

Progress in genetics of type 2 diabetes and diabetic complications

Nobuhiro Shojima , Toshimasa Yamauchi*

Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Keywords

Type 2 diabetes, Diabetic complications, Precision medicine

*Correspondence

Tel.: +81-3-3815-5411

Fax: +81-3-5800-8161

E-mail address:

tyamau@m.u-tokyo.ac.jp

J Diabetes Investig 2023; 14: 503–515

doi: [10.1111/jdi.13970](https://doi.org/10.1111/jdi.13970)

ABSTRACT

Type 2 diabetes results from a complex interaction between genetic and environmental factors. Precision medicine for type 2 diabetes using genetic data is expected to predict the risk of developing diabetes and complications and to predict the effects of medications and life-style intervention more accurately for individuals. Genome-wide association studies (GWAS) have been conducted in European and Asian populations and new genetic loci have been identified that modulate the risk of developing type 2 diabetes. Novel loci were discovered by GWAS in diabetic complications with increasing sample sizes. Large-scale genome-wide association analysis and polygenic risk scores using biobank information is making it possible to predict the development of type 2 diabetes. In the ADVANCE clinical trial of type 2 diabetes, a multi-polygenic risk score was useful to predict diabetic complications and their response to treatment. Proteomics and metabolomics studies have been conducted and have revealed the associations between type 2 diabetes and inflammatory signals and amino acid synthesis. Using multi-omics analysis, comprehensive molecular mechanisms have been elucidated to guide the development of targeted therapy for type 2 diabetes and diabetic complications.

INTRODUCTION

Diabetes is rapidly increasing due to environmental factors such as changes in diet and decreased exercise.¹ Type 2 diabetes and its complications are associated with premature mortality, causing a public health burden worldwide.² The impact of type 2 diabetes on communicable disease has been highlighted. People with type 2 diabetes are at an increased risk of severe COVID-19 infection.

In type 2 diabetes, both genetic and environmental factors affect the onset of the disease, and although the proportion of involvement of these factors varies from case to case, it is thought to develop when the sum of the effects of both factors exceeds a certain threshold. Type 2 diabetes is caused by a combination of genetic factors that influence insulin secretion and insulin resistance, lifestyle such as obesity, overeating, and lack of exercise. It is estimated from twin family studies that the heritability of type 2 diabetic complications is about 40%.^{3–5} By elucidating the molecular mechanisms of genetic factors, we could optimize prevention and treatment, and achieve precision medicine tailored to the individual.

In 2003, about 3 billion base pairs of the human genome were sequenced. Comprehensive analysis methods for sequences using microarrays of single nucleotide polymorphisms have been developed, and disease susceptibility genetic loci have been identified by genome-wide association study (GWAS) on age-related macular degeneration first in 2005.⁶ GWAS is a method of testing the association between the allele frequency of a single-nucleotide polymorphism genotype and disease. Highly reproducible results have been obtained by analyzing a sufficient number of cases and confirming significant differences.^{7–12} Recently, the number of participants for GWAS has exceeded 1 million, and more disease susceptibility genes have been identified using large-scale meta-analysis.^{13–36} Genome-wide level associations with P -values under 5×10^{-8} have been observed in more than 600 genetic loci for diabetes, more than 1,100 loci for obesity.³⁷

In this review, we summarize genetic studies in type 2 diabetes and diabetic complications. We also consider how the polygenic risk score has begun to decipher the mechanisms related to diabetes and its complications. Finally, we explore the multi-omics analysis that enhances genotype information into biology and facilitates new avenues for treatment.

Received 28 November 2022; revised 12 December 2022; accepted 15 December 2022

DEVELOPMENT OF GWAS FOR TYPE 2 DIABETES

Genome-wide association study has shown statistically that the frequency of disease susceptibility gene polymorphisms differs between patient and control groups. It has been suggested that type 2 diabetes differs in genetic factors and pathophysiology depending on the ethnic population. While Europeans develop type 2 diabetes only after becoming obese, East Asians develop diabetes at a lower weight than Europeans due to the different level of insulin secretion.^{38,39} To identify ethnically characteristic type 2 diabetes genes, GWAS using large-scale, high-density SNPs in each ethnic population is required.

An association study on chromosome 10q, which has been mapped as the type 2 diabetes locus, identified variants in *TCF7L2* (T cell factor like 2).⁴⁰ In 2007, the first GWAS for type 2 diabetes included thousands of participants and identified 10 genomic loci.^{7,8,9,41,42} Now the GWAS of type 2 diabetes, mainly from European and Asian populations, has identified more than 600 disease-susceptible loci (Figure 1).

KCNQ1 is a disease susceptibility locus with a relatively high odds ratio of 1.4 identified in the Japanese population in 2008, but it is highly reproducible, including in the European population.^{10,11} *KCNQ1* protein is one of the voltage-dependent potassium channels and has been reported to cause long QT

syndrome by altering repolarization in cardiomyocytes.^{43,44} In 2010, a large-scale GWAS was performed using 459,359 SNPs in 4,470 people with diabetes. The GWAS identified *UBE2E2* and *C2CD4A/B* as disease susceptibility loci for type 2 diabetes.¹³ A ubiquitin-proteasome system has been reported to play a pivotal role in insulin secretion.⁴⁵ In addition, the genetic score using 13 disease susceptibility loci of type 2 diabetes was developed to explain the difference of the risk of diabetes.¹³

A genome-wide association study identified five novel loci in a Han Chinese population,¹⁵⁻¹⁷ and seven novel loci in a South Asian population.¹⁸ One locus was newly identified by imputation with the sequencing information of the 1,000 Genome Project and three new loci of disease susceptibility to type 2 diabetes were identified by large-scale GWAS in a Japanese population.^{19,25} A larger meta-analysis identified seven new disease-susceptible loci of type 2 diabetes, two of which were relatively characteristic of the East Asian population.³¹

LARGE-SCALE GWAS FOR TYPE 2 DIABETES IN JAPANESE POPULATION

To elucidate genetic and environmental factors of type 2 diabetes through the close collaboration of large cohorts of Japan (Biobank Japan, J-MICC, JPHC-NEXT, Tohoku Medical

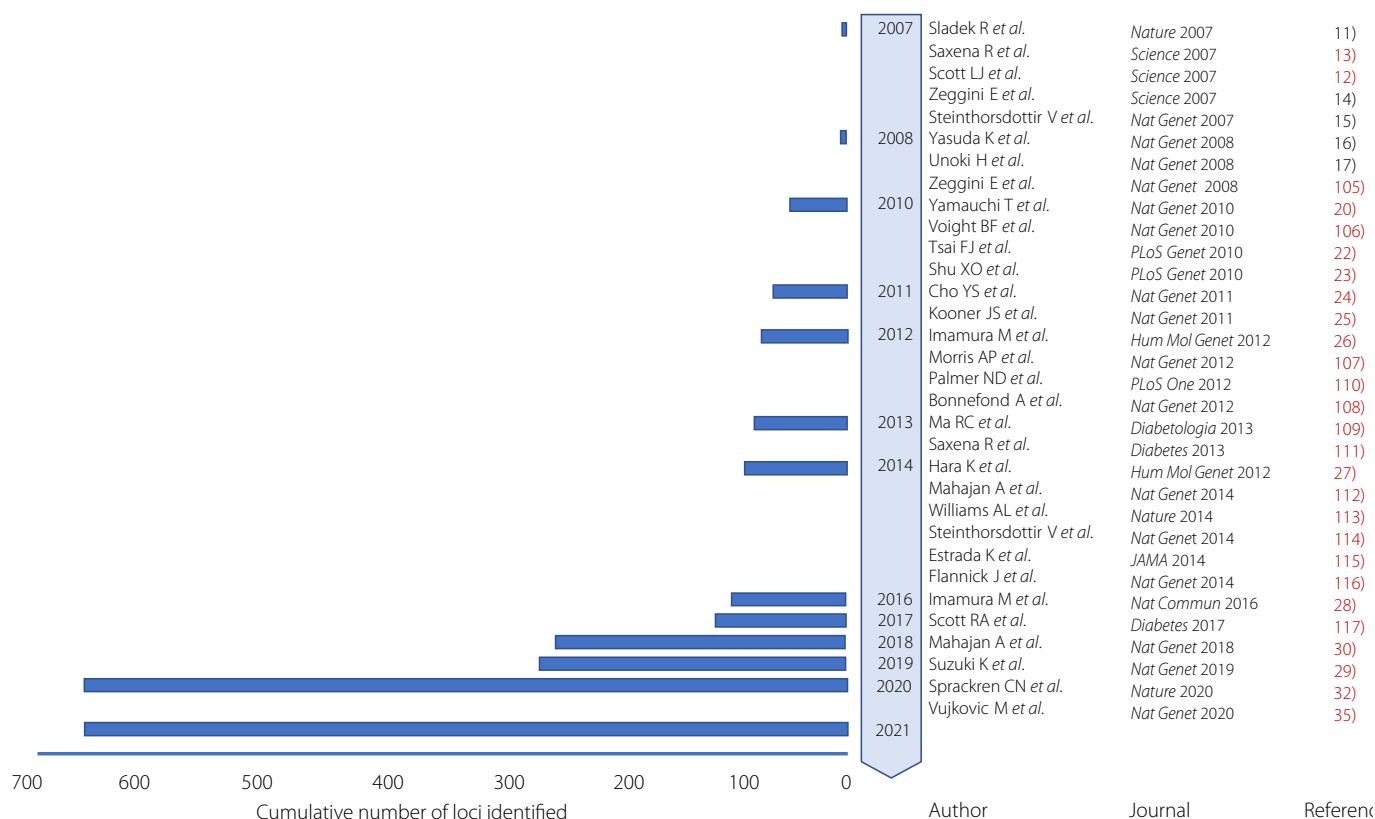


Figure 1 | Timeline of key discoveries of type 2 diabetes genetics by genome-wide association study (GWAS). GWAS identified genetic loci for type 2 diabetes at the genome-wide significance level ($P < 5.0 \times 10^{-8}$). Cumulative count of discovered loci per year are presented.

Megabank) and researchers in the medical and information sciences, it is important to propose a new standard for research on multifactorial diseases. In the Japanese population, a large GWAS meta-analysis was performed in 36,614 case and 155,150 control participants.³⁶ GWAS was conducted on 12 million variants using a reference panel of 1,000 Genomes Project Phase 3 for imputation. Thus 88 regions and 115 independent signals associated with type 2 diabetes were identified. Only eight loci were common with the new loci reported by 900,000 participants in the European population, and the remaining 20 loci were specific to the Japanese population.³³ This shows that no matter how large the GWAS size of the European population becomes, it is not possible to identify the disease susceptibility loci specific to the Japanese population. Pathway analysis revealed that a pathway of maturity-onset diabetes of the young is common to Japanese and European populations. The pathway of insulin secretion regulation was specific to the Japanese population.

In the GWAS in the Japanese population, including Biobank Japan, 15 new missense variants were identified.³⁶ It was newly reported that the variant of incretin-related drug target gene *GLPIR* R131Q (rs3765467 GA/AA genotype) was associated with the risk of developing type 2 diabetes.³⁶ Healthy individuals with the GA/AA genotype had an increased incretin effect and the genotype is expected to be a drug reactivity marker.⁴⁶ The frequency of the GA/AA genotype of rs3765467 is 32% in the Japanese population, but it is extremely rare in the European population and difficult to identify even in GWAS with a 900,000 European population.³³

EAST ASIAN AND TRANS-ETHNIC GWAS FOR TYPE 2 DIABETES

A meta-analysis of a large-scale GWAS of 400,000 people was conducted to identify common type 2 diabetes loci in the East Asian population.³⁵ The number of cases of type 2 diabetes surpassed that of the European GWAS at that time. The Japanese population made the largest contribution to the number of cases. In total, 183 risk loci were associated with type 2 diabetes and 61 loci were novel. In addition to loci that are significant in both European and East Asian populations, loci that are specific to the East Asian population were identified. For example, *TCF7L2* had the same orientation and magnitude of effect in both populations. On the other hand, *DGKD* and *ZNF257* had an effect only in the East Asian population. The effect sizes of the allele adjusted for BMI were highly correlated to those of the allele unadjusted for BMI in the East Asian population. It was speculated that changes in fat distribution may play a larger role in causing type 2 diabetes than BMI in the East Asian population. In addition, the *FTO* region, which is known to be most strongly related to BMI, was similarly attenuated by BMI correction in both East Asian and European populations.

Regarding the *KCNQ1* region that showed the most significant correlation with type 2 diabetes, nine signals were identified in the *KCNQ1* region.³⁵ In the previously reported *ANK1*/

NKX6-3 region, three independent signals were found this time.³⁵ Of these, signal 1 (rs33981001) was adjacent to a variant that regulates the gene expression of *NKX6-3* on the islet of Langerhans of the pancreas. On the other hand, signal 2 (rs62508166) was adjacent to a variant that regulates the gene expression of *ANK1* in subcutaneous adipose tissue and skeletal muscle. It is speculated that two different signals in the same genetic region affect the risk of developing type 2 diabetes by modulating the expression of different genes in different tissues.

Among the loci identified in this study, the locus where the effect on the risk of developing type 2 diabetes differs between men and women was the aldehyde dehydrogenase 2 (*ALDH2*) locus (rs12231737).³⁵ The *ALDH2* locus was strongly associated with the risk of developing type 2 diabetes in men, but not at all in women. It has been reported that the *ALDH2* locus is related to adaptive evolution in the East Asian population.^{47,48} In addition, *ALDH2* protein is an enzyme that regulates alcohol metabolism. The allele was associated with tolerance to alcohol, a high risk of developing type 2 diabetes, obesity, blood pressure, and serum triglycerides.

Trans-ethnic GWAS for type 2 diabetes and diabetic complications was conducted. Vujkovic *et al.*³⁴ investigated type 2 diabetes genetic susceptibility *via* multi-ancestry meta-analysis of 228,499 cases and 1,178,783 controls in the DIAMANTE and Million Veteran Program (MVP), including Biobank Japan and other studies. In total, 568 associations and 318 novel loci for type 2 diabetes were identified. Novel loci in the diabetic complications were reported. The polygenic risk score of diabetic retinopathy was shown to strongly associate with an increased risk of diabetic retinopathy.³⁴

GENOMIC ANALYSIS OF DIABETIC KIDNEY DISEASE

Diabetic kidney disease is a multifactorial disease that develops with the involvement of multiple genetic factors and environmental factors.⁴⁹ The elucidation of genetic factors has been advanced, narrowing several related genetic regions by analysis using family lineage.⁵⁰ As a candidate gene, the erythropoietin gene was associated with diabetic kidney disease. Disease susceptibility genetic loci have been identified in diabetic kidney disease by GWAS. The functions of the identified genes have been elucidated using human kidney, cultured cells, and model mice.⁵¹

In 2009, a GWAS of diabetic nephropathy in type 1 diabetes in the European population reported four loci that did not reach genome-wide levels. In the Japanese population with type 2 diabetes, GWAS identified the *ACACB* (acetyl coenzyme A carboxylase β) locus with a *P*-value of over 5×10^{-8} .⁵² In the Genetics of Nephropathy an International Effort (GENIE) in diabetic nephropathy of type 1 diabetes in European population, *AFF3* (AF4/FMR2 member 3) and *RGMA/MCTP2* (RGM domain family, member A/Multiple C2 domains, transmembrane 2) loci were significantly associated at the genome-wide level⁵³ (Figure 2). *AFF3* protein has been reported to be involved in TGF- β signaling. The *CDC47/SP3* (cell division

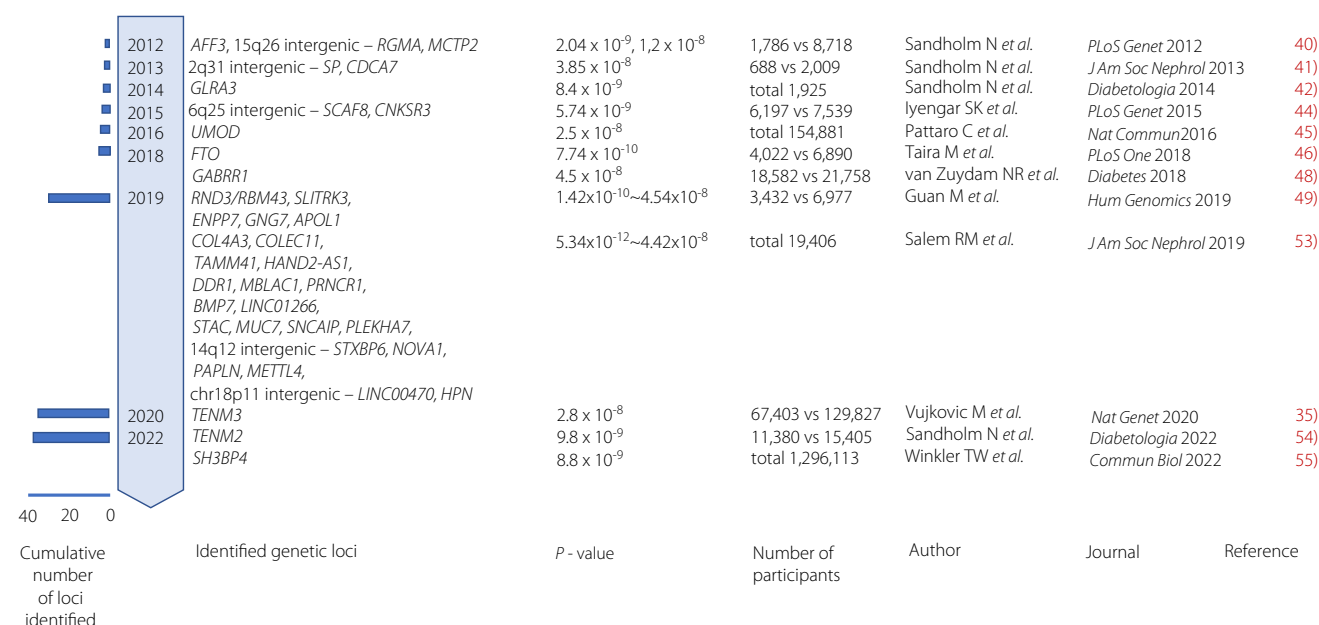


Figure 2 | Timeline of key discoveries of genetics of diabetic kidney disease by genome-wide association study (GWAS). GWAS identified genetic loci for type 2 diabetes at the genome-wide significance level ($P < 5.0 \times 10^{-8}$). Cumulative count of discovered loci per year are presented.

cycle associated 7/transcription factor Sp3) locus was significantly associated with end-stage renal failure in women with type 1 diabetes.⁵⁴ The locus of *GLRA3* (glycine receptor, alpha-3 subunit) is identified to be associated with albuminuria in type 1 diabetes.⁵⁵ The protein encoded by *GLRA3* is a ligand-gated chloride channel responsible for mediating the inhibitory effects of glycine.⁵⁶ Glycine is the predominant amino acid stimulator of glucagon secretion in pancreatic α -cells, acting through the glycine receptors. In the European population, intergenic locus of *SCAF8/CNKSR3* (SR-like carboxyl-terminal domain associated factor 8 gene/connector enhancer of KSR family of scaffold proteins gene) was identified.⁵⁷ The protein encoded by *CNKSR3* is related to aldosterone signals. The *UMOD* locus showed a genome-wide significant association with eGFR (estimated glomerular filtration ratio) among those with diabetes.⁵⁸ Taira *et al.*⁵⁹ identified the *FTO* locus and found that it confers susceptibility to diabetic nephropathy. The *FTO* locus has been reportedly associated with obesity.⁶⁰ The *FTO* variants confer susceptibility to chronic kidney disease through the mechanisms mediated by obesity.

The Surrogate markers for Micro- and Macrovascular hard endpoints for Innovative diabetes Tools (SUMMIT) consortium has conducted an analysis of 40,000 individuals with diabetic kidney disease and type 2 diabetes. Seven phenotypes of diabetic kidney disease were compared based on stage of the disease. In the comparison of individuals with microalbumin urine and those without microalbumin urine, the *GABRR1* (Rho1 subunit of the GABA type A receptor) loci reached a genome-wide significant level.⁶¹ For the association with glomerular

estimated filtration, five loci were significant at genome-wide levels, including the *APOL1* and *ENPP7* locus.⁶² *APOL1* (apolipoprotein L1) G1 and G2 alleles gene contribute to 60% of non-diabetic end-stage kidney disease in African Americans.⁶³ The protein encoded by *ENPP7* is an intestinal phosphodiesterase that converts sphingomyelin to ceramide and phosphocholine. *ENPP7* protein is reported to affect cholesterol absorption.⁶⁴ Studies suggest that the high-density lipoprotein cholesterol levels are risk factors for diabetic kidney disease.⁶⁵

A meta-analysis of GWAS including 19,406 individuals with type 1 diabetes from the Diabetic Nephropathy Collaborative Research Initiative (DNCRI) identified 16 novel loci including *COL4A3*, *COLEC11*, *DDR1*, and *BMP7*. A common missense variant in the *COL4A3* gene showed the strongest association.⁶⁶ Variants in the *COL4A3* gene are responsible for Alport syndrome, a progressive inherited nephropathy. The *COLEC11* gene encodes a collectin-11 protein. *DDR1* encodes the discoidin domain-containing receptor 1. Collectin-11 protein has a role in host defense against pathogens and *DDR1* protein binds collagens. *BMP7* (bone morphogenetic protein 7) is a member of TGF- β superfamily and has a role in renal development and fibrosis. *TENM2* and *TENM3* loci are reported to associate with diabetic nephropathy.^{67,68} *TENM2* (teneurin transmembrane protein 2) is involved in cell–cell adhesion. *TENM3* has been associated with kidney disorders in UK Biobank.⁶⁹

GENETICS OF DIABETIC RETINOPATHY

Diabetic retinopathy causes blindness among individuals with diabetes.^{70,71} Beside candidate gene analysis and linkage

analysis, GWAS were performed in various populations. It was reported that the locus of *CAMK4* (calcium/calmodulin-dependent protein kinase IV) was found to be associated with diabetic retinopathy in a GWAS performed in a Mexican population without reaching genome-wide significance.⁷² Huang *et al.*⁷³ conducted GWAS in a Taiwanese population and identified loci including *ARHGAP22* reaching genome-wide significance (Figure 3). *ARHGAP22* (Rho GTPase-activating protein 22) is implicated in endothelial cell angiogenesis and increased capillary permeability. Grassi *et al.* performed GWAS in individuals with severe diabetic retinopathy in two large cohorts of type 1 diabetes: Genetics of Kidney in Diabetes (GoKinD) and EDIC. The study included 973 cases and assessed them to identify an intergenic locus without reaching genome-wide significance.⁷⁴ Burdon *et al.*⁷⁵ reported genetic variation near *GRB2* (growth factor receptor bound-2) locus associated with sight-threatening diabetic retinopathy. *GRB2* protein was reported to be expressed in all the retina layers, including blood vessels. *GRB2* binds phosphorylated IRS-1 (insulin receptor substrate 1) and activates the MAPK pathway by insulin stimulation. *GRB2* is involved in VEGF signaling. GWAS of severe diabetic retinopathy suggested evidence for the involvement of the locus of *NOX4* (NADPH oxidase 4) in the European population.⁷⁶ *NOX4* protein acts as an oxygen sensor and catalyzes the reduction of molecular oxygen. In a GWAS in European and African cohorts, the locus in an intron of *NVL* (nuclear valosin-containing protein-like) was found to confer susceptibility to diabetic retinopathy in European population. *NVL* is expressed widely in the retina and encodes the AAA (ATPases associated with the cellular activities) superfamily.⁷⁷

Four loci were found to have effects on diabetes retinopathy in GWAS meta-analysis.³⁴ The strongest signal for interaction

to retinopathy was observed for *GJA8*.³⁴ Deletion of *GJA8* gene leads to eye abnormalities and retinopathy in premature infants, inherited cataracts, and visual impairment.^{78,79} *TCF7L2* is found to be a locus associated with both type 2 diabetes and diabetic retinopathy. The product of *SLC18A2* gene is involved in the transport of monoamines. The protein encoded by *SVILP1* has been reported to be associated with thiamine (vitamin B1) metabolism. Imamura *et al.*⁸⁰ conducted a GWAS in the Japanese population and identified novel loci near *STT3B* (Oligosaccharyl transferase Complex Catalytic Subunit B) and *PALM2* (paralemin-2) to be associated with diabetic retinopathy. Deletion of *STT3B* mRNA causes a congenital neurological disease with a lack of visual tracking and the optic nerve hypoplasia.

GENETICS OF DIABETIC NEUROPATHY

Diabetic neuropathy is the most common complication and increases the risk of foot ulcer and lower-extremity amputation.⁸¹ Diabetic neuropathy is presented as peripheral and autonomous neuropathy. Several gene polymorphisms in candidate gene have been studied including *ACE* (angiotensin converting enzyme).⁸²

Meng *et al.*⁸³ conducted GWAS in 699 diabetic foot ulcers in the Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTs; Figure 4). The control was defined as an individual with diabetes who had never been recorded as having a foot ulcer but who had a positive monofilament test result recorded in their e-health records. The *MAPK14* locus reached genome-wide significance with a *P*-value of 2.45×10^{-8} . *MAPK14* protein is one of the p38 MAPKs and plays a role in the cellular response to proinflammatory cytokines. Tang *et al.*⁸⁴ performed GWAS in the Action to Control Cardiovascular Risk in Diabetes (ACCORD) clinical trial

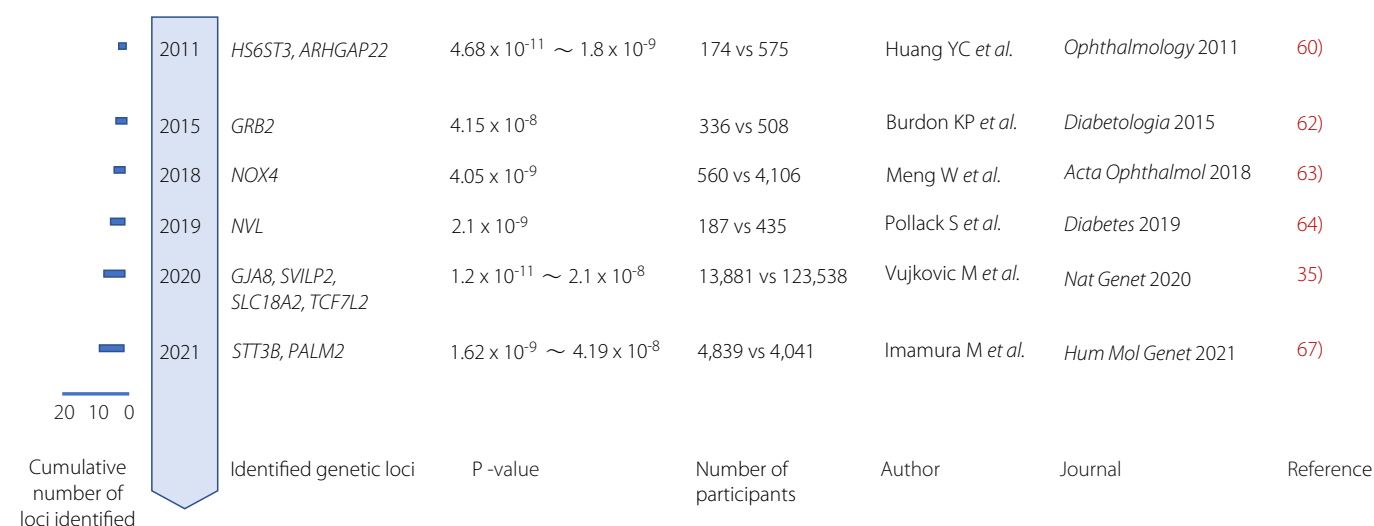


Figure 3 | Timeline of key discoveries of genetics of diabetic retinopathy by genome-wide association study (GWAS). GWAS identified genetic loci for type 2 diabetes at the genome-wide significance level ($P < 5.0 \times 10^{-8}$). Cumulative count of discovered loci per year are presented.

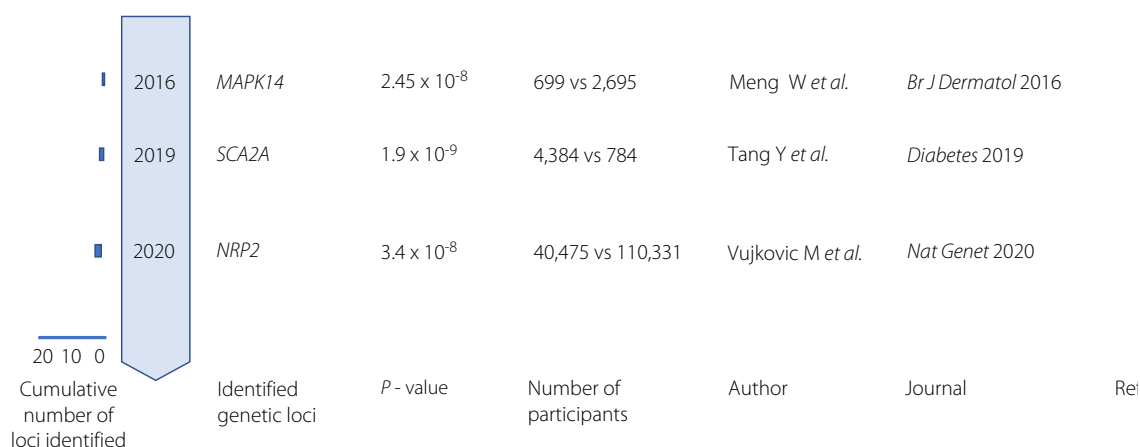


Figure 4 | Timeline of key discoveries of genetics of diabetic neuropathy by genome-wide association study (GWAS). GWAS identified genetic loci for type 2 diabetes at the genome-wide significance level ($P < 5.0 \times 10^{-8}$). Cumulative count of discovered loci per year are presented.

including 4,384 individuals with diabetic peripheral neuropathy. The minor allele of the rs13417782 decreased the odds of diabetic peripheral neuropathy by 36%. This protective allele was associated with a higher expression of a gene *SCN2A*. The protein encoded by *SCN2A* is voltage-gated sodium channel Nav1.2. The *NRP2* locus was identified in GWAS meta-analysis of 40,475 individuals with diabetic neuropathy.⁴⁶ *NRP2* encodes neuropilin-2 which is a receptor involved in angiogenesis and neural regeneration. Recently, metagenomic sampling of diabetic ulcers was performed and microbial dynamics were assessed to determine relationships to clinical events.^{85,86} Larger scale GWAS and a metagenomic study is clearly needed to fully understand the mechanisms of the development of diabetic neuropathy.

POLYGENIC RISK SCORE FOR PREDICTION OF THE DEVELOPMENT OF DIABETES AND THE THERAPEUTIC EFFECTS OF DIABETIC COMPLICATIONS

The polygenic risk score (PRS) is calculated by the weighted sum of risk allele counts of the variants, using their effect sizes from GWASs as their weights.⁸⁷ The power of PRSs has increased due to larger GWAS sample sizes and advanced computational algorithms compared with classical genetic risk score.⁸⁸ The polygenic risk score can be constructed that identifies clinically meaningful genetic predisposition for diseases and the response to their treatments. Large-scale GWAS and biobank information have been used to demonstrate the usefulness of PRS with the significant variants and non-significant weakly influential variants. It was reported that approximately 20% of people with high PRS have a triplicate risk of cardiovascular disease, atrial fibrillation, type 2 diabetes, Crohn's disease, and breast cancer.⁸⁹ It is estimated that the risk in the high-risk group is similar to the risks of developing monogenic diseases with positive genetic testing. The clinical application of PRS for

the prediction of the development of diabetes and diabetic complications is expected (Figure 5).

Genome-wide association studies have been mainly conducted in the European population, and about 80% of the participants are Europeans.^{90–92} In type 2 diabetes, PRS created from data in a cohort of the European population have been reported to be less predictable when adapted to the Japanese population.⁹⁰ Since the frequency of variants involved in the onset and progression of the disease and the linkage around the causative variants differ among ethnic groups, it might not be possible to calculate the effect of the causative variant using GWAS data from different ethnic groups. For the clinical application of research results of genomic medicine and precision medicine, data obtained from large-scale cohorts of each population are important.

In the ACCORD trial about diabetic complications, a subtype with an increased effect of intensive therapy was identified.⁹³ GWAS has been performed in the subtype classified by cluster HbA1c trajectory analysis. PRS using GWAS data were created to identify subtypes. In subtypes classified by PRS, it was reported that intensive therapy inhibited the development of coronary artery disease. In addition, in the ADVANCE (The Action in Diabetes and Vascular Disease: Preterax and Diamicon Modified Release Controlled Evaluation) trial, a multi-PRS combining 10 PRS such as diabetes, obesity, and albuminuria with 598 single-nucleotide polymorphisms was created. Multi-PRS was useful for predicting the development of diabetic complications. In the high-risk group, cardiovascular death was 4.4-fold higher compared with the low-risk group (HR 0.61 [95% CI 0.40, 0.93] for 4.5 year follow-up). In the high-risk group, intensive therapy that strictly lowers plasma glucose levels and blood pressure was shown to have a larger effect of decreasing complications compared with the low-risk group.⁹⁴ The number needed to treat (NNT) was 12 in the high-risk group.

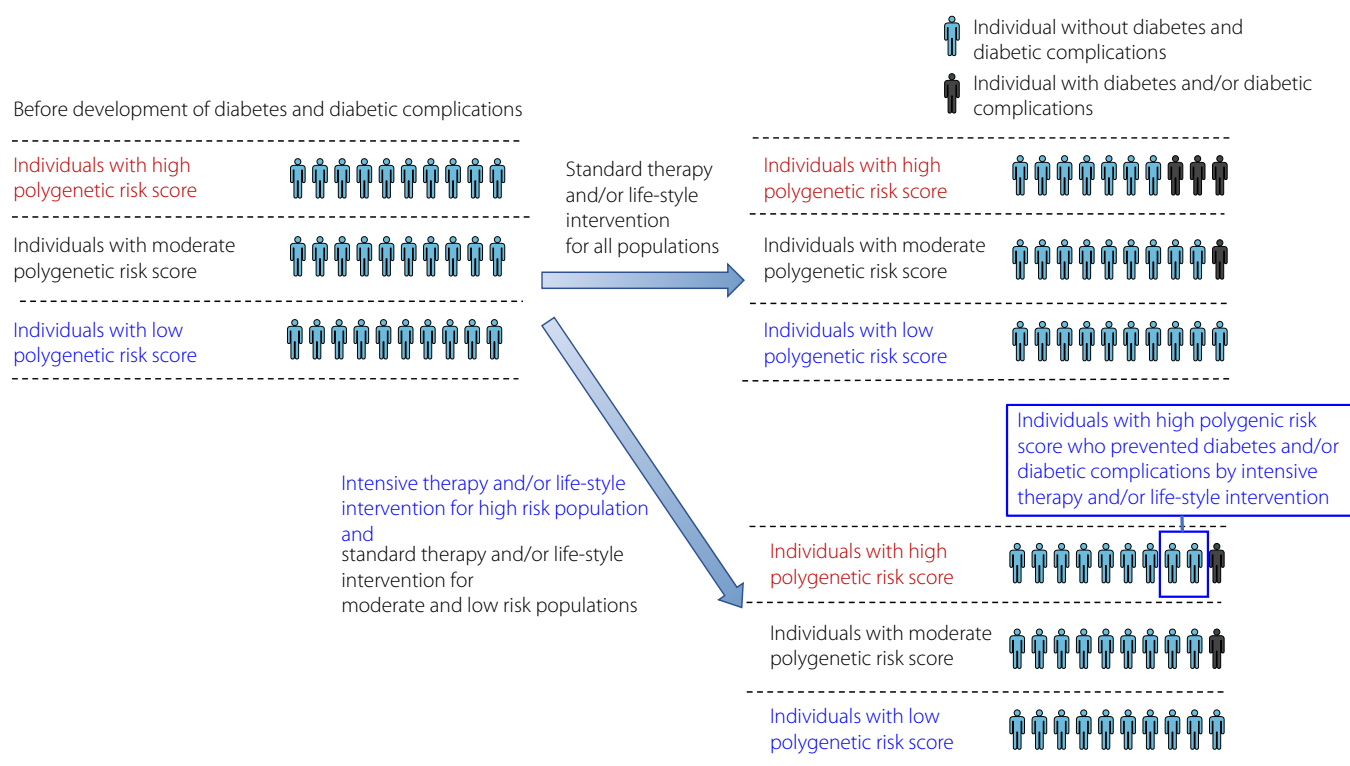


Figure 5 | Predicting diabetes and diabetic complications using polygenic risk score. Intensive therapy and/or life-style intervention for high-risk group reduces the development of diabetes and diabetic complications compared with standard therapy and/or life-style intervention for all groups.

MULTI-OMICS ANALYSIS FOR DIABETES AND DIABETIC COMPLICATIONS

We could now integrate information across multiple data types to gain a comprehensive molecular understanding to guide development of targeted therapy.⁹⁵ In relation to environmental factors, the role of genetic predisposition will be clarified by integrating information at each level of cells and tissues through multi-omics analysis. Multi-omics profiling is an important step toward precision medicine in metabolic diseases.^{96,97} Multi-omics techniques, in particular proteomics and metabolomics together with genomics, have the potential to enhance our understanding of the molecular process that develops type 2 diabetes and influences disease progression (Table 1). Ninety-nine plasma proteins were associated with the incidence of developing type 2 diabetes in 2,940 participants.⁹⁸ Four plasma proteins were associated with pre-diabetes in 439 participants. The biological process concerning these proteins were inflammatory signals, and lipoprotein metabolism.⁹⁹

The progression of diabetic kidney disease and diabetic retinopathy is variable, and little is known about the factors that protect against functional decline. Using targeted proteomics profiling, Niewczas *et al.* identified a kidney risk inflammatory signature consisting of 17 proteins enriched in tumor necrosis factor-receptor superfamily members using

plasma from cohorts with type 1 and type 2. The signature was associated with a 10-year risk of end-stage renal disease.¹⁰⁰ Using untargeted proteomic profiling, Kobayashi *et al.* measured 1,129 proteins by an affinity-based method in the serum of individuals from two cohorts, the Joslin Kidney Study and the Pima Indian Kidney Study. Eleven candidate proteins were identified. The biological process concerning these proteins were a neuron development-related process, organic acid binding, and cell-cell junction assembly.¹⁰¹ Md Dom and colleagues identified plasma proteins associated with protection against diabetic kidney disease progression in individuals with type 1 and type 2 diabetes.¹⁰² It is reported that three plasma proteins, fibroblast growth factor 20, angiopoietin-1, and tumor necrosis factor ligand super family member 12, were elevated.¹⁰² In 1,775 participants with type 2 diabetes, the combined score from the data of 273 urinary proteins measured by mass spectrometry has a high accuracy in predicting the new onset of albuminuria, after adjustments for HbA1c, blood pressure, baseline albuminuria, and eGFR.¹⁰³ Wu *et al.*¹⁰⁴ constructed machine learning-assisted models to identify a biomarker signature that predicts high risk participants with diabetes developing diabetic kidney disease based on clinical information, metabolomics, lipidomics, and single nucleotide polymorphism datasets. The biomarker signature of 10 features was useful

Table 1 | Findings by multi-omics analysis, highlighting gaps in current knowledge about diabetes and diabetic complications

Disease	Omics analysis	Current research	Findings	Future research needs
Diabetes	Genomics	GWAS	Insulin secretion Insulin resistance	Trans-ethnic larger analysis Whole genome sequence
Diabetic complications	Genomics	GWAS	Polygenic risk score for type 2 diabetes Increased inflammatory signals Increased oxidative stress	Identification of rare variant Replication in larger cohort Whole genome sequence
Diabetes	Proteomics	Targeted and non-targeted approaches LC-MS, GC-MS	Polygenic risk score for diabetic complications Increased inflammatory signals Altered lipoprotein metabolism	Identification of rare variant Replication in independent cohort Randomization study design
Diabetic complications	Proteomics	Affinity-based assay Targeted and non-targeted approaches LC-MS, GC-MS	Serum protein signatures for diabetes Increased inflammatory signals Serum and urine protein signatures for diabetic kidney disease	Longitudinal time point sample collection Replication in independent cohort Incorporation of tissue-based investigation Machine learning to select the target molecule
Diabetes	Metabolomics	Affinity-based assay Targeted and non-targeted approaches NMR, LC-MS, GC-MS	Increased BCAA synthesis Increased lipid synthesis	Replication in independent cohort Randomization study design Longitudinal time point sample collection
Diabetic complications	Metabolomics	Targeted and non-targeted approaches NMR, LC-MS, GC-MS	Increased amino acid synthesis Altered HDL particles in diabetic nephropathy	Replication in independent cohort Incorporation of tissue-based investigation Machine learning to select the target molecule

BCAA, branched-chain amino acid; GC-MS, mass spectrometry with gas chromatographic separation; GWAS, genome-wide association study; HDL, high density lipoprotein; LC-MS, mass spectrometry with liquid chromatographic separation; NMR, nuclear magnetic resonance.

when identifying subjects at high risk of renal impairment in type 2 diabetes (AUC (area under the curve) 0.76).¹⁰⁴

Concerning metabolomics analysis, many studies have been conducted and identified the associations between type 2 diabetes and branched-chain amino acid (BCAA), and aromatic amino acid.¹⁰⁵ BCAA was found to be associated with an increased risk of type 2 diabetes during a 12-year follow-up in the Malmo diet and Cancer study.¹⁰⁶ Blood metabolites related to amino acid metabolism and fatty acid metabolism have been reported to be associated with diabetic kidney disease.¹⁰⁵ High-density lipoprotein (HDL) particles and their composition have been associated with higher eGFR in type 2 diabetes.^{107,108} Yoshioka *et al.*¹⁰⁹ performed an untargeted metabolomic analysis of patients with stage G3 diabetic kidney disease and identified increased urinary lysophosphatidylcholine in fast functional decline. It has been reported that blood metabolites related to amino acid metabolism, and fatty acid metabolism were associated with diabetic retinopathy.^{110,111} These metabolites could serve as biomarkers to stratify individuals according to their risk of disease progression. However, the number of participants was small in these studies, which resulted in less consistent findings.

PROSPECTS FOR THE REALIZATION OF PRECISION MEDICINE

Genome-wide association studies have identified genetic predisposition in each stage of kidney damage with indicators of kidney function. Comprehensive analysis of genetic predisposition is shifting from GWAS centered on polymorphism analysis to whole-genome sequencing using next-generation sequencers (Table 1).^{112,113} In diabetes and diabetic complications, more variants specific to race are thought to be found in variants with an allele frequency of <5% and a low frequency of 1% or more. Because rare variants have a strong effect on the onset of disease, they are less frequent due to negative selection pressures, some of which are difficult to identify, and a larger number of cases is considered necessary.

Proteomics and metabolomics studies have been conducted and they have identified the association between type 2 diabetes and inflammatory signals and amino acid synthesis. For a comprehensive understanding of type 2 diabetes, replication in an independent cohort, a randomization study design, and longitudinal sample collection are needed. In addition, proteomics and metabolomics studies showed that inflammatory signals and amino acid synthesis are increased in diabetic complications. To elucidate the genetic predisposition of diabetic kidney disease, it is important to increase the number of cases to be analyzed and to obtain information on phenotypes with high accuracy. Wu *et al.*¹⁰⁴ developed machine learning-assisted models to identify a biomarker signature which was useful when identifying subjects at high risk of renal impairment in type 2 diabetes. Replication in independent cohort and the incorporation of tissue-based investigation, and utilization of machine learning to select the target molecule will elucidate the

mechanistic implications of biomarkers. It is expected that polygenic risk scores and biomarkers developed from multi-omics analysis will improve the predictability of the disease onset, and that the disease can be overcome by the development of new therapeutic drugs utilizing these data.^{114,115}

An atlas of disease risk genes has been created in 180,000 people from Biobank Japan, 360,000 from the UK Biobank, and 140,000 from FinnGen.¹¹⁶ The utilization of a gene atlas will enable the precise classification of diseases. In addition to the optimization and repositioning of existing treatments, new drug discovery such as Adiporon, an agonist of the adiponectin receptor that is greatly involved in the pathophysiology of type 2 diabetes and diabetic complications, is important for the realization of genomic medicine suitable for the lifestyle of individuals.¹¹⁷

CONCLUSION

Many genetic loci related to diabetes and diabetic complications have been discovered, and it is expected that the mechanism of disease onset will be elucidated and its implementation in medical care will progress by utilizing the polygenic risk score and multi-omics analysis. For this challenge to succeed and sustain, it is important to communicate with society.¹¹⁸ Ethical considerations should be addressed so that participants' right to know and not know are adequately protected, and genomic research is truly beneficial.

AUTHOR CONTRIBUTIONS

N.S., and T.Y., wrote the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

Type 2 diabetes GWAS in Japanese population is the result of the kindness and efforts of many collaborators. Dr Ken Suzuki (Manchester University, The University of Tokyo, Osaka University), Prof. Masato Akiyama (Kyushu University), Prof. Yoichiro Kamatani (The University of Tokyo), Dr Momoko Horikoshi (RIKEN), Prof. Yukinori Okada (Osaka University, The University of Tokyo), Dr Takashi Kadowaki (Toranomon Hospital, The University of Tokyo) gave us tremendous guidance, and we were able to obtain genome data from Biobank Japan, Tohoku Medical Megabank Organization, Iwate Tohoku Medical Megabank Organization, and Japan multi-purpose research cohort. In type 2 diabetes GWAS in East Asian population, Dr Momoko Horikoshi became the head of the Japan research group and made great efforts. We would like to take this opportunity to express our sincere gratitude to all our collaborators. This work was supported by grants from the Japan Agency for Medical Research and Development (JP21t-m0424218).

DISCLOSURE

The authors declare no conflict of interest.
Approval of the research protocol: N/A.

Informed consent: N/A.

Registry and the registration no. of the study/trial: N/A.

Animal studies: N/A.

REFERENCES

- World Health Organization. The top 10 causes of death, 2020. Available from: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-cause-of-death> Accessed December 9, 2020.
- Sabanayagam C, Chee ML, Banu R, *et al.* Association of diabetic retinopathy and diabetic kidney disease with all-cause and cardiovascular mortality in a multiethnic Asian population. *JAMA Netw Open* 2019; 2: e191540.
- Newman B, Selby JV, King MC, *et al.* Concordance for type 2 (non-insulin dependent) diabetes mellitus in male twins. *Diabetologia* 1987; 30: 763–768.
- Cole JB, Florenz JC. Genetics of diabetes and diabetic complications. *Nat Rev Nephrol* 2020; 16: 377–390.
- Sandholm N, Groop PH. Genetic basis of diabetic kidney disease and other diabetic complications. *Curr Opin Genet Dev* 2018; 50: 17–24.
- Klein RJ, Zeiss C, Chew EY, *et al.* Complement factor H polymorphism in aggregated macular degeneration. *Science* 2005; 308: 385–389.
- Sladek R, Rocheleau G, Rung J, *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; 445: 881–885.
- Saxena R, Voight BF, Lyssenko V, *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; 316: 1331–1336.
- Steinthorsdottir V, Thorleifsson G, Reynisdottir I, *et al.* A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 2007; 39: 770–775.
- Yasuda K, Miyake K, Horikawa Y, *et al.* Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nat Genet* 2008; 40: 1092–1097.
- Unoki H, Takahashi A, Kawaguchi T, *et al.* SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat Genet* 2008; 40: 1098–1102.
- Zeggini E, Scott LJ, Saxena R, *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008; 40: 638–645.
- Yamauchi T, Hara K, Maeda S, *et al.* A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at UBE2E2 and C2CD4A-C2CD4B. *Nat Genet* 2010; 42: 864–868.
- Voight BF, Scott LJ, Steinthorsdottir V, *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 2010; 42: 579–589.
- Tsai FJ, Yang CF, Chen CC, *et al.* A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. *PLoS Genet* 2010; 6: e1000847.
- Shu XO, Long J, Cai Q, *et al.* Identification of new genetic risk variants for type 2 diabetes. *PLoS Genet* 2010; 6: e1001127.
- 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. *Nat Genet* 2011; 44: 67–72.
- 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. *Nat Genet* 2011; 43: 984–989.
- Imamura M, Maeda S, Yamauchi T, *et al.* A single-nucleotide polymorphism in ANK1 is associated with susceptibility to type 2 diabetes in Japanese populations. *Hum Mol Genet* 2012; 21: 3042–3049.
- Morris AP, Voight BF, Teslovich TM, *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012; 44: 981–990.
- Palmer ND, McDonough CW, Hicks PJ, *et al.* A genome-wide association search for type 2 diabetes genes in African Americans. *PLoS One* 2012; 7: e29202.
- Bonnefond A, Clément N, Fawcett K, *et al.* Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nat Genet* 2012; 44: 297–301.
- Ma RC, Hu C, Tam CH, *et al.* Genome-wide association study in a Chinese population identifies a susceptibility locus for type 2 diabetes at 7q32 near PAX4. *Diabetologia* 2013; 56: 1291–1305.
- Saxena R, Saleheen D, Been LF, *et al.* Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in Sikhs of Punjabi origin from India. *Diabetes* 2013; 62: 1746–1755.
- Hara K, Fujita H, Johnson TA, *et al.* Genome-wide association study identifies three novel loci for type 2 diabetes. *Hum Mol Genet* 2014; 23: 239–246.
- Mahajan A, Go MJ, Zhang W, *et al.* Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet* 2014; 46: 234–244.
- 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: 97–101.
- 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. *Nat Genet* 2014; 46: 294–298.
- Flannick J, Thorleifsson G, Beer NL, *et al.* Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat Genet* 2014; 46: 357–363.
- Estrada K, Aukrust I, Bjørkhaug L, *et al.* Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. *JAMA* 2014; 311: 2305–2314.

31. Imamura M, Takahashi A, Yamauchi T, *et al.* Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. *Nat Commun* 2016; 7: 10531.
32. Scott RA, Scott LJ, Mägi R, *et al.* An expanded genome-wide association study of type 2 diabetes in Europeans. *Diabetes* 2017; 66: 2888–2902.
33. Mahajan A, Taliun D, Thurner M, *et al.* Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018; 50: 1505–1513.
34. Vujkovic M, Keaton JM, Lynch JA, *et al.* Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nat Genet* 2020; 52: 680–691.
35. Spracklen CN, Horikoshi M, Kim YJ, *et al.* Identification of type 2 diabetes loci in 433,540 east Asian individuals. *Nature* 2020; 582: 240–245.
36. Suzuki K, Akiyama M, Ishigaki K, *et al.* Identification of 28 new susceptibility loci for type 2 diabetes in the Japanese population. *Nat Genet* 2019; 51: 379–386.
37. Loos RJF, Yeo GS. The genetics of obesity: from discovery to biology. *Nat Rev Genet* 2022; 23: 120–133.
38. Kodama K, Tojjar D, Yamada S, *et al.* Ethnic differences in the relationship between insulin sensitivity and insulin response. *Diabetes Care* 2013; 36: 1789–1796.
39. Ke C, Narayan KVM, Chan JCN, *et al.* Pathophysiology, phenotypes and management of type 2 diabetes mellitus in Indian and Chinese populations. *Nat Rev Endocrinol* 2022; 18: 413–432.
40. Grant SF, Thorleifsson G, Reynisdottir I, *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 2006; 38: 320–323.
41. Scott LJ, Mohlke KL, Bonnycastle LL, *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007; 316: 1341–1345.
42. Zeggini E, Weedon MN, Lindgren CM, *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 2007; 316: 1336–1341.
43. Wang Q, Curren ME, Splawski I, *et al.* Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nat Genet* 1996; 12: 17–23.
44. Neyroud N, Tesson F, Denjoy I, *et al.* A novel mutation in the potassium channel gene KVLQT1 causes the Jervell and Lange-Nielsen cardioauditory syndrome. *Nat Genet* 1997; 15: 186–189.
45. López-Avalos MD, Duvivier-Kali VF, Xu G, *et al.* Evidence for a role of the ubiquitin-proteasome pathway in pancreatic islets. *Diabetes* 2006; 55: 1223–1231.
46. Sathananthan A, Man CD, Micheletto F, *et al.* Common genetic variation in GLP1R and insulin secretion in response to exogenous GLP-1 in nondiabetic subjects: a pilot study. *Diabetes Care* 2010; 33: 2074–2076.
47. Okada Y, Momozawa Y, Sakaue S, *et al.* Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese. *Nat Commun* 2018; 9: 1631.
48. Yasumizu Y, Sakaue S, Konuma T, *et al.* Genome-wide natural selection signatures are linked to genetic risk of modern phenotypes in the Japanese population. *Mol Biol Evol* 2020; 37: 1306–1316.
49. Harjutsalo V, Katoh S, Sarti C, *et al.* Population-based assessment of familial clustering of diabetic nephropathy in type 1 diabetes. *Diabetes* 2004; 53: 2449–2454.
50. Iyengar SK, Abboud HE, Goddard KA, *et al.* Genome-wide scans for diabetic nephropathy and albuminuria in multiethnic populations. *Diabetes* 2007; 56: 1577–1585.
51. Azushima K, Gurley SB, Coffman TM, *et al.* Modelling diabetic nephropathy in mice. *Nat Rev Nephrol* 2018; 14: 48–56.
52. Maeda S, Kobayashi M, Araki S, *et al.* A single nucleotide polymorphism within the acetyl-coenzyme A carboxylase beta gene is associated with proteinuria in patients with type 2 diabetes. *PLoS Genet* 2010; 6: e1000842.
53. Sandholm N, Salem RM, McKnight AJ, *et al.* New susceptibility loci associated with kidney disease in type 1 diabetes. *PLoS Genet* 2012; 8: e1002921.
54. Sandholm N, McKnight AJ, Salem RM, *et al.* Chromosome 2q31.1 associate with ESRD in women with type 1 diabetes. *J Am Soc Nephrol* 2013; 24: 1537–1543.
55. Sandholm N, Forsblom C, Mäkinen VP, *et al.* Genome-wide association study of urinary albumin excretion rate in patients with type 1 diabetes. *Diabetologia* 2014; 57: 1143–1153.
56. Li C, Liu C, Nissim I, *et al.* Regulation of glucagon secretion in normal and diabetic human islets by gamma-hydroxybutyrate and glycine. *J Biol Chem* 2013; 288: 3938–3951.
57. Iyengar SK, Sedor JR, Freedman BI, *et al.* Genome-wide association and transethnic meta-analysis for advanced diabetic kidney disease: family investigation of nephropathy and diabetes (FIND). *PLoS Genet* 2015; 11: e1005352.
58. Pattaro C, Teumar A, Gorski M, *et al.* Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat Commun* 2016; 7: 10023.
59. Taira M, Imamura M, Takahashi A, *et al.* A variant within the FTO confers susceptibility to diabetic nephropathy in Japanese patients with type 2 diabetes. *PLoS One* 2018; 13: e0208654.
60. Smemo S, Tena JJ, Gamazon ER, *et al.* Obesity-associated variants with FTO form long-range functional connections with IRX3. *Nature* 2014; 507: 371–375.
61. van Zuydam NR, Ahlqvist E, Samdholm N, *et al.* A genome-wide association study of diabetic kidney disease in subjects with type 2 diabetes. *Diabetes* 2018; 67: 1414–1427.

62. Guan M, Keaton JM, Dimitrov L, *et al.* Genome-wide association study identifies novel loci for type 2 diabetes-attributed end-stage kidney disease in African Americans. *Hum Genomics* 2019; 13: 21.
63. Genoveseg FDJ, Ross MD, *et al.* Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 2010; 329: 841–845.
64. Zhang P, Chen Y, Cheng Y, *et al.* Alkaline sphingomyelinase (NPP7) promotes cholesterol absorption by affecting sphingomyelin levels in the gut: a study with NPP7 knockout mice. *Am J Physiol Gastrointest Liver Physiol* 2014; 306: G906–G908.
65. Ceriello A, De Cosmo S, Rossi MC, *et al.* Variability in HbA1c, blood pressure, lipid parameters and serum uric acid and risk of development of chronic kidney disease in type 2 diabetes. *Diabetes Obes Metab* 2017; 19: 1570–1578.
66. Salem RM, Todd JN, Sandholm N, *et al.* Genome-wide association study of diabetic kidney disease highlights biology involved in glomerular basement membrane collagen. *J Am Soc Nephrol* 2019; 30: 2000–2016.
67. Sandholm N, Cole JB, Nair V, *et al.* Genome-wide meta-analysis and omics integration identifies novel gene associated with diabetic kidney disease. *Diabetologia* 2022; 65: 1495–1509.
68. Winkler T, Humaira R, Teumer A, *et al.* Differential and shared genetic effects on kidney function between diabetic and non-diabetic individuals. *Commun Biol* 2023; 5: 580.
69. Canela-Xandri O, Rawlik K, Tenesa A, *et al.* An atlas of genetic associations in UK biobank. *Nat Genet* 2008; 50: 1593–1599.
70. Teo ZL, Tham YC, Yu M, *et al.* Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. *Ophthalmology* 2021; 128: 1580–1591.
71. Thomas R, Halim S, Gurudas S, *et al.* IDF diabetes atlas: a review of studies utilizing retinal photography on the global prevalence of diabetes related retinopathy between 2015 and 2018. *Diab Res Clin Pract* 2019; 157: 10784.
72. Fu YP, Hallman DM, Gonzalez VH, *et al.* Identification for diabetic retinopathy genes through a genome-wide association study among Mexican-Americans from Starr County, Texas. *J Ophthalmol* 2009; 2010: 861291.
73. Huang YC, Lin JM, Lin HJ, *et al.* Genome-wide association study of diabetic retinopathy in a Taiwanese population. *Ophthalmology* 2011; 118: 642–648.
74. Grassi MA, Tikhomirov A, Ramalingam S, *et al.* Genome-wide meta-analysis for severe diabetic retinopathy. *Hum Mol Genet* 2011; 20: 2472–2481.
75. Burden KP, Fogarty RD, Shen W, *et al.* Genome-wide association study for sight-threatening diabetic retinopathy reveals association with genetic variation near the GRB2 gene. *Diabetologia* 2015; 58: 2288–2297.
76. Meng W, Shah KP, Pollack S, *et al.* A genome-wide association study suggests new evidence for an association of the NADPH oxidase 4 (NOX4) gene with severe diabetic retinopathy in type 2 diabetes. *Acta Ophthalmol* 2018; 96: e811–e819.
77. Pollack S, Igo RP, Jensen RA, *et al.* Multiethnic genome-wide association study of diabetic retinopathy using liability threshold modeling of duration of diabetes and glycemic control. *Diabetes* 2019; 68: 441–456.
78. Mackey DS, Bennett TM, Culican SM, *et al.* Exome sequencing identifies novel and recurrent mutations in GJA8 and CRYGD associated with inherited cataract. *Hum Genomics* 2014; 8: 19.
79. Devi RR, Vijayalakshmi P. Novel mutation in GJA8 associated with autosomal dominant congenital cataract and microcornea. *Mol Vis* 2006; 12: 190–195.
80. Imamura M, Takahashi A, Matsunami M, *et al.* Genome-wide association studies identify two novel conferring susceptibility to diabetic retinopathy in Japanese patients with type 2 diabetes. *Hum Mol Genet* 2021; 30: 716–726.
81. Armstrong DG, Boulton AJ, Bus SA, *et al.* Diabetic foot ulcers and their recurrence. *N Engl J Med* 2017; 376: 2367–2375.
82. Wu S, Han Y, Hu Q, *et al.* Effects of common polymorphisms in the MTHFR and ACE genes of diabetic peripheral neuropathy progression: a meta-analysis. *Mol Neurobiol* 2017; 54: 2435–2444.
83. Meng W, Veluchamy A, Hebert HL, *et al.* A genome-wide association study suggests that MAPK14 is associated with diabetic foot ulcer. *Br J Dermatol* 2017; 177: 1664–1670.
84. Tang Y, Lenzini PA, Pop-Busui R, *et al.* A genetic locus on chromosome 2q24 predicting peripheral neuropathy risk in type 2 diabetes: results from the ACCORD and BARI 2D studies. *Diabetes* 2019; 68: 1649–1662.
85. Sloan TJ, Turton JC, Tyson J, *et al.* Examining diabetic heel ulcers through an ecological lens: microbial community dynamics associated with healing and infection. *J Med Microbiol* 2019; 68: 230–240.
86. Kalan LR, Meisel JS, Loesche MA, *et al.* Strain- and species-level variation in the microbiome of diabetic wounds is associated with clinical outcomes and therapeutic efficacy. *Cell Host Microbe* 2019; 25: 641–655.
87. Kullo IJ, Lewis CM, Inouye M, *et al.* Polygenic scores in biomedical research. *Nat Genet* 2022; 223: 524–530.
88. Wong CK, Makalic E, Dite GS, *et al.* Polygenic risk scores for cardiovascular diseases and type 2 diabetes. *PLoS One* 2022; 17: e0278764.
89. Khera AV, Chafn M, Aragam KG, *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018; 50: 1219–1224.
90. Martin AR, Kanai M, Kamatani Y, *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 2019; 51: 584–591.

91. Flannick J, Florez JC. Type 2 diabetes: genetic data sharing to advance complex disease research. *Nat Rev Genet* 2016; 17: 535–549.
92. Udler MS, McCarthy MI, Florez JC. Genetic risk scores for diabetes diagnosis and precision medicine. *Endocr Rev* 2019; 40: 1500–1520.
93. Mariam A, Miller-Atkins G, Pantalone KM, *et al.* A type 2 diabetes subtype responsive to ACCORD intensive glycemia treatment. *Diabetes Care* 2021; 44: 1410–1418.
94. Tremblay J, Haloui M, Attaoua R, *et al.* Polygenic risk scores predict diabetes complications and their response to intensive blood pressure and glucose control. *Diabetologia* 2021; 64: 2012–2025.
95. Haring R, Wallaschofski H. Diving through the omics the case for deep phenotyping and systems epidemiology. *OMICS* 2012; 16: 231–234.
96. Hu C, Jia W. Multi-omics profiling: the way towards precision medicine in metabolic diseases. *J Mol Cell Biol* 2021; 13: 576–593.
97. Suhre K, McCarthy MI, Schwenk JM. Genetics meets proteomics: perspectives for large population-based studies. *Nat Rev Genet* 2021; 22: 9–37.
98. Gudmundsdottir V, Zaghlool SB, Emilsson V, *et al.* Circulating protein signatures and causal candidates for type 2 diabetes. *Diabetes* 2020; 69: 1843–1853.
99. von Toerne C, Huth C, de Las Heras Gala T, *et al.* MASP1, THBS1, GPLD1 and ApoA-IV are novel biomarkers associated with prediabetes: the KORA F4 study. *Diabetologia* 2016; 59: 1882–1892.
100. Niewczasz MA, Pavkov ME, Skupien J, *et al.* A signature of circulating inflammatory proteins and development of end-stage renal disease in diabetes. *Nat Med* 2019; 25: 805–813.
101. Kobayashi H, Looker HC, Satake E, *et al.* Results of untargeted analysis using the SOMAscan proteomics platform indicates novel association to kidney failure in diabetes. *Kidney Int* 2022; 102: 370–381.
102. Md Dom ZI, Satake E, Skupien J, *et al.* Circulating proteins protect against renal decline and progression to end-stage renal disease in patients with diabetes. *Sci Transl Med* 2021; 13: eabd2699.
103. Tofte N, Lindhardt M, Adamova K, *et al.* Early detection of diabetic kidney disease by urinary proteomics and subsequent intervention with spironolactone to delay progression (PRIORITY): a prospective observational study and embedded randomised placebo-controlled trial. *Lancet Diabetes Endocrinol* 2020; 8: 301–312.
104. Wu IW, Tsai TH, Lo CJ, *et al.* Discovering a trans-omics biomarker signature that predisposes high risk diabetic patients to diabetic kidney disease. *NPJ Digit Med* 2020; 5: 166.
105. Wang TJ, Larson MG, Vasan RS, *et al.* Metabolite profiles and the risk of developing diabetes. *Nat Med* 2011; 17: 448–453.
106. Sun Y, Gao HY, Fan ZY, *et al.* Metabolomics signatures in type 2 diabetes: a systematic review and integrative analysis. *J Clin Endocrinol Metab* 2020; 105: 1000–1008.
107. Darshi M, Van Espen B, Sharma K. Metabolomics in diabetic kidney disease: unraveling the biochemistry of a silent killer. *Am J Nephrol* 2016; 44: 92–103.
108. Barrios C, Zierer J, Wurtz P, *et al.* Circulating metabolic biomarkers of renal function in diabetic and non-diabetic populations. *Sci Rep* 2018; 8: 15249.
109. Yoshioka K, Hirakawa Y, Kurano M, *et al.* Lysophosphatidylcholine mediates fast decline in kidney function in diabetic kidney disease. *Kidney Int* 2022; 101: 510–526.
110. Chen L, Cheng CY, Choi H, *et al.* Plasma metabolomic profiling of diabetic retinopathy. *Diabetes* 2016; 65: 1099–1108.
111. Yun JH, Kim JM, Jeon HY, *et al.* Metabolomics profiles associated with diabetic retinopathy in type 2 diabetes patients. *PLoS One* 2020; 15: e0241365.
112. Backman JD, Li AH, Marcketta A, *et al.* Exome sequencing and analysis of 454,787 UK biobank participants. *Nature* 2021; 599: 628–634.
113. Bonnefond A, Froguel P. Rare and common genetic events in type 2 diabetes: what should biologists know? *Cell Metab* 2015; 21: 357–368.
114. Namba S, Konuma T, Wu K, *et al.* A practical guideline of genomics-driven drug discovery in the era of global biobank meta-analysis. *Cell Genomics* 2022; 2: 100190.
115. Iida M, Iwata M, Yamanishi Y, *et al.* Network-based characterization of disease–disease relationships in terms of drugs and therapeutic targets. *Bioinformatics* 2020; 36: i516–i524.
116. Sakaue S, Kanai M, Tanigawa Y, *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet* 2021; 53: 1415–1424.
117. Okada-Iwabu M, Yamauchi T, Iwabu M, *et al.* A small-molecule AdipoR agonist for type 2 diabetes and short life in obesity. *Nature* 2013; 503: 493–499.
118. Polygenic Risk Score Task Force of the International Common Disease Alliance. Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. *Nat Med* 2021; 27: 1876–1884.