

- 23 Gloyn, A. L., Weedon, M. N., Owen, K. R., Turner, M. J., Knight, B. A., Hitman, G. *et al*. Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes* **52**, 568–572 (2003).
- 24 Altshuler, D., Hirschhorn, J. N., Klannemark, M., Lindgren, C. M., Vohl, M. C., Nemesh, J. *et al*. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat. Genet.* **26**, 76–80 (2000).
- 25 Sandhu, M. S., Weedon, M. N., Fawcett, K. A., Wasson, J., Debenham, S. L., Daly, A. *et al*. Common variants in WFS1 confer risk of type 2 diabetes. *Nat. Genet.* **39**, 951–953 (2007).
- 26 Gudmundsson, J., Sulem, P., Steinthorsdottir, V., Bergthorsson, J. T., Thorleifsson, G., Manolescu, A. *et al*. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat. Genet.* **39**, 977–983 (2007).
- 27 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* **50**, 747–751 (2007).
- 28 Yokoi, N., Kanamori, M., Horikawa, Y., Takeda, J., Sanke, T., Furuta, H. *et al*. Association studies of variants in the genes involved in pancreatic beta-cell function in type 2 diabetes in Japanese subjects. *Diabetes* **55**, 2379–2386 (2006).
- 29 Mori, H., Ikegami, H., Kawaguchi, Y., Seino, S., Yokoi, N., Takeda, J. *et al*. 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* **50**, 891–894 (2001).
- 30 Guo, F. & Cavener, D. R. The GCN2 eIF2alpha kinase regulates fatty-acid homeostasis in the liver during deprivation of an essential amino acid. *Cell Metab.* **5**, 103–114 (2007).
- 31 Harding, H. P., Zeng, H., Zhang, Y., Jungries, R., Chung, P., Plesken, H. *et al*. Diabetes mellitus and exocrine pancreatic dysfunction in *Perk^{-/-}* mice reveals a role for translational control in secretory cell survival. *Mol. Cell.* **7**, 1153–1163 (2001).
- 32 Delépine, M., Nicolino, M., Barrett, T., Golamaully, M., Lathrop, G. M. & Julier, C. EIF2AK3, encoding translation initiation factor 2-alpha kinase 3, is mutated in patients with Wolcott-Rallison syndrome. *Nat. Genet.* **25**, 406–409 (2000).
- 33 Sparsø, T., Andersen, G., Nielsen, T., Burgdorf, K. S., Gjesing, A. P., Nielsen, A. L. *et al*. The GCKR rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia* **51**, 70–75 (2008).
- 34 Wu, Y., Li, H., Loos, R. J., Yu, Z., Ye, X., Chen, L. *et al*. 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* **57**, 2834–2842 (2008).
- 35 Lango, H., Palmer, C. N., Morris, A. D., Zeggini, E., Hattersley, A. T., McCarthy, M. I., *et al*. The UK Type 2 Diabetes Genetics Consortium Assessing the combined impact of 18 common genetic variants of modest effect sizes on type 2 diabetes risk. *Diabetes* **57**, 3129–3135 (2008).
- 36 van Hoek, M., Dehgan, A., Witteman, J. C., van Duijn, C. M., Uitterlinden, A. G., Oostra, B. A. *et al*. Predicting type 2 diabetes based on polymorphisms from genome wide association studies: a population-based study. *Diabetes* **57**, 3122–3128 (2008).
- 37 Meigs, J. B., Shrader, P., Sullivan, L. M., McAteer, J. B., Fox, C. S., Dupuis, J. *et al*. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med.* **359**, 2208–2219 (2008).
- 38 Lyssenko, V., Jönsson, A., Almgren, P., Pulizzi, N., Isomaa, B., Tuomi, T. *et al*. Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N. Engl. J. Med.* **359**, 2220–2232 (2008).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)

Variants in *KCNQ1* are associated with susceptibility to type 2 diabetes mellitus

Kazuki Yasuda¹, Kazuaki Miyake², Yukio Horikawa³, Kazuo Hara⁴, Haruhiko Osawa⁵, Hiroto Furuta⁶, Yushi Hirota², Hiroyuki Mori², Anna Jonsson⁷, Yoshifumi Sato⁸, Kazuya Yamagata^{8,27}, Yoshinori Hinokio⁹, He-Yao Wang^{1,27}, Toshihito Tanahashi¹⁰, Naoto Nakamura¹¹, Yoshitomo Oka⁹, Naoko Iwasaki¹², Yasuhiko Iwamoto¹², Yuichiro Yamada^{13,27}, Yutaka Seino^{13,27}, Hiroshi Maegawa¹⁴, Atsunori Kashiwagi¹⁴, Jun Takeda³, Eiichi Maeda¹⁵, Hyoung Doo Shin¹⁶, Young Min Cho¹⁷, Kyong Soo Park¹⁷, Hong Kyu Lee¹⁷, Maggie C Y Ng¹⁸, Ronald C W Ma¹⁸, Wing-Yee So¹⁸, Juliana C N Chan¹⁸, Valeriya Lyssenko⁷, Tiinamajja Tuomi^{19,20}, Peter Nilsson²¹, Leif Groop^{7,19}, Naoyuki Kamatani²², Akihiro Sekine^{23,27}, Yusuke Nakamura²³, Ken Yamamoto²⁴, Teruhiko Yoshida²⁵, Katsushi Tokunaga²⁶, Mitsuo Itakura¹⁰, Hideichi Makino⁵, Kishio Nanjo⁶, Takashi Kadowaki⁴ & Masato Kasuga²

We carried out a multistage genome-wide association study of type 2 diabetes mellitus in Japanese individuals, with a total of 1,612 cases and 1,424 controls and 100,000 SNPs. The most significant association was obtained with SNPs in *KCNQ1*, and dense mapping within the gene revealed that rs2237892 in intron 15 showed the lowest *P* value (6.7×10^{-13} , odds ratio (OR) = 1.49). The association of *KCNQ1* with type 2 diabetes was replicated in populations of Korean, Chinese and European ancestry as well as in two independent Japanese populations, and meta-analysis with a total of 19,930 individuals (9,569 cases and 10,361 controls) yielded a *P* value of 1.7×10^{-42} (OR = 1.40; 95% CI = 1.34–1.47) for rs2237892. Among control subjects, the risk allele of this polymorphism was

associated with impairment of insulin secretion according to the homeostasis model assessment of β -cell function or the corrected insulin response. Our data thus implicate *KCNQ1* as a diabetes susceptibility gene in groups of different ancestries.

In Japan, the prevalence of type 2 diabetes mellitus is increasing rapidly, and more than 10% of individuals over 40 years of age are affected. Relatively few diabetic individuals in Japan are obese, and impairment of insulin secretion often develops before the onset of diabetes¹. As part of a national project designated the Millennium Genome Project in Japan, in 2002 we began a multistage genome-wide association study (GWAS) to identify disease-associated SNPs for type 2 diabetes mellitus using 100,000 SNPs from a collection of

¹Department of Metabolic Disorder, Research Institute, International Medical Center of Japan, Tokyo 162-8655, Japan. ²Division of Diabetes, Metabolism, and Endocrinology, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kobe 650-0017, Japan. ³Department of Diabetes and Endocrinology, Division of Molecule and Structure, Gifu University School of Medicine, Gifu 501-1194, Japan. ⁴Department of Metabolic Diseases, Graduate School of Medicine, University of Tokyo, Tokyo 113-8655, Japan. ⁵Department of Molecular and Genetic Medicine, Ehime University Graduate School of Medicine, Ehime 791-0295, Japan. ⁶First Department of Medicine, Wakayama Medical University, Wakayama 641-8509, Japan. ⁷Department of Clinical Sciences, Diabetes and Endocrinology, University Hospital Malmö, Lund University, S-205 02 Malmö, Sweden. ⁸Department of Metabolic Medicine, Graduate School of Medicine, Osaka University, Osaka 565-0871, Japan. ⁹Division of Molecular Metabolism and Diabetes, Tohoku University Graduate School of Medicine, Sendai 980-8574, Japan. ¹⁰Division of Genetic Information, Institute for Genome Research, University of Tokushima, Tokushima 770-8503, Japan. ¹¹Department of Endocrinology and Metabolism, Kyoto Prefectural University of Medicine, Graduate School of Medical Sciences, Kyoto 602-8566, Japan. ¹²Department of Medicine, Diabetes Center, Tokyo Women's Medical University, Tokyo 162-8666, Japan. ¹³Department of Diabetes and Clinical Nutrition, Kyoto University School of Medicine, Kyoto 606-8501, Japan. ¹⁴Division of Endocrinology and Metabolism, Department of Medicine, Shiga University of Medical Science, Shiga 520-2192, Japan. ¹⁵Clinical Genome Informatics Center, Kobe University Graduate School of Medicine, Kobe 650-0017, Japan. ¹⁶Department of Genetic Epidemiology, SNP Genetics Inc., Seoul 110-834, Korea. ¹⁷Department of Internal Medicine, Seoul National University College of Medicine, Seoul 110-744, Korea. ¹⁸Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Shatin, Hong Kong. ¹⁹Department of Medicine, Helsinki University Hospital, FIN-00300 Helsinki, Finland. ²⁰Folkhaelsan Research Center, FIN-00014 Helsinki, Finland. ²¹Department of Clinical Sciences, Medicine Research Unit, University Hospital Malmö, Lund University, S-205 02 Malmö, Sweden. ²²Division of Genomic Medicine, Department of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo 162-8666, Japan. ²³SNP Research Center, Institute of Physical and Chemical Research (RIKEN), Yokohama 230-0045, Japan. ²⁴Department of Molecular Genetics, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan. ²⁵Genetics Division, National Cancer Center Research Institute, Tokyo 104-0045, Japan. ²⁶Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan. ²⁷Present addresses: Department of Medical Biochemistry, Faculty of Medical and Pharmaceutical Sciences, Kumamoto University, Kumamoto 860-8556, Japan (K. Yamagata), Shanghai Institute of Materia Medica, Chinese Academy of Science, Shanghai 200031, China (H.-Y.W.), Department of Internal Medicine, Akita University School of Medicine, Akita 010-8543, Japan (Y.Y.), Kansai Electric Power Hospital, Osaka 553-0003, Japan (Y. Seino) and Genome Informatics, Center for Genomic Medicine, Graduate School of Medicine and Faculty of Medicine, Kyoto University, Kyoto 606-8501, Japan (A.S.). Correspondence should be addressed to M.K. (kasuga@med.kobe-u.ac.jp).

Received 3 March; accepted 6 June; published online 17 August 2008; doi:10.1038/ng.207

Table 1 Positive SNPs identified in the third screening

| dbSNP ID | Risk allele | Chr. | Gene | Panel 1 (187 cases) | | | Panel 2 (752 cases, 752 controls) | | | Panel 3 (672 cases, 672 controls) | | | | | | |
|------------|-------------|------|---------|---------------------|---------|------------------|-----------------------------------|---------|---------|-----------------------------------|------------------|----------------------|---------|-------------|------------------|----------------------|
| | | | | RAF(DM) | RAF(NC) | P value | Control | RAF(DM) | RAF(NC) | OR (95% CI) | P value | RAF(DM) | RAF(NC) | OR (95% CI) | P value | |
| rs151290 | C | 11 | KCNQ1 | 0.63 | 0.57 | 1.30 (1.03–1.65) | 0.027 | ODG | 0.62 | 0.55 | 1.34 (1.16–1.55) | 7.4×10^{-5} | 0.61 | 0.54 | 1.36 (1.16–1.58) | 1.1×10^{-4} |
| rs163184 | G | 11 | KCNQ1 | 0.51 | 0.43 | 1.33 (1.06–1.67) | 0.015 | JDC | 0.49 | 0.44 | 1.22 (1.06–1.41) | 0.0064 | 0.48 | 0.42 | 1.27 (1.09–1.48) | 0.0021 |
| rs2237895 | C | 11 | KCNQ1 | 0.45 | 0.35 | 1.53 (1.22–1.93) | 2.8×10^{-4} | JDC | 0.42 | 0.33 | 1.49 (1.28–1.73) | 1.4×10^{-7} | 0.42 | 0.33 | 1.45 (1.24–1.70) | 3.4×10^{-6} |
| rs2250402 | C | 15 | E/F2AK4 | 0.20 | 0.27 | 1.45 (1.09–1.93) | 0.011 | JDC | 0.24 | 0.21 | 1.20 (1.01–1.43) | 0.035 | 0.26 | 0.21 | 1.34 (1.11–1.60) | 0.0018 |
| rs2307027 | C | 12 | KRT4 | 0.14 | 0.22 | 1.68 (1.20–2.36) | 0.0024 | ODG | 0.20 | 0.17 | 1.23 (1.02–1.47) | 0.031 | 0.21 | 0.16 | 1.37 (1.12–1.67) | 0.0017 |
| rs3741872 | C | 12 | FAM60A | 0.29 | 0.23 | 1.37 (1.06–1.76) | 0.015 | ODG | 0.29 | 0.24 | 1.29 (1.09–1.52) | 0.0024 | 0.28 | 0.23 | 1.28 (1.07–1.52) | 0.0060 |
| rs74628 | G | 20 | ANGPT4 | 0.56 | 0.64 | 1.38 (1.09–1.74) | 0.0066 | ODG | 0.65 | 0.61 | 1.17 (1.01–1.36) | 0.037 | 0.64 | 0.59 | 1.28 (1.10–1.50) | 0.0018 |
| rs2233647 | G | 6 | SPDEF | 0.92 | 0.86 | 1.87 (1.07–3.27) | 0.026 | ODG | 0.88 | 0.86 | 1.24 (1.00–1.54) | 0.047 | 0.89 | 0.86 | 1.29 (1.02–1.62) | 0.033 |
| rs3785233* | C | 16 | A2BPI | 0.20 | 0.17 | 1.20 (0.90–1.61) | 0.22 | ODG | 0.19 | 0.16 | 1.25 (1.03–1.51) | 0.023 | 0.19 | 0.16 | 1.23 (1.01–1.50) | 0.039 |
| rs2075931 | A | 1 | | 0.71 | 0.64 | 1.37 (1.07–1.75) | 0.013 | ODG | 0.68 | 0.65 | 1.17 (1.01–1.37) | 0.038 | 0.68 | 0.64 | 1.18 (1.00–1.38) | 0.048 |

P values were calculated for allele data. For panel 1, two control groups (ODG, other disease group; JDC, Japanese database control) were used for association studies and the lower P values are listed. RAF(DM) and RAF(NC), risk allele frequencies in cases and controls, respectively. OR, odds ratio for risk allele.
*This SNP was selected for the second stage on the basis of the recessive model (OR = 2.59, CI = 1.20–5.58, P = 0.012).

standard Japanese SNPs² (which we refer to as the JSNP Genome Scan (JGS)), as part of the multi-disease collaborative genome scan (Supplementary Fig. 1 online).

Among 100,000 SNPs genotyped by multiplex PCR-based Invader analysis in the first stage of the study, 82,343 autosomal polymorphisms passed our typing quality control in 187 individuals with diabetes (Supplementary Table 1 online). We then carried out two separate association analyses to compare the 187 individuals with diabetes with two different control groups, which we considered as population controls: one to compare allele frequencies with reference data for 752 individuals representing the general Japanese population deposited in the JSNP database (referred to as the 'JSNP database control' (JDC)), and one to compare allele or genotype frequencies with those of the 752 individuals in the initial panels for the other four disease groups (Alzheimer's disease, gastric cancer, hypertension and asthma) of the national project (referred to as the 'other disease group' (ODG)). The combination of two types of association analysis resulted in the selection of 2,880 SNPs for the second stage of the study. An independent case-control panel (panel 2) was analyzed, and 201 positive SNPs ($P < 0.05$) were selected for the third stage (see Supplementary Table 2a online). Ten SNPs yielded a P value of < 0.05 at the third stage using another case-control panel (panel 3; Table 1 and Supplementary Table 2b). These SNPs showed variable P values in the first stage, suggestive of a limited power of the study design. The most significant association ($P = 3.4 \times 10^{-6}$) was obtained with rs2237895, which is located in intron 15 of *KCNQ1*. Another two SNPs (rs151290 and rs163184) were also located in the same intron, yielding P values of 1.1×10^{-4} and 0.0021, respectively. Panels 2 and 3 combined (panel 2+3) were analyzed for these 10 SNPs, yielding even lower P values for all the SNPs (Supplementary Table 2b). The genotype-based Cochran-Armitage trend test gave P values similar to those based on the allele data (Supplementary Table 2b).

We further analyzed *KCNQ1*, which was the only gene that yielded positive results according to the standard criterion (P value of $< 5 \times 10^{-7}$) recently proposed for GWAS³. The three SNPs of *KCNQ1* that passed the third scan (rs151290, rs163184 and rs2237895) were in moderate linkage disequilibrium (LD) with each other (Fig. 1). The SNP with the lowest P value, rs2237895, yielded D' and r^2 values of 0.54 and 0.12 with rs151290 and 0.83 and 0.46 with rs163184, respectively. We isolated 49 additional SNPs of *KCNQ1* from dbSNP of NCBI and typed them together with the three originally positive SNPs in panel 2+3 (Fig. 1). Among these 52 SNPs, rs2237892, which is also located in intron 15, showed the strongest association with diabetes ($P = 6.7 \times 10^{-13}$), with OR = 1.49 and 95% CI = 1.34–1.66; the P value for the trend test was 1.7×10^{-12} (Table 2). The D' and r^2 values for rs2237895 and rs2237892 were 0.95 and 0.30, respectively.

We also sequenced all the exons and the 47-kb genomic region corresponding to intron 15 of *KCNQ1* in 24 Japanese individuals and identified 212 variations, including three synonymous and two non-synonymous (P448R and G643S) polymorphisms (Supplementary Table 3a online). We then genotyped ten of the newly identified SNPs of intron 15 and the two nonsynonymous polymorphisms in panel 2+3. None of these SNPs showed a stronger association with diabetes than did rs2237892 (Fig. 1 and Supplementary Table 3b).

We next examined the possible association of *KCNQ1* with diabetes in several additional subject panels, including those of other ancestral groups, by genotyping rs2237892, rs2237895 and rs2074196, the three SNPs that showed the strongest association in the original study. Two independent Japanese panels revealed a strong association of these

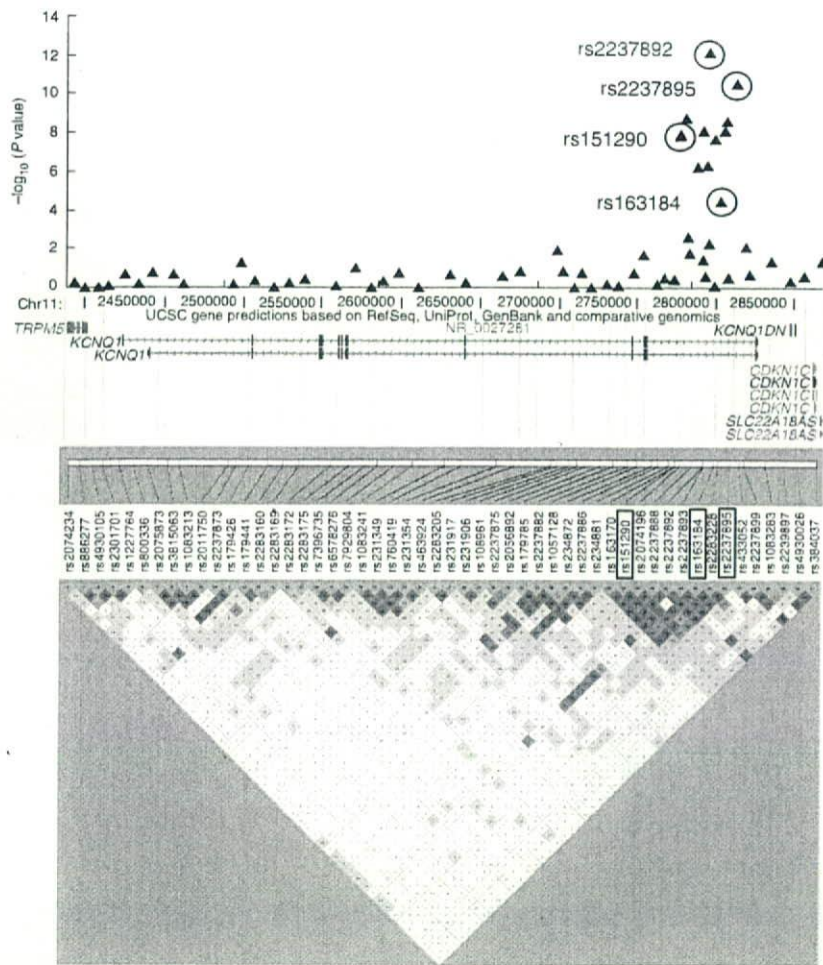


Figure 1 Dense mapping analysis of *KCNQ1*. The top panel shows the association $-\log_{10}(P \text{ value})$ in panel 2+3 for 64 SNPs of *KCNQ1*. The three blue circles represent the positive SNPs in the third screening. The red circle (rs2237892) indicates the SNP showing the most significant association with type 2 diabetes. The upper middle panel shows the physical position of *KCNQ1* and neighboring genes on chromosome 11 (UCSC Genome Browser). The lower middle panel shows the positions and rs numbers of the 52 previously identified SNPs. Blue rectangles indicate the positive SNPs in the third screening. The bottom panel shows a Haploview representation of LD (D') based on genotyping data from control subjects in panel 2+3 ($n = 1,424$).

allele of rs2237892 (CC) showed a significantly lower homeostasis model assessment of β -cell function (HOMA- β)⁴ than did those with the other genotypes (Supplementary Table 5 online). Among nondiabetic subjects of the Botnia prospective cohort (Supplementary Methods online), the corrected insulin response (CIR) at the follow-up visit was significantly lower for individuals with the CC genotype of rs2237892 than for those with the other two genotypes in both an additive and recessive model for this SNP ($P = 0.024$ and 0.010 , respectively; Supplementary Table 5). These results suggested that the risk allele of *KCNQ1* might contribute to diabetes susceptibility by impairing insulin secretion.

The multistage strategy for GWASs has an advantage in the effective elimination of a large number of false-positive results and has proved to be successful⁵. Indeed, we detected the association of several SNPs of *KCNQ1* with diabetes in the JGS, and this association was reproduced in two independent Japanese panels. *KCNQ1*, which encompasses 404 kb, is located at chromosome 11p15.5, not far from a candidate region at 11p13–p12 with suggestive evidence of linkage to type 2 diabetes in two independent studies of affected Japanese sibpairs^{6,7}. We also reproduced the association of *KCNQ1* with diabetes in Chinese and Korean panels, establishing *KCNQ1* as a diabetes susceptibility gene for populations of East Asian descent. We further showed the association to be significant in individuals of European descent. Given that *KCNQ1* was not implicated as a diabetes susceptibility gene in two recent GWASs with individuals of European descent^{8,9}, we examined SNPs of *KCNQ1* in the available datasets (Supplementary Fig. 3 and Supplementary Table 6a,b online). Within the LD block of *KCNQ1* that includes the SNPs associated with diabetes in Japanese, 11 SNPs in the WTCCC dataset⁸ and 9 SNPs in the DGI dataset⁹ had been typed, and none of them had been selected for further analysis. This apparent discrepancy may be due mainly to the allele frequencies of the causative SNPs (the minor allele frequency of rs2237892 was 0.28–0.41 and 0.05–0.07 in populations of East Asian and European descent, respectively). Indeed, in a recent meta-analysis of three GWASs (DGI, WTCCC and FUSION; see URLs section in Methods)¹⁰, the risk alleles of both rs2237892 and rs2074196 identified in the present study were associated with an increased risk of type 2 diabetes ($P = 0.01$ and 0.02 ,

polymorphisms with diabetes (Table 2 and Supplementary Table 4 online); rs2237892, for example, showed allelic P values of 9.6×10^{-10} and 6.9×10^{-10} in the replication 1 and 2 panels, respectively. The three Japanese panels (panel 2+3 and replication 1 and 2), which included a total of 4,378 cases and 4,412 controls, yielded an allelic P value of 2.8×10^{-29} and OR of 1.43 (95% CI = 1.34–1.52) for rs2237892. The association was also reproduced in the replication 3 (Chinese) and replication 4 (Korean) panels; the allelic P values for rs2237892 in these two panels were 1.3×10^{-8} and 1.7×10^{-5} , respectively (Table 2 and Supplementary Table 4). Meta-analysis of the Asian populations yielded a P value of 2.5×10^{-40} and OR of 1.42 (95% CI = 1.34–1.49) for rs2237892. We also examined rs2237892 and rs2074196 in the replication 5 panel (recruited from Sweden), with both SNPs showing a positive association ($P = 7.8 \times 10^{-4}$ and 0.017, respectively). With the inclusion of the replication 5 panel, meta-analysis with a total of 19,930 individuals (9,569 cases and 10,361 controls) yielded a P value of 1.7×10^{-42} and OR of 1.40 (95% CI = 1.34–1.47) for rs2237892 (Table 2 and Supplementary Fig. 2 online).

We next investigated the relation of rs2237892 to clinical phenotype. Among 1,424 individuals with diabetes in panel 2+3, no association was found between this SNP and clinical parameters such as body mass index (BMI) and the level of insulin resistance. Among the 948 control subjects in panel 2+3 whose fasting plasma glucose and insulin levels were available, homozygotes for the risk

Table 2 Association study results for SNPs in *KCNQ1* and type 2 diabetes

| SNP ID | Risk allele | Panel | RAF(DM) | RAF(NC) | P_{allele} | OR | 95% CI | P_{trend} | Meta-analysis OR (95% CI) | P value | |
|---|-------------|---|-------------|---------------------|---|-------------|-----------------------|---|---|-----------------------|---|
| rs2074196 | G | 2+3 (dense mapping) | 0.63 | 0.55 | 1.7×10^{-9} | 1.39 | 1.25 1.54 | 1.8×10^{-9} | 1.34 (1.26–1.42), $P = 4.8 \times 10^{-21}$ | | |
| | | Replication 1 (Japanese) | 0.61 | 0.54 | 1.4×10^{-7} | 1.32 | 1.19 1.46 | 2.1×10^{-7} | | | |
| | | Replication 2 (Japanese) | 0.62 | 0.55 | 4.7×10^{-7} | 1.31 | 1.18 1.46 | 6.2×10^{-7} | | | |
| | | All Japanese (4,378 cases, 4,412 controls) | 0.62 | 0.55 | 4.6×10^{-21} | 1.34 | 1.26 1.42 | 9.8×10^{-21} | | | |
| | | Replication 3 (Chinese) | 0.71 | 0.63 | 1.2×10^{-9} | 1.40 | 1.26 1.56 | 9.8×10^{-10} | | | |
| | | Replication 4 (Korean) | 0.66 | 0.58 | 3.0×10^{-5} | 1.39 | 1.19 1.62 | 2.1×10^{-5} | | | |
| | | All Asian (6,552 cases, 6,621 controls) | 0.64 | 0.57 | 9.9×10^{-32} | 1.35 | 1.28 1.42 | 2.1×10^{-31} | | | |
| | | Replication 5 (European) | 0.96 | 0.95 | 0.017 | 1.23 | 1.04 1.46 | 0.017 | | | |
| | | All | n.a. | n.a. | n.a. | n.a. | n.a. n.a. | n.a. | | | 1.35 (1.28–1.41), $P = 8.6 \times 10^{-34}$ |
| | | rs2237892 | C | 2+3 (dense mapping) | 0.69 | 0.60 | 6.7×10^{-13} | 1.49 | | | 1.34 1.66 |
| Replication 1 (Japanese) | 0.66 | | | 0.59 | 9.6×10^{-10} | 1.39 | 1.25 1.54 | 1.6×10^{-9} | | | |
| Replication 2 (Japanese) | 0.68 | | | 0.60 | 6.9×10^{-10} | 1.41 | 1.26 1.57 | 1.1×10^{-9} | | | |
| All Japanese (4,378 cases, 4,412 controls) | 0.68 | | | 0.59 | 2.8×10^{-29} | 1.43 | 1.34 1.52 | 1.7×10^{-28} | | | |
| Replication 3 (Chinese) | 0.72 | | | 0.65 | 1.3×10^{-8} | 1.38 | 1.24 1.55 | 4.2×10^{-9} | | | |
| Replication 4 (Korean) | 0.69 | | | 0.61 | 1.7×10^{-5} | 1.41 | 1.21 1.65 | 1.0×10^{-5} | | | |
| All Asian (6,552 cases, 6,621 controls) | 0.69 | | | 0.61 | 2.0×10^{-39} | 1.41 | 1.34 1.48 | 2.5×10^{-39} | | | |
| Replication 5 (European) | 0.95 | | | 0.93 | 7.8×10^{-4} | 1.29 | 1.11 1.50 | 7.2×10^{-4} | | | |
| All | n.a. | | | n.a. | n.a. | n.a. | n.a. n.a. | n.a. | 1.40 (1.34–1.47), $P = 1.7 \times 10^{-42}$ | | |
| rs2237895 | C | | | 2+3 (dense mapping) | 0.41 | 0.33 | 3.1×10^{-11} | 1.44 | 1.30 1.61 | 4.0×10^{-11} | 1.34 (1.26–1.43), $P = 1.4 \times 10^{-20}$ |
| | | Replication 1 (Japanese) | 0.38 | 0.33 | 4.5×10^{-5} | 1.25 | 1.12 1.38 | 4.7×10^{-5} | | | |
| | | Replication 2 (Japanese) | 0.41 | 0.34 | 5.8×10^{-8} | 1.35 | 1.21 1.50 | 5.5×10^{-8} | | | |
| | | All Japanese (4,378 cases, 4,412 controls) | 0.40 | 0.33 | 1.3×10^{-20} | 1.34 | 1.26 1.43 | 1.7×10^{-20} | | | |
| | | Replication 3 (Chinese) | 0.40 | 0.34 | 3.5×10^{-5} | 1.25 | 1.12 1.39 | 3.4×10^{-5} | | | |
| | | Replication 4 (Korean) | 0.35 | 0.30 | 3.2×10^{-3} | 1.27 | 1.08 1.49 | 2.7×10^{-3} | | | |
| | | All Asian (6,552 cases, 6,621 controls) | 0.39 | 0.33 | 2.7×10^{-25} | 1.31 | 1.24 1.38 | 2.7×10^{-25} | | | |
| | | Replication 5 (European) | n.a. | n.a. | n.a. | n.a. | n.a. n.a. | n.a. | 1.31 (1.25–1.38), $P = 6.1 \times 10^{-26}$ | | |
| | | All | n.a. | n.a. | n.a. | n.a. | n.a. n.a. | n.a. | | | |

RAF(DM) and RAF(NC), risk allele frequencies in cases and controls, respectively. P_{allele} values were calculated for allele data. OR, odds ratio for risk allele. P_{trend} values were calculated by the Cochran-Armitage trend test. Meta-analysis was performed by the Mantel-Haenszel method (fixed-effects models). n.a., not applicable.

respectively). These results provide further support for *KCNQ1* as a general susceptibility gene for diabetes, and they also highlight the need to extend GWAS to different populations.

Alternative splicing has been found to generate several variants of *KCNQ1* mRNA (see Accession codes section in Methods), but we do not know whether the identified candidate SNPs in intron 15 affect the splicing pattern of the primary transcript. Although neighboring genes seem to be located outside the LD block containing rs2237892, we are not able to exclude completely the possibility that the SNPs identified in the present study affect the expression of other causative genes. We did not find any microRNA harboring rs2237892 in the miRBase database.

KCNQ1 encodes the pore-forming subunit of a voltage-gated K^+ channel (KvLQT1) that is essential for the repolarization phase of the action potential in cardiac muscle¹¹. Mutations in this gene are associated with cardiac diseases such as hereditary long QT syndrome (Romano-Ward syndrome¹² and Jervell and Lange-Nielsen syndrome¹³) and familial atrial fibrillation¹⁴. This K^+ channel is also expressed in other tissues, including brain, adipose tissues and pancreas^{15,16}. The lower HOMA- β or CIR apparent for CC homozygotes of rs2237892 among Japanese and Europeans in the present study may reflect a functional role for this channel in

insulin-producing cells. We examined the abundance of *Kcnq1* mRNA by reverse transcription and real-time PCR analysis in the islets of 12-week-old diabetic KK-Ay mice, which manifested both hyperglycemia and hyperinsulinemia. The amount of the mRNA was significantly increased ($P = 0.0004$) by a factor of 1.6 compared with that in the islets of C57BL/6 control mice (data not shown). The *KCNQ1* protein was previously shown to be expressed in insulin-secreting INS-1 cells, and the *KCNQ1* blocker 293B was found to stimulate insulin secretion in the presence of tolbutamide¹⁷. It is also possible that fine-tuning of the membrane potential by this channel might modulate the survival of pancreatic β cells in the long term. Further studies are necessary to elucidate the precise mechanism by which the risk allele of *KCNQ1* confers susceptibility to diabetes.

We may have missed a substantial number of susceptibility genes in our screening, given that the strategy we adopted seven years ago lacks sufficient analytical power¹⁸ relative to that now achievable as a result of recent progress in genomic studies. The genomic coverage of the SNP set was not robust, in part because the IMS-JST Japanese SNP (JSNP) database was designed to focus on 'gene-centric' SNPs². Several comprehensive studies based on new platforms for GWASs have recently been described, with about ten genes being found to be

reproducibly associated with type 2 diabetes in individuals of European ancestry^{8,9,19–23}. None of these genes showed a positive association in our JGS typing data. Given that some of these genes were recently shown to confer susceptibility to diabetes in Japanese^{24–26}, the lack of association in our study might be due to the limited sample size of the first scan or to weak LD between the SNPs we used and the causative variants; actually, some genes were totally missed in our JGS (Supplementary Table 6c).

In summary, with a comprehensive multistage SNP association study in Japanese, we have identified *KCNQ1* as a previously unreported susceptibility gene as well as several other candidate genes for type 2 diabetes mellitus. Replication studies further confirmed the association of *KCNQ1* with diabetes in individuals of East Asian and European descent. Our findings may provide new insight into the pathophysiology of diabetes as well as a basis for the development of new therapeutic agents.

METHODS

Study participants. We assembled three independent subject panels for multistage genome-wide screening. Panel 1 consisted of 188 cases only, panel 2 of 752 cases and 752 controls and panel 3 of 672 cases and 672 controls. The inclusion criteria for diabetic patients were as follows: (i) age of disease onset of 40 to 55 years, (ii) maximum BMI of $<30 \text{ kg/m}^2$, (iii) insulin treatment not initiated until at least three years after diagnosis and (iv) absence of antibodies to glutamic acid decarboxylase. Most Japanese diabetic individuals have a BMI of $<30 \text{ kg/m}^2$, and we aimed to focus on the most common subtype of type 2 diabetes in Japan. The criteria for controls in panels 2 and 3 were as follows: (i) age of >60 years, (ii) no past history of diagnosis of diabetes and (iii) hemoglobin A_{1c} content of $<5.6\%$. The cases in the three panels and the controls in panels 2 and 3 were recruited at 11 core facilities located in various regions of Japan. Panels 2 and 3 were assembled simultaneously. Genomic DNA was extracted from peripheral blood by standard methods. We also obtained clinical information such as BMI, blood biochemistry (including plasma glucose and insulin levels) and family history of diabetes. The replication panels are described in Supplementary Methods. The clinical characteristics of subjects in each panel are summarized in Supplementary Table 1. The study protocol was approved by the local ethics committee of each institution, and written informed consent was obtained from all participants.

Study design. The general design and power for the multistage screening in the Millennium Genome Project (Supplementary Fig. 1), referred to as the JSNP Genome Scan (JGS), have been described previously¹⁸. In the first stage, 188 individuals with each disease (panel 1 for diabetes) were genotyped for 100,000 SNPs in the IMS-JST JSNP database (see URLs section below)². The coverage of the nucleotide sequences of the RefSeq NM exonic regions (as defined by 5' UTR + CDS (coding sequences) + 3' UTR) achieved by the JSNP 'gene-centric' genome-wide LD mapping is estimated to be $\sim 35\%$, if we assume an average extent of LD of 10 kb for each SNP with a minor allele frequency (MAF) of $>15\%$. We also previously evaluated the power of the first two stages of the JGS by a simulation experiment¹⁸. For example, this analysis would yield a sensitivity of $\sim 13\%$ for SNPs with an odds ratio of 1.5 and a disease-associated genotype frequency of 30%.

One subject did not yield a genotype call for any SNP in the first stage. We then carried out two separate association analyses to compare the 187 diabetic individuals with two different control groups, which we referred to as JDC and ODG, respectively. We did not detect significant population stratification among individuals of the initial panels of the five disease groups by standard methods such as genomic control²⁷ (inflation factor = 1.06 with 1,025 SNPs selected for genomic control analysis). The genotype-based analysis was done with dominant and recessive models. First, SNPs whose MAF was $>10\%$ in the database and which showed either a genotype OR of >1.5 or an allele OR of >1.3 in either association analysis were selected. If multiple SNPs in the same gene with positive association were in strong LD ($r^2 > 0.9$), only one SNP was chosen for the next step to avoid redundancy. A total of 2,880 SNPs for each disease was then selected for the second screening in order of P value; for

diabetes, 2,343 and 1,111 SNPs were selected by the association analyses with ODG and JDC, respectively, with 574 SNPs being selected by both analyses.

In the second stage, an independent case-control panel (panel 2) was analyzed, generating valid data for 2,827 SNPs after a quality check. Thirty-eight SNPs gave no results for all the samples in panel 2, whereas five and three SNPs yielded no data for all case or control samples, respectively, by multiplex PCR-based Invader analysis, and seven probes were not annotated on the updated human genome. The call rate for the 2,827 SNPs was 0.993. A total of 201 positive SNPs ($P < 0.05$) was selected for the third stage of the study on the basis of allelic data (Supplementary Table 2a). In the third stage, another case-control panel (panel 3) was typed; one SNP could not be typed by SSP-PCR-FCS analysis (see below) for any of the subjects in panel 3, with the call rate for the other 200 SNPs being 0.990. The ten positive SNPs ($P < 0.05$; Table 1) were also then analyzed in the combined panels 2 and 3 (panel 2+3, 1,424 cases and 1,424 controls). Panel 2 was genotyped again for these ten SNPs by SSP-PCR-FCS analysis, and the concordance rate with the Invader method used in the second screening was 0.992. The possibility of stratification in panels 2 and 3 was assessed by typing of 28 diabetes-unrelated SNPs followed by (i) comparison of allele and genotype frequencies by the χ^2 test, (ii) principal component analysis or (iii) STRUCTURE analysis (see URLs section below). None of these analyses showed evidence of stratification among cases and controls of panels 2 and 3 (data not shown).

The list of SNPs used for the initial screening and the allele and genotype frequency data for the first and the second stages of the JGS for the five diseases studied in the Millennium Genome Project of Japan, including diabetes, have been deposited in the Genome Medicine Database of Japan (GeMDBJ, see URLs section below).

Dense SNP mapping for *KCNQ1*. We first selected 49 additional SNPs of *KCNQ1* from the dbSNP database of NCBI, with an average interval of ~ 10 kbp, and typed these polymorphisms in panel 2+3 together with the three positive SNPs originally included in the JGS. We sequenced 24 control Japanese subjects for the gene, including all the exons and the putative promoter region (4 kbp upstream from the transcription start site), in order to comprehensively identify genetic variants in Japanese. We also sequenced the regions surrounding the positive SNPs of *KCNQ1*, spanning 47 kbp (intron 15). Ten of the SNPs identified in the 47-kbp region were selected on the basis of LD and MAF ($>10\%$). These 10 SNPs and the two identified nonsynonymous variants were genotyped in panel 2+3. A total of 64 SNPs was thus genotyped for *KCNQ1*, including 18 SNPs in the 35.6-kbp region between rs151290 and rs2237895, with an average interval of 2 kbp (see Supplementary Table 3b).

Typing methods. In the first and second stages of the study, genotyping was done by the multiplex PCR-based Invader assay (Third Wave Technologies) as previously described²⁸. In the third stage and for dense mapping, genome-wide amplified DOP degenerate oligonucleotide-primed (DOP)-PCR templates were genotyped by sequence-specific primer (SSP)-PCR analysis followed by fluorescence correlation spectroscopy (FCS)²⁹. Some SNPs included in dense mapping were therefore re-genotyped in panel 2 by the SSP-PCR-FCS method. Some SNPs were genotyped by real-time PCR analysis with TaqMan probes (Applied Biosystems). For replication panels, we applied either SSP-PCR-FCS or the TaqMan method.

Statistical analysis. In the first screening, we performed two case-control evaluations as described above. We examined allele or genotype (dominant or recessive models) data in 2×2 contingency tables for comparison with ODG, as well as allele data in 2×2 contingency tables for comparison with JDC (for which genotype data were not available). In the second and third screening and dense mapping, we analyzed allele data in 2×2 contingency tables by the χ^2 test. LD and haplotype analyses were done with Haploview 3.31 software³⁰. A P value of <0.05 was considered statistically significant. For ten positive SNPs in the JGS, rs2237892 and rs2074196, genotype-based analyses were also performed by the Cochran-Armitage trend test. Meta-analysis was done by the Mantel-Haenszel method (fixed-effects models) with the "meta" package of the R Project; the P values for heterogeneity among panels joined in the Mantel-Haenszel tests were all >0.05 .

URLs. Genome Medicine Database of Japan, <https://gembdb.nibio.go.jp/dgdb/>; DGI, WTCCC and FUSION, <http://www.well.ox.ac.uk/DIAGRAM/>; miRBase database, <http://microrna.sanger.ac.uk/sequences/>; IMS-JST JSNP database, <http://snp.ims.u-tokyo.ac.jp/>; STRUCTURE analysis, <http://pritch.bsd.uchicago.edu/software.html>.

Accession codes. GenBank: *KCNQ1* mRNA, NM_000218.2 and NM_181798.1.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank all the participants in the project; S. Sugano and S. Tsuji for support and helpful discussion throughout the project; H. Sakamoto, K. Yoshimura and N. Nishida for genotyping and quality control of the data; M. Yamaoka-Sageshima, K. Nagase, D. Suzuki and A. Berglund for technical assistance; and staff of Mitsui Knowledge Industry Inc. (Tokyo) for help with bioinformatics. This work was supported by a grant from the Program for Promotion of Fundamental Studies in Health Sciences of the Pharmaceuticals and Medical Devices Agency (PMDA) of Japan; a grant from the National Institute of Biomedical Innovation (NIBIO) of Japan; grants from the Ministry of Health, Labour, and Welfare of Japan; a Grant-in-Aid for Scientific Research on Priority Areas (C), "Medical Genome Science (Millennium Genome Project)," "Applied Genomics" and "Comprehensive Genomics" from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; and a grant from New Energy and Industrial Technology Development Organization (NEDO). The replication 2 study was supported by a grant from Cooperative Link of Unique Science and Technology for Economy Revitalization (CLUSTER, Tokushima, Japan). The Hong Kong diabetes case-control study was supported by the Hong Kong Research Grants Committee Central Allocation Scheme CUHK 1/04C. The Korean case-control study was supported by a grant from the Korea Health 21 R&D Project of the Ministry of Health and Welfare of the Republic of Korea (00-PJ3-PG6-GN07-001 to K.S.P.). The replication 5 study and Botnia prospective study were supported by Swedish Research Council (Linne grant), Sigrid Juselius Foundation, Folkhaelsan Research Foundation, European Foundation for the Study of Diabetes and Swedish Diabetes Research Foundation.

AUTHOR CONTRIBUTIONS

Principal investigators: K. Yasuda and M.K. Manuscript writing: K. Yasuda, K.M., Y. Horikawa and M.K. Diabetes project planning and design: K. Yasuda, K.M., Y. Hirota, H. Mori, T.Y. and M.K. Ascertainment of study subjects and general data analyses in Japan: K. Yasuda, K.M., Y. Horikawa, K.H., H.O., H.F., Y. Hirota, H. Mori, Y. Sato, K. Yamagata, Y. Hinokio, H.-Y.W., T. Tanahashi, N.N., Y.O., N.I., Y.I., Y.Y., Y. Seino, H. Maegawa, A.K., J.T., E.M., N.K., M.L., H. Makino, K.N., T.K. and M.K. Genotyping and sequencing analyses in Japan: K.M., Y. Horikawa, Y. Hirota, T. Tanahashi, A.S., Y.N., K. Yamamoto, T.Y., K.T. and M.I. Statistical analyses: K.M., Y. Horikawa, Y. Hirota, E.M., T.Y., K.T. and M.I. Genetic analyses in Korea: H.D.S., Y.M.C., K.S.P. and H.K.L. Genetic analyses in Hong Kong: M.C.Y.N., R.C.W.M., W.-Y.S. and J.C.N.C. Genetic analyses in Europe: A.J., V.L., T. Tuomi, P.N. and L.G. Millennium Genome Project Human Genome Variation Team Leader: Y.N. Millennium Genome Project Diabetes Subteam Leader: M.K.

Published online at <http://www.nature.com/naturegenetics/>

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

1. Fukushima, M., Suzuki, H. & Seino, Y. Insulin secretion capacity in the development from normal glucose tolerance to type 2 diabetes. *Diabetes Res. Clin. Pract.* **66**, S37-S43 (2004).

2. Haga, H., Yamada, R., Ohnishi, Y., Nakamura, Y. & Tanaka, T. Gene-based SNP discovery as part of the Japanese Millennium Genome Project: identification of 190,562 genetic variations in the human genome. *J. Hum. Genet.* **47**, 605-610 (2002).
3. NCI-NHGRI Working Group on Replication in Association Studies. Replicating genotype-phenotype associations. *Nature* **447**, 655-660 (2007).
4. Matthews, D.R. *et al.* Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28**, 412-419 (1985).
5. Ozaki, K. *et al.* Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat. Genet.* **32**, 650-654 (2002).
6. Mori, Y. *et al.* Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate loci on 7p and 11p. *Diabetes* **51**, 1247-1255 (2002).
7. Nawata, H. *et al.* Genome-wide linkage analysis of type 2 diabetes mellitus reconfirms the susceptibility locus on 11p13-p12 in Japanese. *J. Hum. Genet.* **49**, 629-634 (2004).
8. Zeggini, E. *et al.* Replication of genome-wide association signals in U.K. samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336-1340 (2007).
9. Saxena, R. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331-1336 (2007).
10. Zeggini, E. *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638-645 (2008).
11. Barhanin, J. *et al.* K(v)LQT1 and IsK (minK) proteins associate to form the I(Ks) cardiac potassium current. *Nature* **384**, 78-80 (1996).
12. Wang, Q. *et al.* Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nat. Genet.* **12**, 17-23 (1996).
13. Neyroud, N. *et al.* A novel mutation in the potassium channel gene KVLQT1 causes the Jervell and Lange-Nielsen cardioauditory syndrome. *Nat. Genet.* **15**, 186-189 (1997).
14. Chen, Y.H. *et al.* *KCNQ1* gain-of-function mutation in familial atrial fibrillation. *Science* **299**, 251-254 (2003).
15. Demolombe, S. *et al.* Differential expression of KVLQT1 and its regulator IsK in mouse epithelia. *Am. J. Physiol. Cell Physiol.* **280**, C359-C372 (2001).
16. Chouabe, C. *et al.* Properties of KVLQT1 K⁺ channel mutations in Romano-Ward and Jervell and Lange-Nielsen inherited cardiac arrhythmias. *EMBO J.* **16**, 5472-5479 (1997).
17. Ullrich, S. *et al.* Effects of I_{Ks} channel inhibitors in insulin-secreting INS-1 cells. *Pflugers Arch.* **451**, 428-436 (2005).
18. Sato, Y. *et al.* Designing a multistage, SNP-based, genome screen for common diseases. *J. Hum. Genet.* **49**, 669-676 (2004).
19. Grant, S.F. *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320-323 (2006).
20. Scott, L.J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341-1345 (2007).
21. Sladek, R. *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881-885 (2007).
22. Steinthorsdottir, V. *et al.* A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat. Genet.* **39**, 770-775 (2007).
23. Frayling, T.M. *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889-894 (2007).
24. Miyake, K. *et al.* Association of TCF7L2 polymorphisms with susceptibility to type 2 diabetes in 4,087 Japanese subjects. *J. Hum. Genet.* **53**, 174-180 (2008).
25. Horikoshi, M. *et al.* Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabetologia* **50**, 2461-2466 (2007).
26. Horikawa, Y. *et al.* Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. *J. Clin. Endocrinol. Metab.* (in the press).
27. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-1004 (1999).
28. Ohnishi, Y. *et al.* A high-throughput SNP typing system for genome-wide association studies. *J. Hum. Genet.* **46**, 471-477 (2001).
29. Bannai, M. *et al.* Single-nucleotide-polymorphism genotyping for whole-genome-amplified samples using automated fluorescence correlation spectroscopy. *Anal. Biochem.* **327**, 215-221 (2004).
30. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265 (2005).

SNPs in *KCNQ1* are associated with susceptibility to type 2 diabetes in East Asian and European populations

Hirofumi Unoki¹, Atsushi Takahashi², Takahisa Kawaguchi³, Kazuo Hara⁴, Momoko Horikoshi⁴, Gitte Andersen⁵, Daniel P K Ng⁶, Johan Holmkvist⁵, Knut Borch-Johnsen^{5,7,8}, Torben Jørgensen⁷, Anneli Sandbæk⁹, Torsten Lauritzen⁹, Torben Hansen⁵, Siti Nurbaya⁶, Tatsuhiko Tsunoda³, Michiaki Kubo¹⁰, Tetsuya Babazono¹¹, Hiroshi Hirose¹², Matsuhiko Hayashi¹³, Yasuhiko Iwamoto¹¹, Atsunori Kashiwagi¹⁴, Kohei Kaku¹⁵, Ryuzo Kawamori¹⁶, E Shyong Tai¹⁷, Oluf Pedersen^{5,7}, Naoyuki Kamatani², Takashi Kadowaki⁴, Ryuichi Kikkawa¹⁴, Yusuke Nakamura¹⁸ & Shiro Maeda¹

We conducted a genome-wide association study using 207,097 SNP markers in Japanese individuals with type 2 diabetes and unrelated controls, and identified *KCNQ1* (potassium voltage-gated channel, KQT-like subfamily, member 1) to be a strong candidate for conferring susceptibility to type 2 diabetes. We detected consistent association of a SNP in *KCNQ1* (rs2283228) with the disease in several independent case-control studies (additive model $P = 3.1 \times 10^{-12}$; OR = 1.26, 95% CI = 1.18–1.34). Several other SNPs in the same linkage disequilibrium (LD) block were strongly associated with type 2 diabetes (additive model: rs2237895, $P = 7.3 \times 10^{-9}$; OR = 1.32, 95% CI = 1.20–1.45, rs2237897, $P = 6.8 \times 10^{-13}$; OR = 1.41, 95% CI = 1.29–1.55). The association of these SNPs with type 2 diabetes was replicated in samples from Singaporean (additive model: rs2237895, $P = 8.5 \times 10^{-3}$; OR = 1.14, rs2237897, $P = 2.4 \times 10^{-4}$; OR = 1.22) and Danish populations (additive model: rs2237895, $P = 3.7 \times 10^{-11}$; OR = 1.24, rs2237897, $P = 1.2 \times 10^{-4}$; OR = 1.36).

Type 2 diabetes affects more than 200 million individuals worldwide, and its prevalence is continuously increasing in many countries, including Japan. Although the precise mechanisms underlying the development and progression of type 2 diabetes have not been fully elucidated, a combination of multiple genetic and environmental factors is considered to contribute to the pathogenesis of the disease¹.

Recent development of SNP typing technology and collation of information regarding LD in the human genome have facilitated genome-wide association studies for investigating genes associated with disease susceptibility across the entire human genome. Recently, genome-wide association studies conducted by several independent research groups in Europe and the United States have identified multiple loci associated with susceptibility to type 2 diabetes, including *TCF7L2* (transcription factor 7-like 2), which had been originally identified by a large-scale association mapping prompted by prior evidence of linkage in that area², *SLC30A8* (solute carrier family 30 member 8), *HHEX* (haematopoietically expressed homeobox), *CDKAL1* (CDK5 regulatory subunit associated protein 1-like 1), *CDKN2A/B* (cyclin-dependent kinase inhibitor 2A/B) and *IGF2BP2* (insulin-like growth factor 2 mRNA-binding protein 2)^{3–7}.

Although these are considered to be loci convincingly associated with susceptibility to type 2 diabetes in populations of European descent, other genes related to susceptibility to the disease are probably still unidentified, particularly those for populations of other ancestries. In order to uncover genetic variants that increase the risk of type 2 diabetes, we conducted a genome-wide association study in Japanese individuals with type 2 diabetes and unrelated controls. We first genotyped 268,068 SNPs, which covered approximately 56% of common SNPs in the Japanese, in 194 individuals with type 2 diabetes and diabetic retinopathy (case 1) and in 1,558 controls (control 1) collected in the BioBank Japan. We compared the allele frequencies of 207,097 successfully genotyped SNPs and selected the

¹Laboratory for Endocrinology and Metabolism, ²Laboratory for Statistical Analysis and ³Laboratory for Medical Informatics, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa 230-0045, Japan. ⁴Department of Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan. ⁵Steno Diabetes Center, Niels Steensens Vej 2, DK-2820 Gentofte, Copenhagen, Denmark. ⁶Department of Community, Occupational and Family Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore. ⁷Research Centre for Prevention and Health, Glostrup University Hospital, 2600 Glostrup, Denmark. ⁸Faculty of Health Science and ⁹Department of General Practice, University of Aarhus, 8000 Aarhus, Denmark. ¹⁰Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa 230-0045, Japan. ¹¹Diabetes Center, Tokyo Women's Medical University, Tokyo 162-8666, Japan. ¹²Health Center and ¹³Department of Internal Medicine, Keio University School of Medicine, Tokyo 160-8582, Japan. ¹⁴Department of Medicine, Shiga University of Medical Science, Otsu, Shiga 520-2192, Japan. ¹⁵Division of Endocrinology and Metabolism, Department of Internal Medicine, Kawasaki Medical School, Kurashiki, Okayama 701-0192, Japan. ¹⁶Department of Medicine, Metabolism and Endocrinology, School of Medicine, Juntendo University, Tokyo 113-8421, Japan. ¹⁷Department of Endocrinology, Singapore General Hospital, Singapore 169608, Singapore. ¹⁸Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan. Correspondence should be addressed to S.M. (smaeda@src.riken.jp).

Received 3 March; accepted 10 June; published online 17 August 2008; doi:10.1038/ng.208

Table 1 Association results for variants analyzed in Japanese individuals with type 2 diabetes and control individuals

| | Stage | Type 2 diabetes | | Control | | P value | OR (95% CI) |
|------------------------------------|----------|-----------------|-------------------|---------|-----------------|-----------------------|------------------|
| | | Freq. | 11/12/22 | Freq. | 11/12/22 | | |
| rs13259803 | 1 | 0.02 | 182/9/0 | 0.06 | 761/87/8 | 0.006 | 0.38 (0.19–0.75) |
| <i>TRPA1</i> , Chr. 8 (C>T) | 2 | 0.11 | 1,083/261/21 | 0.07 | 1,014/139/10 | 3.0×10^{-7} | 1.70 (1.39–2.08) |
| | 3 | 0.12 | 2,526/694/33 | 0.11 | 1,052/256/22 | 0.58 | 1.04 (0.90–1.20) |
| | Combined | 0.23 | 4,139/956/54 | 0.24 | 2,116/491/32 | 0.006 | 0.76 (0.62–0.92) |
| rs612774 | 1 | 0.36 | 80/86/26 | 0.31 | 396/358/77 | 0.05 | 1.26 (1.0–1.59) |
| Chr. 7 (T>G) | 2 | 0.34 | 607/588/163 | 0.27 | 458/323/66 | 3.1×10^{-6} | 1.38 (1.21–1.58) |
| | 3 | 0.31 | 1,526/1,398/317 | 0.32 | 620/570/138 | 0.64 | 0.98 (0.89–1.08) |
| | Combined | 0.35 | 2,711/2,574/500 | 0.29 | 1,126/811/272 | 0.0002 | 1.17 (1.04–1.31) |
| rs4712524 | 1 | 0.46 | 67/77/50 | 0.42 | 293/413/152 | 0.18 | 1.29 (1.16–1.44) |
| <i>CDKAL1</i> , Chr. 6 (A>G) | 2 | 0.47 | 384/655/313 | 0.41 | 444/601/218 | 5.7×10^{-6} | 1.29 (1.16–1.44) |
| | 3 | 0.46 | 980/1,525/744 | 0.42 | 439/653/228 | 0.0002 | 1.19 (1.09–1.31) |
| | Combined | 0.47 | 1,431/2,257/1,107 | 0.42 | 1,176/1,667/598 | 3.3×10^{-10} | 1.22 (1.15–1.31) |
| rs9295475 | 1 | 0.45 | 65/81/47 | 0.41 | 301/409/148 | 0.13 | 1.19 (0.95–1.49) |
| <i>CDKAL1</i> , Chr. 6 (A>G) | 2 | 0.47 | 396/648/302 | 0.40 | 451/604/209 | 1.2×10^{-5} | 1.28 (1.15–1.43) |
| | 3 | 0.46 | 995/1,527/728 | 0.41 | 451/644/225 | 0.0001 | 1.20 (1.09–1.31) |
| | Combined | 0.46 | 1,456/1,986/1,077 | 0.41 | 1,203/1,657/582 | 4.1×10^{-9} | 1.22 (1.14–1.30) |
| rs6769511 | 1 | 0.40 | 67/80/32 | 0.32 | 404/344/98 | 0.003 | 1.44 (1.14–1.82) |
| <i>IGF2BP2</i> , Chr. 3 (T>C) | 2 | 0.38 | 533/619/204 | 0.32 | 528/535/110 | 2.4×10^{-5} | 1.28 (1.14–1.44) |
| | 3 | 0.36 | 1,332/1,504/423 | 0.32 | 621/566/145 | 0.0004 | 1.19 (1.08–1.31) |
| | Combined | 0.37 | 1,932/2,203/659 | 0.32 | 1,553/1,445/353 | 1.3×10^{-9} | 1.23 (1.15–1.31) |
| rs9460546 | 1 | 0.46 | 67/77/50 | 0.42 | 293/414/152 | 0.18 | 1.17 (0.94–1.46) |
| <i>CDKAL1</i> , Chr. 6 (T>G) | 2 | 0.47 | 391/650/307 | 0.41 | 444/604/218 | 3.0×10^{-5} | 1.27 (1.13–1.41) |
| | 3 | 0.47 | 963/1,526/744 | 0.42 | 437/652/228 | 0.0001 | 1.20 (1.10–1.32) |
| | Combined | 0.47 | 1,421/2,253/1,101 | 0.42 | 1,174/1,670/598 | 3.4×10^{-10} | 1.23 (1.15–1.30) |
| rs4376068 | 1 | 0.37 | 82/81/30 | 0.29 | 440/343/76 | 0.003 | 1.42 (1.13–1.79) |
| <i>IGF2BP2</i> , Chr. 3 (A>C) | 2 | 0.35 | 591/607/168 | 0.29 | 584/498/93 | 3.8×10^{-5} | 1.28 (1.14–1.45) |
| | 3 | 0.33 | 1,466/1,417/371 | 0.30 | 658/552/117 | 0.001 | 1.18 (1.07–1.30) |
| | Combined | 0.34 | 2,139/2,108/569 | 0.29 | 1,682/1,393/286 | 2.1×10^{-9} | 1.23 (1.15–1.32) |
| rs2283228 | 1 | 0.65 | 80/84/24 | 0.58 | 266/436/134 | 0.01 | 1.34 (1.06–1.70) |
| <i>KCNQ1</i> , Chr. 11 (A>C) | 2 | 0.64 | 557/638/167 | 0.59 | 370/547/178 | 5.7×10^{-5} | 1.26 (1.13–1.42) |
| | 3 | 0.64 | 1,379/1,592/449 | 0.59 | 426/590/217 | 7.0×10^{-6} | 1.24 (1.13–1.36) |
| | Combined | 0.64 | 2,016/2,314/640 | 0.58 | 1,062/1,573/529 | 3.1×10^{-12} | 1.26 (1.18–1.34) |
| rs10836097 | 1 | 0.11 | 147/33/4 | 0.07 | 544/76/5 | 0.01 | 1.70 (1.15–2.51) |
| <i>C11ORF41</i> , Chr. 11 (A>G) | 2 | 0.09 | 1,084/222/9 | 0.06 | 792/105/0 | 5.3×10^{-5} | 1.62 (1.27–2.05) |
| | 3 | 0.09 | 2,658/534/37 | 0.09 | 1,041/213/11 | 0.86 | 1.01 (0.87–1.19) |

The results include data from stage 1, stage 2 and replication study (3) for 9 SNPs. Risk allele frequency (Freq.) and genotype counts in individuals with type 2 diabetes and control subjects are shown. 11, homozygous of major allele; 12, heterozygous; 22, homozygous of minor allele. P values are for the additive model. Risk allele is denoted in boldface.

8,323 SNPs showing the lowest P values. We then attempted to genotype these 8,323 SNPs in 1,367 individuals with type 2 diabetes and diabetic retinopathy (case 2) and for 1,266 controls (control 2) (stage 2), and successfully obtained data for 6,731 SNPs (the P value distribution in the second test is shown in **Supplementary Fig. 1a** online). The results of principal component analysis⁸ in the stage 1 and 2 samples and HapMap samples revealed that there was no evidence for population stratification between the case and control groups throughout the present tests (**Supplementary Fig. 1b,c**). We selected the 9 SNP loci showing P values < 0.0001 (additive model in stage 2, **Table 1**) and genotyped a third set of cases and controls comprising 3,557 Japanese individuals with type 2 diabetes (cases 3,4,5) and 1,352 controls (controls 3,4). We evaluated the differences in the population structure among these three sets of case and two sets of control groups by Wright's F test. As the results indicated that there was no difference in the population structure among these groups (**Supplementary Table 1b** online), we combined these populations for the third test of case-control study. The third set of analysis identified the significant associations for six SNPs (**Table 1**), including the *CDKAL1* locus at 6p22.3 (rs4712524, rs9295475 and rs9460546), the *IGF2BP2* locus at 3q27.2 (rs6769511 and rs4376068) and the *KCNQ1*

locus at 11p15.5 (rs2283228). The remaining three SNPs (rs13259803, rs612774 and rs10836097) had P values of > 0.05 in the third test and were not further examined. *CDKAL1* and *IGF2BP2* were previously reported as susceptibility genes for type 2 diabetes in the Japanese population⁹. Therefore, we focused on the *KCNQ1* locus, which was highly associated with type 2 diabetes.

Subsequently, we carried out LD mapping using 572 SNPs in the neighboring region of SNP rs2283228 and found that the LD block including this SNP extended approximately 40 kb (30 kb upstream and 10 kb downstream of this SNP; **Fig. 1a,b** and **Supplementary Fig. 2a,b** online). Therefore, we considered the critical region for susceptibility to type 2 diabetes to be present within this 40-kb LD block that corresponds to a part of intron 15 of *KCNQ1*. To further confirm the association, we genotyped the third set of 3,557 cases and 1,352 controls for the 32 additional SNPs within the block and found several SNPs to be significantly associated with type 2 diabetes in these populations ($P < 0.0001$, **Fig. 1c** and **Supplementary Table 2a,b** online). Among them, six SNPs showed stronger associations with type 2 diabetes than rs2283228 (rs2237897, $P = 6.8 \times 10^{-13}$; rs2237896, $P = 3.9 \times 10^{-12}$; rs2299620, $P = 8.3 \times 10^{-11}$; rs2237895, $P = 7.3 \times 10^{-9}$; rs2237892, $P = 1.3 \times 10^{-7}$; rs163171, $P = 4.8 \times 10^{-7}$;

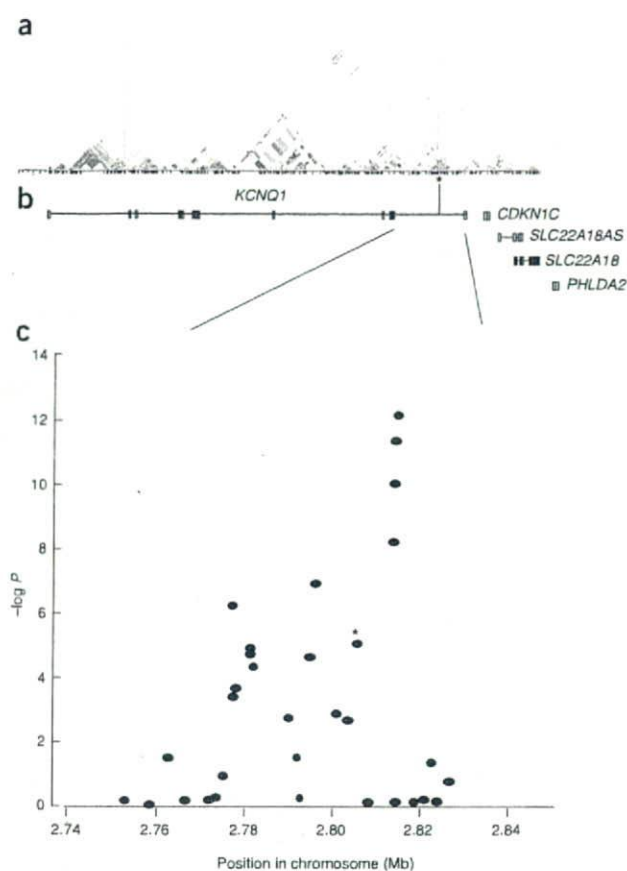


Fig. 1 and Supplementary Table 2a). The genotype distributions of these SNPs in the control subjects were within the Hardy-Weinberg equilibrium (HWE) ($P \geq 0.01$, Supplementary Table 2b). With 207,097 tests performed in the initial genome-wide association study, these associations, except that for rs163171, remained significant after Bonferroni adjustment¹⁰. We also found a significant interactive effect between rs2237897 and rs234844 by means of stepwise logistic regression analysis¹¹ (effect of addition; $P = 0.0015$, Supplementary Table 3 online).

To further validate the association of the *KCNQ1* variants with type 2 diabetes, we examined the association of the above-mentioned SNPs with type 2 diabetes in populations of different ancestries. As shown in Table 2, rs2237895 and rs2237897 are significantly associated with the disease in the Singaporean population of East Asian descent and the Danish population of European descent (Table 2 and Supplementary Table 4 online), although there are significant differences in the allele frequencies within this locus between the East Asian and Danish populations (Supplementary Table 5a and Supplementary Fig. 3a–d online). A combined analysis of these three populations by using a meta-analysis after testing for heterogeneity gave P values of less than 1×10^{-16} for both rs2237895 and rs2237897 (Supplementary Table 5b).

To elucidate the possible involvement of *KCNQ1* in the pathogenesis of type 2 diabetes, we further examined the expression profile of the *KCNQ1* gene. Using reverse transcription polymerase chain reaction (RT-PCR), we found expression of *KCNQ1* in the human heart and pancreas and, to a lesser extent, in the placenta, lung, liver and kidney (Supplementary Fig. 4 online). In addition, we also detected

Figure 1 Schematic view of the association of type 2 diabetes with the variants in the *KCNQ1* region. (a) Pairwise correlation structure at a 0.5-Mb interval (2.4–2.9 Mb on chromosome 11, NCBI Build 36.2). The plot includes pairwise r^2 values from the HapMap release 22 for the JPT population. (b) Exon-intron structure of *KCNQ1* (exons shown as vertical bars). The asterisk indicates the SNP rs2283228 at intron 15. (c) Results of case-control association study for type 2 diabetes in 3,557 Japanese individuals with type 2 diabetes and 1,352 control individuals. The \log_{10} -transformed P values for an additive model are plotted on the y axis.

the expression of *KCNQ1* in a cultured insulin-secreting cell line, RIN-5F (Supplementary Fig. 4).

KCNQ1, located on 11p15.5, encodes the pore-forming α subunit of the $I_{Ks}K^+$ channel, which is expressed mainly in the heart and, to a lesser extent, in the inner ear, stomach, intestine, liver and kidney¹². The $I_{Ks}K^+$ channel enables a K^+ current after electrical depolarization of the cell membrane in the heart^{13,14}. Moreover, mutations in *KCNQ1* have been reported to cause long QT syndrome through the loss of function of the slowly activating K^+ channel in the heart. *KCNQ1* is also a known causative gene for Jervell and Lange-Nielsen syndrome, which is an autosomal recessive trait characterized by hearing loss and cardiac conduction abnormalities¹³. *KCNQ1*-knockout mice have been reported to show cardiac dysfunctions such as prolonged QT interval, hypochlorhydria, hypergastrinaemia, gastric hyperplasia and vacuolation of the parietal cells; the latter four effects suggest that *KCNQ1* also plays some roles in normal gastric acid secretion^{12,15}. In the kidney, *KCNQ1* has been shown to enhance the Na^+ -dependent transport in the proximal tubule, presumably by contributing to a favorable electrical gradient for Na^+ uptake¹⁶. *KCNQ1* is also expressed in the pancreas, including the pancreatic islets; however, the contribution of the *KCNQ1*-encoded protein to the molecular pathogenesis of type 2 diabetes remains to be elucidated. Subjects with a loss-of-function mutation in the gene or homozygous *KCNQ1* null mice have not been reported to show hyperglycaemia or glucose intolerance. Hence, we would expect that increased *KCNQ1* expression in the pancreatic β cells is probably responsible for the development of type 2 diabetes. However, a possibility is still present that a gene near *KCNQ1*, particularly *CDKN1C*, might be a true causal gene (Fig. 1b and Supplementary Table 6 online); further studies are necessary to elucidate the underlying molecular mechanism of the gene and gene polymorphisms for conferring susceptibility to type 2 diabetes.

Our present study has missed most of susceptibility loci previously identified in populations with European descent, except for the *CDKAL1* and *IGF2BP2* loci. As we showed that previously identified loci associated with type 2 diabetes, such as *TCF7L2*, *CDKN2A/B*, *HHEX* and *KCNJ11*, were successfully replicated in our populations⁹, a lack of study power to detect true association in the present stage 1 study seems to be responsible for the discrepancy between the present results and the results in the genome-wide association studies of European-descent populations. However, it is notable that the *KCNQ1* locus was missed in previous genome-wide association studies conducted by groups in Europe and the United States, as those studies are considered to be superior to our study in terms of study power or coverage rate of the human genome. This can probably be explained by the fact that ancestry-related differences were present in the allele frequencies or in the pattern of LD, and that uncovered regions still existed. Therefore, this emphasizes that studies extending genome-wide association analysis beyond the focus of the first rounds of studies, including a genome-wide association meta-analysis recently reported as well as the present study, are needed to search for additional genomic loci.

Table 2 Association results for rs2237897, rs2237895 and rs2283228 in three case-control groups from Japan, Singapore and Denmark

| | Population | Type 2 diabetes | | Control | | P | OR (95% CI) |
|------------------|-------------|-----------------|-----------------|---------|-----------------|-----------------------|------------------|
| | | Freq. | 11/12/22 | Freq. | 11/12/22 | | |
| rs2237897 | | | | | | | |
| (C>T) | Japanese | 0.68 | 1,615/1,574/333 | 0.61 | 476/646/198 | 6.8×10^{-13} | 1.41 (1.29–1.55) |
| | Singaporean | 0.69 | 680/621/132 | 0.65 | 749/747/239 | 2.4×10^{-4} | 1.22 (1.09–1.35) |
| | Dane | 0.97 | 3,643/245/3 | 0.96 | 4,473/405/10 | 1.2×10^{-4} | 1.36 (1.16–1.60) |
| rs2237895 | | | | | | | |
| (A>C) | Japanese | 0.41 | 1,225/1,633/577 | 0.34 | 567/580/152 | 7.3×10^{-9} | 1.32 (1.20–1.45) |
| | Singaporean | 0.38 | 574/709/214 | 0.35 | 801/848/232 | 8.5×10^{-3} | 1.14 (1.03–1.26) |
| | Dane | 0.45 | 1,082/1,854/750 | 0.40 | 1,578/2,203/714 | 3.7×10^{-11} | 1.24 (1.16–1.32) |
| rs2283228 | | | | | | | |
| (A>C) | Japanese | 0.64 | 1,379/1,592/449 | 0.59 | 426/590/217 | 7.0×10^{-6} | 1.24 (1.13–1.36) |
| | Singaporean | 0.64 | 612/667/194 | 0.63 | 692/812/243 | 0.27 | 1.06 (0.96–1.17) |
| | Dane | 0.94 | 3,413/469/15 | 0.92 | 4,154/706/29 | 3.1×10^{-4} | 1.24 (1.10–1.39) |

Risk allele frequency (Freq.) and genotype counts in individuals with type 2 diabetes and control subjects. 11, homozygous of major allele; 12, heterozygous; 22, homozygous of minor allele. P values for additive model. Risk allele is denoted in boldface.

In conclusion, we have replicated the association of *CDKAL1* and *IGF2BP2* with type 2 diabetes in a genome-wide association study. In addition, we have identified a previously unreported locus associated with type 2 diabetes. SNPs in the *KCNQ1* gene were significantly associated with type 2 diabetes in populations of both East Asian and European descent.

METHODS

Study participants. BioBank Japan: For the genome-wide association study, we selected case-control samples (cases 1 and 2, controls 1 and 2) from the subjects enrolled in the BioBank Japan. The subjects were recruited from several medical institutes in Japan, including Fukujiji Hospital, Iizuka Hospital, Iwate Medical University School of Medicine, Juntendo University, National Hospital Organization Osaka National Hospital, Nihon University, Nippon Medical School, Osaka Medical Center for Cancer and Cardiovascular Diseases, Shiga University of Medical Science, The Cancer Institute Hospital of Japanese Foundation for Cancer Research, Tokushukai Hospitals and Tokyo Metropolitan Geriatric Hospital. We selected type 2 diabetes cases from individuals registered as having type 2 diabetes and diabetic retinopathy (case 1, $n = 194$; case 2, $n = 1,367$, **Supplementary Table 1a**). Control groups comprised 1,558 (control 1) and 1,266 individuals (control 2) registered as individuals with diseases other than type 2 diabetes, including bronchial asthma, myocardial infarction, breast cancer, Basedow's disease, cerebral infarction, cerebral aneurism, osteoporosis, heart failure, unstable angina, pollinosis, arteriosclerosis obliterance, emphysema, atopic dermatitis, stomach cancer or liver cirrhosis.

RIKEN case-control study: DNA samples were obtained from peripheral blood samples of 1,630 individuals with type 2 diabetes recruited from the outpatient clinic of the Shiga University of Medical Science, Kawasaki Medical School (case 3: 978 men and 652 women; age, 61.5 ± 11.6 years; duration of diabetes, 10.9 ± 9.3 years; HbA1c, $7.4 \pm 1.6\%$; fasting plasma glucose (FPG), 9.1 ± 3.5 mmol/l; and BMI, 23.7 ± 3.9 kg/m² (all values are expressed as mean \pm s.d.), **Supplementary Table 1a**). Diabetes was diagnosed according to the World Health Organization (WHO) criteria. Type 2 diabetes is clinically defined as disease with a gradual adult onset. Subjects who tested positive for antibodies to glutamic acid decarboxylase (GAD) and those diagnosed with a mitochondrial disease (mitochondrial myopathy, encephalopathy, lactic acidosis or stroke-like episodes (MELAS)) or maturity-onset diabetes of young (MODY) were not included in the case group. Among the 1,630 subjects, 31 had already been enrolled in the Japanese Millennium Project for the study of type 2 diabetes and were therefore excluded from the present analysis. We also examined 1,064 control subjects who were enrolled from an annual health check-up conducted at either the Juntendo University or Keio University

(control 3: 638 men and 426 women; age, 45.5 ± 9.5 years; HbA1c, $4.7 \pm 0.4\%$; fasting plasma glucose, 5.1 ± 0.5 mmol/l; and BMI, 22.9 ± 3.0 kg/m² (all values are expressed as mean \pm s.d.), **Supplementary Table 1a**). We further examined 1,304 independent subjects with type 2 diabetes and diabetic retinopathy who regularly attended outpatient clinics at the Shiga University of Medical Science, Tokyo Women's Medical University, Juntendo University, Kawasaki Medical School, Iwate Medical University, Toride Kyodo Hospital, Kawai Clinic, Osaka City General Hospital, Chiba Tokusuyukai Hospital and Osaka Rosai Hospital (case 4: 774 men and 530 women; age, 61.1 ± 11.2 years; duration of diabetes, 17.4 ± 9.6 years; HbA1c, $7.6 \pm 2.8\%$; fasting plasma glucose, 8.1 ± 2.8 mmol/l; and BMI, 23.7 ± 3.6 kg/m² (all values are expressed as mean \pm s.d.), **Supplementary Table 1a**).

The University of Tokyo study: In this study, we enrolled 654 subjects with type 2 diabetes who visited the outpatient clinic of the Department of Metabolic Diseases, Graduate School of Medicine, The University of Tokyo (Tokyo, Japan) or the outpatient clinic of the Hiroshima Atomic Bomb Casualty Council Health Management Center (Hiroshima, Japan) (case 5: 392 men and 262 women; age, 67.2 ± 8.4 years; duration of diabetes, 13.9 ± 10.2 years; HbA1c, $7.5 \pm 3.6\%$; FPG, 9.5 ± 3.0 mmol/l; and BMI, 23.9 ± 3.3 kg/m² (all values are expressed as mean \pm s.d.), **Supplementary Table 1a**). Exclusion criteria for cases were individuals positive for antibody to GAD or those with diabetes resulting from (i) liver dysfunction, (ii) steroids and other drugs that might raise glucose levels, (iii) malignancy or (iv) monogenic disorder known to cause diabetes. Control subjects were enrolled from an annual health check-up conducted at the Hiroshima Atomic Bomb Casualty Council Health Management Center (control 4: 112 men and 176 women; age, 71.6 ± 6.2 years; HbA1c, $5.2 \pm 0.2\%$; FPG, 5.0 ± 1.1 mmol/l; and BMI 22.6 ± 3.2 kg/m² (all values are expressed as mean \pm s.d.), **Supplementary Table 1a**). All subjects were not enrolled in the Japanese Millennium Project for the study of type 2 diabetes.

Singapore study: We selected Chinese individuals with type 2 diabetes from the Singapore Diabetes Cohort Study (SDCS). All individuals diagnosed with type 2 diabetes at primary care facilities of the National Healthcare Group Polyclinics (NHGPs) in Singapore had been invited to participate in the SDCS since 2004. Of the individuals approached, 91% agreed to participate in the study and formed a part of our SDCS case group. At the time of this genetic study, DNA samples from 1,498 Chinese individuals were available for analysis. The research protocol for SDCS was approved by both the National University of Singapore Institutional Review Board (NUS-012) and the National Healthcare Group Domain-Specific Review Board (C/05/118).

Nondiabetic controls were identified on the basis of the results of oral glucose tolerance test (OGTT) performed on the subjects belonging to the 1998 Singapore National Health Survey (NHS98), as detailed elsewhere¹⁷. The survey

LETTERS

protocol was based on the WHO-recommended model for field surveys of diabetes and other noncommunicable diseases and the WHO MONICA protocol for population surveys. A total of 1,881 nondiabetic Chinese individuals with 2-h OGTT values of <7.8 mmol/l were selected as controls. The research protocol for NHS98 was approved by the Singapore General Hospital Institutional Review Board (#54/2001).

Danish study: The Danish study population comprised the population-based Inter99 sample of middle-aged people sampled at Research Centre for Prevention and Health¹⁸, individuals with type 2 diabetes sampled through the outpatient clinic at Steno Diabetes Center, a population-based group of middle-aged glucose-tolerant subjects recruited from Steno Diabetes Center, and the ADDITION study group sampled through Department of General Practice at University of Aarhus¹⁹. Detailed characteristics of study populations have been previously described²⁰. In total, 4,085 individuals with type 2 diabetes and 5,302 glucose-tolerant control subjects were genotyped for the rs2283228, rs2237895 and rs2237897 polymorphisms using Taqman allelic discrimination (KBioscience). Discordance between 966 random duplicate samples was <0.5% and the genotyping success rate was >96%. All genotype groups obeyed Hardy-Weinberg equilibrium (all $P > 0.4$).

All the participants provided written informed consent. The protocol was approved by the ethics committee of the Institute of Physical and Chemical Research (RIKEN) and of each participating institution.

SNP genotyping. Using the standard protocols, we extracted the genomic DNA from the peripheral blood leukocytes. In the first stage, we genotyped 268,068 SNPs from autosomal chromosomes; these SNPs were selected as the tagging SNPs for the Japanese population from the JSNP²¹ or HapMap database²² by using high-density oligonucleotide arrays (Perlegen Sciences) as described previously. SNPs having call rate >90% and no extreme departure from HWE ($P \geq 10^{-6}$) were passed and used for the association study. In the second stage, genotyping was conducted using Affymetrix GeneChip SNP arrays. We carried out genotyping for the replication study using the multiplex-PCR invader assay²³ or the TaqMan assay, as described previously. The success rates of the multiplex-PCR invader assay and the TaqMan assay were >95%, and >99% agreement existed between the results of genotyping and direct sequencing.

Statistical analysis. Statistical methods for determining the associations and calculating the LD coefficients (r^2) have been described previously²⁴. We performed HWE test according to a method previously described²⁵. The cut-off value for the HWE test in the control groups was 0.000001 for the first stage and 0.01 for the second stage; the SNPs with P values less than the cut-off values of the HWE test were excluded from the analysis. We analyzed the differences between the case and control groups with regard to the genotype distribution or allele frequency in the genome-wide screening (first or second stage) by Fisher's exact test using dominant, recessive and allelic models with autosomal SNPs. The association of the candidate loci with type 2 diabetes in the replication study was evaluated by Armitage test for trend using an additive model, as described previously³. We applied stepwise logistic regression analysis to investigate the combinatorial effects of multiple SNPs¹¹.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank all participating doctors and staff from collaborating institutes for providing DNA samples. We also thank the technical staff of the Laboratory for Endocrinology and Metabolism at the Center for Genomic Medicine for providing technical assistance. This work was partly supported by a grant from the Leading Project of Ministry of Education, Culture, Sports, Science and Technology Japan. The Danish study was supported by Lundbeck Foundation Centre of Applied Medical Genomics in Personalized Disease Prediction, Prevention and Care (LUCAMP), the European Union (EUGENE2, grant no. LSHM-CT-2004-512013) and the Danish Medical Research Council. Funding for Singapore arm of the study was supported by the National Medical Research Council (NMRC/1115/2007). We thank the staff of National Healthcare Group Polyclinics and National University of Singapore for their contributions towards the establishment of the Singapore Diabetes Cohort Study.

AUTHOR CONTRIBUTIONS

Principal investigators: S.M., H.U., Y.N. BioBank project planning and design: Y.N. Collected subjects and participated in the diagnostic evaluations: K.H., M. Horikoshi, T.B., H.H., M. Hayashi, Y.I., A.K., K.K., R. Kawamori, T. Kadowaki, R. Kikkawa. Oversaw a genotyping: M.K. Wrote the paper: S.M., H.U., Y.N., O.P. Statistical analyses: A.T., T. Kawaguchi, T.T., N.K. Principal investigator for Singapore study: D.P.K.N. Singapore study: S.N., E.S.T. Principal investigator for Danish study: O.P. Danish study: G.A., J.H., K.B.-J., T.J., A.S., T.L., T.H.

Published online at <http://www.nature.com/naturegenetics/>

Reprints and permissions information is available online at <http://npng.nature.com/reprintsandpermissions/>

- O'Rahilly, S., Barroso, I. & Wareham, N.J. Genetic factors in type 2 diabetes: the end of the beginning? *Science* **307**, 370–373 (2005).
- Grant, S.F. *et al.* Variant of transcription factor 7-like 2 (*TCF7L2*) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320–323 (2006).
- Sladek, R. *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881–885 (2007).
- Zeggini, E. *et al.* Replication of genome-wide association signals in U.K. samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336–1341 (2007).
- Scott, L.J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).
- Saxena, R. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
- Steinthorsdottir, V. *et al.* A variant in *CDKAL1* influences insulin response and risk of type 2 diabetes. *Nat. Genet.* **39**, 770–775 (2007).
- Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
- Omori, S. *et al.* Association of *CDKAL1*, *IGF2BP2*, *CDKN2A/B*, *HHEX*, *SLC30A8* and *KCNJ11* with susceptibility to type 2 diabetes in a Japanese population. *Diabetes* **57**, 791–795 (2008).
- Skol, A.D., Scott, L.J., Abecasis, G.R. & Boehnke, M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat. Genet.* **38**, 209–213 (2006).
- Cordell, H.J. & Clayton, D.G. A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. *Am. J. Hum. Genet.* **70**, 124–141 (2002).
- Lee, M.P. *et al.* Targeted disruption of the *Kvlqt1* gene causes deafness and gastric hyperplasia in mice. *J. Clin. Invest.* **106**, 1447–1455 (2000).
- Neyroud, N. *et al.* A novel mutation in the potassium channel gene *KVLQT1* causes the Jervell and Lange-Nielsen cardioauditory syndrome. *Nat. Genet.* **15**, 186–189 (1997).
- Moss, A.J. *et al.* Clinical aspects of type-1 long-QT syndrome by location, coding type, and biophysical function of mutations involving the *KCNQ1* gene. *Circulation* **115**, 2481–2489 (2007).
- Casimiro, M.C. *et al.* Targeted disruption of the *Kcnq1* gene produces a mouse model of Jervell and Lange-Nielsen Syndrome. *Proc. Natl. Acad. Sci. USA* **98**, 2526–2531 (2001).
- Vallon, V. *et al.* *KCNQ1*-dependent transport in renal and gastrointestinal epithelia. *Proc. Natl. Acad. Sci. USA* **102**, 17864–17869 (2005).
- Cutler, J., Tan, B.Y., & Chew, S.K. Levels of cardiovascular disease risk factors in Singapore following a national intervention programme. *Bull. World Health Organ.* **79**, 908–915 (2001).
- Jørgensen, T. *et al.* A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99 (1). *Eur. J. Cardiovasc. Prev. Rehabil.* **10**, 377–386 (2003).
- Lauritzen, T. *et al.* Anglo-Danish-Dutch study of intensive treatment in people with screen detected diabetes in primary care. The ADDITION study: proposed trial of the cost-effectiveness of an intensive multifactorial intervention on morbidity and mortality among people with type 2 diabetes detected by screening. *Int. J. Obes.* **24**, (Suppl. 3), S6–S11 (2000).
- Sparsø, T. *et al.* The *GCKR* rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia* **51**, 70–75 (2008).
- Haga, H., Yamada, R., Ohnishi, Y., Nakamura, Y. & Tanaka, T. Gene-based SNP discovery as part of the Japanese Millennium Genome Project: identification of 190,562 genetic variations in the human genome. *J. Hum. Genet.* **47**, 605–610 (2002).
- The international HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
- Ohnishi, Y. *et al.* A high-throughput SNP typing system for genome-wide association studies. *J. Hum. Genet.* **46**, 471–477 (2001).
- Devlin, B. & Risch, N. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29**, 311–322 (1995).
- Wigginton, J.E., Cutler, D.J. & Abecasis, G.R. A note on exact tests of Hardy-Weinberg equilibrium. *Am. J. Hum. Genet.* **76**, 887–893 (2005).

Experimental Studies

Genetic Variations of Mrf-2/Arid5b Confer Risk of Coronary Atherosclerosis in the Japanese Population

Guoqin WANG,¹ MD, Masafumi WATANABE,¹ MD, Yasushi IMAI,¹ MD, Kazuo HARA,² MD, Ichiro MANABE,¹ MD, Koji MAEMURA,¹ MD, Momoko HORIKOSHI,² MD, Takahide KOHRO,³ MD, Eisuke AMIYA,¹ MD, Takao SUGIYAMA,⁴ MD, Takeshi FUJITA,⁵ MD, Takashi KADOWAKI,² MD, Tsutomu YAMAZAKI,⁶ MD, and Ryozo NAGAI,¹ MD

SUMMARY

A phenotypic change of smooth muscle cells (SMCs) is considered to be critical in the pathogenesis of atherosclerotic lesions such as coronary artery disease (CAD). Mrf-2/ARID5B, a member of the AT-rich interaction domain family of transcription factors, is highly expressed in the cardiovascular system and is believed to play essential roles in the phenotypic change of SMCs through its regulation of SMC differentiation. In addition, recent studies on gene-engineered mice suggested that this transcriptional factor is involved in obesity and adipogenesis, which are critical aspects for the pathogenesis of atherosclerosis. Thus, we hypothesized that genetic variations of the Mrf-2 gene might be associated with susceptibility to CAD.

We investigated 11 common genetic variations of Mrf-2 to determine whether they were associated with susceptibility to CAD in 475 CAD subjects and 310 control subjects. The prevalence of homozygotes for the minor allele G of SNP4 (rs2893880) and minor allele G of SNP6 (rs7087507) were significantly more frequent in the control subjects than in patients with CAD ($P = 0.0002$, rs2893880, $P = 0.0058$, rs7087507). Four nearby SNPs (SNP4 to SNP7) (rs2893880, rs10740055, rs7087507 and rs10761600) showed almost complete linkage disequilibrium, and haplotype analysis revealed that the haplotype G (rs2893880)-C (rs10740055)-G (rs7087507)-A (rs10761600) was also significantly negatively associated with susceptibility to CAD ($P = 0.049$). Moreover, these negative disease associations still existed after logistic regression analysis was taken into account to eliminate confounding conventional coronary risk factors.

The results implicate possible disease relevance of the polymorphisms in the Mrf-2 gene with susceptibility to CAD. However, a larger scale prospective study is needed to clarify these findings. (Int Heart J 2008; 49: 313-327)

From the Departments of ¹ Cardiovascular Medicine, ² Diabetes and Metabolism, ³ Translational Research for Healthcare and Clinical Science, Graduate School of Medicine, the University of Tokyo, ⁴ Institute for Adult Diseases, Asahi Life Foundation, ⁵ Hitachi, Ltd., Central Research Laboratory, and ⁶ Department of Clinical Bioinformatics, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan.

Address for correspondence: Masafumi Watanabe, MD, Department of Cardiovascular Medicine, The University of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.

This study was supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and a research grant from the National Institute of Biomedical Innovation, Japan. The study was also supported in part by Special Coordination Funds for Promoting Science and Technology by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

Received for publication March 31, 2008.

Revised and accepted May 1, 2008.

Key words: Coronary artery disease, Atherosclerosis, Mrf-2, ARID5B, Polymorphism, Genetic risk factors

CORONARY artery disease (CAD) is a multifactorial disease and is the leading cause of mortality in industrialized countries. A multitude of environmental influences and genetic factors contribute to the development and outcome of CAD.¹⁾ Single nucleotide polymorphisms (SNPs) are the most common genetic variations in the genome and some are reported to influence susceptibility to atherosclerosis, plaque destabilization, and thrombosis, all of which are fundamental aspects of CAD.^{2,3)}

Vascular smooth muscle is a critical cellular constituent of the blood vessel wall. Phenotypic changes in smooth muscle cells (SMCs) are considered to be critical in the pathogenesis of atherosclerotic lesions such as coronary artery disease (CAD). Our group previously investigated the mechanism of a phenotypic change by establishing an *in vitro* system in which a pluripotent neural crest cell line (MONC-1) can be induced to differentiate into SMCs. By using a differential mRNA display, we isolated a partial cDNA fragment of Mrf-2 (modulator recognition factor) as one of the genes robustly induced during the differentiation progress. Our group was the first to clone the full length of Mrf-2 and reveal the existence of at least two isoforms (α and β) which differ in the N-terminus but share a common DNA-binding domain.⁴⁾

Mrf-2, which is also called as ARID5B or Desrt, is a member of the AT-rich interaction domain (ARID) family of transcription factors. The ARID defines a highly conserved sequence specific DNA binding domain and has been identified in proteins of yeast, *Drosophila*, and mammals.⁵⁻⁷⁾ ARID-encoding genes are involved in a variety of biological processes including embryonic development, tissue specific gene expression, proliferation control, and chromatin remodeling.⁷⁻¹²⁾

We reported Mrf-2 as a nuclear protein, and its mRNA was highly expressed in the aorta, heart and lung.⁴⁾ We also demonstrated that Mrf-2 might also bind to the CArG box, a *cis*-acting element, which was reported to regulate the expression of many smooth muscle cell specific genes, including smooth muscle cell α -actin and SM 22 α . In addition, overexpression of Mrf-2 in 3T3 fibroblast cells induced smooth muscle marker genes and retarded cell proliferation. These results suggested that Mrf-2 might be a potent key regulator for SMC phenotypic change via its regulation of SMC differentiation.

On the other hand, it was reported that targeted disruption of the Mrf-2 gene in mice caused growth retardation and significant reductions in lipid accumulation and weight gain in postnatal and adult life.^{8,13,14)} These observations sug-

gested that Mrf-2 might also play important roles in adipogenesis and obesity, which are also critical aspects for the pathogenesis of atherosclerosis.

Taking together, we therefore hypothesized genetic variations in the Mrf-2 gene may predispose humans to CAD. The aim of the present study was to confirm whether genetic variations of the Mrf-2 gene are associated with susceptibility to CAD.

METHODS

Subject characteristics: This study was approved by the institutional ethics committee of the University of Tokyo, and written informed consent was obtained from all participants.

To examine the associations between Mrf-2 SNPs and susceptibility to CAD, we recruited 475 CAD patients (372 men and 103 women, aged 65.3 ± 8.6 years old) and 310 control subjects (173 men and 137 women, aged 68.1 ± 9.8 years old) as the study population. The CAD patients were consecutively enrolled from the Department of Cardiovascular Medicine of the University of Tokyo Hospital, and the control subjects were recruited consecutively from the Institute for Adult Diseases Asahi Life Foundation, which is located close to the University of Tokyo Hospital. All subjects were of Japanese ancestry and were not first- or second-degree relatives. All female participants were postmenopausal.

The patients with CAD were enrolled from October 1999 to March 2002 and were diagnosed when at least one of the 3 major coronary vessels had more than 50% narrowing on the basis of coronary angiography. The control subjects were recruited in 2000 based on having a normal ECG pattern and having no medical history of coronary artery disease or stroke in the monthly follow-up or health check-up program at the outpatient clinic of the Institute for Adult Diseases Asahi Life Foundation Hospital, Tokyo. The control subjects who had two or more coronary risk factors (DM, hypertension (HT), hypercholesterolemia, smoking) or had a history of chest pain were further evaluated using the Master or Treadmill exercise test, and those who had an ECG abnormality in the additional testing were excluded.

At the point of enrollment, the baseline profiles of both the CAD patients and control subjects were collected, and past medical history and smoking habits were obtained from all study participants. HT was diagnosed according to World Health Organization criteria. Hypercholesterolemia was diagnosed when total plasma cholesterol levels were > 220 mg/dL or when a subject was already being treated with lipid-lowering medication. DM in the CAD group and control group was diagnosed based on HbA1c $> 6.5\%$ or FBS ≥ 126 mg/dL. Fasting venous blood samples were drawn for biochemical and genetic analysis.

Genomic DNA extraction and genotyping of SNPs used in the association study: Venous blood samples were collected in tubes containing Na₂EDTA and applied to genomic DNA extracting columns (Genomix kit, Talent) according to the manufacturer's protocol.

The SNPs were identified based on the sequences reported in the JSNPs (Japanese Single Nucleotide Polymorphisms) database (http://snp.ims.u-tokyo.ac.jp/index_ja.html) or the NCBI GeneBank database (<http://www.ncbi.nih.gov/index.html>). No exon SNP of Mrf-2 was reported in the above databases for the Japanese population, therefore, 11 intron SNPs which covered the whole length of the Mrf-2 gene were almost equally selected from the public databases to detect the associations between genetic variations of Mrf-2 and CAD.

Two JSNPs (nomenclatures of SNP8, SNP11 represent JSNPs 025551 and 025550, respectively) as well as exon 3, exon 4 and their exon-intron boundaries were genotyped by direct sequencing as described in detail previously.¹⁵⁾ Briefly, PCR was performed under standard conditions and the sequencing reactions were performed using a BigDye terminator kit (Applied Biosystems, Foster City, CA), and resolved using an ABI 3700 automated DNA sequencer (Applied Biosystems). The results were integrated using a Sequencher (Gene Codes Corporation, Ann Arbor, MI) and individual SNPs were manually genotyped. Ambiguous base callings were eliminated from further analysis. The primer sequences are available from the authors.

Meanwhile, nine SNPs of the Mrf-2 gene from the public GeneBank database (nomenclatures of SNP1, SNP2, SNP3, SNP4, SNP5, SNP6, SNP7, SNP9 and SNP10 represent rs6415872, rs10821929, rs7901348, rs2893880, rs10740055, rs7087507, rs10761600, rs12357548 and rs10761604, respectively) were selected and genotyped using a MassARRAY system. When using the MassARRAY system, primer extension products were analyzed by chip based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) following PCR amplification as described in detail previously.¹⁶⁾

Statistical analysis: Continuous parameters such as age, body mass index (BMI), total cholesterol (T-chol), high density cholesterol (HDL), and triglycerides (TG) are presented as the mean \pm SD and were compared using Student's *t* test, while characteristics such as sex, HT, hyperlipidemia (HL), DM, and smoking are given as proportions of the entire samples and were compared using the chi-square (χ^2) test. The genotype and allele association for each SNP and CAD were tested using the χ^2 test with 2 and 1 degree (s) of freedom. The odds ratios and 95% CI, with adjustment for classical risk factors of CAD, were calculated by multiple logistic regression analysis. The differences among more than 2 average values were statistically tested using one-way ANOVA. The statistical analyses, except for the linkage disequilibrium and haplotype estimation, were performed using

SPSS 13.0 or Statview v4.0 software.

The statistical significance level was corrected for multiple testing, and we took potential overestimation into account when evaluating the most statistically significant value among those tested under different models. A two-tailed value of $P < 0.05$ was considered to be significant unless otherwise indicated.

Haplotype analysis: To examine the pair-wise linkage disequilibrium (LD) structure, linkage disequilibrium coefficients (D') and r^2 between the SNPs and haplotype frequencies were estimated via the method of maximum likelihood from two-locus genotype data using the EM algorithm under the assumption of Hardy-Weinberg equilibrium.^{17,18)} For the estimation of haplotype frequencies, we selected one of the SNPs as a tagging SNP from every set of SNPs with $r^2 > 0.80$. All haplotypes were jointly tested for association with disease status by performing a $2 \times n \chi^2$ test of independence in a permutation procedure, where " n " indicates the number of haplotypes with a frequency > 0.2 .

The above calculations were carried out using SNPalyze v5.1 standard software (Dynacom, Yokohama, Japan). We considered a $P < 0.05$ to be statistically significant.

RESULTS

Baseline characteristics of the CAD and control populations: The baseline characteristics of the CAD patients and the control subjects are shown in Table I.

Table I. Clinical Characteristics of the Study Subjects

| | CAD subjects (<i>n</i> = 475) | Control subjects (<i>n</i> = 310) | <i>P</i> |
|-------------------|-----------------------------------|---------------------------------------|-----------|
| Males (%) | 372 (78.3%) | 173 (55.8%) | < 0.0001* |
| Age | 65.3 ± 8.6 | 68.1 ± 9.8 | < 0.0001* |
| BMI | 24.1 ± 3.2 | 23.5 ± 4.1 | < 0.0001* |
| Hypertension | 399 (84.0%) | 249 (80.3%) | < 0.016* |
| Hyperlipidemia | 320 (67.3%) | 128 (41.3%) | < 0.0001* |
| T-chol (mg/dL) | 191.6 ± 31.0 | 212.9 ± 47.9 | |
| HDL (mg/dL) | 50.6 ± 14.9 | 54.4 ± 14.7 | |
| TG (mg/dL) | 183.6 ± 230.4 | 136.4 ± 94.0 | |
| Diabetes mellitus | 180 (37.9%) | 42 (13.5%) | < 0.0001* |
| HbA1c (%) | 5.9 ± 1.2 | 6.0 ± 0.6 | |
| Cr (mg/dL) | 0.93 ± 0.68 | 0.88 ± 0.19 | |
| Smoking | 350 (73.7%) | 105 (33.9%) | < 0.0001* |

Continuous parameters such as age, BMI, T-chol, HDL, and TG are presented as the mean ± SD and were compared using Student's *t* test. Characteristics such as sex, HT, HL, DM, and smoking are presented as proportions of the entire sample and were compared using the χ^2 test. A $P < 0.05$ (*) was considered to indicate statistical significance.

Baseline characteristics including age of the CAD patients were calculated at the point of enrollment when patients underwent the coronary angiography. For the control subjects, ages and other clinical characteristics were adopted at the enrollment point in 2000. As indicated, the patients in the CAD group had more conventional risk factors such as being male, smoking, HT, hypercholesterolemia, and DM in comparison to the control subjects. The average age of the control subjects was older than that of the patients with CAD.

Associations between Mrf-2 SNPs and CAD: We identified a total of 17 candidate SNPs composing at least 4 haplotype blocks which covered almost the full length

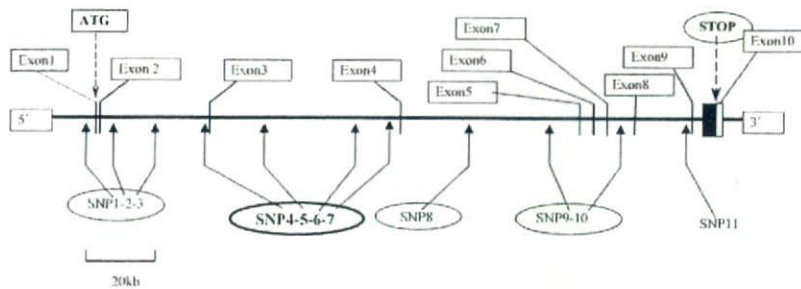


Figure 1. Gene structure and haplotype blocks of Mrf-2. Exons differing from <100 bp to >1000 bp are indicated by lines or closed boxes. Each circle indicates a haplotype block which represents a cluster of SNPs with $D' > 0.8$. SNP1-11 represents the 11 SNPs in database, as shown below:

SNP1: rs6415872 SNP2: rs10821929 SNP3: rs7901348
SNP4: rs2893880 SNP5: rs10740055 SNP6: rs7087507
SNP7: rs10761600 SNP8: JSNP025551 SNP9: rs12357548
SNP10: rs10761604 SNP11: JSNP025550

Table II. Association Between Allele Frequencies of Mrf-2 SNPs and CAD

| | Allele | Allele Frequency | | | P |
|------|--------|------------------|-------------|--------------|--------|
| | | Case | Control | Overall | |
| SNP4 | C | 640 (67.7%) | 389 (62.9%) | 1029 (65.8%) | 0.055 |
| | G | 306 (32.3%) | 229 (37.1%) | 535 (34.2%) | |
| SNP5 | A | 470 (49.5%) | 303 (49.0%) | 773 (49.3%) | 0.86 |
| | C | 480 (50.5%) | 315 (51.0%) | 795 (50.7%) | |
| SNP6 | A | 630 (66.3%) | 380 (61.5%) | 1010 (64.4%) | 0.041* |
| | G | 320 (33.7%) | 238 (38.5%) | 558 (35.6%) | |
| SNP7 | T | 385 (40.5%) | 233 (37.8%) | 618 (39.5%) | 0.29 |
| | A | 565 (59.5%) | 383 (62.2%) | 948 (60.5%) | |

A total of 475 CAD subjects and 310 control subjects were observed. Values are given as numbers of alleles (portions of the whole observed subjects x 2).

*indicates statistical significance.

of the Mrf-2 gene as shown in Figure 1. All polymorphisms were in accordance with Hardy-Weinberg equilibrium both in the CAD and control groups with each minor allele frequency > 5%.

The allele and genotypic frequencies of each SNP are shown in Table II and Table III. The distributions of G-allele in SNP4 (rs2893880) and G-allele in SNP6 (rs7087507) were more frequent in the control group than in the CAD group, and

Table III. Association Between Genotype Frequencies of Mrf-2 SNPs and CAD

| | Genotype | Genotypic Frequency | | <i>P</i> |
|------|----------|---------------------|-------------------|--------------------------------|
| | | Case (n = 475) | Control (n = 310) | |
| SNP4 | CC | 209 (44.2%) | 136 (43.7%) | ¹ CC+CG/GG 0.00020* |
| | CG | 222 (46.9%) | 119 (38.5%) | ² CC/CG/GG 0.00050* |
| | GG | 42 (8.9%) | 55 (17.8%) | ³ CC/CG+GG 0.89 |
| SNP5 | AA | 110 (23.2%) | 84 (27.2%) | ¹ AA+AC/CC 0.20 |
| | AC | 250 (52.6%) | 135 (43.7%) | ² AA/AC/CC 0.050* |
| SNP6 | CC | 115 (24.2%) | 90 (29.1%) | ³ CC+AC/AA 0.13 |
| | AA | 203 (42.7%) | 123 (39.8%) | ¹ AA+AG/GG 0.0058* |
| | AG | 224 (47.2%) | 134 (43.4%) | ² AA/AG/GG 0.022* |
| SNP7 | GG | 48 (10.1%) | 52 (16.8%) | ³ GG+AG/AA 0.38 |
| | TT | 78 (16.4%) | 43 (14.0%) | ¹ TT+AT/AA 0.35 |
| | AT | 168 (35.4%) | 147 (47.7%) | ² AA/AT/TT 0.55 |
| | AA | 229 (48.2%) | 118 (38.3%) | ³ AA+AT/TT 0.40 |

Values are n (%). *P* for genotype frequency are calculated as three models and are unadjusted. The superscripts ¹, ² and ³ represent dominant, codominant, and recessive models, respectively. * indicates statistical significance.

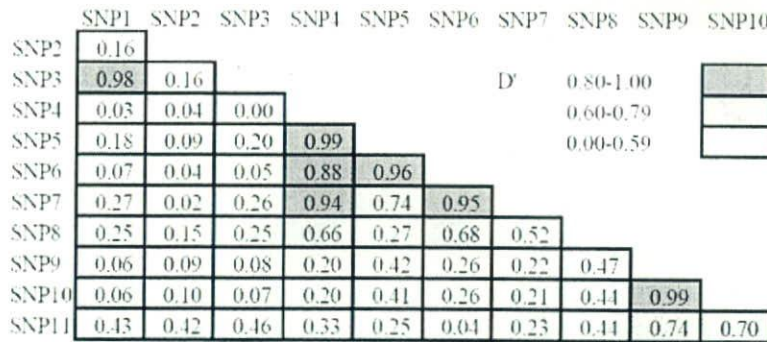


Figure 2. Pairwise linkage disequilibrium (LD). The pairwise LD between SNPs in the Mrf-2 gene and D' are indicated by the color of the block. Block structure is indicated at the upper right of the figure as a closed box.

G allele in SNP6 was significantly more prevalent in the control subjects ($P = 0.041$) (Table II). We also detected statistically significant differences for the genotype distribution of SNP4 (rs2893880) and SNP6 (rs7087507) for both the dominant model ($P = 0.0002$ and 0.0058 , respectively) and the codominant model ($P = 0.0005$ and 0.022 , respectively), and for the codominant model ($P = 0.050$) of SNP5, in the comparison of the CAD group and the control group (Table III).

Linkage disequilibrium between Mrf-2 SNPs and CAD: Linkage disequilibrium (LD) was estimated between pairs of SNPs by using an absolute value of D' ($|D'|$). There was almost complete linkage disequilibrium among SNP4 (rs2893880), SNP5 (rs10740055), SNP6 (rs7087507), and SNP7 (rs10761600). The LD of the above four SNPs covered over 60 kb of the Mrf-2 gene and constituted 1 haplotype block (Figure 2).

We then performed haplotype analysis, which is considered to be a more useful method for detecting associations than the assessment of individual SNPs. Seven major combinations (with a frequency $> 2\%$ for the entire sample) of the haplotype block consisting of SNP4, SNP5, SNP6, and SNP7 are shown in Table IV. Haplotype case-control analysis revealed a significant association between CAD and the haplotype G-C-G-A (SNP4-5-6-7) with a minor G in SNP4, a major C in SNP5, a minor G in SNP6, and a major A in SNP7 ($P = 0.049$). This association was consistent with the observations that GG genotype for SNP4 and GG for SNP6 were negatively associated with susceptibility to CAD.

Logistic regression analysis for the associations of Mrf-2 SNPs with CAD: The CAD group showed more conventional confounding risk factors. In order to

Table IV. Haplotype Analysis for Mrf-2 SNPs and CAD

| SNP4-5-6-7 | Frequency of Major Haplotypes | | | <i>P</i> |
|------------|-------------------------------|--------|---------|----------|
| | Overall | CAD | Control | |
| C-A-A-T | 0.3410 | 0.3432 | 0.3377 | 0.82 |
| G-C-G-A | 0.3123 | 0.2937 | 0.3408 | 0.049* |
| C-A-A-A | 0.1423 | 0.1442 | 0.1397 | 0.80 |
| C-C-A-A | 0.0922 | 0.1031 | 0.0754 | 0.064 |
| C-C-A-T | 0.0413 | 0.0470 | 0.0322 | 0.15 |
| C-C-G-A | 0.0298 | 0.0297 | 0.0300 | 0.97 |
| G-C-A-A | 0.0215 | 0.0206 | 0.0230 | 0.75 |

Seven major haplotypes were detected for SNP4 to SNP7. The G-C-G-A haplotype is negatively associated with susceptibility to CAD. * indicates statistical significance. This association is relevant to the fact that the GG (SNP4) and GG (SNP6) are negatively associated with susceptibility to and severity of CAD.