Jose C. Florez *Editor*

The Genetics of Type 2 Diabetes and Related Traits

Biology, Physiology and Translation

Springer

The Genetics of Type 2 Diabetes and Related Traits

Jose C. Florez
Editor

The Genetics of Type 2 Diabetes and Related Traits

Biology, Physiology and Translation

Springer

Editor
Jose C. Florez
Diabetes Unit
Center for Human Genetic Research
Massachusetts General Hospital
Programs in Metabolism and Medical & Population Genetics
Broad Institute
Department of Medicine
Harvard Medical School
Boston
Massachusetts
USA

ISBN 978-3-319-01573-6 ISBN 978-3-319-01574-3 (eBook)
DOI 10.1007/978-3-319-01574-3

Library of Congress Control Number: 2015958349

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

When Victor Neel termed type 2 diabetes (T2D) "the geneticist's nightmare," (Neel 1962) it was already evident that T2D harbored a familial component; yet its polygenic genetic architecture, the inherent limitations in the techniques available at the time, and the strong (and growing!) influence of environmental determinants made genetic discovery an arduous task. In the 50 years since then, the public health impact of T2D has only skyrocketed, fueled by the changes in modern lifestyle increasingly adopted by developing societies and the expansion of caloric superabundance.

As a result, T2D and its complications represent one of the most serious challenges to public health in the twenty-first century. We live in the midst of a worldwide pandemic that threatens to undermine the significant gains we have made against cardiovascular disease over the last few decades. Despite its status as one of the oldest known endocrinopathies, the availability of a molecular therapy since 1920, and the existence of over a dozen drug classes approved for the management of the disease, we are largely unable to cure it and are losing the population battle in both the developed and developing worlds.

The population and healthcare costs are enormous and only expected to rise. Stoked by the snowballing obesity epidemic, diabetes affects over 29 million Americans, with more than 80 million at high risk. People with Asian, African, and American indigenous ancestry experience a higher risk, with the worldwide prevalence of diabetes expected to hit 500 million by 2030. In the USA alone, diabetes and its complications account for $245 billion annually. With its concomitant life-threatening complications of cardiovascular disease, renal failure, visual loss, and peripheral vascular disease, T2D can undermine the global economy, with a disproportionate burden on underprivileged groups and low- and middle-income countries.

At the crux of our lack of clinical effectiveness lies our incomplete understanding of its pathogenesis, evolution, and metabolic consequences. Definite causal triggers, the interplay between various risk factors, and specific mechanisms that underlie long-term complications remain largely unknown. Thus, T2D has become the quintessential complex disease, with substantial genetic and environmental

components, significant variation in prevalence across ethnic groups, pathophysiological heterogeneity, multiple treatment modalities, and poorly defined interactions with related metabolic phenotypes.

In the midst of many significant advances, key questions remain unanswered: Why do people develop T2D? Why do incidence rates differ across populations drawn from around the world? What causes beta cells to fail? How does obesity influence T2D risk? Can medications be designed that cure, rather than treat, diabetes? If more than one medication is needed, what is the best sequence or combination for specific subgroups?

These pressing questions also represent tantalizing opportunities. Our growing understanding of pathophysiology, the invention and deployment of global technologies that query specific areas of the biological space, and our improved ability to focus on the human as the targeted model system have placed the field of diabetes investigation on the verge of momentous discovery. However, among all valid experimental approaches employed in humans, only two can consistently bypass correlative associations and firmly establish causal inference. Beyond expensive and focused clinical trials, *the genetic method* is unique in ensuring that the exposure of interest (genotype) precedes phenotype, that it is conferred on individuals on a randomized basis at the time of conception, and that it is not in turn affected by the disease process or its treatment. Thus, it can serve as a powerful approach to dissect the nosology of T2D, illuminate its pathogenesis, and identify therapeutic targets through mechanistic insight.

In this volume, we have endeavored to take a contemporary snapshot of a rapidly moving field. In the time passed since Neel's initial cautionary statement, the scientific community has developed methods to measure global genomic variation with great precision, together with the statistical concepts and related analytical techniques that allow us to draw rigorous conclusions. Investigators have coalesced to advance knowledge in a collaborative fashion where needed, introducing appealing notions on the sociology of team science.

All of these ideas are illustrated in this book. The amount of novel information collected here, most of which was simply undreamed of just a decade ago, is staggering. At the same time, and reflecting the dizzying pace of discovery, nascent findings that have emerged most recently may not be fully captured in these chapters, making this reading all the more exciting. Finally, the organization of this work was intended to mirror the collaborative atmosphere that pervades our field, in that every chapter is authored by two or more investigators who hail from different research groups and yet complement each other in style, insight, and perspective.

The initial section, containing seven chapters, centers on fundamental genetic discovery. The initial overview provides a helpful historical viewpoint that will help the nongenetic reader take stock of the chronological evolution of the research enterprise in this area. From the proven effectiveness of genome-wide association studies (GWAS), subsequent contributions touch on the challenges that follow initial associations, the extension of this method to less accessible phenotypes,

and the arrival of next-generation sequencing as the harbinger of discovery focused on rare genetic variation.

The second section expands beyond genetics and illustrates how other data sources can inform initial genetic findings. Leveraging population diversity, correlating genetic associations with physiological measurements, learning from genetic variants that have strong phenotypic effects, and incorporating other key influences such as the modulation of gene expression, environmental factors, and our microbial commensals all help place genetic findings in focus and can lead to additional insight.

The third section makes a fundamental point: genetic association, however robust, is only the beginning of a laborious process. Convincing association signals notwithstanding, in most cases the specific DNA sequences that cause the molecular phenotype have not been identified. Indeed, the polymorphisms identified thus far merely signal genomic regions—at times hundreds of kilobases away from known genes—where an association has been found, but do not necessarily represent the causal variants: further fine-mapping and functional studies must be carried out before the true contribution of these loci to T2D can be accurately assessed. Thus, while we can rapidly and systematically uncover new associations, genetic studies do not circumvent the process of refining the associated loci to find the precise "causal" DNA sequences (causal in the sense of having a direct impact on RNA and/or protein quality or quantity that contributes to the diabetic phenotype). Indeed, variants may exert their molecular effects at remote sites even when they are relatively close to other uninvolved genes. Thus, loci identified by GWAS require in-depth sequencing and functional studies of the cellular and molecular effects of genes in that region. Six successful vignettes are described in this section, illustrating the progress we have made in just a few years.

The final section, comprising seven chapters, attempts to bring our current state of knowledge closer to the clinic, acknowledging both its potential and its limitations. It includes chapters on prediction, interaction of genetic variants with drugs or nutrients, and approaches to prevention or to the inference of causality for clinical relevant questions where randomized clinical trials have not produced conclusive answers or cannot be carried out. The epilogue, authored by a trio of long-standing collaborators who have set the pace for our field and whom many of us consider inspiring mentors, paints a realistic but hopeful vision of the future.

This book would not have been possible without the prescience of Andrea Pillmann at Springer in making the initial suggestion that we undertake this initiative, and without Jutta Lindenborn's patience in managing the editorial process. Over the years I have been fortunate to count on the professionalism and support of a superb publisher such as Springer in a variety of editorial projects, and this was no exception. I am most thankful to so many of my colleagues and friends who took time out of their busy professional lives to share their thoughts through eminently readable and informative chapters. Naturally not everyone who should or could have contributed was able to do so, but we have benefited from their wisdom as well, as this book largely reflects the collective body of knowledge garnered by the community over the past decade. Our remembrance goes to those luminaries

and pioneers who are no longer with us, such as Alan Permutt, Steve Elbein, and Linda Kao. And finally our mind rests in the smart, competent, and energized trainees we have the pleasure of working with, as they represent the bright future for our field: theirs will be the next edition in this fascinating journey of discovery, as we materialize our heartfelt commitment to ameliorate world suffering by improving human health.

Boston, MA Jose C. Florez

Reference

Neel JV (1962) Diabetes mellitus: a “thrifty” genotype rendered detrimental by “progress”? Am J Hum Genet 14:353–362

Contents

Part I
Genetic Discovery

Chapter 1
Historical Overview of Gene Discovery Methodologies in Type 2 Diabetes

Amélie Bonnefond, Alan R. Shuldiner, and Philippe Froguel

Abstract This initial chapter presents a historical snapshot of the various approaches utilized to discover genes implicated in the pathogenesis of type 2 diabetes.

1.1 Introduction

In 1976, James V. Neel described type 2 diabetes as "a geneticist's nightmare" (Neel 1976). Forty years later, this statement holds true despite a plethora of different approaches and methodologies that have been used to discover genetic etiologies of monogenic and polygenic forms of type 2 diabetes (T2D) (Fig. 1.1); still, almost 30 % of patients presenting with putative monogenic diabetes do not have mutations in known causative genes (Vaxillaire et al. 2012), and all T2D-associated genetic variants identified to date explain less than 15 % of T2D heritability (Morris et al. 2012).

A. Bonnefond
CNRS-UMR8199, Lille Pasteur Institute, Lille, France

Lille University, Lille, France

European Genomic Institute for Diabetes (EGID), Lille, France

A.R. Shuldiner
Program for Personalized and Genomic Medicine, and Division of Endocrinology, Diabetes and Nutrition, School of Medicine, University of Maryland, Baltimore, MD, USA

Regeneron Genetics Center, Regeneron Pharmaceuticals Inc., Tarrytown, NY, USA

P. Froguel (✉)
CNRS-UMR8199, Lille Pasteur Institute, Lille, France

Lille University, Lille, France

European Genomic Institute for Diabetes (EGID), Lille, France

Department of Genomics of Common Disease, School of Public Health, Imperial College London, Hammersmith Hospital, London, UK
e-mail: p.froguel@imperial.ac.uk

J.C. Florez (ed.), *The Genetics of Type 2 Diabetes and Related Traits*,
DOI 10.1007/978-3-319-01574-3_1

Historical Overview of Methodology in Type 2 Diabetes

Amélie Bonnefond[1-3], Alan Shuldiner[4-6], Philippe Froguel[1-3,7]

[1]CNRS-UMR8199, Lille Pasteur Institute, Lille, France;
[2]Lille University, Lille, France;
[3]European Genomic Institute for Diabetes(EGID), Lille, France;
[4]Program in Personalized and Genomic Medicine, and Division of Endocrinology, Diabetes and Nutrition, School of Medicine, University of Maryland, Baltimore, Maryland, USA;
[5]Geriatric Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, Maryland, USA;
[6]Present address: RegeneronGenetics Center, Regeneron Pharmaceuticals Inc., Tarrytown, New York, USA;
[7]Department of Genomics of Common Disease, Schoolof Public Health, Imperial College London, Hammersmith Hospital, London, United Kingdom.

Abbreviations:
GWAS, genome-wide association study
NGS, next-generation sequencing
T2D, type 2 diabetes
WES, whole-exomesequencing
WGS, whole-genome sequencing

Fig. 1.1 (continued)

1.2 Family-Based Linkage Analysis

In 1992, the first successful methodology to identify genes involved in diabetes was targeted or genome-wide linkage analysis (using microsatellites or other highly polymorphic genetic markers) in pedigrees (both consanguineous and non-consanguineous) with maturity-onset diabetes of the young (MODY) and in patients with neonatal diabetes mellitus. In combination with Sanger sequencing of candidate genes under the peak(s) of linkage, and more recently wider sequencing approaches, this strategy has enabled the identification of many genes involved in these monogenic forms of diabetes (Fig. 1.1): *GCK* (Froguel et al. 1992; Froguel et al. 1993), *HNF4A* (Bell et al. 1991; Yamagata et al. 1996a), *HNF1A* (Vaxillaire et al. 1995; Yamagata et al. 1996b), *CEL* (Raeder et al. 2006), *BLK* (Borowiec et al. 2009), *TRMT10A* (Igoillo-Esteve et al. 2013), *WFS1* (Inoue et al. 1998; Bonnycastle et al. 2013), *PCBD1* (Simaite et al. 2014), *INS* (Støy et al. 2007), *SLC19A2* (Labay et al. 1999), *EIF2AK3* (Delépine et al. 2000), *FOXP3* (Bennett et al. 2001), *PTF1A* (Sellick et al. 2004), *GLIS3* (Senée et al. 2006), and *IER3IP1* (Poulton et al. 2011). In polygenic forms of T2D, linkage analyses were far less successful, despite huge efforts of the research community. Only two T2D linkage signals were subsequently found to harbor variants in genes reproducibly associated with T2D (Fig. 1.1): *HNF4A* (Silander et al. 2004) and *TCF7L2* (Grant et al. 2006). However, for reasons that are unclear, it is probable that the T2D-associated common variants in these genes do not explain the original linkage signal.

In patients from consanguineous families, homozygosity mapping (namely, the identification of regions of the genome that are homozygous in affected individuals) through DNA arrays in combination with sequencing of candidate genes within

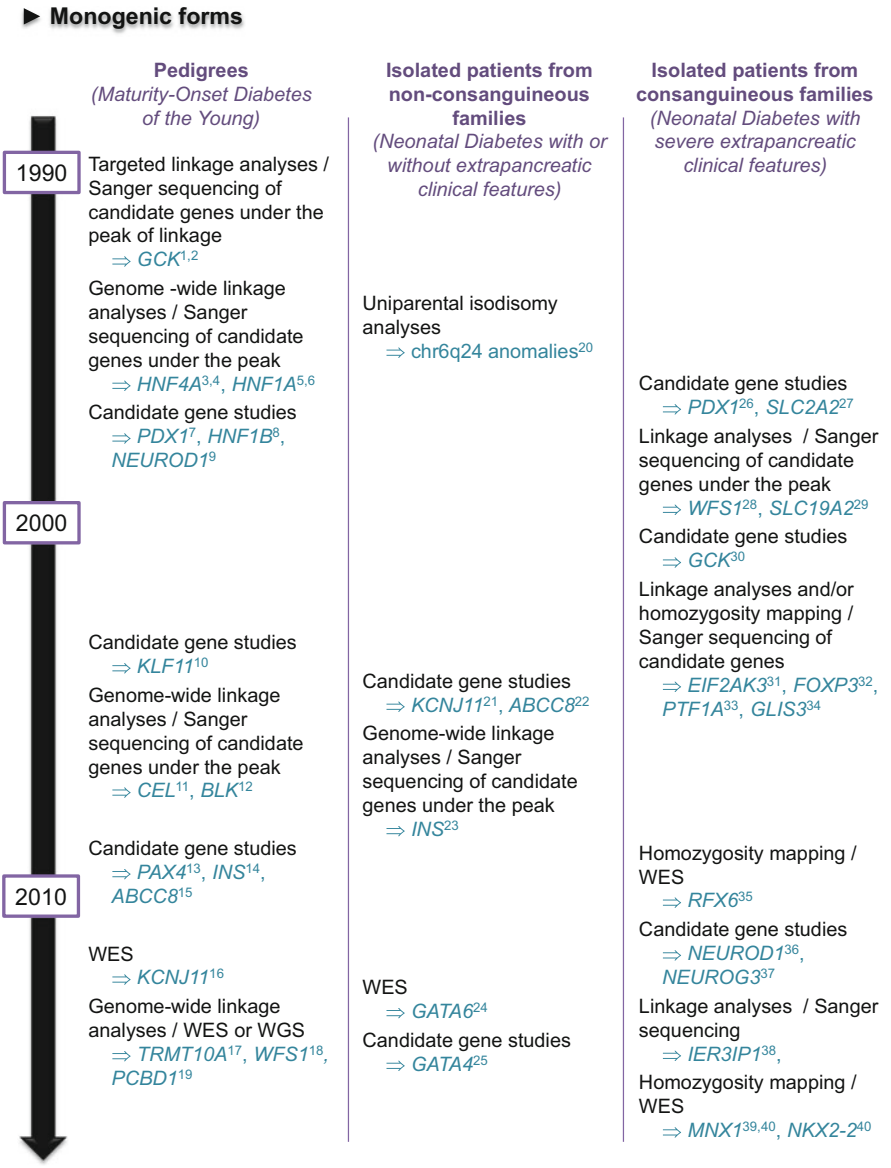

Fig. 1.1 (continued)

homozygous loci has enabled the identification of several genes involved in neonatal diabetes associated with severe extrapancreatic clinical features (Fig. 1.1): *RFX6* (Smith et al. 2010), *MNX1* (Bonnefond et al. 2013; Flanagan et al. 2014) and *NKX2-2* (Flanagan et al. 2014).

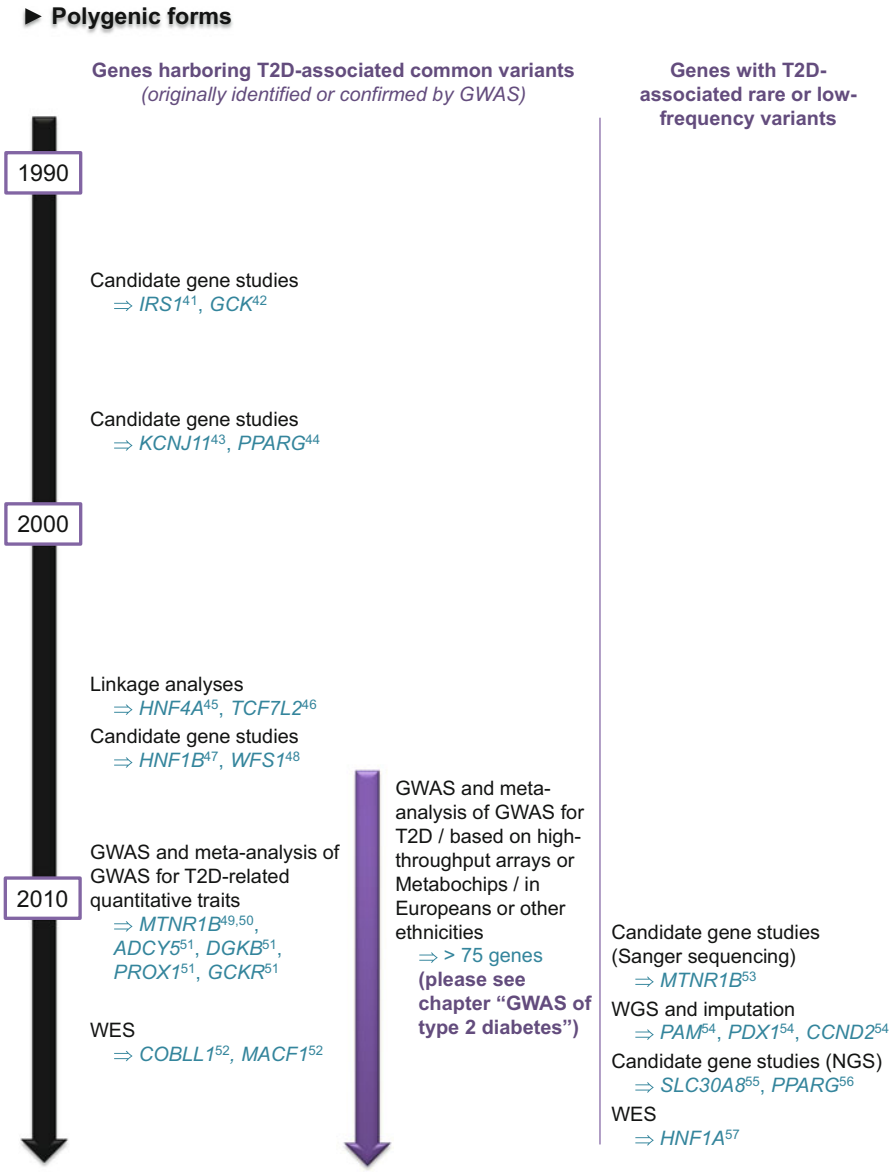

Fig. 1.1 (continued)

References

1. Froguel, P. *et al. Nature* **356,** 162–164 (1992).
2. Froguel, P. *et al.N. Engl. J. Med.* **328,** 697–702 (1993).
3. Bell, G. I. *et al. Proc. Natl. Acad. Sci. U. S. A.* **88,** 1484–1488 (1991).
4. Yamagata, K. *et al. Nature* **384,** 458–460 (1996).
5. Vaxillaire, M. *et al. Nat. Genet.* **9,** 418–423 (1995).
6. Yamagata, K. *et al. Nature* **384,** 455–458 (1996).
7. Stoffers, D. A., Ferrer, J., Clarke, W. L. & Habener, J. F. *Nat. Genet.* **17,** 138–139 (1997).
8. Horikawa, Y. *et al. Nat. Genet.* **17,** 384–385 (1997).
9. Malecki, M. T. *et al. Nat. Genet.* **23,** 323–328 (1999).
10. Neve, B. *et al. Proc. Natl. Acad. Sci. U. S. A.* **102,** 4807–4812 (2005).
11. Raeder, H. *et al. Nat. Genet.* **38,** 54–62 (2006).
12. Borowiec, M. *et al. Proc. Natl. Acad. Sci. U. S. A.* **106,** 14460–14465 (2009).
13. Plengvidhya, N. *et al. J. Clin. Endocrinol. Metab.* **92,** 2821–2826 (2007).
14. Meur, G. *et al. Diabetes* **59,** 653–661 (2010).
15. Bowman, P. *et al. Diabetologia* **55,** 123–127 (2012).
16. Bonnefond, A. *et al. PLoS ONE* **7,** e37423 (2012).
17. Igoillo-Esteve, M. *et al. PLoS Genet.* **9,** e1003888 (2013).
18. Bonnycastle, L. L. *et al. Diabetes* **62,** 3943–3950 (2013).
19. Simaite, D. *et al. Diabetes* **63,** 3557–3564 (2014).
20. Temple, I. K. *et al. Nat. Genet.* **9,** 110–112 (1995).
21. Gloyn, A. L. *et al. N. Engl. J. Med.* **350,** 1838–1849 (2004).
22. Babenko, A. P. *et al. N. Engl. J. Med.* **355,** 456–466 (2006).
23. Støy, J. *et al. Proc. Natl. Acad. Sci. U. S. A.* **104,** 15040–15044 (2007).
24. Lango Allen, H. *et al. Nat. Genet.* **44,** 20–22 (2012).
25. Shaw-Smith, C. *et al. Diabetes* **63,** 2888–2894 (2014).
26. Stoffers, D. A., Zinkin, N. T., Stanojevic, V., Clarke, W. L. & Habener, J. F. *Nat. Genet.* **15,** 106–110 (1997).
27. Santer, R. *et al. Nat. Genet.* **17,** 324–326 (1997).
28. Inoue, H. *et al. Nat. Genet.* **20,** 143–148 (1998).
29. Labay, V. *et al. Nat. Genet.* **22,** 300–304 (1999).
30. Njølstad, P. R. *et al. N. Engl. J. Med.* **344,** 1588–1592 (2001).
31. Delépine, M. *et al. Nat. Genet.* **25,** 406–409 (2000).
32. Bennett, C. L. *et al. Nat. Genet.* **27,** 20–21 (2001).
33. Sellick, G. S. *et al. Nat. Genet.* **36,** 1301–1305 (2004).
34. Senée, V. *et al. Nat. Genet.* **38,** 682–687 (2006).
35. Smith, S. B. *et al. Nature* **463,** 775–780 (2010).
36. Rubio-Cabezas, O. *et al. Diabetes* **59,** 2326–2331 (2010).
37. Rubio-Cabezas, O. *et al. Diabetes* **60,** 1349–1353 (2011).
38. Poulton, C. J. *et al. Am. J. Hum. Genet.* **89,** 265–276 (2011).
39. Bonnefond, A. *et al. Diabetes Metab.* **39,** 276–280 (2013).
40. Flanagan, S. E. *et al. Cell Metab.* **19,** 146–154 (2014).
41. Almind, K. *et al. Lancet* **342,** 828–832 (1993).
42. Stone, L. M., Kahn, S. E., Fujimoto, W. Y., Deeb, S. S. & Porte, D. *Diabetes* **45,** 422–428 (1996).
43. Hani, E. H. *et al. Diabetologia* **41,** 1511–1515 (1998).
44. Deeb, S. S. *et al. Nat. Genet.* **20,** 284–287 (1998).
45. Silander, K. *et al. Diabetes* **53,** 1141–1149 (2004).
46. Grant, S. F. A. *et al. Nat. Genet.* **38,** 320–323 (2006).
47. Winckler, W. *et al. Diabetes* **56,** 685–693 (2007).
48. Sandhu, M. S. *et al. Nat. Genet.***39,** 951–953 (2007).
49. Bouatia-Naji, N. *et al. Nat. Genet.* **41,** 89–94 (2009).
50. Prokopenko, I .*et al. Nat. Genet.* **41,** 77–81 (2009).
51. Dupuis, J. *et al. Nat. Genet.* **42,** 105–116 (2010).
52. Albrechtsen, A. *et al. Diabetologia* **56,** 298–310 (2013).
53. Bonnefond, A. *et al. Nat. Genet.* **44,** 297–301 (2012).
54. Steinthorsdottir, V. *et al. Nat. Genet.* **46,** 294–298 (2014).
55. Flannick, J. *et al. Nat. Genet.* **46,** 357–363 (2014).
56. Majithia, A. R. *et al. Proc. Natl. Acad. Sci. U. S. A.* (2014). doi:10.1073/pnas.1410428111.
57. SIGMA Type 2 Diabetes Consortium *et al. JAMA J. Am. Med. Assoc.* **311,** 2305–2314 (2014).

Fig. 1.1 Historical overview of methodology in type 2 diabetes (including monogenic and polygenic forms)

1.3 Candidate Gene Approaches

Candidate gene studies in which genes thought to be involved in glucose homeostasis are queried for sequence variation have been met with some success. Sequencing in families segregating atypical forms of early-onset nonimmune diabetes or syndromic conditions that include glucose intolerance has resulted in the identification of novel highly penetrant genes causing monogenic diabetes. By contrast, candidate genes for typical T2D involve querying variation in large numbers of unrelated T2D cases and nondiabetic controls to identify sequence variants enriched in the cases. These variants are typically common in the population and exert a small effect on T2D susceptibility. In a few examples, the same gene may harbor variants that have a large effect on function causing monogenic diabetes as well as more common variants that have a smaller effect on function and involved in polygenic T2D. Examples of successful identification of monogenic and polygenic diabetes genes using candidate gene approaches include:

MODY:

1. Key role in pancreatic beta cells: *PDX1* (Stoffers et al. 1997a), *NEUROD1* (Malecki et al. 1999), *KLF11* (Neve et al. 2005), and *PAX4* (Plengvidhya et al. 2007)
2. Genes belonging to a family that includes other genes previously shown to cause MODY, with a putative role in pancreatic beta cells: *HNF1B* (Horikawa et al. 1997)
3. Genes previously found to cause neonatal diabetes mellitus: *INS* (Meur et al. 2010) and *ABCC8* (Bowman et al. 2012)

Neonatal Diabetes Mellitus (Including Syndromic Forms):

1. Key role in pancreatic beta cells: *SLC2A2* (Santer et al. 1997), *KCNJ11* (Gloyn et al. 2004), *ABCC8* (Babenko et al. 2006), and *NEUROG3* (Rubio-Cabezas et al. 2011)
2. Genes previously found to cause MODY: *PDX1* (Stoffers et al. 1997b), *GCK* (Njølstad et al. 2001), and *NEUROD1* (Rubio-Cabezas et al. 2010)
3. Genes belonging to a family that includes other genes previously shown to cause MODY, with a putative role in pancreatic beta cells: *GATA4* (Shaw-Smith et al. 2014)

Polygenic Forms of T2D:

1. Key role in pancreatic beta cells or insulin sensitivity: *IRS1* (Almind et al. 1993), *KCNJ11* (Hani et al. 1998), *PPARG* (Deeb et al. 1998), and *WFS1* (Sandhu et al. 2007)
2. Genes previously found to cause a monogenic form of diabetes: *GCK* (Stone et al. 1996), and *HNF1B* (Winckler et al. 2007)

1.4 Genome-Wide Association Studies (GWAS)

Since 2007, GWAS and meta-analyses of GWAS in Europeans and other ethnicities have been very successful in identifying common single nucleotide polymorphisms (SNPs) significantly associated with T2D (please see Chap. 2). These GWAS have been based on high-throughput DNA microarrays assessing hundreds of thousands to few millions of SNPs across the genome in thousands of T2D cases and nondiabetic controls. GWAS of quantitative traits associated with T2D such as fasting plasma glucose, fasting serum insulin, and 2-h plasma glucose levels during an oral glucose tolerance test have also been quite successful in identifying new T2D-associated loci (Fig. 1.1). More recently, imputation of SNPs known from the 1000 Genomes project but not actually genotyped on the GWAS has boosted the number of analyzed SNPs (>6 million SNPs), enabling the identification of additional T2D-associated loci.

1.5 Next-Generation Sequencing (NGS) Approaches

Most recently, whole exome sequencing (WES) and whole genome sequencing (WGS) have become the most promising methodologies in finding new genes causing monogenic diabetes as well as novel loci associated with T2D risk. WES successfully identified two genes involved in monogenic diabetes (Fig. 1.1): *KCNJ11* in MODY (Bonnefond et al. 2012) (which was known to be mutated in patients with neonatal diabetes) and *GATA6* in neonatal diabetes (Lango Allen et al. 2012). Furthermore, WES performed in 2000 Europeans identified novel T2D-associated common variants in two genes: *COBLL1* and *MACF1* (Albrechtsen et al. 2013). Moreover, WES performed in 3756 Mexicans identified a low-frequency missense variant in *HNF1A* which strongly contributed to T2D risk (SIGMA Type 2 Diabetes Consortium et al. 2014). Finally, WGS of 2630 Icelanders and imputation (either direct imputation based on DNA array genotyping or in silico imputation based on genealogy information) into 11,000 cases and 267,000 controls of Icelandic origin identified T2D-associated low-frequency variants in three genes: *PAM*, *PDX1*, and *CCND2* (Steinthorsdottir et al. 2014).

1.6 Summary and Future Prospects

With increasing knowledge of genetic variation across the human genome, coupled with technological advances to query such variation, dramatic advances have been made in understanding the genetic basis of rare monogenic forms of diabetes as well as more common polygenic T2D. With these insights into the genetic architecture of diabetes, we now understand why family-based linkage analysis

approaches were successful in identifying genes causing monogenic forms of diabetes, while large case-control GWAS approaches were more successful in identifying genes and loci for T2D susceptibility. While small-effect variants identified by GWAS do not contribute substantially to individual risk and are not useful clinically to predict who will develop diabetes, identification of these genes has provided new insights into underlying disease mechanisms. NGS applied to both family-based and population-based approaches promises to unveil even greater granularity of the genetic architecture of diabetes, underlying biological mechanisms, and novel approaches for treatment and prevention.

References

Albrechtsen A, Grarup N, Li Y et al (2013) Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. Diabetologia 56:298–310. doi:10.1007/s00125-012-2756-1

Almind K, Bjørbaek C, Vestergaard H et al (1993) Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin-dependent diabetes mellitus. Lancet 342:828–832

Babenko AP, Polak M, Cavé H et al (2006) Activating mutations in the ABCC8 gene in neonatal diabetes mellitus. N Engl J Med 355:456–466. doi:10.1056/NEJMoa055068

Bell GI, Xiang KS, Newman MV et al (1991) 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 USA 88:1484–1488

Bennett CL, Christie J, Ramsdell F et al (2001) The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet 27: 20–21. doi:10.1038/83713

Bonnefond A, Philippe J, Durand E et al (2012) Whole-exome sequencing and high throughput genotyping identified KCNJ11 as the thirteenth MODY gene. PLoS ONE 7, e37423. doi:10.1371/journal.pone.0037423

Bonnefond A, Vaillant E, Philippe J et al (2013) Transcription factor gene MNX1 is a novel cause of permanent neonatal diabetes in a consanguineous family. Diabetes Metab 39:276–280. doi:10.1016/j.diabet.2013.02.007

Bonnycastle LL, Chines PS, Hara T et al (2013) Autosomal dominant diabetes arising from a Wolfram syndrome 1 mutation. Diabetes 62:3943–3950. doi:10.2337/db13-0571

Borowiec M, Liew CW, Thompson R et al (2009) Mutations at the BLK locus linked to maturity onset diabetes of the young and beta-cell dysfunction. Proc Natl Acad Sci USA 106:14460–14465. doi:10.1073/pnas.0906474106

Bowman P, Flanagan SE, Edghill EL et al (2012) Heterozygous ABCC8 mutations are a cause of MODY. Diabetologia 55:123–127. doi:10.1007/s00125-011-2319-x

Deeb SS, Fajas L, Nemoto M et al (1998) A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. Nat Genet 20:284–287. doi:10.1038/3099

Delépine M, Nicolino M, Barrett T et al (2000) EIF2AK3, encoding translation initiation factor 2-alpha kinase 3, is mutated in patients with Wolcott-Rallison syndrome. Nat Genet 25:406–409. doi:10.1038/78085

Flanagan SE, De Franco E, Lango Allen H et al (2014) Analysis of transcription factors key for mouse pancreatic development establishes NKX2-2 and MNX1 mutations as causes of neonatal diabetes in man. Cell Metab 19:146–154. doi:10.1016/j.cmet.2013.11.021

Froguel P, Vaxillaire M, Sun F et al (1992) Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. Nature 356:162–164. doi:10.1038/356162a0

Froguel P, Zouali H, Vionnet N et al (1993) Familial hyperglycemia due to mutations in glucokinase. Definition of a subtype of diabetes mellitus. N Engl J Med 328:697–702. doi:10.1056/NEJM199303113281005

Gloyn AL, Pearson ER, Antcliff JF et al (2004) Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes. N Engl J Med 350:1838–1849. doi:10.1056/NEJMoa032922

Grant SFA, Thorleifsson G, Reynisdottir I et al (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 38:320–323. doi:10.1038/ng1732

Hani EH, Boutin P, Durand E et al (1998) Missense mutations in the pancreatic islet beta cell inwardly rectifying K+ channel gene (KIR6.2/BIR): a meta-analysis suggests a role in the polygenic basis of Type II diabetes mellitus in Caucasians. Diabetologia 41:1511–1515. doi:10.1007/s001250051098

Horikawa Y, Iwasaki N, Hara M et al (1997) Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. Nat Genet 17:384–385. doi:10.1038/ng1297-384

Igoillo-Esteve M, Genin A, Lambert N et al (2013) tRNA methyltransferase homolog gene TRMT10A mutation in young onset diabetes and primary microcephaly in humans. PLoS Genet 9, e1003888. doi:10.1371/journal.pgen.1003888

Inoue H, Tanizawa Y, Wasson J et al (1998) A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome). Nat Genet 20:143–148. doi:10.1038/2441

Labay V, Raz T, Baron D et al (1999) Mutations in SLC19A2 cause thiamine-responsive megaloblastic anaemia associated with diabetes mellitus and deafness. Nat Genet 22:300–304. doi:10.1038/10372

Lango Allen H, Flanagan SE, Shaw-Smith C et al (2012) GATA6 haploinsufficiency causes pancreatic agenesis in humans. Nat Genet 44:20–22. doi:10.1038/ng.1035

Malecki MT, Jhala US, Antonellis A et al (1999) Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. Nat Genet 23:323–328. doi:10.1038/15500

Meur G, Simon A, Harun N et al (2010) Insulin gene mutations resulting in early-onset diabetes: marked differences in clinical presentation, metabolic status, and pathogenic effect through endoplasmic reticulum retention. Diabetes 59:653–661. doi:10.2337/db09-1091

Morris AP, Voight BF, Teslovich TM et al (2012) Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 44:981–990. doi:10.1038/ng.2383

Neel JV (1976) Diabetes mellitus — a geneticist's nightmare. In: Creutzfeldt W, Köbberling J, Neel JV (eds) The genetics of diabetes mellitus. Springer, Berlin, pp 1–11

Neve B, Fernandez-Zapico ME, Ashkenazi-Katalan V et al (2005) Role of transcription factor KLF11 and its diabetes-associated gene variants in pancreatic beta cell function. Proc Natl Acad Sci USA 102:4807–4812. doi:10.1073/pnas.0409177102

Njølstad PR, Søvik O, Cuesta-Muñoz A et al (2001) Neonatal diabetes mellitus due to complete glucokinase deficiency. N Engl J Med 344:1588–1592. doi:10.1056/NEJM200105243442104

Plengvidhya N, Kooptiwut S, Songtawee N et al (2007) PAX4 mutations in Thais with maturity onset diabetes of the young. J Clin Endocrinol Metab 92:2821–2826. doi:10.1210/jc.2006-1927

Poulton CJ, Schot R, Kia SK et al (2011) Microcephaly with simplified gyration, epilepsy, and infantile diabetes linked to inappropriate apoptosis of neural progenitors. Am J Hum Genet 89:265–276. doi:10.1016/j.ajhg.2011.07.006

Raeder H, Johansson S, Holm PI et al (2006) Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. Nat Genet 38:54–62. doi:10.1038/ng1708

Rubio-Cabezas O, Minton JAL, Kantor I et al (2010) Homozygous mutations in NEUROD1 are responsible for a novel syndrome of permanent neonatal diabetes and neurological abnormalities. Diabetes 59:2326–2331. doi:10.2337/db10-0011

Rubio-Cabezas O, Jensen JN, Hodgson MI et al (2011) Permanent neonatal diabetes and enteric anendocrinosis associated with biallelic mutations in NEUROG3. Diabetes 60: 1349–1353. doi:10.2337/db10-1008

Sandhu MS, Weedon MN, Fawcett KA et al (2007) Common variants in WFS1 confer risk of type 2 diabetes. Nat Genet 39:951–953. doi:10.1038/ng2067

Santer R, Schneppenheim R, Dombrowski A et al (1997) Mutations in GLUT2, the gene for the liver-type glucose transporter, in patients with Fanconi-Bickel syndrome. Nat Genet 17: 324–326. doi:10.1038/ng1197-324

Sellick GS, Barker KT, Stolte-Dijkstra I et al (2004) Mutations in PTF1A cause pancreatic and cerebellar agenesis. Nat Genet 36:1301–1305. doi:10.1038/ng1475

Senée V, Chelala C, Duchatelet S et al (2006) Mutations in GLIS3 are responsible for a rare syndrome with neonatal diabetes mellitus and congenital hypothyroidism. Nat Genet 38: 682–687. doi:10.1038/ng1802

Shaw-Smith C, De Franco E, Lango Allen H et al (2014) GATA4 mutations are a cause of neonatal and childhood-onset diabetes. Diabetes 63:2888–2894. doi:10.2337/db14-0061

SIGMA Type 2 Diabetes Consortium, Estrada K, Aukrust I et al (2014) Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. JAMA 311: 2305–2314. doi:10.1001/jama.2014.6511

Silander K, Mohlke KL, Scott LJ et al (2004) Genetic variation near the hepatocyte nuclear factor-4 alpha gene predicts susceptibility to type 2 diabetes. Diabetes 53:1141–1149

Simaite D, Kofent J, Gong M et al (2014) Recessive mutations in PCBD1 cause a new type of early-onset diabetes. Diabetes 63:3557–3564. doi:10.2337/db13-1784

Smith SB, Qu H-Q, Taleb N et al (2010) Rfx6 directs islet formation and insulin production in mice and humans. Nature 463:775–780. doi:10.1038/nature08748

Steinthorsdottir V, Thorleifsson G, Sulem P et al (2014) Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. Nat Genet 46:294–298. doi:10.1038/ng.2882

Stoffers DA, Ferrer J, Clarke WL, Habener JF (1997a) Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. Nat Genet 17:138–139. doi:10.1038/ng1097-138

Stoffers DA, Zinkin NT, Stanojevic V et al (1997b) Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. Nat Genet 15:106–110. doi:10.1038/ng0197-106

Stone LM, Kahn SE, Fujimoto WY et al (1996) A variation at position -30 of the beta-cell glucokinase gene promoter is associated with reduced beta-cell function in middle-aged Japanese-American men. Diabetes 45:422–428

Støy J, Edghill EL, Flanagan SE et al (2007) Insulin gene mutations as a cause of permanent neonatal diabetes. Proc Natl Acad Sci USA 104:15040–15044. doi:10.1073/pnas.0707291104

Vaxillaire M, Boccio V, Philippi A et al (1995) A gene for maturity onset diabetes of the young (MODY) maps to chromosome 12q. Nat Genet 9:418–423. doi:10.1038/ng0495-418

Vaxillaire M, Bonnefond A, Froguel P (2012) The lessons of early-onset monogenic diabetes for the understanding of diabetes pathogenesis. Best Pract Res Clin Endocrinol Metab 26: 171–187. doi:10.1016/j.beem.2011.12.001

Winckler W, Weedon MN, Graham RR et al (2007) Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. Diabetes 56:685–693. doi:10.2337/db06-0202

Yamagata K, Furuta H, Oda N et al (1996a) Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). Nature 384:458–460. doi:10.1038/384458a0

Yamagata K, Oda N, Kaisaki PJ et al (1996b) Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). Nature 384:455–458. doi:10.1038/384455a0

Chapter 2
Genome-Wide Association Studies of Type 2 Diabetes

Rob Sladek and Inga Prokopenko

Abstract Genome-wide association (GWA) studies represent the single most effective technique for identifying genetic risk loci causing complex diseases. Since the publication of the first GWA studies for type 2 diabetes (T2D) in 2007, nearly 90 statistically robust risk loci have been identified. The T2D risk loci identified by GWA studies contained several genes that are targets of current diabetic therapies; however, the majority of genes in these loci had not previously been implicated in the pathophysiology of T2D. Mechanistic insights about the physiological role of T2D loci in the disease predisposition have been gained from investigation of their contribution into glycemic trait variability in nondiabetic individuals. Current efforts to identify the causative genetic mutations in these loci and the molecular mechanisms through which they exert their effects have the potential to make far-reaching contributions to our understanding of molecular basis of T2D and the development of novel strategies for patient care.

2.1 Introduction

Type 2 diabetes (T2D) is a common, chronic disorder whose prevalence is increasing rapidly across the globe. Like other complex diseases, T2D represents a challenge for genetic studies aiming to uncover the underlying pathophysiological mechanisms. It is predicted that T2D will affect 592 million individuals by 2035 (Federation 2013) in developed and low- and middle-income countries. While the recent increase in T2D prevalence has been attributed to a sedentary "westernized"

R. Sladek
Department of Human Genetics, McGill University, Genome Quebec and McGill University Innovation Centre, 740, Doctor Penfield Avenue, Montreal, QC, Canada
e-mail: robert.sladek@mcgill.ca

I. Prokopenko (✉)
Department of Genomics of Common Disease, School of Public Health, Imperial College London, Hammersmith Hospital, Burlington Danes Building, Du Cane Road, London W12 0NN, UK
e-mail: i.prokopenko@imperial.ac.uk

J.C. Florez (ed.), *The Genetics of Type 2 Diabetes and Related Traits*,
DOI 10.1007/978-3-319-01574-3_2

lifestyle and changes in diet, a significant proportion of heritable factors also contribute to individual susceptibility (Hu 2011).

The strong family clustering and heritability of T2D and related glycemic traits have motivated a large number of studies to identify genetic factors that cause this disease (Permutt et al. 2005; Stumvoll et al. 2005); despite much effort, by late 2006 only three genetic loci had been reproducibly shown to increase T2D risk [reviewed in Majithia and Florez (2009), McCarthy (2008), Stolerman and Florez (2009)]. The earliest attempts to discover T2D-associated genes used either position- or function-based strategies. In a position-based search, genes are identified within families by studying the co-inheritance of the disease with a set of polymorphic markers whose genomic positions are known. Such "linkage studies" usually identify a genomic region (~10 Mbp) that confers genetic risk; the disease-causative mutations are identified by sequencing transcribed and functional elements of all genes in the target region. In function-based approaches, risk association is tested for common genetic variants in candidate genes involved in T2D pathophysiology. In these studies, variants identified in a small number of patients and control subjects are genotyped in larger case-control samples. Both these approaches are characterized by a number of limitations. Linkage analysis is underpowered to detect low penetrance variants, expected to contribute to T2D susceptibility, given its high population prevalence. Candidate gene studies usually were conducted in samples of insufficient size and their findings had low reproducibility as well as difficulty to select good biological candidates.

Positional strategies have identified putative T2D loci in several large chromosome regions (McCarthy 2003) and in a number of specific genes (Horikawa et al. 2000; Meyre et al. 2005; Silander et al. 2004; Hara et al. 2002); however, none of these associations have been convincingly replicated. The candidate-gene approach generated a large number of positive reports, two of which have been confirmed in independent studies (Table 2.2) (Gloyn and McCarthy 2001). The Pro12Ala variant in the peroxisome proliferator-activated receptor gamma (*PPARG*) gene (Deeb et al. 1998; Altshuler et al. 2000; Lohmueller et al. 2003) and the Glu23Lys variant in the potassium inwardly rectifying channel, subfamily J, member 11 (*KCNJ11*) gene were shown to contribute to T2D risk in multiple studies (Gloyn et al. 2003; Laukkanen et al. 2004). Each of these two common variants contributes only modestly (increasing T2D risk by 15–20 % for each susceptibility allele) to the risk of developing common form of diabetes, while rare variants in both these genes cause monogenic diseases such as familial partial lipodystrophy and neonatal diabetes. Interestingly, these variants occur within pharmacological targets for the thiazolidinedione (PPARG) and sulfonylurea compounds (KCNJ11) used to treat T2D.