NOTE *Biochemistry*

Principal Component and Interacting QTL Analyses Identify Novel Gene Loci Associated with Metabolic Traits in Mice

Jun-ichi SUTO[1)]*

[1)]*Division of Animal Sciences, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305–8634, Japan*

(Received 6 December 2009/Accepted 7 March 2010/Published online in J-STAGE 19 March 2010)

ABSTRACT. Previously analyzed F_2 mice between KK/Ta and RR/Sgn strains were further investigated by defining appropriate permutation-derived threshold levels for significance and by searching pairwise gene interactions. In addition, a principal component analysis was conducted to extract a potential parameter that accounts for the joint occurrence of metabolic abnormalities. As a result, one significant interaction, containing novel QTL on chromosome 15, was identified for plasma total-cholesterol levels. For the principal component that potentially accounted for the joint occurrence of metabolic abnormalities, one significant QTL was identified on chromosome 12. This locus was not significant for any single trait. These complex genetic bases could not be disclosed as long as a separate trait was analyzed by traditional single QTL scans.

KEY WORDS: metabolic traits, principal component analysis, quantitative trait loci (QTL).

J. Vet. Med. Sci. 72(8): 1081–1084, 2010

Plasma lipid concentrations are representative quantitative traits that are controlled by multiple genes together with non-genetic environmental effects. Alterations in plasma lipid levels are frequently associated with other metabolic disorders such as non-insulin-dependent diabetes mellitus (NIDDM) and obesity, and are recognized as major risk factors for cardiovascular disease as well [10]. Molecular mechanisms for genetic control of plasma lipid levels have been extensively studied in animal models, but are not yet fully understood. In order to identify genes influencing variation in plasma cholesterol and triglyceride levels, we previously performed a series of quantitative trait loci (QTL) analyses on plasma cholesterol and triglyceride levels in various F_2 intercrosses between mouse strains [13–15]. In one such cross, KK/Ta (hereafter referred to as KK) × RR/Sgn (hereafter referred to as RR) F_2 mice, we identified significant cholesterol QTL on the proximal part of chromosome 9 and significant triglyceride QTL on the mid-part of chromosome 8, but we did not identify any other suggestive loci [15]. In the present study, I newly added data on fasting plasma glucose and body weight at 160 days after birth to previous ones and re-analyzed them with refined statistical protocols. These traits were included in the analysis in order to accomplish more comprehensive evaluation of metabolic abnormalities. The major aims of this study were as follows: (1) To set a genome-wide significance threshold by permutation test; thus, whether additional loci are identified or not was tested for each trait. (2) To perform a pairwise genome scan to reveal possible gene interactions, because gene-to-gene interactions were suggested to be involved in the manifestation of metabolic traits [2]. (3) To asses all metabolic traits simultaneously as a single multivariate character by means of principal component analysis, because a group of metabolic derangements tended to occur together [2]. In man, several genome-wide linkage studies in combination with principal component analysis successfully identified responsible chromosomal regions [1, 3–7, 9]. KK is known to develop NIDDM with moderate obesity. RR is not diabetes-prone, but has elevated plasma lipid levels [15] and is moderately obese; therefore, it will be appropriate to study genetic aspects of overall metabolic derangements in this F_2 set. Identification of novel loci will be some help toward further understanding of complex metabolic traits.

* CORRESPONDENCE TO: SUTO, J., Division of Animal Sciences, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305–8634, Japan.
e-mail: jsuto@affrc.go.jp

As described in the previous paper [15], a total of 145 F_2 female mice were analyzed. They had free access to food [CE-2 (342.2 kcal/100 g, containing 4.4% crude fat), Clea Japan, Tokyo] and water.

Body weight was determined at 160 days after birth. Plasma blood glucose, total-cholesterol, and triglyceride levels were determined at the age of 170 days after 24-hr fasting [15]. Histograms showing distributions of body weight and plasma glucose levels in F_2 mice are shown in Fig. 1.

The vector of the above-mentioned phenotypic measurements of four traits for each individual was handled as a single multivariate character. Therefore, the phenotypic data were analyzed by principal component analysis with SPSS for Windows (release 7.5.1-J, SPSS Inc., Chicago, IL, U.S.A.).

Table 1 gives the eigenvalue and its contribution with respect to the principal component (hereafter referred to as PC) in F_2 mice. Two PCs, in which the eigenvalue was more than 1.0, were successfully extracted. They accounted for about 64% of the variation in metabolic-trait information. In the case of PC1, all coefficients for the variables were positive (Table 2), thus suggesting that PC1 was a factor concerned with the overall severity of metabolic abnormali-

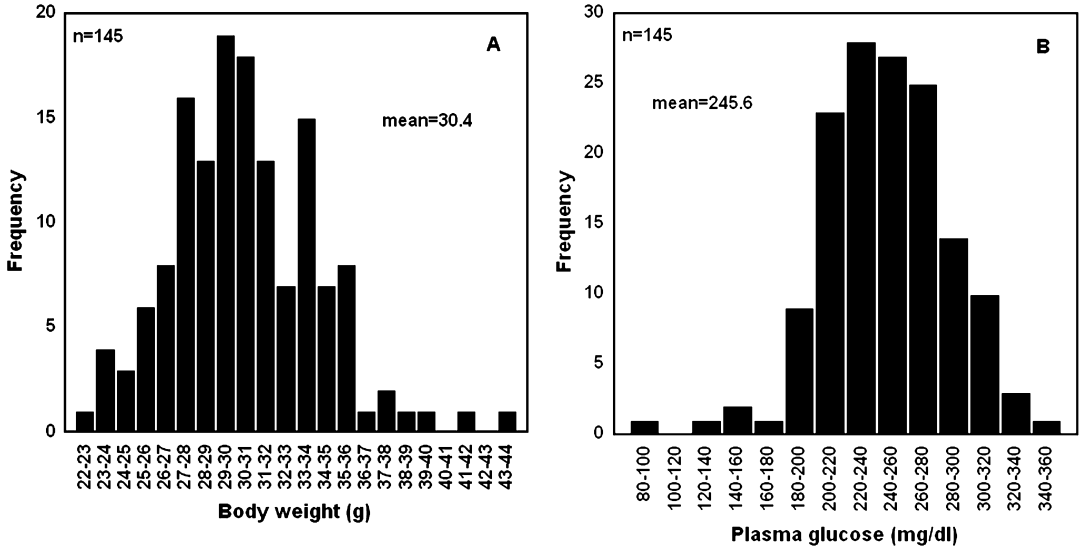

Fig. 1. Histograms showing distributions of body weight (A) and fasting plasma glucose (B) in F_2 mice. The means of the trait values and the number of F_2 mice (n) are also shown.

Table 1. Eigenvalue and its contribution to each Principal Component (PC)

PC	Eigenvalue	Cumulative contribution ratio (%)
1	1.499	37.483
2	1.031	63.260

Table 2. Eigenvector of each Principal Component (PC)

Variables	PC1	PC2
Body weight	0.310	0.790
Fasting glucose	0.606	–0.169
Total-cholesterol	0.511	0.263
Triglyceride	0.525	–0.527

ties about body weight, fasting glucose, and plasma lipids.

First, a genome-wide scan was performed for single QTL by use of Map Manager QTX b20 [12]. Prior to QTL analysis, all trait values were log-transformed to make sure of a normal distribution. The threshold likelihood ratio statistics (LRS) for suggestive (P<0.63), significant (P<0.05), and highly significant (P<0.001) linkages at the genome-wide 5% level were determined by performance of 1,000 permutations for each trait. The LRS was converted to a logarithm of odds (LOD) score by dividing of the LRS by 4.605. Once a significant QTL was identified, the 95% confidence interval (CI) was defined by 1.5 LOD fall.

The permutation-derived threshold LOD scores for suggestive and significant linkages were 1.8 and 3.1–3.3, respectively. In total, 12 single QTL were identified (Table 3). In

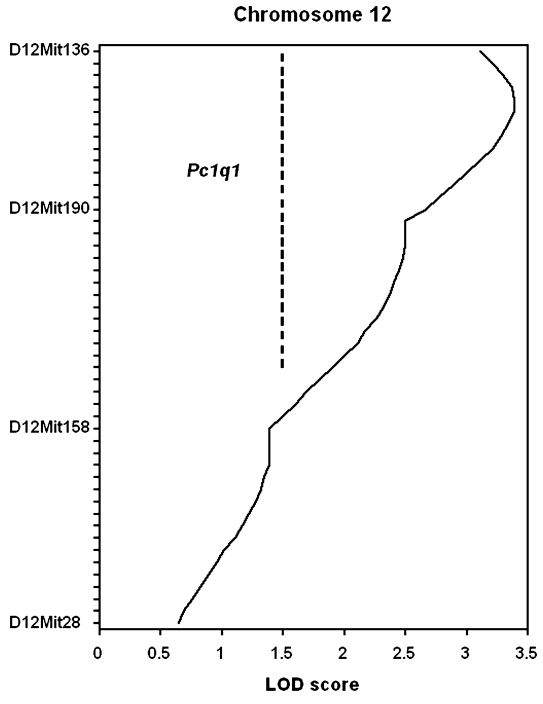

Fig. 2. LOD score plots for *Pc1q1* on chromosome 12. The x-axis represents LOD scores. The y-axis represents microsatellite markers used for this study. The vertical broken line indicates 95% CI for *Pc1q1*.

particular, for PC1, one suggestive QTL was identified on chromosome 11, and one significant QTL was identified on chromosome 12 (Fig. 2). I assigned the locus name *Pc1q1*

Table 3. Identification of QTL

Traits	Chromosome (Peak position, cM)	95% CI[a)]	LOD (Variance, %)[b)]	High allele[c)]	Closest marker	Locus name[d)]	Overlapping QTLs [ref][e)]
Body weight	11 (24)	0–42	2.4 (7)	KK	*D11Mit86*		*Bwdq1* [17]
Fasting glucose	9 (11)	0–37	2.6 (8)	RR	*D9Mit90*		
	12 (18)	0–35	3.1 (10)	RR	*D12Mit136*		
Total-cholesterol	1 (95)	78–*	2.2 (7)	KK	*D1Mit291*		*Cq2* [13], *Cq6* [14]
	3 (45)	0–*	2.0 (6)	RR	*D3Mit29*		*Cq3* [13, 16]
	9 (20)	0–46	6.0 (17)	KK	*D9Mit229*	*Cq5*	*Cq4* [15]
	12 (18)	0–49	2.2 (7)	RR	*D12Mit136*		
	15 (53)	na			*D15Mit190*	*Cq7* *	
Triglyceride	1 (43)	0–68	2.3 (7)	RR	*D1Mit10*		
	8 (30)	17–38	4.4 (13)	KK	*D8Mit205*	*Tgq2*	*Tgq2* [15]
PC1	11 (26)	0–43	3.0 (9)	KK	*D11Mit86*		
	12 (18)	0–35	3.4 (11)	RR	*D12Mit136*	*Pc1q1*	

a) An estimate of the size of a 95% confidence interval (CI) for a QTL, *: distal end of CI could not be defined, na: not applicable, nd: not determined because adjacent marker was distally situated.
b) The amount of the total trait variance which would be explained by a QTL at this locus, in percent.
c) Allele which is associated with higher trait values.
d) Locus name was assigned only to significant QTLs, and * means that this locus was identified only as an interacting QTL.
e) Description about overlapping QTLs for relevant trait was kept to a minimum.

(principal component 1 QTL No. 1) to the latter locus. The 95% CI for *Pc1q1* overlapped with those for two suggestive QTL for fasting glucose and total-cholesterol, and the RR allele was always associated with an increase in trait values at these loci, thus suggesting that they are allelic. It is important to note that *Pc1q1* was not significant QTL for any single traits when the traits were analyzed separately. Therefore, I considered that *Pc1q1* was suggested to indicate the significance of overall metabolic abnormalities. Because of a large 95% CI for *Pc1q1* on chromosome 12 (Fig. 2), it was difficult to look for a plausible candidate gene. Nevertheless, the apolipoprotein B (*Apob*) gene locus could be regarded as one of the candidate loci. *Apob*-deficient mice were reported to have significantly lower body weight and significantly decreased total-cholesterol compared to wild-type control mice, but triglyceride levels did not differ significantly between mutant and control mice when they were fed on a chow diet [8]. Indeed, *Pc1q1* had no significant effect on triglyceride levels.

After single QTL scans, the potential interaction between marker loci was evaluated pairwise. For this analysis, the following two-stage test was used. First, the threshold for significance at the genome-wide 5% level was determined by performance of 1,000 permutations on the interaction model of Map Manager QTX b20, and then the significance of the total effect of the two loci was tested. Second, when pairs of loci had an LOD score exceeding the threshold, the significance of the pairwise interaction was evaluated. According to the Map Manager QTX manual, the interaction effect itself must have a *P* value less than 0.01. In this study, I also employed a stringent criterion, that only the pair of loci with interaction LOD scores of 3 or greater was considered to be a statistically significant interaction [11].

The permutation-derived threshold total LOD scores for association were 7.7–8.3. One significant interaction was identified for total-cholesterol between *D9Mit229* and *D15Mit190* (Fig. 3). The total LOD score association was 9.6, and the interaction LOD score was 3.6. I assigned gene symbol *Cq7* (cholesterol QTL No. 7) to the locus on chromosome 15. Mice homozygous for the KK allele at *D9Mit229* and heterozygous at *D15Mit190* tended to have the highest cholesterol levels. There is a cluster of apolipoprotein L genes in the neighboring region of *Cq7* on chromosome 15, but I have no evidence to support their candidacy. It is important to note that *Cq7* was by no means identified by a single QTL scan. Taken together, these complex genetic bases could not be disclosed as long as separate

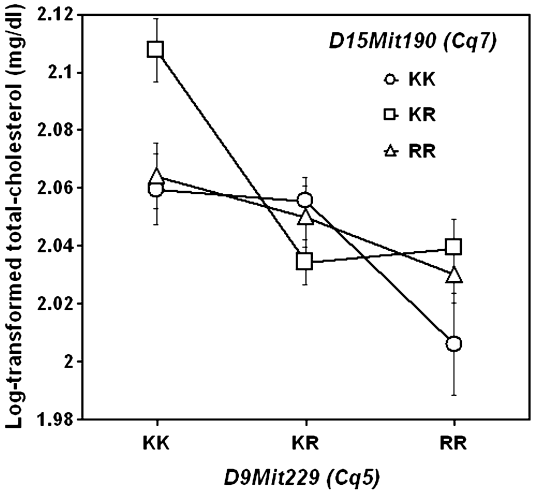

Fig. 3. Identification of interacting QTL for total-cholesterol. Significant interaction was identified between *D9Mit229* (*Cq5*) and *D15Mit190* (*Cq7*). Error bars indicate SE.

trait was analyzed by traditional single QTL scans.

This work was funded in part by a grant from the Ministry of Agriculture, Forestry and Fisheries, Japan.

REFERENCES

1. Arya, R., Blangero, J., Williams, K., Almasy, L., Dyer, T. D., Leach, R. J., O'Connell, P., Stern, M. P. and Duggirala, R. 2002. Factors of insulin resistance syndrome-related phenotypes are linked to genetic locations on chromosomes 6 and 7 in nondiabetic Mexican-Americans. *Diabetes* **51**: 841–847.
2. Bouchard, C. 1995. Genetics and the metabolic syndrome. *Int. J. Obes. Relat. Metab. Disord.* **19**: S52–S59.
3. Cai, G., Cole, S. A., Freeland-Graves, J. H., MacCluer, J. W., Blangero, J. and Comuzzie, A. G. 2004. Principal component for metabolic syndrome risk maps to chromosome 4p in Mexican Americans: the San Antonio Family Heart Study. *Hum. Biol.* **76**: 651–665.
4. Cheng, C. Y., Lee, K. E., Duggal, P., Moore, E. L., Wilson, A. F., Klein, R., Bailey-Wilson, J. E. and Klein, B. E. K. 2009. Genome-wide linkage analysis of multiple metabolic factors: evidence of genetic heterogeneity. *Obesity* (in press).
5. de Andrade, M., Olswold, C., Kardia, S. L. R., Boerwinkle, E. and Turner, S. T. 2002. Multivariate linkage analysis using phenotypes related to the insulin resistance-metabolic disorder. 11th Annual Meeting of the International Genetic Epidemiology Society. *Genet. Epidemiol.* **23**: 275.
6. Edwards, K. L., Hutter, C. M., Wan, J. Y., Kim, H. and Monks, S. A. 2008. Genome-wide linkage scan for the metabolic syndrome: the GENNID study. *Obesity* **16**: 1596–1601.
7. Hamet, P., Merlo, E., Šeda, O., Broeckel, U., Tremblay, J., Kaldunski, M., Gaudet, D., Bouchard, G., Deslauriers, B., Gagnon, F., Antoniol, G., PausovOá, Labuda, M., Jomphe, M., Gossard, F., Tremblay, G., Kirova, R., Tonellato, P., Orlov, S. N., Pintos, J., Platko, J., Hudson, T. J., Rioux, J. D., Kotchen, T. A. and Cowley, Jr. A. W. 2005. Quantitative founder-effect analysis of French Canadian families identifies specific loci contributing to metabolic phenotypes of hypertension. *Am. J. Hum. Genet.* **76**: 815–832.
8. Huang, L. S., Voyiaziakis, E., Markenson, D. F., Sokol, K. A., Hayek, T. and Breslow, J. L. 1995. apoB gene knockout in mice results in embryonic lethality in homozygotes and neural tube defects, male infertility, and reduced HDL cholesterol ester and apo A-I transport rates in heterozygotes. *J. Clin. Invest.* **96**: 2152–2161.
9. Loos, R. J. F., Katzmarzyk, P. T., Rao, D. C., Rice, T., Leon, A. S., Skinner, J. S., Wilmore, J. H., Rankinen, T. and Bouchard, C. 2003. Genome-weide linkage scan for the metabolic syndrome in the HERITAGE family study. *J. Clin. Endocrinol. Metab.* **88**: 5935–5943.
10. Lusis, A. J., Attie, A. D. and Reue, K. 2008. Metabolic syndrome: from epidemiology to systems biology. *Nat. Rev. Genet.* **9**: 819–830.
11. Lyons, M. A., Korstanje, R., Li, R., Sheehan, S. M., Walsh, K. A., Rollins, J. A., Carey, M. C., Paigen, B. and Churchill, G. A. 2005. Single and interacting QTLs for cholesterol gallstones revealed in an intercross between mouse strains NZB and SM. *Mamm. Genome* **16**: 152–163.
12. Manly, K. F., Cudmore, R. H. Jr. and Meer, J. M. 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mamm. Genome* **12**: 930–932.
13. Suto, J., Matsuura, S., Yamanaka, H. and Sekikawa, K. 1999. Quantitative trait loci that regulate plasma lipid concentration in hereditary obese KK and KK-A^y mice. *Biochim. Biophys. Acta* **1453**: 385–395.
14. Suto, J., Takahashi, Y. and Sekikawa, K. 2004. Quantitative trait locus analysis of plasma cholesterol and triglyceride levels in C57BL/6J × RR F_2 mice. *Biochem. Genet.* **42**: 347–363
15. Suto, J. and Sekikawa, K. 2003. Quantitative trait locus analysis of plasma cholesterol and triglyceride levels in KK × RR F_2 mice. *Biochem. Genet.* **41**: 325–341.
16. Suto, J. 2006. Characterization of *Cq3*, a quantitative trait locus that controls plasma cholesterol and phospholipid levels in mice. *J. Vet. Med. Sci.* **68**: 303–309.
17. Suto, J. 2008. Genetic dissection of testis weight in a mouse strain having an extremely large testis: major testis weight determinants are autosomal rather than Y-linked on the basis of comprehensive analyses in Y-chromosome consomic strains. *Proc. Jpn. Acad. Ser. B* **84**: 393–406.