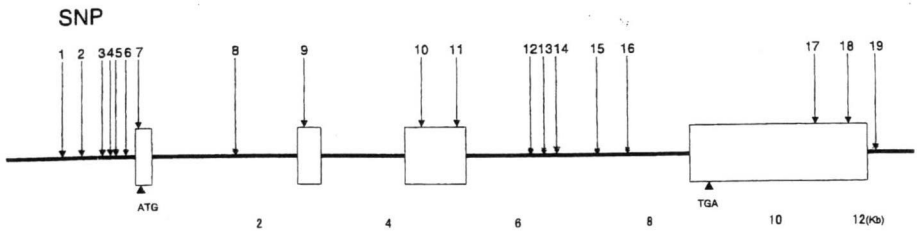

Fig. 1. Polymorphisms of *KLF11* identified in this study. The locations of the polymorphisms described in the text are shown. Nucleotide indicates the location of the SNP relative to the A of ATG of the initiator Met of KLF11.

Table 2. Frequencies of coding SNPs in *KLF11* in controls and in patients with T2DM

SNP No.	Position genome	db SNP ID	Nucleotide change	Exon	minor allele number		
					Controls (n = 563)	Late-onset T2DM (n = 553)	Early-onset T2DM (n = 182)
9	2477	novel	G>A (R29Q)	2	1	3	2
10	3922	novel	C>A (S124F)	3	0	0	1
11	4806	rs11687357	T>A (V395V)	3	26 (n = 96)	21 (n=96)	56

The nucleotide indicates the location of the SNP relative to the A of ATG of initiator Met of KLF11

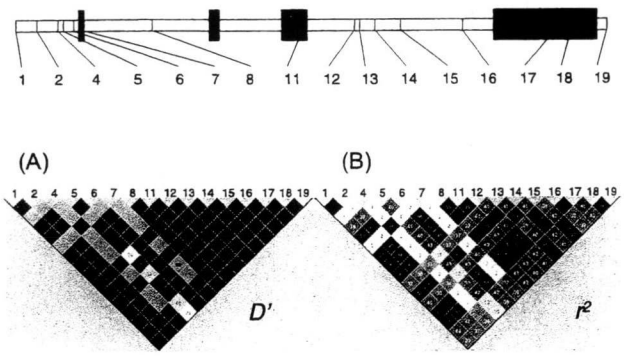

Fig. 2. Pairwise LD of *KLF11* evaluated by D' and r^2. The panel shows a Haploview representation of LD (D' and r^2) based on genotyping data with 16 polymorphisms from control subjects (n = 96) (A) Pairwise combination with LD of $D' = 1$ and LOD≥2, $D' = 1$ and LOD<2, $D'<1$ and LOD≥2, $D'<1$ and LOD<2 is shown with red, blue, pink and white. (B) Pairwise combination with LD of $r^2 = 1$, $0<r^2<1$, $r^2=0$ is shown in black, gray and white.

Evaluation of the pattern of LD in KLF11

Sixteen polymorphisms with frequencies of more than 10% were used to define haplotypes and to evaluate the pattern of LD. As shown in Fig. 2, the single, large LD block appears in this region. Haplotype combinations were estimated with 96 control subjects. Four major haplotypes comprising tag SNPs (SNP-2, -3, -5, -6, -and -13) were inferred in the entire *KLF11* region.

Mutation screening of KLF11 and association study in T2DM patients

All exons and the putative promoter region (–1389 to +162 bp) of *KLF11* were examined in 182 early-onset T2DM patients, 96 late-onset T2DM patients, and 96 control subjects. A total of 3 cSNPs (R29Q, S124F and V395V), two of which, R29Q and S124F, are novel, were found (Table 2). R29Q mutation was found in five T2DM (3 from late-onset and 2 from early-onset) patients and in one control (Fisher's P-value = 0.245). S124F mutation was found in one early-onset T2DM patient and was not found in controls. An association study using tag SNPs in T2DM patients was performed, but no association of *KLF11* variants with T2DM was found (Table 3A & B). No haplotypes were found to be susceptible to T2DM by χ^2 test. Using permutation testing (100 permutations), no significant differences in haplotype frequencies between control and T2DM were found (Table 4).

Functional analysis of KLF11 mutant proteins

The results of alignment by CLUSTALW show that R29 of KLF11 is conserved among human, mouse, and rat, while S124 is not conserved (Fig. 3). The transcriptional activity of the R29Q and S124F mutants of KLF11 were therefore compared with that of wild type

Table 3A. Association study using tag SNPs in *KLF11* in patients with late-onset T2DM and controls

SNP		Frequencies of genotype			Frequencies of alleles		allele		genotype		dominant		recessive	
		Control	T2DM		Control	T2DM	χ^2	*P*	χ^2	*P*	χ^2	*P*	χ^2	*P*
2	W/W	0.747	0.773	W	0.868	0.882	0.159	0.689	0.237	0.888	0.208	0.647	0.006	0.936
	W/M	0.242	0.214	M	0.132	0.117								
	M/M	0.01	0.011											
5	W/W	0.564	0.492	W	0.724	0.725	0.0003	0.984	4.81	0.09	0.755	0.384	2.684	0.101
	W/M	0.320	0.464	M	0.276	0.274								
	M/M	0.115	0.042											
6	W/W	0.734	0.732	W	0.861	0.854	0.036	0.848	0.451	0.797	0.0005	0.982	0.436	0.508
	W/M	0.255	0.244	M	0.138	0.145								
	M/M	0.01	0.023											
13	W/W	0.486	0.505	W	0.707	0.698	0.224	0.636	0.517	0.772	0.042	0.839	0.517	0.472
	W/M	0.418	0.415	M	0.293	0.302								
	M/M	0.094	0.078											

Association study was performed for the SNPs (SNP2, 5, and 6) in the promoter region with 96 subjects each from 553 T2DM and 563 controls by direct sequencing, and for SNP (SNP13) in the 553 T2DM patients and 563 controls by TaqMan assay. M, Mutant; W, wild type.

Table 3B. Association study using tag SNPs in *KLF11* in patients with late-onset T2DM and controls

SNP3	Frequencies of genotype						*P**1
	1 1	1 2	1 3	2 2	2 3	3 3	
Control	0.515	0.252	0.178	0.021	0.021	0.010	0.297
T2DM	0.559	0.178	0.166	0	0.071	0.023	

SNP3	Frequencies of allele			*P**2
	1	2	3	
Control	0.731	0.157	0.110	0.491
T2DM	0.732	0.125	0.142	

SNP3	Frequencies of genotype		*P**3	Frequencies of genotype		*P**4
	2 2	non 2 2		3 3	non 3 3	
Control	0.021	0.978	0.181	0.010	0.989	0.489
T2DM	0	1		0.023	0.976	

SNP3	Frequencies of genotype		*P**5	Frequencies of genotype		*P**6
	2 2+1 2+2 3	1 1+1 3+3 3		3 3+1 3+2 3	1 1+1 2+2 2	
Control	0.294	0.705	0.502	0.210	0.789	0.418
T2DM	0.250	0.750		0.261	0.738	

Association study was performed for SNP3.

Allele 1 indicates 4 CCG repeats. Allele 2 indicates 2 CCG repeats. Allele 3 indicates 5 CCG repeats.

*1 Triallelic varient was analyzed in a 2 × 6 contingency table with genotype frequencies.

*2 Triallelic varient was analyzed in a 2 × 3 contingency table with allele frequencies.

*3 Analyses in the recessive model of 22 genotype vs. others.

*4 Analyses in the recessive model of 33 genotype vs. others.

*5 Analyses in the dominat model of 22 + 12 + 23 genotypes vs. 11 + 13 + 33 genotypes.

*6 Analyses in the dominat model of 33 + 13 + 23 genotypes vs. 11 + 12 + 22 genotypes.

Table 4. Major haplotypes and its combinations in *KLF11* and risk of late-onset T2DM

SNP	Haplotype frequencies		*P*-value
2 3 5 6 13	Control	T2DM	
1 1 1 1 1	0.59	0.57	0.463
1 1 2 1 1	0.127	0.125	0.504
1 3 1 1 2	0.113	0.129	0.985
2 2 2 2 2	0.119	0.113	0.640
Haplotype combination	**Risk (O.R.)**	**95% C.I.**	
1 1 1 1 1/1 1 1 1 1	1.29	0.71–2.35	
1 1 1 1 1/1 1 2 1 1	1.43	0.55–3.75	
1 1 1 1 1/1 3 1 1 2	0.60	0.26–1.36	
1 1 1 1 1/2 2 2 2 2	0.75	0.32–1.77	
1 1 2 1 1/1 1 2 1 1	0.32	0.03–3.19	
1 1 2 1 1/1 3 1 1 2	2.02	0.18–22.75	
1 1 2 1 1/2 2 2 2 2	0.19	0.02–1.67	
1 3 1 1 2/1 3 1 1 2	1.00	0.06–16.3	
1 3 1 1 2/2 2 2 2 2	6.38	0.75–54.13	

The haplotype frequencies were estimated from the genotype of 96 controls and 96 T2DM subjects. The O.R and 95% C.I. of each haplotype combination relative to the other haplotype combinations as a group are shown. The numbers 1 and 2 indicate major allele and minor allele (SNP 2, SNP 5, SNP 6, SNP 13). SNP 3; allele 1 indicates 4 CCG repeats, allele 2 indicates 2 CCG repeats, and allele 3 indicates 5 CCG repeats.

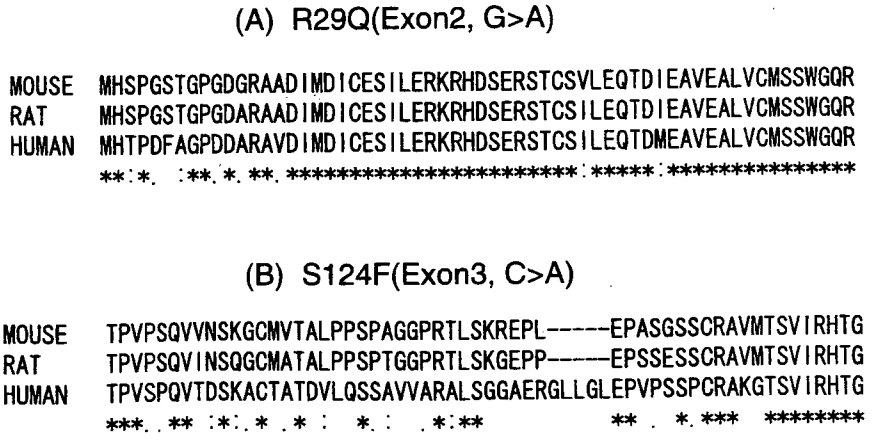

Fig. 3. The results of alignment of KLF11 among human, rat and mouse. A) Red indicates the novel mutation, R29Q, and blue indicates the region of SID. B) Red indicates the novel mutation, S124F, and blue indicates the region of R2. SID: mSin3A interaction domain; R2: transcriptional repression domain * completely conserved,: highly conserved,. moderately conserved.

KLF11. KLF11, catalase1, or Smad7 expression in MIN6-m9 cells was first confirmed, and all of these genes were endogenously expressed in both MIN6-m9 cells (data not shown). Wild-type KLF11 reduced all of these promoter activities to approximately 0.5-fold. Significantly reduced activity of R29Q mutant repression was found in catalase1 promoter, and reduced S124F mutant repression activity was found both in insulin and catalase1 promoter. However, the differences in activity between wild-type and these mutants of KLF11 were very small (Fig. 4).

Identification of the minimal promoter region of human KLF11

To identify the region essential for basal promoter activity of *KLF11*, the relative luciferase activity of the three sizes of 5' deletion reporter constructs was evalu-

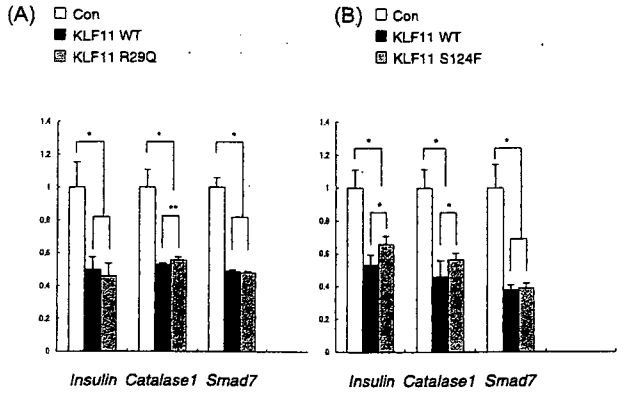

Fig. 4. Transactivation activity of mutant KLF11. Transcription activity with empty vector, wild type, and mutant KLF11 (500 ng) was analyzed by co-transfection assay using reporter vector of Insulin, Catalase-1 and Smad7 promoter-PGL3 (500 ng) and a Renilla luciferase (17 ng) as internal control. (A) The results of R29Q mutant (n = 6) (B) The results of S124F mutant (n = 6) Data are expressed as means ± SD.
*P<0.005, **P<0.05

ated. As shown in Fig. 5, the shortest promoter c struct comprising each of the four major haplotyp had basal activity both in MIN6-m9 cells and Hep cells, suggesting that the basal promoter is involved the region between –250 and +162 bp. The seco longest and longest promoter constructs had stron activity in HepG2 cells, suggesting that the essen enhancer element for promoter activity of *KLF11* incorporated in the region between –896 and –250 while the second longest and longest promoter c structs had similar activity in MIN6-m9 cells. One the promoter polymorphisms, SNP-4, is located at sixth Sp1 binding site from the transcription start s and is not incorporated in one of the major haplotyp 1-3-1-1 (Table 4, Fig. 6). A significant difference promoter activity between this haplotype and the ot three major haplotypes was not detected in the p moter assay in MIN6-m9 cells. One of the ma haplotypes, 2-2-2-2, showed the highest transcriptio activity with the longest and the second longest c structs (1.43 ± 0.046 fold, P = 0.0064; 1.28 ± 0.(fold, P = 0.0069) (Fig. 5).

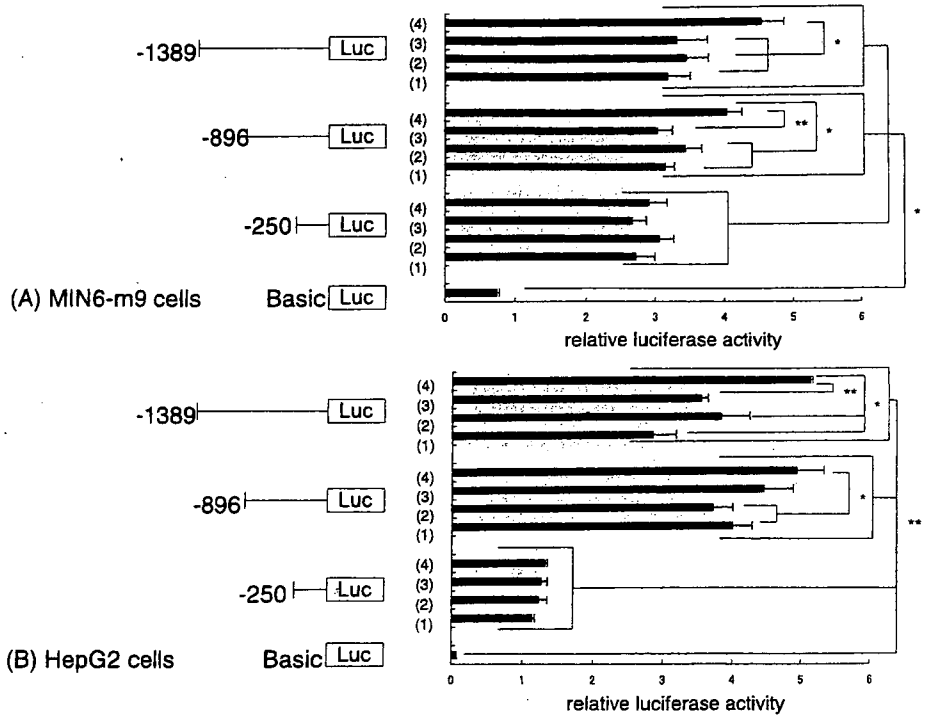

Fig. 5. Identification of *KLF11* promoter region using relative luciferase assays. The four haplotypes of the promoter region indicated by (1) 1-1-1-1 (2) 1-1-2-1 (3) 1-3-1-1 (4) 2-2-2-2. The numerals 1 and 2 indicate major allele and minor allele (S 2, SNP 3, SNP 5, SNP 6). SNP 3, allele 1 indicates 4 CCG repeats, allele 2 indicates 2 CCG repeats, allele 3 indicates 5 C repeats. (A) The result in MIN6-m9 cells (n = 3) (B) The result in HepG2 cells (n = 3). Results are expressed as means ± of a representative of experiments performed in triplicate. *P<0.05, **P<0.005

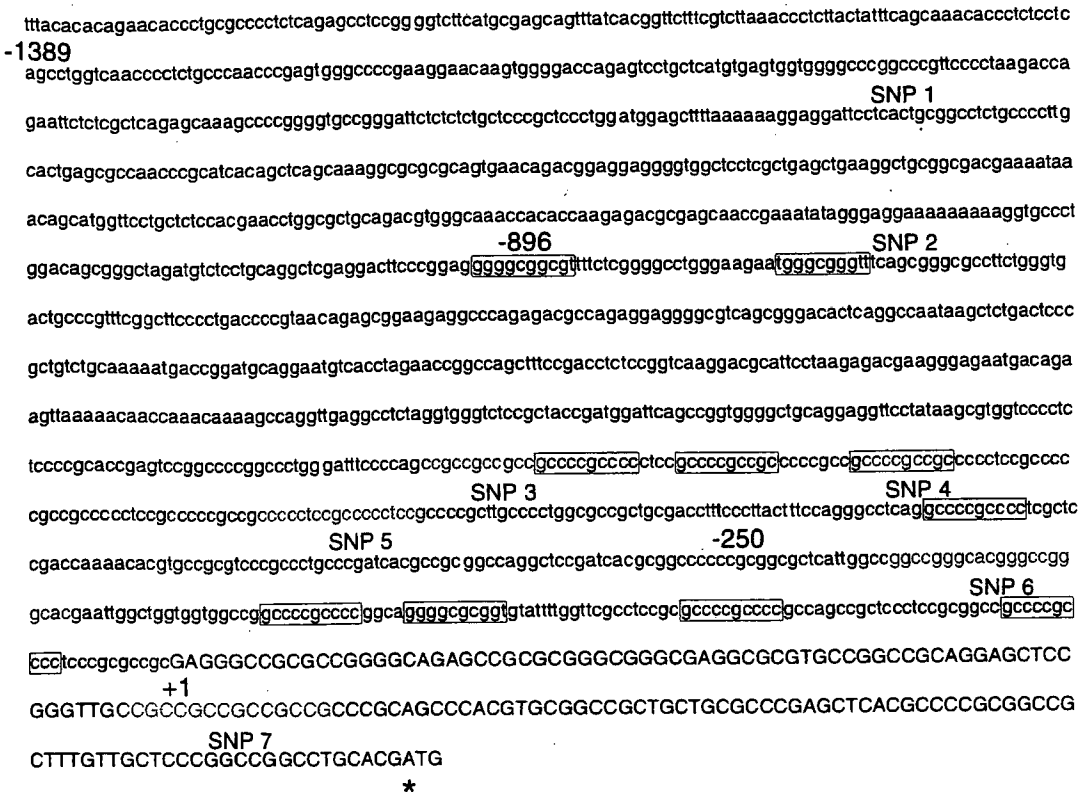

Fig. 6. Nucleotide sequence of the promoter region of *KLF11* from transcription start site to −1389 bp of 5' flanking region. Consensus sequences for Sp1 are boxed. Red color indicates the sites of polymorphisms. Large characters indicate 5' UTR region of *KLF11* and +1 denotes a transcription start site. Asterisk indicates the A of ATG of the initiator Met of KLF11.

Discussion

Genetic analysis of *KLF11* reveals two rare variants, A347S and T220M, which segregate in families with early-onset type 2 diabetes. In addition, analysis of 1,696 type 2 diabetic patients and 1,776 controls found one frequent polymorphic Q62R variant that significantly associates with late-onset type 2 diabetes in North European populations [17]. Florez *et al.* reported that KLF11 Q62R polymorphism is not associated with late-onset type 2 diabetes in 8,676 subjects of northern-European ancestry, and both A347S and T220M rare variants also are not found [22]. Recently, Ma *et al.* reported that neither Q62R nor any other common variant in *KLF11* was associated with late-onset type 2 diabetes in the Pima population [23]. Normoglycemic R62 carriers showed a significantly decreased plasma insulin level at 60 and 120 min after oral glucose load, leanness, and increased insulin sensitivity [17]. These characteristics clearly differ from those of Caucasian patients with T2DM, which are often caused by impaired insulin action mainly due to obesity. Obesity and insulin resistance are closely associated with type 2 diabetes in Caucasians, while type 2 diabetes in Japanese is characterized primarily by pancreatic beta-cell dysfunction [24, 25]. Thus, *KLF11* is a candidate diabetogenic gene in Japanese. Tanahashi *et al.* reported identification of eight variants, but were unable to find any missense mutation or association of common *KLF11* variants in late-onset Type2 diabetes in 1818 Japanese subjects [26].

In the present study, the minimal promoter region identified in this study, 5'UTR, coding region, 3'UTR, and flanking introns were screened for mutations in unrelated Japanese subjects with 182 early-onset and 96 late-onset T2DM patients. Our analysis revealed a novel missense mutation, R29Q (G>A), in exon 2 in two early-onset type 2 diabetes patients and three late-onset T2DM patients in the heterozygous state and S124F in exon 3 in one early-onset type 2 diabetes patient, but none of the A347S, T220M, Q62R, and promoter variant −1659G>C mutations previously

identified in Northern Europeans in *KLF11*.

We examined about 13 kb covering the entire coding region of *KLF11* and identified a total of 19 genetic variations including 6 novel variants. We defined haplotypes based on the LD pattern estimated using the 16 most frequent SNPs, the frequencies of which were more than 10%, and performed an association study with T2DM. No association of *KLF11* variants with T2DM was found and no haplotypes were found to show susceptibility to T2DM. We examined one variant, rs4073397, located 21 kb upstream from the coding region and outside of the LD block of the *KLF11* region, that shows a marginal association with T2DM in Japanese [26], but we could not replicate the association in this study with power similar to that of the former report ($P = 0.268$ and 0.576 for allele and genotype analysis, respectively; OR 1.2; RAF 0.5; $\alpha = 0.05$; $1-\beta = 63.9\%$).

We searched for consensus sequences of transcription factor binding sites in the promoter region by using TFSEARCH. As shown in Fig. 6, no TATA-box and ten Sp1 binding sites were found, but there were no differences in haplotype frequencies of the *KLF11* promoter region between T2DM patients and controls. It is of interest that one of the promoter polymorphisms, SNP-4, is located at the sixth Sp1 binding site from the transcription start site and that this binding site is not incorporated in one of the major haplotypes, but significantly reduced promoter activity of this haplotype was not detected by promoter assay in MIN6-m9 cells. One of the major haplotypes comprising all of the minor alleles showed the highest transcriptional activity with the longest and the second longest constructs. No association of the haplotype with occurrence of T2DM was found, but the findings are inconclusive due to the low statistical power of this study. In any case, the effect of the KLF11 expression level itself on the onset of type 2 diabetes should be slight, if any at all.

RT-PCR demonstrated endogenous KLF11 mRNA expression in whole rat pancreas, human pancreas, and mouse MIN6-m9 cells (data not shown). We therefore examined whether the KLF11 mutant affects transcriptional regulation of the human insulin gene. We found, contrary to Neve's report [17], that wild-type KLF11 inhibited insulin promoter activity as Niu *et al.* reported [18], but no difference in activity between wild and KLF11 mutants was found. Inhibition by wild-type KLF11 was similar in the three beta cell lines, INS-1E, beta-TC3 in the previous study, and MIN6-m9 in this study, indicating stable performance of KLF11 inhibition of insulin promoter activity in rodent beta cells.

KLF11 is involved in the TGF-β pathway, which plays a critical role in the development and homeostasis of exocrine and endocrine pancreas via Smad signaling [16, 27]. KLF11 may affect pancreatic beta cell function by modulating the expression of free radical scavengers such as superoxide dismutase (SOD) 2 and catalase1, recently identified as KLF11 target genes [10]. Oxidative stress is generally believed to be involved in the progression of pancreatic beta cell dysfunction found in T2DM [28, 29]. Since a decreased expression level of these antioxidant enzyme genes might contribute to the occurrence of type 2 diabetes, we examined the effects of the KLF11 R29Q and S124F variants on the expression levels of these target genes.

We then examined whether or not KLF11 mutants affect transcriptional regulation of the catalase1 and Smad7 genes. A significant but small difference in promoter activity between wild-type and R29Q mutant was found in catalase1 only. R29Q mutation is located 12 amino acids upstream of SID (mSin3A interaction domain)/R1 domain, which determines repression activity of KLF11 [10, 30–33]. This mutant of KLF11, although showing somewhat different activities in the catalase1 promoter, showed similar transcriptional activity compared with wild-type KLF11. Although it is inconclusive whether or not R29Q mutation contributes to the development of type 2 diabetes because of the small sample number in this study (Fisher's *P*-value = 0.245), our results suggest that the mutation does not contribute to the development of type 2 diabetes through alteration of the expression level of these target genes. No unusual clinical characteristics were found in subjects with R29Q mutation.

S124F mutation is located between the SID/R1 and R2 domain, and is not conserved among rat, mouse, and human, *i.e.*, proline for rat and mouse and serine for human. Similar data were obtained from S124F mutant of KLF11, which showed somewhat different activities in insulin and catalase1 promoters and similar transcriptional activity compared with wild-type KLF11, suggesting little biological effect. No useful clinical information was available for the patient with S124F mutant.

In this study, we focused on the effects of KLF11 on

beta cell function. In addition, the caveolin-1 gene, which is highly expressed in adipose tissue, is repressed by KLF11 in a cholesterol-dependent manner [34]. The findings that caveolin-1 binding to the insulin receptor stimulates both kinase activity and recruitment of the insulin receptor to lipid rafts at the plasma membrane, and that insulin receptor mutations impairing calveolin binding result in T2DM [35] indicate that caveolin-1 is essential for insulin signaling and suggest that KLF11 might affect insulin sensitivity as well.

In conclusion, we identified two novel missense variants of *KLF11* and the minimal promoter region of *KLF11*, but neither of the variants nor the haplotypes identified in this study was associated with the occurrence of type 2 diabetes in Japanese, although the findings are inconclusive due to the low statistical power of the study. Further investigation with a larger sample number is required to determine the magnitude of the contribution of the R29Q and S124F mutations to the development of diabetes and the role of KLF11 in normal pancreatic β-cell function. Identification of the minimal promoter region of *KLF11* should be useful in further elucidation of its transcriptional regulation mechanism.

Acknowledgments

We thank Tsuchida H, Yokoyama K, Kawada J, Uda I, and Ibe Y for technical assistance. Grants: This work was supported by a Health and Labor Science Research Grant for research on Human Genome and Tissue Engineering from the Japanese Ministry of Health, Labor and Welfare, a Grant-in-Aid for Scientific Research from the Japanese Ministry of Science, Education, Sports, Culture and Technology, and a New Energy and Industrial Technology Development Organization Grant.

References

1. Bieker JJ (2001) Kruppel-like factors: three fingers in many pies. *J Biol Chem* 276: 34355–34358.
2. Lomberk G, Urrutia R (2005) The family feud: turning off Sp1 by Sp1-like KLF proteins. *Biochem J* 392: 1–11.
3. Turner J, Crossley M (1999) Mammalian Kruppel-like transcription factors: more than just a pretty finger. *Trends Biochem Sci* 24: 236–240.
4. Cook T, Gebelein B, Mesa K, Mladek A, Urrutia R (1998) Molecular cloning and characterization of TIEG2 reveals a new subfamily of transforming growth factor-beta-inducible Sp1-like zinc finger-encoding genes involved in the regulation of cell growth. *J Biol Chem* 273: 25929–25936.
5. Kaczynski J, Cook T, Urrutia R (2003) Sp1- and Kruppel-like transcription factors. *Genome Biol* 4: 206.
6. Black AR, Black JD, Azizkhan-Clifford J (2001) Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. *J Cell Physiol* 188: 143–160.
7. Asano H, Li XS, Stamatoyannopoulos G (1999) FKLF, a novel Kruppel-like factor that activates human embryonic and fetal beta-like globin genes. *Mol Cell Biol* 19: 3571–3579.
8. Ou XM, Chen K, Shih JC (2004) Dual functions of transcription factors, transforming growth factor-beta-inducible early gene (TIEG). 2 and Sp3, are mediated by CACCC element and Sp1 sites of human monoamine oxidase (MAO) B gene. *J Biol Chem* 279: 21021–21028.
9. Scohy S, Gabant P, Van Reeth T, Hertveldt V, Dreze PL, Van Vooren P, Riviere M, Szpirer J, Szpirer C (2000) Identification of KLF13 and KLF14 (SP6), novel members of the SP/XKLF transcription factor family. *Genomics* 70: 93–101.
10. Fernandez-Zapico ME, Mladek A, Ellenrieder V, Folch-Puy E, Miller L, Urrutia R (2003) An mSin3A interaction domain links the transcriptional activity of KLF11 with its role in growth regulation. *EMBO J* 22: 4748–4758.
11. Robertson RP, Harmon J, Tran PO, Poitout V (2004) Beta-cell glucose toxicity, lipotoxicity, and chronic oxidative stress in type 2 diabetes. *Diabetes* 53: Suppl 1 S119–124.
12. Massague J, Blain SW, Lo RS (2000) TGFbeta signaling in growth control, cancer, and heritable disorders. *Cell* 103: 295–309.
13. Ellenrieder V, Buck A, Harth A, Jungert K, Buchholz M, Adler G, Urrutia R, Gress TM (2004) KLF11 mediates a critical mechanism in TGF-beta signaling that is inactivated by Erk-MAPK in pancreatic cancer cells. *Gastroenterology* 127: 607–620.
14. Kim SK, Hebrok M (2001) Intercellular signals regulating pancreas development and function. *Genes Dev* 15: 111–127.
15. Sanvito F, Herrera PL, Huarte J, Nichols A, Montesano R, Orci L, Vassalli JD (1994) TGF-beta 1 influences the relative development of the exocrine and endocrine pancreas *in vitro*. *Development* 120: 3451–3462.
16. Smart NG, Apelqvist AA, Gu X, Harmon EB, Topper JN, MacDonald RJ, Kim SK (2006) Conditional expression of Smad7 in pancreatic β cells disrupts TGF-β

signaling and induces reversible diabetes. *Plos Biol* 4: e39.
17. Neve B, Fernandez-Zapico ME, Ashkenazi-Katalan V, Dina C, Hamid YH, Joly E, Vaillant E, Benmezroua Y, Durand E, Bakaher N, Delannoy V, Vaxillaire M, Cook T, Dallinga-Thie GM, Jansen H, Charles MA, Clement K, Galan P, Hercberg S, Helbecque N, Charpentier G, Prentki M, Hansen T, Pedersen O, Urrutia R, Melloul D, Froguel P (2005) Role of transcription factor KLF11 and its diabetes-associated gene variants in pancreatic beta cell function. *Proc Natl Acad Sci USA* 102: 4807–4812.
18. Niu X, Perakakis N, Laubner K, Limbert C, Stahl T, Brendel MD, Bretzel RG, Seufert J, Path G (2007) Human Kruppel-like factor 11 inhibits human proinsulin promoter activity in pancreatic beta cells. *Diabetologia* 50:1433–1441.
19. Gutiérrez-Aguilar R, Froguel P, Hamid YH, Benmezroua Y, Jørgensen T, Borch-Johnsen K, Hansen T, Pedersen O, Neve B (2008) Genetic analysis of KLF11 variants in 5,864 danish individuals; potential effect on insulin resistance and modified signal transducer and activator of transcription-3 binding by promoter variant –1659G>C. *J Clin Endocrinol Metab* 93: 3128–3135.
20. Kuzuya T, Nakagawa S, Satoh J, Kanazawa Y, Iwamoto Y, Kobayashi M, Nanjo K, Sasaki A, Seino Y, Ito C, Shima K, Nonaka K, Kadowaki T (2002) Report of the Committee on the classification and diagnostic criteria of diabetes mellitus. *Diabetes Res Clin Pract* 55: 65–85.
21. Minami K, Yano H, Miki T, Nagashima K, Wang CZ, Tanaka H, Miyazaki JI, Seino S (2000) Insulin secretion and differential gene expression in glucose-responsive and -unresponsive MIN6 sublines. *Am J Physiol Endocrinol Metab* 279: E773–781.
22. Florez JC, Saxena R, Winckler W, Burtt NP, Almgren P, Bengtsson Bostrom K, Tuomi T, Gaudet D, Ardlie KG, Daly MJ, Altshuler D, Hirschhorn JN, Groop L (2006) The Kruppel-like factor 11 (KLF11) Q62R polymorphism is not associated with type 2 diabetes in 8,676 people. *Diabetes* 55: 3620–3624.
23. Ma L, Hanson RL, Que LN, Mack JL, Franks PW, Infante AM, Kobes S, Bogardus C, Baier LJ (2008) Association anlysis of KLF11 variants with type 2 diabetes in Pima Indians. *J Clin Endocrinol Metab* 93: 3644–3649.
24. Fukushima M, Suzuki H, Seino Y (2004) Insulin secretion capacity in the development from normal glucose tolerance to type 2 diabetes. *Diabetes Res Clin Pract* 66: Suppl 1 S37–43.
25. Matsuoka K (2000) Genetic and environmental interaction in Japanese type 2 diabetics. *Diabetes Res Clin Pract* 50: Suppl 2 S17–22.
26. Tanahashi T, Shinohara K, Keshavarz P. Yamaguchi Y, Miyawaki K, Kunika K, Moritani M, Nakamura N, Yoshikawa T, Shiota H, Inoue H, Itakura M (2008) The association of genetic variants in Kruppel-like factor 11 and Type 2 diabetes in the Japanese population. Diabetic Medicine 25: 19–26.
27. Cook T, Urrutia R (2000) TIEG proteins join the Smads as TGF-beta-regulated transcription factors that control pancreatic cell growth. *Am J Physiol Gastrointest Liver Physiol* 278: G513–521.
28. Rehman A, Nourooz-Zadeh J, Moller W, Tritschler H, Pereira P, Halliwell B (1999) Increased oxidative damage to all DNA bases in patients with type II diabetes mellitus. *FEBS Lett* 448: 120–122.
29. Sakuaba H, Mizukami H, Yagihashi N, Wada R, Hanyu C, Yagihashi S (2002) Reduced β cell mass and expression of oxidative stress related DNA damage in the islet of Japanese type II diabetic patients. *Diabetologia* 45: 85–96.
30. Ellenrieder V, Zhang JS, Kaczynski J, Urrutia R (2002) Signaling disrupts mSin3A binding to the Mad1-like Sin3-interacting domain of TIEG2, an Sp1-like repressor. *EMBO J* 21: 2451–2460.
31. Cook T, Gebelein B, Belal M, Mesa K, Urrutia R (1999) Three conserved transcriptional repressor domains are a defining feature of the TIEG subfamily of Sp1-like zinc finger proteins. *J Biol Chem* 274: 29500–29504.
32. Hassig CA, Fleischer TC, Billin AN, Schreiber SL, Ayer DE (1997) Histone deacetylase activity is required for full transcriptional repression by mSin3A. *Cell* 89 341–347.
33. Zhang JS, Moncrieffe MC, Kaczynski J, Ellenrieder V, Prendergast FG, Urrutia R (2001) A conserved alpha-helical motif mediates the interaction of Sp1-like transcriptional repressors with the corepressor mSin3A. *Mol Cell Biol* 21: 5041–5049.
34. Cao S, Fernandez-Zapico ME, Jin D, Puri V, Cook TA, Lerman LO, Zhu XY, Urrutia R, Shah V (2005) KLF11-mediated repression antagonizes SP1/sterol-responsive Element-binding protein-induced transcriptional activation of caveolin-1 in response to cholesterol signaling. *J Biol chem* 280: 1901–1910.
35. Cohen AW, Combs TP, Scherer PE (2003) Role of caveolin and caveolae in insulin signaling and diabetes. *Am J Physiol Endocrinol Metab* 285: E1151–E1160.

Biochemical and Biophysical Research Communications 379 (2009) 499–504

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Hepatic overexpression of dominant negative Mlx improves metabolic profile in diabetes-prone C57BL/6J mice

Katsumi Iizuka [a], Jun Takeda [b], Yukio Horikawa [a,b,*]

[a] Laboratory of Medical Genomics, The Institute for Molecular and Cellular Regulation, Gunma University, Maebashi-shi, 371-8512, Japan
[b] Department of Diabetes and Endocrinology, Gifu University, Graduate School of Medicine, Gifu, 501-1194, Japan

ARTICLE INFO

Article history:
Received 9 December 2008
Available online 31 December 2008

Keywords:
Mlx
Glucose-6-phosphatase
ChREBP
Metabolic Syndrome
Elovl6
Transketolase

ABSTRACT

Mlx and ChREBP form a heterodimer to regulate glucose-mediated gene expression in the liver. This study was performed to determine if the metabolic syndrome might be improved using dominant negative Mlx (dnMlx). An adenovirus bearing dnMlx was constructed and used to test the inhibitory effect of dnMlx on lipogenesis both *in vitro* and *in vivo*. Adenoviral overexpression of dnMlx in rat hepatocytes inhibited expression of glucose-regulated genes, including *Chrebp* and *Transketolase*, which constitute a positive feedback loop in the regulation of *Chrebp* gene expression. Adenoviral overexpression of dnMlx in 25-week-old male C57BL/6J mice reduced hepatic triglyceride contents and improved glucose intolerance by inhibiting expression of *Glucose-6-phosphatase* and *Elovl6* mRNA in addition to lipogenic enzymes. In conclusion, overexpression of dnMlx improves glucose intolerance by inhibiting expression not only of lipogenic enzymes but also other important genes such as *Glucose-6-phosphatase* and *Elovl6*.

Nonalcoholic fatty liver disease is associated with metabolic syndrome and poses increased risk of cardiovascular disease. Development of non-alcoholic fatty liver disease is caused by decreased fatty acid oxidation and/or increased triglyceride synthesis. In the fed state, excess glucose uptake is converted to liver triglyceride storage. Hepatic *de novo* lipogenesis is regulated by insulin and glucose. Insulin and glucose induces glycolytic and lipogenic gene expression by sterol regulatory element binding protein (SREBP1c) and carbohydrate response element binding protein (ChREBP), respectively, [1,2]. Experiments in knockout mice indicate that SREBP1c and ChREBP exhibit *de novo* lipogenesis of 50% and 60%, respectively, [1]. Inhibition of these transcription factors reverses hepatic steatosis in obesity [1]. Thus, SREBP1c and ChREBP coordinately regulate hepatic lipogenesis. ChREBP heterodimerizes with Mlx to bind to the carbohydrate response elements (ChoRE) in the promoter of glycolytic and lipogenic genes [1].

We previously reported that gene deletion of ChREBP in ob/ob mice improved metabolic syndrome compared with ob/ob mice [3]. Consistent with our previous study, adenoviral delivery of ChREBP short hairpin RNA (shRNA) into the liver improved hepatic steatosis and insulin resistance by inhibiting hepatic lipogenesis [4]. In contrast, liver-specific inhibition of SREBP1c was found to improve hepatic steatosis but not glucose intolerance [1]. Liver-specific inhibition of PPARγ also improved hepatic steatosis, but increased glucose intolerance [1]. Clarification of the differences between ChREBP and SREBP1c or PPARγ in the development of the metabolic syndrome thus might facilitate development of clinical treatment options for the condition [1].

In this study, we investigate the mechanism of the improvement by liver-specific inhibition of ChREBP transactivity of glucose intolerance. An adenovirus bearing dominant negative Mlx (dnMlx) was constructed to test the transactivity of ChREBP *in vitro* and *in vivo*. Our findings in Ad-dnMlX-injected mice suggest that drugs disrupting the association between ChREBP and Mlx might be useful in preventing metabolic syndrome.

Materials and methods

Animals, isolation of rat primary hepatocytes, and cell culture. The protocols for all animal experiments were approved by the Institutional Animal Care and Use Committee of Gunma University Medical School (code no. 08-025). Six-week-old male Wistar rats were purchased from Japan SLC (Hamamatsu, Shizuoka, Japan) and male C57BL/6J mice were purchased from Charles River Japan (Yokohama, Kanagawa, Japan). Rat hepatocytes were isolated from 6-week-old male Wistar rats by the collagenase perfusion method. Isolated hepatocytes were suspended in DMEM supplemented with 10% fetal calf serum (FCS), 100 nM insulin, 100 nM dex, 10 nM T_3, and 100 μg/ml pen/strep [5]. Cells were seeded in 6-well plates or 10-cm dishes and grown in a humidified atmosphere of

* Corresponding author. Address: Department of Diabetes and Endocrinology, Division of Molecule and Structure, Gifu University School of Medicine, Gifu, Japan. Fax: +81 58 230 6376.
E-mail address: yhorikaw@gifu-u.ac.jp (Y. Horikawa).

0006-291X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2008.12.100

5% CO_2/95% air at 37 °C. After incubation for 4 h, the medium was replaced with DMEM containing 10 nM T_3.

Construction of plasmid and adenovirus vectors. PCR to construct pENTR vectors was performed using PrimeSTAR DNA polymerase (Takara) and primers (Sigma–Aldrich). A series of Mlx deletion mutants with an N-terminal Flag tag were amplified from mouse liver cDNA (Fig. 1A). PCR fragments (2–245aa, 67–245aa, and 88–245aa) were cloned into the pENTR vector (Invitrogen), yielding pENTR-Mlx1-1, -Mlx2-1, or -Mlx4-1, respectively. Adenovirus vectors bearing Mlx4-1 were constructed by recombining pENTR-Mlx4-1 vectors into pAd/CMV/V5-DEST using LR Clonase II Master Mix (Invitrogen) according to the manufacturer's protocol. The pcDNA6.2 vectors bearing Mlx1-1, Mlx2-1, or Mlx4-1 (pcDNA-Mlx1-1, pcDNA-Mlx2-1, or pcDNA-Mlx4-1, respectively) were constructed in the same manner as the adenovirus. pGL3-Lpk and Ad-daChREBP were the same vector and adenovirus used previously [5]. pGL4-TK-RLuc vector was purchased from Promega.

Treatment with recombinant adenovirus in rat hepatocytes. Rat isolated hepatocytes were cultured in 6-well plates in 2 ml of DMEM. Hepatocytes were infected with adenovirus bearing GFP, Mlx4-1, or dominant active ChREBP (daChREBP) at m.o.i. of 2, 10, or 50 for 2 h, media were removed, and infected cells were incubated in media with 2.5 or 25 mM glucose for 18 h. The cells were then collected and used for RT-PCR analysis.

RNA extraction and quantitative real-time PCR. Total RNA was extracted from hepatocytes or liver samples using an RNeasy Mini Plus kit (Qiagen) according to the manufacturer's protocol [5]. Taqman PCR probes for semi-quantitative RT-PCR were purchased from Applied Biosystems.

Mammalian transfection and reporter assay. Primary hepatocytes were cultured in 6-well plates in 2 ml of DMEM without antibiotics. The cells were transfected with Lipofectamine 2000 (10 μl), pGL3-Lpk (3.6 μg), or the series of pcDNA-Mlx vectors (empty, Mlx1-1, Mlx2-1, or Mlx4-1) (0.4 μg) and the pGL4-TK-Renilla luciferase vector (0.4 μg) [5]. After 24 h of incubation with 2.5 or 25 mM glucose, the cells were collected and used to measure luciferase activity (Dual Luciferase Assay System; Promega) according to the manufacturer's protocol.

Establishment of Ad dominant negative Mlx4-1-infected mice. Twenty-five-week-old male C57BL/6J mice were anesthetized with pentobarbital and injected *via* the tail vein with 3×10^{13} particle/kg body weight (BW) of Ad-GFP or Ad-Mlx4-1 in a final volume of 300 μl of phosphate buffered saline. After 5 days, the mice were sacrificed and the liver was used for extraction of total RNA and liver metabolites.

Measurement of plasma profile and liver metabolites and oral glucose tolerance test. Plasma concentrations of glucose, insulin, and triglyceride were measured as described previously [3]. Total cholesterol, alanine aminotransferase (ALT), and aspartate aminotransferase (AST) were measured using Cholesterol and Transaminase Wako test kits (Wako). Liver glycogen, triglyceride, and cholesterol contents were measured as described previously [3]. Oral glucose tolerance tests at 3 days were performed by oral injection of glucose at a dose of 1 g glucose/kg BW after 18-h fast. Blood glucose levels were measured at the designated times using a FreeStyle Kissei Meter (Kissei, Tokyo, Japan).

Data presentation and statistical methods. All data are expressed as means ± SD. The listed *n* values represent the number of single experiments performed (each experiment was duplicated). Comparisons between two groups were performed by Student's *t*-test, and $P < 0.05$ was considered significant.

Results

Mlx4-1 acts as dominant negative Mlx to inhibit glucose-induced Lpk and Fas mRNA expression

As our and other groups reported previously, inhibition of ChREBP transactivity was found to be beneficial for treating metabolic disorders in ob/ob mice [3,4]. A series of pcDNA6.2-Mlx mutants was constructed as shown in Fig. 1A and screened for dnMlx by analyzing Lpk promoter activity. Consistent with Fig. 1A, Mlx4-1 inhibited glucose-induced Lpk promoter activity by 60% (Fig. 1B). In contrast, Mlx2-1 increased Lpk promoter activity by 60% (Fig. 1B). As with the localization of Mlx mutants, Mlx2-1 was mainly localized in the nucleus, while Mlx4-1 and Mlx1-1 were localized in both the cytosol and nucleus (see Supplementary Fig. S1).

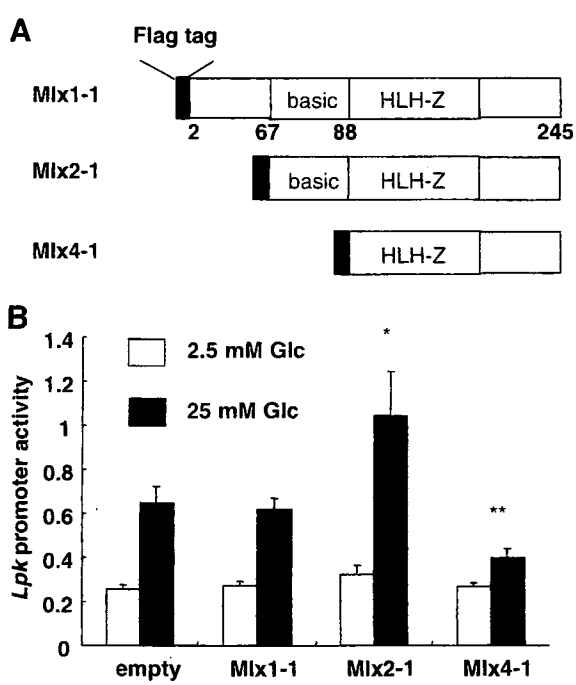

Fig. 1. (A) Schematic representation of wild-type Mlx (Mlx1-1), dominant active Mlx (Mlx2-1), and dominant negative Mlx (Mlx4-1). (B) Effects of Mlx1-1, Mlx2-1, or Mlx4-1 on Lpk promoter activity. Isolated hepatocytes were transfected with 3.6 μg of pGL3-Lpk, 0.4 μg of pGL4-TK-RLuc, and 0.4 μg of pcDNA6.2 empty, Mlx1-1, Mlx2-1, or Mlx4-1 vector using Lipofectamine 2000. Transfected cells were cultured for 24 h and used for analysis of luciferase activity. Data are means ± S.D. ($n = 6$ per group). *, ** $p < 0.05$ *vs.* pcDNA6.2 empty.

Mechanism by which Mlx4-1 inhibits ChREBP-mediated Lpk and Fasn gene expression

In the present study, a deletion mutant, Mlx4-1, was established that inhibits glucose-induced *Lpk* and *Fasn* gene expression. The effects of this mutant on *Lpk* and *Fasn* mRNA gene expression were evaluated. Consistent with Fig. 1B, Mlx4-1 dose-dendently inhibited glucose-induced *Lpk* and *Fasn* mRNA expression (Fig. 2A). When the Mlx4-1 mRNA level was increased 1680-fold, induction of these genes by glucose was completely blocked. The pGL3-3X*Lpk* and 3X*Fasn*-ChoRE vector were then used to clarify whether these inhibitory effects are due to binding of ChREBP/Mlx4-1 to ChoRE. Mlx4-1 completely inhibited ChREBP-regulated *Lpk* and *Fasn* reporter activity in rat hepatocytes (see Supplementary Fig. S2A). The inhibitory effects of Mlx4-1 on glucose-induced expression of *Chrebp* and *Transketolase* (*Tkt*) in rat primary hepatocytes were then examined. *Tkt* is an enzyme that forms xylulose-5-phosphate (Xu-5-P) from two glycolytic intermediates, glyceraldehyde 3-P (GAP) and fructose-6-P (Fru 6-P) (GAP + Fru 6-P ↔ Xu 5-P + erythrose 4-P) in the pentose phosphate pathway [1]. Xu-5-P activates

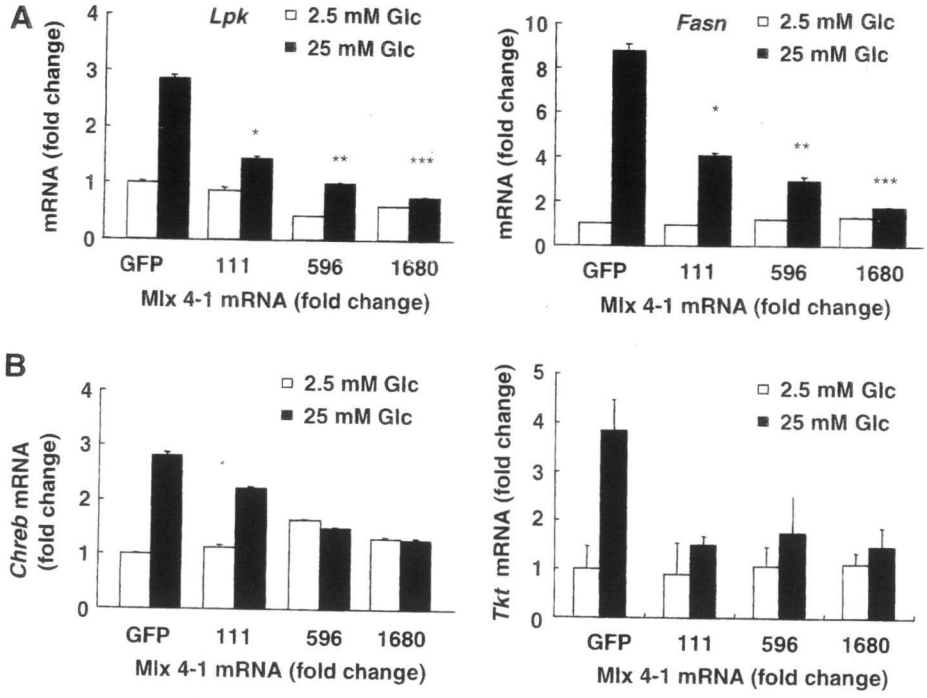

Fig. 2. (A) Overexpression of Mlx4-1 in rat hepatocytes inhibited glucose-induced *Lpk* and *Fasn* mRNA expression in a dose-dependent manner. *, **, *** $p < 0.05$ *vs.* GFP. (B) Mlx4-1 inhibited glucose-mediated *Chrebp* and *Tkt* mRNA induction in rat hepatocytes.

protein phosphatase 2A (PP2A), which in turn activates ChREBP by dephosphorylation [1]. Glucose stimulation was found to increase *Chrebp* and *Tkt* mRNA expression (see Supplementary Fig. S2B). Consistent with these data, a 1680-fold increase in Mlx4-1 completely inhibited glucose-induced *Chrebp* and *Tkt* mRNA expression in rat hepatocytes (Fig. 2B). These findings suggest that ChREBP and Tkt constitute a positive feedback loop involved in the regulation of glucose-mediated gene expression such as *Lpk* and *Fasn*.

Overexpression of Mlx4-1 improves glucose and lipid metabolism in diabetes-prone C57BL/6J mice

C57BL/6J mice develop metabolic syndrome with obesity and diabetes in response to a diet high in carbohydrate and/or fat [6]. Moreover, C57BL/6J mice develop glucose intolerance with age [6]. In this study, 25-week-old male C57BL/6J mice were used as a model diabetes-prone mouse to determine whether overexpression of Mlx4-1 improves glucose intolerance and other metabolic parameters in middle-aged C57BL/6J mice. Mice infected with Ad-GFP were used as controls. To minimize the side effects of adenovirus infection in the liver, all experiments were terminated after 5 days. Ad-Mlx4-1-injected mice appeared normal without hepatomegaly (Table 1). As shown in Table 1, body and tissue weights in Ad-Mlx4-1-infected mice were similar to those in controls. Ad-Mlx 4-1 overexpression did not affect food intake, as shown by stomach weight and food intake (Table 1). As shown in Table 2, AST and ALT concentrations were only slightly changed as compared with previously reported observations [4]. The delivery of 3×10^{13} particles/kg BW of Ad-Mlx4-1 or Ad-GFP adenovirus into C57BL/6J mice was liver-specific (data not shown). Increases of about 170-fold in endogenous Mlx mRNA levels were observed in livers infected with Ad-Mlx4-1 as compared with Ad-GFP (Fig. 3A). In liver overexpressing Ad-Mlx4-1, expression levels of the ChREBP target genes *Lpk* and *Fasn* were decreased by 84% and 65%, respectively, (Fig. 3A), suggesting that Mlx4-1 also functions as a dominant negative form of Mlx *in vivo*.

Liver triglyceride contents in Ad-Mlx4-1-injected mice were decreased by 40% due to decreased hepatic lipogenesis, but liver cholesterol contents were unchanged (Fig. 3B). Plasma lipid profile of Ad-Mlx4-1 mice was similar to that of controls (Table 2). Liver glycogen contents of Ad-Mlx4-1 mice were similar to controls (Fig. 3B). Plasma glucose concentration in Ad-Mlx4-1-injected mice was significantly decreased in both fasted and fed states (Table 2). In contrast, plasma insulin, triglyceride, and total cholesterol

Table 1
Phenotypic characteristics of 5-day-treated Ad-Mlx4-1 C57BL/6J mice.

	Ad-GFP	Ad-Mlx4-1
BW (g)	30.6 ± 2.1	30.4 ± 1.9
Food intake (g)	3.68 ± 0.26	3.60 ± 0.28
Stomach (%)	2.24 ± 0.26	2.04 ± 0.54
Liver (%)	5.8 ± 0.24	6.6 ± 0.73
White adipose tissue (%)	2.4 ± 1.4	1.6 ± 0.6
Kidney (%)	1.36 ± 0.14	1.45 ± 0.03

Data are means ± S.D. ($n = 5$ per group).

Table 2
Metabolic characteristics of 5-day-treated Ad-Mlx4-1 C57BL/6J mice.

	Ad-GFP	Ad-Mlx4-1	*P*-value
5-h fast			
Plasma glucose (mg/dl)	115 ± 6.2	81.4 ± 10.8*	$P = 0.014$
Plasma TG (mg/dl)	97.8 ± 22.1	89.6 ± 16.3	
Plasma T.Chol (mg/dl)	78.7 ± 10.1	81.5 ± 14.0	
18-h fast			
Plasma glucose (mg/dl)	86.2 ± 11	67.8 ± 6.6*	$P = 0.010$
Plasma insulin (ng/ml)	0.63 ± 0.21	0.43 ± 0.07	
Ad. lib. Fed			
Plasma glucose (mg/dl)	178 ± 34	111 ± 7.9*	$P = 0.019$
Plasma insulin (ng/ml)	0.83 ± 0.26	0.80 ± 0.26	
ALT (Units/l)	47.5 ± 16.6	93.9 ± 20.0*	$P = 0.010$
AST (Units/l)	32.3 ± 7.9	87.2 ± 14.2*	$P = 0.005$

Data are means ± S.D. ($n = 5$ per group). *$p < 0.05$ vs GFP. Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase.

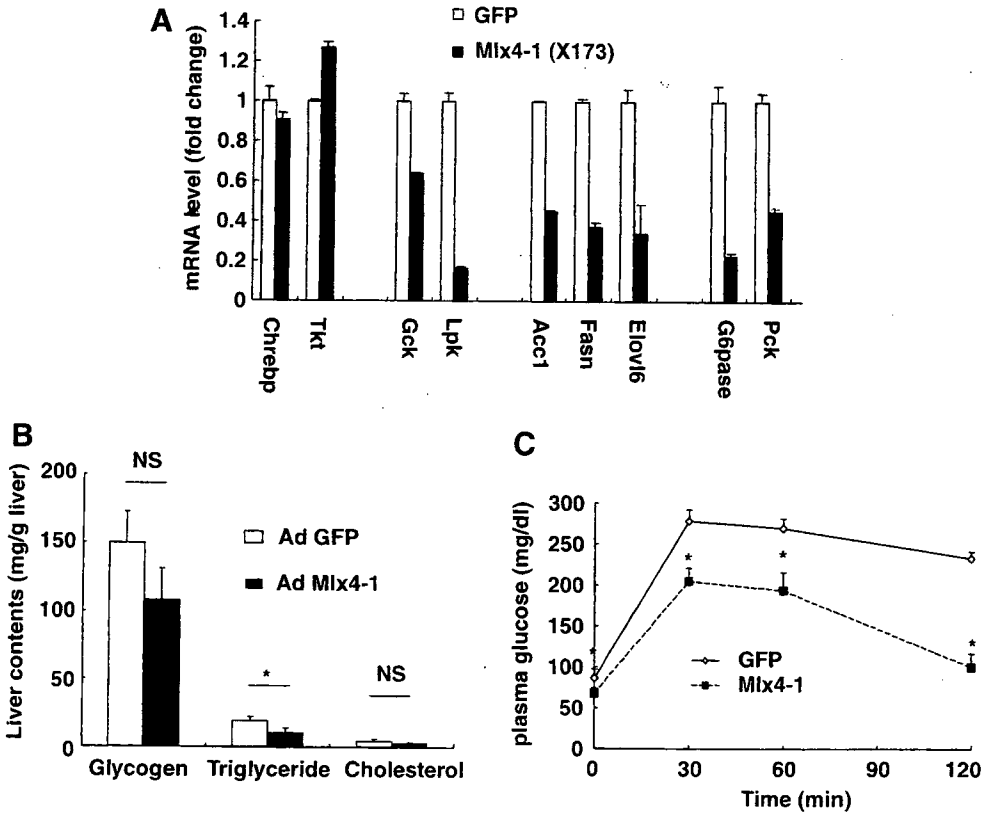

Fig. 3. (A) RT-PCR analysis of glycolytic, lipogenic, gluconeogenic, and lipolytic enzyme mRNA levels in liver of C57BL/6J mice infected with Ad-GFP or Ad-Mlx4-1. Results are means±SD ($n = 5$ per group). $p < 0.05$ vs. Ad-GFP injected C57BL/6J mice. (B) Determination of liver glycogen, triglyceride, and cholesterol contents in Ad-GFP or Ad-Mlx4-1 injected mice. Results are means±SD ($n = 5$ per group) $p < 0.05$ vs. Ad-GFP-injected mice. (C) Glucose tolerance tests (1 g/kg BW) were performed in C57BL/6J mice treated for 3 days with Ad-GFP or Ad-Mlx4-1 adenovirus. Animals were fasted for 18 h before OGTT ($n = 5$ per group). Data are means ± SD ($n = 5$ per group). $p < 0.05$ vs. Ad-GFP-injected C57BL/6J mice.

concentrations were similar to controls (Table 2). Consistent with these observations, oral glucose tolerance test showed that overexpression of Mlx4-1 improved glucose clearance from the body (Fig. 3C). In addition, glycolytic, lipogenic, and gluconeogenic gene mRNA expression levels were measured. As described above, glycolytic genes such as *Lpk* and lipogenic genes such as *Fasn*, *Acc1*, and ELOVL family member 6 (*Elovl6*) were suppressed by overexpression of Ad-Mlx4-1. In the gluconeogenesis pathway, glucose-6-phosphatase (*G-6-Pase*) and phosphoenolpyruvate carboxykinase (*Pepck*) mRNA levels were decreased by 80% and 50%, respectively, (Fig. 3C).

ChREBP induces G-6-Pase gene expression in rat hepatocytes

G-6-Pase gene expression is regulated by glucose signaling [7]. Moreover, Elovl6 is involved in lipid metabolism, and deletion of *Elovl6* improves glucose intolerance in ob/ob mice [8]. As shown in Fig. 3A, *G-6-pase* and *Elovl6* mRNA were down-regulated in the liver of Ad-Mlx4-1-injected mice. We therefore examined whether glucose and ChREBP can induce *G-6-Pase* and *Elovl6* mRNA in rat hepatocytes. Glucose increased hepatic *G-6-Pase* and *Elovl6* mRNA expression in a dose dependent manner (see Supplementary Fig. S4). Moreover, overexpression of dominant active ChREBP increased *G-6-Pase* and *Elovl6* mRNA expression in rat primary hepatocytes (Fig. 4A). Consistent with the findings shown in Fig. 4A, overexpression of dominant negative Mlx4-1 suppressed glucose-induced *G-6-Pase* and *Elovl6* mRNA expression (Fig. 4B). Thus, ChREBP regulates *G-6-Pase* and *Elovl6* gene expression in rat hepatocytes.

Discussion

This study was performed to determine whether hepatic inhibition of ChREBP transactivity can improve glucose intolerance in C57BL/6J mice. Overexpression of dnMlx/Mlx4-1 inhibited glucose-induced *Lpk*, *Fasn*, *Chrebp*, and *Tkt* gene expression in rat hepatocytes. Moreover, adenoviral delivery of dnMlx/Mlx4-1 into the liver improved glucose intolerance. Thus, blocking the association between ChREBP and Mlx is a promising therapeutic strategy to cure metabolic syndrome.

Mlx and ChREBP form the ChREBP-Mlx complex that binds ChoRE in glycolytic and lipogenic gene promoters such as *Lpk* and *Fasn*. Mlx4-1 is a dominant negative form of Mlx without the N-terminal and DNA-binding domains (Fig. 1A). Mlx4-1 is localized in both the cytosol and nucleus (see Supplementary Fig. S1). In contrast, Mlx2-1 is a dominant active form lacking only the N-terminal domain and is localized mainly in the nucleus. Consistent with our data, some groups have reported that the N-terminal domain of ChREBP has an important role in functions of ChREBP such as nuclear translocation and glucose sensitivity [9]. These observations suggest that the N-terminal domain of Mlx plays a role in determining glucose sensitivity and localization.

dnMlx/Mlx4-1 inhibits glucose-mediated *Lpk* and *Fasn* mRNA expression in rat hepatocytes. Towle et al. reported that Mlx plays an important role in glucose regulation of lipogenic enzymes, using their original Mlx mutant lacking only the DNA-binding domain, and obtained results compatible with ours [10]. Unlike their mutant, we constructed a mutant lacking both the N-terminal and the DNA-binding domains. Our observations suggest that the

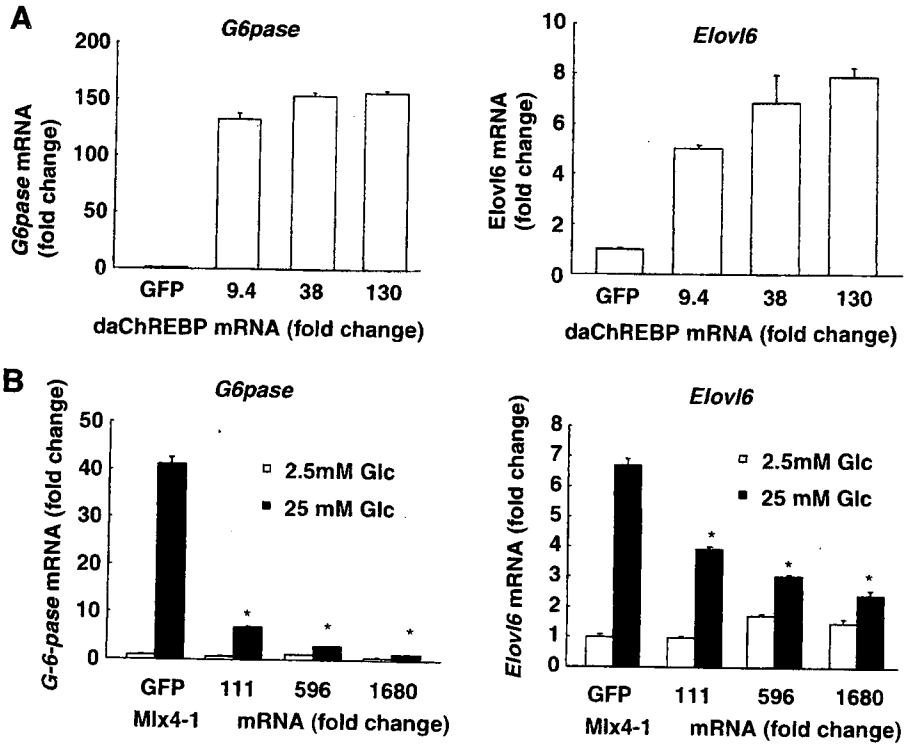

Fig. 4. (A) Adenoviral overexpression of dominant active ChREBP (daChREBP) induced *G-6-Pase* and *Elovl6* mRNA expression in rat hepatocytes. Adenovirus expressing either GFP or daChREBP was transduced into rat hepatocytes at m.o.i. of 2, 10, or 50. As control, a recombinant adenovirus expressing GFP was used at m.o.i. of 50. Two h after infection, cells were kept in DMEM including 2.5 mM glucose for an additional 18 h. Total RNA was then extracted from hepatocytes and RT-PCR analysis was performed. Values represent means±S.D. $^{*}p < 0.05$ *vs.* Ad-GFP. (B) Adenoviral overexpression of Mlx4-1 inhibited glucose-stimulated *G-6-Pase* and *Elovl6* mRNA expression in rat hepatocytes. Adenovirus overexpressing either GFP or Mlx4-1 was transduced into rat hepatocytes at m.o.i. of 2, 10, or 50. As control, a recombinant adenovirus expressing GFP was used at m.o.i. of 50. Two h after infection, cells were kept in DMEM including glucose concentrations of 2.5 or 25 mM for an additional 18 h. Total RNA was then extracted from hepatocytes and RT-PCR analysis was performed. Values represent means±S.D. $^{*}p < 0.05$ *vs.* Ad-GFP.

DNA-binding domain is critical in transactivity of Mlx. Moreover, dnMlx/Mlx4-1 inhibited activation of *Chrebp* and *Tkt* mRNA expression by glucose. In the pentose pathway, glucose is converted to Xu-5-P, an activator of ChREBP [1]. As Xu-5-P is produced by Tkt, ChREBP and Tkt constitutes a positive feedback loop by increasing *Tkt* and *ChREBP* mRNA in rat hepatocytes. It is surprising that glucose increases *Chrebp* mRNA expression in rat hepatocytes and the mechanism remains unknown. Using dnMlx-expressing adenovirus vector, ChREBP was found to directly regulate *Chrebp* gene expression. However, *Tkt* and *Chrebp* mRNA levels were not decreased in the livers of Ad-Mlx4-1-injected mice. In contrast, overexpression of daChREBP caused 40- and 14-fold increases in endogenous *Chrebp* and *Tkt* mRNA levels, respectively, (data not shown). Thus, while glucose regulates *Tkt* and *ChREBP* mRNA expression in rat hepatocytes, other factors may regulate their gene expression more potently *in vivo*.

We and other groups have independently reported that gene deletion or shRNA of *Chrebp* improves metabolic disorders in ob/ob mice [3,4]. However, complete deletion of *Chrebp* causes massive hepatomegaly due to liver glycogen accumulation [3]. Moderate inhibition of ChREBP transactivity thus might improve the symptoms of metabolic syndrome without hepatomegaly. In rat hepatocytes, Mlx4-1 functions only at higher glucose concentrations. In contrast, *Chrebp* gene deletion or *Chrebp* shRNA works at both lower and higher glucose concentrations. Since this is a unique characteristic of dnMlx/Mlx4-1, we tested 25-week-old diabetes-prone C57BL/6J mice. The C57BL/6J mouse is a model of diet-induced obesity and diabetes [6]. As C57BL/6J mice develop glucose intolerance with age similarly to humans, middle-aged C57BL/6J mice were used to test the effects of dnMlx on improvement of glucose intolerance. Overexpression of dnMlx/Mlx4-1 improved glucose intolerance and liver lipid contents without hepatomegaly. These observations suggest that drugs disrupting the association between ChREBP and Mlx may be promising for preventing the development of metabolic syndrome.

dnMlx/Mlx4-1 inhibited expression of genes encoding the glycolytic enzyme *Lpk*, lipogenic enzymes *Acc1*, *Fasn*, and *Elovl6*, and gluconeogenic enzyme *G-6-Pase*, although the precise mechanism by which inhibition of ChREBP transactivity improves metabolic syndrome remains unknown. Elovl6 catalyzes conversion of palmitate to stearate [8]. Consistent with Fig. 4B, dnMlx inhibited glucose-induced *Elovl6* gene expression in primary hepatocytes [11]. Gene deletion of *Elovl6* in high-fat-loaded mice or ob/ob mice improved glucose intolerance by modifying hepatic fatty acid composition, but did not improve hepatic steatosis [8]. We did not analyze hepatic fat composition in Ad-Mlx4-1-injected mice, but fat composition in Chrebp$^{-/-}$ mice was similar to that in normal controls (unpublished data). Moreover, glucose clearance in *Elovl6* knockout mice was similar to that in controls [8]. Thus, *Elovl6* mRNA may contribute to improvement of glucose intolerance, even though its effect in Ad-dnMlx/Mlx4-1 mice is limited.

DnMlx inhibited glucose induction of *G-6-pase* mRNA expression. Acute inhibition of the G-6-Pase system lowers the plasma glucose concentration in rats [12]. These observations suggest that G-6-Pase plays an important role in glucose homeostasis in rodents. Some groups have reported the possibility that ChREBP binds to ChoRE in the rat *G-6-Pase* promoter region [7]. Fig. 4A and B indicates that glucose and ChREBP strongly regulate *G-6-pase* gene expression in rat hepatocytes. As SREBP1c did not regulate *G-6-Pase* mRNA expression, and deletion of SREBP1c did not improve glucose intolerance, the difference in *G-6-pase* gene regulation be-

tween ChREBP and SREBP may underlie the improvement of glucose intolerance by ChREBP but not by SREBP [1].

In conclusion, experiments using dnMlx/Mlx4-1 indicate that ChREBP is regulated by *Chrebp* and *Tkt* gene expression, and forms a positive feedback loop in rat hepatocytes. Moreover, overexpression of dominant negative Mlx improves glucose intolerance in diabetes-prone C57BL/6J mice with metabolic syndrome. These observations suggest that drugs capable of dissociating the ChREBP-Mlx complex may be a promising approach to treatment of metabolic syndrome including glucose intolerance.

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (K. Iizuka), the ONO Medical Research Foundation (K. Iizuka), the Kao Research Council for the Study of Healthcare Science (K. Iizuka), and in part by a New Energy and Industrial Technology Development Organization grant (Y. Horikawa). We apologize to authors whose work was not been cited because of limited space.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2008.12.100.

References

[1] K. Iizuka, Y. Horikawa, ChREBP: a glucose-activated transcription factor involved in the development of metabolic syndrome, Endocr. J. 55 (2008) 617–624.

[2] H. Shimano, Sterol regulatory element binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes, Prog. Lipid. Res. 40 (2001) 439–452.

[3] K. Iizuka, B. Miller, K. Uyeda, Deficiency of carbohydrate-activated transcription factor ChREBP prevents obesity and improves plasma glucose control in leptin-deficient (ob/ob) mice, Am. J. Physiol. Endocrinol. Metab. 291 (2006) E358–364.

[4] R. Dentin, F. Benhamed, I. Hainault, V. Fauveau, F. Foufelle, J.R. Dyck, J. Girard, C. Postic, Liver-specific inhibition of ChREBP improves hepatic steatosis and insulin resistance in ob/ob mice, Diabetes 55 (2006) 2159–2170.

[5] K. Iizuka, Y. Horikawa, Regulation of lipogenesis via BHLHB2/DEC1 and ChREBP feedback looping, Biochem. Biophys. Res. Commun. 374 (2008) 95–100.

[6] A.E. Petro, J. Cotter, D.A. Cooper, J.C. Peters, S.J. Surwit, R.S. Surwit, Fat, carbohydrate, and calories in the development of diabetes and obesity in the C57BL/6J mouse, Metabolism 53 (2004) 454–457.

[7] K.B. Pedersen, P. Zhang, C. Doumen, M. Charbonnet, D. Lu, C.B. Newgard, J.W. Haycock, A.J. Lange, D.K. Scott, The promoter for the gene encoding the catalytic subunit of rat glucose-6-phosphatase contains two distinct glucose-responsive regions, Am. J. Physiol. Endocrinol. Metab. 292 (2007) E788–801.

[8] T. Matsuzaka, H. Shimano, N. Yahagi, T. Kato, A. Atsumi, T. Yamamoto, N. Inoue, M. Ishikawa, S. Okada, N. Ishigaki, H. Iwasaki, Y. Iwasaki, T. Karasawa, S. Kumadaki, T. Matsui, M. Sekiya, K. Ohashi, A.H. Hasty, Y. Nakagawa, A. Takahashi, H. Suzuki, S. Yatoh, H. Sone, H. Toyoshima, J. Osuga, N. Yamada, Crucial role of a long-chain fatty acid elongase, Elovl6, in obesity-induced insulin resistance, Nat. Med. 13 (2007) 1193–1202.

[9] M.V. Li, B. Chang, M. Imamura, N. Poungvarin, L. Chan, Glucose-dependent transcriptional regulation by an evolutionaly conserved glucose-sensing module, Diabetes 55 (2006) 1179–1189.

[10] L. Ma, L.N. Robinson, H.C. Towle, ChREBP·Mlx is the principal mediator of glucose-induced gene expression in the liver, J. Biol. Chem. 281 (2006) 28721–28730.

[11] Y. Wang, D. Botolin, J. Xu, B. Christian, E. Mitchell, B. Jayaprakasam, M.G. Nair, J.M. Peters, J.V. Busik, L.K. Olson, D.B. Jump, Regulation of hepatic fatty acid elongase and desaturase expression in diabetes and obesity, J. Lipid. Res. 47 (2006) 2028–2204.

[12] J.C. Parker, M.A. VanVolkenburg, C.B. Levy, W.H. Martin, S.H. Burk, Y. Kwon, C. Giragossian, T.G. Gant, P.A. Carpino, R.K. McPherson, P. Vestergaard, J.L. Treadway, Plasma glucose levels are reduced in rats and mice treated with an inhibitor of glucose-6-phosphate translocase, Diabetes 47 (1998) 1630–1636

Journal of Human Genetics (2009), 1–6

www.nature.com/jhg

ORIGINAL ARTICLE

Construction of a prediction model for type 2 diabetes mellitus in the Japanese population based on 11 genes with strong evidence of the association

Kazuaki Miyake[1], Woosung Yang[2], Kazuo Hara[3], Kazuki Yasuda[4], Yukio Horikawa[5], Haruhiko Osawa[6], Hiroto Furuta[7], Maggie CY Ng[8], Yushi Hirota[1], Hiroyuki Mori[1], Keisuke Ido[2,20], Kazuya Yamagata[9,21], Yoshinori Hinokio[10], Yoshitomo Oka[10], Naoko Iwasaki[11], Yasuhiko Iwamoto[11], Yuichiro Yamada[12,22], Yutaka Seino[12,23], Hiroshi Maegawa[13], Atsunori Kashiwagi[13], He-yao Wang[4,24], Toshihito Tanahashi[14], Naoto Nakamura[15], Jun Takeda[5], Eiichi Maeda[2], Ken Yamamoto[16], Katsushi Tokunaga[17], Ronald CW Ma[8], Wing-Yee So[8], Juliana CN Chan[8], Naoyuki Kamatani[18], Hideichi Makino[6], Kishio Nanjo[7], Takashi Kadowaki[3] and Masato Kasuga[1,19]

Prediction of the disease status is one of the most important objectives of genetic studies. To select the genes with strong evidence of the association with type 2 diabetes mellitus, we validated the associations of the seven candidate loci extracted in our earlier study by genotyping the samples in two independent sample panels. However, except for *KCNQ1*, the association of none of the remaining seven loci was replicated. We then selected 11 genes, *KCNQ1*, *TCF7L2*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *SLC30A8*, *HHEX*, *GCKR*, *HNF1B*, *KCNJ11* and *PPARG*, whose associations with diabetes have already been reported and replicated either in the literature or in this study in the Japanese population. As no evidence of the gene–gene interaction for any pair of the 11 loci was shown, we constructed a prediction model for the disease using the logistic regression analysis by incorporating the number of the risk alleles for the 11 genes, as well as age, sex and body mass index as independent variables. Cumulative risk assessment showed that the addition of one risk allele resulted in an average increase in the odds for the disease of 1.29 (95% CI=1.25–1.33, $P=5.4\times10^{-53}$). The area under the receiver operating characteristic curve, an estimate of the power of the prediction model, was 0.72, thereby indicating that our prediction model for type 2 diabetes may not be so useful but has some value. Incorporation of data from additional risk loci is most likely to increase the predictive power.
Journal of Human Genetics advance online publication, 27 February 2009; doi:10.1038/jhg.2009.17

Keywords: gene–gene interaction; genome-wide association study; prediction model; single nucleotide polymorphism (SNP); type 2 diabetes mellitus

[1]Division of Diabetes, Metabolism and Endocrinology, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kobe, Japan; [2]Clinical Genome Informatics Center, Kobe University Graduate School of Medicine, Kobe, Japan; [3]Department of Metabolic Diseases, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [4]Department of Metabolic Disorder, Research Institute, International Medical Center of Japan, Tokyo, Japan; [5]Division of Molecule and Structure, Department of Diabetes and Endocrinology, Gifu University School of Medicine, Gifu, Japan; [6]Department of Molecular and Genetic Medicine, Ehime University Graduate School of Medicine, Ehime, Japan; [7]First Department of Medicine, Wakayama Medical University, Wakayama, Japan; [8]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Shatin, Hong Kong; [9]Department of Metabolic Medicine, Graduate School of Medicine, Osaka University, Osaka, Japan; [10]Division of Molecular Metabolism and Diabetes, Tohoku University Graduate School of Medicine, Sendai, Japan; [11]Department of Medicine, Diabetes Center, Tokyo Women's Medical University, Tokyo, Japan; [12]Department of Diabetes and Clinical Nutrition, Kyoto University School of Medicine, Kyoto, Japan; [13]Division of Endocrinology and Metabolism, Department of Medicine, Shiga University of Medical Science, Shiga, Japan; [14]Division of Genetic Information, Institute for Genome Research, University of Tokushima, Tokushima, Japan; [15]Department of Endocrinology and Metabolism, Kyoto Prefectural University of Medicine, Graduate School of Medical Sciences, Kyoto, Japan; [16]Department of Molecular Genetics, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan; [17]Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [18]Division of Genomic Medicine, Department of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo, Japan and [19]Research Institute, International Medical Center of Japan, Tokyo, Japan
Correspondence: Dr M Kasuga, Research Institute, International Medical Center of Japan, 1-21-1 Toyama, Shinjuku-ku, Tokyo 162-8655, Japan.
E-mail: kasuga@ri.imcj.go.jp
[20]Current address: Information Center for Medical Sciences, Tokyo Medical and Dental University, Tokyo, Japan.
[21]Current address: Faculty of Medical and Pharmaceutical Sciences, Department of Medical Biochemistry, Kumamoto University, Kumamoto, Japan.
[22]Current address: Department of Internal Medicine, Akita University School of Medicine, Akita, Japan.
[23]Current address: Kansai Electric Power Hospital, Osaka, Japan.
[24]Current address: Shanghai Institute of Materia Medica, Chinese Academy of Science, Shanghai, China.
Received 22 December 2008; revised 25 January 2009; accepted 5 February 2009

INTRODUCTION

Genome-wide association studies (GWASs) have identified novel susceptibility genes for type 2 diabetes mellitus in Caucasians.[1–5] *TCF7L2*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *SLC30A8* and *HHEX* have been widely replicated as susceptibility genes for type 2 diabetes in Asian populations[6–12] as well as in populations of European ancestry.[13,14] We recently identified *KCNQ1* as a novel susceptibility gene, as well as seven other candidate susceptibility loci in a multistage GWAS for type 2 diabetes in the Japanese population, in which a total of 1612 cases and 1424 controls and 100 000 single nucleotide polymorphisms (SNPs) were included.[15] *KCNQ1* was found to confer risk of type 2 diabetes with a relatively large effect size in Asian populations (odds ratio (OR) for Japanese, Chinese and Korean individuals of 1.42),[15] which was similar to that demonstrated earlier for *TCF7L2* in the Japanese population.[6]

Follow-up of GWASs includes analysis of second-tier genes, meta-analysis for specific populations, as well as analysis of gene–gene or gene–environment interactions. A large-scale meta-analysis[16] and an analysis of gene–gene interaction for susceptibility genes[17] have been performed for type 2 diabetes in populations of European ancestry.

In this study, we attempted to confirm in independent subject panels of Japanese and Hong Kong Chinese individuals the associations of the seven candidate susceptibility loci that we identified in addition to *KCNQ1* in our GWAS of type 2 diabetes.[15] However, as described in this article, we failed to replicate the associations of the seven loci with diabetes. We then attempted to extract genes with strong evidence of the associations with diabetes, and selected 11 genes, including *KCNQ1*. As we did not detect any gene–gene interaction between the 11 genes, we then attempted to construct a prediction model for this disease by using the data from the 11 genes, as well as age, gender and body mass index (BMI) as independent variables to obtain a comprehensive understanding of the genetic background of diabetes in the Japanese population.

MATERIALS AND METHODS

Validation of the results from a multistage GWAS in the Japanese population

Study subjects. We assembled two independent subject panels for our replication study: replication-Japanese and replication-Chinese. The 1000 cases and 1000 controls for the replication-Japanese panel were recruited by the Study Group of the Millennium Genome Project for Diabetes Mellitus. The inclusion criteria for diabetic patients were (i) an age at disease onset of 30–60 years and (ii) the absence of antibodies to GAD. Types of diabetes other than type 2 were excluded on the basis of clinical data. The criteria for controls included (i) an age of >50 years, (ii) no past history of a diagnosis of diabetes and (iii) an HbA_{1c} content of $<5.8\%$.

For the replication-Chinese panel, subjects of southern Han Chinese ancestry, who resided in Hong Kong, were recruited. The cases consisted of 1416 individuals with type 2 diabetes selected from the Prince of Wales Hospital Diabetes Registry;[5,18] 626 of these subjects had early-onset diabetes (age at diagnosis of <40 years) and a positive family history, whereas the remaining 790 patients were randomly selected from the registry. Patients with classic type 1 diabetes with acute ketotic presentation or a continuous requirement for insulin within 1 year of diagnosis were excluded. The controls consisted of 1577 subjects with normal glucose tolerance (fasting plasma glucose concentration of $<6.1\,mmol\,l^{-1}$); 596 of these individuals were recruited either from the general population participating in a community-based screening program for cardiovascular risk or from hospital staff, whereas the remaining 981 subjects were recruited from a population-based screening program for cardiovascular risk in adolescents.[19] The clinical characteristics of the subjects in each panel are summarized in Supplementary Table 1A. The study protocol was approved by the local ethics committee of each institution. Written informed consent was obtained from each subject.

Study design and statistical analysis. For the validation of the results from our earlier multistage GWAS,[15] seven SNPs (rs2250402, rs2307027, rs3741872, rs574628, rs2233647, rs3785233 and rs2075931) were genotyped in the two panels either by sequence-specific primer–PCR analysis followed by fluorescence correlation spectroscopy[20] or by real-time PCR analysis with TaqMan probes (Applied Biosystems, Foster City, CA, USA). Differences in allele frequency between cases and controls for each SNP were evaluated by χ^2 with one degree of freedom. Meta-analysis was performed by the Mantel–Haenszel method (fixed-effects models) with the 'meta' package of the R-Project (http://www.r-project.org). A P-value of <0.05 was considered statistically significant.

Examination of gene–gene interaction and construction of a prediction model

Study subjects. In total, 2424 cases and 2424 controls of the Japanese population obtained by combining the second and third screening panels in our original study[15] and the replication-Japanese panel of this study were included in this analysis (analysis-panel). The criteria for the second and third screening panels were described in the earlier report.[15] The clinical characteristics of the subjects are summarized in Supplementary Table 1B.

Selection of the loci included in this study. Prediction of the phenotypes on the basis of genetic polymorphisms should include the genetic data from the loci with strong evidence of the association. Starting from 15 genes described in earlier reports, we selected 11 genes with strong evidence of the association on the basis of the data in the literature and on the results of the replication experiments in this study. Process of the selection of the 11 genes will be described in detail in Results.

Statistical methods. Multiplicative gene–gene interaction was evaluated for each pair of the 11 genes using an interaction term in addition to the terms for the pair of the genes in the logistic regression model. The genotypes for each locus were coded by 0, 1 and 2. Correction for multiple testing was performed by Bonferroni's method.

As there was no evidence for the presence of gene–gene interactions, we attempted to construct a phenotype prediction model by incorporating the number of risk alleles for the 11 loci as an independent variable in addition to age, gender and BMI. The Cochran–Armitage test was used to examine the trend of the increase in the odds by increasing the number of the risk alleles. To construct a prediction model, the log of odds was expressed by the linear combination of the independent variables. Coefficients for the variables were estimated by the logistic regression analysis after making disease (cases) or nondisease (controls) as the dependent variable. Using the coefficients estimated by the logistic regression analysis, we constructed a phenotype prediction model. To evaluate the prediction model, receiver operating characteristic (ROC) curves[21] for the sensitivity and specificity of the prediction model with or without adjustment for age, sex and BMI were generated, and the area under the curve (AUC) was calculated from the ROC curve.

RESULTS

Validation of the results from a multistage GWAS in the Japanese population

We identified earlier 10 loci associated with type 2 diabetes by three-staged GWAS starting from 100 000 SNPs. Among the 10 loci, 3 SNPs were located in an intron of *KCNQ1*, and the association of this gene with diabetes was confirmatory.[15] To validate the other seven loci for the association with type 2 diabetes, we analyzed them in two independent replication panels of Japanese and Han-Chinese individuals (Table 1, Supplementary Table 2). Only one SNP, rs2250402, which is located in *EIF2AK4*, was found to be significantly associated in the replication-Japanese panel (P=0.039, OR=1.17, 95% CI=1.01–1.36). However, neither this SNP (P=0.41, OR=1.05) nor any of the other six SNPs showed such an association in the replication-Chinese panel. Meta-analyses for these SNPs showed that rs2307027 in *KRT4* and rs3785233 in *A2BP1* yielded P-values of <0.05 and ORs between 1.12 and 1.13 (Table 1). When the original second and third screening

Table 1 Association study for the candidate susceptibility genes for type 2 diabetes selected by multistage screening in the Japanese population

SNP ID	*Chr*	*Gene*	*Risk allele*	*Panel*	*RAF (DM)*	*RAF (NC)*	*P*	*OR*	*95% CI*
rs2250402	15	*EIF2AK4*	C	Replication-Japanese	0.23	0.20	0.04	1.17	1.01–1.36
				Replication-Chinese	0.24	0.23	0.41	1.05	0.93–1.19
				Meta-analysis			0.05	1.10	1.00–1.20
rs2307027	12	*KRT4*	C	Replication-Japanese	0.18	0.17	0.17	1.12	0.95–1.32
				Replication-Chinese	0.14	0.13	0.16	1.11	0.96–1.29
				Meta-analysis			0.05	1.12	1.00–1.25
rs3741872	12	*FAM60A*	C	Replication-Japanese	0.25	0.24	0.18	1.11	0.96–1.28
				Replication-Chinese	0.23	0.22	0.21	1.08	0.96–1.22
				Meta-analysis			0.07	1.09	0.99–1.20
rs574628	20	*ANGPT4*	G	Replication-Japanese	0.60	0.61	0.46	0.95	0.84–1.08
				Replication-Chinese	0.65	0.65	0.59	1.03	0.93–1.15
				Meta-analysis			0.96	1.00	0.92–1.08
rs2233647	6	*SPDEF*	G	Replication-Japanese	0.86	0.87	0.70	0.97	0.81–1.16
				Replication-Chinese	0.94	0.93	0.54	1.07	0.87–1.31
				Meta-analysis			0.90	1.01	0.88–1.16
rs3785233	16	*A2BP1*	C	Replication-Japanese	0.18	0.16	0.19	1.12	0.95–1.32
				Replication-Chinese	0.13	0.12	0.10	1.14	0.97–1.34
				Meta-analysis			0.04	1.13	1.01–1.27
rs2075931	1	Intergenic	A	Replication-Japanese	0.67	0.66	0.85	1.01	0.89–1.16
				Replication-Chinese	0.73	0.74	0.27	0.94	0.84–1.05
				Meta-analysis			0.48	0.97	0.89–1.06

Abbreviations: Chr, chromosome; OR, odds ratio for risk allele frequency.
Assignment of risk alleles was based on the original study.[15] Numbers of cases versus control subjects in the replication-Japanese and replication-Chinese panels were 1000 versus 1000 and 1416 versus 1577, respectively. RAF (DM) and RAF (NC) denote risk allele frequencies in cases and controls, respectively. *P* values were calculated for allele frequency. Meta-analysis was performed by the Mantel–Haenszel method (fixed-effects models). *P*-values for the test of heterogeneity among panels joined in the Mantel–Haenszel tests were all >0.05.

panels were included in the meta-analyses, these two loci, as well as the SNPs in *EIF2AK4* (rs2250402) and *FAM60A* (rs3741872), gave *P*-values of <0.001 and ORs between 1.15 and 1.18 (Supplementary Table 3). However, the *P*-values did not reach the proposed significance of GWAS ($=5\times10^{-7}$).

Selection of polymorphisms for the prediction model

To construct a reliable prediction model for diabetes, polymorphisms with strong evidence of association should be used. From the previous literature, we selected 15 genes (including one intergenic marker), that is, *SLC30A8*, *HHEX*, *LOC387761*, *EXT2*, *CDKN2A/B*, *GCKR*, *IGF2BP2*, *CDKAL1*, *FTO*,[1–5] *TCF7L2*,[22] *KCNJ11*,[23] *PPARG*,[24] *WFS1*,[25] *HNF1B*[26] and *KCNQ1*,[15] as candidate genes to be included in both gene–gene interaction analysis and construction of a prediction model. Starting from 23 SNPs in these 15 genes, we selected 11 SNPs in 11 genes according to the following process. There is sufficient evidence of the associations of *KCNQ1* and *TCF7L2* genes with diabetes as supported by replication studies in the Japanese population.[6,15,27] In addition, *SLC30A8*, *HHEX*, *CDKN2A/B*, *IGF2BP2* and *CDKAL1* associated with the disease in the European population were found in our earlier study to be associated with the disease in the Japanese population as well.[7–9]

To further extract genes with strong evidence of the association with diabetes, we attempted to replicate the associations reported earlier using our own data (analysis panel with 2424 cases and 2424 controls). For the 19 SNPs in *SLC30A8*, *HHEX*, *LOC387761*, *EXT2*, *CDKN2A/B*, *GCKR*, *IGF2BP2*, *CDKAL1*, *FTO*, *TCF7L2*, *KCNJ11*, *PPARG* and *KCNQ1*, we extracted genotyping data from our earlier studies[6–9,15,27–29] and, if necessary, genotyped additional subjects to obtain a data set for 2424 cases and 2424 controls of the Japanese population (analysis panel). The SNPs in *WFS1* (rs6446482, rs734312) and *HNF1B* (rs7501939, rs4430796) were genotyped for this study in the same individuals. SNPs with *P*-values for the test of deviation from the Hardy–Weinberg equilibrium of <0.01 were excluded for further analysis. When two SNPs were located in the same genomic region, the one with the lower *P*-value for the association test was selected for further analysis. *GCKR*, for which we earlier reported the marginal association with type 2 diabetes,[7] was found to be associated with the disease in this enlarged Japanese panel ($P=1.7\times10^{-5}$; Supplementary Table 4). *KCNJ11* and *PPARG*, which have been included in the genes associated with diabetes in Caucasians, showed marginal associations ($P=0.066$ and $P=0.075$, respectively; Supplementary Table 4) in our panel. Two SNPs in *WFS1* and two SNPs in *HNF1B* were newly genotyped in the analysis panel. Although no association was apparent between *WFS1* and type 2 diabetes, both SNPs in *HNF1B* exhibited *P*-values of <0.05 (Supplementary Table 4). From these data, we included 11 SNPs in 11 genes as described above for the source of genotype data to be analyzed in both the examination of gene–gene interaction and the prediction of phenotypes.

Gene–gene interaction

We evaluated multiplicative gene–gene interaction for each pair of the 11 loci as described in Materials and methods. Two combinations, rs1801282 (*PPARG*) ×rs1470579 (*IGF2BP2*) (nominal $P=0.0025$) and rs1801282×rs3802177 (*SLC30A8*) (nominal $P=0.018$), showed *P*-values of less than 0.05 (Supplementary Figure 1). However, these *P*-values were not significant when Bonferroni's correction for multiple testing was applied (significance level, $0.05/55=9.1\times10^{-4}$). Although *PPARG* and *IGF2BP2* are located on the same chromosome (3p25 and 3q28, respectively), it is unlikely that loci on different arms of the same chromosome show significant linkage disequilibrium. *SLC30A8* is located on a different chromosome (8q24.11) from

PPARG. The reason why nominal *P*-values of these combinations showed less than 0.05 may be because of the low minor allele frequency of rs1801282.

Cumulative risk assessment for type 2 diabetes on the basis of susceptibility genes

As there was no evidence of gene–gene interaction between 11 SNPs of 11 genes, *SLC30A8*, *HHEX*, *CDKN2A/B*, *GCKR*, *IGF2BP2*, *CDKAL1*, *TCF7L2*, *KCNJ11*, *PPARG*, *KCNQ1* and *HNF1B*, they were included in the prediction model as independent variables with the additive effect (additive effect in the liability and multiplicative effect in the odds) without interaction terms. Effective numbers of cases and controls whose genotypes for the 11 loci were successfully obtained were 2316 and 2370, respectively. The Cochran–Armitage trend test gave a *P*-value of 4.7×10^{-56} for the trend in the increase in the odds for cases relative to controls with an increasing number of risk alleles for the 11 susceptibility loci (Supplementary Table 5). We then estimated ORs for type 2 diabetes in subjects with different numbers of risk alleles on the basis of the multiplicative model by logistic regression analysis with adjustment for age, sex and BMI. The ORs for type 2 diabetes in subjects with 7–18 risk alleles in comparison with those harboring 0–6 risk alleles are shown in Figure 1. An increase of one risk allele resulted in an average increase in the odds of 1.29 (95% CI=1.25–1.33, $P=5.4\times10^{-53}$, logistic regression analysis).

To predict disease status for type 2 diabetes in a given individual, we constructed a prediction model on the basis of the number of risk alleles or the liability value calculated from the number of risk alleles as well as age, sex and BMI. The coefficients to calculate the liability value were estimated with the logistic regression model. To estimate the predictive power of the model, we generated ROC curves as described in Materials and methods. The AUC was 0.63 when only the number of risk alleles was used for the prediction. When age, sex and BMI were also included, the AUC increased to 0.72 (Figure 2). Meanwhile, an AUC value for the ROC curve based on only age, sex and BMI was 0.68, which was better than that based on only the number of risk alleles (data now shown). The model incorporating age, sex and BMI as well as the number of risk alleles thus showed moderate power for the prediction of type 2 diabetes. The best accuracy was 0.66 at the threshold between non-diabetic and diabetic status of 0.52 (non-diabetic status=0, diabetic status=1), for which the specificity and the sensitivity were 0.71 and 0.61, respectively.

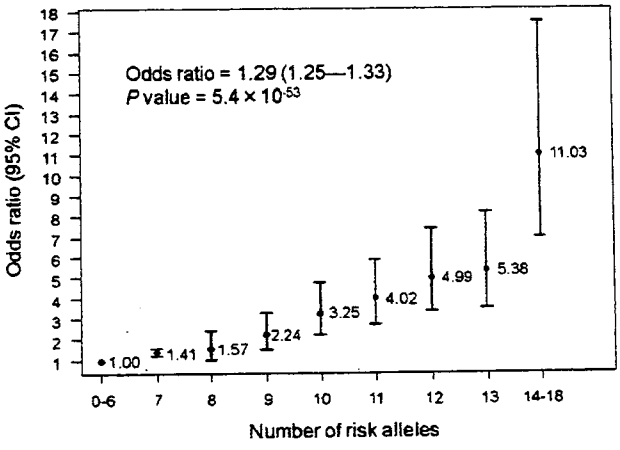

Figure 1 Odds ratios for subjects with different numbers of risk alleles for 11 susceptibility loci for type 2 diabetes. The cumulative effect of the 11 loci on type 2 diabetes was tested by counting the number of risk alleles associated with type 2 diabetes with a logistic regression model with adjustment for age, sex and BMI. The ORs for subjects with each number of risk alleles are expressed relative to individuals with 0–6 risk alleles.

DISCUSSION

By the validation of the results from our multistage GWAS, we detected only marginal associations of *EIF2AK4*, *KRT4* and *A2BP1* with type 2 diabetes in meta-analyses with two subject panels of Japanese or Chinese individuals. Relations of *KRT4* (keratin 4 gene) and *A2BP1* (ataxin-2-binding protein 1 gene, also known as *FOX1*) to glucose or lipid metabolism are unknown. Deletion of *EIF2AK4* (eukaryotic translation initiation factor 2 alpha kinase 4 gene, also known as *GCN2*) in mice resulted in liver steatosis during leucine deprivation as a result of unrepressed expression of lipogenic genes.[30] The functionally related gene, *EIF2AK3* (also known as *PERK* or *PEK*), has been shown to cause diabetes mellitus both in humans (Wolcott–Rallison syndrome, OMIM604032) and in rodent models.[31,32] Taken together, *EIF2AK4* may be a good candidate for the diabetes susceptibility gene. The sample size required for a statistical power of 0.80 with equal numbers of cases and controls is 10 505 when the frequency of the risk allele, OR and type I error probability are assumed to be 0.20, 1.10 (the value for *EIF2AK4* in the meta-analysis in Table 1) and 0.05, respectively. Further studies of these genes in other Asian populations as well as in other ethnic groups are needed for confirmation of their association with type 2 diabetes. Given this uncertainty, we did not include these genes in the assessments of cumulative risk and gene–gene interaction.

Among tens of type 2 diabetes susceptibility genes identified by recent GWASs in Caucasians, the associations of six genes, that is, *TCF7L2*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *SLC30A8* and *HHEX*, have been replicated in Asian populations as well as in populations of European ancestry. A recent meta-analysis in Japanese subjects also supported the associations.[12] In this study, we performed replication

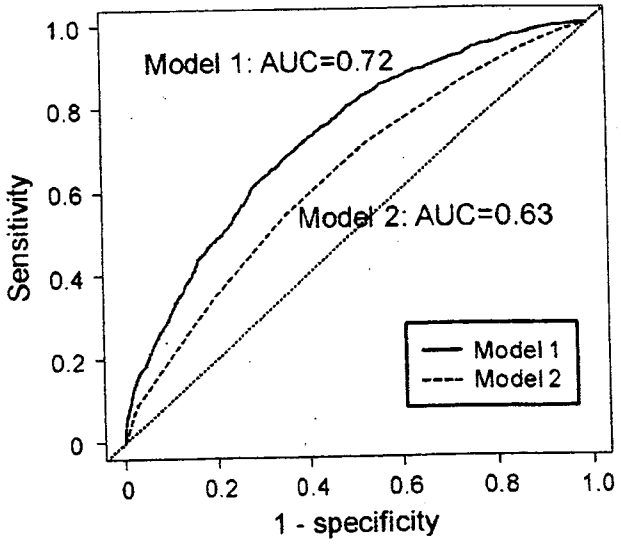

Figure 2 ROC curves for the prediction model on the basis of the number of risk alleles for 11 susceptibility loci for type 2 diabetes. The prediction model for type 2 diabetes was constructed using the logistic regression model, and ROC curves for the model were generated. In model 1, the number of risk alleles was used as an independent variable together with age, sex and BMI as covariates, whereas only the number of risk alleles was used as an independent variable in model 2.

study, and, on the basis of the results, we added five more genes, that is, *KCNJ11*, *PPARG*, *GCKR*, *KCNQ1* and *HNF1B*, for the cumulative risk assessment for type 2 diabetes. Thus, the SNPs of *HNF1B*, which were earlier associated with type 2 diabetes in Chinese as well as in Caucasians,[26] showed the association with the disease in the Japanese population in this study. In addition, the C allele of rs780094 in *GCKR* was associated with increased risk of type 2 diabetes in this study, which is consistent with a recent study in Caucasians.[33] The associations of *KCNJ11* and *PPARG* with diabetes were marginal in this study; however, they were included for the prediction model, as the associations were replicated in some studies of Caucasians.

Our gene–gene interaction analysis showed no significant interaction for any of the 55 possible pairs of genes when corrected for multiple testing. When the significance level was set at 0.05, two pairs were judged to be significant. However, such gene–gene interactions were not supported from the functional point of view. A large-scale study may provide more convincing evidence for such interactions.

As no confirmatory evidence for gene–gene interaction was observed, we treated the 11 genes as independent variables in the prediction model. The addition of one risk allele was estimated to increase the odds by an average of 1.29 according to the multiplicative model. This value is similar to that (1.24) estimated for type 2 diabetes in Caucasians.[17] Two earlier cumulative risk assessments for type 2 diabetes in Asian populations with relatively small numbers of associated loci yielded values of 1.17 and 1.24 for the fold increase in risk for each additional risk allele.[11,34] In our prediction model for type 2 diabetes, the AUC for the ROC curve was lower than that in the earlier study[17] based on 15 loci in Caucasians (0.72 and 0.86, respectively). However, the number of loci in our study (11 loci) was lower than that in the study for Caucasians. The inclusion of additional loci in our model should improve its ability to predict type 2 diabetes in Asian populations. Several reports of the prediction of type 2 diabetes using ~18 loci were recently described for populations of European ancestry.[35–38] A prediction based on 18 loci gave an AUC value of 0.80 for the ROC curve,[35] whereas the corresponding values for a population-based prospective study were 0.68,[36] 0.615[37] and 0.75.[38] They concluded that genetic variations associated with diabetes had a small effect on the ability to predict the development of type 2 diabetes as compared with clinical characteristics alone. In fact, the AUC value (0.72) based on both the genetic variations and the clinical characteristics was slightly better than that based on only the clinical characteristics (0.68). We admit that the evidence of the association with diabetes is a little weaker for *KCNJ11* and *PPARG* in the Japanese population than for the other nine genes. If *KCNJ11* and *PPARG* were excluded from the analysis, the AUC for the ROC curve in the prediction model incorporating age, sex and BMI remained unchanged at 0.72, probably because of the relatively large effects of *KCNQ1* and *TCF7L2*.

Finally, our prediction model for type 2 diabetes achieved limited success even though it has some value. Given that GWASs for diabetes in Asians have not been as extensive as those in Caucasians, many risk loci for diabetes in Asians remain most likely to be undiscovered. Considering that the average increase in OR conferred by each additional risk allele was similar between Caucasians and Japanese, incorporation of data from additional risk loci is most likely to increase the predictive power.

ACKNOWLEDGEMENTS

We thank all the subjects who joined this project; Sumio Sugano and Shoji Tsuji for support and helpful discussion throughout the project; and Megumi Yamaoka-Sageshima for technical assistance. This work was supported by a grant from the Program for Promotion of Fundamental Studies in Health Sciences of the Pharmaceuticals and Medical Devices Agency (PMDA) of Japan; a grant from the National Institute of Biomedical Innovation (NIBIO) of Japan; grants from the Ministry of Health, Labour and Welfare of Japan; a Grant-in-Aid for Scientific Research on Priority Areas (C), 'Medical Genome Science (Millennium Genome Project)', 'Applied Genomics', and 'Comprehensive Genomics', from the Ministry of Education, Culture, Sports, Science and Technology of Japan; and a grant from the Cooperative Link of Unique Science and Technology for Economy Revitalization (CLUSTER, Tokushima, Japan). The Hong Kong diabetes case–control study was supported by the Hong Kong Research Grants Committee Central Allocation Scheme CUHK 1/04C.

1 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).
2 Saxena, R., Voight, B. F., Lyssenko, V., Burtt, N. P., de Bakker, P. I., Chen, H. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
3 Zeggini, E., Weedon, M. N., Lindgren, C. M., Frayling, T. M., Elliott, K. S., Lango, H. *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336–1341 (2007).
4 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).
5 Steinthorsdottir, V., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Jonsdottir, T., Walters, G. B. *et al.* A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat. Genet.* **39**, 770–775 (2007).
6 Miyake, K., Horikawa, Y., Hara, K., Yasuda, K., Osawa, H., Furuta, H. *et al.* Association of TCF7L2 polymorphisms with susceptibility to type 2 diabetes in 4087 Japanese subjects. *J. Hum. Genet.* **53**, 174–180 (2008).
7 Horikawa, Y., Miyake, K., Yasuda, K., Enya, M., Hirota, Y., Yamagata, K. *et al.* Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. *J. Clin. Endocrinol. Metab.* **93**, 3136–3141 (2008).
8 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. *Diabe* **50**, 2461–2466 (2007).
9 Furukawa, Y., Shimada, T., Furuta, H., Matsuno, S., Kusuyama, A., Doi, A. *et al.* Polymorphisms in the IDE-KIF11-HHEX gene locus are reproducibly associated with type 2 diabetes in a Japanese population. *J. Clin. Endocrinol. Metab.* **93**, 310–314 (2008).
10 Omori, 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).
11 Ng, M. C., Park, K. S., Oh, B., Tam, C. H., Cho, Y. M., Shin, H. D. *et al.* Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2 and FTO in type 2 diabetes and obesity in 6719 Asians. *Diabetes* **57**, 2226–2233 (2008).
12 Tabara, Y., Osawa, H., Kawamoto, R., Onuma, H., Shimizu, I., Miki, T. *et al.* Replication study of candidate genes associated with type 2 diabetes based on genome-wide screening. *Diabetes* **58**, 493–498 (2009)10.2337/db07-1785..
13 Grarup, N., Rose, C. S., Andersson, E. A., Andersen, G., Nielsen, A. L., Albrechtsen, A. *et al.* Studies of association of variants near the HHEX, CDKN2A/B, and IGF2BP2 genes with type 2 diabetes and impaired insulin release in 10705 Danish subjects: validation and extension of genome-wide association studies. *Diabetes* **56**, 3105–3111 (2007).
14 Cauchi, S., Proença, C., Choquet, H., Gaget, S., De Graeve, F., Marre, M. *et al.* 2008 Analysis of novel risk loci for type 2 diabetes in a general French population: the D.E.S.I.R. study. *J. Mol. Med.* **86**, 341–348 (2008).
15 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).
16 Zeggini, E., Scott, L. J., Saxena, R., Voight, B. F., Marchini, J. L., Hu, T. *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638–645 (2008).
17 Cauchi, S., Meyre, D., Durand, E., Proença, C., Marre, M., Hadjadj, S. *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* **3**, e2031 (2008).
18 Yang, X., So, W. Y., Kong, A. P., Ho, C. S., Lam, C. W., Stevens, R. J. *et al.* Development and validation of stroke risk equation for Hong Kong Chinese patients with type 2 diabetes: the Hong Kong Diabetes Registry. *Diabetes Care* **30**, 65–70 (2007).
19 Ozaki, R., Qiao, Q., Wong, G. W., Chan, M. H., So, W. Y., Tong, P. C. *et al.* Overweight, family history of diabetes and attending schools of lower academic grading are independent predictors for metabolic syndrome in Hong Kong Chinese adolescents. *Arch. Dis. Child.* **92**, 224–228 (2007).
20 Bannai, M., Higuchi, K., Akesaka, T., Furukawa, M., Yamaoka, M., Sato, K. *et al.* Single-nucleotide-polymorphism genotyping for whole-genome-amplified samples

using automated fluorescence correlation spectroscopy. *Anal. Biochem.* **327**, 215–221 (2004).
21 Sing, T., Sander, O., Beerenwinkel, N. & Lengauer, T. ROCR: visualizing classifier performance in R. *Bioinformatics* **21**, 3940–3941 (2005).
22 Grant, S. F., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Manolescu, A., Sainz, J. *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320–323 (2006).
23 Gloyn, A. L., Weedon, M. N., Owen, K. R., Turner, M. J., Knight, B. A., Hitman, G. *et al.* Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes* **52**, 568–572 (2003).
24 Altshuler, D., Hirschhorn, J. N., Klannemark, M., Lindgren, C. M., Vohl, M. C., Nemesh, J. *et al.* The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat. Genet.* **26**, 76–80 (2000).
25 Sandhu, M. S., Weedon, M. N., Fawcett, K. A., Wasson, J., Debenham, S. L., Daly, A. *et al.* Common variants in WFS1 confer risk of type 2 diabetes. *Nat. Genet.* **39**, 951–953 (2007).
26 Gudmundsson, J., Sulem, P., Steinthorsdottir, V., Bergthorsson, J. T., Thorleifsson, G., Manolescu, A. *et al.* Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat. Genet.* **39**, 977–983 (2007).
27 Horikoshi, M., Hara, K., Ito, C., Nagai, R., Froguel, P. & Kadowaki, T. A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. *Diabetologia* **50**, 747–751 (2007).
28 Yokoi, N., Kanamori, M., Horikawa, Y., Takeda, J., Sanke, T., Furuta, H. *et al.* Association studies of variants in the genes involved in pancreatic beta-cell function in type 2 diabetes in Japanese subjects. *Diabetes* **55**, 2379–2386 (2006).
29 Mori, H., Ikegami, H., Kawaguchi, Y., Seino, S., Yokoi, N., Takeda, J. *et al.* The Pro12→Ala substitution in PPAR-gamma is associated with resistance to development of diabetes in the general population: possible involvement in impairment of insulin secretion in individuals with type 2 diabetes. *Diabetes* **50**, 891–894 (2001).
30 Guo, F. & Cavener, D. R. The GCN2 eIF2alpha kinase regulates fatty-acid homeostasis in the liver during deprivation of an essential amino acid. *Cell Metab.* **5**, 103–114 (2007).
31 Harding, H. P., Zeng, H., Zhang, Y., Jungries, R., Chung, P., Plesken, H. *et al.* Diabetes mellitus and exocrine pancreatic dysfunction in Perk$^{-/-}$ mice reveals a role for translational control in secretory cell survival. *Mol. Cell.* **7**, 1153–1163 (2001).
32 Delépine, M., Nicolino, M., Barrett, T., Golamaully, M., Lathrop, G. M. & Julier, C. EIF2AK3, encoding translation initiation factor 2-alpha kinase 3, is mutated in patients with Wolcott–Rallison syndrome. *Nat. Genet.* **25**, 406–409 (2000).
33 Sparsø, T., Andersen, G., Nielsen, T., Burgdorf, K. S., Gjesing, A. P., Nielsen, A. L. *et al.* The GCKR rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia* **51**, 70–75 (2008).
34 Wu, Y., Li, H., Loos, R. J., Yu, Z., Ye, X., Chen, L. *et al.* Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8 and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. *Diabetes* **57**, 2834–2842 (2008).
35 Lango, H., Palmer, C. N., Morris, A. D., Zeggini, E., Hattersley, A. T., McCarthy, M. I., *et al.* The UK Type 2 Diabetes Genetics Consortium Assessing the combined impact of 18 common genetic variants of modest effect sizes on type 2 diabetes risk. *Diabetes* **57**, 3129–3135 (2008).
36 van Hoek, M., Dehgan, A., Witteman, J. C., van Duijn, C. M., Uitterlinden, A. G., Oostra, B. A. *et al.* Predicting type 2 diabetes based on polymorphisms from genome wide association studies: a population-based study. *Diabetes* **57**, 3122–3128 (2008).
37 Meigs, J. B., Shrader, P., Sullivan, L. M., McAteer, J. B., Fox, C. S., Dupuis, J. *et al.* Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med.* **359**, 2208–2219 (2008).
38 Lyssenko, V., Jonsson, A., Almgren, P., Pulizzi, N., Isomaa, B., Tuomi, T. *et al.* Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N. Engl. J. Med.* **359**, 2220–2232 (2008).

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