

When is it MODY? Challenges in the Interpretation of Sequence Variants in MODY Genes

Sara Althari¹ and Anna L. Gloyn^{1,2,3}

¹ Oxford Centre for Diabetes Endocrinology & Metabolism, University of Oxford, UK. ² Wellcome Trust Centre for Human Genetics, University of Oxford, UK. ³ Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK. Address correspondence to: Anna L. Gloyn, Oxford Centre for Diabetes Endocrinology and Metabolism, Churchill Hospital, Headington, Oxford OX3 7LE, UK, e-mail: anna.gloyn@dr1.ox.ac.uk

Manuscript submitted June 18, 2015; accepted July 24, 2015

■ Abstract

The genomics revolution has raised more questions than it has provided answers. Big data from large population-scale resequencing studies are increasingly deconstructing classic notions of Mendelian disease genetics, which support a simplistic correlation between mutational severity and phenotypic outcome. The boundaries are being blurred as the body of evidence showing monogenic disease-causing alleles in healthy genomes, and in the genomes of individuals with increased common complex disease risk, continues to grow. In this review, we focus on the newly emerging challenges which pertain to the interpretation of sequence variants in genes implicated in the pathogenesis of maturity-onset diabetes of the young (MODY), a presumed monogenic form of diabetes characterized by Mendelian inheritance. These challenges highlight the complexities surrounding the assignments of pathogenicity, in particular to rare protein-

altering variants, and bring to the forefront some profound clinical diagnostic implications. As MODY is both genetically and clinically heterogeneous, an accurate molecular diagnosis and cautious extrapolation of sequence data are critical to effective disease management and treatment. The biological and translational value of sequence information can only be attained by adopting a multitude of confirmatory analyses, which interrogate variant implication in disease from every possible angle. Indeed, studies which have effectively detected rare damaging variants in known MODY genes in normoglycemic individuals question the existence of a single gene mutation scenario: does monogenic diabetes exist when the genetic culprits of MODY have been systematically identified in individuals without MODY?

Keywords: type 2 diabetes · MODY · HNF1A · GCK · GWAS · variant causality · monogenic disease · mutant allele · next generation sequencing

1. What is maturity-onset diabetes of the young (MODY)?

Maturity-onset diabetes of the young (MODY) (MIM606391) is a rare disorder that results in nonketotic hyperglycemia in young adults as a consequence of altered pancreatic beta-cell function, primarily through reduced glucose-stimulated insulin secretion [1, 2]. It is characterized by a Mendelian autosomal dominant mode of inheritance, early age of onset (typically <25 years), and non-insulin dependency. It is esti-

mated to comprise 1-2% of all diabetes cases. However, prevalence rates of MODY are inaccurate and underestimated globally as it is frequently undiagnosed or misclassified as type 1 diabetes (T1D) or type 2 diabetes (T2D) because of overlapping clinical features [3-6]. In terms of genetic etiology, nearly 80% of all MODY cases are attributable to highly penetrant heterozygous mutations in one of three genes, namely the transcription factors hepatocyte nuclear factor-1 alpha (*HNF1A*) and 4-alpha (*HNF4A*), and the glycolytic enzyme glucokinase (*GCK*) (**Table 1**) [5-10]. Mutations in another transcription factor gene, *HNF1B*, account

Abbreviations:

| | |
|--------------|--|
| 1,5AG | 1,5 anhydroglucitol |
| 4PBA | 4-phenyl butyric acid |
| ABCC8 | ATP-binding cassette transporter, subfamily C, member 8 |
| Atlantic DIP | Atlantic Diabetes in Pregnancy cohort |
| BMI | body mass index |
| BRCA | breast cancer 1 (2), early onset |
| Cas9 | CRISPR associated protein 9 |
| CEL | carboxyl ester lipase |
| CHI | congenital hyperinsulinism of infancy |
| CI | confidence interval |
| CRISPR | clustered regularly interspaced short palindromic repeat |
| CRP | C-reactive protein |
| DMS2 | Diabetes in Mexico Study 2 |
| EMSA | electrophoretic mobility shift assay |
| ER | endoplasmic reticulum |
| GCK | glucokinase |
| GKRP | glucokinase regulatory protein |
| GLUT-2 | glucose transporter 2 |
| GWAS | genome-wide association studies |
| Hb1Ac | glycated hemoglobin |
| HDL | high density lipoprotein |
| HGMD | Human Gene Mutation Database |
| HNF1 | hepatocyte nuclear factor 1 |
| hPSC | human pluripotent stem cells |
| hsCRP | high-sensitivity C-reactive protein |
| IGTT | intra-peritoneal glucose tolerance tests |
| Indel | insertion/deletion |
| INS | insulin |
| iPS | induced pluripotent stem cells |
| KCNJ11 | potassium channel, inwardly rectifying, subfamily J, member 11 |
| Kir6.2 | potassium channel, inwardly rectifying |
| LDL | low-density lipoprotein |
| LDLR | low-density lipoprotein receptor |
| LoF | loss-of-function |
| MI | myocardial infarction |
| MODY | maturity-onset diabetes of the young |
| NANOG | homeobox transcription factor nanog |
| NEUROD1 | neuronal differentiation 1 |
| NGS | next-generation sequencing |
| OCT4 | octamer-binding protein 4 |
| OGTT | oral glucose tolerance test |
| OHA | oral hypoglycaemic agent |
| OR | odds ratio |
| PDX1 | pancreatic and duodenal homeobox 1 |
| PNDM | permanent neonatal diabetes |
| PPARG | peroxisome proliferator-activated receptor gamma |
| RCAD | renal cysts and diabetes |
| SIFT | sorting intolerant from tolerant |
| SIGMA | Slim Initiative in Genomic Medicine for the Americas |
| SNV | single nucleotide variant |
| SOX2 | SR γ (sex-determining region Y)-box 2 |
| SSEA-4 | stage-specific embryonic antigen 4 |
| SUR1 | sulphonylurea receptor 1 subunit of the pancreatic beta-cell ATP-sensitive potassium channel |
| TNDM | transient neonatal diabetes |
| TRA-1-60 | tumor resistance antigen 1-60 |
| UPR | unfolded protein response |

for a minority of phenotypically distinct MODY cases, characterized by profound renal complications which often progress to complete kidney failure demanding renal replacement [11] (**Table 1**). Although more genes have been implicated in the pathogenesis of MODY, most are either exceptionally rare—reported once and detected only in a handful of families—or are more commonly responsible for neonatal diabetes, such as genes encoding the Kir6.2 and sulphonylurea 1 (SUR1) subunits of the pancreatic beta-cell ATP-sensitive potassium channel, *KCNJ11* and *ABCC8* [5, 7, 8, 10, 12] (**Table 1**).

MODY is classically defined as a monogenic form of diabetes, although it is clinically, metabolically, and genetically heterogeneous as molecular dysfunction of each presumed MODY-causal gene manifests in a distinct clinical phenotype [8, 10, 13]. Each genetic subtype of MODY demands different management and treatment strategies. Thus, an accurate genetic diagnosis is the key to adequate treatment, and entails proper interpretation and assessment of known and novel variants detected in both the research and clinical sphere. In a post-next generation sequencing (NGS) age, we are faced with the challenge of making meaningful extrapolations and obtaining translational value from the enormous pools of genomic data that are being generated at an unprecedentedly rapid pace. Mendelian and other presumed monogenic disorders, such as MODY, are no exception to this challenge, despite their seemingly straightforward genetics, biology, and pathophysiology. In this article, we highlight some of these challenges in the context of MODY genetics as we review the latest efforts in variant identification, interpretation, and functional annotation, which we believe warrant a thorough re-evaluation of the existing classification scheme of diabetes.

2. Who has MODY? Discovery of genetic culprits and estimations of prevalence

Since the first clinical case report that described young-onset non-insulin dependent diabetes in three families [1], a substantial amount of progress has been made in defining the genetic architecture of MODY. Early on, this was achieved by linkage analysis in large multigenerational pedigrees. In 1992, *GCK* was identified on chromosome 7p as the first MODY gene via linkage studies in British and French families [14-16]. Shortly thereafter, microsatellite marker evaluation and positional cloning identified *HNF1A* at chromo-

Table 1. Genes implicated in the pathogenesis of maturity-onset diabetes of the young (MODY)

| MODY sub-type | Distinguishing features | Other forms of monogenic diabetes | Other disease phenotypes |
|--|--|---|--|
| GCK-MODY | Heterozygous inactivating mutations result in asymptomatic mild hyperglycemia from birth (5.5-8 mmol/l). Small OGTT increment (< 3 mmol/l). No insulin or pharmacological treatment required. Slight alterations in glycemic control with age. | Homozygous or compound heterozygous inactivating mutations cause permanent neonatal diabetes (PNDM). Carriers of heterozygous inactivating mutations are at risk of developing gestational diabetes. Heterozygous activating mutations cause hyperinsulinemic hypoglycemia. | |
| <i>Transcription factor MODY</i> | | | |
| HNF1A-MODY | Typically presents in early adulthood/adolescence, though age of onset dependent on mutation position within the gene. High sensitivity to treatment with low-dose sulphonylurea tablets. Progressive beta-cell dysfunction and increasing penetrance with age. Large glucose increments in OGTT (> 5 mmol/l). Elevated high-density lipoprotein (HDL) cholesterol levels. | | Hepatocellular adenoma (heterozygous inactivating mutations) |
| HNF4A-MODY | Macrosomia (birth weight increase by 800 g compared with non-carriers). Reduced HDL-cholesterol, triglycerides, apolipoproteins A1 and A2, and increased low-density lipoprotein (LDL) cholesterol concentrations. | ~10% heterozygous mutation carries present with diazoxide-responsive congenital hyperinsulinemic hypoglycemia in the first week of life | |
| HNF1B-MODY | Requires early insulin therapy. Non-diabetic cystic renal disease. Abnormalities in renal structure and development, and genital tract abnormalities in female carriers. Low birth weight. | Mutations may cause transient neonatal diabetes (TNDM) | Renal cysts and diabetes (RCAD) in heterozygous mutation carries |
| IPF1-MODY | Very rare. Heterozygous mutation reported in a single family. | Homozygous recessive mutations cause neonatal diabetes. | Pancreatic agenesis in homozygotes |
| NEUROD1-MODY | Rare. Heterozygous mutations reported in five families to date. | Two homozygous loss-of-function mutations in two patients with PNDM and neurological abnormalities have been reported. | |
| <i>Potassium channel (K_{ATP}) MODY</i> | | | |
| KCNJ11-MODY | Glycemic control achieved with high-dose sulphonylurea therapy. | Heterozygous activating mutations are the commonest cause of PNDM or TNDM. Recessive loss-of-function mutations manifest in congenital hyperinsulinism of infancy (CHI). | |
| ABCC8-MODY | Glycemic control achieved with high-dose sulphonylurea therapy. | Heterozygous activating mutations are the commonest cause of PNDM or TNDM. Recessive loss-of-function mutations manifest in CHI. | |
| <i>Rare forms of MODY</i> | | | |
| CEL-MODY | Rare. Heterozygous single-base deletions in two families and a short 3-repeat allele in one. | | Fecal elastase deficiency. Exocrine pancreatic insufficiency. |
| INS-MODY | Rare. Heterozygous co-segregating missense variants identified in 6 European families. | Heterozygous missense and autosomal dominant misfolding mutations cause PNDM. | |

some 12q in MODY families [17, 18]. Similar efforts also led to the discovery of *HNF4A* in several MODY pedigrees [19-21]. The aforementioned findings made by linkage mapping facilitated variant identification in those genes which seemed to be responsible for familial MODY phenotypes. Since then, >620 *GCK* mutations, >414 *HNF1A* mutations, and >103 *HNF4A* mutations in over 1440, 1240, and 170 European Caucasian families, respectively, have been reported [6, 22, 23].

Yet, the true population prevalence of MODY remains ambiguous and geographically variable because of varying methods of clinical ascertainment and patient recruitment worldwide [4-6, 24]. For example, in Italy, France, and Spain, where pediatric testing for fasting blood glucose is routine practice, mild hyperglycemia as a result of GCK-MODY (5.4-8.3 mmol/l) is more frequently detected resulting in higher reported prevalence [6, 7, 25-29]. These reports are consistent with the observation that the onset of GCK-MODY is much earlier than other MODY subtypes, particularly transcription factor MODY that is characterized by progressive beta-cell dysfunction with onset between the second and fifth decade [4]. In contrast to Southern Europe, HNF1A-MODY is more commonly diagnosed in the United Kingdom and Scandinavia, where blood glucose tests are mainly reserved for referred patients for whom a diagnosis of diabetes has already been made, but rarely conducted in asymptomatic individuals [4, 5]. Efforts have been made to approximate the minimum population prevalence of MODY among adults and children in European countries, albeit inconclusively, as most were neither systematic nor up to scale [3, 5, 30-33]. The most recent attempt, a community-based cross-sectional study by Kropff and colleagues, estimated the prevalence of HNF1A-MODY in the UK to be 84 cases per million (95% CI 36-131). The study also estimated that nearly 90% of HNF1A-MODY cases in the UK, of which there are at least 5,000-6,000, remain undiagnosed, which was in close agreement with a preceding study by Shields and colleagues [5, 33].

However, to determine MODY prevalence accurately, more large-scale population-based assessments and molecular-genetic screens of patients with diabetes are needed. Chakera and colleagues have conducted such a study using a subset of Euro-Caucasian female patients from the Atlantic Diabetes in Pregnancy (Atlantic DIP) cohort (n = 5,500) [24, 34]. Based on their analysis, it was estimated that the prevalence of GCK-MODY would be 1.1 in 1,000 (95% CI 0.03-0.29) among white

Europeans [24]. Other groups have evaluated the frequency of MODY in youth populations by genetically screening individuals from large pediatric diabetes cohorts. Some groups avoided classical MODY presentations and selection based on referral patterns to offset ascertainment bias. Based on such studies, the minimum prevalence rates of MODY among children in Norway, Poland, Germany, and the United States have been estimated to be at 3.1/100,000, 4.2-4.6/100,000, 2.39/100,000, and 2.1/100,000, respectively [9, 35-37]. Most comprehensive genetics-based assessments of MODY frequency and etiology have been conducted in white European and North American populations, thus providing more accurate representations of these specific experiences. In African and Asian populations, for instance, where MODY has been identified, it is not as well characterized. This disparity is attributable to several factors that include limited access and high cost associated with population-wide genetic analyses of diabetic cohorts, variable levels of awareness, and variable discriminatory clinical criteria in populations in which common polygenic diabetes is astoundingly high, as is the case among South Asians and Africans [6, 38-40].

3. Efforts to find the missed and misdiagnosed MODY cases

Apart from enhancing epidemiological data, there is great therapeutic and prognostic value in providing an accurate diagnosis of MODY. Genetically discriminating MODY from other types of diabetes, as well as MODY subtypes from each other, is of vital importance, as each is associated with a discrete etiology and a unique set of complications with different implications for treatment and management [8, 41, 42].

3.1 Why an accurate molecular diagnosis of MODY matters

Patients with undiagnosed MODY often follow glycemic control plans specific to either T1D or T2D, including exogenous insulin, which could be avoided with a proper clinical and genetic diagnosis [4, 43]. There are profound negative impacts on quality of life associated with insulin administration, particularly from a psychosocial standpoint [44]. Individuals prescribed insulin treatment express a number of concerns and fears, which include stress and anxiety related to self-management, weight gain, stigma, and hindrance to normal social functioning [44].

Individuals diagnosed with HNF1A-MODY are highly responsive to the low-dose oral hypoglycemic agent (OHA) sulphonylurea, and can avoid insulin injections entirely [45]. In fact, early studies testing the efficacy of sulphonylurea treatment reported a five times greater glycemic control response than metformin and a four times greater response than insulin therapy as measured by lowering fasting glucose in HNF1A-MODY subjects using BMI- and glycemia-matched T2D controls [45, 46]. The case, in terms of responsiveness to low-dose sulphonylurea tablets, is similar for patients with MODY as a result of mutations in *HNF4A* [32]. Transcription factor MODY patients are typically born normoglycemic, and develop progressive beta-cell dysfunction in adolescence or early adulthood (<25 years) [42, 43]. HNF1A-MODY is associated with micro- and macrovascular complications similar to those developed by T1D and T2D patients [47].

GCK-MODY is another example of how a genetic diagnosis can spare affected individuals ineffective and burdensome treatment. Unlike transcription factor MODY, GCK-MODY patients have lifelong, asymptomatic mild fasting hyperglycemia, usually present from birth, with only minor deterioration with age [7, 8, 48] (Table 1). Glucose is regulated in affected individuals, and insulin response is maintained, at elevated set-points [49, 50]. Micro- and macrovascular complications are rare in GCK-MODY patients, and according to recent cross-sectional studies, are less frequent than in T1D and T2D individuals, and comparable to healthy controls [51].

As a result of this homeostatic control and lack of severe complications associated with the GCK-MODY subtype, patients usually require no treatment [52]. A recent cross-sectional study evaluated the impact of pharmacological treatment, either insulin or oral hypoglycemic agents, on glycemic control in 799 UK GCK-MODY patients, and found no difference in HbA1c between patients on or off treatment (48 vs. 46 mmol/mol; 6.5% vs. 6.4%; $p = 0.11$) [53]. Stride and colleagues performed a longitudinal study to assess the pharmacotherapeutic impact of stopping treatment in GCK-MODY patients, who had been receiving OHAs or insulin, on glycemia. They found no significant alterations in Hb1Ac between treatment statuses as well (mean change: -0.68 mmol/mol (95% CI: -2.97, 1.61)) [53].

The identification of mutations in *GCK* or *HNF4A* also has important implications for pregnancy management. Inherited mutations in *HNF4A* are associated with an 800 g increase in birth weight. Therefore, babies of affected parents

are 50% more likely to develop macrosomia [54]. Additionally, about 10% of *HNF4A* mutation carriers experience diazoxide-responsive hyperinsulinemic hypoglycemia in the first week of life [55] (Table 1). The opposite is true with respect to macrosomia in the case of GCK-MODY, as babies that do not inherit the heterozygous *GCK* allele are at risk of excessive birth weight because of increased fetal insulin secretion and insulin-mediated growth in response to maternal hyperglycemia [56, 57]. Conversely, babies that have either developed a *de novo* *GCK* mutation or inherited only one allele paternally have a birth weight reduced by nearly 500 g as a direct effect of reduced fetal insulin secretion [56, 57]. Babies that do inherit heterozygous inactivating *GCK* mutations from GCK-MODY-affected mothers are born with a normal weight as they have the same elevated set-point for glucose regulation [57].

3.2 Current MODY detection methods

Although the cost of high throughput genetic sequencing is rapidly dropping, whole exome/whole genome applications are still too expensive to sequence patients indiscriminately in the molecular diagnostic setting. Mutation screening of known MODY genes by Sanger sequencing is widely available for patients selected with classic MODY presentations, with 99% sensitivity for heterozygous base mutations [4, 6, 12, 43]. These patients are selected on the basis of clinical criteria, including early age of onset, positive family history, endogenous insulin secretion, non-obesity, and absence of pancreatic autoantibodies [6, 8, 12, 42]. Alone, these criteria are not reliable and sensitive enough, as only 50% of European cases diagnosed with MODY satisfy them, which means that we are still missing a large proportion of cases [5].

Limitations associated with these selection criteria were highlighted in a recent study in which MODY genes, *HNF1A*, *HNF4A*, and *GCK*, were sequenced in a large group of European patients clinically diagnosed as having either T1D or early-onset T2D [41]. This sequencing effort showed that 25% of the subjects with apparent T1D or T2D before the age of 30 were in fact MODY cases with pathogenic variants in one of the three MODY genes. Thus, Thanabalasingham and colleagues concluded that broadening the current selection criteria to include all patients diagnosed with C-peptide-positive diabetes for molecular diagnostic sequencing before the age of 30 increases sensitivity, potentially sparing more misdiagnosed individuals unnecessary treatment [41].

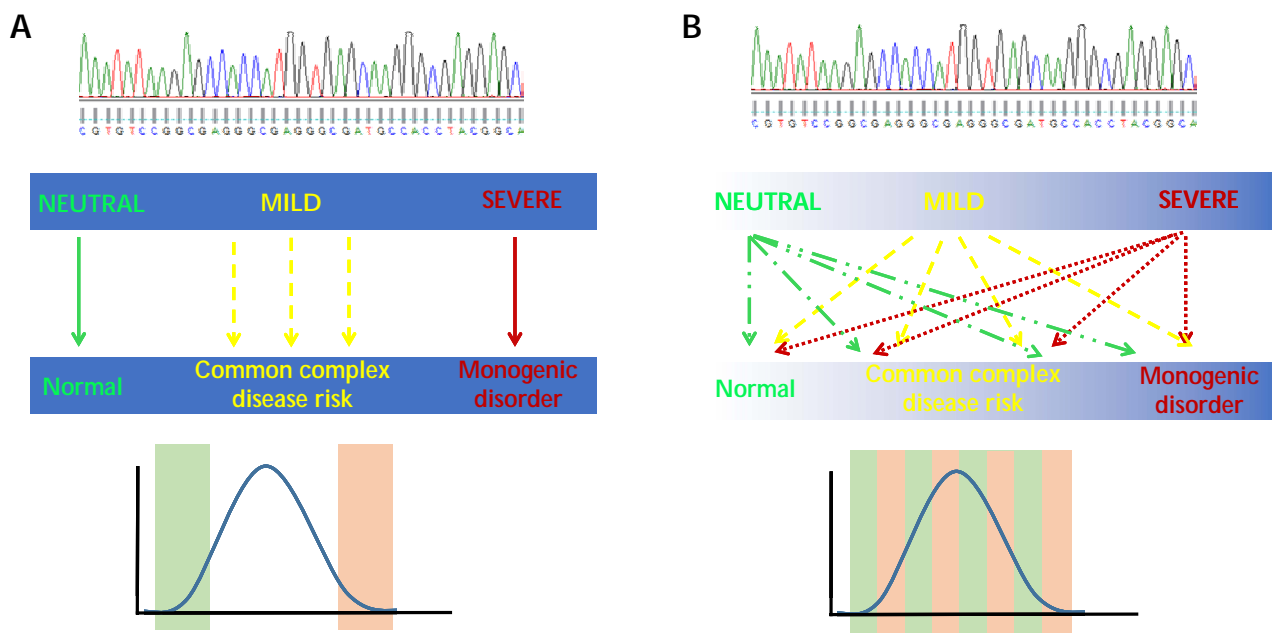


Figure 1. Relationship between mutational severity in known Mendelian disease-causing genes and clinical phenotype. (A) Traditional, deterministic understanding of Mendelian disease genetics based on biased modes of sample ascertainment. **(B)** Spectrum-based understanding of Mendelian disease genetics, as revealed by large unbiased population-wide resequencing efforts. This approach encompasses scenarios of incompletely penetrant, monogenic, disease-causing alleles in healthy individuals and, for example, monogenic disorders falsely attributed to a neutral variant in a known risk gene, while the disorder is actually caused by an allele in a different gene.

3.3 How biomarkers help to find and classify MODY

Efforts centered on biomarker utility have emerged within the last 5 years, with the aim of refining the imperfect MODY selection criteria that qualify patients for genetic screening. Glucose increments in oral glucose tolerance tests can be used to discriminate between HNF1A and GCK-MODY [6]. Patients with HNF1A-MODY show larger glucose increments (>5 mmol/l) as a consequence of a more compromised insulin secretory capacity compared to GCK-MODY where glucose increments are modest (< 3 mmol/l) [10]. In a recent study, Pal and colleagues evaluated the discriminatory capacity of serum 1,5 anhydroglucitol (1,5 AG) levels in UK subjects with HNF1A-MODY, GCK-MODY, T1D, T2D, and latent autoimmune diabetes in adults [58]. They found that 1,5 AG levels were significantly higher in GCK-MODY than in any other group ($p \leq 0.0003$), with greatest discriminatory accuracy for GCK-MODY from HNF1A-MODY subjects (C-statistic: 0.86; 86% sensitivity and 84% specificity for 1,5 AG >7.5 µg/ml) [58].

Findings from genome-wide association studies (GWAS) have inspired research into the potential of another set of candidate biomarkers, namely highly sensitive C-reactive protein (hsCRP) in serum and glycan profiles, to discriminate HNF1A-MODY from other types of diabetes [59-63]. Levels of hsCRP were initially tested in a small UK cohort, and then replicated in two large European studies, which included about 700 HNF1A-MODY cases [60-62]. These studies showed that hsCRP levels are significantly reduced in HNF1A-MODY cases, which is consistent with the fact that the *CRP* gene is under the regulation of HNF1A, and that HNF1A-MODY cases could be discriminated from other types of diabetes [60-62]. In fact, levels of hsCRP discriminated HNF1A-MODY from T2D with high sensitivity and specificity as studies have shown that 70-80% of HNF1A-MODY patients have <0.5 mg/l hsCRP, whilst that was the case in only <20% of type 2 diabetics [60, 61, 64]. GWAS efforts revealed that HNF1A also plays a role in posttranslational glycoprotein modifications, in particular via upregulating fucosylation of the antennary sections of these proteins [59]. This generated the hypothesis that loss-of-function mu-

tations in *HNF1A* must be accompanied by alterations in plasma glycan profiles, which can be used as a discriminatory diagnostic marker [63]. Indeed, a pilot study and a large replication, comparing the DG9-glycan index (the ratio of fucosylated to nonfucosylated triantennary glycans) of *HNF1A*-MODY patients with other diabetes subtypes, reported good discrimination from T1D and T2D (C-statistic ≥ 0.90) [63].

Until we are financially, technically, and analytically equipped to interpret both coding and noncoding sequence variants identified in patients unselected for extreme phenotypes, diagnostic clinical biomarkers appear to be the most attractive and cost-effective adjuncts for patient prioritization. When high-throughput, next-generation sequencing technologies are affordable and widely applied in clinics, these biomarkers will likely serve more as complementary validation, and aid in sequence interpretation and assessments of variant-level pathogenicity.

4. The challenge: when does a sequence variant cause MODY?

Focusing on the 'functional genome', the subset of the human genome enriched for rare, highly penetrant, protein-altering variants, has helped to unravel the pathogenesis of numerous monogenic disorders, and to understand the impact of rare alleles on complex disease risk [65-67]. Sufficiently powered positional cloning studies have been highly successful in identifying protein-coding variations in genes underlying inherited monogenic diseases. Nowadays, in the research setting, targeted exome sequencing has become the primary approach for readily and cost-effectively detecting rare coding variants with large effect sizes in the human genome [68, 69]. A large fraction of nonsynonymous variants detected by exome sequencing are functionally significant, and likely to be harmful as reflected by their low frequency which is due to purifying natural selection [69-71]. However, one of the major challenges associated with exome-based efforts has been to distinguish between disease-causing alleles and those that impact protein function, and between potentially functional alleles and the vast majority of neutral variants [68, 72]. Another layer of complexity is added when considering that most detected mutations are novel, i.e., neither previously reported nor characterized. In fact, 60-65% of MODY variants identified in *GCK*, *HNF1A*, and *HNF4A* are novel, and many tend to cluster geographically, making it ever more challenging to assign patho-

genicity with confidence to individual variants [6, 73].

4.1 Incomplete penetrance: monogenic disease variant, no monogenic disease

Data from whole exome-sequencing studies are increasingly blurring the genetic boundaries previously thought to be discrete between rare, Mendelian, monogenic disorders and common, complex, polygenic ones, as well as between genomes of healthy and diseased individuals (**Figure 1**) [72, 74, 75]. Empirical data from large-scale, high throughput resequencing studies have unmasked the presence of nearly 100 classic loss-of-function (LoF) variants (splice site single nucleotide variants (SNVs), insertions/deletions (indels), nonsense truncating mutations) per typical healthy human genome, mostly in the heterozygous state [68, 75]. Additional resequencing efforts have estimated up to 400 damaging mutations per individual; a number that is expected to rise as methods of rare variant ascertainment continue to improve [72].

Rare coding MODY variants in normoglycemic individuals. To date, the majority of genetic sequencing studies have drastically overestimated the penetrance of rare protein-coding variants—especially in the context of monogenic disorders—by employing inherently biased methods of sample ascertainment [74, 76]. Most variant effect size estimates have been based on analysis of the exomes of either exceptionally high-risk individuals or those who lie on the extreme ends of a phenotypic distribution (**Figure 1A**) [69, 74, 76]. Although this approach has been instrumental in identifying disease variants, it is misleading as, by default, it dismisses every scenario that does not involve full penetrance of disease alleles, i.e., it fails to capture events along the phenotypic spectrum (**Figure 1B**) [68]. Consequently, erroneous or exaggerated assignments of variant pathogenicity are made, especially in the clinical diagnostic setting, which may result in prescribing treatment and intervention regimens to an increasingly high number of low-risk individuals [74]. By studying well-phenotyped, population-based, case-control cohorts, in which cases are not preselected on the basis of satisfying classical MODY diagnostic criteria, for example, we could mitigate the upward bias, and remedy inflated effect size estimations, as has been demonstrated in the cases of *BRCA1* and *BRCA2* in breast cancer studies [74, 77].

A very recent study showed that following these design guidelines is helpful in characterizing the spectrum of low-frequency alleles in the genes under investigation in the general population [74]. By targeted sequencing of the three MODY-causing genes (*GCK*, *HNF1A*, *HNF4A*), in addition to four of the rarer genes implicated in MODY (*PDX1*, *INS*, *NEUROD1*, and *HNF1B*), in well phenotyped, randomly ascertained individuals from population-based cohorts, Flannick and colleagues found that presumed MODY-causing mutations are incompletely penetrant in normal euglycemic carriers. The variants in the healthy individuals were either annotated as causing MODY, and classified as pathogenic in the Human Gene Mutation Database (HGMD) [65], or met strict criteria for ascribing *de novo* pathogenicity by being rare, conserved, and predicted to be protein-damaging based on *in silico* tools [74].

Studies like this, which have reliably exposed putative disease-causing or harmful alleles in the genome of normal healthy humans, impose many challenges on variant interpretation and pathogenicity assignments. Most notably, such findings refute the notion that Mendelian disease variants are simpler to interpret—particularly in the medical sequencing setting—because of their deterministic nature. On the contrary, classic Mendelian single gene mutations are as likely to exert mild or neutral effects as common polymorphisms that influence phenotypic diversity or, collectively, common complex diseases (**Figure 1B**). As has been shown, rare, protein-altering, and possibly pathogenic variants in *HNF1A*, for instance, do not necessarily give rise to a MODY phenotype; they can also either contribute to T2D predisposition or exist innocuously in the genomes of healthy individuals.

Low-frequency-coding MODY variants associated with increased T2D risk. A recent study by the SIGMA consortium also nicely demonstrated the complexities outlined above. This study was among the first systematic, high-resolution screens for sequence variation in protein-coding regions in a large population unselected for an extreme monogenic clinical phenotype. The consortium reported the detection of a low-frequency missense-coding variant in *HNF1A* (c.1522G>A (p.E508K)) after sequencing whole exomes of Mexican and US Latino T2D cases (n = 1794) and diabetes-free controls (n = 1962) [78]. This variant was observed in 2.1% of cases and 0.36% non-diabetic controls (OR 5.48; 95% CI 2.83-10.61; $p = 4.4 \times 10^{-7}$); it was significantly associated with a 5-fold increase in T2D

prevalence for carriers, and a frequency of 1% in the population [78]. The association of the p.E508K variant in *HNF1A* with a T2D phenotype was replicated in the T2D-GENES (Type 2 Diabetes Genetic Exploration by Next-Generation Sequencing in Multi-Ethnic Samples) Latino group and in a *de novo* genotyping data set from self-identified indigenous Mexicans (DMS2) (OR 4.16; 95% CI 2.93-8.38). There was no evidence of p.E508K in exomes from large multi-ethnic data sets, other than in Latino samples from these consortia cohorts, confirming its exclusivity to Mexican/Mexican American ethnic groups [78].

Taken together, these findings showcase the genetic complementarity between rare and common disorders. Indeed, well-studied Mendelian genetics (as in MODY) can invaluablely inform and accelerate our understanding of common complex disorders (such as T2D), for which genome-wide association efforts have provided extensive catalogues of common variants with low to modest effect sizes. Importantly, the above study addresses some of the past limitations associated with evaluating the implication of variants in disease. For instance, we could not have concluded that the role of the p.E508K *HNF1A* allele in the pathogenesis of MODY is minor (i.e. not sufficient to cause MODY in isolation) without the detection of this allele in a state of incomplete penetrance in individuals with elevated T2D risk and the biological follow-up provided by the study group [78].

With respect to allelic penetrance, it is also worth mentioning that the identified variant lies in exon 8 which is only present in the longest of the three *HNF1A* genetic isoforms, termed *HNF1A(A)*. Harries and colleagues have shown that the *HNF1A* isoforms exhibit differential expression, with *HNF1A(A)* (exons 1-10) being the major isoform in liver and kidney, but the minor one in the pancreatic islet, where expression of the two shorter isoforms, *HNF1A(B)* (exons 1-7) and *HNF1A(C)* (exons 1-6), dominates [79]. The same study revealed variation in age of onset of *HNF1A*-MODY as a consequence of mutation position. Individuals with mutations in exons 8-10, present exclusively in *HNF1A(A)*, developed MODY about 7 years later than those with mutations harbored by the first 6 exons present in all three isoforms [79]. The penetrance of the p.E508K Latino variant could be influenced by this isomeric expression trend, which poses an additional challenge to the interpretation of the functional relevance and clinical significance of *HNF1A* alleles. For example, a protein-altering variation in *HNF1A(A)* may not necessarily translate to a beta-cell effect, yet it

may nonetheless show clear effects on liver biomarkers, such as hsCRP and glycosylation, which have a unique discriminatory capacity for HNF1A-MODY. In such instances, it is important to consider mutation position and isoform-specific expression patterns alongside biochemical phenotypes for making well-informed evaluations of variant pathogenicity.

4.2 Assessing variant causality: contextualize with caution

In the light of these and other emerging challenges, MacArthur and colleagues have recommended a set of guidelines to assist investigators in cautiously and reliably defining the role of sequence variants in human diseases [68]. In essence, these guidelines call for applying more systematic and unbiased approaches to identification, molecular characterization, and definition of the clinical relevance of sequence variants implicated in disease, in particular those of large effects, influencing both severe monogenic disorders and complex disease risk (**Figure 1**) [68]. As previously mentioned, acquiring translational value from sequencing studies requires the random recruitment of a large number of participants. Ideally, the study cohorts consist of multiple ethnicities with matched controls for whom clinical phenotypic data are available. This procedure enables a decreased impact of population stratification. Furthermore, a rigorous assessment of conservation, novelty, and frequency should be performed on the variant. Family-based segregation data and detection of the same *de novo* mutation in multiple individuals from the same cohort are necessary criteria for determining causality, but they should not be considered as definitive [68]. The latter are rare in the known MODY-causal genes, *GCK*, *HNF1A*, and *HNF4A*, but expected to increase considerably as more and more individuals unselected for a positive family history and classic MODY phenotype are sequenced in the research and clinical diagnostic settings [73]. Of note, *a priori* functional and clinical information is problematic as it is rife with false positive findings; it is thus inconclusive and often misleading. Therefore, with each investigation, researchers should conduct a host of complementary functional and statistical analyses, explicitly report results by highlighting limitations and ambiguities, and update relevant mutational catalogues and variant databases (**Figure 2**) [68]. Transparency in data sharing, critical evaluation of variant-level implication in disease, confirmatory experimental assays, and

an evaluation of pharmacotherapeutic targetability of variant-harboring proteins, will help to accelerate annotation efforts and to realize the translational potential of population-sequencing endeavors [68, 80].

5. Establishing the functional consequences of variants in known MODY genes

Our ability to identify disease-relevant sequence variants exceeds our ability to interpret their true biological effects by far. This so-called interpretative gap is a major roadblock in the path bridging basic research to the clinical arena [81]. Genetic findings are rife with false positives, variants of unknown clinical significance, and sloppy assessments of disease liability. Proper annotation and functional follow-up of identified sequence variants, in ways that allow us to experimentally validate disruption and phenotypic recapitulation, is paramount to closing the gap (**Figure 2**) [68, 81]. Defining the actionable potential of sequence alleles and making accurate assignments of pathogenicity require employment of comprehensive and scalable molecular characterization pipelines. These must be meticulously designed and customized to the disease genes and disease phenotype under interrogation. For example, it would be futile to assess the function of naturally occurring HNF1A-MODY variants in a mouse model, because heterozygous *HNF1A* mutations have no effect on glycemic status in mice (in contrast to humans) as their insulin-secretory function and glucose tolerance are normal [82]. Even within the mouse, there are inconsistencies with phenotypic consequences in humans when comparing global and tissue-specific genetic manipulations [83]. As such, decisions regarding the most suitable model systems—cell lines or animals—depend primarily on which most effectively recapitulate human disease pathology.

5.1 Using *in vitro* assays to uncover the functional impact of mutant alleles

For MODY genes, *HNF1A* and *GCK*, there are sets of well-established, multi-tiered, reproducible assays to functionally characterize mutations identified in diabetic patients. For transcription factor MODY, experiments designed to interrogate the capacity for DNA binding and transactivation are most valuable. In the case of the gene encoding glucokinase enzyme, taking identified variants

through a series of kinetic analyses is the most relevant approach for making sound extrapolations regarding function and protein stability [22, 84].

A study by Galán and colleagues showed the various *in vitro* techniques that can be utilized to characterize the molecular function of HNF1A proteins. They also characterized *HNF1A* variants detected in MODY probands; these variants exhibited co-segregation with the disease phenotype, and were absent in healthy controls [85]. They identified a number of missense variants from an HNF1A-specific mutation screen in patients who were selected for familial early-onset hyperglycemia. These variants were recreated via site-directed mutagenesis, and transfected along with wild-type HNF1A cDNA constructs into cell lines, Min6 (mouse insulinoma-derived pancreatic beta-cells) and Cos7 (monkey kidney-derived). Since HNF1A regulates the expression of transcripts in pancreatic islets and hepatocytes, they co-transfected different tissue-specific target promoters in each cell line to test the ability of the mutants to drive reporter gene expression [85]. These were either promoters expressed exclusively in pancreatic beta cells (HNF-4A P2 promoter and insulin), liver (beta-fibrinogen), or in both tissue types (GLUT-2). The direct impact of the clinical variants on HNF1A protein expression, stability, and function was thoroughly evaluated using immunoblot (protein expression), luciferase reporter (transactivation) analyses, and electrophoretic mobility shift assays (EMSA) (DNA binding) [85].

There is much value in performing complementary assays and testing multiple variables in parallel, in particular when the aim is to acquire an objective and systematic understanding of the functional consequences of sequence variants. For example, although Min6 cells, which express endogenous HNF1A protein, are more physiologically similar to pancreatic beta-cells, the endogenous protein may mask the effect of transfected mutants in a way that would not naturally occur in carriers [85]. For this reason, it is also worth assessing the mutants in a system that is less noisy, such as Cos7 cells, which lack endogenous *HNF1A*, thereby providing an opportunity for more straightforward interpretations of mutant effects. Galán and colleagues found that DNA binding and transactivation analyses correlated across different promoters and cell types. However, the *HNF1A* variants exerted differential, often unexpected effects, which led to speculation that these effects might underlie the complex and progressive nature of the HNF1A-MODY phenotype.

The study which identified the rare nonsynonymous missense *HNF1A* allele, p.E508K, through exome sequencing of a large Mexican/Mexican American case-control cohort, exploited similar methods to annotate the variant functionally [78]. As p.E508K lies within the transactivation domain of HNF1A, investigators from the SIGMA Type 2 Diabetes consortium performed a thorough set of transcriptional activation assays. These assays compared the transfected mutant's ability to induce luciferase reporter gene expression via multiple promoter sites (HNF-4A P2, GLUT-2, rat albumin) and two different cell lines (HeLa and Min6), against wild-type HNF-1A, three known HNF1A-MODY mutations (p.P447L, p.R229Q, p.P379fsdelCT), and a wild-type-like-functioning HNF1A mutant reported in a single patient with T2D (p.M490T) [78]. Consistently, across all promoters and cell types, they found that the variant was less transcriptionally active than the wild-type protein ($p < 0.0001$) and p.M490T. However, the variant showed much greater transactivation than the three variants previously identified in HNF1A-MODY families [78]. In their EMSA DNA-binding analysis, the ability of the p.E508K mutant protein to bind rat albumin promoter was indistinguishable from that of the wild-type protein. A third complementary approach to characterize the function of the mutant protein was to assess the intracellular localization in the cell lines, using a previously reported cytosol-retained variant, p.Q466X, as control [78]. Based on these experimental and statistically evaluated data on a single variant, along with phenotypic characteristics of the p.E508K carriers from a large population, a more confident assessment of pathogenicity is possible. Another benefit of this study is the specification of limitations, including imperfect exome coverage, and avoidance of an ascertainment bias which allowed for a reliable assessment of allelic penetrance (**Figure 2**) [78]. Overall, the authors were able to conclude that p.E508K impacts protein function, albeit not as severely as other *bona fide* HNF1A-MODY mutations. This is in line with the association of the variant with T2D risk elevation in the studied population, rather than MODY causation.

There is also a well-established series of biochemical assays for kinetically characterizing variants identified in GCK-MODY cases. Of the 620 MODY mutations which have been identified and reported in *GCK*, about 80 of them have been functionally characterized [22]. For MODY-specific heterozygous inactivating mutations in *GCK*, the enzyme's maximal specific activity (K_{cat}) is usually

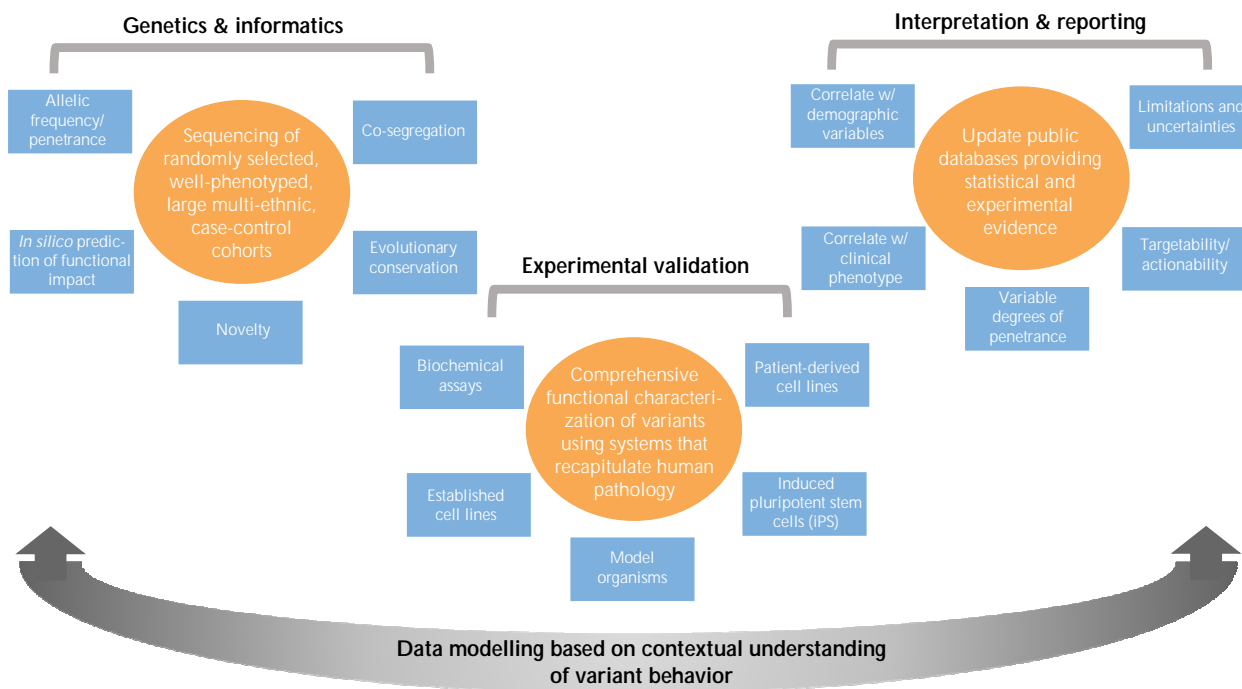


Figure 2. Recommended guidelines for a systematic and objective approach to identifying, characterizing, and assessing the pathogenicity of rare protein-coding genetic variants.

decreased accompanied by various trends in $S_{0.5}$ (glucose affinity at half maximal activity), ATP K_m (ATP concentration at half maximal activity), and Hill coefficient values [22]. The phosphorylating capacity of the glucokinase enzyme is normally diminished in proteins harboring MODY-associated variants. This is a functional defect reflected by the mild hyperglycemia in GCK-MODY patients, which is due to reduced glucose processing [86]. However, kinetic dysfunction *in vitro* does not directly correlate with the clinical phenotype, as the wild-type *GCK* allele in mutation-carrying GCK-MODY patients plays a compensatory role, and is upregulated by glucose at a posttranslational level [87]. Other enzyme characteristics such as thermostability and binding with glucokinase regulatory protein (GKRP) often explain mutational mechanisms in cases where kinetic data are inconsistent [84, 88-90]. GCK-MODY mutations have been catalytically assessed in physiologically relevant cell lines such as Min6 [90], and the enzyme's kinetics have been characterized in pancreatic islets as a function of glucose concentration [87]. Moreover, unlike HNF1A-MODY mutations, global and beta-cell-specific heterozygous knockouts of *GCK* in mice recapitulate the moder-

ate early-onset diabetic phenotype and reduced glucose-stimulated insulin secretory response in GCK-MODY patients [91-93].

Adopting careful and comprehensive approaches that combine clinical, statistic, bioinformatic, and experimental validation, as illustrated in **Figure 2**, is key to ascribing sequence-level pathogenicity with high accuracy. Transparent reporting of findings is equally pivotal, as this allows medics and researchers to make independent interpretations and assessments of the published data and the way in which they were generated. Recent familial/functional studies on glucokinase have shown that analyzing variants, which were previously reported and characterized as pathogenic in this combinatorial manner, sometimes showed the opposite effect [94, 95]. Steele and colleagues found that a heterozygous *GCK* missense variant, p.T342P, which was previously reported as pathogenic loss-of-function, neither demonstrated impaired catalytic activity nor co-segregation with the disease phenotype in the two families in which it was identified [95]. The variant was also classified as benign by three different *in silico* prediction tools, SIFT, Polyphen, and Mutation Taster [95]. The p.T342P substitution was

not inherited by relatives with an overtly diabetic or elevated fasting glucose phenotype, and other relatives in whom the allele was identified were euglycemic (<5.5 mmol/l fasting glucose) [95].

Another molecular and functional effort, which aimed to characterize 6 heterozygous missense variants identified in combination in 3 individuals referred for GCK-MODY genetic testing, also refuted some classically held notions regarding interpretation and pathogenicity assignment of heterozygous inactivating missense *GCK* variants [94]. Beer and colleagues found that two mutations, p.R43H and p.G68H, present in *trans* and independently reported to be GCK-MODY causal, were associated with a moderate hyperglycemic phenotype rather than permanent neonatal diabetes mellitus (PNDM), the disease phenotype typically associated with *trans* GCK-inactivating mutations [94]. A range of functional assessments, including kinetic and thermostability tests as well as evaluations of the behavior of mutation-harboring GCK in the presence of glucokinase activators and GKRPs, revealed that R43H-GCK instability accounted for the pathogenic effect, although the effect was catalytically wild-type-like. Whereas, p.G68D, which was inherited from an unaffected mother, neither displayed instability nor abnormal kinetics other than mild activation [94]. The familial and functional evidence of p.G68D neutrality challenges previous reports of p.G68D pathogenicity. In sum, these studies emphasize the need for thorough and critical evaluation of identified mutations before defining and reporting these as clinically significant and likely to be disease-relevant.

5.2 A new dawn for molecular characterization of mutant alleles

Human-derived cellular systems. The human pancreas remains largely inaccessible, and researchers continue to rely primarily on organ donation and transplantation systems. Therefore, human pancreatic beta-cell lines [96, 97] and patient-derived induced pluripotent stem cells (iPS) [98-100], are by far the most informative models for studying MODY and the human-specific phenotypic manifestations. Cellular phenotyping in patient-derived iPS cells provides the necessary and highly sought after contextual information (e.g. haplotypes and genetic background) which are otherwise absent from other systems or difficult to access from a technical and analytical point of view. Patient-derived stem cells also allow for reliable assessments of pharmacological drug targeting and re-

sponse, in addition to the opportunity to perform genetic rescue experiments which provide unparalleled insight into molecular mechanism and causality. In the context of MODY, most disease-causing genes play critical regulatory and developmental roles in more than one tissue type, such as *HNF1B* in the kidney and *GCK*, *HNF1A*, and *HNF4A* in the hepatocyte. Therefore, patient-derived iPS cells provide an opportunity for simultaneous investigations of tissue-specific effects via unique maturation switches [101]. The insulin-producing pancreatic beta-cells, recently developed *in vitro* from human pluripotent stem cells (hPSCs), also provide an ideal genetic and physiological setting for such interrogations, not to mention their immense potential in cell replacement therapy [102, 103].

Hua and colleagues recently published a study demonstrating the utility of patient-derived iPS cell systems [100]. They were able to successfully generate iPS cells from skin fibroblasts of two GCK-MODY patients heterozygous for hypomorphic missense variants (p.G299R and p.E256K). In this study, late differentiation and maturation of these cells into a C-peptide-positive human pancreatic beta-cell lineage, expressing all of the appropriate genetic markers, was achieved *in vivo* by transplanting them under the kidney capsules of immune-deficient mice [100]. Intraperitoneal glucose tolerance tests (IGTT) were performed on the transplanted mice and, as observed in humans, mice with GCK-MODY patient-derived beta-cells demonstrated lower responsiveness to elevated glucose levels, with otherwise normal insulin response to other secretagogues, and normal production and processing of insulin. Interestingly, cells that did not undergo maturation *in vivo* (only differentiated *in vitro*) failed to secrete insulin in response to elevated ambient levels of glucose, hinting at the importance of replicating a physiological, three-dimensional environment [100]. The *GCK* locus in the GCK-MODY patient-derived cells was rescued by homologous recombination, effectively restoring glucose responsiveness to that observed in controls. These findings shed light on the molecular mechanism by which glucokinase deficiency specifically contributes to a defective beta-cell phenotype. Further gene dosage analyses under the same conditions, looking at homozygous biallelic *GCK* mutations, yielded a markedly low number of beta-cells, suggesting an important role in beta-cell proliferation.

Teo and colleagues also describe the *in vitro* development of human iPS cells derived from skin biopsies of MODY patients, each with a mutation

in *HNF1A*, *HNF1B*, *HNF4A*, *GCK*, or *CEL*. These were morphologically and functionally identical to human embryonic stem cells and other human pluripotent stem cells, expressing pluripotency markers OCT4, SOX2, NANOG, SSEA-4, and TRA-1-60 [101]. Since then, a study was published in which investigators performed a thorough characterization of reprogrammed human iPS cells derived from primary fibroblasts of patients with HNF1A-MODY [98]. As more findings emerge from these unique pre-clinical disease-modeling systems, we will be able both to confirm and enhance our knowledge of the pathophysiological basis of MODY and the precise biological mechanisms underlying beta-cell dysfunction in the various subtypes.

Another successful patient-derived stem cell endeavor was described, although it recapitulated specifically the phenotype of a rare syndromic form of diabetes known as Wolfram syndrome (MIM222300) [99, 104]. Autosomal recessive or compound heterozygous mutations in the wolframin gene, *WFS1*, which encodes an ER transmembrane protein, give rise to Wolfram syndrome [104]. The most prominent clinical features of Wolfram syndrome include juvenile insulin-dependent diabetes, optic atrophy, and deafness [105, 106]. Shang and colleagues created insulin-producing C-peptide-positive beta-cells from skin fibroblasts of patients with Wolfram syndrome and control subjects [99]. After differentiation, the mutations in *WFS1* inducing the insulin-dependent diabetic phenotype in the Wolfram patients yielded characteristically low insulin content and increased activity of the unfolded protein response (UPR) [99]. Upon exposing the cells to the chemical chaperone, 4-phenyl butyric acid (4PBA), UPR pathway activity was reduced and insulin content was adjusted to levels similar to those in control cells, suggesting 4PBA as a novel therapeutic candidate [99]. Moreover, Shang and colleagues performed sophisticated *WFS1* locus rescue experiments in patient-derived beta-cells by introducing wild-type *WFS1* cDNA via lenti-viral transduction and antibiotic selection of cells expressing stably integrated sequence. This procedure restored insulin levels and maintained normal insulin response to secretagogues, even under conditions of endoplasmic reticulum stress, in which cells otherwise demonstrated impaired insulin processing [99].

High-throughput mutation screening and genome-editing applications. The patient-tissue-based techniques outlined above are indeed ideal for candidate gene and variant analyses. However, we

are still in the early stages of exploiting these methods efficiently, and they remain at low throughput and limited in scale. Until they are optimized, scalable, and better understood, high-throughput screens currently represent the most attractive approach for functionally interrogating large numbers of alleles at once [68]. Melnikov and colleagues have reported on multiple validated methodologies for massively parallel high-throughput mutational scanning of single-substitution libraries of tens of thousands of regulatory and protein-coding variants [107, 108]. A recent study on rare nonsynonymous missense variants identified in *PPARG* from exome sequencing of a large multi-ethnic case-control type 2 diabetes cohort implemented a novel quantitative high-throughput differentiation assay [109].

Majithia and colleagues combined bioinformatic prediction, systematic molecular characterization, drug response, and clinical phenotypic data to aid the interpretation of variants in *PPARG*, and accurately determine their association with type 2 diabetes risk. The variants identified in this study were packaged into lenti-viral constructs to ensure efficient integration by transduction, and were evaluated on the basis of their ability to rescue differentiation of human preadipocytes *in vitro* [109]. The investigators evaluated the impact of the variants on differentiation by combining high-content microscopy and customized image analysis [109]. Interestingly, they found that individuals from their cohort who were heterozygous carriers of variants in *PPARG*, which showed significantly reduced function in their characterization pipeline, had a sevenfold increase in T2D risk. In contrast, when the investigators analyzed all of the rare variants identified in *PPARG*, without filtering for *in vitro* dysfunction, they did not observe a significant association with elevated T2D risk [109].

A study published earlier this year employed similar techniques for characterizing variants in low-density lipoprotein receptor (*LDLR*) [110]. Thormaehlen and colleagues leveraged the exomes of a large European cohort of early-onset myocardial infarction (MI) cases and MI-free controls to identify rare missense variants in *LDLR* associated with premature MI. Their aim was to discriminate disruptive, functional, and potentially pathogenic alleles from the majority of neutral missense variants [110]. This was achieved by an approach that involved correlation between multiple complementary datasets: *in silico* prediction, locus-specific *a priori* clinical information, systematic cell-based functional profiling of identified alleles, and available phenotypic information. The

impact of naturally occurring missense alleles on LDLR protein function was assessed *in vitro* via uptake of fluorescently labeled LDL using microscopy coupled with a customized quantitative multiparametric image analysis pipeline [110].

While these large-scale, high-throughput functional studies are still in a proof-of-concept stage, requiring rigorous validation from phenotypic, familial co-segregation and bioinformatic data, most investigators rely on high-resolution microscopy coupled with multi-parametric, automated image analysis tools [109, 110]. Looking ahead, and scaling up to thousands of *in silico*-designed variant libraries [108], we predict enormous potential for cell-sorting techniques optimized to discriminate on the basis of mutational severity coupled with next-generation sequencing technology of tagged libraries. This is especially promising for transcription factor MODY, which can be easily interrogated based on the intensity of fluorophore reporters in biologically relevant cell lines.

The future is even brighter for genome-scale editing and functional screening via RNA-guided CRISPR (clustered regularly interspaced short palindromic repeat)-associated Cas9 nuclease technology [111-113]. CRISPR-Cas9-mediated screens allow for a wide variety of targeted, sequence-specific perturbations to achieve unbiased complementary cellular phenotype readouts with incredibly high efficiency, including knockdown, knockout, and activation [113]. Indeed, current *in vitro* methods limit biological interrogations that pertain to systemic variability. Genomic editing by Cas9 provides an opportunity to answer some of the most crucial experimental and clinical questions, such as whether the cellular phenotype from a patient-derived sample, for instance, is solely attributable to the effect of a sequence variant, or perhaps influenced by regulatory modifiers or the greater genetic background upon which the variant is located. Moreover, with CRISPR-Cas9 technology, *in vivo* levels of gene dosage and expression can be easily and efficiently replicated unlike the majority of artificial cell-based manipulations, which generally involve overexpression or partial silencing of the gene(s) of interest.

Another layer of variant-level inquiry via Cas9 could potentially involve parallel comparisons of mutation rescue between a mutation-harboring, patient-derived system and an artificial CRISPR-engineered replica. The seemingly unlimited application of Cas9-mediated genome editing is indeed able to substantially accelerate translational efforts on the genomic and precision medicine front, though not without simultaneously unraveling a host of new and unexpected molecular and functional genomic phenomena.

6. Conclusions

The rapid reduction in the cost of high-throughput genomic technologies is inevitably going to be associated with increased application and data generation in the research and clinical arena alike, uncovering a host of new analytical and ethical obstacles. Already, genetic sequencing data have refuted the longstanding notion that Mendelian disease variants pose less of an interpretive challenge than variants in polygenic diseases. The genetics of MODY perfectly exemplify the extent to which variant-phenotype mapping is more complex and nuanced in the monogenic disease context than previously assumed.

Careful interpretation of sequence data by accounting for demographic variables, molecular function in relevant experimental systems, clinical and biochemical phenotypes, and allelic penetrance, collectively assist in making more reliable assessments of variant pathogenicity. Developing such a contextual understanding of mutation behavior will better position scientists and clinicians to evaluate the implication of novel and incidentally detected variants in known disease genes, especially in the medical sequencing setting.

Acknowledgments: ALG is a Wellcome Trust Senior Research Fellow in Basic Biomedical Research. Work in the Gloyd lab is funded by the Wellcome Trust (095101/Z/10/Z) and the Oxford NIHR Biomedical Research Centre.

Disclosures: The authors report no conflict of interests.

References

1. **Tattersall RB.** Mild familial diabetes with dominant inheritance. *Q J Med* 1974. 43(170):339-357.
2. **Tattersall RB, Fajans SS.** A difference between the inheritance of classical juvenile-onset and maturity-onset type diabetes of young people. *Diabetes* 1975. 24(1):44-53.
3. **Ledermann HM.** Is maturity onset diabetes at young age (MODY) more common in Europe than previously assumed? *Lancet* 1995. 345(8950):648.
4. **Thanabalasingham G, Owen KR.** Diagnosis and management of maturity onset diabetes of the young (MODY). *BMJ* 2011. 343:d6044.
5. **Shields BM, Hicks S, Shepherd MH, Colclough K, Hattersley AT, Ellard S.** Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia* 2010. 53(12):2504-2508.
6. **Colclough K, Saint-Martin C, Timsit J, Ellard S, Bel-**

- lanne-Chantelot C. Clinical utility gene card for: Maturity-onset diabetes of the young. *Eur J Hum Genet* 2014. 22(9):E1-E6.
7. Fajans SS, Bell GI, Polonsky KS. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N Engl J Med* 2001. 345(13):971-80.
 8. Murphy R, Ellard S, Hattersley AT. Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nat Clin Pract Endocrinol Metab* 2008. 4(4):200-213.
 9. Pihoker C, Gilliam LK, Ellard S, Dabelea D, Davis C, Dolan LM, Greenbaum CJ, Imperatore G, Lawrence JM, Marcovina SM, et al. Prevalence, characteristics and clinical diagnosis of maturity onset diabetes of the young due to mutations in HNF1A, HNF4A, and glucokinase: results from the SEARCH for Diabetes in Youth. *J Clin Endocrinol Metab* 2013. 98(10):4055-4062.
 10. McDonald TJ, Ellard S. Maturity onset diabetes of the young: identification and diagnosis. *Ann Clin Biochem* 2013. 50(Pt 5):403-415.
 11. Bingham C, Hattersley AT. Renal cysts and diabetes syndrome resulting from mutations in hepatocyte nuclear factor-1beta. *Nephrol Dial Transplant* 2004. 19(11):2703-2708.
 12. Ellard S, Bellanne-Chantelot C, Hattersley AT, European Molecular Genetics Quality Network MODY Group. Best practice guidelines for the molecular genetic diagnosis of maturity-onset diabetes of the young. *Diabetologia* 2008. 51(4):546-553.
 13. Hattersley AT. Maturity-onset diabetes of the young: clinical heterogeneity explained by genetic heterogeneity. *Diabet Med* 1998. 15(1):15-24.
 14. Froguel P, Vaxillaire M, Sun F, Velho G, Zouali H, Butel MO, Lesage S, Vionnet N, Clement K, Fougerousse F, et al. Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. *Nature* 1992. 356(6365):162-164.
 15. Hattersley AT, Turner RC, Permutt MA, Patel P, Tanizawa Y, Chiu KC, O'Rahilly S, Watkins PJ, Wainscoat JS. Linkage of type 2 diabetes to the glucokinase gene. *Lancet* 1992. 339(8805):1307-1310.
 16. Velho G, Froguel P, Clement K, Pueyo ME, Rako-toambinina B, Zouali H, Passa P, Cohen D, Robert JJ. Primary pancreatic beta-cell secretory defect caused by mutations in glucokinase gene in kindreds of maturity onset diabetes of the young. *Lancet* 1992. 340(8817):444-448.
 17. Vaxillaire M, Boccio V, Philippi A, Vigouroux C, Terwilliger J, Passa P, Beckmann JS, Velho G, Lathrop GM, Froguel P. A gene for maturity onset diabetes of the young (MODY) maps to chromosome 12q. *Nat Genet* 1995. 9(4):418-423.
 18. Yamagata K, Oda N, Kaisaki PJ, Menzel S, Furuta H, Vaxillaire M, Southam L, Cox RD, Lathrop GM, Boriraj VV, et al. Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). *Nature* 1996. 384(6608):455-458.
 19. Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI. Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* 1996. 384(6608):458-460.
 20. Fajans SS, Bell GI, Bowden DW, Halter JB, Polonsky KS. Maturity-onset diabetes of the young. *Life Sci* 1994. 55(6):413-422.
 21. Bell GI, Xiang KS, Newman MV, Wu SH, Wright LG, Fajans SS, Spielman RS, Cox NJ. Gene for non-insulin-dependent diabetes mellitus (maturity-onset diabetes of the young subtype) is linked to DNA polymorphism on human chromosome 20q. *Proc Natl Acad Sci U S A* 1991. 88(4):1484-1488.
 22. Osbak KK, Colclough K, Saint-Martin C, Beer NL, Bellanne-Chantelot C, Ellard S, Gloyon AL. Update on mutations in glucokinase (GCK), which cause maturity-onset diabetes of the young, permanent neonatal diabetes, and hyperinsulinemic hypoglycemia. *Hum Mutat* 2009. 30(11):1512-1526.
 23. Colclough K, Bellanne-Chantelot C, Saint-Martin C, Flanagan SE, Ellard S. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha and 4 alpha in maturity-onset diabetes of the young and hyperinsulinemic hypoglycemia. *Hum Mutat* 2013. 34(5):669-685.
 24. Chakera AJ, Spyer G, Vincent N, Ellard S, Hattersley AT, Dunne FP. The 0.1% of the population with glucokinase monogenic diabetes can be recognized by clinical characteristics in pregnancy: the Atlantic Diabetes in Pregnancy cohort. *Diabetes Care* 2014. 37(5):1230-1236.
 25. Costa A, Bescos M, Velho G, Chevre J, Vidal J, Ses-milo G, Bellanne-Chantelot C, Froguel P, Casamitjana R, Rivera-Fillat F, et al. Genetic and clinical characterisation of maturity-onset diabetes of the young in Spanish families. *Eur J Endocrinol* 2000. 142(4):380-386.
 26. Lorini R, Klersy C, d'Annunzio G, Massa O, Minuto N, Iafusco D, Bellanne-Chantelot C, Frongia AP, Toni S, Meschi F, et al. Maturity-onset diabetes of the young in children with incidental hyperglycemia: a multi-center Italian study of 172 families. *Diabetes Care* 2009. 32(10):1864-1866.
 27. Velho G, Blanche H, Vaxillaire M, Bellanne-Chantelot C, Pardini VC, Timsit J, Passa P, Deschamps I, Robert JJ, Weber IT, et al. Identification of 14 new glucokinase mutations and description of the clinical profile of 42 MODY-2 families. *Diabetologia* 1997. 40(2):217-224.
 28. Lehto M, Wipemo C, Ivarsson SA, Lindgren C, Lipsanen-Nyman M, Weng J, Wibell L, Widen E, Tuomi T, Groop L. High frequency of mutations in MODY and mitochondrial genes in Scandinavian patients with familial early-onset diabetes. *Diabetologia* 1999. 42(9):1131-1137.
 29. Sagen JV, Bjorkhaug L, Molnes J, Raeder H, Grevle L, Sovik O, Molven A, Njolstad PR. Diagnostic screening of MODY2/GCK mutations in the Norwegian MODY Registry. *Pediatr Diabetes* 2008. 9(5):442-449.
 30. Schober E, Rami B, Grabert M, Thon A, Kapellen T, Reinehr T, Holl RW, DPV-Wiss Initiative of the German Working Group for Paediatric Diabetology. Phenotypic aspects of maturity-onset diabetes of the young (MODY diabetes) in comparison with Type 2 diabetes mellitus (T2DM) in children and adolescents: experience from a large multicentre database. *Diabet Med* 2009. 26(5):466-473.
 31. Eide SA, Raeder H, Johansson S, Midthjell K, Sovik O, Njolstad PR, Molven A. Prevalence of HNF1A (MODY3) mutations in a Norwegian population (the HUNT2 Study). *Diabet Med* 2008. 25(7):775-781.
 32. Pearson ER, Pruhova S, Tack CJ, Johansen A, Castleden HA, Lumb PJ, Wierzbicki AS, Clark PM, Lebl J, Pedersen O, et al. Molecular genetics and phenotypic characteristics of MODY caused by hepatocyte nuclear factor

- 4alpha mutations in a large European collection. *Diabetologia* 2005. 48(5):878-885.
33. **Kropff J, Selwood MP, McCarthy MI, Farmer AJ, Owen KR.** Prevalence of monogenic diabetes in young adults: a community-based, cross-sectional study in Oxfordshire, UK. *Diabetologia* 2011. 54(5):1261-1263.
 34. **O'Sullivan EP, Avalos G, O'Reilly M, Denny MC, Gaffney G, Dunne F, Atlantic DIP collaborators.** Atlantic Diabetes in Pregnancy (DIP): the prevalence and outcomes of gestational diabetes mellitus using new diagnostic criteria. *Diabetologia* 2011. 54(7):1670-1675.
 35. **Fendler W, Borowiec M, Baranowska-Jazwiecka A, Szadkowska A, Skala-Zamorowska E, Deja G, Jarosz-Chobot P, Techmanska I, Bautembach-Minkowska J, Mysliwiec M, et al.** Prevalence of monogenic diabetes amongst Polish children after a nationwide genetic screening campaign. *Diabetologia* 2012. 55(10):2631-2635.
 36. **Neu A, Feldhahn L, Ehehalt S, Hub R, Ranke MB, DIARY group Baden-Wurtemberg.** Type 2 diabetes mellitus in children and adolescents is still a rare disease in Germany: a population-based assessment of the prevalence of type 2 diabetes and MODY in patients aged 0-20 years. *Pediatr Diabetes* 2009. 10(7):468-473.
 37. **Irgens HU, Molnes J, Johansson BB, Ringdal M, Skriverhaug T, Undlien DE, Sovik O, Joner G, Molven A, Njolstad PR.** Prevalence of monogenic diabetes in the population-based Norwegian Childhood Diabetes Registry. *Diabetologia* 2013. 56(7):1512-1519.
 38. **Kousta E, Ellard S, Allen LI, Saker PJ, Huxtable SJ, Hattersley AT, McCarthy MI.** Glucokinase mutations in a phenotypically selected multiethnic group of women with a history of gestational diabetes. *Diabet Med* 2001. 18(8):683-684.
 39. **Xu JY, Dan QH, Chan V, Wat NM, Tam S, Tiu SC, Lee KF, Siu SC, Tsang MW, Fung LM, et al.** Genetic and clinical characteristics of maturity-onset diabetes of the young in Chinese patients. *Eur J Hum Genet* 2005. 13(4):422-427.
 40. **Hwang JS, Shin CH, Yang SW, Jung SY, Huh N.** Genetic and clinical characteristics of Korean maturity-onset diabetes of the young (MODY) patients. *Diabetes Res Clin Pract* 2006. 74(1):75-81.
 41. **Thanabalasingham G, Pal A, Selwood MP, Dudley C, Fisher K, Bingley PJ, Ellard S, Farmer AJ, McCarthy MI, Owen KR.** Systematic assessment of etiology in adults with a clinical diagnosis of young-onset type 2 diabetes is a successful strategy for identifying maturity-onset diabetes of the young. *Diabetes Care* 2012. 35(6):1206-1212.
 42. **Bellanne-Chantelot C, Levy DJ, Carette C, Saint-Martin C, Riveline JP, Larger E, Valero R, Gautier JF, Reznik Y, Sola A, et al.** Clinical characteristics and diagnostic criteria of maturity-onset diabetes of the young (MODY) due to molecular anomalies of the HNF1A gene. *J Clin Endocrinol Metab* 2011. 96(8):E1346-E1351.
 43. **Shepherd M, Shields B, Ellard S, Rubio-Cabezas O, Hattersley AT.** A genetic diagnosis of HNF1A diabetes alters treatment and improves glycaemic control in the majority of insulin-treated patients. *Diabet Med* 2009. 26(4):437-441.
 44. **Pouwer F, Hermanns N.** Insulin therapy and quality of life. A review. *Diabetes Metab Res Rev* 2009. 25(Suppl 1):S4-S10.
 45. **Pearson ER, Liddell WG, Shepherd M, Corral RJ, Hattersley AT.** Sensitivity to sulphonylureas in patients with hepatocyte nuclear factor-1alpha gene mutations: evidence for pharmacogenetics in diabetes. *Diabet Med* 2000. 17(7):543-545.
 46. **Pearson ER, Starkey BJ, Powell RJ, Gribble FM, Clark PM, Hattersley AT.** Genetic cause of hyperglycaemia and response to treatment in diabetes. *Lancet* 2003. 362(9392):1275-1281.
 47. **Isomaa B, Henricsson M, Lehto M, Forsblom C, Karanko S, Sarelin L, Hagblom M, Groop L.** Chronic diabetic complications in patients with MODY3 diabetes. *Diabetologia* 1998. 41(4):467-473.
 48. **Prisco F, Iafusco D, Franzese A, Sulli N, Barbetti F.** MODY 2 presenting as neonatal hyperglycaemia: a need to reshape the definition of "neonatal diabetes"? *Diabetologia* 2000. 43(10):1331-1332.
 49. **Byrne MM, Sturis J, Clement K, Vionnet N, Pueyo ME, Stoffel M, Takeda J, Passa P, Cohen D, Bell GI, et al.** Insulin secretory abnormalities in subjects with hyperglycemia due to glucokinase mutations. *J Clin Invest* 1994. 93(3):1120-1130.
 50. **Guenat E, Seematter G, Philippe J, Temler E, Jequier E, Tappy L.** Counterregulatory responses to hypoglycemia in patients with glucokinase gene mutations. *Diabetes Metab* 2000. 26(5):377-384.
 51. **Steele AM, Shields BM, Wensley KJ, Colclough K, Ellard S, Hattersley AT.** Prevalence of vascular complications among patients with glucokinase mutations and prolonged, mild hyperglycemia. *JAMA* 2014. 311(3):279-286.
 52. **Pearson ER, Velho G, Clark P, Stride A, Shepherd M, Frayling TM, Bulman MP, Ellard S, Froguel P, Hattersley AT.** beta-cell genes and diabetes: quantitative and qualitative differences in the pathophysiology of hepatic nuclear factor-1alpha and glucokinase mutations. *Diabetes* 2001. 50(Suppl 1):S101-S107.
 53. **Stride A, Shields B, Gill-Carey O, Chakera AJ, Colclough K, Ellard S, Hattersley AT.** Cross-sectional and longitudinal studies suggest pharmacological treatment used in patients with glucokinase mutations does not alter glycaemia. *Diabetologia* 2014. 57(1):54-56.
 54. **Pearson ER, Boj SF, Steele AM, Barrett T, Stals K, Shield JP, Ellard S, Ferrer J, Hattersley AT.** Macrosomia and hyperinsulinaemic hypoglycaemia in patients with heterozygous mutations in the HNF4A gene. *Plos Med* 2007. 4(4):e118.
 55. **Flanagan SE, Kapoor RR, Mali G, Cody D, Murphy N, Schwahn B, Sihanidou T, Banerjee I, Akcay T, Rubio-Cabezas O, et al.** Diazoxide-responsive hyperinsulinemic hypoglycemia caused by HNF4A gene mutations. *Eur J Endocrinol* 2010. 162(5):987-992.
 56. **Hattersley AT, Beards F, Ballantyne E, Appleton M, Harvey R, Ellard S.** Mutations in the glucokinase gene of the fetus result in reduced birth weight. *Nat Genet* 1998. 19(3):268-270.
 57. **Spyer G, Macleod KM, Shepherd M, Ellard S, Hattersley AT.** Pregnancy outcome in patients with raised blood glucose due to a heterozygous glucokinase gene mutation. *Diabet Med* 2009. 26(1):14-18.
 58. **Pal A, Farmer AJ, Dudley C, Selwood MP, Barrow BA, Klyne R, Grew JP, McCarthy MI, Gloyn AL, Owen KR.** Evaluation of serum 1,5 anhydroglucitol levels as a clinical test to differentiate subtypes of diabetes. *Diabetes Care* 2010. 33(2):252-257.

59. **Lauc G, Essafi A, Huffman JE, Hayward C, Knezevic A, Kattla JJ, Polasek O, Gornik O, Vitart V, Abrahams JL, et al.** Genomics meets glycomics—the first GWAS study of human N-Glycome identifies HNF1alpha as a master regulator of plasma protein fucosylation. *Plos Genet* 2010. 6(12):e1001256.
60. **Thanabalasingham G, Shah N, Vaxillaire M, Hansen T, Tuomi T, Gasperikova D, Szopa M, Tjora E, James TJ, Kokko P, et al.** A large multi-centre European study validates high-sensitivity C-reactive protein (hsCRP) as a clinical biomarker for the diagnosis of diabetes subtypes. *Diabetologia* 2011. 54(11):2801-2810.
61. **McDonald TJ, Shields BM, Lawry J, Owen KR, Gloyn AL, Ellard S, Hattersley AT.** High-sensitivity CRP discriminates HNF1A-MODY from other subtypes of diabetes. *Diabetes Care* 2011. 34(8):1860-1862.
62. **Owen KR, Thanabalasingham G, James TJ, Karpe F, Farmer AJ, McCarthy MI, Gloyn AL.** Assessment of high-sensitivity C-reactive protein levels as diagnostic discriminator of maturity-onset diabetes of the young due to HNF1A mutations. *Diabetes Care* 2010. 33(9):1919-1924.
63. **Thanabalasingham G, Huffman JE, Kattla JJ, Novokmet M, Rudan I, Gloyn AL, Hayward C, Adamczyk B, Reynolds RM, Muzinic A, et al.** Mutations in HNF1A result in marked alterations of plasma glycan profile. *Diabetes* 2013. 62(4):1329-1337.
64. **Owen KR.** Monogenic diabetes: old and new approaches to diagnosis. *Clin Med* 2013. 13(3):278-281.
65. **Stenson PD, Ball EV, Howells K, Phillips AD, Mort M, Cooper DN.** The Human Gene Mutation Database: providing a comprehensive central mutation database for molecular diagnostics and personalized genomics. *Hum Genomics* 2009. 4(2):69-72.
66. **Ponting CP, Hardison RC.** What fraction of the human genome is functional? *Genome Res* 2011. 21(11):1769-1776.
67. **Rands CM, Meader S, Ponting CP, Lunter G.** 8.2% of the Human genome is constrained: variation in rates of turnover across functional element classes in the human lineage. *Plos Genet* 2014. 10(7):e1004525.
68. **MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, Adams DR, Altman RB, Antonarakis SE, Ashley EA, et al.** Guidelines for investigating causality of sequence variants in human disease. *Nature* 2014. 508(7497):469-476.
69. **Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J.** Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011. 12(11):745-755.
70. **Kryukov GV, Pennacchio LA, Sunyaev SR.** Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. *Am J Hum Genet* 2007. 80(4):727-739.
71. **Zuk O, Schaffner SF, Samocha K, Do R, Hechter E, Kathiresan S, Daly MJ, Neale BM, Sunyaev SR, Lander ES.** Searching for missing heritability: designing rare variant association studies. *Proc Natl Acad Sci U S A* 2014. 111(4):E455-E464.
72. **Xue Y, Chen Y, Ayub Q, Huang N, Ball EV, Mort M, Phillips AD, Shaw K, Stenson PD, Cooper DN, et al.** Deleterious- and disease-allele prevalence in healthy individuals: insights from current predictions, mutation databases, and population-scale resequencing. *Am J Hum Genet* 2012. 91(6):1022-1032.
73. **Stanik J, Dusatkova P, Cinek O, Valentinova L, Huckova M, Skopkova M, Dusatkova L, Stanikova D, Pura M, Klimes I, et al.** De novo mutations of GCK, HNF1A and HNF4A may be more frequent in MODY than previously assumed. *Diabetologia* 2014. 57(3):480-484.
74. **Flannick J, Beer NL, Bick AG, Agarwala V, Molnes J, Gupta N, Burt NP, Florez JC, Meigs JB, Taylor H, et al.** Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes. *Nat Genet* 2013. 45(11):1380-1385.
75. **MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, et al.** A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 2012. 335(6070):823-828.
76. **Goring HH, Terwilliger JD, Blangero J.** Large upward bias in estimation of locus-specific effects from genome-wide scans. *Am J Hum Genet* 2001. 69(6):1357-1369.
77. **Begg CB.** On the use of familial aggregation in population-based case probands for calculating penetrance. *J Natl Cancer Inst* 2002. 94(16):1221-1226.
78. **SIGMA Type 2 Diabetes Consortium, Estrada K, Aukrust I, Bjorkhaug L, Burt NP, Mercader JM, Garcia-Ortiz H, Huerta-Chagoya A, Moreno-Macias H, Walford G, et al.** Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. *JAMA* 2014. 311(22):2305-2314.
79. **Harries LW, Ellard S, Stride A, Morgan NG, Hattersley AT.** Isoforms of the TCF1 gene encoding hepatocyte nuclear factor-1 alpha show differential expression in the pancreas and define the relationship between mutation position and clinical phenotype in monogenic diabetes. *Hum Mol Genet* 2006. 15(14):2216-2224.
80. **Plenge RM, Scolnick EM, Altshuler D.** Validating therapeutic targets through human genetics. *Nat Rev Drug Discov* 2013. 12(8):581-594.
81. **Cutting GR.** Annotating DNA variants is the next major goal for human genetics. *Am J Hum Genet* 2014. 94(1):5-10.
82. **Pontoglio M, Sreenan S, Roe M, Pugh W, Ostrega D, Doyen A, Pick AJ, Baldwin A, Velho G, Froguel P, et al.** Defective insulin secretion in hepatocyte nuclear factor 1alpha-deficient mice. *J Clin Invest* 1998. 101(10):2215-2222.
83. **Boj SF, van Es JH, Huch M, Li VS, Jose A, Hatzis P, Mokry M, Haegebarth A, van den Born M, Chambon P, et al.** Diabetes risk gene and Wnt effector Tcf7l2/TCF4 controls hepatic response to perinatal and adult metabolic demand. *Cell* 2012. 151(7):1595-1607.
84. **Raimondo A, Chakera AJ, Thomsen SK, Colclough K, Barrett A, De Franco E, Chatelas A, Demirbilek H, Akcay T, Alawneh H, et al.** Phenotypic severity of homozygous GCK mutations causing neonatal or childhood-onset diabetes is primarily mediated through effects on protein stability. *Hum Mol Genet* 2014. 23(24):6432-6440.
85. **Galan M, Garcia-Herrero CM, Azriel S, Gargallo M, Duran M, Gorgojo JJ, Andia VM, Navas MA.** Differential effects of HNF-1alpha mutations associated with familial young-onset diabetes on target gene regulation. *Mol Med* 2011. 17(3-4):256-265.
86. **Davis EA, Cuesta-Munoz A, Raoul M, Buettger C, Sweet I, Moates M, Magnuson MA, Matschinsky FM.** Mutants of glucokinase cause hypoglycaemia- and hyperglycaemia syndromes and their analysis illuminates fundamental quantitative concepts of glucose homeostasis. *Diabetologia*

1999. 42(10):1175-1186.
87. **Liang Y, Najafi H, Smith RM, Zimmerman EC, Magnuson MA, Tal M, Matschinsky FM.** Concordant glucose induction of glucokinase, glucose usage, and glucose-stimulated insulin release in pancreatic islets maintained in organ culture. *Diabetes* 1992. 41(7):792-806.
 88. **Burke CV, Buettger CW, Davis EA, McClane SJ, Matschinsky FM, Raper SE.** Cell-biological assessment of human glucokinase mutants causing maturity-onset diabetes of the young type 2 (MODY-2) or glucokinase-linked hyperinsulinaemia (GK-HI). *Biochem J* 1999. 342(Pt 2):345-352.
 89. **Gloyn AL, Odili S, Zelent D, Buettger C, Castleden HA, Steele AM, Stride A, Shiota C, Magnuson MA, Lorini R, et al.** Insights into the structure and regulation of glucokinase from a novel mutation (V62M), which causes maturity-onset diabetes of the young. *J Biol Chem* 2005. 280(14):14105-14113.
 90. **Arden C, Trainer A, de la Iglesia N, Scougall KT, Gloyn AL, Lange AJ, Shaw JA, Matschinsky FM, Agius L.** Cell biology assessment of glucokinase mutations V62M and G72R in pancreatic beta-cells: evidence for cellular instability of catalytic activity. *Diabetes* 2007. 56(7):1773-1782.
 91. **Postic C, Shiota M, Niswender KD, Jetton TL, Chen Y, Moates JM, Shelton KD, Lindner J, Cherrington AD, Magnuson MA.** Dual roles for glucokinase in glucose homeostasis as determined by liver and pancreatic beta cell-specific gene knock-outs using Cre recombinase. *J Biol Chem* 1999. 274(1):305-315.
 92. **Terauchi Y, Sakura H, Yasuda K, Iwamoto K, Takahashi N, Ito K, Kasai H, Suzuki H, Ueda O, Kamada N, et al.** Pancreatic beta-cell-specific targeted disruption of glucokinase gene. Diabetes mellitus due to defective insulin secretion to glucose. *J Biol Chem* 1995. 270(51):30253-30256.
 93. **Grupe A, Hultgren B, Ryan A, Ma YH, Bauer M, Stewart TA.** Transgenic knockouts reveal a critical requirement for pancreatic beta cell glucokinase in maintaining glucose homeostasis. *Cell* 1995. 83(1):69-78.
 94. **Beer NL, Osbak KK, van de Bunt M, Tribble ND, Steele AM, Wensley KJ, Edghill EL, Colcough K, Barrett A, Valentinova L, et al.** Insights into the pathogenicity of rare missense GCK variants from the identification and functional characterization of compound heterozygous and double mutations inherited in cis. *Diabetes Care* 2012. 35(7):1482-1484.
 95. **Steele AM, Tribble ND, Caswell R, Wensley KJ, Hattersley AT, Gloyn AL, Ellard S.** The previously reported T342P GCK missense variant is not a pathogenic mutation causing MODY. *Diabetologia* 2011. 54(8):2202-2205.
 96. **Ravassard P, Hazhouz Y, Pechberty S, Bricout-Neveu E, Armanet M, Czernichow P, Scharfmann R.** A genetically engineered human pancreatic beta cell line exhibiting glucose-inducible insulin secretion. *J Clin Invest* 2011. 121(9):3589-3597.
 97. **Scharfmann R, Pechberty S, Hazhouz Y, von Bulow M, Bricout-Neveu E, Grenier-Godard M, Guez F, Rachdi L, Lohmann M, Czernichow P, et al.** Development of a conditionally immortalized human pancreatic beta cell line. *J Clin Invest* 2014. 124(5):2087-2098.
 98. **Stepniewski J, Kachamakova-Trojanowska N, Ogrocki D, Szopa M, Matlok M, Beilharz M, Dyduch G, Malecki MT, Jozkowicz A, Dulak J.** Induced pluripotent stem cells as a model for diabetes investigation. *Sci Rep* 2015. 5:8597.
 99. **Shang L, Hua H, Foo K, Martinez H, Watanabe K, Zimmer M, Kahler DJ, Freeby M, Chung W, LeDuc C, et al.** beta-cell dysfunction due to increased ER stress in a stem cell model of Wolfram syndrome. *Diabetes* 2014. 63(3):923-933.
 100. **Hua H, Shang L, Martinez H, Freeby M, Gallagher MP, Ludwig T, Deng L, Greenberg E, Leduc C, Chung WK, et al.** iPSC-derived beta cells model diabetes due to glucokinase deficiency. *J Clin Invest* 2013. 123(7):3146-3153.
 101. **Teo AK, Windmueller R, Johansson BB, Dirice E, Njolstad PR, Tjora E, Raeder H, Kulkarni RN.** Derivation of human induced pluripotent stem cells from patients with maturity onset diabetes of the young. *J Biol Chem* 2013. 288(8):5353-5356.
 102. **Pagliuca FW, Millman JR, Gurtler M, Segel M, Van Dervort A, Ryu JH, Peterson QP, Greiner D, Melton DA.** Generation of functional human pancreatic beta cells in vitro. *Cell* 2014. 159(2):428-439.
 103. **Rezania A, Bruin JE, Arora P, Rubin A, Batushansky I, Asadi A, O'Dwyer S, Quiskamp N, Mojibian M, Albrecht T, et al.** Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat Biotechnol* 2014. 32(11):1121-1133.
 104. **Urano F.** Wolfram syndrome iPS cells: the first human cell model of endoplasmic reticulum disease. *Diabetes* 2014. 63(3):844-846.
 105. **Barrett TG, Bunday SE, Macleod AF.** Neurodegeneration and diabetes: UK nationwide study of Wolfram (DID-MOAD) syndrome. *Lancet* 1995. 346(8988):1458-1463.
 106. **Marshall BA, Permutt MA, Paciorekowski AR, Hoekel J, Karzon R, Wasson J, Viehover A, White NH, Shimony JS, Manwaring L, et al.** Phenotypic characteristics of early Wolfram syndrome. *Orphanet J Rare Dis* 2013. 8:64.
 107. **Melnikov A, Murugan A, Zhang X, Tesileanu T, Wang L, Rogov P, Feizi S, Gnirke A, Callan CG Jr, Kinney JB, et al.** Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat Biotechnol* 2012. 30(3):271-277.
 108. **Melnikov A, Zhang X, Rogov P, Wang L, Mikkelsen TS.** Massively parallel reporter assays in cultured mammalian cells. *J Vis Exp* 2014. 90.
 109. **Majithia AR, Flannick J, Shahinian P, Guo M, Bray MA, Fontanillas P, Gabriel SB, GoT2D Consortium, NHGRI JHS/FHS Allelic Spectrum Project, SIGMA T2D Consortium, 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 U S A* 2014. 111(36):13127-13132.
 110. **Thormaehlen AS, Schuberth C, Won HH, Blattmann P, Joggerst-Thomalla B, Theiss S, Asselta R, Duga S, Merlini PA, Ardissino D, et al.** Systematic cell-based phenotyping of missense alleles empowers rare variant association studies: a case for LDLR and myocardial infarction. *PLoS Genet* 2015. 11(2):e1004855.
 111. **Hsu PD, Lander ES, Zhang F.** Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 2014. 157(6):1262-1278.
 112. **Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM.** RNA-guided human ge-

-
- nome engineering via Cas9. *Science* 2013. 339(6121):823-826.
113. **Shalem O, Sanjana NE, Zhang F.** High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* 2015. 16(5):299-311.