BMJ Open Diabetes Research & Care

Association of *G6PD* variants with hemoglobin A1c and impact on diabetes diagnosis in East Asian individuals

Aaron Leong,[1,2] Victor Jun Yu Lim,[3] Chaolong Wang,[4,5] Jin-Fang Chai,[3] Rajkumar Dorajoo,[5] Chew-Kiat Heng,[6,7] Rob M van Dam,[3] Woon-Puay Koh,[8] Jian-Min Yuan,[9,10] Jost B Jonas,[11,12] Ya Xing Wang,[12] Wen-Bin Wei,[13] Jianjun Liu,[5,14] Dermot F Reilly,[15] Tien-Yin Wong,[16,17] Ching-Yu Cheng,[16,17] Xueling Sim [3]

To cite: Leong A, Lim VJY, Wang C, *et al*. Association of *G6PD* variants with hemoglobin A1c and impact on diabetes diagnosis in East Asian individuals. *BMJ Open Diab Res Care* 2020;**8**:e001091. doi:10.1136/bmjdrc-2019-001091

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjdrc-2019-001091).

AL and VJYL contributed equally.

Received 2 December 2019
Revised 20 January 2020
Accepted 14 February 2020

For numbered affiliations see end of article.

Correspondence to
Dr Xueling Sim; ephsx@nus.edu.sg and Dr Aaron Leong; asleong@partners.org

ABSTRACT

Objective Hemoglobin A1c (HbA1c) accuracy is important for diabetes diagnosis and estimation of overall glycemia. The *G6PD*-Asahi variant which causes glucose-6-phosphate dehydrogenase (G6PD) deficiency has been shown to lower HbA1c independently of glycemia in African ancestry populations. As different *G6PD* variants occur in Asian ancestry, we sought to identify Asian-specific *G6PD* variants associated with HbA1c.

Research design and methods In eight Asian population-based cohorts, we performed imputation on the X chromosome using the 1000 Genomes reference panel and tested for association with HbA1c (10 005 East Asians and 2051 South Asians). Results were meta-analyzed across studies. We compared the proportion of individuals classified as having diabetes/pre-diabetes by fasting glucose ≥100 mg/dL or HbA1c ≥5.7% units among carriers and non-carriers of HbA1c-associated variants.

Results The strongest association was a missense variant (*G6PD*-Canton, rs72554665, minor allele frequency=2.2%, effect in men=−0.76% unit, 95% CI −0.88 to −0.64, $p=1.25\times10^{-27}$, n=2844). Conditional analyses identified a secondary distinct signal, missense variant (*G6PD*-Kaiping, rs72554664, minor allele frequency=1.6%, effect in men=−1.12 % unit, 95% CI −1.32 to −0.92, $p=3.12\times10^{-15}$, $p_{conditional\_Canton}=7.57\times10^{-11}$). Adjusting for glucose did not attenuate their effects. The proportion of individuals with fasting glucose ≥100 mg/dL did not differ by carrier status of *G6PD*-Canton (p=0.21). Whereas the proportion of individuals with HbA1c ≥5.7% units was lower in carriers (5%) compared with non-carriers of *G6PD*-Canton (30%, p=0.03).

Conclusions We identified two *G6PD* variants in East Asian men associated with non-glycemic lowering of HbA1c. Carriers of these variants are more likely to be underdiagnosed for diabetes or pre-diabetes than non-carriers if screened by HbA1c without confirmation by direct glucose measurements.

Significance of this study

What is already known about this subject?

- Hemoglobin A1c (HbA1c) accuracy is important for diabetes classification and estimation of overall glycemia.
- A recent genome-wide association study meta-analysis of HbA1c identified an African-specific, missense variant in *G6PD* (Asahi rs1050828) that lowered HbA1c in African-Americans independently of glycemia. Carriers of this variant were likely to be underdiagnosed for diabetes if HbA1c were used as the sole diagnostic criterion.
- *G6PD* variants are diverse globally and occur quite commonly in Asia; yet, the effects of Asian-specific *G6PD* variants on HbA1c remain unclear.

What are the new findings?

- Through association analysis of the X chromosome with HbA1c in 12 056 Asians, we identified associations of two distinct low-frequency variants that lowered HbA1c independently of glycemia, *G6PD*-Canton (rs72554665, effect in men=−0.76% unit, 95% CI −0.88 to −0.64) and *G6PD*-Kaiping (rs72554664, effect in men=−1.12% unit, 95% CI −1.32 to −0.92).
- Carriers of these Asian-specific *G6PD* variants are more likely to be underdiagnosed for diabetes or pre-diabetes than non-carriers if they are screened by HbA1c without confirmation by direct glucose measurements.

How might these results change the focus of research or clinical practice?

- The use of both fasting glucose and HbA1c in combination for diabetes diagnosis is paramount when screening ethnically diverse populations.

INTRODUCTION

Hemoglobin A1c (HbA1c), a biomarker that reflects average glycemia over the previous 3 months,[1] is routinely used to estimate glycemic control in patients with diabetes, and more recently, diagnose diabetes.[2] However, HbA1c can also be influenced by non-glycemic factors, including genetic variants that affect erythrocyte turnover.[3 4] A number of genome-wide association study (GWAS) meta-analyses

have identified African-specific variants in *G6PD* that are associated with HbA1c in African-American and Hispanic individuals in the USA.[3 5 6] One of these is a missense variant (*G6PD*-Asahi, rs1050828, minor allele frequency (MAF)=11% in African-Americans) that lowers HbA1c by 0.81% (95% CI −0.96 to −0.66, −8.9 mmol/mol, 95% CI −10.6 to −7.2, $p=8.23\times10^{-135}$),[3] and causes glucose-6-phosphate dehydrogenase (G6PD) deficiency, an X linked disease characterized by hemolysis in response to infections, and certain foods or drugs.[7]

G6PD deficiency is widespread across malaria-endemic countries with a global prevalence of 4.9%.[8] While sub-Saharan African and the Arabian Peninsula have higher prevalence estimates of the disease (up to 32.5%) compared with Central and Southeast Asia (less than 20%), the majority of G6PD-deficient individuals reside in Asia.[9] Furthermore, the genetic diversity of *G6PD* variants varies globally, exhibiting high heterogeneity in Asia-Pacific regions.[10] While several *G6PD* variants have been reported to occur in Asia, their effects on HbA1c and implications on the clinical utility of HbA1c remain unclear. Notably, many Asian countries have adopted the standard threshold of HbA1c ≥6.5% (48 mmol/mol) to diagnose diabetes.[2 11]

We hypothesized that *G6PD* variants occurring in Asian ancestry individuals affect the diagnostic accuracy of HbA1c. To our knowledge, previous GWAS to identify HbA1c-associated genetic variants included relatively smaller samples of Asian ancestry, and excluded analysis of the X chromosome due to its non-diploid nature and lack of analytical methods, limiting the discovery of Asian-specific genetic variants. Here, to discover Asian-specific *G6PD* variants, we performed imputation on the chromosome X (ChrX) with the 1000 Genomes reference panel and tested the association of variants with HbA1c in eight Asian studies.

RESEARCH DESIGN AND METHODS

Cohort studies included in meta-analysis

We included genome-wide array data from the following eight studies: Multi-Ethnic Cohort (MEC), Singapore Prospective Study Program (SP2), Living Biobank, Singapore Malay Eye Study (SiMES), the Singapore Chinese Eye Study (SCES), Singapore Indian Eye Study (SINDI), Singapore Chinese Health Study (SCHS-MI), and Beijing Eye Study (BES, online supplementary table 1 and online supplementary appendix).

Glycemic trait measurements and exclusion criteria

We excluded individuals with self-reported physician-diagnosed diabetes, use of diabetes medication, or undiagnosed diabetes (fasting glucose (FG) ≥126 mg/dL (7 mmol/L) or random glucose (RG) ≥200 mg/dL (11.1 mmol/L) or HbA1c ≥6.5% (48 mmol/mol) where available). FG was measured after an overnight 8–12 hours' fast using enzymatic methods (ADVIA 2400; Bayer Diagnostics and Siemens Healthcare Diagnostics). HbA1c was reported in National Glycohemoglobin Standardization Program (NGSP) percent.

ChrX array genotype calling, quality control and imputation

We performed quality control (QC) on ChrX genotypes within each array and only considered the samples that passed autosomal QC. We performed reclustering and genotype calling on the intensity files using GenomeStudio V.2.0. Variants on pseudoautosomal region (PAR) were called from clusters generated by females and males. Those on non-PAR were called from female-only generated clusters. We mapped variants to the hg19 reference genome, forward stand, and excluded those that were unmapped, had a call rate <0.95, or a Hardy-Weinberg equilibrium p value $<10^{-6}$. We further excluded samples with >5% missingness on ChrX or chromosome Y or if the reported sex did not match the genetic sex.

A second round of variant QC was performed to ensure that physical positions were matched to the 1000 Genomes Phase 3 (1000G) reference panel,[12] alleles were on the forward strand, palindromic variants with MAF >0.4 were excluded, and allele frequencies of Chinese and Indians were within 0.2 from 1000G East Asians (EAS) and South Asians (SAS), respectively, and Malays were within 0.3 from 1000G EAS (online supplementary table 2). The final set of ChrX variants was prephased using SHAPEITv2.[13] Heterozygous haploid variants in males were set as missing and reimputed. The phased genotypes were then imputed using Michigan Imputation Server[14] with the full 1000G Phase 3v5 as the reference panel.

Accuracy of ChrX data imputation

To maximize sample set by study, we prioritized samples genotyped on the Illumina arrays. We evaluated the accuracy of imputation by checking the concordance in a subset of samples with both imputed genotypes and exome sequencing: SINDI (599 men and 518 women), SP2 (841 men and 1055 women) within the Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) consortium,[15 16] and the Living Biobank (1228 men and 1220 women).[17] The remaining exomes not previously genotyped in T2D-GENES (Chinese: 156 men and 76 women; Indians: 229 men and 336 women) were used for association with HbA1c. Generally, men had higher concordance rate (>99%) than women (>97%) for all overlapping variants (online supplementary table 3). The imputed calls for the two most common *G6PD* mutations in Singapore,[18] *G6PD*-Canton and *G6PD*-Kaiping, were highly concordant with the exome calls, with >98.6% and 100% concordance for Canton and Kaiping, respectively, in Living Biobank. In SP2, concordance with *G6PD*-Canton was >98.1% across the Illumina 1M and 610 arrays.

Association tests and meta-analysis

We tested imputed ChrX variants for association with HbA1c by study, in men and women separately. We regressed trait values with age, age^2 and the first three principal components to obtain residuals. We then

regressed inverse-normalized residuals against the variants and incorporated empirical kinship matrices using RVTESTS.[19] We repeated all regression analyses with the untransformed trait values to obtain effect size change.

To combine association results from sex-stratified analyses, we first excluded variants that were poorly imputed (imputation Rsq <0.3), had a minor allele count (AC) <5 or association test SE >10 within sex-stratified analyses. We assumed that non-PAR variants in men were homozygous diploid when combining with women in an additive model. Using METAL, we combined p values from inverse-normalized trait values and weighted them by sample size.[20] We combined effect sizes from untransformed traits using inverse-variance weights. Meta-analyses were performed within each ancestry (EAS: BES, Living Biobank, SCES, SCHS-MI, SiMES, SP2, MEC; SAS: SINDI) in variants that were present in two or more studies. To control for residual population structure, genomic control (GC) was applied at the study level and meta-analysis.

We used p value $<5\times10^{-8}$ in the sex-stratified or sex-combined meta-analyses to declare genome-wide significance. Association results were visualized using regional LocusZoom plots.[21] QC and analysis workflows of the association and meta-analysis can be found in online supplementary figure 1. We used dbNSFP[22] and the Variant Effect Predictor[23] to compile function predictions, gene annotations (eg, coding, regulatory, upstream) and consequences (eg, non-sense, missense, intronic) to annotate the variants.

To determine whether associations were independent of glycemia, we performed association testing with (1) FG/RG and (2) HbA1c adjusted for FG/RG where available, using both inverse-normalized traits and raw traits.

Implications of HbA1c-associated variants on undiagnosed diabetes and pre-diabetes classification

To investigate the impact of HbA1c-associated variants on undiagnosed diabetes and pre-diabetes classification, we first excluded individuals with a prior diagnosis of diabetes or medication use, and then compared the proportion of individuals classified as having pre-diabetes or undiagnosed diabetes by FG across the genotype categories at HbA1c-associated variants. Of the eight studies, only three (SP2, MEC, Living Biobank) measured both HbA1c and FG (n=2566). To classify impaired FG or undiagnosed diabetes, we used the American Diabetes Association (ADA) biochemical criteria of HbA1c ≥5.7% (39 mmol/mol) or FG ≥100 mg/dL (5.6 mmol/L),[2] respectively. In a sensitivity analysis, we used the WHO criteria of HbA1c ≥6.0% (42 mmol/mol) or FG ≥110 mg/dL (6.1 mmol/L).[11]

RESULTS

Summary of association results on ChrX

We performed a meta-analysis of ChrX association with HbA1c on 10 005 EAS and 2051 SAS from eight studies. The GC inflation factors ranged from 0.965 to 1.069 in men and from 0.894 to 1.327 in women (likely due to low sample size in one of the studies). After applying study-level GC correction, GC inflation factors were less than 1.0 in sex-stratified and sex-combined meta-analyses. Thirty-six variants, all located within 2.2 Mb around *G6PD,* reached genome-wide significance in EAS men. None reached genome-wide significance in EAS women (online supplementary table 4 and online supplementary figure 2). In EAS sex-combined meta-analysis, 35 of these 36 variants remained significant. In SAS, none reached genome-wide significance.

Strongest association with HbA1c at *G6PD*-Canton in men

The lead variant in men, rs72554665, was a low-frequency missense variant, *G6PD*-Canton (MAF=2.2%, $p=1.25\times10^{-27}$) (table 1). This variant lowered HbA1c by −0.76% (95% CI −0.88 to −0.64), 8.3 mmol/mol (95% CI −9.6 to −7.0), and explained 6.9% of the variance in men. In women, *G6PD*-Canton showed suggestive association (MAF=1.8%, $p=1.40\times10^{-7}$); each copy of the minor allele lowered HbA1c by −0.39% (−4.3 mmol/mol) and accounted for 1.0% of the variance. The lead variant in women was an intergenic variant, rs148112010 (MAF=1.6%, $p=5.36\times10^{-8}$, table 1) which was in moderate linkage disequilibrium (LD) with *G6PD*-Canton ($r^2=0.52$ in 1000G EAS).

In the sex-combined meta-analysis, the lead variant was at ChrX:153734352 (MAF=2.0%, $p=1.71\times10^{-30}$), with evidence of heterogeneity between men and women ($p_{het}=6.7\times10^{-4}$) (table 1). This variant was in complete LD ($r^2=1.00$) with *G6PD*-Canton. Conditioning on *G6PD*-Canton abolished its association ($p_{conditional\_Canton}=0.86$, table 1). Direction of effect and strength of association of *G6PD*-Canton were consistent across studies in both men and women (online supplementary figure 3 and online supplementary table 5).

Conditional analysis on *G6PD*-Canton identified a second distinct signal: *G6PD*-Kaiping

To identify distinct signals at *G6PD*, we conditioned on the strongest signal—*G6PD*-Canton. At a Bonferroni corrected threshold of 1.40×10^{-5} computed from 3576 variants in the 2.5 Mb region centered on *G6PD*-Canton, 11 variants remained significant in men (online supplementary table 6) and 10 in the sex-combined analyses. In women, none was significant on conditioning on *G6PD*-Canton.

The lead secondary signal in men and the sex-combined analysis was rs190054725, an intronic variant on *MPP1* and located 279 kb from *G6PD*-Canton ($r^2=0.0001$, $MAF_{men}=0.9\%$, $effect_{conditional\_Canton}=-1.33\%$, 95% CI −1.57 to −1.09, −14.5 mmol/mol, 95% CI −17.2 to −11.9, $p=7.29\times10^{-13}$, $p_{conditional\_Canton}=1.07\times10^{-14}$, online supplementary figure 4). Joint conditional analyses on *G6PD*-Canton and rs190054725 abolished all signals ($p_{conditional}>1.40\times10^{-5}$). Of the remaining 10 variants that remained significant in the conditional analysis, one was a missense *G6PD* variant, rs72554664, *G6PD*-Kaiping ($MAF_{men}=1.6\%$, $effect_{conditional\_Canton}=-1.12\%$, 95% CI −1.35 to −0.88, −12.2 mmol/mol, 95%

Table 1 Lead variants associated with HbA1c in East Asians

						Single-variant association								
Analysis	**Variant name**	**Chr**	**Position**	**EA/NEA**	**Model**	**n**	**Studies, n**	**EAF**	**Effect (%)**	**SE**	**Effect (mmol/mol)**	**SE**	**P value**	**P_{het}**
Men only	rs72554665 (*G6PD*-Canton)	X	153760484	A/C	Men	2844	5	0.022	−0.38	0.03	−4.23	0.29	1.25×10^{-27}	5.13×10^{-04}
					Women	2836	6	0.018	−0.22	0.04	−2.47	0.42	1.40×10^{-07}	
					Sex combined	5680	11	0.020	−0.33	0.02	−3.65	0.24	3.00×10^{-30}	
Women only	rs148112010	X	154409817	G/A	Men	2844	5	0.026	−0.31	0.02	−3.45	0.24	6.19×10^{-27}	4.29×10^{-04}
					Women	3947	8	0.016	−0.18	0.03	−1.95	0.35	5.39×10^{-08}	
					Sex combined	6791	13	0.020	−0.27	0.02	−2.96	0.20	1.24×10^{-28}	
Sex combined	–	X	153734352	A/G	Men	2844	5	0.022	−0.39	0.03	−4.27	0.30	1.35×10^{-27}	6.70×10^{-04}
					Women	2836	6	0.018	−0.23	0.04	−2.52	0.42	9.16×10^{-08}	
					Sex combined	5680	11	0.020	−0.34	0.02	−3.69	0.24	1.71×10^{-30}	
Conditioned on *G6PD*-Canton (rs72554665)														
Men only	rs72554665 (*G6PD*-Canton)	X	153760484	A/C	–	–	–	–	–	–	–	–	–	
					–	–	–	–	–	–	–	–	–	–
					–	–	–	–	–	–	–	–	–	
Women only	rs148112010	X	154409817	G/A	Men	2844	5	0.03	−0.17	0.06	−1.83	0.64	0.730	–
					Women	2836	6	0.02	0.12	0.11	1.37	1.18	0.632	
					Sex combined	5680	11	0.02	−0.10	0.05	−1.11	0.56	0.923	
Sex combined	–	X	153734352	A/G	Men	2844	5	0.02	0.47	0.72	5.19	7.92	0.789	–
					Women	2836	6	0.02	−1.10	0.68	−12.15	7.47	0.994	
					Sex combined	5680	11	0.02	−0.36	0.49	−3.98	5.43	0.856	

Samples included in association analyses are type 2 diabetes free: no self-reported diabetes and not on diabetes medication use and (fasting glucose <126 mg/dL (7mmol/L) or random glucose <200 mg/dL (11.1mmol/L) where available) and HbA1c <6.5% (48 mmol/mol).
Three different lead variants were identified from association analysis in men only (*G6PD*-Canton/rs72554665), women only (rs148112010) and sex combined (ChrX:153734352). Associations of rs148112010 and ChrX:153734352 with HbA1c disappeared upon conditioning on *G6PD*-Canton (rs72554665).
Effect allele frequency denotes sample size weighted allele frequency across all studies. P values are obtained from sample size weighted meta-analysis using derived inverse-normalized residuals of HbA1c (%) after adjustment for age, age^2 and first three principal components. Effect and SE shown are per allele effect for untransformed HbA1c trait values, assuming that males are homozygous diploid in non-pseudoautosomal region. P_{het} refers to test of heterogeneity between men and women.
As hemizygous males were coded as '2' and males with no variant were coded as '0', the difference in HbA1c between hemizygous males and males with no variant is estimated to be two times the beta effect estimate from regression models (ie, −0.38×2=−0.76)%.
Chr, chromosome; EA, effect allele; EAF, effect allele frequency; n, sample size; NEA, non-effect allele.

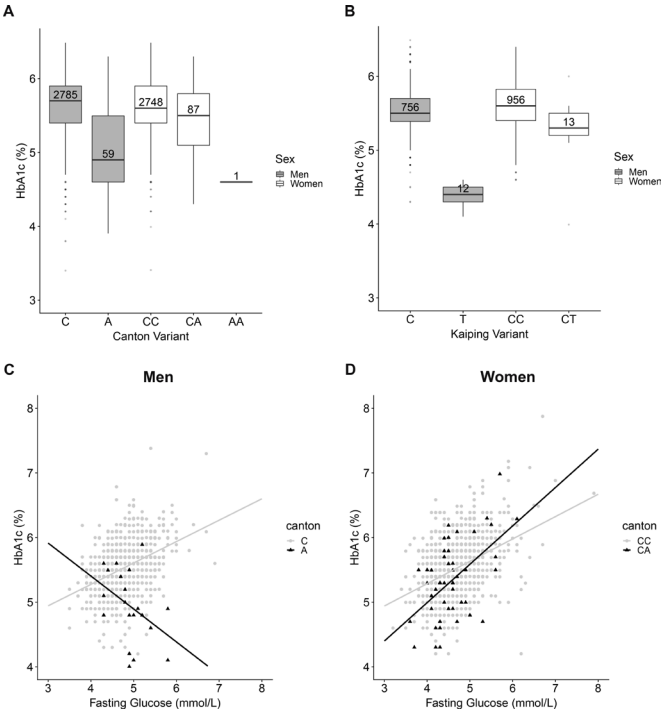

Figure 1 Distribution of HbA1c (%) by sex and *G6PD* variant carrier status, and fasting glucose (mmol/L). (A) Boxplot of HbA1c by sex and *G6PD*-Canton (rs72554665). (B) Boxplot of HbA1c by sex and *G6PD*-Kaiping (rs72554664). (C) Scatterplot of HbA1c and fasting glucose among men by *G6PD*-Canton carrier status. (D) Scatterplot of HbA1c and fasting glucose among women by *G6PD*-Canton carrier status. HbA1c 3%=9 mmol/mol; 4%=20 mmol/mol; 5%=31 mmol/mol; 6%=42 mmol/mol; 7%=53 mmol/mol; 8%=64 mmol/mol. HbA1c, hemoglobin A1c.

CI −14.8 to −9.6, $p=3.12\times10^{-15}$, $p_{conditional\_Canton}=7.57\times10^{-11}$, online supplementary figure 5), which was in moderate LD with rs190054725 ($r^2=0.62$). Neither association at the two missense variants attenuated on reciprocal conditioning ($r^2=0$, D′=1 in 1000G EAS). When conditioned on both *G6PD*-Kaiping and *G6PD*-Canton, none of the variants remained significant (p>0.1; online supplementary table 7).

Other *G6PD* variant associations with HbA1c

We extracted the association statistics from our meta-analyses for all *G6PD* allelic variants reported in Online Mendelian Inheritance in Man (OMIM).[24] Only four of these were polymorphic: *G6PD*-Canton, *G6PD*-Kaiping, *G6PD*-RFLP (rs2230037), and *G6PD*-Viangchan (rs137852327; online supplementary table 8). *G6PD* variants identified through previous GWAS were either monomorphic in our Asian individuals[3 5 6] or not associated with HbA1c (p>0.02).[25] *G6PD*-Viangchan, a missense variant reported in Cambodians, Thais, Laotians, and Malaysia's Malays,[26 27] showed suggestive association with HbA1c in SiMES Malay men (MAF=0.9%, $effect_{men}=-0.64\%$, 95% CI −0.92 to −0.36, −7.0 mmol/mol, 95% CI −10.1 to −3.9, $p=7.21\times10^{-5}$). Due to a low imputation info score, *G6PD*-Viangchan could not be confirmed in Living Biobank, and a meta-analysis with exome sequence data did not reach genome-wide significance (MAF=0.8%, $effect_{men}=-0.62\%$, 95% CI −0.85 to −0.39, −6.8 mmol/mol, 95% CI −9.3 to −4.3, $p=3.44\times10^{-6}$).

No association with glucose measurements at *G6PD* locus

Both *G6PD*-Canton and *G6PD*-Kaiping variants were not associated with FG/RG (p>0.04; online supplementary table 9). Effect estimates on HbA1c did not attenuate on adjusting for FG/RG (online supplementary table 10), suggesting that these genetic effects were glucose independent.

Relationship between HbA1c and FG by *G6PD*-Canton genotypes

Compared with women, the difference in mean HbA1c by *G6PD*-Canton genotypes was more pronounced in men (figure 1A,B and online supplementary table 11). We generated scatterplots of HbA1c on FG. HbA1c increased with FG in male non-carriers (Pearson's r=0.37) but not in carriers (Pearson's r=−0.33, online supplementary tables 12 and 13, figure 1C,D). To determine whether the genetic effect varied by FG, we performed interaction analyses within men and women separately, regressing HbA1c on FG, *G6PD*-Canton genotype, and the interaction term, FG**G6PD*-Canton. Genetic effects differed by sex ($p=3.32\times10^{-5}$), indicating that the HbA1c-lowering effect of *G6PD*-Canton increased with higher FG. The relationship between HbA1c and FG did not differ by genotype in women (p=0.29). Due to low allele count (AC=13), similar analyses were not performed for *G6PD*-Kaiping.

Proportion of pre-diabetes/undiagnosed diabetes by *G6PD*-Canton genotype classes

In 2566 individuals with no prior diagnosis of diabetes or medication use, and both HbA1c and FG measured, 738 men (69.0%) and 1035 women (69.1%) were classified as having normoglycemia (FG <100 mg/dL (5.6 mmol/L) and HbA1c <5.7% (39 mmol/mol). The remaining had pre-diabetes/undiagnosed diabetes defined by FG ≥100 mg/dL (5.6 mmol/L) or HbA1c ≥5.7% (39 mmol/mol), ADA criteria). In men, the proportion of FG-defined pre-diabetes/undiagnosed diabetes was similar between carriers and non-carriers (p=0.21) but the proportion of HbA1c-defined diabetes/undiagnosed diabetes was lower in carriers (307 non-carriers (30%) and 1 carrier (5%), p=0.03; figure 2A,B and online supplementary table 14). In women, there were no significant differences (p>0.1) (figure 2C,D and online supplementary table 14). When using WHO criteria for pre-diabetes/diabetes, none of the male carriers of *G6PD*-Canton were classified as having pre-diabetes/diabetes.

DISCUSSION

While there have been extensive discovery efforts on the genetics of HbA1c in European ancestry, recent

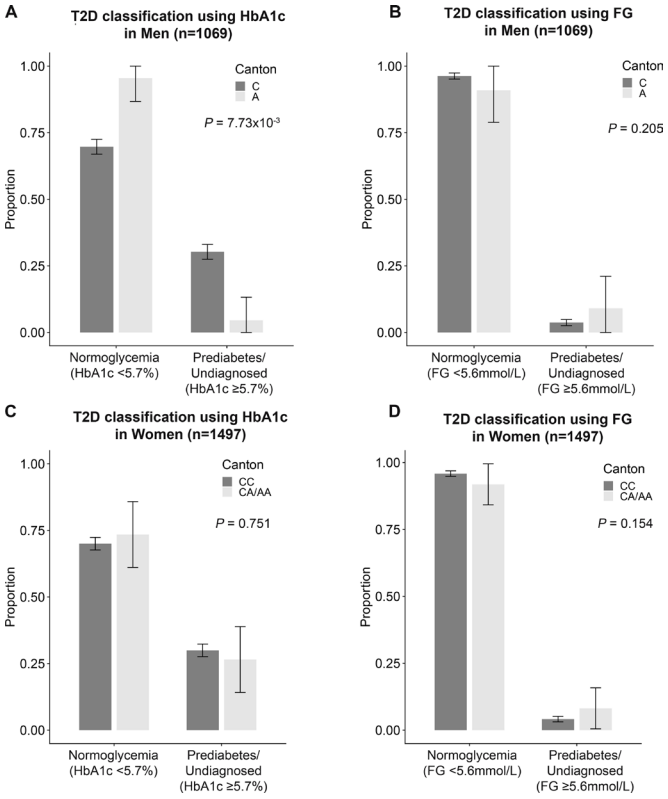

Figure 2 Proportions of individuals by pre-diabetes/diabetes status using American Diabetes Association (ADA) guidelines for fasting glucose ≥100 mg/dL (5.6 mmol/L) and HbA1c ≥5.7% (39 mmol/mol) in men and women separately. (A) HbA1c thresholds in men. (B) Fasting glucose threshold in men. (C) HbA1c thresholds in women. (D) Fasting glucose threshold in women. P values are from Fisher's exact test comparing normoglycemia with pre-diabetes/undiagnosed diabetes. FG, fasting glucose; HbA1c, hemoglobin A1c; T2D, type 2 diabetes.

studies assessing the impact of genetics on the diagnostic accuracy of HbA1c in Asians are limited. Here, we undertook a large-scale association analysis of variants on the X chromosome in 10 005 EAS and 2051 SAS from eight Asian studies to identify Asian-specific *G6PD* variants associated with HbA1c. We identified associations of two distinct low-frequency *G6PD* variants, *G6PD*-Canton and *G6PD*-Kaiping, which lowered HbA1c independently of glycemia. Additionally, we showed that diabetes or pre-diabetes would be underdiagnosed in carriers of these *G6PD* variants if screened only by HbA1c without confirmation by direct glucose measurements. These findings highlight the importance of genetic discovery efforts to uncover ancestry-specific variants with direct clinical implications on diagnostic tests which are routinely performed around the world.

G6PD deficiency confers a relative resistance against *Plasmodium falciparum* malaria.[28] This balancing selection pressure leads to the accumulation of *G6PD* alleles in malaria-endemic countries. Howes *et al* used geostatistical modeling to map the prevalence of G6PD deficiency in malaria-endemic countries and showed that G6PD deficiency is spatially heterogenous across these countries.[10] While occurring globally, most people with G6PD deficiency reside in Asia.[9] *G6PD*-Canton (MAF=2.2%) and *G6PD*-Kaiping (MAF=1.6%) have been classified by WHO as class II variants that are associated with severe acute hemolysis due to low enzymatic activity (<10%).[29]

By contrast, the common African-specific *G6PD* variant (rs1050828, *G6PD*-Asahi, MAF=11%) has been classified by WHO as a class III variant with partial enzymatic activity (10%–60% enzymatic activity) that is generally only associated with intermittent hemolysis.[29] Among these three variants, *G6PD*-Kaiping has the lowest frequency and the largest HbA1c-lowering effect. In Singapore, a country with three major Asian ethnicities (Chinese, Malay and Indian), we identified three distinct low-frequency *G6PD* variants that differed in their effects on HbA1c: *G6PD*-Canton in Chinese, *G6PD*-Kaiping in Chinese,[18] *G6PD*-Viangchan in Malays,[26] and none in Indians. These three variants are specific to Asians and monomorphic in other ancestries.

Prevailing diagnostic thresholds for HbA1c are conservative relative to FG.[2] Our findings suggest that the diagnostic sensitivity of HbA1c for diabetes or pre-diabetes is likely poorer in carriers of these *G6PD* variants than non-carriers, supporting the use of FG over HbA1c for diabetes screening of individuals with known G6PD deficiency. We recognize that not every country with G6PD deficiency has universal screening for the disease. In countries without universal screening, asymptomatic carriers are unlikely aware of their carrier status. Moreover, despite HbA1c having several advantages compared with FG (eg, measurement in the non-fasting state, capturing of overall glycemia, and lower biological variability), NGSP-certified HbA1c testing may not be readily accessible in some parts of the world and is generally more costly more than glucose measurements. Thus, the use of both FG and HbA1c in combination for diabetes diagnosis is paramount in individuals with unknown carrier status, especially when screening ethnically diverse populations that include people with African, Arabian or Asian ancestry.

We observed a slightly lower than expected effect in females (sex heterogeneity p_{Canton}=0.00051, $p_{Kaiping}$=0.071). In an additive model, we would have expected the difference in HbA1c between hemizygous males and those without the variant to be double the difference between heterozygous females and those without the variant due to random X inactivation. Heterogeneity of genetic effects by sex could be due to X linked dosage compensation or skewed inactivation in heterozygous females.[30] It is also possible that females with one normal copy of the gene maintain adequate enzymatic activity and do not have any or only have subclinical disease, in keeping with an X linked recessive genetic disorder. Alternatively, other non-glycemic factors (eg, iron status) could have contributed to heterogeneity of effects by sex.

Our study has several strengths. This study represents one of the largest genetic discovery efforts of HbA1c for the X chromosome in multiethnic Asian ancestry.

Singapore is a relatively young multiethnic country with immigrants from China, India, and the Malaysia and Indonesia archipelagos. While our study samples are predominantly Chinese, and may not represent the full complement of genetic effects on HbA1c in the geographically diverse Asian populations, the association of these low-frequency *G6PD* variants with HbA1c helped inform the unique genetic architecture of HbA1c in individuals from EAS and SAS ancestries. These *G6PD* variants likely cause clinically meaningful reductions in HbA1c, highlighting the potential impact of genetics on the diagnostic classification of diabetes by HbA1c.

We acknowledge limitations. We did not distinguish type 1 and type 2 diabetes, though HbA1c is used to estimate glycemia in both types of diabetes. As continuous glucose monitoring data were not available, we were unable to estimate the concordance between measured HbA1c and HbA1c estimated from continuous glucose monitoring average glucose.[31 32] Nevertheless, continuous glucose monitoring is not routinely used to detect pre-diabetes or diabetes, and so comparing the diagnostic classification by HbA1c to FG is more aligned with current screening practices. As we did not have measures of G6PD enzymatic activity or other blood traits in our study, we were unable to formally test for mediating mechanisms underlying the relationship between *G6PD*-Canton and *G6PD*-Kaiping with HbA1c. Nevertheless, both of these *G6PD* variants have been previously classified by the WHO as class II variants that cause severe G6PD deficiency (<10% residual enzymatic levels and intermittent hemolysis), in keeping with shortened erythrocytic life span as the mechanism involved in lowering HbA1c.[9 10 29 33] While we were able to demonstrate the effect of *G6PD*-Canton on reclassification of diabetes and pre-diabetes status, larger samples are needed to perform a similar analysis for *G6PD*-Kaiping. As this investigation focused solely on the X chromosome, we recognize that variation in the rest of the genome was not examined and sequence data would have strengthened the discovery of HbA1c-related variants in the low allele frequency spectrum, especially in these populations. The limited exome sequence data in smaller sample set were either included in the meta-analysis as independent samples or used to verify the ChrX imputation that we had undertaken using array genotypes. However, previously reported HbA1c-associated loci in Asian ancestry, alone or in combination, have effect sizes several times smaller than *G6PD* variants, and likely play a much smaller role in the overall contribution of genetics to underdiagnosis of diabetes by HbA1c.

In addition to diagnostic testing, HbA1c is routinely used in the management of diabetes to assess glycemic control, disease progression, prediction of diabetes-related complications, and response to treatment. We posit that HbA1c remains a reasonable biomarker to assess change in glycemia over time for both carriers and non-carriers. However, in a treat-to-A1C-target paradigm, physicians may be falsely reassured by low HbA1c values and undertreat or delay treatment escalation in carriers. These carriers may have higher overall glycemia and higher rates of diabetes-related complications compared with non-carriers. We note that the HbA1c-lowering effect of *G6PD*-Canton variant seemed to increase with higher FG within the non-diabetic range in men, implying that the HbA1c-lowering effect of variant could be greater in the diabetic range. Considering the small sample size, this association needs to be interpreted with caution. Future work could focus on replicating this finding in independent samples with both sexes. Further research is also required to determine whether genetic effects of *G6PD* variants on HbA1c are larger in people with diabetes, and whether rates of diabetes-related complications, delayed diagnoses or undertreatment differ by genotype in real-world clinical settings.

In this study, we identified low-frequency Asian-specific *G6PD* variants associated with non-glycemic reductions in HbA1c. Carriers of these *G6PD* variants could have delayed diagnosis or underdetection of diabetes. Given the genetic diversity of *G6PD* variants globally, and the growing use of HbA1c as a diagnostic test worldwide, individual *G6PD* genotype could be considered in precision medicine approaches to diabetes screening in ethnically diverse populations, even if the proportion of affected individuals is small. Both HbA1c and FG ought to be used in combination for diabetes diagnosis in carriers of these *G6PD* variants to avoid potential disparities in care.

Author affiliations
[1]Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA
[2]Harvard Medical School, Boston, Massachusetts, USA
[3]Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[4]Department of Epidemiology and Biostatistics, Key Laboratory for Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[5]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore
[6]Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[7]Khoo Teck Puat–National University Children's Medical Institute, National University Health System, Singapore
[8]Health Services and Systems Research, Duke NUS Medical School, Singapore
[9]Division of Cancer Control and Population Sciences, UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[10]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[11]Department of Ophthalmology, Medical Faculty Mannheim, University of Heidelberg, Heidelberg, Baden-Württemberg, Germany
[12]Ophthalmology and Visual Sciences Key Laboratory, Beijing Institute of Ophthalmology, Beijing Tongren Hospital, Capital Medical University, Beijing, China
[13]Beijing Key Laboratory of Intraocular Tumor Diagnosis and Treatment, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China
[14]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[15]Genetics, Merck Sharp and Dohme IA, Kenilworth, New Jersey, USA
[16]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore
[17]Ophthalmology and Visual Sciences Academic Clinical Program (Eye ACP), Duke–NUS Medical School, Singapore

Acknowledgements The authors thank all the investigators, staff members, and study participants for their contributions to all the participating studies. The authors are grateful to the NUH Tissue Repository for tissue archival services and the Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore for genotyping (SCES, SiMES, SINDI, MEC, SP2, Living Biobank and SCHS-MI).

Contributors AL, VJYL, and XS conceptualized, planned and conducted the study, interpreted the results, and wrote the manuscript. VJYL, CW, RD, and JFC performed the data analyses. CKH, RMvD, WPK, JMY, JBJ, YXW, WBW, JL, DFR, TYW, and CYC were involved in the study design, sample collection and data generation (phenotype, genotype). All authors reviewed and approved the manuscript for submission. XS and AL had full access to all the data in the study and had final responsibility for the decision to submit for publication. XS and AL are the guarantors of the work.

Funding The Diabetic Cohort (DC), Multi-Ethnic Cohort (MEC) and Living Biobank was supported by grants from the Ministry of Health, Singapore, the National University of Singapore and the National University Health System, Singapore. In addition, genotyping for Living Biobank was funded by the Agency for Science, Technology and Research, Singapore, and Merck Sharp & Dohme Corp., Whitehouse Station, NJ, USA. The Singapore Prospective Study Program (SP2) were supported by the individual research grant and clinician scientist award schemes from the National Medical Research Council (NMRC) and the Biomedical Research Council (BMRC) of Singapore. The Singapore Chinese Eye Study (SCES) and the Singapore Malay Eye Study (SiMES) are supported by the National Medical Research Council (NMRC), Singapore (grants 0796/2003, 1176/2008, 1149/2008, STaR/0003/2008, 1249/2010, CG/SERI/2010, CIRG/1371/2013, and CIRG/1417/2015), and Biomedical Research Council (BMRC), Singapore (08/1/35/19/550 and 09/1/35/19/616). The Singapore Chinese Health Study was supported by the National Institutes of Health, USA (R01 CA144034 and UM1 CA182876), the nested case–control study of myocardial infarction by the Singapore National Medical Research Council (NMRC 1270/2010) and genotyping by the HUJ-CREATE Programme of the National Research Foundation, Singapore (Project No 370062002). The Beijing Eye Study (BES) was funded by the National Natural Science Foundation of China (NSFC, No 81170890 and 81570835).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by the National University of Singapore International Review Board (08-013) and conducted according to the Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Summary-level statistics will be publicly available at http://blog.nus.edu.sg/agen/summary-statistics/hba1c-chrx-2020/

ORCID iD

Xueling Sim http://orcid.org/0000-0002-1233-7642

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;134:317–26.
2 American Diabetes Association. 2. Classification and Diagnosis of Diabetes: *Standards of Medical Care in Diabetes-2018*. *Diabetes Care* 2018;41:S13–27.
3 Wheeler E, Leong A, Liu C-T, *et al*. 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.
4 Leong A, Wheeler E. Genetics of HbA1c: a case study in clinical translation. *Curr Opin Genet Dev* 2018;50:79–85.
5 Moon J-Y, Louie TL, Jain D, *et al*. 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;42:1784–91.
6 Sarnowski C, Leong A, Raffield LM, *et al*. Impact of rare and common genetic variants on diabetes diagnosis by hemoglobin A1c in Multi-Ancestry cohorts: the Trans-Omics for precision medicine program. *Am J Hum Genet* 2019;105:706–18.
7 Motulsky AG, Stamatoyannopoulos G. Clinical implications of glucose-6-phosphate dehydrogenase deficiency. *Ann Intern Med* 1966;65:1329–34.
8 Nkhoma ET, Poole C, Vannappagari V, *et al*. The global prevalence of glucose-6-phosphate dehydrogenase deficiency: a systematic review and meta-analysis. *Blood Cells Mol Dis* 2009;42:267–78.
9 Howes RE, Piel FB, Patil AP, *et al*. G6Pd deficiency prevalence and estimates of affected populations in malaria endemic countries: a geostatistical model-based MAP. *PLoS Med* 2012;9:e1001339.
10 Howes RE, Dewi M, Piel FB, *et al*. Spatial distribution of G6PD deficiency variants across malaria-endemic regions. *Malar J* 2013;12:418.
11 WHO. *Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus: abbreviated report of a who consultation*, 2011.
12 Auton A, Brooks LD, Durbin RM, *et al*. A global reference for human genetic variation. *Nature* 2015;526:68–74.
13 Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2013;10:5–6.
14 Das S, Forer L, Schönherr S, *et al*. Next-Generation genotype imputation service and methods. *Nat Genet* 2016;48:1284–7.
15 Fuchsberger C, Flannick J, Teslovich TM, *et al*. The genetic architecture of type 2 diabetes. *Nature* 2016;536:41–7.
16 Flannick J, Mercader JM, Fuchsberger C, *et al*. Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature* 2019;570:71–6.
17 Dou J, Sun B, Sim X, *et al*. Estimation of kinship coefficient in structured and admixed populations using sparse sequencing data. *PLoS Genet* 2017;13:e1007021.
18 Quak SH, Saha N, Tay JS. Glucose-6-Phosphate dehydrogenase deficiency in Singapore. *Ann Acad Med Singapore* 1996;25:45–8.
19 Zhan X, Hu Y, Li B, *et al*. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics* 2016;32:1423–6.
20 Willer CJ, Li Y, Abecasis GR. Metal: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26:2190–1.
21 Pruim RJ, Welch RP, Sanna S, *et al*. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336–7.
22 Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Hum Mutat* 2013;34:E2393–402.
23 McLaren W, Gil L, Hunt SE, *et al*. The Ensembl variant effect predictor. *Genome Biol* 2016;17:122.
24 Hamosh A, Scott AF, Amberger JS, *et al*. Online Mendelian inheritance in man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* 2005;33:D514–7.
25 De Vita G, Alcalay M, Sampietro M, *et al*. Two point mutations are responsible for G6PD polymorphism in Sardinia. *Am J Hum Genet* 1989;44:233–40.
26 Ainoon O, Yu YH, Amir Muhriz AL, *et al*. Glucose-6-Phosphate dehydrogenase (G6PD) variants in Malaysian Malays. *Hum Mutat* 2003;21:101.
27 Bancone G, Menard D, Khim N, *et al*. Molecular characterization and mapping of glucose-6-phosphate dehydrogenase (G6PD) mutations in the greater Mekong subregion. *Malar J* 2019;18:20.
28 Luzzatto L, Arese P, Favism AP. Favism and glucose-6-phosphate dehydrogenase deficiency. *N Engl J Med* 2018;378:60–71.
29 WHO. Glucose-6-Phosphate dehydrogenase deficiency. who Working group. *Bull World Health Organ* 1989;67:601–11.
30 Sidorenko J, Kassam I, Kemper KE, *et al*. The effect of X-linked dosage compensation on complex trait variation. *Nat Commun* 2019;10:3009.
31 Malka R, Nathan DM, Higgins JM. Mechanistic modeling of hemoglobin glycation and red blood cell kinetics enables personalized diabetes monitoring. *Sci Transl Med* 2016;8:359ra130.
32 Bergenstal RM, Beck RW, Close KL, *et al*. Glucose management indicator (GMI): a new term for estimating A1c from continuous glucose monitoring. *Diabetes Care* 2018;41:2275–80.
33 Yoshida A, Beutler E, Motulsky AG. Human glucose-6-phosphate dehydrogenase variants. *Bull World Health Organ* 1971;45:243–53.