Genome, Exome, and Targeted Next-Generation Sequencing in Neonatal Diabetes

Elisa De Franco, PhD, Sian Ellard, PhD, FRCPath*

KEYWORDS

• Next-generation sequencing • Gene discovery • Genetic testing • Neonatal diabetes

KEY POINTS

- Next-generation sequencing has revolutionized the approach to genetic testing and research.
- The 3 main applications of next-generation sequencing technology are targeted gene panels and exome and genome sequencing.
- Neonatal diabetes is a genetically and clinically heterogeneous disease, which means that genetic testing and research of new causes of the disease are challenging.
- A targeted gene panel has been developed to test all the known causes of neonatal diabetes in a single test. Early comprehensive testing has changed the way patients with neonatal diabetes are managed.
- Exome sequencing is a powerful tool to identify novel disease genes. In neonatal diabetes, it has led to the identification of 2 novel causes: mutations in *GATA6* and *STAT3*.
- Genome sequencing is the most comprehensive test available, and it was used to identify mutations in a novel enhancer that cause pancreatic agenesis.

INTRODUCTION TO NEONATAL DIABETES

Neonatal diabetes diagnosed before 6 months is a rare disease (approximate incidence of 1:100,000 live births[1]) that reflects severe β-cell dysfunction (**Fig. 1**). Two separate studies[2,3] have shown that diabetes diagnosed before 6 months of age is most likely to have a monogenic cause rather than being caused by autoimmunity.

The authors declare no conflicts of interest. S. Ellard is a Wellcome Trust Senior Investigator.
Institute of Biomedical and Clinical Science, University of Exeter Medical School, Barrack Road, Exeter EX2 5DW, UK
* Corresponding author. University of Exeter Medical School, RILD level 3, Royal Devon & Exeter Hospital, Barrack Road, Exeter EX2 5DW, UK.
E-mail address: sian.ellard@nhs.net

Pediatr Clin N Am 62 (2015) 1037–1053
http://dx.doi.org/10.1016/j.pcl.2015.04.012 **pediatric.theclinics.com**
0031-3955/15/$ – see front matter

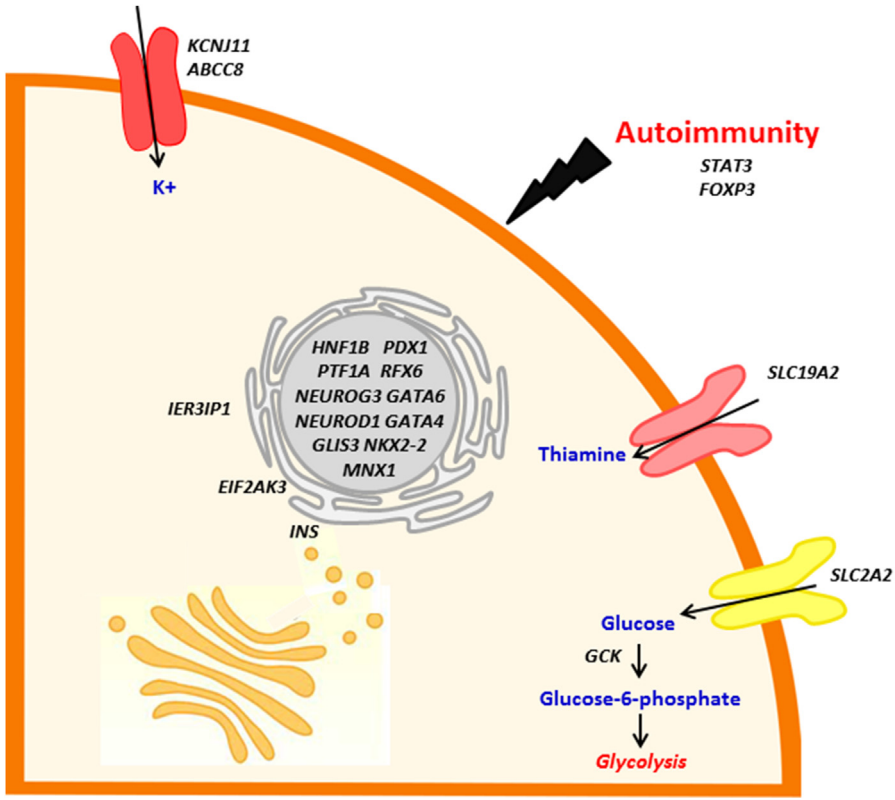

Fig. 1. The β cell and genes causing neonatal diabetes. Gene names are reported in black font. KCNJ11, ABCC8, SLC19A2, and SLC2A2 are transmembrane channels. FOXP3 and STAT3 are involved in the immune response. HNF1B, PDX1, PTF1A, RFX6, NEUROG3, GATA6, NEUROD1, GATA4, GLIS3, NKX2-2, and MNX1 are transcription factors that regulate genes in the nucleus. EIF2AK3 and IER3IP1 regulate protein trafficking in the endoplasmic reticulum. Mutations in the *INS* gene cause neonatal diabetes either by resulting in absence of insulin or by producing a defective insulin protein that accumulates in the endoplasmic reticulum and is not secreted in the blood stream. For genes encoding proteins acting within the β cell, the position of the gene name indicates the intracellular location of the protein. Substrates and transported molecules are indicated in blue. Biological processes are indicated in red.

Neonatal diabetes is a clinically and genetically heterogeneous disease. To date there are 23 different genetic causes of neonatal diabetes that identify different clinical subtypes of the disease (De Franco and colleagues, *submitted for publication* and[4]) (see **Fig. 1**, **Table 1**).

The most common causes of neonatal diabetes are mutations in the genes encoding the subunits of the voltage-dependent potassium channel *ABCC8* and *KCNJ11*.[8,9,27] Correct function of the potassium channel is necessary for secretion of insulin in response to glucose levels. Approximately 40% of patients with neonatal diabetes have a potassium channel gene mutation.[27,50] Patients with mutations in these two genes are sensitive to sulfonylurea treatment, and their glycemic control can be greatly improved switching from insulin to sulfonylurea therapy.[51,52] This

finding has led to international guidelines suggesting immediate referral for genetic testing after a clinical diagnosis of neonatal diabetes.[53] Mutations in *KCNJ11* and *ABCC8* can cause transient neonatal diabetes, permanent neonatal diabetes, or DEND (developmental delay, epilepsy, and neonatal diabetes) syndrome.[8,9,27,28,54]

Clinically neonatal diabetes can be divided into 3 broad categories:

- Transient neonatal diabetes (The diabetes remits and eventually relapses later in life.)
- Permanent neonatal diabetes (The diabetes does not remit.)
- Syndromic neonatal diabetes (Neonatal diabetes is one of the clinical features characterizing a syndrome.)

The most common causes of transient neonatal diabetes are methylation abnormalities resulting in overexpression of paternally expressed genes at the 6q24 locus[5–7] and mutations in *ABCC8* or *KCNJ11* (see **Table 1**).[28,54] Patients with a transient form of neonatal diabetes are diagnosed with hyperglycemia in the first 6 months of life; the diabetes then remits, and in most cases it relapses later in life.

Isolated insulin-requiring permanent neonatal diabetes is caused by mutations in the *INS* and *GCK* genes.[17,25,26] Mutations in 18 genes are known to cause syndromic neonatal diabetes (see **Table 1**), in which neonatal diabetes is just one of the features of the clinical spectrum that defines a particular condition. Because neonatal diabetes is diagnosed in the first 6 months of life, in most cases it is the presenting feature of the syndrome; additional clinical features will sequentially appear later in life. For this reason, a differential clinical diagnosis in the first 6 months of life is often difficult and can only be achieved months or even years after the first presentation with neonatal diabetes.

INTRODUCTION TO NEXT-GENERATION SEQUENCING

The term *next-generation sequencing* collectively refers to the high throughput DNA sequencing technologies that are able to sequence many DNA sequences in a single reaction (ie, in parallel). The advent of next-generation sequencing enables DNA sequencing at several orders of magnitude greater than was possible using the Sanger method developed in the 1970s. The Sanger methodology permits sequencing of a maximum of a few hundred nucleotides in a single reaction with each nucleotide being sequenced (or read by the instrument) just once. In contrast, next-generation sequencing allows entire exomes or genomes to be sequenced in a single test with each nucleotide being independently read multiple times (**Fig. 2**).

The introduction of next-generation sequencing technologies on the market in 2005[55] has resulted in the possibility to sequence entire exomes and genomes much more quickly and at a much lower cost. A widely quoted example is the first human genome sequence that took 13 years and cost nearly £2 billion, compared with the current cost of approaching £1000 for a genome sequence obtained in just 2 days. Next-generation sequencing technologies are now extensively used both for new disease gene discovery and for improving diagnostic genetic tests for known diseases.

Preparation of samples for next-generation sequencing usually includes fragmentation of DNA, ligation of adapters, and, in most cases, amplification via polymerase chain reaction. Several kits for library preparation are commercially available that allow for automation of the process and preparation of multiple samples in parallel.

The most widely used applications of next-generation sequencing are targeted analysis of a panel of genes and exome sequencing with genome sequencing

Table 1
Genetic causes of neonatal diabetes

Gene	Mode of Inheritance	Neonatal Diabetes Phenotype	Additional Features	Frequency in NDM Patients (De Franco et al, *Submitted*) (%)	References
6q24	—	Transient	Intrauterine growth retardation, macroglossia, umbilical hernia, neurologic features (rare)	11.1	Gardner et al,[5] Temple et al,[6] Temple & Shield[7]
ABCC8	Dominant/recessive	Transient, permanent	Developmental delay with/without epilepsy	14.7	Babenko et al,[8] Proks et al[9]
EIF2AK3	Recessive	Permanent	Skeletal dysplasia, liver dysfunction	7.5	Delepine et al,[10] Rubio-Cabezas et al[11]
FOXP3	X-linked	Permanent	Eczema, enteropathy, other autoimmune features	1.4	Chatila et al[12]
GATA4	Dominant	Transient, permanent	Exocrine insufficiency, congenital heart malformations	0.4	D'Amato et al,[13] Shaw-Smith et al[14]
GATA6	Dominant	Transient, permanent	Exocrine insufficiency, congenital heart malformation, neurologic defects, hypothyroidism, gut and hepatobiliary malformations	2.8	Lango Allen et al,[15] De Franco et al[16]
GCK	Recessive	Permanent	—	2.9	Njolstad et al,[17] Barbetti et al[18]
GLIS3	Recessive	Permanent	Hypothyroidism	0.9	Dimitri et al,[19] Senee et al[20]
HNF1B	Dominant	Transient	Exocrine insufficiency, renal cysts	0.2	Edghill et al,[21] Yorifuji et al[22]
IER3IP1	Recessive	Permanent	Microcephaly, epilepsy	0.1	Abdel-Salam et al,[23] Poulton et al[24]
INS	Dominant/recessive	Transient, permanent	—	10.8	Garin et al,[25] Stoy et al[26]

KCNJ11	Dominant	Transient, permanent	Developmental delay with/without epilepsy	23.5	Gloyn et al,[27] Gloyn et al[28]
MNX1	Recessive	Permanent	Sacral agenesis, neurologic defects	0.1	Flanagan et al[29]
NEUROD1	Recessive	Permanent	Cerebellar hypoplasia, sensorineural deafness, visual impairment	0.3	Rubio-Cabezas et al[30]
NEUROG3	Recessive	Permanent	Congenital malabsorptive diarrhea	0.2	Rubio-Cabezas et al[31]
NKX2-2	Recessive	Permanent	Corpus callosum agenesis	0.2	Flanagan et al[29]
PDX1	Recessive	Permanent	Exocrine insufficiency	0.6	Schwitzgebel et al,[32] Stoffers et al,[33] Thomas et al,[34] De Franco et al,[35] Nicolino et al[36]
PTF1A	Recessive	Permanent	Exocrine insufficiency, cerebellar agenesis (only for coding mutations)	2.2	Al-Shammari et al,[37] Sellick et al,[38] Tutak et al,[39] Weeden et al[40]
RFX6	Recessive	Permanent	Intestinal atresia and/or malrotation, gall bladder agenesis	0.1	Smith et al,[41] Spiegel et al[42]
SLC19A2	Recessive	Permanent	Thiamine-responsive megaloblastic anemia, sensorineural deafness	0.7	Bay et al,[43] Bergmann et al,[44] Mandel et al,[45] Shaw-Smith et al[46]
SLC2A2	Recessive	Transient	Hepatorenal glycogen accumulation, renal dysfunction, impaired utilization of glucose and galactose	0.6	Sansbury et al[47]
STAT3	Dominant	Permanent	Autoimmune enteropathy, thyroid dysfunction, pulmonary disease, juvenile-onset arthritis	0.4	Flanagan et al[4]
ZFP57	Recessive	Transient	Intrauterine growth retardation	1.2	Mackay et al,[48] Mackay & Temple[49]

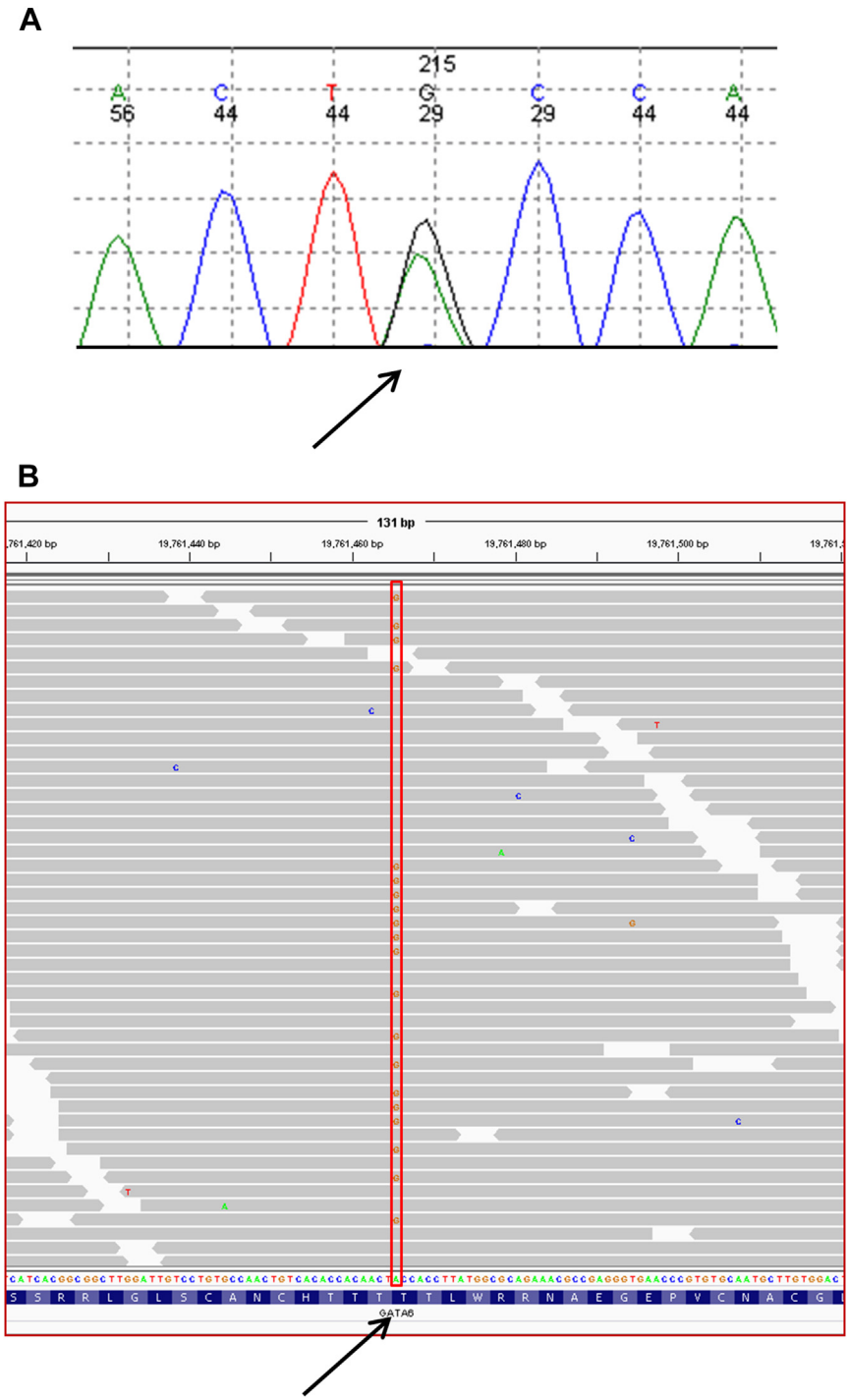

Fig. 2. With Sanger sequencing each nucleotide is sequenced (or read by the instrument) once. With next-generation sequencing entire exomes or genomes can be sequenced in a single test, and each position is covered by multiple reads. (*A*) Sanger sequencing trace of a mutation in *GATA6*. (*B*) The same mutation is detected by multiple reads with a next-generation sequencing assay. Black arrows indicate the mutation.

becoming more popular as prices decrease. Details on these technologies and their applications in neonatal diabetes are discussed in the following sections.

TARGETED NEXT-GENERATION SEQUENCING

Method

Next-generation sequencing technologies provide the potential for simultaneous analysis of all the genes known to cause a disease in a single assay at a similar cost to testing a few genes by Sanger sequencing. A widely used application is targeted next-generation sequencing of a given set of genes. Various methods have been successfully used to target specific genomic regions, currently the most commonly used is hybridization capture.[56–58] With this approach, several marked oligonucleotides (or baits) with sequences complementary to the targeted regions are used to capture the genes in the panel.

The error rate for next-generation sequencing is estimated at 1%[59]; therefore, multiple reads are required to obtain equivalent sensitivity to Sanger sequencing. The minimum depth of coverage needed (number of reads per base) will depend on the reason for testing (eg, clinical diagnostic vs prescreen before exome analysis). For clinical diagnostic testing the minimum read depth required is 30 reads, which enables reliable detection of heterozygous single nucleotide variants. Detection of small (<30 base pairs) insertions and deletions has proven more difficult, and optimization of the process is still ongoing. Various specific methods for detection of larger deletions and duplications have been developed (reviewed in[60,61]).

The use of targeted next-generation sequencing gene panels has become very common for genetic testing of genetically heterogeneous diseases, such as breast cancer,[62] polycystic kidney disease,[63] Bardet-Biedl/Alström syndrome,[58] and retinal disease.[57]

Targeted Next-Generation Sequencing in Neonatal Diabetes

To date mutations in 22 genes are known to cause neonatal diabetes (see **Fig. 1**, **Table 1**) (De Franco and colleagues, *submitted for publication*, and[4]). Three targeted next-generation sequencing assays have been developed for testing of monogenic forms of diabetes, including one developed by the authors.[64–66]

The targeted panel test developed in Exeter, United Kingdom uses a capture-in-solution system with baits for 48 genes known to cause monogenic forms of diabetes (eg, maturity-onset diabetes of youth, lipodystrophy), including the 22 known neonatal diabetes genes. This test can detect single nucleotide mutations and small deletions/insertions as well as large deletions and duplications.[65]

Because approximately 40% of patients diagnosed with neonatal diabetes have a mutation in a potassium channel gene (De Franco and colleagues, *submitted for publication*) and can be successfully treated with sulfonylurea tablets instead of insulin,[51,52] early comprehensive genetic testing in neonatal diabetes is of the utmost importance for patients' clinical management.

A recent study has evaluated the impact of early comprehensive genetic testing in a large cohort of patients with neonatal diabetes (De Franco and colleagues, *submitted for publication*). Comprehensive genetic analysis, including targeted next-generation sequencing and a methylation assay to detect 6q24 methylation abnormalities (**Fig. 3**), was performed on 1020 patients diagnosed with neonatal diabetes before 6 months of age and referred from 79 countries over 14 years.

De Franco and colleagues showed that a genetic diagnosis could be identified in more than 80% of patients with neonatal diabetes. As expected, the genetic causes of neonatal diabetes were very different in patients born to nonconsanguineous and

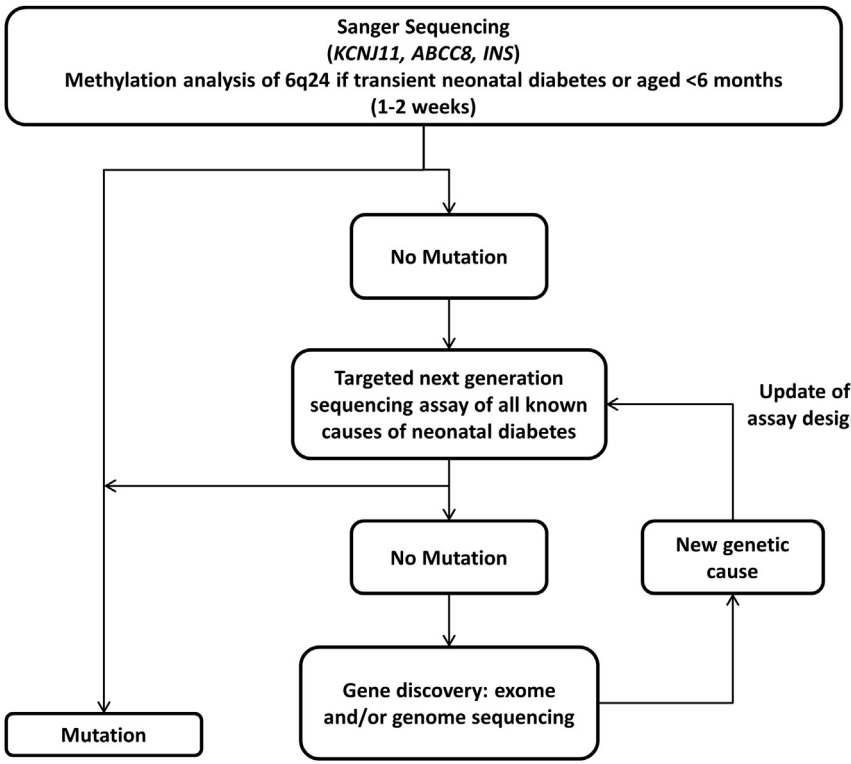

Fig. 3. Current genetic testing pipeline for neonatal diabetes referrals in the Exeter Genetics laboratory.

consanguineous parents. Mutations in *KCNJ11* and *ABCC8* were common in the non-consanguineous cohort, accounting for 46% of cases, but were present in only 12% of the patients in the consanguineous group.

The authors also reported that the median time from diagnosis of neonatal diabetes to referral for genetic testing has decreased from more than 5 years in 2000 to less than 3 months in 2012, indicating that now patients are more likely to be referred for genetic testing when neonatal diabetes is the only clinical feature present, before development of other clinical characteristics suggestive of their specific neonatal diabetes subtypes. This change means that comprehensive testing of all the known causes of neonatal diabetes results in these patients receiving a genetic diagnosis before development of the full clinical spectrum.

This has important implications for the patients with one of the neonatal diabetes subtypes for which an alternative treatment is available (potassium channel,[51,52] Thiamine-responsive megaloblastic anemia[67]) and for patients with syndromic forms of neonatal diabetes for whom clinicians will be aware of the likely development of specific additional features before these present. The authors conclude that the availability of a comprehensive genetic test has resulted in a paradigm shift in genetic testing for neonatal diabetes: although traditionally genetic testing was used only to confirm a diagnosis made by a set of clinical features, now genetic testing is a first-line investigation that makes the diagnosis and guides decisions for clinical care of patients.

EXOME SEQUENCING

Method

Exome sequencing allows simultaneous investigation of the less than 2% of the human genome (the exons) that encodes for proteins. About 80% of the disease-causing mutations are predicted to be located in a protein-coding part of the genome, thus making exome sequencing an attractive strategy to investigate the genetic basis of Mendelian diseases.

The exonic sequences are generally selected from genomic DNA by hybridization capture. Multiple capture systems are available, and the number of targeted exons is approximately 200,000, representing approximately 95% of known genes (**Table 2**).

Selection of the coding regions of the genome for sequencing is at the same time the main advantage and main disadvantage for exome sequencing when compared with whole-genome sequencing. In fact selection of the exonic sequences means that exome sequencing is currently cheaper and produces far less data to analyze than whole-genome sequencing. At the same time, the target selection process is subjected to different efficiency depending on the genomic region (for example, GC rich regions are captured far less efficiently than other parts of the genome); this results in uneven coverage (different number of reads) of the different targets, affecting the ability to detect variants in some parts of the genome.

Typically, between 20,000 and 50,000 variants are identified per exome sequenced.[68] Filtering and prioritizing strategies are needed to reduce this number to a small subset of variants that are most likely to be pathogenic. The filtering steps applied to exome sequencing data account for qualitative requirements, predicted effect of the variant on the protein, and whether the variant has been previously identified. Generally, these steps leave 150 to 500 nonsynonymous or splice-site variants to be considered as potentially pathogenic.[69] This number is generally too large to allow follow-up of all the variants, and additional prioritization strategies are needed. These strategies are generally based on the likely inheritance pattern of the disease (eg, looking for recessive mutations in linkage interval or de novo mutations in apparently sporadic disease).

Exome Sequencing in Neonatal Diabetes

More than 100 genes causing Mendelian disorders have been identified using exome sequencing,[70] including 2 novel causes of syndromic neonatal diabetes: *GATA6*[15] and *STAT3*.[4]

Identification of mutations in GATA6 as the most common cause of pancreatic agenesis

Neonatal diabetes caused by pancreatic agenesis is an extremely rare condition characterized by insulin-dependent diabetes and pancreatic exocrine insufficiency requiring enzyme supplementation therapy.[15] Before the identification of mutations in *GATA6*, only recessive mutations in 2 pancreatic developmental factors, *PDX1* and *PTF1A*, were known to cause pancreatic agenesis in humans. Mutations in

Table 2
Comparison between exome and genome sequencing

Technology	Amount of Data	Number of Variants	Cost ($)
Exome sequencing	50 Mb	~25,000	500
Genome sequencing	200 Gb	~3,000,000	5000

PDX1 had been described in 4 cases with isolated agenesis of the pancreas.[32–34] Mutations in *PTF1A* had been reported in 4 families in which affected individuals had both pancreatic and cerebellar agenesis.[37–39]

Lango Allen and colleagues[15] investigated a cohort of 27 patients with pancreatic agenesis and noticed that most patients with syndromic pancreatic agenesis were born to unaffected, unrelated parents. This finding suggested that the mutation causing the condition was most likely sporadic. To investigate this hypothesis, the investigators performed exome sequencing of 2 probands with pancreatic agenesis and a congenital heart malformation and their unaffected, unrelated parents with the objective of identifying de novo mutations (present in the patients but not inherited from either parent). After exclusion of common variants and variants that were unlikely to be pathogenic (eg, synonymous variants that do not lead to changes in the amino acid sequence), a single de novo variant was confirmed in each patient. Both variants, a missense mutation and a frameshift deletion, affected the coding region of the developmental transcription factor gene *GATA6*.

GATA6 is a transcription factor involved in early embryonic development of multiple organs, including the pancreas.[71] Traditionally, identification of novel causes of neonatal diabetes was guided by a candidate gene approach, based on the phenotype of mouse models. These studies were not suggestive of a role of Gata6 in pancreatic development in rodents,[71,72] and for this reason investigation of *GATA6* in patients with neonatal diabetes had not been considered before. In this case exome sequencing led to the identification of a novel disease gene and gave unexpected insights into human pancreatic development.

Lango Allen and colleagues[15] then sequenced *GATA6* in 25 additional patients with pancreatic agenesis and identified mutations in 13 additional cases. The investigators concluded that heterozygous mutations in *GATA6* are the most common cause of pancreatic agenesis.

A subsequent study[16] looked at the contribution of *GATA6* mutations in patients with neonatal diabetes but no reported exocrine insufficiency. The results showed that mutations in this gene cause a broad phenotypic spectrum of diabetes, from complete pancreatic agenesis to adult-onset diabetes without exocrine pancreatic insufficiency.

Identification of activating STAT3 mutations as a cause of early onset multiorgan autoimmune disease

In some cases diabetes diagnosed before 6 months can be caused by mutations in a single gene causing a polyendocrinopathy syndrome characterized by severe early autoimmunity leading to β-cell destruction. The most common of these conditions is IPEX syndrome (immune dysregulation, polyendocrinopathy, enteropathy, X-linked), which is caused by mutations in the *FOXP3* gene.[12] Identification of the genes causing these conditions is crucial to understand the mechanisms involved in the pathogenesis of more common autoimmune diseases.

In order to identify the gene causing early onset poly-autoimmunity, Flanagan and colleagues[4] performed exome sequencing of a proband/parents trio for a patient diagnosed with diabetes at 2 weeks and early onset additional autoimmune conditions (autoimmune hypothyroidism diagnosed at 3 years and celiac disease diagnosed at 17 months). A single de novo mutation in the transcription factor gene *STAT3* was identified. Sequencing of *STAT3* in 63 additional patients (24 with early onset autoimmune disease and 39 with neonatal diabetes) identified mutations in 4 further individuals.

Functional studies on the mutated STAT3 protein showed that the changes identified in patients with the early onset poly-autoimmunity phenotype were all activating

mutations, whereas mutations resulting in decreased activity of the STAT3 protein have been previously described as a cause of hyper immunoglobulin E syndrome.[73] The investigators proposed a mechanism in which *STAT3* activating mutations lead to early autoimmunity by impairing the development of regulatory T cells.[4]

In this study the use of exome sequencing lead to the identification of a novel cause of neonatal diabetes and gave important insights into the complex mechanisms leading to autoimmunity. This knowledge can be exploited to better understand the basis of more common autoimmune diseases, such as type 1 diabetes.

GENOME SEQUENCING

Method

Genome sequencing allows the analysis of approximately the entire genomic sequence (~98%[74]), without prior selection of specific regions. Each genome sequenced produces about 200 Gb of data with 3 to 4 million single nucleotide variants expected to be detected in each individual (see **Table 2**).

Genome sequencing is considered the most comprehensive test currently available[75] and presents some technical advantages compared with whole-exome sequencing: It is more sensitive and accurate for detecting structural variation (such as insertions, deletions, and translocations) because it does not rely on capture of a subset of regions and it allows a more even coverage throughout the genome. Another advantage of genome sequencing is the possibility to investigate nonexonic regulatory regions that are missed by exome sequencing. The main obstacles to the use of whole-genome sequencing for diagnostics have been the relatively high cost (which is rapidly falling) and the enormous amount of data produced, resulting in challenging data analysis (see **Table 2**). Most of the studies reporting the use of whole-genome sequencing so far have limited the initial variants analysis to the part of the genome encoding for proteins and have proceeded to the investigation of the noncoding variants just when a causing mutation could not be identified in the exome.

Genome Sequencing in Neonatal Diabetes

The use of genome sequencing in patients with isolated pancreatic agenesis recently led to the identification of mutations in a previously unrecognized regulatory element of the *PTF1A* gene.[40]

Biallelic mutations affecting the gene encoding for the transcription factor *PTF1A* are a known cause of pancreatic and cerebellar agenesis, with 4 families reported so far.[37–39] Weedon and colleagues[40] used whole-genome sequencing to identify the genetic cause of isolated pancreatic agenesis. The investigators studied 3 consanguineous pedigrees, which included multiple affected individuals, suggesting a recessive pattern of inheritance. Linkage analysis in the 3 families highlighted a single shared region on chromosome 10, including the *PTF1A* gene. No coding mutation segregating with the disease was identified.

Genome sequencing was subsequently performed in 2 probands; analysis was performed prioritizing coding variants for the initial analysis, but no likely cause was found. The investigators then concentrated on molecular changes affecting genomic regions regulating early pancreatic development. A single shared homozygous variant located in a highly conserved region approximately 25 kb from the *PTF1A* locus was identified. Sequencing analysis of the putative regulatory element in 19 additional patients with pancreatic agenesis identified a mutation in 8 of them.

Functional studies showed that the regulatory element is indeed a previously unrecognized *PTF1A* enhancer (a genomic element enhancing gene transcription), which is

selectively active during pancreatic development.[40] These results probably explain why patients with mutations in the *PTF1A* distal enhancer do not present the severe cerebellar phenotype associated with mutations affecting the *PTF1A* gene.[37–39]

In this study the application of genome sequencing has been crucial to identify the genetic cause of pancreatic agenesis in 10 families and uncover the role of a previously unsuspected regulatory element necessary for normal pancreatic development in humans.

CONCLUDING REMARKS

Next-generation sequencing applications are now widely used both for diagnostic genetic testing and for identification of novel causes of genetic conditions. Neonatal diabetes, being a genetically heterogeneous Mendelian disorder, has greatly benefitted from the application of next-generation sequencing technologies both in diagnostic and research settings.

Impact of Next-Generation Sequencing in Diagnosis of Neonatal Diabetes

There are 23 known causes of neonatal diabetes that identify different clinical subtypes of the disease (De Franco and colleagues, *submitted for publication*, and[4]), including isolated permanent neonatal diabetes, transient neonatal diabetes, and complex syndromes whereby neonatal diabetes is often the presenting feature (eg, Wolcott-Rallison syndrome). Traditional genetic testing for neonatal diabetes requires accurate clinical information regarding the patients' phenotype to allow selection of a small number of genes to test. Because patients are now referred soon after diagnosis with neonatal diabetes, this approach is limited by the clinical features present at the moment of genetic testing.

Targeted next-generation sequencing allows comprehensive analysis of all the genes known to cause neonatal diabetes in a single test. A genetic diagnosis can be identified in more than 80% of patients (De Franco and colleagues, *submitted for publication*), including approximately 40% of patients who have the genetic subtypes (mutations in *KCNJ11* and *ABCC8*) treatable with high-dose sulfonylurea instead of insulin.

For patients with syndromic forms of neonatal diabetes, hyperglycemia in the first 6 months of life is often the presenting feature of the disease. The other phenotypic features characterizing the disease can often present months or years after the initial diagnosis of neonatal diabetes. The most common of these conditions is Wolcott-Rallison syndrome, a recessive disease characterized by neonatal diabetes, skeletal dysplasia, and liver dysfunction.[13,14] For these patients an early genetic diagnosis predicts the future development of additional clinical features and raises awareness of the potential life-threatening complications (De Franco and colleagues, *submitted for publication*).

In the context of neonatal diabetes, next-generation sequencing allowed a shift in the paradigm of genetic testing: the genetic investigation is not merely confirmatory anymore, but it makes the diagnosis and guides clinical management of the patients.

Impact of Next-Generation Sequencing for Identification of Novel Causes of Neonatal Diabetes

Applications of next-generation sequencing technologies, exome sequencing in particular, have led to the identification of many novel causes of different genetic conditions.

Before the introduction of next-generation sequencing, identification of disease-causing genes was focused on analysis of candidate genes, selected by observations coming from animal model experiments, and linkage studies that are possible only when multiple affected individuals are available. Strategies based on exome sequencing are not biased by the prior knowledge on biological function suggested by experiments on animal models. For this reason discoveries of novel causes of disease by next-generation sequencing can sometimes highlight previously unrecognized roles for known genes. This circumstance is the case for the identification of mutations in *GATA6* as the most common cause of pancreatic agenesis in humans by exome sequencing.[15] Mutations in *GATA6* account for approximately 50% of pancreatic agenesis cases, but experiments looking at the development of the pancreas in mouse models had not suggested a role of this gene in pancreatic development.[71,72]

Identification of a genetic defect causing extreme phenotypes, such as early onset multiorgan autoimmunity, has important implications in the study of complex diseases, such as type 1 diabetes. The use of exome sequencing recently led to the identification of activating mutations in *STAT3* as a cause of early onset poly-autoimmunity, highlighting a fundamental role of this transcription factor in immune system regulation.[4]

Genome sequencing is considered the most comprehensive of the genetic tests currently available[75] as it allows investigation of complex genomic rearrangements, copy number variation, and variants in intronic and regulatory regions as well as in the coding parts of the genome. A combination of linkage studies and whole-genome sequencing in 2 consanguineous families with isolated pancreatic agenesis led to the identification of mutations in a previously unknown enhancer regulating expression of *PTF1A* during development of the pancreas.[40]

The introduction of next-generation sequencing has greatly expanded the potential of the strategies used to identify novel causes of genetic diseases. The research of novel causes of neonatal diabetes is a perfect example of how applications of next-generation sequencing technologies, coupled with appropriate analysis strategies, are powerful tools to successfully identify novel disease-causing genes.

ACKNOWLEDGMENTS

The authors thank all their colleagues in Exeter for their contributions in developing of the next-generation sequencing capacity for monogenic diabetes.

REFERENCES

1. Iafusco D, Massa O, Pasquino B, et al. Minimal incidence of neonatal/infancy onset diabetes in Italy is 1:90,000 live births. Acta Diabetol 2012;49(5):405–8.
2. Edghill EL, Dix RJ, Flanagan SE, et al. HLA genotyping supports a nonautoimmune etiology in patients diagnosed with diabetes under the age of 6 months. Diabetes 2006;55(6):1895–8.
3. Iafusco D, Stazi MA, Cotichini R, et al. Permanent diabetes mellitus in the first year of life. Diabetologia 2002;45(6):798–804.
4. Flanagan SE, Haapaniemi E, Russell MA, et al. Activating germline mutations in STAT3 cause early-onset multi-organ autoimmune disease. Nat Genet 2014;46(8):812–4.
5. Gardner RJ, Mackay DJ, Mungall AJ, et al. An imprinted locus associated with transient neonatal diabetes mellitus. Hum Mol Genet 2000;9(4):589–96.

6. Temple IK, Gardner RJ, Robinson DO, et al. Further evidence for an imprinted gene for neonatal diabetes localised to chromosome 6q22-q23. Hum Mol Genet 1996;5(8):1117–21.
7. Temple IK, Shield JP. Transient neonatal diabetes, a disorder of imprinting. J Med Genet 2002;39(12):872–5.
8. Babenko AP, Polak M, Cave H, et al. Activating mutations in the ABCC8 gene in neonatal diabetes mellitus. N Engl J Med 2006;355(5):456–66.
9. Proks P, Arnold AL, Bruining J, et al. A heterozygous activating mutation in the sulphonylurea receptor SUR1 (ABCC8) causes neonatal diabetes. Hum Mol Genet 2006;15(11):1793–800.
10. Delepine M, Nicolino M, Barrett T, et al. EIF2AK3, encoding translation initiation factor 2-alpha kinase 3, is mutated in patients with Wolcott-Rallison syndrome. Nat Genet 2000;25(4):406–9.
11. Rubio-Cabezas O, Patch AM, Minton JA, et al. Wolcott-Rallison syndrome is the most common genetic cause of permanent neonatal diabetes in consanguineous families. J Clin Endocrinol Metab 2009;94(11):4162–70.
12. Chatila TA, Blaeser F, Ho N, et al. JM2, encoding a fork head-related protein, is mutated in X-linked autoimmunity-allergic disregulation syndrome. J Clin Invest 2000;106(12):R75–81.
13. D'Amato E, Giacopelli F, Giannattasio A, et al. Genetic investigation in an Italian child with an unusual association of atrial septal defect, attributable to a new familial GATA4 gene mutation, and neonatal diabetes due to pancreatic agenesis. Diabet Med 2010;27(10):1195–200.
14. Shaw-Smith C, De Franco E, Allen HL, et al. GATA4 mutations are a cause of neonatal and childhood-onset diabetes. Diabetes 2014;63(8):2888–94.
15. Lango Allen H, Flanagan SE, Shaw-Smith C, et al. GATA6 haploinsufficiency causes pancreatic agenesis in humans. Nat Genet 2012;44(1):20–2.
16. De Franco E, Shaw-Smith C, Flanagan SE, et al. GATA6 mutations cause a broad phenotypic spectrum of diabetes from pancreatic agenesis to adult-onset diabetes without exocrine insufficiency. Diabetes 2013;62(3):993–7.
17. Njolstad PR, Sovik O, Cuesta-Munoz A, et al. Neonatal diabetes mellitus due to complete glucokinase deficiency. N Engl J Med 2001;344(21):1588–92.
18. Barbetti F, Cobo-Vuilleumier N, Dionisi-Vici C, et al. Opposite clinical phenotypes of glucokinase disease: description of a novel activating mutation and contiguous inactivating mutations in human glucokinase (GCK) gene. Mol Endocrinol 2009;23(12):1983–9.
19. Dimitri P, Warner JT, Minton JA, et al. Novel GLIS3 mutations demonstrate an extended multisystem phenotype. Eur J Endocrinol 2011;164(3):437–43.
20. Senee V, Chelala C, Duchatelet S, et al. Mutations in GLIS3 are responsible for a rare syndrome with neonatal diabetes mellitus and congenital hypothyroidism. Nat Genet 2006;38(6):682–7.
21. Edghill EL, Bingham C, Ellard S, et al. Mutations in hepatocyte nuclear factor-1beta and their related phenotypes. J Med Genet 2006;43(1):84–90.
22. Yorifuji T, Kurokawa K, Mamada M, et al. Neonatal diabetes mellitus and neonatal polycystic, dysplastic kidneys: phenotypically discordant recurrence of a mutation in the hepatocyte nuclear factor-1beta gene due to germline mosaicism. J Clin Endocrinol Metab 2004;89(6):2905–8.
23. Abdel-Salam GM, Schaffer AE, Zaki MS, et al. A homozygous IER3IP1 mutation causes microcephaly with simplified gyral pattern, epilepsy, and permanent neonatal diabetes syndrome (MEDS). Am J Med Genet A 2012;158A(11):2788–96.

24. Poulton CJ, Schot R, Kia SK, et al. Microcephaly with simplified gyration, epilepsy, and infantile diabetes linked to inappropriate apoptosis of neural progenitors. Am J Hum Genet 2011;89(2):265–76.
25. Garin I, Edghill EL, Akerman I, et al. Recessive mutations in the INS gene result in neonatal diabetes through reduced insulin biosynthesis. Proc Natl Acad Sci U S A 2010;107(7):3105–10.
26. Stoy J, Edghill EL, Flanagan SE, et al. Insulin gene mutations as a cause of permanent neonatal diabetes. Proc Natl Acad Sci U S A 2007;104(38):15040–4.
27. Gloyn AL, Pearson ER, Antcliff JF, et al. Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. N Engl J Med 2004;350(18):1838–49.
28. Gloyn AL, Reimann F, Girard C, et al. Relapsing diabetes can result from moderately activating mutations in KCNJ11. Hum Mol Genet 2005;14(7):925–34.
29. Flanagan SE, De Franco E, Lango Allen H, et al. Analysis of transcription factors key for mouse pancreatic development establishes NKX2-2 and MNX1 mutations as causes of neonatal diabetes in man. Cell Metab 2014;19(1):146–54.
30. Rubio-Cabezas O, Minton JA, Kantor I, et al. Homozygous mutations in NEUROD1 are responsible for a novel syndrome of permanent neonatal diabetes and neurological abnormalities. Diabetes 2010;59(9):2326–31.
31. Rubio-Cabezas O, Jensen JN, Hodgson MI, et al. Permanent neonatal diabetes and enteric anendocrinosis associated with biallelic mutations in NEUROG3. Diabetes 2011;60(4):1349–53.
32. Schwitzgebel VM, Mamin A, Brun T, et al. Agenesis of human pancreas due to decreased half-life of insulin promoter factor 1. J Clin Endocrinol Metab 2003;88(9):4398–406.
33. Stoffers DA, Zinkin NT, Stanojevic V, et al. Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. Nat Genet 1997;15(1):106–10.
34. Thomas IH, Saini NK, Adhikari A, et al. Neonatal diabetes mellitus with pancreatic agenesis in an infant with homozygous IPF-1 Pro63fsX60 mutation. Pediatr Diabetes 2009;10(7):492–6.
35. De Franco E, Shaw-Smith C, Flanagan SE, et al. Biallelic PDX1 (insulin promoter factor 1) mutations causing neonatal diabetes without exocrine pancreatic insufficiency. Diabet Med 2013;30(5):e197–200.
36. Nicolino M, Claiborn KC, Senee V, et al. A novel hypomorphic PDX1 mutation responsible for permanent neonatal diabetes with subclinical exocrine deficiency. Diabetes 2010;59(3):733–40.
37. Al-Shammari M, Al-Husain M, Al-Kharfy T, et al. A novel PTF1A mutation in a patient with severe pancreatic and cerebellar involvement. Clin Genet 2011;80(2):196–8.
38. Sellick GS, Barker KT, Stolte-Dijkstra I, et al. Mutations in PTF1A cause pancreatic and cerebellar agenesis. Nat Genet 2004;36(12):1301–5.
39. Tutak E, Satar M, Yapicioglu H, et al. A Turkish newborn infant with cerebellar agenesis/neonatal diabetes mellitus and PTF1A mutation. Genet Couns 2009;20(2):147–52.
40. Weedon MN, Cebola I, Patch AM, et al. Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis. Nat Genet 2014;46(1):61–4.
41. Smith SB, Qu HQ, Taleb N, et al. Rfx6 directs islet formation and insulin production in mice and humans. Nature 2010;463(7282):775–80.
42. Spiegel R, Dobbie A, Hartman C, et al. Clinical characterization of a newly described neonatal diabetes syndrome caused by RFX6 mutations. Am J Med Genet 2011;155A(11):2821–5.

43. Bay A, Keskin M, Hizli S, et al. Thiamine-responsive megaloblastic anemia syndrome. Int J Hematol 2010;92(3):524–6.
44. Bergmann AK, Sahai I, Falcone JF, et al. Thiamine-responsive megaloblastic anemia: identification of novel compound heterozygotes and mutation update. J Pediatr 2009;155(6):888–92.e1.
45. Mandel H, Berant M, Hazani A, et al. Thiamine-dependent beriberi in the "thiamine-responsive anemia syndrome". N Engl J Med 1984;311(13):836–8.
46. Shaw-Smith C, Flanagan SE, Patch AM, et al. Recessive SLC19A2 mutations are a cause of neonatal diabetes mellitus in thiamine-responsive megaloblastic anaemia. Pediatr Diabetes 2012;13(4):314–21.
47. Sansbury FH, Flanagan SE, Houghton JA, et al. SLC2A2 mutations can cause neonatal diabetes, suggesting GLUT2 may have a role in human insulin secretion. Diabetologia 2012;55(9):2381–5.
48. Mackay DJ, Callaway JL, Marks SM, et al. Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. Nat Genet 2008;40(8):949–51.
49. Mackay DJ, Temple IK. Transient neonatal diabetes mellitus type 1. Am J Med Genet C Semin Med Genet 2010;154C(3):335–42.
50. Ellard S, Flanagan SE, Girard CA, et al. Permanent neonatal diabetes caused by dominant, recessive, or compound heterozygous SUR1 mutations with opposite functional effects. Am J Hum Genet 2007;81(2):375–82.
51. Codner E, Flanagan S, Ellard S, et al. High-dose glibenclamide can replace insulin therapy despite transitory diarrhea in early-onset diabetes caused by a novel R201L Kir6.2 mutation. Diabetes Care 2005;28(3):758–9.
52. Pearson ER, Flechtner I, Njolstad PR, et al. Switching from insulin to oral sulfonylureas in patients with diabetes due to Kir6.2 mutations. N Engl J Med 2006;355(5):467–77.
53. Hattersley A, Bruining J, Shield J, et al. ISPAD clinical practice consensus guidelines 2006-2007. The diagnosis and management of monogenic diabetes in children. Pediatr Diabetes 2006;7(6):352–60.
54. Flanagan SE, Patch AM, Mackay DJ, et al. Mutations in ATP-sensitive K+ channel genes cause transient neonatal diabetes and permanent diabetes in childhood or adulthood. Diabetes 2007;56(7):1930–7.
55. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005;437(7057):376–80.
56. Johansson H, Isaksson M, Sorqvist EF, et al. Targeted resequencing of candidate genes using selector probes. Nucleic Acids Res 2011;39(2):e8.
57. O'Sullivan J, Mullaney BG, Bhaskar SS, et al. A paradigm shift in the delivery of services for diagnosis of inherited retinal disease. J Med Genet 2012;49(5):322–6.
58. Redin C, Le Gras S, Mhamdi O, et al. Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alstrom syndromes. J Med Genet 2012;49(8):502–12.
59. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol 2008;26(10):1135–45.
60. Duan J, Zhang JG, Deng HW, et al. Comparative studies of copy number variation detection methods for next-generation sequencing technologies. PloS One 2013;8(3):e59128.
61. Wang H, Nettleton D, Ying K. Copy number variation detection using next generation sequencing read counts. BMC Bioinformatics 2014;15:109.
62. Morgan JE, Carr IM, Sheridan E, et al. Genetic diagnosis of familial breast cancer using clonal sequencing. Hum Mutat 2010;31(4):484–91.

63. Rossetti S, Hopp K, Sikkink RA, et al. Identification of gene mutations in autosomal dominant polycystic kidney disease through targeted resequencing. J Am Soc Nephrol 2012;23(5):915–33.
64. Bonnefond A, Philippe J, Durand E, 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 2014;37(2):460–7.
65. Ellard S, Lango Allen H, De Franco E, et al. Improved genetic testing for monogenic diabetes using targeted next-generation sequencing. Diabetologia 2013; 56(9):1958–63.
66. Gao R, Liu Y, Gjesing AP, et al. Evaluation of a target region capture sequencing platform using monogenic diabetes as a study-model. BMC Genet 2014;15(1):13.
67. Neufeld EJ, Fleming JC, Tartaglini E, et al. Thiamine-responsive megaloblastic anemia syndrome: a disorder of high-affinity thiamine transport. Blood Cells Mol Dis 2001;27(1):135–8.
68. Gilissen C, Hoischen A, Brunner HG, et al. Disease gene identification strategies for exome sequencing. Eur J Hum Genet 2012;20(5):490–7.
69. Vissers LE, de Ligt J, Gilissen C, et al. A de novo paradigm for mental retardation. Nat Genet 2010;42(12):1109–12.
70. Rabbani B, Mahdieh N, Hosomichi K, et al. Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders. J Hum Genet 2012; 57(10):621–32.
71. Morrisey EE, Tang Z, Sigrist K, et al. GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. Genes Dev 1998; 12(22):3579–90.
72. Watt AJ, Zhao R, Li J, et al. Development of the mammalian liver and ventral pancreas is dependent on GATA4. BMC Dev Biol 2007;7:37.
73. Holland SM, DeLeo FR, Elloumi HZ, et al. STAT3 mutations in the hyper-IgE syndrome. N Engl J Med 2007;357(16):1608–19.
74. Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. Nature 2014;511(7509):344–7.
75. Lupski JR, Reid JG, Gonzaga-Jauregui C, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. N Engl J Med 2010;362(13): 1181–91.