

## Common and rare forms of diabetes mellitus: towards a continuum of diabetes subtypes

Jason Flannick<sup>1,2</sup>, Stefan Johansson<sup>3,4</sup> and Pål R. Njølstad<sup>3,5</sup>

Abstract | Insights into the genetic basis of type 2 diabetes mellitus (T2DM) have been difficult to discern, despite substantial research. More is known about rare forms of diabetes mellitus, several of which share clinical and genetic features with the common form of T2DM. In this Review, we discuss the extent to which the study of rare and low-frequency mutations in large populations has begun to bridge the gap between rare and common forms of diabetes mellitus.

We hypothesize that the perceived division between these diseases might be due, in part, to the historical ascertainment bias of genetic studies, rather than a clear distinction between disease pathophysiologies. We also discuss possible implications of a new model for the genetic basis of diabetes mellitus subtypes, where the boundary between subtypes becomes blurred.

### Sulfonylureas

A class of antidiabetic medication that increases insulin secretion by blockage of the ATP-sensitive potassium channel in the pancreas.

<sup>1</sup>Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 415 Main Street, Cambridge, Massachusetts 02142, USA.

<sup>2</sup>Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, USA.

<sup>3</sup>K.G. Jebsen Center for Diabetes Research, The Department of Clinical Science, University of Bergen, Jonas Lies veg 87, N-5020 Bergen, Norway.

<sup>4</sup>Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Jonas Lies veg 65, N-5021 Bergen, Norway.

<sup>5</sup>Department of Pediatrics, Haukeland University Hospital, Jonas Lies veg 65, N-5021 Bergen, Norway.

Correspondence to P. R. N. [pal.njolstad@uib.no](mailto:pal.njolstad@uib.no)

doi:10.1038/nrendo.2016.50  
Published online 15 Apr 2016

Type 2 diabetes mellitus (T2DM) affects ~400 million people worldwide<sup>1</sup> and imposes a substantial public health burden owing to a range of diabetic complications and the inadequacy of current therapies. As T2DM is a heritable disorder, the possibility of studying human genetics to provide insight into disease biology, treatment and prevention has received much attention. However, the large environmental component of T2DM susceptibility poses challenges for identifying informative mutations.

In contrast to T2DM, much more is known regarding the pathophysiology of other forms of diabetes mellitus. Type 1 diabetes mellitus (T1DM) is caused by autoimmune-mediated destruction of  $\beta$  cells<sup>2</sup>, a pathogenesis (presumably) distinct from that of T2DM. However, many rarer forms of diabetes mellitus do share clinical similarities with T2DM, including maturity-onset diabetes of the young (MODY)<sup>3</sup>, neonatal diabetes mellitus (NDM)<sup>4</sup>, mitochondrial diabetes mellitus<sup>5</sup> and certain multiorgan syndromes such as Wolcott–Rallison syndrome and Wolfram syndrome<sup>6</sup>. For many of these conditions, the genetic causes are known and have led, in some cases, to diagnostic screening tools or effective treatments<sup>7</sup>.

As diagnoses of diabetes mellitus are based on clinical presentation rather than specific molecular defect(s), a longstanding hypothesis has been that the genetic aetiologies of common and rare forms of diabetes mellitus might not be wholly distinct<sup>8–10</sup>. An obvious question is, thus: to what extent can knowledge about rare forms of diabetes mellitus inform our understanding of T2DM and vice versa? Here, we discuss how genetic studies have begun to elucidate the links between these

disorders. We first review the clinical presentations of these conditions and classic approaches to dissect their genetic bases. We then discuss studies that combine features of these traditional approaches: studies of rare variants (traditionally analysed for rare diseases) for association with common T2DM; and studies of large populations (traditionally analysed for common diseases) for mutations causing rare forms of diabetes mellitus. In our view, genetic evidence now supports a unified risk model for rare and common forms of diabetes mellitus; we explore the consequences for future research, diagnosis and treatment.

### Clinical presentation

#### T2DM

T2DM presents as chronic hyperglycaemia, typically after 40 years of age, and is thought to result from the progressive failure of pancreatic  $\beta$  cells to secrete sufficient insulin to meet metabolic demands<sup>11</sup>. Obesity and age are major risk factors for T2DM, as they result in insulin resistance that typically precedes diabetes mellitus<sup>12</sup>. Most drugs for T2DM, such as insulin, metformin and sulfonylureas, are primarily intended to lower blood glucose levels, but have, thus far, been unable to change the progressive course of T2DM or adequately reduce the risk of late-stage complications associated with diabetes mellitus<sup>13</sup>.

#### Rare forms of diabetes mellitus

A number of much rarer, monogenic conditions can also result in nonautoimmune diabetes mellitus (FIG. 1). The most notable, MODY, is characterized by

**Key points**

- The common disease of type 2 diabetes mellitus (T2DM) is genetically distinct from, but clinically similar to, rarer forms of diabetes mellitus
- Today, understanding of the pathophysiology of rare forms of diabetes mellitus is greater than that of the common form of T2DM
- As genetic studies examine rarer variants in large populations, higher penetrance variants have been found for T2DM even as previously 'Mendelian' variants have been found to have lower penetrances
- The increasing overlap between regions associated with T2DM and genes or variants relevant to monogenic diseases suggests an overlapping disease aetiology, possibly including distinct T2DM subtypes
- We propose that a unified diabetes mellitus risk model, spanning variants of all frequencies and penetrances, might provide future insights into understanding or treatment of all forms of diabetes mellitus

autosomal-dominant inheritance and nonketotic hyperglycaemia due to  $\beta$ -cell failure; typical onset is between 6 months and 35 years of age<sup>7</sup>. The Online Mendelian Inheritance in Man (OMIM) database currently lists 14 MODY subtypes, of which four account for 80–90% of diagnosed cases<sup>14</sup>. Patients with MODY have different disease trajectories dependent on subtype; for example, patients with GCK–MODY have only mild fasting hyperglycaemia with minimal late-stage complications<sup>15</sup>, whereas patients with HNF1A–MODY or HNF4A–MODY develop progressive  $\beta$ -cell dysfunction and microvascular complications<sup>16</sup>. Age at onset can also vary by subtype; for example, GCK–MODY can present at any age, whereas HNF1A–MODY can present into the fifth decade of life<sup>17</sup>.

Three other rare classes of diabetes mellitus are NDM, syndromic diabetes mellitus and mitochondrial diabetes mellitus. NDM typically presents before 6 months of age and can be either transient (TNDM) or permanent (PNDM); patients are characterized by intrauterine growth restriction, low birth weight and  $\beta$ -cell failure<sup>7</sup>. Syndromic diabetes mellitus encompasses a class of diseases in which features other than diabetes mellitus dominate the clinical phenotype, such as Wolfram syndrome, which is characterized by optic atrophy, deafness and neurological abnormalities<sup>18</sup>, or syndromes of severe insulin resistance that include insulin signalling defects and lipodystrophies<sup>19</sup>; some forms of MODY or NDM could also be classified as syndromic diabetes mellitus, such as HNF1B–MODY with kidney and pancreatic malformations<sup>20</sup> or CEL–MODY with exocrine pancreas dysfunction<sup>21</sup>. Finally, mitochondrial diabetes mellitus, which exhibits maternal inheritance, gradual  $\beta$ -cell dysfunction and sensorineural hearing loss, has been discussed as a potential additional form of monogenic diabetes mellitus<sup>5</sup>.

### Genetic studies

#### Rare forms of diabetes mellitus

Compared with T2DM, rare forms of diabetes mellitus are distinguished by causal mutations within a single gene and are, thus, often called 'monogenic diabetes mellitus'. These forms exhibit Mendelian inheritance patterns, with large effect sizes of causal variants and

minimal environmental contributions (BOX 1). To identify variants for monogenic diseases, linkage mapping within large families with a history of diabetes mellitus, followed by sequencing to localize signals to individual genes, has been the dominant approach for several decades. More than 80 disease genes, most of which have been thoroughly reviewed elsewhere<sup>7,22,23</sup>, have been identified via this approach (FIG. 2).

In the past 5 years, next-generation sequencing (NGS) has provided an alternative strategy to identify monogenic disease mutations (BOX 2). NGS has been used to successfully screen panels of known genes in many diagnostic laboratories, demonstrating a sensitivity >90%<sup>24–26</sup> and in some cases identifying putative disease variants in genes that would otherwise not have been tested<sup>25,27</sup>. Whole-exome sequencing (WES) approaches, which sequence all protein-coding regions in patient genomes<sup>28</sup>, have also successfully identified causal variants in known genes<sup>29,30</sup>. Novel genes have been comparatively less frequently identified via such approaches, although successes have been observed for *GATA6* and pancreatic agenesis<sup>31</sup>; *KCNJ11* (REF. 32) and *APPL1* (REF. 33) for MODY; and *CDKN1C*<sup>34</sup> and *PRPS1* (REF. 35) for syndromic forms of diabetes mellitus characterized by intrauterine growth restriction.

Many of our insights into the disease mechanisms of diabetes mellitus are due to knowledge of genes associated with monogenic forms. The  $K_{ATP}$  channel, which has an essential role in glucose-stimulated insulin secretion, was shown to have inactivating mutations, which cause hyperinsulinaemia and hypoglycaemia<sup>36</sup>, as well as activating mutations, which can variously cause NDM<sup>37</sup> or MODY<sup>32</sup>. In the same pathway, the role of glucokinase (GCK) as a glucose sensor was confirmed when homozygous mutations, which cause complete GCK deficiency, were shown to cause NDM due to impaired glucose-dependent ATP production and inhibition of  $K_{ATP}$  channel closure<sup>38</sup>. In other pathways, the identification of MODY mutations in genes encoding liver-enriched transcription factors active during embryogenesis (*HNF1A*, *HNF4A* and *HNF1B*) clarified observations that several MODY subtypes are in fact early-onset multiorgan diseases<sup>39–41</sup>, whereas the identification of MODY mutations in *CEL* (expressed in pancreatic acinar cells) showed that cell types other than the  $\beta$  cell might be implicated in diabetes mellitus<sup>42</sup>.

This improved understanding of the aetiology and pathogenesis of monogenic diabetes mellitus has, in many cases, led to improved patient care. Many countries have introduced screening programs for all children diagnosed with diabetes mellitus before 6 months of age<sup>43</sup>, as a diagnosis of activating  $K_{ATP}$  channel mutations can usually enable patients to safely discontinue insulin and start high-dose sulfonylureas<sup>44,45</sup>. Patients diagnosed with a mutation in *HNF1A* or *HNF4A* are similarly highly sensitive to sulfonylurea treatment<sup>46</sup>. Patients who carry a mutation in *GCK*, who further display mild fasting hyperglycaemia (5.4–8.3 mmol/l, HbA<sub>1c</sub> range 5.8–7.6%) from birth but with mild deterioration with ageing<sup>47</sup>, do not require any treatment at all, as the increased fasting glucose

**Mendelian**

A trait whose inheritance pattern obeys the laws proposed by Gregor Mendel; characteristic of a disease caused by mutations in a single gene.

**Linkage**

An approach for identifying disease genes by testing for co-segregation between DNA segments and phenotype in a familial study design.

**Next-generation sequencing (NGS).** A family of technologies for relatively inexpensive and high-throughput sequencing of DNA.

**Genome-wide association studies (GWAS).** An approach for identifying disease genes by testing many common genetic variants in different individuals for associations with a trait.

**Expression quantitative trait loci (eQTL).** A genomic region that is associated with variation in expression levels of mRNA.

‘set-point’ for GCK–MODY is seldom sufficient to cause microvascular complications<sup>15</sup>. Insulin therapy is usually required for other conditions, however, such as HNF1B–MODY or the majority of mitochondrial diabetes mellitus conditions.

**Common variants for T2DM**

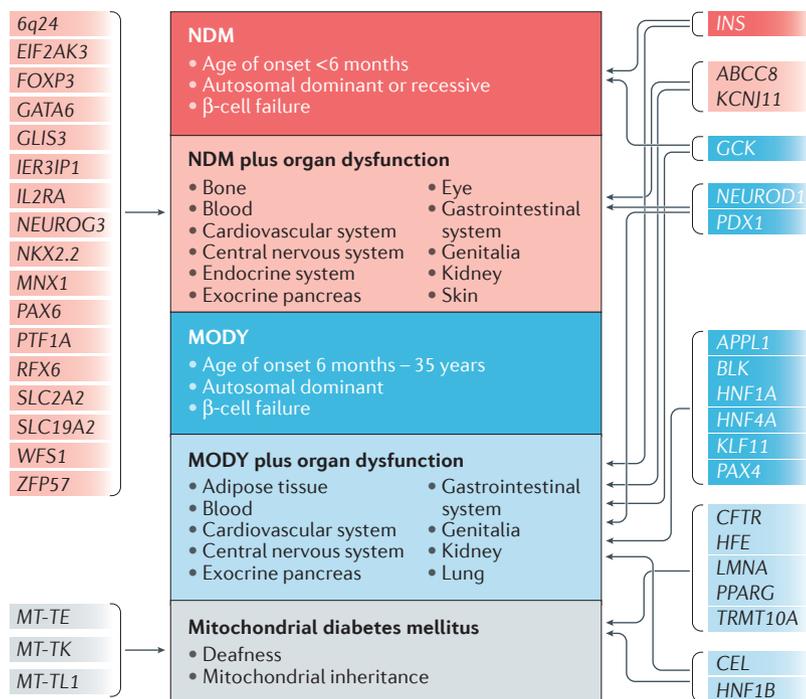
Unlike monogenic diabetes mellitus, T2DM lacks a clear inheritance pattern, being characterized instead by familial aggregation; for example, a twofold to threefold sibling relative risk and a heritability of 30–70%<sup>48,49</sup>. A substantial amount of T2DM risk is environmental, which leads to much smaller genetic effect sizes than those for monogenic forms (BOX 1). Consequently, linkage scans have yielded few signals for T2DM<sup>50</sup>.

By contrast, genome-wide association studies (GWAS) of common variants have proven far more successful than linkage studies for mapping T2DM loci (the major findings are reviewed elsewhere<sup>51,52</sup>). In brief, sample sizes now exceed 100,000 (REFS 53,54), spanning multiple ethnicities and identifying around 100 genomic loci (FIG. 2). All variants affect risk modestly (with relative risks usually around 1.1–1.2 (REF. 53)) and explain a significant but small amount (10–15%) of T2DM heritability. These findings are increasingly interpreted as consistent with a polygenic model for the genetic basis

of T2DM<sup>55,56</sup>, which is characterized by possibly thousands of distinct genetic risk factors<sup>53</sup>. Most associations map to noncoding regions of the genome, preferentially acting through regulatory elements such as pancreatic islet enhancers<sup>57–59</sup>.

As GWAS associations implicate genomic regions that span hundreds of kilobases, less is known about the causal genes or variants responsible for these associations. Efforts to localize signals to individual variants have had limited success, yielding ‘credible sets’ of candidate variants that are sometimes manageable but usually ambiguous<sup>54,60</sup>. Additional data and experiments have in some cases suggested promising genes or variants at these loci. For the *FTO* locus, for example, a series of studies used cross-species conservation<sup>61</sup> or chromatin conformation capture<sup>62,63</sup> to link the GWAS association to the distant *IRX3* transcript; a later comprehensive investigation of the same locus used epigenomics, transcript expression, motif conservation and network analysis to suggest an *IRX3*-mediated regulatory mechanism underlying adipocyte thermogenesis that was subsequently validated by directed perturbations, mouse models and endogenous CRISPR–Cas9 genome editing<sup>62</sup>. Parent-of-origin effects and differential imprinting patterns have also been used to gain insight into effector transcripts (for the *KLF14* locus<sup>64</sup>), the stage of development that mediates disease risk (for the *KCNQ1* locus<sup>65</sup>) or tissue-specific effects (for the *GRB10* locus<sup>66</sup>). Additional insights into causal genes or variants have come from cross-species comparisons at the *PPARG* locus<sup>67</sup>, overlap with chromatin immunoprecipitation binding sites at the *MTNR1B* locus<sup>68</sup>, or expression quantitative trait loci (eQTL) mapping at the *ZMIZ1* locus<sup>69</sup>. For most loci, however, the causal genes or biological mechanism remain to be elucidated.

For a subset of common variant signals, a gene implicated in monogenic diabetes mellitus lies nearby and is, therefore, a strong candidate for causality. For example, an early candidate gene study of *HNF1A* identified a common variant private to the Canadian Oji-Cree population with a moderate (twofold) effect on T2DM risk<sup>70</sup>; additional common variant associations for T2DM were identified from candidate gene studies of *PPARG*<sup>71</sup>, a gene implicated in a severe form of insulin-resistant diabetes mellitus and lipodystrophy, and *KCNJ11*, which is known for its role in PNDM<sup>72</sup>. Early T2DM GWAS identified further associations near *HNF1B*, *HNF1A* and *WFS1* (REF. 73). As studies have increased in size, the overlap between common variant signals and genes associated with monogenic diabetes mellitus has continued to increase: as of 2015, >10% of common variant signals for T2DM localized within 250 kb of a gene associated with monogenic diabetes mellitus, with additional overlap observed for glucose or insulin common variant associations (FIG. 2). Furthermore, a substantial number of signals localize within the same regulatory networks affected by mutations linked to monogenic diabetes mellitus, as demonstrated by the enrichment of T2DM or glucose GWAS associations within binding sites for transcription factors implicated in MODY<sup>57,58,68</sup>.



**Figure 1 | Monogenic forms of diabetes mellitus.** Shown are genes implicated in a restricted set of monogenic diabetes mellitus subtypes (discussed in the article). Each class of disorder is shown as a box in the central column: neonatal diabetes mellitus (NDM, red); maturity-onset diabetes of the young (MODY, blue); and mitochondrial diabetes mellitus (grey), while syndromic forms are divided based on closer similarity to NDM or MODY (light blue). The major clinical features of each disorder are described in each box. NDM and MODY with organ dysfunction can have one or more organs affected. Genes are shown as boxes in the outer two columns, with arrows connecting to the disorders for which the genes are implicated. Genes implicated in more than one disorder have arrows to each. CNS, central nervous system.

**Box 1 | Distinction between rare and common forms of diabetes mellitus**

Rare and common forms of diabetes mellitus are characterized by vastly different 'genetic architectures', defined as the number, frequencies and effect sizes of causal variants in a population<sup>55</sup>. Rare forms are characterized by highly penetrant, often assumed deterministic, variants within a single gene, although incomplete penetrance can occur and lead to overestimation of risk<sup>88</sup>. Although heterogeneity is often high, with numerous causal variants, the combined allele frequency is very low, with individual mutations very rare or even private to a familial lineage.

The genetic architecture of type 2 diabetes mellitus is far less certain than for rare forms of diabetes mellitus. In general, 'rare variant' models, in which very low frequency variants explain the bulk of genetic risk of the disease, 'common variant' models, in which common variants explain the bulk of risk, and intermediate models are all possible<sup>55</sup>. Rare variant models have been justified on the basis of the modest effects of common variants identified thus far<sup>113–115</sup>. Common variant models have, conversely, been justified based on the substantial amount of disease risk that common variants explain when analysed in aggregate<sup>53,56,116,117</sup>.

Although perhaps academic, genetic architecture greatly determines the degree to which genetic studies can inform on disease risk and mechanisms. Rare coding variants with large effect clearly implicate disease mechanisms, as they suggest a specific gene for which a perturbation can have a substantial phenotypic effect; loss-of-function protective mutations can further suggest a directional perturbation with a clinical benefit<sup>87,118</sup>. Furthermore, disease risk can be substantially elevated in individuals who carry large-effect mutations, which enables risk prediction or preventive measures. By contrast, common variant associations of weak effect are less valuable for risk prediction<sup>119</sup> and can be challenging to localize to individual genes or disease mechanisms<sup>62,120</sup>. Links between rare and common forms of diabetes mellitus, therefore, provide opportunities to translate insights from well-understood to less-understood disease forms.

**Low frequency variants for T2DM**

Findings from T2DM GWAS have, thus, shown a clear genetic link between common and monogenic forms of diabetes mellitus<sup>9–10</sup>, although the properties of variants differ greatly between disease forms. Variants for monogenic diabetes mellitus are rare and result in large, possibly 'deterministic', increases in disease risk. GWAS variants are generally common and result in modest effects on disease risk. The increasing scale of NGS has, in the past 2–3 years, enabled studies to test whether a broader 'allelic series' — consisting of low frequency or rare variants with moderate-to-large effect on risk — contributes to T2DM.

The studies most successful at identifying lower frequency associations (that is, a frequency of  $\leq 5\%$ ) have used the Illumina<sup>®</sup> (Illumina, Inc., USA) Human Exome Array, which includes ~70–80% of coding variants with a frequency  $>0.5\%$  in the European population<sup>74</sup>. The first study, in 2013, analysed ~8,000 individuals and identified five independent low-frequency variants, each explaining 0.4–0.9% of genetic variance for fasting proinsulin levels or insulinogenic index<sup>75</sup>. Two 2015 studies of 33,000–61,000 individuals identified a series of low-frequency variants in *G6PC2*, as well as a variant in *GLPIR*, associated with fasting glucose levels<sup>76,77</sup>. An additional low-frequency variant in *AKT2* was associated with fasting insulin levels<sup>78</sup>. A striking finding from a separate study, which used a similar genotyping array (the MetaboChip<sup>79</sup>) in the isolated Inuit population, was a protein-truncating variant in *TBC1D4* for which homozygous carriers have a 10-fold increased risk of T2DM in addition to increased postprandial levels of glucose and insulin<sup>80</sup>. Despite these positive examples,

however, the largest exome array study conducted to date for T2DM identified only one statistically significant low-frequency association (a previously reported variant in *PAM*<sup>75,81</sup>), even though numerous common variants far exceeded stringent significance thresholds<sup>74</sup>.

Genome-wide sequencing studies have yielded a few additional associations. Although the first WES study identified no low-frequency or rare variants associated with T2DM in 2,000 Danish individuals<sup>82</sup>, a similar study of ~4,000 Mexican and Latino individuals identified one low-frequency variant in *HNFL1A* with a fivefold increased risk of T2DM<sup>83</sup>. In a larger study that imputed variation from ~2,600 Icelandic whole genomes into ~278,000 individuals<sup>81</sup>, four rare or low-frequency variants were associated with T2DM: two in *PAM*, which were reported in the first exome array study as associated with insulinogenic index<sup>75</sup>, and one in each of *CCND2* and *PDX1*, with twofold effect sizes on T2DM risk. Sequencing of more heterogeneous samples (13,000 multiethnic exome sequences and 2,600 Northern European genome sequences with imputation into 44,000 additional Europeans<sup>74</sup>) did not identify any additional low-frequency or rare variants associated with T2DM, apart from the previously reported *PAM* and *CCND2* signals.

Relative to genome-wide studies, targeted sequencing studies of selected genes have yielded more promising findings. In most cases, restriction of analysis to predicted functional variation has been necessary to demonstrate association. When 36 *MTNR1B* variants with frequencies below 0.1% were assessed for melatonin binding and signalling capacity, partial or total loss-of-function variants were associated with a 5.6-fold increased risk of T2DM<sup>84</sup>. Similarly, nine rare *PPARG* variants that resulted in defects in adipocyte differentiation were associated with a 7.2-fold increase in T2DM risk<sup>85</sup>, and 11 rare variants that impaired transcriptional activity, DNA binding or nuclear localization of *HNFL1A* were associated with a sixfold increase in T2DM risk<sup>86</sup>. In perhaps the largest targeted gene study to date, beginning with exon sequencing of Finnish individuals and culminating in analysis of 150,000 individuals of five ethnicities, a series of 12 predicted protein-truncating rare variants in *SLC30A8* (which also harbours a common missense variant with a modest effect on T2DM risk) were associated with a threefold reduction in T2DM risk<sup>87</sup>.

Many of the rare or low-frequency variants identified in these studies are located near GWAS signals or within genes linked to monogenic diabetes mellitus (FIG. 2), which supports the 'allelic series' hypothesis for T2DM risk. Two sequencing studies have examined this hypothesis further. The first study sequenced seven *MODY* genes in ~4,000 individuals from the general population and demonstrated an excess of predicted deleterious mutations in individuals with diabetes mellitus<sup>88</sup>. The second study analysed the sequences of 82 genes associated with monogenic diabetes mellitus in 13,000 individuals and demonstrated a strong enrichment of predicted deleterious alleles in individuals with diabetes mellitus<sup>74</sup>. Enrichment increased with

**Allelic series**

A number of alleles of a gene or chromosome region with a range of phenotypic and/or molecular effects; useful to infer a genetic–phenotypic 'dose-response' curve.

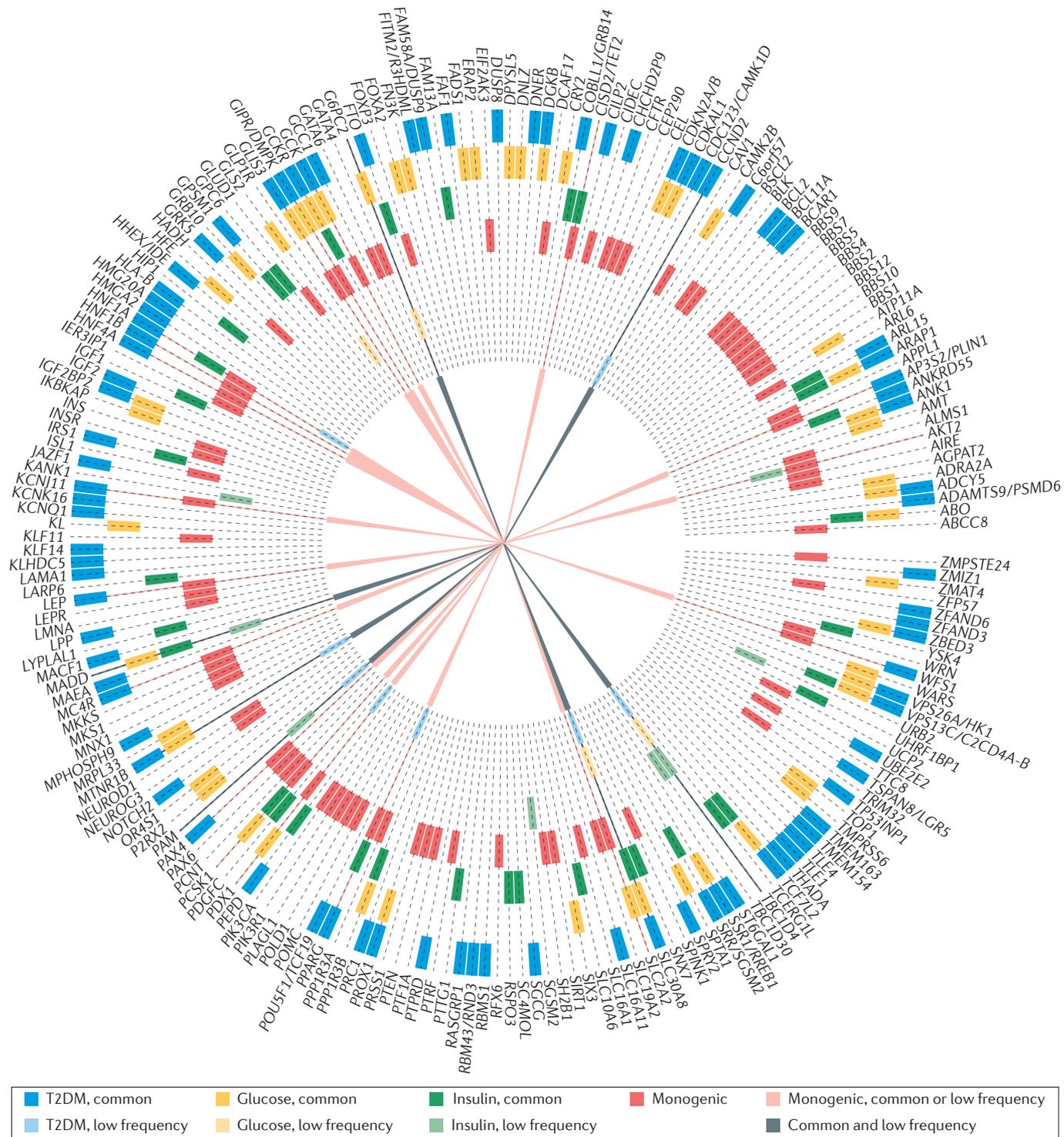


Figure 2 | Variants associated with monogenic diabetes mellitus, T2DM or levels of glucose or insulin. Shown is a list of reported associations between genes and diabetes mellitus or related traits. Associations are labelled based on the gene through which the associated variant is hypothesized to act (for variants identified in genome-wide association studies (GWAS), in many cases the gene is uncertain). The outer three circles show common variant associations for type 2 diabetes mellitus (T2DM) (blue), fasting or 2-h glucose levels (yellow), and fasting insulin or proinsulin levels (green). The middle circle (red) shows variants reported for monogenic disorders due to abnormal glucose or insulin levels, or adipocyte function. The inner three circles show higher-effect variants (in most cases of low frequency) associated with insulin (light green), glucose (light yellow), or T2DM (light blue). Pink lines are drawn through genes for which associations have been reported for a monogenic disorder as well as with a common or low-frequency variant. Grey lines are drawn through genes for which associations with both a higher-effect and common variant association have been reported.

## Box 2 | Strategies to identify disease genes and mutations

Technology greatly determines the resolution with which disease-relevant genomic regions can be mapped. The traditional means to identify genes for monogenic disorders has been through linkage scans, in which a small number of polymorphic markers are assessed for co-segregation with disease in pedigrees, followed by positional cloning, in which the mapped locus is sequenced and analysed for causal variants and effector genes<sup>121,122</sup>. Linkage studies have good power for monogenic diseases but have proven less successful for common diseases such as type 2 diabetes mellitus (T2DM)<sup>50</sup>.

By contrast, genome-wide association studies (GWAS), motivated by the common disease–common variant hypothesis (CDCV)<sup>123</sup>, compare variant frequencies between unrelated affected individuals and controls. Shared haplotypes of causal variants, which result from rapid historical population growth, enable the study of nearly all common variants in the genome through linkage disequilibrium with a fairly small number of genotyped variants (<1 million)<sup>122</sup>. With sample sizes now exceeding hundreds of thousands, GWAS have detected numerous common variant associations with T2DM, although most have a small effect on disease risk<sup>53</sup>.

Both linkage and GWAS approaches have limited resolution; sequencing of implicated regions is necessary to identify the actual causal variant. In the past 5 years, next-generation sequencing has made possible higher resolution studies that sequence entire patient exomes or genomes. Whole-exome sequencing (WES) has proven to be a powerful alternative to linkage studies to detect novel Mendelian disease genes<sup>28</sup>, leading to large-scale collaborative mapping efforts<sup>124–126</sup> and affecting diagnostic settings<sup>127</sup>. For common diseases, WES and whole-genome sequencing have been applied to tens of thousands of samples<sup>74</sup>, with the goal to characterize rarer variants than have been analysed by common variant GWAS<sup>113</sup>. As WES is still much more expensive than array-based technologies, ‘next-generation’ arrays (such as the Illumina<sup>®</sup> (Illumina, Inc., USA) Human Exome Array) have been used to target subsets of low-frequency variants in larger sample sizes.

predicted molecular severity of the variant or predicted relevance to diabetes mellitus of the affected gene<sup>74</sup>. In each of these studies, analyses ruled out undiagnosed monogenic diabetes mellitus as an explanation for the associations.

### Variant effects in the population

NGS studies of T2DM have, thus, suggested that rare or low-frequency alleles of moderate effect for T2DM coexist alongside common alleles of low effect for T2DM, as well as alongside rare alleles of high penetrance for monogenic diabetes mellitus. Plausibly, upon re-examination, variants previously reported to cause monogenic diabetes mellitus might, in fact, have more modest effects on diabetes mellitus risk. For example, linkage or familial studies of monogenic diabetes mellitus discover individuals with severe forms of the disease; these individuals are enriched not only for disease alleles<sup>89</sup> but also for many other risk factors, which, when unaccounted for, can increase risk estimates for the general population<sup>90</sup>. Other biases, such as unrecognized confounding or ‘winner’s curse’, can also lead to overestimated association strengths<sup>91–93</sup>. Benign variants within disease genes can also be falsely classified as molecularly damaging and, indeed, large-scale sequencing efforts have shown that putatively healthy individuals carry a substantial number of such variants<sup>94–96</sup>.

A study published in 2013 investigated the extent to which variants in genes implicated in MODY affected diabetes mellitus risk in the general population<sup>88</sup>. Despite stringent bioinformatics filters, ~1% of individuals were found to carry variants potentially pathogenic

for MODY; however, none developed MODY and there was, at most, a weak trend towards the development of diabetes mellitus later in life<sup>88</sup>. Similarly, an exome sequencing study showed that variants reported in the literature to cause MODY were only modestly associated with diabetes mellitus risk, with effect sizes of the order of those for common variants identified from GWAS<sup>74</sup>.

To illustrate specific examples of variants that might have lower penetrance than previously assumed, we examined a set of mutations initially reported as associated with monogenic diabetes mellitus but subsequently observed in large exome sequencing datasets. We intersected variants for the five most common MODY subtypes (*GCK*, *HNF1A*, *HNF4A*, *HNF1B* and *INS*), as reported in the [ClinVar](#) database, with those identified in a T2DM exome sequencing study of ~17,000 individuals ([Type 2 Diabetes Genetics Portal](#)). We then followed a protocol used in clinical diagnostic laboratories<sup>97</sup> and guidelines provided by the American College of Medical Genetics<sup>98</sup>, together with a manual examination of the original published evidence of pathogenicity, to assign each variant to one of five classes: class 1, benign; class 2, likely benign; class 3, variants of unknown significance (VUS); class 4, likely pathogenic; and class 5, pathogenic.

A total of 15 variants (cumulative allele frequency of 0.5%) in the T2DM exome sequencing study were labelled as either pathogenic or likely pathogenic in the [ClinVar](#) database (TABLE 1). Our manual assessment based on current protocols, however, showed that only three were assessed as pathogenic, three as probably pathogenic, and the remaining nine as either VUS or likely benign. For many variants that we reclassified as VUS or likely benign, the original report of pathogenicity provided little information on whether the variant segregated with the disease in the families of the probands. Few of these variants were, or have been since, tested for functionality. Thus, in this small sample, almost two-thirds of variants initially reported as fully penetrant for MODY have, based on the best-practices used today, very limited evidence for being pathogenic.

Indeed, the variants reclassified as VUS or likely benign were, in aggregate, observed at equal frequencies in individuals with and without diabetes mellitus ([Type 2 Diabetes Genetics Portal](#)). Of the nine variants, more than half were observed in individuals without diabetes mellitus and three were observed in at least five individuals without diabetes mellitus. By contrast, five of the six variants reclassified as pathogenic or likely pathogenic were observed exclusively in individuals with diabetes mellitus. This example, while not statistically significant, suggests that modern criteria for variant classification might improve upon those originally used to identify MODY mutations. However, the presence of pathogenic or likely pathogenic variants in individuals with reported T2DM, but not MODY, suggests that even modern criteria can overstate the penetrance of mutations classified as pathogenic for monogenic diseases, and that even more stringent criteria (such as functional characterization<sup>83,85</sup>) might be necessary to establish pathogenicity.

### Probands

In a familial study, the first affected family member who seeks attention for a disease presentation.

Table 1 | Analysis of variants originally reported as causal for MODY

Variant information			Assessment of pathogenicity		Frequencies in 17,000 exomes		
Gene/MODY subtype	cDNA change	Protein change	ClinVar	Manual	MAF	MAC cases	MAC controls
<b>Manual evaluation: All pathogenic and likely pathogenic</b>							
GCK/MODY2	c.683C>T	p.T229M	Pathogenic	5: Pathogenic	3E-05	1	0
GCK/MODY2	c.676G>A	p.V227M	Pathogenic	5: Pathogenic	3E-05	1	0
GCK/MODY2	c.1240A>G	p.K415E	Likely pathogenic	5: Pathogenic	3E-05	1	0
GCK/MODY2	c.1160C>T	p.A388V	Likely pathogenic	4: Likely pathogenic	2.7E-05	1	0
HNF1A/MODY3	c.965A>G	p.Y322C	Likely pathogenic	4: Likely pathogenic	0.0003	6	4
HNF4A/MODY1	c.991C>T	p.R331C	Likely pathogenic	4: Likely pathogenic	3E-05	1	0
Total	–	–	–	–	0.0005	11	4
<b>Manual Evaluation: All VUS or benign</b>							
HNF1A/MODY3	c.92G>A	p.G31D	Pathogenic	3: VUS	0.0004	7	8
HNF4A/MODY1	c.406C>T	p.R136W	Pathogenic	3: VUS	6.5E-05	2	0
GCK/MODY2	c.440G>A	p.G148D	Likely pathogenic	3: VUS	3E-05	1	0
HNF1A/MODY3	c.1424C>T	p.P475L	Likely pathogenic	3: VUS	3E-05	1	0
HNF1A/MODY3	c.1745A>G	p.H582R	Likely pathogenic	3: VUS	6E-05	1	1
HNF1A/MODY3	c.1748G>A	p.R583Q	Likely pathogenic	3: VUS	0.0002	2	5
HNF1B/MODY5	c.140C>T	p.P47L	Likely pathogenic	3: VUS	5.9E-05	2	0
HNF4A/MODY1	c.505G>A	p.V169I	Likely pathogenic; uncertain	2: Likely benign	0.004	65	60
INS/Hyperpro-insulinaemia	c.266G>A	p.R89H	Pathogenic*	3: VUS*	3.3E-05	0	1
Total	–	–	–	–	0.0047	81	75
<b>Clinvar: All pathogenic and likely pathogenic</b>							
Total	–	–	–	–	0.0052	92	79

\*As a representative example of variants initially reported as pathogenic for maturity-onset diabetes of the young (MODY), but which might in fact have reduced penetrance, we intersected variants reported in the ClinVar database as causal for the five most common MODY subtypes (GCK, HNF1A, HNF4A, HNF1B and INS) with those identified in a type 2 diabetes mellitus exome sequencing study of ~17,000 individuals and we then re-analysed the original reported evidence for pathogenicity according to a modern protocol used in clinical diagnostic laboratories<sup>97</sup>. The table above shows the 15 analysed variants, together with their modern classification and frequency information from the exome sequencing study. Gene/MODY subtype: the reported MODY phenotype; cDNA change/protein change: the effect on the coding transcript or protein; ClinVar: the classification as reported in the ClinVar database; Manual: the classification using modern criteria; minor allele frequency (MAF): the frequency across 17,000 exomes; minor allele count (MAC) cases/controls: the number of case and control observations in the 17,000 exomes. VUS, variant of unknown significance.

### A continuum of genetic effects

In light of these findings, a series of alleles ranging from fully-penetrant to weak effect, from low frequency to high frequency, and from coding to regulatory, might lie within or near many genes or pathways relevant to diabetes mellitus. Although T2DM and monogenic diabetes mellitus might be caused by variants at extreme ends of this spectrum, elucidation of the full allelic series might make it impossible to establish a clear division between these classes of disease.

Variants in *HNF1A* provide an excellent example of this hypothesis. *HNF1A* encodes a master transcription factor (HNF-1A) that regulates genes expressed in the liver, kidney and pancreas<sup>99,100</sup>. The protein structure of HNF-1A consists of distinct domains for dimerization, DNA binding and transactivation<sup>101</sup>. Most mutations in *HNF1A* that are associated with MODY result in an altered DNA binding domain of the encoded protein, causing a near-complete impairment of HNF-1A transactivation activity (FIG. 3). By contrast, common coding variants within *HNF1A*

(for example, encoding p.I27L and p.A98V) modestly increase risk of T2DM and have only a minimal effect on transactivation activity<sup>83</sup>. In between these two extremes, a low-frequency *HNF1A* variant (encoding p.E508K) identified in Mexican and Latino populations has intermediate phenotypic and molecular effects<sup>83</sup>.

Modest effects on T2DM risk have also been observed for variants that alter regulation of the transcriptional network involving HNF-1A. Variants in HNF-1A binding sites are enriched for T2DM associations, which suggests that perturbation of genes regulated by HNF-1A can cause less extreme forms of diabetes mellitus<sup>57,58</sup>. In addition, noncoding common variants upstream of *HNF1A* have also been associated with T2DM<sup>73</sup>, which suggests that regulation of *HNF1A* itself might also predispose to T2DM. Notably, common variants near *HNF1A* associated with other traits such as lipid levels<sup>102</sup> or coronary artery disease<sup>103</sup> are distinct from those associated with T2DM, which suggests that



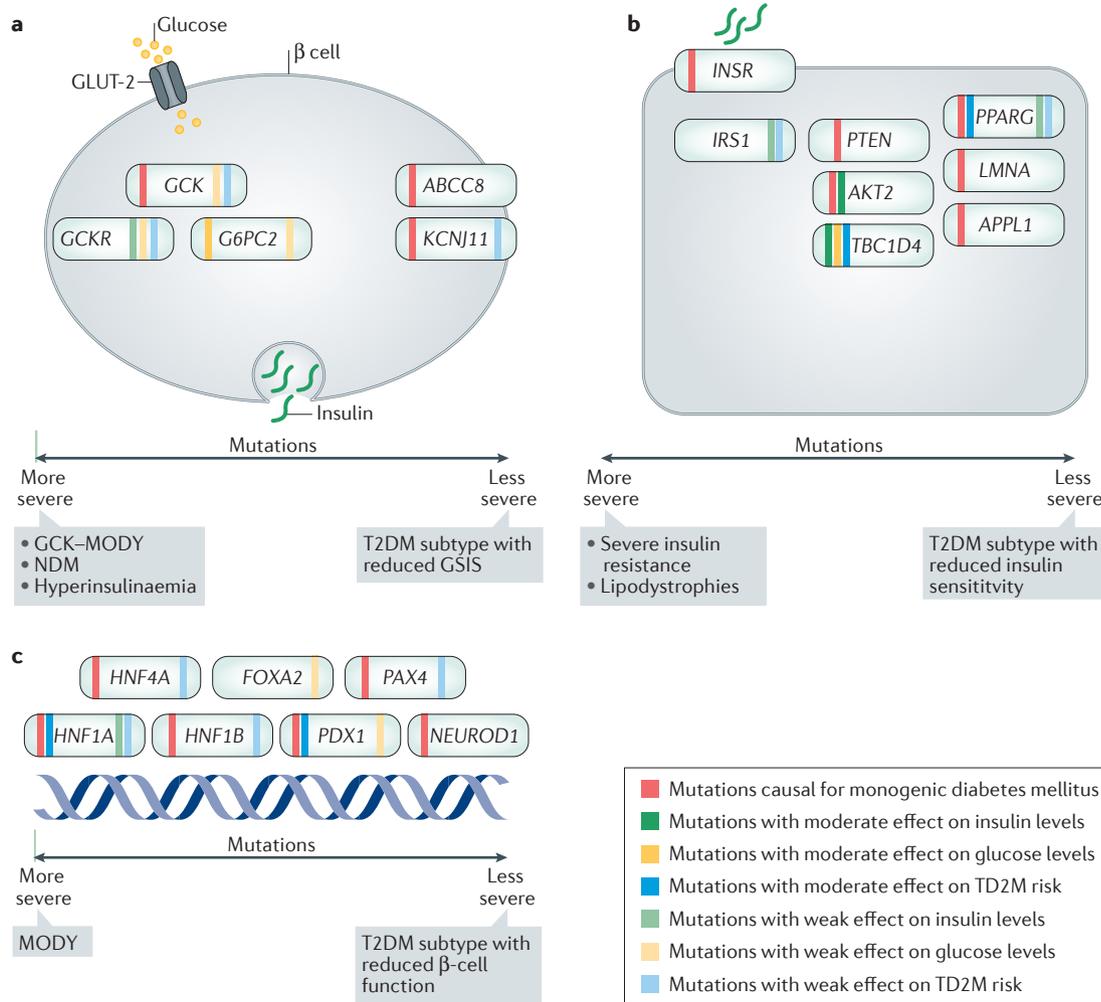
**Figure 3 | An allelic series in *HNF1A*.** *HNF1A* contains rare variants that cause maturity-onset diabetes in the young (MODY), as well as variants within or nearby the gene with smaller effects on diabetes mellitus risk. Shown is a schematic of the gene, with the distinct protein domains labelled at the top. Regions are shaded according to the likelihood that a variant within causes MODY. Below the schematic are shown some variants for which functional classification or phenotypic association is available. Variants with phenotypic association data are labelled and coloured as causing type 2 diabetes mellitus (T2DM), or no disease, while other variants are labelled according to predicted molecular effect. At the bottom, all *HNF1A* variants found by intersecting the [ClinVar](#) database with variants found in the [Type 2 Diabetes Genetics Portal](#) (see TABLE 1) are shown. POU, bipartite DNA-binding domain found in eukaryotic transcription factors; VUS, variants of unknown significance. Modified with permission from Sigma Type 2 Diabetes Consortium. Association of a low-frequency variant in *HNF1A* with type 2 diabetes in a Latino population. *JAMA* 311, 2305–2314 (2014) Copyright © (2014) American Medical Association. All rights reserved.

multiple disease-relevant regulatory pathways might involve HNF-1A. A range of possible molecular perturbations of *HNF1A*, as well as genes upstream or downstream of it within multiple regulatory networks, can, thus, result in a range of phenotypes relevant to diabetes mellitus.

On the basis of overlap between genes associated with monogenic diabetes mellitus and those associated with T2DM, multiple pathways relevant to diabetes mellitus could have properties similar to those observed for *HNF1A*. Three such pathways are highlighted in FIG. 4. One pathway consists of genes responsible for glucose-stimulated insulin secretion. Whereas carriers of severe mutations might develop NDM or MODY, carriers of less severe mutations might exhibit a T2DM subtype characterized primarily by insufficient glucose secretion, with potential response to sulfonylureas or progression to minimal complications (similar to patients with *GCK* mutations). A second pathway controls insulin signalling in muscle and fat. Whereas carriers of loss-of-function variants might develop severe insulin resistance or lipodystrophies, carriers of

less severe mutations might exhibit a T2DM subtype characterized primarily by insulin resistance, defects in insulin response or partially degenerative conditions of adipose tissue. A final pathway consists of the transcriptional network of  $\beta$  cells, which regulates multiple genes, influences  $\beta$ -cell growth and differentiation, and affects internal metabolic processes. Whereas carriers of loss-of-function variants might develop MODY, carriers of less severe mutations might exhibit a T2DM subtype characterized primarily by inadequate transcriptional response or decreased  $\beta$ -cell mass and renewal.

Under this model, T2DM and monogenic diabetes mellitus would be better viewed as extreme points on a landscape of phenotypes, rather than fundamentally distinct conditions. Many phenotypes considered T2DM might in fact be diabetes mellitus subtypes<sup>104</sup> caused by mutations in specific pathways. Such phenotypes might more so resemble monogenic disorders caused by large-effect mutations in the same pathway, rather than subtypes of T2DM due predominantly to mutations in different pathways.



**Figure 4 | A unified model of diabetes mellitus risk.** Increasingly, diabetes mellitus-associated monogenic, low-frequency and common variants cluster within not only individual genes but within pathways as well. Shown are subsets of three such pathways for which multiple genes carry a series of variants: **a** | glucose-stimulated insulin secretion (GSIS), **b** | insulin signalling and **c** |  $\beta$ -cell transcriptional regulation. Variants reported as associated with monogenic diabetes mellitus (red), type 2 diabetes mellitus (T2DM; blue: moderate effect, light blue: weak effect), glucose (yellow: moderate effect, light yellow: weak effect) or insulin (green: moderate effect, light green: weak effect). The gradients underneath each pathway illustrate the hypothesis that, rather than being distinct conditions, monogenic disorders caused by mutations in one of these pathways might in reality be extreme points on a phenotypic continuum. The current broad classification of T2DM might disguise disease subtypes characterized by presentations similar to those of monogenic forms, such as conditions of moderately impaired GSIS (variants in **a**), moderately impeded insulin signalling (variants in **b**), and reduced  $\beta$ -cell mass or function (variants in **c**). Potential implications of this model are discussed in TABLE 2. GLUT-2, glucose transporter type 2; MODY, maturity-onset diabetes of the young; NDM, neonatal diabetes mellitus.

**Future directions**

Links have been hypothesized between T2DM and monogenic diabetes mellitus for some time<sup>8-10</sup>, helping in some cases to understand diabetes mellitus pathophysiology or treatment responses<sup>44,46</sup>. Today, however, the diseases continue to be studied somewhat independently, within partially overlapping research communities and with different experimental approaches. A unified model of diabetes mellitus risk, encompassing both common and rare forms of the disease, could have implications for the study and treatment of each (TABLE 2).

Toward the goal of diagnosis and study of monogenic forms of diabetes mellitus, we see several areas that a unified diabetes mellitus risk model could

influence. First, a unified model predicts that many pathways will harbour highly penetrant mutations for monogenic diabetes mellitus, and skilled analysis of whole-exome or whole-genome sequence data will, thus, be necessary to diagnose many families. Multiple mutations, some expected to be regulatory, will need to be analysed simultaneously. Analysis of causal mutations could assign higher likelihood to those present in genes implicated in T2DM, but incorporating this idea into diagnostics will require statistical sophistication. In some cases, families thought to harbour Mendelian diseases might need reconsideration for more complex inheritance patterns, possibly approaching those of earlier-onset forms of T2DM.

Second, large genetic or functional datasets should be increasingly used as essential resources to calibrate the frequencies and effects of variants observed from diagnostic sequencing. Public datasets such as the [Exome Variant Server](#) or the [ExAC Browser](#) have already been recommended for use as frequency filters on potential causal variants<sup>98</sup>. However, under a unified risk model, many variants with appreciable frequencies in these datasets will in fact contribute to risk of diabetes mellitus, some with significant effect. Public datasets established for diabetes mellitus, such as those in the [Type 2 Diabetes Genetics Portal](#), might be better incorporated into diagnostic procedures for diabetes mellitus, as they will enable a more accurate estimate of the association with diabetes mellitus. Ultimately, however, molecular or cellular assays might become necessary in many diagnostic settings, as they could be the only means to obtain sufficient precision in the estimate of functional and, hence, phenotypic effect. If so, the degree to which such assays reflect human physiology, measurable in part by correlations between assay readouts and phenotypic associations<sup>84–86</sup>, will need to be communicated, and a balance between accuracy and turnaround time, addressed potentially by new techniques to screen possible mutations in advance of diagnosis<sup>105,106</sup>, will need to be struck.

Finally, standards for reporting variants that cause monogenic diabetes mellitus should be revisited<sup>107</sup>. Some considerations apply in all clinical genetics settings<sup>108</sup>, such as ascertainment bias in the original reporting study and

difficulty in classifying a variant as functional. However, a unified model implies that, even when variants are functional and causal for monogenic diabetes mellitus, they will sometimes be nondeterministic and sometimes increase risk only modestly and later in life. New guidelines will need to be developed to accurately represent the condition for which a patient has increased risk, as well as the quantitative degree of increased risk. These complex risk factors will need to be communicated clearly and consistently to patients, possibly through increased collaboration between diagnostic laboratories and centralized databases for reporting and storing clinical genetic results.

In addition to these considerations for monogenic forms of diabetes mellitus, we anticipate several aspects in which a unified risk model could also affect efforts to understand or treat the common disease of T2DM. In the near future, analysis of genes linked to monogenic diabetes mellitus might be the most profitable means to gain genetic insights into T2DM. Relative to the few rare variant associations identified by exome-wide and genome-wide analyses<sup>74</sup>, the associations within monogenic diabetes mellitus genes have been striking<sup>74</sup>. A unified model of diabetes mellitus risk would predict that an allelic series exists within many monogenic genes, and a more narrow focus on these genes might prioritise variants for study that stringent genome-wide significance thresholds might disregard. Logical candidates for study beyond monogenic genes might be those within the same biological pathway; the association of a common

Table 2 | Implications of a unified model for diabetes mellitus risk

Area	Current practice	Future approaches?
<b>Monogenic forms of diabetes mellitus</b>		
Sequencing patients with disease	Targeted gene panels	Whole-exome or whole-genome sequencing
	Mendelian inheritance pattern	More complex inheritance patterns
	Simple filters on individual variants	Multi-gene analysis, prioritise T2DM regions
Variant interpretation	Phenotype-agnostic or in-house 'control' datasets	Databases of samples ascertained for disease
	Bioinformatic predictors and small-scale functional assays	High-throughput functional assay 'look-up' tables
Disease risk reporting	Qualitative likelihood of pathogenicity with 'deterministic' pathogenic effects	Quantitative estimates of both pathogenicity and risk
	Risk for monogenic disease	Risk for spectrum of diabetes mellitus subtypes
<b>T2DM</b>		
Identification of novel genes or disease mechanisms	Unbiased genome-wide or exome-wide association studies	Prior likelihood of association for genes in monogenic pathways
	Enrichment or aggregation tests within monogenic genes	Dissect and characterize signals within monogenic genes
Ascertainment of samples for genetic studies	Ascertain individuals with diabetes mellitus and matched controls	Ascertain based on a spectrum of clinical features
	Exclude individuals with potential monogenic diabetes mellitus	Group samples or genes on the basis of clinical or functional similarity
Targeted medications for patients with T2DM	Little genetic information available to target treatments	MODY therapeutic targets with moderate-effect T2DM variants might be precision medicine candidates

Genetic evidence now supports an 'allelic series' across numerous genes and pathways involved in both rare and common forms of diabetes mellitus. A model in which these diseases are viewed as different subtypes of a continuous spectrum of disorders might have implications as to how we study, diagnose and treat diabetes mellitus. MODY, maturity-onset diabetes in the young; T2DM, type 2 diabetes mellitus.

## CRISPR/Cas9

A technique for precise and efficient editing of genetic information within a cell.

Greenlandic variant within *TBC1D4* (REF. 80), a phosphorylation target of a gene (*AKT2*) in which variants cause dominant disorders of severe insulin resistance<sup>109</sup> and lipodystrophy<sup>110</sup>, is an example of how an allelic series can span multiple genes in the same pathway.

Association studies for T2DM might also improve power to detect association by stratifying patients according to phenotypic subtypes. Under a unified model, analysing T2DM as a single phenotype might introduce substantial genetic heterogeneity<sup>104</sup>, whereas stratifying patients based on phenotypic similarity with rarer disorders might reduce heterogeneity. One means to achieve this stratification might be to measure characteristics related to monogenic disease symptoms, such as renal function (relevant to HNF1B–MODY), C-reactive protein levels (HNF1A–MODY), or body mass distribution (relevant to lipodystrophies). Genes could also be stratified for analysis on the basis of functional similarity to those implicated in a monogenic disorder. For example, adipocyte differentiation has proven an effective assay to characterize variants in *PPARG*<sup>85</sup>, a gene implicated in lipodystrophies. With the advent of new genome-editing technologies such as CRISPR/Cas9 (REF. 111), unbiased genome-wide knockout screens could be performed to identify additional genes relevant to adipocyte differentiation for subsequent genetic analysis.

Finally, medications used to treat monogenic disorders might be explored for efficacy in subtypes of T2DM. For example, the p.E508K variant in HNF-1A,

which has a fivefold effect on T2DM risk, is present in ~2% of Mexicans with T2DM. Carriers of this mutation might have a subtype of T2DM closely related to HNF1A–MODY, a monogenic diabetes mellitus subtype for which patients can show sensitivity to sulfonylureas<sup>112</sup>. Future research could assess whether screening for the p.E508K variant could guide personalized treatment for diabetes mellitus in the Mexican population. As other low-frequency mutations are discovered in other fairly well-understood disease pathways, these mutations might offer additional candidates to inform personalized medicine.

## Conclusions

On the basis of these considerations, we believe that reconciling knowledge of monogenic diabetes mellitus with findings from T2DM association studies should be a high priority in the diabetes mellitus research community. Building an accurate, unified model of diabetes mellitus risk would provide a valuable framework to organize our current knowledge of diabetes mellitus pathophysiology, enable future findings to be placed in a richer context than is available today, and provide valuable ideas to catalyse future research, treatments and diagnostics. At a minimum, we anticipate that researchers who today study solely common or rare forms of diabetes mellitus will be well-served to be familiar with advances across both fields, as data sharing and unified analysis will be a fertile research area for the foreseeable future.

- World Health Organization. Global status report on noncommunicable diseases. [online], <http://www.who.int/nmh/publications/ncd-status-report-2014/en/> (2014).
- Anderson, M. S. & Bluestone, J. A. The NOD mouse: a model of immune dysregulation. *Annu. Rev. Immunol.* **23**, 447–485 (2005).
- Fajans, S. S. & Bell, G. I. MODY: history, genetics, pathophysiology, and clinical decision making. *Diabetes Care* **34**, 1878–1884 (2011).
- Polak, M. & Cave, H. Neonatal diabetes mellitus: a disease linked to multiple mechanisms. *Orphanet J. Rare Dis.* **2**, 12 (2007).
- Maassen, J. A. *et al.* Mitochondrial diabetes: molecular mechanisms and clinical presentation. *Diabetes* **53**, S105–S109 (2004).
- Hattersley, A., Bruining, J., Shield, J., Njolstad, P. & Donaghue, K. C. The diagnosis and management of monogenic diabetes in children and adolescents. *Pediatr. Diabetes* **10** (Suppl. 12), 33–42 (2009).
- Molven, A. & Njolstad, P. R. Role of molecular genetics in transforming diagnosis of diabetes mellitus. *Expert Rev. Mol. Diagn.* **11**, 313–320 (2011).
- McCarthy, M. I. & Hattersley, A. T. Learning from molecular genetics: novel insights arising from the definition of genes for monogenic and type 2 diabetes. *Diabetes* **57**, 2889–2898 (2008).
- Klupa, T., Skupien, J. & Malecki, M. T. Monogenic models: what have the single gene disorders taught us? *Curr. Diab. Rep.* **12**, 659–666 (2012).
- Tallapragada, D. S., Bhaskar, S. & Chandak, G. R. New insights from monogenic diabetes for 'common' type 2 diabetes. *Front. Genet.* **6**, 251 (2015).
- Kahn, S. E., Cooper, M. E. & Del Prato, S. Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future. *Lancet* **383**, 1068–1083 (2014).
- DeFronzo, R. A., Bonadonna, R. C. & Ferrannini, E. Pathogenesis of NIDDM. A balanced overview. *Diabetes Care* **15**, 318–368 (1992).
- Srinivasan, S. & Florez, J. C. Therapeutic challenges in diabetes prevention: we have not found the 'exercise pill'. *Clin. Pharmacol. Ther.* **98**, 162–169 (2015).
- Shields, B. M. *et al.* Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia* **53**, 2504–2508 (2010).
- Steele, A. M. *et al.* Prevalence of vascular complications among patients with glucokinase mutations and prolonged, mild hyperglycemia. *JAMA* **311**, 279–286 (2014).
- Isomaa, B. *et al.* Chronic diabetic complications in patients with MODY3 diabetes. *Diabetologia* **41**, 467–473 (1998).
- Gardner, D. S. & Tai, E. S. Clinical features and treatment of maturity onset diabetes of the young (MODY). *Diabetes Metab. Syndr. Obes.* **5**, 101–108 (2012).
- Barrett, T. G. & Bundey, S. E. Wolfram (DIDMOAD) syndrome. *J. Med. Genet.* **34**, 838–841 (1997).
- Seemple, R. K., Savage, D. B., Cochran, E. K., Gorden, P. & O'Rahilly, S. Genetic syndromes of severe insulin resistance. *Endocr. Rev.* **32**, 498–514 (2011).
- Heidet, L. *et al.* Spectrum of *HNF1B* mutations in a large cohort of patients who harbor renal diseases. *Clin. J. Am. Soc. Nephrol.* **5**, 1079–1090 (2010).
- Raeder, H. *et al.* Mutations in the *CEL* VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. *Nat. Genet.* **38**, 54–62 (2006).
- Murphy, R., Ellard, S. & Hattersley, A. T. Clinical implications of a molecular genetic classification of monogenic  $\beta$ -cell diabetes. *Nat. Clin. Pract. Endocrinol. Metab.* **4**, 200–213 (2008).
- Vaxillaire, M. & Froguel, P. Monogenic diabetes in the young, pharmacogenetics and relevance to multifactorial forms of type 2 diabetes. *Endocr. Rev.* **29**, 254–264 (2008).
- Gao, R. *et al.* Evaluation of a target region capture sequencing platform using monogenic diabetes as a study-model. *BMC Genet.* **15**, 13 (2014).
- Ellard, S. *et al.* Improved genetic testing for monogenic diabetes using targeted next-generation sequencing. *Diabetologia* **56**, 1958–1963 (2013).
- Bonnefond, A. *et al.* Highly sensitive diagnosis of 43 monogenic forms of diabetes or obesity through one-step PCR-based enrichment in combination with next-generation sequencing. *Diabetes Care* **37**, 460–467 (2014).
- Alkorta-Aranburu, G. *et al.* Phenotypic heterogeneity in monogenic diabetes: the clinical and diagnostic utility of a gene panel-based next-generation sequencing approach. *Mol. Genet. Metab.* **113**, 315–320 (2014).
- Bamshad, M. J. *et al.* Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* **12**, 745–755 (2011).
- Bonnefond, A. *et al.* Molecular diagnosis of neonatal diabetes mellitus using next-generation sequencing of the whole exome. *PLoS ONE* **5**, e13630 (2010).
- Johansson, S. *et al.* Exome sequencing and genetic testing for MODY. *PLoS ONE* **7**, e38050 (2012).
- Lango Allen, H. *et al.* *GATA6* haploinsufficiency causes pancreatic agenesis in humans. *Nat. Genet.* **44**, 20–22 (2012).
- Bonnefond, A. *et al.* Whole-exome sequencing and high throughput genotyping identified *KCNJ11* as the thirteenth MODY gene. *PLoS ONE* **7**, e37423 (2012).
- Prudente, S. *et al.* Loss-of-function mutations in *APPL1* in familial diabetes mellitus. *Am. J. Hum. Genet.* **97**, 177–185 (2015).
- Kerns, S. L. *et al.* A novel variant in *CDKN1C* is associated with intrauterine growth restriction, short stature, and early-adulthood-onset diabetes. *J. Clin. Endocrinol. Metab.* **99**, E2117–E2122 (2014).
- Al-Maawali, A. *et al.* Prenatal growth restriction, retinal dystrophy, diabetes insipidus and white matter disease: expanding the spectrum of PRPS1-related disorders. *Eur. J. Hum. Genet.* **23**, 310–316 (2015).
- Thomas, P. M. *et al.* Mutations in the sulfonyleurea receptor gene in familial persistent hyperinsulinemic hypoglycemia of infancy. *Science* **268**, 426–429 (1995).
- Gloyn, A. L. *et al.* Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. *N. Engl. J. Med.* **350**, 1838–1849 (2004).
- Njolstad, P. R. *et al.* Neonatal diabetes mellitus due to complete glucokinase deficiency. *N. Engl. J. Med.* **344**, 1588–1592 (2001).
- Lindner, T. H. *et al.* A novel syndrome of diabetes mellitus, renal dysfunction and genital malformation associated with a partial deletion of the pseudo-POU domain of hepatocyte nuclear factor-1 $\beta$ . *Hum. Mol. Genet.* **8**, 2001–2008 (1999).

40. Haldorsen, I. S. *et al.* Lack of pancreatic body and tail in *HNF1B* mutation carriers. *Diabet. Med.* **25**, 782–787 (2008).
41. Rubio-Cabezas, O. *et al.* Homozygous mutations in *NEUROD1* are responsible for a novel syndrome of permanent neonatal diabetes and neurological abnormalities. *Diabetes* **59**, 2326–2331 (2010).
42. Hattersley, A. T. Beyond the beta cell in diabetes. *Nat. Genet.* **38**, 12–13 (2006).
43. Irgens, H. U. *et al.* Prevalence of monogenic diabetes in the population-based Norwegian Childhood Diabetes Registry. *Diabetologia* **56**, 1512–1519 (2013).
44. Sagen, J. V. *et al.* Permanent neonatal diabetes due to mutations in *KCNJ11* encoding Kir6.2: patient characteristics and initial response to sulfonylurea therapy. *Diabetes* **53**, 2713–2718 (2004).  
**Reference 44 is a proof of concept study that described the first patient with permanent neonatal diabetes mellitus (PNDM) who successfully switched from insulin to sulfonylurea therapy based on the molecular diagnosis of KCNJ11-induced PNDM.**
45. Pearson, E. R. *et al.* Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations. *N. Engl. J. Med.* **355**, 467–477 (2006).  
**Reference 45 changed clinical practice. In a large series of patients, Pearson et al. showed that most patients with KCNJ11 mutations can safely switch from insulin to oral sulfonylurea therapy.**
46. Pearson, E. R. *et al.* Genetic cause of hyperglycaemia and response to treatment in diabetes. *Lancet* **362**, 1275–1281 (2003).  
**Reference 46 established that patients with HNF1A-MODY (MODY3) have marked sulfonylurea sensitivity and should be treated accordingly.**
47. Chakera, A. J. *et al.* Recognition and management of individuals with hyperglycaemia because of a heterozygous glucokinase mutation. *Diabetes Care* **38**, 1383–1392 (2015).
48. Hemminki, K., Li, X., Sundquist, K. & Sundquist, J. Familial risks for type 2 diabetes in Sweden. *Diabetes Care* **33**, 293–297 (2010).
49. Almgren, P. *et al.* Heritability and familiarity of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia* **54**, 2811–2819 (2011).
50. Guan, W., Pluzhnikov, A., Cox, N. J. & Boehnke, M. Meta-analysis of 23 type 2 diabetes linkage studies from the International Type 2 Diabetes Linkage Analysis Consortium. *Hum. Hered.* **66**, 35–49 (2008).
51. Bonnefond, A. & Froguel, P. Rare and common genetic events in type 2 diabetes: what should biologists know? *Cell Metab.* **21**, 357–368 (2015).
52. Mohlke, K. L. & Boehnke, M. Recent advances in understanding the genetic architecture of type 2 diabetes. *Hum. Mol. Genet.* **24**, R85–R92 (2015).
53. Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990 (2012).  
**Reference 53, a large meta-analysis, provides evidence supporting a major role for common low-effect variants in type 2 diabetes mellitus susceptibility.**
54. DIAbetes Genetics Replication Meta-analysis Consortium *et al.* Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.* **46**, 234–244 (2014).
55. Agarwala, V., Flannick, J., Sunyaev, S., GoT2D Consortium & Altshuler, D. Evaluating empirical bounds on complex disease genetic architecture. *Nat. Genet.* **45**, 1418–1427 (2013).
56. Stahl, E. A. *et al.* Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat. Genet.* **44**, 483–489 (2012).
57. Maurano, M. T. *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190–1195 (2012).  
**Reference 57 demonstrated the first systematic enrichment of regulatory DNA for common variant GWAS associations.**
58. Pasquali, L. *et al.* Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat. Genet.* **46**, 136–143 (2014).
59. Stitzel, M. L. *et al.* Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. *Cell Metab.* **12**, 443–455 (2010).
60. The Wellcome Trust Case Control Consortium *et al.* Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* **44**, 1294–1301 (2012).
61. Ragvin, A. *et al.* Long-range gene regulation links genomic type 2 diabetes and obesity risk regions to *HHEX*, *SOX4*, and *IRX3*. *Proc. Natl Acad. Sci. USA* **107**, 775–780 (2010).
62. Claussnitzer, M. *et al.* *FTO* obesity variant circuitry and adipocyte browning in humans. *N. Engl. J. Med.* **373**, 895–907 (2015).  
**Reference 62 describes a comprehensive and successful strategy for investigating non-coding complex trait associations by combining bioinformatics and genomics methods.**
63. Smemo, S. *et al.* Obesity-associated variants within *FTO* form long-range functional connections with *IRX3*. *Nature* **507**, 371–375 (2014).
64. Small, K. S. *et al.* Identification of an imprinted master trans regulator at the *KLF14* locus related to multiple metabolic phenotypes. *Nat. Genet.* **43**, 561–564 (2011).
65. Travers, M. E. *et al.* Insights into the molecular mechanism for type 2 diabetes susceptibility at the *KCNO1* locus from temporal changes in imprinting status in human islets. *Diabetes* **62**, 987–992 (2013).
66. Prokopenko, I. *et al.* A central role for *GRB10* in regulation of islet function in man. *PLoS Genet.* **10**, e1004235 (2014).
67. Claussnitzer, M. *et al.* Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. *Cell* **156**, 343–358 (2014).
68. Gaulton, K. J. *et al.* Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nat. Genet.* **47**, 1415–1425 (2015).
69. van de Bunt, M. *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 Genet.* **11**, e1005694 (2015).
70. Hegele, R. A., Cao, H., Harris, S. B., Hanley, A. J. & Zinman, B. The hepatic nuclear factor-1a G319S variant is associated with early-onset type 2 diabetes in Canadian Oji-Cree. *J. Clin. Endocrinol. Metab.* **84**, 1077–1082 (1999).
71. Altshuler, D. *et al.* The common PPAR $\gamma$  Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat. Genet.* **26**, 76–80 (2000).  
**Reference 71 describes the first association between a common variant and type 2 diabetes mellitus.**
72. Gloyn, A. L. *et al.* Large-scale association studies of variants in genes encoding the pancreatic  $\beta$ -cell  $K_{ATP}$  channel subunits Kir6.2 (*KCNJ11*) and SUR1 (*ABCC8*) confirm that the *KCNJ11* E23K variant is associated with type 2 diabetes. *Diabetes* **52**, 568–572 (2003).
73. Voight, B. F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.* **42**, 579–589 (2010).  
**Reference 73 extended the number of genes known to harbour variants for both common type 2 diabetes mellitus and monogenic forms of diabetes mellitus.**
74. Fuchsberger, C. *et al.* Genetic architecture of type 2 diabetes from sequence based analyses. *Nature* (in the press).  
**Reference 74, a comprehensive series of sequencing studies, investigates the contribution of rare and low-frequency variants to the genetic architecture of type 2 diabetes mellitus.**
75. Huyghe, J. R. *et al.* Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. *Nat. Genet.* **45**, 197–201 (2013).  
**Reference 75 describes the first systematic screen of lower-frequency coding variants in large cohorts of patients with type 2 diabetes mellitus.**
76. Wessel, J. *et al.* Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. *Nat. Commun.* **6**, 5897 (2015).
77. Mahajan, A. *et al.* Identification and functional characterization of *G6PC2* coding variants influencing glycemic traits define an effector transcript at the *G6PC2-ABCB11* locus. *PLoS Genet.* **11**, e1004876 (2015).
78. Manning, A. K. *et al.* A low frequency *AKT2* coding variant enriched in the Finnish population is associated with fasting insulin levels. Presented at the 60th Annual Meeting of The American Society of Human Genetics (2014).
79. Voight, B. F. *et al.* The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet.* **8**, e1002795 (2012).
80. Moltke, I. *et al.* A common Greenlandic *TBC1D4* variant confers muscle insulin resistance and type 2 diabetes. *Nature* **512**, 190–193 (2014).  
**Reference 80 demonstrated the power of studying isolated populations in complex trait analysis, identifying a population-specific variant that confers very high risk for type 2 diabetes mellitus.**
81. Steinthorsdottir, V. *et al.* Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nat. Genet.* **46**, 294–298 (2014).
82. Lohmueller, K. E. *et al.* Whole-exome sequencing of 2,000 Danish individuals and the role of rare coding variants in type 2 diabetes. *Am. J. Hum. Genet.* **93**, 1072–1086 (2013).
83. The SIGMA Type 2 Diabetes Consortium *et al.* Association of a low-frequency variant in *HNF1A* with type 2 diabetes in a Latino population. *JAMA* **311**, 2305–2314 (2014).  
**Reference 83 provides strong support for the existence of an allelic series of high-to-low effect diabetes variants in HNF1A.**
84. Bonnefond, A. *et al.* Rare *MTNR1B* variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nat. Genet.* **44**, 297–301 (2012).  
**Reference 84 demonstrated the ability of functional assays to discriminate disease alleles from benign alleles.**
85. Majithia, A. R. *et al.* Rare variants in *PPARG* with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. *Proc. Natl Acad. Sci. USA* **111**, 13127–13132 (2014).
86. Laeya, A. *et al.* Functional investigations of the monogenic diabetes gene *HNF1A* identify rare variants as risk factors for type 2 diabetes in the general population. *Diabetes* **64** (Suppl 1) A460 (2015).
87. Flannick, J. *et al.* Loss-of-function mutations in *SLC30A8* protect against type 2 diabetes. *Nat. Genet.* **46**, 357–363 (2014).  
**Reference 87 was the first study to pinpoint genetic variants that confer strong protective effects against type 2 diabetes mellitus.**
88. Flannick, J. *et al.* Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes. *Nat. Genet.* **45**, 1380–1385 (2013).  
**Reference 88 used population based sequencing to yield unbiased estimates of disease risk associated with rare coding variants in Mendelian disease genes.**
89. Guey, L. T. *et al.* Power in the phenotypic extremes: a simulation study of power in discovery and replication of rare variants. *Genet. Epidemiol.* **35**, 236–246 (2011).
90. Antoniou, A. *et al.* Average risks of breast and ovarian cancer associated with *BRCA1* or *BRCA2* mutations detected in case Series unselected for family history: a combined analysis of 22 studies. *Am. J. Hum. Genet.* **72**, 1117–1130 (2003).
91. Terwilliger, J. D. & Weiss, K. M. Confounding, ascertainment bias, and the blind quest for a genetic ‘fountain of youth’. *Ann. Med.* **35**, 532–544 (2003).
92. Goring, H. H., Terwilliger, J. D. & Blangero, J. Large upward bias in estimation of locus-specific effects from genomewide scans. *Am. J. Hum. Genet.* **69**, 1357–1369 (2001).
93. Gail, M. & Chatterjee, N. In *Proceedings of the Second Seattle Symposium in Biostatistics* (Seattle, 2004).
94. MacArthur, D. G. *et al.* A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**, 823–828 (2012).
95. Kohane, I. S., Hsing, M. & Kong, S. W. Taxonomizing, sizing, and overcoming the incidentalome. *Genet. Med.* **14**, 399–404 (2012).
96. Sulem, P. *et al.* Identification of a large set of rare complete human knockouts. *Nat. Genet.* **47**, 448–452 (2015).
97. Plon, S. E. *et al.* Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum. Mutat.* **29**, 1282–1291 (2008).
98. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).  
**Reference 99 is highly influential and provides recommendations for variant interpretation in rare disorders.**

99. Boj, S. F., Parrizas, M., Maestro, M. A. & Ferrer, J. A transcription factor regulatory circuit in differentiated pancreatic cells. *Proc. Natl Acad. Sci. USA* **98**, 14481–14486 (2001).
100. Odom, D. T. *et al.* Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**, 1378–1381 (2004).
101. Baumhueter, S. *et al.* HNF-1 shares three sequence motifs with the POU domain proteins and is identical to LF-B1 and APF. *Genes Dev.* **4**, 372–379 (1990).
102. Teslovich, T. M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
103. Erdmann, J. *et al.* New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* **41**, 280–282 (2009).
104. Groop, L. & Pociot, F. Genetics of diabetes — are we missing the genes or the disease? *Mol. Cell. Endocrinol.* **382**, 726–739 (2014).
105. Melnikov, A. *et al.* Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* **30**, 271–277 (2012).
106. Melnikov, A., Rogov, P., Wang, L., Gnirke, A. & Mikkelsen, T. S. Comprehensive mutational scanning of a kinase *in vivo* reveals substrate-dependent fitness landscapes. *Nucleic Acids Res.* **42**, e112 (2014).
107. Cooper, G. M. Parlez-vous VUS? *Genome Res.* **25**, 1423–1426 (2015).
108. MacArthur, D. G. *et al.* Guidelines for investigating causality of sequence variants in human disease. *Nature* **508**, 469–476 (2014).
109. George, S. *et al.* A family with severe insulin resistance and diabetes due to a mutation in *AKT2*. *Science* **304**, 1325–1328 (2004).
110. Tan, K. *et al.* Analysis of genetic variation in Akt2/PKB- $\beta$  in severe insulin resistance, lipodystrophy, type 2 diabetes, and related metabolic phenotypes. *Diabetes* **56**, 714–719 (2007).
111. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR–Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
112. Pearson, E. R., Liddell, W. G., Shepherd, M., Corral, R. J. & Hattersley, A. T. Sensitivity to sulphonylureas in patients with hepatocyte nuclear factor-1 $\alpha$  gene mutations: evidence for pharmacogenetics in diabetes. *Diabet. Med.* **17**, 543–545 (2000).
113. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009). **Reference 114 is a highly cited discussion of potential explanations for the ‘missing heritability’ that was often debated after the first wave of GWAS.**
114. McClellan, J. & King, M. C. Genetic heterogeneity in human disease. *Cell* **141**, 210–217 (2010).
115. Cirulli, E. T. & Goldstein, D. B. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat. Rev. Genet.* **11**, 415–425 (2010).
116. The International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
117. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
118. Cohen, J. C., Boerwinkle, E., Mosley, T. H. Jr & Hobbs, H. H. Sequence variations in *PCSK9*, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* **354**, 1264–1272 (2006). **Reference 119 is one of the most influential papers to demonstrate the power of protective loss-of-function mutations to suggest therapeutic targets.**
119. Jostins, L. & Barrett, J. C. Genetic risk prediction in complex disease. *Hum. Mol. Genet.* **20**, R182–R188 (2011).
120. Musunuru, K. *et al.* From noncoding variant to phenotype via *SORT1* at the 1p13 cholesterol locus. *Nature* **466**, 714–719 (2010).
121. Risch, N. J. Searching for genetic determinants in the new millennium. *Nature* **405**, 847–856 (2000).
122. Altshuler, D., Daly, M. J. & Lander, E. S. Genetic mapping in human disease. *Science* **322**, 881–888 (2008). **Reference 123 is a thorough review of the history of genetic mapping.**
123. Reich, D. E. & Lander, E. S. On the allelic spectrum of human disease. *Trends Genet.* **17**, 502–510 (2001). **Reference 124 is responsible for much of the motivation behind GWAS.**
124. Bamshad, M. J. *et al.* The Centers for Mendelian Genomics: a new large-scale initiative to identify the genes underlying rare Mendelian conditions. *Am. J. Med. Genet. A* **158A**, 1523–1525 (2012).
125. Beaulieu, C. L. *et al.* FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. *Am. J. Hum. Genet.* **94**, 809–817 (2014).
126. Wright, C. F. *et al.* Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* **385**, 1305–1314 (2015).
127. Yang, Y. *et al.* Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* **369**, 1502–1511 (2013).

#### Acknowledgements

Stefan Johansson has received a grant from the Novo Nordisk Foundation (#15OC0016108). Pål R. Njølstad has received grants from the European Research Council (AdG; #293574), the Research Council of Norway (#240413), the KG Jebsen Foundation (Translational Research Center), the University of Bergen (Diabetes Group), and the Helse Vest (Strategic Fund).

#### Author contributions

J.F., S.J. and P.R.N. researched data for the article, made substantial contributions to discussions of the content, wrote the article and reviewed and/or edited the manuscript before submission.

#### Competing interests statement

The authors declare no competing interests.

#### FURTHER INFORMATION

OMIM: <http://www.ncbi.nlm.nih.gov/omim>  
 ClinVar: <http://www.ncbi.nlm.nih.gov/clinvar/>  
 Exome Variant Server: <http://evs.gs.washington.edu/EVS/>  
 ExAC Browser: <http://exac.broadinstitute.org/>  
 Type 2 Diabetes Genetics Portal: <http://www.type2diabetesgenetics.org/home/portallHome>

ALL LINKS ARE ACTIVE IN THE ONLINE PDF