

Inheritance Patterns of Transcript Levels in F₁ Hybrid Mice

Xiangqin Cui,^{*,1} Jason Affourtit,[†] Keith R. Shockley,[†] Yong Woo[†] and Gary A. Churchill[†]

^{*}Department of Biostatistics, Section on Statistical Genetics, Department of Medicine, Genetic and Translational Medicine Division, University of Alabama, Birmingham, Alabama 35294 and [†]The Jackson Laboratory, Bar Harbor, Maine 04609

Manuscript received May 2, 2006

Accepted for publication July 31, 2006

ABSTRACT

Genetic analysis of transcriptional regulation is a rapidly emerging field of investigation that promises to shed light on the regulatory networks that control gene expression. Although a number of such studies have been carried out, the nature and extent of the heritability of gene expression traits have not been well established. We describe the inheritance of transcript levels in liver tissue in the first filial (F₁) generation of mice obtained from reciprocal crosses between the common inbred strains A/J and C57BL/6J. We obtain estimates of genetic and technical variance components from these data and demonstrate that shrinkage estimators can increase detectable heritability. Estimates of heritability vary widely from transcript to transcript, with one-third of transcripts showing essentially no heritability (<0.01) and one-quarter showing very high heritability (>0.50). Roughly half of all transcripts are differentially expressed between the two parental strains. Most transcripts show an additive pattern of inheritance. Dominance effects were observed for 20% of transcripts and a small number of transcripts were identified as showing an overdominance mode of inheritance. In addition, we identified 314 transcripts with expression levels that differ between the reciprocal F₁ animals. These genes may be related to maternal effect.

MICROARRAY technology has enabled the measurement of transcript levels for thousands of genes simultaneously. It has been used to identify gene expression changes in response to treatments or over time (BROWN and BOTSTEIN 1999; PARMIGIANI *et al.* 2003; CHOUDHURI 2004). However, there is growing interest in the application of microarray technology to investigate the genetic control of transcript levels. The expression level of each gene in a genetically segregating population can be treated as a quantitative trait for genetic analysis (DARVASI 2003). For example, SCHADT *et al.* (2003) conducted microarray experiments on 111 mice from an intercross population and mapped 4339 quantitative trait loci (QTL) that are associated with variation in gene expression. MONKS *et al.* (2004) profiled 167 individuals in 15 human families using microarray and estimated the heritability of genes that are differentially expressed in the children. They found that 31% of these genes are heritable and median heritability is 0.34. Additional studies have been carried out in yeast (BREM *et al.* 2002), mouse (HITZEMANN *et al.* 2003; BYSTRYKH *et al.* 2005; CHESLER *et al.* 2005), rat (HUBNER *et al.* 2005), and human (MORLEY *et al.* 2004) mapping populations.

Another aspect of the genetic analysis of gene expression as quantitative traits is the characterization of the pattern of genetic inheritance in F₁ hybrids (GIBSON

and WEIR 2005). GIBSON *et al.* (2004) compared the gene expression between inbred parents and hybrid F₁ *Drosophila* using microarray and found that 33% of the genes were different between F₁'s and parents. Interestingly, only 25% of transcripts were found different between the two inbred parents and only 2% of these transcripts show additive patterns of inheritance. AUGER *et al.* (2005) analyzed the expression of 30 maize genes using Northern blotting and found that 20 genes exhibited significant dominance effects. These studies indicate that inheritance patterns of gene expression show extensive nonadditivity. VUYLSTEKE *et al.* (2005) studied three *Arabidopsis* accessions and two pairs of reciprocal F₁ hybrids using cDNA arrays. They found that ~30% of genes analyzed are differentially expressed between each pair of the three accessions and only ~10% of the genes show significant dominance effects (VUYLSTEKE *et al.* 2005). A shortcoming of these studies is that pooled biological samples were often used without biological replication and statistical inferences are based on technical variance only. In this study, we compare transcript levels in liver samples from inbred mouse strains A/J (A) and C57BL/6J (B) and their F₁ hybrid progenies from reciprocal crosses (A × B and B × A) using three independent biological samples from each group and two technical replicates of each biological sample. This design allows us to partition the total variation in each transcript into genetic, biological (but nongenetic), and measurement components. These inferences apply to the populations of inbred mice. We

¹Corresponding author: RPHB 327, 1530 3rd Ave. S., Birmingham, AL 35294-0022. E-mail: xcui@uab.edu

applied a shrinkage estimator for each variance component and demonstrate that this improves the estimated heritability of gene expression.

MATERIALS AND METHODS

Animals and tissue collection: Mice were obtained from The Jackson Laboratory and maintained on standard (6% fat) chow diet in a clean mouse room facility on a 14:10 light:dark cycle. F₁ mice were obtained from reciprocal crosses between parental strains A and B. Three female mice were sampled from each strain (A, B, A × B, and B × A) for microarray analysis. Mice were fasted starting at 7:00 AM on the day of collection and collections were completed between 10:00 and 11:30 AM. All mice were 8 weeks (± 4 days) at the time of tissue collection. Two tissue samples were collected from each mouse and stored directly in RNAlater (Ambion, Austin, TX) following dissection and later homogenized in TRIzol (Invitrogen, Carlsbad, CA).

Microarray processing: Total RNA was isolated by standard TRIzol methods according to the manufacturer's protocols, and the quality was assessed using a 2100 Bioanalyzer instrument and a RNA 6000 Nano LabChip assay (Agilent Technologies, Palo Alto, CA). Following reverse transcription with an oligo(dT)-T7 primer, double-stranded cDNA was synthesized with the Superscript double-stranded cDNA synthesis custom kit (Invitrogen). In an *in vitro* transcription (IVT) reaction with T7 RNA polymerase, the cDNA was linearly amplified and labeled with biotinylated nucleotides (Enzo Diagnostics, Farmingdale, NY). Fifteen micrograms of biotin-labeled and fragmented cRNA was then hybridized onto MOE430v2.0 GeneChip arrays (Affymetrix, Santa Clara, CA) for 16 hr at 45°. Posthybridization staining and washing were performed according to manufacturer's protocols using the Fluidics Station 450 instrument (Affymetrix). Finally, the arrays were scanned with a GeneChip Scanner 3000 laser confocal slide scanner. The images were quantified using GeneChip Operating Software (GCOS) v1.2 (Affymetrix).

Statistical analysis of microarray data: *Data preprocessing:* Probe intensity data from all 24 arrays were read into the R software environment (<http://www.R-project.org>) directly from .CEL files using the R/affy package (GAUTIER *et al.* 2004), which was also used to extract and manipulate probe level data to assess data quality and to create expression summary measures. Normalization was carried out using the robust multiarray average (RMA) method (IRIZARRY *et al.* 2003) to form one expression measure for each probe set on each array. RMA processing was performed on all probe intensity data sets together. Briefly, the RMA method adjusted the background of perfect match (PM) probes, applied a quantile normalization of the corrected PM values, and calculated final expression measures using the median polish algorithm.

Statistical analyses of the expression summaries were conducted using the R/maanova package version 2.3.0 (<http://www.jax.org/staff/churchill/labsite/software/Rmaanova>) in the R environment with models and tests specified as below.

Probe set-specific ANOVA models: To identify transcripts that are differentially expressed across the 12 mice in the experiment, we fitted a fixed-effect ANOVA model to the expression data from each probe set using the model

$$y_{jk} = \mu + S_j + \varepsilon_{jk}, \quad (1)$$

where μ is the overall probe set mean, S_j ($j = 1, \dots, 12$) is the deviation from the overall mean in the j th mouse, and ε_{jk} is the deviation of the k th ($k = 1, 2$) measurement from the mean of

mouse j . This model ignores the strain information and treats the two samples from each mouse as independent technical replicates.

To estimate effects of each strain and test for overall genetic effects, a mixed-ANOVA model was fitted to each gene:

$$y_{ijk} = \mu + G_i + M_{j(i)} + \varepsilon_{jk}. \quad (2)$$

Here, the S_j effect in Equation 1 is decomposed into a fixed genetics (strain) effect, G_i ($i = 1, 2, 3, 4$) and a random mouse effect $M_{j(i)}$ ($j = 1, 2, 3$) nested within strain G_i . This model was also used to identify transcripts that show significant variation among mice within strain by testing the mouse effect $M_{j(i)}$ in model (2).

Estimating variance components: We estimated the variance components for strain ($\hat{\sigma}_g^2$), mouse ($\hat{\sigma}_m^2$), and technical replicates ($\hat{\sigma}_e^2$) by treating all effects except μ as random in model (2). These variance components were estimated using the restricted maximum-likelihood (REML) method (SEARL *et al.* 1992; WITKOVSKY 2002). The estimated variance estimates were then stabilized using a shrinkage estimator based on Stein's method (CUI *et al.* 2005). For estimating the mouse variance in each strain, the data were split into four subsets corresponding to the four strains. The mouse variance for each strain was estimated using the same model as model (1) except that the sample effect S_j ($j = 1, 2, 3$ here) was treated as random.

Estimating heritability: Heritability was computed using the estimated strain, mouse, and technical variance components for each transcript as

$$h^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_m^2 + \hat{\sigma}_e^2}. \quad (3)$$

Testing for strain effect, mouse effect, and strain contrasts: To test the overall difference among the 12 mice, the variation explained by the S_j term in model (1) was compared to the technical variance using the shrinkage-based F_s -test (CUI *et al.* 2005). The empirical distribution of the F_s -statistic was established through permutation analysis, where the rows of the design matrix corresponding to the S_j term were shuffled 1000 times randomly while the data were kept unchanged (KENNEDY and CADE 1996; CUI *et al.* 2005). The F_s -statistics calculated from the permutations were pooled across genes that are not significant (XIE *et al.* 2005) at the nominal 0.1 level according to a conservative gene-specific F-test to form one overall empirical distribution. The percentile of the F_s from observed data in the empirical distribution provides an estimate of the P -value for each gene. To test the natural biological variation after accounting for strain effect, the $M_{j(i)}$ term in model (2) was tested against the technical variance using the F_s -test. To test the strain effect (G_i) in model (2), the mouse effect $M_{j(i)}$ was treated as a random effect and the variance explained by the strain effect was tested against the combination of mouse variance and technical variance using the F_s -test. Differences between parental strains, between reciprocal F₁ animals, and between the F₁ and parental animals were tested by contrasting the appropriate components of the strain effect (G_i) in model (2), using a shrinkage-based t -test (CUI *et al.* 2005). The permutation for these later tests is carried out with the constraint that data from technical replicates must remain paired in the permuted data.

Gene lists were generated using a false discovery rate (FDR) of 0.05 (BENJAMINI and HOCHBERG 2000) unless otherwise specified. The fraction of differentially expressed transcripts was estimated using the method mix-o-matic (ALLISON *et al.* 2002). Briefly, the distribution of P -values is modeled as a mixture of a uniformly distributed component representing

the nondifferential transcripts and a beta-distributed component representing the differentially expressed transcripts. The proportion of nondifferential genes (λ_0) is estimated using maximum likelihood. A web implementation of this method at <http://services.ssg.uab.edu/mixomatic/> was used to analyze the P -values obtained from each test. The λ_0 is very similar to the π_0 obtained from the *qvalue* package (STOREY 2002, 2003) at <http://www.bioconductor.org/packages/bioc/stable/src/contrib/html/>.

Estimation of the additive and dominance effects: The expression level of each gene within each strain was estimated using generalized least squares for the G_i term in model (2) after estimating the variance components for the residual (ϵ_{jk}) and mouse [$M_{j(i)}$] terms (SEARL *et al.* 1992). The additive effect, a , is estimated as half of the observed difference between the parental strains. The dominance effect, d , was estimated as the difference between the F_1 and the midparent values.

Statistical tests for additive, dominance, and overdominance effects: The additive effect was tested ($H_0: a = 0$) by contrasting the two parental strains. The dominance effect was tested ($H_0: d = 0$) by contrasting the F_1 and the mean of the two parental strains. Overdominance effect was identified when the F_1 expression value fell outside the range of the two parental strains and was significantly different from both parental strains. All tests were based on F_s with permutation-based critical values and significance level of FDR of 0.05.

Cluster analysis: The hierarchical clustering analysis was based on the correlation distance ($1 - r^2$) of the gene expression among samples (WU *et al.* 2003). Bootstrapping (100 times) was used to evaluate the cluster stability (KERR and CHURCHILL 2001) and a consensus tree was used to summarize the common features of the bootstrapped trees (MARGUSH and McMORRIS 1981). Clades that appeared in $>80\%$ of the bootstrapped trees were shown in the consensus tree.

SNP effect assessment: The allele-specific probe list for the Affymetrix MOE430v2.0 chip was downloaded from <http://arrayanalysis.mbni.med.umich.edu/MBNIUM.html>. This list contains SNPs for many mouse strains on the basis of information in the dbSNP database (build 126). SNPs between strains C57BL/6J and A/J were used to identify the specific SNP-containing probes and these were removed from the Affymetrix chip description file (CDF) using the R package ProbeFilter 1.4.0. Probe sets with fewer than four probes remaining were discarded from further analysis. The .CEL files were preprocessed again using the customized CDF file.

Determination of overrepresented biological categories: Overrepresented categories within lists of differential expressed genes were identified by testing for association with gene ontology "biological process" terms in Expression Analysis Systematic Explorer (EASE) (HOSACK *et al.* 2003) or gene product relationships available in a curated database of biological networks (Ingenuity Pathways Analysis) (<http://www.ingenuity.com>). Enrichment of pathway members among differentially expressed transcripts was assessed with the one-tailed Fisher's exact test for 2×2 contingency tables (Ingenuity) or a conservative adjustment to the Fisher's exact test result (EASE score) that weights significance according to the number of transcripts identified in a list. In the EASE analysis, only transcripts mapping to a unique EntrezGene identification were included for analysis. Because biological themes supported by a single gene in Fisher's exact test are not stable (HOSACK *et al.* 2003), only results containing multiple genes were explored further.

Real-time PCR: Total RNA (1 μ g) was reverse transcribed, employing standard random hexamer priming methods and Superscript III enzyme (Invitrogen) according to the manufacturer's protocols. Diluted reaction products were then used in a subsequent PCR reaction containing Taqman Universal

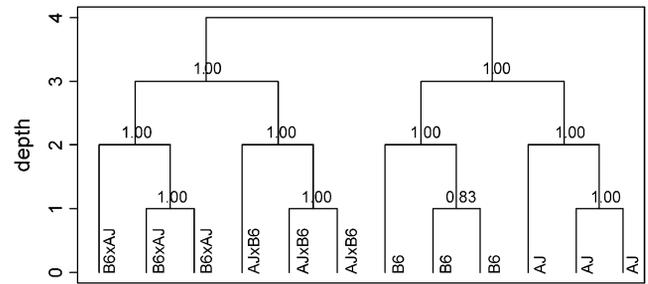


FIGURE 1.—Hierarchical cluster of the 12 mice samples. The samples were clustered on the basis of 6148 genes significantly differently among the 12 samples at FDR 0.001. The number at each branch represents the percentage of times the subbranches cluster together in bootstrapping analysis. Depth is the level of branching.

PCR master mix (Applied Biosystems). Gene-specific primers and probe sets were obtained from the Applied Biosystems Assay on Demand service and used according to manufacturer's protocols. Real-time PCR reactions were performed using the ABI PRISM 7900HT sequence detection system (Applied Biosystems) with recommended thermal cycling protocols and 40 cycles of amplification. Threshold cycle (C_t) values were determined using the supplied sequence detection system (SDS v2.2) software package. Three replicated reactions were conducted for each of the 12 samples used in the microarray study. The *GUS* gene was used as the internal control for all samples.

RESULTS

Overall sample difference and sample clustering: To provide a picture of the overall relationship of gene expression among the 12 mice (3 from each of the four strains: A, B, $A \times B$, and $B \times A$) profiled in this experiment, we first fit an ANOVA model (Equation 1) to each transcript, treating the 12 mice as independent samples without considering their strain identities. A shrinkage-based F_s -test (CUI *et al.* 2005) was applied to each of the 45,000 transcripts present on the array. A total of 6143 transcripts were identified as significantly different among these 12 mice at the high-stringency significance level of FDR 0.001 (BENJAMINI and HOCHBERG 2000). On the basis of the estimated expression level of these transcripts we clustered the 12 mice using a hierarchical clustering method (WU *et al.* 2003). As expected, the genetically defined groups cluster together (Figure 1). The two reciprocal F_1 hybrids cluster together and the two inbred parental strains cluster together, which is similar to what was described in the *Drosophila* study (GIBSON *et al.* 2004).

Natural variation among mice within strain: Some of the differentially expressed transcripts among the 12 samples can be attributed to nonheritable natural variation among individual mice. To identify these transcripts we tested the mouse effect after accounting for strain effect in model (2) and found 1754 transcripts

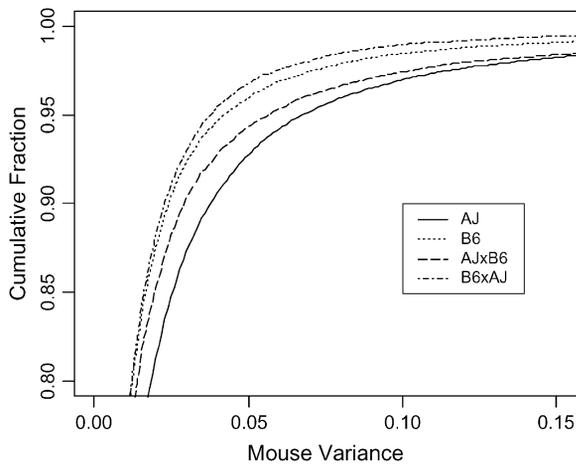


FIGURE 2.—The cumulative distribution function of the shrinkage estimator for the mouse variance from each of the four strains.

(FDR < 0.05) with significant biological variation among mice within a strain. Gene ontology analysis of these genes shows that they are enriched for fatty acids metabolism, amino acid metabolism, and hormone related processes, which indicates that some genes in these biological processes have high natural variation among individual mice with same genotype.

F₁ hybrids are often more uniform in phenotype than inbred individuals (CROW 1998), possibly due to the buffering effects of heterozygosity (HARTMAN *et al.* 2001; HARTWELL 2004). To test whether the gene expression of hybrids has this characteristic, we estimated and compared the variation of gene expression in F₁ mice to variation within the parental strains. Cumulative plots of within-strain variance confirm that the B × A hybrid mice show less individual variation in gene expression than either of the two parental strains (Figure 2). However, the F₁ hybrid mice from the A × B cross show more variation than the parental strain B. Thus gene expression in F₁ hybrids does not necessarily resemble the heterosis phenomena associated with whole-organism phenotypes (BIRCHLER *et al.* 2003). It is curious that the two reciprocal hybrids differ in this regard. Further investigation may be required to confirm that this is a general phenomenon or that it may be particular to these individuals.

Heritability of gene expression: The variation of gene expression can be decomposed into heritable and non-heritable variation. The latter can be decomposed into the within-strain mouse variation and measurement error variation. The fraction of the total variation explained by genotype is known as the broad sense heritability (FALCONER 1986). We calculated the heritability for each transcript on the basis of the estimated variance for genotype, mouse, and measurement. Due to the small number of biologically independent samples variance estimates are not very accurate. We applied a shrinkage adjustment to improve the estimation (CUI

et al. 2005). The shrinkage adjustment is adaptive to the distribution of each variance component. After shrinkage the spread of the measurement variance was reduced substantially, reflecting a level of variation that is consistent with statistical estimation error. The strain and the mouse variances had minimal shrinkage due to their markedly heterogeneous variance distributions (Figure 3, A and B). Both of these variance components have bimodal distributions with about one-third of genes having estimates close to 0 (<10⁻⁶). However, only 8% of genes have both strain and mouse variance components close to 0. About 60% of the genes show both strain and mouse variance components >10⁻⁶ and the correlation (on the standard deviation scale) of these two variance components is 0.47. The high correlation between strain and mouse variance components indicates that genes that have high within-strain variation also tend to have high variation across strains.

We computed heritability as the ratio of between-strain variance to the total variance for both the raw and the shrinkage estimates of variance. The application of shrinkage increased the heritability in many genes. For example, 6503 (14% of the total) transcripts showed heritability >50% without shrinkage. With shrinkage, the number increased to 8131 (18% of the total transcripts). The median heritability of all the transcripts increased to 22% from 16%. Figure 3C shows the histograms of the heritability from transcripts with heritability >1% (~75% of the transcripts present on the chip). The overall median heritability (22%) obtained here is larger than the 11% estimated from the mouse brain expression profile mapping experiment using recombinant inbred (RI) lines (CHESLER *et al.* 2005).

To identify the heritable transcripts, we tested for the strain effect on the basis of the combination of biological variation and technical variation in a mixed-effect ANOVA model (2). We identified 8979 transcripts as significantly heritable (FDR < 0.05) among the four strains of mice. The median heritability of these genes is ~70%. This gene list has 312 overlaps with the gene list for natural biological variation, which indicates that although these genes are naturally variable, the genetic effect is still significant after the nongenetic variation is accounted for. The heritable genetic effect was decomposed into three contrasts among the strains: the contrast between the two parental strains, the contrast between the two F₁ strains, and the contrast between the mean of the two parental strains and the mean of the two F₁ strains. These contrasts yield 8950, 357, and 67 significant transcripts, respectively, at FDR of 0.05. Treating the two reciprocal F₁ hybrids as the same in the third contrast may result in reduced power due to the large difference between the two hybrids. To avoid this problem, we also compared the two hybrids with the midparent value separately. The A × B hybrid gave 317 significant genes while the B × A hybrid gave only 24 significant genes.

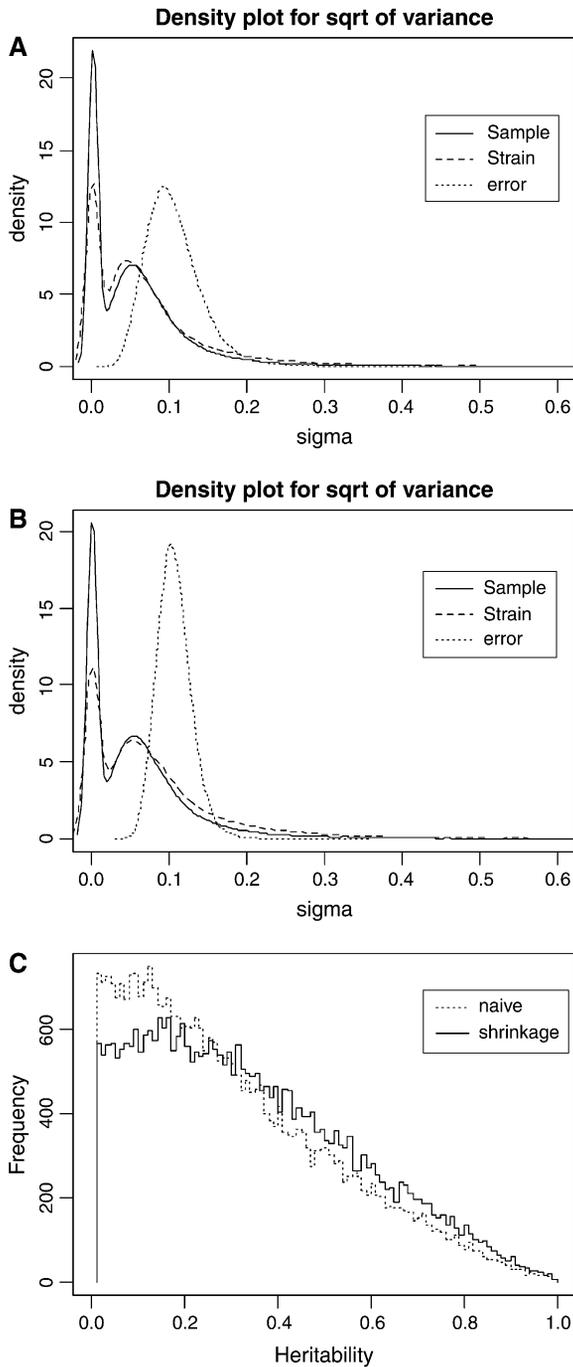


FIGURE 3.—Estimated variance components and heritability. (A) Naive variance components estimated using data from each gene only; (B) shrinkage estimators of the variance components; (C) histogram of heritability. The heritability was computed as the ratio between strain variance and the total variance.

Additive and dominance effects: The inheritance pattern, the difference of gene expression between parental strains and F_1 hybrids, can be decomposed into an additive effect (a) and a dominance effect (d) as in conventional quantitative genetics (FALCONER 1986). For purely additive effects ($d = 0$), the phenotype of the F_1 will be equal to the midparent values. When a dominance

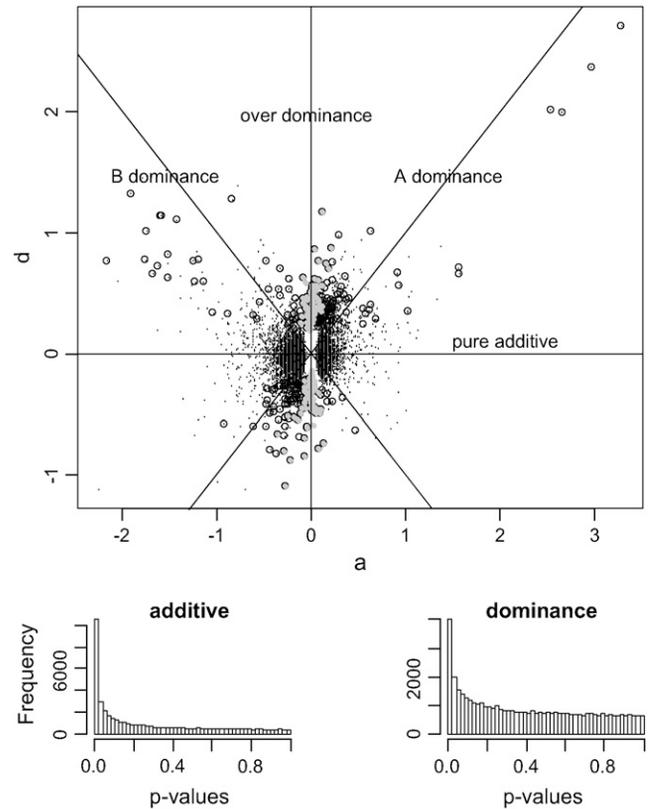


FIGURE 4.—Estimating and testing for additive and dominance effects. (Top) The estimated additive (a) and dominance (d) effects for the F_1 hybrid $A \times B$. Genes that show significant additive effect are represented by solid dots and genes that show significant dominance effect are represented by open circles. Genes on the horizontal line show pure additive effect. Genes on the vertical line show pure dominant effect (overdominance). The line with slope of 1 represents A allele dominant and the line with slope of -1 represents B allele dominant. Shaded diamonds represent genes that show significant overdominance effect. Significance level is FDR 0.05 for all tests. (Bottom) Histograms of P -values obtained from testing for additive and dominance effects in the top. The excess of low P -values indicates the presence of significant genes.

effect is present ($d \neq 0$), the F_1 gene expression differs from the midparent value by an amount d . Due to the presence of technical and biological variation, the estimated a and d effects are never exactly 0. We rely on statistical tests that contrast the parental and F_1 trait values as described earlier to test for the a and d effects. Figure 4 shows the inheritance of the significant genes from some of the tests for the $A \times B$ hybrid. A large number of transcripts are significant for the additive effect and some of them are also significant for the dominance effect. Transcripts that are significant for the overdominance effects tend to have small additive effects and therefore are centered around the $a = 0$ axis.

The numbers of significant genes from testing for the additive and dominance effects are also represented by the excess of small P -values in the P -value histograms (Figure 4). Under the assumption that the P -values

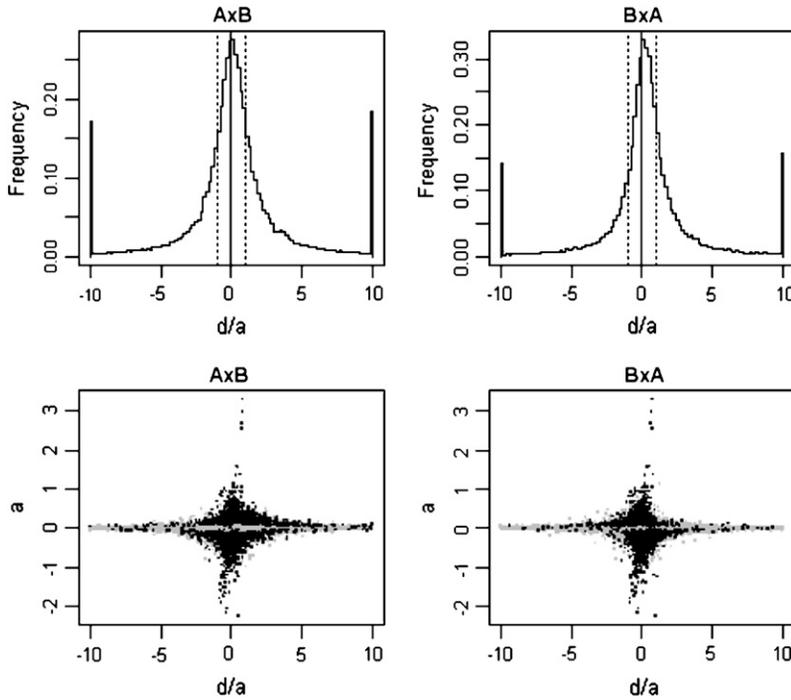


FIGURE 5.—Degree of dominance. The dominance effect, d , is computed as the difference between the mean of a F_1 hybrid and the mean of the parental strains. The additive effect a is computed as half of the difference between the two parental strains. (Top) Histograms of the degree of dominance for the two F_1 hybrids. The solid vertical line indicates no dominant effect. The two dotted vertical lines represent that F_1 is equal to one of the parental strains (full dominant). Left dotted line represents AJ and the right dotted line represents B6. The spikes at -10 and 10 represent the frequency of all transcripts with $d/a \leq -10$ and $d/a \geq 10$, respectively. (Bottom) The relationship between the additive effect and the degree of dominance. Small additive effects cause a large degree of dominance represented by the tight distribution along the $y = 0$ line. Solid points represent the genes that show significant genetic effect at significance level of FDR 0.01 and the shaded points represent genes with no significant genetic effect.

would be uniformly distributed from 0 to 1 if there were no true difference among transcripts, the presence of excessively small P -values would indicate the presence of true differences. The histogram of P -values from testing the dominance effect shows an excess of small P -values but much less than those from the tests for additive effects.

Degree of dominance: To describe the relative size of the dominance effect compared to that of the additive effect in gene expression, we computed the ratio of d/a as a measure of nonadditivity (GIBSON *et al.* 2004). A purely additive inheritance model will have $d/a = 0$ and a fully dominant trait will have $d/a = 1$. If the absolute value is >1 , the effect is considered to be overdominance. To illustrate the degree of dominance, we plotted the histograms of estimated d/a for each F_1 hybrid strain (Figure 5). It appears that many transcripts have a d/a ratio >1 , indicating overdominance. However, this pattern may also be explained by variation in the estimates of the dominance ratios that have associated errors. The excess of transcripts with large dominance ratios may be due to the small additive effect estimates that occur by chance. This is evident when we plotted the additive effect against the degree of dominance. Most transcripts with large d/a ratios have small a estimates (Figure 5) and these genes are less significant for the additive effect. Therefore, formal statistical tests are needed to determine whether the dominance or overdominance effects are significant by considering both the size of the effect and the variation from both biological and technical sources.

Overdominance: Overdominance is commonly associated with hybrid vigor but may also have a negative effect

on fitness (JOHNSON 1980; PAN *et al.* 2005). To declare overdominance, we required that a transcript mean expression in at least one hybrid F_1 lies outside of the range of the parental strains and it should be significantly different from both parents at a significance level of FDR 0.05. We found 158 significantly overdominant transcripts for $A \times B$ hybrids (Figure 4) and most of these have a relatively small additive effect. The test for overdominance among the $B \times A$ hybrids identified only 18 transcripts with significant overdominance effect. The small number of significant genes with overdominance effect further indicates that using just the d/a ratio is not a reliable method to identify overdominant genes.

Estimating the proportion of differentially expressed transcripts: The length of a significant gene list depends on the choice of the significance levels, multiple testing adjustment, the power of the study design, and the size of the actual effects. To obtain an estimate for the percentage of differentially expressed genes independent of the significance level, we applied the mix-o-matic analysis to all P -values obtained from testing individual transcripts. As shown in Figure 4, the number of differentially expressed transcripts from each test is represented by the excess of small P -values as compared to the uniform distribution expected under the complete null. The proportion of differentially expressed genes is estimated as the proportion of the beta component in the mixture of a beta and a uniform distribution using maximum likelihood in the mix-o-matic analysis (ALLISON *et al.* 2002). On the basis of this analysis, 73% of transcripts are differentially expressed among the 12 mice and 68% of transcripts are differentially expressed

TABLE 1
Comparison of results before and after the removal of SNP-containing probes

Tests	Significance level (FDR)	snp-free-ps (snp-ps)		
		Before	After	Common
<i>F</i> -test for samples	0.001	5601 (542)	5623 (509)	5599 (460)
Mouse effect	0.05	1618 (136)	1526 (128)	1526 (114)
Strain effect	0.05	8186 (793)	7855 (766)	7854 (648)
A <i>vs.</i> B	0.05	8149 (801)	7655 (792)	7655 (672)
F ₁ (A × B) <i>vs.</i> F ₁ (B × A)	0.05	292 (22)	217 (16)	217 (9)
F ₁ 's <i>vs.</i> parents	0.05	60 (7)	60 (5)	60 (4)
AB dominance	0.05	295 (22)	525 (46)	295 (18)
BA dominance	0.05	19 (4)	19 (2)	19 (2)
AB overdominance	0.05	149 (9)	122 (9)	122 (4)
BA overdominance	0.05	18 (0)	17 (1)	17 (0)

Eight snp-ps probe sets were removed from testing because there were fewer than four probes remaining. Snp-free-ps, significant probe sets that do not contain any known SNP in any probes; snp-ps, significant probe sets that contain an SNP in one or more probes.

among the four strains after accounting for both biological and technical variation. About 23% of transcripts are differentially expressed among individual mice within strain and about half (56%) of the transcripts are differentially expressed between the two parental strains. Dominance effects were estimated to be present in ~20% of transcripts. Compared with the relative small proportion of transcript with dominance effects, the large proportion of transcripts with additive effects suggests that most transcripts show an additive pattern of inheritance, with F₁ levels similar to the midparent value. The tests for overdominance were not subjected to the estimation for the proportion of differentially expressed transcripts because they were selected with a combination of tests and the assumption of uniform *P*-value distribution under the complete null may not hold. For testing the two F₁ hybrids, 31% of transcripts were estimated to be different between the two reciprocal F₁'s, indicating that the expression of a large number of genes may be affected by the maternal environment.

SNP effect: Probe designs for the Affymetrix array are based on the reference C57BL/6J sequence and it is plausible that strain polymorphisms could alter their hybridization properties. Thus there is a potential to observe differences in hybridization intensity even if the mRNA levels are the same in the two strains. To evaluate the effect of SNPs in this study, we identified individual probes containing known SNP(s) between A/J and C57BL/6J in the dbSNP database. A total of 3453 probe sets (~10% of the total probe sets on the chip) were found to contain such SNPs. We reconstructed the Affymetrix CDF file to exclude the individual probes that contain any SNP from these probe sets. Eight probe sets were completely removed from further analysis because they had ≤3 probes remaining. We reanalyzed the data and compared the results with those from our original analysis (Table 1). The comparison showed

only very minor or no changes in the clustering of samples, the heritability, and the variance components. These global characteristics may not be sensitive to changing a small subset (10%) of probe sets. We also compared the expression fold change between A/J and C57BL/6J for the SNP-containing probe sets and found that ~50% of them increased and 50% of them decreased. This result suggests that SNPs have not significantly reduced the expression measurements for the A/J strain samples.

The number and identity of significant genes from each test after removing the SNP-containing probes were largely unchanged, with allowance for some numerical differences in borderline calls that result from the permutation analysis (Table 1). For almost all tests, ~10% of significant probe sets contain SNPs, which is consistent with the overall rate of SNPs in probe sets on the chip. Therefore, there is no enrichment of SNP-containing probe sets, which might be the case if SNPs contribute heavily to the expression difference between the two strains. However, some differences are observed between the identities of the significant SNP-containing probe sets. For example, the contrast of the two parental strains identified 801 and 792 of the 3453 SNP-containing probe sets as significant before and after removing the SNP-containing probes. The 648 probes sets identified in common between these two analyses showed similar estimated fold change values. The 129 probe sets identified only in the first analysis show reduced fold change after removing SNP-containing probes, while the 120 genes significant only in the second analyses show increased fold change (Figure 6). Dramatic fold change differences were observed for only a few probe sets (the points far away from the identity line). These results indicate that SNPs in individual probes have a relatively small impact on the overall comparison of gene expression changes between strains.

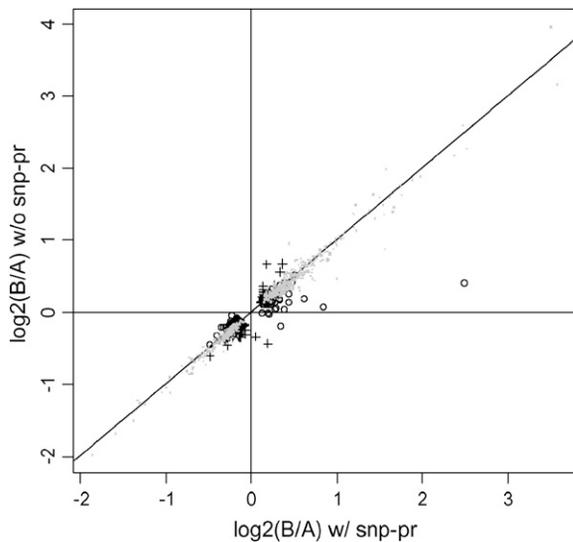


FIGURE 6.—SNP effect on fold change between two parental strains. The fold changes (base 2 logarithm) between C57BL/6J (B) and A/J (A) are compared before and after removing the SNP probes (snp-pr) for probe sets that are significantly different between the two parental strains. The shaded points represent the 672 common significant probe sets from the two analyses. The open squares represent the 129 probe sets significant in the analysis only before removing SNP probes. The cross signs (+) represent the 120 probe sets significant in the analysis only after removing SNP probes. The vertical and horizontal lines represent no fold change between the two parental strains before and after removing SNP-containing probes, respectively. The diagonal line represents the same fold change.

Overrepresented biological categories: To identify biological processes that may be overrepresented in each comparison, we performed statistical tests for overrepresentation of functional categories determined using the EASE analysis platform and the known functional relationships available in the Ingenuity Pathways Knowledge Base (see MATERIALS AND METHODS). Multiple processes, including components of metabolism (including metabolism of proteins, DNA, RNA, phosphates, cofactors, and pyruvate), transcription and RNA processing, cell division, and biosynthesis pathways, were induced in B relative to A. Literature-supported pathways involving the expression of genes related to the metabolism of arachidonic acid, fatty acids, tryptophan, xenobiotic signaling, and amino groups were statistically overrepresented in gene lists that had lower expression in B compared to A. Other processes were expressed at higher levels in A relative to B, including probe sets mapping to genes associated with ion transport, protein localization, metabolism of alcohols and monosaccharides, biogenesis, and signaling. Transcripts associated with the literature-supported processes of N-glycan biosynthesis, the phototransduction pathway, and linoleic acid metabolism were overrepresented in gene lists, showing greater expression on A compared to B. These differences are summarized in supple-

mental Tables S1 and S2 at <http://www.genetics.org/supplemental/>.

Tests for the presence of significant dominance effects contrasting the parental strains and F₁ hybrids were conducted separately for each hybrid. For the A × B hybrid, genes that showed lower expression in the hybrid than the midparent value were associated with macromolecule (protein) biosynthesis. For genes that showed higher expression in the A × B hybrid, categories related to metabolism (fatty acids, amino acids, and amino sugars), signaling (IL-10, IL-6, Integrin, and B cell receptor), and localization were overrepresented. In the test of dominance for the B × A hybrid, probe sets involved in protein localization were more highly expressed in the parental strains and probes sets associated with antigen presentation and the metabolism of fatty acids were more highly expressed in the B × A hybrids.

Differentially expressed transcripts expressed at a higher level when the expression of A × B was higher than B × A included bile acid biosynthesis, arginine and proline metabolism, and death receptor signaling. Ubiquinone biosynthesis and oxidative phosphorylation were overrepresented when the expression of A × B was lower than B × A.

Real-time PCR validation: To validate the results from the microarray measurements, we chose six transcripts that show differential expression between the two parental strains and the expression in F₁'s is more similar to that in parental strain A. The same RNA samples used in the microarray experiment were assayed via real-time PCR with three technical replicates per RNA sample. The results showed that the relative expression among the strains for each transcript was similar to microarray results in both the direction of the difference and the relationship between F₁ hybrids and the parental inbred line (Figure 7). Only the *Cap1* transcript did not confirm the microarray results. Its real-time PCR did not show any difference among the strains while the microarray results showed large differences. This inconsistency could be explained by the fact that the primers used in the real-time PCR experiments are located at the upstream, exon 2/3, of the transcript while the Affymetrix probe set hybridizes to the 3'-untranslated region.

DISCUSSION

Comparing with similar studies in other organisms: In this study we compared the gene expression of two parental strains and their reciprocal F₁ hybrids in mouse liver. We tested the additive, dominance, and overdominance effects and identified lists of genes that are significant for each effect. At an FDR of 0.05, 19.9, 0.8, and 0.4% of all transcripts showed significant additive, dominance, and overdominance effects, respectively. Compared with other studies on gene expression

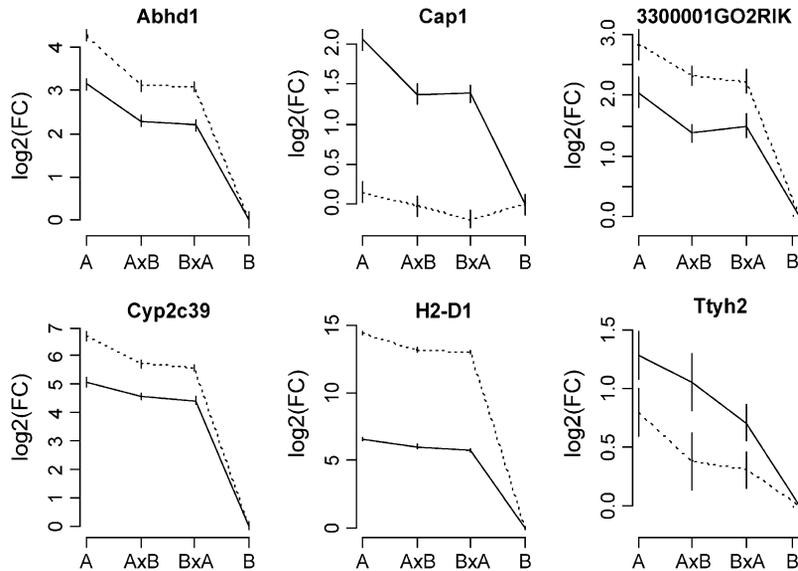


FIGURE 7.—RT-PCR validation of six genes. The \log_2 -fold change of each strain (A, A \times B, B \times A, and B) relative to the parental strain B is plotted for both the microarray (solid lines) and the RT-PCR (dotted lines) results. The standard errors are represented by the vertical bars.

inheritance (GIBSON *et al.* 2004; AUGER *et al.* 2005; VUYLSTEKE *et al.* 2005), our significant gene lists appeared to be substantially shorter in general (Table 2). One obvious reason for the discrepancy is that these studies examine different organisms and different tissues, which are likely to have a large impact on expression patterns (WANG *et al.* 2006). One potential explanation for the differences between the various studies is the effect of the design of these experiments. Our mouse study used three biological replicates per strain and two technical replicates per sample, while most others used one pool of many samples with more technical replicates. Because of the large number of technical replications, the measurements in some of these studies appear to be very precise and the tests appear to be more powerful, especially when replicated spots were treated as independent replicates in the Arabidopsis study

(VUYLSTEKE *et al.* 2005). In studies without true biological replication the tests for differential expression depend on the technical variation in the microarray assay and the scope of the inference is limited. With biological replicates it is possible to construct tests that account for individual (nongenetic) variation in gene expression; the resulting tests have lower power but the conclusions apply to the population of animals from which the sample was obtained (CHURCHILL 2002). For many genes, there is a substantial biological variance component and thus it can be expected that fewer genes will be detected.

One striking feature of Table 2 we would like to draw attention to is the dramatically different criteria used to select differentially expressed genes, which makes any direct result comparison across experiments impossible. Some studies even employed different significance

TABLE 2
Result comparison among similar studies on gene expression inheritance

	Organism			
	Arabidopsis	Maize	Drosophila	Mouse
Technology:	Ara 6k cDNA array	Northern blotting	Agilent long-oligo array	Affymetrix MG430v2.0
Probe no./gene no.:	6008/4876	30/30	21929/12017	45000/39000
Pool or biorep:	Pool (%)	Pool	Pool (%)	Biorep (%)
P1 vs. P2	27–37 ^a	11–15	25	19.9
F ₁ vs. 1/2(P1 + P2)	6.4–21 ^b	19–20	33	0.8
F ₁ vs. F ₁ '	3.7–6.4 ^c	NT	9–20	0.8
Overdominance	9 ^d	NT	5 ^e	0.4
Significance level	Vary	Nominal, 0.05	Bonferroni, 0.05	FDR, 0.05

Biorep, biological replicate; NT, not tested.

^a One to two false positives in each gene list.

^b FDR < 0.001.

^c FDR < 0.01.

^d 99.8% per gene confidence interval.

^e Based on fold change: F₁ is at least 1.25-fold greater or less than both parental lines.

criteria for different gene lists. One way to avoid inconsistency of significance criteria is to estimate the proportion of differentially expressed genes independent of significance levels using the *P*-value distributions. For example, the mix-o-matic algorithm (ALLISON *et al.* 2002) estimates the proportion of genes that are differentially expressed on the basis of the distribution of the *P*-values from individual gene tests. It provides an overall estimate of the proportion of genes that are differentially expressed. In our study, although the significance gene list for additive effect at an FDR of 0.05 is only 19.8%, the mix-o-matic estimation showed that >50% of transcripts are differentially expressed between the two parental strains. However, the identities of these differentially expressed genes are not known. It is not possible to conduct a functional analysis of these 50% transcripts. In addition, this approach requires the assumption of uniformity under the null, which may not hold in some tests. For example, the null hypothesis for testing the overdominance effect is that the expression of F₁ hybrids is within the range of the two parental strains. Under this null hypothesis, the *P*-values will probably not follow a uniform distribution. Therefore, we cannot estimate the proportion of genes that show overdominance effect using the mix-o-matic method.

Another striking difference across these studies is in the degree of nonadditivity, which can be assessed by comparing the lengths of significant gene lists obtained from testing the additive and dominance effects within each study. The fly study (GIBSON *et al.* 2004) showed a very high degree of nonadditivity, with ~40% of genes showing dominance or overdominance effects *vs.* only 25% of genes showing additive effects. In contrast, our mouse study identified 19.9% of genes that were significant for additive effects but only <1% of genes that were significant for dominance or overdominance. The findings from the Arabidopsis study (VUYLSTEKE *et al.* 2005) fall between these extremes. Liver tissue is composed largely of tetraploid cells, which may explain the low levels of nonadditivity observed in this study. Another potential cause could be the difference in genetic divergence between the parental inbred strains used in different studies. It has been suggested that in crosses between more distantly related strains there should be greater differences in expression and that the expression of the F₁ hybrid will diverge more from the midparent values (ENARD *et al.* 2002; LANDRY *et al.* 2005; LEMOS *et al.* 2005). The two inbred mouse strains are relatively close (BECK *et al.* 2000) while the two *Drosophila* inbred strains may be more distant. The larger number of genes with additive effects compared with those with dominance effects in our study suggests that most of the difference of gene expression between the two parental strains may derive from *cis*-regulatory variation. This is consistent with the finding that most of the expression QTL (eQTL) obtained in expression

mapping studies in mouse are *cis*-eQTL (SCHADT *et al.* 2003; BYSTRYKH *et al.* 2005; CHESLER *et al.* 2005).

After significant gene lists are obtained, it is common to run some classification or gene category analysis to examine which biological processes or pathways are overrepresented in the significant gene lists. If we compare the results from these similar studies, it is easy to identify some common terms among these studies. For example, "metabolism," "signaling," and "transportation" appear to be overrepresented in the significant gene lists for nonadditive inheritance pattern in multiple studies even though the studies differ in many aspects as discussed above, which indicates that there are some common processes or pathways showing dominance in gene expression inheritance. These results are consistent with the finding that genes involved in the metabolism and stress response are more likely to show nonadditive inheritance due to their role in responding to environmental stress (KRISTENSEN *et al.* 2005). However, it should be pointed out that these terms are abundant in the array annotations. The overrepresentation of these terms in different studies may not be the direct result of underline biology.

Heritability of gene expression: In classical quantitative genetics, phenotypic variation is decomposed into genetic and environmental variances (FALCONER 1986). For inbred lines, the environmental variance can be estimated using the within-strain variation among genetically identical individuals. In this study we used biological replicates of inbred lines and repeated (technical) measurements to further decompose the environmental variance of gene expression into a biological component and a measurement component. Genetic and environmental variation was estimated from the variation between and within strains, respectively. The broad-sense heritability was estimated for the expression of each transcript. About one-quarter of the transcripts have heritability <1% but 18% of transcripts show heritability >50% and the median heritability of all genes is 20%. This heritability is substantially higher than the 11% found in the mouse brain gene expression from mouse RI mapping lines (CHESLER *et al.* 2005). The application of shrinkage to improve variance component estimation in our study improved the overall heritability and resulted in more transcripts with high heritability. The median heritability increased from 16 to 20% with shrinkage. If we examine only the genes that have heritability >1%, the median heritability after shrinkage is 31%, which is similar to the 0.34 estimated from human family cell lines (MONKS *et al.* 2004).

The authors are especially indebted to Sandy Daigle for help with RNA isolation and to the Gene Expression Service in the Microchemistry Department at The Jackson Laboratory for help with RNA quality and microarray analyses. We thank Manhong Dai at University of Michigan for help with custom CDF files. The Jackson Laboratory is American Association for Laboratory Animal Science accredited and all animal procedures were approved by The Jackson Laboratory

Animal Care and Use Committee. This study was supported by National Institutes of Health grants HL55001 (G.A.C.) and HL072241 (Jurgen Naggert) and by the National Cancer Institute core grant CA34196 (The Jackson Laboratory).

LITERATURE CITED

- ALLISON, D. B., G. L. GADBURY, M. S. HEO, J. R. FERNANDEZ, C. K. LEE *et al.*, 2002 A mixture model approach for the analysis of microarray gene expression data. *Comput. Stat. Data Anal.* **39**: 1–20.
- AUGER, D. L., A. D. GRAY, T. S. REAM, A. KATO, E. H. COE *et al.*, 2005 Nonadditive gene expression in diploid and triploid hybrids of maize. *Genetics* **169**: 389–397.
- BECK, J. A., S. LLOYD, M. HAFEZPARAST, M. LENNON-PIERCE, J. T. EPPIG *et al.*, 2000 Genealogies of mouse inbred strains. *Nat. Genet.* **24**: 23–25.
- BENJAMINI, Y., and Y. HOCHBERG, 2000 On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J. Educ. Behav. Stat.* **25**: 60–83.
- BIRCHLER, J. A., D. L. AUGER and N. C. RIDDLE, 2003 In search of the molecular basis of heterosis. *Plant Cell* **15**: 2236–2239.
- BREM, R. B., G. YVERT, R. CLINTON and L. KRUGLYAK, 2002 Genetic dissection of transcriptional regulation in budding yeast. *Science* **296**: 752–755.
- BROWN, P. O., and D. BOTSTEIN, 1999 Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21**: 33–37.
- BYSTRYKH, L., E. WEERSING, B. DONTJE, S. SUTTON, M. T. PLETCHER *et al.*, 2005 Uncovering regulatory pathways that affect hematopoietic stem cell function using 'genetical genomics.' *Nat. Genet.* **37**: 225–232.
- CHESLER, E. J., L. LU, S. SHOU, Y. QU, J. GU *et al.*, 2005 Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat. Genet.* **37**: 233–242.
- CHOUHDURI, S., 2004 Microarrays in biology and medicine. *J. Biochem. Mol. Toxicol.* **18**: 171–179.
- CHURCHILL, G. A., 2002 Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.* **32**(Suppl. 2): 490–495.
- CROW, J. F., 1998 90 years ago: the beginning of hybrid maize. *Genetics* **148**: 923–928.
- CUI, X., J. T. HWANG, J. QIU, N. J. BLADES and G. A. CHURCHILL, 2005 Improved statistical tests for differential gene expression by shrinking variance components estimates. *Biostatistics* **6**: 59–75.
- DARVASI, A., 2003 Genomics: gene expression meets genetics. *Nature* **422**: 269–270.
- ENARD, W., P. KHAITOVICH, J. KLOSE, S. ZOLLNER, F. HEISSIG *et al.*, 2002 Intra- and interspecific variation in primate gene expression patterns. *Science* **296**: 340–343.
- FALCONER, D. S., 1986 *Introduction to Quantitative Genetics*. Longman Scientific & Technical, Harlow, UK.
- GAUTIER, L., L. COPE, B. M. BOLSTAD and R. A. IRIZARRY, 2004 Affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**: 307–315.
- GIBSON, G., and B. WEIR, 2005 The quantitative genetics of transcription. *Trends Genet.* **21**: 616–623.
- GIBSON, G., R. RILEY-BERGER, L. HARSHMAN, A. KOPP, S. VACHA *et al.*, 2004 Extensive sex-specific nonadditivity of gene expression in *Drosophila melanogaster*. *Genetics* **167**: 1791–1799.
- HARTMAN, J. L., B. GARVIK and L. HARTWELL, 2001 Cell biology—principles for the buffering of genetic variation. *Science* **291**: 1001–1004.
- HARTWELL, L., 2004 GENETICS: robust interactions. *Science* **303**: 774–775.
- HITZEMANN, R., B. MALMANGER, C. REED, M. LAWLER, B. HITZEMANN *et al.*, 2003 A strategy for the integration of QTL, gene expression, and sequence analyses. *Mamm. Genome* **14**: 733–747.
- HOSACK, D. A., G. DENNIS, JR., B. T. SHERMAN, H. C. LANE and R. A. LEMPICKI, 2003 Identifying biological themes within lists of genes with EASE. *Genome Biol.* **4**: R70.
- HUBNER, N., C. A. WALLACE, H. ZIMDAHL, E. PETRETTO, H. SCHULZ *et al.*, 2005 Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat. Genet.* **37**: 243–253.
- IRIZARRY, R. A., B. M. BOLSTAD, F. COLLIN, L. M. COPE, B. HOBBS *et al.*, 2003 Summaries of affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**: e15.
- JOHNSON, W., 1980 Metabolic interference and the + – heterozygote, a hypothetical form of simple inheritance which is neither dominant nor recessive. *Am. J. Hum. Genet.* **32**: 374–386.
- KENNEDY, P. E., and B. S. CADE, 1996 Randomization tests for multiple regression. *Commun. Stat. Simul.* **25**: 923–936.
- KERR, M. K., and G. A. CHURCHILL, 2001 Bootstrapping cluster analysis: assessing the reliability of conclusions from microarray experiments. *Proc. Natl. Acad. Sci. USA* **98**: 8961–8965.
- KRISTENSEN, T. N., P. SORENSEN, M. KRUIHOFFER, K. S. PEDERSEN and V. LOESCHKE, 2005 Genomewide analysis on inbreeding effects on gene expression in *Drosophila melanogaster*. *Genetics* **171**: 157–167.
- LANDRY, C. R., P. J. WITTKOPP, C. H. TAUBES, J. M. RANZ, A. G. CLARK *et al.*, 2005 Compensatory cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*. *Genetics* **171**: 1813–1822.
- LEMONS, B., C. D. MEIKLEJOHN, M. CACERES and D. L. HARTL, 2005 Rates of divergence in gene expression profiles of primates, mice, and flies: stabilizing selection and variability among functional categories. *Evol. Int. J. Org. Evol.* **59**: 126–137.
- MARGUSH, T., and F. R. McMORRIS, 1981 Consensus n-trees. *Bull. Math. Biol.* **43**: 239.
- MONKS, S. A., A. LEONARDSON, H. ZHU, P. CUNDIFF, P. PIETRUSIAK *et al.*, 2004 Genetic inheritance of gene expression in human cell lines. *Am. J. Hum. Genet.* **75**: 1094–1105.
- MORLEY, M., C. M. MOLONY, T. M. WEBER, J. L. DEVLIN, K. G. EWENS *et al.*, 2004 Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**: 743–747.
- PAN, H. J., P. REIFSNYDER, D. E. VANCE, Q. XIAO and E. H. LEITER, 2005 Pharmacogenetic analysis of rosiglitazone-induced hepatosteatosis in new mouse models of type 2 diabetes. *Diabetes* **54**: 1854–1862.
- PARMIGIANI, G., E. S. GARRETT, R. A. IRIZARRY and S. L. ZEGER, 2003 *The Analysis of Gene Expression Data—Methods and Software*. Springer, Berlin/Heidelberg, Germany/New York.
- SCHADT, E. E., S. A. MONKS, T. A. DRAKE, A. J. LUSIS, N. CHE *et al.*, 2003 Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**: 297–302.
- SEARL, S. R., G. CASELLA and C. E. MCCULLOCH, 1992 *Variance Components*. John Wiley & Sons, New York.
- STOREY, J. D., 2002 A direct approach to false discovery rates. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**: 479–498.
- STOREY, J. D., 2003 The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann. Stat.* **31**: 2013–2035.
- VUYLSTEKE, M., F. VAN EEUWIJK, P. VAN HUMMELEN, M. KUIPER and M. ZABEAU, 2005 Genetic analysis of variation in gene expression in *Arabidopsis thaliana*. *Genetics* **171**: 1267–1275.
- WANG, J., L. TIAN, H. S. LEE, N. E. WEI, H. JIANG *et al.*, 2006 Genome-wide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* **172**: 507–517.
- WITKOVSKY, V., 2002 MATLAB algorithm mixed.m for solving Henderson's mixed model equations. Technical report (<http://www.mathpreprints.com>).
- WU, H., M. K. KERR, X. CUI and G. A. CHURCHILL, 2003 MAANOVA, a software package for the analysis of spotted cDNA microarray experiments, pp. 313–341 in *The Analysis of Gene Expression Data: An Overview of Methods and Software*, edited by G. PARMIGIANI, E. S. GARRETT, R. A. IRIZARRY and S. L. ZEGER. Springer, New York.
- XIE, Y., W. PAN and A. B. KHODURSKY, 2005 A note on using permutation-based false discovery rate estimates to compare different analysis methods for microarray data. *Bioinformatics* **21**: 4280–4288.