nature genetics

Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease

Norbert Hubner[1], Caroline A Wallace[2,7], Heike Zimdahl[1,7], Enrico Petretto[2,7], Herbert Schulz[1], Fiona Maciver[2], Michael Mueller[2], Oliver Hummel[1], Jan Monti[1], Vaclav Zidek[3], Alena Musilova[3], Vladimir Kren[3,4], Helen Causton[2], Laurence Game[2], Gabriele Born[1], Sabine Schmidt[1], Anita Müller[1], Stuart A Cook[2], Theodore W Kurtz[5], John Whittaker[6], Michal Pravenec[3,4] & Timothy J Aitman[2]

Integration of genome-wide expression profiling with linkage analysis is a new approach to identifying genes underlying complex traits. We applied this approach to the regulation of gene expression in the BXH/HXB panel of rat recombinant inbred strains, one of the largest available rodent recombinant inbred panels and a leading resource for genetic analysis of the highly prevalent metabolic syndrome. In two tissues important to the pathogenesis of the metabolic syndrome, we mapped *cis*- and *trans*-regulatory control elements for expression of thousands of genes across the genome. Many of the most highly linked expression quantitative trait loci are regulated in *cis*, are inherited essentially as monogenic traits and are good candidate genes for previously mapped physiological quantitative trait loci in the rat. By comparative mapping we generated a data set of 73 candidate genes for hypertension that merit testing in human populations. Mining of this publicly available data set is expected to lead to new insights into the genes and regulatory pathways underlying the extensive range of metabolic and cardiovascular disease phenotypes that segregate in these recombinant inbred strains.

Determining the molecular basis of natural phenotypic variation, including interindividual susceptibility to common diseases, is a central challenge of postgenome genetics. The availability of genome sequences and genome-scale technologies has enabled new strategies for identifying genes underlying complex phenotypes and has greatly accelerated progress in this field.

The high heritability of variation in gene expression[1] has suggested that identification of the genetic determinants of gene expression may give insights into the molecular basis of complex traits. One justification for studying the genetics of gene expression is that transcript abundance may act as an intermediate phenotype between genomic DNA sequence variation and more complex whole-body phenotypes. Mapping of gene expression levels as quantitative trait loci (called eQTLs) has been undertaken in yeast[2,3] and more recently in mammals[4,5] and has shown that the approach is feasible. As a tool for studying disease phenotypes, however, aside from a single study in F_2 mice[4], the approach has not been extensively tested.

For the past 50 years, the rat has been a leading model for the study of common, complex human diseases[6,7]. The availability of the rat genome sequence[8] has made feasible studies of gene expression at the level of the genome alongside well-characterized rat phenotypes. The spontaneously hypertensive rat (SHR) is a widely studied model of human hypertension and also has many features of the metabolic syndrome[9–13]. In the early 1980s, the SHR strain was crossed with the normotensive Brown Norway (BN) strain to generate the BXH/HXB panel of recombinant inbred (RI) strains[14,15]. Rodent RI panels are powerful and permanent resources for genetic mapping that offer the opportunity to accumulate genetic and physiological data over time[16]. A further advantage of RI strains, as with chromosome substitution strains[17,18], is the ability to study genetically identical biological replicates, which increases trait heritability by reducing environmental variance[19].

Here we applied integrated gene expression profiling and linkage analysis to the regulation of gene expression in fat and kidney tissue in the BXH/HXB panel of rat RI strains. We found that these RI strains are a suitable genetic system for large-scale identification of positional candidates and regulatory pathways for previously mapped physiological QTLs (called pQTLs). By comparative mapping, we compiled a data set of candidate genes for investigating the molecular basis of human hypertension. By identifying hundreds of robustly mapped *cis*- and *trans*-acting eQTLs in a model system with large numbers of existing pQTLs, we generated a unique

[1]Max-Delbrück-Center for Molecular Medicine, Berlin-Buch 13125, Germany. [2]MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College, London W12 0NN, UK. [3]Institute of Physiology, Czech Academy of Sciences and Centre for Applied Genomics, 142 20 Prague 4, Czech Republic. [4]Institute of Biology and Medical Genetics, Charles University, 120 00 Prague 2, Czech Republic. [5]University of California, San Francisco, California 94143-0134, USA. [6]Department of Epidemiology and Public Health, Imperial College, London W2 1PG, UK. [7]These authors contributed equally to this work. Correspondence should be addressed to M.P. (pravenec@biomed.cas.cz) or T.J.A. (t.aitman@csc.mrc.ac.uk).

Published online 13 February 2005; doi:10.1038/ng1522

and accessible resource to test the hypothesis that genetic variation in gene expression has a key role in the molecular evolution of complex physiological and pathophysiological phenotypes.

RESULTS

Linkage analysis of expression profiles in RI strains

We analyzed genome-wide expression data of 15,923 transcripts collected from fat and kidney of 30 RI strains and the SHR and BN progenitor strains. To assess variability within RI strains and to generate a robust set of data, we analyzed gene expression in fat and kidney from four independent rats from each RI strain and from four or five rats from each progenitor strain. Including parental progenitor strains, we carried out 259 individual array hybridizations.

We first compared gene expression profiles generated from fat and kidney tissue from the parental strains (SHR and BN). Of the 15,923 probe sets present on the array, 2,046 and 1,553 detected differential expression ($P < 0.05$) in fat and kidney, respectively, between SHR and BN parental strains. We next carried out genome-wide linkage analysis for the expression data generated in the RI strains. Because genetic regulation of gene expression may be detected for genes that are not differentially expressed between the parental strains, we carried out the linkage analysis for expression profiles from all 15,923 probe sets without filtering.

We generated likelihood ratio statistic (LRS) values and established empirical genome-wide significance by permutation testing. At genome-wide significance of $P = 0.05$, we detected 3,520 and 4,530 linkages in fat and kidney, respectively. To account for multiple testing of 15,923 expression phenotypes, we evaluated the false discovery rate (FDR) at several levels of significance. Although the FDR is relatively high at $P = 0.05$, the expected number of true linkages is 2,644 in fat and 2,917 in kidney (**Supplementary Table 1** online). At a more stringent level of $P = 10^{-3}$, the FDR is ~4% in both tissues, corresponding to 509 and 761 expected true positive linkages in fat and kidney, respectively. An independent test for linkage between marker and transcript using the Wilcoxon-Mann-Whitney test showed that 65–68% of these linkages at $P = 10^{-3}$ are common to the two analyses (**Supplementary Table 2** online).

To validate differential expression detected by microarray and to determine whether interstrain differences in gene expression could be due to DNA sequence variants affecting probe binding, we measured mRNA abundance in parental and RI strains by quantitative RT-PCR and sequenced the target region of a subset of *cis*-regulated transcripts. We found significant sequence differences that could affect probe binding for only 1 of the 15 transcripts that we examined: this transcript, in the major histocompatibility complex (MHC), showed significant sequence variation between SHR and BN that accounted for the apparent differential expression detected on the microarray (data not shown). We observed strong concordance between the microarray and RT-PCR data for differential expression between the parental strains (**Supplementary Fig. 1** online). In addition, we confirmed eQTL linkages for nine transcripts showing strong genome-wide linkage ($P < 10^{-3}$; **Supplementary Table 3** online), including those showing small relative changes between genotypic classes.

Characterization of *cis*- and *trans*-acting eQTLs

Linkages of individual expression phenotypes to multiple tightly linked markers give rise to artificially inflated numbers of eQTLs. We removed this redundancy using a custom algorithm (**Supplementary Fig. 2** online), delineating a data set of 2,118 and 2,490 nonredundant eQTLs at genome-wide significance level of $P < 0.05$ in fat and kidney, respectively (**Table 1**). We examined which of these eQTLs were regulated in *cis* or in *trans* by defining a *cis*-acting eQTL as having a linkage peak within 10 Mbp of the physical location of the probe set (**Supplementary Fig. 3** online). The proportion of eQTLs regulated in *cis* or in *trans* varied substantially in accordance with the genome-wide significance of the eQTL linkage. At a genome-wide threshold of $P < 0.05$, 60–65% of the eQTLs were regulated in *trans* in both tissues. At a higher significance level ($P \leq 10^{-4}$), however, 85–100% of eQTLs were regulated in *cis* (**Table 1**). These results may reflect the large gene effects of some *cis*-acting eQTLs and the probable oligo- or polygenic *trans*-acting influences on gene expression.

A small proportion (~15%) of the eQTLs detected independently in kidney and fat were common to both tissues. These 311 eQTLs (**Table 1**) can be considered replicated linkages and probably reflect common regulatory mechanisms that are shared between the two tissues. A very high proportion of these shared linkages are *cis*-acting (70% at $P < 0.05$). This suggests that the preponderance of *trans*-acting eQTLs observed in the distinct fat and kidney data sets belong to tissue-specific networks for control of gene regulation.

Some of the most significantly linked *cis*-acting eQTLs detected in the RI strains also showed marked differential expression in the parental strains (**Fig. 1** and **Table 2**). Although a number of these eQTLs were shared between kidney and fat (*e.g.*, *Cd36* and *Ilf3*), others showed tissue-specific differences in regulation of gene expression (*Sah*, also called *Sa*, and the transcribed sequence 1388437_at). Many of these *cis*-acting eQTLs showed similar distributions of gene expression in the parental strains and in the RI strains when separated according to marker genotype at the linkage peak (**Fig. 1**). This is suggestive of a monogenic, or near monogenic, *cis*-acting regulation of

Table 1 Number of eQTLs detected in fat and kidney data sets at different genome-wide thresholds of significance

	Fat					Kidney					Shared[a]				
Threshold	*Cis*	*Trans*	Total	Unknown[b]	Total	*Cis*	*Trans*	Total	Unknown[b]	Total	*Cis*	*Trans*	Total	Unknown[b]	Total
$P < 0.05$	622	1,211	1,833	285	2,118	800	1,251	2,051	439	2,490	222	44	266	45	311
$P \leq 0.01$	392	268	660	94	754	514	295	809	140	949	138	18	156	27	183
$P \leq 10^{-3}$	189	46	235	36	271	270	56	326	58	384	75	9	84	13	97
$P \leq 10^{-4}$	97	18	115	22	137	129	10	139	29	168	45	3	48	7	55
$P \leq 10^{-5}$	49	4	53	8	61	62	4	66	6	72	23	0	23	2	25
$P \leq 10^{-6}$	18	0	18	2	20	28	3	31	3	34	5	0	5	3	8

[a]For each threshold of significance, overlapping eQTLs between fat and kidney data sets were identified using the following criteria: (i) the eQTLs are linked to the same probe set and (ii) the eQTLs have the same set of genetic markers identifying the locus or they have both the same genetic marker and *P*-value threshold at the peak of linkage. [b]These eQTLs could not be defined as *cis* or *trans* because the physical location of either the probe set or the peak marker is unknown.

gene expression for these genes and for others with linkage at similar levels of significance (**Table 2**). Nevertheless, 675 of all linked transcripts showed linkage to two or more loci, emphasizing the generally complex nature of regulation of gene expression.

SNP detection and SNP frequency in eQTL genes

To identify DNA sequence variants that may represent candidate quantitative trait nucleotides underlying eQTLs, we generated sequence data for seven of the most statistically significant *cis*-acting eQTL genes (*Pik3c3*, *Myh9*, *Kif1c*, *Aox1*, *Ascl3*, *Dgat2* and *Gnpat*). Although variations in all parts of the gene can affect transcript levels, we focused on variations in the promoter or cDNA that might directly influence transcriptional activity or mRNA stability.

We identified sequence variants in the exons or upstream regulatory regions in six of the seven genes. As a first step towards assessing the functional importance of these sequence variants, we determined the

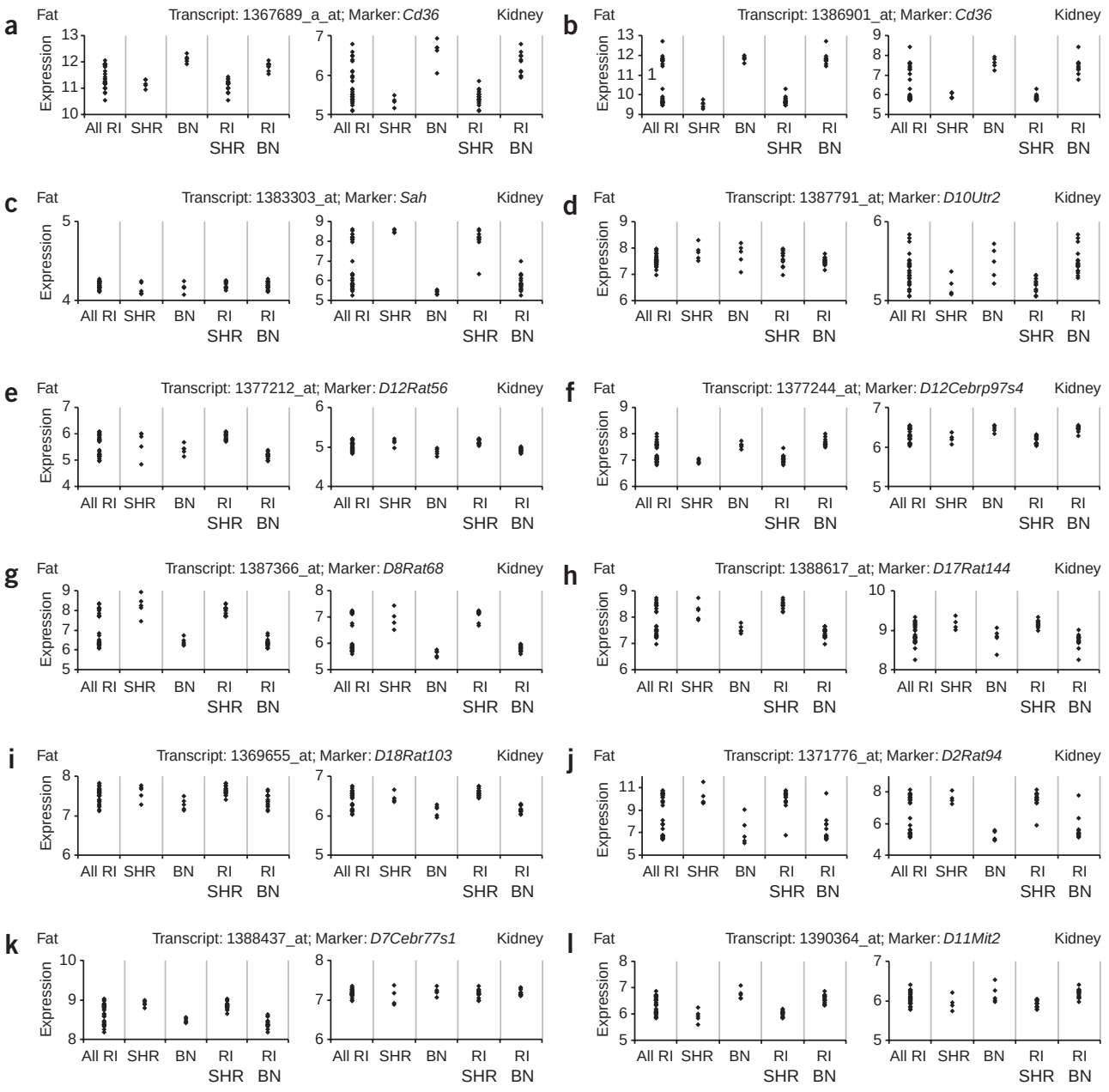

Figure 1 Expression values of parental and RI strains for 12 transcripts in kidney and fat. The first column in each plot shows expression levels for all 30 RI strains. The second and third columns show expression levels for the replicates from each parental strain (SHR and BN, respectively). The fourth and fifth columns show expression values for RI strains by SHR and BN marker genotype, respectively. RMA-normalized expression values are shown on the *y* axis. Detailed properties for all genes depicted are given in **Table 2**. These genes, all with *cis*-acting eQTLs at genome-wide significance level of $P \leq 5 \times 10^{-4}$ in either kidney or fat, were selected according to the following criteria: (i) four genes previously reported to show *cis*-acting differential expression or association with an SHR phenotype (**a–d**); (ii) four genes with highly significant *cis*-acting eQTLs in both tissues (**e–h**); (iii) two genes with highly significant eQTLs and with biological relevance in kidney tissue (**i–j**); (iv) two genes with highly significant eQTLs in fat tissue (**k–l**). Gene names and linkage statistics are given in **Table 2**.

Table 2 Linkage statistics and biological properties of *cis*-regulated genes

Transcript	Gene symbol	Gene name	Transcript physical position (Mb)	Genetic marker at peak of linkage	Marker physical position (Mb)	*P* value* kidney	*P* value* fat
1367689_a_at	*Cd36*	Cd36 antigen	13.51	*Cd36*	13.51	3.0×10^{-6}	1.0×10^{-6}
1386901_at	*Cd36*	Cd36 antigen	13.51	*Cd36*	13.51	$<10^{-6}$	1.0×10^{-6}
1383303_at	*Sah*	SAH (Sah) gene, complete cds	178.22	*Sah*	178.17	1.0×10^{-5}	>0.05
1387791_at	*Ace*	Angiotensin 1 converting enzyme 1	95.39	*D10Utr2*	95.58	2.1×10^{-4}	>0.05
1377212_at	–	Transcribed sequences	3.57	*D12Rat56*	3.02	1.0×10^{-6}	1.0×10^{-6}
1377244_at	–	Similar to zinc finger protein 95 (LOC304275), mRNA	9.73	*D12Cebrp97s4*	5.55	1.5×10^{-5}	1.0×10^{-6}
1387366_at	*Ilf3*	Interleukin enhancer binding factor 3	20.5	*D8Rat68*	19.46	$<10^{-6}$	$<10^{-6}$
1388617_at	–	Similar to RIKEN cDNA 2010012D11 (LOC361239), mRNA	37.15	*D17Rat144*	37.97	3.0×10^{-6}	3.0×10^{-6}
1369655_at	*Pik3c3*	Phosphatidylinositol 3-kinase	22.56	*D18Rat103*	20.16	$<10^{-6}$	0.025
1371776_at	*Pik3r1*	Phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1	32.52	*D2Rat94*	31.1	1.4×10^{-4}	8.6×10^{-4}
1388437_at	–	Transcribed sequences	115.17	*D7Cebr77s1*	115.04	>0.05	$<10^{-6}$
1390364_at	–	Transcribed sequences	32.78	*D11Mit2*	30.89	>0.05	$<10^{-6}$

**P* value from eQTL Reaper linkage analysis in RI strains.

allele status for all sequence variants in three additional inbred strains by resequencing and compared the results to the reference rat genome sequence. Only one of these genes, *Pik3c3*, had a distribution of allelic variants (one promoter SNP and one silent substitution in exon 19) that was common to all of the studied hypertensive strains (SHR/Ola, SHR/Mdc and SHRSP/Mdc) and distinct from that of the normotensive strains (BN.Lx/Cub, BN/SsNHsd/Mcwi and WKY/Mdc). Further sequencing of *Pik3c3* from genomic DNA detected 19 SNPs distributed throughout the gene on two distinct haplotypes (**Supplementary Fig. 4** online).

Pik3c3 showed allelic segregation with hypertension across strains, is one of the most statistically significant ($P < 10^{-6}$) *cis*-acting eQTL genes in our data set and resides within the blood pressure QTL BP46 (pQTL ID 107) and in a chromosome 18 SHR congenic strain (SHR.BN-*D18Rat32*/*D18Rat12*; pQTL ID 105; **Supplementary Table 4** online). *Pik3c3* was upregulated by a factor of 1.4, as shown by quantitative RT-PCR, in kidneys from the SHR strain compared with those from both the BN and the chromosome 18 congenic strains (both $P < 0.05$), confirming *cis*-regulated control of *Pik3c3* expression in the SHR strain. Kidneys transplanted from congenic SHR.BN-*D18Rat32*/*D18Rat12* rats into SHR rats led to a highly significant drop in blood pressure (by 12 mmHg) compared with kidneys transplanted from SHR rats into other SHR rats (data not shown). These data indicate that the congenic kidney is sufficient to mediate blood pressure changes and support the idea that *Pik3c3* is a good candidate for involvement with hypertension in the SHR strain.

The finding of SNPs in six of the seven *cis*-acting eQTL genes that we studied led us to investigate the SNP frequency at the level of the genome compared with the frequency in eQTL genes. We first inspected the number of SNPs identified in a previously published data set of rat cDNAs[20]. For the 8,986 genes on the RAE230A array with Ensembl IDs, SNPs were detected between the stroke-prone SHR (SHRSP) and the BN reference sequences in 2,092 genes (23.3%), which we considered the observed SNP rate across the genome in this SNP data set. We then determined the number of genes in this data set that contained SNPs in which we detected *cis*- and *trans*-acting eQTLs and found significant enrichment for SNPs in the *cis*-regulated eQTL genes compared with either the *trans*-regulated eQTL genes or the observed rate across the genome (**Table 3**).

Relationship of eQTLs to pQTLs

We addressed how the eQTL data can be used to identify candidate genes underlying SHR phenotypes, many of which have been analyzed by QTL mapping in experimental crosses over the past 15 years. *Cis*-acting eQTLs are good candidates for these pQTLs because they show strain-specific differences in gene expression that are under the control of DNA sequence variants in or close to the gene itself.

To illustrate the overlap between individual *cis*-acting eQTLs and previously mapped SHR pQTLs, we plotted the locations of the most stringently mapped ($P \leq 10^{-4}$) *cis*-acting eQTLs against the chromosomal locations of known SHR pQTLs (**Fig. 2**). This generated a data set of genes that have *cis*-acting sequence variants that mediate interstrain differences in gene transcription or mRNA stability (**Fig. 2**; details of the pQTLs and lists of the *cis*-acting eQTLs with $P \leq 10^{-4}$ (FDR $\leq$ 1% in both tissues) are given in **Supplementary Tables 4** and **5** online). The sequence variants may also underlie

Table 3 Percentage of transcripts with sequence variants

eQTL significance	Total number of eQTLs	Total number of eQTLs containing SNPs	Total number of *cis*-acting eQTLs	Number of *cis*-acting eQTLs containing SNPs	Total number of *trans*-acting eQTLs	Number of *trans*-acting eQTLs containing SNPs
$P < 0.05$	1,510	429 (28.4%)[a]	548	204 (37.2%)[a,b]	1,096	272 (24.8%)[b]
$P < 0.001$	238	87 (36.6%)[a]	199	80 (40.2%)[a,b]	50	11 (22.0%)[b]

[a]Significant difference ($P < 0.001$ by χ^2 test) from overall proportion of genes with SNPs across the genome (23.3%). [b]Significant difference ($P < 0.001$ by χ^2 test) between *cis*- and *trans*-acting eQTL genes. The observed percentage of transcripts with sequence variants is significantly different ($P < 0.001$) from the expected percentage across the genome (23.3%) due to increased sequence variability in *cis*-regulated transcripts. The sum of *cis*- and *trans*-regulated transcripts is more than the total number because some transcripts show both *cis* and *trans* regulation.

(patho)physiological phenotypes mapped as pQTLs in crosses between SHR, BN and other strains (**Fig. 2**). These *cis*-regulated eQTL genes merit testing in genetic and functional assays as positional candidates for colocalizing SHR pQTLs and potentially for related human QTLs. Other *cis*-acting eQTLs may be considered 'orphan' eQTLs, candidate genes for pQTLs that have yet to be mapped.

Trans-acting eQTLs represent loci that influence expression of genes or transcripts remote from the eQTL itself. Coincidental mapping of *trans*-acting eQTLs for multiple transcripts to the same chromosomal location, as observed on chromosomes 3 and 17 (**Fig. 3**), may represent a shared regulatory transcriptional control mechanism by a single gene at the eQTL. The locations of *trans*-acting eQTLs in relation to known SHR pQTLs (**Fig. 3**), together with the locations of *cis*-acting eQTLs (**Fig. 2**), may point to genes and regulatory pathways underlying individual SHR pQTLs.

Comparative analysis of blood pressure QTLs

To explore the applicability of the detected fat and kidney *cis*-acting eQTLs to human hypertension, we formed a data set of 255 *cis*-acting eQTL genes with FDR $\leq$ 5% that were contained within rat pQTLs for blood pressure and left ventricle and cardiac mass previously mapped in the SHR strain (**Supplementary Table 4** online). We used the Ensembl Ensmart database to map Affymetrix probe set identifiers of each *cis*-acting eQTL to rat Ensembl genes and then to identify the putative human orthologs. We determined the physical location of each human ortholog on the human genome sequence and compared

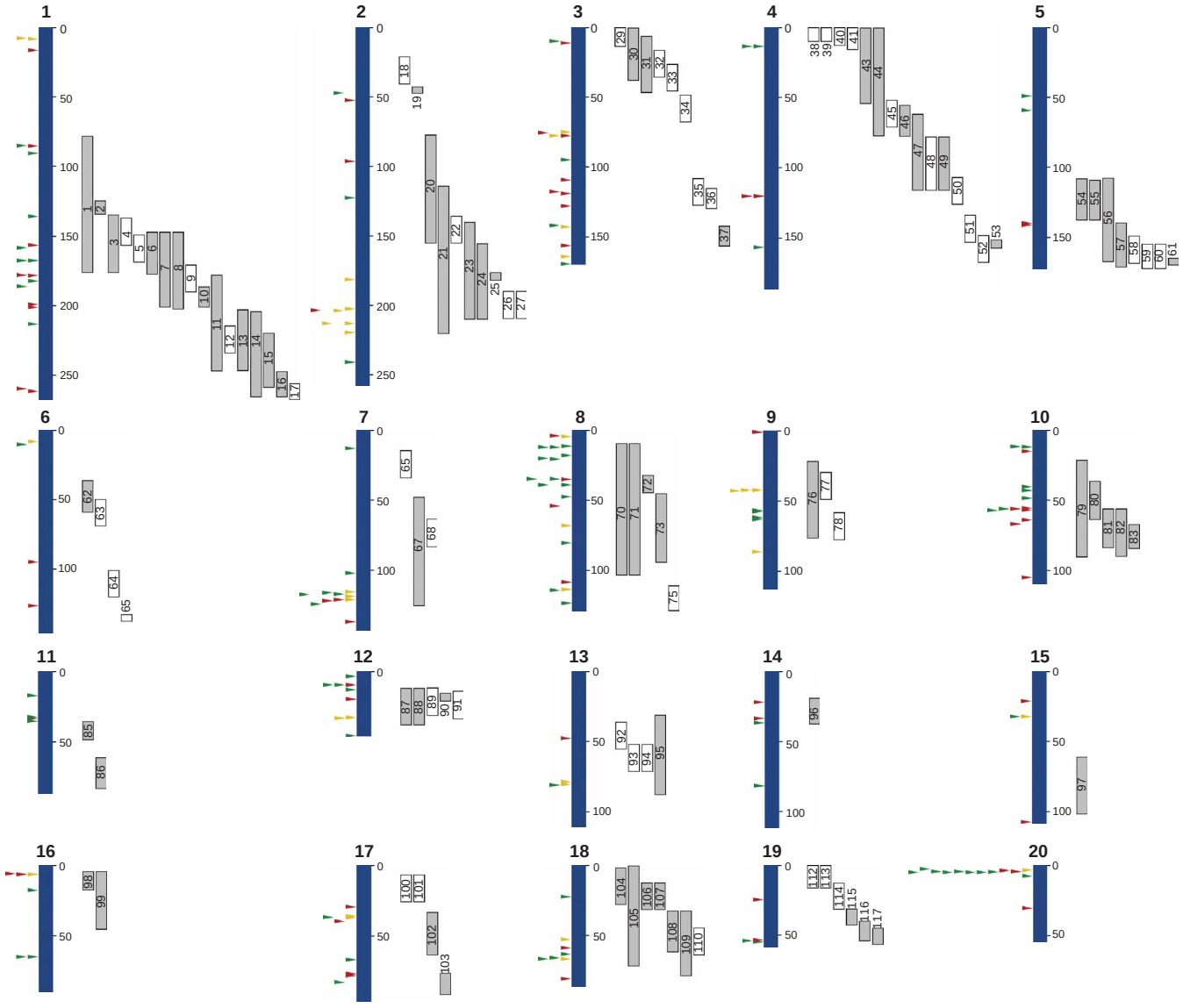

Figure 2 Locations of *cis*-acting eQTLs and previously mapped SHR pQTLs. Chromosomes are shown in blue. The arrowheads on the left of each chromosome (yellow, fat; red, kidney; green, shared) represent the location of the probe set for each *cis*-acting eQTL with $P \leq 10^{-4}$. Previously identified pQTLs in the SHR strains and in the RI strains are shown on the right of each chromosome. Gray boxes represent pQTLs for which both flanking markers are mapped. White boxes represent pQTLs for which incomplete flanking marker information was available, in which case the flanking marker(s) are estimated to be 10 Mbp from the linkage peak. The numbers on the pQTL bars correspond to pQTL information reported in **Supplementary Table 4** online. Some pQTLs in **Supplementary Table 4** online do not appear in this diagram owing to a lack of physical mapping information for the genetic markers that define the pQTL. Gene names, relative changes in parental strains and public database references for probe sets are given in **Supplementary Table 5** online. Scale in Mbp.

these locations with the locations of previously mapped human blood pressure QTLs (**Supplementary Table 6** online). This yielded a set of 73 human orthologs of *cis*-regulated rat eQTL genes that are contained within human blood pressure QTLs and are candidate genes for human hypertension (**Table 4**).

DISCUSSION

Identification of genes that underlie complex (polygenic) traits remains a challenge despite the availability of the genome sequences of humans and other species[21,22]. We previously combined linkage analysis with microarray-based expression profiling to identify a defective gene, *Cd36*, that underlies complex disease phenotypes in the SHR strain[12,23]. Other investigators have successfully used this approach in the study of both monogenic and complex traits[24–26]. But the relative paucity of successful studies using this strategy, and the nature of the *Cd36* genomic deletion[27] that led to detection of dysregulated *Cd36* expression on the microarray, raised questions about the generality of this approach.

Here we applied global gene expression profiling and linkage analysis to study the regulation of gene expression in fat and kidney

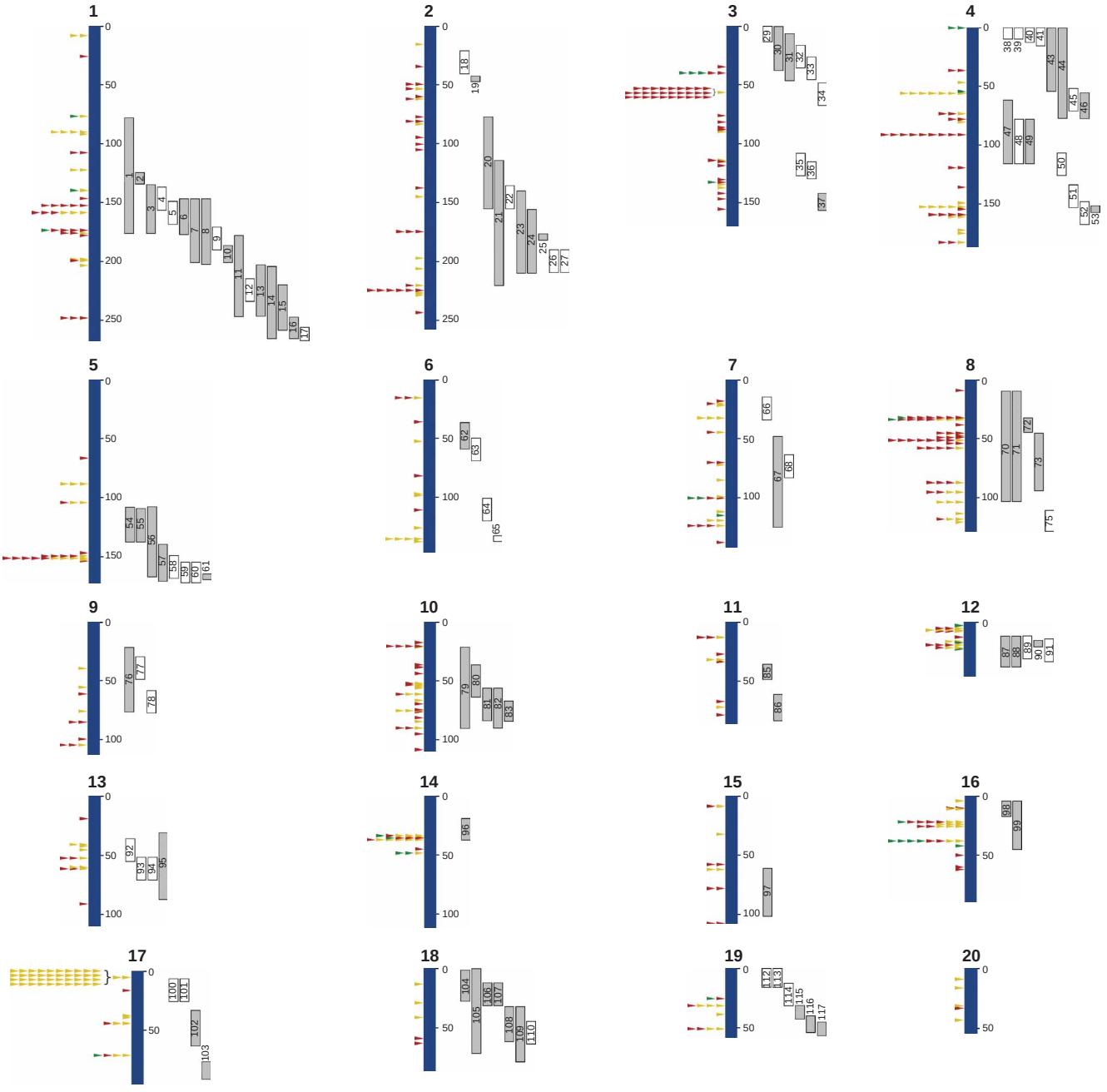

Figure 3 Location of *trans*-acting eQTLs and previously mapped SHR pQTLs. pQTLs are shown as in **Figure 2**. The arrowheads on the left of each chromosome (colors as in **Fig. 2**) represent the locations of the markers at the linkage peak for each *trans*-regulated eQTL with $P \leq 10^{-2}$. This threshold of significance was chosen because *trans*-acting eQTLs are expected to have smaller effects than *cis*-acting eQTLs[4]. Gene names, relative changes in parental strains and public database references for probe sets are given in **Supplementary Table 8** online. Scale in Mbp.

Table 4 Comparative analysis of rat and human blood pressure and blood pressure–related QTLs

Rat tissue	Rat eQTL probe set	Rat gene	Rat chr.	Overlapping rat BP QTLs (ID)	Putative human ortholog	Human chr.	Overlapping human BP QTLs[a]
Shared	1387812_at	*Pace4*	1	BpQTLcluster1 (1)	*PCSK6*	15	BP8_H, BP_H_K1[b]
	1376780_at	–	1	BpQTLcluster1 (1), Lvm2 (3)	*C6orf119*	15	BP32_H, BP_H_K1[b]
	1377614_at	–	1	Bp26 (8), Bp (Huang et al) (9)	NP_775889	16	BP33_H, BP27_H
	1389264_at	–	7	Lvm7 (Huang et al) (67)	NP_620152	22	BP55_H, BP38_H
	1371725_at	–	7	Lvm7 (Huang et al) (67)	*MYH9*	22	BP55_H, BP38_H
	1373838_at	*Fut4*	8	Bp35 (71)	*FUT4*	11	BP31_H
	1390185_at	*Dcps*	8	Bp35 (71), Bp22 (72)	*DCPS*	11	BP22_H
	1371442_at	*Hyou1*	8	Bp35 (71), Bp62 (73)	*HYOU1*	11	BP31_H, BP22_H
	1380286_x_at	–	8	Bp35 (71)	Q9C0D2	11	BP31_H
	1398460_at	–	8	Bp35 (71)	Q9C0D2	11	BP31_H
	1377245_a_at	–	8	Bp35 (71)	Q9C0D2	11	BP31_H
	1375988_at	–	8	Bp35 (71), Bp22 (72)	NP_659451	11	BP22_H
	1382105_at	*Gnb5*	8	Bp35 (71), Bp62 (73)	*GNB5*	15	BP32_H
	1370176_at	*Als2cr3*	9	Bp53 (76)	*ALS2CR3*	2	BP23_H, BP46_H
	1374196_at	*Lancl1*	9	Bp53 (76), Bp108 (78)	*LANCL1*	2	BP46_H
	1387376_at	*Aox1*	9	Bp53 (76)	*AOX1*	2	BP23_H, BP46_H
	1371803_at	*Gm2a*	10	BpQTLcluster9 (79)	*GM2A*	5	BP21_H
	1398309_at	*Pigl*	10	BpQTLcluster9 (79)	*PIGL*	17	BP16_H, BP15_H, BP34_H
	1370930_at	*Kif1c*	10	BpQTLcluster9 (79), Bp45 (82)	*KIF1C*	17	BP15_H, BP34_H
	1368285_at	*Shbg*	10	BpQTLcluster9 (79)	*SHBG*	17	BP15_H, BP34_H
	1373034_at	–	11	BpQTLcluster10 (85)	*WRB*	21	BP12_H, BP_H_K_2[b]
Kidney	1375983_at	–	1	Cm7 (12), Bp42 (13), Map1 (14), Km1 (15)	*MAMDC2*	9	BP_H_K_3[b]
	1371615_at	*Dgat2*	1	BpQTLcluster1 (1), Lvm2 (3), Bp44 (6), Bp26 (8)	*DGAT2*	11	BP31_H
	1383303_at	*Sah*	1	Bp26 (8), Bp (Huang et al) (9)	*SAH*	16	BP27_H
	1388567_at	–	1	Bp26 (8), Bp (Huang et al) (9)	*THUMPD1*	16	BP27_H
	1389300_at	–	1	Bp26 (8), Bp (Huang et al) (9)	–	16	BP27_H
	1386985_at	*Gstm1*	2	Bp105 (21), BpQTLcluster3 (23), Bp101 (24), Bp16 (26), Bp63 (27)	*GSTM1*	1	BP45_H
	1372782_a_at	*Ampd2*	2	Bp105 (21), BpQTLcluster3 (23), Bp101 (24), Bp16 (26), Bp63 (27)	*AMPD2*	1	BP45_H
	1387067_at	–	2	Bp105 (21), BpQTLcluster3 (23), Bp101 (24), Bp19 (25)	NP_958800	1	BP5_H
	1389142_at	–	3	Cm17 (35)	*SQRDL*	15	BP32_H
	1388178_at	*Ncoa3*	3	Bp20 (37)	*NCOA3*	20	BP29_H, BP_H_K_4[b]
	1373204_at	–	4	Bp79 (44), Bp21 (46), BpQTLcluster5 (47)	NP_060957	7	BP48_H
	1389475_at	*Smo*	4	Bp79 (44), Bp21 (46)	*SMO*	7	BP48_H
	1372980_at	–	4	Bp79 (44), Bp21 (46)	NP_848657	7	BP48_H
	1370401_at	*Ly6h*	7	Lvm7 (Huang et al) (67)	*LY6H*	8	BP11_H
	1367917_at	*Cyp2d26*	7	Lvm7 (Huang et al) (67)	*CYP2D7P1*	22	BP55_H, BP38_H
	1370249_at	*Bzrp*	7	Lvm7 (Huang et al) (67)	*BZRP*	22	BP55_H, BP38_H
	1376501_at	–	7	Lvm7 (Huang et al) (67)	*ARHGAP8*	22	BP55_H
	1372897_at	*Plod2*	8	Bp35 (71)	*PLOD2*	3	BP24_H
	1374933_at	*Mcam*	8	Bp35 (71), Bp62 (73)	*MCAM*	11	BP31_H, BP22_H
	1369665_a_at	*Il18*	8	Bp35 (71), Bp62 (73)	*IL18*	11	BP31_H
	1376921_at	–	8	Bp35 (71), Bp22 (72)	NP_060017	11	BP22_H
	1377457_a_at	–	8	Bp35 (71), Bp22 (72)	*SORL1*	11	BP31_H, BP22_H
	1390710_x_at	–	8	Bp35 (71), Bp22 (72)	*SORL1*	11	BP31_H, BP22_H
	1376110_at	–	8	Bp35 (71), Bp62 (73)	*RPP25*	15	BP32_H
	1373887_at	*Sf3b1*	9	Bp53 (76)	*SF3B1*	2	BP23_H, BP46_H
	1376285_at	–	9	Bp53 (76), Bp113 (77)	*GULP1*	2	BP23_H, BP37_H, BP46_H
	1390063_at	–	10	BpQTLcluster9 (79)	*MFAP3*	5	BP21_H
	1368574_at	*Adra1b*	10	BpQTLcluster9 (79)	*ADRA1B*	5	BP_H_K_5[b]
	1373665_at	–	10	BpQTLcluster9 (79)	NP_079375	17	BP15_H, BP34_H
	1371645_at	–	10	BpQTLcluster9 (79), Bp45 (82)	*SDF2*	17	BP16_H, BP34_H
	1377353_a_at	–	10	BpQTLcluster9 (79), Bp45 (82)	*TNFSF12*	17	BP15_H, BP34_H
	1367989_at	*Slc2a4*	10	BpQTLcluster9 (79), Bp45 (82)	*SLC2A4*	17	BP15_H, BP34_H
	1372064_at	–	10	BpQTLcluster9 (79), Bp45 (82)	*CXCL16*	17	BP15_H, BP34_H
	1372397_at	–	13	Bp80 (95)	*RPL34P1*	1	BP5_H
	1376259_at	*PRKCQ*	17	Lvm6 (103)	*PRKCQ*	10	BP52_H
	1388656_at	–	18	Bp46 (107)	*UBE2D2*	5	BP21_H

Table 4 continued on following page

Table 4 Continued

Rat tissue	Rat eQTL probe set	Rat gene	Rat chr.	Overlapping rat BP QTLs (ID)	Putative human ortholog	Human chr.	Overlapping human BP QTLs[a]
Fat	1375516_at	–	1	BpQTLcluster1 (1), Lvm2 (3), Bp44 (6), Bp26 (8)	*NDUFC2*	11	BP31_H
	1389300_at	–	1	Bp26 (8), Bp (Huang et al) (9)	–	16	BP27_H
	1369866_at	*LOC56825*	2	Bp105 (21), BpQTLcluster3 (23), Bp101 (24), Bp16 (26), Bp63 (27)	–	1	BP45_H
	1373243_at	–	2	Bp105 (21), BpQTLcluster3 (23), Bp101 (24), Bp19 (25)	*PMVK*	1	BP5_H
	1398960_at	–	5	Cm (Hamet et al) (89), Bp142 (91)	*COPS5*	8	BP49_H
	1370377_at	*Cyp2d26*	7	Lvm7 (Huang et al) (67)	*CYP2D7P1*	22	BP55_H, BP38_H
	1375940_a_at	–	7	Lvm7 (Huang et al) (67)	NP_078957	22	BP55_H, BP38_H
	1389229_at	–	8	Bp35 (71)	*ACPL2*	3	BP24_H
	1390206_at	–	8	Bp35 (71), Bp22 (72)	NP_060017	11	BP22_H
	1390827_at	*P34*	8	Bp35 (71), Bp62 (73)	NP_078942	15	BP32_H
	1372500_at	*Tmod2*	8	Bp35 (71), Bp62 (73)	*TMOD2*	15	BP32_H
	1371951_at	*Fhl2*	9	Bp53 (76), Bp113 (77)	*FHL2*	2	BP46_H
	1372646_at	–	9	Bp53 (76), Bp113 (77)	NP_115787	2	BP46_H
	1373417_at	–	10	BpQTLcluster9 (79), Bp45 (82)	NP_689979	17	BP15_H, BP34_H
	1368304_at	*Fmo3*	13	Bp80 (95)	*FMO3*	1	BP5_H
	1371732_at	–	13	Bp80 (95)	*DPT*	1	BP5_H

[a]Full details of rat and human genes and pQTLs are given in **Supplementary Table 6** online. [b]BP_H_K*x* are newly designated QTLs[50]. BP, blood pressure.

tissue in the BXH/HXB panel of rat RI strains. We used this panel to map the genetic determinants of gene expression in the SHR strain for 15,923 genes in two of the key tissues in the pathophysiology of the metabolic syndrome. After assessing genome-wide significance, removing redundancy and accounting for multiple testing using FDR, we found more than 1,000 eQTLs in each tissue, of which several hundred were common to both kidney and fat. These eQTLs represent a large source of candidate genes for the scores of pQTLs that have been mapped in the SHR strain.

As previously observed[2–5,28], a powerful feature of this experimental design is the ability to discriminate between *cis*- and *trans*-acting influences on gene expression. In keeping with previous results[4], we found that, at relatively low levels of significance ($P < 0.05$), a high proportion of detected eQTLs, ~65%, result from *trans*-acting regulators of gene expression. At higher levels of significance ($P \leq 10^{-4}$), 85–100% of eQTLs arise from *cis*-acting regulation and have larger, mainly monogenic effects on gene expression. This suggests that, owing to sequence variation in the gene itself, *cis*-acting regulation is more easily detected than variation in genes with secondary effects on transcription of other genes and underscores the more complex nature of *trans* regulation.

Investigation of sequence variation in eQTL genes showed that the SNP frequency was much higher in genes with *cis*-acting eQTLs than in genes with *trans*-acting eQTLs or than the rate observed across the genome (**Table 3**). This observation has a number of possible explanations. First, because the SHR and SHRSP strains are closely related, and both are genetically distant from the BN strain, the increase in polymorphisms detected in *cis*-regulated genes could reflect identification of causative nucleotide variants that underlie *cis*-acting control of gene expression. Second, *cis*-regulated genes may lie in SNP-rich chromosomal regions, and SNPs in these genes could therefore be considered markers of chromosomal regions of phylogenetic diversity between SHR-related strains and the BN strain. If this is the case, a haplotype map of the rat could point to chromosomal regions containing a high density of *cis*-regulated genes.

Cis-acting eQTLs are of particular interest as positional candidate genes for pQTLs (**Fig. 2** and **Supplementary Tables 4** and **5** online). Among many other *cis*-acting eQTLs, our data showed *cis*-acting control of gene expression for *Cd36* and *Sah*, in which *cis*-acting control of gene expression or significant intragene sequence variation has been documented[23,29]. *Cd36* is represented twice on the microarray by non-overlapping probe sets. Our data identified strong *cis*-acting eQTLs for both probe sets ($P \approx 10^{-6}$) in fat and in kidney (**Table 2** and **Supplementary Table 5** online). One of the probe sets (1386901_at) is derived from sequence in the 3′ untranslated region that is deleted from the SHR genome[23,27], and the second (1367689_a_ at) is derived from sequence in the *Cd36* coding region and uses oligonucleotides that do not differ in sequence between the SHR and BN strains[27]. This confirms that we can detect the chromosomal deletion found previously with cDNA microarrays[23] and also shows, using probe set 1367689_a_at, that the *cis*-acting eQTL for *Cd36* cannot simply be attributed to strain differences in probe affinity for *Cd36* mRNA.

For *Sah*, our data showed highly significant differential expression in kidney ($P = 10^{-9}$) between SHR and BN parental strains and strong *cis* linkage in the RI strains ($P = 10^{-5}$; **Fig. 1c** and **Table 2** and **Supplementary Table 5** online). The genotype-dependent bimodal distribution of *Sah* expression in the RI strains (**Fig. 1c**) indicates essentially monogenic *cis* regulation, as previously reported[29]. Although *Sah* is now known not to be a primary determinant of hypertension in the SHR strain[30,31], the detection of this *cis*-acting eQTL in kidney but not in fat demonstrates the ability of this system to identify tissue-specific differences in gene expression, as shown also for many other genes in our data set.

In this study, we investigated seven of the most statistically significant *cis*-regulated eQTL genes from kidney as positional candidates for involvement in hypertension. Nucleotide variations in one of these genes, *Pik3c3*, formed a haplotype that associated with hypertension in a small range of hypertensive and normotensive strains (**Supplementary Fig. 4** online). *Pik3c3* resides in both an SHR blood pressure QTL and an SHR congenic segment that carries a hypertension gene on this region of chromosome 18 (refs. 32,33). Additional physiological and biochemical data (**Supplementary Note** online) support the idea that *Pik3c3* is a candidate for involvement in hypertension in the SHR strain. Existing functional data should similarly encourage investigation of other *cis*-acting eQTL genes (**Table 2** and **Supplementary Table 5** online), such as *Pik3r1*, the regulatory subunit of PI3K, and *Ace*, which has been extensively investigated as a candidate for involvement in cardiovascular phenotypes[34,35].

Eleven of the detected *cis*-acting eQTLs are located in or close to the MHC on chromosome 20 (**Fig. 2**). Because we observed cDNA sequence variation between the SHR and BN strains in the single MHC gene that we tested, which accounted for apparent differential expression between strains, several of the apparent *cis*-acting eQTLs in the MHC probably have a similar basis. Because we found no significant sequence variation in an additional 15 non-MHC genes, however, we do not believe that this accounts for more than a very small proportion of all *cis*-acting eQTLs.

The detection in this study of known *cis*-acting regulation of gene expression raises the question of whether this approach is of value in identifying the genes that underlie pQTLs. The SHR defect in *Cd36* has been conclusively shown in complementation studies to result in pathophysiological phenotypes[36], whereas upregulation in SHR *Sah* gene expression, detected here and elsewhere, is not now believed to contribute to the hypertensive phenotype[30,31]. Several other examples now exist in which a causal relationship has definitively been shown between *cis*-acting eQTLs and functional or physiological phenotypes[2,4,26]. The available data therefore suggest that the combined expression and linkage approach may be useful for pQTL gene identification, particularly when applied on a genome scale.

The *trans*-acting eQTLs that we mapped represent transcripts whose abundance is regulated by loci remote from the genomic locus of each of these genes. Yvert *et al.* analyzed in detail a set of *trans*-acting eQTLs and defined clusters of significantly coregulated genes with functional effects on yeast biology[3]. We found several large groups of genes, up to 43 in a single group (**Fig. 3**), with colocalizing *trans*-acting eQTLs, suggestive of coregulation by a common gene in the eQTL. Some of these groups of *trans*-acting eQTLs overlap with pQTLs, suggesting that they may play a part in mediating the development of SHR phenotypes. Functional and *in silico* analyses of the genes in these groups may advance understanding of the regulatory pathways that underlie these phenotypes.

To show how mining of our data set may be applied to the study of human disease phenotypes, we identified a set of 73 robustly mapped *cis*-regulated rat eQTL genes with FDR $< 5\%$ that lie in SHR blood pressure–related pQTLs and whose human orthologs reside in QTLs for human hypertension (**Table 4** and **Supplementary Table 6** online). These genes are good candidates for underlying human hypertension, although for many there is little functional information. Some of the genes for which functional information does exist show circumstantial association with hypertension. For example, glutathione *S*-transferase mu-type 1 (*Gstm1*) has a strong *cis* linkage ($P < 10^{-5}$) in kidney tissue, is differentially expressed by a factor of $\sim$2.4 fold between the parental strains and lies in a cluster of blood pressure QTLs on rat chromosome 2. *Gstm1* has also been implicated previously in hypertension in genetic studies in the SHRSP strain[37] and in humans[38], suggesting that it may have a role in hypertensive rat strains other than SHR and in humans. Like *Gstm1*, other genes in this data set play a part in cellular resistance to oxidative stress or exhibit different biological characteristics that make them plausible positional candidates for human hypertension (**Supplementary Note** online).

Previous studies taking the combined expression and linkage approach used simple eukaryotes, transformed cell lines from healthy individuals or segregating populations that are no longer available for additional phenotyping[2–5]. Here we applied the same strategy to study, in two different mammalian tissues, regulation of gene expression in the BXH/HXB panel of rat RI strains. RI strains have a number of advantages for this type of study, including the ability to make measurements in multiple genetically identical animals from the same strain to increase trait heritability and the ability to accumulate new phenotypes over time. In addition, the continued breeding of these strains, which are publicly available on a collaborative basis (from M.P. and V.K.), the public availability of the expression data set in this study and of the linkage maps generated in these strains[39] give value to these results outside the context of the immediate findings presented here. Alternative study designs in consomic or congenic strains could provide similar long-term resources and should complement the approach presented here[16–18]. In our study, we mapped genetic determinants of gene expression in a single strain combination, SHR $\times$ BN. But the coincidental mapping of physiological phenotypes in several crosses[23,40], taken together with the data for *Pik3c3* showing association between allelic variants and hypertension across strains, suggests that the eQTLs identified in this study will probably be relevant to SHR phenotypes mapped under a variety of environmental conditions or in strain combinations other than SHR $\times$ BN.

The value of our results is enhanced by the extremely detailed studies (over a 30-year period) of hypertension, other components of the metabolic syndrome and further unrelated SHR phenotypes[9,12,15,23,36,41–43]. Many of these phenotypes have been subjected to genetic analysis in the BXH/HXB RI panel or are amenable to such analysis. Any phenotype that segregates in the SHR $\times$ BN strain combination can potentially be analyzed by this approach, either using the RI strain expression data set from this study or by generating new expression data sets in other tissues or studying additional transcripts in this RI panel.

METHODS

Strains and tissue. We produced a set of RI strains by inbreeding between members of the F_2 generation resulting from the cross of the two highly inbred strains: BN (BN.Lx/Cub) and SHR (SHR/Ola)[14]. We used 30 RI strains (BXH and HXB) at F_{60}. We housed rats in an air-conditioned animal facility and allowed them free access to standard laboratory chow and water. All experiments were done in agreement with the Animal Protection Law of the Czech Republic (311/1997) and were approved by the Ethics Committee of the Institute of Physiology, Czech Academy of Sciences, Prague. We killed rats at 6 weeks of age. We collected tissues from four unfasted males of each RI strain and from four or five rats from each parental strain between 9:00am and 10:00am, froze them in liquid nitrogen and stored them at -80 °C.

Preparation of labeled cRNA and hybridization. We extracted total RNA from retroperitoneal fat pads[12] and from whole kidney from four or five rats of each strain using Trizol reagent (Invitrogen) and purified it using an RNeasy Mini kit (Qiagen) in accordance with the manufacturer's protocol. We synthesized double-stranded cDNA from total RNA without pooling samples, synthesized biotinylated cRNA from cDNA using the MEGAscript T7 kit (Ambion) and nucleotide analogs (Perkin Elmer) in the kidney samples and using Bioarray High Yield RNA Transcript Labelling Kit (Enzo Diagnostics) in fat. We hybridized 15 μg of the fragmented cRNA samples to rat expression Affymetrix RAE 230A GeneChips arrays in accordance with the Affymetrix protocol.

Analysis of expression data. We computed gene expression summary values for Affymetrix GeneChip data using the robust multichip average (RMA) algorithm[44], which uses background adjustment, quantile normalization and summarization. Statistical comparison of expression data in the parental strains was by student's *t*-test (two-tailed) without correction for multiple testing. To determine relative changes, we back-transformed raw RMA output values to the raw intensity scale (anti-log).

Validation of microarray gene expression data. We used quantitative real-time PCR (TaqMan) to compare mRNA levels of 16 transcripts in kidneys that show a range of relative changes in expression between the SHR and BN progenitors. To validate the linkage data, we also measured mRNA levels by quantitative real-time PCR for nine *cis*-regulated transcripts across all RI strains. We reverse-transcribed DNA-free total RNA (2 g) with oligo(dT) primers (Gibco-BRL), Superscript II reverse transcriptase (Gibco-BRL) and dNTP (Boehringer

Mannheim) in 40 µl of reaction buffer (Gibco-BRL). We designed primers and probes using Primer Express 1.0 (Applied Biosystems). TaqMan analysis used an Applied Biosystems 7700 system (Perkin Elmer). We normalized expression levels to 18S rRNA expression by using the $2^{-\Delta\Delta CT}$ method.

Map construction. We constructed a linkage map of 1,011 autosomal markers for all chromosomes using MAPMAKER/Exp. 3.0 (ref. 45) using two- and four-point linkage analysis and published marker genotypes[39,46,47]. We corrected the map based on known physical positions of markers and optimized it by 19 iterative steps using multipoint linkage analysis. Map details are given in **Supplementary Table 7** online. We retrieved physical map positions of genetic markers from Ensembl or by alignment of available marker sequences to the rat genome using blastn. Markers that could not be mapped using blastn but that were located between physically anchored markers were placed on the physical map by interpolation. Eleven percent of all genetic markers could not be placed on the physical map and were used only for the linkage analysis.

Mapping of eQTLs. We derived mean expression values from the four replicates for each RI strain and each tissue after application of the Nalimov outlier test at $P < 10^{-3}$. We carried out genome-wide linkage analysis for each of the 15,923 expression traits and the 1,011 genetic markers using the eQTL Reaper program (K.F. Manly; University of Tennessee Health Science Center, Memphis, Tennessee), which generates an LRS as a measure of the significance of a possible eQTL. eQTL Reaper establishes genome-wide significance by permutation test, estimating an empirical genome-wide probability for observing a given LRS score by chance[48]. For each probe set, permutations are carried out until an LRS greater than that for the real data is observed or until 1,000,000 permutations have been done, conditional on at least 1,000 permutations being done. The permutation procedures for the eQTL Reaper program correct, for individual expression phenotypes, for multiple testing across genetic markers to give a genome-wide corrected *P* value.

Calculation of FDR. The eQTL Reaper genome-wide *P* value accounts for multiple testing across genetic markers but not for multiple testing across the 15,923 expression measurements. We estimated the number of falsely discovered linkages at a given genome-wide significance level by calculating the q value[49], defined as the minimum positive FDR for a fixed significance threshold (**Supplementary Table 1** online).

Definition of *cis*- and *trans*-acting eQTLs. *Cis*-acting regulatory variants are polymorphisms located at or near a gene that influence mRNA levels of that gene. We defined *cis*-acting eQTLs as eQTLs that map within 10 Mbp of the physical location of the probe set on the genomic sequence (20 Mbp total window size; **Supplementary Fig. 3** online). Other eQTLs were defined as acting in *trans*. We obtained physical locations of probe sets from Affymetrix or Ensembl.

SNP detection in eQTL genes. To identify DNA sequence variants that could underlie eQTLs, we generated sequence data for the cDNA sequence and 2–5 kb upstream of exon 1 for seven of the most statistically significant *cis*-regulated eQTL genes from the kidney data set. We carried out direct sequencing from PCR-amplified cDNA or genomic DNA using primers designed by Primer 3.0 from Ensembl annotations. We purified PCR products by shrimp alkaline phosphatase (Promega) and exonuclease I (Promega) treatment and sequenced them directly on an ABI 3730 Sequencer (Applied Biosystems). To determine allele status in three additional inbred strains, we resequenced SNPs in promoters, exons and exon-intron boundaries on genomic DNA from WKY/Mdc, SHR/Mdc and SHRSP/Mdc strains.

URLs. ArrayExpress is available at http://www.ebi.ac.uk/arrayexpress/.

Accession numbers. dbSNP: unique SNP identifiers, ss35032354–ss35032394. ArrayExpress: scanned microarray data, E-AFMX-7.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank H. Banks, N. Cooley, F. Rahman, M. Gerhardt, H. Kistel, S. Blachut and R. Sarwar for technical assistance; K. Manly for providing the eQTL Reaper software; and Affymetrix for donation of microarrays. We acknowledge funding to T.J.A. from the MRC Clinical Sciences Centre, from the British Heart Foundation and from a Wellcome Trust Cardiovascular Functional Genomics initiative; to N.H. from the German Ministry for Science and Education (National Genome Research Network); to M.P. and to V.K. from the Grant Agency of the Czech Republic; to M.P. and T.J.A. from the Wellcome Trust Collaborative Research Initiative Grant; to T.W.K. from the US National Institutes of Health; and to T.W.K. and M.P. from a Fogarty International Research Collaboration Award. M. Pravenec is an International Research Scholar of the Howard Hughes Medical Institute.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 29 November 2004; accepted 26 January 2005
Published online at http://www.nature.com/naturegenetics/

1. Cheung, V.G. *et al.* Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat. Genet.* **33**, 422–425 (2003).
2. Brem, R.B., Yvert, G., Clinton, R. & Kruglyak, L. Genetic dissection of transcriptional regulation in budding yeast. *Science* **296**, 752–755 (2002).
3. Yvert, G. *et al.* Trans-acting regulatory variation in Saccharomyces cerevisiae and the role of transcription factors. *Nat. Genet.* **35**, 57–64 (2003).
4. Schadt, E.E. *et al.* Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**, 297–302 (2003).
5. Morley, M. *et al.* Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**, 743–747 (2004).
6. Jacob, H.J. & Kwitek, A.E. Rat genetics: attaching physiology and pharmacology to the genome. *Nat. Rev. Genet.* **3**, 33–42 (2002).
7. Wallace, C.A. & Aitman, T.J. The rat comes clean. *Nat. Genet.* **36**, 441–442 (2004).
8. Gibbs, R.A. *et al.* Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493–521 (2004).
9. Okamoto, K. *Spontaneous Hypertension: Its Pathogenesis and Complications* (Springer, Berlin, 1972).
10. Rao, R.H. Insulin resistance in spontaneously hypertensive rats: difference in interpretation based on insulin infusion rate or on plasma insulin in glucose clamp studies. *Diabetes* **42**, 1364–1371 (1993).
11. Hulman, S., Falkner, B. & Freyvogel, N. Insulin resistance in the conscious spontaneously hypertensive rat: euglycemic hyperinsulinemic clamp study. *Metabolism* **42**, 14–18 (1993).
12. Aitman, T.J. *et al.* Quantitative trait loci for cellular defects in glucose and fatty acid metabolism in hypertensive rats. *Nat. Genet.* **16**, 197–201 (1997).
13. Grundy, S.M., Brewer, H.B. Jr., Cleeman, J.I., Smith, S.C. Jr. & Lenfant, C. Definition of metabolic syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. *Circulation* **109**, 433–438 (2004).
14. Pravenec, M., Klir, P., Kren, V., Zicha, J. & Kunes, J. An analysis of spontaneous hypertension in spontaneously hypertensive rats by means of new recombinant inbred strains. *J. Hypertens.* **7**, 217–221 (1989).
15. Pravenec, M. *et al.* Genetic analysis of metabolic syndrome in the spontaneously hypertensive rat. *Physiol. Res.* **53** (Suppl 1), 15–23 (2004).
16. Silver, L.M. *Mouse Genetics: Concepts and Applications* (Oxford University Press, New York, 1995).
17. Cowley, A.W. Jr., Roman, R.J. & Jacob, H.J. Application of chromosomal substitution techniques in gene–function discovery. *J. Physiol.* **554**, 46–55 (2004).
18. Singer, J.B. *et al.* Genetic dissection of complex traits with chromosome substitution strains of mice. *Science* **304**, 445–448 (2004).
19. Belknap, J.K., Mitchell, S.R., O'Toole, L.A., Helms, M.L. & Crabbe, J.C. Type I and type II error rates for quantitative trait loci (QTL) mapping studies using recombinant inbred mouse strains. *Behav. Genet.* **26**, 149–160 (1996).
20. Zimdahl, H. *et al.* A SNP map of the rat genome generated from cDNA sequences. *Science* **303**, 807 (2004).
21. Glazier, A.M., Nadeau, J.H. & Aitman, T.J. Finding genes that underlie complex traits. *Science* **298**, 2345–2349 (2002).
22. Abiola, O. *et al.* The nature and identification of quantitative trait loci: a community's view. *Nat. Rev. Genet.* **4**, 911–916 (2003).
23. Aitman, T.J. *et al.* Identification of *Cd36*(*Fat*) as an insulin-resistance gene causing defective fatty acid and glucose metabolism in hypertensive rats. *Nat. Genet.* **21**, 76–83 (1999).
24. Berge, K.E. *et al.* Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science* **290**, 1771–1775 (2000).
25. Lawn, R.M. *et al.* The Tangier disease gene product ABC1 controls the cellular apolipoprotein–mediated lipid removal pathway. *J. Clin. Invest.* **104**, R25–R31 (1999).
26. Karp, C.L. *et al.* Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. *Nat. Immunol.* **1**, 221–226 (2000).
27. Glazier, A.M., Scott, J. & Aitman, T.J. Molecular basis of the Cd36 chromosomal deletion underlying SHR defects in insulin action and fatty acid metabolism. *Mamm. Genome* **13**, 108–113 (2002).
28. Jansen, R.C. & Nap, J.-P. Genetical Genomics: the added value from segregation. *Trends Genet.* **17**, 388–391 (2001).

29. Frantz, S.A. *et al.* Successful isolation of a rat chromosome 1 blood pressure quantitative trait locus in reciprocal congenic strains. *Hypertension* **32**, 639–646 (1998).
30. Hubner, N., Lee, Y.A., Lindpaintner, K., Ganten, D. & Kreutz, R. Congenic substitution mapping excludes Sa as a candidate gene locus for a blood pressure quantitative trait locus on rat chromosome 1. *Hypertension* **34**, 643–648 (1999).
31. Frantz, S., Clemitson, J.R., Bihoreau, M.T., Gauguier, D. & Samani, N.J. Genetic dissection of region around the Sa gene on rat chromosome 1: evidence for multiple loci affecting blood pressure. *Hypertension* **38**, 216–221 (2001).
32. Kovacs, P., Voigt, B. & Kloting, I. Novel quantitative trait loci for blood pressure and related traits on rat chromosomes 1, 10, and 18. *Biochem. Biophys. Res. Commun.* **235**, 343–348 (1997).
33. Pravenec, M., Zidek, V., Kren, V., St Lezin, E. & Kurtz, T.W. Genetic isolation of a quantitative trait locus on chromosome 18 associated with blood pressure and salt sensitivity in the SHR. *Am. J. Hypertens.* **14**, 82A (2001).
34. Cambien, F. *et al.* Deletion polymorphism in the gene for angiotensin–converting enzyme is a potent risk factor for myocardial infarction. *Nature* **359**, 641–644 (1992).
35. Lindpaintner, K. *et al.* A prospective evaluation of an angiotensin–converting–enzyme gene polymorphism and the risk of ischemic heart disease. *N. Engl. J. Med.* **332**, 706–711 (1995).
36. Pravenec, M. *et al.* Transgenic rescue of defective Cd36 ameliorates insulin resistance in spontaneously hypertensive rats. *Nat. Genet.* **27**, 156–158 (2001).
37. McBride, M.W. *et al.* Microarray analysis of rat chromosome 2 congenic strains. *Hypertension* **41**, 847–853 (2003).
38. Rice, T. *et al.* Genome-wide linkage analysis of systolic and diastolic blood pressure: the Quebec Family Study. *Circulation* **102**, 1956–1963 (2000).
39. Jirout, M. *et al.* A new framework marker-based linkage map and SDPs for the rat HXB/BXH strain set. *Mamm. Genome* **14**, 537–546 (2003).
40. Kato, N. *et al.* Complete genome searches for quantitative trait loci controlling blood pressure and related traits in four segregating populations derived from Dahl hypertensive rats. *Mamm. Genome* **10**, 259–265 (1999).
41. Pravenec, M. *et al.* Mapping of quantitative trait loci for blood pressure and cardiac mass in the rat by genome scanning of recombinant inbred strains. *J. Clin. Invest.* **96**, 1973–1978 (1995).
42. Bottger, A. *et al.* Quantitative trait loci influencing cholesterol and phospholipid phenotypes map to chromosomes that contain genes regulating blood pressure in the spontaneously hypertensive rat. *J. Clin. Invest.* **98**, 856–862 (1996).
43. Printz, M.P., Jirout, M., Jaworski, R., Alemayehu, A. & Kren, V. Genetic Models in Applied Physiology. HXB/BXH rat recombinant inbred strain platform: a newly enhanced tool for cardiovascular, behavioral, and developmental genetics and genomics. *J. Appl. Physiol.* **94**, 2510–2522 (2003).
44. Irizarry, R.A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).
45. Lander, E.S. *et al.* MAPMAKER–an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**, 174–181 (1987).
46. Pravenec, M. *et al.* A genetic linkage map of the rat derived from recombinant inbred strains. *Mamm. Genome* **7**, 117–127 (1996).
47. Pravenec, M. *et al.* HXB/Ipcv and BXH/Cub recombinant inbred strains of the rat: strain distribution patterns of 632 alleles. *Folia Biol. (Praha)* **45**, 203–215 (1999).
48. Churchill, G.A. & Doerge, R.W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971 (1994).
49. Storey, J.D. A direct approach to false discovery rates. *J. R. Stat Soc (B)* **64**, 479–498 (2002).
50. Krushkal, J. *et al.* Genome-wide linkage analyses of systolic blood pressure using highly discordant siblings. *Circulation* **99**, 1407–1410 (1999).