Systems genetics of susceptibility to obesity-induced diabetes in mice

Richard C. Davis,[1]* Atila van Nas,[2]* Lawrence W. Castellani,[1] Yi Zhao,[1] Zhiqiang Zhou,[1] Pingzi Wen,[1] Suzanne Yu,[1] Hongxiu Qi,[1] Melenie Rosales,[1] Eric E. Schadt,[5] Karl W. Broman,[6] Miklós Péterfy,[1,7] and Aldons J. Lusis[1,2,3,4]

Departments of [1]Medicine, [2]Human Genetics, and [3]Microbiology, Immunology and Molecular Genetics, [4]Molecular Biology Institute, David Geffen School of Medicine, University of California, Los Angeles, California; [5]Department of Genetics, Rosetta Inpharmatics, Merck, Seattle, Washington; [6]Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, Wisconsin; and [7]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California

Davis RC, van Nas A, Castellani LW, Zhao Y, Zhou Z, Wen P, Yu S, Qi H, Rosales M, Schadt EE, Broman KW, Péterfy M, Lusis AJ. Systems genetics of susceptibility to obesity-induced diabetes in mice. *Physiol Genomics* 44: 1–13, 2012. First published October 18, 2011; doi:10.1152/physiolgenomics.00003.2011.—Inbred strains of mice are strikingly different in susceptibility to obesity-driven diabetes. For instance, deficiency in leptin receptor (*db/db*) leads to hyperphagia and obesity in both C57BL/6 and *DB*A/2 mice, but only on the *DB*A/2 background do the mice develop beta-cell loss leading to severe diabetes, while C57BL/6 mice are relatively resistant. To further investigate the genetic factors predisposing to diabetes, we have studied leptin receptor-deficient offspring of an F2 cross between C57BL/6J (*db/+*) males and *DB*A/2J females. The results show that the genetics of diabetes susceptibility are enormously complex and a number of quantitative trait loci (QTL) contributing to diabetes-related traits were identified, notably on chromosomes 4, 6, 7, 9, 10, 11, 12, and 19. The Chr. 4 locus is likely due to a disruption of the *Zfp69* gene in C57BL/6J mice. To identify candidate genes and to model coexpression networks, we performed global expression array analysis in livers of the F2 mice. Expression QTL (eQTL) were identified and used to prioritize candidate genes at clinical trait QTL. In several cases, clusters of eQTLs colocalized with clinical trait QTLs, suggesting a common genetic basis. We constructed coexpression networks for both 5 and 12 wk old mice and identified several modules significantly associated with clinical traits. One module in 12 wk old mice was associated with several measures of hepatic fat content as well as with other lipid- and diabetes-related traits. These results add to the understanding of the complex genetic interactions contributing to obesity-induced diabetes.

leptin receptor; leptin signaling; C57BLKS; Zfp69; coexpression networks

INBRED STRAINS OF MICE EXHIBIT variable susceptibility to Type 2 diabetes. A classic approach demonstrating the role of genetic background on the development of obesity-associated diabetes has been the introgression of diabetes (*db*) mutation (9) into various inbred strains (4). The *db* mutation disrupts leptin signaling by abrogating the function of leptin receptor and results in hyperphagia and obesity regardless of strain background. In contrast, inbred mouse strains exhibit profound metabolic differences in response to obesity induced by *db*. Some strains, such as *DB*A/2 (*DB*A)-*db/db* and C57BLKS (BKS)-*db/db*, develop severe diabetes, whereas others, including C57BL6 (B6)-*db/db* and 129-*db/db*, are resistant to the disease (15). Although the role of genetic factors in the development of obesity-associated diabetes has long been appreciated, the genes responsible for differences in diabetes susceptibility remain largely uncharacterized.

In the present study, we examined obese (*db/db*) F2 mice from a cross between B6 and *DB*A to identify loci that contribute to diabetic progression, insulin resistance, and lipid accumulation in the liver. We characterized two cohorts of mice: one at 5 wk of age prior to anticipated islet depletion and another cohort at 12 wk when BKS *db/db* begin to show decreased insulin production but prior to the onset of significant diabetic complications. For the loci contributing to diabetes-related traits that were identified, we examined candidate genes for *cis*-acting variations in gene expression. Additionally, we modeled liver gene coexpression networks to help identify mechanisms contributing to metabolic dysfunction in obesity-induced diabetes.

METHODS

Animals. We carried out an F2 intercross between the inbred strains *DB*A/2 and C57BL/6. These parental mice were obtained from The Jackson Laboratory (Bar Harbor, ME). The male C57BL/6 parents carried heterozygous deficiency in the leptin receptor (*db +/−*) and F1 progeny were selected for the presence of the mutant allele. Among F2 progeny, only those with homozygous deficiency in leptin receptor (*db/db*) were selected for further phenotype and genotype analysis. Animals were housed in vivaria accredited by the Association for Assessment and Accreditation of Laboratory Animal Care, and all procedures were approved by the UCLA Institutional Animal Care and Use Committee. The breeding colony was maintained on a 12 h light-dark cycle with lights on from 6 AM to 6 PM, and animals were fed a chow diet with 6% fat by weight.

Body composition and liver fat content by quantitative nuclear magnetic resonance. Animals were measured for total body fat mass and lean mass by nuclear magnetic resonance (NMR) using the Bruker Minispec with software from Echo Medical Systems (Houston, TX) (23). All F2 mice were measured at 4 wk of age. Animals in the 12 wk cohort were measured again on the day prior to euthanasia. For measurement of liver fat content, a 0.5–1.0 g sample was collected at euthanasia and analyzed using an EchoMRI 3-in-1 quantitative NMR system, also from Echo Medical Systems (23).

Plasma lipids and proteins. Plasma was collected from mice that were fasted for 4–5 h and bled 2–3 h after the beginning of the light cycle from the retro-orbital plexus under isoflurane anesthesia. Total cholesterol, HDL cholesterol, free cholesterol, triglycerides, and FFA concentrations were determined in triplicate as described previously (7, 19).

Levels of insulin in the plasma were measured by RBM (Austin, TX) using their RodentMAP quantitative, multiplexed immunoassay system. The samples, stored at −80°C, were thawed at room temper-

ature, vortexed, and spun at 13,000 *g* for 5 min for clarification. An aliquot of each sample was introduced into one of the capture microsphere multiplexes of the RodentMAP and incubated at room temperature for 1 h. Multiplexed cocktails of biotinylated, reporter antibodies were then added. Following incubation for 1 h at room temperature, multiplexes were developed using an excess of streptavidin-phycoerythrin and further incubation for 1 h at room temperature. Fluorescent intensities were determined using a Luminex 100 instrument, and protein levels determined based on eight point concentration standards and controls included on each microtiter plate.

Genotyping. Genomic DNA was isolated by standard procedures. Genotyping for the *db* mutation was carried out essentially as described by Horvat and Bünger (8) except that the *Acc*I digestion was carried out at 55°C for 1.5 h to increase efficiency of cutting. Genome-wide analysis of single nucleotide polymorphism (SNP) genotypes for quantitative trait locus (QTL) mapping was carried out using the Affymetrix GeneChip Mouse Mapping 5K SNP panel (14). This panel carries >2,800 informative makers between C57BL/6 and *DB*A/2 (average spacing: 0.5–0.6 cM) making it quite suitable for resolving recombination breakpoints in the F2 population. This informative set was filtered for those with Affymetrix quality scores >0.9 and showing 10–90% heterozygosity among F2 animals except on chromosome 4 where marker distribution was distorted by selection for homozygosity of the leptin receptor knockout.

QTL mapping. Phenotypic traits were log transformed to normalize the trait distribution. Logarithm of the odds (LOD) score distributions were calculated using the scanone function in R/qtl (http://www.rqtl.org) (http://www.cran.r-project.org) (2). Sex and age were treated as interactive covariates as needed (24). In most cases, we carried out separate analysis of the 5 wk and 12 wk cohorts using sex as an interactive covariate. Permutations of *n* = 1,000 were carried out to determine thresholds for significance at *P* < 0.05 as a means to account for the multiple testing of many markers for association with each trait.

Multiple QTL analysis. To examine for multiple interacting QTLs that might influence diabetes-related phenotypes, these traits were analyzed using the model selection approach described by Manichaikul et al. (17). Because this approach has not yet been extended to consider QTL-by-sex interactions, we focused on sex-specific analyses.

Expression arrays. Total liver RNA preparation and array hybridizations were performed at Rosetta Inpharmatics/Merck Pharmaceuticals using the Affymetrix GeneChip Mouse Genome 430 2.0 Array that assays for levels of 32,640 transcripts. The microarray data from this study have been deposited in the Gene Expression Omnibus with accession number GSE30140.

Gene expression quantitative trait mapping. Expression QTL (eQTL) were determined by using the transcript abundance of each gene as a quantitative trait and correlating them to genetic markers to map gene eQTL (24). The eQTL were calculated utilizing the scanone function in R/qtl (http://www.rqtl.org) (2). Sex and age were treated as interactive covariates as needed. One thousand permutations of *n* = 1,000 randomly chosen gene transcripts resulted in a significance threshold of LOD >6.2 (*P* < 0.05).

Weighted gene coexpression network analysis. The construction of gene coexpression networks and modules was performed using the R/Weighted Gene Co-Expression Network Analysis (WGCNA) (13). Duplicate probes and controls were removed for analysis as well as probe sets with the lowest intensity [lower quartile of the robust multiarray average (RMA)], which was equivalent to RMA >3.77. The remainder of the probes were filtered on the highest coefficient of variation (>0.08) (three standard deviations above the mean). The resulting total of 15,397 genes was used to construct the networks in four groups: males aged *week 5* (*n* = 92), males aged *week 12* (*n* = 99), females aged *week 5* (*n* = 119) and females aged *week 12* (*n* = 109). To choose a particular power β, we used the scale-free topology

criterion as described (13, 26), which led us to power β = 4 for each of the groups.

Analysis of coexpression network correlation to metabolic and genetic traits. To determine correlation of traits with gene coexpression modules, the module "eigengene" was correlated to the metabolic trait values, and the module gene significance was determined. The module eigengene is the first principal component of gene expression of each module and explained most of the variance (55–70%) in gene expression of each module in this study. Module significance was calculated as the correlation between the module eigengene and clinical trait. Modules correlated with at least 4 or more traits at R >0.4 were further analyzed for QTL detection. We treated the eigengene as a trait to identify loci controlling the expression of all the genes within each module. We used the scanone function in R/qtl (2) and treated sex as an interactive covariate. The significance threshold was based on a permutation of *n* = 1,000 of randomly chosen probes relative to the number of probes of each module tested (13).

RESULTS

Experimental design. We constructed a cross between strains *DB*A and B6 on the background of the *db* mutation to examine susceptibility to obesity-induced diabetes. Mice were studied at 5 wk of age, prior to anticipated islet depletion (*n* = 221), and at 12 wk of age when BKS *db/db* mice exhibit reduced insulin production (*n* = 218).

Glucose and insulin levels in F2 mice. Surprisingly, many of the F2 *db/db* animals at both 5 wk and 12 wk showed far more advanced diabetes in terms of plasma insulin and glucose levels than was anticipated from levels seen at these ages in B6 *db/db* and BKS *db/db* animals (5). For example, at 5 wk of age (Fig. 1*A*), while most F2 *db/db* animals have low plasma glucose and insulin levels, ~23% showed elevated glucose levels (>300 mg/dl) indicative of outright diabetes at an age well before such glucose levels are observed in BKS *db/db*. At 12 wk (Fig. 1*B*), this trend came to predominate with most F2 animals showing insulin resistance (high insulin, low glucose) or diabetes (low insulin, high glucose) or elevated insulin and glucose, suggestive of a transition between insulin resistance and diabetes. This progression was generally more extreme in F2 animals than BKS *db/db* and indicates that the impact of the *db/db* mutation is much more severe among F2 animals, causing a diabetic phenotype at a significantly earlier age. This strong diabetic phenotype appearing in a large fraction of the F2 animals is compatible with the hypothesis that there are many contributing loci.

Consistent with observations in BKS *db/db* mice, male F2 animals were more severely affected. Female F2 mice at 12 wk showed wide variation in insulin and glucose levels, with many animals showing highly elevated plasma insulin and a wide range of plasma glucose levels. By contrast, F2 males were predominately diabetic with few animals having the elevated insulin levels observed commonly among females (Fig. 1*B*).

Insulin and glucose levels were also measured at 4–5 wk in the 12 wk cohort (Data given in Supplemental Table ST2).[1] As observed with the animals euthanized at 5 wk, the 12 wk cohort, when measured at 4–5 wk, fell into a single cluster with insulin <65 uIU per ml and glucose <350 mg/dl. A few 5 wk animals trended to higher insulin, consistent with incipient insulin resistance, while another group had elevated glucose

---

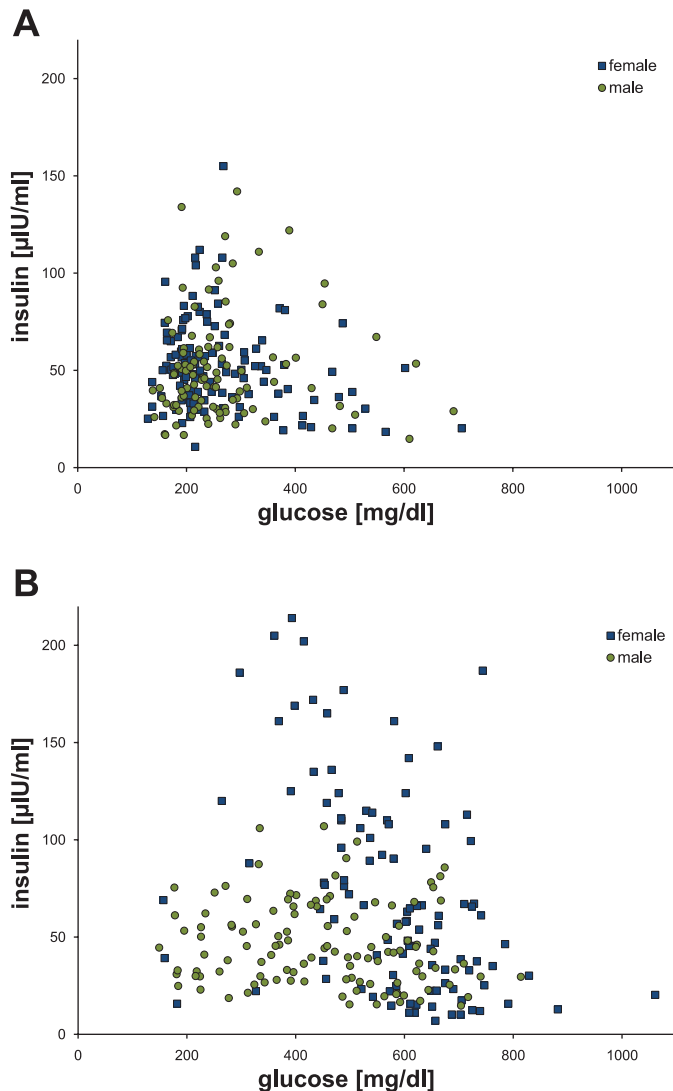[1] The online version of this article contains supplemental material.

Fig. 1. Insulin and glucose levels in F2 mice. The plot shows plasma levels of insulin and glucose in 5 wk cohort (*A*) and 12 wk cohort (*B*). In the 5 wk cohort, most animals clustered around glucose 200 mg/dl and insulin of 60 uIU/ml, whereas the 12 wk female mice showed a wide spread of glucose and insulin with >25% showing insulin levels >100 uIU/ml. By contrast, while most 12 wk males were hyperglycemic (>300 mg/dl), few expressed plasma insulin >100 uIU/ml. Females (blue squares); males (green circles). Because mapping was carried out using log-transformed data for normalization, plots of the log-normalized data are shown in Supplemental Fig. SF1.

lin sensitivity and, perhaps, greater capacity for insulin production. Islet RNAs from these animals have been isolated, and we anticipate future expression analysis of this tissue in relation to insulin and glucose levels as well as diabetes susceptibility.

Adiposity and liver fat correlations with diabetes-related traits in F2 mice. While all animals in the cross were genetically obese, there was considerable heterogeneity among F2 animals in body weight and adiposity as determined by NMR. And, in fact, adiposity among 5 wk animals was strongly and positively correlated with plasma glucose levels for both males and females (Table 1). For 12 wk males, this correlation is lost, perhaps due to diabetic wasting in these animals.

We had previously noted (5) that the diabetes-prone BKS *db/db* mice have much lower hepatic triglyceride content than the more resistant B6 *db/db* mouse. Those results were in contrast to humans, where hepatic steatosis is commonly associated with obesity and Type 2 diabetes. Therefore, in this present study, we assessed total lipid content in the livers of 5 wk and 12 wk animals. As shown in Table 1, young animals showed little correlation between liver lipids and other measures of gluco-homeostasis except for the significant correlation with overall adiposity in 5 wk females. However, consistent with the result in BKS *db/db* animals, in 12 wk males there is a significant negative correlation of liver fat to plasma glucose with a similar trend in females. The lesser correlation in 12 wk females may reflect the much greater capacity for insulin production that allows females to evade the extreme plasma glucose that would otherwise be produced.

QTL for clinical traits. To identify loci contributing to diabetic phenotypes in this cross, F2 animals were genotyped for a set of SNPs, ~2,800 of which are informative between B6 and *DB*A, resulting in an average marker spacing of ~0.5 cM. Using the SNP genotype data and phenotype data collected for all F2 mice we carried out QTL mapping for diabetes-related traits.

Peak markers correlated with clinical traits. Table 2 summarizes the location of significant diabetes-related QTLs. The table includes QTLs that reached genome-wide significance based on 1,000X permutations of the genotype data and that also showed significant ($P < 0.05$) differences in the trait between F2 animals homozygous for the B6 allele compared with animals homozygous for the *DB*A allele of the peak SNP. The "higher allele" indicates which strain-genotype has the higher trait value, while "*P* value" indicates the significance of that difference. Strikingly, higher values for almost all the traits derive from *DB*A, consistent with *DB*A alleles increasing most traits associated with diabetes susceptibility including hyperglycemia, hyperlipidemia, and several measures of adiposity. Only for the subcutaneous fat pad-mass QTL on chromosome 9 are higher values associated with the B6 allele. Supplemental Table ST1 contains an expanded set of QTL including those that do not reach genome-wide significance or that do not achieve a significant difference in trait value between F2 animals homozygous for the B6 allele compared with animals homozygous for the *DB*A allele of the peak SNP.

While all F2 animals were deficient in leptin receptor and thus genetically obese, there was genetically determined variation in body weight leading to a significant QTL for body weight on Chr. 7 (Fig. 2) in 12 wk animals as well as suggestive QTL on chromosome 12. However, mapping of

only, suggesting that, in this subpopulation, hyperglycemia occurred without a preceding hyperinsulinemia phase. Consistent with this, all animals from this hyperglycemic cluster at 5 wk showed increased glucose at 12 wk, but stable or decreased insulin. Interestingly, the hyperinsulinemic group at 5 wk generally showed no increase in insulin levels at 12 wk but had increased glucose levels, consistent with these animals being the most insulin resistant and thus producing higher insulin output at an early age. The failure of these insulin levels to increase with increasing age suggests that this subpopulation had a limited ability to increase insulin production in the face of escalating glucose levels. By contrast, almost all animals with highly elevated insulin at 12 wk came from the normoglycemic/insulinemic cluster at 4 wk consistent with higher insu-

Table 1. *Trait correlations among male and female F2 db/db mice at 5 wk (above diagonal) and 12 wk (below diagonal)*

	Glucose		Insulin		Adiposity		Liver Fat %	
	♂	♀	♂	♀	♂	♀	♂	♀
Glucose								
Corr.			$R = 0.041$	$R = -0.242$	$R = 0.41$	$R = 0.41$	$R = -0.009$	$R = -0.01$
P value			$(P = 0.6834)$	$(P = 0.008)$	$(P = 1.89e\text{-}5)$	$(P = 5.04e\text{-}06)$	$(P = 0.9299)$	$(P = 0.9121)$
Insulin								
Corr.	$R = -0.077$	$R = -0.12$			$R = 0.238$	$R = 0.032$	$R = -0.010$	$R = 0.072$
P value	$(P = 0.4312)$	$(P = 0.1982)$			$(P = 0.0166)$	$(P = 0.7296)$	$(P = 0.9247)$	$(P = 0.4374)$
Adiposity								
Corr.	$R = 0.04$	$R = 0.332$	$R = -0.098$	$R = -0.039$			$R = 0.083$	$R = -0.225$
P value	$(P = 0.6817)$	$(P = 0.0003)$	$(P = 0.3163)$	$(P = 0.6769)$			$(P = 0.4)$	$(P = 0.0052)$
Liver fat %								
Corr.	$R = -0.287$	$R = -0.135$	$R = 0.036$	$R = 0.093$	$R = 0.118$	$R = 0.149$		
P value	$(P = 0.003)$	$(P = 0.148)$	$(P = 0.7172)$	$(P = 0.3165)$	$(P = 0.2387)$	$(P = 0.1087)$		

Correlation coefficients and *P* values are shown in boldface for correlations with $P < 0.05$. Traits are: Glucose (plasma glucose, mg/dl), insulin (plasma insulin, uIU/ml), adiposity [total body fat mass/(total body fat mass + total body lean mass) determined by NMR] and liver fat (%) [liver fat mass/(liver fat-mass + liver lean mass) determined by NMR]. Roman type above diagonal shows 5 wk cohort values; 12 wk cohort values below diagonal are italicized.

whole body adiposity, as determined by NMR, revealed only a QTL on chromosome 11 (Fig. 2). By contrast, fat accumulation in liver had a strong QTL on chromosome 12 (Fig. 2), perhaps underlying the suggestive body weight QTL at the same location.

Unexpectedly, there were no strong QTLs for plasma insulin (Supplemental Fig. SF3A) or glucose (Supplemental Fig. SF3B). The highest peak for plasma insulin levels (LOD 4.2) was observed on distal chromosome 12, roughly coinciding with the liver fat QTL (Fig. 2) and the secondary QTL for body weight (Fig. 2).

In 5 wk animals, a QTL for plasma glucose (LOD 4.0) (Supplemental Fig. SF3A) was observed on distal chromosome 5, roughly coinciding with the second highest peak for insulin QTL (Supplemental Fig. SF3A). And for 12 wk animals there is a secondary glucose QTL (LOD 3) (Supplemental Fig. SF3B) on central chromosome 2 that roughly coincides with a QTL for insulin (LOD 3.3) seen in both 5 and 12 wk animals (Supplemental Fig. SF3A). We had hoped to use the same mapping strategy to identify loci governing the changes in insulin and glucose between 5 wk and 12 wk in the 12 wk

cohort. Unfortunately, no QTLs for the change in insulin, glucose, or glucose-to-insulin ratio reached threshold of significance in either males or females.

Interestingly, for 12 wk animals there also appeared coincident LOD peaks for glucose, insulin, and plasma triglycerides present on distal chromosome 4 and central chromosome 7 (Fig. 3). For triglycerides and glucose these peaks are prominent but do not achieve genome-wide significance. For insulin, the peaks are discernable but not prominent. However, the coincidence of peaks for all three traits suggested the presence of genetic variation at these loci that contributes to the biological modulation of each phenotype. We used covariate analysis to provide evidence of common genetic loci underlying the apparently coincident QTLs for body weight and plasma levels of insulin, glucose, and triglycerides. The results suggested that glucose and triglycerides were strongly interdependent, while the insulin and body weight phenotypes showed no relationship to the other two traits. Based on this, we created an eigenvector reflecting the combined triglyceride and glucose phenotypes using the "moduleEigengenes" function in the WGCNA package (13). The LOD-plot mapping data for the combined glu-

Table 2. *Locations for diabetes-related QTL in BXDF2 db/db mouse cross*

Phenotype	Age, wk	Chromosome	LOD	*P* Value	Peak Position, bp	1.5 LOD Support Interval, bp		Higher Allele
Body fat, %	12	11	9.81	4.23E-05	102,463,995	101,812,775	106,069,315	DBA
Body weight, g	12	7	6.13	6.37E-05	39,311,621	30,434,146	59,479,351	DBA
Body weight, g	12	12	4.82	1.07E-01	87,049,428	63,880,020	92,357,757	DBA
Gonadal fat, g	12	7	5.91	9.13E-04	39,311,621	30,434,146	66,321,841	DBA
Gonadal fat, g	12	10	9.76	9.88E-06	79,594,580	70,812,220	86,902,245	DBA
GluTri Eigenvector	12	4	5.16		127,166,519	119,874,699	141,459,187	
GluTri Eigenvector	12	7	5.83		75,551,627	66,891,086	97,767,679	
Leptin	5	6	6.20	9.57E-06	36,710,046	17,535,219	64,395,902	DBA
Leptin	12	6	8.58	1.98E-07	24,023,461	17,311,494	47,688,539	DBA
Leptin	12	19	5.22	1.51E-05	31,095	31,095	36,468,044	DBA
Liver fat, g	12	12	5.88	3.91E-03	67,359,386	15,208,531	98,588,121	DBA
Mesenteric fat, g	12	19	7.02	4.49E-08	24,385,880	31,095	31,819,037	DBA
Subcutaneous fat, g	12	9	5.11	5.75E-04	123,402,293	121,716,696	126,414,413	B6

Quantitative trait locus (QTL) logarithm of the odds (LOD) scores were calculated using R/qtl and sex as an interactive covariate. The table includes QTLs that reached genome-wide significance based on 1,000X permutations of the genotype data and that showed significant ($P < 0.05$) differences in the trait between F2 animals homozygous for the B6 allele compared with animals homozygous for the DBA allele of the peak single nucleotide polymorphism (SNP). The "higher allele" indicates which strain-genotype has the higher trait value, and "*P* value" indicates the significance of that difference. "Peak position" reflects the base-pair location of the SNP having the strongest linkage, while "LOD" is the LOD score for that SNP. "1.5 LOD Support Interval" indicates the positions of SNPs bracketing the LOD peak where the LOD score drops by 1.5 LOD from the peak SNP. The full set of diabetes-related QTLs are given in Supplemental Table ST1.
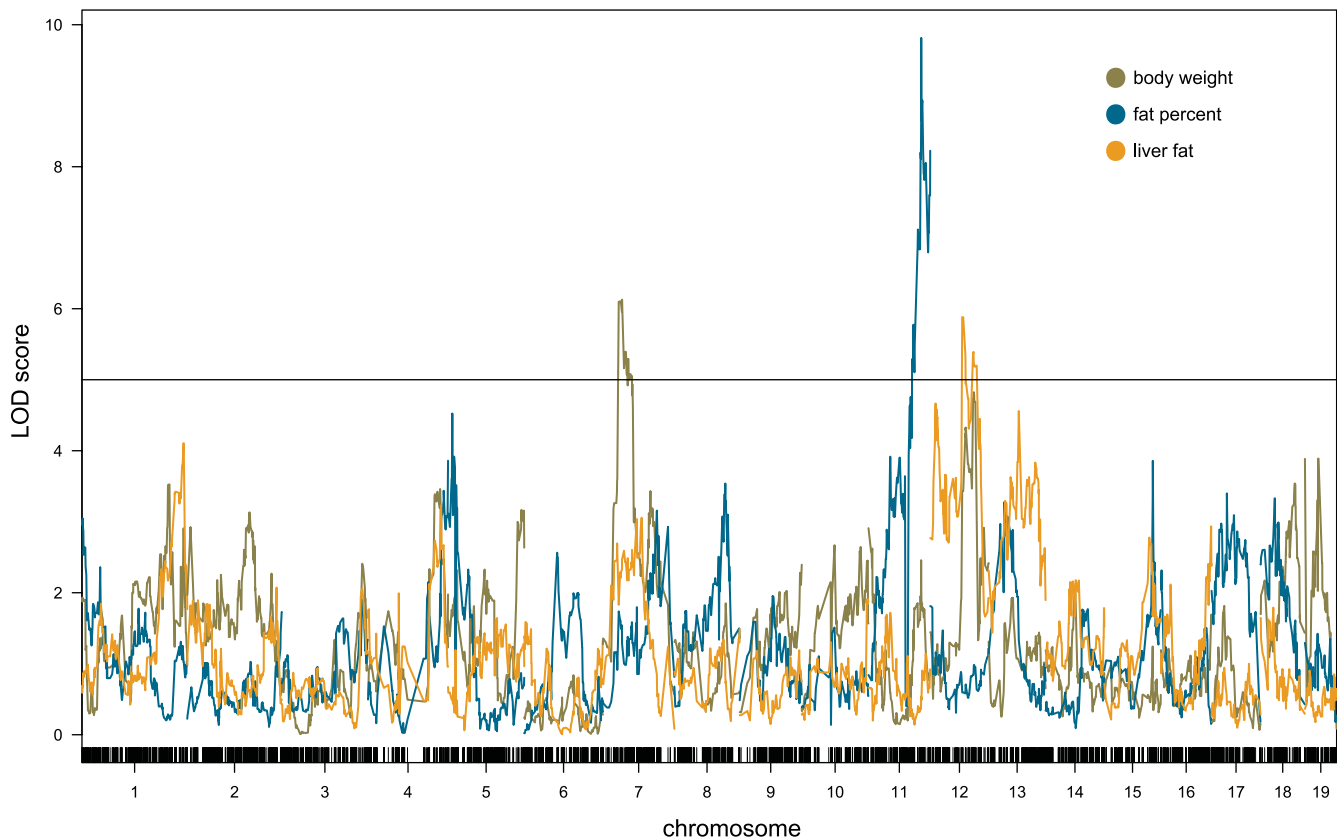
Fig. 2. Quantitative trait loci (QTL) at 12 wk for body weight (g) (tan), adiposity (blue), and liver fat (%) (orange) in F2 mice with sex as an interactive covariate. The logarithm of the odds (LOD) score for genome-wide significance as determined by $n = 1,000$ permutations of the genotype data is indicated by the horizontal line ($P = 0.05$).

cose-triglyceride eigenvector reached genome-wide significance for both the Chr. 4 (LOD = 5.2) and Chr. 7 (LOD = 5.8) peaks (Fig. 3, *A* and *B*). Thus, the eigenvector approach enhanced our ability to map these key diabetes-associated traits that individually failed to achieve genome-wide significance. Other investigators have successfully used a similar approach [e.g., Mahler et al. (16) and Lan et al. (12)], although the enhancement in LOD scores are modest. Limitations of the procedure are discussed in Ref. 16.

Systems genetics analyses of the F2 mice. Because we and others have previously observed a correlation between diabetes susceptibility and hepatic lipid accumulation (5, 11), we carried out comprehensive gene expression analysis in liver samples from the F2 mice for network analysis and correlation mapping of gene expression and clinical traits. Altogether, expression arrays were performed on total RNA isolated from livers of 221 F2 mice at 5 wk of age and 218 F2 mice at 12 wk of age.

To look for clusters of genes regulated by common site, we plotted the location of all eQTLs with LOD >4.3 relative to the location of the encoded gene in the 5 wk (Fig. 4*A*) and 12 wk cohorts (Fig. 4*B*). Local or cis-eQTL are observed along the diagonal. These represent genes that show differential expression among F2 mice due to genetic variation between the parental strains in, or near, the structural gene. Points that are off-diagonal represent genes regulated by genetic loci distant from the structural gene (trans-eQTLs). For both cohorts, there are clusters of trans-eQTLs on chromosome 12 (represented by

the dense vertical column of eQTLs on that chromosome) consistent with the mapped location for the "greenyellow" module discussed above. Similarly, there is a trans-eQTL cluster on distal chromosome 4, consistent with the insulin-glucose-triglyceride eigenvector-QTL. Thus, trans-eQTLs show a nonrandom distribution, suggesting a common regulatory factor that underlies expression variation for multiple genes in the cluster. Moreover, the trans-eQTL clusters appear to be correlated with complex phenotypes that map to the same location.

Another use of this expression data was to identify candidate genes for the diabetes-associated QTLs above. For instance, the peak marker for the chromosome 4 eigenvector QTL occurred at 127 Mb. To identify possible genetic variants that might underlie this QTL we looked for cis-eQTLs in the liver expression profiles that coincided with the eigenvector QTL. For this, we selected genes with significant (LOD >4.2) colocalizing eQTL that also lie within 10 Mb of the eigenvector QTL (cis-eQTL). These genes and associated eQTLs are listed in Table 3. Significantly, the gene list for the chromosome 4 (Table 3*A*) eigenvector QTL includes the zinc finger domain transcription factor Zfp69 that has previously been associated with the obesity-induced diabetes locus Nidd/SJL (20). We carried out a similar search for cis-eQTLs that colocalize with the chromosome 7 eigenvector QTL (Table 3*B*). While there are many strong cis-eQTLs in the region, a literature search for the corresponding genes failed to reveal any previous association with diabetes. Nor are any Type 2
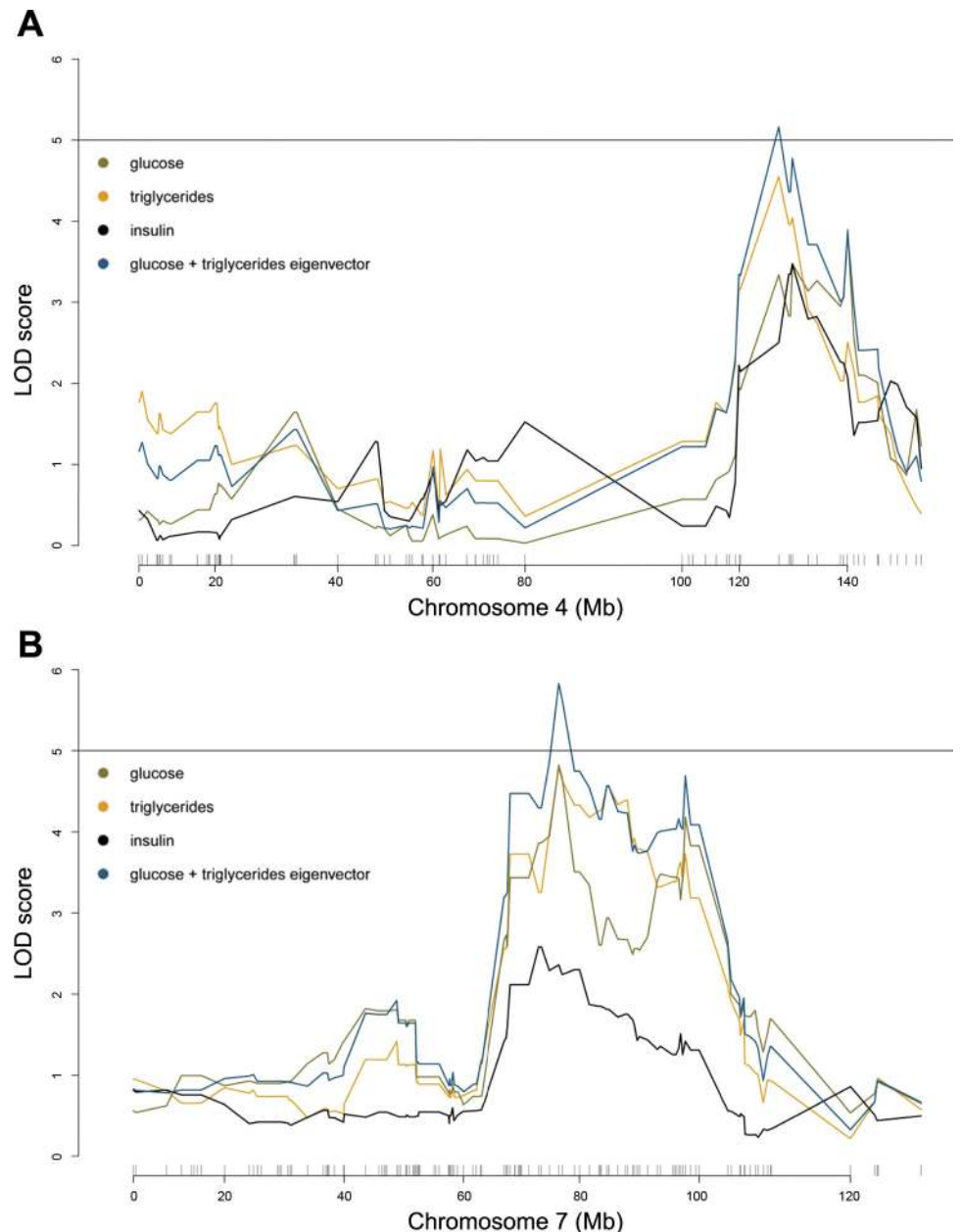
Fig. 3. QTL in the 12 wk cohort (with sex as an interactive covariate) on chromosomes 4 (*A*) and 7 (*B*) for plasma insulin (uIU/g) (black), plasma glucose (mg/dl) (tan), and plasma triglycerides (mg/dl) (orange) of F2 mice. The QTL for the eigenvector for combined traits of glucose and triglycerides (blue) has peaks that reach genome-wide significance. Full genome LOD plots for both 5 wk and 12 wk cohorts are shown for each trait in Supplemental Fig. SF3. The LOD score for genome-wide significance as determined by $n = 1,000$ permutations of the genotype data is indicated by the horizontal line $P = 0.05$.

diabetes-related QTL identified for this region (3). This suggests that the Chr. 7 QTL represents a novel diabetes locus that contributes to plasma insulin, glucose, and triglyceride levels in the obese mouse. Supplemental Table ST3 lists genes with trans-eQTL located near the eigenvector QTLs on Chr. 4 and Chr. 7. These are candidates for genes whose expression is regulated by the genetic variant underlying the eigenvector QTL.

Some QTL may not arise from differences in mRNA levels. For these, identity-by-descent (IBD) approaches may help to narrow the intervals for the underlying genes (See Supplemental Fig. SF5). For instance, much of the glucose-triglyceride eigenvector QTL on chromosome 4 (118–135 Mbp; 95% confidence interval) occurs over non-IBD regions including the gene locus for Zfp69 at 120 Mbp (Supplemental Fig. SF5B). By contrast, QTL regions at 129.8–132 and 134.2–135 Mbp

appear to be IBD between B6 and *DB*A, suggesting that gene variations in these regions are unlikely to contribute to this QTL.

To identify candidate genes that might underlie QTL for the other diabetes-related traits, we looked for genes with strong cis-acting eQTL that colocalized with the trait QTL. Some of these are shown in Table 4. Among these genes are numerous zinc-finger proteins and other genes likely to play roles in regulating gene transcription. Also prominent are several examples of genes with roles in lipid metabolism including Apolipoproteins A-II and C-II.

Gene coexpression modules in liver associated with clinical traits. Gene expression data were further analyzed using WGCNA (13) to identify clusters of genes showing highly correlated expression profiles. Of the 23,000 genes on the arrays, ~15,000 had sufficient signal quality to be used for
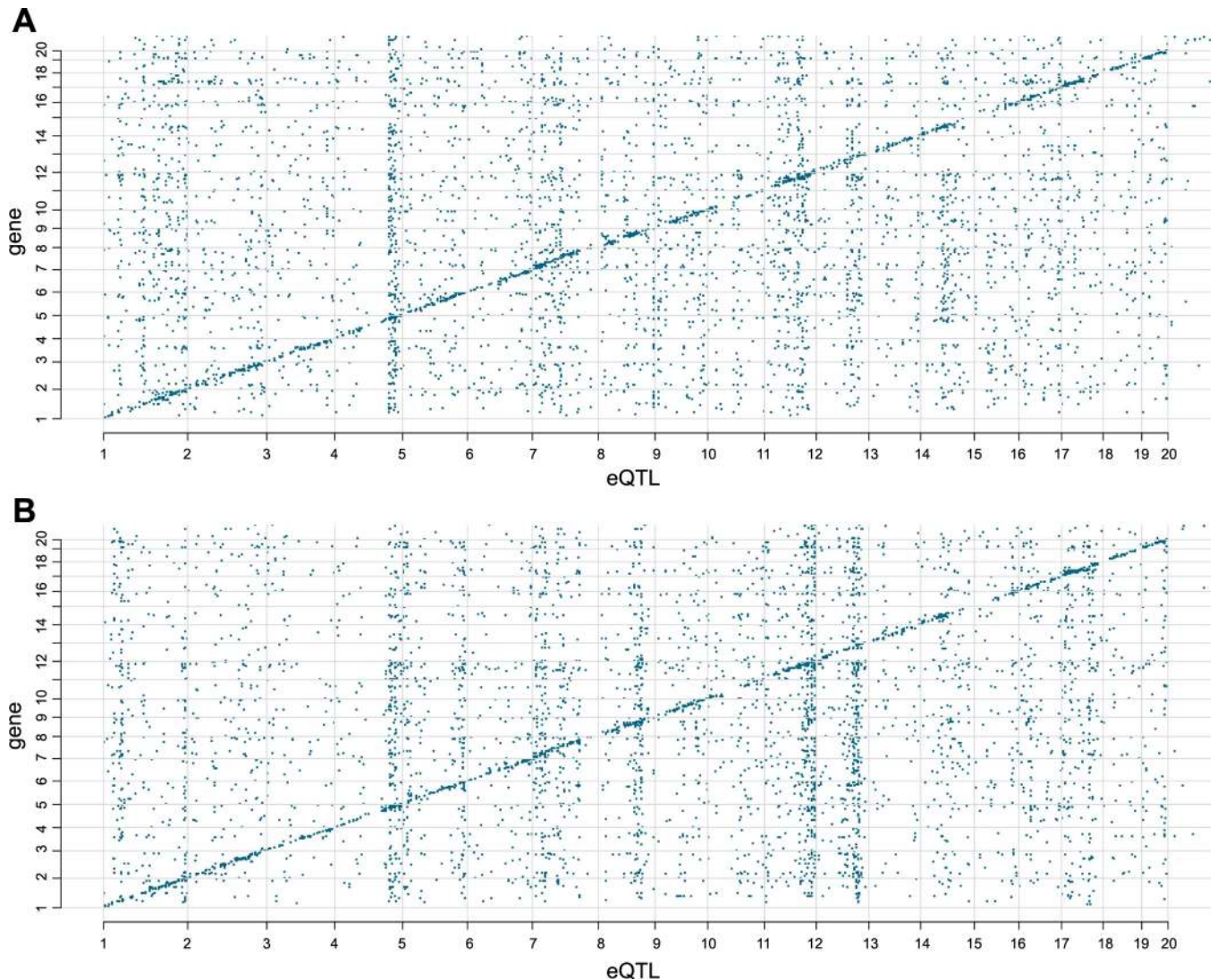
Fig. 4. Distribution of significant expression quantitative trait loci (eQTLs) for animals at wk 5 (*A*) and wk 12 (*B*). This plot shows the genomic position of eQTL (LOD >4.3) along the *x*-axis and position of coding sequence for the corresponding transcripts on the *y*-axis. Axis numbers indicate chromosome ID. Points for local eQTLs are on the diagonal while off-axis points correspond to distal- or trans-eQTLs. Regions with higher concentrations or "hotspots" of trans-eQTLs appear as vertical clusters such as those evident on distal chromosome 4 in both cohorts.

WGCNA. This analysis was carried out separately for males and females at 5 wk and at 12 wk of age and, for each population, identified ~50 distinct coexpression modules containing 50 to several hundred genes. Module assignments for all groups are detailed in Supplemental Table ST4.

Correlation of the each gene module in male F2 *db/db* mice at 12 wk of age with diabetes-related phenotypes, including plasma glucose, insulin, and triglycerides, body weight, liver fat, and size of fat depots, is shown in Fig. 5*A*. The individual coexpression modules are identified by an arbitrary color on the left of the figure and the specific clinical traits are identified across the bottom of the figure. Positive correlation between a trait and a gene module is shown in red, and negative correlation is shown in green, with stronger correlations indicated by increased color intensity. In the online version of the figure, the correlation coefficient and *P* value is shown for each module and trait. Similar sets of correlated coexpression modules were observed with 5 wk males and with female mice (Supplemental Fig. SF4).

A few of the modules show strong correlation across numerous relevant phenotypes. For instance the greenyellow module in 12 wk males shows strong positive correlation with plasma cholesterol and with various measures of adiposity as well as with fat in the liver. Weaker positive correlation was seen with plasma insulin. The same module showed weak negative correlation with plasma glucose and triglycerides (Fig. 5*B*). These correlations suggest that genes in this module contribute to, or are driven by, the diabetic state.

To better characterize these common pathways, we carried out functional annotation clustering analysis using the DAVID platform. Consistent with its correlation to adiposity in the liver and other diabetes-related traits, the greenyellow module for 12 wk male liver shows enrichment for lipid biosynthetic processes, glucose catabolism, and steroid biosynthesis and metabolism (Table 5). Genes for all modules are included in Supplementary Table ST1.

Table 3. *Genes located within 10 Mb of peak SNP for eigenvector glucose-triglyceride QTL with local eQTL LOD >3.0*

SNP			Probe		Gene		Significance		
ID	Chr.	Position	ID	Chr.	EntrezID	Symbol	LOD	P Value	FDR
A: Chromosome 4									
rs13477991	4	132735018	1451212_at	4	70012	Ccdc21	61.34	NA	NA
rs13477972	4	127166519	1438329_at	4	384059	Tlr12	52.95	NA	NA
rs13477991	4	132735018	1422591_at	4	27224	Tceb3	44.69	0.000	0.000
rs13477940	4	117712824	1425201_a_at	4	68180	Hyi	43.05	0.000	0.000
rs13477991	4	132735018	1447703_x_at	4	68040	Zfp593	42.95	NA	NA
rs6319522	4	119591509	1458274_at	4	381549	Zfp69	37.20	NA	NA
rs13477991	4	132735018	1450240_a_at	4	269589	Sytl1	24.52	NA	NA
rs13477991	4	132735018	1423305_at	4	56219	Extl1	23.61	0.000	0.000
rs13477999	4	134408830	1424531_a_at	4	21401	Tcea3	23.51	0.000	0.000
rs13477999	4	134408830	1435312_at	4	71904	Paqr7	21.28	NA	NA
rs6355453	4	129755364	1452937_s_at	4	66264	Ccdc28b	20.84	NA	NA
rs13477991	4	132735018	1439590_at	4	381560	Xkr8	20.20	NA	NA
rs13477999	4	134408830	1417300_at	4	100340	Smpdl3b	19.39	0.000	0.000
rs13477972	4	127166519	1420612_s_at	4	19244	Ptp4a2	19.29	0.000	0.000
rs13477980	4	129400117	1436905_x_at	4	16792	Laptm5	18.63	NA	NA
rs13477991	4	132735018	1460218_at	4	23833	Cd52	16.63	NA	NA
rs13477999	4	134408830	1419471_a_at	4	18221	Nudc	14.93	0.000	0.000
rs13477943	4	118875991	1456912_at	4	195522	Zfp691	14.83	NA	NA
rs13477999	4	134408830	1420059_at	4	76824	2410166I05Rik	12.41	0.000	0.000
rs6355453	4	129755364	1450476_at	4	12802	Cnr2	12.12	NA	NA
rs13477999	4	134408830	1417797_a_at	4	69073	1810019J16Rik	10.58	0.000	0.000
rs13477972	4	127166519	1455587_at	4	230766	BC030183	10.57	NA	NA
rs13477938	4	117178080	1452752_at	4	56401	Lepre1	9.92	NA	NA
rs13477980	4	129400117	1438683_at	4	242687	Wasf2	9.26	NA	NA
rs6355453	4	129755364	1435904_at	4	214150	Eif2c3	9.11	NA	NA
rs13477980	4	129400117	1436482_a_at	4	20970	Sdc3	8.50	NA	NA
rs6319522	4	119591509	1417636_at	4	14664	Slc6a9	7.58	0.001	0.019
rs13477980	4	129400117	1416911_a_at	4	68050	6330407G11Rik	6.61	0.000	0.000
rs13477980	4	129400117	1437226_x_at	4	17357	Marcksl1	6.55	NA	NA
rs13477980	4	129400117	1448246_at	4	433759	Hdac1	6.52	NA	NA
rs6355453	4	129755364	1424956_at	4	230793	Ahdc1	6.46	0.000	0.000
rs6319522	4	119591509	1438226_at	4	230696	AU022252	6.30	NA	NA
rs13477972	4	127166519	1448644_at	4	67898	Pef1	6.04	NA	NA
rs13477991	4	132735018	1417615_a_at	4	67025	Rpl11	5.98	0.002	0.035
rs13477972	4	127166519	1433748_at	4	503610	Zdhhc18	5.42	NA	NA
rs6319522	4	119591509	1416852_a_at	4	26562	Ncdn	5.35	0.001	0.019
rs13477972	4	127166519	1433775_at	4	97130	C77080	5.28	NA	NA
rs13477980	4	129400117	1425139_at	4	230784	Sesn2	4.25	0.016	0.164
rs13477980	4	129400117	1424140_at	4	74246	Gale	4.15	0.009	0.111
rs13477999	4	134408830	1416630_at	4	15903	Id3	4.00	0.015	0.156
rs13477980	4	129400117	1439533_at	4	68233	1700125D06Rik	3.96	NA	NA
rs13477980	4	129400117	1456054_a_at	4	80912	Pum1	3.74	NA	NA
rs13477999	4	134408830	1417397_at	4	20544	Slc9a1	3.68	0.075	0.449
rs6319522	4	119591509	1450782_at	4	22417	Wnt4	3.66	NA	NA
rs13477999	4	134408830	1436898_at	4	71514	Sfpq	3.62	NA	NA
rs13477972	4	127166519	1425054_a_at	4	76799	2510006D16Rik	3.06	0.059	0.379
B: Chromosome 7									
rs13479264	7	48196494	1438477_a_at	7	73724	Mcee	33.29	NA	NA
rs13479319	7	62351357	1422554_at	7	66647	Ndnl2	20.69	0.000	0.000
rs13479304	7	58497977	1434738_at	7	272396	Tarsl2	15.25	NA	NA
rs13479366	7	73977283	1425192_at	7	207952	Klhl25	12.54	0.000	0.000
rs6164834	7	57884429	1432096_at	7	20646	Snrpn	8.73	0.000	0.000
rs13479416	7	88013305	1430404_at	7	75771	4833416J08Rik	6.08	0.001	0.019
rs13479336	7	66321841	1432543_a_at	7	50794	Klf13	3.92	0.051	0.345
rs13479132	7	11881289	1457707_at	7	244049	Mctp2	3.92	NA	NA
rs13479360	7	72159921	1452288_at	7	233315	Mtmr10	3.69	NA	NA
rs13479300	7	57129185	1434273_at	7	100038347	BC034069	3.57	NA	NA
rs13479416	7	88013305	1434093_at	7	75099	Lysmd4	3.53	NA	NA
rs13479132	7	11881289	1454088_at	7	76673	5330411O13Rik	3.48	NA	NA
rs6296859	7	61613731	1445040_at	7	52419	D7Ertd783e	3.28	NA	NA

eQTL, expression QTL; FDR, false discovery rate; NA, not available.

Table 4. *Cis-acting eQTL that colocalize with diabetes-related clinical trait QTL*

| F2 Mouse Cohort | Clinical Trait | cQTL | | | Colocalizing eQTL | | Functional Pathways |
		Chr.	Position	LOD	Entrez Gene ID	LOD	Colocalizing Gene
Week 5	fat %	1	41626841	3.94			Thioredoxin fold
					68833	10.17	phosducin-like 3
					98258	21.66	thioredoxin domain containing 9
	body weight	1	175559118	3.90			Fatty acid metabolic process
					619326	33.29	RIKEN cDNA 9130409I23 gene
					11807	4.28	apolipoprotein A-II
					13244	7.02	degenerative spermatocyte homolog 1 (*Drosophila*)
							Lipoprotein metabolic process
					67292	7.46	phosphatidylinositol glycan anchor biosynthesis, class C
					67556	29.99	phosphatidylinositol glycan anchor biosynthesis, class M
							Oxidation reduction
					226646	28.19	NADH dehydrogenase (ubiquinone) Fe-S protein 2
					56752	34.35	aldehyde dehydrogenase 9, subfamily A1
					55990	50.73	flavin containing monooxygenase 2
					69051	23.62	pyrroline-5-carboxylate reductase family, member 2
	insulin	5	133710611	4.15	22697	10.63	zinc finger and SCAN domain containing 21
					56314	4.18	zinc finger protein 113
					72154	11.80	zinc finger protein 157
					269704	2.79	zinc finger protein 664
					24135	7.56	zinc finger protein 68
Week12	body weight	12	87049428	4.82			Mitochondrion
					66272	4.32	synaptojanin 2 binding protein
					217830	49.39	RIKEN cDNA 9030617O03 gene
					15512	3.35	heat shock protein 2
					74316	4.96	iron-sulfur cluster assembly 2 homolog (*S. cerevisiae*)
	body weight	7	39311621	6.13	11813	4.51	apolipoprotein C-II
	fat percent	11	102463995	9.81	19300	2.81	ATP-binding cassette, subfamily D (ALD), member 4
	triglycerides	4	127166519	4.55	381549	37.20	zinc finger protein 69
					195522	14.83	zinc finger protein 691
					21401	23.51	transcription elongation factor A (SII), 3
					68040	42.95	zinc finger protein 593
	liver fat	12	67359386	5.88	110095	9.33	liver glycogen phosphorylase
					12192	19.76	zinc finger protein 36, C3H type-like 1
					207965	94.22	hypothetical protein LOC100048447; predicted gene 71
					68055	8.45	ATP synthase, H^+ transporting, mitochondrial F0 complex, subunits
					73991	6.97	predicted gene 3086; atlastin GTPase 1; similar to Spg3a protein

cQTL, clinical trait QTL.

Genetic drivers of coexpression modules. To identify genetic loci associated with the diabetes- and obesity-related greenyellow module for 12 wk male liver, we carried out mapping analysis for the eigenvector of expression values for module genes. Figure 6 shows the resulting LOD plot. For this module, there is a single strong QTL on distal chromosome 12 that appears to coincide with above noted QTL peaks for body weight, liver fat, and plasma insulin. Thus, this locus appears to contain genetic variation that is a primary regulator of expression for genes in the greenyellow module and that is associated with modulation of several diabetes-related traits that also map to this locus. These results suggest multiple genes relevant to this complex set of diabetes-related phenotypes are regulated by a common genetic variant on chromosome 12.

DISCUSSION

Among ~220 F2 progeny that were phenotyped at 5 wk of age and a similar number that were phenotyped at 12 wk, we identified significant and suggestive QTLs for diabetes-related traits of body weight, adiposity, fat pad weights, liver lipid content, and plasma levels of glucose, insulin, triglycerides, and leptin (Table 2). For each QTL in this table, we indicate the trait affected, as well as the chromosomal position and QTL strength. In addition, we show which allele, B6 or *DB*A, gives the higher trait value ("higher allele"). Strikingly, it is the *DB*A allele that engenders higher values for all the QTLs listed except the one for subcutaneous fat on Chromosome 9. This suggests either that *DB*A contributes many alleles leading to enhanced diabetes susceptibility but each with a small effect or that there are interactions between the *DB*A and B6 alleles responsible for the enhanced susceptibility.

That pronounced diabetic phenotypes were seen in a high fraction of F2 animals, even in the 5 wk cohort, is consistent with the existence of multiple strong predisposing loci in the parental strains. For instance, approximately a quarter of F2 males showed plasma glucose levels >300 mg/dl at 5 wk of age. However, the hypothesis of a few strong underlying loci is difficult to reconcile with the paucity of strong diabetes-related

A: module−trait relationships for male week 12

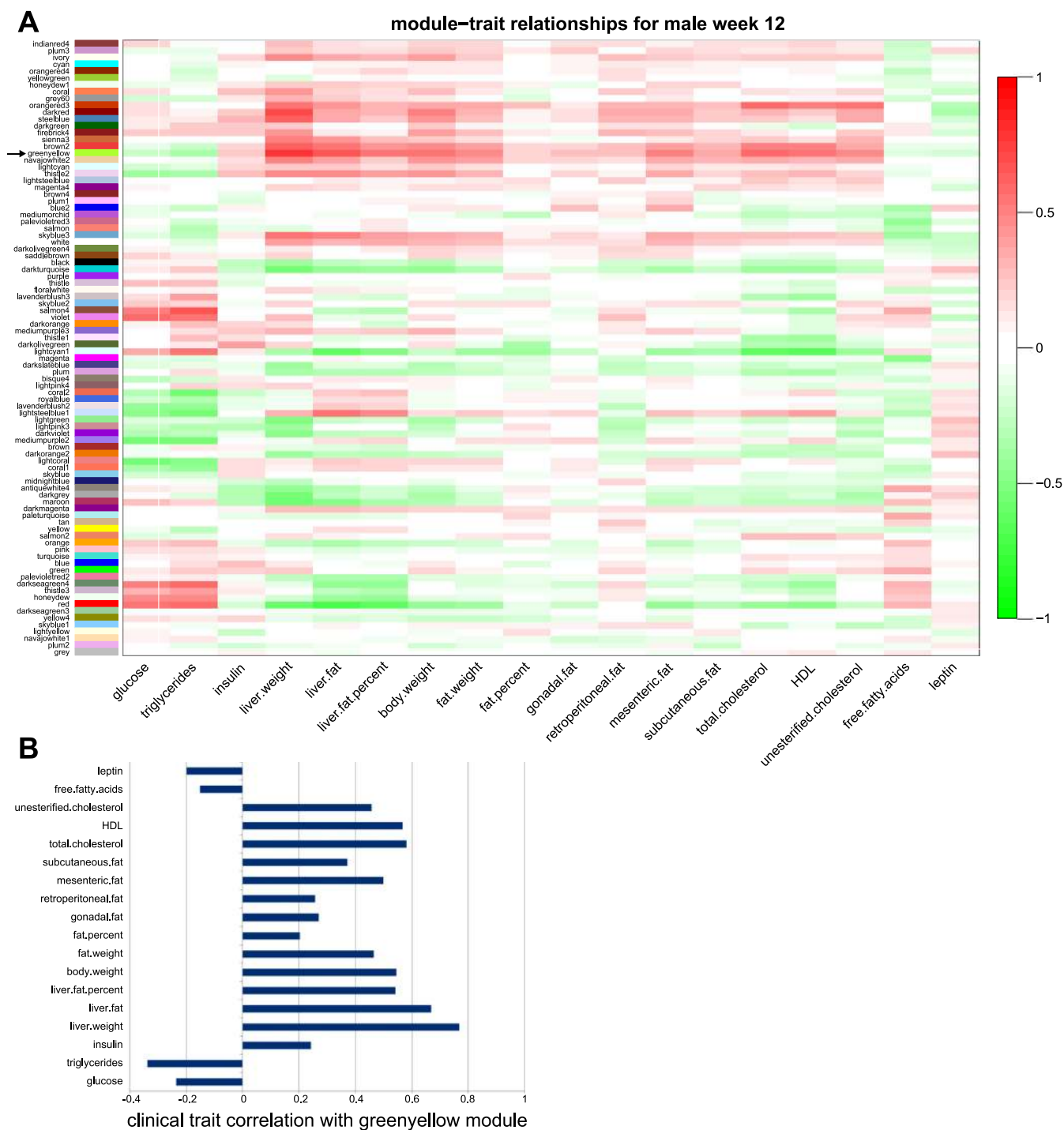**B**: clinical trait correlation with greenyellow module

Fig. 5. *A*: weighted gene coexpression network analysis of livers of 12 wk male F2 mice. The network analysis identified ~90 gene modules with highly correlated gene expression across the population of ~250 F2 mice. These modules are identified by arbitrary colors, indicated along the *y*-axis. Clinical traits including plasma insulin, glucose, and measures of adiposity are listed across the *x*-axis. Correlation of each module with these traits is indicated by the intensity of red (positive correlation) or green (negative correlation) within the table and a scale relating color intensity to correlation coefficient is shown at the *right*. The "greenyellow" module (arrow) showed negative correlation with glucose and triglycerides and positive correlation with insulin and multiple measures of obesity and fatty liver. Module-trait correlation tables for all groups (male and female, 5 wk and 12 wk) are shown in supplementary data. (Supplemental Fig. SF4). For these figures, the actual correlation coefficient and associated *P* value are shown within each color block. *B*: correlation of diabetes-related traits in 12 wk F2 males with greenyellow coexpression module in liver. The statistical threshold significance for modules correlated with traits was $R > 0.4$ ($P < 1e-6$). The greenyellow module was positively correlated with unesterified cholesterol, HDL, total cholesterol, mesenteric fat, fat weight, body weight, liver fat percent, liver fat, and liver weight.

Table 5. *DAVID GO analysis of greenyellow expression module in 12 wk male liver*

Category	Term	P Value
GOTERM_BP_FAT	GO:0008610~lipid biosynthetic process	2.48E-05
KEGG_PATHWAY	mmu00983:drug metabolism	7.97E-05
SP_PIR_KEYWORDS	lipid synthesis	1.30E-04
GOTERM_BP_FAT	GO:0034655~nucleobase, nucleoside, nucleotide, and nucleic acid catabolic process	6.36E-04
GOTERM_BP_FAT	GO:0034656~nucleobase, nucleoside, and nucleotide catabolic process	6.36E-04
SP_PIR_KEYWORDS	transferase	8.16E-04
GOTERM_BP_FAT	GO:0019318~hexose metabolic process	1.05E-03
GOTERM_BP_FAT	GO:0009119~ribonucleoside metabolic process	1.26E-03
GOTERM_BP_FAT	GO:0044270~nitrogen compound catabolic process	1.42E-03
GOTERM_BP_FAT	GO:0006694~steroid biosynthetic process	1.44E-03
SP_PIR_KEYWORDS	steroid biosynthesis	1.85E-03
GOTERM_BP_FAT	GO:0005996~monosaccharide metabolic process	2.00E-03
KEGG_PATHWAY	mmu00480:glutathione metabolism	2.78E-03
GOTERM_BP_FAT	GO:0008202~steroid metabolic process	3.24E-03
GOTERM_BP_FAT	GO:0009116~nucleoside metabolic process	3.29E-03
PIR_SUPERFAMILY	PIRSF000503:glutathione transferase	4.06E-03
GOTERM_BP_FAT	GO:0006007~glucose catabolic process	4.29E-03
GOTERM_BP_FAT	GO:0019320~hexose catabolic process	4.29E-03
GOTERM_BP_FAT	GO:0046365~monosaccharide catabolic process	4.66E-03

Selected for enrichments with *P* value <0.005.

QTLs. Alternatively, the presence of many weaker diabetogenic loci could also account for the enhanced susceptibility observed in F2 animals. And, if no single locus accounts for a large fraction of the variance, this would also explain the absence of strong QTL. This hypothesis, where many diabetogenic loci are scattered over the whole *DB*A genome is also consistent with the high frequency of diabetes that we observed in F2 *db/db* animals.

Recently, Yang et al. (25) have provided evidence that the genomes of modern inbred strains have common and overlapping ancient origins from wild-derived strains, most frequently, *Mus musculus domesticus*. This was based on a set of indicator SNPs found in wild-caught examples of the musculus, domesticus, and castaneus substrains. Despite these common origins, inbred strains such as B6 and *DB*A/2 carry hundreds of thousands of additional SNPs that distinguish the individual strains and presumably underlie their many phenotypic differences. We have previously characterized the enhanced diabetes susceptibility observed in C57BLKS (BKS) *db/db* mice compared with B6 *db/db* mice (5) and have also shown that BKS is genetically identical to B6 >70% of the genome, with most of the rest due to large contaminating blocks recently derived from *DB*A/2 plus a small contribution from an uncharacterized strain (6). This initially led us to hypothesize that the greater diabetes susceptibility seen in BKS compared with B6 mice most likely derives from the blocks of *DB*A/2 DNA carried by BKS. Surprisingly however, most of the QTL in Table 2 occur in regions of the genome that BKS inherits from B6. The chromosome 4 eigenvector QTL coincides with a region that derives from the unidentified mouse strain in BKS, and, while the 95% confidence interval for the liver fat QTL on chromosome 12 overlaps a *DB*A region in BKS, the QTL peak is in a B6-derived region. Only the QTL on chromosome 9 (at 123 Mb) for subcutaneous fat is located in a region corresponding to the *DB*A-derived block that spans from 116 to 126 Mb (6). Moreover, it is the B6 allele that contributes to higher subcutaneous fat in the present cross, suggesting that the corresponding locus in the *DB*A-derived DNA block of BKS is unlikely to enhance BKS diabetes susceptibility. These results are in

stark contrast to our expectations that we would observe a few strong diabetogenic loci in the present cross and that some, if not most, of these would be located in regions of BKS that derive from *DB*A, thereby accounting for increased diabetes susceptibility of that strain. Ultimately, our failure to detect strong QTLs for diabetic phenotypes in the BXD *db/db* F2 population where such phenotypes are highly prevalent suggests either a very complex set of contributing genetic loci or the presence of nongenetic factors that significantly impact the phenotypes. In either case, there appears to be insufficient power in the present data to infer likely genetic regions of BKS that underlie the sensitivity of that strain.

The observed weakness of QTLs for insulin and glucose may also partially result from the fact that neither insulin nor glucose levels increase in direct proportion to disease progression: Early in classic Type 2 diabetes, insulin levels rise to compensate insulin resistance that is induced or amplified by severe obesity. Later, with failure of the pancreatic islets, insulin levels fall. Similarly, glucose levels may be repressed until late in disease when insulin levels decrease. Moreover, altered levels of glucose, insulin, and triglycerides may also reflect maladaptive changes in metabolism driven by epigenetic changes such as altered DNA methylation, histone modifications, or microRNA levels associated with the diabetic state (reviewed by Villeneuve and Natarajan in Ref. 24a; 24). The impact of these factors on the ability to detect QTL can only be compounded by the strong sexual dimorphism that was observed and the surprisingly early disease progression. Finally, while we are unaware of any environmental factor that would have induced wholesale diabetes in this population, it is possible that some additional unknown, nongenetic factor underlies the unanticipated elevation in diabetic response among the F2 animals.

The zinc finger domain transcription factor Zfp69 was recently associated with the obesity-induced diabetes locus Nidd/ SJL by Scherneck et al. (20). Working with data from an (NZO × SJL) N2 backcross the authors demonstrated QTLs for body weight and blood glucose on Chr. 4 near the gene for Zfp69. Subsequent work with interval-specific congenic strains
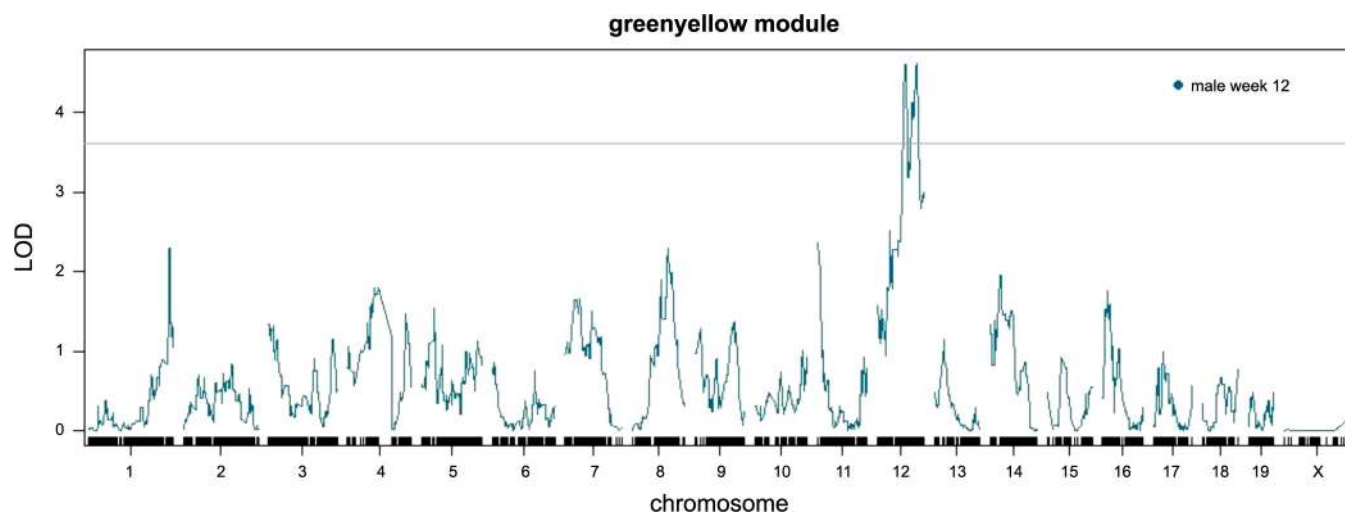
Fig. 6. Mapping of the eigenvector for genes in the greenyellow module for 12 wk male F2 animals with sex as an interactive covariate. The LOD score for the module eigenvector is plotted as a function of position across the genome. Genome-wide significance was determined by measuring the maximum LOD score achieved by an equal number of randomly selected genes. The horizontal line represents the LOD score exceeded by <5% of 1,000 such randomly selected gene sets ($P < 0.05$).

localized the effect to a 2 Mb region containing 10 genes, of which Zfp69 was shown to be disrupted by insertion of retro-transposon sequences in C57BL/6 and NZO strains, protecting them from the diabetogenic effects of the nondisrupted allele found in other strains such as SJL, NON, and NZB. In obese mice lacking the gene for leptin, the presence of nondisrupted Zfp69 strongly increases plasma glucose and triglycerides. Correspondingly, the human ortholog is more highly expressed in human diabetics than healthy controls, and this correlation is significant only in obese but not lean individuals (20, 21). Thus, the presence of the disrupted allele in C57BL/6 mice and the strong cis-eQTL for expression of this gene among F2 mice in our cross is consistent with the hypothesis that disruption of Zfp69 underlies the Chr. 4 QTL associated with glucose and triglycerides. In fact, a gene effect calculation showed that Zfp69 explains 19.5% of the variance.

Previously, Stoehr et al. (22) identified body weight QTLs in an F2 cross between C57BL/6J and BTBR where leptin signaling was disrupted by selecting for mice deficient in leptin (ob/ob). They named these loci as "modifiers of obesity," *Moo1* on chromosome 2 and *Moo2* on chromosome 13, each explaining ~13–14% of the variance in body weight at 10 wk of age. The dominant alleles for both Moo1 and Moo2 were contributed by BTBR, thus explaining the absence of these loci among the body weight QTLs in the present cross. Similarly, the *DB*A-determined body weight QTLs in our cross were not observed in the cross between C57BL/6 and BTBR.

Analysis of gene expression in the BTBR × B6 F2 *ob/ob* cross also detected an association between lipogenic gene expression in liver and diabetes-related traits (12). In particular, they mapped the "diabetes mRNA cluster" for the combined expression of *SCD1*, *FAS*, *GPAT*, and *PEPCK* to loci on chromosomes 2, 7, and 9 that were associated with plasma insulin levels. This is in contrast to the greenyellow module in the present study that is enriched in lipogenic genes and that is strongly correlated with diabetes-related traits but which maps strongly only to chromosome 12. This suggests that while hepatic lipogenesis has significant correlation to diabetes in

both crosses, the genetic variations that modulate associated lipogenic gene expression are different between BTBR and *DB*A/2. And more generally, an improved understanding of the relationship between diabetes and hepatic lipogenesis more likely lies in understanding the associated metabolic pathways than in the multitude of specific mutations that have the potential to perturb those pathways.

A related aspect of obesity-induced diabetes between mouse strains that is of particular interest is the relationship between liver fat content and the diabetic phenotype. In our previous studies of BKS and B6 mice, we observed an inverse correlation between liver fat and diabetes (5). A similar observation was made by Lan et al. (11) in their study of BTBR and B6 carrying the *ob/ob* mutation. In our present study, we also observed a striking correlation between liver fat content and plasma insulin (positive) and plasma glucose (negative), consistent with the concept that uptake of fat by liver helps protect against the development of diabetes. Lan et al. (11) proposed that this protective effect might arise from increased hepatic lipogenesis that diverts nutrient overload to fat deposits in the liver, nutrients that otherwise would contribute to gluconeogenesis and hyperglycemia. This argument was supported by several published experiments where activation of SREBP1 increased hepatic steatosis and decreased hyperglycemia.

Interestingly, the same argument does not seem to apply to whole body fat deposition. We found that whole body lipids were only weakly correlated with liver lipids, suggesting that there are independent underlying genetic determinants for these two phenotypes. Similarly, the relationship between whole body fat as measured by NMR and plasma glucose and insulin was less striking than the correlations seen with liver lipids. Only in young animals did we see a strong correlation of plasma glucose to body fat. This correlation was completely lost in older males although 12 wk females retained a weaker correlation. In the 5 wk male mice, there was a modest positive correlation of body fat with insulin levels that was lost in older animals. The weak association of total body lipids with these diabetic phenotypes may simply reflect the extreme obesity of

all the F2 animals due to the *db* mutation with little possible further impact due to background variations in body fat. On the other hand, lipogenesis is suppressed in adipose tissue of these obese animals (11), and there may be factors that lead tissue-specific regulation of the lipogenic pathway. For instance, there is evidence of selective binding of SREBP isoforms and co-regulatory proteins to promoters for lipid metabolic genes (1), and this binding appears dependent on metabolic factors other than SREBP protein concentration.

In humans, obesity is associated with increased risk of diabetes as well as fat deposition in liver and other organs (10). Obesity contributes to insulin resistance and, hence, increased demand for insulin, and it is hypothesized that toxicity from lipid accumulation in the pancreas directly leads to islet depletion and diabetes (10). Human genome-wide association studies have identified numerous loci contributing to diabetes, but these nevertheless explain a small part of the genetic component (27).

In the present study, we have performed genetic and systems-based analysis of a cross between C57BL/6J and *DB*A/2J mice with a *db/db* mutation. The results emphasize the genetic complexity of Type 2 diabetes and suggest a set of candidate loci, genes, and pathways that add to the body of information about the underlying mechanisms.

ACKNOWLEDGMENTS

GRANTS

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: R.C.D., E.E.S., and A.J.L. conception and design of research; R.C.D., L.W.C., Z.Z., P.W., S.Y., H.Q., and M.R. performed experiments; R.C.D., A.V.N., L.W.C., Y.Z., Z.Z., P.W., S.Y., H.Q., K.W.B., and M.P. analyzed data; R.C.D., A.V.N., Y.Z., E.E.S., K.W.B., M.P., and A.J.L. interpreted results of experiments; R.C.D., A.V.N., Y.Z., and K.W.B. prepared figures; R.C.D., A.V.N., M.P., and A.J.L. drafted manuscript; R.C.D., A.V.N., K.W.B., M.P., and A.J.L. edited and revised manuscript; R.C.D., M.P., and A.J.L. approved final version of manuscript.

REFERENCES

1. **Bennett MK, Seo YK, Datta S, Shin DJ, Osborne TF.** Selective binding of sterol regulatory element-binding protein isoforms and co-regulatory proteins to promoters for lipid metabolic genes in liver. *J Biol Chem* 283: 15628–15637, 2008.

2. **Broman KW, Wu H, Sen S, Churchill GA.** R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19: 889–890, 2003.

3. **Clee SM, Attie AD.** The genetic landscape of type 2 diabetes in mice. *Endocr Rev* 28: 48–83, 2007.

4. **Coleman DL, Hummel KP.** Symposium IV: Diabetic syndrome in animals. Influence of genetic background on the expression of mutations at the diabetes locus in the mouse II Studies on background modifiers. *Isr J Med Sci* 11: 708–713, 1975.

5. **Davis RC, Castellani LW, Hosseini M, Ben-Zeev O, Mao HZ, Weinstein MM, Jung DY, Jun JY, Kim JK, Lusis AJ, Peterfy M.** Early hepatic insulin resistance precedes the onset of diabetes in obese C57BLKS-*db/db* mice. *Diabetes* 59: 1616–1625, 2010.

6. **Davis RC, Schadt EE, Cervino AC, Peterfy M, Lusis AJ.** Ultrafine mapping of SNPs from mouse strains C57BL/6J, *DB*A/2J, and C57BLKS/J for loci contributing to diabetes and atherosclerosis susceptibility. *Diabetes* 54: 1191–1199, 2005.

7. **Hedrick CC, Castellani LW, Warden CH, Puppione DL, Lusis AJ.** Influence of mouse apolipoprotein A-II on plasma lipoproteins in transgenic mice. *J Biol Chem* 268: 20676–20682, 1993.

8. **Horvat S, Bünger L.** Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) assay for the mouse leptin receptor (Lepr(*db*)) mutation. *Lab Anim* 33: 380–384, 1999.

9. **Hummel KP, Dickie MM, Coleman DL.** Diabetes, a new mutation in the mouse. *Science* 153: 1127–1128, 1966.

10. **Kusminski CM, Shetty S, Orci L, Unger RH, Scherer PE.** Diabetes and apoptosis: lipotoxicity. *Apoptosis* 14: 1484–1495, 2009.

11. **Lan H, Rabaglia ME, Stoehr JP, Nadler ST, Schueler KL, Zou F, Yandell BS, Attie AD.** Gene expression profiles of nondiabetic and diabetic obese mice suggest a role of hepatic lipogenic capacity in diabetes susceptibility. *Diabetes* 52: 688–700, 2003.

12. **Lan H, Stoehr JP, Nadler ST, Schueler KL, Yandell BS, Attie AD.** Dimension reduction for mapping mRNA abundance as quantitative traits. *Genetics* 164: 1607–1614, 2003.

13. **Langfelder P, Horvath S.** WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9: 559, 2008.

14. **Laurie CC, Nickerson DA, Anderson AD, Weir BS, Livingston RJ, Dean MD, Smith KL, Schadt EE, Nachman MW.** Linkage disequilibrium in wild mice. *PLoS Genet* 3: e144, 2007.

15. **Leiter EH, Coleman DL, Hummel KP.** The influence of genetic background on the expression of mutations at the diabetes locus in the mouse. III. Effect of H-2 haplotype and sex. *Diabetes* 30: 1029–1034, 1981.

16. **Mahler M, Most C, Schmidtke S, Sundberg JP, Li R, Hedrich HJ, Churchill GA.** Genetics of colitis susceptibility in IL-10-deficient mice: backcross versus F2 results contrasted by principal component analysis. *Genomics* 80: 274–282, 2002.

17. **Manichaikul A, Moon JY, Sen S, Yandell BS, Broman KW.** A model selection approach for the identification of quantitative trait loci in experimental crosses, allowing epistasis. *Genetics* 181: 1077–1086, 2009.

18. **Mao HZ, Roussos ET, Peterfy M.** Genetic analysis of the diabetes-prone C57BLKS/J mouse strain reveals genetic contribution from multiple strains. *Biochim Biophys Acta* 1762: 440–446, 2006.

19. **Puppione DL, Charugundla S.** A microprecipitation technique suitable for measuring alpha-lipoprotein cholesterol. *Lipids* 29: 595–597, 1994.

20. **Scherneck S, Nestler M, Vogel H, Bluher M, Block MD, Berriel Diaz M, Herzig S, Schulz N, Teichert M, Tischer S, Al-Hasani H, Kluge R, Schurmann A, Joost HG.** Positional cloning of zinc finger domain transcription factor Zfp69, a candidate gene for obesity-associated diabetes contributed by mouse locus Nidd/SJL. *PLoS Genet* 5: e1000541, 2009.

21. **Scherneck S, Vogel H, Nestler M, Kluge R, Schurmann A, Joost HG.** Role of zinc finger transcription factor zfp69 in body fat storage and diabetes susceptibility of mice. *Results Probl Cell Differ* 52: 57–68, 2010.

22. **Stoehr JP, Byers JE, Clee SM, Lan H, Boronenkov IV, Schueler KL, Yandell BS, Attie AD.** Identification of major quantitative trait loci controlling body weight variation in ob/ob mice. *Diabetes* 53: 245–249, 2004.

23. **Taicher GZ, Tinsley FC, Reiderman A, Heiman ML.** Quantitative magnetic resonance (QMR) method for bone and whole-body-composition analysis. *Anal Bioanal Chem* 377: 990–1002, 2003.

24. **van Nas A, Ingram-Drake L, Sinsheimer JS, Wang SS, Schadt EE, Drake T, Lusis AJ.** Expression quantitative trait loci: replication, tissue- and sex-specificity in mice. *Genetics* 185: 1059–1068, 2010.

24a. **Villeneuve LM, Natarajan R.** The role of epigenetics in the pathology of diabetic complications. *Am J Physiol Renal Physiol* 299: F14–F25, 2010.

25. **Yang H, Wang JR, Didion JP, Buus RJ, Bell TA, Welsh CE, Bonhomme F, Yu AH, Nachman MW, Pialek J, Tucker P, Boursot P, McMillan L, Churchill GA, de Villena FP.** Subspecific origin and haplotype diversity in the laboratory mouse. *Nat Genet* 43: 648–655, 2011.

26. **Zhang B, Horvath S.** A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4: Article17, 2005.

27. **Zhong H, Beaulaurier J, Lum PY, Molony C, Yang X, Macneil DJ, Weingarth DT, Zhang B, Greenawalt D, Dobrin R, Hao K, Woo S, Fabre-Suver C, Qian S, Tota MR, Keller MP, Kendziorski CM, Yandell BS, Castro V, Attie AD, Kaplan LM, Schadt EE.** Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet* 6: e1000932, 2010.