A Genome-Wide Association Study of Treated A1C
A Genetic Needle in an Environmental Haystack?

Jose C. Florez

Glycated hemoglobin (A1C), expressed as a percentage of total hemoglobin, results from the nonenzymatic glycation of hemoglobin molecules. Because this irreversible process is directly proportional to intracellular glucose concentrations (which are generally correlated with extracellular glucose levels), A1C can be used as a marker of average ambient glycemia over the mean life span of erythrocytes. Absent alterations in hemoglobin turnover or erythrocyte physiology, A1C adequately captures glucose homeostasis in the organism. Therefore, it has been widely adopted as an indicator of glycemic control in the treatment of diabetic patients (1), and it has recently been proposed as a diagnostic test for diabetes (2).

Variation in A1C is subject to environmental and genetic determinants. Its heritability (the proportion of the variance explained by the familial contribution to the trait, integrating both genetic and early shared environmental components) nears 40% in the general population (3) and 60% in twin studies (4). The distribution of A1C displays a long tail to the right, populated by levels of people whose hyperglycemia signals diabetes. Pharmacological treatment can significantly reduce A1C, providing the rationale for near normalization of A1C as a treatment goal (1). This is particularly true for insulin as a therapeutic modality: barring hypoglycemia and other practical limitations, insulin can be gradually raised until the target A1C has been reached. Thus, presumably higher insulin doses should be able to overcome most endogenous (i.e., genetic) obstacles to lowering A1C. In the setting of treated A1C, therefore, the environmental contribution to the trait acquires a much larger role than that in the native state, and one might suspect that it could overwhelm any genetic component (Fig. 1).

It is therefore intriguing that A1C levels are significantly correlated in monozygotic twins whether they are concordant for type 1 diabetes or not (4): in a discordant twin pair one twin is treated with insulin, whereas the other one isn't, and thus this degree of correlation suggests that genetic contributors to A1C may be detectable despite the superimposition of a strong environmental modifier. Rigorous estimates of heritability of treated A1C, however, are not available.

In this issue, the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) investigators describe a laborious attempt to uncover genetic determinants of A1C as a measure of glycemic control in treated individuals (5). The DCCT randomized subjects with type 1 diabetes to conventional and intensive treatment (n = 667 and 637, respectively) and demonstrated the benefit of intensive insulin therapy to prevent diabetes complications (6). Of note, although insulin therapy in the intensive arm targeted specific goals, participants randomized to the conventional arm were generally directed to simply avoid symptoms; this distinction is important because the presence of genetic variants that raise A1C (or resist its lowering by insulin) might identify subjects earmarked for more aggressive treatment regimens, thus obliterating the very genetic effect under investigation. Such masking by treatment may be a concern in the intensive arm, but it is much less so in the conventional arm where participants did not adjust their insulin dose by glucose self-monitoring in as many as 82% of their visits.

Paterson et al. report a genome-wide association (GWA) study for variants associated with glycemic control in DCCT/EDIC. The primary trait was the mean of quarterly A1C values measured over 6.5 years of active study. They first performed a full GWA study of 841,000 directly genotyped or 1.7 million imputed (best-guess genotypes based on known correlations in the human genome) single nucleotide polymorphisms (SNPs) for association with mean A1C, stratified by intervention arm as well as in the combined cohort. This yielded 233 SNPs with $P < 10^{-4}$, which were then tested for association with quarterly A1C using repeated measures; 13 SNPs reached a P value $< 10^{-7}$ in this second stage. The 13 SNPs were then tested for association with mean daily glucose and baseline C-peptide (to confirm a glycemic mechanism), with A1C repeated measures in the combined cohort (adjusting for treatment group and testing for interaction with treatment) and with glycemic complications including coronary calcium, neuropathy, hypoglycemia, and time to renal or retinal complications (in each arm separately). It should be noted that none of these tests (with the exception of the separate GWA studies conducted in the intensive and conventional arms) are independent because they were performed in the same individuals.

The authors find that one SNP (rs1358030) ~200 kb from the *SORCS1* gene in chromosome 10 is associated with A1C at $P = 7 \times 10^{-10}$, a value that exceeds the prespecified empirically determined genome-wide significance threshold (5×10^{-8}) (7). Although the association was most evident in the conventional arm, a consistent nominal association ($P = 0.01$) could also be detected in

From the Massachusetts General Hospital/Harvard Medical School, Center for Human Genetic Research and Diabetes Research Center (Diabetes Unit), Boston, Massachusetts.
Corresponding author: Jose C. Florez, jcflorez@partners.org.
DOI: 10.2337/db09-1636

See accompanying original article, p. 539.

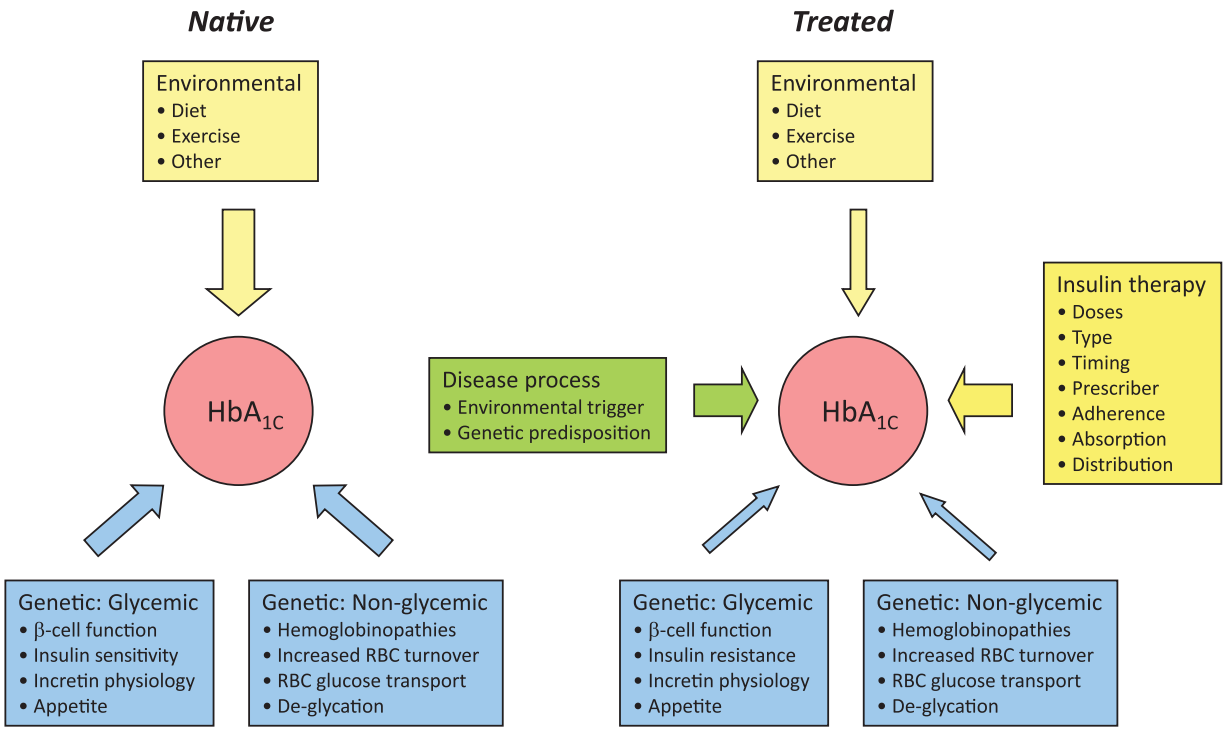

FIG. 1. Simplified theoretical model illustrating the potential contributors to A1C levels in the native (*left*) or treated (*right*) states. The heritability of native A1C is estimated to be ~50%, indicating an approximately even contribution from environmental and genetic factors. Among the genetic components, processes relevant to glycemic physiology and total hemoglobin levels are both operational; the relative contribution of each is the subject of active investigation. In diabetes, the disease process (with distinct interacting genetic and environmental determinants dependent on the type) leads to marked elevations in A1C and, therefore, a large contribution to its variance. Most of that elevation can be overcome with treatment, which represents a strong environmental modifier whose genetic component is limited to endogenous pharmacokinetic and pharmacodynamic parameters. In this scenario, uncovering genetic contributors to A1C becomes much more challenging; how these relate to the genetic variants that underlie native A1C levels is also more difficult to ascertain. RBC, erythrocyte.

the intensive arm. This effect seemed to be glycemia dependent because the same SNP was also associated with mean glucose ($P = 2 \times 10^{-5}$). Because GWA studies query hundreds of thousands of genetic variants, type 1 error must be minimized through stringent significance thresholds and convincing replication; the authors therefore sought a similar replication cohort in the Genetics of Kidneys in Diabetes (GoKinD) Study, a case-control collection of patients with type 1 diabetes with and without diabetic nephropathy. In the control group (851 patients with type 1 diabetes and normal renal function), rs1358030 again showed a nominal association with A1C ($P = 0.01$), which was not seen among patients with nephropathy. Of the other top signals (those with very low *P* values but not quite genome-wide significant), a region in chromosome 9 deserves mention. SNPs in the *BNC2* gene were associated with A1C in DCCT/EDIC ($P \sim 10^{-7}$ in the conventional arm) and also with A1C in a separate meta-analysis of >35,000 nondiabetic individuals performed by the Meta-Analysis of Glucose and Insulin-related traits Consortium (MAGIC) investigators ($P \sim 10^{-4}$). Interestingly, the *SORCS1* SNP did not replicate in MAGIC, and the *BNC2* SNPs did not replicate in GoKinD. SNPs in both regions were associated with diabetes complications in the expected direction: *SORCS1* with hypoglycemia (and less robustly with both retinopathy and nephropathy) and *BNC2* with microvascular end points.

Several questions emerge from these results: *1*) Are the associations real? That is, can we trust they signify true genetic effects despite the strong environmental modifier? *2*) Do they provide biological insight? *3*) What do they mean in clinical practice?

The authors have tried everything possible, given the limitations of their unique cohort, to answer the first question in the affirmative. The distribution of association *P* values does not reveal any systematic bias. As outlined above, one key concern in an interventional study is whether insulin doses may have been raised blindly in carriers of resistant variants, abolishing the genetic effect: adjustment for mean insulin dose did not significantly change the results. The *P* values for association approach or exceed genome-wide significance (for a single scan: two were conducted here), and some evidence of replication is welcome. Consistent associations with other traits in the same cohort are helpful and demonstrate internal consistency but do not provide independent evidence of replication. It is puzzling that some associations seem to replicate in one independent sample but not the other: while power is limiting in GoKinD, it should not be in MAGIC. In either case, it appears that at the very least the top signals reported here merit further study.

Biologically, the gene closest to rs1358030 is *SORCS1*. As in all association studies, whether rs1358030 represents the causal SNP or (more likely) is simply correlated with an ungenotyped causal variant in the region awaits detailed fine-mapping and functional studies; therefore, *SORCS1* may not be the gene that gives rise to the phenotype. Nevertheless, several features make it an interesting candidate: Attie and colleagues have described a quantitative trait locus for fasting insulin in the syntenic region in mice (8), and independent evidence has been obtained in rats (9). There is modest evidence of genetic association in two human studies (10,11), albeit not widely replicated and with variants not correlated to the one

reported here. From a mechanistic standpoint, the *SORCS1* protein product binds to platelet-derived growth factor-BB, which is required for the recruitment of pericytes to the vascular endothelium. Alterations in that process could lead to impaired insulin availability in relevant tissues. With regard to diabetes complications, *SORCS1* is expressed in both the kidney and the retina.

Assuming that the genetic association is real and biologically relevant, the final question relates to its significance in clinical practice. "Glycemic control" as captured by A1C is a heterogeneous trait influenced by individual responsiveness to insulin (which in turn may be affected by other variables such as presence of autoantibodies, BMI, insulin resistance, type of insulin, or mode of administration), timing of injections, environmental factors (e.g., diet and exercise), subject adherence, and prescriber decisions. On the other hand, the key end point of glycemia is easily and cheaply measured, resistance can be overcome by simply increasing insulin doses, and many of the contributing factors mentioned above are likely unrelated to this genetic signal; therefore, how this information could be incorporated by the practitioner is not readily apparent. Thus, the main contributions of this work may relate to the development of analytical methods (e.g., illustrating the advantages of repeated longitudinal measures of a quantitative variable as a more stable end point than a single measurement) and the identification of a pathway potentially relevant to glycemic physiology that may spawn further research.

ACKNOWLEDGMENTS

J.C.F. is supported by a Physician Scientist Development Award by the Massachusetts General Hospital and a Clinical Scientist Development Award from the Doris Duke Charitable Foundation.

J.C.F. has received consulting honoraria from Merck, bioStrategies, XOMA, and Daiichi-Sankyo, and he has been a paid invited speaker at internal scientific seminars hosted by Pfizer and Alnylam Pharmaceuticals. He is an investigator in the Meta-Analysis of Glucose and Insulin-related traits Consortium (MAGIC), which contributed to the article reviewed here. No other potential conflicts of interest relevant to this article were reported.

REFERENCES

1. American Diabetes Association. Executive summary: standards of medical care in diabetes–2009. Diabetes Care 2009;32(Suppl. 1):S6–S12
2. International Expert Committee. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes.Diabetes Care 2009;32:1327–1334
3. Meigs JB, Panhuysen CI, Myers RH, Wilson PW, Cupples LA. A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: the Framingham Offspring Study. Diabetes 2002;51:833–840
4. Snieder H, Sawtell PA, Ross L, Walker J, Spector TD, Leslie RD. HbA(1c) levels are genetically determined even in type 1 diabetes: evidence from healthy and diabetic twins. Diabetes 2001;50:2858–2863
5. Paterson AD, Waggott D, Boright AP, Hosseini SM, Shen E, Sylvestre M-P, Wong I, Bharaj B, Cleary PA, Lachin JM, MAGIC (Meta-Analyses of Glucose and Insulin-related traits Consortium), Below JE, Nicolae D, Cox NJ, Canty AJ, Sun L, Bull SB, Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Research Group. A genome-wide association study identifies a novel major locus for glycemic control in type 1 diabetes, as measured by both A1C and glucose. Diabetes 2010;59:539–549
6. The Diabetes Control and Complications Trial. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 1993;329:977–986
7. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 2008;32:381–385
8. Clee SM, Yandell BS, Schueler KM, Rabaglia ME, Richards OC, Raines SM, Kabara EA, Klass DM, Mui ET, Stapleton DS, Gray-Keller MP, Young MB, Stoehr JP, Lan H, Boronenkov I, Raess PW, Flowers MT, Attie AD. Positional cloning of *Sorcs1*, a type 2 diabetes quantitative trait locus. Nat Genet 2006;38:688–693
9. Granhall C, Park HB, Fakhrai-Rad H, Luthman H. High-resolution quantitative trait locus analysis reveals multiple diabetes susceptibility loci mapped to intervals <800 kb in the species-conserved Niddm1i of the GK rat. Genetics 2006;174:1565–1572
10. Goodarzi MO, Lehman DM, Taylor KD, Guo X, Cui J, Quiñones MJ, Clee SM, Yandell BS, Blangero J, Hsueh WA, Attie AD, Stern MP, Rotter JI. SORCS1: a novel human type 2 diabetes susceptibility gene suggested by the mouse. Diabetes 2007;56:1922–1929
11. Florez JC, Manning AK, Dupuis J, McAteer J, Irenze K, Gianniny L, Mirel DB, Fox CS, Cupples LA, Meigs JB. A 100K genome-wide association scan for diabetes and related traits in the Framingham Heart Study: replication and integration with other genome-wide datasets. Diabetes 2007;56:3063–3074