Available online at www.sciencedirect.com

SCIENCE DIRECT

Genomics 85 (2005) 591–599

GENOMICS

www.elsevier.com/locate/ygeno

Identification of candidate genes in the type 2 diabetes modifier locus using expression QTL

Hiroshi Yaguchi[a,b], Katsuhiko Togawa[a,b], Maki Moritani[a], Mitsuo Itakura[a,*]

[a]*Division of Genetic Information, Institute for Genome Research, University of Tokushima, Tokushima 770-8503, Japan*
[b]*Division of Pharmacology, Drug Safety, and Metabolism, Otsuka Pharmaceutical Factory, Inc., Tokushima 772-8601, Japan*

Received 10 December 2004; accepted 20 January 2005

Abstract

To identify new genetic determinants relevant to type 2 diabetes (T2D), diabetic F2 progeny were generated by intercrossing F1 mice obtained from a cross of BKS.Cg-*Lepr*^{db}+/+*m* and DBA/2, and T2D-related phenotypes were measured. In the F2 population, increased susceptibility to diabetes and obesity was observed. We also detected the major quantitative trait loci (QTL) modifying the severity of diabetes on chromosome 9, where peaks of logarithm of odds (LOD) overlapped for three traits. To identify candidate genes in the QTL intervals, we combined "expression QTL" (eQTL), taking mRNA levels as quantitative traits, and "interstrain sequence variations, including cSNPs." As a result, four genes were identified from cosegregation of clinical QTL with eQTL and 13 genes were found from interstrain cSNPs as candidates in the T2D modifier QTL. Our combined approach shows the acceleration of the discovery of candidate genes in the QTL of interest, spanning several megabases.

Keywords: Type 2 diabetes; QTL analysis; Expression QTL; Candidate genes

Quantitative trait loci (QTL) mapping in mice has revealed hundreds of chromosomal regions containing genes affecting polygenic traits. Once QTL are identified, the selection of candidate genes can usually begin. The process of selecting candidate genes relies on a wealth of information gained through molecular approaches and several genomic resources, i.e., the mouse consensus genome sequence, interstrain haplotype information, interstrain polymorphic information, and human–mouse orthologs. Despite the useful bioinformatic tools, it remains difficult to identify suitable candidate genes in the QTL region and few complex-trait genes have been discovered so far [1]. Novel methods to identify the candidate genes in the QTL region need to be developed.

Schadt et al. have performed a genome-wide linkage analysis of gene expression and identified genetic regions (loci) that can account for variation in the levels of gene expression [2]. They term these loci "eQTL" for expression quantitative trait loci against "cQTL" for the classical or clinical trait loci such as height, weight, and blood glucose (BG) concentrations.

In this study, we present two approaches to identify candidate genes responsible for the effect of cQTL: one is based on eQTL analysis and the second on the interstrain coding single-nucleotide polymorphisms (cSNPs). By combining these two approaches, we effectively identified candidate genes in the type 2 diabetes (T2D) modifier locus using F2 progeny by intercrossing F1 mice obtained from a cross of BKS.Cg-*Lepr*^{db}+/+*m* (BKS-*db*/+*m*) and DBA/2 (D2) mice. Seventeen candidates were found from a total of 106 genes in the target QTL region of approximately 10 Mb.

Results

Characterization of D2BKS F2-db/db mice

The diabetes (*db*) mutation is a result of a point mutation in the leptin receptor gene, *Lepr*. The ligand, leptin, is a key

* Corresponding author. Fax: +81 88 633 9455.
E-mail address: itakura@genome.tokushima-u.ac.jp (M. Itakura).

0888-7543/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2005.01.006

Table 1
Values for phenotypes for parental strains and F2 mice

	Mouse strain				
	DBA	BKS	BKS-*db*/+*m*	BKS-*db*/*db*	F2-*db*/*db*
Sex	Female	Female	Female	Female	Female
Number of mice	15	20	12	18	113
Blood glucose at 0 min in ipGTT (mM)	4.2 ± 0.2	3.8 ± 0.2	3.9 ± 0.2	10.5 ± 0.8	18.2 ± 1.0
Blood glucose at 30 min in ipGTT (mM)	10.7 ± 0.5	17.0 ± 0.8	13.0 ± 0.7	29.2 ± 1.6	39.7 ± 0.6
Blood glucose at 60 min in ipGTT (mM)	11.4 ± 0.4	10.2 ± 0.5	11.9 ± 0.8	31.7 ± 1.7	39.9 ± 0.6
Blood glucose at 120 min in ipGTT (mM)	7.8 ± 0.3	5.3 ± 0.2	7.3 ± 0.4	32.6 ± 2.3	39.2 ± 0.7
Body weight at 9 weeks of age (g)	18.6 ± 0.3	18.2 ± 0.2	19.2 ± 0.3	31.6 ± 0.5	36.7 ± 0.4
Parametrial fat pad weight (mg)	350.8 ± 26.5	87.1 ± 6.6	377.9 ± 39.0	1874.1 ± 66.6	2831.9 ± 56.1

Data are shown as means ± SE.

weight control hormone and mice homozygous for the *db* mutation become remarkably obese and hyperglycemic [3]. Compared to the two parental strains and the heterozygotes for the *db* mutation (BKS-*db*/+*m*), homozygotes for the *db* mutation (BKS-*db*/*db* and F2-*db*/*db*) exhibited higher body weights, BG concentration in ipGTT, and parametrial fat pad weight (Table 1 and Fig. 1).

Also, there were marked mean differences between BKS-*db*/*db* and F2-*db*/*db* for all traits examined. Body weight and BG concentration (0 min in ipGTT) in F2-*db*/*db* were higher than those in BKS-*db*/*db* ($p < 0.01$, 36.7 ± 0.4 versus 31.6 ± 0.5 g, and $p < 0.05$, 18.2 ± 1.0 versus 10.5 ± 0.8 mM, respectively). This tendency was maintained after glucose injection in ipGTT. Parametrial fat pad weight in F2-*db*/*db* was significantly higher than that in BKS-*db*/*db* ($p < 0.01$, 2831.9 ± 56.1 versus 1874.1 ± 66.6 mg). Fig. 1 illustrates the distribution of body weight and BG concentration (0 min in ipGTT). In F2-*db*/*db*, body weight was distributed over a range of 25–48 g in a unimodal fashion. For BG concentrations, F2-*db*/*db* exhibited a much broader bimodal distribution than BKS-*db*/*db*.

cQTL analysis and confidence QTL region

Characterization of D2BKS F2-*db*/*db* mice strongly indicated that the introgression of the D2 background genetically predisposed the F2 population to diabetes and obesity. To identify the chromosomal loci of modifiers determining increased susceptibility to diabetes and obesity, a genome-wide scan was performed to map QTL for six clinical traits, including BG concentrations in ipGTT at 0, 30, 60, and 120 min; body weight; and parametrial fat pad weight. In this analysis, a cQTL was found on chromosome 9 (Chr. 9) with high lod scores for the three traits ipGTT at 120 min, body weight, and parametrial fat pad weight (Fig. 1A). The maximum lod score and genetic variance were 5.67 and 21% (BG concentrations at 120 min in ipGTT), 3.35 and 13% (body weight), and 4.22 and 16% (parametrial fat pad weight), respectively. No significant linkage was observed for BG concentrations at 0 min (LOD 1.97) and 30 min (LOD 0.87) in ipGTT, and weak linkage was found for 60 min (LOD 2.65). The effect on the latter part of ipGTT may suggest this cQTL is not responsible for the glucose-induced initial insulin response but is responsible for insulin resistance in muscle, adipocytes, and liver.

Bootstrap analysis estimated the confidence QTL regions for ipGTT at 120 min (top in Fig. 2B), body weight (middle in Fig. 2B), and parametrial fat pad weight (bottom in Fig. 2B). The highest single peak was not consistently observed at the same location for the three traits, i.e., the high peaks for body weight and parametrial fat pad weight were not reproducible by the peak for BG concentration at 120 min in

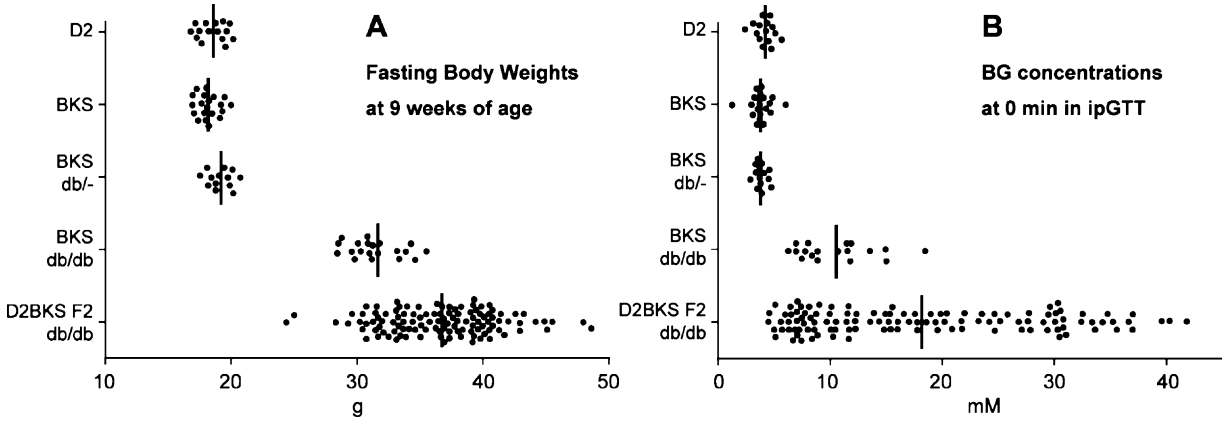

Fig. 1. Scatter plots of (A) fasting body weights at 9 weeks of age and (B) BG concentrations at 0 min in ipGTT. Plots display measurements from D2 ($n = 15$), BKS ($n = 20$), BKS-*db*/− ($n = 12$), BKS-*db*/*db* ($n = 18$), and F2-*db*/*db* ($n = 113$). Bold lines show the mean value.