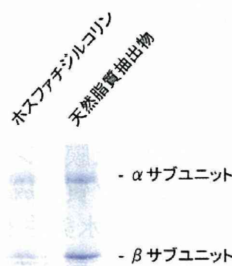

図1　大腸菌無細胞タンパク質合成技術によるチャネル分子調製
発現系に添加するリポソームの脂質種を天然脂質抽出物に変更したことで発現量が著しく上昇した。

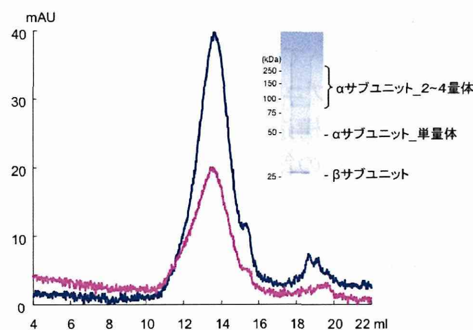

図2　大腸菌無細胞タンパク質合成技術により調製したチャネル分子の精製
ゲル濾過クロマトグラフィーで単分散性ピークを示す高純度標品の単離に成功。

可溶化工程では、数種の界面活性剤について検討を行い、可溶化量と精製時の安定化を両立した効果的な界面活性剤としてドデシルマルトシド（DDM）を見出した。さらに、αサブユニットのＮ末端にヒスチジンタグ付き可溶性タンパク質を融合することで、可溶化量の改善が認められた。また、プロテオリポソーム上に共発現させたαとβ両サブユニットも、αサブユニット単体を発現させたものと同様の高い可溶化量を示すことを確認した。

アフィニティカラム精製では、2種のアフィニティカラムを併用することで、高純度かつ高収量の試料調製が可能となった。この際、複数種の精製タグをそれぞれ連結し、発現量と可溶性を指標に最適なタグを選択した。さらに、αとβサブユニットを共発現させたものでは、αサブユニット側のヒスチジンタグでの精製により、高純度のα/βサブユニット複合体として精製出来ることを確認した。加えて、単独で精製したものよりも凝集傾向が軽減され、安定的に精製できることが判明した。

一方、このチャネル分子の精製において危惧される精製過程での断片化について、αサブユニット内の断片化部位を質量分析等で特定し、この部位を除去した変異体を作製して、断片化による収量低下を抑えることが出来た。このように、発現から精製までの諸条件とコンストラクトの最適化によって、単分散性のゲル濾過カラム溶出ピークを示す（図2）、α/βサブユニット複合体の高純度標品が調製可能となった。

D．考察

膜タンパク質にはそれぞれ適切な脂質環境があり、生細胞での局在化や機能発揮に重要な役割を果たしていることが知られている。このようなことから、無細胞合成タンパク質合成系では、目的の膜タンパク質に適した脂質を選択することが重要である。本課題において、大腸菌無細胞タンパク合成に標準的に用いているホスファチジルコリンから天然脂質抽出物に変更したことで、発現量が改善したこともこのような理由によるものと考えられる。今後、さらに脂質種を詳細検討し、さらなる発現量改善も期待できる。

一般的に、膜タンパク質の可溶性には、膜タンパク質の膜外領域（水溶性領域）の寄与が大きいと考えられている。本課題において、チャネルαサブユニットのN末端（細胞外領域）へ比較的大きな可溶性タンパク質を付加したことでチャネルの可溶性が改善したのは、その為であると推定される。断片化が顕著な領域を切除したαサブユニット変異体についても、高い可溶化量を維持することができた。また、βサブユニットとの複合体形成により精製過程での凝集傾向が改善されたことについて、αサブユニットとβサブユニットとの膜貫通領域や膜外領域を介した相互作用によって安定性を高めていることが主因であると考えられるが、βサブユニットとの複合体形成によるN端側およびC端側それぞれの水溶性領域の拡大による効果も一因である可能性がある。

βサブユニットとの複合体形成による水溶性領域の拡大は、結晶化でのcrystal contactにおいて非常に有効であることから、このαサブユニット/βサブユニット複合体を、立体構造解析に向けた結晶化スクリーニングでの第一候補と位置付けている。

E．結論

本研究により、大腸菌無細胞タンパク質合成系にてKCNQ1/KCNE複合体のより安定な発現条件と精製条件を見いだし、結晶構造解析達成に向けて大きく前進した。結晶構造解析が達成されれば、KCNQ1の糖尿病発症への関わり方を解明する糸口となることが期待される。

F．健康危険情報

非該当

G．研究発表

1. 論文発表

Tetrameric interaction of the ectoenzyme CD38 on the cell surface enables its catalytic and raft-Association activities. Hara-Yokoyama, M., Kukimoto-Niino, M., Terasawa, K., Harumiya, S., Podyma-Inoue, K. A., Hino, N., Sakamoto, K., Itoh, S., Hashii, N., Hiruta, Y., Kawasaki, N., Mishima-Tsumagari, C., Kaitsu, Y., Matsumoto, T., Wakiyama, M., Shirouzu, M., Kasama, T., Takayanagi, H., Utsunomiya-Tate, N., Takatsu, K., Katada, T., Hirabayashi, Y., Yokoyama, S. and Yanagishita, M., *Structure*, **20** (9), 1585-1595 (2012).

Rotation mechanism of *Enterococcus hirae* V_1-ATPase based on asymmetric crystal structures. Arai, S., Saijo, S., Suzuki, K., Mizutani, K., Kakinuma, Y., Ishizuka-Katsura, Y., Ohsawa, N., Terada, T., Shirouzu, M., Yokoyama, S., Iwata, S., Yamato, I. and Murata, T., *Nature*, **493**(7434), 703-707(2013)

2. 学会発表

「構造解析を目指した膜タンパク質の生産」～無細胞タンパク質合成系を中心として～，染谷友美，白水美香子，横山茂之，日本薬学会関東支部　第３７回学術講演会（2012年12月）

H．知的財産権の出願・登録状況
（予定を含む。）

1. 特許取得

無し

2. 実用新案登録

無し

研究成果の刊行に関する一覧表

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Fujita H, Hara K, Shojima N, Horikoshi M, Iwata M, Hirota Y, Tobe K, Seino S, Kadowaki T.	Variations with modest effects have an important role in the genetic background of type 2 diabetes and diabetes-related traits.	J Hum Genet	57(12)	776-9	2012
Iwata M, Maeda S, Kamura Y, Takano A, Kato H, Murakami S, Higuchi K, Takahashi A, Fujita H, Hara K, Kadowaki T, Tobe K.	Genetic risk score constructed using 14 susceptibility alleles for type 2 diabetes is associated with the early onset of diabetes and may predict the future requirement of insulin injections among Japanese individuals.	Diabetes Care.	35(8)	1763-70	2012
Kodama K, Horikoshi M, Toda K, Yamada S, Hara K, Irie J, Sirota M, Morgan AA, Chen R, Ohtsu H, Maeda S, Kadowaki T, Butte AJ.	Expression-based genome-wide association study links the receptor CD44 in adipose tissue with type 2 diabetes.	Proc Natl Acad Sci U S A	109(18)	7049-54	2012
Waki H, Yamauchi T, Kadowaki T.	The epigenome and its role in diabetes.	Curr Diab Rep	12	673-685	2012
脇　裕典	Keystone Symposia, Genetic and Molecular Basis of Obesity and Body Weight Regulation (J7), Pathogenesis of Diabetes (J8)に参加して	肥満研究	Vol. 18 No. 1	61-65	2012

脇　裕典	脂肪細胞分化特異的エンハンサーのゲノムワイド解析と分化におけるNFIの役割	糖尿病学2012		35-43	2012
脇　裕典、山内敏正、門脇　孝	白色脂肪細胞のエピジェネティックス	医学のあゆみ	242(12)	918-923	2012
高本偉碩	Wntシグナルと膵内分泌細胞分化	内分泌・糖尿病・代謝内科	36(3)	204-210	2013
Hara-Yokoyama M, Kukimoto-Niino M, Terasawa K, Harumiya S, Podyma-Inoue K A, Hino N, Sakamoto K, Itoh S, Hashii N, Hiruta Y, Kawasaki N, Mishima-Tsumagari C, Kaitsu Y, Matsumoto T, Wakiyama M, Shirouzu M, Kasama T, Takayanagi H, Utsunomiya-Tate N, Takatsu K, Katada T, Hirabayashi Y, Yokoyama S, Yanagishita M.	Tetrameric interaction of the ectoenzyme CD38 on the cell surface enables its catalytic and raft-Association activities.	Structure	20	1585-1595	2012
Arai S, Saijo S, Suzuki K, Mizutani K, Kakinuma Y, Ishizuka-Katsura Y, Ohsawa N, Terada T, Shirouzu M, Yokoyama S, Iwata S, Yamato I, Murata T.	Rotation mechanism of Enterococcus hirae V1-ATPase based on asymmetric crystal structures.	Nature	493	703-707	2013

Journal of Human Genetics (2012) 57, 776–779
© 2012 The Japan Society of Human Genetics All rights reserved 1434-5161/12
www.nature.com/jhg

ORIGINAL ARTICLE

Variations with modest effects have an important role in the genetic background of type 2 diabetes and diabetes-related traits

This article has been corrected since Advance Online Publication, and a corrigendum is also printed in this issue.

Hayato Fujita[1], Kazuo Hara[1], Nobuhiro Shojima[1], Momoko Horikoshi[1], Minoru Iwata[2], Yushi Hirota[3], Kazuyuki Tobe[2], Susumu Seino[3] and Takashi Kadowaki[1]

The aim of the present study was to explore the role of variations with modest effects (previously identified by a large-scale meta-analysis in European populations) in the genetic background of type 2 diabetes (T2D) and diabetes-related traits in a Japanese population. We enrolled 2632 Japanese subjects with T2D and 2050 non-diabetic subjects. We analyzed nine single-nucleotide polymorphisms (SNPs), including rs340874 (*PROX1*), rs4607517 (*GCK*), rs2191349 (*DGKB-TMEM195*), rs7034200 (*GLIS3*), rs10885122 (*ADRA2A*), rs174550 (*FADS1*), rs11605924 (*CRY2*), rs10830963 (*MTNR1B*) and rs35767 (*IGF1*). rs340874 (*PROX1*) and rs174550 (*FADS1*) were significantly associated with T2D ($P=0.0078$, OR: 1.12; and $P=0.0071$, OR: 1.12, respectively). Subjects with more risk alleles related to nine SNPs had an increased risk of T2D ($P=0.0017$), as well as a higher fasting plasma glucose level ($P=0.018$), higher HbA_{1c} level ($P=0.013$) and lower HOMA-β ($P=0.033$) compared with subjects who had fewer risk alleles. We identified a significant association of a SNP of *FADS1* and a SNP near *PROX1* with T2D in a Japanese population. The present findings suggest that inclusion of SNPs with a tendency to increase the disease risk captured more of the genetic background of T2D than that revealed by only assessing significant SNPs.
Journal of Human Genetics (2012) **57,** 776–779; doi:10.1038/jhg.2012.110; published online 20 September 2012

Keywords: common disease; single-nucleotide polymorphism; type 2 diabetes

INTRODUCTION

Type 2 diabetes (T2D) is a complex disorder characterized by hyperglycemia that results from impaired pancreatic β-cell function and decreased action of insulin on its target tissues. T2D is one of the most common polygenic disorders, in which each genetic variation has a partial and additive effect.[1] Over 50 susceptibility loci for T2D have been identified through genome-wide association studies (GWAS),[2–4] most of which were conducted in European populations. *KCNQ1*, *UBE2E2* and *C2CD4A/B* are among the few T2D loci discovered in non-European populations.[5–7]

Recently, new genetic loci were found to be associated with fasting glucose homeostasis in a meta-analysis of 21 GWAS of European non-diabetic subjects (the Meta-Analyses of Glucose and Insulin-related traits Consortium: MAGIC).[4] The contribution of these new loci should be investigated further, because it is well known that there are significant differences in the frequencies of some genetic variations among different ethnic groups. Possibly because of the modest size of the effect of most variations, however, it is hard to replicate the results of studies that have analyzed large samples in other populations. In fact, only a handful of loci (*CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *TCF7L2*, *SLC30A8*, *HHEX*, *KCNJ11* and *IRS-1*) have been confirmed to be associated with T2D in Japanese populations.[6,8–10]

In the meantime, arguments against the common disease–common variant hypothesis have been proposed, suggesting that rare variants with a large effect have a substantial role in the genetic architecture of common diseases. So far, we only have limited examples of rare variants conferring a markedly increased risk of common diseases. It has also recently been proposed that a very large number of common variants with effects that cannot be detected by studies employing a conventional sample size may be involved in the genetic architecture of common diseases, which is the so-called 'infinitesimal' hypothesis.[11]

Even if individual small effects are not detectable by a single SNP (single-nucleotide polymorphism) analysis, we may be able to detect their cumulative effect by combined SNP analysis.

Accordingly, the aim of the present study was to determine whether the loci identified in the MAGIC study comprise part of

[1]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [2]First Department of Internal Medicine, Toyama University, Toyama, Japan and [3]Department of Internal Medicine, Division of Diabetes, Metabolism and Endocrinology, Kobe University Graduate School of Medicine, Kobe, Japan
Correspondence: Dr K Hara, Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-0033, Japan.
E-mail: thara-tky@umin.ac.jp
Received 11 April 2012; revised 10 July 2012; accepted 8 August 2012; published online 20 September 2012

the genetic component of diabetes-related traits and T2D in a Japanese population by using combined SNP analysis.

MATERIALS AND METHODS

Cases and controls

Written informed consent was obtained from all the participants. We enrolled 1200 subjects with T2D aged 65.5 ± 9.5 years (body mass index (BMI): 24.3 ± 3.7 kg m^{-2} (mean ± s.d.)) who attended the University of Tokyo Hospital (Tokyo, Japan). The control group consisted of 855 persons aged 69.5 ± 6.8 years (BMI: 24.5 ± 3.2 kg m^{-2}) who participated in an annual health check conducted by the Hiroshima Atomic Bomb Casualty Council Health Management Center (Hiroshima, Japan). We also enrolled 722 subjects with T2D aged 64.9 ± 11.1 years (BMI: 24.6 ± 3.9 kg m^{-2}) and 763 controls aged 66.7 ± 10.6 years (BMI: 22.7 ± 3.3 kg m^{-2}) who attended the University of Toyama Hospital (Toyama, Japan) and 710 patients aged 60.9 ± 10.2 years (BMI: 24.1 ± 3.7 kg m^{-2}) and 432 controls aged 75.2 ± 8.0 years (BMI: 21.5 ± 3.6 kg m^{-2}) from the University of Kobe (Kobe, Japan).

Diabetes was diagnosed according to the World Health Organization criteria.[12] Exclusion criteria included positivity for antibody to glutamic acid decarboxylase or diabetes secondary to (i) liver dysfunction, (ii) steroids or other drugs that might elevate glucose levels, (iii) malignancy or (iv) any monogenic disorder known to cause diabetes. Inclusion criteria for the control group were as follows: (i) $HbA_{1c} < 6.2\%$, (ii) fasting plasma glucose $< 7.0\ mmol\ l^{-1}$ and (iii) postprandial plasma glucose $< 11.1\ mmol\ l^{-1}$. HbA_{1c} was estimated according to the method of the Japan Diabetes Society and converted to the corresponding National Glycohemoglobin Standardization Program (NGSP) value.

Genotyping

We genotyped SNPs that showed a significant association with diabetes-related traits or phenotypes in the MAGIC study. SNPs with a minor allele frequency of less than 0.05 in the Japanese population (HapMap JPT) were excluded from this study. We also excluded SNPs for which the association with T2D has already been reported in samples from Tokyo or Toyama.[7,13–15] As a result, a total of nine SNPs (rs340874, rs4607517, rs2191349, rs7034200, rs10885122, rs174550, rs11605924, rs10830963 and rs35767) were genotyped.

SNP genotyping was performed by using the TaqMan assay (Applied Biosystems, Foster City, CA, USA). The success rates of the assay was >95%.

Statistical analysis

In the non-diabetic subjects, the genotype distributions of all SNPs were in Hardy–Weinberg equilibrium ($P \geqslant 0.05$). Using the non-diabetic subjects, we performed a quantitative trait analysis to assess the association between each SNP and indexes of glucose homeostasis (fasting plasma glucose, fasting insulin, HbA_{1c}, HOMA-β and HOMA-IR). Differences of the genotype distribution between the diabetic and non-diabetic groups were also assessed by logistic regression analysis with an additive model. We performed the above analyses separately on groups from the University of Tokyo, University of Toyama and University of Kobe, and then performed joint analyses.

We analyzed the combined effect of multiple SNPs. Risk alleles were selected from the MAGIC study and we calculated how many risk alleles there were in each sample for the nine SNPs analyzed in this study. Samples with 12 or more risk alleles were classified as the high-risk group, samples with 9–11 risk alleles were classified as the medium-risk group and samples with 8 or fewer risk alleles were classified as the low-risk group. Differences in the distribution of the cases and controls between the risk groups were analyzed. Using the control samples, we analyzed differences of diabetes-related traits between the risk groups. After excluding two SNPs (*PROX1* and *FADS1*) that were significantly associated with T2D in the present study, we also analyzed the combined influence of the remaining seven SNPs. In this analysis, the high-risk group was defined as having 10 or more risk alleles, the medium-risk group had 7–9 risk alleles and the low-risk group had 6 or fewer risk alleles. We assessed the differences in the prevalence of T2D by using a logistic regression model and diabetes-related traits by using a linear regression model. We calculated odds ratios per one risk-group increase for T2D and effect sizes per 1 risk-group increase for diabetes-related traits. Effect sizes and s.e. were calculated for groups from the University of Tokyo, University of Toyama and University of Kobe, after which combined analysis was performed.

As dependent variables in the linear regression models, we used log-transformed data for fasting insulin, HOMA-IR and HOMA-β, whereas untransformed data were used for fasting plasma glucose and HbA_{1c}.

Multiple corrections were done by the Benjamini–Hochberg method.[16] A false discovery rate (FDR) of 0.05 was employed as the significance threshold.

These analyses were performed with R software (http://www.R-project.org).

RESULTS

We found that 2 SNPs (rs340874 and rs174550) were significantly associated with T2D (Table 1). In the case of rs340874 (*PROX1*), a minor allele was associated with an increased risk of T2D ($P = 0.0078$, FDR-adjusted $P = 0.035$, OR: 1.12, 95% CI: 1.03–1.22). For rs174550 (*FADS1*), a major allele was associated with an increased risk of T2D ($P = 0.0071$, FDR-adjusted $P = 0.035$, OR: 1.12, 95% CI: 1.03–1.22). In addition, rs2191349 (*DGKB-TMEM195*) was nominally associated with T2D ($P = 0.035$, FDR-adjusted $P = 0.11$, OR: 1.10, 95% CI: 1.01–1.21). No associations were observed between T2D and the remaining loci in our study population.

In the non-diabetic subjects, there was a significant association with HbA_{1c} for rs4607517 (near *GCK*, $P = 0.00056$, FDR-adjusted $P = 0.0050$) (Table 2). However, none of the loci were significantly associated with fasting plasma glucose, fasting plasma insulin, HOMA-IR or HOMA-β (Supplementary Tables 1–4).

The higher risk groups had an increased risk of T2D ($P = 0.0017$, OR: 1.17, 95% CI: 1.06–1.29) in the combined analysis (Table 3). It is noteworthy that the higher risk groups had an increased risk of T2D even after excluding the two SNPs (near *PROX1* and at *FADS1*) that were significantly associated with T2D in the individual SNP analyses ($P = 0.011$, OR: 1.14, 95% CI: 1.03–1.27) (Table 3). The higher risk groups had higher fasting plasma glucose levels and HbA_{1c} levels ($P = 0.018$ and $P = 0.013$, respectively) (Table 4). The higher risk groups had a lower HOMA-β ($P = 0.033$), but no differences of HOMA-IR or fasting plasma insulin levels were found between the risk groups.

DISCUSSION

In the present study, we analyzed nine SNPs that were previously shown to be associated with diabetes-related traits by the MAGIC study. We found that a SNP of *FADS1* and a SNP near *PROX1* were significantly associated with an increased risk of T2D in a Japanese population. We also identified an association between a SNP of *GCK* and elevation of HbA_{1c} in a non-diabetic Japanese population.

To our knowledge, this is the first study to identify the association of T2D with a SNP at *FADS1*. In the MAGIC study, *FADS1* was associated with fasting plasma glucose levels within the physiological range, but not with pathological glucose levels.[4] However, this was not the case in the present study because *FADS1* was associated with susceptibility to T2D in our Japanese population.

The present study is the first to identify a SNP near *PROX1* related to susceptibility to T2D in a Japanese population. The contribution of a SNP near *PROX1* to the pathogenesis of T2D has now been demonstrated across different ethnicities because the risk allele in the present study was directionally consistent with those found in other populations.[4,17]

In the combined SNP analysis, the higher risk groups had an increased risk of T2D. This suggests that these SNPs had an additive effect on the risk of T2D. Unexpectedly, an increased risk of T2D was

Table 1 Association of the SNPs with type 2 diabetes

Chr	Gene	SNP	Allele (risk/other)[a]	No. (cases) No. (controls)	RAF (cases)[b] RAF (controls)[b]	I^{2c}	OR (95% CI)	P-value (FDR-adjusted P)[d]
1	*PROX1*	rs340874	G/A	2576	0.39	0	1.12 (1.03–1.22)	0.0078 (0.035)
				2013	0.37			
7	*GCK*	rs4607517	A/G	2571	0.21	0	0.96 (0.87–1.07)	0.49 (0.88)
				2001	0.21			
7	*DGKB-TMEM195*	rs2191349	T/G	2564	0.70	0	1.10 (1.01–1.21)	0.035 (0.11)
				1992	0.68			
9	*GLIS3*	rs7034200	A/C	2565	0.42	0	1.09 (1.00–1.18)	0.053 (0.119)
				1972	0.40			
10	*ADRA2A*	rs10885122	G/T	2561	0.91	0	1.00 (0.86–1.15)	0.96 (0.98)
				1994	0.91			
11	*FADS1*	rs174550	T/C	2550	0.64	0	1.12 (1.03–1.22)	0.0071 (0.035)
				1984	0.61			
11	*CRY2*	rs11605924	A/C	2588	0.18	30	1.00 (0.90–1.11)	0.98 (0.98)
				2001	0.18			
11	*MTNR1B*	rs10830963	G/C	2592	0.42	0	0.99 (0.91–1.08)	0.80 (0.98)
				2017	0.42			
12	*IGF1*	rs35767	C/T	2557	0.65	0	1.01 (0.93–1.10)	0.80 (0.98)
				1986	0.65			

Abbreviations: CI, confidence interval; FDR, false discovery rate; OR, odds ratio; SNP, single-nucleotide polymorphism.
[a]The risk alleles were determined according to the results in European populations.[4]
[b]RAF denotes the risk allele frequency.
[c]Heterogeneity was assessed by using the I^2 index.
[d]FDR-adjusted *P*-value for multiple hypotheses, using the Benjamini–Hochberg method.

Table 2 Association of the SNPs with HbA_{1c}

Chr	Gene	SNP	Alleles (effect/ other)[a]	No.	Effect (%) (s.e.)[b]	P-value (FDR-adjusted P)[c]
1	*PROX1*	rs340874	G/A	1837	0.013 (0.017)	0.450 (0.59)
7	*GCK*	rs4607517	A/G	1830	0.040 (0.012)	0.00056 (0.0050)
7	*DGKB-TMEM195*	rs2191349	T/G	1822	0.001 (0.015)	0.94 (0.94)
9	*GLIS3*	rs7034200	A/C	1816	0.008 (0.015)	0.060 (0.23)
10	*ADRA2A*	rs10885122	G/T	1825	0.027 (0.016)	0.10 (0.23)
11	*FADS1*	rs174550	T/C	1824	0.015 (0.010)	0.13 (0.23)
11	*CRY2*	rs11605924	A/C	1843	0.009 (0.012)	0.47 (0.59)
11	*MTNR1B*	rs10830963	G/C	1851	0.011 (0.017)	0.52 (0.59)
12	*IGF1*	rs35767	C/T	1831	0.016 (0.010)	0.11 (0.23)

Abbreviations: FDR, false discovery rate; SNP, single-nucleotide polymorphism.
[a]The effect allele was determined according to the results in European populations.[4]
[b]Per-allele effect and s.e. for quantitative traits were estimated.
[c]FDR-adjusted *P*-value for multiple hypotheses, using the Benjamini–Hochberg method.

Table 3 Type 2 diabetes in the risk groups

SNPs[a]	No. (low risk)[b]	No. (medium risk)[c]	No. (high risk)[d]	I^{2e}	OR (95% CI)[f]	P-value
9	701	2241	1007	0	1.17 (1.06–1.29)	0.0017
7	588	2584	900	0	1.14 (1.03–1.27)	0.011

Abbreviations: CI, confidence interval; OR, odds ratio; SNP, single-nucleotide polymorphism.
[a]Number of SNPs used to calculate the genetic risk. Nine SNPs were used to calculate the genetic risk (above). Seven SNPs were used after removing two SNPs that were significantly associated with type 2 diabetes in the individual SNP analyses (below).
[b]The low-risk group had 8 or fewer risk alleles around the nine SNPs (above), or 6 or fewer risk alleles around the seven SNPs (below).
[c]The medium-risk group had 9–11 risk alleles around the nine SNPs (above), or 7–9 risk alleles around the seven SNPs (below).
[d]The high-risk group had 12 or more risk alleles around the nine SNPs (above), or 10 or more risk alleles around the seven SNPs (below).
[e]Heterogeneity was assessed by using the I^2 index.
[f]Odds ratios per one risk-group increase were calculated.

also found even after we excluded the two SNPs that were significantly associated with T2D. Subjects with multiple risk alleles in the MAGIC study had a higher fasting plasma glucose, higher HbA_{1c} and lower HOMA-β, even though few or none of their SNPs were significantly associated with diabetes-related traits in the individual analyses. These results suggest that SNPs with an effect that is too small to detect in individual SNP analyses have a role in the genetic architecture of T2D and support the infinitesimal hypothesis. It seems that we are able to detect their cumulative effect by combined SNP analysis.

In this study, correction for multiple testing was done by the Benjamini–Hochberg method. This controls the FDR and is less conservative than methods that control the family-wise error rate (FWER), such as Bonferroni's correction. The seven SNPs with no significant association in the individual analyses had a significant influence on T2D in the combined analysis, which suggests that methods controlling the FWER may be too conservative and methods controlling the FDR may be preferable in such studies. Similarly, the standard GWAS threshold of 5×10^{-8} may be too strict. It could be useful to calculate a genetic risk score by using all SNPs beyond the nominal statistical thresholds in GWAS and determine whether the genetic score predicts the phenotype in another sample, but further investigation of this point is needed.

The higher risk groups had lower HOMA-β, but there were no differences of HOMA-IR or fasting plasma insulin levels between the

Table 4 Diabetes-related traits in the risk groups.

Phenotype	No. (low risk)[a]	No. (medium risk)[b]	No. (high risk)[c]	I^{2d}	Effect (s.e.)[e]	P-value
FPG ($mmol\,l^{-1}$)[f]	226	685	288	0	0.054 (0.023)	0.018
HbA_{1c} (%)	275	837	350	0	0.026 (0.011)	0.013
Fasting insulin ($pmol\,l^{-1}$)	227	687	286	0	−0.022 (0.023)	0.34
HOMA-IR	219	658	279	0	−0.010 (0.024)	0.67
HOMA-β	219	657	279	0	−0.051 (0.024)	0.033

[a]The low-risk group had 8 or fewer risk alleles around the nine SNPs.
[b]The medium-risk group had 9–11 risk alleles around the nine SNPs.
[c]The high-risk group had 12 or more risk alleles around the nine SNPs.
[d]Heterogeneity was assessed by using the I^2 index.
[e]Effect sizes per one risk-group increase were calculated.
[f]Fasting plasma glucose.

risk groups. This was consistent with the results of the MAGIC study, which showed that most of the nine SNPs were associated with HOMA-β and only two SNPs were associated with HOMA-IR or fasting insulin.

We could not replicate all of the MAGIC study findings. One possible reason is that the linkage disequilibrium (LD) relationships of disease-susceptibility alleles may differ between populations. Because the LD block in the *CRY2* regions is much larger in CEU than in JPT according to the HapMap database, it is possible that the causative SNPs under investigation are in different LD blocks in the Japanese population but are in the same LD in an European population. The fact that Japanese subjects with T2D tend to be leaner and less hyperinsulinemic than Europeans may be one of the reasons why SNPs associated with fasting insulin levels or HOMA-IR were not replicated.

Because our sample size did not have enough power to detect susceptibility genes with modest effect sizes, it cannot be denied that the other loci investigated in the current study could have some role in predisposing Japanese to T2D. In the quantitative analysis of the association between each SNP and diabetes-related traits except for HbA_{1c}, only 1344 samples were included, which is one of the reasons we could not replicate the associations between SNPs and diabetes-related traits, except for that between *GCK* and HbA_{1c}.

In conclusion, the present study was the first to demonstrate that a SNP of *FADS1* contributes to susceptibility to T2D. We also identified a significant association between T2D and a SNP near *PROX1*. Furthermore, our results indicated that analysis of SNPs with a small contribution to disease risk could capture more of the genetic background of T2D than assessment of only the significant SNPs.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the participating doctors and staff from collaborating institutions for providing DNA samples. We also thank N Miyama and M Yamaguchi for technical assistance. Furthermore, we thank the technical staff at the Laboratory of the First Department of Internal Medicine, Faculty of Medicine, Toyama University; and Division of Diabetes, Metabolism and Endocrinology, Department of Internal Medicine, Kobe University Graduate School of Medicine for their technical assistance.

1 Stumvoll, M., Goldstein, B. J. & van Haeften, T. W. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* **365**, 1333–1346 (2005).
2 Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D. *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881–885 (2007).
3 Scott, L. J., Mohlke, K. L., Bonnycastle, L. L., Willer, C. J, Li, Y, Duren, W. L. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).
4 Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A. U. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–U32 (2010).
5 Yasuda, K., Miyake, K., Horikawa, Y., Hara, K., Osawa, H., Furuta, H. *et al.* Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nat. Genet.* **40**, 1092–1097 (2008).
6 Unoki, H., Takahashi, A., Kawaguchi, T., Hara, K., Horikoshi, M., Andersen, G. *et al.* SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat. Genet.* **40**, 1098–1102 (2008).
7 Yamauchi, T., Hara, K., Maeda, S., Yasuda, K., Takahashi, A., Horikoshi, M. *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.* **42**, 864–868 (2010).
8 Takeuchi, F., Serizawa, M., Yamamoto, K., Fujisawa, T., Nakashima, E., Ohnaka, K. *et al.* Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. *Diabetes* **58**, 1690–1699 (2009).
9 Mori, S., Tanaka, Y., Takahashi, A., Hirose, H., Kashiwagi, A., Kaku, K *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).
10 Ohshige, T., Iwata, M., Omori, S., Tanaka, Y., Hirose, H., Kaku, K. *et al.* Association of new loci identified in European genome-wide association studies with susceptibility to type 2 diabetes in the Japanese. *PLoS One* **6**, e26911 (2011).
11 Gibson, G. Rare and common variants: twenty arguments. *Nat. Rev. Genet.* **13**, 135–145 (2011).
12 Alberti, K. G. & Zimmet, P. Z. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet. Med.* **15**, 539–553 (1998).
13 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).
14 Horikoshi, M., Hara, K., Ito, C., Shojima, N., Nagai, R., Ueki, K. *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).
15 Iwata, M., Maeda, S., Kamura, Y., Takano, A., Kato, H., Murakami, S. *et al.* Genetic risk score constructed using 14 susceptibility alleles for type 2 diabetes is associated with the early onset of diabetes and may predict the future requirement of insulin injections among Japanese individuals. *Diabetes Care* **35**, 1763–1770 (2012).
16 Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. Ser. B* **57**, 289–300 (1995).
17 Hu, C., Zhang, R., Wang, C., Wang, J., Ma, X., Hou, X. *et al.* Variants from GIPR, TCF7L2, DGKB, MADD, CRY2, GLIS3, PROX1, SLC30A8 and IGF1 are associated with glucose metabolism in the Chinese. *PLoS One* **5**, e15542 (2010).

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

Genetic Risk Score Constructed Using 14 Susceptibility Alleles for Type 2 Diabetes Is Associated With the Early Onset of Diabetes and May Predict the Future Requirement of Insulin Injections Among Japanese Individuals

Minoru Iwata, md, phd[1]
Shiro Maeda, md, phd[2]
Yutaka Kamura, md[1]
Atsuko Takano, md, phd[3]
Hiromi Kato, md, phd[4]
Shihou Murakami, md, phd[5]
Kiyohiro Higuchi, md, phd[6]
Atsushi Takahashi, phd[7]
Hayato Fujita, md[8]
Kazuo Hara, md, phd[8]
Takashi Kadowaki, md, phd[8]
Kazuyuki Tobe, md, phd[1]

OBJECTIVE—We evaluated the clinical usefulness of a genetic risk score (GRS) based on 14 well-established variants for type 2 diabetes.

RESEARCH DESIGN AND METHODS—We analyzed 14 SNPs at *HHEX*, *CDKAL1*, *CDKN2B*, *SLC30A8*, *KCNJ11*, *IGF2BP2*, *PPARG*, *TCF7L2*, *FTO*, *KCNQ1*, *IRS-1*, *GCKR*, *UBE2E2*, and *C2CD4A/B* in 1,487 Japanese individuals (724 patients with type 2 diabetes and 763 control subjects). A GRS was calculated according to the number of risk alleles by counting all 14 SNPs (T-GRS) as well as 11 SNPs related to β-cell function (β-GRS) and then assessing the association between each GRS and the clinical features.

RESULTS—Among the 14 SNPs, 4 SNPs were significantly associated with type 2 diabetes in the present Japanese sample ($P < 0.0036$). The T-GRS was significantly associated with type 2 diabetes ($P = 5.9 \times 10^{-21}$). Among the subjects with type 2 diabetes, the β-GRS was associated with individuals receiving insulin therapy ($\beta = 0.0131$, SE = 0.006, $P = 0.0431$), age at diagnosis ($\beta = -0.608$, SE = 0.204, $P = 0.0029$), fasting serum C-peptide level ($\beta = -0.032$, SE = 0.0140, $P = 0.022$), and C-peptide index ($\beta = -0.031$, SE = 0.012, $P = 0.0125$).

CONCLUSIONS—Our data suggest that the β-GRS is associated with reduced β-cell functions and may be useful for selecting patients who should receive more aggressive β-cell–preserving therapy.

Diabetes Care 35:1763–1770, 2012

From the [1]First Department of Internal Medicine, Faculty of Medicine, Toyama University, Toyama, Japan; the [2]Laboratory for Endocrinology and Metabolism, RIKEN Center for Genomic Medicine, Yokohama, Kanagawa, Japan; the [3]Division of Endocrinology and Metabolism, Department of Internal Medicine, Saiseikai Takaoka Hospital, Takaoka, Toyama, Japan; the [4]Division of Endocrinology and Metabolism, Department of Internal Medicine, Shakaihoken Takaoka Hospital, Takaoka, Toyama, Japan; the [5]Division of Endocrinology and Metabolism, Department of Internal Medicine, Toyama Rosai Hospital, Uozu, Toyama, Japan; the [6]Department of Internal Medicine, Itoigawa General Hospital, Itoigawa, Niigata, Japan; the [7]Laboratory for Statistical Analysis, RIKEN Center for Genomic Medicine, Yokohama, Kanagawa, Japan; and the [8]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan.
Corresponding author: Minoru Iwata, iwamino-tym@umin.ac.jp.
Received 13 October 2011 and accepted 27 March 2012.
DOI: 10.2337/dc11-2006
This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc11-2006/-/DC1.
A slide set summarizing this article is available online.
© 2012 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See http://creativecommons.org/licenses/by-nc-nd/3.0/ for details.

Type 2 diabetes affects nearly 300 million individuals worldwide, and its prevalence continues to increase in many countries, including Japan (1). Although the precise mechanisms underlying the development and progression of type 2 diabetes have not been elucidated, a combination of multiple genetic and/or environmental factors contribute to the pathogenesis of the disease (2,3). Impaired insulin secretion and insulin resistance, the two main pathophysiological mechanisms leading to type 2 diabetes, have a significant genetic component (4).

Recent studies have confirmed ~40 genetic loci associated with type 2 diabetes (5); most of these loci were discovered in genome-wide association studies (6–16), with the exception of *PPARG* (17), *KCNJ11* (18), and *WFS1* (19), which were identified using candidate gene approaches, and *TCF7L2*, which was discovered using a linkage-positional cloning strategy (20). Among them, many loci (at least 10), such as *MTNR1B*, *SLC30A8*, *THADA*, *TCF7L2*, *KCNQ1*, *CAMK1D*, *CDKAL1*, *IGF2BP2*, *HNF1B*, and *CENTD2*, have been shown to be associated with impaired β-cell functions, whereas only a few loci such as *PPARG*, *IRS1*, and *FTO* have been associated with insulin resistance (13).

Although the molecular mechanisms responsible for the susceptibility effect can be well assigned for some loci, such as those at *KCNJ11* and *SLC30A8*, the mechanisms by which most genetic loci contribute to the development of type 2 diabetes are not understood.

Recently, the construction of a genetic risk score (GRS) using information on these diabetes susceptibility loci has been shown to be useful for evaluating the risk of the development of type 2 diabetes in individuals (21–26). However, the currently available genetic information is

obviously insufficient for predicting the development of type 2 diabetes, and little is known about the detailed relationship between the GRS and the clinical features of type 2 diabetes. In the current study, we selected 14 well-replicated and well-established genetic variants associated with type 2 diabetes in the Japanese population (25,27–32) and constructed a GRS, which may predict mechanism (β-cell function and insulin resistance) of diabetes development, to evaluate the possibility that currently available genetic information can be translated into clinical practice.

Table 1—*Clinical characteristics of the participants*

	Type 2 diabetic	Control	*P*
n	724	763	
Sex (male/female)	451/273	359/404	<0.0001*
Age (years)	64.9 ± 11.1	72.5 ± 9.0	<0.001
Duration of diabetes (years)	13.6 ± 9.1		
Age at diagnosis (years)	51.4 ± 11.6		
Self-reported family history of diabetes (%)	55.7		
BMI (kg/m^2)	24.5 ± 3.9	22.7 ± 3.3	<0.0001
Maximum BMI (kg/m^2)	27.4 ± 4.3	24.7 ± 3.1	<0.0001
Waist circumference (cm)			
Males	87.1 ± 9.5	85.3 ± 8.1	<0.05
Females	87.4 ± 11.7	81.2 ± 9.1	<0.0001
FPG (mmol/L)	7.60 ± 1.88	5.33 ± 0.58	<0.0001
A1C (%)	7.53 ± 1.25	5.54 ± 0.25	<0.0001
eGFR (mL/min)	74.8 ± 21.6	72.3 ± 16.3	<0.05
HOMA-β (%)[a]	37.5 ± 40.2	60.8 ± 41.2	<0.0001
HOMA-IR (mol · μU/L^2)[a]	2.17 ± 1.60	1.24 ± 0.78	<0.0001
F-CPR (ng/mL)[b]	1.65 ± 0.85	1.49 ± 0.61	<0.0001
F-CPI[b]	1.25 ± 0.71	1.56 ± 0.61	<0.0001
Complications (%)			
Diabetic nephropathy	39.3		
Diabetic retinopathy	42.3		
Treatment of diabetes (%)			
Diet alone	13.4		
Using oral hypoglycemic agents	73.9		
Sulfonylureas	46.7		
Thiazolidinediones	20.3		
Biguanides	28.5		
α-Glucosidase inhibitor	33.7		
Glinide	4.9		
Using insulin	31.4		
Presence of hypertension (%)[c]	76.1	63.2	<0.0001*
Systolic blood pressure (mmHg)	130 ± 16	130 ± 18	0.5755
Diastolic blood pressure (mmHg)	75 ± 11	76 ± 11	<0.01
Presence of dyslipidemia (%)[d]	78.9	64.9	<0.0001*
LDL cholesterol (mg/dL)	112 ± 26	119 ± 29	<0.0001
HDL cholesterol (mg/dL)	53.9 ± 16.3	60.6 ± 16.5	<0.0001
Triglycerides (mg/dL)	122 ± 74	109 ± 67	<0.0001

Data are means ± SD. The value for A1C (%) was estimated as a National Glycohemoglobin Standardization Program (NGSP) equivalent value (%) calculated using the following formula: A1C (%) = A1C (Japan Diabetes Society) (%) + 0.4%. CPI was calculated using the following equation: CPI = (F-CPR/FPG) × 100. eGFR, estimated glomerular filtration rate. *Pearson χ^2 test. [a]HOMA-β and -IR were calculated in all participants except for those treated with insulin therapy. [b]F-CPR and CPI were calculated in all participants except for those with serum creatinine level >1.5 mg/dL. [c]Determination of hypertension was defined as systolic blood pressure ≥130 mmHg or diastolic blood pressure ≥85 mmHg or having been treated for hypertension. [d]Determination of dyslipidemia was defined as serum LDL cholesterol ≥120 mg/dL, serum triglycerides ≥150 mg/dL, or HDL cholesterol <40 mg/dL or having been treated for dyslipidemia.

RESEARCH DESIGN AND METHODS

—All patients with type 2 diabetes who regularly attended the outpatient clinics in five hospitals—University of Toyama Hospital (Toyama, Japan), Shakaihoken Takaoka Hospital, Saiseikai Takaoka Hospital, Nanto City Hospital (Nanto, Japan), and Asahi General Hospital (Asahi-machi, Japan)—were asked to participate in this study. Among them, informed consent was obtained from 724 patients between January 2008 and December 2009, and these 724 patients were enrolled in the current study as case subjects (62.3% male, mean ± SD age 64.9 ± 11.1 years, and A1C 7.5 ± 1.3%) (Table 1). We also enrolled control individuals (n = 763) selected from subjects who had undergone an annual health check-up at the Itoigawa General Hospital (Itoigawa, Japan), Aoi Hospital (Tonami, Japan), Amenithy Tsukioka Hospital (Toyama, Japan), Hida City Hospital (Hida, Japan), Sakurai Hospital (Kurobe, Japan), Hokuriku chuo Hospital (Oyabe, Japan), and the above five hospitals. The inclusion criteria for the nondiabetic control subjects were as follows: *1*) >50 years of age, *2*) A1C values <6.0%, *3*) no family history of type 2 diabetes in first- and second-degree relatives, and *4*) no past history of a diagnosis of diabetes. Diabetes was diagnosed based on the 1998 American Diabetes Association criteria (33). The exclusion criteria for the case subjects with diabetes were diabetes caused by *1*) liver dysfunction, *2*) steroids and other drugs that might increase glucose levels, *3*) malignancy, *4*) monogenic disorders known to cause diabetes, and *5*) individuals who tested positive for anti-GAD antibody. Characteristics of the participants are presented in Table 1.

We also performed an examination of another cohort for the association of GRS with type 2 diabetes (homeostasis model assessment [HOMA] of β-cell function or HOMA of insulin resistance [HOMA-IR]), which was conducted in Tokyo University, Tokyo, Japan (30) (type 2 diabetes cases, n = 1,182, 59.6% male, age 65.3 ± 9.5 years, and A1C 7.7 ± 1.6%; nondiabetic subjects, n = 859, 44.4% male, age 69.5 ± 6.8 years, and A1C 5.6 ± 0.2%) (Supplementary Table 1). The inclusion criteria for the nondiabetic control subjects and the exclusion criteria for the case subjects with diabetes were identical between the two studies, except for the age of control individuals >60 years in the Tokyo University study.

Collection of clinical information

We obtained clinical information including the current BMI, maximum BMI, family history of diabetes, age at diagnosis, blood chemistry (including plasma glucose, insulin level, serum C-peptide, and serum creatinine) at fasting state, diabetes complications, and use of antidiabetes drugs.

Patients who were required to inject >10 units of insulin a day continuously were regarded as undergoing insulin therapy.

Diabetic nephropathy was defined as having a urinary albumin-to-creatinine ratio ≥30 mg/gCr, determined in at least two consecutive overnight samples collected over a 3- to 6-month period. Patients diagnosed as having a urinary tract infection, other glomerular diseases, or gross hematuria were excluded.

All patients underwent ophthalmologic examinations, including funduscopic examination. We defined nonproliferative diabetic retinopathy, proliferative diabetic retinopathy, and a history of photocoagulation or vitrectomy as indicating the presence of diabetic retinopathy. All the study procedures were approved by the ethics committee of the University of Toyama, and informed consent was obtained from all of the participants.

Genotyping assay

Genomic DNA was extracted from peripheral blood (QIAamp DNA blood kit; QIAGEN, Hilden, Germany). We selected 14 single nucleotide polymorphisms (SNPs) at genetic loci that had been previously shown to be robustly associated with type 2 diabetes in seven recent studies performed in Japanese populations (25,27–32). The following SNPs were examined: in *KCNJ11* (rs5219), in *HHEX* (rs1111875), in *CDKAL1* (rs7756992), near *CDKN2B* (rs10811661), in *SLC30A8* (rs13266634), in *IGF2BP2* (rs4402960), in *PPARG* (rs1801282), in *TCF7L2* (rs7903146), in *FTO* (rs8050136), near *IRS-1* (rs2943641), in *GCKR* (rs780094), in *UBE2E2* (rs7612463), in *C2CD4A-C2CD4B* (rs7172432), and in *KCNQ1* (rs2237892). The genotyping of these SNPs was performed using TaqMan SNP Genotyping assays (Applied Biosystems, Foster City, CA) or a multiplex-PCR-invader assay as described previously (34,35).

The success rates for these assays were >95%, and the concordance rate, based on duplicate comparisons in 763 control participants and 724 type 2 diabetic patients, was 99.4%. A tagging approach to detect all variations completely covering each genomic region has not been used. Although no apparent deviations in the genotype distributions from Hardy-Weinberg equilibrium (HWE) were observed for all of the SNPs ($P \geq 0.001$) (6), some of them had borderline results for the HWE test (rs13266634 in control; rs2237892 in control) (Supplementary Table 2).

Construction of GRS

We combined the information on the 14 SNPs using an allele count model (21). To construct the GRS, we summed the number of risk alleles of all 14 SNPs included in this study in each individual, assuming an equal and additive effect of each allele (T-GRS). The T-GRS was distributed normally in both the control and the diabetic subjects.

We further classified these 14 genetic variants into two categories: *1)* 11 β-cell function–related SNPs (rs1111875 in *HHEX*, rs7756992 in *CDKAL1*, rs10811661 in *CDKN2B*, rs13266634 in *SLC30A8*, rs4402960 in *IGF2BP2*, rs7903146 in *TCF7L2*, rs780094 in *GCKR*, rs7612463 in *UBE2E2*, rs7172432 in *C2CD4A/B*, rs2237892 in *KCNQ1*, and rs5219 in *KCNJ11*) and *2)* three insulin resistance/obesity-related variants (rs1801282 in *PPARG*, rs8050136 in *FTO*, and rs2943641 in *IRS-1*), based on previously reported information (13). We then calculated the GRS of the β-cell function–related SNPs (β-GRS) and the insulin resistance and obesity-related SNPs (R-GRS). The β-GRS and R-GRS were also distributed normally in both the control and diabetic groups.

Statistical analysis

Differences in clinical features, such as the insulin secretory capacity and age at the time of the diagnosis of diabetes, between the risk allele groups were determined using ANOVA and multiple regression analysis after adjustments for related covariables. Results with *P* values <0.05 were considered statistically significant.

We performed HWE tests according to the method described by Nielsen et al. (36). The proportions of genotypes for each SNP were compared between the type 2 diabetic case and the nondiabetic control subjects using a multiple logistic regression analysis with or without adjustments for age, sex, and BMI. The allele-specific odds ratios (ORs) were calculated using logistic regression with or without adjustments for age, sex, and BMI. Variables with skewed distributions were logarithmically (natural) transformed for further analyses. Quantitative trait analyses were performed using a multiple linear regression analysis with or without adjustments for related covariables. Bonferroni correction was applied to correct for multiple testing errors, and $P < 0.0036$ (0.05 divided by 14: the total number of SNPs studied) was considered significant.

The effects of the GRS on the clinical features and quantitative metabolic traits were examined by calculating the β values for the risk allele score using linear generalized estimating equations. *P* values <0.05 were considered statistically significant for this analysis.

The statistical analyses were performed using JMP for Windows version 8.00 software (SAS Institute, Cary, NC). The power of the sample size for the current study to identify the association of previously reported SNP loci with type 2 diabetes was calculated using "CaTS power calculator for genetic studies" software (http://www.sph.umich.edu/csg/abecasis/CaTS/).

RESULTS

Associations of each of the 14 SNPs with type 2 diabetes and quantitative metabolic traits

Among the 14 SNPs from 14 loci, 4 SNPs (rs7756992 in *CDKAL1*, rs10811661 near *CDKN2B*, rs13266634 in *SLC30A8*, and rs2237892 in *KCNQ1*) were found to be significantly associated with type 2 diabetes (Supplementary Table 3) ($P = 1.7 \times 10^{-5}$, 7.5×10^{-6}, 2.8×10^{-3}, and 1.4×10^{-7}, respectively) after adjustments for age, sex, and BMI; the association of rs2237892 in *KCNQ1* was the strongest in the present Japanese sample, as reported previously (16,28). rs4402960 in *IGF2BP2*, rs2943641 near *IRS-1*, rs780094 in *GCKR*, and rs5219 in *KCNJ11* showed a nominal association with type 2 diabetes ($P = 0.010$, $P = 0.028$, $P = 0.013$, and $P = 0.033$, respectively), and rs7172432 in *C2CD4A/B* tended to be associated with type 2 diabetes ($P = 0.073$). As for rs7903146 in *TCF7L2*, rs1111875 in *HHEX*, rs1801282 in *PPARG*, rs8050136 in *FTO*, and rs7612463 in *UBE2E2*, we were unable to detect any SNPs that were significantly associated with type 2 diabetes in the present Japanese sample ($P = 0.659$, 0.773, 0.997, 0.187, and 0.207, respectively). The effect directions of the above-mentioned SNPs, with the exception of rs1111875 in *HHEX*, were consistent with those in previous reports (OR >1, $P = 9.2 \times 10^{-4}$, binomial test) (16,28).

We next studied the associations of the T-GRS (equivalent to the sum of the risk alleles of the 14 SNPs studied here), the β-GRS (equivalent to the sum of the 11 β-cell function–related genes), and the R-GRS (equivalent to the sum of the three obesity and insulin resistance-related genes) with the development of type 2

diabetes. The T-GRS and β-GRS were significantly associated with the development of type 2 diabetes (T-GRS OR 1.26 [95% CI 1.20–1.33], $P = 5.9 \times 10^{-21}$ [Supplementary Fig. 1]; β-GRS 1.26 [1.20–1.33], $P = 1.1 \times 10^{-19}$ [Supplementary Table 3]; and R-GRS, nominally associated with the development of type 2 diabetes, 1.18 [1.02–1.37], $P = 0.024$ [Supplementary Table 3]). We further determined that when all of the participants were stratified according to the β-GRS (high-risk genetic group [H]-β-GRS ≥12; intermediate risk [I], 12 > β-GRS ≥ 10; and low risk [L]-β-GRS <10) or the R-GRS (H-R-GRS ≥5; I, 5 > R-GRS ≥ 4; and L-R-GRS <4 (Supplementary Table 4), the risk of developing diabetes in the H-β-GRS and the H-R-GRS groups ($n = 108$) was 6.2-fold higher than in the L-β-GRS and the L-R-GRS groups ($n = 78$) (Supplementary Fig. 2). Interestingly, an effect of the R-GRS was only seen in the L-β-GRS group (OR 1.43 [95% CI 1.06–1.95], $P = 0.02$) and not in the H-β-GRS groups (1.17 [0.85–1.61], $P = 0.34$) (Supplementary Fig. 2), suggesting that the β-GRS has a predominant effect on conferring susceptibility to type 2 diabetes over the R-GRS. To statistically evaluate the interaction between β-GRS and R-GRS, we performed a stepwise logistic regression analysis using strategies of both forward selection (addition of each parameter) and backward selection (starting from all parameters). The results indicated that significant interaction was observed when we added β-GRS to R-GRS ($P < 0.001$), whereas the effect of addition of R-GRS to β-GRS was modest ($P = 0.03$).

We next examined the associations of each genetic variant with quantitative metabolic traits related to type 2 diabetes. None of the SNPs had a significant effect on the HOMA-β or HOMA-IR by themselves, but the β-GRS and R-GRS showed stronger association with the HOMA-β ($P = 0.025$) and HOMA-IR ($P = 0.0004$), respectively, than single SNP alone, in control individuals and patients with type 2 diabetes who were not treated with medications (Table 2). We further examined the association of the three types of GRS with type 2 diabetes and quantitative traits in a previously published independent cohort, which was conducted in Tokyo University. In this cohort, the association between T-GRS and type 2 diabetes (OR 1.18 [95% CI 1.13–1.24], $P = 2.08 \times 10^{-12}$) and β-GRS and HOMA-β (β of ln-HOMA-β = −0.0377, SE = 0.0103, $P = 0.0003$) was statistically significant, whereas the association of the R-GRS with HOMA-IR did not reach a statistically significant level (β of ln-HOMA-IR = 0.0294, SE = 0.0290, $P = 0.3120$) (Supplementary Table 5).

Investigation of combined effects of GRS on the clinical features of type 2 diabetes

We next examined the association of the T-GRS with clinical features, such as the maximum BMI, the age at the time of diagnosis, and the individuals presently receiving insulin therapy (Supplementary Table 6). Significant inverse correlations were observed between the T-GRS and the maximum BMI (β of maximum BMI −0.225 [95% CI −0.367 to −0.083], $P = 0.002$) and the age at diagnosis (β of age at diagnosis −0.663 [−1.048 to −0.278], $P = 0.0008$). We also found that the individuals receiving insulin therapy were positively associated with the T-GRS (β of insulin therapy 0.249 [0.025–0.473], $P = 0.029$).

We then divided all the participants into three approximately equally sized strata according to the T-GRS: L-T-GRS, I-T-GRS, and H-T-GRS genetic groups, as described in Supplementary Table 4. The characteristics of the three groups are shown in Table 3. In the H-T-GRS group, the duration of diabetes was significantly longer ($P < 0.01$) and the current BMI was lower ($P < 0.05$) than those in the L-T-GRS group (Table 3). We next studied the association of the T-GRS with clinical features such as the maximum BMI, the age at the time of diagnosis, and the percentage of individuals receiving insulin therapy (Table 3). We found that the maximum BMI in the H-T-GRS group (27.1 ± 4.2) was significantly lower than that in the L-T-GRS group (28.5 ± 4.6) ($P < 0.01$). In addition, the age at the time of the diagnosis of diabetes in the H-T-GRS group (49.8 ± 12.4 years) was significantly younger than that in the L-T-GRS group (52.5 ± 11.4 years) ($P < 0.001$) after adjustments for sex and the maximum BMI (Table 3). The percentage of individuals receiving insulin therapy in the H-T-GRS group (34.9%) was greater than that in the L-T-GRS group (22.7%) ($P < 0.05$) after adjustments for age, sex, current BMI, duration of diabetes, class of antihyperglycemic drugs, and present HbA_{1c} level.

We next examined the associations of the genetic risk score of β-cell function–related SNPs (β-GRS) with the clinical features (Table 4). We found that the β-GRS was associated with individuals receiving insulin therapy (β of insulin therapy 0.0131 [95% CI 0.0004–0.0259], $P = 0.0431$) and a younger age at diagnosis (β of age at diagnosis −0.608 [−1.008 to −0.208], $P = 0.0029$). Furthermore, we found a significant inverse correlation between the β-GRS and β-cell function–related parameters including the fasting serum C-peptide (F-CPR) (β of serum C-peptide −0.036 [−0.065 to −0.007], $P = 0.0140$) and the C-peptide index (CPI) (β of CPI −0.031 [−0.056 to −0.005], $P = 0.0179$) after adjustments for age, sex, BMI, duration of diabetes, class of antihyperglycemic drugs, fasting plasma glucose, the presence of diabetic nephropathy, and the presence of diabetic retinopathy. We also examined the association of T-GRS with these parameters, but as expected the β-GRS had stronger effects on basal insulin secretion than the T-GRS (Supplementary Table 6). The R-GRS was not associated with any parameters (Table 4).

We further tried, as much as possible, to include all information of European study–derived type 2 diabetes variants in the GRS. Overall, the 36 SNP GRS constructed with the 14 SNPs and additional 22 SNPs, however, did not show stronger association with each metabolic trait than the original T-, β-, and R-GRS in this study (Supplementary Tables 7 and 8).

CONCLUSIONS—In the current study, we examined 14 SNP loci, which were robustly shown to be susceptibility loci for type 2 diabetes in the Japanese population, and constructed a GRS to evaluate the usefulness of this genetic information in clinical practice. We found that most SNPs (13 of 14) showed a directionally consistent association with the results of previous reports (6–16), and constructed GRS (T-GRS) showed a much stronger association with type 2 diabetes than any of the single SNPs alone. The T-GRS was also associated with age at the time of the diagnosis of diabetes. Additionally, we found that a β-GRS, consisting of eleven β-cell function–related SNPs, was associated with requirement of insulin therapy and a reduced basal insulin secretion level in Japanese patients with type 2 diabetes.

Currently, >40 loci have been confirmed as susceptibility loci for type 2 diabetes in populations of European origin (5), but the integration of this information can only explain ~10% of type 2 diabetes

Table 2—*Association of the 14 SNPs with quantitative traits related to glucose metabolism in control subjects and diabetic subjects*

	Control[a]		Control and type 2 diabetic subjects without medication[b]	
	HOMA-IR	HOMA-β	HOMA-IR	HOMA-β
rs5219, *KCNJ11*				
Effect (SE)	0.039 (0.037)	−1.800 (2.039)	0.042 (0.039)	−2.531 (1.868)
P	0.297	0.378	0.281	0.178
rs7903146, *TCF7L2*				
Effect (SE)	−0.050 (0.095)	6.088 (5.229)	0.068 (0.100)	9.183 (4.794)
P	0.599	0.245	0.501	0.056
rs1111875, *HHEX*				
Effect (SE)	−0.037 (0.040)	−1.915 (2.192)	−0.022 (0.042)	−1.743 (2.010)
P	0.354	0.383	0.597	0.386
rs13266634, *SLC30A8*				
Effect (SE)	0.015 (0.035)	−1.283 (1.928)	−0.003 (0.038)	−1.464 (1.789)
P	0.673	0.506	0.937	0.413
rs7756992, *CDKAL1*				
Effect (SE)	−0.053 (0.036)	−1.629 (1.963)	−0.046 (0.038)	−0.937 (1.804)
P	0.137	0.407	0.219	0.604
rs10811661, *CDKN2B*				
Effect (SE)	0.036 (0.036)	−0.878 (2.001)	0.046 (0.039)	−1.443 (1.863)
P	0.328	0.661	0.24	0.439
rs4402960, *IGF2BP2*				
Effect (SE)	0.009 (0.040)	0.308 (2.170)	−0.013 (0.041)	−0.230 (1.991)
P	0.813	0.887	0.75	0.908
rs2237892, *KCNQ1*				
Effect (SE)	0.027 (0.035)	0.552 (1.947)	−0.008 (0.038)	−0.447 (1.833)
P	0.441	0.777	0.833	0.807
rs780094, *GCKR*				
Effect (SE)	0.030 (0.038)	−0.025 (2.091)	0.022 (0.040)	−0.275 (1.927)
P	0.43	0.99	0.585	0.887
rs7612463, *UBE2E2*				
Effect (SE)	−0.057 (0.052)	−4.372 (2.831)	−0.032 (0.055)	−4.768 (2.648)
P	0.271	0.123	0.568	0.072
rs7172432, *C2CD4A/B*				
Effect (SE)	−0.007 (0.037)	−3.208 (2.023)	0.003 (0.039)	−2.740 (1.880)
P	0.851	0.113	0.947	0.145
rs2943641, *IRS-1*				
Effect (SE)	0.153 (0.061)	3.989 (3.340)	0.128 (0.065)	2.993 (3.131)
P	0.012	0.233	0.051	0.339
rs1801282, *PPARG*				
Effect (SE)	0.097 (0.114)	5.243 (6.262)	0.110 (0.123)	5.903 (5.903)
P	0.394	0.403	0.372	0.318
rs8050136, *FTO*				
Effect (SE)	0.117 (0.048)	3.885 (2.632)	0.138 (0.050)	3.500 (2.402)
P	0.015	0.14	0.006	0.146
T-GRS				
Effect (SE)	0.017 (0.011)	−0.656 (0.630)	0.016 (0.012)	−0.868 (0.584)
P	0.15	0.298	0.179	0.138
β-GRS				
Effect (SE)	0.003 (0.012)	−1.213 (0.669)	0.002 (0.013)	−1.388 (0.618)
P	0.775	0.07	0.856	0.025[c]
R-GRS				
Effect (SE)	0.125 (0.035)	3.873 (1.917)	0.131 (0.037)	3.368 (1.763)
P	3.0×10^{-4}*	0.044	4.0×10^{-4}†	0.056

Results of linear regression analyses. The effect size corresponds to the β-coefficient (SE) per copy of the type 2 diabetes risk allele and was calculated using a linear regression analysis. [a]n = 763 (adjusted for sex, age, and BMI). [b]n = 860 (adjusted for age, sex, BMI, and disease status). [c]P = 0.03687 after 100,000 permutations, P = 0.03108 after Bonferroni correction. *P = 0.00633 after 100,000 permutations, P = 0.00469 after Bonferroni correction. †P = 0.00568 after 100,000 permutations, P = 0.00425 after Bonferroni correction.

Table 3—*Clinical characteristics of the three groups according to the T-GRS of 14 SNPs in patients with type 2 diabetes*

GRS	Low	Intermediate	High	P (ANOVA)	P (multivariate)*
No. of risk alleles	≤13	14–15	≥16		
n	176	244	304		
Sex (male/female)	113/63	155/89	183/121	0.607	
Age (years)	64.95 ± 11.53	65.29 ± 10.23	64.46 ± 11.52	0.684	
BMI (kg/m^2)	25.25 ± 4.27	24.43 ± 3.34	24.21 ± 4.02	0.016	
Self-reported family history of diabetes (%)	51.46	60.67	54.03	0.137	
Duration of diabetes (years)	12.49 ± 8.65	12.72 ± 8.52	14.82 ± 14.82	0.006	
FPG (mmol/L)	7.48 ± 1.89	7.62 ± 1.73	7.66 ± 1.99	0.628	
A1C (%)	7.53 ± 1.26	7.48 ± 1.22	7.56 ± 1.26	0.783	
Diabetic nephropathy (%)	42.44	38.66	37.59	0.650	
Diabetic retinopathy (%)	43.14	40.61	43.11	0.824	
Insulin secretagogue (%)	50.57	50.00	52.96	0.764	
Age at diagnosis (years)	52.49 ± 11.42	52.65 ± 10.49	49.81 ± 12.43	0.008	<0.001
Maximum BMI (kg/m^2)	28.52 ± 4.63	27.05 ± 3.93	27.13 ± 4.23	0.001	0.002
Insulin requirement (%)	22.73	32.79	34.87	0.018	0.044

Data are means ± SD or n/n unless otherwise indicated. The value for A1C (%) was estimated as a National Glycohemoglobin Standardization Program (NGSP) equivalent value (%) calculated using the following formula: A1C (%) = A1C (Japan Diabetes Society [JDS]) (%) + 0.4%. *P value for comparison of adjusted data. Age at diagnosis was adjusted for sex and maximum BMI. Maximum BMI was adjusted for sex. The ratio of insulin therapy was adjusted for age, sex, BMI, duration of diabetes, class of antihyperglycemic drug, and A1C level.

heritability; therefore, currently available genetic information for type 2 diabetes is likely insufficient for predicting disease progression and has failed to provide a significant impact on human health care in the general population to date.

Previously, several groups investigated the impact of a GRS for loci related to β-cell function and reported that the GRS was significantly associated with glucose-stimulated insulin secretion (GSIS) in nondiabetic subjects or the subjects with impaired glucose tolerance (4,26,37,38). However, the effects of the GRS on clinical features have not been evaluated in patients with type 2 diabetes. In the current study, we demonstrated for the first time that a β-GRS was associated with individuals receiving insulin therapy and possessing reduced basal insulin secretion, as evaluated using the F-CPR or CPI, among type 2 diabetic subjects. Thus, the β-GRS may be useful for predicting future reductions in basal insulin secretion, resulting in the need for insulin injections to control plasma glucose levels. Interestingly, the association of β-GRS with the reduction in basal insulin secretion long after the onset (e.g., >10 years) was independent of confounding factors, such as age, sex, BMI, duration of diabetes, presence of microvascular complications, and the use of insulin secretagogues. Of note, since the presence of microvascular complications reflects chronic hyperglycemia, the declining β-cell function in individuals with higher β-GRS may not be a consequence of the relatively longer terms of hyperglycemia. Thus, we think that the evaluation of β-GRS at an earlier

Table 4—*Association of β-GRS and R-GRS with quantitative metabolic traits and clinical information in diabetic subjects*

GRS and index	β†	SE	P	Covariables
β-GRS				
F-CPR	−0.036	0.014	0.0140	Age, sex, BMI, duration of diabetes, class of antihyperglycemic drug, FPG, the presence of diabetic nephropathy, and the presence of diabetic retinopathy
CPI	−0.031	0.0123	0.0179	Age, sex, BMI, duration of diabetes, class of antihyperglycemic drug, the presence of diabetic nephropathy, and the presence of diabetic retinopathy
Age at diagnosis	−0.608	0.204	0.0029	Sex and maximum BMI
Maximum BMI	−0.263	0.075	0.0004	Sex
Insulin requirement	0.013	0.006	0.0431	Age, sex, BMI, duration of diabetes, class of antihyperglycemic drug, and A1C
R-GRS				
F-CPR	0.018	0.042	0.676	Age, sex, BMI, duration of diabetes, class of antihyperglycemic drug, FPG, the presence of diabetic nephropathy, and the presence of diabetic retinopathy
CPI	0.015	0.037	0.696	Age, sex, BMI, duration of diabetes, class of antihyperglycemic drug and the presence of diabetic nephropathy, and the presence of diabetic retinopathy
Age at diagnosis	−0.891	0.599	0.137	Sex and maximum BMI
Maximum BMI	0.196	0.222	0.380	Sex
Insulin requirement	0.013	0.019	0.507	Age, sex, BMI, duration of diabetes, class of antihyperglycemic drug, and A1C

CPI was calculated using the following equation: CPI = [F-CPR/FPG] × 100. F-CPR and CPI were calculated in all diabetic subjects except for those with serum creatinine level >1.5 mg/dL. †Regression coefficient adjusted for covariables.

stage of the disease may be useful, and patients with a higher β-GRS should be strongly encouraged to receive specialized therapies, such as intensive lifestyle modifications and/or the earlier introduction of β-cell–preserving therapy, such as the use of glucagon-like peptide 1 receptor agonists or medications that ameliorate insulin resistance.

In the current study, we were able to replicate the previously reported associations of 8 of the 14 loci in a Japanese population (4 significantly [$P < 0.0036$] and 4 modestly [$P < 0.05$]) (25,27–32). As for rs7903146 in *TCF7L2*, rs1111875 in *HHEX*, rs1801282 in *PPARG*, rs8050136 in *FTO*, and rs7612463 in *UBE2E2*, which were reported to be associated with type 2 diabetes in previous Japanese reports (25,28,30), we were unable to detect any SNPs that were significantly associated with type 2 diabetes in the present Japanese sample (P = 0.659, 0.773, 0.997, 0.187, and 0.207, respectively). However, since the effect directions of most of the SNP loci (13 of 14) were consistent with the results of previous reports and the estimated study power was 15−81% for the 6 unreplicated SNPs (Supplementary Table 2), the lack of replication might be explained by the insufficient power of the current study. In the quantitative trait analyses using control individuals and type 2 diabetic patients with no medications, we did not observe any significant association between each of the single SNPs and glycemic traits, but the β-GRS and R-GRS showed stronger association with the HOMA-β (P = 0.025) and HOMA-IR (P = 0.0004), respectively, indicating that the constructed β-GRS and R-GRS in the current study were appropriate and useful for evaluating the genetic effects on susceptibility to the disease or on related quantitative traits, even among a relatively small study population. The association of the three types of GRS with the quantitative traits could also be consistently observed in an independent cohort, which was conducted in Tokyo University (28,30), further validating the usefulness of the GRS.

Since HOMA indices have some limitations as indicators of β-cell functions or peripheral insulin sensitivity, evaluation of other independent measures of insulin secretion or resistance, such as 2-h glucose and insulin measurements, is required to confirm our findings. We demonstrated that the β-GRS was associated with a reduced basal insulin secretion in diabetic subjects with an average disease duration of 13.6 years. We also observed that the β-GRS was inversely associated with the GSIS determined by disposition index (β of ln–disposition index −0.102 [95% CI −0.006 to −0.194], P = 0.038, after adjustments for age, sex, FPG, and BMI) at the onset of diabetes (n = 134 [unpublished results]); therefore, the β-GRS may be involved in the GSIS at the onset of diabetes, and a further reduction in basal insulin secretion in patients with a higher β-GRS long after the onset of diabetes (>10 years) may contribute to the need for insulin injections. A cohort study involving a larger number of subjects is needed to clarify this point.

In conclusion, we have shown that the β-GRS, as determined using eleven β-cell function–related loci, is associated with a lower basal insulin secretion and the percentage of individuals requiring insulin therapy among Japanese subjects with type 2 diabetes. These results suggest that the evaluation of β-GRS at an earlier stage of the disease may be useful, and patients with a higher β-GRS should receive specialized therapy, including guidance regarding intensive lifestyle modifications and β-cell–preserving therapy.

Acknowledgments—This work was partly supported by a grant from the Ministry of Education, Culture, Sports, Science and Technology, Japan, to S.M.

No potential conflicts of interest relevant to this article were reported.

M.I. wrote the manuscript and researched data. S.M. researched data, wrote the manuscript, and edited the manuscript. Y.K. researched data. A.Takan., H.K., S.M., and K.H. contributed to discussion. A.Takah. researched data. H.F. researched data. K.H. researched data and reviewed the manuscript. T.K. reviewed the manuscript. K.T. wrote the manuscript and reviewed and edited the manuscript. K.T. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Parts of this study were presented in abstract form at the 70th Scientific Sessions of the American Diabetes Association, Orlando, Florida, 25–29 June 2010.

The authors are grateful to all the subjects who took part in this study. The authors thank Dr. Sachie Asamizu of Shakaihoken Takaoka Hospital; Dr. Rie Oka and Dr. Susumu Miyamoto of Hokuriku Cho Hospital; Dr. Kunimasa Yagi of the University of Kanazawa; Dr. Isao Usui, Dr. Manabu Ishiki, Dr. Toshiyasu Sasaoka, Dr. Chikaaki Kobashi, Dr. Katsuya Yamazaki, Dr. Masaharu Urakaze, Dr. Shiho Fujisaka, Dr. Satoko Senda, Dr. Hikari Suzuki, and Dr. Yu Yamazaki of Toyama University; Dr. Rie Temaru and Dr. Mariko Ikubo of Nanto City Hospital; Dr. Naoji Akagawa and Dr. Yasuo Fukushima of Asahi General Hospital; Dr. Kazuko Taki of Amenithy Tsukioka Hospital; Dr. Yasuhumi Igarashi of Aoi Hospital; Dr. Shigeki Sawasaki of Hida City Hospital; Dr. Hisae Honoki of Saiseikai Takaoka Hospital; and Dr. Hirohumi Oda of Sakurai Hospital for recruiting the study subjects. The authors thank Dr. Momoko Horikoshi of the University of Tokyo for researching data. The authors also thank the technical staff at the Laboratory for Endocrinology and Metabolism, RIKEN Center for Genomic Medicine, for technical assistance.

References

1. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 2004;27:1047–1053
2. O'Rahilly S, Barroso I, Wareham NJ. Genetic factors in type 2 diabetes: the end of the beginning? Science 2005;307:370–373
3. Chauhan G, Spurgeon CJ, Tabassum R, et al. Impact of common variants of PPARG, KCNJ11, TCF7L2, SLC30A8, HHEX, CDKN2A, IGF2BP2, and CDKAL1 on the risk of type 2 diabetes in 5,164 Indians. Diabetes 2010;59:2068–2074
4. Stancáková A, Kuulasmaa T, Paananen J, et al. Association of 18 confirmed susceptibility loci for type 2 diabetes with indices of insulin release, proinsulin conversion, and insulin sensitivity in 5,327 nondiabetic Finnish men. Diabetes 2009; 58:2129–2136
5. McCarthy MI. Genomics, type 2 diabetes, and obesity. N Engl J Med 2010;363: 2339–2350
6. Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445:881–885
7. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat Genet 2007;39:770–775
8. Saxena R, Voight BF, Lyssenko V, et al.; Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316:1331–1336
9. Zeggini E, Weedon MN, Lindgren CM, et al.; Wellcome Trust Case Control Consortium (WTCCC). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007;316:1336–1341
10. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2

diabetes in Finns detects multiple susceptibility variants. Science 2007;316:1341–1345
11. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 2007;316:889–894
12. Zeggini E, Scott LJ, Saxena R, et al.; Wellcome Trust Case Control Consortium. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638–645
13. Voight BF, Scott LJ, Steinthorsdottir V, et al.; MAGIC investigators; GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579–589
14. Dupuis J, Langenberg C, Prokopenko I, et al.; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; Anders Hamsten on behalf of Procardis Consortium; MAGIC investigators. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 2010;42:105–116
15. Unoki H, Takahashi A, Kawaguchi 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
16. Yasuda K, Miyake K, Horikawa Y, et al. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. Nat Genet 2008;40:1092–1097
17. Altshuler D, Hirschhorn JN, Klannemark M, et al. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 2000;26:76–80
18. Gloyn AL, Weedon MN, Owen KR, 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 2003;52:568–572
19. Sandhu MS, Weedon MN, Fawcett KA, et al. Common variants in WFS1 confer risk of type 2 diabetes. Nat Genet 2007;39:951–953
20. Grant SF, Thorleifsson G, Reynisdottir I, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 2006;38:320–323
21. Weedon MN, McCarthy MI, Hitman G, et al. Combining information from common type 2 diabetes risk polymorphisms improves disease prediction. PLoS Med 2006;3:e374
22. Cauchi S, Meyre D, Durand E, et al. Post genome-wide association studies of novel genes associated with type 2 diabetes show gene-gene interaction and high predictive value. PLoS ONE 2008;3:e2031
23. Lango H, Palmer CN, Morris AD, et al.; 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 2008;57:3129–3135
24. Lyssenko V, Jonsson A, Almgren P, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. N Engl J Med 2008;359:2220–2232
25. Miyake K, Yang W, Hara K, et al. Construction of a prediction model for type 2 diabetes mellitus in the Japanese population based on 11 genes with strong evidence of the association. J Hum Genet 2009;54:236–241
26. 't Hart LM, Simonis-Bik AM, Nijpels G, et al. Combined risk allele score of eight type 2 diabetes genes is associated with reduced first-phase glucose-stimulated insulin secretion during hyperglycemic clamps. Diabetes 2010;59:287–292
27. Takeuchi F, Serizawa M, Yamamoto K, et al. Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. Diabetes 2009;58:1690–1699
28. Yamauchi T, Hara K, Maeda S, 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
29. Horikawa Y, Miyake K, Yasuda K, et al. Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. J Clin Endocrinol Metab 2008;93:3136–3141
30. Horikoshi M, Hara K, Ito C, et al. Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. Diabetologia 2007;50:2461–2466
31. Omori S, Tanaka Y, Takahashi A, et al. Association of CDKAL1, IGF2BP2, CDKN2A/B, HHEX, SLC30A8, and KCNJ11 with susceptibility to type 2 diabetes in a Japanese population. Diabetes 2008;57:791–795
32. Tabara Y, Osawa H, Kawamoto R, et al. Replication study of candidate genes associated with type 2 diabetes based on genome-wide screening. Diabetes 2009;58:493–498
33. 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(Suppl. 1):S5–S20
34. Maeda S, Tsukada S, Kanazawa A, et al. Genetic variations in the gene encoding TFAP2B are associated with type 2 diabetes mellitus. J Hum Genet 2005;50:283–292
35. Báez S, Tsuchiya Y, Calvo A, et al. Genetic variants involved in gallstone formation and capsaicin metabolism, and the risk of gallbladder cancer in Chilean women. World J Gastroenterol 2010;16:372–378
36. Nielsen DM, Ehm MG, Weir BS. Detecting marker-disease association by testing for Hardy-Weinberg disequilibrium at a marker locus. Am J Hum Genet 1998;63:1531–1540
37. Pascoe L, Frayling TM, Weedon MN, et al.; RISC Consortium. Beta cell glucose sensitivity is decreased by 39% in nondiabetic individuals carrying multiple diabetes-risk alleles compared with those with no risk alleles. Diabetologia 2008;51:1989–1992
38. Haupt A, Staiger H, Schäfer SA, et al. The risk allele load accelerates the age-dependent decline in beta cell function. Diabetologia 2009;52:457–462

Expression-based genome-wide association study links the receptor *CD44* in adipose tissue with type 2 diabetes

Keiichi Kodama[a,b], Momoko Horikoshi[c], Kyoko Toda[d,1], Satoru Yamada[e,1], Kazuo Hara[c,1], Junichiro Irie[e,f,1], Marina Sirota[a,b], Alexander A. Morgan[a,b], Rong Chen[a,b], Hiroshi Ohtsu[g], Shiro Maeda[h], Takashi Kadowaki[c], and Atul J. Butte[a,b,2]

[a]Division of Systems Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA 94305; [b]Lucile Packard Children's Hospital, Palo Alto, CA 94304; Departments of [c]Metabolic Diseases and [g]Clinical Trial Data Management, Graduate School of Medicine, University of Tokyo, Tokyo 113-8655, Japan; [d]Division of Basic Research, Biomedical Laboratory, Kitasato Institute Hospital, Kitasato University, Tokyo 108-8642, Japan; [e]Diabetes Center, Kitasato Institute Hospital, Tokyo 108-8642, Japan; [f]Department of Internal Medicine, Keio University School of Medicine, Tokyo 160-8582, Japan; and [h]Laboratory for Endocrinology and Metabolism, Center for Genomic Medicine, RIKEN, Yokohama City, Kanagawa 230-0045, Japan

Edited by Alan D. Attie, University of Wisconsin, Madison, WI, and accepted by the Editorial Board March 19, 2012 (received for review September 7, 2011)

Type 2 diabetes (T2D) is a complex, polygenic disease affecting nearly 300 million people worldwide. T2D is primarily characterized by insulin resistance, and growing evidence has indicated the causative link between adipose tissue inflammation and the development of insulin resistance. Genetic association studies have successfully revealed a number of important genes consistently associated with T2D to date. However, these robust T2D-associated genes do not fully elucidate the mechanisms underlying the development and progression of the disease. Here, we report an alternative approach, gene expression-based genome-wide association study (eGWAS): searching for genes repeatedly implicated in functional microarray experiments (often publicly available). We performed an eGWAS across 130 independent experiments (totally 1,175 T2D case-control microarrays) to find additional genes implicated in the molecular pathogenesis of T2D and identified the immune-cell receptor *CD44* as our top candidate ($P = 8.5 \times 10^{-20}$). We found *CD44* deficiency in a diabetic mouse model ameliorates insulin resistance and adipose tissue inflammation and also found that anti-CD44 antibody treatment decreases blood glucose levels and adipose tissue macrophage accumulation in a high-fat, diet-fed mouse model. Further, in humans, we observed CD44 is expressed in inflammatory cells in obese adipose tissue and discovered serum CD44 levels were positively correlated with insulin resistance and glycemic control. *CD44* likely plays a causative role in the development of adipose tissue inflammation and insulin resistance in rodents and humans. Genes repeatedly implicated in publicly available experimental data may have unique functionally important roles in T2D and other complex diseases.

bioinformatics | meta-analysis | integration | obesity | hyperglycemia

Type 2 diabetes (T2D) is a common multifactorial disease characterized by hyperglycemia primarily resulting from peripheral insulin resistance, and growing functional evidence has indicated the causative link between adipose tissue inflammation and the development of insulin resistance (1, 2). In the past decade, a number of genetic genome-wide association studies (GWAS) have revealed 40 loci consistently associated with susceptibility to T2D and have rapidly expanded our knowledge of the genetic architecture of this disease (3–13). However, the genes located in or near these loci do not fully elucidate the tissue-specific molecular mechanisms underlying the development of T2D.

A large number of experiments using genome-wide gene-expression microarray measurements have been also performed over the past decade; however, there has been little success in fully identifying functionally important genes in the pathogenesis of T2D. Because a large number of genes are often detected as significant in each microarray experiment, it may be hard to subselect optimal candidates from individual studies for further verification. A combination of genome-wide data from two or more experiments has been performed for obesity (14–16), T2D (14), and other multifactorial disorders (17–20). Several of these methods have used microarray technology to focus on candidate genes already implicated in a region of a congenic or model animal (accelerated positional candidate identification). More recently, investigators have applied microarrays to genetics by considering gene expression levels as quantitative traits (expression quantitiative trait loci, eQTLs) and finding relations between gene variants and transcripts, with successful application to the identification of genes and targets for T2D (21, 22). The need for large numbers of simultaneously acquired genetic and gene-expression measurements within a single study makes this approach less scalable.

We suspected that the large number of molecular measurements from experimental results that are now publicly available, because of requirements from journals and funding agencies (23), could be used as an alternative scalable source of data. The strategy of finding commonly implicated genes across related—but deliberately varied—experimental conditions has been theorized to yield less overfit, potentially more generalizable causal factors (24), and publicly available data could be used as a source of these varied experimental vantage points for a condition. In this report, we propose the application of a gene expression-based genome-wide association study (eGWAS), a meta-analysis method for computing the likelihood of finding repeated differential expression for every gene across a large number of case and control microarray experiments, compared with expected. Our hypothesis is that those genes most repeatedly implicated across a large set of experimental representations of T2D can serve as data-driven causal T2D genes and candidates for validation. This approach is only feasible because many of these source raw experimental results are publicly available; here, we integrated 130 independent microarray experiments for T2D. In this case, our T2D candidates were found independent of any knowledge about insulin signaling, glucose, or lipid metabolism. For our top candidate gene, identification was followed by confirmatory functional studies using mouse models and samples from human subjects. (Fig. S1).

Author contributions: K.K., M.H., K.H., T.K., and A.J.B. designed research; K.K., M.H., K.T., S.Y., K.H., J.I., S.M., T.K., and A.J.B. performed research; M.S., A.A.M., and R.C. contributed new analytic tools; K.K., M.H., K.H., H.O., S.M., T.K., and A.J.B. analyzed data; and K.K., M. H., K.H., T.K., and A.J.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. A.D.A. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

[1]K.T., S.Y., K.H., and J.I. contribute equally to this work.

[2]To whom correspondence should be addressed. E-mail: abutte@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1114513109/-/DCSupplemental.

MEDICAL SCIENCES

Results

eGWAS Identifies *CD44* as a Functional Candidate Gene for T2D. We carried out an eGWAS for T2D by using 130 independent microarray experiments, totaling 1,175 samples collected from public repositories (Fig. 1, Tables S1 and S2, and *SI Materials and Methods*). We ranked all 24,898 genes by the likelihood that repeated differential expression for that gene was due to chance, then controlled for multiple-hypothesis testing. To overview which molecular functions are most shared in the highest-ranked genes in our T2D eGWAS, we took the top 127 genes (Table S3; Bonferroni threshold, $P < 2.0 \times 10^{-6}$) from our eGWAS and then estimated the enrichment of Gene Ontology (GO) terms. Interestingly, "receptor activity" and "receptor binding" functions were the most implicated of the top-ranked genes (receptor activity, 38%; receptor binding, 19%) (Fig. S2). These activities suggested that a number of top-ranked genes on our list are involved in intra- and intertissue signaling cascades in the development of T2D (1).

Our top-most T2D candidate gene was *CD44* (Fig. 1: χ^2 analysis, $P = 8.5 \times 10^{-20}$; Fig. S3: Fisher's exact test, $P = 6.1 \times 10^{-17}$; Fig. S4: weighted Z-method), markedly differentially expressed in experiments studying diabetes in adipose tissue compared with other tissues (Fig. S5). *CD44* is located on chromosome 11p13 and codes for a cell-surface glycoprotein, an immunological cell (macrophage/T-cell) receptor, involved in inflammatory cell migration and activation. Interestingly, one of the known ligands for *CD44*, secreted phosphoprotein 1 [SPP1; also known as osteopontin (OPN)], a Th1 cytokine secreted by immunological cells (macrophages), was also included in the top-ranked genes (Fig. 1: χ^2 analysis, $P = 1.3 \times 10^{-11}$; Fig. S3: Fisher's exact test, $P = 3.8 \times 10^{-10}$). Recent studies have indicated that obese adipose tissue is hallmarked with chronic, low-grade inflammation, and that inflammation plays a central role in the development of insulin resistance (1, 2). Although the contributions of the *CD44* encoded protein to the molecular pathogenesis of T2D have not yet been reported, SPP1 was reported as a link between adipose tissue inflammation (stromal infiltration by inflammatory cells) and the development of insulin resistance in a murine model of diet-induced obesity (25). Furthermore, the expression profile of *CD44* and *SPP1* are coordinately dysregulated, especially in adipose tissue (Fig. S6; coordinate dysregulation rate = 0.90). These findings suggest that CD44 might have a key role in mediating obesity-induced adipose tissue inflammation and insulin resistance.

***CD44* Expression Increases in Obese Adipose Tissue.** High-fat feeding in C57BL/6J mice leads to the development of obesity, adipose inflammation, and insulin resistance (25, 26). To examine whether the *CD44* mRNA transcript is expressed in adipose tissue and modulated by obesity, C57BL/6J mice were maintained either on a normal-fat diet (NFD; 12% of total calories from fat) or high-fat diet (HFD; 60% of total calories from fat) for 16 wk ($n = 8$ per group). Compared with the NFD group, mice fed a HFD gained 37% more weight after a 16-wk feeding period (29.9 ± 0.5 g versus 40.9 ± 1.6 g; $P = 5.8 \times 10^{-8}$). Epididymal white adipose tissue (EWAT) was removed from these mice to analyze *CD44* mRNA expression levels. Feeding a HFD resulted in a significant 11.3-fold increase of *CD44* mRNA levels in adipose tissue compared with NFD (Fig. 2*A*). To establish the presence of CD44 protein in adipose tissue, an immunohistochemical localization of CD44 was performed on EWAT isolated from HFD mice. We found that CD44 was abundantly expressed in inflammatory cells within adipose tissue (Fig. 2*B*). These results clearly indicate that $CD44^+$ cells accumulated in EWAT of diet-induced obese mice. In addition, we confirmed that the expression levels of *SPP1* mRNA in adipose tissue in the HFD group was significantly higher than that in the NFD mice (0.005 ± 0.003 versus 0.06 ± 0.02; $P < 0.05$), similar to previous reports (25). Interestingly, we also found that *CD44* mRNA expression level was positively correlated with the *SPP1* mRNA expression level in the HFD mice (Fig. 2*C*; $r = 0.78$, $P = 0.02$), suggesting that CD44 and SPP1 may be closely related in obese adipose tissue.

***CD44* Deficiency Ameliorates Adipose Tissue Inflammation and Insulin Resistance.** To next determine the contribution of CD44 to the development of adipose tissue inflammation and insulin resistance, we fed male $CD44^{-/-}$ and diabetes-prone C57BL/6J ($CD44^{+/+}$) mice with either a HFD ($n = 16$ per group) or a NFD ($n = 10$ per group) for 12 wk and performed immunohistochemical analysis and metabolic measurements on these mice. There were no significant differences in body weights between $CD44^{-/-}$ and $CD44^{+/+}$ mice after feeding a NFD or a HFD (NFD: $CD44^{-/-}$ 28.5 ± 0.5 g versus $CD44^{+/+}$ 29.5 ± 0.5 g; HFD: 38.1 ± 1.5 g versus 39.5 ± 1.2 g). In $CD44^{+/+}$ mice that were fed a HFD, we frequently observed the accumulation of inflammatory cells (macrophages) forming crown-like structures (CLSs) surrounding adipocytes in obese visceral adipose tissue. However, $CD44^{-/-}$ mice fed a HFD exhibited strikingly less macrophage infiltration into the stroma of adipose tissue compared with $CD44^{+/+}$ mice fed a HFD (Fig. 3*A*). Fasting blood glucose levels were significantly lower in $CD44^{-/-}$ mice fed a HFD compared with the diabetes-prone $CD44^{+/+}$ mice fed a HFD (Fig. 3*B*). Glucose tolerance tests also indicated that $CD44^{-/-}$ mice fed a HFD were significantly more efficient in their ability to clear intraperitoneally injected glucose than $CD44^{+/+}$ mice fed a HFD (Fig. 3*C*, solid lines) despite similar insulin secretory responses after the injection of glucose. Furthermore, insulin sensitivity, as measured by insulin tolerance test, showed that $CD44^{-/-}$ mice fed a HFD were significantly more efficient at insulin-mediated suppression of blood glucose than $CD44^{+/+}$ mice fed a HFD (Fig. 3*D*, solid lines).

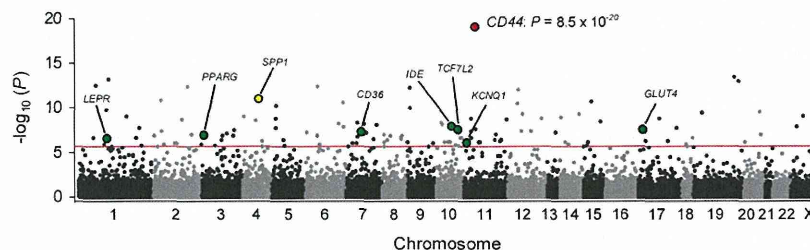

Fig. 1. eGWAS for T2D using a χ^2 analysis. Plot of -log$_{10}$ (*P* value) (*y* axis) by chromosomal position (*x* axis). *P* values for each gene were calculated from our eGWAS across 130 microarray experiments with 1,175 T2D case-control microarray samples (591 T2D cases and 584 controls) as the likelihood of finding repeated differential expression compared with expected using a χ^2 analysis, or a Fisher's exact test (Fig. S3). Our top gene, *CD44*, showed a significant differential expression in 78 experiments ($P = 8.5 \times 10^{-20}$). The red line indicates the Bonferroni threshold ($P = 2.0 \times 10^{-6}$). The green dots indicate several well known T2D-susceptibility genes.

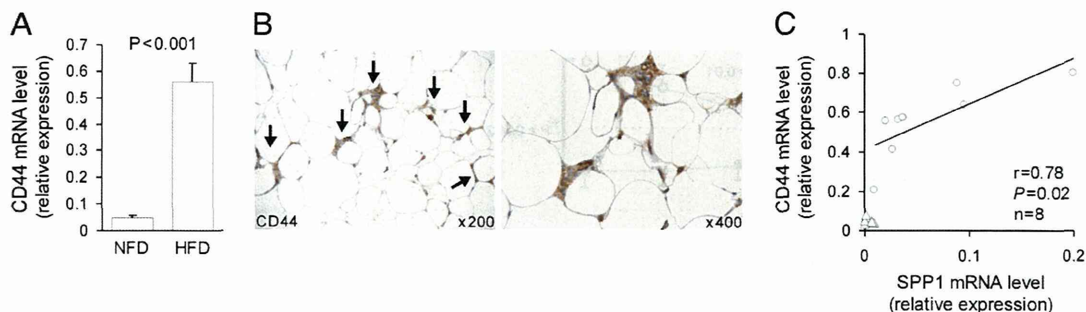

Fig. 2. CD44 expression in adipose tissue of obese mice. (*A*) *CD44* mRNA expression levels in epididymal adipose tissues in C57BL/6J mice fed either a NFD or HFD ($n = 8$ per group). (*B*) CD44 immunoreactivity (arrows, DAB chromogen; brown) in epididymal adipose tissues from C57BL/6J mice fed a HFD. Sections were counterstained with hematoxylin (blue). (*C*) *CD44* and *SPP1* gene expression profiles in the HFD ($n = 8$; circles) and NFD ($n = 8$; triangles) groups. Correlation between *CD44* and *SPP1* mRNA expression in the HFD group (circles) was analyzed by using the Pearson's correlation test. Gene expression was monitored by using real-time RT-PCR and normalized to expression of *GAPDH* mRNA.

Interestingly, even in the mice fed a NFD, the ameliorative effects of *CD44* deficiency on insulin sensitivity was observed in both glucose and insulin tolerance tests (Fig. 3 *C* and *D*, dashed lines). These results that we obtained with CD44-deficient mice confirm that CD44 molecules are essential for macrophage recruitment and inflammation in adipose tissue and the development of insulin resistance in diet-induced obese mice.

CD44 Blockade Decreases Blood Glucose Levels and Adipose Macrophage Infiltration. Several in vivo studies have shown that anti-CD44 monoclonal antibody (CD44 mAb) treatment exhibits robust antiinflammatory effects in animal models of immune-mediated diseases (27–30). We therefore sought to investigate whether CD44 blockade might demonstrate a therapeutic effect on T2D. We performed daily i.p. injections of CD44 mAb in diabetic model mice for 1 wk and found that blood glucose levels and adipose macrophage accumulation were significantly reduced in CD44 mAb treated mice compared with isotype-control treated mice, despite continuing on the HFD and similar body weight increase during the treatment (Fig. 4 *A* and *B*). When adipose inflammation was quantified as the average number of CLSs per low power field, EWAT from CD44 mAb treated mice contained significantly fewer inflammatory cells compared with control treated mice (2.4 ± 0.5 versus 5.4 ± 0.7; $P = 0.0005$). Collectively, these effects of CD44 mAb clearly show that CD44 molecules are required for the recruitment of macrophages into obese adipose tissue and the maintenance of inflammatory reactions there, and the CD44 receptor may be useful as a therapeutic target for T2D.

To gain additional insight into the clinical importance of CD44 in obese fat, we performed an immunohistochemical analysis of CD44 in omental adipose tissue in human obesity. Consistent with our mouse model observation, we discovered that $CD44^+$ cells infiltrated into the stroma of adipose tissue in human obese subjects, suggesting that CD44 molecules may mediate macrophage migration into obese adipose tissue in humans (Fig. 5*A*).

Soluble CD44 shed from cell surfaces exists in normal human serum. To estimate the relevance between CD44 protein and glucose homeostasis in human subjects, we evaluated the relationship between serum levels of CD44 and metabolic traits in human and found that serum CD44 was positively correlated with glycemic control and insulin resistance as estimated through HbA1c ($n = 55, r = 0.49, P < 0.001$) and HOMA-IR ($n = 55, r = 0.29$, $P = 0.03$) (Fig. 5 *B* and *C*). We then classified the 55 subjects into two groups according to the WHO criteria (31): hyperglycemia ($n = 21$: "diabetes mellitus" + "impaired glucose regulation") and normoglycemia ($n = 34$: "Normal Glucose Tolerance"), and found that the serum levels of CD44 were significantly higher in the hyperglycemic group than the normoglycemic group (246.9 ± 15.0 versus 209.5 ± 9.8 ng/mL; $P = 0.02$). These results suggest that CD44 protein may be released from insulin-resistant and diabetic tissues into circulation in humans.

Discussion

Adipose tissue inflammation is thought to be a pivotal event leading to the metabolic syndrome, insulin resistance, and T2D.

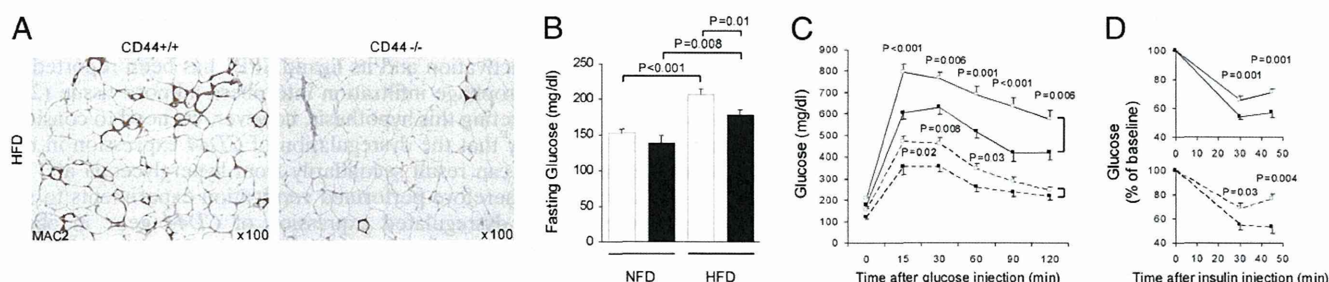

Fig. 3. Histological and metabolic analyses of wild-type $CD44^{+/+}$ and $CD44^{-/-}$ mice. (*A*) Inflammatory cell (macrophage) content determined by immunohistochemical staining for Mac-2 (DAB, brown; hematoxylin, blue) in epididymal adipose tissues from $CD44^{-/-}$ and $CD44^{+/+}$ mice fed a HFD. (*B–D*) Metabolic measurements on $CD44^{+/+}$ (open bars and symbols; diabetes-prone) and $CD44^{-/-}$ (filled bars and symbols) mice fed either a HFD ($n = 16$ per group; solid lines) or a NFD ($n = 10$ per group; dashed lines). (*B*) Fasting blood glucose. (*C*) Glucose tolerance tests [i.p. glucose (2 g/kg body weight)] after a 14-h overnight fast. Venous blood was obtained for measurement of blood glucose at 0, 15, 30, 60, 90, and 120 min after the injection. (*D*) Insulin tolerance tests [i.p. insulin (1.0 unit/kg body weight)] after a 4-h fast. Venous blood was obtained for measurement of blood glucose at 0, 30, and 45 min after the injection.

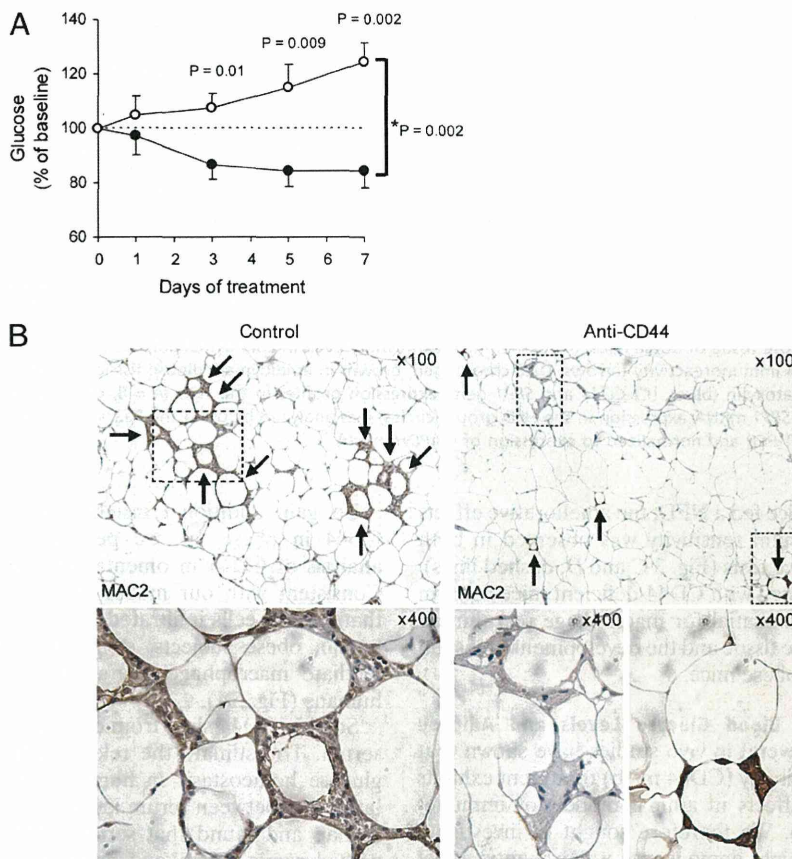

Fig. 4. Anti-CD44 antibody treatment for diabetic mice. (*A*) HFD-fed C57BL/6J mice were injected intraperitoneally with purified rat anti-mouse CD44 (IM7; 553131, BD Pharmingen) ($n = 6$; filled circles) or purified rat IgG2b, κ isotype control (A95-1; 559478, BD Pharmingen) ($n = 8$; open circles) for 1 wk (100 μg at day 0 and 50 μg at day 1–7). Morning blood glucose was measured at day 0, 1, 3, 5, and 7 during the treatment. The effect of anti-CD44 treatment on blood glucose levels was evaluated with two-way repeated measures ANOVA (*P; treatment × time). Comparisons between two groups were performed by using the two-tailed Welch's *t* test. Data are represented as mean ± SE. (*B*) Epididymal adipose tissues from control and anti-CD44 antibody-treated mice were analyzed for inflammatory cell (macrophage) content by using a Mac-2 antibody (magnified as indicated). Arrows indicate crown-like structures (CLSs) surrounding individual adipocytes.

Genome-wide experimental methods to identify disease genes, such as association studies (GWAS), linkage studies (GWL), and eQTL analyses, have been performed for T2D by many researchers to date, and these methods have revealed a number of loci to be linked with T2D (3–14, 21, 32–35). However, these genetically mapped loci do not fully account for the tissue-specific mechanisms underlying the development of T2D. In this report, we proposed an alternative methodology, eGWAS, computing the likelihood of finding repeated differential expression of a gene in disease-related tissues by using thousands of case-control microarray samples. To detect additional genes functionally implicated in the molecular pathogenesis of T2D, we successfully performed an eGWAS to identify T2D candidate genes and verified our top candidate gene, *CD44*, an immunological cell receptor, plays a significant role in the development of adipose tissue inflammation and insulin resistance in mouse models and human subjects.

In our T2D eGWAS, we identified in total 127 genes as significantly repeatedly dysregulated with *P* values $<2.0 \times 10^{-6}$ (under the Bonferroni-corrected threshold) and rediscovered several genes that have been shown to be important in T2D pathogenesis, including *TCF7L2*, *PPARG*, *KCNQ1*, *IDE*, *CD36*, *GLUT4*, and *LEPR*, supporting the validity of our methodology. Of the 127 genes, we found that more than one-half were implicated in the "receptor" or "ligand" activity by using the GO term enrichment analysis (Fig. S2). Interestingly, *SPP1*, encoding a ligand for the CD44 receptor, was shown to be included also in our top-ranked gene list (Fig. 1 and Table S3). Furthermore, our eGWAS provided the prediction of tissue specificity of gene expression by calculating a distribution of scores for each gene across tissues, indicating that *CD44* mRNA expression was more highly up-regulated in adipose tissue in diabetes than other tissues (Fig. S5). These analysis results led us to the speculation that our top-most candidate gene, *CD44*, encoding an immune-cell surface receptor, may be implicated in adipose tissue inflammation causing insulin resistance in obesity, given the fact that the CD44 receptor is known to regulate immune-cell migration and activation and its ligand SPP1 has been reported to mediate macrophage infiltration into obese adipose tissue (25). When interpreting this hypothesis, however, we need to consider the possibility that the dysregulation of *CD44* expression in relevant tissues can result secondarily from hyperglycemia and diabetes. We therefore performed verification experiments to see whether the dysregulated expression of *CD44* gene in obese adipose tissue can be accepted as the cause of insulin resistance.

In our functional tests for the *CD44* gene products, we showed that knocking out the receptor *CD44* leads to striking reduction in immune-cell infiltration into visceral adipose tissue and improvements in insulin sensitivity in mouse models, and anti-CD44 monoclonal antibody treatment decrease the blood glucose levels and visceral adipose tissue macrophages in diabetic obese mice. In addition, we showed that higher serum levels of