Title

Multiple Nonglycemic Genomic Loci Are Newly Associated With Blood Level of Glycated Hemoglobin in East Asians

Journal

Diabetes, 63(7)

ISSN

0012-1797

Authors

Chen, Peng
Takeuchi, Fumihiko
Lee, Jong-Young
et al.

Publication Date

2014-07-01

DOI

10.2337/db13-1815

Peer reviewed

Peng Chen,[1] Fumihiko Takeuchi,[2] Jong-Young Lee,[3] Huaixing Li,[4] Jer-Yuarn Wu,[5,6] Jun Liang,[7] Jirong Long,[8] Yasuharu Tabara,[9] Mark O. Goodarzi,[10] Mark A. Pereira,[11] Young Jin Kim,[3] Min Jin Go,[3] Daniel O. Stram,[12] Eranga Vithana,[13–15] Chiea-Chuen Khor,[1,14,16,17] Jianjun Liu,[1,16] Jiemin Liao,[13,14] Xingwang Ye,[4] Yiqin Wang,[4] Ling Lu,[4] Terri L. Young,[15,18] Jeannette Lee,[1] Ah Chuan Thai,[19] Ching-Yu Cheng,[1,13,14,20] Rob M. van Dam,[1,19] Yechiel Friedlander,[21] Chew-Kiat Heng,[17] Woon-Puay Koh,[1,22] Chien-Hsiun Chen,[5,6] Li-Ching Chang,[5] Wen-Harn Pan,[5] Qibin Qi,[23] Masato Isono,[2] Wei Zheng,[8] Qiuyin Cai,[8] Yutang Gao,[24] Ken Yamamoto,[25] Keizo Ohnaka,[26] Ryoichi Takayanagi,[27] Yoshikuni Kita,[28] Hirotsugu Ueshima,[29] Chao A. Hsiung,[30] Jinrui Cui,[10] Wayne H.-H. Sheu,[31–33] Jerome I. Rotter,[34] Yii-Der I. Chen,[34] Chris Hsu,[12] Yukinori Okada,[35,36] Michiaki Kubo,[37] Atsushi Takahashi,[36] Toshihiro Tanaka,[35,38] Frank J.A. van Rooij,[39] Santhi K. Ganesh,[40] Jinyan Huang,[41] Tao Huang,[41] Jianmin Yuan,[42] Joo-Yeon Hwang,[43] CHARGE Hematology Working Group, Myron D. Gross,[44] Themistocles L. Assimes,[45] Tetsuro Miki,[46] Xiao-Ou Shu,[8] Lu Qi,[47,48] Yuan-Tson Chen,[5,49] Xu Lin,[4] Tin Aung,[13] Tien-Yin Wong,[13,14] Yik-Ying Teo,[1,13,16,50–52] Bong-Jo Kim,[3] Norihiro Kato,[2] and E-Shyong Tai[1,19,22]

Multiple Nonglycemic Genomic Loci Are Newly Associated With Blood Level of Glycated Hemoglobin in East Asians

Diabetes 2014;63:2551–2562 | DOI: 10.2337/db13-1815

Glycated hemoglobin A_{1c} (HbA_{1c}) is used as a measure of glycemic control and also as a diagnostic criterion for diabetes. To discover novel loci harboring common variants associated with HbA_{1c} in East Asians, we conducted a meta-analysis of 13 genome-wide association studies (GWAS; *N* = 21,026). We replicated our findings in three additional studies comprising 11,576 individuals of East Asian ancestry. Ten variants showed associations that reached genome-wide significance in the discovery data set, of which nine (four novel variants at *TMEM79* [*P* value = 1.3×10^{-23}], *HBS1L/MYB* [8.5×10^{-15}], *MYO9B* [9.0×10^{-12}], and *CYBA* [1.1×10^{-8}] as well as five variants at loci that had been previously identified [*CDKAL1*, *G6PC2/ABCB11*, *GCK*, *ANK1*, and *FN3KI*]) showed consistent evidence of association in replication data sets. These variants explained 1.76% of the variance in HbA_{1c}. Several of these variants (*TMEM79*, *HBS1L/MYB*, *CYBA*, *MYO9B*, *ANK1*, and *FN3K*) showed no association with either blood glucose or type 2 diabetes. Among individuals with nondiabetic levels of fasting glucose (<7.0 mmol/L) but elevated HbA_{1c} (≥6.5%), 36.1% had HbA_{1c} <6.5% after adjustment for these six variants. Our East Asian GWAS meta-analysis has identified novel variants associated with HbA_{1c} as well as demonstrated that the effects of known variants are largely transferable across ethnic groups. Variants affecting erythrocyte parameters rather than glucose metabolism may be relevant to the use of HbA_{1c} for diagnosing diabetes in these populations.

Glycated hemoglobin A_{1c} (HbA_{1c}) is formed through a nonenzymatic reaction between glucose and hemoglobin. After formation, HbA_{1c} remains and accumulates primarily in erythrocytes throughout its life span. The blood level of HbA_{1c} reflects the average blood glucose level over ~90 days. Genome-wide association studies (GWAS) have identified variants at multiple loci that are associated with HbA_{1c}. In several instances, the presence of these variants is associated with altered glucose homeostasis, e.g., variants in or near solute carrier family 30 (zinc transporter); member 8 (*SLC30A8*) (1); transcription factor 7-like 2 (*TCF7L2*) (2); glucose-6-phosphatase, catalytic, 2 (*G6PC2*); glucokinase (*GCK*); melatonin receptor 1B (*MTNR1B*) (3); and CDK5 regulatory subunit associated protein 1-like 1 (*CDKAL1*) (4). Thus GWAS of HbA_{1c} may uncover variants that are relevant to the regulation of blood glucose or to the pathogenesis of type 2 diabetes (T2D) and complement GWAS for other glycemic traits.

HbA_{1c} is also affected by pathways that are not associated with the regulation of blood glucose. For example, in addition to the associations with HbA_{1c}, several variants close to hemochromatosis (*HFE*); transmembrane protease, serine 6 (*TMPRSS6*); ATPase, class VI, type 11A/tubulin, γ-complex-associated protein 3 (*ATP11A/TUBGCP3*); ankyrin 1, erythrocytic (*ANK1*); spectrin, α, erythrocytic 1 (*SPTA1*); and hexokinase 1 (*HK1*) also showed suggestive or definitive associations with erythrocyte parameters (5–10). Several of these genes were also known to harbor rare variants that cause hereditary anemias (11–13). These

raise the possibility that the impact of these variants on HbA_{1c} relates to their effects on erythrocyte half-life. Alternatively, variants near fructosamine 3 kinase (*FN3K*) may act through their effects on protein deglycation (14). These effects become particularly relevant now in that HbA_{1c} has been adopted as a diagnosis criterion of diabetes, because it exhibits less intraindividual variability than either fasting glucose or 2-h post challenge glucose after an oral glucose tolerance test and does not require fasting (15,16). It is recognized that the glucose and HbA_{1c} criteria are not completely concordant and that genetic variants result in a significant reclassification of individuals based on HbA_{1c} criteria when compared with fasting glucose criteria (17).

To date, with one exception (4), all GWAS for HbA_{1c} have been conducted in populations of European ancestry. We have already demonstrated that genetic association studies in different ethnic groups offer opportunities

[1]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore
[2]National Center for Global Health and Medicine, Tokyo, Japan
[3]Center for Genome Science, National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, Republic of Korea
[4]Key Laboratory of Nutrition and Metabolism, Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China
[5]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan
[6]School of Chinese Medicine, China Medical University, Taichung, Taiwan
[7]Department of Endocrinology, Xuzhou Central Hospital, Xuzhou Clinical School of Xuzhou Medical College, Affiliated Hospital of Southeast University, Xuzhou, Jiangsu, China
[8]Vanderbilt Epidemiology Center and Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN
[9]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Sakyo-ku, Kyoto, Japan
[10]Division of Endocrinology, Diabetes, and Metabolism, Cedars-Sinai Medical Center, Los Angeles, CA
[11]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN
[12]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA
[13]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore
[14]Department of Ophthalmology, National University of Singapore, Singapore
[15]Neuroscience and Behavioural Disorders (NBD) Program, Duke-National University of Singapore Graduate Medical School, Singapore
[16]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore
[17]Department of Paediatrics, National University of Singapore, Singapore
[18]Duke Eye Center, Duke University Medical Center, Durham, NC
[19]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[20]Centre for Quantitative Medicine, Office of Clinical Sciences, Duke-National University of Singapore Graduate Medical School, Singapore
[21]Hebrew University, School of Public Health, Jerusalem, Israel
[22]Duke-National University of Singapore Graduate Medical School, Singapore
[23]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY
[24]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China
[25]Division of Genomics, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan
[26]Department of Geriatric Medicine, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
[27]Department of Medicine and Bioregulatory Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan
[28]Department of Health Science, Shiga University of Medical Science, Seta Tsukinowa-cho, Otsu, Japan
[29]Department of Health Science, and Center for Epidemiologic Research in Asia, Shiga University of Medical Science, Seta Tsukinowa-cho, Otsu, Japan
[30]Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan
[31]Division of Endocrine and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan
[32]School of Medicine, National Yang-Ming University, Taipei, Taiwan
[33]School of Medicine, National Defense Medical Center, Taipei, Taiwan
[34]Institute for Translational Genomics and Biomedical Sciences, Los Angeles Biomedical Research Institute, Harbor-University of California, Los Angeles Medical Center, Torrance, CA
[35]Department of Human Genetics and Disease Diversity, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan
[36]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
[37]Laboratory for Genotyping Development, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
[38]Laboratory for Cardiovascular Diseases, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan
[39]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands
[40]Departments of Internal Medicine and Human Genetics, University of Michigan, Ann Arbor, MI
[41]Department of Epidemiology, Harvard School of Public Health, Boston, MA
[42]University of Pittsburgh Cancer Institute, Pittsburgh, PA
[43]Center for Genome Science, National Institute of Health, Chungcheongbuk-do, Republic of Korea
[44]Department of Laboratory Medicine and Pathology, Medical School, University of Minnesota, Minneapolis, MN
[45]Department of Medicine, Stanford University School of Medicine, Stanford, CA
[46]Department of Geriatric Medicine, Ehime University Graduate School of Medicine, Toon, Ehime, Japan
[47]Department of Nutrition, Harvard School of Public Health, Boston, MA
[48]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA
[49]Department of Pediatrics, Duke University Medical Center, Durham, NC
[50]NUS Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore
[51]Life Sciences Institute, National University of Singapore, Singapore
[52]Department of Statistics and Applied Probability, National University of Singapore, Singapore

Corresponding author: E-Shyong Tai, e_shyong_tai@nuhs.edu.sg.

Received 28 November 2013 and accepted 8 March 2014.

This article contains Supplementary Data online at http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db13-1815/-/DC1.

P.C., F.T., J.-Y.L., H.L., J.-Y.W., J.Lian., J.Lo., Y.T., M.O.G., M.D.G., T.L.A., T.M., X-O.S., L.Q., Y.-T.C., X.L., T.A., T.-Y.W., Y.-Y.T., B.-J.K., N.K., and E-S.T. contributed equally to this work.

to identify novel loci that harbor variants encoding susceptibility to T2D (18,19). Furthermore, some hereditary anemias are more common in Asians, which may also affect HbA_{1c} (20). In this study, we sought to find common genetic variants at novel loci that are associated with blood HbA_{1c} level in the consortium of the Asian Genetic Epidemiology Network (AGEN; http://www.agenconsortium.org/), which consists of East Asian cohorts.

RESEARCH DESIGN AND METHODS

We conducted a meta-analysis of data from 16 cohorts comprising 32,602 individuals of East Asian ancestry (Table 1). The study was carried out in two stages. Stage 1 involved the meta-analysis of GWAS of 19,017 Chinese, Korean, Japanese, and Malay individuals, as well as an in silico lookup of all index single nucleotide polymorphisms (SNPs) with *P* value $<10^{-4}$ and their proxies (in total 198 SNPs) in 2,009 subjects of Chinese ancestry of the Singapore Chinese Health Study (SCHS) of diabetes. This gave a sample size of 21,026 for the discovery stage. In stage 2, we selected five genome-wide significant (*P* value $\leq 5 \times 10^{-8}$) SNPs at potential novel loci from stage 1 and two SNPs that reached suggestive levels of significance at the fatty acid desaturase 2 (*FADS2*) and proteasome (prosome, macropain) 26S subunit, non-ATPase, 13 (*PSMD13*) genes in view of their known impact on blood lipid levels (21) and hematologic traits (22). We then carried out de novo genotyping of these SNPs in 9,592 Japanese individuals. In addition, 48 out of these 198 index SNPs were also genotyped in a study of 1,984 Chinese individuals of the TaiChi study. The findings from all these studies were then meta-analyzed.

Most of the studies were population-based cross-sectional studies in adults. For the studies that used a case-control design to study genetic association with diabetes, only nondiabetic control individuals were included in this study. For other case-control studies, the association testing was done in cases and controls separately. The average HbA_{1c} across the participating studies varied between 4.8 and 5.8%, while the SDs were less than 0.4%. All subjects provided written informed consent. Detailed information of each study can be found in Table 1 and the Supplementary Data.

SNP Genotyping, Quality Control, and Imputation

Our stage 1 cohorts were genotyped on Affymetrix (Santa Clara, CA) and Illumina SNP arrays. We applied stringent sample and SNP quality control (QC) within each study separately. As a result, the QC criteria varied slightly across these studies, but the following steps have been largely applied (Supplementary Table 1). Firstly, we excluded samples that showed cryptic relatedness, were population outliers, or showed inconsistency between clinical and genetic genders. Secondly, we removed SNPs with minor allele frequency (MAF) <1%, genotype call rate <0.95, or Hardy-Weinberg equilibrium *P* value $\leq 1 \times 10^{-6}$.

We imputed our post-QC array genotypes up to the HapMap phase 2 haplotypes (23) using several widely used

Table 1—Demographics of the participant cohorts

Study name	*n*	Design	Ethnicity	Age, mean (SD)	Male, %	BMI, mean (SD)	HbA_{1c} %, mean (SD)	HbA_{1c} IFCC,[1] mean (SD)	λ_{GC}
Stage 1	21,026						5.7 (0.6)	39 (6.6)	
CRC	640	Population	Chinese	43.9 (7.7)	24.7	22.4 (2.2)	5.5 (0.4)	37 (4.4)	1.020
KARE	7,696	Population	Korean	51.6 (8.8)	46.6	24.5 (3.8)	5.6 (0.4)	38 (4.4)	1.052
CAGE-NCGM	323	Population	Japanese	63.0 (6.1)	53.6	22.9 (2.9)	5.2 (0.3)	33 (3.3)	1.001
NHAPC	2,507	Population	Chinese	58.5 (6.0)	42.4	24.3 (3.6)	5.7 (0.4)	39 (4.4)	1.005
SCES	1,580	Population	Chinese	57.7 (9.4)	42.2	23.5 (3.5)	5.8 (0.3)	40 (3.3)	1.002
SiMES	1,727	Population	Malay	57.6 (11.2)	33.9	25.8 (5.1)	5.7 (0.4)	39 (4.4)	1.007
SP2–610	797	Population	Chinese	47.4 (10.6)	17.0	22.2 (3.7)	5.6 (0.4)	38 (4.4)	1.027
SP2–1M	777	Population	Chinese	46.7 (10.1)	61.5	22.8 (3.4)	5.6 (0.4)	38 (4.4)	1.006
SP2–550	266	Population	Chinese	48.3 (12.2)	66.0	23.3 (3.5)	5.6 (0.4)	38 (4.4)	0.992
TWSC	920	Population	Chinese	50.0 (17.8)	50.8	23.6 (3.5)	5.2 (0.4)	33 (4.4)	1.010
SCHS-CHD	1,024	MI controls	Chinese	60.1 (8.0)	67.6	22.9 (3.1)	5.7 (0.4)	39 (4.4)	0.998
	457	MI cases	Chinese	59.8 (7.8)	63.0	22.6 (2.9)	5.7 (0.4)	39 (4.4)	1.008
SBCS	303	BC controls	Chinese	53.0 (8.4)	0.0	24.7 (5.1)	5.8 (0.4)	40 (4.4)	1.012
SCHS-DB	2,009	T2D controls	Chinese	55.2 (7.1)	46.7	22.7 (3.1)	5.5 (0.3)	37 (3.3)	—
Stage 2	11,576								
TaiChi	1,984	Population	Chinese	68.6 (9.0)	49.7	24.3 (3.4)	5.8 (0.3)	40 (3.3)	1.09
CAGE-Fukuoka	4,880	Population	Japanese	63.8 (5.8)	46.2	22.7 (2.7)	4.8 (0.2)	29 (2.2)	NA
JMGP	4,712	Population	Japanese	59.5 (14.3)	35.6	22.8 (3.1)	5.5 (0.4)	37 (4.4)	NA

The participant cohorts are listed in stage 1 (genome-wide association test) and stage 2 (de novo genotyping), respectively. Age, BMI, and HbA_{1c} are given as the mean value and SE in each cohort. λ_{GC} is the inflation factor calculated as per genomic control. CAGE, Cardiovascular Genomic Epidemiology; CRC, Cardiometabolic Risk in Chinese; JMGP, Japanese Millenium Genome Project; KARE, Korea Association Resource; NCGM, National Center for Global Health and Medicine; NHAPC, Nutrition and Health of Aging Population in China; SBCS, Shanghai Breast Cancer Study; SCES, Singapore Chinese Eye Study; SCHS-CHD, Singapore Chinese Health Study of Coronary Heart Disease; SCHS-DB, Singapore Chinese Health Study of Diabetes Mellitus; SiMES, Singapore Malay Eye Study; SP2, Singapore Progressive Study Program; TWSC, Taiwan Super Control Study. [1]The International Federation of Clinical Chemistry unit for HbA_{1c} is mmol/mol.

software packages. For the cohorts of Chinese, Korean, and Japanese ancestry, HapMap haplotypes of the Japanese in Tokyo, Japan (JPT) and Chinese in Beijing, China (CHB) were used as the reference panel in the imputation. For Malays, the combined African Yoruba in Ibadan, Nigeria (YRI); Utah residents with ancestry from Northern and Western Europe (CEU); and JPT+CHB panels were used. The combined HapMap panel resulted in about 1.9 million SNPs in the imputed genotypes, while there were 2.4 million in the imputation data when the JPT+CHB panel was used. Imputed SNPs of MAF <1%, Hardy-Weinberg equilibrium P value $\leq 1 \times 10^{-6}$, or poor imputation quality (IMPUTE info <0.5, BEAGLE allelic R^2 <0.5, or MACH R_{sq} <0.3) were removed (24–26).

In stage 2, the follow-up SNPs were genotyped on three platforms. The two Japanese cohorts were genotyped using the TaqMan system (Life Technologies Corporation, Carlsbad, CA). The TaiChi cohort was genotyped using the Illumina Cardio-MetaboChip (San Diego, CA). The same SNP QC criteria as in stage 1 were applied to the two Japanese studies, while similar sample QC and SNP QC as in stage 1 were applied to TaiChi study (Supplementary Table 1).

Phenotype Definition

The traditional HbA_{1c} unit (percentage of HbA_{1c} in total hemoglobin) was used in the analysis. Subjects with diabetes were excluded. Diabetes was diagnosed if the subject gave a history of physician-diagnosed diabetes, was taking medication for diabetes, or had fasting glucose ≥7 mmol/L or HbA_{1c} ≥6.5% (48 mmol/mol) in those studies where fasting glucose was not available. The HbA_{1c} measurements were fitted into a linear model that adjusted for the sex and BMI of the individuals. The residuals were then normalized to have a mean of 0 and an SD of 1 using inverse-normal transformation.

Association With HbA_{1c} and Meta-analysis

We tested the association of the SNPs with the normalized HbA_{1c} residuals using linear regression assuming additive effects of the dosage of the effect allele. The linear regression was evaluated by several widely used software packages (Supplementary Table 1). Study-specific covariates, such as principal components and sample recruitment sites, were also considered as required (Supplementary Data).

The results from separate studies were combined using the fixed-effect scheme weighted by the inverse of the SE as implemented in METAL (27). Genomic control was applied within METAL.

RESULTS

In total, we identified nine loci harboring variants associated with HbA_{1c} levels in East Asian populations. Among them, four loci were associated with HbA_{1c} for the first time. The fixed-effect meta-analysis showed no significant evidence of heterogeneity for most of the index SNPs in the cohorts we have recruited. Although the heterogeneity test for variants near the *G6PC2*/ATP-binding cassette, subfamily B (MDR/TAP), member 11 (*ABCB11*) locus reached a borderline level of statistical significance, the extent of heterogeneity was moderate (Supplementary Table 2).

Known Associated Loci Replicated in East Asian Populations

In stage 1, the maximal inflation factor was 1.052 in the Korea Association Resource (KARE) study, while those of the other studies were all ~1, indicating that population stratification was unlikely to confound our findings (Table 1). Supplementary Fig. 1 shows the quantile–quantile plots with and without the SNPs known to be associated with HbA_{1c}. This displayed a marked deviation from the null hypothesis of no association that persisted after all known SNPs associated with HbA_{1c} were removed. In this stage, we had ~81% power to discover genomic variants that explain 0.2% phenotype variance at a significance level of $P = 5 \times 10^{-8}$.

The associations between index SNPs at novel and known loci are presented in Table 2. Ten loci showed associations with HbA_{1c} with P values that met the criteria for genome-wide significance (P value $\leq 5 \times 10^{-8}$) (Fig. 1). Of these, five were at loci that were known to harbor variants associated with HbA_{1c} levels. The index SNPs at these known loci were rs7772603 in the *CDKAL1* gene region, rs3755157 at the *G6PC2/ABCB11* locus, rs1799884 near the *GCK* gene, rs4737009 in the *ANK1* gene region, and rs1046875 near the *FN3K* gene. We investigated the linkage disequilibrium (LD) between our top SNPs, and the reported index SNPs at these loci identified in populations of European ancestry (Supplementary Table 3). At the *GCK* and *FN3K* loci, our index SNPs were exactly the same as or were in perfect LD with the index SNPs reported in Europeans. In fact, the SNP identified in populations of European ancestry at both loci (rs1046896 near the *FN3K* locus and rs730497 near the *GCK* locus) (3) also showed genome-wide significant associations in our meta-analysis (P value = 3.4×10^{-13} and 1.3×10^{-18} after genomic control). For *ANK1*, we replicated the primary signal rs4737009 reported in populations of European ancestry (P value = 1.3×10^{-15}) (3). This same study identified a second variant (rs6474359) in the *ANK1* region, which was not in LD with the primary signal at this locus. This second SNP showed only a minor association with HbA_{1c} in our study (P value = 9.3×10^{-3}), with an opposite direction of effect compared with Europeans.

The index SNP identified in European populations at *G6PC2/ABCB11* (rs552976) did not show any association with HbA_{1c} in our meta-analysis (P value = 0.54). However, we did identify an association with rs3755157 near this locus that reached genome-wide significance. Although there was no evidence of LD between this and rs552976 in either HapMap panels of European (CEU) or Asian (JPT and CHB) ancestry (Supplementary Table 3), rs3755157 did show a genome-wide significant association with fasting glucose in populations of European

Table 2—Association of the top hits of stage 1 and stage 2 in East Asians

SNP	Gene	Chr	Base pair position	Alleles	Stage	Risk allele frequency	Effect (SE)	*P* value	*n*
Novel loci									
rs6684514*	TMEM79	1	154,522,080	G/A	1	0.75	0.09 (0.01)	1.1E-15	20,831
					2	0.78	0.09 (0.02)	2.1E-09	9,494
					1+2	0.76	0.09 (0.01)	1.3E-23	30,325
rs9399137*	HBS1L/MYB	6	135,460,711	T/C	1	0.72	0.06 (0.01)	1.9E-08	20,535
					2	0.65	0.07 (0.01)	7.6E-08	9,501
					1+2	0.69	0.07 (0.01)	8.5E-15	30,036
rs1467311	9q31.2	9	109,576,753	G/A	1	0.22	0.07 (0.01)	2.9E-08	20,845
					2	0.24	0.01 (0.01)	3.5E-01	11,474
					1+2	0.23	0.04 (0.01)	1.0E-06	32,319
rs540078	PSMD13	11	244,256	T/C	1	0.46	0.04 (0.01)	6.2E-06	20,865
					2	0.39	0.01 (0.01)	4.1E-01	9,485
					1+2	0.44	0.03 (0.01)	3.2E-05	30,350
rs174570	FADS2	11	61,353,788	C/T	1	0.52	0.05 (0.01)	5.4E-07	20,639
					2	0.59	0.03 (0.01)	3.8E-02	11,524
					1+2	0.55	0.04 (0.01)	2.0E-07	32,163
rs9933309*	CYBA	16	87,372,433	C/T	1	0.64	0.08 (0.01)	3.3E-08	11,015
					2	0.63	0.05 (0.03)	1.1E-01	1,983
					1+2	0.63	0.07 (0.01)	1.1E-08	12,998
rs11667918*	MYO9B	19	17,093,499	C/T	1	0.61	0.06 (0.01)	1.9E-10	20,835
					2	0.65	0.04 (0.01)	3.7E-03	9,516
					1+2	0.62	0.06 (0.01)	9.0E-12	30,351
Known loci									
rs3755157*	G6PC2/ABCB11	2	169,500,417	T/C	1	0.34	0.07 (0.01)	2.8E-11	20,630
rs7772603*	CDKAL1	6	20,773,925	C/T	1	0.42	0.06 (0.01)	3.5E-08	19,156
rs1799884*	GCK	7	44,195,593	T/C	1	0.19	0.12 (0.01)	1.5E-22	20,874
rs4737009*	ANK1	8	41,749,562	A/G	1	0.50	0.09 (0.01)	1.3E-15	20,558
					2	0.54	0.05 (0.03)	1.4E-01	1,984
					1+2	0.51	0.08 (0.01)	1.1E-15	22,542
rs1046875*	FN3K	17	78,278,715	A/G	1	0.49	0.08 (0.01)	1.6E-14	20,871
					2	0.52	0.11 (0.03)	2.6E-04	1,984
					1+2	0.49	0.08 (0.01)	3.8E-17	22,855

The index SNPs were grouped into novel loci that were first discovered in our study and known loci that have been reported in previous publications. Gene refers to the most relevant gene within each locus. The cytoband of 9q31.2 was designated to rs1467311 since no gene was found in the 400 Kb flanking region nearby. Alleles are given as the effect allele/other allele. Effect and SE are the risk and its SE, respectively. For the novel loci, association result were given for the stage 1, stage 2, and the meta-analyzed stage 1 and stage 2, indicated as 1, 2, and 1+2, respectively. Chr, chromosome number. *The index SNPs showed genome-wide significance in stage 1.

ancestry (Supplementary Tables 4 and 5). We conducted association tests with and without conditioning on the European SNP rs552976 in Singapore cohorts, including 5,147 Chinese and Malays. *P* values for rs3755157 were 1.5×10^{-4} and 1.4×10^{-4} with and without conditioning on rs552976, respectively. In addition, the index SNP at the *CDKAL1* locus (rs7772603) was different from rs7747752 identified in the Korean cohort that formed part of this analysis. However, these two SNPs showed moderate LD in the JPT+CHB panel, and rs7747752 showed an association with HbA_{1c} that was in the same direction and similar magnitude as that reported previously in the Korean population (4), with a suggestive degree of statistical significance in our study (*P* value = 1.1×10^{-5}). This suggests that associations with HbA_{1c} for these two variants may represent the same association signal.

We further compared the direction of the effect for known HbA_{1c}-associated variants between European and East Asian populations (Supplementary Table 6). Of the 17 index SNPs associated with HbA_{1c} in previous publications, rs1800562 (*HFE*) is monomorphic in East Asians, while rs16926246 (*HK1*) was excluded because of the poor imputation quality. Most of the remaining 15 index SNPs showed consistent effects between Europeans and East Asians (Fig. 2). The only exception was rs6474359 at the *ANK1* locus, for which each copy of the T allele was associated with higher HbA_{1c} level by 0.06% in Europeans (*P* value = 1.2×10^{-8}) but lower HbA_{1c} level by 0.08 SD of the normalized HbA_{1c}, which was equivalent to 0.05% HbA_{1c} (*P* value = 9.3×10^{-3}), assuming a normal distribution of HbA_{1c} in our study.

Novel Associated Loci Revealed in East Asian Populations

The other five genome-wide significant hits in stage 1 were all novel associations (Table 2). These five variants and two additional SNPs (rs540078 at *PSMD13* and rs174570 at *FADS2*) were examined in stage 2. Of these, data for the top SNP at cytochrome b-245, α polypeptide (*CYBA*; rs9933309) was available only in the TaiChi study, which consists of 1,984 Chinese. Four of these seven SNPs were

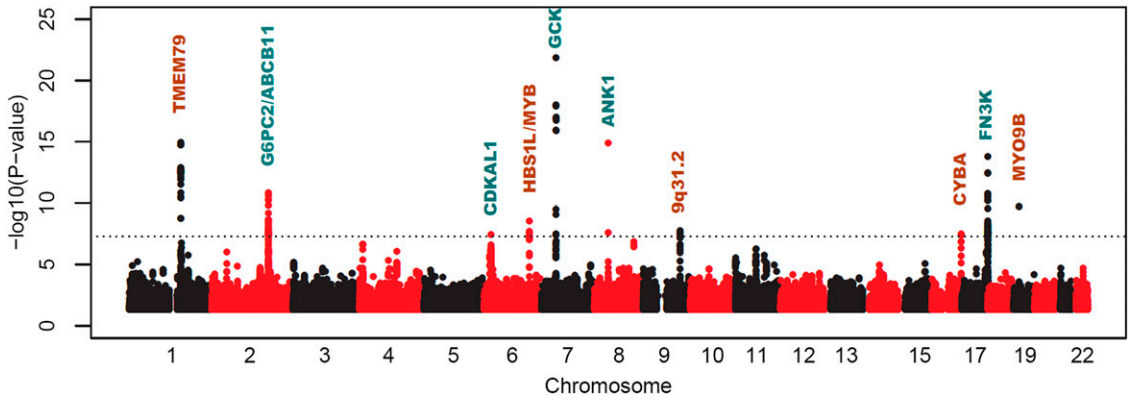

Figure 1—Manhattan plot of genome-wide meta-analysis of stage 1 cohorts. The $-\log_{10}$ of the association *P* values (*y*-axis) are plotted against the genomic coordinates (*x*-axis). The horizontal line in the plot indicates the genome-wide significance (5×10^{-8}). The most relevant gene of each signal was labeled on the top of it, with the novel loci presented in brown and known loci in blue.

replicated successfully in our stage 2 cohorts with consistent effect directions and showed stronger associations after combining the results of stage 2 and stage 1 with *P* values less than 5×10^{-8}. They were variants close to the transmembrane protein 79 (*TMEM79*), Hsp70 subfamily B suppressor 1-like protein/v-myb avian myeloblastosis viral oncogene homolog (*HBS1L/MYB*), *CYBA*, and myosin IXB (*MYO9B*) loci (Table 2). The remaining SNPs,

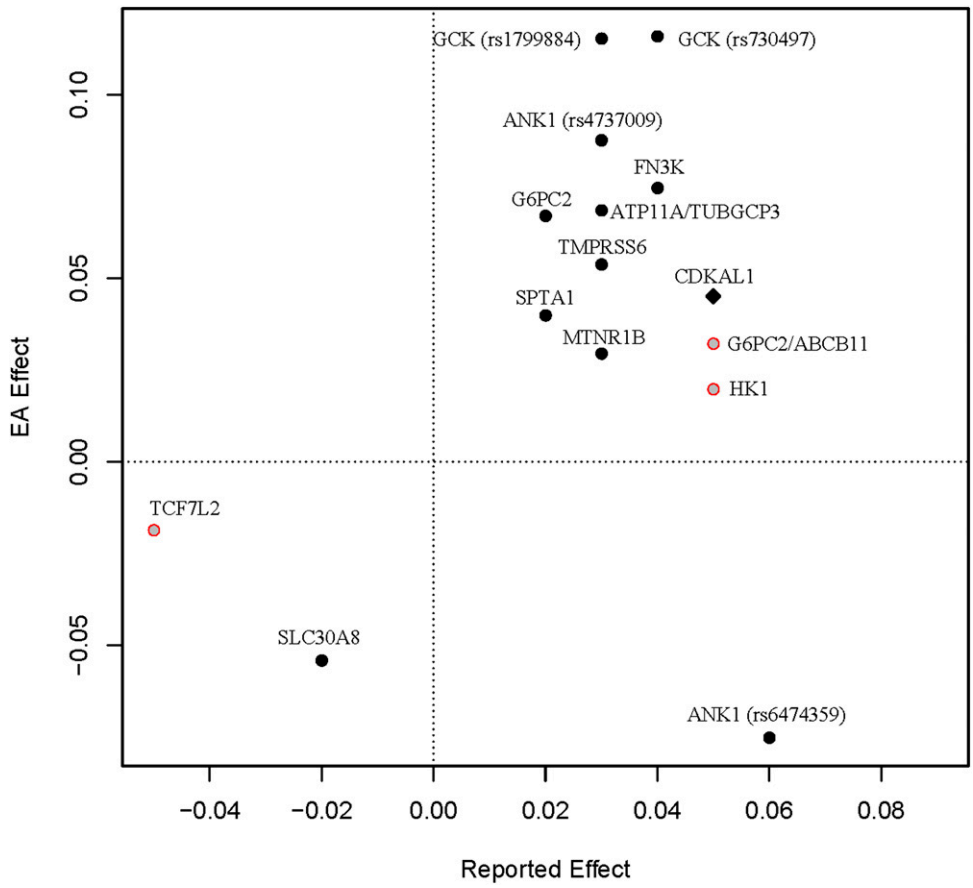

Figure 2—The bivariate plot of the effect directions in our meta-analysis and in previous reports. For the index SNPs in the previous GWAS, we plotted the reported effect (*x*-axis) versus the AGEN effect (*y*-axis). Whenever available, we used the effects reported by Soranzo et al. (3). The solid dots represent the SNPs that were significant (*P* value ≤0.05) in our study, while the hollow red dots represent the insignificant ones. The *CDKAL1* top SNP reported in Koreans was specially marked by a diamond. EA, East Asian.

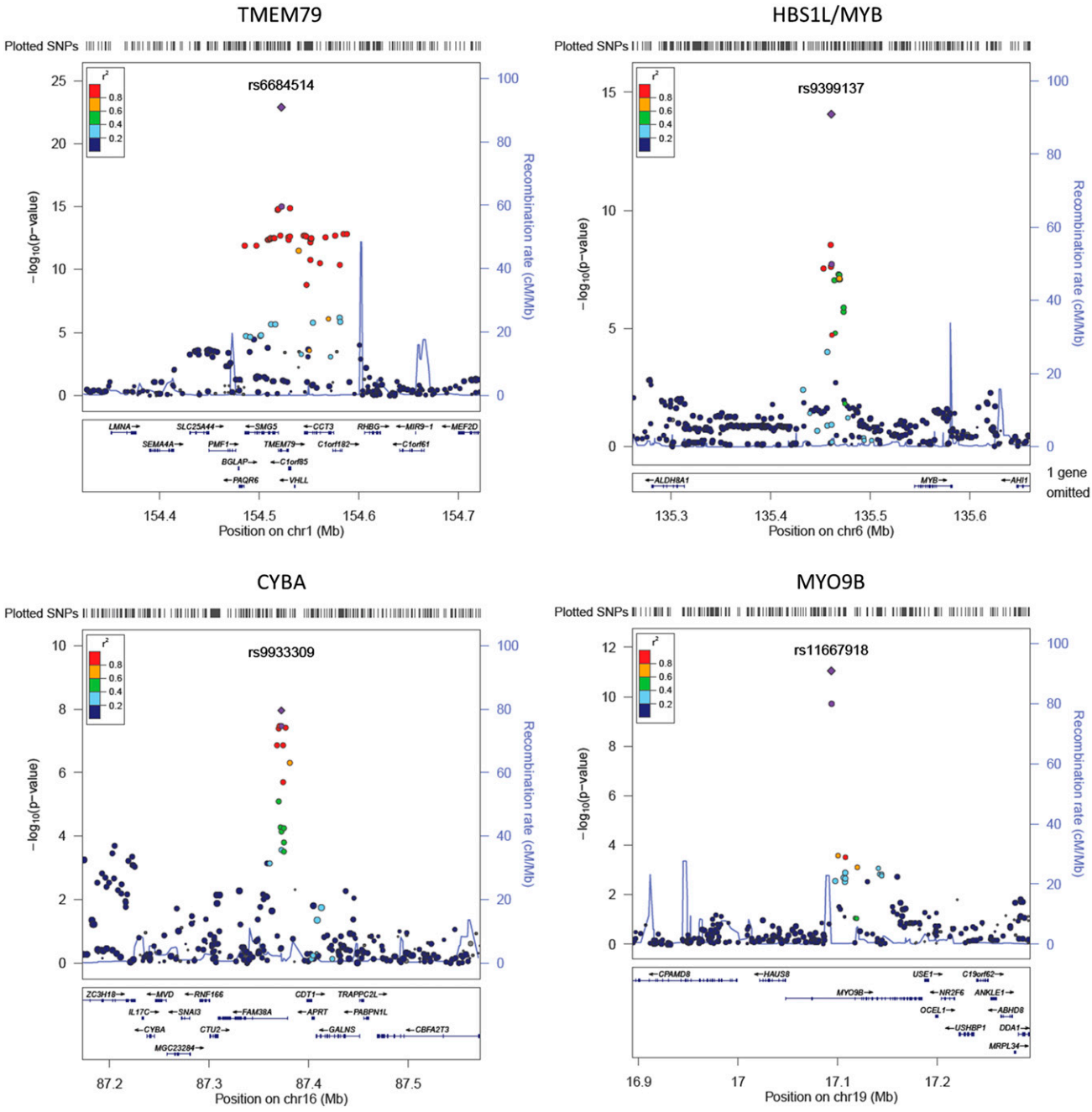

Figure 3— Regional association plots of the novel loci. The $-\log_{10}$ of association *P* values are plotted against the genomic coordinates. The index SNPs are indicated in purple, with circles for stage 1 and squares for stage 1+2. Other SNPs are colored from red to blue as per their LD with the index SNP. Chr, chromosome.

rs1467311 at 9q31.2 (stage 1+2 meta-analysis *P* value = 1.0×10^{-6}), rs174570 in *FADS2* (*P* value = 2.0×10^{-7}), and rs540078 in *PSMD13* (*P* value = 3.2×10^{-5}) did not show genome-wide significant association with HbA_{1c} in stage 1 and stage 2 meta-analysis. The regional association signals at the novel loci and the known loci were presented in Fig. 3 and Supplementary Fig. 2, respectively.

For the novel loci identified in East Asians, the MAF of the index SNPs in HapMap JPT+CHB panel were comparable to those in CEU panel (Supplementary Table 7). In most cases, the direction of effect was the same in AGEN as it was in the study of European ancestry, with the exception of rs6684514 at *TMEM79*. However, this SNP showed only a nominal degree of statistical significance in Europeans (*P* value = 4.0×10^{-2}).

Association With Fasting Glucose and Erythrocyte Traits

We also looked up the associations between the variants that reached genome-wide significance with glucose-related

traits in Europeans (28–30) and Asians (19,31) (Supplementary Tables 4 and 5). SNPs near *GCK* and *G6PC2* were associated with blood glucose in both Europeans and East Asians. SNPs at the *CDKAL1* locus were associated with T2D in East Asians. Among the known loci, *FN3K* and *ANK1* were not associated with glucose level and T2D in either East Asians or Europeans. For the novel SNPs associated with HbA_{1c} that we identified in this study, none showed a statistically significant association with glucose levels except rs9399137 near the *HBS1L/MYB* loci, which showed an association with fasting glucose with borderline statistical significance that did not survive Bonferroni correction. We next examined the associations between these variants and red-cell-associated parameters in Europeans (6,8) and East Asians (32) (Supplementary Tables 8 and 9). Variants close to *TMEM79*, *HBS1/MYB*, *CYBA*, and *ANK1* showed statistically significant associations with red-cell-associated parameters in at least one population.

Phenotype Variance Explained by the Associated Loci

We estimated the phenotype variance explained (PVE) by the genome-wide significant index SNPs in our meta-analysis. The PVE in each cohort was calculated by fitting the raw HbA_{1c} in a linear regression model on the allele dosages of the index SNPs under investigation. The adjusted r-squared from the linear regression model was used as the estimation of PVE. The estimates were obtained for the novel, known loci and all loci separately. The estimates from separate studies were combined using a sample-size-weighted scheme.

In Singapore Chinese Eye Study (SCES), Singapore Malay Eye Study (SiMES), and the three Singapore Progressive Study Program (SP2) cohorts, the known loci and the novel loci explained 1.01% and 0.75% of the total HbA_{1c} variance, respectively, while all the index SNPs as a whole explained 1.76% HbA_{1c} variance.

Reclassification of Diabetes Diagnosis Using HbA_{1c}

Finally, we calculated the proportion of the samples that were reclassified by adjusting for the allele dosage of index SNPs at the six loci that did not show association with glucose or T2D in populations of either European ancestry (28–30) or Asian ancestry (19,31). These included *TMEM79*, *HBS1L/MYB*, *CYBA*, *MYO9B*, *ANK1*, and *FN3K*.

This reclassification analysis was done in 15,150 individuals with fasting glucose measurements available, which came from SP2, KARE, Nutrition and Health of Aging Population in China (NHAPC), National Center for Global Health and Medicine (NCGM), Cardiometabolic Risk in Chinese (CRC), and TaiChi. Individuals with known diabetes (defined by having diabetic history or using diabetic medication) were removed. In the remaining subjects, undiagnosed diabetes was defined as those having fasting glucose ≥7 mmol/L, whereas the nondiabetic individuals had fasting glucose <7 mmol/L. We adjusted the raw HbA_{1c} levels using a linear regression, including the allele dosages of the six SNPs as covariates. We then classified individuals into those with and without diabetes based on HbA_{1c} ≥6.5% (15). We compared the concordance between these three methods for diagnosing diabetes (fasting glucose, HbA_{1c}, and HbA_{1c} adjusted for these six SNPs) (Table 3).

Table 3—Individual reclassification using raw HbA_{1c} or adjusted HbA_{1c}

	Adjusted HbA_{1c}			
	<6.5	≥6.5	Subtotal	Total
Fasting plasma glucose ≥7				
HbA_{1c} <6.5	57	1	58	181
HbA_{1c} ≥6.5	3	120	123	
Fasting plasma glucose <7				
HbA_{1c} <6.5	13,813	40	13,853	14,329
HbA_{1c} ≥6.5	172	304	476	

The reclassification analysis was done in cohorts with fasting glucose measurement. Individuals with known diabetic history or diabetic medication were removed. In the remaining subjects, undiagnosed diabetes was defined as those having fasting glucose ≥7 mmol/L, whereas the nondiabetic individuals had fasting glucose <7 mmol/L. We adjusted the raw HbA_{1c} levels using a linear regression, including the allele dosages of the six SNPs as covariates. We then classified individuals into those with and without diabetes based on HbA_{1c} ≥6.5% (15). We compared the concordance between these three methods for diagnosing diabetes (fasting glucose ≥7.0 mmol/L, HbA_{1c} ≥6.5%, and HbA_{1c} adjusted for these six SNPs ≥6.5%). HbA_{1c} was adjusted on the nonglycemic index SNPs. The subtotal is the number of individuals in each HbA_{1c} category. The reclassification rate in each HbA_{1c} category is the ratio of the number of individuals reclassified after adjustment for the six SNPs to the subtotal. The total is the number of individuals in each fasting plasma glucose category.

One hundred eighty-one subjects had undiagnosed diabetes based on fasting glucose, of which 123 had HbA_{1c} ≥6.5%, of whom three had adjusted HbA_{1c} <6.5%. Fifty-eight individuals had HbA_{1c} <6.5%, of which one had adjusted HbA_{1c} ≥6.5%.

In contrast, 14,96 individuals were nondiabetic, with fasting glucose <7.0 mmol/L. Of these, 476 were classified as having diabetes based on HbA_{1c} ≥6.5%. However, 172 of the 476 individuals had adjusted HbA_{1c} <6.5%; 14,493 had HbA_{1c} <6.5%. Of these, 40 had adjusted HbA_{1c} ≥6.5%. As such, some nonconcordance was observed when diabetes was diagnosed based on fasting plasma glucose ≥7.0 mmol/L and HbA_{1c} ≥6.5%. Among those with fasting plasma glucose ≥7.0 mmol/L, very little of the nonconcordance was explained by these six variants. However, among those with fasting plasma glucose <7.0 mmol/L, 36.13% of those with elevated HbA_{1c} had HbA_{1c} <6.5% after these six variants were taken into consideration.

In populations of European ancestry, the equivalent proportion was 15.18% (3).

DISCUSSION

In this East Asian GWAS meta-analysis, we showed that most of the variants identified in populations of European ancestry had similar effects in East Asians. Variants close

to the *G6PC2/ABCA11*, *GCK*, *ANK1*, and *FN3K* that were first identified in populations of European ancestry also showed associations that reached genome-wide significance in our study. However, at the *G6PC2/ABCB11*, our index SNP was not in LD with the index SNP (rs552976) identified in populations of European ancestry. The latter was not associated with HbA_{1c} in our study (*P* value = 0.54). Therefore, rs3755157 could represent an additional signal at this locus, particularly given that we did not observe obvious attenuation of the associations in the analysis conditioning on rs552976. Another SNP at this locus (rs560887) was associated with fasting glucose in Europeans (28) but not with HbA_{1c} in our study (*P* value = 0.10). The majority of the other known SNPs showed similar direction of effect and effect size in our population as in populations of European Ancestry. The exception was a variant at the *ANK1* locus with showed an association in the opposite direction to that observed in populations of European ancestry. We are not able to explain this finding at this time. However, we would point out that the association in our study, although in a different direction, was far from reaching genome-wide significance. Larger sample sizes will be required to formally test for heterogeneity of effect of this variant in populations of European and East Asian ancestry.

We identified variants at or close to four novel loci associated with HbA_{1c} in this meta-analysis. *TMEM79* is a transmembrane protein that is highly expressed in liver, erythrocytes, and adipose tissue (The European Bioinformatics Institute, Expression Atlas database, http://www.ebi.ac.uk/gxa). The index SNP rs6684514 was also associated with mean corpuscular hemoglobin concentration in a GWAS conducted in Japanese (32). Although this suggests that this SNP may exert its effects in HbA_{1c} through its effects on erythrocyte biology, we cannot exclude the possibility that this variant could alter glucose regulation. *TMEM79* is downregulated by a high-fat diet in adipose tissue in mice (33). *TMEM79* is also differentially methylated in adipose tissue in twins discordant for T2D. However, this association did not survive multiple testing (34). The index SNP at the *TMEM79* locus rs6684514 was a common missense variant (Val147Met). However, this was predicted as benign by sorting intolerant from tolerant (SIFT) (35) and polymorphism phenotyping (PolyPhen) (36). We also looked for other functional variants in the nearby genes. Three missense SNPs were found to be in high LD (r^2 = 0.97; MAF = 0.22) with rs6684514. They were rs10908495 and rs10908496 at the *C1orf85* locus and rs2230194 at the *CCT3* locus. However, their *P* values were all less than 3.7×10^{-9} in stage 1. Haplotype analysis showed that there were only two common haplotypes (frequency >0.01) in Chinese cohorts of SCES and SP2. They were A-C-A-A and G-T-G-G for rs6684514-rs10908495-rs10908496-rs2230194, while G-T-G-G carried all the alleles associated with elevated HbA_{1c}. Hence, we are unable to discriminate the genuine causal variant among these four missense variants. rs2230194 did not exist in the combined reference panel of HapMap CEU, JPT+CHB, and YRI and hence was absent in Malay cohort of SiMES.

The index SNP rs9399137 is located in the intergenic region between *HBS1L* and *MYB* genes and resides in a LD block (*HMIP 2*), which contains SNPs associated with various hematology traits (5–8,22,37,38), including HbA_2 (39) and HbF (40). Another SNP in the same LD block was associated with mean corpuscular hemoglobin, erythrocyte count, and other hematology traits significantly in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium (6). The index SNPs reported by CHARGE were in moderate/high LD with rs9399137 (r^2 ranges from 0.45 to 1.00 in HapMap release 22 CEU panel). This observation was also supported by recent GWAS in African Americans (9,41). A study has shown a distal regulatory element may exist here to control the expression of *MYB* gene (42), which plays an important role in erythrocyte formation and differentiation.

MYO9B codes for a single-head myosin isoform that was found to be expressed primarily in peripheral blood white cells (43). The product of the gene is important for actin remodeling of epithelial enterocytes, hence it has been associated with several inflammatory bowel diseases, such as Crohn disease and celiac disease, in which the impairment of the enterocyte function explained part of the disease syndrome (44). *MYO9B* function could have an impact on both glycemic and nonglycemic determinants of HbA_{1c}. Celiac disease can result in anemia (45) due to various nutrient deficiencies. This can alter red cell turnover and thus HbA_{1c} level. In addition, celiac disease is also associated with lower BMI and a lower prevalence of T2D (46). Actually, the T allele of our index SNP rs11667918 was associated with lower BMI in another meta-analysis conducted by AGEN (*P* value = 0.03) (47), as well as lower HbA_{1c} level in our study. Furthermore, intestinal permeability may have a role in the pathogenesis of type 1 diabetes (T1D) (48). *MYO9B*'s expression is altered by a hydrolyzed casein diet, which is also able to prevent diabetes in the DP-BB mouse, a mouse model of T1D (49). Variants at this locus showed an association with T1D in a Spanish population (50). However, this finding was not replicated in Dutch and British populations (51). Another SNP rs2279008 at *MYO9B* was associated with human height (52), which was weakly associated with T2D (53), but this SNP is not in LD with our index SNP rs11667918 in both CEU and JPT+CHB panels.

Although, the association of the index SNP of *CYBA* rs9933309 in our replication cohort was not statistically significant, this SNP was only genotyped in a small number of individuals as part of the replication cohorts due to limitations of resources. However, the association after meta-analysis did reach GWAS significance, and the same SNP showed an association with HbA_{1c} in the Meta-Analyses of Glucose and Insulin-Related Traits Consortium

(MAGIC) meta-analysis (3), with an effect direction consistent with that found in our study and *P* value of 1.2×10^{-4} (Supplementary Table 7). For this reason, we believe that this variant is very likely to represent a true positive finding in our study. *CYBA* encodes p22phox, a subunit of NADPH oxidase. NADPH oxidase produces superoxide after catalyzing the reaction from NADPH to $NADPH^+$. In phagocytes, these superoxides are released to kill bacteria and fungi. Since p22phox is the primary component of this microbicidal system, mutations in *CYBA* gene were reported to cause immunodeficiency diseases, such as the recessive chronic granulomatous disease (54). Of relevance to glucose metabolism, higher NADPH oxidase activity, together with higher levels of its subunits, including p22phox, was found to be induced by high environmental glucose in a T2D rat model (55). This increased oxidative stress has been observed in both animal (56) and human (57) β-cell models. In fact, the β-cell is particularly susceptible to the effects of oxidative stress because of the relatively lower expression level of superoxide dismutase, which is protective against oxidation damage, as compared with other tissues (58). It has been argued that high-glucose-induced superoxide was the major cause of β-cell dysfunction and death (59). Thus variants at the *CYBA* locus may result in oxidative stress in the β-cell, leading to glucose intolerance. However, this variant did not show any association with glycemic traits or T2D in any of the populations examined (Supplementary Table 3). Instead, it showed an association with mean cell hemoglobin concentration and mean corpuscular volume in Japanese population, as well as hemoglobin level in a population of European ancestry (Supplementary Table 8 and 9).

It is noteworthy that six of the nine variants that showed genome-wide significant associations with HbA_{1c} showed no evidence of association with glucose traits or T2D. This included three out of four of the variants at novel loci identified in our study (*TMEM79*, *HBS1L/MYB*, and *MYO9B*). Several of these showed strong associations with red-cell-associated parameters (*TMEM79*, *HBS1/MYB*, *ANK1*, and *CYBA*). Another (FNK3) is thought to influence HbA_{1c} by impacting the deglycation of proteins (14) rather than through an impact on glucose metabolism. Approximately 3.3% of the populations studied with fasting plasma glucose <7.0 mmol/L would have been diagnosed as having diabetes based on HbA_{1c} ≥6.5%. Over one-third of this nonconcordance could be explained by these six variants. These findings may have implications for the diagnosis of diabetes using HbA_{1c} in East Asians. It may be that the threshold for diagnosing diabetes based on HbA_{1c} should be slightly higher in these populations. However, an evaluation of appropriate thresholds for the diagnosis of diabetes really requires a careful examination of the relationship between HbA_{1c} and microvascular complications associated with diabetes in these populations. Most of the studies included in this analysis did not have data on microvascular complications, and this is outside the scope of this study. These variants may also contribute to the observation that in individuals with diabetes, Southeast and East Asians exhibit HbA_{1c} that is ~0.2–0.5% higher compared with Caucasians with similar mean blood glucose level (60).

In conclusion, common genetic variants associated with HbA_{1c} levels in populations of European ancestry have similar effects on HbA_{1c} levels in East Asians. We identified four novel associated genomic loci in our East Asian populations. Existing data point to an effect that is mediated by the effect of these variants on nonglycemic factors that affect the erythrocyte for at least three of these variants. However, we cannot conclusively exclude the possibility that they may have an impact on glucose regulation. These findings may have implication on the use of HbA_{1c} to diagnose diabetes in East Asian populations.

Acknowledgments. The authors thank the subjects for their participation and collaborators for their efforts in this study. The complete acknowledgments and the list of collaborators can be found in the Supplementary Data.

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

Author Contributions. P.C., J.-Y.L., and M.O.G. analyzed and interpreted the data and drafted or revised the manuscript. F.T., H.L., Y.T., Y.J.K., M.J.G., C.-H.C., L.-C.C., Q.Q., M.D.G., T.M., and Y.-Y.T. analyzed and interpreted the data. J.-Y.W., J.Lian., J.Lo., E.V., C.-C.K., J.Liu, X.Y., Y.W., L.L., T.L.Y., A.C.T., C.-Y.C., W.-H.P., M.I., K.O., Y.K., H.U., Y.O., M.K., A.T., T.T., F.J.A.v.R., S.K.G., T.L.A., Y.-T.C., T.A., and T.-Y.W. performed data acquisition. M.A.P., D.O.S., J.Liao, K.Y., J.C., C.H., and J.-Y.H. analyzed and interpreted the data and performed data acquisition. J.Le. drafted or revised the manuscript. R.M.v.D., Y.F., C.-K.H., W.-P.K., Q.C., Y.G., C.A.H., W.H.-H.S., J.I.R., Y.-D.I.C., J.Y., T.L.A., and L.Q. performed data acquisition and drafted or revised the manuscript. W.Z., X.-O.S., B.-J.K., and E-S.T. performed the conception and design of the work and data acquisition and drafted or revised the manuscript. R.T., T.H., and X.L. performed the conception and design of the work. J.H. analyzed and interpreted data and performed the conception and design of the work. N.K. analyzed and interpreted the data, performed data acquisition, and drafted or revised the manuscript. E-S.T. is the guarantor of this work and, as such, had 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.

References

1. Paré 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

2. 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

3. Soranzo N, Sanna S, Wheeler E, et al. 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

4. 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

5. Ferreira MA, Hottenga JJ, Warrington NM, et al. Sequence variants in three loci influence monocyte counts and erythrocyte volume. Am J Hum Genet 2009;85:745–749

6. Ganesh SK, Zakai NA, van Rooij FJ, et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. Nat Genet 2009;41:1191–1198
7. Kullo IJ, Ding K, Jouni H, Smith CY, Chute CG. A genome-wide association study of red blood cell traits using the electronic medical record. PLoS ONE 2010;5
8. Soranzo N, Spector TD, Mangino M, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat Genet 2009;41:1182–1190
9. Li J, Glessner JT, Zhang H, et al. GWAS of blood cell traits identifies novel associated loci and epistatic interactions in Caucasian and African-American children. Hum Mol Genet 2013;22:1457–1464
10. van der Harst P, Zhang W, Mateo Leach I, et al. Seventy-five genetic loci influencing the human red blood cell. Nature 2012;492:369–375
11. Zoller H, Theurl I, Koch RO, McKie AT, Vogel W, Weiss G. Duodenal cytochrome b and hephaestin expression in patients with iron deficiency and hemochromatosis. Gastroenterology 2003;125:746–754
12. Guillem F, Lawson S, Kannengiesser C, Westerman M, Beaumont C, Grandchamp B. Two nonsense mutations in the TMPRSS6 gene in a patient with microcytic anemia and iron deficiency. Blood 2008;112:2089–2091
13. Bianchi M, Magnani M. Hexokinase mutations that produce nonspherocytic hemolytic anemia. Blood Cells Mol Dis 1995;21:2–8
14. Delpierre G, Collard F, Fortpied J, Van Schaftingen E. Fructosamine 3-kinase is involved in an intracellular deglycation pathway in human erythrocytes. Biochem J 2002;365:801–808
15. American Diabetes Association. Standards of medical care in diabetes—2010. Diabetes Care 2010;33(Suppl. 1):S11–S61
16. World Health Organization. Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus. Diabetes Res Clin Pract 2011;93:299–309
17. Soranzo N. Genetic determinants of variability in glycated hemoglobin (HbA(1c)) in humans: review of recent progress and prospects for use in diabetes care. Curr Diab Rep 2011;11:562–569
18. Kooner JS, Saleheen D, Sim X, et al.; DIAGRAM; MuTHER. Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat Genet 2011;43:984–989
19. 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
20. Fucharoen S, Winichagoon P. Haemoglobinopathies in southeast Asia. Indian J Med Res 2011;134:498–506
21. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707–713
22. Gieger C, Radhakrishnan A, Cvejic A, et al. New gene functions in megakaryopoiesis and platelet formation. Nature 2011;480:201–208
23. Frazer KA, Ballinger DG, Cox DR, et al.; International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–861
24. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316:1341–1345
25. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39:906–913
26. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet 2009;84:210–223
27. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–2191
28. 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 [published correction appears in Nat Genet 2010; 42:464]. Nat Genet 2010;42:105–116
29. 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
30. Saxena R, Hivert MF, Langenberg C, et al.; GIANT consortium; MAGIC investigators. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat Genet 2010;42:142–148
31. Kim YJ, Go MJ, Hu C, et al.; MAGIC consortium. Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. Nat Genet 2011;43:990–995
32. Kamatani Y, Matsuda K, Okada Y, et al. Genome-wide association study of hematological and biochemical traits in a Japanese population. Nat Genet 2010; 42:210–215
33. Hageman RS, Wagener A, Hantschel C, Svenson KL, Churchill GA, Brockmann GA. High-fat diet leads to tissue-specific changes reflecting risk factors for diseases in DBA/2J mice. Physiol Genomics 2010;42:55–66
34. Ribel-Madsen R, Fraga MF, Jacobsen S, et al. Genome-wide analysis of DNA methylation differences in muscle and fat from monozygotic twins discordant for type 2 diabetes. PLoS ONE 2012;7:e51302
35. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4:1073–1081
36. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–249
37. Okada Y, Hirota T, Kamatani Y, et al. Identification of nine novel loci associated with white blood cell subtypes in a Japanese population. PLoS Genet 2011;7:e1002067
38. Nuinoon M, Makarasara W, Mushiroda T, et al. A genome-wide association identified the common genetic variants influence disease severity in beta0-thalassemia/hemoglobin E. Hum Genet 2010;127:303–314
39. Menzel S, Garner C, Rooks H, Spector TD, Thein SL. HbA2 levels in normal adults are influenced by two distinct genetic mechanisms. Br J Haematol 2013; 160:101–105
40. Thein SL, Menzel S, Lathrop M, Garner C. Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. Hum Mol Genet 2009; 18(R2):R216–R223
41. Chen Z, Tang H, Qayyum R, et al.; BioBank Japan Project; CHARGE Consortium. Genome-wide association analysis of red blood cell traits in African Americans: the COGENT Network. Hum Mol Genet 2013;22:2529–2538
42. Wahlberg K, Jiang J, Rooks H, et al. The HBS1L-MYB intergenic interval associated with elevated HbF levels shows characteristics of a distal regulatory region in erythroid cells. Blood 2009;114:1254–1262
43. Wirth JA, Jensen KA, Post PL, Bement WM, Mooseker MS. Human myosin-IXb, an unconventional myosin with a chimerin-like rho/rac GTPase-activating protein domain in its tail. J Cell Sci 1996;109:653–661
44. Wolters VM, Xu W, Zhao X, et al. Replication of genetic variation in the MYO9B gene in Crohn's disease. Hum Immunol 2011;72:592–597
45. Baydoun A, Maakaron JE, Halawi H, Abou Rahal J, Taher AT. Hematological manifestations of celiac disease. Scand J Gastroenterol 2012;47:1401–1411
46. Kabbani TA, Kelly CP, Betensky RA, et al. Patients with celiac disease have a lower prevalence of non-insulin-dependent diabetes mellitus and metabolic syndrome. Gastroenterology 2013;144:912–917
47. Wen W, Cho YS, Zheng W, et al.; Genetic Investigation of ANthropometric Traits (GIANT) Consortium. Meta-analysis identifies common variants associated with body mass index in east Asians. Nat Genet 2012;44:307–311
48. Visser J, Rozing J, Sapone A, Lammers K, Fasano A. Tight junctions, intestinal permeability, and autoimmunity: celiac disease and type 1 diabetes paradigms. Ann N Y Acad Sci 2009;1165:195–205

49. Visser JT, Lammers K, Hoogendijk A, et al. Restoration of impaired intestinal barrier function by the hydrolysed casein diet contributes to the prevention of type 1 diabetes in the diabetes-prone BioBreeding rat. Diabetologia 2010;53:2621–2628
50. Santiago JL, Martínez A, Núñez C, et al. Association of MYO9B haplotype with type 1 diabetes. Hum Immunol 2008;69:112–115
51. Persengiev S, Koeleman BP, Downes K, et al. Association analysis of myosin IXB and type 1 diabetes. Hum Immunol 2010;71:598–601
52. Lango Allen H, Estrada K, Lettre G, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 2010;467:832–838
53. Lawlor DA, Ebrahim S, Davey Smith G. The association between components of adult height and Type II diabetes and insulin resistance: British Women's Heart and Health Study. Diabetologia 2002;45:1097–1106
54. Köker MY, van Leeuwen K, de Boer M, et al. Six different CYBA mutations including three novel mutations in ten families from Turkey, resulting in autosomal recessive chronic granulomatous disease. Eur J Clin Invest 2009;39:311–319
55. Kim YK, Lee MS, Son SM, et al. Vascular NADH oxidase is involved in impaired endothelium-dependent vasodilation in OLETF rats, a model of type 2 diabetes. Diabetes 2002;51:522–527
56. Ihara Y, Toyokuni S, Uchida K, et al. Hyperglycemia causes oxidative stress in pancreatic beta-cells of GK rats, a model of type 2 diabetes. Diabetes 1999;48:927–932
57. Sakuraba H, Mizukami H, Yagihashi N, Wada R, Hanyu C, Yagihashi S. Reduced beta-cell mass and expression of oxidative stress-related DNA damage in the islet of Japanese Type II diabetic patients. Diabetologia 2002;45:85–96
58. Grankvist K, Marklund SL, Täljedal IB. CuZn-superoxide dismutase, Mn-superoxide dismutase, catalase and glutathione peroxidase in pancreatic islets and other tissues in the mouse. Biochem J 1981;199:393–398
59. Inoguchi T, Nawata H. NAD(P)H oxidase activation: a potential target mechanism for diabetic vascular complications, progressive beta-cell dysfunction and metabolic syndrome. Curr Drug Targets 2005;6:495–501
60. Wolffenbuttel BH, Herman WH, Gross JL, Dharmalingam M, Jiang HH, Hardin DS. Ethnic differences in glycemic markers in patients with type 2 diabetes. Diabetes Care 2013;36:2931–2936