Incorporating Mouse Genome
Mammalian Genome
Genes and Phenotypes

Combining gene expression QTL mapping and phenotypic spectrum analysis to uncover gene regulatory relationships

Lei Bao,[1,2] Lai Wei,[3,4] Jeremy L. Peirce,[2,4] Ramin Homayouni,[2,4] Hongqiang Li,[1,2] Mi Zhou,[1,2] Hao Chen,[5] Lu Lu,[2,4] Robert W. Williams,[2,4,6] Lawrence M. Pfeffer,[3] Dan Goldowitz,[2,4] Yan Cui[1,2]

[1]Department of Molecular Sciences, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA
[2]Center of Genomics and Bioinformatics, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA
[3]Department of Pathology and Laboratory Medicine, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA
[4]Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA
[5]Department of Pharmacology, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA
[6]Department of Pediatrics, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA

Received: 30 November 2005 / Accepted: 21 February 2006

Abstract

Gene expression QTL (eQTL) mapping can suggest candidate regulatory relationships between genes. Recent advances in mammalian phenotype annotation such as mammalian phenotype ontology (MPO) enable systematic analysis of the phenotypic spectrum subserved by many genes. In this study we combined eQTL mapping and phenotypic spectrum analysis to predict gene regulatory relationships. Five pairs of genes with similar phenotypic effects and potential regulatory relationships suggested by eQTL mapping were identified. Lines of evidence supporting some of the predicted regulatory relationships were obtained from biological literature. A particularly notable example is that promoter sequence analysis and real-time PCR assays support the predicted regulation of protein kinase C epsilon (*Prkce*) by cAMP responsive element binding protein 1 (*Creb1*). Our results show that the combination of gene eQTL mapping and phenotypic spectrum analysis may provide a valuable approach to uncovering gene regulatory relations underlying mammalian phenotypes.

Introduction

Uncovering the gene networks underlying complex phenotypes is an important goal of mammalian genetics and genomics. Mammalian phenotypes are often polygenic, involving multiple interacting genes. Many allele-phenotype relationships have been discovered by studying the effects of naturally occurring genetic variations or experimentally induced mutations (Badano and Katsanis 2002; Gkoutos et al. 2004; McKusick 1998; Nadeau et al. 2001; Rikke and Johnson 1998; Skarnes et al. 2004; Smith et al. 2005; Stenson et al. 2003). Such allele-phenotype relationships provide crucial insights into the genetic basis of phenotypes. Recently, the mammalian phenotype ontology (MPO) project (Smith et al. 2005) has emerged as an effort to standardize the description and annotation of phenotypes and provide a central depository of the known allele-phenotype relationships accumulated during the past decades. MPO uses a hierarchically structured vocabulary to describe phenotypes. The structure of MPO is a directed acyclic graph (DAG) in which child nodes represent more detailed classifications of the phenotypes than those represented by their parent nodes. Genes (alleles) are annotated to the nodes according to the experimentally derived allele-phenotype relationships. Collectively, the MPO nodes associated with a gene represent the phenotypic spectrum of that gene. Transcriptional regulation is central to variation in many phenotypes, so regulatory interactions between two genes may result in significant similarities in the spectrum of phenotypes the genes affect. Similarly, similarities in the phenotypes of different genes often suggest functional relationships between those genes (Asthana et al. 2004; Bao and Sun 2002; Clare and King 2002; Gunsalus et al. 2005; King et al. 2003). By using yeast phenotype data and a supervised learning approach,

Correspondence to: Yan Cui, Department of Molecular Sciences, University of Tennessee Health Science Center, 858 Madison Avenue, Memphis, TN 38163, USA; E-mail: ycui2@utmem.edu

Clare and King (2002) were able to predict functions of 83 uncharacterized yeast genes with an estimated accuracy of 80%. By studying the early embryogenesis of *C. elegans*, Gunsalus et al. (2005) found that proteins clustered by phenotypic similarity are more likely to interact physically with one another. Therefore, a pair of genes annotated with similar phenotypes when disrupted may be considered candidates for involvement in the same pathways or in some type of interaction, including transcriptional regulation. Obviously, independent data are required for further evaluation of the candidate regulatory relationships derived from phenotypic spectrum analysis.

Gene expression QTL (eQTL) mapping provides another approach to identify candidate regulatory relations (Bing and Hoeschele 2005; Brem and Kruglyak 2005; Brem et al. 2002, 2005; Bystrykh et al. 2005; Chesler et al. 2005; de Koning and Haley 2005; de Koning et al. 2005; Doss et al. 2005; Hubner et al. 2005; Jansen and Nap 2001; Kirst et al. 2004; Li and Burmeister 2005; Li et al. 2005, 2006 Mehrabian et al. 2005; Monks et al. 2004; Morley et al. 2004; Ronald et al. 2005; Schadt et al. 2003, 2005; Storey et al. 2005; Wayne and McIntyre 2002; Yvert et al. 2003; Zhu et al. 2004). In this approach a population is genotyped at a number of markers, and variations in gene expression levels in a tissue or structure are measured across the population using microarrays. These gene expression levels are then mapped as quantitative traits, and QTL regulating gene expression are identified. The genes in a QTL interval are potential regulators for the genes whose expression levels are influenced by this QTL. It is notable that the regulatory relationship we are referring to here does not have to be direct interactions in the usual sense; they can be indirect regulatory response to changes in the system that result from a genetic polymorphism.

In this work we combined phenotypic analysis and eQTL mapping to identify the regulatory relations underlying mammalian phenotypes. Currently, tissue-specific (forebrain and hematopoietic stem cell) gene expression data are available for mouse BXD recombinant inbred (RI) strains (Bystrykh et al. 2005; Chesler et al. 2005). Two classes of phenotypes related to these tissues, behavior/neurologic phenotypes and hematopoietic system phenotypes, were analyzed. We selected genes with similar phenotypic spectra based on the phenotype ontology annotations, and then assessed whether a regulatory relationship between the genes is supported by the eQTL mapping result. We predicted five candidate regulatory relations, two of which are supported by lines of evidence in the published literature. Transcription factor binding site analysis and real-time PCR assays further support the predicted regulation of *Prkce* (protein kinase C, epsilon) by *Creb1* (cAMP responsive element binding protein 1) in PC12 cells.

Material and methods

QTL mapping and significance testing. All genotype and gene expression data were obtained from BXD RI strains generated from C57BL/6J × DBA/2J progenitors. The genotypes of these BXD strains were characterized using 3795 markers that have been carefully error-checked (Williams et al. 2001; Wiltshire et al. 2003). Two published microarray data sets were used: gene expression data of forebrains of 32 BXD strains (Chesler et al. 2005) and that of hematopoietic stem cells from 30 BXD strains (Bystrykh et al. 2005). The forebrain data of two strains (BXD67 and BXD68) that were not fully inbred were excluded from the analysis. The original microarray data were processed using the Robust Multichip Average (RMA) method of background correction, quantile normalization, and summarization of cell signal intensity (Bolstad et al. 2003). All the data are available from the Genenetwork website at http://www.genenetwork.org.

Genome-wide QTL mapping was carried out for each transcript in the forebrain or hematopoietic stem cell data set using a standard marker regression protocol (Manly et al. 2001). Briefly, the regression model estimates the additive effects of alleles:

$$y_i = b_0 + b_1 x_i + e_i$$

where y_i is the expression trait, x_i is the marker genotype, and e_i is the random effect for the ith individual. The largest likelihood ratio statistic (LRS) of each transcript across the genome was determined and the corresponding empirical genome-adjusted p value (Churchill and Doerge 1994) was estimated using 5000 independent permutations of the original transcript trait values. An eQTL was considered significant if the $p < 0.05$. The QTL intervals were estimated using the 1.5 log-of-odds[1] (LOD) rule (Dupuis and Siegmund 1999; Lander and Botstein 1989).

Similarity between phenotypic spectra. Mammalian phenotype (MP) annotations of mouse genes were retrieved from the Mammalian Phenotype Browser (May 18, 2005, data freeze) (Smith et al.

[1]The relation between log of odds (LOD) and the likelihood ratio statistic (LRS) is: LOD = LRS/4.61.2.

2005). For the forebrain data set, we used the annotations descending from the MP term "behavior/neurological phenotype" (MP:0005386). For the hematopoietic stem cell data set, we used those annotations descending from of the MP term "hematopoietic system phenotype" (MP:0005397). Noting that annotations were usually assigned at the most specific levels, we recursively included annotations implied by the structure of the mammalian phenotype ontology, so that if a MP term was associated with a gene, then all the ancestors of the MP term should also be associated with that gene. The phenotypic similarity between two genes was computed using the following formula:

$$\text{PhenoSim}(i,j) = \sum_{c \in \text{CMP}} S(c) \quad (1)$$

where CMP is the set of all the common MP terms associated with the two genes i and j, satisfying the criterion that no MP term is the ancestor of the other MP terms in this set. $S(c)$ is the score of a MP term c and is defined as

$$S(c) = -\log\left(\frac{N_c}{N}\right) \quad (2)$$

where N_c is the number of genes associated with the MP term c and N is the number of genes associated with the root MP term of a selected branch of the MPO DAG (i.e., MP:0005386 or MP:0005397). The score decreases as the MP term approaches the root, and the root MP term has the minimum score of zero (Lord et al. 2003). Based on this scoring method, the more specific MP terms two genes share, the higher the phenotypic similarity between the two genes.

Transcription factor binding site analysis. The 1000-bp upstream regions of the target genes were extracted from the mouse genome annotation database (Kent et al. 2002). The MATCH program (Kel et al. 2003) was used to assess whether there were any putative transcription factor binding sites for predicted modulators. The computed matrix similarity scores were compared with the preset cutoffs that minimize the sum of false-positive and false-negative error rates.

Biological reagents and cell culture. PC12 and NIH 3T3 cells were obtained from ATCC. Forskolin was obtained from Fisher (No. 053155). PC12 cells were plated at 1×10^4 cells/cm^2 every five days in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% DCS (Hyclone Laboratories, Logan, UT), 100 μg/ml penicillin G, and 100 μg/ml streptomycin.

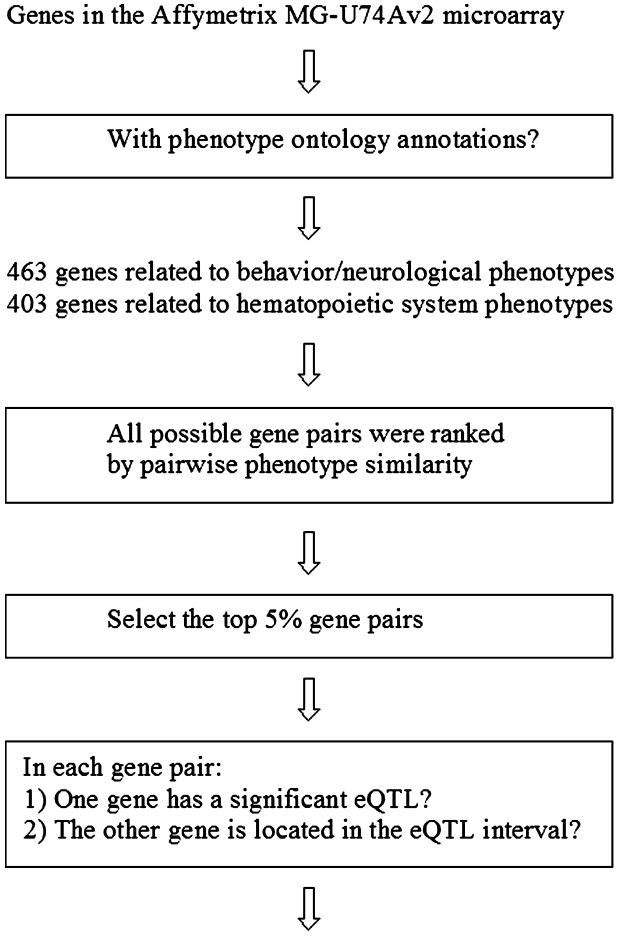

Fig. 1. The procedure for identifying gene regulatory relationships.

Quantitative Real-Time PCR. Total RNA was isolated from PC12 cells using TRIzol reagent (Invitrogen, Carlsbad, CA). Quantitative real-time PCR (RT-PCR) was performed on a SmartCycler (Cepheid, Sunnyvale, CA) using the AccessQuick™ RT-PCR system (Promega, Madison, WI) and SYBR green I (Molecular Probes, Eugene, OR) according to the manufacturers' instructions. The following forward and reverse primers were used for each gene respectively: *Prkce*, 5′-AGGAGATTGACTGGGTACTG-3′, 5′-AGTGTAAGTATTGGCTCTTC-3′; β-*actin*, 5′-AAGGAGATTACTGCTCTGGC-3′, 5′-ACATC TGCTGGAAGGTGGAC-3′. Reverse transcription was performed at 48°C for 45 min. RT-PCR cycling parameters were as follows: denaturation at 95°C for 2 min, amplification at 94°C for 30 sec, and 62°C for 30 sec for 35 cycles. The product size was initially monitored by agarose gel electrophoresis and melting curves were analyzed to control for specificity of

Table 1. Predicted regulatory relationships

Target gene symbol	*Target gene description*	*eQTL p value*[a]	*Regulator gene symbol*	*Regulator gene description*	*Common MP term*
Smn1	survival motor neuron 1	0.02	*Cpe*	carboxypeptidase E	MP:0000745 tremors MP:0001431 abnormal feeding behavior MP:0003313 abnormal locomotor activation
Prkce	protein kinase C, epsilon	0.01	*Creb1*	cAMP responsive element binding protein 1	MP:0002065 abnormal fear/anxiety-related behavior MP:0002069 abnormal feeding/drinking behavior
Avpr2	arginine vasopressin receptor 2	0.05	*Prkce*	protein kinase C, epsilon	MP:0001422 abnormal drinking behavior
Tfrc	transferrin receptor	0.03	*Epb4.1*	erythrocyte protein band 4.1	MP:0001585 hemolytic anemia
Ada	adenosine deaminase	0.03	*Hmgb3*	high mobility group box 3	MP:0002124 abnormal red blood cell MP:0002398 abnormal bone marrow cell Morphology/development MP:0005460 abnormal leucopoiesis

[a]Empirical p value of the eQTL from 5000 permutation tests.

PCR reactions. The relative abundance of gene expression was normalized to the expression of the housekeeping gene β-*actin*. The relative units were calculated from a standard curve, plotting three different concentrations against the PCR cycle number at which the measured intensity reaches a fixed value (with a tenfold increment equivalent to approximately 3.1 cycles).

Results and discussion

There were 463 genes on the U74Av2 microarray that were annotated to behavior/neurologic phenotypes and 403 annotated to hematopoietic system phenotypes. The procedure of our analysis is shown in Fig. 1. We computed the phenotypic similarity scores between each gene pair and selected the top 5% of gene pairs with the highest similarity scores for further analysis. For a pair of genes i and j with high phenotypic similarity, if the expression trait of gene i maps to the genomic region where gene j is located, then gene j is considered a candidate regulator of gene i and vice versa. This strategy resulted in the discovery of three regulatory relationships underlying behavior/neurologic phenotypes and two underlying hematopoietic system phenotypes (Table 1). The list of eQTL controlling expression traits of genes in the top 5% gene pairs can be found in the supplementary tables.

Using our method, we predicted that *Creb1* is a regulator of protein kinase C epsilon (*Prkce*) (Table 1). Some independent lines of evidence also support this prediction (Fig. 2). *Creb1* encodes a transcription factor that binds to the cAMP-responsive element and induces transcription in response to stimulation of the cAMP pathway. Using sequence analysis, we identified a putative *Creb1* binding site in the promoter region of *Prkce* (Table 2). In addition, a recent study has described an interaction between cAMP and Prkce pathways (Parada et al. 2005). Parada and colleagues found that hyperalgesia is enhanced by the recruitment of the cAMP/PKCepsilon signaling pathway in addition to the usual cAMP/PKA pathway (Parada et al. 2005).

To test the predicted regulatory relationship between *Creb1* and *Prkce* experimentally, quantitative real-time PCR assays were performed for *Prkce* using RNA from untreated and Forskolin -treated (25 μM, 1 h) PC12 cells. Forskolin is a well-known activator of *Creb1*. Expression of *Prkce* (Fig. 3) in forskolin-treated cells was approximately 40% of that in untreated cells ($p = 0.02$, Student's t-test). Consistent with our prediction, this result suggests that *Creb1* influences the expression of *Prkce*. To test the specificity of these results, we performed the same PCR analysis in the NIH 3T3 cells. In contrast to the negative regulation of *Prkce* by *Creb1* in PC12 cells, similar *Prkce* mRNA levels were detected in untreated and forskolin-treated (25 μM, 1 h) 3T3 cells (Fig. 3) ($p = 0.4$, Student's t-test), indicating that the regulatory relationship between *Prkce* and *Creb1* is cell-type specific. Many genes have upstream cAMP

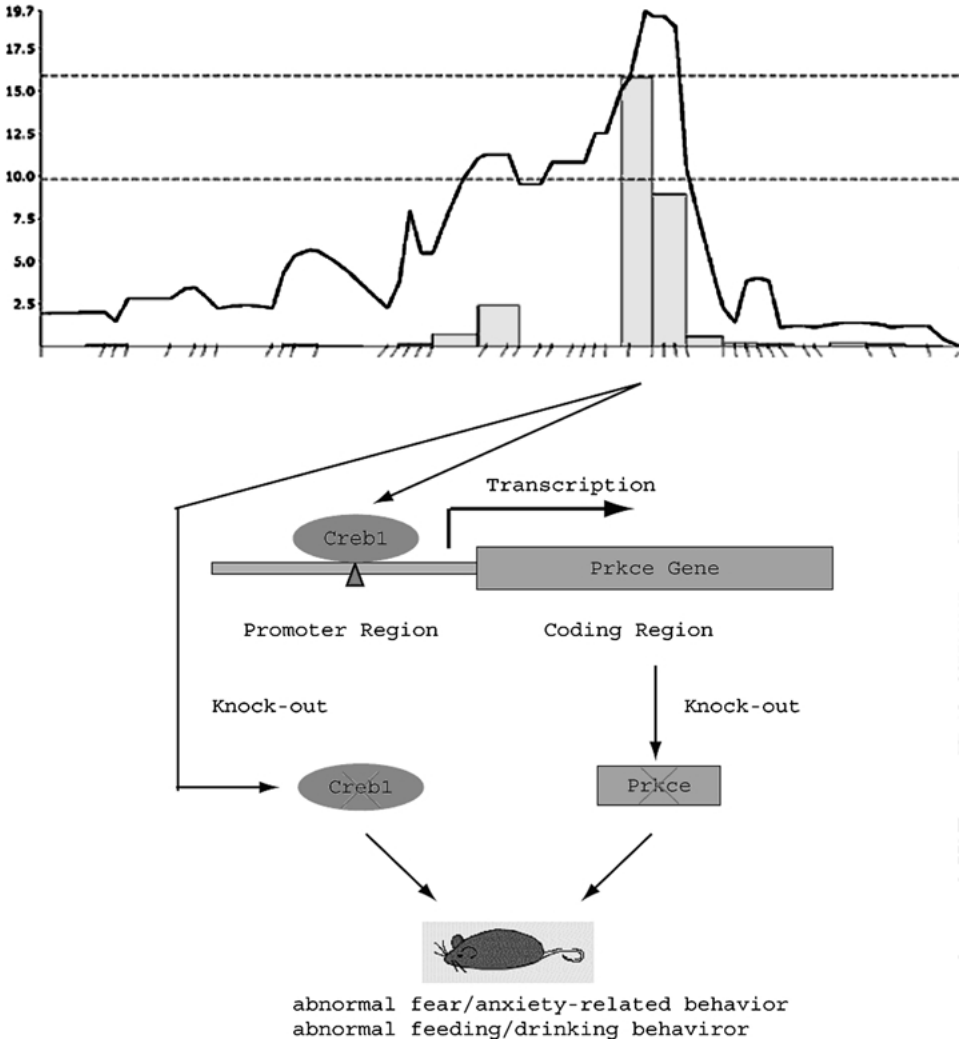

Fig. 2. Evidence showing that *Creb1* regulates expression of *Prkce*. **Upper panel**: The *Creb1* gene is located in the eQTL interval that influences *Prkce* expression. **Middle panel**: Promoter sequence analysis identifies a putative *Creb1* binding site in the promoter region of *Prkce*. **Lower panel**: Mutagenesis studies show that disruption of either *Creb1* or *Prkce* results in both abnormal feeding/drinking behavior (MP:0002069) and abnormal fear/anxiety-related behavior (MP:0002065) (Smith et al. 2005).

response elements. A recent genome-wide Creb1 occupancy study showed that about 17% of human genes possess cAMP response elements; however, only less than 2% of these genes actually respond to cAMP agonist forskolin in a tissue-specific manner (Zhang et al. 2005). Thus, about 0.34% of the human genes might be induced by cAMP in a specific cellular context. Assuming comparable effects of Creb1 in mouse, the chance of observing forskolin-induced expression change (like that we observed in *Prkce*) of a randomly selected gene would be very small.

Prkce belongs to a family of serine- and threonine-specific protein kinases, PKC, which can be activated by calcium and the second messenger diacylglycerol. It has been shown to be involved in a myriad of cellular functions such as neurite induction (Zeidman et al. 2002), apoptosis (Basu et al. 2002), cardioprotection from ischemia (McCarthy et al. 2005), and heat shock response (Wu et al. 2003). Knockout studies in mice suggest that Prkce is important for lipopolysaccharide (LPS)-mediated signaling in activated macrophages and may also

Table 2. Sequence analysis predicts a *Creb1* binding site in the *Prkce* promoter region

Target gene symbol	*Regulatory gene symbol*	*Matrix accession*	*Position*[a]	*Orient*	*Match score*	*minSUM cutoff*[b]	*minFN cutoff*[c]
Prkce	*Creb1*	V$CREB_Q3	−643	–	0.999	0.92	0.89

Match program (Kel et al. 2003) was used to determine putative *Creb1* binding sites in the promoter regions of *Prkce*.
[a]Nucleotide position upstream of the target gene's transcription start site.
[b]The minSUM cutoff is preset by the Match program to minimize the sum of false-positive and false-negative rates.
[c]The minFN cutoff is preset by the Match program to tolerate less than 10% false-negative rate.

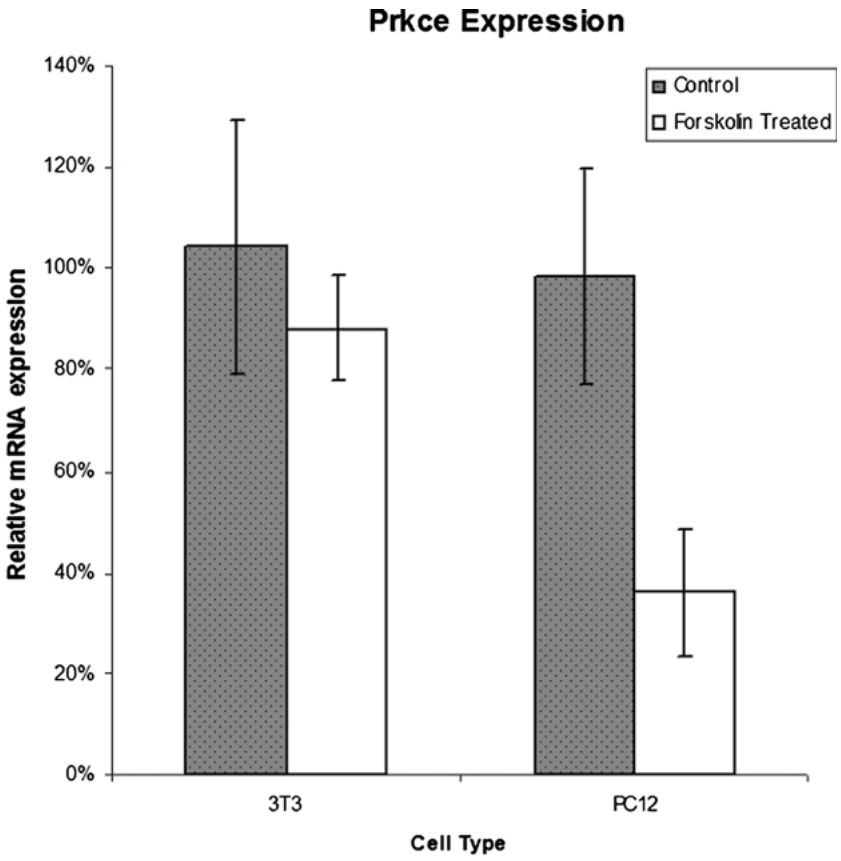

Fig. 3. *Creb1* influences the expression of *Prkce*. Quantitative real-time PCR assays for *Prkce* were performed on RNA prepared from 3T3 and PC12 cells that were either not treated or treated with 25 μM forskolin for 1 h. The mRNA was from three independent experiments. Real-time PCR was performed in triplicate. The results were normalized to β-*actin* expression and were expressed as percentage relative to the level of expression in the absence of forskolin treatment.

play a role in controlling anxiety-like behavior (Aksoy et al. 2004; Gruber et al. 2005; Khasar et al. 1999). Creb1 is a member of a transcription factor family consisting of CREB, CREM, and ATF (reviewed in Mayr and Montminy 2001; Shaywitz and Greenberg 1999). While PKC was found to phosphorylate and activate *Creb1*, a direct link between *Prkce* and *Creb1* is not clearly established. In this work, we found that forskolin treatment, which is known to activate *Creb1*, inhibited expression of *Prkce*. These results indicate that the *Creb1* may be part of an autoregulatory feedback pathway that downregulates *Prkce* expression to control activation of *Creb1*. Interestingly, the regulation of *Prkce* by *Creb1* was found in PC12 cells but not in NIH3T3 cells. Further studies will be needed to characterize the cell-type specificity of this regulation.

We also predicted that *Prkce* is a regulator of *Avpr2*, a type 2 arginine vasopressin receptor (Table 1). Interestingly, DeCoy et al. (1995) found that *Prkce* is involved in regulating the arginine vasopressin effect on salt transport and that downregulation of *Prkce* enhances and sustains arginine vasopressin-stimulated Na^+ absorption.

Identification of gene regulatory relationships underlying mammalian phenotypes is crucially important for uncovering the molecular basis of allele-phenotype relationships. Phenotypic similarities between genes have been shown to be highly suggestive of functional relationships (Asthana et al. 2004; Clare and King 2002; Gunsalus et al. 2005; King et al. 2003). The recently developed mammalian phenotype ontology makes the previously unformatted biological knowledge of phenotypes and their genetic causes ''computable'' by standardizing the description system (Smith et al. 2005). Our approach combines computational analysis of phenotype similarity and eQTL mapping to predict regulatory relationships related to two classes of mammalian phenotypes: ''behavior/neurological phenotype'' (MP:0005386) and ''hematopoietic system phenotype'' (MP:0005397). Currently, 34 high-level classes of phenotypes are included in the Mammalian Phenotype Browser (http://www.informatics.jax.org/searches/MP_form.shtml). These high-level phenotype classes are the immediate child nodes of the root. Among them, 21 phenotype classes are related to specific tissue types: adipose tissue phenotypes, behavior/neurologic phenotypes, cardiovascular system phenotypes, craniofacial phenotypes, digestive/alimentary, endocrine/exocrine gland phenotypes, hearing/ear phenotypes, hematopoietic system phenotypes, immune system phenotypes, limbs/digits/tail phenotypes, liver/biliary

system phenotypes, muscle phenotypes, nervous system phenotypes, pigmentation phenotypes, renal/urinary system phenotypes, reproductive system phenotypes, respiratory system phenotypes, skeleton phenotypes, skin/coat/nails phenotypes, taste/olfaction phenotypes, and vision/eye phenotypes. Our study focused on the two classes of phenotypes mentioned because microarray data from related tissues are available for a genetic reference population. In principle, the same genome-wide genetic study of gene expression can be carried out for any tissue. Thus, our method can be used to discover gene regulatory relationships underlying these phenotypes when the data for the related tissues become available. It should be noted that eQTL mapping indicates only transcriptional regulatory relationships. Other functional relationships (e.g., translational control) cannot be identified by this method. Furthermore, our method relies on known MP annotations of genes, hence limiting its application to uncharacterized gene pairs.

To some extent, MPO depends on its constructors' interpretation of the structure of biological knowledge. Changes in the MPO structure may affect the score of phenotypic similarity. However, as long as the structural changes do not affect the set of common MP terms shared by a gene pair, the score will remain unchanged. Our scoring scheme is robust against the changes in MPO structure between the root and the common MP terms, thus alleviate the uncertainty from the subjective factors of MPO.

Acknowledgments

The authors thank the reviewers for helpful suggestions. This work was supported by a PhRMA Foundation grant (YC); NIH grants HD52472 (DG), AA14425 (LL), AA13499 (RWW), and CA73753 (LMP); and by funds from the Muirhead Chair Endowment at the University of Tennessee Health Science Center (LMP).

References

1. Aksoy E, Goldman M, Willems F (2004) Protein kinase C epsilon: a new target to control inflammation and immune-mediated disorders. Int J Biochem Cell Biol 36, 183–188
2. Asthana S, King OD, Gibbons FD, Roth FP (2004) Predicting protein complex membership using probabilistic network reliability. Genome Res 14, 1170–1175
3. Badano JL, Katsanis N (2002) Beyond Mendel: an evolving view of human genetic disease transmission. Nat Rev Genet 3, 779–789
4. Bao L, Sun Z (2002) Identifying genes related to drug anticancer mechanisms using support vector machine. FEBS Lett 521, 109–114
5. Basu A, Lu D, Sun B, Moor AN, Akkaraju GR, et al. (2002) Proteolytic activation of protein kinase C-epsilon by caspase-mediated processing and transduction of antiapoptotic signals. J Biol Chem 277, 41850–41856
6. Bing N, Hoeschele I (2005) Genetical genomics analysis of a yeast segregant population for transcription network inference. Genetics 170, 533–542
7. Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19, 185–193
8. Brem RB, Kruglyak L (2005) The landscape of genetic complexity across 5,700 gene expression traits in yeast. Proc Natl Acad Sci USA 102, 1572–1577
9. Brem RB, Yvert G, Clinton R, Kruglyak L (2002) Genetic dissection of transcriptional regulation in budding yeast. Science 296, 752–755
10. Brem RB, Storey JD, Whittle J, Kruglyak L (2005) Genetic interactions between polymorphisms that affect gene expression in yeast. Nature 436, 701–703
11. 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
12. 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
13. Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138, 963–971
14. Clare A, King RD (2002) Machine learning of functional class from phenotype data. Bioinformatics 18, 160–166
15. de Koning DJ, Haley CS (2005) Genetical genomics in humans and model organisms. Trends Genet 21, 377–381
16. de Koning D-J, Carlborg O, Haley CS (2005) The genetic dissection of immune response using gene-expression studies and genome mapping. Vet Immunol Immunopathol 105, 343
17. DeCoy DL, Snapper JR, Breyer MD (1995) Anti sense DNA down-regulates proteins kinase C-epsilon and enhances vasopressin-stimulated Na+ absorption in rabbit cortical collecting duct. J Clin Invest 95, 2749–2756
18. Doss S, Schadt EE, Drake TA, Lusis AJ (2005) Cis-acting expression quantitative trait loci in mice. Genome Res 15, 681–691
19. Dupuis J, Siegmund D (1999) Statistical methods for mapping quantitative trait loci from a dense set of markers. Genetics 151, 373–386
20. Gkoutos GV, Green EC, Mallon AM, Hancock JM, Davidson D (2004) Building mouse phenotype ontologies. Pac Symp Biocomput , 178–189
21. Gruber T, Thuille N, Hermann-Kleiter N, Leitges N, Baier G (2005) Protein kinase Cepsilon is dispensable for TCR/CD3-signaling. Mol Immunol 42, 305–310

22. Gunsalus KC, Ge H, Schetter AJ, Goldberg DS, Han J-DJ, et al. (2005) Predictive models of molecular machines involved in Caenorhabditis elegans early embryogenesis. Nature 436, 861
23. 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
24. Jansen RC, Nap JP (2001) Genetical genomics: the added value from segregation. Trends Genet 17, 388–391
25. Kel AE, Gossling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, et al. (2003) MATCH: A tool for searching transcription factor binding sites in DNA sequences. Nucleic Acids Res 31, 3576–3579
26. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, et al. (2002) The human genome browser at UCSC. Genome Res 12, 996–1006
27. Khasar SG, Lin YH, Martin A, Dadgar J, McMahon T, et al. (1999) A novel nociceptor signaling pathway revealed in protein kinase C epsilon mutant mice. Neuron 24, 253–260
28. King OD, Lee JC, Dudley AM, Janse DM, Church GM, et al. (2003) Predicting phenotype from patterns of annotation. Bioinformatics 119 Suppl 1, i183–189
29. Kirst M, Myburg AA, De Leon JP, Kirst ME, Scott J, et al. (2004) Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus. Plant Physiol 135, 2368–2378
30. Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121, 185–199
31. Li H, Lu L, Manly KF, Chesler EJ, Bao L, et al. (2005) Inferring gene transcriptional modulatory relations: a genetical genomics approach. Hum Mol Genet 14, 1119–1125
32. Li H, Chen H, Bao L, Manly KF, Chesler EJ, et al. (2006) Integrative genetic analysis of transcription modules: towards filling the gap between genetic loci and inherited traits. Hum Mol Genet 15, 481–492
33. Li J, Burmeister M (2005) Genetical genomics: combining genetics with gene expression analysis. Hum Mol Genet 14, R163–R169
34. Lord PW, Stevens RD, Brass A, Goble CA (2003) Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation. Bioinformatics 19, 1275–1283
35. Manly KF, Cudmore RH Jr, Meer JM (2001) Map Manager QTX, cross-platform software for genetic mapping. Mamm Genome 12, 930–932
36. Mayr B, Montminy M (2001) Transcriptional regulation by the phosphorylation-dependent factor CREB. Nat Rev Mol Cell Biol 2, 599–609
37. McCarthy J, McLeod CJ, Minners J, Essop MF, Ping P, et al. (2005) PKCepsilon activation augments cardiac mitochondrial respiratory post-anoxic reserve—a putative mechanism in PKCepsilon cardioprotection. J Mol Cell Cardiol 38, 697–700
38. McKusick VA (1998) *Mendelian Inheritance in Man. A Catalog of Human Genes and Genetic Disorders* (Baltimore, MD: Johns Hopkins University Press)
39. Mehrabian M, Allayee H, Stockton J, Lum PY, Drake TA, et al. (2005) Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits. Nat Genet 37, 1224–1233
40. Monks SA, Leonardson A, Zhu H, Cundiff P, Pietrusiak P, et al. (2004) Genetic inheritance of gene expression in human cell lines. Am J Hum Genet 75, 1094–1105
41. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, et al. (2004) Genetic analysis of genome-wide variation in human gene expression. Nature 430, 743–747
42. Nadeau JH, Balling R, Barsh G, Beier D, Brown SD, et al. (2001) Sequence interpretation. Functional annotation of mouse genome sequences. Science 291, 1251–1255
43. Parada CA, Reichling DB, Levine JD (2005) Chronic hyperalgesic priming in the rat involves a novel interaction between cAMP and PKCepsilon second messenger pathways. Pain 113, 185–190
44. Rikke BA, Johnson TE (1998) Towards the cloning of genes underlying murine QTLs. Mamm Genome 9, 963–968
45. Ronald J, Brem RB, Whittle J, Kruglyak L (2005) Local regulatory variation in *Saccharomyces cerevisiae*. PLoS Genet 1, e25
46. 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
47. 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
48. Shaywitz AJ, Greenberg ME (1999) CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals. Annu Rev Biochem 68, 821–861
49. Skarnes WC, von Melchner H, Wurst W, Hicks G, Nord AS, et al. (2004) A public gene trap resource for mouse functional genomics. Nat Genet 36, 543–544
50. Smith CL, Goldsmith CA, Eppig JT (2005) The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. Genome Biol 6, R7
51. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, et al. (2003) Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat 21, 577–581
52. Storey JD, Akey JM, Kruglyak L (2005) Multiple locus linkage analysis of genomewide expression in yeast. PLoS Biol 3, e267
53. Wayne ML, McIntyre LM (2002) Combining mapping and arraying: An approach to candidate gene identification. Proc Natl Acad Sci USA 99, 14903–14906
54. Williams RW, Gu J, Qi S, Lu L (2001) The genetic structure of recombinant inbred mice: high-resolution

consensus maps for complex trait analysis. Genome Biol 2:RESEARCH0046
55. Wiltshire T, Pletcher MT, Batalov S, Barnes SW, Tarantino LM, et al. (2003) Genome-wide single-nucleotide polymorphism analysis defines haplotype patterns in mouse. Proc Natl Acad Sci USA 100, 3380–3385
56. Wu JM, Xiao L, Cheng XK, Cui LX, Wu NH, et al. (2003) PKC epsilon is a unique regulator for hsp90 beta gene in heat shock response. J Biol Chem 278, 51143–51149
57. 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
58. Zeidman R, Troller U, Raghunath A, Pahlman S, Larsson C (2002) Protein kinase Cepsilon actin-binding site is important for neurite outgrowth during neuronal differentiation. Mol Biol Cell 13, 12–24
59. Zhang XM, Odom DT, Koo SH, Conkright MD, Canettieri G, et al. (2005) Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. Proc Natl Acad Sci USA 102, 4459–4464
60. Zhu J, Lum PY, Lamb J, GuhaThakurta D, Edwards SW, et al. (2004) An integrative genomics approach to the reconstruction of gene networks in segregating populations. Cytogenet Genome Res 105, 363–374