Combined Expression Trait Correlations and Expression Quantitative Trait Locus Mapping

Hong Lan[1☯], Meng Chen[2☯], Jessica B. Flowers[1,3], Brian S. Yandell[2,4], Donnie S. Stapleton[1], Christine M. Mata[1], Eric Ton-Keen Mui[1], Matthew T. Flowers[1], Kathryn L. Schueler[1], Kenneth F. Manly[5], Robert W. Williams[5], Christina Kendziorski[6], Alan D. Attie[1*]

1 Department of Biochemistry, University of Wisconsin, Madison, Wisconsin, United States of America, **2** Department of Statistics, University of Wisconsin, Madison, Wisconsin, United States of America, **3** Department of Nutritional Sciences, University of Wisconsin, Madison, Wisconsin, United States of America, **4** Department of Horticulture, University of Wisconsin, Madison, Wisconsin, United States of America, **5** Departments of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, Tennessee, United States of America, **6** Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, Wisconsin, United States of America

Coordinated regulation of gene expression levels across a series of experimental conditions provides valuable information about the functions of correlated transcripts. The consideration of gene expression correlation over a time or tissue dimension has proved valuable in predicting gene function. Here, we consider correlations over a genetic dimension. In addition to identifying coregulated genes, the genetic dimension also supplies us with information about the genomic locations of putative regulatory loci. We calculated correlations among approximately 45,000 expression traits derived from 60 individuals in an F_2 sample segregating for obesity and diabetes. By combining the correlation results with linkage mapping information, we were able to identify regulatory networks, make functional predictions for uncharacterized genes, and characterize novel members of known pathways. We found evidence of coordinate regulation of 174 G protein–coupled receptor protein signaling pathway expression traits. Of the 174 traits, 50 had their major LOD peak within 10 cM of a locus on Chromosome 2, and 81 others had a secondary peak in this region. We also characterized a Riken cDNA clone that showed strong correlation with stearoyl-CoA desaturase 1 expression. Experimental validation confirmed that this clone is involved in the regulation of lipid metabolism. We conclude that trait correlation combined with linkage mapping can reveal regulatory networks that would otherwise be missed if we studied only mRNA traits with statistically significant linkages in this small cross. The combined analysis is more sensitive compared with linkage mapping alone.

Citation: Lan H, Chen M, Flowers JB, Yandell BS, Stapleton DS, et al. (2006) Combined expression trait correlations and expression quantitative trait locus mapping. PLoS Genet 2(1): e6.

Introduction

Biological systems are tightly controlled such that the concentrations of components are precisely balanced in an active biological pathway. The balances are maintained in part by coordinately regulated expression levels of the genes involved in the biological processes. Thus, genes that are coregulated across various physiologic conditions are likely to have similar functions. This creates a context by which the function of unannotated genes can be predicted.

One of the greatest challenges to biologists today is to functionally characterize genes identified by the genome sequencing projects. Traditionally, gene functions are inferred based on similarities in primary sequences, functional domains, or structures of encoded proteins [1]. In contrast to these qualitative approaches, we and others are using quantitative methods to predict gene function and discover gene regulatory networks [2,3].

Eisen and colleagues [4] were among the first to use microarray gene expression profiles for gene function prediction. They used hierarchical clustering algorithms to analyze time-course gene expression profiles of budding yeast and human fibroblasts and found that grouping of transcripts with similar profiles across a time dimension could accurately identify genes of similar function. Recently, Zhang and colleagues [3] generated microarray expression data for nearly 40,000 mRNAs in 55 mouse tissues and found that coordinated expression across a tissue dimension can contribute to the prediction of gene function. Although extremely informative, these approaches do not establish lines of causation among correlated mRNAs and therefore cannot provide the origin of a particular gene regulatory network.

Genetics establishes a one-way line of causation between genotype and phenotype. Brem et al. [5] used microarray data

Editor: Greg Gibson, North Carolina State University, United States of America

Received August 12, 2005; **Accepted** December 6, 2005; **Published** January 20, 2006

A previous version of this article appeared as an Early Online Release on December 7, 2005 (DOI: 10.1371/journal.pgen.0020006.eor).

DOI: 10.1371/journal.pgen.0020006

Abbreviations: BP, Biological Process; eQTL, expression quantitative trait loci; FDR, false discovery rate; GABA, gamma-aminobutyric acid; GO, Gene Ontology; GPCR, G protein–coupled receptor; LXR, liver-X-receptor; RMA, Robust Multi-array Average

* To whom correspondence should be addressed. E-mail: attie@biochem.wisc.edu

☯ These authors contributed equally to this work.

Synopsis

In order to annotate gene function and identify potential members of regulatory networks, the authors explore correlation of expression profiles across a genetic dimension, namely genotypes segregating in a panel of 60 F_2 mice derived from a cross used to explore diabetes in obese mice. They first identified 6,016 seed transcripts for which they observe that the gene expression is linked to a particular region of the genome. Then they searched for transcripts whose expression is highly correlated with the seed transcripts and tested for enrichment of common biological functions among the lists of correlated transcripts. They found and explored the properties of 1,341 sets of transcripts that share a particular "gene ontology" term. Thirty-eight seeds in the G protein–coupled receptor protein signaling pathway were correlated with 174 transcripts, all of which are also annotated as G protein–coupled receptor protein signaling pathway and 131 of which share a regulatory locus on Chromosome 2. The authors note many of these findings would have been missed by simple expression quantitative trait loci analysis without the correlation step. The approach was used to identify a common set of genes involved in lipid metabolism.

as phenotypes for genetic mapping in a segregating yeast sample. This approach has inspired the mapping of expression quantitative trait loci (eQTL) in mammalian species [6–12]. Each study identified *cis* and *trans* traits. A *cis* trait is one that genetically maps to the physical location of the gene encoding its mRNA, suggesting that variation at the locus is responsible for the heritable changes in gene expression. A *trans* trait maps to a region distinct from its physical location and thus implies the location of a potential regulator acting in *trans*. Theoretically, when multiple traits map together in *trans* (i.e., overlapping eQTL peaks within an approximately 10-cM region), we can hypothesize that they are coregulated by a common regulator encoded by a gene locus underlying the eQTL. Unfortunately, in practice, because mRNAs are typically regulated by multiple factors, individual *trans*-acting eQTL exert a relatively weak genetic signal and are difficult to detect in a small sample size. This limits the power to detect metabolic regulators.

Schadt et al. [9] surveyed 23,574 genes using Rosetta oligonucleotide arrays. More recent studies used Affymetrix U74A arrays, or RAE230A arrays, which covered about one third of all the transcripts [7,8,11]. The transcripts chosen for measurement were biased toward known genes. Here, we measured the expression levels of the transcripts on the Affymetrix MOE430 Set arrays, which include many genes of unknown function.

We studied an F_2 segregating for obesity and diabetes to expose key components of gene regulatory networks. To this end, we combined correlation analysis with linkage mapping. Rather than considering correlation across a time or tissue dimension, we evaluated correlation across a genetic dimension, as created by a segregating F_2 sample. An F_2 represents an allelic block permutation. Our combined approach increases the power to identify metabolic regulatory networks.

Results

We generated an F_2-*ob/ob* sample from the C57BL/6J (B6) and BTBR founder strains. The B6 and BTBR strains, when made obese, differ in diabetes susceptibility [13]; B6-*ob/ob* mice are diabetes resistant, whereas BTBR-*ob/ob* mice develop severe diabetes. The F_2 mice were genotyped for 194 markers (average marker interval of approximately 10 cM). A subset of 60 mice was chosen using a selective phenotyping algorithm developed in our group [14]. From each liver, 45,037 expression measurements were obtained using the Affymetrix MOE430 microarrays. According to the December 2004 gene annotation table, 14,487 (32.2%) of the transcripts have unique RefSeq protein identifiers, and 16,466 (36.6%) have been annotated to at least one Gene Ontology (GO) Biological Process (BP). Altogether, about two thirds of the transcripts that we measured were not annotated in either of these two sources.

Mapping of Metabolic Regulatory Loci

After processing the Affymetrix CEL files with Robust Multi-array Average (RMA) [15], we used interval mapping to screen the 45,037 traits. Transcripts with LOD scores of 3.4 or greater at one or more locations in the genome were recorded as mapping transcripts. Use of this LOD threshold roughly corresponds to a genomewide type I error rate of 0.08 [16] and a false discovery rate (FDR) of 0.48 (see Materials and Methods). The high FDR is not surprising, considering that adjustments for multiple tests across transcripts have not been made [17]. We nevertheless keep this LOD threshold since we consider this step as only an initial screen; the threshold reduces the computational burden significantly while maintaining a high true-positive rate. This procedure yielded 6,016 mapping transcripts.

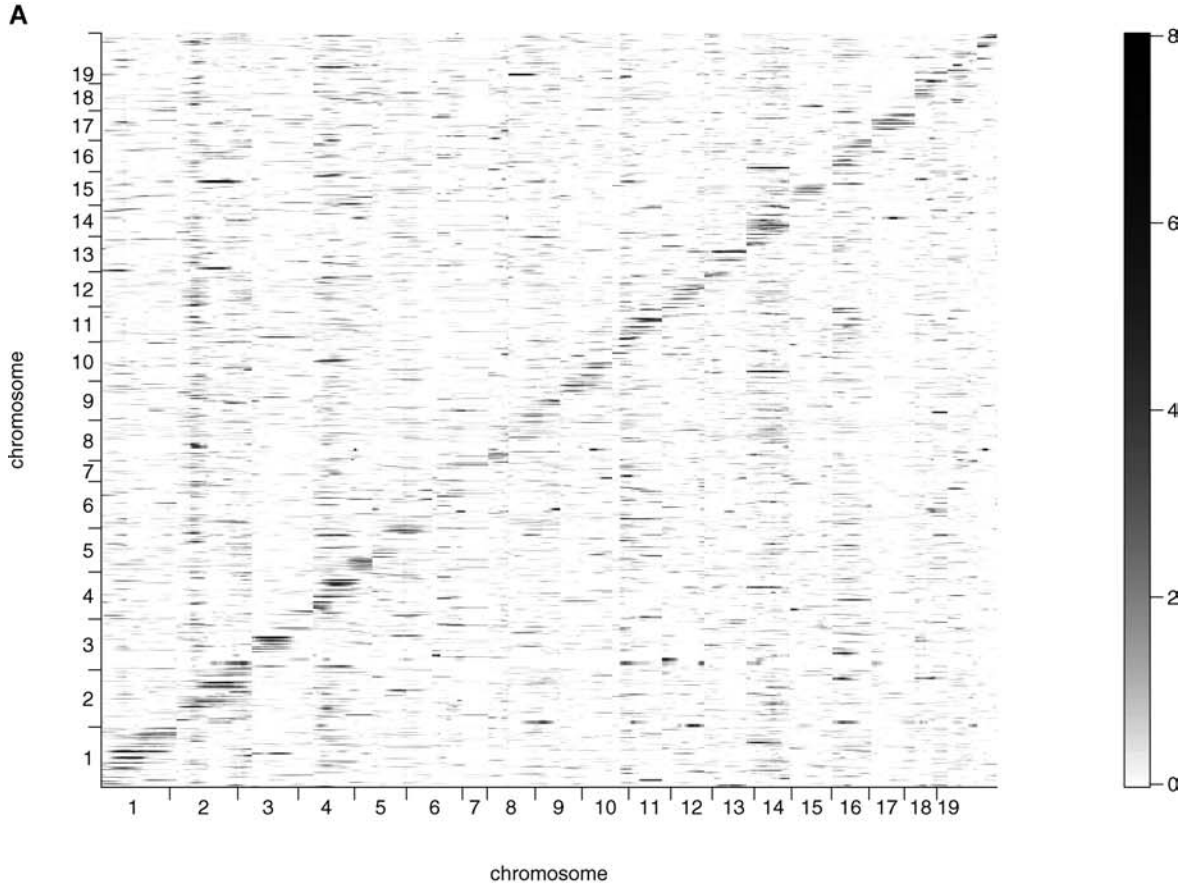
We plotted the mapping locations of these expression traits against their physical locations (Figure 1A). Of the 6,016 mapping transcripts, 723 were classified as *cis* (the diagonal spots in Figure 1A) and the rest were classified as *trans*. The *cis* traits are listed in Table S1. Rearrangement of Figure 1A using hierarchical clustering of the LOD profiles enabled us to identify the putative locations of metabolic regulators (Figure 1B). We identified 15 regions with the highest number of mapping transcripts, and for each region, we tested the list of mapping traits for enrichment in a common functional annotation, using the GO database and a Hypergeometric calculation. For each region, we examined up to 50 GO terms,

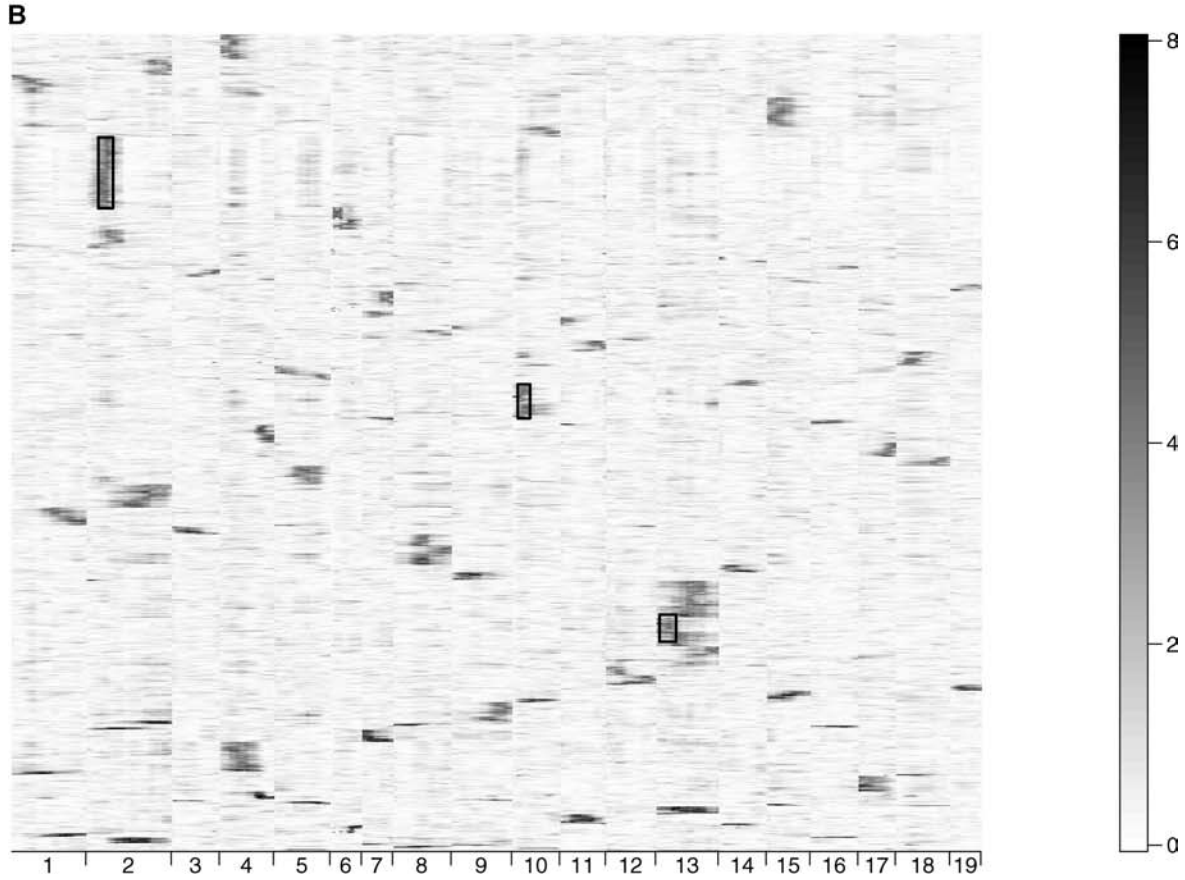
Figure 1. Expression QTL in the (B6 × BTBR) F_2-*ob/ob* Cross

(A) Expression QTL that regulate gene expression in the (B6 × BTBR) F_2-*ob/ob* cross. The physical locations of the transcripts are organized on the y-axis. The chromosome regions (x-axis) to which those transcripts are mapped were obtained using the *scanone* function in the R/qtl package with 5-cM intervals. The gray scale reflects the strength of the linkage signals (LOD scores greater than 8 are scaled to 8). Transcripts appearing on the diagonal are inferred to be *cis*-regulated. The chromosomes were concatenated to form a 2,300-Mb genome, starting from Chromosome 1 and ending with Chromosome 19.

(B) Global display of mapping patterns of linkage clusters. The y-axis shows the 6,016 mapping transcripts. The x-axis shows the physical locations to which these transcripts map. The transcripts are grouped based on the hierarchical clustering of linkage mapping patterns across the genome. Darker areas indicate the regions to which traits are comapped or coregulated. The boxes correspond to the hot spots on Chromosomes 2, 10, and 13. The gray scale reflects the strength of the linkage signals (LOD scores greater than 8 are scaled to 8).

DOI: 10.1371/journal.pgen.0020006.g001


A
chromosome
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
chromosome
8
6
4
2
0


B
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
8
6
4
2
0

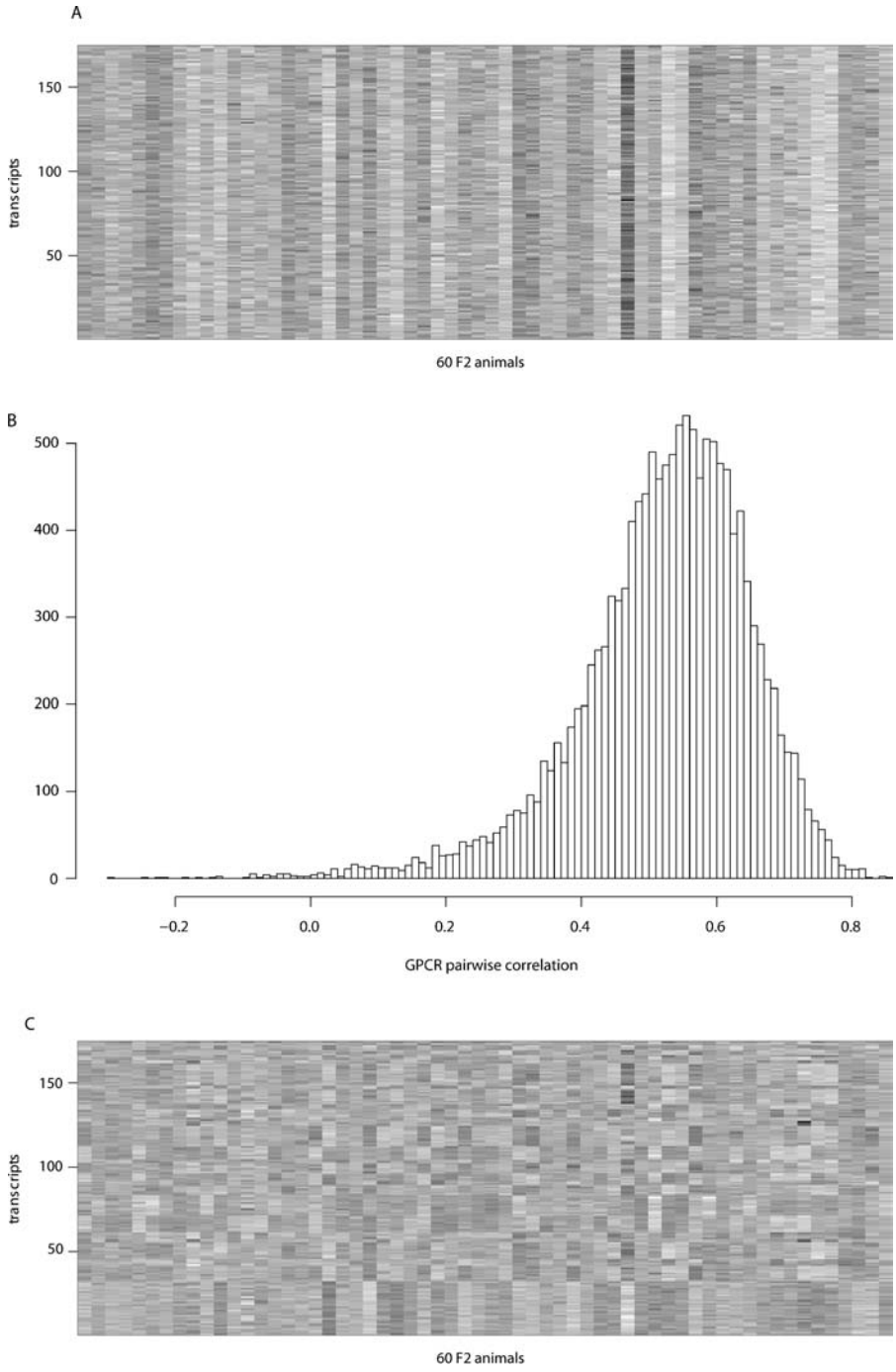


Figure 2. Coordinate Regulation of Genes for GPCR Protein Signaling Pathways

(A) GPCR protein signaling pathway genes show coordinated changes across the 60 F_2 mice. The rescaled expression levels of 174 distinct transcripts (y-axis) that belong to the GPCR pathway. The x-axis displays the 60 F_2 mice. The plot shows that the subset of GPCR traits identified by our trait correlation-GO analysis has vertical patterns that indicate coordinate expression across the 60 F_2 mice ($P = 1e10^{-4}$ by permutation test).

(B) Histogram of pairwise correlation coefficients of GPCR traits identified with different seeds.

(C) Scaled expression levels of traits from one realization of the permuted lists. Thirty-eight lists of correlated traits were randomly sampled from a total of 1,341 correlated trait lists. The list sizes were adjusted to match those of the GPCR lists. The expression values were scaled to have mean intensity = 0 and variance = 1, as in the GPCR list in Figure 2A.

DOI: 10.1371/journal.pgen.0020006.g002

each of which involved more than ten genes. A region is said to be enriched for a particular GO term if the resulting Hypergeometric *P*-value is ≤ 0.01. Using these criteria, we found modest functional enrichment in each region (results not shown). However, when we followed up and examined all eQTLs that have similar maximum LOD positions (whether statistically significant or not), we found clear evidence of enrichment (Figure 3B). We performed the same procedure

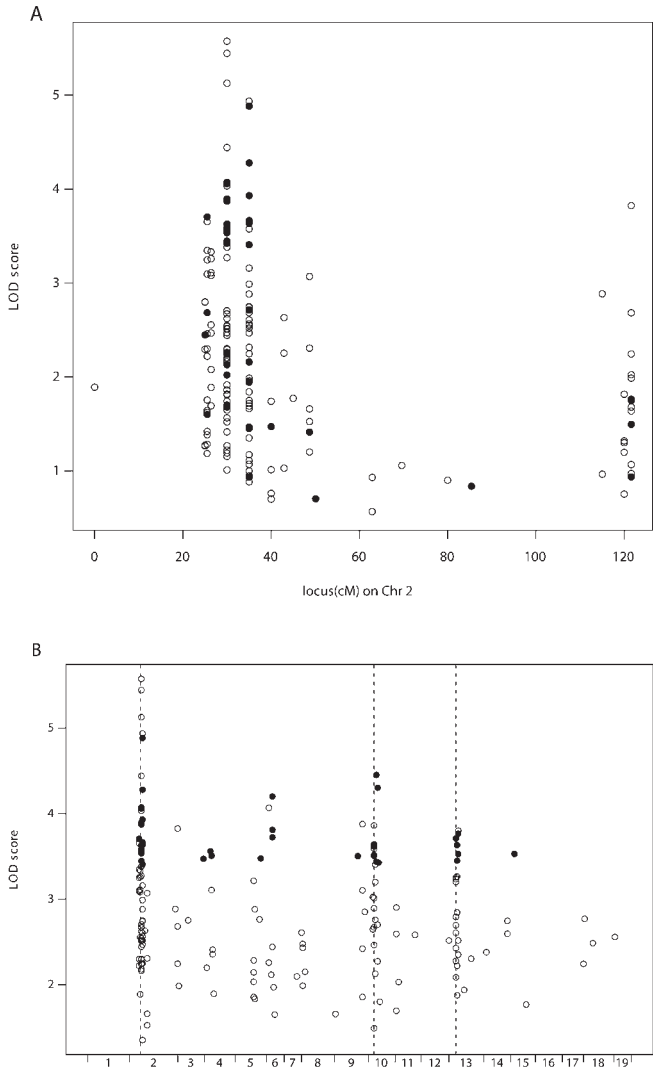


Figure 3. An Expression Quantitative Trait Locus on Chromosome 2 Regulates Many GPCR Protein Pathway Genes

(A) Comapping of GPCR pathway traits to loci on Chromosome 2. The figure shows the peaks on Chromosome 2 versus LOD score for 174 GPCR protein signaling pathway traits. The closed circles correspond to 38 seed traits.

(B) Comapping of GPCR pathway traits. The figure shows the magnitude (y-axis) versus location of LOD peaks (x-axis is scaled by cM) for 174 GPCR protein signaling pathway traits. The closed circles correspond to the 38 seed traits. The interval mapping LOD profiles for all 38 seeds are included in Figure S2.

DOI: 10.1371/journal.pgen.0020006.g003

on randomized data, preserving the correlation structure among the traits. This enrichment test did not reveal results as significant as were observed in the original data.

Combining Mapping and Correlation Analysis Improves Sensitivity to Detect Regulatory Loci

We developed a strategy to explore metabolic regulatory networks on a genomewide scale that incorporates correlated traits that by themselves do not show statistically significant linkage. Considering each of the 6,016 mapping transcripts as a "seed," we computed the Pearson correlation coefficients *(r)* across the 60 F_2 mice for the 45,037 expression measurements and constructed a list of traits having $r \geq 0.7$ (FDR = $1.52e10^{-6}$). We then tested for enrichment of any GO BP. Among the 6,016 seed transcripts, 1,341 produced lists of traits enriched for at least one GO term. The lists identified by the seed traits and GO terms are given in Table S2.

Among all of the lists from the 1,341 seed transcripts, we identified 862 unique GO BP terms (15,726 in total). We traced back the lists of correlated traits that are enriched for each unique GO term and concatenated all such lists. Different seeds may have identified overlapping traits enriched for the same GO term. These redundant copies of traits were removed after concatenation. By doing so, we associated a collection of distinct traits with each of the 862 unique GO terms. The full archive of lists is available in Table S3.

To illustrate how a gene mapping approach alone is insufficient, consider a locus on Chromosome 2 that regulates in *trans* multiple ion transport genes (Figure 1B, box), including six genes encoding subunits of the mitochondrial F_1 or F_0 ATPase of the proton channel of the electron transport system. It is known that there are 24 other ATP synthase subunits in the genome. None of the 24 expression traits exceeded the LOD threshold of 3.4, but many of them shared the linkage peak on Chromosome 2 (Figure S1). These results exposed a limitation of eQTL mapping; without inspection of the list of correlated *trans* traits, biological inference, and reconsideration of the 24 statistically insignificant LOD profiles, we would have missed evidence for coregulation of an entire family of mitochondrial respiratory genes. To automate this procedure, we consider augmenting eQTL mapping with trait correlation analysis.

A Regulator Locus for G Protein–Coupled Receptor Genes

Using the correlation analysis described above, we identified 38 seeds that are in the G protein–coupled receptor (GPCR) protein signaling pathway and found 174 distinct traits correlated with these 38 seeds. Dramatically, all 174 traits were annotated as belonging to the GPCR protein signaling pathway GO term (Table S4). We created a heat plot to visualize the correlation patterns across multiple lists for this pathway by first rescaling the $\log_2$ intensity values of each trait across all the 60 F_2 mice so that the mean = 0 and variance = 1. This rescaling step is necessary because some genes are overexpressed across all the F_2 mice and others are underexpressed, which fails to reveal the coordinate regulation pattern across traits. We plotted the rescaled expression levels of the transcripts across the 60 F_2 mice (Figure 2A). A striking feature of Figure 2A is the monochromatic columns, which reflect the coordinate changes in gene expression across the genetic dimension represented by the F_2 sample. The traits from different seeds are not all highly correlated with each other (Figure 2B). However, the fact that they are enriched for the same GO term implies that they are coordinately regulated.

To ensure that the vertical patterns that were preserved across multiple lists were indeed functionally relevant, we randomly sampled 38 lists from the total of 1,341 lists and adjusted the sizes to be the same as those of the GPCR lists. We concatenated traits from those 38 randomly sampled lists. The vertical patterns shown in Figure 2A are no longer evident (Figure 2C). This was repeated 10,000 times. The vertical patterns in Figure 2A were shown to be statistically significant ($P = 1e10^{-4}$; see Materials and Methods).

It is important to note that many of the traits correlated to

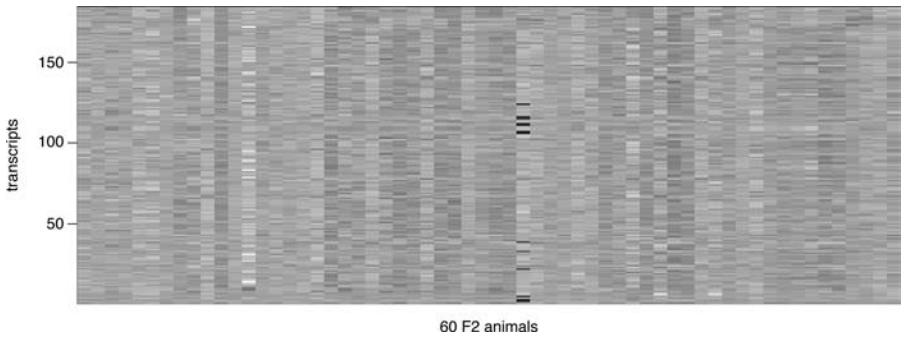


Figure 4. Lipid Metabolism Pathway Genes Show Coordinated Expression across the F_2 Mice

The rescaled expression levels of 184 distinct transcripts (y-axis) that belong to the lipid metabolism pathway are shown for each of the 60 F_2 mice (x-axis). Plot shows that the subset of lipid metabolism traits identified by our trait correlation-GO analysis has preserved vertical patterns ($P = 4e10^{-4}$ by permutation test).

DOI: 10.1371/journal.pgen.0020006.g004

seed traits did not show significant linkage across the 60 F_2 mice and would have been overlooked by linkage mapping alone. When we mapped the GPCR seed with the most significant linkage, we discovered that the expression levels of this GPCR gene were influenced by a regulatory locus on Chromosome 2 at 30 cM, between *D2Mit297* and *D2Mit9* (Figure 3A). Of the 174 traits, 50 had their major LOD peak within 10 cM of this locus, and 81 others had a secondary peak in this region (Figure 3A). While many of these LOD scores do not exceed our significance threshold, the strong agreement suggests that most are driven by this same regulator or closely linked regulators, thus comprising a regulatory network. The LOD profiles for the 38 seed traits are shown in Figure S2. There appeared to be secondary regulators on Chromosomes 10 and 13 (Figure 3B). Similar evidence was suggested by inspecting the chromosome regions as shown in the boxes in Figure 1B. We picked out all transcripts, whether or not they have LOD scores exceeding 3.4, that share the same LOD peaks on the regions of Chromosome 2, 10, or 13 and tested for BP GO enrichment. We found GPCR protein signaling pathway in the enriched list with very small *P*-values. In short, we found that trait correlation combined with linkage mapping can reveal regulatory networks that would otherwise be missed if we studied only mRNA traits with statistically significant linkages in this small cross. The combined analysis is more sensitive compared with linkage mapping only.

A Regulator Locus of Lipid Metabolism Genes

The strategy given above can be applied to many metabolic pathways. We found evidence of a regulatory pathway

Table 1. Top 20 Gene Expression Traits Correlated with *Scd1* Are All Lipid Metabolism Genes

Probe Set ID	Gene Symbol	Description	Correlation	Chr[a]	Position (Mb)	2-QTL Fit[b]	D2Mit263	D5Mit240
1415965_at	*Scd1*	Stearoyl-coenzyme A desaturase 1	1.000	19	44	0.000	0.003	0.005
1415964_at	*Scd1*	Stearoyl-coenzyme A desaturase 1	0.921	19	44	0.000	0.003	0.049
1441881_x_at	*3110032G18Rik*	*RIKEN cDNA 3110032G18* gene	0.884	5	123	0.000	0.000	0.125
1422432_at	*Dbi*	Diazepam binding inhibitor	0.830	1	120	0.000	0.014	0.008
1417963_at	*Pltp*	Phospholipid transfer protein	0.825	2	165	0.001	0.084	0.008
1456424_s_at	*Pltp*	Phospholipid transfer protein	0.823	2	165	0.000	0.029	0.004
1417404_at	*Elovl6*	ELOVL family member 6, elongation of long-chain fatty acids (yeast)	0.806	3	130	0.003	0.119	0.011
1415840_at	*Elovl5*	ELOVL family member 5, elongation of long-chain fatty acids (yeast)	0.792	9	78	0.000	0.081	0.003
1423828_at	*Fasn*	Fatty acid synthase	0.791	11	121	0.001	0.038	0.024
1422185_a_at	*Dia1*	Diaphorase 1 (NADH)	0.789	15	84	0.010	0.035	0.097
1430307_a_at	*Mod1*	Malic enzyme, supernatant	0.773	13	61	0.018	0.071	0.081
1437211_x_at	*Elovl5*	ELOVL family member 5, elongation of long-chain fatty acids (yeast)	0.766	9	78	0.001	0.007	0.045
1417403_at	*Elovl6*	ELOVL family member 6, elongation of long-chain fatty acids (yeast)	0.756	3	130	0.008	0.122	0.022
1419499_at	*Gpam*	Glycerol-3-phosphate acyltransferase, mitochondrial	0.743	19	54	0.001	0.044	0.018
1455976_x_at	*Dbi*	Diazepam binding inhibitor	0.742	1	120	0.000	0.005	0.099
1428714_at	*Pgrmc2*	Progesterone receptor membrane component 2	0.742	3	41	0.010	0.004	0.794
1451457_at	*Sc5d*	Sterol-C5-desaturase (fungal ERG3, delta-5-desaturase) homolog *(S. cerevisae)*	0.741	9	42	0.014	0.039	0.091
1425303_at	*Gck*	Glucokinase	0.738	11	6	0.002	0.021	0.075
1426264_at	*Dlat*	Dihydrolipoamide S-acetyltransferase (E2 component of pyruvate dehydrogenase complex)	0.735	9	51	0.003	0.051	0.041
1428082_at	*Acsl5*	Acyl-CoA synthetase long-chain family member 5	0.734	19	55	0.007	0.043	0.080

[a]Physical location of the gene encoding the mRNA.

[b]*P*-values for overall fit of two-QTL model and type 3–adjusted test for each QTL.

DOI: 10.1371/journal.pgen.0020006.t001

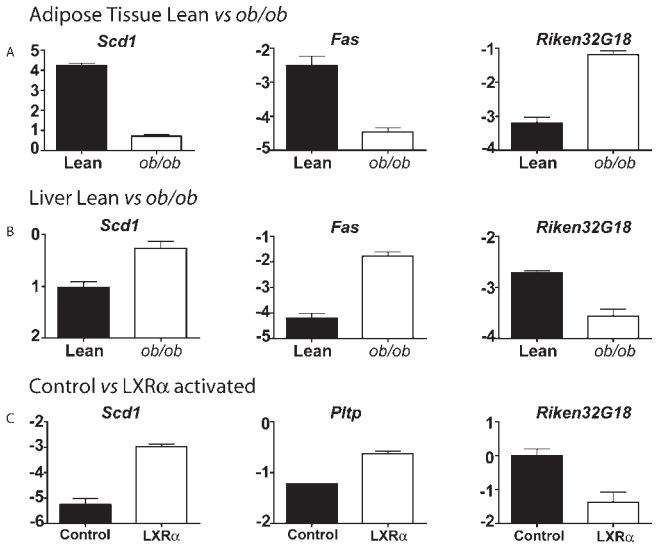


Figure 5. Regulation of *Riken32G18* Inversely Parallels the Regulation of Lipogenic Genes

(A) Metabolic regulation of *Riken32G18*. Expression of lipogenic genes and *Riken32G18* in adipose tissues of lean versus obese mice.
(B) Expression of lipogenic genes and *Riken32G18* in livers of lean versus obese mice.
(C) Expression of LXRα target genes and *Riken32G18* in mouse livers treated with LXRα agonist T0901317.
DOI: 10.1371/journal.pgen.0020006.g005

controlling lipid metabolism (Figure 4; $P = 4e10^{-4}$). These expression patterns were combined from lists using 71 different seeds. As was the case for the GPCR transcripts, these show evidence of coregulation across the 60 F_2 mice.

We have previously studied eQTL for stearoyl-CoA desaturase-1 *(Scd1)* [2], an important gene for lipid metabolism and insulin sensitivity [18,19]. When we used *Scd1* expression as a seed, the traits that were most closely correlated with *Scd1* were clearly enriched in lipid metabolism genes (Table S5).

The traits that are highly correlated also map to the same genomic locations, even though their respective genes are in separate locations (i.e., they are being regulated in *trans*). Major QTL peaks for most of the 20 lipid metabolism traits were found on Chromosomes 2 *(D2Mit263)* and 5 *(D5Mit240);* to a lesser extent, they were found by loci on several other chromosomes. We found strong evidence that these two loci are jointly predictive for all 20 traits (Table 1).

Two genes that ranked high in Table 1, *RIKEN cDNA 3110032G18* (*Riken32G18* hereafter; $r = 0.868$) and diazepam binding inhibitor ($r = 0.816$), did not seem to have obvious functional relevance with *Scd1*. We used this opportunity to test the predictive power of trait correlations.

Diazepam Binding Inhibitor Is Functionally Related to *Scd1*

The correlation between diazepam binding inhibitor *(Dbi)* and *Scd1* is stronger than that for the classic lipid biosynthesis genes such as fatty acid synthase, malic enzyme, and fatty acyl-CoA ligase. Interval mapping of the expression levels of *Dbi* and *Scd1* showed highly similar patterns (data not shown), suggesting that they are regulated by common loci. To understand the biological basis for the predicted coregulation, we first sought to use in silico approaches to investigate this gene. The *Dbi* gene symbol is synonymous with acyl-CoA binding protein *(Acbp)*. ACBP is a 10-kDa cytosolic protein that binds long-chain fatty acyl-CoAs and has a high affinity for oleoyl-CoA, the product of the SCD1 reaction [20]. Our data analysis predicts that a common regulator maintains a fixed ratio of mRNA for SCD1 and for a protein that binds its product, oleoyl-CoA. Recently, *Dbi* was reported to be regulated by sterol regulatory element binding protein-1c (SREBP-1c), a known master regulator of lipogenic enzymes [21]. Perhaps one or more of the regulators we have mapped are coactivators or corepressors of SREBP-1c, analogous to PPARγ coactivator-1β [22]. Thus, the prediction from our trait correlation analysis that *Dbi* is functionally related to *Scd1* was confirmed by recently published experimental data.

Prediction of *Riken32G18* as a Putative Lipid Metabolism Gene

The mRNA most closely correlated to *Scd1* is *Riken32G18*, which maps to the same location on Chromosome 2 as *Scd1* expression but with almost twice the LOD score. Notably, *Riken32G18* is physically located on Chromosome 5 and *Scd1* is on Chromosome 19; thus, the two genes are likely coregulated in *trans* by the Chromosome 2 locus.

We hypothesized that *Riken32G18* plays a role in lipid metabolism with a close functional relationship with *Scd1*. There is no information available from the online resources to functionally link *Riken32G18* to *Scd1* or to lipid metabolism. In order to test its functional relationship with lipid metabolism, we asked if *Riken32G18* expression is responsive to conditions that regulate expression of lipid metabolism genes.

From our previous results, we know that lipogenic genes are expressed at a lower level in *ob/ob* adipose tissue than in lean adipose tissue [23]. Whereas *Scd1* and fatty acid synthase *(Fas)* were expressed at a lower level in *ob/ob* adipose tissue, *Riken32G18* expression was higher in obese adipose tissue. Thus, *Riken32G18* appeared to be regulated reciprocally with lipogenic genes (Figure 5A). In contrast to adipose tissue, lipogenic genes are more highly expressed in livers of *ob/ob* mice than in lean mice [24]. Again, *Riken32G18* was regulated in the opposite direction; i.e., its expression was reduced in *ob/ob* liver (Figure 5B). Finally, liver-X-receptor (LXR) α activation increases genes in fatty acid metabolism in the liver. Once again, *Riken32G18* expression was down-regulated by LXRα activation (Figure 5C). We conclude that *Riken32G18* is very likely a novel gene involved in lipid metabolism. It is coregulated with lipid metabolism genes across the 60 F_2 mice and reciprocally regulated with such genes across physiologic conditions.

Discussion

Macromolecules and metabolites in cells interact with one another across a critical concentration range. The functional concentration must be within boundaries defined by the kinetics and thermodynamics of these interactions. One way cells can maintain the critical concentration range for molecules within a pathway is to maintain them under the control of a common regulatory system. A well-characterized example is the regulation of lipogenic enzymes by SREBP-1c, a master regulator of virtually all enzymes critical to fatty acid

synthesis [25]. Pathways that feed into lipogenesis are also coregulated with lipogenic enzymes.

Whereas Eisen et al. [4] studied gene expression correlation over a time dimension and Zhang et al. [3] studied gene expression correlation over a tissue dimension, we performed our studies over a genetic dimension. In addition to identifying coregulated genes, the genetic dimension also supplies us with information about the genomic location of putative regulatory loci. Our method involved a two-step approach. First, we applied a genomewide linkage analysis to identify mapping transcripts using a liberal filter to ensure a high true-positive rate. In the second step, we performed trait correlation analysis, using seeds within linkage groups.

Our approach revealed biologically meaningful relationships among the traits. In particular, we found that traits that are correlated over a genetic dimension are frequently enriched for a physiologically meaningful biological function. For example, we found a gene whose expression was highly correlated with those of lipid metabolism genes and with *Scd1* expression as a seed trait. The prediction that this gene is involved in lipid metabolism was borne out by recent studies showing that this protein binds to oleoyl-CoA, the product of the SCD1 reaction. Moreover, the eQTL for this gene maps to the same location on distal Chromosome 2 as did other lipid metabolism eQTL.

The distal region of Chromosome 2 containing the cluster of lipid metabolism traits is quite interesting as several QTL for obesity and related traits have been mapped to this region [13,26–30]. Furthermore, Estrada-Smith et al. [31] have used congenic mouse lines derived from B6 and CAST/Ei strains to identify a 17-Mb region of distal Chromosome 2 that regulates plasma levels of triglycerides. This region directly overlaps with the location of our fatty acid metabolism cluster and our fasting plasma insulin locus.

We also found an entirely unannotated gene highly correlated with lipid metabolism genes. Analysis of tissues from obese animals or animals treated with an LXRα agonist revealed that this gene is regulated inversely under these physiologic conditions with lipid metabolism genes, providing evidence that it participates in lipid metabolism or perhaps is a negative regulator of lipid metabolism.

Another example of coordinated expression was the GPCR protein signaling pathway genes. Here, we found that *all* of the 174 transcripts that were highly correlated with 38 GPCR seeds encoded GPCR pathway proteins and most mapped to the same locus on Chromosome 2. Fewer than 10% of the transcripts are olfactory receptors. Among the other correlating genes are gamma-aminobutyric acid (GABA) receptor signaling pathway genes, neuropeptide signaling pathway genes, and cAMP signaling genes (Table S6).

The GPCR protein signaling pathway cluster overlaps with a region on proximal Chromosome 2 that was narrowed down to an 8-Mb region regulating body weight differences in the B6 and BTBR strains [32] and a 10-Mb region underlying obesity-related traits in the B6 and CAST/Ei strains [31]. Several of the GPCRs in the cluster are related to obesity traits in mice when mutated or overexpressed [33–36], and others have been proposed as candidate genes for human obesity [37–39] (Figure S3A). An additional subset of the GPCR cluster represents GABA receptor subunits (Figure S3B). Alterations in the GABA system are related to the obesity phenotype seen in obese *(fa/fa)* Zucker rats [40] and humans with Prader-Willi syndrome [41,42].

Although we have focused our own analysis on lipid metabolism, our database, which is available through WebQTL (www.genenetwork.org) [43], can provide insights into the regulation of metabolic pathways expressed in the liver. A researcher can start with a particular mRNA of interest as a seed trait and quickly identify genes whose expression is highly correlated with the seed trait. By looking for overlapping loci that control the expression of the seed and additional genes of interest, a regulatory pathway can be hypothesized.

The particular implementation of the approach can and should be study specific. For example, there are a number of thresholds that must be determined, each affecting the number of transcripts considered and the overall FDR. The level of tolerable FDR depends on a number of factors. We accepted a very high FDR in our first stage of transcript mapping as we desired only a mild filtering of the data to reduce the computational burden. We have recently developed statistical methods that provide better control of FDR for eQTL mapping [17]. These could be used if a more stringent filtering was desired at this stage. A more conservative filtering in the second stage was used here by imposing a relatively high threshold for the correlation coefficient. Although a correlation coefficient threshold as low as 0.4 would still correspond to an acceptable FDR of 0.05, we chose the more conservative 0.7. We have reported stage-specific FDRs as they were used to guide our analysis. However, we note that the FDR associated with the entire procedure is not known. Indeed, it is an extremely difficult problem to precisely define, not to mention determine, overall FDR for studies involving multiple types of data and multiple analysis methods. Doing so remains an open question of importance to both statisticians and biologists.

In summary, here we used two complementary approaches. First, we started with single QTL mapping to obtain seeds. Second, we used the seeds to obtain correlated traits. Finally, we confirmed genetic architecture for the correlated traits. As we have argued, the combined approach is more sensitive to retrieve biologically meaningful information from the data set. If we just did QTL mapping, those traits with non-significant LOD scores would have been missed. If we just performed correlation analysis, we would have lost the causation information that comes from the genetic linkage analysis. Our approach provides an important organizing step in the analysis of expression networks. It yields sets of related traits that can be further analyzed using multiple trait mapping [44] and causal networks [45], as these methods are most effective after screening to obtain a modest number of closely related traits. Moreover, our approach increases the power to formally uncover members of gene expression networks.

Materials and Methods

Animals. The power of a genetic mapping study depends on the heritability of the trait, the number of individuals included in the analysis, and the genetic dissimilarity among them. In this study, we selected 60 (B6 × BTBR) F_2-*ob/ob* mice (29 males and 31 females) based on the selective phenotyping algorithm that we developed [14]. The algorithm selects a subset of F_2 individuals from a mapping panel based on genotype data. The selection achieves substantial improvements in sensitivity compared to a random sample of the same size [14]. The 60 mice used in this study were a subset of F_2 mice

segregating for phenotypes associated with obesity and diabetes [13]. The F_1 mice used to generate the F_2 sample were all derived from crosses in which the B6 parent was male and the BTBR parent was female. The framework map consists of 194 microsatellite markers, with an average spacing of 10 cM.

Microarrays. Liver total RNA was extracted from frozen tissue samples with RNAzol reagent (Tel-Test, Friendswood, Texas, United States). Crude RNA samples were purified with RNeasy mini-columns (Qiagen, Valencia, California, United States) before being subjected to microarray and RT-PCR studies. The RNA samples were processed according to the Affymetrix Expression Analysis Technical Manual. A total of 60 MOE430A and MOE430B arrays were used to monitor the expression levels of approximately 45,000 genes or ESTs. The data were processed with MAS5.0 to generate cell intensity files. Quantitative expression levels of all the transcripts were estimated using the RMA algorithm for normalization [15].

Identifying transcripts with linkages. To identify traits that have linkages in the 60 F_2 mice, we used standard interval mapping implemented in R/qtl [46] to map each of the 45,037 unique probe sets at 5-cM resolution. Transcripts with LOD scores of 3.4 or greater at one or more locations in the genome were recorded as transcripts with linkages, or "seeds." Using maximum LOD of 3.4 or greater as a threshold roughly corresponds to genomewide type I error rate of 0.08 [16]. We permuted the animal labels five times and for each permutation ran R/qtl interval mapping on a 5-cM grid. The loci-specific LOD scores from the five permutations were pooled together to estimate empirical *P*-values These *P*-values were then used for calculation of q-values and estimation of FDR [47]. The FDR corresponding to a LOD score of 3.4 was 0.48. This resulted in 6,016 mapping traits. We found some regions of the genome with strong sex-by-genotype interactions. The results reported here are in regions with little or no evidence for such interactions.

Hierarchical clustering of the LOD score profiles. LOD score profiles were obtained for each transcript using R/qtl interval mapping with 5-cM intervals. This yielded, for each trait, 465 interval mapping testing points in the entire genome. Treating each LOD score profile as a point in the 465-dimensional space, we calculated pairwise Euclidean distances between LOD score profiles and performed hierarchical clustering based on those distances. Clustering the LOD profiles has the advantage that transcripts having similar mapping patterns across the genome will be grouped together.

Expression correlation. Using each of the 6,016 transcripts as a seed, Pearson correlation coefficients were calculated against the log-transformed expression intensities (raw intensities were normalized by RMA) of all the 45,037 transcripts in 60 F_2 mice. We obtained a 6,016-by-45,037 matrix of correlation coefficients. For each seed, we collected the traits with a correlation coefficient of 0.7 or greater, which forms the list of correlated traits for that seed. For each pair of transcripts, we computed its Pearson correlation together with the *P*-value. We then computed the corresponding q-values to determine the FDR [47]. A correlation of 0.7 corresponds to an FDR of $1.52e10^{-6}$.

Test for enrichment. Enriched functional groups, defined as GO categories corresponding to BP, were identified from each list with *P*-values computed from a Hypergeometric calculation, carried out using a custom variant of "GOHyperG" in the "GOstats" package (version 1.1.3; Bioconductor Core Team 2004; http://www.bioconductor.org). Interpretation of these *P*-values is not straightforward since many dependent hypotheses are tested. Furthermore, the Hypergeometric calculation tends to result in small *P*-values when groups with few transcripts are considered. It has been suggested that one consider only GO nodes with small *P*-values and a reasonable number of genes (ten or more) [48]. For each list, we examined up to 50 GO terms, each of which involves more than ten genes and has a Hypergeometric *P*-value of ≤ 0.01.

Creating the coordinated expression plot. From the lists of correlated traits, we identified 15,726 BP GO terms in total, among which 862 are unique. For each of the unique GO terms, we traced back the lists of correlated traits that were enriched ($P = 0.01$), and we concatenated all such lists. To create the heat plot, we first normalized the log intensity values of each trait across all the 60 F_2 mice to mean = 0 and variance = 1. This step is very important because some traits were highly overexpressed or underexpressed across all the F_2 mice. We then plotted the rescaled expression values of transcripts across the 60 F_2 mice.

Permutation test for coordinated expression. We developed a statistic to assess the extent of coordinate expression. We first collapsed all the members in a list into a pseudo-member by taking the average across the transcripts in that list. We then computed pairwise correlations among the pseudo-members. The final measure of coordinate expression was the average of those pairwise correlations. For a given GO term, we sampled the same number of lists from the total of 1,341 lists 10,000 times and matched the list sizes to those in the GO term of consideration. We computed the statistic in each random sample. The empirical *P*-value was calculated as the frequency of the samples where the statistic exceeded the one calculated using the observed data. The *P*-value for GPCR protein signaling pathway set was $1e10^{-4}$, and it was $4e10^{-4}$ for the lipid metabolism set.

Genetic architecture of GO groups. Subsequent QTL analyses were performed on traits within GO groups using one- and two-QTL scans with R/qtl. Possible interacting effects of sex were examined and found not to be significant for the GPCR and lipid metabolism groups reported here (data not shown).

Real-time quantitative PCR for *Riken32G18* and lipogenic genes. First-strand cDNA was synthesized from 1 μg of total RNA using Super Script II Reverse Transcriptase (GIBCO BRL, San Diego, California, United States) primed with a mixture of oligo-dT and random hexamers. Reactions lacking the reverse transcriptase served as a control for amplification of genomic DNA. Representative genes encoding enzymes in lipid metabolism pathways were studied together with *Riken32G18,* including *Scd1, Fas,* and phospholipid transfer protein *(Pltp).* The housekeeping gene β-*actin* was used as a normalization control.

LXRα agonist treatment of mice. BTBR male 7-wk-old mice were fed a chow diet (Harlan 7001 4% fat rodent chow) supplemented with DMSO vehicle or DMSO + T0901317 (0.025% by weight; Tularik, San Francisco, California, United States) for 7 d. T0901317 is a non-steroidal LXR agonist that has previously been shown to potently transactivate LXR-regulated genes in mice [49]. After a 4-h fast, liver samples were collected and frozen in liquid nitrogen.

Supporting Information

Figure S1. LOD Profiles of ATP Synthase Subunit mRNA Traits

The traits exhibiting a linkage peak on Chromosome 2 are grouped together on top rows.

Found at DOI: 10.1371/journal.pgen.0020006.sg001 (477 KB PDF).

Figure S2. LOD Profiles of 38 Seed Traits That Produce the 174 Correlated Traits Belonging to the GPCR Pathway

Found at DOI: 10.1371/journal.pgen.0020006.sg002 (495 KB PDF).

Figure S3. LOD Profiles of GPCR eQTL

(A) LOD profiles of GPCR protein signaling pathway mRNA traits related to obesity. The LOD profiles were generated on WebQTL. The empirical *P*-values are generated by permutation test (1,000 permutations); the lower line indicates suggestive linkage and the upper line indicates significant linkage. The green line indicates BTBR alleles increase the trait values; the red line indicates that B6 alleles increase the trait values.

(B) LOD profiles of GPCR protein signaling pathway mRNA traits related to obesity, from the GABA receptor subunit family. The LOD profiles were generated on WebQTL. The empirical *P*-values are generated by permutation test (1,000 permutations); the lower line indicates suggestive linkage and the upper line indicates significant linkage. The green line indicates BTBR alleles increase the trait values; the red line indicates that B6 alleles increase the trait values.

Found at DOI: 10.1371/journal.pgen.0020006.sg003 (3.2 MB PDF).

Table S1. *cis* Traits Inferred by Using LOD Greater Than 3.4

Columns A, B, and C show the Affymetrix probe set IDs, gene names, and the gene symbols of the *cis* traits. Columns D and E show the physical chromosome and position (Mb) of the traits. Column F shows the maximum LOD scores of the *cis* traits.

Found at DOI: 10.1371/journal.pgen.0020006.st001 (46 KB TXT).

Table S2. Enriched GO Categories for Expression Traits That Are Correlated with the Seed Traits

Using each of the 6,016 transcripts as a seed, Pearson correlation coefficients were calculated against the log-transformed expression intensities (raw intensities were normalized by RMA) of all of the 45,037 transcripts in the 60 F_2 mice. For each seed, we collected the traits with correlation coefficients of 0.7 or greater, which form the list of correlated traits for that seed. Enriched functional groups, defined as GO categories corresponding to BP, were identified from each list with *P*-values computed using a custom variant of "GOHyperG" in the "GOstats" package (Bioconductor Core Team 2004; http://www.bioconductor.org). For each list, we examined up to

50 GO terms, each of which involves more than ten genes and has a *P*-value of ≤ 0.01. Column A contains the Affymetrix probe set identifiers. Column B contains the names of the genes. Columns C through AZ are the GO terms identified, and Columns BA through CX are the identifiers of these GO terms.

Found at DOI: 10.1371/journal.pgen.0020006.st002 (948 KB TXT).

Table S3. GO Enrichments of Traits That Are Correlated with Seed Traits

From the lists of traits correlated with the seed traits in Table S2, we identified 15,726 BP GO terms in total, among which 862 are unique. For each of the unique GO terms, we traced back the lists of correlated traits that were enriched ($P = 0.01$), and we concatenated all such lists. Columns A and B show the GO category identifiers and the names, respectively. Columns C and D show the probe set IDs and gene names of the correlated traits. Column E shows whether ("T" or "F") the traits were annotated as belonging to the corresponding GO terms in the public database.

Found at DOI: 10.1371/journal.pgen.0020006.st003 (8.3 MB TXT).

Table S4. Descriptions of the 174 GPCR Protein Signaling Pathway Traits

Columns A, B, and C show the Affymetrix probe set IDs, gene symbols, and gene titles of the 174 gene traits. Columns D and E show the physical chromosomes and positions (Mb) of the 174 traits. Columns F, G, and H show the chromosomes, positions (Mb), and LOD scores of the maximum LOD peaks of those 174 traits.

Found at DOI: 10.1371/journal.pgen.0020006.st004 (12 KB TXT).

Table S5. Top 100 Traits Whose Expression Levels Are Correlated with Those of *Scd1* in the 60 F_2 Mice

Column A shows the Affymetrix probe set IDs. Columns B and C show the gene symbols and the titles of the genes. Columns D and E show the genomic locations of these traits. Column F shows the Pearson correlation coefficients. Column G shows the unadjusted *P*-values for the correlations.

Found at DOI: 10.1371/journal.pgen.0020006.st005 (7 KB TXT).

Table S6. Descriptions of the Gene Ontology "BP" Terms That Are Enriched in the 174 GPCR Protein Signaling Pathway Traits

Columns A and B show the GO category identifiers and the names. Column C shows the enrichment *P*-values (rounded to five decimal places) calculated using Hypergeometric distribution. Column D shows the GO category size.

Found at DOI: 10.1371/journal.pgen.0020006.st006 (1 KB TXT).

Accession Numbers

The original genotypes, physiologic phenotypes, and microarray-derived gene expression data are available in WebQTL (http://www.genenetwork.org; group: B6BTBRF2 [43]) and in GEO (http://www.ncbi.nlm.nih.gov/geo; accession number GSE3330).

The Gene Ontology (www.geneontology.org) accession numbers are GPCR protein signaling pathway (0007186) and *Scd1* (1415965_at).

Acknowledgments

We thank Mark Craven, Ping Wang, Keith Noto, Dan Klass, and Elias Chaibub Neto for technical assistance and helpful discussions. The authors are grateful to Yanhua Qu and Jintao Wang for depositing our data into WebQTL. This research was supported in part by American Diabetes Association grant 7–03-IG-01, National Institutes of Health (NIH)/National Institute of Diabetes and Digestive and Kidney Diseases grants 5803701 and 66369, and grant HL56593. We are grateful to the Center for Inherited Disease Research Genotyping Laboratory for their genotyping under an NIH-supported project.

Author contributions. HL, JBF, BSY, MTF, CK, and ADA conceived and designed the experiments. HL, JBF, DSS, ETKM, MTF, and KLS performed the experiments. HL, MC, JBF, BSY, CMM, MTF, CK, and ADA analyzed the data. HL, JBF, BSY, KFM, and RWW contributed reagents/materials/analysis tools. HL, MC, JBF, BSY, CK, and ADA wrote the paper.

Competing interests. The authors have declared that no competing interests exist. ■

References

1. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, et al. (2004) Integrative annotation of 21,037 human genes validated by full-length cDNA clones. PLoS Biol 2: e162. DOI: 10.1371/journal.pbio.0020162
2. Lan H, Stoehr JP, Nadler ST, Schueler KL, Yandell BS, et al. (2003) Dimension reduction for mapping mRNA abundance as quantitative traits. Genetics 164: 1607–1614.
3. Zhang W, Morris QD, Chang R, Shai O, Bakowski MA, et al. (2004) The functional landscape of mouse gene expression. J Biol 3: 21.
4. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95: 14863–14868.
5. Brem RB, Yvert G, Clinton R, Kruglyak L (2002) Genetic dissection of transcriptional regulation in budding yeast. Science 296: 752–755.
6. Damerval C, Maurice A, Josse JM, de Vienne D (1994) Quantitative trait loci underlying gene product variation: A novel perspective for analyzing regulation of genome expression. Genetics 137: 289–301.
7. Bystrykh L, Weersing E, Dontje B, Sutton S, Pletcher MT, et al. (2005) Uncovering regulatory pathways that affect hematopoietic stem cell function using 'genetical genomics.' Nat Genet 37: 225–232.
8. Hubner N, Wallace CA, Zimdahl H, Petretto E, Schulz H, et al. (2005) Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. Nat Genet 37: 243–253.
9. Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, et al. (2003) Genetics of gene expression surveyed in maize, mouse and man. Nature 422: 297–302.
10. Yvert G, Brem RB, Whittle J, Akey JM, Foss E, et al. (2003) Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. Nat Genet 35: 57–64.
11. Chesler EJ, Lu L, Shou S, Qu Y, Gu J, et al. (2005) Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. Nat Genet 37: 233–242.
12. Lan H, Rabaglia ME, Stoehr JP, Nadler ST, Schueler KL, et al. (2003) Gene expression profiles of nondiabetic and diabetic obese mice suggest a role of hepatic lipogenic capacity in diabetes susceptibility. Diabetes 52: 688–700.
13. Stoehr JP, Nadler ST, Schueler KL, Rabaglia ME, Yandell BS, et al. (2000) Genetic obesity unmasks nonlinear interactions between murine type 2 diabetes susceptibility loci. Diabetes 49: 1946–1954.
14. Jin C, Lan H, Attie AD, Churchill GA, Bulutuglo D, et al. (2004) Selective phenotyping for increased efficiency in genetic mapping studies. Genetics 168: 2285–2293.
15. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4: 249–264.
16. Dupuis J, Siegmund D (1999) Statistical methods for mapping quantitative trait loci from a dense set of markers. Genetics 151: 373–386.
17. Kendziorski CM, Chen M, Yuan M, Lan H, Attie AD (2005) Statistical methods for expression quantitative trait loci (eQTL) mapping. Biometrics (in press).
18. Cohen P, Miyazaki M, Socci ND, Hagge-Greenberg A, Liedtke W, et al. (2002) Role for stearoyl-CoA desaturase-1 in leptin-mediated weight loss. Science 297: 240–243.
19. Ntambi JM, Miyazaki M, Stoehr JP, Lan H, Kendziorski CM, et al. (2002) Loss of stearoyl-CoA desaturase-1 function protects mice against adiposity. Proc Natl Acad Sci U S A 99: 11482–11486.
20. Frolov A, Schroeder F (1998) Acyl coenzyme A binding protein. Conformational sensitivity to long chain fatty acyl-CoA. J Biol Chem 273: 11049–11055.
21. Sandberg MB, Bloksgaard M, Duran-Sandoval D, Duval C, Staels B, et al. (2005) The gene encoding acyl-CoA-binding protein is subject to metabolic regulation by both sterol regulatory element-binding protein and peroxisome proliferator-activated receptor alpha in hepatocytes. J Biol Chem 280: 5258–5266.
22. Lin J, Yang R, Tarr PT, Wu PH, Handschin C, et al. (2005) Hyperlipidemic effects of dietary saturated fats mediated through PGC-1beta coactivation of SREBP. Cell 120: 261–273.
23. Nadler ST, Stoehr JP, Schueler KL, Tanimoto G, Yandell BS, et al. (2000) The expression of adipogenic genes is decreased in obesity and diabetes mellitus. Proc Natl Acad Sci U S A 97: 11371–11376.
24. Shimomura I, Matsuda M, Hammer RE, Bashmakov Y, Brown MS, et al. (2000) Decreased IRS-2 and increased SREBP-1c lead to mixed insulin resistance and sensitivity in livers of lipodystrophic and ob/ob mice. Mol Cell 6: 77–86.
25. Horton JD, Goldstein JL, Brown MS (2002) SREBPs: Activators of the complete program of cholesterol and fatty acid synthesis in the liver. J Clin Invest 109: 1125–1131.
26. Jerez-Timaure NC, Kearney F, Simpson EB, Eisen EJ, Pomp D (2004) Characterization of QTL with major effects on fatness and growth on mouse chromosome 2. Obes Res 12: 1408–1420.
27. Mehrabian M, Wen PZ, Fisler J, Davis RC, Lusis AJ (1998) Genetic loci

controlling body fat, lipoprotein metabolism, and insulin levels in a multifactorial mouse model. J Clin Invest 101: 2485–2496.
28. Taylor BA, Phillips SJ (1997) Obesity QTLs on mouse chromosomes 2 and 17. Genomics 43: 249–257.
29. Farahani P, Fisler JS, Wong H, Diament AL, Yi N, et al. (2004) Reciprocal hemizygosity analysis of mouse hepatic lipase reveals influence on obesity. Obes Res 12: 292–305.
30. Lembertas AV, Perusse L, Chagnon YC, Fisler JS, Warden CH, et al. (1997) Identification of an obesity quantitative trait locus on mouse chromosome 2 and evidence of linkage to body fat and insulin on the human homologous region 20q. J Clin Invest 100: 1240–1247.
31. Estrada-Smith D, Castellani LW, Wong H, Wen PZ, Chui A, et al. (2004) Dissection of multigenic obesity traits in congenic mouse strains. Mamm Genome 15: 14–22.
32. Stoehr JP, Byers JE, Clee SM, Lan H, Boronenkov IV, et al. (2004) Identification of major quantitative trait loci controlling body weight variation in ob/ob mice. Diabetes 53: 245–249.
33. Weiland TJ, Voudouris NJ, Kent S (2004) The role of CCK2 receptors in energy homeostasis: Insights from the CCK2 receptor-deficient mouse. Physiol Behav 82: 471–476.
34. Kushi A, Sasai H, Koizumi H, Takeda N, Yokoyama M, et al. (1998) Obesity and mild hyperinsulinemia found in neuropeptide Y-Y1 receptor-deficient mice. Proc Natl Acad Sci U S A 95: 15659–15664.
35. Ohki-Hamazaki H, Watase K, Yamamoto K, Ogura H, Yamano M, et al. (1997) Mice lacking bombesin receptor subtype-3 develop metabolic defects and obesity. Nature 390: 165–169.
36. Huang XF, Yu Y, Zavitsanou K, Han M, Storlien L (2005) Differential expression of dopamine D2 and D4 receptor and tyrosine hydroxylase mRNA in mice prone, or resistant, to chronic high-fat diet-induced obesity. Brain Res Mol Brain Res 135: 150–161.
37. Fang YJ, Thomas GN, Xu ZL, Fang JQ, Critchley JA, et al. (2005) An affected pedigree member analysis of linkage between the dopamine D2 receptor gene TaqI polymorphism and obesity and hypertension. Int J Cardiol 102: 111–116.
38. Rosmond R, Bouchard C, Bjorntorp P (2002) Association between a variant at the GABA(A)alpha6 receptor subunit gene, abdominal obesity, and cortisol secretion. Ann N Y Acad Sci 967: 566–570.
39. Vionnet N, Hani EH, Lesage S, Philippi A, Hager J, et al. (1997) Genetics of NIDDM in France: Studies with 19 candidate genes in affected sib pairs. Diabetes 46: 1062–1068.
40. Blasi C (2000) Influence of benzodiazepines on body weight and food intake in obese and lean Zucker rats. Prog Neuropsychopharmacol Biol Psychiatry 24: 561–577.
41. Ebert MH, Schmidt DE, Thompson T, Butler MG (1997) Elevated plasma gamma-aminobutyric acid (GABA) levels in individuals with either Prader-Willi syndrome or Angelman syndrome. J Neuropsychiatry Clin Neurosci 9: 75–80.
42. Lucignani G, Panzacchi A, Bosio L, Moresco RM, Ravasi L, et al. (2004) GABA A receptor abnormalities in Prader-Willi syndrome assessed with positron emission tomography and [11C]flumazenil. Neuroimage 22: 22–28.
43. Wang J, Williams RW, Manly KF (2003) WebQTL: Web-based complex trait analysis. Neuroinformatics 1: 299–308.
44. Jiang C, Zeng Z-B (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140: 1111–1127.
45. Schadt EE, Lamb J, Yang X, Zhu J, Edwards S, et al. (2005) An integrative genomics approach to infer causal associations between gene expression and disease. Nat Genet 37: 710–717.
46. Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19: 889–890.
47. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 100: 9440–9445.
48. Gentleman R (2005) Using GO for Statistical Analyses. Bioconductor Vignettes.
49. Schultz JR, Tu H, Luk A, Repa JJ, Medina JC, et al. (2000) Role of LXRs in control of lipogenesis. Genes Dev 14: 2831–2838.