, Vallon V. Probing SGLT2 as a therapeutic target for diabetes: basic physiology and consequences. *Diabetes & vascular disease research* 2015; **12**(2): 78-89.
100. Greenfield JR, Miller JW, Keogh JM, et al. Modulation of blood pressure by central melanocortinergic pathways. *The New England journal of medicine* 2009; **360**(1): 44-52.
101. Flannick J, Thorleifsson G, Beer NL, et al. Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nature genetics* 2014; **46**(4): 357-63.
102. Analysis of predicted loss-of-function variants in UK Biobank identifies variants protective for disease. 2018.
103. Pearson ER, Flechtner I, Njolstad PR, et al. Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations. *The New England journal of medicine* 2006; **355**(5): 467-77.
104. McCarthy MI, Hattersley AT. Learning from molecular genetics: novel insights arising from the definition of genes for monogenic and type 2 diabetes. *Diabetes* 2008; **57**(11): 2889-98.
105. Oral EA, Simha V, Ruiz E, et al. Leptin-replacement therapy for lipodystrophy. *The New England journal of medicine* 2002; **346**(8): 570-8.
106. Zhou K, Donnelly L, Yang J, et al. Heritability of variation in glycaemic response to metformin: a genome-wide complex trait analysis. *Lancet Diabetes Endocrinol* 2014; **2**(6): 481-7.
107. Majithia AR, Tsuda B, Agostini M, et al. Prospective functional classification of all possible missense variants in PPARG. *Nature genetics* 2016; **48**(12): 1570-5.
108. Agostini M, Schoenmakers E, Beig J, et al. A Pharmacogenetic Approach to the Treatment of Patients with PPARG Mutations. *Diabetes* 2018.
109. Maruthur NM, Gribble MO, Bennett WL, et al. The pharmacogenetics of type 2 diabetes: a systematic review. *Diabetes care* 2014; **37**(3): 876-86.
110. Srinivasan S, Kaur V, Chamarthi B, et al. TCF7L2 Genetic Variation Augments Incretin Resistance and Influences Response to a Sulfonylurea and Metformin: The Study to Understand the Genetics of the Acute Response to Metformin and Glipizide in Humans (SUGAR-MGH). *Diabetes care* 2018; **41**(3): 554-61.
111. GoDarts, Group UDPS, Wellcome Trust Case Control C, et al. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nature genetics* 2011; **43**(2): 117-20.
112. Zhou K, Yee SW, Seiser EL, et al. Variation in the glucose transporter gene SLC2A2 is associated with glycemic response to metformin. *Nature genetics* 2016; **48**(9): 1055-9.
113. Rotroff DM, Yee SW, Zhou K, et al. Genetic Variants in CPA6 and PRPF31 are Associated with Variation in Response to Metformin in Individuals with Type 2 Diabetes. *Diabetes* 2018.
114. Myocardial Infarction G, Investigators CAEC. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *The New England journal of medicine* 2016.
115. Larsson M, Caraballo R, Ericsson M, et al. Identification of a small molecule that stabilizes lipoprotein lipase in vitro and lowers triglycerides in vivo. *Biochemical and biophysical research communications* 2014; **450**(2): 1063-9.
116. Geldenhuys WJ, Aring D, Sadana P. A novel Lipoprotein lipase (LPL) agonist rescues the enzyme from inhibition by angiopoietin-like 4 (ANGPTL4). *Bioorganic & medicinal chemistry letters* 2014; **24**(9): 2163-7.
117. Gaudet D, Alexander VJ, Baker BF, et al. Antisense Inhibition of Apolipoprotein C-III in Patients with Hypertriglyceridemia. *The New England journal of medicine* 2015; **373**(5): 438-47.
118. Graham MJ, Lee RG, Brandt TA, et al. Cardiovascular and Metabolic Effects of ANGPTL3 Antisense Oligonucleotides. *The New England journal of medicine* 2017; **377**(3): 222-32.
119. Liu DJ, Peloso GM, Yu H, et al. Exome-wide association study of plasma lipids in >300,000 individuals. *Nature genetics* 2017.
120. Sattar N, Preiss D, Murray HM, et al. Statins and risk of incident diabetes: a collaborative meta-analysis of randomised statin trials. *Lancet* 2010; **375**(9716): 735-42.

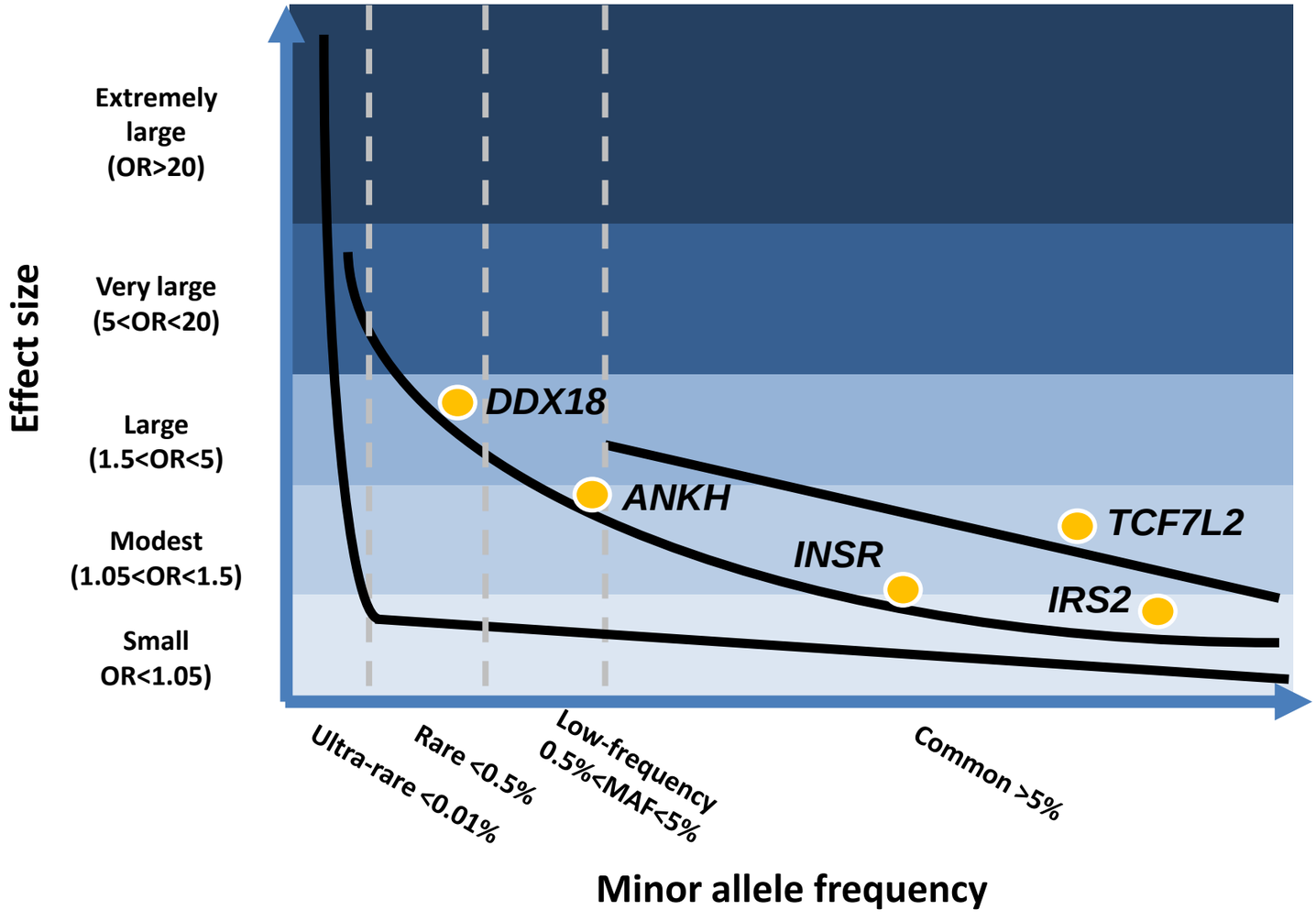
121. Swerdlow DI, Preiss D, Kuchenbaecker KB, et al. HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials. *Lancet* 2015; **385**(9965): 351-61.
122. Lotta LA, Sharp SJ, Burgess S, et al. Association Between Low-Density Lipoprotein Cholesterol-Lowering Genetic Variants and Risk of Type 2 Diabetes: A Meta-analysis. *JAMA : the journal of the American Medical Association* 2016; **316**(13): 1383-91.
123. Schmidt AF, Swerdlow DI, Holmes MV, et al. PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *Lancet Diabetes Endocrinol* 2017; **5**(2): 97-105.
124. Ference BA, Robinson JG, Brook RD, et al. Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. *The New England journal of medicine* 2016; **375**(22): 2144-53.
125. Blazing MAobotI-Ii. New-onset Diabetes Mellitus in the IMPROVE-IT Trial. *European Society of Cardiology Congress 2015* 2015.
126. de Carvalho LSF, Campos AM, Sposito AC. Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Inhibitors and Incident Type 2 Diabetes: A Systematic Review and Meta-analysis With Over 96,000 Patient-Years. *Diabetes care* 2018; **41**(2): 364-7.
127. FDA. Guidance for Industry. Diabetes Mellitus — Evaluating Cardiovascular Risk in New Antidiabetic Therapies to Treat Type 2 Diabetes 2008.
128. Lipska KJ, Krumholz HM. Is Hemoglobin A1c the Right Outcome for Studies of Diabetes? *JAMA : the journal of the American Medical Association* 2017; **317**(10): 1017-8.
129. Marso SP, Bain SC, Consoli A, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. *The New England journal of medicine* 2016; **375**(19): 1834-44.
130. Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and Cardiovascular Outcomes in Type 2 Diabetes. *The New England journal of medicine* 2016; **375**(4): 311-22.
131. Scott RA, Freitag DF, Li L, et al. A genomic approach to therapeutic target validation identifies a glucose-lowering GLP1R variant protective for coronary heart disease. *Science translational medicine* 2016; **8**(341): 341ra76.
132. Emdin CA, Klarin D, Natarajan P, et al. Genetic Variation at the Sulfonylurea Receptor, Type 2 Diabetes, and Coronary Heart Disease. *Diabetes* 2017; **66**(8): 2310-5.
133. Palmer SC, Mavridis D, Nicolucci A, et al. Comparison of Clinical Outcomes and Adverse Events Associated With Glucose-Lowering Drugs in Patients With Type 2 Diabetes: A Meta-analysis. *JAMA : the journal of the American Medical Association* 2016; **316**(3): 313-24.
134. Tabak AG, Jokela M, Akbaraly TN, Brunner EJ, Kivimaki M, Witte DR. Trajectories of glycaemia, insulin sensitivity, and insulin secretion before diagnosis of type 2 diabetes: an analysis from the Whitehall II study. *Lancet* 2009; **373**(9682): 2215-21.
135. Kengne AP, Beulens JW, Peelen LM, et al. Non-invasive risk scores for prediction of type 2 diabetes (EPIC-InterAct): a validation of existing models. *Lancet Diabetes Endocrinol* 2014; **2**(1): 19-29.
136. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *The New England journal of medicine* 2008; **359**(21): 2208-19.
137. Talmud PJ, Cooper JA, Morris RW, et al. Sixty-five common genetic variants and prediction of type 2 diabetes. *Diabetes* 2015; **64**(5): 1830-40.
138. Shigemizu D, Abe T, Morizono T, et al. The construction of risk prediction models using GWAS data and its application to a type 2 diabetes prospective cohort. *PloS one* 2014; **9**(3): e92549.
139. Farmer A, Fox R. Diagnosis, classification, and treatment of diabetes. *Bmj* 2011; **342**: d3319.
140. Thomas NJ, Jones SE, Weedon MN, Shields BM, Oram RA, Hattersley AT. Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *Lancet Diabetes Endocrinol* 2018; **6**(2): 122-9.
141. Ahlqvist E, Storm P, Karajamaki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* 2018.
142. Li L, Cheng WY, Glicksberg BS, et al. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. *Science translational medicine* 2015; **7**(311): 311ra174.

143. McCarthy MI. Painting a new picture of personalised medicine for diabetes. *Diabetologia* 2017; **60**(5): 793-9.
144. Dewey FE, Murray MF, Overton JD, et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science* 2016; **354**(6319).
145. Abul-Husn NS, Manickam K, Jones LK, et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science* 2016; **354**(6319).
146. Collins FS, Varmus H. A new initiative on precision medicine. *The New England journal of medicine* 2015; **372**(9): 793-5.

Formatted: Font: (Default) Times New Roman

Fig 1

Genetic mapping of diabetes susceptibility in the era of large biobanks



Sample size:

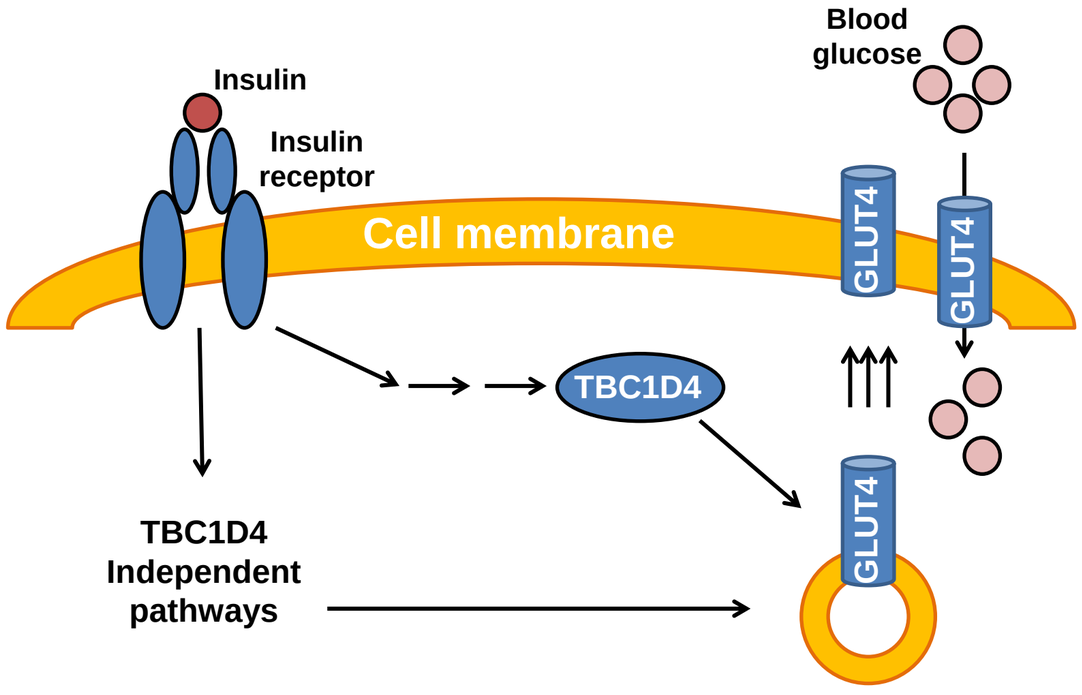
Early GWAS, N ~ 10,000s

Biobank-size GWAS , N ~ 100,000s

Meta-analysis of large biobanks , N ~ 1,000,000s

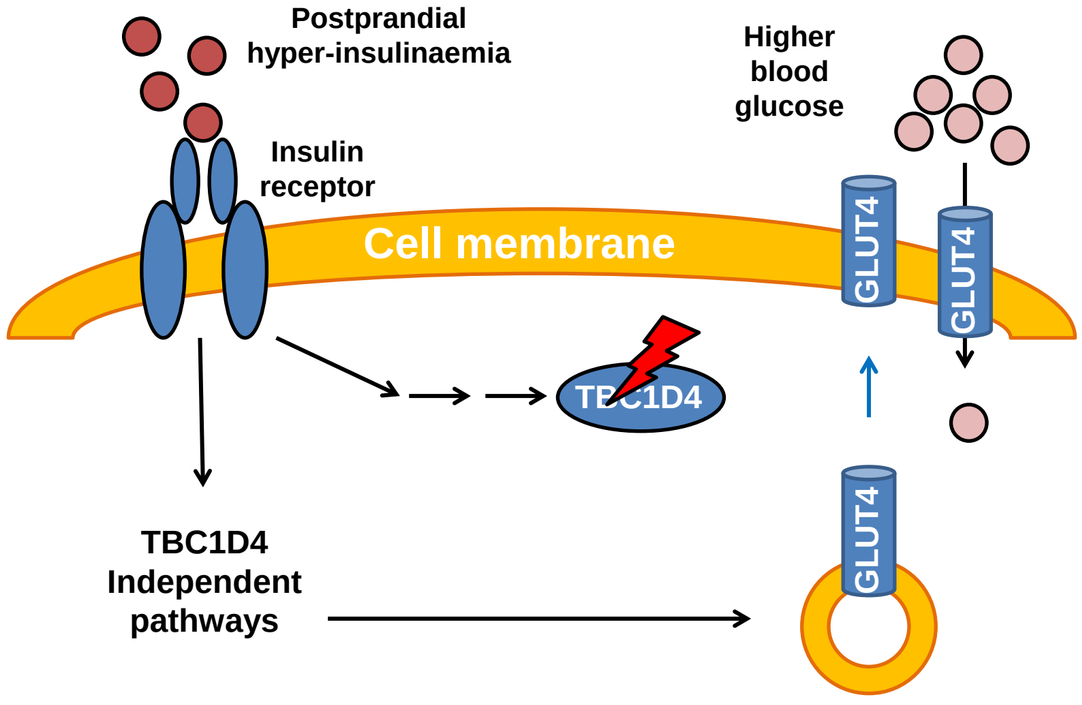
Fig 2

Normal skeletal muscle GLUT4 translocation



Skeletal Myocyte

Impaired skeletal muscle GLUT4 translocation



Reduced GLUT4 translocation and glucose uptake

Skeletal Myocyte

Fig 3

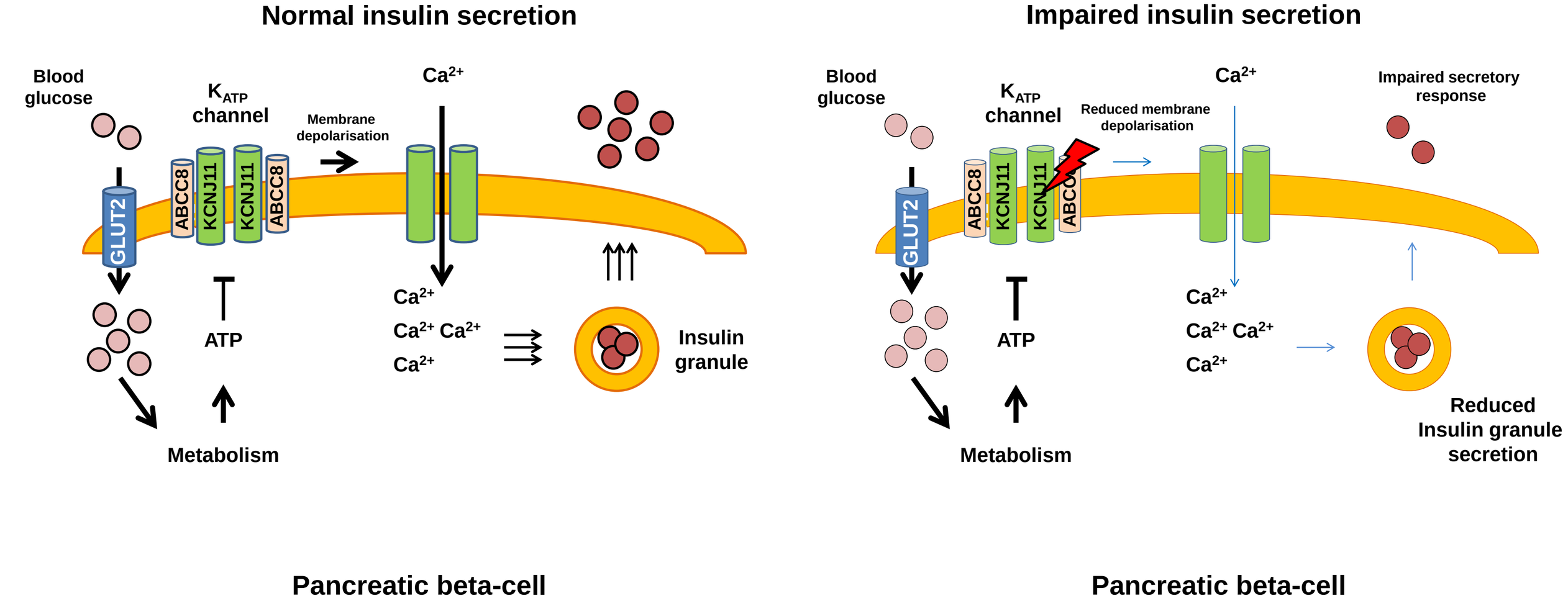
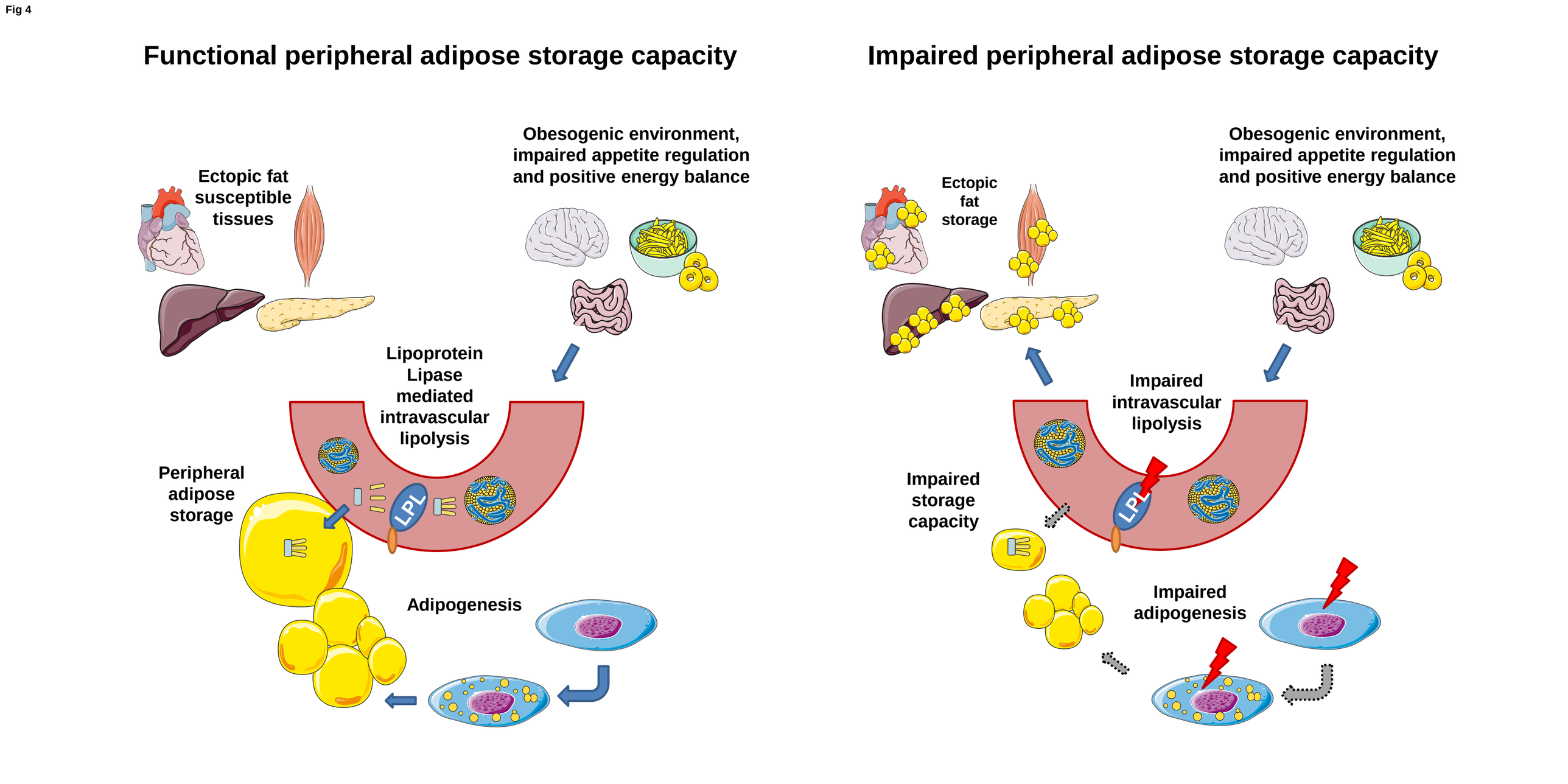

Normal insulin secretion
Blood glucose
K_{ATP} channel
Ca^{2+}
Membrane depolarisation
GLUT2
ABCC8
KCNJ11
KCNJ11
ABCC8
ATP
Ca^{2+}
Ca^{2+} Ca^{2+}
Ca^{2+}
Insulin granule
Metabolism
Pancreatic beta-cell
Impaired insulin secretion
Blood glucose
K_{ATP} channel
Ca^{2+}
Reduced membrane depolarisation
Impaired secretory response
GLUT2
ABCC8
KCNJ11
KCNJ11
ATP
Ca^{2+}
Ca^{2+} Ca^{2+}
Ca^{2+}
Reduced Insulin granule secretion
Metabolism
Pancreatic beta-cell

Fig 4


Functional peripheral adipose storage capacity
Ectopic fat susceptible tissues
Obesogenic environment, impaired appetite regulation and positive energy balance
Lipoprotein Lipase mediated intravascular lipolysis
LPL
Peripheral adipose storage
Adipogenesis
Impaired peripheral adipose storage capacity
Ectopic fat storage
Obesogenic environment, impaired appetite regulation and positive energy balance
Impaired intravascular lipolysis
LPL
Impaired storage capacity
Impaired adipogenesis

Editor

Editor comment 1 (EC1): "We feel that addressing the reviewers' concerns will considerably strengthen the article, particularly those of reviewer 1. Some editors raised concerns about the heavy emphasis on findings that are not yet available in peer reviewed journals. While it is fine to use data from these sources, we would suggest that you use them as you would a conference abstract and clearly signpost when data are sourced from the bioRchiv and that these are not peer reviewed. If, in the meantime, the articles have been accepted for publication, of course these comments become irrelevant. With this in mind, please also add a sentence to your search strategy about how you found papers on BioRchiv, was this a systematic search?"

EC1, Authors' reply: We agree and have now amended our review and references to pre-publication articles accordingly. Please also see response to reviewer 1, comment 1. Following the Editor's and first reviewer's concern, we have downplayed the relevance of pre-print publications and now state more clearly when we refer to them in the text. One of the articles has since been published in Nature Genetics. We now also provide the search strategy that was used to systematically identify relevant articles in BioRxiv.

EC2: "We wondered if you might also like to comment on recent contributions of genetic studies to classifications of disease beyond the monogenic forms of diabetes. Notably, Ahlqvist et al was recently published in TLDE and other clustering studies might be worth considering. This might fit with pathways rather than being a separate section."

EC2, Authors' reply: We agree that it is important to cite these approaches. The interesting article by Ahlqvist et al. has emerged after submission and we have now discussed and referenced this (P 11, L 420 of the new manuscript).

EC3: "Although we recognise that for many patients with type 2 diabetes their disease and obesity are intrinsically linked, we suggest that you shorten the discussion on obesity and lipids to make space for the additional suggestions by from reviewers."

EC3, Authors' reply: We have amended the manuscript to accommodate this suggestion (see detailed responses below).

Reviewer #1

Reviewer 1 comment 1 (R1C1): "this reviewer is surprised that this review is largely focused on a paper that has not been yet reviewed and published (only a preprint). In this regard the 250 loci for T2D should be cautiously presented especially for a large audience of the Lancet. The preprints are made to disseminate new ideas and to increase debate quality but this reviewer believes that only published peer reviewed papers should be presented in state of art review in particular in high IF journals."

R1C1, Authors' reply: We agree with this reviewer and editor that we needed to better highlight the non-peer reviewed nature of the BioRxiv articles and have amended the text of our article accordingly. One of the two manuscripts by Mahajan et al. has now been published in Nature Genetics and we reference it accordingly. The other has been reviewed and resubmitted and at this stage results are expected to not change dramatically. Because this large-scale meta-analysis a) substantially expands current knowledge, b) has been conducted by a team of internationally recognised experts with an outstanding track record in this area, and c) uses methodological standards that are accepted and go beyond what is expected for a discovery effort such as this one, we do think that reference to these results is appropriate and important and increases the timeliness and topicality of this review.

R1C2: "The abstract poorly summarizes the paper and should be rewritten. The last sentence is pure hope but not based on any robust evidence (that it may work) so far."

R1C2, Authors' reply: We have revised the abstract to better reflect the structure of the paper.

R1C3: "p 2: The 130 new loci should not be presented as granted. This introduction especially the last paragraph is too technical for geneticists only. This reviewer is not convinced that outside this small world people would understand what a posterior probability means. It is only a copy and paste of the preprint which is not the spirit of such a review for a large audience of readers."

R1C3, Authors' reply: We have revised the text to improve clarity and facilitate the understanding for a general and clinical audience, for instance by removing reference to the concept of posterior probability.

R1C4: "p3: the first paragraph is confusing as it does not differentiate GWAS analyses in various ethnicities and DNA sequencing of T2D genes (or exomes) in specific outbred populations. The DIAMANTE project is only a project and should not be presented as such in this review."

R1C4, Authors' reply: We have revised the text to reflect this suggestion and to more clearly distinguish findings arising from exome sequencing versus GWAS studies (P 3, L 83). We have trimmed the text referring to DIAMANTE, now mentioned only briefly but still highlighting the large scale and multi-ethnic nature of this important effort (P 3, L 92).

R1C5: "P 3 last paragraph: the ADCY3 story relates to food intake and obesity first and only secondarily to diabetes. The mechanism is likely to be very different from other T2D genes. In addition the ADY3 mutations, apart from the Inuit mutation were found in consanguineous populations. Of note these consanguinous populations are not isolated (such as Greenland Inuits who are 45K people) and constitute 20% of mankind. The study of consanguinous families was useful for monogenic diabetes but has not been correctly studied so far by GWAS and exome sequencing."

R1C5, Authors' reply: We agree with the reviewer that the ADCY3 example is relatively peripheral to the theme of this article and have shortened the corresponding text considerably (P 3, L 118), while still referring to the parallel publication based on the elegant study of consanguineous families.

R1C6: "p5: the pathway analysis part is very weak and too general in its findings. Many groups work (eg in Oxford, Lille, Malmo) on insulin secretion modulating genes found by GWAS and this quest should be presented. BTW the recent controversy on melatonin receptor opens the debate about genes directly or indirectly impairing insulin secretion (through brain effects) The last paragraph is very self centered on authors' own research. This reviewer has nothing against referring to fat deposition GWAS genes but it is rather far from T2D physiology (even if there is an overlap that should be presented with care). The lipodystrophy hypothesis is very interesting but deserves more explanation. Figure 2 is rather biased as seems to pretend that fat deposition is THE mechanism by which T2D develops in response to gene variants which is not proven. To this reviewer's knowledge recent papers from Gloyn/McCarthy, Bonnefond etc were strongly in favor of GWAS associated genes mostly expressed and functional in pancreatic beta cells. The fat deposition genes are more relevant to inflammation and cardiometabolic diseases. This part of the paper should be fixed."

R1C6, Authors' reply: It was not our intention to centre this part too strongly on fat deposition and following the reviewer's suggestion we now present a more balanced view of the different pathways contributing to T2D pathophysiology, clarifying that fat deposition is a specific contributory mechanism but not the only one (P 5, L 150-208). We have also drawn additional figures depicting other relevant mechanisms in T2D aetiology, including impaired GLUT4 translocation and insulin secretion, aiming to provide a broader set of relevant mechanistic examples (Figures 2-3). Since these latter mechanisms have been discussed elsewhere, we defer to the Editor on the final decision of whether these additional figures add to the review and should be included in the final version.

R1C7: "p7: it is too early to compare the status of the 250 loci (in preprint for half!) and PCSK9 with drugs in the market. BTW this part (in particular the last paragraph) confuses GWAS analyses (the preprint paper and all others before) and targeted/exome sequencing for rare mutations (protective of LoF). Very confusing indeed."

R1C7, Authors' reply: Thank you for bringing this to our attention; we have revised the text to distinguish the more clearly between GWAS versus sequencing studies (e.g. P 8, L 293-299). It was not our intention to draw a direct comparison between newly emerging GWAS loci and PCSK9. We merely refer to this as a successful example of rapid translation (PCSK9 variants discovered by gene sequencing that prompted the development of new cholesterol

lowering drugs have been discovered around the time of first GWAS of T2D), because this supports the value of genomic discovery approaches and precision medicine more generally.

R1C8: "p8: drug response and safety: the paper forgot the most important papers in the field that are related to metformin complications (Pearson's papers). This reviewer is obliged to say that according to the title the paper is not on lipids, fat and PCSK9 but on diabetes genes !"

R1C8, Authors' reply: We completely agree about the importance of Prof Pearson's papers on metformin and these are now cited and discussed (P 9, L 326-337). We also trimmed the discussion of examples of lipid-lowering drugs to stay closer to the main focus of the article (P 9, L 364).

R1C9: "p10: genomics in the clinic: authors have a negative view of GWAS outcomes in clinic. It is their right although recent evidence from genomic Risk Scores show hope that soon absolute risk may be assessed by genetics. Authors totally ignore prospective studies of incident risk (means not in case control studies) that have illustrated the interest and limits of SNPs in the prediction of incident cases (see Swedish, French and US studies). In addition on top of genetics metabonomics offers strong hope for accurate prediction of incident cases (which BTW is THE question in clinic). See UK Lolipop, French Desir, Dutch studies..."

R1C9, Authors' reply: This comment highlights the need to clarify our position and avoid any misunderstandings. Our comments on the relevance of GWAS findings for clinical prediction are meant to solely apply to the example of T2D, for which good, clinically accepted and used non-invasive and invasive predictive markers (i.e. HbA1c or glucose) already exist. These comments can and should not be applied to other outcomes of equally high clinical importance that are currently poorly predicted and/ or for which genetic prediction may play an important role even in the near future.

In the light of the reviewer's comment we have now revised this section (P 11, L 393) to also make it clearer that results are *not* based on case-control but prospective evidence (cohort and case-cohort).

Subgroups of the population where existing prediction algorithms work less well, and genetics adds *relatively* more (e.g. leaner and/ or younger individuals) are at low *absolute* risk and hence from a population or universal clinical provider perspective, the possible benefits of testing for genotypes or other markers in these groups has to be weighed against their low absolute risk. We have modified the text to reflect these considerations (P 11, L 407).

R1C10: "The fact that Regeneron/GSK contribute to sequence patients genome does not mean that it will provide key insight on T2D genetics. Indeed their interest is more on LoF "protective " mutations in important genes found in general population modifying important phenotypes (lipids, glucose...). Thus wait and see..."

R1C10, Authors' reply: We accept the views of the reviewer and have amended that section accordingly (P 11, L 436-453). As reviewer 3 highlights, "genomic studies for the identification of relevant genomic regions and validation of new therapeutic targets show tremendous promise for transformative clinical impact". As academia-industry partnerships

delivering large-scale genomics are becoming more prevalent, the challenges and opportunities arising from such efforts for T2D deserve some mention in our view.

Reviewer #3:

Reviewer 3 comment 1 (R3C1): "The authors reviewed the literature and have discussed several points concerning the hereditary basis of type 2 diabetes, pathways involved in diabetes identified through genomic studies, the genomics and therapeutics of type 2 diabetes, and the clinical relevance and future outlook. Genome-wide association studies have implicated many genomic regions, and there has been considerable advancement in the understanding of the mechanisms underlying these associations. Genomic studies for the identification of relevant genomic regions and validation of new therapeutic targets show tremendous promise for transformative clinical impact. Therefore, the authors focused on recent developments and offer insights into clinical translation of genetic findings."

R3C1, Authors' reply: We are grateful to the reviewer for thoroughly reviewing our work and providing important suggestions for improvements.

R3C2: "Major comments: Page 5, lines 24-43.
Shungin et al. identified several loci associated with body fat distribution with stronger effects in one sex than the other (Nature. 2015;518(7538):187-96). This finding should be mentioned because variants with sex-specific effects are interesting and important for considering the genetic basis of diabetes."

R3C2, Authors' reply: We agree and now cite this relevant article in the new version of the manuscript (P 5, L 191). However, we had to considerably shorten this section and the discussion of body fat distribution following the requests of the editor and other reviewers, so were unable to discuss details of any sex specific effects.

R3C3: "Page 8, lines 9-18.
The limitations of pharmacogenetics study design should be addressed. Recently, Srinivasan et al. conducted a novel prospective pharmacogenetic clinical trial (SUGAR-MGH), which revealed that a TCF7L2 variant associated with T2D influences the acute response to both glipizide and metformin in non-diabetic participants (Diabetes Care. 2018;41(3):554-61). The study design of SUGAR-MGH is very important because it is free of the uncontrolled nature of retrospective clinical data sets. The authors should describe variants at not only ATM and SLC2A2 but also TCF7L2 loci that were associated with response to diabetes drugs in Table 1."

R3C3, Authors' reply: Thank you for raising this; we have cited this interesting work and also added *TCF7L2* to the table (P 9, L 329).

R3C4: "Page 8, lines 19-32.
A systematic bioinformatics approach would be useful for repurposing of approved drugs in diabetes. For example, Imamura et al. identified two genes, GSK3<beta> and JUN, whose products directly interact with those of multiple biological T2D susceptibility genes, using a bioinformatics approach (Nature communications. 2016;7:10531). While therapeutic drugs for diseases other than diabetes targeting

GSK3<beta> and JUN were under clinical trials, these compounds could also be potential treatments for T2D. The authors may want to cite this article."

R3C4, Authors' reply: We thank the reviewer for this suggestion and now cite the article as a possible way forward in genetic-evidence driven re-purposing (P 9, L 347).

R3C5: "Minor comments:
Figure 1.
The authors should clarify the meaning of the number next to each given approach (e.g. ~10,000s). Does it indicate the total number of variants genotyped?"

R3C5, Authors' reply: We have modified the figure to clarify that numbers refer to the sample size (see Figure 1).

R3C6: "Page 10, lines 7-21.
Regarding diabetes prediction models, the authors should mention not only the clinical utility of a genetic risk score but also machine learning applications. For example, Shigemizu et al. developed a predictive model for T2D that consisted of nine SNPs selected using a Bayes Factor and lasso method with three clinical risk factors (age, gender, and BMI) and conducted a two-stage study (training and test sets) in a prospective cohort in Japan. The predictive model exhibited a 1.5% increase in the AUC over the clinical risk factors alone (PLoS One. 2014;9(3):e92549)."

R3C6, Authors' reply: This is an interesting point and we have modified the text to reference advanced statistical modelling approaches for improved prediction (P 11, L 402).

Reviewer #4:

Reviewer 4 comment 1 (R4C1): "I really liked this review. I think the last two sections in particular focus on the translational/clinical side of T2D genetics in a way I haven't seen too often before. Below are hopefully constructive comments:"

R4C1, Authors' reply: We would like to thank the reviewer for assessing our work so carefully and for the helpful and constructive comments. We have made changes to the manuscript in response to these suggestions and feel that the manuscript has greatly improved as a result.

R4C2: "Section 1 (global view)
In general I found the first two sections to need the most work. They seemed relatively unfocused and not necessarily making an argument leading up to the last two sections, which I felt to be quite strong. The shift in fact was quite abrupt. Maybe the two sections could state the overriding hypothesis of the review (which to me is how T2D genomics can impact translational and clinical research), frame the findings in the first two sections as such, and use their results as a foundation to support the major conclusions (e.g., to evaluate drug targets you need a lot of data and quantitative traits; for genetic risk prediction you need a lot of multi-ethnic data). I realize this is high level so below I'll call out specific areas that struck me."

R4C2, Authors' reply: Thank you for this suggestion; we have re-structured the first two sections in light of these suggestions and provide more detailed responses to related changes in our answers to the following comments below.

R4C3: "Very early on in Section 1 the authors state the reported (strong) ORs and lower frequencies from the recent T2D GWAS as evidence of much higher effect variants. It's important to mention (in a review like this) that these variants are of a different ilk than those previously reported: they haven't been validated in individual cohorts or even directly genotyped to make sure the imputation is accurate. Additionally, with the extreme case/control imbalance rare variant test statistics may not be well calibrated; a sentence suggesting this caution would be appropriate (particularly since future studies are going to only exacerbate this issue). Related to the first comment, I think at some point imputation quality should be addressed for these newer studies of much rarer variants."

R4C3, Authors' reply: We agree with these insights and have now added these points and downplayed our original comments about the newly identified rare variants (P 2, L 56).

R4C4: "The genetic isolate section seemed a little unfocused. It was framed as helping address the global basis of T2D, but then it goes into detail on three genes, the mechanism of which isn't really related to populations in which they are discovered. Mechanistic understanding of genes could come from variants discovered in any population; the fact that these variants were discovered in isolates is kind of orthogonal to the main content. The point of why isolates are useful is that large effect variants may (by chance) rise to high frequency in the isolate, which in some ways means that

individual high effect variants are better powered for detection in isolates. But this point doesn't come through with the micro focus on gene mechanisms. I personally think the gene mechanisms are well-covered in countless reviews and this one could focus more on why isolates can help the global epidemic of T2D. Restating this more broadly, the "global view" in the title doesn't really come through"

R4C4, Authors' reply: We agree with this comment and try to now provide a higher level "global view" by giving more prominence to the discussion of the overall utility, advantages and insights provided by studying population isolates (P 3, L 102). We have trimmed the discussion of specific mechanisms accordingly.

R4C5: "Is it worth mentioning the T2D knowledge portal, which is funded by five pharmaceutical companies and the NIH/FNIH Accelerating Medicines Partnership? T2D has a unique commitment to data sharing."

R4C5, Authors' reply: We completely agree and we now reference the T2D Knowledge Portal as a very important resource for T2D researchers (P 2, L 38).

R4C6: "How were the lines drawn in Figure 1? Is there a reason the bottom curve has a sharp point at 0.01%? If not, I would make it smooth. At minimum, some justification for this figure beyond pure heuristics would be useful (e.g., show actual discovered MAF/OR combinations)?"

R4C6, Authors' reply: We have drawn the figure using Microsoft pptx. Figures will be redrawn and improved by the Lancet's team for the final version. Following this comment, we have now added exemplar genetic variants to the graph and made it clear that the graph follows the empirical results of previous GWAS and sequencing studies (see Figure 1 and its legend).

R4C7: "Section 2
The idea about the trade-off between sample size and degree of refinement of phenotype ascertainment was very interesting and something I had not seen clearly articulated before. It wasn't really developed though. This whole section was extremely vague and didn't tell me a whole lot. I would love to see it refocused on the first sentence, and to tell me how different studies had chosen one of these trade-offs rather than the other, and what each could tell us about T2D. When do you do one and not the other? What findings were from one but not the other?"

R4C7, Authors' reply: We thank the reviewer for this helpful comment and have amended the new version accordingly (P 5, L 152-167).

R4C8: "Should PPARG be mentioned in the second subsection?"

R4C8, Authors' reply: We agree and have now added PPARG to this section (P 6, L 203).

R4C9: "The multi-omic data section I didn't really follow at all. It sets up a bunch of technologies, but then there are two paragraphs on metabolomics based discovery. Where are the other technologies? How did these prioritize pathways? A nested association is not a pathway. I think there is a case to be made that multi-omics is a major future direction, but this section seems to only cover metabolomics. In general, I found this section to be really weak. I think the two ideas (trade-offs in how secondary traits are studied, multiomics) are interesting to address, but they are not (at least upon my reading) actually addressed."

R4C9, Authors' reply: We have amended the section to now include a wider and more structured discussion of "-omics" approaches (P 6, L 211-254).

R4C10: "Should Mendelian randomization be covered, particularly given the clinical bent of the paper?"

R4C10, Authors' reply: While we agree that genetic approaches to causality are an interesting topic but given that this has been reviewed and covered extensively elsewhere, we think that this is beyond the scope of this review. However, we have now included clear reference and a brief statement to bring this topic to an interested reader's attention (P 6, L 229).

R4C11: "I would downplay the pathway/mechanism findings in general. They are overcovered by every other T2D review, and pruning them could focus this on the interesting ideas it broaches that have not been covered elsewhere."

R4C11, Authors' reply: We have shortened the sections about specific mechanisms to be able to expand the topics that have not been extensively covered by previous reviews as suggested (e.g. section on ADCY3, P3 L 118).

R4C12: "Sections 3 & 4
I really liked these sections, as mentioned above. My only suggestion is to refocus sections 1 and 2 so that they set up this section even more. I don't have any major comments here."

R4C12, Authors' reply: We thank the reviewer for these supportive comments.

R4C13: "My only minor comment (maybe a question): is this review supposed to only cover genomics of T2D in the UK? All of the future data mentioned is in UK, and similar US efforts like the MVP and All of Us are not mentioned. The explicit caveat that UK citizens have less to worry about with respect to discrimination made me think perhaps it is UK focused, but given that the title of the first section is "A global view of T2D genomics" the focus on the UK seems odd."

R4C13, Authors' reply: We agree and present a more balanced discussion of these topics in the new version, including specific reference to the MVP and All of Us efforts (P 12, L 455-463).