Physiol Genomics 42: 456–468, 2010.
First published June 15, 2010; doi:10.1152/physiolgenomics.00110.2009.

Sex-specific gene expression in the BXD mouse liver

Daniel M. Gatti,[1] Ni Zhao,[1] Elissa J. Chesler,[2] Blair U. Bradford,[1] Andrey A. Shabalin,[3] Roumyana Yordanova,[4] Lu Lu,[5] and Ivan Rusyn[1]

[1]*Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill, North Carolina;* [2]*Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee;* [3]*Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, North Carolina;* [4]*Bristol-Myers Squibb Research and Development, Princeton, New Jersey; and* [5]*Department of Anatomy and Neurobiology, University of Tennessee College of Medicine, Memphis, Tennessee*

Submitted 1 July 2009; accepted in final form 8 June 2010

Gatti DM, Zhao N, Chesler EJ, Bradford BU, Shabalin AA, Yordanova R, Lu L, Rusyn I. Sex-specific gene expression in the BXD mouse liver. *Physiol Genomics* 42: 456–468, 2010. First published June 15, 2010; doi:10.1152/physiolgenomics.00110.2009.— Differences in clinical phenotypes between the sexes are well documented and have their roots in differential gene expression. While sex has a major effect on gene expression, transcription is also influenced by complex interactions between individual genetic variation and environmental stimuli. In this study, we sought to understand how genetic variation affects sex-related differences in liver gene expression by performing genetic mapping of genomewide liver mRNA expression data in a genetically defined population of naive male and female mice from C57BL/6J, DBA/2J, B6D2F1, and 37 C57BL/6J × DBA/2J (BXD) recombinant inbred strains. As expected, we found that many genes important to xenobiotic metabolism and other important pathways exhibit sexually dimorphic expression. We also performed gene expression quantitative trait locus mapping in this panel and report that the most significant loci that appear to regulate a larger number of genes than expected by chance are largely sex independent. Importantly, we found that the degree of correlation within gene expression networks differs substantially between the sexes. Finally, we compare our results to a recently released human liver gene expression data set and report on important similarities in sexually dimorphic liver gene expression between mouse and human. This study enhances our understanding of sex differences at the genome level and between species, as well as increasing our knowledge of the molecular underpinnings of sex differences in responses to xenobiotics.

BXD recombinant inbred mouse; microarray; gene expression quantitative trait locus; genetical genomics

PHYSIOLOGICAL DIFFERENCES between the sexes have their roots in sex differences in gene expression (37, 58). Although the sexes possess identical autosomes and differ by a single Y chromosome, genes exhibiting expression differences are broadly distributed throughout the genome. In the liver, many of these differences are driven primarily by the pituitary growth hormone (GH) (59), which acts through signal transducer and activator of transcription (STAT)5b (13, 29) and hepatocyte-enriched nuclear factor (HNF)4α (30). These transcription factors, through direct and indirect mechanisms, are thought to regulate expression of many steroid and xenobiotic metabolizing enzymes, leading to observed patterns of sexually dimorphic expression in a number of important biological pathways (58).

Address for reprint requests and other correspondence: I. Rusyn, Dept. of Environmental Sciences and Engineering, Univ. of North Carolina at Chapel Hill, Chapel Hill, NC 27599 (e-mail: iir@unc.edu).

Several investigators have examined sex differences in gene expression in the mouse liver with microarrays. Rinn et al. (41) surveyed the expression of 13,977 genes in male and female Swiss-Webster mice in the hypothalamus, kidney, liver, and reproductive tissues. Using criteria of fold change (FC) between sexes > 3 and $P < 0.001$, they found six genes to be sexually dimorphic in the liver, *Cyp2b13*, *Cyp3a16*, *Cyp4a12*, *Hsd3b5m*, *Elovl3*, and *Keg1*. Clodfelter et al. (12) studied liver gene expression changes in *Stat5b* knockout and wild-type mice, finding 1,603 differentially regulated genes, with 850 being male- and 753 female biased ($P < 0.05$ and FC > 1.5). A large study consisting of 344 mice comprising an F_2 cross between C57B/6J.apoE$^{-/-}$ and C3H/HeJ.apoE$^{-/-}$ strains (~50% from each sex) produced two reports (57, 61) that examined sexually dimorphic gene expression in adipose tissue, brain, liver, and muscle. It was reported that 9,250 genes are dimorphic in the liver ($P < 0.01$ and FC > 1). In an effort to uncover regulatory pathways driven by genetic variation, the authors also performed gene expression quantitative trait locus (eQTL) mapping to find genomic loci responsible for transcriptional regulation of clusters of genes. The authors reported that one locus on medial chromosome (Chr) 5 regulated the expression of ~80 genes, and the genes in this cluster were enriched for cell cycle, cell proliferation, and DNA replication. In addition, van Nas et al. (55) also illustrated that there are differences in both expression levels and gene coexpression between males and females. While these and other previous investigations into sex differences in gene expression in the liver made crucial contributions to the knowledge of sex-specific gene expression patterns in the liver, only a few studies examined the genetic basis of these effects (55, 57).

Even though sex-independent polymorphic local and distant QTLs, including several loci that control expression of large numbers of genes, were identified with the eQTL approach in the BXD panel (19) and confirmed in an independent cohort of inbred strains (20), relatively little is known about the extent of genetic control of sex differences in liver gene expression networks in the mouse, or in other species. We hypothesize that sex differences in liver xenobiotic metabolizing functions are largely due to variation in the expression of steroid metabolizing enzymes. We examined differences in constitutive liver gene expression between males and females across a genetically diverse population of BXD recombinant inbred (RI) mice derived from the C57BL/6J and DBA/2J strains. We examined both genes and pathways that are differentially regulated or differentially coexpressed between sexes in the mouse. eQTL mapping was performed separately for each sex, and while sex-specific eQTLs were observed, no sex-specific eQTL hot-

spots of transcriptional regulation were identified. Finally, we compared sexually dimorphic genes between murine and human liver by drawing a comparison to the data from Schadt et al. (44), who reported on liver gene expression in over 400 human subjects.

METHODS

Gene Expression Data

The details of mouse breeding, housing, RNA isolation, and gene expression are described elsewhere (19). Tissue collection (synchronized with respect to the time of day) was conducted at the University of Tennessee at Memphis and approved by the Institutional Animal Care and Use Committee. Briefly, gene expression in the livers of 37 strains of male and female C57BL/6J × DBA/2J (BXD) RI mice, the C57BL/6J and DBA/2J parentals, and the C57BL/6J × DB2/2J F1 generation (B6D2F1) mice was measured with the Agilent (Santa Clara, CA) G4121A microarray (20,868 transcripts). Additional details regarding animals, microarrays, data acquisition, processing, and analyses can be found at WebQTL.org (**http://webqtl.org/dbdoc/LV_G_0106_B.html**). The data have been deposited in the Gene Expression Omnibus (GSE17522).

Gene Expression Data Analysis

The raw data files from the Agilent scanner were uploaded into the University of North Carolina Microarray Database (**http://genome.unc.edu**), and a custom normalization was applied to the $\log_2$ of the Lowess normalized Red/Green ratio of background-subtracted intensities. A nested, mixed analysis of variance (ANOVA) model (where batch, sex, and strain were considered as main effects and sex by strain was an interaction term) was fit to the data, and least-squares means were produced for each sex and strain along with *P* values for batch, strain, sex, and strain by sex interaction. The Significance Analysis of Function and Expression (SAFE) tool (5) was used to find significant differentially regulated Gene Ontology (GO) categories, as well as Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Rather than employing an arbitrary significance cutoff, SAFE uses the entire data set and array permutation to produce empirical *P* values for each category or pathway. Significant categories were selected based on *q* values of ≤0.25, a 25% false discovery rate (FDR) for the entire list.

Sex-specific correlated gene clusters. The Spearman correlation of each transcript with all 20,867 other transcripts on the array was calculated for males and females independently. Transcripts with correlations ≥0.7 in either males or females were retained separately for each transcript as highly correlated clusters. This threshold corresponds to a *P* value of 3.44×10^{-7}, which is equivalent to a Bonferroni-corrected *P* value (for 20,868 transcripts) of 0.0072. Gene sets with absolute mean differences in correlation between sexes ≥0.35 were selected as sex-specific coexpression clusters. Both cutoffs were selected to produce reasonably short lists of interpretable results. The highly correlated sex-specific gene clusters from the above analysis were searched for GO and KEGG pathway enrichment with Fisher's exact test. Significant categories were selected with *q* values ≤0.1.

Transcription factor binding site enrichment. oPOSSUM (28) was used to search for enrichment in transcription factor binding sites (TFBSs) among sets of genes. The options used, in the order that they appear on the website, were vertebrate taxonomic supergroup, using the top 10% of conserved regions, matrix threshold = 80%, upstream/downstream sequence = 5,000/5,000, Z-score cutoff ≥ 10. oPOSSUM does not provide *P* values for results.

Gene Expression Quantitative Trait Locus Mapping

eQTL mapping was performed by using single marker mapping in FastMap (21), which uses sample permutation to produce per-gene significance thresholds. Heritability was estimated with the strain intraclass correlation. This calculation was performed on variance components obtained with restricted maximum likelihood estimation in a random-effects ANOVA model implemented with SAS v9.1.2. In RI lines, this analysis can be performed by using the replicates within strain and sex as detailed in Reference 10. Within each strain, these replicates should be similar if the gene expression trait is heritable, or if it has a high proportion of variance explained by genetic factors relative to technical and environmental noise. The top quartile (5,217) of transcripts with the highest heritabilities ($h^2 > 0.136$) were kept. Mapping was performed on these 5,217 transcripts by using a set of 3,795 markers downloaded from WebQTL (**http://www.genenetwork.org**) for the BXD mice. Permutations (1,000) of the expression data were employed to establish significance cutoffs for each gene. Mapping was performed with the strain means, the male data, the female data, and the strain differences between males and females.

Mapping with the strain differences allowed us to test sex by marker interactions as follows. For a given transcript, we observe the expression G_{ij}, where $i = 1 \ldots n$ indexes the mouse strains and $j = 0, 1$ denotes the sex. Let a (fixed) genotype be denoted as $S_i = 0,1$, $i = 1 \ldots n$, as it does not depend on the sex *j*.

The full model is given by:

$$G_{ij} = b_0 + b_1 \times S_i + b_2 \times j + b_3 \times (j \times S_i) \quad (1)$$

where b_0 is the constant term, b_1 is the genotype effect, b_2 is the sex effect, and b_3 is the interaction term.

For each sex separately ($j = 0$ and $j = 1$) the model takes the form:

$$G_{i0} = b_0 + b_1 \times S_i \text{ (Males)}$$

$$G_{i0} = (b_0 + b_2) + (b_1 + b_3) \times S_i \text{ (Females)}$$

The difference is:

$$G_{i1} - G_{i0} = b_2 + b_3 \times S_i \quad (2)$$

The model (*Eq. 2*) contains S_i as the regressor, estimated on the sex differences ($G_{i1} - G_{i0}$). This model produces the same estimate of b_3 as in the full model (*Eq. 1*). The permutation scheme for the full model (*Eq. 1*) involves permuting S_i relative to G_{ij}, keeping the strain pairs (G_{i0}, G_{i1}) intact in order to correctly estimate the empirical *P* value (11). To select a significance threshold that accounts for multiple testing across transcripts, we applied the *q* value algorithm (50) to the maximum *P* value for each transcript and selected all transcripts with $q < 0.1$. We also investigated whether polymorphisms that lie within the probe sequence produce false local eQTLs (14). Since C57BL/6J is the reference strain used to create probe sequences, probes querying the DBA/2J allele should have lower probe intensities among local eQTLs (2), creating false local eQTLs. We tested for this by performing Fisher's exact test, using a 2 × 2 table with the high-expression allele (C57BL/6J or DBA/2J) on one axis and SNP-in-probe/no-SNP-in-probe eQTL on the other, and found no evidence of an effect ($P = 0.359$). To assess a stringency of eQTL band analysis, eQTL bands were derived from the above analysis repeated 100 times, permuting the markers each time and counting the number of eQTL hotspots of each size. The probability of seeing an eQTL band of a given size was assessed from the empirical distribution of the number of eQTL bands of each size produced in all 100 permutations.

RESULTS

Characterization of Sexually Dimorphic Liver Genes in BXD Mouse Panel

Sex differences in gene expression were identified with a two-way (sex and strain of each sample) ANOVA model

Table 1. *Number of sexually dimorphic genes in BXD mouse liver*

	q Value	*P* Value	Male-Biased	Female-Biased	Total
BXD recombinant inbred panel	0.01	0.0038	1,554	1,750	3,304
	0.05	0.0321	2,504	3,103	5,607
	0.10	0.0823	3,156	4,042	7,198
Yang et al. (61)	NA	0.01	4,441	4,809	9,250

BXD, C57BL/6J × DBA/2J; NA, not applicable.

(Supplemental Table S1).[1] *P* values for sex and strain were filtered separately, using *q* values (50) of 0.01, 0.05, and 0.1, which represent the FDR of the resulting gene lists, corresponding to FDRs of 1%, 5%, and 10%. At the smallest *q* value, we find a similar number of differentially expressed genes as that in a previous report (61) and replicate the finding that there are more female-biased genes than male-biased genes (Table 1). To further explore the differences, we plotted the mean expression value for each gene in males versus that in females (Fig. 1) and observed that, while there are more female-biased genes, there is also a large cluster of male-biased genes (Fig. 1, circle). The genes in this cluster ($q \leq 0.01$, $0 \leq$ male ≤ 2.6, $-0.5 \leq$ female ≤ 1.5) are significantly enriched for olfactory receptor activity (GO: 0004984, $P = 9.6 \times 10^{-22}$), which consists of G protein-coupled receptors involved in signaling.

As expected, some of the most sexually dimorphic genes (Table 2) are located on the X and Y chromosomes, including inactive X specific transcripts (*Xist*), DEAD box polypeptide 3, Y-linked (*Ddx3y*), and eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked (*Eif2s3y*). However, most genes that exhibit considerable fold difference in expression (>2) between sexes reside on autosomes. Although most sexually dimorphic genes were found on the autosomes, there was no statistically significant enrichment for sexually dimorphic genes on autosomal chromosomes compared with the sex chromosomes. We have compared our results with two independent microarray studies that examined sex differences in liver gene expression in mice. The study of Clodfelter et al. (12) was conducted on *Stat5b*-null and wild-type mice. Using a threshold of $P \leq 0.01$ for sex bias in both data sets, we found that 51.5% (410 of 796 genes from that study) were sexually dimorphic in both studies. A similar comparison with a study that used a much larger number of mice in an F_2 cross (61) yields a 27.2% overlap (9,250 genes were reported as sexually dimorphic with $P \leq 0.01$, of which 8,087 mapped to Agilent 4121A microarray and 2,196 matched our results). There was no significant GO category or KEGG pathway enrichment ($q \leq 0.1$) among the genes that replicate between the three studies.

We also examined the differential expression of several cytochrome *P*-450 (CYP) enzymes previously reported to be sexually dimorphic. Consistent with previous reports (42), we found that liver expression of *Cyp2c55*, *Cyp2d9*, *Cyp4a12*, *Cyp7b1*, and *Cyp8b1* is male biased and expression of *Cyp2a4*, *Cyp2b10*, *Cyp2b13*, *Cyp2c40*, *Cyp3a16*, and *Cyp3a41* is female biased. However, in contrast to some previous observations, we find that *Cyp2j13* is male biased and *Cyp2e1* is female biased in the BXD mouse panel.

To find transcription factors responsible for sexually dimorphic gene expression, we submitted lists of male- and female-biased genes to oPOSSUM (28), a web-based tool that searches for TFBS enrichment in lists of genes. In females, Hnf1α was significant (Z-score = 16.98) and was found in 395 transcripts.

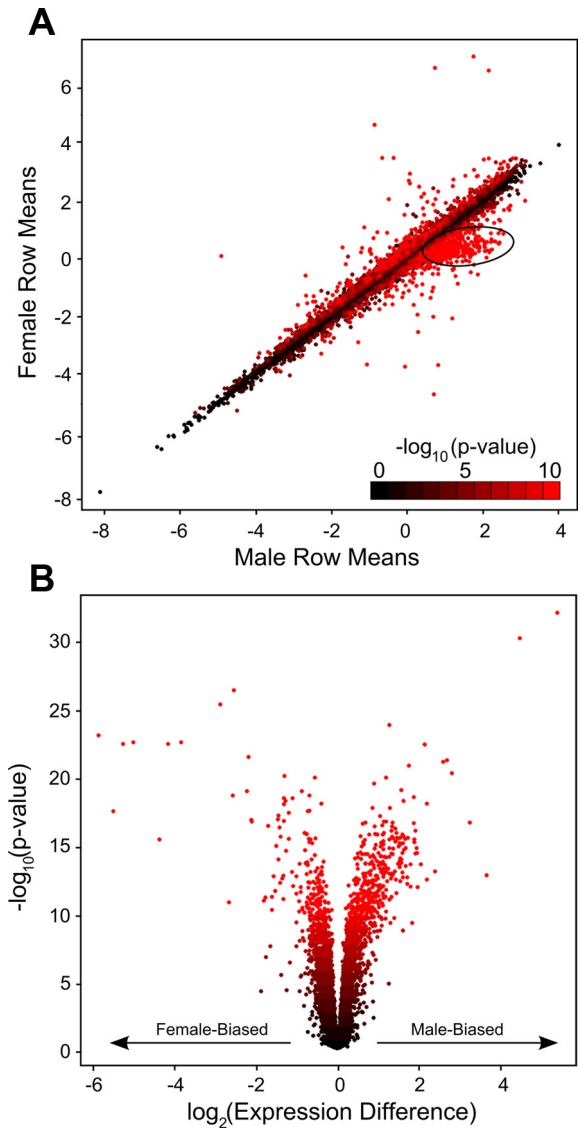


Fig. 1. Comparison in expression of 20,868 genes in livers of male and female C57BL/6J × DBA/2J (BXD) recombinant inbred mice. *A*: mean expression ($\log_2$) of each gene in all males is plotted vs. that in all females, colored by the statistical significance (see color bar *inset*) of differential expression. *B*: volcano plot of differential expression vs. $-\log_{10}(P$ value) for the sex effect from the ANOVA analysis. Positive values indicate higher expression in males.

[1] Supplemental Material for this article is available online at the Journal website.

Table 2. *Genes with significant differential expression [mean log_2(diff) >2] between sexes*

Probe ID	Gene	Gene Name	Sex Bias	Mean log_2(diff)	Chr	Ref. 12	Ref. 61
A_51_P145453	*Xist*	Inactive X specific transcripts	F	5.9	X		
A_51_P479715	*Eif2s3y*	Eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked	M	5.4	Y	*	
A_51_P269404	*Fmo3*	Flavin containing monooxygenase 3	F	5.5	1	*	*
A_51_P426513	*Sult3a1*	Sulfotransferase family 3A, member 1	F	5.3	10	*	*
A_51_P114722	*Hao3*	Hydroxyacid oxidase 3	F	5.0	3	*	*
A_51_P182370	*Cyp2b10*	Cytochrome *P*-450, 2b10	F	2.5	7	*	*
A_51_P184750	*Cyp2b10*	Cytochrome *P*-450, 2b10	F	2.9	7	*	*
A_51_P402994	*Ddx3y*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	M	4.5	Y	*	*
A_51_P480533	*Sult2a2*	Sulfotransferase family 2A, dehydro-epiandrosterone-preferring, mem. 2	F	4.4	7		
A_51_P467076	*Cyp2b9*	Cytochrome *P*-450, family 2, subfamily b, polypeptide 9	F	4.2	7		
A_51_P492339	*Cyp2b13*	Cytochrome *P*-450, family 2, subfamily b, polypeptide 13	F	3.8	7	*	
A_51_P192886	*Cml5*	Camello-like 5	M	3.7	6		
A_51_P324636	*Elovl3*	Elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 3	M	3.3	19	*	*
A_51_P496162	*Hsd3b5*	Hydroxy-δ-5-steroid dehydrogenase, 3β- and steroid δ-isomerase 5	M	2.8	3	*	*
A_51_P461429	*Cyp7b1*	Cytochrome *P*-450, family 7, subfamily b, polypeptide 1	M	2.7	3	*	*
A_51_P238576	*Cyp4a14*	cytochrome *P*-450, family 4, subfamily a, polypeptide 14	F	2.7	4	*	*
A_51_P352005	*Hsd3b4*	Hydroxy-δ-5-steroid dehydrogenase, 3β- and steroid δ-isomerase 4	M	2.6	3		
A_51_P290535	*Abcd2*	ATP-binding cassette, subfamily D (ALD), member 2	F	2.6	15	*	*
A_51_P252002	*Susd4*	Sushi domain containing 4	M	2.2	1	*	
A_51_P299528	*Ust3m*	Putative integral membrane transport protein; 3′ UTR	F	2.2	19	*	
A_51_P165111	*Slco1a1*	Solute carrier organic anion transporter family, member 1a1	M	2.4	6	*	*
A_51_P249297	*Abcd2*	ATP-binding cassette, subfamily D (ALD), member 2	F	2.2	15	*	
A_51_P103137	*Slco1a1*	Solute carrier organic anion transporter family, member 1a1	M	2.2	6		*
A_51_P319213	Atp6v0d2	ATPase, H^+ transporting, lysosomal V0 subunit D2	F	2.1	4	*	*
A_51_P374464	*Gstp1*	Glutathione *S*-transferase, pi 1	M	2.1	19	*	*
A_51_P496045	*Sorbs1*	RIKEN cDNA 9530039L23 gene	M	2.1	19		
A_51_P414072	*Prom1*	Prominin 1	F	2.1	5	*	*

In comparison of the overlap in sexually dimorphic genes with other studies, asterisk represents concordance with the results reported in Ref. 12 or 61. See Supplemental Table S1 for values of all genes.

Hnf1α was more highly expressed in females ($P = 7.67 \times 10^{-4}$). In males, Pax5 was significant (Z-score = 11.54) and was found in 165 transcripts. *Pax5* did not show differential expression between sexes ($P = 0.6$).

While the examination of expression differences between individual genes increases our understanding of physiological differences between the sexes, it is often informative to search for functionally related gene sets that are differentially expressed. SAFE (5) is a permutation-based software tool that assesses the significance of differentially expressed gene sets, in this case between males and females, using the entire data set rather than a predefined "significant" gene list. We found 89 significant KEGG pathways ($q \leq 0.10$, Supplemental Table S2a) of which many are reflective of xenobiotic metabolism pathways such as Drug Metabolism-Cytochrome *P*-450 (KEGG 00982, $q = 0.01$; Fig. 2*A*, Supplemental Table S3a), Retinol Metabolism (KEGG 00830, $q = 0.01$), Glutathione Metabolism (KEGG 00480, $q = 0.05$), and Fatty Acid Metabolism (KEGG 00071, $q = 0.08$). Similarly, of 45 significant GO Molecular Function categories ($q \leq 0.25$, Supplemental Table S2b) many are also highly relevant to xenobiotic metabolism in the liver, such as Oxidoreductase Activity (GO: 0016491, $q = 0.03$; Fig. 2*B*, Supplemental Table S3b), Cysteine-type Endopeptidase Activity (GO: 0004197, $q = 0.04$), Iron Ion Binding (GO: 0005506, $q = 0.05$), Heme Binding (GO: 0020037, $q = 0.08$), Sulfotransferase Activity (GO: 0008146, $q = 0.13$), and Glutathione Transferase Activity (GO: 004364, $q = 0.14$). There were 30 significant GO Biological Process categories ($q \leq 0.25$; Supplemental Table S2c). The most significant ($q = 0.11$) were GO: 0006766 (Vitamin Metabolic Process; Fig. 2*C*, Supplemental Table S3c) and GO: 0051186 (Cofactor Metabolic Process). There were eight significant GO Cellular Component categories ($q \leq 0.25$; Supplemental Table S2d). The most significant categories were Vesicular Fraction (GO: 042598, $q = 0.09$; Fig. 2*D*, Supplemental Table S3d) and Microsome (GO: 0005792, $q = 0.09$), both of which include a large number of sexually dimorphic CYPs that are located on the endoplasmic reticulum. Also included in the list was Peroxisome (GO: 0005777, $q = 0.1$), which contains the female-biased hydroxyacid oxidase (*Hao*)3 and ATP-binding cassette D2 (*Abcd2*), as well as the male-biased *Hao1*.

Comparison of Sex Differences in Liver Gene Expression Between Mouse and Human

Sex differences in human liver gene expression have been largely unexplored (58). We compared the sexually dimorphic genes in the BXD liver data set with a recently published large human liver gene expression data set (44) obtained from the Gene Expression Omnibus. Expression measurements for over 39,000 transcripts in over 400 human samples (both male and female) were taken with a custom microarray. We determined that 7,342 transcripts were sexually dimorphic in this data set (Student's *t*-test between sexes for each transcript at $q \leq 0.10$ and $p \leq 0.03$). Murine orthologs could be mapped with high

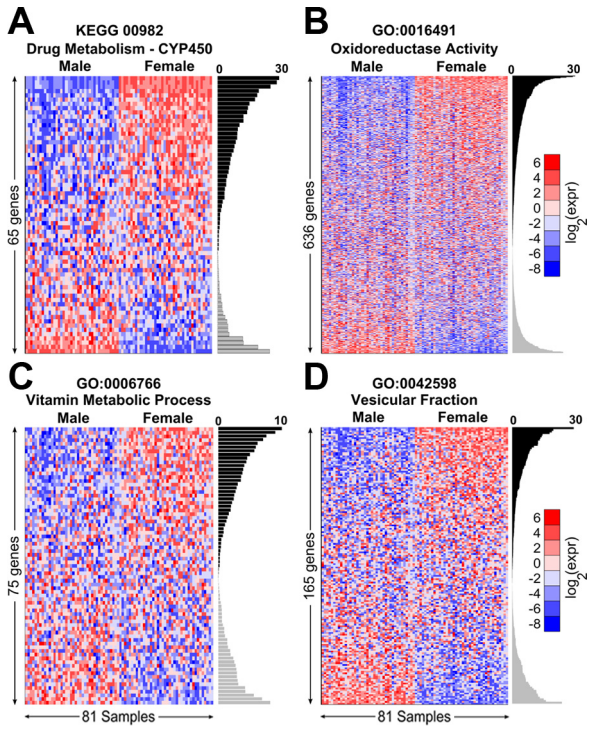


Fig. 2. KEGG and Gene Ontology (GO) categories that exhibit most significant sex bias in mouse liver gene expression. Representative KEGG pathway (*A*), GO Molecular Function (*B*), GO Biological Pathway (*C*), and GO Cellular Component (*D*) categories that are highly significant for sex bias are shown as heat maps of expression for all genes (no. of genes is shown) in each category, several significantly differentially regulated. In the heat maps, red indicates high expression and blue indicates low expression (see color bar *insets*). Genes are sorted by the $-\log_{10}(P$ value) for the sex effect (from the ANOVA analysis), which is plotted next to each heat map with genes expressed more highly in females colored black and those more highly expressed in males colored gray.

certainty only for 2,850 of these genes (due to rather insufficient annotation of the custom array used in the human study), and 1,258 (44.1%) of these were also sexually dimorphic in the mouse. We compared the direction of expression difference between sexes in this subset of 1,258 transcripts (Supplemental Table S4) and found that 747 (59.4%) genes were sex biased in the same direction in mouse and human and 511 (40.6%) were biased in the opposite direction. The genes that exhibited unidirectional sex bias between species were enriched in GO Biological Process categories Oxidation Reduction (GO: 0055514, $P = 3.05 \times 10^{-9}$, $q = 5.60 \times 10^{-6}$) and Glutathione Metabolic Process (GO: 0006749, $P = 1.22 \times 10^{-4}$, $q = 0.045$) and GO Molecular Function categories Monooxygenase Activity (GO: 0016491, $P = 1.68 \times 10^{-7}$, $q = 1.48 \times 10^{-4}$) and Catalytic Activity (GO: 0003824 $P = 2.47 \times 10^{-5}$, $q = 7.24 \times 10^{-3}$).

Gene Expression Quantitative Trait Locus Analysis

Quantitative trait locus (QTL) mapping is a statistical technique that finds associations between phenotype and genotype in a genetically segregating population (32). Here, we performed eQTL mapping on the male and female data separately to determine whether genetic polymorphisms are responsible for variation in transcript abundance. First, we filtered the transcript expression data by heritability, retaining the top 5,217 transcripts with heritabilities greater than the 75th percentile ($h^2 > 0.136$). We performed eQTL mapping on the male and female data separately and also on the mean and difference of the male and female data. There were 1,638 significant eQTLs ($q \leq 0.1$, $P \leq 0.079$) found for liver gene expression in males (Supplemental Table S4a) and 2,076 ($q \leq 0.1$, $P \leq 0.105$) in females (Supplemental Table S4b). Performing eQTL mapping on the mean of the male and female data uses all of the data and increases the power to detect additive effects. Using the full data set, we found 2,925 significant eQTL ($q \leq 0.1$, $P \leq 0.0005$) (Supplemental Table S4c). The difference between the male and female data produced only five significant eQTLs ($q \leq 0.1$, $P \leq 0.199$), which suggests either that genotype by sex interactions are rare in the BXD strains or that this population is not of sufficient size.

First, we explored differences in patterns of eQTL locations between sexes by plotting the genomic locations of each eQTL versus the transcript location (Fig. 3, *A* and *B*). The strong diagonal band represents local eQTLs (43), for which the eQTL is located within 1 Mb of the transcript. Points off the diagonal represent distant eQTLs (43) for which the eQTL is located farther than 1 Mb from the transcript or on another chromosome. A histogram of the number of significant eQTLs associated with each genomic marker is also shown for each sex (Fig. 3, *D* and *E*). Markers associated with a larger number of transcripts than expected by chance, termed eQTL bands, may lie near genes that regulate the expression of clusters of genes. Legitimate concern has been expressed over the validity of eQTL bands, and a permutation approach to assess their significance has been proposed (7). We applied this approach (see METHODS) and colored the eQTL bands by their empirical *P* value. In males, there are four significant ($P \leq 0.05$) eQTL bands on Chr 2, 7, 12, and 17 and in females four eQTL bands on Chr 1, 2, 12, and 17. The genomic regions covered by the eQTL bands on Chr 2, 12, and 17 overlap in males and females and contain 25, 33, and 20 transcripts in common, respectively.

Second, to further explore the similarities in genetic control of liver gene expression between sexes, we compared the eQTL overlap between males, females, and all of the data as well as the overlap of local and distant eQTLs (Fig. 4, *A–C*). It is evident that a larger number of local eQTLs are shared between the two sexes than distant eQTLs. The strongest distant eQTL band on distal Chr 12 is shared between sexes and has been previously reported (19). There are 87 transcripts regulated by this locus in the male data and 45 transcripts in the female data; however, the two sets of genes are not identical. The union consists of 101 genes, and the intersection of these two sets consists of only 33 genes. Several GO categories are significantly overrepresented in the union of these genes: Response to Cytokine Stimulus (GO: 0034097, $P = 0.0050$), Calcium-Mediated Signaling (GO: 0019722, $P = 0.0070$), Serine-Type Endopeptidase Inhibitor Activity (GO: 0004867, $P = 0.0064$), and Cyclin-Dependent Protein Kinase Activity (GO: 0004693, $P = 0.0098$). A clade of serine protease inhibitors contains a putative regulator for these genes, as discussed in Reference 20. GO and KEGG pathway enrichment analysis for other sex-specific eQTL bands on Chr 1, 2, 7, and 17 did not produce significant categories or pathways ($q \leq 0.1$). No TFBS enrichment ($P \leq 0.05$) exists for genes in theses eQTL bands.

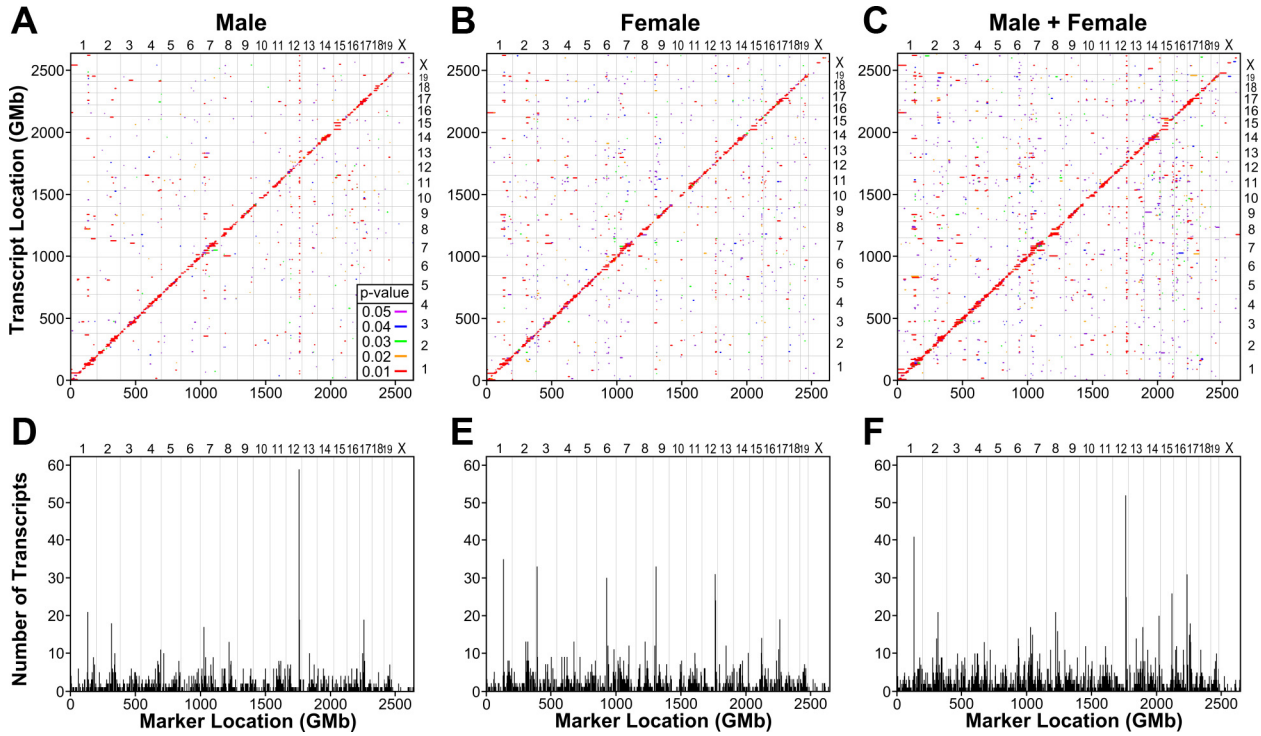


Fig. 3. Sex-specific transcriptome maps reveal differences and similarities in the genetic regulation of gene expression in mouse liver. *A* and *B*: male (*A*) and female (*B*) transcriptome maps in which significant ($P < 0.05$) gene expression quantitative trait locus (eQTL) locations are shown. Genomic locations of each single nucleotide polymorphism (SNP) marker (*x*-axis) and each transcript (*y*-axis) are plotted. Each cross represents the location of the maximum QTL for a particular gene. Locally regulated genes are located along the 45° lines, while the vertical lines correspond to the loci that regulate distant genes. The color of each symbol corresponds to the significance of the eQTL (color bar). *D–F*: eQTL histograms of the male (*D*) female (*E*), and combined (*F*) data, counting the number of significant eQTL associated with each marker, with bands colored by permutation-based *P* value (see *inset* in *A*).

To examine whether the transcripts that have significant eQTLs in both sexes are identically regulated in males and females, we plotted the male versus female eQTL locations for 1,030 transcripts (of 1,152) with known genomic locations (Fig. 4*D*). For 757 (73.5%) of the transcripts, the eQTL is in the same location (diagonal band in Fig. 4*D*). For 273 (26.5%) transcripts, one sex had a single significant eQTL while the other had two, one at the same location and another elsewhere (off-diagonal points in Fig. 4*D*), suggesting that sex is a factor in the regulation of these transcripts. Still, for all transcripts that had a single significant eQTL in one sex and two in the other, we found that either sex had a less significant ($P \geq 0.05$) second eQTL. We examined the direction of the eQTL effect for each allele in the 757 eQTLs that replicate between males and females. We found that transcript expression was higher when the C57BL/6J allele was present for 394 (52%) eQTL and that expression was higher when the DBA/2J allele was present for 363 (48%) eQTLs. In all cases, the effect size was in the same direction in both sexes (Supplemental Fig. S1).

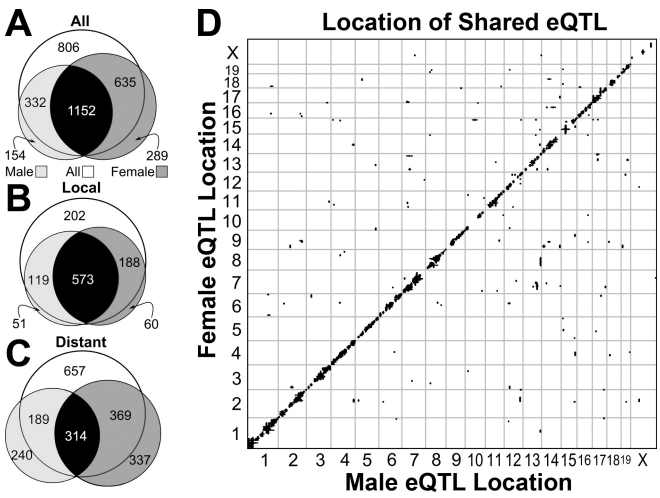


Fig. 4. Distribution of liver eQTLs between sexes in BXD recombinant inbred mice. *A–C*: Venn diagrams displaying the number of intersecting eQTLs in males and females for all (*A*), local (*B*), and distant (*C*) eQTLs. Numbers beside the circles indicate total number of eQTL in each sex. *D*: location of shared eQTL in males plotted vs. the eQTL location in females for the same transcript.

Third, we compared the sex-specific eQTL locations obtained from BXD RI mice with those from an F_2 cross study of Wang et al. (57). While the comparisons of eQTL results are challenging because of the varying genetic backgrounds between populations, differences in microarray probe sequences, experimental conditions, and statistical techniques, and likely environmental differences inherent in different animal facilities, it is likely that eQTLs that overlap may serve as strong candidate loci for further investigation (39). Wang et al. (57) reported mouse liver eQTLs for 8,349 transcripts ($P < 0.001$), and we subselected the 3,855 most significant ($P < 10^{-6}$) eQTLs for 2,667 transcripts. An eQTL was deemed reproducible between two studies if the maximal eQTL interval intersected, and we found that ~25% of local and 8.5% of distant

Table 4. *Genes with significant differential correlation between sexes, sorted by the number of correlated genes*

Probe ID	Gene	Gene Name	Sex Bias	Chr	No. of Genes
A_51_P402909	*Ncf2*	neutrophil cytosolic factor 2	F	1	164
A_51_P110576	*Lgmn*	legumain	F	12	158
A_51_P188704	*Unc13a*	unc-13 homolog A (*C. elegans*)	M	8	153
A_51_P229990	*Rsb30*	RAB30, member RAS oncogene family	M	7	148
A_51_P155972	*Clec14a*	C-type lectin domain family 14, member a	F	12	147
A_51_P470432	*Pdlim4*	PDZ and LIM domain 4	F	11	145
A_51_P324636	*Elovl3*	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 3	M	19	140
A_51_P183439	*C77080*	expressed sequence C77080	M	4	139
A_51_P233338	*Olfr46*	olfactory receptor 46	M	7	136
A_51_P115049	*Mpzl2*	myelin protein zero-like 2	M	9	133
A_51_P338040	*Vopp1*	vesicular, overexpressed in cancer, prosurvival protein 1	F	6	127
A_51_P388325	*Clgn*	calmegin	M	8	122
A_51_P365964	*Tor1a*	torsin family 1, member A (torsin A)	F	2	117
A_51_P248394	*Olfr178*	olfactory receptor 178	M	16	113
A_51_P379080	*Adam24*	a disintegrin and metallopeptidase domain 24 (testase 1)	F	8	111
A_51_P357341	*Clic1*	chloride intracellular channel 1	F	17	110
A_51_P221781	*Zfp33b*	zinc finger protein 33B	M	5	108
A_51_P277376	*Slc8a1*	solute carrier family 8 (sodium/calcium exchanger), member 1	F	17	106
A_51_P473086	*Tbxa2r*	thromboxane A_2 receptor	M	10	103
A_51_P513011	*Rad9b*	RAD9 homolog B (*S. cerevisiae*)	M	5	101
A_51_P391754	*Soat1*	sterol *O*-acyltransferase 1	F	1	99
A_51_P483366	*9230114N12Rik*	9230114N12Rik	M	16	97
A_51_P151828	*Flna*	filamin, α	F	X	97
A_51_P243641	*XM_994259.2*	*Mus musculus* myeloid/lymphoid or mixed-lineage leukemia 2	M	15	96
A_51_P372992	*4632428N05Rik*	RIKEN cDNA 4632428N05 gene	F	10	94
A_51_P485472	*4933430H15Rik*	RIKEN cDNA 4933430H15 gene	F	3	94
A_51_P152845	*Trim24*	tripartite motif-containing 24	M	6	93

No. of genes, number of genes this transcript correlates with within a sex at Spearman correlation ≥ 0.7.

eQTLs satisfied these criteria (Table 3). This result is consistent with previous observations that local eQTLs are more reproducible than distant eQTLs (39).

It has been reported that genetic polymorphisms that are located within transcript probe sequences may alter apparent transcript intensity in an allele-specific manner and hence create false local eQTLs (2, 14). To investigate this in our data, we intersected the microarray probe sequences with a high-density single nucleotide polymorphism (SNP) data set (52). We did not find evidence for this in our data (see METHODS). Transcripts with SNPs in the probe sequences are identified in Supplemental Tables S5a, b, and c.

Gene Correlation Exhibits Distinct Patterns from Gene Expression Between Sexes

While it is clear that many genes differ in expression between the sexes, there is also a more subtle type of variation that involves variability in the correlation among genes within each sex, often without pronounced difference in expression. The sex-specific clusters of such differentially correlated genes may represent complex networks of transcriptional regulation that are important in physiology and pathophysiology.

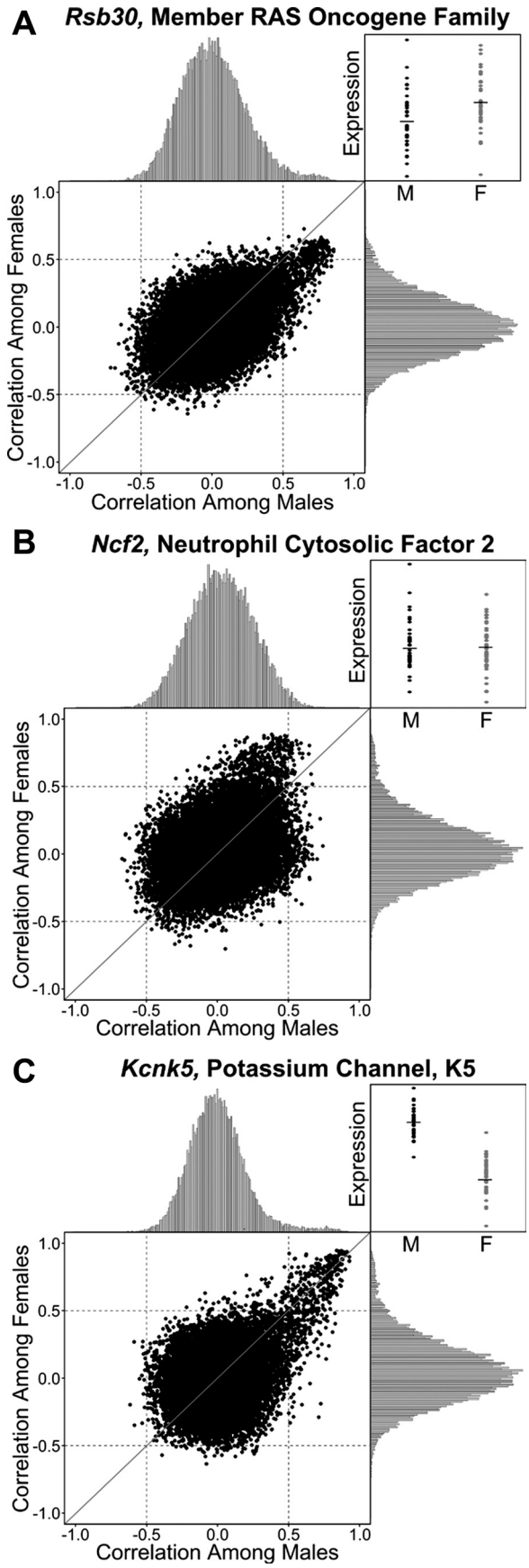
Table 3. *Comparison of BXD eQTLs with BXH eQTLs from Wang et al. (57)*

	Male	Female	All
BXD eQTLs	1,638	2,076	2,925
Local	743	821	1,118
Transcripts in common	315	336	428
Match (% of local)	172 (23.1)	172 (21.0)	233 (21.5)
Distant	718	1,020	1,493
Transcripts in common	172	194	272
Match (% of distant)	24 (3.3)	24 (2.3)	24 (1.7)

eQTL, gene expression quantitative trait locus. Columns do not sum because transcripts without known genomic locations cannot be called local or distant.

On a genomewide level, we examined the extent of correlation in gene expression between sexes. We selected genes with Spearman correlation ≥ 0.7, an arbitrary cutoff that reflects a balance between lower values, which quickly produced so many genes that analysis proved computationally intractable, and higher values, which yield only a small set of genes. We searched for genes exhibiting three types of patterns. The first type contains genes with stronger correlations in males than in females (i.e., male-biased correlations). Member of RAS oncogene family *Rsb30* (Fig. 5*A*) is a representative gene that does not differ in expression between sexes yet exhibits strong correlation in males. The genes that are highly correlated (>0.7) with *Rsb30* in males do not show significant enrichment for any GO category or KEGG pathway. However, the GO categories with the lowest *P* values ($P < 0.0023$) relate to male sexual development, suggesting that there may be some relevance to liver metabolism in males. The second type of pattern, female-biased correlations, is exemplified by neutrophil cytosolic factor (*Ncf*)2 gene (Fig. 5*B*). *Ncf2* is involved in superoxide production, and the genes highly correlated (>0.7) with *Ncf2* in females are enriched for Immune Response (GO: 0006955, $P = 8.24 \times 10^{-11}$). Finally, there are genes that have an equal level of correlation in both sexes regardless of expression levels, exemplified by potassium channel K5 (*Kcnk5*) gene (Fig. 5*C*).

To establish whether genes that have the highest difference in expression between sexes (Table 2) also would exhibit strong correlation, we selected genes that had the largest number of correlations with other genes in one of the sexes



(Table 4). It is noteworthy that there is only one gene (*Elovl3*) that is both differentially expressed between sexes and exhibits strong correlation in only one sex (males).

There are also differentially correlated gene clusters in which a set of genes is highly correlated (≥0.7) with each other in one sex but not in the other (Fig. 6). We focused on the genes that had strong correlation in one sex but considerably weaker correlation in the other (mean absolute correlation difference ≥0.35 between the sexes). There were 33 differentially correlated clusters in males consisting of 254 unique genes (Fig. 6, *A–C*) and 28 clusters in females containing 332 unique genes (Fig. 6, *D–F*). The male-specific clusters are related through a complex network of correlations (Fig. 6*A*) that are neither primarily driven by expression differences between sexes (Fig. 6*B*) nor correlated in females (Fig. 6*C*). The female clusters exhibit a similar pattern in the opposite direction (Fig. 6, *D–F*). Finally, we searched for clusters of genes that are strongly correlated (≥0.7) in both sexes. There were 2,639 genes with high correlation in both sexes (Fig. 6, *G* and *I*), and most of these genes were also differentially expressed between sexes (Fig. 6*H*).

DISCUSSION

Sex-Dependent Differences in Gene Expression in Liver: Physiological and Toxicological Significance

In this study, we compared liver gene expression of males and females in a panel of BXD RI mice in an effort to understand whether genetic background affects sexually dimorphic gene expression. Our results show that sex differences in liver gene expression are largely determined by the differences in endogenous steroid metabolism between sexes, and to a lesser extent by the differences in genetic polymorphisms. Indeed, the role for liver steroid metabolism in sex differentiation is well established, and it is known that the neonatal exposure to androgens is a crucial imprinting mechanism (24, 25). The physiological process of sexual differentiation requires differentiation of hepatic processing of steroids in fetal liver, and this differentiation results in an epiphenomenon of sex differences in xenobiotic processing.

It has been long known that females exhibit a higher incidence of drug-induced liver injury (34). There are significant pharmacokinetic differences between the sexes that are likely due to significant differential expression in CYPs, glutathione *S*-transferases (Gst), sulfotransferases (Sult), UDP glucuronosyltransferases (Ugt), and ATP-binding cassette (ABC) transporters (49). Consistent with previous work, we found a large number of differentially expressed CYPs, including the previously reported female-biased genes *Cyp2a4* (54), *Cyp2b9*, and *Cyp3a16* (58) and the male-biased *Cyp1a2* (58) and *Cyp2d9* (54). Overall, 39 of the 69 CYPs on the array (56%) were

Fig. 5. Differential correlation of gene expression occurs between males and females, independent of expression level differences. *A*: the correlation of *Rsb30* with all other transcripts on the array in male samples is plotted against the correlation in female samples. *Rsb30* is more highly correlated with other genes in males than in females, despite little difference in expression ($P = 0.005$). *B*: *Ncf2* exhibits the opposite pattern, with more correlation among females than males, again with little difference in expression ($P = 0.645$). *C*: *Kcnk5* is highly correlated with many genes in both males and females, despite having a large expression difference between sexes ($P = 1.9 \times 10^{-18}$).

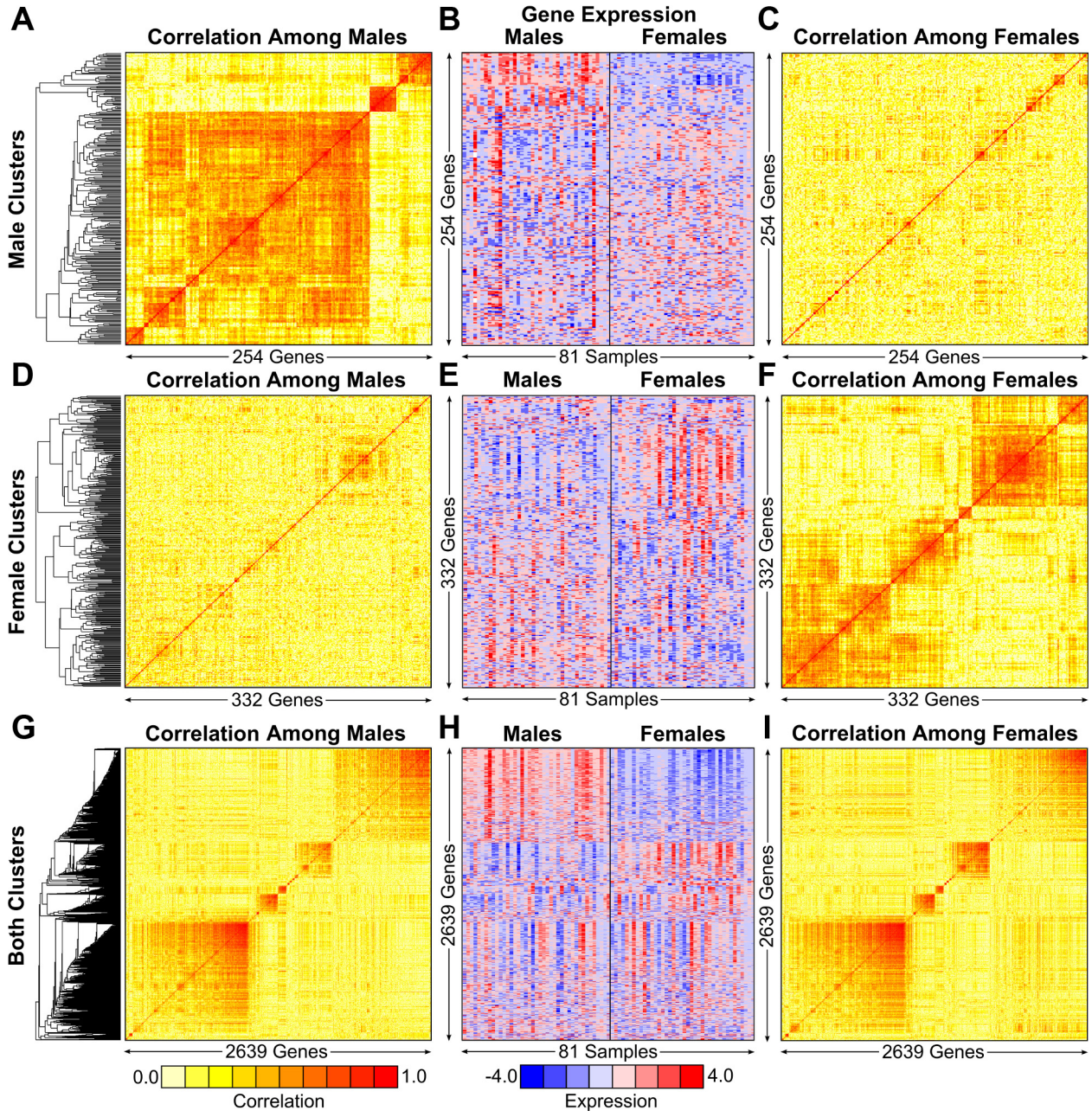


Fig. 6. Correlation analysis of gene expression reveals sex-specific clusters. The gene-gene correlation of clusters of genes more highly correlated in males than in females is plotted in *A*, clustered by correlation within males. *B* shows a gene expression heat map for these same genes, separated by sex and clustered as in *A*. Note that the clusters are not dominated by expression differences between the sexes. *C* shows the correlation of these genes in females. Note that there is strong correlation in the males (bright yellow clusters) and little to no correlation in the females (red). *D–F* show the same types of data for genes more highly correlated in females than in males. In this case, there is little correlation for these genes among males (red in *D*), while there is strong intergene correlation in females (yellow in *F*). *G–I* show clusters that are highly correlated in both sexes, evidenced by the yellow blocks in both *G* and *I*. In this case, many, but not all, of the clusters arise because of expression differences between the sexes.

differentially expressed, with 26 CYPs biased toward high female expression and 13 male biased. However, we also found *Cyp2j13* to be male biased and *Cyp2e1* female biased, suggesting that sex differences of important xenobiotic enzymes may also be affected by a complex interaction with genetics.

Higher expression of most Sult genes was found in females, including *Sult3a1* (3), *Sult1a1*, *Sult1c1*, *Sult1c2*, *Sult1d1*, and *Sult1e1*. In agreement with previous reports, we find that *Gstp1* is male biased and *Gsta3* is female biased (36). We also find that *Gstm7* is male biased and that most differentially expressed Gsts are female biased (11 vs. 23). We found that seven of nine (77.8%) Ugt genes on the array were differentially expressed at q values ≤ 0.01. Ugt genes were found to be exclusively male biased, which is consistent with the finding that Ugt expression is regulated by male GH production in the liver (8). ABC transporters affect the disposition of many endo- and xenobiotics. About 44% of these genes on the array (out of 50) were more highly expressed in females than in

males, whereas only three, including *Abcg2* (35), were more highly expressed in males. Similarly, male bias was observed in expression of epoxide hydrolases 1 and 2 (*Ephx1*, *Ephx2*) and critical cofactor supply genes hexose-6-phosphate dehydrogenase (*H6pd*) and nicotinamide nucleotide adenyltransferase 1 (*Nmnat1*).

The retinol binding protein (*Rbp1*), alcohol dehydrogenase 1 (*Adh1*), and retinol dehydrogenases 6 and 9 (*Rdh6*, *Rdh9*) exhibit strong female bias in liver gene expression. Estrogen is known to control retinoic acid metabolism in many tissues (33), and there is considerable overlap between alcohol and retinol metabolism (27). Animal models for the study of alcoholism suggest that, in general, females drink more alcohol than males and there are sex-related differences in drinking patterns (15). Interestingly, *Rbp1* has been proposed as a locally regulated candidate gene for alcohol preference in the mouse based on the meta-analysis of the brain gene expression data (38). Furthermore, *Rbp1* expression in the liver shows a strong negative correlation (−0.783) with alcohol preference in BXD strains (WebQTL Phenotype no. 10158; Ref. 23), further supporting the hypothesis that *Rbp1* may play a role in alcohol preference through its involvement in alcohol metabolism in multiple tissues. However, *Rbp1* did not have a significant liver eQTL in either sex and did not show significant differential correlation. Likewise, previous QTL studies examining alcohol preference in BXD and AXB mice have not revealed a QTL for alcohol preference in either sex that maps to the genomic location of *Rbp1* (15). Thus, while additional studies that address the role of Rbp1 in alcohol preference and metabolism may yield additional information, it is also possible that the genetic predisposition for elevated alcohol consumption extends to thousands of transcripts forming many functional groups (38), which may render a candidate-gene approach futile.

It should be noted that not all sex-specific gene expression differences found in the mouse translate to humans. While the mapping of mouse microarray probes to human array probes limited the comparison, we found that many genes are biased in the same direction in each sex, suggesting that the mouse is an appropriate model for studying human liver disease and function. Still, even though both human and mouse sex-specific liver gene expression differences are prominently enriched with xenobiotic metabolism, important gene-directed distinctions exist. One example is flavin-containing monooxygenase 3 (*Fmo3*), a key player in metabolism of numerous substrates, including nicotine, tertiary amines, drugs, carbamates, and organophosphates (9). It is a prominent female-biased gene in the mouse liver, but in humans *Fmo3* is known to be expressed in livers of both sexes (31). Thus extrapolations between studies and species should be exercised with caution.

Transcriptional Control of Sex Differences in Liver Gene Expression

The sex difference in expression of many drug metabolism and other liver-expressed genes has been reported to be largely regulated by the temporal pattern of plasma GH release by the pituitary gland (13, 58). These differences are most pronounced in rodents, where plasma GH profiles are highly pulsatile in males but are nearly continuous in females. Hnf4α and Stat5b are key transcription factors that respond to sexually dimorphic GH secretion pattern responsible for upregulation of male-biased genes (60). Our study did not identify enrichment in binding sites for GR, Stat5b, or HNF4α among sexually dimorphic genes in the liver of the BXD mouse panel. While Stat5b and Hnf4α undoubtedly play a major role, and can be identified in comparisons of male and female mice of the same strain, our result suggests that the regulation of sex-dependent gene expression on the population level may be complex, and both direct and indirect linkages in the network of regulators that maintain sex-dependent liver homeostasis are yet to be determined. Instead, we found that HNF1α was enriched among genes more highly expressed in females and PAX5 was enriched among genes with high expression in males. *Hnf1α* knockouts survive only until weaning, suggesting that *Hnf1α* is critical for normal liver development and function (40). In our data, *Hnf1α* was more highly expressed in females ($P = 7.67 \times 10^{-4}$). *Hnf1α* upregulates the expression of several acute-phase proteins (4), and we found several, such as albumin ($P = 5.55 \times 10^{-09}$), α-1-antitrypsin ($P = 4.29 \times 10^{-10}$), angiotensin ($P = 3.49 \times 10^{-15}$), and insulin-like growth factor I ($P = 7.45 \times 10^{-08}$) that are more highly expressed in females. However, the cholesterol oxidizing genes that are regulated by *Hnf1α*, *Cyp7a1* ($P = 0.0129$), and *Cyp27a* ($P = 0.0828$) are not differentially expressed between the sexes. *Pax5* did not show differential expression between sexes ($P = 0.633$), and its function has primarily been studied in the central nervous system and B cell development (1). *Pax5* has been shown to regulate gene involved in cell signaling, adhesion, and migration in B cells (46), but further research is needed to validate whether it is involved in the same functions in the murine liver.

Genetic Component of Sex-Specific Gene Regulation

It is likely that mechanisms other than growth and sex hormones may also contribute to regulation of sex-specific hepatic gene expression. QTL mapping has been adapted to gene expression data to find expression QTLs that represent genomic loci responsible for differential transcriptional regulation (45). The sex-specific eQTL hotspots discovered in our study provide some evidence for the sex-specific differences in genetic control that may be also important. We have assessed the existence of eQTLs in males and females by considering gene expression data from each sex separately. While three loci identified as putative regulatory hotspots are shared between males and females, including the strongest regulatory locus on Chr 12 (19, 20), there were few additional significant sex-specific putative regulatory loci.

Ghazalpour et al. (22) also performed eQTL mapping using a C57BL/6J × DBA/2J intercross but did not include all eQTL results in their publication, so we were not able to perform a detailed comparison with their study. It was reported by the authors that significantly enriched eQTL bands on Chr 3, 6, 16, and 19 were identified in females. These bands did not reproduce in our study, and there are several possible explanations for the discord. The microarrays used, while both based on Agilent technology, were different in probes, transcripts, as well as regions of overlapping transcript. Most importantly, whereas the BXD mice in our study were fed a normal chow diet and were killed at a mean of 70 days, a high-fat diet and a 16-mo study length were employed by Ghazalpour et al. (22).

In addition, Yang et al. (61) examined the interaction of sex with genetic background in liver gene expression. It is of

relevance that our study was based on a cross between C57BL/6J and DBA/2J strains, whereas Yang et al. (61) used an intercross between C57BL/6J and C3H/HeJ. The genomes of these three strains differ at 37.2% of known SNPs [**http://compgen.unc.edu/DisplayIntervals/DisplayIntervals.html**], which may at least partially account for the lack of overlap. This also means that only 62.8% of the mouse genome has been explored in these two studies and that much work remains to be done in studying the effect of genetics on sex differences. Collectively, since the number of genes regulated by the QTLs in a sex-specific manner is relatively small and is poorly reproducible between studies, we posit that it is unlikely that genetic polymorphisms play a major role in controlling sex-specific gene regulatory networks.

Our data show that there were several hundred sex-biased eQTLs and slightly more than 2,925 total eQTLs out of 5,217 transcripts that were probed. This raises the question of how transcripts without significant eQTL are regulated. First, it is plausible that most genes may be regulated via signaling and germline effects may be fairly narrow. Second, the genetic diversity of the BXD panel used in this study is rather limited, and other panels should be used to detect additional regulatory loci. Finally, the modest sample size of the BXD panel limits the power and allows detection of eQTLs with large effect sizes. As gene expression is a complex trait, it is expected that genes with heritable expression may also be regulated by multiple loci.

Correlation Analysis Reveals Sex-Specific Gene Expression Networks

While differential expression plays a large role in sexual dimorphism, it was also shown that there is a more subtle, yet important, role played by varying degrees of differential correlation among gene expression networks between the sexes (56). Correlated genes represent functionally related expression networks that are often under common transcriptional regulation (10). Several of the clusters, such as the male immune response cluster and the female ribosomal cluster, show functional coherence that has been annotated in GO or KEGG. However, many of the clusters discovered here have not been previously reported and represent coexpression networks that may help to explain differences between male and female liver metabolism.

One of the male-specific correlated clusters with the largest differential correlation has a mean absolute correlation of 0.68 in males and 0.30 in females. It is significantly enriched in acute-phase response genes (GO: 0006953, $q = 4.6 \times 10^{-8}$) and contains three serum amyloid A (Saa) genes (*Saa1*, *Saa2*, *Saa3*) along with 16 other highly correlated genes. With the exception of the Saa genes, which are located together on mouse Chr 7, the rest of the genes in this cluster are not colocated in the mouse genome. The Saa genes produce acute-phase proteins in response to inflammatory stimuli (16) and are associated with insulin resistance in mice (47). Several other genes in this cluster have also been implicated in inflammation or the immune system, including lipocalin (*Lcn*)2 (63), intercellular adhesion molecule (*Icam*)1 (26), orosomucoid (*Orm*)1 (48), orosomucoid (*Orm*)2 (6), LPS-induced TNF-α factor (*Litaf*) (53), and β-galactoside α2,6 sialyltransferase (*St6gal*)1 (62). Even though other genes in this cluster have been not implicated in the inflammatory response, their strong correlation may suggest that they may also be involved in immune response in liver.

In females, one of the most differentially correlated clusters (mean absolute male cor. = 0.33, mean absolute female cor. = 0.66) is significantly enriched in ribosomal genes (GO: 0003755, $q = 4.9 \times 10^{-6}$) and contains 12 genes, including ribosomal proteins L19, L21, S14, and S27. Another cluster of 11 genes (mean absolute male cor. = 0.34, mean absolute female cor. = 0.69) is enriched for antigen processing and presentation of exogenous peptide antigens (GO: 0002478, $q = 2.47 \times 10^{-6}$), and similar to observation in the male clusters, many genes that highly correlate may have yet undiscovered functional relationships.

The genes that are strongly correlated (≥0.7) in both sexes form two large clusters (Fig. 6, *G* and *I*, *top right* and *bottom left*). The bottom left cluster, consisting of 620 genes and exhibiting no sex bias in expression, shows no significant GO category or KEGG pathway enrichment ($q < 0.05$). The top right cluster, consisting of 740 genes with male-biased expression, shows enrichment for ion channel activity (GO: 0005216, $q = 6.02 \times 10^{-4}$), a category that includes calcium, sodium, and potassium channel transporters. These genes show significant TFBS enrichment for several transcription factors, including myeloid zinc finger 1 (Mzf1) ($P = 3.09 \times 10^{-4}$), which is known to activate calcium-activated potassium channels (51). The third largest cluster consists of 245 genes with a female-biased expression pattern and shows enrichment for immune response (GO: 0006955, $q = 8.32 \times 10^{-12}$). These genes show TFBS enrichment for transcription factor SpiB ($P = 7.14 \times 10^{-5}$) and signal transducer and activator of transcription Stat6 ($P = 1.6 \times 10^{-3}$). SpiB is known to be associated with the transcriptional regulation of immunity (17), and Stat6 controls immune response in many organs, including the liver (18).

In conclusion, this study assessed the hepatic gene expression networks between two sexes in a BXD RI mouse population. Unlike studies carried out in only one strain, the genetic diversity in the BXD population, combined with a comparison to human data, sheds light on the global sexually dimorphic gene expression that is highly relevant to the human population. Studies in mouse populations may allow us to generate testable hypothesis with regard to gene regulation networks and modes of action of toxicants. Also, our findings have several implications for toxicological studies. It is not only the genes that show significant sex differences that are important in assessing toxic responses in males and females; a large number of functionally related genes with small differences between sexes could also contribute to many sex-biased phenotypes. Caution should be taken when applying the results from a study that was carried out only in male subjects (which is very common in toxicological studies) and applying the conclusions to the whole population. Finally, sexually dimorphic gene expression, at least partly, may be due to the genetic control. Other genomic information, like genome structure similarities and identical by descent regions, when combined with eQTL mapping, can facilitate the process of finding candidate regulatory genes.

GRANTS

This study was funded, in part, by grants from the National Institutes of Health (R01-ES-015241, R01-AA-016258, U01-AA-014425, and P42-ES-005948). E. J. Chesler and R. Yordanova were supported by the Office of Biological and Environmental Research, U.S. Department of Energy, under Contract DE-AC05-00OR22725 with UT-Battelle, LLC.

DISCLOSURES

The authors declare that they have no competing financial interests.

REFERENCES

1. **Adams B, Dorfler P, Aguzzi A, Kozmik Z, Urbanek P, Maurer-Fogy I, Busslinger M.** Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. *Genes Dev* 6: 1589–1607, 1992.
2. **Alberts R, Terpstra P, Li Y, Breitling R, Nap JP, Jansen RC.** Sequence polymorphisms cause many false cis eQTLs. *PLoS ONE* 2: e622, 2007.
3. **Alnouti Y, Klaassen CD.** Tissue distribution and ontogeny of sulfotransferase enzymes in mice. *Toxicol Sci* 93: 242–255, 2006.
4. **Armendariz AD, Krauss RM.** Hepatic nuclear factor 1-alpha: inflammation, genetics, and atherosclerosis. *Curr Opin Lipidol* 20: 106–111, 2009.
5. **Barry WT, Nobel AB, Wright FA.** Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinformatics* 21: 1943–1949, 2005.
6. **Baumann H, Held WA, Berger FG.** The acute phase response of mouse liver. Genetic analysis of the major acute phase reactants. *J Biol Chem* 259: 566–573, 1984.
7. **Breitling R, Li Y, Tesson BM, Fu J, Wu C, Wiltshire T, Gerrits A, Bystrykh LV, de Haan G, Su AI, Jansen RC.** Genetical genomics: spotlight on QTL hotspots. *PLoS Genet* 4: e1000232, 2008.
8. **Buckley DB, Klaassen CD.** Mechanism of gender-divergent UDP-glucuronosyltransferase mRNA expression in mouse liver and kidney. *Drug Metab Dispos* 37: 834–840, 2009.
9. **Cashman JR.** Human flavin-containing monooxygenase: substrate specificity and role in drug metabolism. *Curr Drug Metab* 1: 181–191, 2000.
10. **Chesler EJ, Lu L, Shou S, Qu Y, Gu J, Wang J, Hsu HC, Mountz JD, Baldwin NE, Langston MA, Threadgill DW, Manly KF, Williams RW.** Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat Genet* 37: 233–242, 2005.
11. **Churchill GA, Doerge RW.** Naive application of permutation testing leads to inflated type I error rates. *Genetics* 178: 609–610, 2008.
12. **Clodfelter KH, Holloway MG, Hodor P, Park SH, Ray WJ, Waxman DJ.** Sex-dependent liver gene expression is extensive and largely dependent upon signal transducer and activator of transcription 5b (STAT5b): STAT5b-dependent activation of male genes and repression of female genes revealed by microarray analysis. *Mol Endocrinol* 20: 1333–1351, 2006.
13. **Clodfelter KH, Miles GD, Wauthier V, Holloway MG, Zhang X, Hodor P, Ray WJ, Waxman DJ.** Role of STAT5a in regulation of sex-specific gene expression in female but not male mouse liver revealed by microarray analysis. *Physiol Genomics* 31: 63–74, 2007.
14. **Doss S, Schadt EE, Drake TA, Lusis AJ.** Cis-acting expression quantitative trait loci in mice. *Genome Res* 15: 681–691, 2005.
15. **Fattore L, Altea S, Fratta W.** Sex differences in drug addiction: a review of animal and human studies. *Womens Health* 4: 51–65, 2008.
16. **Gabay C, Kushner I.** Acute-phase proteins and other systemic responses to inflammation. *N Engl J Med* 340: 448–454, 1999.
17. **Gallant S, Gilkeson G.** ETS transcription factors and regulation of immunity. *Arch Immunol Ther Exp (Warsz)* 54: 149–163, 2006.
18. **Gao B.** Cytokines, STATs and liver disease. *Cell Mol Immunol* 2: 92–100, 2005.
19. **Gatti D, Maki A, Chesler EJ, Kirova R, Kosyk O, Lu L, Manly KF, Williams RW, Perkins A, Langston MA, Threadgill DW, Rusyn I.** Genome-level analysis of genetic regulation of liver gene expression networks. *Hepatology* 46: 548–557, 2007.
20. **Gatti DM, Harrill AH, Wright FA, Threadgill DW, Rusyn I.** Replication and narrowing of gene expression quantitative trait loci using inbred mice. *Mamm Genome* 20: 437–446, 2009.
21. **Gatti DM, Shabalin AA, Lam TC, Wright FA, Rusyn I, Nobel AB.** FastMap: fast eQTL mapping in homozygous populations. *Bioinformatics* 25: 482–489, 2009.
22. **Ghazalpour A, Doss S, Sheth SS, Ingram-Drake LA, Schadt EE, Lusis AJ, Drake TA.** Genomic analysis of metabolic pathway gene expression in mice. *Genome Biol* 6: R59, 2005.
23. **Gill K, Liu Y, Deitrich RA.** Voluntary alcohol consumption in BXD recombinant inbred mice: relationship to alcohol metabolism. *Alcohol Clin Exp Res* 20: 185–190, 1996.
24. **Gustafsson JA, Ingelman-Sundberg M.** Regulation and properties of a sex-specific hydroxylase system in female rat liver microsomes active on steroid sulfates. I. General characteristics. *J Biol Chem* 249: 1940–1945, 1974.
25. **Gustafsson JA, Stenberg A.** Neonatal programming of androgen responsiveness of liver of adult rats. *J Biol Chem* 249: 719–723, 1974.
26. **Gutierrez-Ramos JC, Bluethmann H.** Molecules and mechanisms operating in septic shock: lessons from knockout mice. *Immunol Today* 18: 329–334, 1997.
27. **Gyamfi MA, Kocsis MG, He L, Dai G, Mendy AJ, Wan YJ.** The role of retinoid X receptor alpha in regulating alcohol metabolism. *J Pharmacol Exp Ther* 319: 360–368, 2006.
28. **Ho Sui SJ, Fulton DL, Arenillas DJ, Kwon AT, Wasserman WW.** oPOSSUM: integrated tools for analysis of regulatory motif over-representation. *Nucleic Acids Res* 35: W245–W252, 2007.
29. **Holloway MG, Cui Y, Laz EV, Hosui A, Hennighausen L, Waxman DJ.** Loss of sexually dimorphic liver gene expression upon hepatocyte-specific deletion of Stat5a-Stat5b locus. *Endocrinology* 148: 1977–1986, 2007.
30. **Holloway MG, Miles GD, Dombkowski AA, Waxman DJ.** Liver-specific hepatocyte nuclear factor-4alpha deficiency: greater impact on gene expression in male than in female mouse liver. *Mol Endocrinol* 22: 1274–1286, 2008.
31. **Koukouritaki SB, Simpson P, Yeung CK, Rettie AE, Hines RN.** Human hepatic flavin-containing monooxygenases 1 (FMO1) and 3 (FMO3) developmental expression. *Pediatr Res* 51: 236–243, 2002.
32. **Lander ES, Botstein D.** Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121: 185–199, 1989.
33. **Li XH, Kakkad B, Ong DE.** Estrogen directly induces expression of retinoic acid biosynthetic enzymes, compartmentalized between the epithelium and underlying stromal cells in rat uterus. *Endocrinology* 145: 4756–4762, 2004.
34. **Maddrey WC.** Drug-induced hepatotoxicity: 2005. *J Clin Gastroenterol* 39: S83–S89, 2005.
35. **Merino G, van Herwaarden AE, Wagenaar E, Jonker JW, Schinkel AH.** Sex-dependent expression and activity of the ATP-binding cassette transporter breast cancer resistance protein (BCRP/ABCG2) in liver. *Mol Pharmacol* 67: 1765–1771, 2005.
36. **Mitchell AE, Morin D, Lakritz J, Jones AD.** Quantitative profiling of tissue- and gender-related expression of glutathione *S*-transferase isoenzymes in the mouse. *Biochem J* 325: 207–216, 1997.
37. **Mode A, Gustafsson JA.** Sex and the liver—a journey through five decades. *Drug Metab Rev* 38: 197–207, 2006.
38. **Mulligan MK, Ponomarev I, Hitzemann RJ, Belknap JK, Tabakoff B, Harris RA, Crabbe JC, Blednov YA, Grahame NJ, Phillips TJ, Finn DA, Hoffman PL, Iyer VR, Koob GF, Bergeson SE.** Toward understanding the genetics of alcohol drinking through transcriptome meta-analysis. *Proc Natl Acad Sci USA* 103: 6368–6373, 2006.
39. **Peirce JL, Li H, Wang J, Manly KF, Hitzemann RJ, Belknap JK, Rosen GD, Goodwin S, Sutter TR, Williams RW, Lu L.** How replicable are mRNA expression QTL? *Mamm Genome* 17: 643–656, 2006.
40. **Pontoglio M, Barra J, Hadchouel M, Doyen A, Kress C, Bach JP, Babinet C, Yaniv M.** Hepatocyte nuclear factor 1 inactivation results in hepatic dysfunction, phenylketonuria, and renal Fanconi syndrome. *Cell* 84: 575–585, 1996.
41. **Rinn JL, Rozowsky JS, Laurenzi IJ, Petersen PH, Zou K, Zhong W, Gerstein M, Snyder M.** Major molecular differences between mammalian sexes are involved in drug metabolism and renal function. *Dev Cell* 6: 791–800, 2004.
42. **Rinn JL, Snyder M.** Sexual dimorphism in mammalian gene expression. *Trends Genet* 21: 298–305, 2005.
43. **Rockman MV, Kruglyak L.** Genetics of global gene expression. *Nat Rev Genet* 7: 862–872, 2006.
44. **Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, Zhu J, Millstein J, Sieberts S, Lamb J, Guhathakurta D, Derry J, Storey JD, Avila-Campillo I, Kruger MJ, Johnson JM, Rohl CA, van Nas A, Mehrabian M, Drake TA, Lusis AJ, Smith RC, Guengerich FP, Strom SC, Schuetz E, Rushmore TH,**

Ulrich R. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 6: e107, 2008.
45. **Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH.** Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422: 297–302, 2003.
46. **Schebesta A, McManus S, Salvagiotto G, Delogu A, Busslinger GA, Busslinger M.** Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity* 27: 49–63, 2007.
47. **Scheja L, Heese B, Zitzer H, Michael MD, Siesky AM, Pospisil H, Beisiegel U, Seedorf K.** Acute-phase serum amyloid A as a marker of insulin resistance in mice. *Exp Diabetes Res* 2008: 230837, 2008.
48. **Schmidt WM, Spiel AO, Jilma B, Wolzt M, Muller M.** In-vivo effects of simvastatin and rosuvastatin on global gene expression in peripheral blood leucocytes in a human inflammation model. *Pharmacogenet Genomics* 18: 109–120, 2008.
49. **Schwartz JB.** The influence of sex on pharmacokinetics. *Clin Pharmacokinet* 42: 107–121, 2003.
50. **Storey JD, Tibshirani R.** Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 100: 9440–9445, 2003.
51. **Sun G, Tomita H, Shakkottai VG, Gargus JJ.** Genomic organization and promoter analysis of human KCNN3 gene. *J Hum Genet* 46: 463–470, 2001.
52. **Szatkiewicz JP, Beane GL, Ding Y, Hutchins L, Pardo-Manuel DF, Churchill GA.** An imputed genotype resource for the laboratory mouse. *Mamm Genome* 19: 199–208, 2008.
53. **Tang X, Metzger D, Leeman S, Amar S.** LPS-induced TNF-alpha factor (LITAF)-deficient mice express reduced LPS-induced cytokine: evidence for LITAF-dependent LPS signaling pathways. *Proc Natl Acad Sci USA* 103: 13777–13782, 2006.
54. **Tullis KM, Krebs CJ, Leung JY, Robins DM.** The regulator of sex-limitation gene, rsl, enforces male-specific liver gene expression by negative regulation. *Endocrinology* 144: 1854–1860, 2003.
55. **van Nas A, Guhathakurta D, Wang SS, Yehya N, Horvath S, Zhang B, Ingram-Drake L, Chaudhuri G, Schadt EE, Drake TA, Arnold AP, Lusis AJ.** Elucidating the role of gonadal hormones in sexually dimorphic gene coexpression networks. *Endocrinology* 150: 1235–1249, 2009.
56. **Voy BH, Scharff JA, Perkins AD, Saxton AM, Borate B, Chesler EJ, Branstetter LK, Langston MA.** Extracting gene networks for low-dose radiation using graph theoretical algorithms. *PLoS Comput Biol* 2: e89, 2006.
57. **Wang S, Yehya N, Schadt EE, Wang H, Drake TA, Lusis AJ.** Genetic and genomic analysis of a fat mass trait with complex inheritance reveals marked sex specificity. *PLoS Genet* 2: e15, 2006.
58. **Waxman DJ, Holloway MG.** Sex differences in the expression of hepatic drug metabolizing enzymes. *Mol Pharmacol* 76: 215–228, 2009.
59. **Waxman DJ, O'Connor C.** Growth hormone regulation of sex-dependent liver gene expression. *Mol Endocrinol* 20: 2613–2629, 2006.
60. **Wiwi CA, Gupte M, Waxman DJ.** Sexually dimorphic P450 gene expression in liver-specific hepatocyte nuclear factor 4alpha-deficient mice. *Mol Endocrinol* 18: 1975–1987, 2004.
61. **Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ.** Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res* 16: 995–1004, 2006.
62. **Zeng J, Joo HM, Rajini B, Wrammert JP, Sangster MY, Onami TM.** The generation of influenza-specific humoral responses is impaired in ST6Gal I-deficient mice. *J Immunol* 182: 4721–4727, 2009.
63. **Zhang J, Wu Y, Zhang Y, Leroith D, Bernlohr DA, Chen X.** The role of lipocalin 2 in the regulation of inflammation in adipocytes and macrophages. *Mol Endocrinol* 22: 1416–1426, 2008.