Title

A Genome-Wide Association Study Identifies Blood Disorder—Related Variants Influencing Hemoglobin A1c With Implications for Glycemic Status in U.S. Hispanics/Latinos

Journal

Diabetes Care, 42(9)

ISSN

0149-5992

Authors

Moon, Jee-Young
Louie, Tin L
Jain, Deepti
et al.

Publication Date

2019-09-01

DOI

10.2337/dc19-0168

Peer reviewed

A Genome-Wide Association Study Identifies Blood Disorder–Related Variants Influencing Hemoglobin A_{1c} With Implications for Glycemic Status in U.S. Hispanics/Latinos

Diabetes Care 2019;42:1784–1791 | https://doi.org/10.2337/dc19-0168

Jee-Young Moon,[1] Tin L. Louie,[2] Deepti Jain,[2] Tamar Sofer,[2,3] Claudia Schurmann,[4] Jennifer E. Below,[5] Chao-Qiang Lai,[6] M. Larissa Aviles-Santa,[7] Gregory A. Talavera,[8] Caren E. Smith,[6] Lauren E. Petty,[5] Erwin P. Bottinger,[9] Yii-Der Ida Chen,[10] Kent D. Taylor,[10] Martha L. Daviglus,[11] Jianwen Cai,[12] Tao Wang,[1] Katherine L. Tucker,[13] José M. Ordovás,[6,14] Craig L. Hanis,[5] Ruth J.F. Loos,[4] Neil Schneiderman,[15] Jerome I. Rotter,[10] Robert C. Kaplan,[1,16] and Qibin Qi[1]

OBJECTIVE

We aimed to identify hemoglobin A_{1c} (HbA_{1c})-associated genetic variants and examine their implications for glycemic status evaluated by HbA_{1c} in U.S. Hispanics/Latinos with diverse genetic ancestries.

RESEARCH DESIGN AND METHODS

We conducted a genome-wide association study (GWAS) of HbA_{1c} in 9,636 U.S. Hispanics/Latinos without diabetes from the Hispanic Community Health Study/Study of Latinos, followed by a replication among 4,729 U.S. Hispanics/Latinos from three independent studies.

RESULTS

Our GWAS and replication analyses showed 10 previously known and novel loci associated with HbA_{1c} at genome-wide significance levels ($P < 5.0 \times 10^{-8}$). In particular, two African ancestry–specific variants, *HBB*-rs334 and *G6PD*-rs1050828, which are causal mutations for sickle cell disease and *G6PD* deficiency, respectively, had ∼10 times larger effect sizes on HbA_{1c} levels (β = −0.31% [−3.4 mmol/mol]) and −0.35% [−3.8 mmol/mol] per minor allele, respectively) compared with other HbA_{1c}-associated variants (0.03–0.04% [0.3–0.4 mmol/mol] per allele). A novel Amerindian ancestry–specific variant, *HBM*-rs145546625, was associated with HbA_{1c} and hematologic traits but not with fasting glucose. The prevalence of hyperglycemia (prediabetes and diabetes) defined using fasting glucose or oral glucose tolerance test 2-h glucose was similar between carriers of *HBB*-rs334 or *G6PD*-rs1050828 HbA_{1c}-lowering alleles and noncarriers, whereas the prevalence of hyperglycemia defined using HbA_{1c} was significantly lower in carriers than in noncarriers (12.2% vs. 28.4%, $P < 0.001$). After recalibration of the HbA_{1c} level taking *HBB*-rs334 and *G6PD*-rs1050828 into account, the prevalence of hyperglycemia in carriers was similar to noncarriers (31.3% vs. 28.4%, $P = 0.28$).

CONCLUSIONS

This study in U.S. Hispanics/Latinos found several ancestry-specific alleles associated with HbA_{1c} through erythrocyte-related rather than glycemic-related pathways. The potential influences of these nonglycemic-related variants need to be considered when the HbA_{1c} test is performed.

[1]*Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY*
[2]*Department of Biostatistics, University of Washington, Seattle, WA*
[3]*Division of Sleep and Circadian Disorders, Department of Medicine, Brigham and Women's Hospital, Boston, MA*
[4]*The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY*
[5]*Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX*
[6]*Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, MA*
[7]*National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD*
[8]*Graduate School of Public Health, San Diego State University, San Diego, CA*
[9]*Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, NY*
[10]*Institute for Translational Genomics and Population Sciences, Harbor-UCLA Medical Center, Torrance, CA*
[11]*Institute for Minority Health Research, University of Illinois at Chicago, Chicago, IL*

Glycated hemoglobin (HbA_{1c}) results from nonenzymatic and mostly irreversible chemical modification by glucose of hemoglobin molecules carried in erythrocytes. HbA_{1c} reflects the average concentration of blood glucose over the average life span of an erythrocyte (~3 months in humans) and indicates glycemic status over a longer period compared with fasting glucose (1). Thus, HbA_{1c} is used both as a measure of glycemic control and as a diagnostic criterion for diabetes (2).

HbA_{1c} levels are heritable, with a heritability of ~50% (3). Previous genome-wide association studies (GWAS) have identified several diabetes- and glucose metabolism–related genetic loci associated with HbA_{1c} (4–7). In addition, GWAS have identified several erythrocyte-related genetic loci that are associated with HbA_{1c} (4,7–9). HbA_{1c} levels attributed to nonglycemic-related genetic variants may not reflect glycemic status, and this implication has been noted in diabetes screening and diagnosis (4,7,9,10). For example, the common *G6PD* variant rs1050828, which affects red blood cell (RBC) life span, was recently found to be associated with lower HbA_{1c} in African Americans, and it was estimated that 650,000 African Americans with diabetes would be missed when screened by HbA_{1c} if this genetic information was not taken into account (10). Another study reported that African Americans with sickle cell trait, who are heterozygous for an abnormal hemoglobin allele and usually have no symptoms of sickle cell disease, had lower HbA_{1c} levels compared with those without sickle cell trait (11).

Prior GWAS of HbA_{1c} levels have been largely conducted in populations of European ancestry (4,5,12) and East Asians (6,7,13). A recent transethnic genome-wide meta-analysis included individuals from other ethnic groups (e.g., African American, South Asian) (10), but no GWAS of HbA_{1c} levels has been conducted in U.S. Hispanics/Latinos. Understanding genetic determinants of HbA_1 in this population is of public health and clinical importance because U.S. Hispanics/Latinos, the largest minority group in the U.S., are disproportionally affected by diabetes (14) and also have poorer diabetes management compared with non-Hispanic whites (15). Moreover, the diverse U.S. Hispanic/Latino population, admixture of African, European, and Amerindian ancestries, may offer opportunities to identify novel ancestry-specific alleles affecting HbA_{1c} levels (16). Therefore, this study conducted a GWAS of HbA_{1c} in U.S. Hispanics/Latinos of diverse backgrounds using data from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) and replicated novel genetic variants associated with HbA_{1c} levels using data from other Hispanic/Latino cohorts. We also examined the potential implications of nonglycemic-related HbA_{1c} variants for glycemic status evaluated using HbA_{1c}.

RESEARCH DESIGN AND METHODS

Study Participants

The HCHS/SOL is a population-based study of 16,415 Hispanic/Latino adults, aged 18–74 years, living in four U.S. metropolitan areas (Bronx, NY; Chicago, IL; Miami, FL; and San Diego, CA) (17,18). A comprehensive battery of interviews relating to personal and family characteristics and health status and behaviors, and a clinical assessment with a blood draw, were conducted at an in-person baseline clinic visit during 2008–2011. The Starr County Health Study (SCHS) is a population-based study of 1,980 Mexican-American adults (aged ≥20 years) in Starr County, TX. Survey collection has been previously described (19). The Boston Puerto Rican Health Study (BPRHS) is a longitudinal cohort of 1,500 Puerto Rican adults (aged 45–75 years) living in the greater Boston, MA, area. Baseline data from this ongoing study have been described elsewhere (20). The Bio*Me* Biobank is an ongoing hospital- and outpatient-based population research study that has enrolled more than 34,000 participants since September 2007. This is a electronic medical records–linked biobank that integrates research data and clinical care information for consented patients at the Mount Sinai Medical Center, which serves diverse local communities of upper Manhattan (21).

The analysis excluded participants with diabetes (self-reported, on antihyperglycemic medications, HbA_{1c} ≥6.5% [48 mmol/mol], fasting glucose ≥200 mg/dL, or glucose ≥200 mg/dL after oral glucose tolerance test [OGTT]), those with self-reported history of major blood abnormalities (self-reported, if known), and those who had received a blood transfusion 3 months before HbA_{1c} measures. The GWAS of HbA_{1c} levels described here included 9,636 U.S. Hispanics/Latinos from the HCHS/SOL (discovery study) and 4,777 U.S. Hispanics/Latinos from three replication studies, the SCHS (n = 395), the BPRHS (n = 832), and the Bio*Me* (n = 3,550). Characteristics of study participants are reported in Supplementary Table 1. The study was approved by the Institutional Review Boards at all participating institutions, and all participants gave written informed consent.

HbA_{1c}, Plasma Glucose, Liver, and Hematologic Measures

In the HCHS/SOL, HbA_{1c} was measured in EDTA whole blood using a Tosoh G7 automated high-performance liquid chromatography analyzer (Tosoh Bioscience, San Francisco, CA). In the SCHS, HbA_{1c} was measured in EDTA whole blood using the DCA Vantage Analyzer point of care device (Siemens, Malvern, PA) following standard protocols. In the BPRHS, a Tosoh G7 automated

[12]*Department of Biostatistics and Collaborative Studies Coordinating Center, University of North Carolina, Chapel Hill, NC*
[13]*Department of Biomedical and Nutritional Sciences, University of Massachusetts Lowell, Lowell, MA*
[14]*IMDEA Food Institute, Campus de Excelencia Internacional Universidad Autónoma de Madrid–Consejo Superior de Investigaciones Científicas, Madrid, Spain*
[15]*Department of Psychology, University of Miami, Miami, FL*
[16]*Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA*
Corresponding author: Qibin Qi, qibin.qi@einstein.yu.edu
Received 24 January 2019 and accepted 24 May 2019
This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc19-0168/-/DC1.

high-performance liquid chromatography analyzer was used to measure HbA_{1c}. In the Bio*Me*, data on HbA_{1c} were extracted from electronic medical records of participants at the time of enrollment. In the HCHS/SOL, plasma glucose (fasting and 2 h OGTT glucose) was measured using a hexokinase enzymatic method (Roche Diagnostics Corporation, Indianapolis, IN). Hemogram and platelet count were measured using a Sysmex XE-2100 instrument (Sysmex America, Mundelein, IL). Serum iron and unsaturated iron binding capacity (UIBC) were measured on a Roche Modular P chemistry analyzer using a Fe reagent kit and a UIBC reagent kit, respectively (Roche Diagnostics), and ferritin was measured in serum with Roche reagents on a Cobas 6000 Analyzer (Roche Diagnostics) using a particle-enhanced immunoturbidimetric assay. Total iron binding capacity was calculated as the sum of serum iron and UIBC, and transferrin saturation was calculated as the percentage of serum iron in total iron binding capacity. Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) were measured in serum on a Roche Modular P Chemistry Analyzer using an α-ketoglutaratic enzymatic method (Roche Diagnostics). Suspected nonalcoholic fatty liver disease was defined by the elevated aminotransferase levels as AST $>$31 IU/mL or ALT $>$40 IU/mL for men and AST or ALT $>$31 IU/mL for women (22).

Genotyping and Imputation

In the HCHS/SOL, genotyping was performed with an Illumina custom array (15041502 B3), which consists of the Illumina Omni 2.5M array (HumanOmni 2.5-8v1-1) plus ~150,000 custom single nucleotide polymorphisms (SNPs), with the quality control performed at the HCHS/SOL Genetic Analysis Center (16). Genome-wide imputation was performed with 1000 Genomes Project (1000G) phase 1 worldwide reference panel (v3, released March 2012) using SHAPEIT2 and IMPUTE2 software, as described previously (16). In the SCHS, genome-wide SNPs were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 and imputed to the 1000G phase 1 reference data (19). In the BPRHS, genome-wide SNPs were genotyped using the Affymetrix Axiom Genome-Wide LAT Array and imputed to the 1000G phase 1 reference data (23). In the Bio*Me*, genome-wide SNPs were genotyped with the Illumina OmniExpressExome (Bio*Me*-OMNI) (n = 1,867) or the Illumina Multi-Ethnic Global BeadChip (Bio*Me*-MEGA) array (n = 1,683) and imputed to the 1000G phase 3 reference data.

Statistical Analysis

Linear mixed-effects regressions were used to test genome-wide SNP-HbA_{1c} associations among 9,636 individuals without diabetes using data from HCHS/SOL. Correlations between individuals were accounted for by incorporating covariance matrices corresponding to genetic relatedness (kinship), household, and census block group as random effects. The model also included field center, age, sex, the first five principal components to adjust for ancestry (16), and sampling weights (24). Similar linear regression models were used to test associations between 18 potentially novel SNPs and HbA_{1c} ($P < 1 \times 10^{-5}$), adjusting for age, sex, top principal components, center, and relatedness (if appropriate), using data from the three replication cohorts. Inverse variance fixed-effect meta-analyses were performed to combine the results of replication studies and to combine these with results from HCHS/SOL.

We examined the potential implications of two variants, *HBB*-rs334 and *G6PD*-rs1050828, on the screening of hyperglycemia (prediabetes and diabetes) using HbA_{1c} among 10,470 HCHS/SOL participants without diagnosed diabetes (those with self-reported diabetes or antidiabetic medication use were not included in this analysis). We reestimated the effects of *HBB*-rs334 and *G6PD*-rs1050828 on HbA_{1c} levels, respectively, because the additional samples included the natural right-end of HbA_{1c} distribution, instead of cutoff at 6.5% (48 mmol/mol) in our GWAS analysis, but still without prominent influences by self-awareness or medication. Then, measured HbA_{1c} levels were adjusted to account for genetic variants using the re-estimated effects in the following equation:

$$\begin{aligned} \mathrm{HbA1c}_{\mathrm{adjusted}}(\%) &= \\ &\mathrm{HbA1c}_{\mathrm{measured}}(\%) + 0.33 \\ &\times \#\text{ of T-alleles in } \textit{HBB}\text{-rs334} + 0.35 \\ &\times \#\text{ of A-alleles in } \textit{G6PD}\text{-rs1050828} \end{aligned} \quad \text{Eq. 1}$$

where a hemizygous AO of rs1050828 in the X chromosome for men was coded to have two copies of the A allele. A survey logistic regression was used to compare the prevalence of hyperglycemia defined using fasting glucose, 2-h glucose, or HbA_{1c} between carriers and noncarriers of *HBB*-rs334 T allele or *G6PD*-rs1050828 A allele.

We also constructed an unweighted genetic risk score (GRS) based on five SNPs (*ANK1*-rs4737010, *HK1*-rs72805692, *TMPRSS6*-rs855791, *HBM*-rs145546625, and *TMC6*-rs2748424) that might be associated with HbA_{1c} through the erythrocytic pathway according to a previous GWAS of HbA_{1c} (10), by summing the HbA_{1c}-raising alleles. The associations of this GRS with HbA_{1c} level and prevalence of hyperglycemia were examined, taking into account the complex design of the HCHS/SOL study.

All analyses were performed using R 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria) or SAS 9.4 (SAS Institute, Cary, NC) software.

RESULTS

GWAS of HbA_{1c}

In the discovery GWAS in the HCHS/SOL, we identified 12 genome-wide significant associations with HbA_{1c} ($P < 5.0 \times 10^{-8}$), 6 of which were previously known loci (*G6PC2*, *GCK*, *ANK1*, *HK1*, *FN3K*, and *TMPRSS6*) (Fig. 1 and Table 1). We then sought replication of 18 potentially novel independent SNPs with suggestive associations with HbA_{1c} ($P < 1.0 \times 10^{-5}$) in three independent studies of Hispanics/Latinos in the U.S. (Supplementary Tables 2 and 3). Of note, *G6PD* and *HBB* variants (10,11) were not known to be associated with HbA_{1c} when we selected SNPs for replication in this study. Four SNPs, including rs334 at *HBB* ($P = 2.7 \times 10^{-40}$ in HCHS/SOL), rs145546625 at *HBM* ($P = 8.2 \times 10^{-9}$), rs145546625 at *TMC6* ($P = 2.0 \times 10^{-15}$), and rs1050828 at *G6PD* ($P = 9.6 \times 10^{-132}$), were robustly associated with HbA_{1c} in the combined replication analyses with nominal significance (all $P \leq 0.05$) (Table 1 and Supplementary Table 2). No significant heterogeneity was observed across cohorts. Regional plots for these four novel HbA_{1c} loci are shown in Supplementary Fig. 1. The effect sizes of *HBB* rs334 ($\beta = -0.31\%$ [-3.4 mmol/mol] per minor allele) and *G6PD* rs1050828 ($\beta = -0.35\%$ [-3.8 mmol/mol] per minor allele) on HbA_{1c} were ~10 times larger compared with other HbA_{1c}-related variants (Table 1).

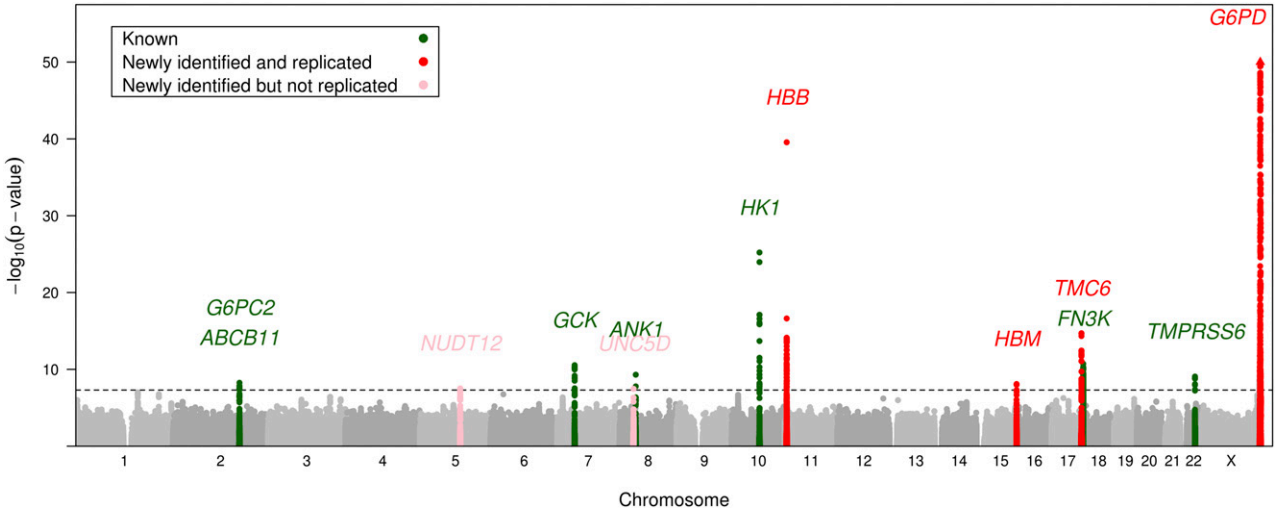

Figure 1—Manhattan plot for GWAS of HbA_{1c} in the HCHS/SOL. A total of 11,510,031 SNPs with MAF >0.01 were tested in 9,636 U.S. Hispanics/Latinos without diabetes from the HCHS/SOL. Six previously known loci (green), four newly identified and replicated loci (red), and two newly identified but not replicated loci (pink) associated with HbA_{1c} at the genome-wide significance level. $P < 5.0 \times 10^{-8}$ (dashed line).

In analyses stratified by Hispanic/Latino background, associations between these four SNPs and HbA_{1c} were consistent across groups with no significant heterogeneity (all P for heterogeneity ≥0.14) (Supplementary Table 4). Given that SNP rs1050828 is located at the X chromosome, we further examined the association by sex, and results were consistent between men and women. In addition, a sensitivity analysis excluding participants with iron deficiency showed similar associations between these SNPs and HbA_{1c} (Supplementary Table 5).

Conditional Analysis

Although our lead SNP rs145546625 at *TMC6* has not been reported before, three SNPs at the same locus (rs2748427, rs761772, and rs2073285) were identified in recent HbA_{1c} GWAS (10,12,13), reaching genome-wide significance in our study as well (Supplementary Table 6 and Supplementary Fig. 2). In the joint analyses of our lead SNP with rs2748427 or rs761772, which are in moderate-to-high linkage disequilibrium (LD) with the lead SNP (r^2 = 0.70, 0.53), our lead SNP showed attenuated but still suggestively significant signals (P = 2.8×10^{-5}, P = 3.1×10^{-5}), whereas the signals of the other SNPs became null (P = 0.59, P = 0.78) (Supplementary Table 6). The joint analysis with rs2073285 in weak LD (r^2 = 0.10) showed that the lead SNP was still at genome-wide significance (P = 1.1×10^{-11}), and rs2073285 was at near suggestive significance (P = 6.2×10^{-5}). Expanding the conditional analysis to ~500 kb of our lead SNP, we found a suggestive but not genome-wide significant secondary signal at rs80149164 (P = 5.3×10^{-7}) (Supplementary Fig. 2).

HbA_{1c}-Related SNPs and Hematologic Traits

We then examined associations of these newly identified four HbA_{1c} variants with diabetes-related traits, hematologic traits (25,26), iron traits (27), and liver-related traits in the HCHS/SOL to explore potential mechanisms underlying observed SNP-HbA_{1c} relationships. These newly identified HbA_{1c} SNPs were significantly associated with hematologic traits rather than glycemic traits

Table 1—Summary of genetic variants associated with HbA_{1c} in U.S. Hispanics/Latinos

				Allele		HCHS/SOL			
SNP	Nearest gene	Chr	Position	Effect	Other	EAF	β (SE)	*P* value	Replication *P* value*
Previously known loci before 2017									
rs557462	*G6PC2/ ABCB11*	2	169777595	C	T	0.182	−0.03 (0.01)	5.7×10^{-9}	
rs2971670	*GCK*	7	44226101	T	C	0.202	0.04 (0.01)	2.9×10^{-11}	
rs4737010	*ANK1*	8	41630447	A	G	0.263	0.03 (0.01)	5.1×10^{-10}	
rs72805692	*HK1*	10	71099109	G	A	0.066	−0.10 (0.01)	6.1×10^{-26}	
rs2256339	*FN3K*	17	80693281	T	A	0.475	−0.03(0.01)	1.9×10^{-11}	
rs855791	*TMPRSS6*	22	37462936	A	G	0.438	0.03 (0.01)	8.7×10^{-10}	
Additional loci									
rs334	*HBB*	11	5248232	A	T	0.013	−0.31 (0.02)	2.7×10^{-40}	0.004
rs145546625	*HBM*	16	220583	T	C	0.066	0.06 (0.01)	8.2×10^{-9}	0.050
rs2748424	*TMC6*	17	76124865	G	C	0.179	0.05 (0.01)	2.0×10^{-15}	0.008
rs1050828	*G6PD*	X	153764217	A	G	0.020	−0.35 (0.01)	9.6×10^{-132}	0.003

β, effect size for each effect allele of SNP on HbA_{1c} (%); Chr, chromosome; EAF, effect allele frequency; Position, in GRCh37/hg19. *Fixed effect meta-analysis of SCHS, BPRHS, Bio*Me*-Omni, and Bio*Me*-MEGA.

or liver-related traits (Supplementary Table 7).

Specifically, *HBB*-rs334 (A-to-T, Glu7Val) is a causal mutation for sickle cell trait (heterozygous of the T allele) and sickle cell disease (homozygous of the T allele), producing abnormal β-globin in hemoglobin. In line with this, the minor T allele (HbA_{1c}-lowering allele) was associated with lower hematocrit ($P = 1.3 \times 10^{-10}$), mean corpuscular volume (MCV) ($P = 1.1 \times 10^{-22}$), and mean corpuscular hemoglobin (MCH) ($P = 1.3 \times 10^{-5}$), and higher mean corpuscular hemoglobin concentration ($P = 3.6 \times 10^{-16}$), which are hematologic characteristics for sickle cell trait and disease. *G6PD*-rs1050828 (G-to-A, Val98-Met) is a causal mutation for glucose-6-phosphate dehydrogenase (G6PD) deficiency, resulting in the premature breakdown of RBCs. The minor A allele was associated with lower RBC count ($P = 1.4 \times 10^{-19}$), higher MCV ($P = 3.7 \times 10^{-13}$), higher MCH ($P = 4.9 \times 10^{-11}$), lower RBC distribution width ($P = 3.7 \times 10^{-29}$), higher iron ($P = 3.0 \times 10^{-5}$), and higher transferrin saturation ($P = 1.1 \times 10^{-6}$), which are hematologic characteristics for G6PD deficiency and subsequent iron overload from chronic anemia. In addition, the minor T allele of *HBM*-rs145546625 was associated with lower MCV and MCH.

HbA_{1c}-Related SNPs and Screening for Diabetes and Prediabetes

Given the much larger effect sizes of *HBB*-rs334 and *G6PD*-rs1050828 on HbA_{1c} compared with other HbA_{1c}-related variants and their nonglycemic-related features (strong associations with hematologic traits rather than glycemic traits), we then examined the influences of these two variants on hyperglycemia screening using HbA_{1c}.

In 10,470 participants without diagnosed diabetes (self-reported diabetes or antidiabetic medication use), 224 participants were carriers of the *HBB*-rs334 T allele, of which 222 individuals were heterozygous (sickle cell trait) and two individuals were homozygous of T allele (sickle cell disease) (Supplementary Table 8). Among 309 carriers of the *G6PD*-rs1050828 A allele in X chromosome, 91 individuals were homozygous of AA for women and AO for men, and 218 women were heterozygous. With exclusion of the two individuals with sickle cell disease by rs334 hereinafter, the A-to-T mutation of *HBB*-rs334 lowered HbA_{1c} by 0.33% (3.6 mmol/mol). The G-to-A mutation of *G6PD*-rs1050828 lowered HbA_{1c} by 0.35% (3.8 mmol/mol), with no significant difference by sex.

We then classified individuals into two groups, noncarriers and carriers of the *HBB*-rs334 T allele or *G6PD*-rs1050828 A allele. Carriers of HbA_{1c}-lowering alleles tended to have lower HbA_{1c} levels ($P < 0.001$) as expected, but fasting glucose and 2-h glucose levels did not differ significantly between groups ($P = 0.14$ and $P = 0.50$, respectively) (Supplementary Table 9). At the same fasting glucose levels, carriers tended to have lower measured HbA_{1c} than noncarriers, whereas after genetic recalibration by Eq. 1, HbA_{1c} levels in respect to fasting glucose became comparable between carriers and noncarriers (Fig. 2 and Supplementary Fig. 3).

Carriers and noncarriers had comparable prevalence of hyperglycemia (prediabetes and undiagnosed diabetes) defined by fasting glucose ≥100 mg/dL or 2-h glucose ≥140 mg/dL (carriers vs. noncarriers: 21.2% vs. 25.4%, 18.6% vs. 21.8%; $P = 0.10$, $P = 0.17$, respectively), whereas carriers had lower prevalence of hyperglycemia defined by HbA_{1c} ≥5.7% (39 mmol/mol [12.2% vs. 28.4%], $P < 0.001$) compared with noncarriers (Table 2). After recalibration, there was no significant difference in the prevalence of hyperglycemia defined by genetically adjusted HbA_{1c} between carriers and noncarriers (31.3% vs. 28.4%, $P = 0.28$). Results were similar but not significant when we examined the potential implications of these genetic variants in the diabetes screen using HbA_{1c} levels, which might be due to the small number of individuals with undiagnosed diabetes among carriers (Supplementary Table 10).

Although other nonglycemic-related genetic variants have much smaller effects on HbA_{1c} compared with *HBB*-rs334 and *G6PD*-rs1050828, a combined effect may have a considerable impact on HbA_{1c} for the hyperglycemia screening. Hence, we constructed an unweighted GRS of five potentially erythrocytic genetic variants. As expected, the GRS showed an additive effect on HbA_{1c} levels (Supplementary Fig. 4). We classified individuals into two groups by a cutoff of 5th percentile of the GRS. Thus, the comparison between these two groups was comparable to that between carriers (~5% of the study population) and noncarriers (~95% of the study population) of *HBB* or *G6PD* variants. Individuals with the bottom 5% of the GRS had lower HbA_{1c} levels compared with those with higher GRS (mean 5.38% [35 mmol/mol] vs. 5.50% [37 mmol/mol]; $P = 1.2 \times 10^{-7}$). The prevalence of hyperglycemia defined by fasting glucose or 2-h glucose was comparable between the two groups, whereas the prevalence of hyperglycemia defined by HbA_{1c} was lower in the lower GRS group than in the higher GRS group (21.9% vs. 28.1%) (Supplementary Table 11).

CONCLUSIONS

This study, the first GWAS of HbA_{1c} to date among Hispanics/Latinos in the U.S., showed multiple previously known as well as novel loci associated with HbA_{1c} at genome-wide significance level ($P < 5.0 \times 10^{-8}$). In particular, two blood disorder–related variants, *HBB*-rs334 and *G6PD*-rs1050828, showed ~10-fold larger effect sizes (0.3–0.4% [3.3–4.4 mmol/mol] per allele) on HbA_{1c} compared with other variants (0.03–0.04% [0.3–0.4 mmol/mol] per allele). The *G6PD* variant was recently identified in a GWAS of HbA_{1c} in African Americans (10), and sickle cell trait (determined by the *HBB*-rs334) was reported to be associated with HbA_{1c} in African Americans (11).

We found that the HbA_{1c} level was significantly lower by 0.33% (3.6 mmol/mol; 95% CI 0.24–0.42 [2.6–4.6 mmol/mol]) per A allele of *HBB*-rs334 in Hispanics/Latinos, comparable to the reported difference in HbA_{1c} between African Americans with and without sickle cell trait (0.29% [3.2 mmol/mol]; 95% CI 0.23–0.35 [2.5–3.8 mmol/mol]) (11). The missense A allele of rs334 is observed in African ancestry at a frequency of 0.1, is very rare in Amerindian ancestry at a frequency of 0.01, and does not appear in other ancestries from 1000G (25,28), which comaps with malaria risk. The homozygotes of A allele have sickle cell anemia, a lifelong disease by chronic hemolytic anemia, caused by the rigid and sickling form of RBCs formed by polymerization of hemoglobin S especially at low oxygen concentrations (29). Although the heterozygotes (AT; sickle cell trait) are asymptomatic, they have increased complications, such as urinary tract infection, splenic

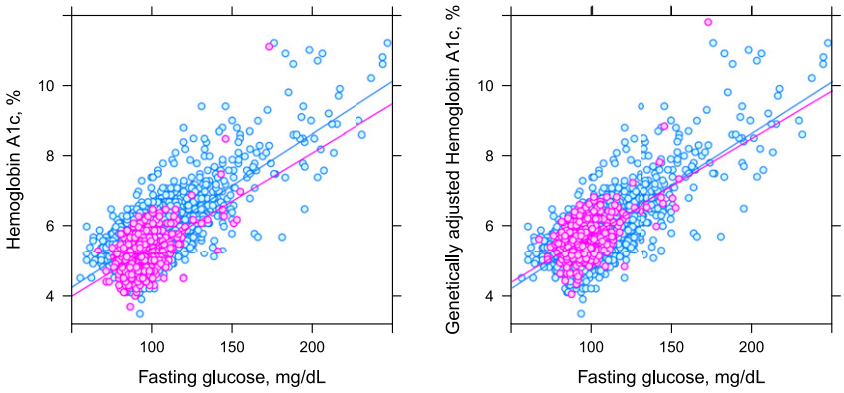

Figure 2—Scatterplots of measured HbA_{1c} and genetically adjusted HbA_{1c} against fasting glucose in carriers and noncarriers of *HBB*-rs334 or *G6PD*-rs1050828 minor alleles. Zoomed-in plots among the individuals without diagnosed diabetes. Red points indicate carriers and blue points indicate noncarriers of *HBB*-rs334 or *G6PD*-rs1050828 minor alleles.

infarction, or sudden death, when exposed to strenuous exercise, high altitudes, dehydration, or low oxygen levels. Lacy et al. (11) hypothesized that lower HbA_{1c} in people with sickle cell trait might be due to the shorter life span of the RBCs and accordingly less opportunity for glycation of hemoglobin. In line with this, our data in the HCHS/SOL demonstrated that this variant was not associated with fasting glucose but was associated with small RBCs (microcytes), characterized by lower MCV and MCH. Because microcytes may be more susceptible to oxidative stress, the life span of RBCs might be shortened (30). These data suggest that *HBB*-rs334 may influence HbA_{1c} through hematologic mechanisms independent of blood glucose. On the contrary, it has also been postulated that lower HbA_{1c} in individuals with sickle cell trait might be due to assay interference by hemoglobin S compared with those without sickle cell trait (11,31).

The X chromosome–linked *G6PD*-rs1050828 is another erythrocytic variant with a relatively large effect on HbA_{1c} in our study of U.S. Hispanics/Latinos (−0.35% [−3.8 mmol/mol] per A allele) as well as in a previous study of African Americans (−0.40% [−4.4 mmol/mol] per A allele for men and −0.34% [−3.7 mmol/mol] for women) (10). The missense A allele in rs1050828 is African-ancestry specific (frequency 0.13), is rarely found in Amerindian ancestry (frequency 0.01), and is not seen in other ancestries (28). *G6PD* (G6PD) catalyzes the reaction of the pentose phosphate pathway and in turn generates the reduced form of glutathione as an antioxidant, which protects RBCs against oxidative stress (32). G6PD deficiency reduces G6PD activity, results in a premature breakdown of RBCs, and aggravates to hemolytic anemia when triggered by certain foods (fava beans), drugs, or infection. This explains the survival advantage of *G6PD* deficiency against malaria because it uses RBCs as a host cell (32). However, the shortened life span of RBCs may result in lower HbA_{1c} levels, regardless of blood glucose levels (10). Further, the regulation via the erythrocytic pathway is supported by the significant associations with RBC count, red cell distribution width, MCH concentration, MCV, and higher iron levels, potentially from chronic hemolysis (33), but no association with other glycemic traits in our study.

TMC6-rs2748424 in the upstream of transmembrane channel-like 6 (TMC6, also known as epidermodysplasia verruciformis 1 [EVER1]) is a common variant observed across all ancestries (0.19–0.46 of minor allele frequency) (28). Three SNPs (rs2748427, rs761772, and rs2073285) near our lead SNP rs2748424 were previously identified in Japanese, European, and non-Hispanic/Latino transethnic populations (10,13). All of these three previously reported SNPs were at genome-wide significance in our Hispanic/Latino populations. SNPs rs2748427 and rs761772 showed moderate-to-high LD with our lead SNP, whereas rs2073285 showed weak LD (r^2 = 0.1) with our lead SNP. Of note, only our lead SNP rs2748424 remained at genome-wide significance level in the conditional analysis, suggesting a potential causal signal represented by this SNP. In addition, our lead SNP and two other SNPs (rs2748427 and rs761772) in LD were related to hematologic traits, including the immature fraction of reticulocytes (34), RBCs, MCH, and MCV (35) rather than glycemic markers such as fructosamine and glycated albumin (13,36), although there were no significant associations with hematologic traits in our study. It has been hypothesized that *TMC6* genetic variants may affect HbA_{1c}, independent of blood glucose levels,

Table 2—Prevalence of hyperglycemia estimated based on fasting glucose, 2-h glucose, and HbA_{1c} in carriers and noncarriers of *HBB*-rs334 or *G6PD*-rs1050828 minor alleles without diagnosed diabetes

	Prevalence of hyperglycemia, % (95% CI)		
	Carrier	Noncarrier	*P* value
Fasting glucose ≥100 mg/dL	21.2 (16.6–25.8)	25.4 (24.0–26.8)	0.10
Post-OGTT glucose ≥140 mg/dL	18.6 (14.2–23.0)	21.8 (20.5–23.2)	0.17
Measured HbA_{1c} ≥5.7% (39 mmol/mol)	12.2 (8.9–15.4)	28.4 (27.1–29.7)	<0.001
Fasting glucose ≥100 mg/dL, post-OGTT glucose ≥140 mg/dL, or HbA_{1c} ≥5.7% (39 mmol/mol)	31.9 (26.5–37.2)	46.6 (45.0–48.2)	<0.001
Genetically adjusted HbA_{1c} ≥5.7% (39 mmol/mol)*	31.3 (26.0–36.5)	28.4 (27.1–29.7)	0.28
Fasting glucose ≥100 mg/dL, post-OGTT glucose ≥140 mg/dL, or genetically adjusted HbA_{1c} ≥5.7% (39 mmol/mol)*	42.0 (36.0–48.0)	46.6 (45.0–48.2)	0.13

Prevalence estimates were computed while accounting for the complex study design of HCHS/SOL. *Hyperglycemia defined with genetically adjusted HbA_{1c}.

through erythrocyte life span, iron handling, or glucose concentration difference across the erythrocyte membrane (13).

Lastly, to the best of our knowledge, *HBM*-rs145546625 was discovered and replicated in this study of Hispanics/Latinos for the first time. *HBM*-rs145546625 originated from Amerindian ancestry (frequency of T allele 0.08) (28). This Amerindian-specific variant is located upstream of hemoglobin-alpha2 (*HBA2*) and downstream of hemoglobin-mu (*HBM*), in high LD ($r^2 > 0.99$) with a splice-site variant rs148323035 in *HBM* (25). There is evidence suggesting that these Amerindian ancestral variants may influence HbA_{1c} levels through blood cell biology (25), but more study is warranted to elucidate how this variant may influence HbA_{1c} through the erythrocyte pathway.

The findings of the current study provide evidence supporting the significant influences of nonglycemic-related genetic variants on diabetes screening using HbA_{1c} tests, consistent with two previous studies (10,11). Specifically, individuals in this study carrying the HbA_{1c}-lowering alleles of *HBB*-rs334 or *G6PD*-rs1050828 had a lower prevalence of hyperglycemia (prediabetes and diabetes) defined by HbA_{1c} levels compared with noncarriers, whereas there were no differences in the prevalence of hyperglycemia defined by fasting glucose or post-OGTT glucose between carriers and noncarriers. In addition to findings in previous studies (10,11), our current analysis using genetically adjusted HbA_{1c} levels demonstrated that this underestimation could be recalibrated when genetic information was taken into account. Although our study did not test influences of nonglycemic-related genetic variants on the prediction of diabetes and related complications such as stroke and coronary artery disease, these potential implications have been reported previously (37).

Taken together, these findings support the current American Diabetes Association recommendations that hemoglobin variants need to be considered when evaluating the HbA_{1c} tests (38). However, it should be noted that the clinical interpretation of HbA_{1c} may need to consider conditions such as shortened erythrocyte life span, hemolytic anemia, microcytosis, iron deficiency, iron overload, vitamin B_{12}, and folate, which may impact the HbA_{1c} levels from the changes in the opportunity of glycation, hemoglobin levels, and glycation rate (27,33,38,39). On the other hand, the increased glycated proteins from high glucose concentrations may increase the oxidative stress or alter erythrocyte conditions (40). More investigations are needed to evaluate the potential use of genetic adjustment with HbA_{1c} levels among individuals without symptoms but at high risk of these blood abnormities (e.g., those with African ancestry and family history of sickle cell disease/trait).

In summary, this study found multiple genetic variants associated with HbA_{1c} levels in U.S. Hispanics/Latinos, particularly two African ancestry–specific variants at *HBB* and *G6PD* and an Amerindian ancestry–specific variant at *HBM*. Further analysis suggested that these variants may influence HbA_{1c} through erythrocyte-related pathways rather than glycemic-related pathways. Our findings expand the understandings of erythrocyte-related genetic determinants of HbA_{1c} and their implications for evaluation of HbA_{1c} tests of glycemic status in U.S. Hispanics/Latinos.

Acknowledgments. The authors thank the participants and staffs of HCHS/SOL, SCHS, Boston Puerto Rican Health Study, and Bio*Me* Biobank for their important contributions.

Funding. HCHS/SOL was performed as a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC-65233), University of Miami (N01-HC-65234), Albert Einstein College of Medicine (N01-HC-65235), Northwestern University (N01-HC-65236), and San Diego State University (N01-HC-65237). The following institutes/centers/offices contributed to the HCHS/SOL first funding period through a transfer of funds to the NHLBI: the National Institute on Minority Health and Health Disparities, the National Institute on Deafness and Other Communication Disorders, the National Institute of Dental and Craniofacial Research, the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), the National Institute of Neurological Disorders and Stroke, and the National Institutes of Health Office of Dietary Supplements. The Genetic Analysis Center at the University of Washington was supported by NHLBI contracts (HHSN268201300005C AM03) and National Institute of Dental and Craniofacial Research contracts (HHSN268201300005C MOD03). Genotyping efforts were supported by NHLBI HSN 26220/20054C, National Center for Advancing Translational Sciences Clinical and Translational Science Institute grant UL1-TR-000123, and NIDDK Diabetes Research Center (DRC) grant DK-063491. SCHS was supported in part by Division of Intramural Research, National Institute of Allergy and Infectious Diseases grant AI-085014, NIDDK grants DK-020595, DK-073541, and DK-085501, and NHLBI grant HL-102830 from the National Institutes of Health and funds from the University of Texas Health Science Center at Houston. Genotyping services were provided by the Center for Inherited Disease Research, which is funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. BRPHS was funded by NHLBI grant P50-HL-105185 and National Institute on Aging grant P01-AG-023394. The Andrea and Charles Bronfman Philanthropies supported the Bio*Me* Biobank program. Q.Q. is supported by NHLBI grants K01-HL-129892, R01-HL-060712, and R01-HL-140976, and by NIDDK grants R01-DK-119268 and R01-DK-120870.

The funding sources had no role in the study design or execution, data analysis, manuscript writing, or manuscript submission.

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

Author Contributions. J.-Y.M. and Q.Q. performed the literature search, analyzed and interpreted data, and drafted the manuscript. T.L.L., D.J., T.S., and T.W. performed data analyses in the HCHS/SOL. C.S., J.E.B., C.-Q.L., C.E.S., L.E.P., E.P.B., K.L.T., J.M.O., C.L.H., and R.J.F.L. contributed to replication data. M.L.A.-S., G.A.T., Y.-D.I.C., K.D.T., M.L.D., J.C., N.S., J.I.R., and R.C.K. contributed to data generation in the HCHS/SOL. Q.Q. designed the study. All authors contributed to data interpretation and editing and reviewing the manuscript. Q.Q. is the guarantor of this work and, as such, has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract and poster form at the 77th Scientific Sessions of the American Diabetes Association, San Diego, CA, 9–13 June 2017.

References

1. Mortensen HB, Christophersen C. Glucosylation of human haemoglobin a in red blood cells studied in vitro. Kinetics of the formation and dissociation of haemoglobin A1c. Clin Chim Acta 1983;134:317–326

2. The 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. Pilia G, Chen WM, Scuteri A, et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. PLoS Genet 2006;2:e132

4. Soranzo N, Sanna S, Wheeler E, et al.; WTCCC. Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways [published correction appears in Diabetes 2011;60:1050–1051]. Diabetes 2010;59:3229–3239

5. Franklin CS, Aulchenko YS, Huffman JE, et al. The TCF7L2 diabetes risk variant is associated with HbA(1)(C) levels: a genome-wide association meta-analysis. Ann Hum Genet 2010;74:471–478

6. Ryu J, Lee C. Association of glycosylated hemoglobin with the gene encoding CDKAL1 in the Korean Association Resource (KARE) study. Hum Mutat 2012;33:655–659

7. Chen P, Takeuchi F, Lee JY, et al. Multiple nonglycemic genomic loci are newly associated with blood level of glycated hemoglobin in East Asians. Diabetes 2014;63:2551–2562
8. Pare G, Chasman DI, Parker AN, et al. Novel association of HK1 with glycated hemoglobin in a non-diabetic population: a genome-wide evaluation of 14,618 participants in the Women's Genome Health Study. PLoS Genet 2008;4: e1000312
9. Syreeni A, Sandholm N, Cao J, et al.; DCCT/EDIC Research Group; FinnDiane Study Group. Genetic determinants of glycated hemoglobin in type 1 diabetes. Diabetes 2019;68:858–867
10. Wheeler E, Leong A, Liu CT, et al.; EPIC-CVD Consortium; EPIC-InterACT Consortium; Lifelines Cohort Study. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: a transethnic genome-wide meta-analysis. PLoS Med 2017;14:e1002383
11. Lacy ME, Wellenius GA, Sumner AE, et al. Association of sickle cell trait with hemoglobin A1c in African Americans. JAMA 2017;317:507–515
12. Prins BP, Kuchenbaecker KB, Bao Y, et al. Genome-wide analysis of health-related biomarkers in the UK Household Longitudinal Study reveals novel associations. Sci Rep 2017;7:11008
13. Hachiya T, Komaki S, Hasegawa Y, et al. Genome-wide meta-analysis in Japanese populations identifies novel variants at the TMC6-TMC8 and SIX3-SIX2 loci associated with HbA1c. Sci Rep 2017;7:16147
14. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and trends in diabetes among adults in the United States, 1988-2012. JAMA 2015;314:1021–1029
15. Stark Casagrande S, Fradkin JE, Saydah SH, Rust KF, Cowie CC. The prevalence of meeting A1C, blood pressure, and LDL goals among people with diabetes, 1988-2010. Diabetes Care 2013;36:2271–2279
16. Conomos MP, Laurie CA, Stilp AM, et al. Genetic diversity and association studies in US Hispanic/Latino populations: applications in the Hispanic Community Health Study/Study of Latinos. Am J Hum Genet 2016;98:165–184
17. Lavange LM, Kalsbeek WD, Sorlie PD, et al. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol 2010;20:642–649
18. Sorlie PD, Aviles-Santa LM, Wassertheil-Smoller S, et al. Design and implementation of the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol 2010;20:629–641
19. Below JE, Gamazon ER, Morrison JV, et al. Genome-wide association and meta-analysis in populations from Starr County, Texas, and Mexico City identify type 2 diabetes susceptibility loci and enrichment for expression quantitative trait loci in top signals. Diabetologia 2011;54:2047–2055
20. Tucker KL, Mattei J, Noel SE, et al. The Boston Puerto Rican Health Study, a longitudinal cohort study on health disparities in Puerto Rican adults: challenges and opportunities. BMC Public Health 2010;10:107
21. Gottesman O, Kuivaniemi H, Tromp G, et al. The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. Genet Med 2013;15:761–771
22. Clark JM, Brancati FL, Diehl AM. The prevalence and etiology of elevated aminotransferase levels in the United States. Am J Gastroenterol 2003;98:960–967
23. Blanco-Rojo R, Delgado-Lista J, Lee YC, et al. Interaction of an S100A9 gene variant with saturated fat and carbohydrates to modulate insulin resistance in 3 populations of different ancestries. Am J Clin Nutr 2016;104:508–517
24. Pfefferman D. The role of sampling weights when modeling survey data. Int Stat Rev 1993; 61:317–337
25. Hodonsky CJ, Jain D, Schick UM, et al. Genome-wide association study of red blood cell traits in Hispanics/Latinos: the Hispanic Community Health Study/Study of Latinos. PLoS Genet 2017;13:e1006760
26. Schick UM, Jain D, Hodonsky CJ, et al. Genome-wide association study of platelet count identifies ancestry-specific loci in Hispanic/Latino Americans. Am J Hum Genet 2016;98: 229–242
27. Raffield LM, Louie T, Sofer T, et al. Genome-wide association study of iron traits and relation to diabetes in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL): potential genomic intersection of iron and glucose regulation? Hum Mol Genet 2017; 26:1966–1978
28. Auton A, Brooks LD, Durbin RM, et al.; 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526: 68–74
29. Kato GJ, Piel FB, Reid CD, et al. Sickle cell disease. Nat Rev Dis Primers 2018;4:18010
30. Vives Corrons JL, Miguel-Garcia A, Pujades MA, et al. Increased susceptibility of microcytic red blood cells to in vitro oxidative stress. Eur J Haematol 1995;55:327–331
31. Rohlfing C, Hanson S, Little RR. Measurement of hemoglobin A1c in patients with sickle cell trait. JAMA 2017;317:2237
32. Guindo A, Fairhurst RM, Doumbo OK, Wellems TE, Diallo DA. X-linked G6PD deficiency protects hemizygous males but not heterozygous females against severe malaria. PLoS Med 2007; 4:e66
33. Fibach E, Rachmilewitz EA. Iron overload in hematological disorders. Presse Med 2017;46: e296–e305
34. Astle WJ, Elding H, Jiang T, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. Cell 2016;167: 1415–1429.e19
35. Kanai M, Akiyama M, Takahashi A, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat Genet 2018;50:390–400
36. Loomis SJ, Li M, Maruthur NM, et al. Genome-wide association study of serum fructosamine and glycated albumin in adults without diagnosed diabetes: results from the Atherosclerosis Risk in Communities Study. Diabetes 2018; 67:1684–1696
37. Au Yeung SL, Luo S, Schooling CM. The impact of glycated hemoglobin (HbA1c) on cardiovascular disease risk: a Mendelian randomization study using UK Biobank. Diabetes Care 2018;41:1991–1997
38. American Diabetes Association. 6. Glycemic targets: *Standards of Medical Care in Diabetes—2018*. Diabetes Care 2018;41(Suppl. 1):S55–S64
39. Cohen RM, Franco RS, Khera PK, et al. Red cell life span heterogeneity in hematologically normal people is sufficient to alter HbA1c. Blood 2008;112:4284–4291
40. Singh VP, Bali A, Singh N, Jaggi AS. Advanced glycation end products and diabetic complications. Korean J Physiol Pharmacol 2014;18:1–14