Common variants at the *GCK*, *GCKR*, *G6PC2–ABCB11* and *MTNR1B* loci are associated with fasting glucose in two Asian populations

F. Takeuchi · T. Katsuya · S. Chakrewarthy · K. Yamamoto · A. Fujioka ·
M. Serizawa · T. Fujisawa · E. Nakashima · K. Ohnaka · H. Ikegami · T. Sugiyama ·
T. Nabika · A. Kasturiratne · S. Yamaguchi · S. Kono · R. Takayanagi · Y. Yamori ·
S. Kobayashi · T. Ogihara · A. de Silva · R. Wickremasinghe · N. Kato

Received: 15 July 2009 / Accepted: 6 October 2009 / Published online: 25 November 2009
© Springer-Verlag 2009

Abstract

Aims/hypothesis To test fasting glucose association at four loci recently identified or verified by genome-wide association (GWA) studies of European populations, we performed a replication study in two Asian populations.

Methods We genotyped five common variants previously reported in Europeans: rs1799884 (*GCK*), rs780094 (*GCKR*), rs560887 (*G6PC2–ABCB11*) and both rs1387153 and rs10830963 (*MTNR1B*) in the general Japanese ($n=4,813$) and Sri Lankan ($n=2,319$) populations. To identify novel

Electronic supplementary material The online version of this article (doi:10.1007/s00125-009-1595-1) contains supplementary material, which is available to authorised users.

F. Takeuchi · M. Serizawa · N. Kato (✉)
Department of Gene Diagnostics and Therapeutics,
Research Institute, International Medical Center of Japan,
1-21-1 Toyama, Shinjuku-ku,
Tokyo 162-8655, Japan
e-mail: nokato@ri.imcj.go.jp

T. Katsuya
Department of Clinical Gene Therapy,
Osaka University Graduate School of Medicine,
Suita, Japan

S. Chakrewarthy · A. Kasturiratne · A. de Silva ·
R. Wickremasinghe
Faculty of Medicine, University of Kelaniya,
Kelaniya, Sri Lanka

K. Yamamoto
Department of Molecular Genetics,
Medical Institute of Bioregulation, Kyushu University,
Fukuoka, Japan

A. Fujioka
Amagasaki Health Medical Foundation,
Amagasaki, Japan

T. Fujisawa
Department of Geriatric Medicine,
Osaka University Graduate School of Medicine,
Suita, Japan

E. Nakashima
Division of Endocrinology and Diabetes,
Department of Internal Medicine,
Nagoya University Graduate School of Medicine,
Nagoya, Japan

E. Nakashima
Department of Metabolism and Endocrine Internal Medicine,
Chubu Rosai Hospital,
Nagoya, Japan

K. Ohnaka
Department of Geriatric Medicine,
Graduate School of Medical Sciences, Kyushu University,
Fukuoka, Japan

H. Ikegami
Department of Endocrinology, Metabolism and Diabetes,
Kinki University School of Medicine,
Osaka-Sayama, Japan

T. Sugiyama
Institute for Adult Diseases, Asahi Life Foundation,
Tokyo, Japan

T. Nabika
Department of Functional Pathology,
Shimane University School of Medicine,
Izumo, Japan

variants, we further examined genetic associations near each locus by using GWA scan data on 776 non-diabetic Japanese samples.

Results Fasting glucose association was replicated for the five single nucleotide polymorphisms (SNPs) at $p < 0.05$ (one-tailed test) in South Asians (Sri Lankan) as well as in East Asians (Japanese). In fine-mapping by GWA scan data, we identified in the *G6PC2-ABCB11* region a novel SNP, rs3755157, with significant association in Japanese ($p = 2.6 \times 10^{-8}$) and Sri Lankan ($p = 0.001$) populations. The strength of association was more prominent at rs3755157 than that of the original SNP rs560887, with allelic heterogeneity detected between the SNPs. On analysing the cumulative effect of associated SNPs, we found the per-allele gradients ($\beta = 0.055$ and 0.069 mmol/l in Japanese and Sri Lankans, respectively) to be almost equivalent to those reported in Europeans.

Conclusions/interpretation Fasting glucose association at four tested loci was proven to be replicable across ethnic groups. Despite this overall consistency, ethnic diversity in the pattern and strength of linkage disequilibrium certainly exists and can help to appreciably reduce potential causal variants after GWA studies.

Keywords Asians · Association study · Ethnicity · Fasting plasma glucose · Polymorphisms

Abbreviations

CEU Utah residents with northern and western European ancestry from the Centre d'Etude du Polymorphisme Humain collection

S. Yamaguchi
Internal Medicine III, Shimane University School of Medicine,
Izumo, Japan

S. Kono
Department of Preventive Medicine,
Graduate School of Medical Sciences, Kyushu University,
Fukuoka, Japan

R. Takayanagi
Department of Medicine and Bioregulatory Science,
Graduate School of Medical Sciences, Kyushu University,
Fukuoka, Japan

Y. Yamori
Mukogawa Women's University,
Institute for World Health Development,
Nishinomiya, Japan

S. Kobayashi
Shimane University Hospital,
Izumo, Japan

T. Ogihara
Osaka General Medical Center,
Osaka, Japan

FPG	Fasting plasma glucose
GWA studies	Genome-wide association studies
JPT	Japanese in Tokyo
LD	Linkage disequilibrium
RAF	Risk allele frequency
SNP	Single nucleotide polymorphism

Introduction

Fasting plasma glucose (FPG) levels are associated with the future risk of type 2 diabetes and cardiovascular diseases [1, 2] and are tightly regulated despite considerable variation in food intake [3]. It has been reported that genetic effects explain 54.8% of the variance of glucose levels in a European population [4]. Recent progress in complex-trait genetics has allowed the identification of loci regulating FPG levels [5–20].

Several loci influencing FPG levels have been identified or verified by genome-wide association (GWA) studies of Europeans; these include glucokinase (*GCK*) [5–7], glucokinase regulatory protein (*GCKR*) [8–10, 13], glucose-6-phosphatase catalytic subunit 2 (*G6PC2*), the ATP-binding cassette, subfamily B (*MDR/TAP*), member 11 (*ABCB11*) [14–17], and melatonin receptor 1B (*MTNR1B*) [16–18].

All the associations were originally identified in populations of European ancestry. While some studies have shown reproducible associations [9, 11, 12, 19, 20], it remains to be further defined to what degree loci discovered in Europeans will show an association in populations of different ancestries. In addition, to localise the variant(s) responsible for an association signal, we need to generate a comprehensive list of potential causal variants in the regions of interest, i.e. to conduct fine-mapping after GWA studies. As discussed elsewhere [21], this fine-mapping will be challenging and genetic information from populations of different ancestries is expected to be useful [21–24].

Apart from assessing the previously identified variants in two Asian populations, Japanese of East Asian ancestry and Sri Lankan of South Asian ancestry, we also explored index single nucleotide polymorphism (SNP) markers, which either tag the SNPs attaining a locus-wise significance level in the GWA scan of Japanese or were previously reported in Europeans. This was done to advance the fine-mapping of the associated loci [5–20].

Methods

Study populations

A replication study of the previously identified variants was performed in the general Japanese and Sri Lankan

populations (Electronic supplementary material [ESM] Table 1, [ESM Study samples for continuous traits](#)), using 5,456 Japanese samples (including 4,813 non-diabetic participants) consecutively enrolled in a population-based setting as described elsewhere [25] and 3,012 Sri Lankan samples (including 2,319 non-diabetic participants) who had participated in the baseline survey of the Ragama Health Study [26] in Sri Lanka. Complementary to this replication study, we organised genetic studies of FPG levels as part of an ongoing GWA scan for cardiometabolic disorders among the Japanese population ([ESM Study samples for continuous traits](#)). We used 776 population-based, non-diabetic Japanese samples for preliminary screening of association with FPG levels. Then, the association signals were examined in the general populations mentioned above. In addition to quantitative trait analysis, type 2 diabetes associations were tested for index SNPs at *G6PC2-ABCB11* in a Japanese case–control study panel comprising 5,629 cases and 6,406 controls as previously reported [27], and in a Sri Lankan case–control study panel ([ESM Study samples for type 2 diabetes case–control studies](#)). All participants from these different studies provided written informed consent and the local Ethics Committees approved the protocols.

Type 2 diabetes was diagnosed according to the WHO criteria as described in [ESM Study samples for type 2 diabetes case–control studies](#).

SNP genotyping and quality control

In the replication study, samples were genotyped using the TaqMan assay (Life Technologies Japan, Tokyo, Japan) for five SNPs from four gene loci previously identified in European-descent populations [5–10, 13–18]. These included *GCK* (rs1799884), *GCKR* (rs780094), *G6PC2-ABCB11* (rs560887) and *MTNR1B* (rs1387153 and rs10830963).

In the GWA scans, genotyping was performed with a bead array (Infinium HumanHap550; Illumina, San Diego, CA, USA) as described elsewhere [27] (ESM Fig. 1, [ESM SNP genotyping](#), [ESM Quality control of the GWA scan data](#)). After the GWA scan, three additional SNPs in the *G6PC2-ABCB11* region, rs483234, rs3755157 and rs853778, were genotyped with the TaqMan assay for follow-up.

Statistical analysis

SNP association analysis SNPs were tested for association with FPG levels by using linear regression analysis in the additive genotype model ([ESM SNP-based association analysis](#)). A *p* value of <0.05 was considered statistically significant. For an association to be considered significant, it had to involve the same risk allele as that reported in

Europeans and was accordingly assessed with a one-tailed test. To assess the proportion of variance for FPG that could be explained by a SNP, we calculated the coefficient of determination R^2 . The per-allele gradients, which correspond to the increase in FPG levels by additional ‘high FPG’ alleles of associated SNPs, were calculated in the linear regression model (including age, sex and BMI as covariates) as previously reported [14, 18] ([ESM Evaluation of cumulative effect of multiple loci on FPG](#)). We used PLINK (<http://pngu.mgh.harvard.edu/~purcell/plink/>), the R software (version 2.8.1; www.r-project.org) and the rmeta package (<http://cran.r-project.org>) for association test and meta-analysis (websites accessed 15 October 2009).

Haplotype analysis In the *G6PC2-ABCB11* region, we selected SNPs attaining a locus-wise significance level ($p < 0.002$ by Bonferroni’s correction for 23 SNPs genotyped in the relevant region) or reported in European studies, inferring the haplotypes using PLINK [28] and PHASE [29] software (http://depts.washington.edu/ventures/UW_Technology/Express_Licenses/PHASEv2.php). We then tested which haplotypes were strongly associated with the trait. In parallel, haplotypes were inferred from the genotype data of the SNPs in HapMap (www.hapmap.org) Utah residents with northern and western European ancestry from the Centre d’Etude du Polymorphisme Humain collection (CEU) and Japanese in Tokyo (JPT) categories using HaploView software (www.broad.mit.edu/mpg/haploview/) [30] and in South Asians from the Human Genome Diversity Panel (<http://hagsc.org/hgdp/files.html>) [31]. Haplotype-tagging SNPs were selected and characterised in the large study panels.

Stepwise regression analysis for testing of independent associations To test the most likely explanation for the signal of association among the index SNPs and their genotyped correlates, we performed stepwise linear regression analysis for FPG levels by forward selection ([ESM Index SNPs showing an independent association](#)). If two SNPs simultaneously included in the model each attained significance ($p < 0.05$), they could have independent associations. Further, when two haplotype classes that are distant in the phylogeny have an opposite effect and are tagged by two SNPs showing independent associations, the haplotype classes are presumably linked to different causative variants, thus implying allelic heterogeneity ([ESM Haplotype explaining index association](#)).

Cross-population filtering of causal variants To appreciably narrow the location of potential causal variants, we closely inspected subsets of SNPs and haplotypes shared by multiple ethnicities. We partitioned all the HapMap SNPs

Fig. 1 Plots of FPG association, and LD and SNP partitioning for the *G6PC2-ABCB11* region in Japanese (a, d, g, j), Sri Lankan (b, e, h, k) and Europeans (c, f, i, l). Association results for Europeans are drawn from the published studies [14, 17]. a–c Bar graphs of all genotyped SNPs that passed the quality control (ESM Table 6) in the Japanese GWA scan (a) and those in the published European GWA scan [14] (c) with $-\log_{10}(p \text{ values})$ for FPG plotted against chromosome position in Mb. Red diamonds, p values for the genotypes of general populations ($n=4,813$ in Japanese, $n=2,319$ in Sri Lankan) and meta-analysis data ($>12,000$ in the ENGAGE consortium [17]). d–f Genomic location of *G6PC2* and *ABCB11* genes with intron and

exon structure (NCBI Build 36) in the relevant populations. g–i WGAViewer (<http://people.genome.duke.edu/~dg48/WGAViewer/whatis.php>) [36] plot of LD (r^2) for all SNPs across the regions for the HapMap populations JPT (g) and CEU (i), and for the Human Genome Diversity Panel, ethnic groups belonging to South Asia (h). j–l SNPs in the LD block were partitioned into four subsets using the extent of LD with lead SNPs and/or haplotype-tagging SNPs, i.e. rs560887, rs483234, rs3755157 and rs853778 (see Methods). Red circles, SNP with $0.8 \leq r^2 \leq 1.0$ to the index SNP; orange circles, SNP with $0.6 \leq r^2 < 0.8$ to the index SNP; green bars, intervals where causative variants are most likely to be located

located in the *G6PC2-ABCB11* region (Fig. 1) so that SNPs in the strongest linkage disequilibrium (LD) with one index SNP (e.g. rs3755157) rather than with other index SNPs (e.g. rs560887, rs483234 and rs853778) were grouped into a bin of rs3755157 correlates. We then narrowed target intervals by investigating a subset (or subsets) of variants that could show a consistent pattern of trait association across different ethnic groups (ESM Narrowing target intervals in fine-mapping).

Results

Association with FPG and metabolic traits at four loci

Significant ($p < 0.05$ by one-tailed test) association was replicated for all five SNPs from four tested loci in the Sri Lankan and Japanese populations (Table 1). Together with the previous reports in Europeans and Chinese [8, 11, 12, 17, 18, 20], we performed meta-analysis of FPG associa-

Table 1 Association of SNPs with fasting plasma glucose level

SNP	Neighbouring gene(s)	Alleles		Japanese panel (<i>n</i> =4,813)			Sri Lankan panel (<i>n</i> =2,319)			Europeans	
		FPG+ ^a	FPG- ^a	Allele frequency ^b	Per-allele effect ^c	<i>p</i> value	Allele frequency ^b	Per-allele effect ^c	<i>p</i> value	Allele frequency ^d	Per-allele effect ^{c,e}
Reported in European studies											
rs780094	<i>GCKR</i>	G	A	0.44	0.032 (0.012, 0.052)	0.002	0.80	0.074 (0.035, 0.113)	2.1×10 ⁻⁴	0.62	0.067 (0.045, 0.090)
rs560887	<i>G6PC2-ABCB11</i>	G	A	0.97	0.103 (0.044, 0.162)	6.4×10 ⁻⁴	0.91	0.050 (-0.005, 0.105)	0.075	0.67	0.064 (0.056, 0.072)
rs1799884	<i>GCK</i>	A	G	0.18	0.075 (0.049, 0.101)	1.1×10 ⁻⁸	0.12	0.076 (0.028, 0.123)	0.002	0.20	0.062 (0.048, 0.076)
rs1387153	<i>MTNR1B</i>	T	C	0.41	0.058 (0.038, 0.078)	9.7×10 ⁻⁹	0.38	0.036 (0.005, 0.068)	0.024	0.28	0.07 (0.05, 0.08)
rs10830963	<i>MTNR1B</i>	G	C	0.42	0.056 (0.036, 0.075)	2.9×10 ⁻⁸	0.45	0.064 (0.033, 0.094)	3.6×10 ⁻⁵	0.30	0.072 (0.062, 0.082)
Tested in addition ^f											
rs483234	<i>G6PC2-ABCB11</i>	A	G	0.51	0.046 (0.026, 0.065)	3.9×10 ⁻⁶	0.42	0.043 (0.012, 0.074)	0.007	0.70	–
rs3755157	<i>G6PC2-ABCB11</i>	T	C	0.38	0.057 (0.037, 0.078)	2.6×10 ⁻⁸	0.16	0.069 (0.027, 0.111)	0.001	0.07	–
rs853778	<i>G6PC2-ABCB11</i>	A	G	0.40	0.044 (0.024, 0.064)	1.5×10 ⁻⁵	0.35	0.026 (-0.006, 0.058)	0.117	0.46	–

Type 2 diabetes participants were excluded from the analysis; FPG association of each SNP was tested using linear regression models with adjustment for BMI, age and sex

^a FPG-increasing (+) and decreasing (–)

^b Of FPG-increasing allele

^c Effect (95% CI) (mmol/l)

^d Of FPG-increasing allele in HapMap CEU panel

^e Association results were drawn from the previous studies: rs560887, rs4607517 (in substitution for rs1799884, $r^2 = 1$ in CEU) and rs10830963 [17]; rs1260326 (in substitution for rs780094, $r^2 = 0.93$ in CEU) [8]; rs1387153 [18]

^f In the *G6PC2-ABCB11* region

tions to compare the effect sizes among different ethnic groups (ESM Fig. 2). Among the four loci, significant cross-population heterogeneity was detected for rs1387153 (*MTNR1B*; $p=0.03$). The variance for FPG that was explained by the associated SNPs totalled 2% in both the Japanese and Sri Lankan populations (ESM Table 2). Per-allele gradients in the two Asian populations ($\beta=0.055$ and 0.069 mmol/l in Japanese and Sri Lankans, respectively) were almost equivalent to those reported in populations of European descent ($\beta=0.07$ mmol/l) [18] (ESM Fig. 3, [ESM Evaluation of cumulative effect of multiple loci on FPG](#)).

Besides FPG levels, we analysed the relationship of SNPs with lipid traits (ESM Tables 3 and 4). Notably, rs780094 (*GCKR*) significantly and consistently modulated triacylglycerol levels in both ethnic groups ($p=2.2\times 10^{-10}$ in Japanese, $p=1.4\times 10^{-4}$ in Sri Lankan populations), where glucose-increasing alleles were associated with lower triacylglycerol levels as previously reported [8, 9, 11, 13]. Furthermore, glucose-increasing alleles at rs1799884 (*GCK*) and rs10830963 (*MTNR1B*) were significantly associated with reduced beta cell function (HOMA-B; $p=0.037$ for rs1799884, $p=2.6\times 10^{-4}$ for rs10830963 in the Sri Lankan population), with no appreciable effect on fasting insulin or insulin sensitivity (ESM Table 5).

Refinement of genetic association in the *G6PC2-ABCB11* region

In fine-mapping with Japanese GWA scan data, we identified in the *G6PC2-ABCB11* region a novel associated SNP rs3755157, which was proven to be independent of the

SNPs previously reported by GWA studies in Europeans [14–17].

In our GWA scan, multiple and significant SNPs were found in the *G6PC2-ABCB11* region ($p=0.0004$ to 0.002 ; Fig. 1, ESM Table 6) but not in the other candidate regions (ESM Tables 7–9). We therefore performed a detailed investigation of the *G6PC2-ABCB11* region. With reference to the LD and haplotype data (ESM Tables 10–12), we chose four haplotype-tagging SNPs (rs3755157, rs483234, rs853778 and rs560887) for genotyping the general Japanese population ($n=4,813$), which resulted in concordant evidence of associations between the tests of individual SNPs and those of haplotypes (Tables 1 and 2, [ESM Haplotype explaining index association](#)). The most significantly associated haplotype, class 5 (frequency = 0.35 , $p=2.8\times 10^{-7}$) (Table 2), was almost unequivocally tagged by rs3755157, which showed the strongest association by SNP-based test in the general Japanese population ($p=2.6\times 10^{-8}$; Table 1).

We then performed a stepwise linear regression (for FPG levels) to test whether one of the four haplotype-tagging SNPs was necessary and sufficient to explain the association signal (ESM Tables 13–17, ESM Fig. 4, [ESM Index SNP showing an independent association](#)). The FPG association remained significant ($p<0.05$) when two haplotype-tagging SNPs, rs3755157 and rs560887, were included in the regression model (ESM Table 13). This independent association had gone unnoticed among more significant associations of SNPs that were in strong LD with a leading SNP, rs560887, among Europeans [14, 15]. Thus the presence of a novel SNP, rs3755157, and of allelic heterogeneity (ESM Fig. 5)

Table 2 Fasting glucose association according to haplotypes in the *G6PC2-ABCB11* region

Haplotype class	Tested SNPs ^a				Japanese panel ^b			Sri Lankan panel ^c			Europeans
	rs560887	rs483234	rs3755157	rs853778	Frequency	Effect (mmol/l)	<i>p</i> value	Frequency	Effect (mmol/l)	<i>p</i> value	Frequency ^d
1	G ^a	G	C	G	0.48	−0.047	2.3×10^{-6}	0.56	−0.041	9.6×10^{-3}	0.26
2	G ^a	G	C	A ^a	0	–	–	0.02	−0.029	0.65	0
3	G ^a	A ^a	C	G	0.09	0.004	0.81	0.08	0.035	0.23	0.28
4	G ^a	A ^a	C	A ^a	0.01	0.048	0.31	0.09	0.030	0.28	0.07
5	G ^a	A ^a	T ^a	A ^a	0.35	0.054	2.8×10^{-7}	0.16	0.072	9.8×10^{-4}	0.08
6	G ^a	A ^a	T ^a	G	0.02	0.053	0.12	0	–	–	0
7	A	A ^a	C	A ^a	0.02	−0.098	2.5×10^{-3}	0.08	−0.078	8.1×10^{-3}	0.28

FPG association was tested with adjustment for BMI, age and sex

^a Allele increasing fasting glucose

^b 4,792 complete observations

^c 2,306 complete observations

^d Haplotype frequency estimated in the HapMap CEU panel

Table 3 Association of FPG-altering SNPs with type 2 diabetes in the *G6PC2–ABCB11* region

SNP	rs560887 (FPG-increasing allele: G)	rs483234 (FPG-increasing allele: A)	rs3755157 (FPG-increasing allele: T)
Japanese (JPN) ^a			
Frequency, cases (<i>n</i> =5,629)	0.974	0.508	0.380
Frequency, controls (<i>n</i> =6,406)	0.969	0.490	0.360
OR (95% CI)	1.20 (1.03–1.41)	1.07 (1.02–1.13)	1.09 (1.03–1.15)
<i>p</i> value for trend	0.019	0.0056	0.0017
Sri Lankan (SL) ^b			
Frequency, cases (<i>n</i> =599)	0.917	0.422	0.169
Frequency, controls (<i>n</i> =515)	0.896	0.408	0.159
OR (95% CI)	1.28 (0.95–1.73)	1.06 (0.89–1.26)	1.08 (0.86–1.36)
<i>p</i> value for trend	0.08	0.51	0.52
JPN and SL combined			
OR (95% CI)	1.22 (1.06–1.40)	1.07 (1.02–1.13)	1.09 (1.03–1.15)
European-descent			
OR (95% CI)	0.93 (0.89–0.97) ^c	1.05 (0.99–1.13) ^d	1.00 (0.91–1.09) ^d
<i>p</i> value	0.0017	0.12	0.93

Type 2 diabetes association was tested with the Cochran–Armitage trend test in the case–control analysis

^a SNPs were genotyped in a Japanese case–control study panel independently of the general Japanese population [27]

^b In the Sri Lankan population, 515 controls are part of those used for FPG association analysis, whereas 599 cases were independent participants

^c Results for 18,236 cases and 64,453 controls from a previous study [17]

^d Results for 4,549 cases and 5,579 controls from the DIAGRAM consortium [37]

has become evident in the *G6PC2–ABCB11* region for the first time, as a result of comparing the GWA scan data between European and Japanese populations.

In the *G6PC2–ABCB11* region, *G6PC2* and *ABCB11* are both biologically plausible candidate genes [15, 32, 33]. During fine-mapping, we attempted to partition the LD block into intervals, each containing SNPs strongly correlated with an index SNP, in the hope that correlation coefficients r^2 would reflect phylogenetic closeness once the index SNPs were selected from a reasonably dense set of SNP markers. This partitioning approach helped to prioritise the target interval for fine-mapping, thereby reducing the potential candidate variants to manageable proportions. For the *G6PC2–ABCB11* region, the target intervals were estimated to be 14 kb (in the *ABCB11* gene) for rs3755157 and 14 kb (in the *G6PC2* gene and between the genes) for rs560887 when the LD threshold was set at $r^2 \geq 0.6$ in the HapMap data (Fig. 1, ESM Table 18, ESM Narrowing target intervals in fine-mapping).

Concordance of association for FPG levels and type 2 diabetes

In addition to the quantitative trait analysis of FPG, we performed case–control analysis of type 2 diabetes for three

of four haplotype-tagging SNPs, rs3755157, rs483234 and rs560887, in the *G6PC2–ABCB11* region and confirmed a significant ($p < 0.05$) association in a relatively large study panel comprising 5,629 cases and 6,406 controls. The strongest association was found for rs3755157 (OR 1.09, 95% CI 1.03–1.15, $p = 1.7 \times 10^{-3}$; Table 3), which was in good agreement with the FPG association mentioned above. To confirm the consistency of associations with increases in FPG levels and the risk of type 2 diabetes, we examined the changes in risk allele frequency (RAF) between the diabetes subgroup and non-diabetic participants in quartiles of FPG levels stratified in the general populations for unbiased estimates (ESM Fig. 6, ESM Table 19). In the Japanese and Sri Lankan populations, the RAF in the diabetes subgroup reached the second highest quartile, supporting the concordant association for FPG levels and the risk of type 2 diabetes in the *G6PC2–ABCB11* region.

Discussion

The present study has proven that common variant loci influencing FPG levels are reproducible in two populations of Asian descent, Japanese (East Asians) and Sri Lankan

(South Asians). To our knowledge, this is the first study investigating the genetic associations with FPG and related metabolic traits at four candidate loci, *GCK*, *GCKR*, *G6PC2-ABCB11* and *MTNR1B*, in South Asians, who are known to have high prevalence of type 2 diabetes [34]. The combined impact of associated SNPs is almost equivalent across the ethnic groups despite some cross-population diversity in the effect size of individual loci (ESM Figs 2 and 3, ESM Table 2). Other novel aspects of the present study include a fine-mapping approach using Japanese GWA scan data and consistent associations of FPG and type 2 diabetes in the *G6PC2-ABCB11* region.

According to genome-wide patterns of SNPs examined in the Human Genomic Diversity Panel [31], much of sub-Saharan Africa, Europe, South and Central Asia (including Sri Lanka), and East Asia appear to be homogeneous and individuals from these populations can be distinguished from each other. Although limited in the number of examples, our study has provided evidence supporting the importance of human genetic diversity in complex disease studies. For instance, beside replicating FPG association at four candidate loci in two Asian populations, our data also clarified the genetic architecture of the *G6PC2-ABCB11* region with regard to ethnic diversity. Using the GWA scan data, we found a novel SNP, rs3755157, to be a leading SNP among Japanese and independent of a leading SNP, rs560887, in Europeans (ESM Table 13–15).

As a fine-mapping approach, we listed HapMap SNPs having the strongest r^2 (in the range of $r^2 \geq 0.6$) with each of the index SNPs in the *G6PC2-ABCB11* region (Fig. 1, ESM Table 18). We performed cross-population filtering, which appreciably decreased the number of potential causal variants from 79 to 8 in the *G6PC2-ABCB11* region (ESM [Narrowing target intervals in fine-mapping](#)). The novel SNP rs3755157 and its correlated SNPs are located in the 3'-side (introns) of the *ABCB11* gene. While four different mRNAs, two alternatively spliced variants and two unspliced forms, are known to be transcribed from the *ABCB11*, it is possible that the potential causal variant(s) will influence the selective production of any of the 3'-side mRNA variants or the alteration of mRNA expression. Thus, closer inspection of subsets of SNP haplotypes shared by multiple ethnicities may allow us to appreciably narrow the field of potential causal variants before starting in-depth resequencing and functional follow-up studies, as demonstrated for the *G6PC2-ABCB11* region. During the preparation of our manuscript, replication of the *G6PC2* association was also reported in a Chinese population [19], where four SNPs were selected from the HapMap database so as to tag common variations near the *G6PC2* gene. Although the index SNP (rs560887) originally detected in Europeans was not tested, three (of four) SNPs appeared to

show significant association in the Chinese population, in agreement with our findings in Japanese.

Our data also verified concordance of association for FPG levels and type 2 diabetes risk by using a systematic study design; i.e. unbiased estimates with stratification of general populations plus large-scale case-control studies involving 12,035 Japanese and 1,114 Sri Lankan samples (Table 3, ESM Fig. 6). It has been debated whether the genetic determinants regulating FPG levels in physiological states differ from those increasing type 2 diabetes risk. Some studies report that carriers of glucose-increasing alleles at three loci (*MTNR1B*, *GCK* and *GCKR*) show a higher risk of type 2 diabetes [8, 17, 18], although there is no significant association between *G6PC2-ABCB11* variants and type 2 diabetes in populations of European descent [14, 17, 35]. In this context, our data not only supported the concordant association of *G6PC2-ABCB11* variants for FPG and type 2 diabetes in two Asian populations, but also indicated that genetic determinants regulating FPG levels could, at least in part, differ from those increasing type 2 diabetes risk (ESM Fig. 6, [ESM Consistent association of fasting glucose and type 2 diabetes in the *G6PC2-ABCB11* region](#)). It is likely that genetic susceptibility for FPG levels increases type 2 diabetes risk in the population at large, but that some diabetic patients will develop the disease independently of a predisposition to elevated FPG levels.

In summary, despite the overall reproducibility of FPG association across the populations, ethnic diversity in allele frequencies led to the discovery of allelic heterogeneity in the *G6PC2-ABCB11* region. The diversity in the LD pattern also helped to reduce the probable causative variants in the corresponding region. The prevalence of the phenomena described here in human complex trait genetics is another research area warranting investigation. For applicable cases, the use of ethnic diversity in genetic studies can constitute an efficient approach subsequent to GWA scan.

Acknowledgements This work was supported by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation Organization (NIBIO) and the Manpei Suzuki Diabetes Foundation. Support also came from the Ministry of Education, Culture, Sports, Science and Technology of Japan. We acknowledge the outstanding contributions of the International Medical Center of Japan (IMCJ) employees, who provided technical and infrastructural support for this work. Above all, we thank the patients and study participants who made this work possible and who give it value. We also thank all the people who continuously support the Hospital-based Cohort Study at IMCJ, the Amagasaki Study and the Kyushu University Fukuoka Cohort Study in Japan, and the Ragama Health Study in Sri Lanka. We also thank C. Makibayashi and the many physicians of the Amagasaki Medical Association, as well as M. Makaya, T. Mizoue, H. Janaka de Silva, U. Ranawaka and

other staff at the University of Kelaniya for their help with collection of DNA samples and accompanying clinical information. The DNA samples of type 2 diabetes cases used for this research were partly provided by the Leading Project for Personalized Medicine in the Ministry of Education, Culture, Sports, Science and Technology, Japan. The GWA study conducted by NIBIO GWA Study Group was organised to clarify the pathogenesis of diabetes and associated metabolic disorders as well as cardiovascular complications. The collaborating institutions that constitute the NIBIO GWA Study Group are: International Medical Center of Japan; Kyushu University; Osaka University; Nagoya University; Kinki University; Shimane University; Tohoku University; the Institute for Adult Diseases, Asahi Life Foundation; Chubu Rosai Hospital; Amagasaki Health Medical Foundation; collaborating groups in the Amagasaki Medical Association; and collaborating groups in the Kyushu region.

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

References

1. Tirosh A, Shai I, Tekes-Manova D et al (2005) Normal fasting plasma glucose levels and type 2 diabetes in young men. *N Engl J Med* 353:1454–1462
2. Coutinho M, Gerstein HC, Wang Y, Yusuf S (1999) The relationship between glucose and incident cardiovascular events. A metaregression analysis of published data from 20 studies of 95,783 individuals followed for 12.4 years. *Diabetes Care* 22:233–240
3. Xiang AH, Wang C, Peters RK, Trigo E, Kjos SL, Buchanan TA (2006) Coordinate changes in plasma glucose and pancreatic beta-cell function in Latino women at high risk for type 2 diabetes. *Diabetes* 55:1074–1079
4. Pilia G, Chen WM, Scuteri A et al (2006) Heritability of cardiovascular and personality traits in 6, 148 Sardinians. *PLoS Genet* 2:e132
5. Weedon MN, Frayling TM, Shields B et al (2005) Genetic regulation of birth weight and fasting glucose by a common polymorphism in the islet cell promoter of the glucokinase gene. *Diabetes* 54:576–581
6. Rose CS, Ek J, Urhammer SA et al (2005) A -30G>A polymorphism of the beta-cell-specific glucokinase promoter associates with hyperglycemia in the general population of whites. *Diabetes* 54:3026–3031
7. Weedon MN, Clark VJ, Qian Y et al (2006) A common haplotype of the glucokinase gene alters fasting glucose and birth weight: association in six studies and population-genetics analyses. *Am J Hum Genet* 79:991–1001
8. Vaxillaire M, Cavalcanti-Proenca C, Dechaume A et al (2008) The common P446L polymorphism in GCKR inversely modulates fasting glucose and triglyceride levels and reduces type 2 diabetes risk in the DESIR prospective general French population. *Diabetes* 57:2253–2257
9. Orho-Melander M, Melander O, Guiducci C et al (2008) Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and C-reactive protein but lower fasting glucose concentrations. *Diabetes* 57:3112–3121
10. Sparso T, Andersen G, Nielsen T et al (2008) The GCKR rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinemia, and reduced risk of type 2 diabetes. *Diabetologia* 51:70–75
11. Tam CH, Ma RC, So WY et al (2009) Interaction effect of genetic polymorphisms in glucokinase (GCK) and glucokinase regulatory protein (GCKR) on metabolic traits in healthy Chinese adults and adolescents. *Diabetes* 58:765–769
12. Qi Q, Wu Y, Li H et al (2009) Association of GCKR rs780094, alone or in combination with GCK rs1799884, with type 2 diabetes and related traits in a Han Chinese population. *Diabetologia* 52:834–843
13. Saxena R, Voight BF, Lyssenko V et al (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316:1331–1336
14. Bouatia-Naji N, Rocheleau G, van Lommel L et al (2008) A polymorphism within the G6PC2 gene is associated with fasting plasma glucose levels. *Science* 320:1085–1088
15. Chen WM, Erdos MR, Jackson AU et al (2008) Variations in the G6PC2/ABCB11 genomic region are associated with fasting glucose levels. *J Clin Invest* 118:2620–2628
16. Sabatti C, Service SK, Hartikainen AL et al (2009) Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* 41:35–46
17. Prokopenko I, Langenberg C, Florez JC et al (2009) Variants in MTNR1B influence fasting glucose levels. *Nat Genet* 41:77–81
18. Bouatia-Naji N, Bonnefond A, Cavalcanti-Proenca C et al (2009) A variant near MTNR1B is associated with increased fasting plasma glucose levels and type 2 diabetes risk. *Nat Genet* 41:89–94
19. Hu C, Zhang R, Wang C et al (2009) A genetic variant of G6PC2 is associated with type 2 diabetes and fasting plasma glucose level in the Chinese population. *Diabetologia* 52:451–456
20. Rönn T, Wen J, Yang Z et al (2009) A common variant in MTNR1B, encoding melatonin receptor 1B, is associated with type 2 diabetes and fasting plasma glucose in Han Chinese individuals. *Diabetologia* 52:830–833
21. McCarthy MI, Hirschhorn JN (2008) Genome-wide association studies: potential next steps on a genetic journey. *Hum Mol Genet* 17:R156–R165
22. Saccone NL, Saccone SF, Goate AM et al (2008) In search of causal variants: refining disease association signals using cross-population contrasts. *BMC Genet* 9:58
23. Waters KM, Le Marchand L, Kolonel LN et al (2009) Generalizability of associations from prostate cancer genome-wide association studies in multiple populations. *Cancer Epidemiol Biomarkers Prev* 18:1285–1289
24. Helgason A, Palsson S, Thorleifsson G et al (2007) Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet* 39:218–225
25. Tsuchihashi-Makaya M, Serizawa M, Yanai K et al (2009) Gene-environment interaction regarding alcohol-metabolizing enzymes in the Japanese general population. *Hypertens Res* 32:207–213
26. Dassanayake A, Kasturiratne A, Rajindrajith S et al (2009) Prevalence and risk factors for non-alcoholic fatty liver disease (NAFLD) among adults in an urban Sri Lankan population. *J Gastroenterol Hepatol* 24:1284–1288
27. Takeuchi F, Serizawa M, Yamamoto K et al (2009) Confirmation of multiple risk loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. *Diabetes* 58:1690–1699
28. Purcell S, Neale B, Todd-Brown K et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575
29. Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989
30. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265
31. Li JZ, Absher DM, Tang H et al (2008) Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 319:1100–1104
32. Petrolonis AJ, Yang Q, Tummino PJ et al (2004) Enzymatic characterization of the pancreatic islet-specific glucose-6-phosphatase-related protein (IGRP). *J Biol Chem* 279:13976–13983

33. Lang T, Haberl M, Jung D et al (2006) Genetic variability, haplotype structures, and ethnic diversity of hepatic transporters MDR3 (ABCB4) and bile salt export pump (ABCB11). *Drug Metab Dispos* 34:1582–1599
34. Ramaraj R, Chellappa P (2008) Cardiovascular risk in South Asians. *Postgrad Med J* 84:518–523
35. Reiling E, van't Riet E, Groenewoud MJ et al (2009) Combined effects of single-nucleotide polymorphisms in GCK, GCKR, G6PC2 and MTNR1B on fasting plasma glucose and type 2 diabetes risk. *Diabetologia* 52:1866–1870
36. Ge D, Zhang K, Need AC et al (2008) WGAViewer: software for genomic annotation of whole genome association studies. *Genome Res* 18:640–643
37. Zeggini E, Scott LJ, Saxena R et al (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 40:638–645