Confirmation of Multiple Risk Loci and Genetic Impacts by a Genome-Wide Association Study of Type 2 Diabetes in the Japanese Population

Fumihiko Takeuchi,[1,2] Masakuni Serizawa,[3] Ken Yamamoto,[4] Tomomi Fujisawa,[5] Eitaro Nakashima,[6,7] Keizo Ohnaka,[8] Hiroshi Ikegami,[9] Takao Sugiyama,[10] Tomohiro Katsuya,[5] Makoto Miyagishi,[3] Naoki Nakashima,[11] Hajime Nawata,[12] Jiro Nakamura,[6] Suminori Kono,[13] Ryoichi Takayanagi,[14] and Norihiro Kato[3]

OBJECTIVE—To identify novel type 2 diabetes gene variants and confirm previously identified ones, a three-staged genome-wide association study was performed in the Japanese population.

RESEARCH DESIGN AND METHODS—In the stage 1 scan, we genotyped 519 case and 503 control subjects with 482,625 single nucleotide polymorphism (SNP) markers; in the stage 2 panel comprising 1,110 case subjects and 1,014 control subjects, we assessed 1,456 SNPs ($P < 0.0025$, stage 1); additionally to direct genotyping, 964 healthy control subjects formed the in silico control panel. Along with genome-wide exploration, we aimed to replicate the disease association of 17 SNPs from 16 candidate loci previously identified in Europeans. The associated and/or replicated loci (23 SNPs; $P < 7 \times 10^{-5}$ for genome-wide exploration and $P < 0.05$ for replication) were examined in the stage 3 panel comprising 4,000 case subjects and 12,569 population-based samples, from which 4,889 nondiabetic control subjects were preselected. The 12,569 subjects were used for overall risk assessment in the general population.

RESULTS—Four loci—1 novel with suggestive evidence (*PEPD* on 19q13, $P = 1.4 \times 10^{-5}$) and three previously reported—were identified; the association of *CDKAL1*, *CDKN2A/CDKN2B*, and *KCNQ1* were confirmed ($P < 10^{-19}$). Moreover, significant associations were replicated in five other candidate loci: *TCF7L2*, *IGF2BP2*, *SLC30A8*, *HHEX*, and *KCNJ11*. There was substantial overlap of type 2 diabetes susceptibility genes between the two populations, whereas effect size and explained variance tended to be higher in the Japanese population.

CONCLUSIONS—The strength of association was more prominent in the Japanese population than in Europeans for more than half of the confirmed type 2 diabetes loci. *Diabetes* 58: 1690–1699, 2009

The predisposition to and the course of type 2 diabetes vary according to ethnic group (1–3). In Japan, the incidence of type 2 diabetes has increased recently and is now comparable to that of other countries; this is supposedly attributable to the gradual spread of Western habits, such as consuming a high-fat diet, and the lower insulin secretory capacity of Japanese subjects (4,5). Recent technological developments have allowed the successful identification of gene regions involved in the development of type 2 diabetes in genome-wide association (GWA) studies (6–17). Several susceptibility gene loci identified by GWA studies to date have been used to obtain reproducible evidence of disease association in different populations of European descent and Asians, but not necessarily in African Americans (18–24). A number of GWA studies on type 2 diabetes have been conducted on populations of European descent (6–12). Two GWA scans in the Japanese population simultaneously reported the discovery of type 2 diabetes susceptibility gene (*KCNQ1*) variants in non-European populations; this result was also obtained in Scandinavian samples (25,26). Thus far, the replicated associations for a limited number of candidate genes have broadly indicated the tendency of interethnic similarity. Even though the common (or cosmopolitan) effect of type 2 diabetes risk variants is known, the extent to which the causation of this disease differs or overlaps between populations remains unknown. Here, besides comparing the genetic associations between European-descent and Japanese populations, we aimed to identify new genetic variants using a three-staged GWA study design.

RESEARCH DESIGN AND METHODS

Detailed characteristics of the subjects enrolled in each stage are described in the supplementary information and in supplementary Table S1, which is available in an online appendix at http://diabetes.diabetesjournals.org/cgi/content/full//db08-1494/DC1. Briefly, patients and unaffected control subjects analyzed in stages 1 and 2 were enrolled depending on whether they met certain uniform criteria. Type 2 diabetes was diagnosed according to 1999 World Health Organization criteria. All stage 1 and 2 control subjects (≥55

FIG. 1. Flow chart summarizing the multistage design and study aims. (A high-quality digital representation of this figure is available in the online issue.)

years of age at examination) had normal glucose tolerance. The stage 3 samples comprised 4,000 type 2 diabetes case subjects derived from the Biobank Japan project (http://biobankjp.org/) (27) and 12,569 subjects randomly selected from residents aged 50–74 years in the general population. The 12,569 subjects were used as a population panel; this panel contained 4,889 nondiabetic subjects who met the following conditions: age ≥55 years, A1C ≤5.0%, no previous and/or current treatment for diabetes, and absence of renal failure (serum creatinine <3.0 mg/dl). In stage 3, these 4,889 control subjects were used in a replication study wherein their genotypes were compared with those of 4,000 patients. In addition to the samples genotyped here, to boost the power of the GWA scan, we incorporated genotype frequencies in the general Japanese population ($n = 964$) derived from the Genome Medicine Database of Japan (GeMDBJ; http://gemdbj.nibio.go.jp), which was used as an in silico control panel. A flowchart summarizing the multistage design and study aims is shown in Fig. 1.

Stage 1 genome-wide scan and quality control. Genotyping was performed with the Infinium HumanHap550 BeadArray (Illumina), which interrogated 555,352 SNPs (supplementary information). The average call rate was 96.9% for the case and control subjects. Data cleaning and analysis were performed using PLINK software (28). Samples with a genotype call rate of <90% were excluded, as were outliers with respect to the number of heterozygous SNPs, duplicates or relatives of another sample, or ethnic outliers. We excluded SNPs for the following reasons: *1*) GenTrain genotype quality score <0.53, *2*) genotype call rate <0.95, *3*) genotype call rate <0.99 and minor allele frequency (MAF) <0.05, *4*) significant ($P < 10^{-6}$) deviation from the Hardy-Weinberg equilibrium in the control subjects, or *5*) MAF <0.001 (supplementary Table S2); the remaining 482,625 SNPs were analyzed.

Analysis of stage 1 genotype data. Ethnicity was verified for 1,022 samples (519 case and 503 control subjects) in the stage 1 panel with reference to data from HapMap populations (29) (see supplementary information). Type 2 diabetes association was tested with the Cochran-Armitage trend test in the stage 1 panel and an additional panel of 964 random control subjects. We pooled the genotype counts for combining multiple panels. To detect and correct population stratification and unnoticed differences in data processing

between facilities, the test statistic was adjusted using Eigenstrat software (30) and the genomic-control method (31). The significance level for the first-stage scan was set to $P < 0.0025$; significant SNPs were additionally chosen using Fisher's χ^2 test ($P < 0.0025$) to combine the association results with the P value at the same locus in our previous affected sib-pair scan (32). A total of 1,811 SNPs surpassed the stage 1 threshold, and we removed redundant SNPs that were in mutual strong linkage disequilibrium ($r^2 > 0.9$) before proceeding to stage 2 (see supplementary information and supplementary Fig. S1 and Table S2 for detailed analysis).

Stage 2 genotyping and analysis. Stage 2 genotyping was performed with iPLEX (Sequenom) and GoldenGate (Illumina) assays. Quality control was conducted as described in stage 1, and 1,456 SNPs were successfully genotyped. We calculated P values with the trend test by combining 1,517 nondiabetic control subjects with 964 random control subjects similar to stage 1. The significance level for the second-stage scan was set to $P < 7 \times 10^{-5}$ in the comparison between 1,629 case subjects and 2,481 control subjects (i.e., the stage 1 + 2 panels and the 964 random control subjects). A total of 30 SNPs representing 17 unique loci remained significant.

Replication of previously reported SNPs. Along with genome-wide exploration, type 2 diabetes association was tested in the stage 1 and 2 panels using the HumanHap550 BeadArray, iPLEX assay, or TaqMan method (Applied Biosystems) for 17 SNPs from 16 candidate loci previously identified by GWA studies in populations of European descent (6–17). These included *IGF2BP2* (rs4402960), *PPARG* (rs1801282), *CDKAL1* (rs7754840 and rs7756992), *SLC30A8* (rs13266634), *CDKN2A/CDKN2B* (rs10811661), *HHEX* (rs1111875), *TCF7L2* (rs7903146), *EXT2* (rs3740878), *KCNJ11* (rs5219), *FTO* (rs8050136), *JAZF1* (rs864745), *CDC123-CAMK1D* (rs12779790), *TSPAN8-LGR5* (rs7961581), *THADA* (rs7578597), *ADAMTS9* (rs4607103), and *NOTCH2-ADAM30* (rs10923931). The significant SNPs (trend test, $P < 0.05$) were further analyzed in the stage 3 panel with the TaqMan method. Despite finding significant association for *CDKAL1* and *CDKN2A/CDKN2B* in the stage 1 and 2 panels, we proceeded with rs4712523 instead of rs7754840 and rs7756992 (*CDKAL1*) and with rs2383208 instead of rs10811661 (*CDKN2A/CDKN2B*) in the GWA scans from stage 1 to stage 3; this decision was made considering the

strong linkage disequilibrium between the SNPs in each of the corresponding loci.

Stage 3 genotyping and analysis. The stage 3 design involved the replication of association and the estimation of effect sizes in the GWA scan and/or replication study of previously reported SNPs. For an association to be considered significant in the case-control comparison (4,000 case vs. 4,889 nondiabetic control subjects), it had to involve the same risk allele as that in the previous stages, and it was accordingly assessed with a one-tailed test. For each SNP locus confirmed in stage 3, the association of additional independent SNPs or haplotypes in the locus was also tested (supplementary information). Moreover, to assess the risk of diabetes and pre-diabetes in the general population from the combination of SNPs robustly confirmed both in populations of European descent and in our panel, multiple regression analysis was performed with the logarithm of A1C (log A1C) as a response variable (supplementary information), using the entire 12,569-subject population-based sample.

Meta-analysis of other type 2 diabetes case-control studies in the Japanese population. In addition, for SNPs with robustly confirmed association in populations of European descent, we performed a meta-analysis by combining our stage 1 + 2 (rs1801282, rs7756992, and rs8050136) or stage 1 + 2 + 3 results (the remaining seven SNPs shown in supplementary Figure S2) with those of previous Japanese studies conducted by three other groups (19–21,33–36). According to Woolf's test (37), the heterogeneity among the studies in the Japanese population was insignificant ($P > 0.05$), with the exception of *PPARG* rs1801282 ($P = 0.0012$), for which the observed heterogeneity is supposedly attributable to low allele frequency. Thus, we pooled genotype counts across the studies to form a combined dataset for the Japanese population, and we estimated the effect sizes of individual loci. We used the rmeta package for R software (http://www.r-project.org) for the analysis.

Moreover, to compare the explained variance between the Japanese population and populations of European descent, we calculated the coefficient of determination R^2 for the loci confirmed in our replication study. Here, R^2 is the square of the correlation between the genotypes of an SNP coded by the number of risk alleles (0, 1, and 2) and the disease status (0 and 1) (supplementary information).

RESULTS

GWA scans. Of 482,625 SNPs that passed quality control in stage 1, genotypes were obtained for an average of 99.8% markers for each subject. The subjects were enrolled from regions of Japan with no strong population stratification (38), and although some variance inflation partly attributable to the subtle subpopulation structure was apparent, such confounding influences could be sufficiently removed using Eigenstrat (30) and genomic-control adjustment (31). A total of 1,456 markers were assessed in the stage 2 panel (Fig. 1 and supplementary Fig. S1 and Table S2).

After the second-stage scan, 30 SNPs representing 17 unique loci attained the arbitrarily defined statistical significance ($P < 7 \times 10^{-5}$) (supplementary Table S3). We used one SNP each from these 17 loci in the third-stage scan. Of 17 SNPs, 4 reached the significance threshold of $P < 0.003$ (= 0.05/17) with Bonferroni correction.

The current GWA study showed strong and highly consistent evidence for disease association of SNPs from *CDKAL1*, *CDKN2A/CDKN2B*, and *KCNQ1* (Fig. 2 and Table 1 and supplementary Tables S4 and S5). Although these three loci had already been reported in previous GWA studies (8,11,25,26), here they were identified as part of our genome-wide exploration. *CDKAL1* is among the best-replicated susceptibility loci. Significant association has also been detected in a region on chromosome 9p, near *CDKN2A/CDKN2B*. Moreover, strong association signals were observed in the intron of *KCNQ1* on chromosome 11p15.5, which is in agreement with the results of two previous GWA scans in the Japanese population (25,26).

Stage 2 genotyping provided evidence suggestive of a new association on chromosome 19q13. Several SNPs located in the vicinity of the *PEPD* (peptidase D) gene showed the tendency of replicated association in stages 1 and 2 (supplementary Table S3). Significant association was further replicated in a relatively large case-control study on the stage 3 panel (rs10425678, $P = 0.002$), but it did not attain genome-wide significance (i.e., $P = 1.4 \times 10^{-5}$ for all stages and $P = 2.1 \times 10^{-6}$ when the number of control subjects was increased by adding 964 random control subjects) (supplementary Table S4). Given the modest strength of association ($R^2 = 0.0017$, see below) assumed for this locus, the association still needs to be established.

Replication of previously reported SNPs. Of 16 candidate loci tested for replication in the Japanese population, 7 were found to be associated with type 2 diabetes (Table 1). However, no significant association was observed for SNPs from the remaining nine loci (*FTO*, *PPARG*, *EXT2*, *JAZF1*, *CDC123-CAMK1D*, *TSPAN8-LGR5*, *ADAMTS9*, and *NOTCH2-ADAM30* in the stage 1 + 2 panel and *THADA* in the stage 1 + 2 + 3 panel). Notably, the originally reported SNPs or those in complete linkage disequilibrium showed the strongest statistical evidence of association in the seven confirmed loci, where the linkage disequilibrium relations and haplotype patterns appear to be similar but not identical between European-descent and Japanese populations (supplementary Figs. S3 and S4).

Besides *KCNQ1* and the 16 candidate loci prioritized here, we investigated the disease association of two candidate gene SNPs—rs734312 in *WFS1* (39) and rs7501939 in *TCF2* (40)—based on the genotype data of our stage 1 panel ($n = 1,022$) and 964 random control subjects (supplementary Table S6). In some instances, it appeared that the sample size was not sufficient to detect the presumed odds ratio (OR) (supplementary Table S7). Nevertheless, except for rs12779790 in *CDC123-CAMK1D* and rs3740878 in *EXT2*, in the majority of instances, the ORs were consistent with those previously reported. Furthermore, we analyzed seven previously reported SNPs with suggestive evidence of an association in the Japanese population (25), but none attained nominal significance in our first-stage scan (supplementary Table S8).

Ethnic differences in genetic effects on type 2 diabetes. With regard to the candidate gene SNPs robustly confirmed in GWA studies conducted on Japanese and European-descent populations, we compared the risk allele frequency and OR between the meta-analysis dataset of the Japanese population and that of populations of European descent (8–10) (Fig. 3). The OR was consistently higher in the Japanese population for all SNPs except rs5219 in *KCNJ11*. Among the confirmed loci, *CDKAL1*, *CDKN2A/CDKN2B*, *SLC30A8*, and *HHEX* showed a significant difference in the ORs between European-descent and Japanese populations ($P < 0.05$, Woolf's test) (supplementary Table S6). However, the risk allele frequency fluctuated between the two ethnic groups, and the strength of association differed accordingly; this is because an SNP with an risk allele frequency of ~0.5 and a higher OR can give rise to stronger association signals. Thus, whereas *TCF7L2* was shown as the strongest susceptibility locus in populations of European descent (41), its association is estimated to be much weaker in the Japanese population because of the low risk allele frequency. In contrast, the results of the meta-analysis showed that the *CDKN2A/CDKN2B* and *CDKAL1* loci had the strongest associations in the Japanese population;
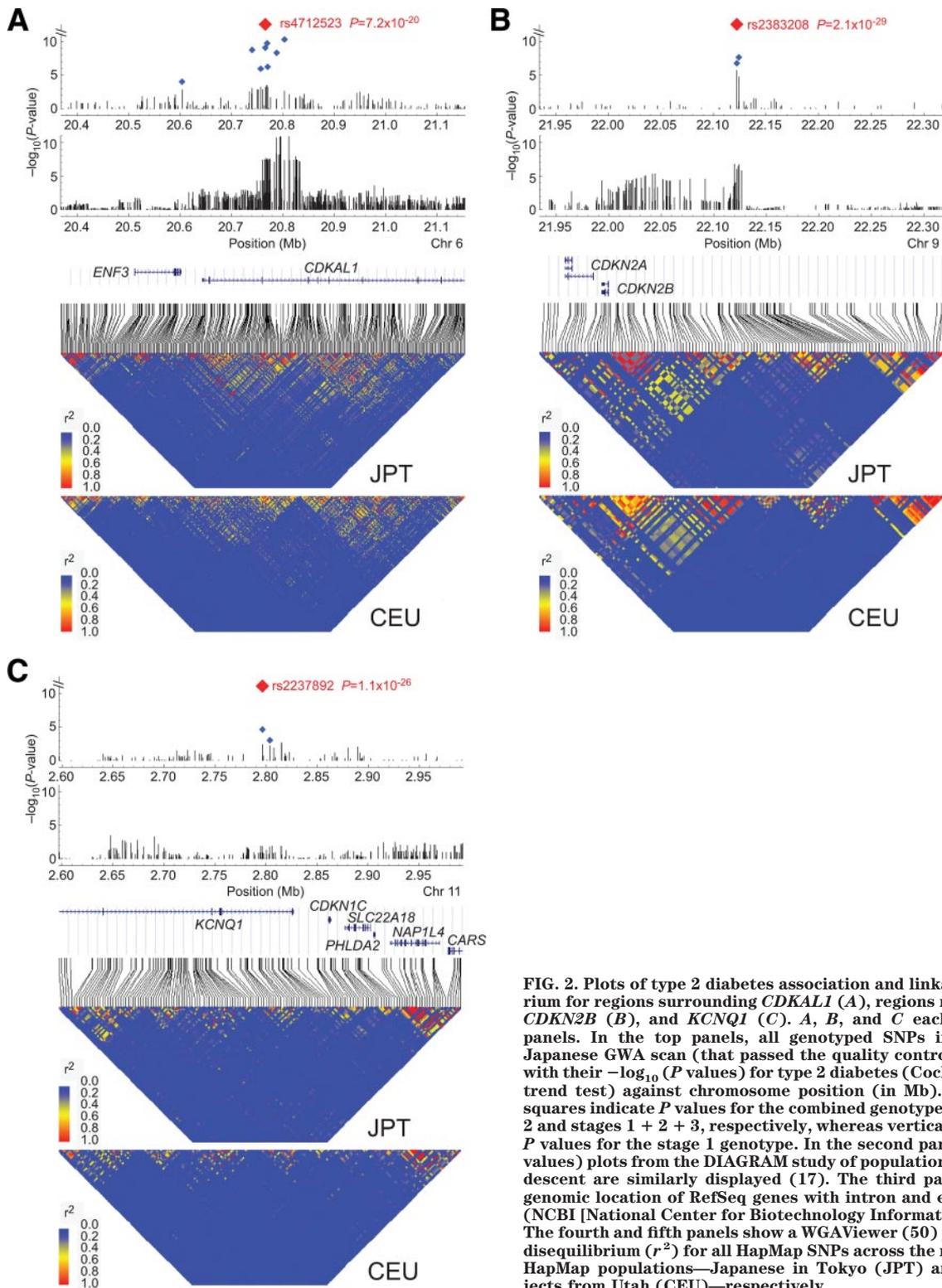
FIG. 2. Plots of type 2 diabetes association and linkage disequilibrium for regions surrounding *CDKAL1* (*A*), regions near *CDKN2A/ CDKN2B* (*B*), and *KCNQ1* (*C*). *A*, *B*, and *C* each contain five panels. In the top panels, all genotyped SNPs in the current Japanese GWA scan (that passed the quality control) are plotted with their $-\log_{10}$ (*P* values) for type 2 diabetes (Cochran-Armitage trend test) against chromosome position (in Mb). Blue and red squares indicate *P* values for the combined genotypes of stages 1 + 2 and stages 1 + 2 + 3, respectively, whereas vertical bars indicate *P* values for the stage 1 genotype. In the second panels, $-\log_{10}$ (*P* values) plots from the DIAGRAM study of populations of European descent are similarly displayed (17). The third panels show the genomic location of RefSeq genes with intron and exon structure (NCBI [National Center for Biotechnology Information] Build 35). The fourth and fifth panels show a WGAViewer (50) plot of linkage disequilibrium (r^2) for all HapMap SNPs across the regions for the HapMap populations—Japanese in Tokyo (JPT) and CEPH subjects from Utah (CEU)—respectively.

indeed, we obtained the highest number of hits for these loci.

Next, we compared the strength of association for the seven confirmed loci between Japanese and European-descent populations and calculated R^2 as the proportion of phenotypic variance explained by an SNP (see RESEARCH DESIGN AND METHODS). In Fig. 3, we illustrate the curves corresponding to $R^2 = 0.008$, 0.004, and 0.002, for which the total sample size of case and control subjects required

to attain 80% power is $n = 4,300$, 8,600, and 17,200 at a significance level of $P = 5 \times 10^{-7}$ (which is the significance threshold generally required in GWA tests), and $n = 1,000$, 2,000, and 3,900 at a level of $P = 0.05$. Based on R^2 measurements using the meta-analysis data, the associations of five of seven replicated loci are stronger in the Japanese population than in populations of European descent. For the *CDKAL1* locus, for example, one-fourth of the sample size necessary in populations of European descent is

TABLE 1
Type 2 diabetes susceptibility loci identified or tested for replication in the current Japanese study

rs no.*	Chromosome	Position (bp)	Region	Risk allele/ nonrisk allele	Control risk allele proportion	Stage 1 + 2 (1,629 case subjects/1,517 control subjects)	
						OR (95% CI)	P trend
Identified in this GWA scan							
rs4712523	6	20,765,543	CDKAL1	G/A	0.407	1.38 (1.25–1.52)	8.0E-10
rs2383208	9	22,122,076	CDKN2A/B	A/G	0.553	1.31 (1.18–1.45)	1.6E-07
rs2237892	11	2,796,327	KCNQ1	C/T	0.594	1.25 (1.13–1.39)	2.3E-05
rs10425678	19	38,669,236	PEPD	C/T	0.261	1.23 (1.10–1.37)	3.6E-04
Replication of previously-reported SNPs							
rs10923931	1	120,230,001	NOTCH2-ADAM30	T/G	0.020	1.17 (0.83–1.65)	0.3821
rs7578597	2	43,644,474	THADA	T/C	0.990	1.95 (1.03–3.67)	0.0392
rs1801282	3	12,368,125	PPARG	C/G	0.969	1.00 (0.75–1.34)	0.9741
rs4607103	3	64,686,944	ADAMTS9	C/T	0.594	1.09 (0.99–1.21)	0.0902
rs4402960	3	186,994,389	IGF2BP2	T/G	0.310	1.15 (1.04–1.28)	0.0098
rs7754840	6	20,769,229	CDKAL1	C/G	0.392	1.42 (1.28–1.57)	1.7E-10
rs7756992	6	20,787,688	CDKAL1	G/A	0.448	1.35 (1.23–1.50)	4.6E-09
rs864745	7	27,953,796	JAZF1	T/C	0.789	1.08 (0.95–1.22)	0.2456
rs13266634	8	118,253,964	SLC30A8	C/T	0.570	1.18 (1.06–1.30)	0.0015
rs10811661	9	22,124,094	CDKN2A/B	T/C	0.555	1.35 (1.21–1.49)	2.2E-08
rs12779790	10	12,368,016	CDC123-CAMK1D	G/A	0.151	0.98 (0.85–1.13)	0.7984
rs1111875	10	94,452,862	HHEX	C/T	0.275	1.19 (1.07–1.33)	0.0011
rs7903146	10	114,748,339	TCF7L2	T/C	0.035	1.42 (1.10–1.84)	0.0073
rs5219	11	17,366,148	KCNJ11	T/C	0.355	1.22 (1.09–1.35)	2.5E-04
rs3740878	11	44,214,378	EXT2	A/G	0.633	1.01 (0.91–1.12)	0.8849
rs7961581	12	69,949,369	TSPAN8-LGR5	C/T	0.202	1.12 (0.99–1.27)	0.0751
rs8050136	16	52,373,776	FTO	A/C	0.203	1.11 (0.98–1.26)	0.0915

Continued on following page

sufficient to obtain the same level of statistical significance in the Japanese population. This is true for *CDKN2A/CDKN2B*, *HHEX*, and *SLC30A8*, in which <50% of the sample size seems to be sufficient for significance in the Japanese population. However, *TCF7L2* shows an opposite trend in this regard.

Combined genetic risk of type 2 diabetes. Despite the small value of explained variance (R^2) at each risk locus, it is assumed that knowledge about multiple-risk loci could allow us to identify individuals with accumulated genetic risk (42). To this end, a GWA study in Finns (10) investigated the combined risk of type 2 diabetes based on 10 associated loci by logistic regression analysis of the resampled dataset. The total variance explained by 10 loci in Finns is $R^2 = 0.030$, which is equivalent to the value for 7 loci obtained here (see DISCUSSION). Likewise, in a simulated population, we arranged the individuals in the order of the risk estimated by logistic regression, sorted them into 20 equal-sized groups (5% in each), and calculated the actual proportion of affected individuals in each group. We found a 3.7-fold variation in type 2 diabetes prevalence from the lowest to highest estimated risk groups for the combination of seven associated loci in our study (Fig. 4). The receiver operating characteristic curve was also depicted for the combined SNPs as a measure of sensitivity and specificity (supplementary Fig. S5).

Moreover, for risk assessment in the general population, we performed multiple regression analysis using A1C as a surrogate quantitative phenotype to estimate the unbiased effect size of individual loci (supplementary Table S9) and evaluated the combined risk from multiple loci in 12,569 population-based samples (Table 2 and supplementary information). Then, the estimated risk was compared with the actual A1C value and the disease classification of diabetes or pre-diabetes (supplementary information). In the multiple regression analysis, significant association ($P < 0.005$) was observed for all seven loci tested in accordance with the results for the case-control study (Table 1 and supplementary Table S4). As shown in Fig. 4, 5% of male subjects with the highest estimated risk are 2.3 times more likely to suffer from diabetes than those with the lowest estimated risk; the risk is 5.2 times in female subjects, indicating the potential existence of sex difference in the genetic risk of type 2 diabetes (supplementary Fig. S6). Moreover, notably, SNP genotypes alone exerted more exaggerated effects on the increase in genetic risk in diabetes compared with pre-diabetes (Table 2).

DISCUSSION

Conducting GWA studies on a wider range of populations, including East Asians, has recently gained importance because of the discovery of new type 2 diabetes susceptibility variants mapping to the *KCNQ1* gene simultaneously reported in two Japanese studies (25,26). Both studies were, however, initiated some years ago, and they are, by current standards, considered to be modest with regard to the coverage of common SNPs (21 and 56% in HapMap) and number of case subjects (187 and 194 subjects, respectively) in the first-stage scan. Therefore, we conducted another GWA study on the Japanese population with greater coverage of common SNPs (87% of all phase

TABLE 1
Continued

Stage 3 (4,000 case subjects/4,889 control subjects)†		All combined (5,629 case subjects/6,406 control subjects)†		OR (95% CI) reported in Europeans (14,586 case subjects/17,968 control subjects)
OR (95% CI)	P trend‡	OR (95% CI)	P trend	
1.23 (1.16–1.30)	4.0E-12	1.27 (1.21–1.33)	7.2E-20	1.12 (1.08–1.16)
1.33 (1.26–1.42)	4.8E-22	1.34 (1.27–1.41)	2.1E-29	1.20 (1.14–1.25)
1.36 (1.28–1.45)	8.0E-23	1.33 (1.27–1.41)	1.1E-26	1.18 (1.03–1.33)§
1.10 (1.03–1.18)	0.0020	1.14 (1.07–1.20)	1.4E-05	1.03 (0.97–1.09)§
—	—	—	—	1.13 (1.08–1.17)‖
0.98 (0.73–1.31)	0.55	1.13 (0.87–1.47)	0.35	1.15 (1.10–1.20)‖
—	—	—	—	1.14 (1.08–1.20)
—	—	—	—	1.09 (1.06–1.12)‖
1.14 (1.07–1.21)	2.5E-05	1.14 (1.08–1.21)	1.0E-06	1.14 (1.11–1.18)
—	—	—	—	1.12 (1.08–1.16)
—	—	—	—	1.26 (1.18–1.34)§
—	—	—	—	1.10 (1.07–1.13)‖
1.24 (1.17–1.31)	5.8E-13	1.22 (1.16–1.28)	1.8E-14	1.12 (1.07–1.16)
—	—	—	—	1.2 (1.14–1.25)
—	—	—	—	1.11 (1.07–1.14)‖
1.21 (1.13–1.29)	2.6E-09	1.21 (1.15–1.28)	6.7E-12	1.13 (1.09–1.17)
1.59 (1.38–1.83)	5.3E-11	1.54 (1.36–1.74)	7.6E-12	1.37 (1.31–1.43)
1.02 (0.96–1.08)	0.3008	1.07 (1.01–1.13)	0.0149	1.14 (1.10–1.19)
—	—	—	—	1.20 (1.11–1.30)¶
—	—	—	—	1.09 (1.06–1.12)‖
—	—	—	—	1.17 (1.12–1.22)

Results for one SNP each selected from the individual chromosomal regions in the GWA scans are shown in the table (see supplementary Table S4 for details and supplementary Table S5 for the results of logistic regression adjusted for BMI. The final P value was assessed by pooling genotype counts for each SNP from all stages tested (without including 964 random control subjects from GeMBDJ). In two regions, chromosome 6p22.3 (*CDKAL1*) and 19p13 (*PEPD*), the haplotype class showed more significant association than the individual SNP (see supplementary Information). *In the stage 3 panel, we genotyped rs4712523 instead of rs7754840 ($r^2 = 0.96$) or rs7756992 ($r^2 = 0.65$) in CDKAL1, and rs2383208 instead of rs10811661 ($r^2 = 0.89$) near CDKN2A/B, with the aim of determining the SNP(s) with the strongest association in the Japanese population. †In stage 3 of the replication study on previously reported SNPs, after the confirmation of significant association in 4,000 subjects and 4,889 preselected control subjects, we further characterized 7,680 subjects (who comprised the rest of the 12,569 population-based samples) (see RESEARCH DESIGN AND METHODS and Fig. 1). Thus, for the corresponding SNPs, 5,395 control subjects were reselected from the entire population-based samples and used for the final association analysis in stage 3, which increased the total number of control subjects across the three stages to 6,912. ‡One-tailed test for association was performed in the direction consistent with stage 1 + 2 data; §for 4,549 case and 5,579 control subjects derived from the DIAGRAM consortium of Zeggini et al. (17); ‖for ~60,000 total samples from Zeggini et al. (17); ¶for 3,278 case and 3,508 control subjects from Sladek et al. (6).

1 + 2 HapMap variants [MAF ≥0.05] in CHB (Chinese in Beijing) + JPT (Japanese in Tokyo) and a larger number of case subjects (519 subjects) and unaffected control subjects (503 subjects) in addition to random control subjects in the first-stage scan. Four loci (three previously reported and one novel) were identified via the multistage scans. For the top three loci (*KCNQ1*, *CDKN2A/CDKN2B*, and *CDKAL1*) the OR (>1.25) and MAF (0.41–0.45 in the control subjects) were higher in the Japanese population than in populations of European descent. In addition to the nomination of four susceptibility loci (*KCNQ1*, *CDKN2A/CDKN2B*, *CDKAL1*, and *PEPD*), the current study replicated the significant association of five other loci (*TCF7L2*, *IGF2BP2*, *SLC30A8*, *HHEX*, and *KCNJ11*) previously reported in populations of European descent (6–17) and provided an unbiased estimate of the risk from the confirmed disease genotype.

Empirical studies suggest that the genetic effects of individual causal risk alleles underlying complex genetic diseases such as type 2 diabetes are modest, with most genotype relative risks in the range of 1.1–2.0 (43). Indeed, we observed this to be true for loci that were robustly implicated in the development of type 2 diabetes by GWA scans and/or extensive candidate gene approaches in populations of European descent. Currently, the number of loci has increased to almost 20 (as listed in supplementary Table S6), and in most cases, except for *TCF7L2* and *KCNQ1*, the OR is estimated to be between 1.09 and 1.20.

The current study provides, via genome-wide exploration and replication analysis of some a priori selected loci, significant evidence for the overall tendency toward a stronger association in Japanese rather than European-descent populations at least for alleles with a cosmopolitan effect. The tendency for higher OR in Asians than in Europeans was previously reported for the *CDKAL1* locus (22). Currently, it remains unknown whether the penetrance for a genotype of interest differs considerably between Japanese and European-descent populations. With regard to genetic effects, four of seven confirmed loci have demonstrated significantly higher OR in the Japanese population ($P = 4.1 \times 10^{-5}$ to 0.024) (supplementary Table S6). To simplify the situation, we have further assessed the strength of association for individual SNPs by measuring R^2, which is scaled against OR and risk allele frequency in
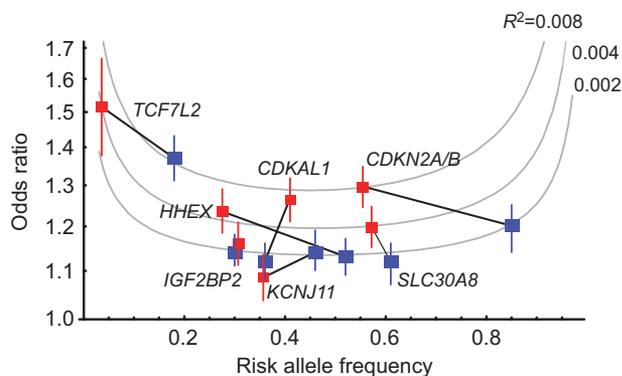
FIG. 3. Comparison of the strength of association for seven confirmed type 2 diabetes loci between Japanese and European-descent populations. For the Japanese population, we estimated ORs and their 95% CIs (red solid squares and vertical lines, respectively) for each locus based on our meta-analysis involving four Japanese case-control studies (supplementary Fig. S2). For populations of European descent, on the other hand, the corresponding values (blue solid squares and vertical lines) were derived from the published data (8–10). The association of an SNP with type 2 diabetes is measured by the coefficient of determination (R^2), which represents the ability to detect association signals using the Cochran-Armitage trend test.

Fig. 3. We found that despite the limited number of SNPs tested here, the same level of statistical significance is often detectable in the Japanese population with a much smaller sample size than that in populations of European descent (supplementary Table S7). Theoretically, the stringency of ascertaining control subjects could lead to some bias in effect size (44). In this respect, in addition to the multistage case-control study, an extensive analysis of associated loci in the general population was conducted, which is the strength of the current study. We used the population-based samples ($n = 12,569$) in stage 3 to
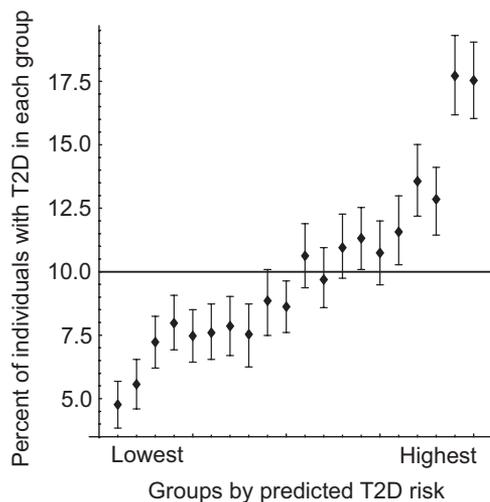
FIG. 4. Estimation of the increase in type 2 diabetes risk from the combination of seven susceptibility variants previously identified and robustly replicated in the current study. We used case and control subjects with complete data from all stages of our study ($n = 12,105$). First, the risk for the genotypes of an SNP was estimated by logistic regression. Then, the multilocus risk for an individual was assessed as the sum of the risks for his/her genotype at seven SNPs. We simulated a population with 10% prevalence by bootstrap sampling. In the simulated population, we arranged the individuals in the order of their multilocus risk, sorted them into 20 equal-sized groups, and calculated the actual prevalence in each group. Means and 95% CIs of the groupwise prevalence were estimated based on 1,000 bootstrap sampling trials and are plotted in the figure. No significant gene-gene interaction was observed between the seven SNPs by multiple logistic regression analysis. T2D, type 2 diabetes.

investigate the effect of control selection criteria on OR in a case-control comparison and found that the ORs in our meta-analysis were almost comparable to those estimated in the general Japanese population (supplementary Table S10). Moreover, with regard to ethnic diversity, linkage disequilibrium in *CDKAL1* and *KCNJ11* is stronger in East Asians (JPT + CHB), whereas linkage disequilibrium in *IGFBP2* and *HHEX* tends to be stronger in Europeans (CEU [Centre d'Etude du Polymorphisme Humain (CEPH) subjectsfrom Utah]) (Supplementary Figure S4); thus, besides the issue of power, the results of the GWA scans in the Japanese population (or East Asians) seem to be useful in terms of interethnic comparison of association signals, which may enhance the power of fine-mapping efforts designed to identify the causal variants (45).

The tendency of stronger genetic association in the Japanese population is also supported by the concomitant evaluation of multilocus effects. When assuming an additive model, the combined risk of type 2 diabetes can be measured by the sum of the R^2 values of individual loci. For example, the total variance explained by the seven loci depicted in Fig. 3 is 0.030 in the Japanese population and 0.018 in populations of European descent. It remains unknown whether these findings reflect higher heritability of type 2 diabetes in Japanese than in European-descent populations. Because little data are available on the estimation of heritability in the Japanese or East Asian populations, further studies are required to obtain the standardized measures of heritability across different populations by taking into account potential sources of heterogeneity, such as the degree of westernization of lifestyle.

Suggestive evidence of association was identified for SNPs in the *PEPD* gene. *PEPD* plays an important role in collagen metabolism, and some extracellular matrix constituents such as collagen IV have been shown to have a profound impact on insulin secretion (46). Moreover, enhanced collagen degradation via *PEPD* activity has been reported in diabetic patients (47). Although there is evidence suggestive of association at *PEPD* in all three stages, the current GWA study by itself could not confirm or refute the evidence; no significant association was found in the previously reported Diabetes Genetics Replication and Meta-Analysis (DIAGRAM) data from Europeans (risk allele frequency = 0.52, OR = 1.03) (Table 1) and in the initial screening data of the JSNP (Japanese Single Nucleotide Polymorphisms) scan in the Japanese population (187 casevs. 752 random control subjects; $P = 0.18$ at rs2241380, which is in complete linkage disequilibrium with rs10425678 in *PEPD*; $r^2 = 1.0$) (25).

The number of genes that could account for an appreciable population-attributable fraction of common diseases is under debate (48). Although the current study detected and/or replicated a total of nine susceptibility loci, including *PEPD* in the Japanese population, a substantial number of SNPs showing some extent of association signals in the first-stage scan remain to be investigated, as reflected by the wide distribution of replicated SNPs with unexamined "gaps" in the lower-left part of the Q-Q plot (supplementary Fig. S7). The ORs corresponding to such unexamined SNPs mostly fall in the range of 1.10–1.25. To assess the statistical power in our GWA scan, we simulated the frequency at which a disease-associated SNP could surpass the cutoff level of the first two stages (stages 1 and 2) (supplementary information and supplementary Table S11). In the current experimental setting, it is likely that >50% of the susceptibility loci

TABLE 2
Combined risk of diabetes and pre-diabetic status based on seven confirmed loci, age, BMI, and sex in the general Japanese population

		Diabetes		Pre-diabetes		Diabetes + Pre-diabetes	
	A1C (%)	RR versus population average (95% CI)	Prevalence	RR versus population average (95% CI)	Prevalence	RR versus population average (95% CI)	Prevalence
Male							
Whole population	5.29 ± 0.88	1.00	0.16	1.00	0.07	1.00	0.23
Highest risk group (5%) assessed by							
All predictors	5.48 ± 0.87	1.65 (1.29–1.97)	0.27	1.34 (0.73–1.83)	0.09	1.56 (1.26–1.78)	0.36
SNP genotypes	5.57 ± 1.12	1.67 (1.32–2.06)	0.27	0.92 (0.44–1.40)	0.07	1.45 (1.16–1.73)	0.34
Age and BMI*	5.44 ± 0.78	1.16 (0.87–1.46)	0.19	1.95 (1.39–2.60)	0.14	1.40 (1.16–1.65)	0.33
Lowest risk group (5%) assessed by							
All predictors	4.98 ± 0.73	0.46 (0.26–0.74)	0.08	0.50 (0.20–0.90)	0.04	0.47 (0.33–0.70)	0.11
SNP genotypes	5.11 ± 0.74	0.72 (0.39–0.92)	0.12	0.71 (0.30–1.10)	0.05	0.72 (0.46–0.86)	0.17
Age and BMI*	4.98 ± 0.77	0.46 (0.30–0.73)	0.08	0.40 (0.10–0.60)	0.03	0.44 (0.27–0.63)	0.10
Female							
Whole population	5.17 ± 0.60	1.00	0.07	1.00	0.06	1.00	0.13
Highest risk group (5%) assessed by							
All predictors	5.55 ± 0.96	3.09 (2.36–3.73)	0.22	2.05 (1.37–2.60)	0.13	2.61 (2.10–2.96)	0.35
SNP genotypes	5.37 ± 0.88	2.30 (1.60–2.78)	0.17	1.17 (0.73–1.80)	0.07	1.78 (1.41–2.10)	0.24
Age and BMI*	5.42 ± 0.78	2.26 (1.71–2.78)	0.16	1.95 (1.34–2.53)	0.12	2.12 (1.73–2.46)	0.28
Lowest risk group (5%) assessed by							
All predictors	4.91 ± 0.43	0.16 (0.00–0.32)	0.01	0.14 (0.00–0.28)	0.01	0.15 (0.04–0.26)	0.02
SNP genotypes	5.02 ± 0.45	0.45 (0.16–0.73)	0.03	0.80 (0.38–1.22)	0.05	0.61 (0.35–0.83)	0.08
Age and BMI*	4.94 ± 0.36	0.24 (0.08–0.64)	0.02	0.19 (0.00–0.37)	0.01	0.22 (0.09–0.47)	0.03

Data are the means ± SD, unless otherwise indicated. Relative risk (RR) is calculated as the ratio of the prevalence in 5% of people with the highest or lowest risk to the prevalence in the whole population. In this study, the combined disease risk for each individual was assessed using the regression for A1C (see supplementary information). Subjects with self-reported diabetes or with A1C ≥6.1 were classified as diabetic, and those who were not under antidiabetic medication and with $5.6 \leq A1C < 6.1$ were classified as pre-diabetic. The actual A1C level and the distribution by diabetic status for each 5% subgroup of the risk group are illustrated in supplementary Fig. S6. *For reference, diabetes and/or pre-diabetes risk was assessed using the participant's age and BMI alone as predictors.

with modest but substantial effects (OR = 1.2–1.3) were unidentified. For example, though not statistically significant, the association of *PPARG* in the Japanese population showed an OR (P = 0.06, OR = 1.18 at rs1801282) similar to that in populations of European descent in a meta-analysis, including the current study (supplementary Table S4). Increasing the sample size of the stage 1 panel and/or the number of SNPs genotyped in the second-stage scan would allow us to discover more susceptibility variants, including new population-specific loci, in the Japanese population.

The incidence of type 2 diabetes is escalating to epidemic proportions globally, with a higher acceleration rate in non-European populations (49). The integration of GWA study results, i.e., a meta-analysis (17), for both European-descent and non-European populations is necessary for a comprehensive understanding of the genetics of type 2 diabetes, and it will lead to the efficient use of genomic information based on ethnic diversity in clinical research.

REFERENCES

1. Ehtisham S, Crabtree N, Clark P, Shaw N, Barrett T. Ethnic differences in insulin resistance and body composition in United Kingdom adolescents. J Clin Endocrinol Metab 2005;90:3963–3969
2. Wong J, Molyneaux L, Zhao D, Constantino M, Gray RS, Twigg SM, Xu ZR, Yue DK. Different accelerators to early-onset type 2 diabetes: a comparison of Anglo-Celtic and Chinese patients. J Diabetes Complications 2008;22:389–394
3. Stevens J, Truesdale KP, Katz EG, Cai J. Impact of body mass index on incident hypertension and diabetes in Chinese Asians, American whites, and American blacks: the People's Republic of China Study and the Atherosclerosis Risk in Communities Study. Am J Epidemiol 2008;167:1365–1374
4. Fukushima M, Usami M, Ikeda M, Nakai Y, Taniguchi A, Matsuura T, Suzuki H, Kurose T, Yamada Y, Seino Y. Insulin secretion and insulin sensitivity at different stages of glucose tolerance: a cross-sectional study of Japanese type 2 diabetes. Metabolism 2004;53:831–835
5. Nakanishi S, Okubo M, Yoneda M, Jitsuiki K, Yamane K, Kohno N. A comparison between Japanese-Americans living in Hawaii and Los Angeles and native Japanese: the impact of lifestyle westernization on diabetes mellitus. Biomed Pharmacother 2004;58:571–577
6. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445:881–885
7. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–678
8. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney AS; Wellcome Trust Case Control Consortium (WTCCC), McCarthy MI, Hattersley AT. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007;316:1336–1341
9. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Boström K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Råstam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjögren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316:1331–1336
10. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316:1341–1345
11. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, Styrkarsdottir U, Gretarsdottir S, Emilsson V, Ghosh S, Baker A, Snorradottir S, Bjarnason H, Ng MC, Hansen T, Bagger

Y, Wilensky RL, Reilly MP, Adeyemo A, Chen Y, Zhou J, Gudnason V, Chen G, Huang H, Lashley K, Doumatey A, So WY, Ma RC, Andersen G, Borch-Johnsen K, Jorgensen T, van Vliet-Ostaptchouk JV, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Rotimi C, Gurney M, Chan JC, Pedersen O, Sigurdsson G, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet 2007;39:770–775
12. Salonen JT, Uimari P, Aalto JM, Pirskanen M, Kaikkonen J, Todorova B, Hyppönen J, Korhonen VP, Asikainen J, Devine C, Tuomainen TP, Luedemann J, Nauck M, Kerner W, Stephens RH, New JP, Ollier WE, Gibson JM, Payton A, Horan MA, Pendleton N, Mahoney M, Meyre D, Delplanque J, Froguel P, Luzzatto O, Yakir B, Darvasi A. Type 2 diabetes whole-genome association study in four populations: the DiaGen Consortium. Am J Hum Genet 2007;81:338–345
13. Florez JC, Manning AK, Dupuis J, McAteer J, Irenze K, Gianniny L, Mirel DB, Fox CS, Cupples LA, Meigs JB. A 100K genome-wide association scan for diabetes and related traits in the Framingham Heart Study: replication and integration with other genome-wide datasets. Diabetes 2007;56:3063–3074
14. Hanson RL, Bogardus C, Duggan D, Kobes S, Knowlton M, Infante AM, Marovich L, Benitez D, Baier LJ, Knowler WC. A search for variants associated with young-onset type 2 diabetes in American Indians in a 100K genotyping array. Diabetes 2007;56:3045–3052
15. Hayes MG, Pluzhnikov A, Miyake K, Sun Y, Ng MC, Roe CA, Below JE, Nicolae RI, Konkashbaev A, Bell GI, Cox NJ, Hanis CL. Identification of type 2 diabetes genes in Mexican Americans through genome-wide association studies. Diabetes 2007;56:3033–3044
16. Rampersaud E, Damcott CM, Fu M, Shen H, McArdle P, Shi X, Shelton J, Yin J, Chang YP, Ott SH, Zhang L, Zhao Y, Mitchell BD, O'Connell J, Shuldiner AR. Identification of novel candidate genes for type 2 diabetes from a genome-wide association scan in the Old Order Amish: evidence for replication from diabetes-related quantitative traits and from independent populations. Diabetes 2007;56:3053–3062
17. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Boström KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jørgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjögren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ; Wellcome Trust Case Control Consortium, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638–645
18. Frayling TM, McCarthy MI. Genetic studies of diabetes following the advent of the genome-wide association study: where do we go from here? Diabetologia 2007;50:2229–2233
19. Omori S, Tanaka Y, Takahashi A, Hirose H, Kashiwagi A, Kaku K, Kawamori R, Nakamura Y, Maeda S. Association of CDKAL1, IGF2BP2, CDKN2A/B, HHEX, SLC30A8, and KCNJ11 with susceptibility to type 2 diabetes in a Japanese population. Diabetes 2008;57:791–795
20. Horikoshi M, Hara K, Ito C, Shojima N, Nagai R, Ueki K, Froguel P, Kadowaki T. Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. Diabetologia 2007;50:2461–2466
21. Horikawa Y, Miyake K, Yasuda K, Enya M, Hirota Y, Yamagata K, Hinokio Y, Oka Y, Iwasaki N, Iwamoto Y, Yamada Y, Seino Y, Maegawa H, Kashiwagi A, Yamamoto K, Tokunaga K, Takeda J, Kasuga M. Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. J Clin Endocrinol Metab 2008;93:3136–3141
22. Wu Y, Li H, Loos RJ, Yu Z, Ye X, Chen L, Pan A, Hu FB, Lin X. Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. Diabetes 2008;57:2834–2842
23. Ng MC, Park KS, Oh B, Tam CH, Cho YM, Shin HD, Lam VK, So WY, Cho YS, Kim HL, Lee HK, Chan JC, Cho NH. Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2 diabetes and obesity in 6,719 Asians. Diabetes 2008;57:2226–2233
24. Lewis JP, Palmer ND, Hicks PJ, Sale MM, Langefeld CD, Freedman BI,

Divers J, Bowden DW. Association analysis in African Americans of European-derived type 2 diabetes single nucleotide polymorphisms from whole-genome association studies. Diabetes 2008;57:2220–2225

25. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, Hirota Y, Mori H, Jonsson A, Sato Y, Yamagata K, Hinokio Y, Wang HY, Tanahashi T, Nakamura N, Oka Y, Iwasaki N, Iwamoto Y, Yamada Y, Seino Y, Maegawa H, Kashiwagi A, Takeda J, Maeda E, Shin HD, Cho YM, Park KS, Lee HK, Ng MC, Ma RC, So WY, Chan JC, Lyssenko V, Tuomi T, Nilsson P, Groop L, Kamatani N, Sekine A, Nakamura Y, Yamamoto K, Yoshida T, Tokunaga K, Itakura M, Makino H, Nanjo K, Kadowaki T, Kasuga M. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet 2008;40:1092–1097

26. Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, Andersen G, Ng DP, Holmkvist J, Borch-Johnsen K, Jørgensen T, Sandbaek A, Lauritzen T, Hansen T, Nurbaya S, Tsunoda T, Kubo M, Babazono T, Hirose H, Hayashi M, Iwamoto Y, Kashiwagi A, Kaku K, Kawamori R, Tai ES, Pedersen O, Kamatani N, Kadowaki T, Kikkawa R, Nakamura Y, Maeda S. SNPs in KCNQ1 are associated with susceptibilityto type 2 diabetes in East Asian and European populations. Nat Genet 2008;40:1098–1102

27. Nakamura Y. The BioBank Japan Project. Clin Adv Hematol Oncol 2007;5:696–697

28. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575

29. The International HapMap Consortium, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–861

30. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. NatGenet 2006;38:904–909

31. Devlin B, Roeder K. Genomic control for association studies. Biometrics 1999;55:997–1004

32. Nawata H, Shirasawa S, Nakashima N, Araki E, Hashiguchi J, Miyake S, Yamauchi T, Hamaguchi K, Yoshimatsu H, Takeda H, Fukushima H, Sasahara T, Yamaguchi K, Sonoda N, Sonoda T, Matsumoto M, Tanaka Y, Sugimoto H, Tsubouchi H, Inoguchi T, Yanase T, Wake N, Narazaki K, Eto T, Umeda F, Nakazaki M, Ono J, Asano T, Ito Y, Akazawa S, Hazegawa I, Takasu N, Shinohara M, Nishikawa T, Nagafuchi S, Okeda T, Eguchi K, Iwase M, Ishikawa M, Aoki M, Keicho N, Kato N, Yasuda K, Yamamoto K, Sasazuki T. Genome-wide linkage analysis of type 2 diabetes mellitus reconfirms the susceptibility locus on 11p13–p12 in Japanese. J Hum Genet 2004;49:629–634

33. Hayashi T, Iwamoto Y, Kaku K, Hirose H, Maeda S. Replication study for the association of TCF7L2 with susceptibility to type 2 diabetes in a Japanese population. Diabetologia 2007;50:980–984

34. Horikoshi M, Hara K, Ito C, Nagai R, Froguel P, Kadowaki T. A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. Diabetologia 2007;50:747–751

35. Miyake K, Horikawa Y, Hara K, Yasuda K, Osawa H, Furuta H, Hirota Y, Yamagata K, Hinokio Y, Oka Y, Iwasaki N, Iwamoto Y, Yamada Y, Seino Y, Maegawa H, Kashiwagi A, Yamamoto K, Tokunaga K, Takeda J, Makino H, Nanjo K, Kadowaki T, Kasuga M. Association of TCF7L2 polymorphisms with susceptibility to type 2 diabetes in 4,087 Japanese subjects. J Hum Genet 2008;53:174–180

36. Mori H, Ikegami H, Kawaguchi Y, Seino S, Yokoi N, Takeda J, Inoue I, Seino Y, Yasuda K, Hanafusa T, Yamagata K, Awata T, Kadowaki T, Hara K, Yamada N, Gotoda T, Iwasaki N, Iwamoto Y, Sanke T, Nanjo K, Oka Y, Matsutani A, Maeda E, Kasuga M. The Pro12→Ala substitution in PPAR-gamma is associated with resistance to development of diabetes in the general population: possible involvement in impairment of insulin secretion in individuals with type 2 diabetes. Diabetes 2001;50:891–894

37. Woolf B. On estimating the relation between blood group and disease. Ann Intern Med 1955;19:251–253

38. Yamaguchi-Kabata Y, Nakazono K, Takahashi A, Saito S, Hosono N, Kubo M, Nakamura Y, Kamatani N. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. Am J Hum Genet 2008;83:445–456

39. Sandhu MS, Weedon MN, Fawcett KA, Wasson J, Debenham SL, Daly A, Lango H, Frayling TM, Neumann RJ, Sherva R, Blech I, Pharoah PD, Palmer CN, Kimber C, Tavendale R, Morris AD, McCarthy MI, Walker M, Hitman G, Glaser B, Permutt MA, Hattersley AT, Wareham NJ, Barroso I. Common variants in WFS1 confer risk of type 2 diabetes. Nat Genet 2007;39:951–953

40. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifs-son G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, Adeyemo A, Chen Y, Zhou J, So WY, Tong PC, Ng MC, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Tres A, Fuertes F, Ruiz-Echarri M, Asin L, Saez B, van Boven E, Klaver S, Swinkels DW, Aben KK, Graif T, Cashy J, Suarez BK, van Vierssen Trip O, Frigge ML, Ober C, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Palmer CN, Rotimi C, Chan JC, Pedersen O, Sigurdsson G, Benediktsson R, Jonsson E, Einarsson GV, Mayordomo JI, Catalona WJ, Kiemeney LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet 2007;39:977–983

41. Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, Weitgasser R, Nejjari C, Patsch W, Chikri M, Meyre D, Froguel P. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. J Mol Med 2007;85:777–782

42. Lango H; UK Type 2 Diabetes Genetics Consortium, Palmer CN, Morris AD, Zeggini E, Hattersley AT, McCarthy MI, Frayling TM, Weedon MN. Assess-ing the combined impact of 18 common genetic variants of modest effect sizes on type 2 diabetes risk. Diabetes 2008;57:3129–3135

43. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk to disease from genome-wide association studies. Genome Res 2007;17:1520–1528

44. Garner C. The use of random controls in genetic association studies. Hum Hered 2006;61:22–26

45. McCarthy MI. Casting a wider net for diabetes susceptibility genes. Nat Genet 2008;40:1039–1040

46. Yuan J, Li T, Yin XB, Guo L, Jiang X, Jin W, Yang X, Wang E. Character-ization of prolidase activity using capillary electrophoresis with tris(2,2′-bipyridyl)ruthenium(II) electrochemiluminescence detection and application to evaluate collagen degradation in diabetes mellitus. Anal Chem 2006;78:2934–2938

47. Kaido T, Yebra M, Cirulli V, Rhodes C, Diaferia G, Montgomery AM. Impact of defined matrix interactions on insulin production by cultured human beta-cells: effect on insulin content, secretion, and gene transcription. Diabetes 2006;55:2723–2729

48. Yang Q, Khoury MJ, Friedman J, Little J, Flanders WD. How many genes underlie the occurrence of common complex diseases in the population? Int J Epidemiol 2005;34:1129–1137

49. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004;27:1047–1053

50. Ge D, Zhang K, Need AC, Martin O, Fellay J, Urban TJ, Telenti A, Goldstein DB. WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 2008;18:640–643