

# Predicting Diabetes: Our Relentless Quest for Genomic Nuggets

**DIABETES EPIDEMIC**—The latest estimates from the Center for Disease Control and Prevention indicate that in 2010 approximately 26 million American adults had diabetes and 79 million had prediabetes (1). African Americans and other ethnic groups continue to suffer higher rates of diabetes than whites. Worldwide, diabetes affects 285 million adults (2). Type 2 diabetes accounts for ~95% of all cases. The exact reasons for the diabetes epidemic, and its predilection for certain ethnic groups, are unknown. However, interactions between genetic predisposition and environmental triggers (or accelerants) are generally presumed to underlie the etiology of diabetes (3–5) (Fig. 1). The best known environmental risk factors are dietary habits, physical inactivity, and obesity; interventions that ameliorate these risk factors prevent the development of type 2 diabetes (6,7).

By contrast, knowledge of the genetic basis of diabetes is incomplete, despite Herculean efforts (8–12). Genome-wide association studies have accelerated the discovery of single-nucleotide polymorphisms (SNPs) at numerous loci. Comparison of the frequencies of these SNPs in case-control studies has enabled the calculation of the odds of their association with specific disease phenotypes. To date, genome-wide studies have added more than 4,000 SNPs involving some 200 diseases, including >30 diabetes-related SNPs (diabetoSNPs). The analysis of diabetoSNPs has intrinsic appeal as a tool for diabetes prediction, and could also yield potential clues to ethnic disparities in the susceptibility to type 2 diabetes. Because the diabetoSNPs individually confer modest effects, investigators have adopted an approach based on cumulative genetic risk score (GRS) at several loci to improve sensitivity (13–16). Using available information on the relative odds of diabetes per risk allele (11,12), investigators can further calculate a weighted GRS.

**GENETIC PREDICTION**—In this issue of *Diabetes Care*, Cooke et al. (17), using such an approach, compared the cumulative GRS for 17 type 2 diabetes risk variants in a cross-sectional population

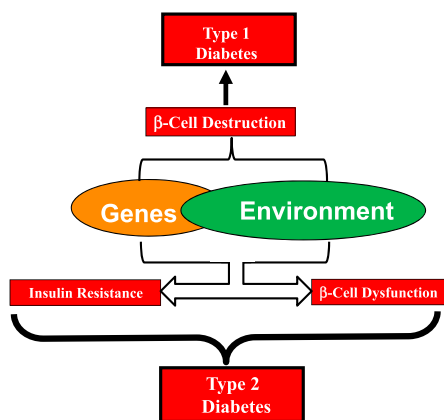
comprising 2,652 African American patients with type 2 diabetes and 1,393 nondiabetic control subjects. The authors found association between type 2 diabetes risk and cumulative GRS in the unweighted and weighted data set, and after adjusting for BMI. Notably, 5 of the 17 risk alleles had nominally significant association with type 2 diabetes, the strongest effect being observed for the rs7903146 SNP at the *TCF7L2* locus. After controlling for the latter, the GRS no longer predicted diabetes risk. Thus, the authors concluded that a GRS based on their panel of 17 European-derived risk variants did not predict type 2 diabetes status in African Americans, after excluding *TCF7L2* risk variant rs7903146.

The present report by Cooke et al. (17) is significant in that it replicates the known association of the rs7903146 SNP at *TCF7L2* with diabetes in African Americans. The bulk of genetic risk variants for type 2 diabetes have been derived from populations of European ancestry (10–14,16,18), with limited primary or replicative data for African populations (15,19). A major strength of the present report is the authentication of African ancestry of the study subjects using admixture analysis.

The promise of genetic risk scoring for diabetes can be evaluated in the framework of three perspectives. First is the potential for robust prediction of diabetes risk. Second is the prospect of designing targeted preventive and therapeutic interventions (personalized medicine). Thirdly, increased knowledge could provide genomic clues to ethnic disparities in diabetes. Regarding robustness of prediction, results from the Framingham Offspring Study showed that clinical risk assessment (using age, sex, family history, BMI, fasting glucose level, systolic blood pressure, high-density lipoprotein cholesterol level, and triglyceride level) performed as well as cumulative genotype score at 18 loci in predicting incident type 2 diabetes during 28 years of follow-up of initially normoglycemic subjects (14). Also, cumulative genotype score at 34 loci did not add significantly to clinical risk factors in predicting progression from impaired glucose tolerance to type

2 diabetes among the multiethnic cohort enrolled in the Diabetes Prevention Program (15). One current limitation is the incomplete framework from which GRS is constructed. For example, the 17 SNPs studied in the present report (17) represent just about half of the >30 diabetoSNPs identified to date. Even the latter do not represent all possible risk loci, and important information on structural variants that might increase diabetes risk is often lacking. Thus, current experience renders the promise of robust genetic prediction and personalized diabetes intervention a distant hope.

**ETHNIC DISPARITIES**—As noted by Cooke et al. (17), risk information on all 17 SNPs was obtained from European descendants. The replication in an African American population is informative; however, allelic variation is always a concern when applying the SNP panel to persons of African and other non-European ancestry (18,19). It is plausible that as more risk alleles from diverse populations are added to the panel, novel markers could emerge. For example, putative rare alleles with major effects that currently remain unrecognized could come to light. Until such new discoveries are made, what we know is that numerous previous reports together with the present report (17) convincingly identify the rs7903146 polymorphism in *TCF7L2* as a major genomic marker for type 2 diabetes in human beings. The individual association of other SNPs with diabetes is rather modest, barely grazing nominal statistical significance in many instances. Thus, chance associations become increasingly likely regarding marginal SNPs, particularly in studies of limited sample size. Other limitations include the near-universal lack of information on gene-gene interactions among the risk genotypes and limited data on gene-environment interactions (20,21) that could modify diabetes risk. Also lacking is information on correlative physiological measures, such as insulin sensitivity and  $\beta$ -cell function, which could provide a hint into underlying mechanisms of the diabetes risk conferred by these alleles.



**Figure 1**—Schematic of the pathogenesis of diabetes. Genetic and environmental factors, acting via complex immunological mechanisms, result in  $\beta$ -cell destruction that leads to type 1 diabetes. Gene-environment interactions also underlie susceptibility to type 2 diabetes, the pathophysiological hallmarks of which include insulin resistance and  $\beta$ -cell dysfunction.

**SURPRISING PATTERNS**—Despite these limitations, Cooke et al. (17) report nominally significant diabetes associations for five SNPs at *ADAMTS9*, *WFS1*, *CDKAL1*, *JAZF1*, and *TCF7L2* among their African American subjects. The *ADAMTS9* gene has been implicated in tumorigenesis; *WFS1* encodes wolframin, a transmembrane protein that is expressed in pancreas, brain, and insulinoma  $\beta$ -cell lines; *CDKAL1* encodes a methyltransferase of unknown function; and *JAZF1* encodes a zinc finger nuclear protein that functions as a transcriptional repressor. *TCF7L2*, the gene most strongly associated with type 2 diabetes, encodes a transcription factor in the *Wnt* signaling pathway that is involved in  $\beta$ -cell survival (22). Carriers of the *TCF7L2* risk allele have been reported to show decreased glucose-stimulated insulin secretion and defective insulin processing (20). Several other diabetoSNPs have functional implications that cluster around pancreatic growth, cell survival, insulin gene expression and protein processing (23). Remarkably, with a few exceptions, SNPs along the insulin signaling pathways have not featured prominently among the diabetes-associated risk alleles. Thus, genomic clues to the prevalent phenotype of increased insulin resistance, particularly among African Americans, have been largely elusive in genome-wide scans for diabetoSNPs.

Another surprising finding from studies in ethnically diverse populations (15,18–20) has been the lack of major

ethnic-specific risk alleles that would explain the disparities in the prevalence of type 2 diabetes. Data from the Diabetes Prevention Program (15) indicate that the allelic frequencies of the *TCF7L2* polymorphism are roughly similar in whites (TT 11%, Tc 45%, cc 44%) and blacks (TT 10%, Tc 43%, cc 47%). Such genomic concordance between the races underscores the importance of environmental factors (Fig. 1) in the etiology of ethnic disparities in type 2 diabetes. Thus, efforts to understand and address those factors (and unravel how they interact with genetic predisposition) constitute a dominant strategy for containing the diabetes epidemic while simultaneously assuaging ethnic disparities (24,25).

In conclusion, genome-wide studies have added valuable scientific data to our repertoire of diabetes knowledge. However, there have been few genomic nuggets that enable a more robust prediction of diabetes than is achieved by using common environmental risk factors and none that clarify the peculiar ethnic proclivities of type 2 diabetes. The latter realization ought to temper enthusiasm for the indiscriminate use of genetic testing for diabetes.

SAMUEL DAGOGO-JACK, MD, MBBS, MSC

From the Division of Endocrinology, Diabetes and Metabolism and Clinical Research Center, University of Tennessee Health Science Center, Memphis, Tennessee.

Corresponding author: Samuel Dagogo-Jack, sdj@uthsc.edu.

DOI: 10.2337/dc11-2106

© 2012 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

**Acknowledgments**—S.D.-J. is supported in part by grants from the National Institutes of Health (DK67269, DK62203, and DK48411).

No potential conflicts of interest relevant to this article were reported.

**References**

1. Centers for Disease Control and Prevention. Number of Americans with diabetes rises to nearly 26 million. More than a third of adults estimated to have prediabetes [press release], 2011. Available from <http://www.cdc.gov/media/releases/2011/p0126-diabetes.html>. Accessed 20 October 2011
2. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes

- for 2010 and 2030. *Diabetes Res Clin Pract* 2010;87:4–14
3. Dagogo-Jack S. Ethnic disparities in type 2 diabetes: pathophysiology and implications for prevention and management. *J Natl Med Assoc* 2003;95:774–789, 779–789
4. Egede LE, Dagogo-Jack S. Epidemiology of type 2 diabetes: focus on ethnic minorities. *Med Clin North Am* 2005;89:949–975, viii
5. Xiang AH, Li BH, Black MH, et al. Racial and ethnic disparities in diabetes risk after gestational diabetes mellitus. *Diabetologia* 2011;54:3016–3021
6. Knowler WC, Barrett-Connor E, Fowler SE, et al.; Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002; 346:393–403
7. Tuomilehto J, Lindström J, Eriksson JG, et al.; Finnish Diabetes Prevention Study Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* 2001;344:1343–1350
8. Tanizawa Y, Riggs AC, Dagogo-Jack S, et al. Isolation of the human LIM/homeodomain gene islet-1 and identification of a simple sequence repeat polymorphism [corrected]. *Diabetes* 1994;43:935–941
9. Ehm MG, Karnoub MC, Sakul H, et al.; American Diabetes Association GENNID Study Group. Genetics of NIDDM. Genomewide search for type 2 diabetes susceptibility genes in four American populations. *Am J Hum Genet* 2000;66:1871–1881
10. McCarthy MI, Zeggini E. Genome-wide association studies in type 2 diabetes. *Curr Diab Rep* 2009;9:164–171
11. Dupuis J, Langenberg C, Prokopenko I, et al.; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; Anders Hamsten on behalf of Procardis Consortium; MAGIC investigators. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010;42:105–116
12. Voight BF, Scott LJ, Steinthorsdottir V, et al.; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 2010;42:579–589
13. Weedon MN, McCarthy MI, Hitman G, et al. Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. *PLoS Med* 2006;3:e374
14. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 2008;359:2208–2219
15. Hivert MF, Jablonski KA, Perreault L, et al.; DIAGRAM Consortium; Diabetes Prevention Program Research Group. Updated ge-

- netic score based on 34 confirmed type 2 diabetes Loci is associated with diabetes incidence and regression to normoglycemia in the diabetes prevention program. *Diabetes* 2011;60:1340–1348
16. van Hoek M, Dehghan A, Witteman JC, et al. Predicting type 2 diabetes based on polymorphisms from genome-wide association studies: a population-based study. *Diabetes* 2008;57:3122–3128
  17. Cooke JN, Ng MCY, Palmer ND, et al. Genetic risk assessment of type 2 diabetes-associated polymorphisms in African Americans. *Diabetes Care* 2012;35:287–292
  18. Lewis JP, Palmer ND, Hicks PJ, et al. Association analysis in african americans of European-derived type 2 diabetes single nucleotide polymorphisms from whole-genome association studies. *Diabetes* 2008;57:2220–2225
  19. Palmer ND, Hester JM, An SS, et al. Re-sequencing and analysis of variation in the TCF7L2 gene in African Americans suggests that SNP rs7903146 is the causal diabetes susceptibility variant. *Diabetes* 2011;60:662–668
  20. McCaffery JM, Jablonski KA, Franks PW, et al.; Diabetes Prevention Program Research Group. TCF7L2 polymorphism, weight loss and proinsulin:insulin ratio in the diabetes prevention program. *PLoS ONE* 2011;6:e21518 [Epub Jul 26, 2011]
  21. Fisher E, Boeing H, Fritsche A, Doering F, Joost HG, Schulze MB. Whole-grain consumption and transcription factor-7-like 2 (TCF7L2) rs7903146: gene-diet interaction in modulating type 2 diabetes risk. *Br J Nutr* 2009;101:478–481
  22. Shu L, Sauter NS, Schulthess FT, Matveyenko AV, Oberholzer J, Maedler K. Transcription factor 7-like 2 regulates beta-cell survival and function in human pancreatic islets. *Diabetes* 2008;57:645–653
  23. Kirchhoff K, Machicao F, Haupt A, et al. Polymorphisms in the TCF7L2, CDKAL1 and SLC30A8 genes are associated with impaired proinsulin conversion. *Diabetologia* 2008;51:597–601
  24. Florez JC, Jablonski KA, Bayley N, et al.; Diabetes Prevention Program Research Group. TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. *N Engl J Med* 2006;355:241–250
  25. Dagogo-Jack S, Edeoga C, Nyenwe E, Chapp-Jumbo E, Wan J. Pathobiology of Prediabetes in a Biracial Cohort (POP-ABC): design and methods. *Ethn Dis* 2011;21:33–39