

Timothy M. Frayling and Andrew T. Hattersley



# Physiology Helps GWAS Take a Step Closer to Mechanism



*Diabetes* 2014;63:1836–1837 | DOI: 10.2337/db14-0130

Genome-wide association studies (GWAS) have been extremely successful at identifying replicable associations between common genetic variants and type 2 diabetes risk. The latest studies, including 35,000 European (1), 7,000 East Asian (2), 5,500 South Asian (3), and most recently 3,800 Latin American (4) and 6,000 Japanese (5) type 2 diabetes cases, bring the total number of associated variants to more than 70. There is strong evidence that many of the associated genetic variants lie in or close to genes important in type 2 diabetes etiology (e.g., the regions of the genome identified by GWAS are enriched for monogenic diabetes genes, such as *HNF1A*, *HNF1B*, and *PPARG*, and small noncoding regions of the genome [enhancers] critical for islet-specific gene expression [6]). Nevertheless, the field has not moved from genetic associations to improved understanding of biology as quickly or as often as hoped.

In this issue, Dimas et al. (7) present data that move the field a step closer to mechanisms. They tested the hypothesis that a systematic analysis of insulin secretion and insulin resistance measures in nondiabetic individuals would improve the understanding of the intermediate mechanisms by which genetic variants predispose to type 2 diabetes. They performed the most extensive analysis yet to group variants into categories based on their likely intermediate mechanism. The authors combined data from thousands of individuals with fasting-, oral-, and intravenous-based measures of insulin secretion and resistance. This approach has been used before for a smaller number of loci and individuals (8), but here the authors added a statistical clustering approach to provide the most robust categorization of the type 2 diabetes variants. This clustering analysis successfully binned 16 type 2 diabetes risk variants into four broad groups. Four variants fitted a clear insulin resistance pattern, two reduced insulin secretion with fasting hyperglycemia, nine reduced insulin secretion with normal fasting glycemia, and one altered insulin processing. A further 20 variants did not fit a clear physiological category, probably because

of their relatively weaker effects on type 2 diabetes risk. These findings move the field forward by providing functional biologists with more information about where to start their experiments.

It could be argued that there are few surprises among the findings. For example, the variants categorized as insulin resistance, those near *PPARG*, *IRS1*, *GCKR*, and *KLF11*, either lie near genes with clear roles in insulin resistance (*PPARG*, *IRS1*) or are associated with insulin resistance-related measures (*GCKR* [9] and *KLF11* [10]). Likewise, studies have established that diabetes risk alleles in or near *TCF7L2* and *SLC30A8* are associated with reduced insulin secretory capacity in response to a glucose challenge (8). Nevertheless, the data provide the strongest evidence yet that those genetic variants in or near (but not necessarily functioning through) the *PROX1*, *TMEM*, *CDKAL1*, *CDKN2A/B*, *THADA*, *HHEX/IDE*, and *ADCY5* genes operate primarily through an insulin secretory defect. Dimas et al. also highlight the intriguing pattern of associations observed with the variant near *ARAP1*. Previously noted in a genome-wide study of proinsulin levels (11), Dimas et al. categorize this variant as “insulin processing.” Most type 2 diabetes risk alleles are associated with raised proinsulin levels, and this is in keeping with the epidemiologic associations (12). In contrast, the type 2 diabetes risk allele in the *ARAP1* locus is associated with reduced proinsulin levels relative to insulin levels in the fasting state. The underlying explanation of this paradoxical finding is still not known.

The study reemphasizes some old questions and raises some new questions. Notably, why is it so difficult to assign an intermediate physiological mechanism to alleles clearly associated with type 2 diabetes? Despite using up to 58,000 individuals with fasting measures, 11,000 with oral glucose tolerance tests (OGTTs), and 4,600 with intravenous-based measures including 2,600 with euglycemic-hyperinsulinemic clamps, Dimas et al. were only able to group 16 of the 37 strongest type 2 diabetes risk variants into recognizable categories. Intuitively, a genetic

risk variant should be associated with the trait that leads to type 2 diabetes more strongly than diabetes, but this rarely seems to be the case. Sample size is unlikely to be to blame—most of the diabetes risk alleles studied were discovered using less than 12,000 cases. One of the most likely explanations is measurement error. Insulin and, to a lesser extent, glucose vary within individuals much more than type 2 diabetes status. Imprecision in measuring intermediate physiology will reduce statistical power to detect effects (13), and it is noticeable that 12 of the 16 classifiable variants are among the 22 with the strongest odds ratios for type 2 diabetes. It may be that studies of 12,000 individuals at the ~5% extreme end of the population with the poorest  $\beta$ -cell function and highest insulin resistance (those with diabetes) may be much more powerful than studies of 12,000 individuals from the remaining 95% of the distribution.

The explanation of the unclassified associations may not be just lack of statistical power. It is also an intriguing possibility that some variants may not play a role in altering physiology in normal individuals (at least as assessed by the submaximal tests performed in these studies), but may be important in influencing initial  $\beta$ -cell mass or the rate of deterioration of  $\beta$ -cell function once diabetes develops. These possible alternatives may explain why some of the variants lying close to the known monogenic  $\beta$ -cell genes, including *HNF1A*, *HNF1B*, *WFS1*, and *KCNJ11*, were not grouped into clear physiological categories.

Also worth highlighting are the common genetic variants noticeable by their absence. Variants in or near the *G6PC2* and *MADD* (11) genes are among those with the strongest effects on fasting glycemia and proinsulin levels, respectively, and yet are not associated (even nominally) with type 2 diabetes. Hence, these variants were not included in Dimas et al., but further study of these variants and genes could improve knowledge of  $\beta$ -cell function. Since the GWAS finding, the *G6PC2* gene (also called islet-specific glucose-6-phosphatase-related protein [*IGRP*]) has been the subject of renewed interest (14).

These large studies of subtle physiological effects also allow examination of the rather crude tools available to examine intermediate traits in large numbers of individuals. It is clear that in normoglycemic individuals the vast majority of variation in derived homeostasis model assessment of  $\beta$ -cell function is explained by fasting glucose, and this model adds little. Interestingly, there is not a clear increase in precision when intravenous glucose tolerance tests are used rather than OGTT-derived indices. It would be interesting to examine if the more sophisticated modeling of OGTT data using deconvolution of C-peptide could give new insights (15).

Perhaps the most important general message emerging from Dimas et al. (7) is that the type 2 diabetes GWAS field needs scientists from other areas—those with expertise in physiology, cell biology, and functional biology—to carefully inspect the 70 loci, the associated phenotypes, and the nearby genes. By way of example, follow-up of the

variants and genes in the loci labeled as *PROX1*, *TMEM*, *CDKAL1*, *CDKN2A/B*, *THADA*, *HHEX/IDE*, and *ADCY5* should target insulin secretion mechanisms.

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

## References

- Morris AP, Voight BF, Teslovich TM, et al.; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of Anthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;44:981–990
- Cho YS, Chen CH, Hu C, et al.; DIAGRAM Consortium; MuTHER Consortium. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in East Asians. *Nat Genet* 2012;44:67–72
- Kooner JS, Saleheen D, Sim X, et al.; DIAGRAM Consortium; MuTHER Consortium. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat Genet* 2011;43:984–989
- Williams AL, Jacobs SB, Moreno-Macias H et al.; SIGMA Type 2 Diabetes Consortium. Sequence variants in *SLC16A11* are a common risk factor for type 2 diabetes in Mexico. *Nature* 2014;506:97–101
- Hara K, Fujita H, Johnson TA, et al.; DIAGRAM Consortium. Genome-wide association study identifies three novel loci for type 2 diabetes. *Hum Mol Genet* 2014;23:239–246
- Pasquali L, Gaulton KJ, Rodríguez-Seguí SA, et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat Genet* 2014;46:136–143
- Dimas AS, Lagou V, Barker A, et al.; MAGIC Investigators. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* 2014;63:2158–2171
- Ingelsson E, Langenberg C, Hivert MF, et al.; MAGIC Investigators. Detailed physiologic characterization reveals diverse mechanisms for novel genetic loci regulating glucose and insulin metabolism in humans. *Diabetes* 2010;59:1266–1275
- Dupuis J, Langenberg C, Prokopenko I, et al.; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; 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
- Small KS, Hedman AK, Grundberg E, et al.; GIANT Consortium; MAGIC Investigators; DIAGRAM Consortium; MuTHER Consortium. Identification of an imprinted master trans regulator at the *KLF14* locus related to multiple metabolic phenotypes. *Nat Genet* 2011;43:561–564
- Strawbridge RJ, Dupuis J, Prokopenko I, et al.; DIAGRAM Consortium; GIANT Consortium; MuTHER Consortium; CARDIOGRAM Consortium; C4D Consortium. Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. *Diabetes* 2011;60:2624–2634
- Wareham NJ, Byrne CD, Williams R, Day NE, Hales CN. Fasting proinsulin concentrations predict the development of type 2 diabetes. *Diabetes Care* 1999;22:262–270
- Utzscheider KM, Prigeon RL, Tong J, et al. Within-subject variability of measures of beta cell function derived from a 2 h OGTT: implications for research studies. *Diabetologia* 2007;50:2516–2525
- O'Brien RM. Moving on from GWAS: functional studies on the *G6PC2* gene implicated in the regulation of fasting blood glucose. *Curr Diab Rep* 2013;13:768–777
- Mari A, Schmitz O, Gastaldelli A, Oestergaard T, Nyholm B, Ferrannini E. Meal and oral glucose tests for assessment of beta-cell function: modeling analysis in normal subjects. *Am J Physiol Endocrinol Metab* 2002;283:E1159–E1166