

## Identification of five genetic variants as novel determinants of type 2 diabetes mellitus in Japanese by exome-wide association studies

Yoshiji Yamada<sup>1,2</sup>, Jun Sakuma<sup>2,3,4</sup>, Ichiro Takeuchi<sup>2,4,5</sup>, Yoshiki Yasukochi<sup>1,2</sup>, Kimihiko Kato<sup>1,6</sup>, Mitsutoshi Oguri<sup>1,7</sup>, Tetsuo Fujimaki<sup>8</sup>, Hideki Horibe<sup>9</sup>, Masaaki Muramatsu<sup>10</sup>, Motoji Sawabe<sup>11</sup>, Yoshinori Fujiwara<sup>12</sup>, Yu Taniguchi<sup>12</sup>, Shuichi Obuchi<sup>13</sup>, Hisashi Kawai<sup>13</sup>, Shoji Shinkai<sup>14</sup>, Seiji Mori<sup>15</sup>, Tomio Arai<sup>16</sup> and Masashi Tanaka<sup>17</sup>

<sup>1</sup>Department of Human Functional Genomics, Advanced Science Research Promotion Center, Mie University, Tsu, Japan

<sup>2</sup>CREST, Japan Science and Technology Agency, Kawaguchi, Japan

<sup>3</sup>Computer Science Department, College of Information Science, University of Tsukuba, Tsukuba, Japan

<sup>4</sup>RIKEN Center for Advanced Intelligence Project, Tokyo, Japan

<sup>5</sup>Department of Computer Science, Nagoya Institute of Technology, Nagoya, Japan

<sup>6</sup>Department of Internal Medicine, Meitoh Hospital, Nagoya, Japan

<sup>7</sup>Department of Cardiology, Kasugai Municipal Hospital, Kasugai, Japan

<sup>8</sup>Department of Cardiovascular Medicine, Inabe General Hospital, Inabe, Japan

<sup>9</sup>Department of Cardiovascular Medicine, Gifu Prefectural Tajimi Hospital, Tajimi, Japan

<sup>10</sup>Department of Molecular Epidemiology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan

<sup>11</sup>Section of Molecular Pathology, Graduate School of Health Care Sciences, Tokyo Medical and Dental University, Tokyo, Japan

<sup>12</sup>Research Team for Social Participation and Community Health, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan

<sup>13</sup>Research Team for Promoting Support System for Home Care, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan

<sup>14</sup>Research Team for Social Participation and Health Promotion, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan

<sup>15</sup>Center for Promotion of Clinical Investigation, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan

<sup>16</sup>Department of Pathology, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan

<sup>17</sup>Department of Clinical Laboratory, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan

**Correspondence to:** Yoshiji Yamada, **email:** yamada@gene.mie-u.ac.jp

**Keywords:** diabetes mellitus, fasting plasma glucose, blood glycosylated hemoglobin, polymorphism, exome-wide association study

**Received:** March 31, 2017

**Accepted:** June 09, 2017

**Published:** July 17, 2017

**Copyright:** Yamada et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### ABSTRACT

**We performed exome-wide association studies to identify single nucleotide polymorphisms that either influence fasting plasma glucose level or blood hemoglobin A<sub>1c</sub> content or confer susceptibility to type 2 diabetes mellitus in Japanese. Exome-wide association studies were performed with the use of Illumina Human Exome-12 DNA Analysis or Infinium Exome-24 BeadChip arrays and with 11,729 or 8635 subjects for fasting plasma glucose level or blood hemoglobin A<sub>1c</sub> content, respectively, or with 14,023 subjects for type 2 diabetes mellitus (3573 cases, 10,450 controls). The relation of genotypes of 41,265 polymorphisms to fasting plasma glucose level or blood hemoglobin A<sub>1c</sub> content was examined by linear regression analysis. After Bonferroni's correction, 41 and 17 polymorphisms were significantly ( $P < 1.21 \times 10^{-6}$ ) associated with fasting plasma glucose level or blood hemoglobin A<sub>1c</sub> content,**

respectively, with two polymorphisms (rs139421991, rs189305583) being associated with both. Examination of the relation of allele frequencies to type 2 diabetes mellitus with Fisher's exact test revealed that 87 polymorphisms were significantly ( $P < 1.21 \times 10^{-6}$ ) associated with type 2 diabetes mellitus. Subsequent multivariable logistic regression analysis with adjustment for age and sex showed that four polymorphisms (rs138313632, rs76974938, rs139012426, rs147317864) were significantly ( $P < 1.44 \times 10^{-4}$ ) associated with type 2 diabetes mellitus, with rs138313632 and rs139012426 also being associated with fasting plasma glucose and rs76974938 with blood hemoglobin A<sub>1c</sub>. Five polymorphisms—rs139421991 of *CAT*, rs189305583 of *PDCL2*, rs138313632 of *RUFY1*, rs139012426 of *LOC100505549*, and rs76974938 of *C21orf59*—may be novel determinants of type 2 diabetes mellitus.

## INTRODUCTION

Type 2 diabetes mellitus (DM) is a major cause of nephropathy, retinopathy, and neuropathy as well as cardiovascular disease and stroke [1, 2]. The heritability of type 2 DM has been estimated to be 50% to 60% [3]. Genome-wide association studies (GWASs) and meta-analyses thereof have identified >80 susceptibility loci for type 2 DM in individuals of European [4–9] or African [10] ancestry, in East Asians [11], or in multiple ethnic groups [12]. Genetic variants identified in these previous studies typically have a minor allele frequency (MAF) of  $\geq 5\%$  and a small individual effect size. Given that these common variants explain only a fraction of the heritability of type 2 DM, low-frequency ( $0.5\% \leq \text{MAF} < 5\%$ ) or rare ( $\text{MAF} < 0.5\%$ ) variants with larger effect sizes are also thought to contribute to the genetic architecture of this condition [13]. Among Japanese, GWASs have identified *KCNQ1* [14, 15], *UBE2E2*, *C2CD4A-B* [16], *ANK1* [17], *MIR129-LEP*, *GPSMI1*, and *SLC16A13* [18] as susceptibility genes for type 2 DM, and a recent meta-analysis identified an additional seven susceptibility loci [19]. Genetic variants, including low-frequency and rare variants, that influence fasting plasma glucose (FPG) levels and blood glycosylated hemoglobin (HbA<sub>1c</sub>) content or which contribute to predisposition to type 2 DM in Japanese remain to be identified definitively, however.

We have now performed exome-wide association studies (EWASs) with the use of exome array-based genotyping methods to identify single nucleotide polymorphisms (SNPs)—especially low-frequency or rare coding variants with moderate to large effect sizes—that influence FPG levels and blood HbA<sub>1c</sub> content or which confer susceptibility to type 2 DM in Japanese. We used Illumina arrays that provide coverage of functional SNPs in entire exons including low-frequency and rare variants.

## RESULTS

### Characteristics of subjects

The characteristics of subjects are shown in Table 1. Age, the frequency of men, body mass index, and the

prevalence of hypertension, dyslipidemia, chronic kidney disease, and hyperuricemia as well as systolic and diastolic blood pressure, serum concentrations of triglycerides, creatinine, and uric acid, FPG level, and blood HbA<sub>1c</sub> content were greater, whereas estimated glomerular filtration rate and the serum concentration of high density lipoprotein (HDL)-cholesterol were lower, in subjects with type 2 DM than in controls.

### EWAS for FPG concentration

We examined the relation of genotypes of 41,265 SNPs that passed quality control to FPG levels in 11,729 subjects by linear regression analysis. A Manhattan plot for the EWAS is shown in Supplementary Figure 1A. After Bonferroni's correction, 41 SNPs were significantly [ $P < 1.21 \times 10^{-6}$  ( $0.05/41,265$ )] associated with FPG concentration (Table 2).

### EWAS for blood HbA<sub>1c</sub> content

We examined the relation of genotypes of 41,265 SNPs to blood HbA<sub>1c</sub> content in 8635 subjects by linear regression analysis. A Manhattan plot for the EWAS is shown in Supplementary Figure 1B. After Bonferroni's correction, 17 SNPs were significantly ( $P < 1.21 \times 10^{-6}$ ) associated with blood HbA<sub>1c</sub> content (Table 3). SNPs rs139421991 [G/A (R320Q)] of *CAT* and rs189305583 [C/T (V69I)] of *PDCL2* were significantly associated with both FPG level and blood HbA<sub>1c</sub> content.

### EWAS for type 2 DM

The EWAS for type 2 DM was performed with 14,023 subjects (3573 individuals with type 2 DM, 10,450 controls). We examined the relation of allele frequencies of 41,265 SNPs to type 2 DM with Fisher's exact test. A Manhattan plot for the EWAS is shown in Supplementary Figure 1C. After Bonferroni's correction, 87 SNPs were significantly ( $P < 1.21 \times 10^{-6}$ ) associated with type 2 DM (Supplementary Table 1). The genotype distributions of these SNPs were in Hardy-Weinberg equilibrium ( $P > 0.001$ ) among controls (Supplementary Table 2).

**Table 1: Characteristics of the 14,023 study subjects**

| Characteristic   | Type 2 DM    | Controls    | P       |
|--|--------------|-------------|---------|
| No. of subjects  | 3573         | 10,450      |         |
| Age (years)  | 66.2 ± 11.5  | 59.5 ± 13.8 | <0.0001 |
| Sex (male/female, %)                                       | 66.8/33.2    | 52.7/47.3   | <0.0001 |
| Body mass index (kg/m <sup>2</sup> )                       | 24.2 ± 3.8   | 23.0 ± 3.3  | <0.0001 |
| Current or former smoker (%)                               | 38.5         | 35.8        | 0.0067  |
| Hypertension (%)   | 76.9         | 45.6        | <0.0001 |
| Systolic blood pressure (mmHg)                             | 141 ± 26     | 128 ± 22    | <0.0001 |
| Diastolic blood pressure (mmHg)                            | 78 ± 15      | 76 ± 13     | <0.0001 |
| Dyslipidemia (%)   | 75.4         | 56.7        | <0.0001 |
| Serum triglycerides (mmol/L)                               | 1.68 ± 1.25  | 1.38 ± 0.92 | <0.0001 |
| Serum HDL-cholesterol (mmol/L)                             | 1.31 ± 0.41  | 1.57 ± 0.46 | <0.0001 |
| Serum LDL-cholesterol (mmol/L)                             | 3.08 ± 0.97  | 3.13 ± 0.81 | 0.0159  |
| Chronic kidney disease (%)                                 | 34.5         | 20.4        | <0.0001 |
| Serum creatinine (μmol/L)                                  | 94.9 ± 123.5 | 73.1 ± 64.4 | <0.0001 |
| Estimated GFR (mL min <sup>-1</sup> 1.73 m <sup>-2</sup> ) | 68.8 ± 24.1  | 73.4 ± 19.1 | <0.0001 |
| Hyperuricemia (%)  | 22.4         | 15.8        | <0.0001 |
| Serum uric acid (μmol/L)                                   | 336 ± 101    | 325 ± 92    | <0.0001 |
| Fasting plasma glucose (mmol/L)                            | 8.81 ± 3.66  | 5.38 ± 0.77 | <0.0001 |
| Blood HbA <sub>1c</sub> (%)                                | 7.15 ± 1.73  | 5.48 ± 0.39 | <0.0001 |

Quantitative data are means ± SD and were compared between subjects with type 2 DM and controls with the unpaired Student's *t* test. Categorical data were compared between the two groups with Fisher's exact test. Based on Bonferroni's correction, a *P* value of <0.0028 (0.05/18) was considered statistically significant. Hypertension: systolic blood pressure of ≥140 mmHg, diastolic blood pressure of ≥90 mmHg, or taking of anti-hypertensive medication. Dyslipidemia: serum triglyceride concentration of ≥1.65 mmol/L, serum high density lipoprotein (HDL)-cholesterol concentration of <1.04 mmol/L, serum low density lipoprotein (LDL)-cholesterol concentration of ≥3.64 mmol/L, or taking of anti-dyslipidemic medication. Chronic kidney disease: estimated glomerular filtration rate (eGFR) of <60 mL min<sup>-1</sup> 1.73 m<sup>-2</sup>, where eGFR (mL min<sup>-1</sup> 1.73 m<sup>-2</sup>) = 194 × (age in years)<sup>-0.287</sup> × (serum creatinine in mg/dL)<sup>-1.094</sup> (× 0.739 if female). Hyperuricemia: serum concentration of uric acid of >416 μmol/L.

The relation of these 87 SNPs to type 2 DM was examined further by multivariable logistic regression analysis with adjustment for age and sex (Supplementary Table 3). Four SNPs—rs138313632 [T/G (S705A)] of *RUFY1*, rs76974938 [C/T (D67N)] of *C21orf59*, rs139012426 [G/C (S1242T)] of *LOC100505549*, and rs147317864 [C/T (A262T)] of *TRABD2B*—were significantly [*P* < 1.44 × 10<sup>-4</sup> (0.05/348)] associated with type 2 DM (Table 4). The minor *G*, *T*, and *C* alleles of rs138313632, rs76974938, and rs139012426, respectively, were protective against type 2 DM, whereas the minor *T* allele of rs147317864 was a risk factor for this condition. SNPs rs138313632 of *RUFY1* and rs139012426 of *LOC100505549* were significantly associated with both FPG level and type 2 DM, whereas rs76974938 of

*C21orf59* was significantly associated with both blood HbA<sub>1c</sub> content and type 2 DM.

#### Relation of SNPs to FPG level or blood HbA<sub>1c</sub> content

We examined the relation of genotypes of identified SNPs to the FPG level or blood HbA<sub>1c</sub> content by one-way analysis of variance (ANOVA). The 41 SNPs identified in the EWAS for FPG level, including the two SNPs (rs138313632, rs139012426) also found to be associated with type 2 DM, were all significantly [*P* < 0.0012 (0.05/43)] associated with FPG level. The remaining two SNPs associated with type 2 DM (rs76974938, rs147317864) were not significantly related to the FPG level (Table 5).

**Table 2: The 41 SNPs significantly ( $P < 1.21 \times 10^{-6}$ ) associated with FPG level in the EWAS**

| Gene                | dbSNP       | Nucleotide (amino acid) substitution <sup>a</sup> | Chromosome: position | MAF (%) | P (genotype)           |
|---------------------|-------------|---|----------------------|---------|------------------------|
| <i>OR4F6</i>        | rs141569282 | G/A (A117T)                                       | 15: 101806068        | 1.7     | $1.07 \times 10^{-12}$ |
| <i>SLC35F3</i>      | rs140011243 | T/G (C144G)                                       | 1: 234231563         | 0.5     | $1.56 \times 10^{-11}$ |
| <i>RUFY1</i>        | rs138313632 | T/G (S705A)                                       | 5: 179609505         | 0.5     | $9.91 \times 10^{-11}$ |
| <i>KARS</i>         | rs201151665 | A/T (M29L)  | 16: 75647967         | 0.4     | $3.98 \times 10^{-10}$ |
| <i>IFITM5</i>       | rs146230729 | G/T (P31T)  | 11: 299400           | 0.5     | $6.30 \times 10^{-10}$ |
| <i>CADMI</i>        | rs561567580 | G/C (R7P)   | 11: 115504375        | 0.4     | $6.76 \times 10^{-10}$ |
| <i>PPP1R9B</i>      | rs113281588 | G/C (G311A)                                       | 17: 50149582         | 0.5     | $1.57 \times 10^{-9}$  |
| <i>MUC17</i>        | rs78010183  | A/T (T1305S)                                      | 7: 101035329         | 1.8     | $1.63 \times 10^{-9}$  |
| <i>CCDC166</i>      | rs75368383  | T/C (K187R)                                       | 8: 143707454         | 0.5     | $1.70 \times 10^{-9}$  |
| <i>SCAMP4</i>       | rs150715312 | A/G (K69E)  | 19: 1912252          | 0.6     | $1.74 \times 10^{-9}$  |
| <i>PLEC</i>         | rs201654895 | A/G (M3874V)                                      | 8: 143918282         | 0.4     | $2.21 \times 10^{-9}$  |
| <i>LGR5</i>         | rs117324318 | G/A   | 12: 71440257         | 0.4     | $2.68 \times 10^{-9}$  |
| <i>CCDC114</i>      | rs140189114 | C/T (G632R)                                       | 19: 48297095         | 0.4     | $3.82 \times 10^{-9}$  |
| <i>NLRC3</i>        | rs116433328 | G/C (M286I)                                       | 16: 3564079          | 0.4     | $6.74 \times 10^{-9}$  |
| <i>CECR2</i>        | rs201989565 | G/A   | 22: 17548704         | 0.4     | $1.12 \times 10^{-8}$  |
| <i>DUS2</i>         | rs202069030 | G/C (R51S)  | 16: 68023050         | 0.4     | $1.14 \times 10^{-8}$  |
| <i>SIGLEC1</i>      | rs201950990 | A/C (V69G)  | 20: 3706550          | 0.5     | $2.11 \times 10^{-8}$  |
| <i>ALKBH1</i>       | rs200168197 | A/C (V329G)                                       | 14: 77673996         | 0.4     | $2.72 \times 10^{-8}$  |
| <i>TNFRSF4</i>      | rs150516264 | A/G (L98P)  | 1: 1213069           | 0.4     | $2.95 \times 10^{-8}$  |
| <i>LOC100505549</i> | rs139012426 | G/C (S1242T)                                      | 18: 57648519         | 0.4     | $3.64 \times 10^{-8}$  |
| <i>ADAD2</i>        | rs149894736 | C/T (P107L)                                       | 16: 84191550         | 0.7     | $3.70 \times 10^{-8}$  |
| <i>YBEY</i>         | rs200145138 | C/G (L148V)                                       | 21: 46297572         | 0.4     | $3.84 \times 10^{-8}$  |
| <i>PRKCDBP</i>      | rs11544766  | C/G (T68S)  | 11: 6320274          | 0.4     | $5.38 \times 10^{-8}$  |
| <i>MYLIP</i>        | rs201021082 | T/C (V17A)  | 6: 16129372          | 0.4     | $5.47 \times 10^{-8}$  |
| <i>MGAT3</i>        | rs201417286 | T/G (V200G)                                       | 22: 39487946         | 0.3     | $6.54 \times 10^{-8}$  |
| <i>B3GNT6</i>       | rs559157215 | A/G (H301R)                                       | 11: 77040453         | 0.3     | $7.38 \times 10^{-8}$  |
| <i>KNDC1</i>        | rs146093427 | A/G (N546D)                                       | 10: 133195723        | 0.4     | $8.57 \times 10^{-8}$  |
| <i>TNXB</i>         | rs141190850 | T/C (D677G)                                       | 6: 32095823          | 0.3     | $8.62 \times 10^{-8}$  |
| <i>PCDHAC1</i>      | rs185216314 | A/G (Q479R)                                       | 5: 140968202         | 0.1     | $1.40 \times 10^{-7}$  |
| <i>CSPG4</i>        | rs137981794 | T/C (D1936G)                                      | 15: 75676712         | 0.4     | $1.50 \times 10^{-7}$  |
| <i>DNAJB2</i>       | rs148615702 | C/G (Q235E)                                       | 2: 219284715         | 0.3     | $1.63 \times 10^{-7}$  |
| <i>FGD3</i>         | rs116496123 | G/T   | 9: 93034624          | 0.7     | $1.71 \times 10^{-7}$  |
| <i>AK8</i>          | rs150636539 | G/A (P328S)                                       | 9: 132792773         | 0.3     | $2.04 \times 10^{-7}$  |
| <i>LAMB3</i>        | rs202068754 | A/C (V753G)                                       | 1: 209623605         | 0.4     | $2.06 \times 10^{-7}$  |
| <i>GDPD3</i>        | rs200801803 | G/A   | 16: 30112132         | 0.3     | $2.95 \times 10^{-7}$  |
| <i>CAT</i>          | rs139421991 | G/A (R320Q)                                       | 11: 34456720         | 0.3     | $3.01 \times 10^{-7}$  |
| <i>OAS3</i>         | rs62623451  | G/A (A49T)  | 12: 112938675        | 0.3     | $3.13 \times 10^{-7}$  |
| <i>NOTCH1</i>       | rs201053795 | T/C (T970A)                                       | 9: 136509794         | 0.6     | $3.43 \times 10^{-7}$  |
| <i>PDCL2</i>        | rs189305583 | C/T (V69I)  | 4: 55580834          | 0.1     | $4.24 \times 10^{-7}$  |
| <i>SYNM</i>         | rs200549249 | G/A (G235E)                                       | 15: 99105903         | 0.2     | $7.54 \times 10^{-7}$  |
| <i>SNTB1</i>        | rs145615160 | A/G (Y57H)  | 8: 120811675         | 0.5     | $1.09 \times 10^{-6}$  |

The relation of genotypes of SNPs to FPG level was examined by linear regression analysis. <sup>a</sup>Major allele/minor allele.

**Table 3: The 17 SNPs significantly ( $P < 1.21 \times 10^{-6}$ ) associated with blood HbA<sub>1c</sub> content in the EWAS**

| Gene            | dbSNP       | Nucleotide (amino acid) substitution <sup>a</sup> | Chromosome: position | MAF (%) | P (genotype)           |
|-----------------|-------------|---|----------------------|---------|------------------------|
| <i>PTCHD3</i>   | rs77473776  | T/G (Q186K)                                       | 10: 27413695         | 30.6    | $3.59 \times 10^{-38}$ |
| <i>C21orf59</i> | rs76974938  | C/T (D67N)  | 21: 32609946         | 2.4     | $1.68 \times 10^{-27}$ |
| <i>TNC</i>      | rs138406927 | C/T (A1096T)                                      | 9: 115064848         | 2.1     | $4.60 \times 10^{-27}$ |
| <i>KRR1</i>     | rs17115182  | G/A (P43S)  | 12: 75508405         | 7       | $5.17 \times 10^{-22}$ |
| <i>ZNF43</i>    | rs149604219 | G/A (A93V)  | 19: 21809741         | 3       | $7.87 \times 10^{-17}$ |
| <i>PRCP</i>     | rs2229437   | T/G (E133D)                                       | 11: 82853252         | 6.4     | $9.82 \times 10^{-12}$ |
| <i>CYP4F12</i>  | rs609636    | G/A (D76N)  | 19: 15678288         | 2.3     | $4.74 \times 10^{-11}$ |
|                 | rs1917321   | A/C   | 11: 49356208         | 35.5    | $3.33 \times 10^{-10}$ |
| <i>LGALS14</i>  | rs72480733  | G/A (R27H)  | 19: 39705987         | 36.7    | $2.19 \times 10^{-8}$  |
| <i>ANKRD26</i>  | rs12572862  | C/G (L1219V)                                      | 10: 27033374         | 21.7    | $1.35 \times 10^{-7}$  |
| <i>CAT</i>      | rs139421991 | G/A (R320Q)                                       | 11: 34456720         | 0.3     | $1.69 \times 10^{-7}$  |
| <i>ZKSCAN3</i>  | rs13201752  | A/G (K200E)                                       | 6: 28363350          | 36.5    | $2.60 \times 10^{-7}$  |
|                 | rs10451497  | C/T   | 19: 15093055         | 31.2    | $5.29 \times 10^{-7}$  |
| <i>OR8H3</i>    | rs61751933  | C/T (T16M)  | 11: 56122419         | 18.3    | $7.54 \times 10^{-7}$  |
| <i>PDCL2</i>    | rs189305583 | C/T (V69I)  | 4: 55580834          | 0.1     | $8.56 \times 10^{-7}$  |
| <i>HIVEP1</i>   | rs200286173 | A/G (Y374C)                                       | 6: 12120916          | 0.2     | $9.93 \times 10^{-7}$  |
| <i>RSL24D1</i>  | rs200023487 | A/G (V28A)  | 15: 55192832         | 0.1     | $1.20 \times 10^{-6}$  |

The relation of genotypes of SNPs to blood HbA<sub>1c</sub> content was examined by linear regression analysis. <sup>a</sup>Major allele/minor allele.

**Table 4: Relation of SNPs to type 2 DM as determined by multivariable logistic regression analysis**

| Gene                | SNP         |              | Dominant                                 |                         | Recessive |             | Additive 1                               |                         | Additive 2 |             |
|---------------------|-------------|--------------|--|-------------------------|-----------|-------------|--|-------------------------|------------|-------------|
|                     |             |              | P  | OR (95% CI)             | P         | OR (95% CI) | P  | OR (95% CI)             | P          | OR (95% CI) |
| <i>RUFY1</i>        | rs138313632 | T/G (S705A)  | <b><math>4.20 \times 10^{-8}</math></b>  | 0.20 (0.09–0.39)        | ND        |             | <b><math>4.20 \times 10^{-8}</math></b>  | 0.20 (0.09–0.39)        | ND         |             |
| <i>C21orf59</i>     | rs76974938  | C/T (D67N)   | <b><math>1.00 \times 10^{-23}</math></b> | 0.33 (0.26–0.41)        | ND        |             | <b><math>1.00 \times 10^{-23}</math></b> | 0.33 (0.26–0.41)        | ND         |             |
| <i>LOC100505549</i> | rs139012426 | G/C (S1242T) | <b><math>2.28 \times 10^{-6}</math></b>  | 0.22 (0.09–0.44)        | ND        |             | <b><math>2.28 \times 10^{-6}</math></b>  | 0.22 (0.09–0.44)        | ND         |             |
| <i>TRABD2B</i>      | rs147317864 | C/T (A262T)  | <b><math>2.25 \times 10^{-5}</math></b>  | $1.21 \times 10^8$ (ND) | ND        |             | <b><math>2.25 \times 10^{-5}</math></b>  | $1.21 \times 10^8$ (ND) | ND         |             |

Multivariable logistic regression analysis was performed with adjustment for age and sex. Based on Bonferroni's correction, P values of  $< 1.44 \times 10^{-4}$  (0.05/348) were considered statistically significant and are shown in bold. OR, odds ratio; CI, confidence interval; ND, not determined.

**Table 5: Relation of SNPs to FPG level**

| SNP                                     |              | FPG (mmol/L) |             | P                              |
|---|--------------|--------------|-------------|--------------------------------|
| Associated with FPG level and type 2 DM |              |              |             |                                |
| rs138313632                             | T/G (S705A)  | <i>TT</i>    | <i>TG</i>   |                                |
|   |              | 6.25 ± 2.45  | 4.92 ± 1.66 | <b>9.91 × 10<sup>-11</sup></b> |
| rs139012426                             | G/C (S1242T) | <i>GG</i>    | <i>GC</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.96 ± 1.83 | <b>3.64 × 10<sup>-8</sup></b>  |
| Associated with FPG level               |              |              |             |                                |
| rs141569282                             | G/A (A117T)  | <i>GG</i>    | <i>GA</i>   |                                |
|   |              | 6.38 ± 2.65  | 5.50 ± 0.90 | <b>1.07 × 10<sup>-12</sup></b> |
| rs140011243                             | T/G (C144G)  | <i>TT</i>    | <i>TG</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.87 ± 1.63 | <b>1.56 × 10<sup>-11</sup></b> |
| rs201151665                             | A/T (M29L)   | <i>AA</i>    | <i>AT</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.81 ± 1.14 | <b>3.98 × 10<sup>-10</sup></b> |
| rs146230729                             | G/T (P31T)   | <i>GG</i>    | <i>GT</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.98 ± 1.72 | <b>6.30 × 10<sup>-10</sup></b> |
| rs561567580                             | G/C (R7P)    | <i>GG</i>    | <i>GC</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.96 ± 1.74 | <b>6.76 × 10<sup>-10</sup></b> |
| rs113281588                             | G/C (G311A)  | <i>GG</i>    | <i>GC</i>   |                                |
|   |              | 6.22 ± 2.44  | 5.00 ± 1.73 | <b>1.57 × 10<sup>-9</sup></b>  |
| rs78010183                              | A/T (T1305S) | <i>AA</i>    | <i>AT</i>   |                                |
|   |              | 6.25 ± 2.47  | 5.63 ± 1.39 | <b>1.63 × 10<sup>-9</sup></b>  |
| rs75368383                              | T/C (K187R)  | <i>TT</i>    | <i>TC</i>   |                                |
|   |              | 6.25 ± 2.44  | 5.02 ± 1.69 | <b>1.70 × 10<sup>-9</sup></b>  |
| rs150715312                             | A/G (K69E)   | <i>AA</i>    | <i>AG</i>   |                                |
|   |              | 6.25 ± 2.45  | 5.15 ± 1.70 | <b>1.74 × 10<sup>-9</sup></b>  |
| rs201654895                             | A/G (M3874V) | <i>AA</i>    | <i>AG</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.96 ± 1.19 | <b>2.21 × 10<sup>-9</sup></b>  |
| rs117324318                             | G/A          | <i>GG</i>    | <i>GA</i>   |                                |
|   |              | 6.24 ± 2.44  | 5.01 ± 1.73 | <b>2.68 × 10<sup>-9</sup></b>  |
| rs140189114                             | C/T (G632R)  | <i>CC</i>    | <i>CT</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.96 ± 1.75 | <b>3.82 × 10<sup>-9</sup></b>  |
| rs116433328                             | G/C (M286I)  | <i>GG</i>    | <i>GC</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.93 ± 1.77 | <b>6.74 × 10<sup>-9</sup></b>  |
| rs201989565                             | G/A          | <i>GG</i>    | <i>GA</i>   |                                |
|   |              | 6.24 ± 2.44  | 5.00 ± 1.76 | <b>1.12 × 10<sup>-8</sup></b>  |
| rs202069030                             | G/C (R51S)   | <i>GG</i>    | <i>GC</i>   |                                |
|   |              | 6.24 ± 2.44  | 4.91 ± 1.74 | <b>1.14 × 10<sup>-8</sup></b>  |

(Continued)

|             | SNP          | FPG (mmol/L) |             | P                             |
|-------------|--------------|--------------|-------------|-------------------------------|
| rs201950990 | A/C (V69G)   | AA           | AC          |                               |
|             |              | 6.24 ± 2.44  | 5.08 ± 1.79 | <b>2.11 × 10<sup>-8</sup></b> |
| rs200168197 | A/C (V329G)  | AA           | AC          |                               |
|             |              | 6.24 ± 2.44  | 4.97 ± 1.84 | <b>2.72 × 10<sup>-8</sup></b> |
| rs150516264 | A/G (L98P)   | AA           | AG          |                               |
|             |              | 6.24 ± 2.44  | 5.03 ± 1.81 | <b>2.95 × 10<sup>-8</sup></b> |
| rs149894736 | C/T (P107L)  | CC           | CT          |                               |
|             |              | 6.24 ± 2.44  | 5.25 ± 1.57 | <b>3.70 × 10<sup>-8</sup></b> |
| rs200145138 | C/G (L148V)  | CC           | CG          |                               |
|             |              | 6.24 ± 2.44  | 4.98 ± 1.92 | <b>3.84 × 10<sup>-8</sup></b> |
| rs11544766  | C/G (T68S)   | CC           | CG          |                               |
|             |              | 6.24 ± 2.44  | 5.01 ± 1.77 | <b>5.38 × 10<sup>-8</sup></b> |
| rs201021082 | T/C (V17A)   | TT           | TC          |                               |
|             |              | 6.24 ± 2.43  | 4.98 ± 1.89 | <b>5.47 × 10<sup>-8</sup></b> |
| rs201417286 | T/G (V200G)  | TT           | TG          |                               |
|             |              | 6.24 ± 2.44  | 4.98 ± 1.85 | <b>6.54 × 10<sup>-8</sup></b> |
| rs559157215 | A/G (H301R)  | AA           | AG          |                               |
|             |              | 6.24 ± 2.43  | 4.90 ± 1.87 | <b>7.38 × 10<sup>-8</sup></b> |
| rs146093427 | A/G (N546D)  | AA           | AG          |                               |
|             |              | 6.24 ± 2.44  | 5.09 ± 1.80 | <b>8.57 × 10<sup>-8</sup></b> |
| rs141190850 | T/C (D677G)  | TT           | TC          |                               |
|             |              | 6.24 ± 2.43  | 4.91 ± 1.92 | <b>8.62 × 10<sup>-8</sup></b> |
| rs185216314 | A/G (Q479R)  | AA           | AG          |                               |
|             |              | 6.21 ± 2.41  | 8.38 ± 6.21 | <b>1.40 × 10<sup>-7</sup></b> |
| rs137981794 | T/C (D1936G) | TT           | TC          |                               |
|             |              | 6.24 ± 2.44  | 5.06 ± 1.88 | <b>1.50 × 10<sup>-7</sup></b> |
| rs148615702 | C/G (Q235E)  | CC           | CG          |                               |
|             |              | 6.24 ± 2.44  | 4.86 ± 2.00 | <b>1.63 × 10<sup>-7</sup></b> |
| rs116496123 | G/T          | GG           | GT          |                               |
|             |              | 6.24 ± 2.46  | 5.36 ± 1.87 | <b>1.71 × 10<sup>-7</sup></b> |
| rs150636539 | G/A (P328S)  | GG           | GA          |                               |
|             |              | 6.24 ± 2.44  | 4.93 ± 1.88 | <b>2.04 × 10<sup>-7</sup></b> |
| rs202068754 | A/C (V753G)  | AA           | AC          |                               |
|             |              | 6.24 ± 2.44  | 5.10 ± 1.86 | <b>2.06 × 10<sup>-7</sup></b> |
| rs200801803 | G/A          | GG           | GA          |                               |
|             |              | 6.24 ± 2.44  | 4.88 ± 1.98 | <b>2.95 × 10<sup>-7</sup></b> |
| rs139421991 | G/A (R320Q)  | GG           | GA          |                               |
|             |              | 6.17 ± 2.35  | 7.43 ± 3.08 | <b>3.01 × 10<sup>-7</sup></b> |

(Continued)

|                           | SNP         | FPG (mmol/L) |             | <i>P</i>                      |
|---------------------------|-------------|--------------|-------------|-------------------------------|
| rs62623451                | G/A (A49T)  | <i>GG</i>    | <i>GA</i>   | <b>3.13 × 10<sup>-7</sup></b> |
|                           |             | 6.24 ± 2.44  | 5.03 ± 1.85 |                               |
| rs201053795               | T/C (T970A) | <i>TT</i>    | <i>TC</i>   | <b>3.43 × 10<sup>-7</sup></b> |
|                           |             | 6.24 ± 2.45  | 5.27 ± 1.87 |                               |
| rs189305583               | C/T (V69I)  | <i>CC</i>    | <i>CT</i>   | <b>4.24 × 10<sup>-7</sup></b> |
|                           |             | 6.21 ± 2.41  | 8.26 ± 5.49 |                               |
| rs200549249               | G/A (G235E) | <i>GG</i>    | <i>GA</i>   | <b>7.54 × 10<sup>-7</sup></b> |
|                           |             | 6.23 ± 2.44  | 4.76 ± 1.12 |                               |
| rs145615160               | A/G (Y57H)  | <i>AA</i>    | <i>AG</i>   | <b>1.09 × 10<sup>-6</sup></b> |
|                           |             | 6.24 ± 2.44  | 5.26 ± 2.15 |                               |
| Associated with type 2 DM |             |              |             |                               |
| rs76974938                | C/T (D67N)  | <i>CC</i>    | <i>CT</i>   | 0.0758                        |
|                           |             | 6.21 ± 2.46  | 6.46 ± 2.27 |                               |
| rs147317864               | C/T (A262T) | <i>CC</i>    |             | ND                            |
|                           |             | 6.21 ± 2.46  |             |                               |

Data were compared among genotypes by one-way ANOVA. Based on Bonferroni's correction, *P* values of <0.0012 (0.05/43) were considered statistically significant and are shown in bold. ND, not determined.

The 17 SNPs identified in the EWAS for blood HbA<sub>1c</sub> content, including the one SNP (rs76974938) also found to be associated with type 2 DM, were all significantly [*P* < 0.0025 (0.05/20)] associated with blood HbA<sub>1c</sub> content by one-way ANOVA. The remaining three SNPs associated with type 2 DM (rs138313632, rs139012426, rs147317864) were not significantly related to blood HbA<sub>1c</sub> content (Table 6).

### Relation of identified SNPs to phenotypes examined in previous GWASs

We examined the genes, chromosomal loci, and SNPs identified in the present study to DM-related phenotypes examined in previous GWASs deposited in a public database [GWAS Catalog (<http://www.ebi.ac.uk/gwas>)]. *LGR5* has been previously shown to be related to type 2 DM, whereas *TNXB* has been previously associated with type 1 DM and *PTCHD3* with fasting insulin-related traits (Supplementary Table 4). The remaining 54 SNPs identified in the present study have not been previously found to be related to DM-related phenotypes.

### Network analysis of the genes identified in the present study

We performed network analysis of the top ten (high scores) genes that have been shown to be associated with type 2 DM selected from DisGeNET database ([http://](http://www.disgenet.org/web/DisGeNET)

[www.disgenet.org/web/DisGeNET](http://www.disgenet.org/web/DisGeNET)) and four genes (*CAT*, *PDCL2*, *RUFY1*, and *C21orf59*) identified in the present study by the use of Cytoscape version 3.4.0 software (<http://www.cytoscape.org/>). Given that *LOC100505549* protein has not been characterized, it could not be examined. The network analysis showed that *CAT*, *PDCL2*, *RUFY1*, and *C21orf59* have potential indirect interactions with several genes previously shown to be associated with type 2 DM (Supplementary Figure 2).

## DISCUSSION

We have now shown that two SNPs—rs139421991 [G/A (R320Q)] of *CAT* and rs189305583 [C/T (V69I)] of *PDCL2*—were significantly associated with both FPG levels and blood HbA<sub>1c</sub> content; two SNPs—rs138313632 [T/G (S705A)] of *RUFY1* and rs139012426 [G/C (S1242T)] of *LOC100505549*—were significantly associated with both FPG levels and type 2 DM; and one SNP, rs76974938 [C/T (D67N)] of *C21orf59*, was significantly associated with both blood HbA<sub>1c</sub> content and type 2 DM. These five SNPs may thus be novel determinants of type 2 DM. Given that FPG levels are affected by meals of the day before examination, we selected the SNPs and genes that were associated with both FPG levels and blood HbA<sub>1c</sub> or type 2 DM.

The catalase gene (*CAT*) is located at chromosomal region 11p13 (NCBI Gene, <https://www.ncbi.nlm.nih.gov/gene/>) and is expressed in various tissues and organs



**Table 6: Relation of SNPs to blood HbA<sub>1c</sub> content**

| SNP   |              | Blood HbA <sub>1c</sub> (%) |             |             | P                        |
|---|--------------|-----------------------------|-------------|-------------|--------------------------|
| Associated with blood HbA <sub>1c</sub> content and type 2 DM |              |                             |             |             |                          |
| rs76974938  | C/T (D67N)   | CC                          | CT          |             |                          |
|   |              | 6.04 ± 1.25                 | 5.47 ± 0.68 |             | <1.0 × 10 <sup>-23</sup> |
| Associated with blood HbA <sub>1c</sub> content               |              |                             |             |             |                          |
| rs77473776  | T/G (Q186K)  | TT                          | TG          | GG          |                          |
|   |              | 6.03 ± 1.28                 | 5.85 ± 1.12 | 5.47 ± 0.63 | <1.0 × 10 <sup>-23</sup> |
| rs138406927   | C/T (A1096T) | CC                          | CT          |             |                          |
|   |              | 6.04 ± 1.25                 | 5.44 ± 0.61 |             | <1.0 × 10 <sup>-23</sup> |
| rs17115182  | G/A (P43S)   | GG                          | GA          |             |                          |
|   |              | 6.04 ± 1.25                 | 5.48 ± 0.57 |             | 5.20 × 10 <sup>-22</sup> |
| rs149604219   | G/A (A93V)   | GG                          | GA          |             |                          |
|   |              | 5.95 ± 1.21                 | 5.44 ± 0.51 |             | 7.87 × 10 <sup>-17</sup> |
| rs2229437   | T/G (E133D)  | TT                          | TG          | GG          |                          |
|   |              | 5.89 ± 1.14                 | 6.15 ± 1.44 | 6.26 ± 1.32 | 6.94 × 10 <sup>-11</sup> |
| rs609636  | G/A (D76N)   | GG                          | GA          |             |                          |
|   |              | 6.01 ± 1.24                 | 5.31 ± 0.50 |             | 4.74 × 10 <sup>-11</sup> |
| rs1917321   | A/C          | AA                          | AC          | CC          |                          |
|   |              | 5.87 ± 1.09                 | 6.02 ± 1.26 | 6.10 ± 1.37 | 1.11 × 10 <sup>-9</sup>  |
| rs72480733  | G/A (R27H)   | GG                          | GA          | AA          |                          |
|   |              | 5.85 ± 1.10                 | 5.94 ± 1.23 | 6.09 ± 1.30 | 9.51 × 10 <sup>-8</sup>  |
| rs12572862  | C/G (L1219V) | CC                          | CG          | GG          |                          |
|   |              | 5.92 ± 1.21                 | 6.06 ± 1.33 | 6.18 ± 1.34 | 8.99 × 10 <sup>-7</sup>  |
| rs139421991   | G/A (R320Q)  | GG                          | GA          |             |                          |
|   |              | 5.89 ± 1.14                 | 6.64 ± 1.50 |             | 1.69 × 10 <sup>-7</sup>  |
| rs13201752  | A/G (K200E)  | AA                          | AG          | GG          |                          |
|   |              | 6.05 ± 1.24                 | 5.90 ± 1.20 | 5.87 ± 1.09 | 3.47 × 10 <sup>-7</sup>  |
| rs10451497  | C/T          | CC                          | CT          | TT          |                          |
|   |              | 6.02 ± 1.21                 | 5.93 ± 1.27 | 5.78 ± 0.92 | 2.03 × 10 <sup>-6</sup>  |
| rs61751933  | C/T (T16M)   | CC                          | CT          | TT          |                          |
|   |              | 6.04 ± 1.26                 | 5.88 ± 1.16 | 5.91 ± 1.04 | 8.58 × 10 <sup>-7</sup>  |
| rs189305583   | C/T (V69I)   | CC                          | CT          |             |                          |
|   |              | 5.90 ± 1.17                 | 7.00 ± 2.87 |             | 8.56 × 10 <sup>-7</sup>  |
| rs200286173   | A/G (Y374C)  | AA                          | AG          |             |                          |
|   |              | 5.90 ± 1.17                 | 6.99 ± 2.46 |             | 9.93 × 10 <sup>-7</sup>  |
| rs200023487   | A/G (V28A)   | AA                          | AG          |             |                          |
|   |              | 5.90 ± 1.17                 | 7.04 ± 2.55 |             | 1.20 × 10 <sup>-6</sup>  |

(Continued)

| SNP                       |              | Blood HbA <sub>1c</sub> (%) |           | P      |
|---------------------------|--------------|-----------------------------|-----------|--------|
| Associated with type 2 DM |              |                             |           |        |
| rs138313632               | T/G (S705A)  | <i>TT</i>                   |           |        |
|                           |              | 5.90 ± 1.18                 |           | ND     |
| rs139012426               | G/C (S1242T) | <i>GG</i>                   | <i>GC</i> |        |
|                           |              | 5.90 ± 1.18                 | 8.2       | 0.0504 |
| rs147317864               | C/T (A262T)  | <i>CC</i>                   |           |        |
|                           |              | 6.02 ± 1.24                 |           | ND     |

Data were compared among genotypes by one-way ANOVA. Based on Bonferroni's correction, *P* values of <0.0025 (0.05/20) were considered statistically significant and are shown in bold. ND, not determined.

including the pancreas (The Human Protein Atlas, <http://www.proteinatlas.org>). Catalase catalyzes the breakdown of hydrogen peroxide into oxygen and water. Inherited catalase deficiency has been associated with an increased prevalence of type 2 DM in Hungarians [20–22]. The frequency of various *CAT* mutations has been found to be increased in individuals with DM, especially in females with type 2 DM, and such inherited catalase deficiency is associated with an early onset of type 2 DM [20]. We have now shown that rs139421991 [G/A (R320Q)] of *CAT* was significantly associated with both FPG levels and blood HbA<sub>1c</sub> content, with the minor *A* allele being related to an increase in these parameters. This association of *CAT* with type 2 DM may be attributable to the role of the encoded protein in the metabolism of hydrogen peroxide and oxidative stress, although the underlying molecular mechanism remains to be determined.

The phosducin like 2 gene (*PDCL2*) is located at chromosome 4q12 (NCBI Gene) and is expressed at a high level in testis (The Human Protein Atlas). The PDCL2 protein belongs to the phosducin family [23]. PDCL1 has been shown to be essential for G protein signaling as a result of its role in folding and assembly of the Gβγ dimer. PDCL2 and PDCL3 likely assist in the folding of actin, tubulin, and proteins that activate cell cycle progression [24]. We have now shown that rs189305583 [C/T (V69I)] of *PDCL2* was significantly associated with both the FPG concentration and blood HbA<sub>1c</sub> content, with the minor *T* allele being related to increases in these parameters. Given that G protein–coupled receptor signaling promotes insulin secretion and the proliferation of pancreatic β cells [25], the association *PDCL2* with FPG levels and blood HbA<sub>1c</sub> content may be attributable to an effect of PDCL2 on such signaling, although the molecular mechanism remains unclear.

The RUN and FYVE domain containing 1 gene (*RUFY1*) is located at chromosomal region 5q35.3 (NCBI Gene) and is expressed in various tissues and organs including the pancreas (The Human Protein Atlas). The

*RUFY1* protein binds to phosphatidylinositol 3-phosphate and promotes early endosomal trafficking including the tethering and fusion of vesicles through interactions with small GTPases such as Rab4, Rab5, and Rab14 [26]. We have now shown that rs138313632 [T/G (S705A)] of *RUFY1* was significantly associated with both FPG concentration and type 2 DM, with the minor *G* allele being related to a decreased FPG level and a reduced risk for type 2 DM. Given that small GTPases enhances insulin granule exocytosis [27], the association of *RUFY1* with both FPG levels and type 2 DM may be attributable to an effect of the encoded protein on insulin secretion.

The uncharacterized *LOC100505549* gene is located at chromosome 18q21.31 (NCBI Gene). The function of *LOC100505549* remains unknown. We have now shown that rs139012426 [G/C (S1242T)] of *LOC100505549* was significantly associated with both FPG levels and type 2 DM, with the minor *C* allele being related to a decreased FPG concentration and a reduced risk of type 2 DM, although the functional relevance of this association remains unknown.

The chromosome 21 open reading frame 59 gene (*C21orf59*) is located at chromosome 21q22.11 (NCBI Gene) and is expressed in various tissues and organs including the pancreas (The Human Protein Atlas). The *C21orf59* protein promotes dynein arm assembly in motile cilia, and mutations in *C21orf59* cause ciliary dyskinesia [28]. Ciliopathies are associated with pancreatic defects that manifest mostly as cysts originating from ductal cells. Ciliary proteins have been suggested to influence insulin secretion and energy regulation [29, 30]. We have now shown that rs76974938 [C/T (D67N)] of *C21orf59* was significantly associated with both blood HbA<sub>1c</sub> content and type 2 DM, with the minor *T* allele being related to a decreased blood HbA<sub>1c</sub> content and a reduced risk for type 2 DM. Given that *C21orf59* may activate ciliary function and that cilia influence insulin secretion [29, 30], the association of *C21orf59* with blood HbA<sub>1c</sub> content and type 2 DM might reflect an effect of this gene on insulin

secretion, although the underlying molecular mechanism remains unclear.

In previous GWASs of type 2 DM in the Japanese population [14–19], the MAF of identified SNPs ranged from 2% to 48% and the odds ratio (OR) from 0.38 to 1.70. In a meta-analysis of GWASs for type 2 DM in East Asian populations [11] and in a trans-ancestry meta-analysis of GWASs for type 2 DM [12], the OR ranged from 1.06 to 1.10 or from 1.08 to 1.13, respectively. In our study, we identified four SNPs associated with type 2 DM, with the MAF and OR in a dominant model of logistic regression analysis for rs138313632, rs76974938, rs139012426, and rs147317864 being 0.5% and 0.20, 2.4% and 0.33, 0.4% and 0.22, and 0.2% and  $1.21 \times 10^8$ , respectively. Both rs138313632 and rs76974938 were thus low-frequency variants with a moderate effect size, whereas rs139012426 and rs147317864 were rare variants with a moderate to large effect size.

Two SNPs (MAF and differences in FPG level or blood HbA<sub>1c</sub> content between genotypes, respectively, shown in parentheses) associated with both FPG concentration and blood HbA<sub>1c</sub> content—rs139421991 of *CAT* (0.3%, 17.0%, 11.3%) and rs189305583 of *PDCL2* (0.1%, 24.8%, 15.7%)—were rare variants with a large effect size; two SNPs associated with both FPG levels and type 2 DM—rs138313632 of *RUFY1* (0.5%, 21.3%, 22.4%) and rs139012426 of *LOC100505549* (0.4%, 20.5%, 28.0%)—were rare or low-frequency variants with a large effect size; and one SNP associated with both blood HbA<sub>1c</sub> content and type 2 DM, rs76974938 of *C21orf59* (2.4%, 4.3%, 9.4%), was a low-frequency variant with a moderate effect size. Among the remaining 37 SNPs associated with FPG levels, 25 SNPs were rare variants with a large effect size and 12 SNPs were low-frequency variants with a moderate to large effect size. Among the remaining 14 SNPs associated with blood HbA<sub>1c</sub> content, two SNPs were rare variants with a large effect size, three SNPs were low-frequency variants with a moderate to large effect size, and nine SNPs were common variants with a small to moderate effect size (Supplementary Table 5).

There are several limitations to the present study.

(i) Given that our results were not replicated, they will require validation in other subject panels or in other ethnic groups. (ii) There is a possibility that some control individuals are prediabetic. (iii) It is possible that SNPs identified in the present study are in linkage disequilibrium with other polymorphisms in other nearby genes that are actually determinants of FPG levels, blood HbA<sub>1c</sub> content, or the development of type 2 DM. (iv) One SNP associated with type 2 DM was not significantly related to FPG level or blood HbA<sub>1c</sub> content, a discrepancy that may be attributable to the effects of medical treatment. (v) The biological or functional evidence of the association of the identified SNPs with FPG level, blood HbA<sub>1c</sub> content, or type 2 DM remains to be determined. Because of lack

of experiments for functional analyses, the association of the SNPs identified in the present study with type 2 DM, FPG levels, or blood HbA<sub>1c</sub> content should be interpreted carefully.

In conclusion, we have identified five SNPs—rs139421991 [G/A (R320Q)] of *CAT*, rs189305583 [C/T (V69I)] of *PDCL2*, rs138313632 [T/G (S705A)] of *RUFY1*, rs139012426 [G/C (S1242T)] of *LOC100505549*, and rs76974938 [C/T (D67N)] of *C21orf59*—as novel determinants of type 2 DM. We also identified 37, 14, or one SNPs as candidate determinants of FPG levels, blood HbA<sub>1c</sub> content, and type 2 DM, respectively. Determination of genotypes for these SNPs may prove informative for assessment of the genetic risk for type 2 DM in Japanese.

## MATERIALS AND METHODS

### Study subjects

A total of 14,023 individuals was examined. The subjects were recruited as described previously [31].

Type 2 DM was defined according to the criteria of the World Health Organization as described previously [32–34]. Subjects with type 2 DM had an FPG level of  $\geq 6.93$  mmol/L (126 mg/dL) or a blood HbA<sub>1c</sub> content of  $\geq 6.5\%$  or were taking antidiabetes medication. We thus examined 3573 subjects with type 2 DM and 10,450 controls. Individuals with type 1 DM, maturity-onset diabetes of the young, DM associated with mitochondrial diseases or single-gene disorders, pancreatic diseases, or other metabolic or endocrinologic diseases were excluded from the study. Those taking medications that may cause secondary DM were also excluded. The control subjects had an FPG level of  $< 6.05$  mmol/L (110 mg/dL), a blood HbA<sub>1c</sub> content of  $< 6.2\%$ , and no history of DM or of having taken antidiabetes medication. Autopsy cases were excluded from controls.

The study protocol complied with the Declaration of Helsinki and was approved by the Committees on the Ethics of Human Research of Mie University Graduate School of Medicine, Hirosaki University Graduate School of Medicine, Tokyo Metropolitan Institute of Gerontology, and participating hospitals. Written informed consent was obtained from all subjects or families of the deceased subjects.

### EWASs

Methods for sample collection and extraction of genomic DNA have been described previously [31]. EWASs for FPG concentration and blood HbA<sub>1c</sub> content included 11,729 and 8635 subjects, respectively, whereas that for type 2 DM included 14,023 individuals (3573 subjects with type 2 DM, 10,450 controls). Data for FPG levels were obtained from subjects who had

fasted overnight. Data for blood HbA<sub>1c</sub> content were obtained from subjects with type 2 DM or impaired glucose tolerance or from those who had annual health checkup. The EWASs were performed with the use of a HumanExome-12 v1.1 or v1.2 DNA Analysis BeadChip or Infinium Exome-24 v1.0 BeadChip (Illumina, San Diego, CA, USA). Detailed information of the exome arrays and methods of quality control have been described previously [31]. Genotype data were examined for population stratification by principal components analysis [35] (Supplementary Figure 3). A total of 41,265 SNPs passed quality control and was subjected to analysis.

### Statistical analysis

The relation of genotypes of SNPs to FPG level or blood HbA<sub>1c</sub> content in the EWASs was examined by linear regression analysis. For analysis of characteristics of the study subjects, quantitative and categorical data were compared between individuals with type 2 DM and controls with the unpaired Student's *t* test or Fisher's exact test, respectively. Allele frequencies were estimated by the gene counting method, and Fisher's exact test was used to identify departure from Hardy-Weinberg equilibrium. The relation of allele frequencies of SNPs to type 2 DM in the EWAS was examined with Fisher's exact test. To compensate for multiple comparisons of genotypes with FPG level or blood HbA<sub>1c</sub> content or of allele frequencies with type 2 DM, we applied Bonferroni's correction for statistical significance of association. Given that 41,265 SNPs were analyzed, the significance level was set at  $P < 1.21 \times 10^{-6}$  ( $0.05/41,265$ ) for the EWASs. Quantile-quantile plots for *P* values of genotypes in the EWASs for FPG level or blood HbA<sub>1c</sub> content or for those of allele frequencies in the EWAS for type 2 DM are shown in Supplementary Figure 4. The inflation factor ( $\lambda$ ) was 1.02 for FPG level, 1.03 for blood HbA<sub>1c</sub> content, and 1.26 for type 2 DM. Multivariable logistic regression analysis was performed with type 2 DM as a dependent variable and independent variables including age, sex (0, woman; 1, man), and genotype of each SNP. A detailed method of analysis has been described previously [31]. The relation of genotypes of identified SNPs to FPG level or blood HbA<sub>1c</sub> content was examined by one-way ANOVA. Bonferroni's correction was also applied to other statistical analysis as indicated to compensate for multiple comparisons. Statistical tests were performed with JMP Genomics version 6.0 software (SAS Institute, Cary, NC, USA).

### Author contributions

Y. Yamada contributed to conception and design of the study; to acquisition, analysis, and interpretation of the data; and to drafting of the manuscript. J. Sakuma, I. Takeuchi, and Y. Yasukochi contributed to analysis and

interpretation of the data as well as to revision of the manuscript. K. Kato, M. Oguri, T. Fujimaki, H. Horibe, M. Muramatsu, M. Sawabe, Y. Fujiwara, Y. Taniguchi, S. Obuchi, H. Kawai, S. Shinkai, S. Mori, and T. Arai contributed to acquisition of the data and to revision of the manuscript. M. Tanaka contributed to acquisition, analysis, and interpretation of the data as well as to revision of the manuscript. All authors approved submission of the final version of the article for publication.

### CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

### FUNDING

This work was supported by CREST (grant number, JPMJCR1302), Japan Science and Technology Agency (to Y. Yamada, J. Sakuma, and I. Takeuchi) and by Japan Society for the Promotion of Science KAKENHI grants JP15H04772 (to Y. Yamada), JP25242062 (to M. Tanaka), and JP16H01872 (to M. Tanaka).

### REFERENCES

1. Ismail-Beigi F. Clinical practice. Glycemic management of type 2 diabetes mellitus. *N Engl J Med.* 2012; 366:1319–1327.
2. Sarwar N, Gao P, Seshasai SR, Gobin R, Kaptoge S, Di Angelantonio E, Ingelsson E, Lawlor DA, Selvin E, Stampfer M, Stehouwer CD, Lewington S, Pennells L, et al, and Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet.* 2010; 375:2215–2222.
3. Almgren P, Lehtovirta M, Isomaa B, Sarelin L, Taskinen MR, Lyssenko V, Tuomi T, Groop L, and Botnia Study Group. Heritability and familiarity of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia.* 2011; 54:2811–2819.
4. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007; 447:661–678.
5. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature.* 2007; 445:881–885.
6. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Mägi R, Morris AP, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet.* 2010; 42:105–116.

7. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, McCulloch LJ, Ferreira T, Grallert H, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet.* 2010; 42:579–589.
8. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segrè AV, Steinthorsdottir V, Strawbridge RJ, Khan H, Grallert H, Mahajan A, Prokopenko I, Kang HM, Dina C, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet.* 2012; 44:981–990.
9. Scott RA, Lagou V, Welch RP, Wheeler E, Montasser ME, Luan J, Mägi R, Strawbridge RJ, Rehnberg E, Gustafsson S, Kanoni S, Rasmussen-Torvik LJ, Yengo L, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet.* 2012; 44:991–1005.
10. Ng MC, Shriner D, Chen BH, Li J, Chen WM, Guo X, Liu J, Bielinski SJ, Yanek LR, Nalls MA, Comeau ME, Rasmussen-Torvik LJ, Jensen RA, et al. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS Genet.* 2014; 10:e1004517.
11. Cho YS, Chen CH, Hu C, Long J, Ong RT, Sim X, Takeuchi F, Wu Y, Go MJ, Yamauchi T, Chang YC, Kwak SH, Ma RC, et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nat Genet.* 2011; 44:67–72.
12. Mahajan A, Go MJ, Zhang W, Below JE, Gaulton KJ, Ferreira T, Horikoshi M, Johnson AD, Ng MC, Prokopenko I, Saleheen D, Wang X, Zeggini E, et al, and Diabetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, and Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium, and South Asian Type 2 Diabetes (SAT2D) Consortium, and Mexican American Type 2 Diabetes (MAT2D) Consortium, and Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in multi-Ethnic Samples (T2D-GENES) Consortium. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet.* 2014; 46:234–244.
13. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, et al. Finding the missing heritability of complex diseases. *Nature.* 2009; 461:747–753.
14. Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, Andersen G, Ng DP, Holmkvist J, Borch-Johnsen K, Jørgensen T, Sandback A, Lauritzen T, Hansen T, et al. SNPs in *KCNQ1* are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat Genet.* 2008; 40:1098–1102.
15. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, Hirota Y, Mori H, Jonsson A, Sato Y, Yamagata K, Hinokio Y, Wang HY, et al. Variants in *KCNQ1* are associated with susceptibility to type 2 diabetes mellitus. *Nat Genet.* 2008; 40:1092–1097.
16. Yamauchi T, Hara K, Maeda S, Yasuda K, Takahashi A, Horikoshi M, Nakamura M, Fujita H, Grarup N, Cauchi S, Ng DP, Ma RC, Tsunoda T, et al. A genome-wide association study in the Japanese population identifies susceptibility loci for type 2 diabetes at *UBE2E2* and *C2CD4A-C2CD4B*. *Nat Genet.* 2010; 42:864–868.
17. Imamura M, Maeda S, Yamauchi T, Hara K, Yasuda K, Morizono T, Takahashi A, Horikoshi M, Nakamura M, Fujita H, Tsunoda T, Kubo M, Watada H, et al. A single-nucleotide polymorphism in *ANK1* is associated with susceptibility to type 2 diabetes in Japanese populations. *Hum Mol Genet.* 2012; 21:3042–3049.
18. Hara K, Fujita H, Johnson TA, Yamauchi T, Yasuda K, Horikoshi M, Peng C, Hu C, Ma RC, Imamura M, Iwata M, Tsunoda T, Morizono T, et al. Genome-wide association study identifies three novel loci for type 2 diabetes. *Hum Mol Genet.* 2014; 23:239–246.
19. Imamura M, Takahashi A, Yamauchi T, Hara K, Yasuda K, Grarup N, Zhao W, Wang X, Huerta-Chagoya A, Hu C, Moon S, Long J, Kwak SH, et al. Genome-wide association studies in the Japanese population identify seven novel loci for type 2 diabetes. *Nat Commun.* 2016; 7:10531.
20. Góth L, Nagy T. Acatlasemia and diabetes mellitus. *Arch Biochem Biophys.* 2012; 525:195–200.
21. Góth L, Eaton JW. Hereditary catalase deficiencies and increased risk of diabetes. *Lancet.* 2000; 356:1820–1821.
22. Vitai M, Fátai S, Rass P, Csordás M, Tarnai I. Simple PCR heteroduplex, SSCP mutation screening methods for the detection of novel catalase mutations in Hungarian patients with type 2 diabetes mellitus. *Clin Chem Lab Med.* 2005; 43:1346–1350.
23. Lou X, Bao R, Zhou CZ, Chen Y. Structure of the thioredoxin-fold domain of human phosphatidylethanolamine 3-phosphate phosphatase. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2009; 65:67–70.
24. Willardson BM, Howlett AC. Function of phosphatidylethanolamine 3-phosphate phosphatase in G protein signaling and chaperone-assisted protein folding. *Cell Signal.* 2007; 19:2417–2427.
25. Ahrén B. Islet G protein-coupled receptors as potential targets for treatment of type 2 diabetes. *Nat Rev Drug Discov.* 2009; 8:369–385.
26. Yang J, Kim O, Wu J, Qiu Y. Interaction between tyrosine kinase ETK and a RUN domain- and FYVE domain-containing protein RUFY1. A possible role of ETK in regulation of vesicle trafficking. *J Biol Chem.* 2002; 277:30219–30226.
27. Wang Z, Thurmond DC. Mechanisms of biphasic insulin-granule exocytosis—roles of the cytoskeleton, small GTPases and SNARE proteins. *J Cell Sci.* 2009; 122:893–903.
28. Austin-Tse C, Halbritter J, Zariwala MA, Gilberti RM, Gee HY, Hellman N, Pathak N, Liu Y, Panizzi JR, Patel-King RS, Tritschler D, Bower R, O’Toole E, et al. Zebrafish ciliopathy screen plus human mutational analysis identifies

- C21orf59 and CCDC65 defects as causing primary ciliary dyskinesia. *Am J Hum Genet.* 2013; 93:672–686.
29. Oh EC, Vasanth S, Katsanis N. Metabolic regulation and energy homeostasis through the primary cilium. *Cell Metab.* 2015; 21:21–31.
  30. Lee H, Song J, Jung JH, Ko HW. Primary cilia in energy balance signaling and metabolic disorder. *BMB Rep.* 2015; 48:647–654.
  31. Yamada Y, Sakuma J, Takeuchi I, Yasukochi Y, Kato K, Oguri M, Fujimaki T, Horibe H, Muramatsu M, Sawabe M, Fujiwara Y, Taniguchi Y, Obuchi S, et al. Identification of STXBP2 as a novel susceptibility locus for myocardial infarction in Japanese individuals by an exome-wide association study. *Oncotarget.* 2017; 8:33527–33535. <https://doi.org/10.18632/oncotarget.16536>.
  32. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care.* 2003; 26:S5–S20.
  33. Kuzuya T, Nakagawa S, Satoh J, Kanazawa Y, Iwamoto Y, Kobayashi M, Nanjo K, Sasaki A, Seino Y, Ito C, Shima K, Nonaka K, Kadowaki T, et al. Report of the Committee on the classification and diagnostic criteria of diabetes mellitus. *Diabetes Res Clin Pract.* 2002; 55:65–85.
  34. World Health Organization, International Diabetes Federation. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia: report of a WHO/IDF consultation. Geneva, Switzerland: World Health Organization; 2006. pp. 1–46.
  35. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006; 38:904–909.