Microarray Expression Profiling Identifies Genes with Altered Expression in HDL-Deficient Mice

Matthew J. Callow,[1,4] Sandrine Dudoit,[2] Elaine L. Gong,[1] Terence P. Speed,[3] and Edward M. Rubin[1]

[1]*Genome Sciences Department, Lawrence Berkeley National Laboratory, Berkeley, California, 94720, USA;* [2]*Department of Biochemistry, Stanford University, Stanford, California 94305, USA;* [3]*Department of Statistics and Program in Biostatistics, University of California, Berkeley, California 94720-3860, USA*

Based on the assumption that severe alterations in the expression of genes known to be involved in high-density lipoprotein (HDL) metabolism may affect the expression of other genes, we screened an array of >5000 mouse expressed sequence tags for altered gene expression in the livers of two lines of mice with dramatic decreases in HDL plasma concentrations. Labeled cDNA from livers of apolipoprotein AI (apoAI)-knockout mice, scavenger receptor BI (SR-BI) transgenic mice, and control mice were cohybridized to microarrays. Two-sample *t* statistics were used to identify genes with altered expression levels in the knockout or transgenic mice compared with control mice. In the SR-BI group we found nine array elements representing at least five genes that were significantly altered on the basis of an adjusted *P* value < 0.05. In the apoAI-knockout group, eight array elements representing four genes were altered compared with the control group (adjusted *P* < 0.05). Several of the genes identified in the SR-BI transgenic suggest altered sterol metabolism and oxidative processes. These studies illustrate the use of multiple-testing methods for the identification of genes with altered expression in replicated microarray experiments.

Although a strong inverse relationship between high density lipoprotein (HDL) cholesterol levels and atherosclerosis susceptibility has been known for ~50 years, the underlying mechanisms of HDL function have been poorly defined. Significant recent progress has been made in identifying important factors involved in the metabolism of HDL in vivo through the use of transgenic and gene knockout mice. Apolipoprotein AI (apoAI) and scavenger receptor BI (SR-BI) are pivotal in HDL metabolism and have been functionally characterized through the use of gene manipulation techniques. The apoAI knockout (Williamson et al. 1992; Plump et al. 1996) and the SR-BI transgenic mice (Wang et al. 1998; Ueda et al. 1999) are two mouse models with extremely low HDL cholesterol levels, but each of these proteins acts via different mechanisms to affect HDL cholesterol delivery to the liver.

SR-BI, representing the first HDL receptor to be well-defined at a molecular level, expresses mostly in the adrenals and liver and has been shown to participate in cholesterol uptake (Acton et al. 1996) as well as in cholesterol efflux (Ji et al. 1997). Transgene and adenovirus-mediated overexpression of SR-BI in mice results in a virtual disappearance of plasma HDL (Kozarsky et al. 1997; Wang et al. 1998; Ueda et al. 1999), whereas SR-BI knockout mice have increased HDL cholesterol, which suggests that SR-BI may affect plasma HDL cholesterol concentrations by promoting hepatic cholesterol uptake. Despite the reported phenotypes that arise from altered hepatic expression of SR-BI, no studies have revealed the effects that increased expression of this gene may have on the expression of other hepatic genes.

Mice with reduced apoAI plasma levels caused by the inactivation of the murine apoAI gene have also been produced and studied (Williamson et al. 1992; Plump et al. 1997). Animals that are homozygous for the inactivated apoAI gene experience a reduction in total plasma cholesterol by ~70% when compared with controls. Very little is known about the effects of apoAI deficiency on the expression of other genes in the liver; however, one study showed altered cholesterol 7α-hydroxylase expression (Plump et al. 1997) and others have shown that apoAI deficiency can affect SR-BI-protein levels in the adrenals (Sun et al. 1999).

Microarray technology has been used in the past for the analysis of expression changes in single-celled organisms (Shalon et al. 1996; DeRisi et al. 1997; Lashkari et al. 1997; Chu et al. 1998), mammalian cell cultures (DeRisi et al. 1996; Schena et al. 1996; Iyer et al. 1999), and human and mouse tissue (Alon et al. 1999; Perou et al. 1999; Friddle et al. 2000) with success in identifying groups of correlated genes. This technology has thus far not been applied extensively to the identification of differential gene expression induced by single-gene effects in complex tissues, such as the

[4]Corresponding author.
E-MAIL mjcallow@lbl.gov; FAX (510) 486-4229.
Article and publication are at www.genome.org/cgi/doi/10.1101/gr.147200.

10:2022–2029 ©2000 by Cold Spring Harbor Laboratory Press ISSN 1088-9051/00 $5.00; www.genome.org

mammalian liver, although recently McNeish et al. (2000) reported several genes with altered expression levels in the livers of ABC-1 knockout mice using oligonucleotide arrays. In this study, to identify in an unbiased manner genes that respond to alterations in hepatic HDL delivery, we have performed microarray expression profiling of liver transcripts from apoAI knockout and SR-BI transgenic mice. Our analysis provides a methodological framework for studies that investigate differential gene expression through the use of microarrays. From the analysis of our data, we identified several known and novel expressed sequences in each HDL modified group to be altered. Of particular note, genes involved in oxidative processes and sterol metabolism were altered in expression in the livers of SR-BI transgenic mice.

RESULTS

Experimental Outline

To identify genes that were altered in each of our two test cases, we first isolated RNA from the livers of apoAI-knockout mice, SR-BI transgenic mice, C57Bl/6 mice, and FVB mice (eight mice per group). The RNA from each individual control and test mouse was labeled with the Cy5 dye. For the Cy3 reference we prepared a pool of RNA from either the C57Bl/6 mice (apoAI-knockout experiments) or the FVB mice (SR-BI transgenic experiments).

Ranking Array Elements: Using *t* Statistics as a Measure of Differential Expression

To determine those elements that were altered in the transgenic or knockout group we calculated *t* statistics comparing gene expression in the genetically modified group with the appropriate control group. For each gene, the numerator of the *t* statistic estimates the difference in the mean expression ratio of the two groups and the denominator provides an estimate of the precision of this difference. Figure 1 displays a normal quantile–quantile (q–q) plot of our 5600 observed *t* statistics, that is, plots of the ranked *t* statistics against the corresponding quantiles of the standard normal distribution. The q–q plot for SR-BI shows a gradual deviation from the line expected with a large number of replicates, under the null hypothesis of equal expression for all genes. Nine elements deviating from the line had absolute *t* statistic values greater than six, with seven elements having positive *t* scores (indicating overexpression) and two elements having negative *t* scores (indicating underexpression). Three of the elements represented the β-globin gene and two represented glutathione-*S*-transferase (GST).

In the apoAI-knockout study, it can be seen from Figure 1B that eight array elements with *t* statistics greater than an absolute value of six deviated markedly

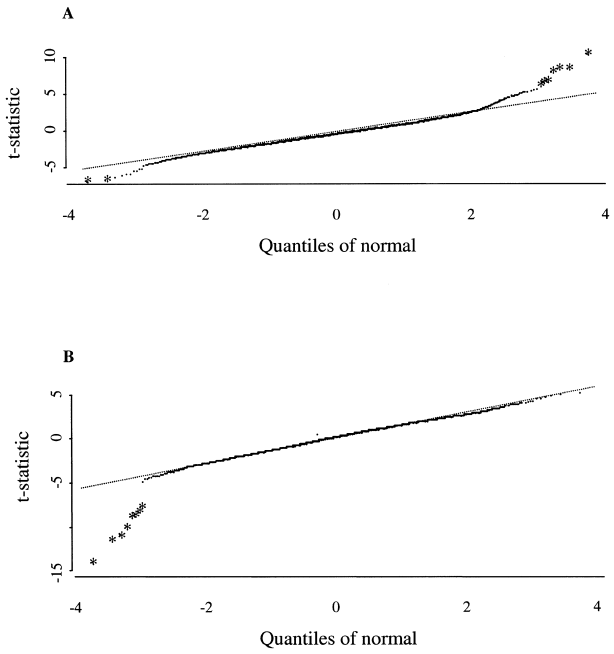

Figure 1 Normal quantile–quantile plot of *t* statistics calculated from the normalized log ratios of the SR-B1 transgenic group (*A*) and the apoAI-knockout group (*B*). Array elements with adjusted *P* values <0.05 are displayed with an asterisk.

from the expected line under the null hypothesis. These eight elements represent four genes: apoAI, apoCIII, sterol C5 desaturase, and one novel expressed sequence tag (EST) AA080005. It was not apparent from this display of the data that any genes were increased in expression in the apoAI group, in contrast with the finding in the SR-BI transgenic group.

In using any measure of difference between the transgenic/knockout and control groups it is necessary to consider that thousands of comparisons are being performed and so the possibility of obtaining extreme *t* statistic values by chance increases. The normal q–q plots provide a useful visual analysis of the data to reveal whether the *t* statistics we observe are likely to represent real differences between the two groups under comparison. The points that are well off the line are likely to correspond to those genes whose expression levels are different between the two groups, which informally corrects for the large number of test statistics. The points on the line are displaying chance variation.

Using *t* Statistics Avoids Filtering for Low-Intensity Elements

One question that arises is whether or not low-intensity spots should be removed because of the unreliability of the measurements. The relationship between *t* statistics and overall spot intensity is displayed in Figure 2. In the analysis of the apoAI-knockout mice, the eight elements shifted in the normal q–q plot were

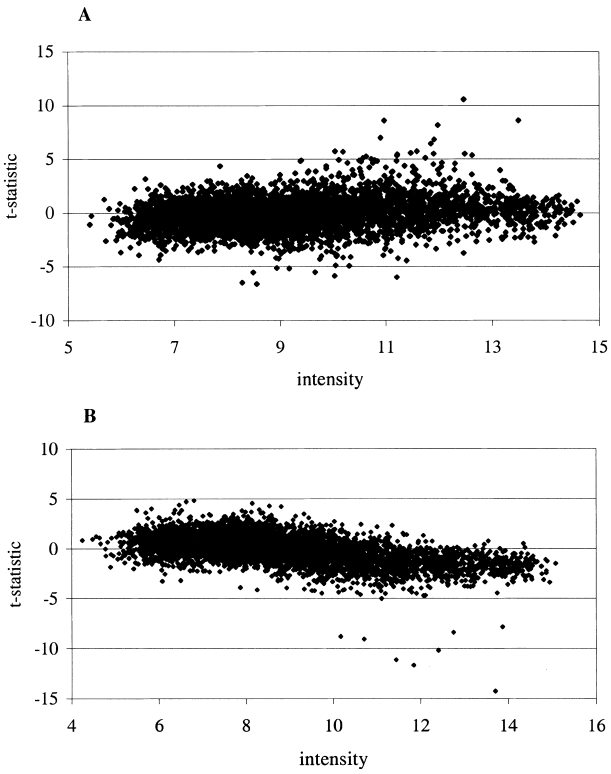

Figure 2 *t* statistics for 5600 array elements plotted versus intensity. (*A*) SR-B1 transgenic group, (*B*) apoAI-knockout group. Intensity values and *t* statistics were determined as described in methods.

again clearly separated from the other elements. All eight elements appeared in the top 50% of intensities. In contrast, the relationship of intensity to *t* statistic in the SR-BI transgenic group did not show such a clear demarcation. It is interesting to note that two of the underexpressed genes with the lowest (negative) *t* statistics were also observed to fall within the lower half of the intensity plot.

The analysis of the relationship of intensity of the elements to *t* statistic shows that filtering for intensity is not necessary and that elements must exhibit larger differences in the means of the ratios (numerator of the *t* statistic) to overcome increased variability (denominator of the *t* statistic) if they are to achieve large absolute *t* statistics.

Genes that Change in SR-BI Transgenic

Adjusted *P* values were computed to obtain a more precise assessment of the statistical significance of the results and to account for multiple comparisons. Table 1A lists those genes in the SR-BI transgenic analysis with the largest (absolute value) *t* statistics and their adjusted *P* values. After elimination of replicates (several array elements representing the same gene) we identified five genes with an adjusted *P* value <0.05. Three of these genes showed increased expression and two showed decreased expression. An additional array element was identified with an adjusted *P* value of 0.0002; however, its identity could not be confirmed because of a mixed population of ESTs. We also identified a further 11 elements that represent seven genes that had adjusted *P* values ≤0.2. As expected, an increase in levels of SR-BI message was observed in the transgenic mice relative to the FVB control mice. Other genes that were overexpressed included GST and β-globin. Cytochrome P450 2B10, also known as testosterone 16-α-hydroxylase, was suppressed and an EST (AI746730) with similarity to rat peroxisomal long chain α-hydroxy acid oxidase, was also decreased in expression ratio in the SR-BI relative to the control FVB mice.

We used real-time quantitative PCR to confirm many of the changes that were identified in Table 1. Primers were designed to the expressed sequences for Cytochrome P450 2B10, β-globin, GST and the EST AI746730 and were used to determine the relative differences in the specific target to 18s RNA levels in the total RNA sample. Figure 3 shows that GST and β-globin changes were comparable in size when determined by quantitative PCR or microarray hybridization but microarray hybridization tended to underrepresent the change relative to quantitative PCR. Changes of 10-fold (Cyp2B10) to 70-fold (AI746730) were observed using quantitative PCR, whereas expression changes by microarray hybridization were approximately fourfold for both Cyp2B10 and AI746730.

Genes that Change in apoAI-Knockout Mice

For the apoAI-knockout experiment, after elimination of replicates, we identified four genes with an adjusted *P* value <0.05. As expected, the largest ratios were obtained for apoAI with an apparent 20-fold decrease in levels of RNA.

ApoCIII was also underexpressed in the knockout group compared with the control group. The ApoCIII gene is located at the apoAI locus, so the decrease in expression may represent either a decrease in apoCIII expression arising from lack of apoAI or genetic variation contributing to the altered expression of apoCIII. Even though the apoAI-knockout mice had been bred to the C57Bl/6 line for 10 generations, the apoAI/apoCIII locus would be of the 129 genetic background—the genetic background of the embryonic stem cells used for the production of the knockout. To address this issue we performed quantitative RT-PCR analysis of RNA collected from the livers of 129P1/ReJ mice and compared the relative expression of apoCIII in these mice to the C57Bl/6 mice. Figure 4 shows the differences in apoCIII RNA levels by quantitative PCR of the C57Bl/6 strain, apoAI knockout, and the 129P1/ReJ strain. Levels of apoCIII RNA were fivefold less in the 129 strain compared with the C57Bl/6 strain and threefold less in the apoAI knockout, which suggests

Table 1A. Genes Identified from the Microarrays with (*A*) Adjusted *P* Values <0.05 or (*B*) with Adjusted *P* Values <0.2 but >0.05

Gene	Accession no.	Fold change	Adjusted *P* value
SR-B1 transgenic:			
SR-B1	AI323419	+11	0.0036
Glutathione-*S*-transferase	AA462433	+2.5	0.0023
β-globin	AB020015	+1.7	0.0177
Cytochrome P450 2B10	AI323886	−5.0	0.0319
EST AI746730	AI746730	−3.5	0.0407
ApoAI knockout:			
ApoAI	AA821910	−20	0.0002
ApoCIII	AI327016	−2.5	0.0002
Sterol C5 desaturase	W65233	−2.5	0.0005
EST AA080005	AA080005	−1.9	0.0017

Table 1B.

Gene	Accession no.	Fold change	Adjusted *P* value
SR-B1 transgenic:			
α-globin	M26895	+1.8	0.1543
Surfactant protein B	S78114	+1.5	0.1249
Id	M31885	+1.5	0.1265
EST AI642071	AI642071	+1.3	0.1340
Tubulin α 4	NM_009447	+1.5	0.1890
Cytochrome P450 2A4	NM_009997	−1.9	0.1739
2-oxoglutarate carrier protein	AI099441	−1.5	0.1051
ApoAI knockout:			
none			

Adjusted *P* values were calculated as described in Methods. Fold change in expression was determined from the antilog of the difference between the average of the log ratios of the control group and the knockout or transgenic group; +, increased expression; −, decreased expression in the genetically modified group. The array element with the lowest *P* value is represented when multiple elements were identified for a gene.

that the decrease in apoCIII observed between the C57Bl/6 strain and the apoAI knockout was caused by genetic polymorphism rather than to apoAI deficiency.

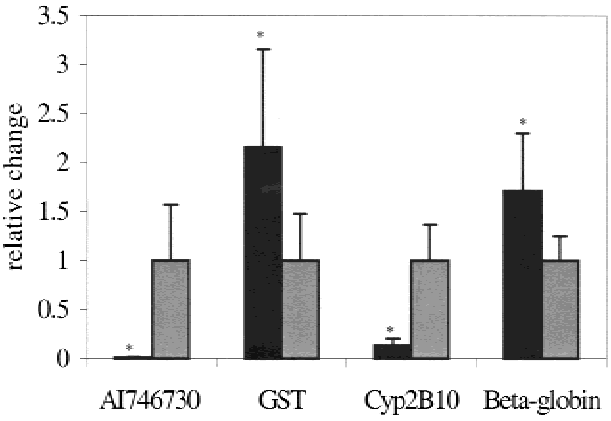

Figure 3 Quantitative RT-PCR of liver mRNA for EST AI746730, glutathione*S*-transferase (GST), Cytochrome P4502B10 (cyp2B10) and β-globin. cDNA was synthesized using random primers and amplified with gene-specific primers or primers to 18s RNA. Results are expressed relative to expression in the FVB control (adjusted to a value of 1) after normalization for 18s content of the sample. (* *P* value <0.05 by two-tailed *t* test). Solid bars, SR-BI transgenic; shaded bars, FVB control.

We also identified sterol C5 desaturase to be decreased in the apoAI knockout by 2.5-fold. Sterol C5 desaturase catalyzes one of the terminal steps in cholesterol synthesis by converting Δ-7-cholestenol to dehydrocholesterol (Taton et al. 2000). A novel EST (AA080005) that shares similarity to a family of

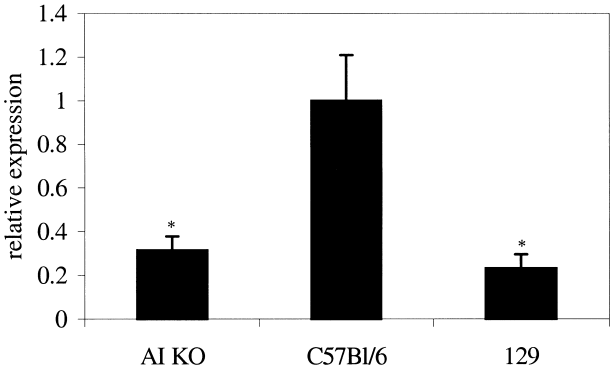

Figure 4 Quantitative RT-PCR for apolipoprotein CIII (apoCIII). cDNA was synthesized using random primers and amplified with apoCIII-specific primers or primers to 18s RNA. Results are expressed relative to C57Bl/6 (adjusted to a value of 1) after normalization for 18s content of the sample. (* *P* value <0.05 by two-tailed *t* test).

ATPases was decreased in expression relative to the C57Bl/6 control by less than twofold.

DISCUSSION

The genetically modified mouse models presented in this study are characterized by diminished HDL levels. In one case (apoAI knockout), decreased production of HDL results in decreased HDL levels, whereas in the other (SR-BI transgenic), increased clearance of HDL is the cause of reduced steady-state plasma HDL concentrations. We have performed expression profiling and data analysis to enhance the possibility of detecting subtle yet potentially significant changes in gene expression in two models of reduced HDL. This was achieved by comparing arrays in which cDNA from both control mice and mutant mice were cohybridized with the same pool of control reference RNA, followed by a determination of *t* statistics for each element on the array.

One of the most prominent changes in the SR-BI transgenic liver was the decrease in RNA levels for Cytochrome P450 2B10 and, to a lesser extent, Cytochrome P450 2A4. Both of these enzymes participate in the catabolism of testosterone (Honkakoski et al. 1992), which suggests that steroid catabolism may be altered in the livers of SR-BI transgenic mice. Members of the Cytochrome P450 2B family have also been shown to be suppressed by intracellular sterols (Kocarek et al. 1998), which suggests that the decrease we observed in Cytochrome P450 2B10 might represent a response to increased importation of sterols in the livers of the overexpressing SR-BI transgenics. The reduced expression of the cytochromes noted is consistent with there being a feedback loop in response to the increased removal of circulatory steroid hormones and sterols caused by the overexpression of SR-BI.

A novel EST (AI746730), sharing sequence similarity to rat peroxisomal long chain α-hydroxy-acid oxidase, was significantly altered in the SR-BI transgenic livers. Although no information is available about the function of this gene, its sequence homology suggests that it may have similar properties to the hydroxy-acid-oxidase enzyme in its oxidation of substrates such as long-chain fatty acids. Consistent with this is the finding that GST, an enzyme induced by lipid peroxides (Singhal et al. 1992; Khan et al. 1995; Tjalkens et al. 1998), was increased in the SR-BI transgenics. Furthermore, the oxoglutarate carrier protein, which facilitates the transport of glutathione into the mitochondria (Chen & Lash 1998), was also observed to be decreased in the SR-BI transgenic (adjusted $0.05 < P < 0.2$). Overall, this pattern of gene expression in the SR-BI transgenic suggests an imbalance in the delivery and/or formation of oxidative products within the livers of the transgenic mice that overexpress SR-BI.

The effect of SR-BI overexpression on globin expression was surprising and it is difficult to reconcile this observation with what is known about the regulation of globin gene expression. The adult liver is not believed to be a site of active erythrogenesis; however, ESTs for β-globin have been identified in cDNA libraries of diverse nonerythrogenic tissues, including liver. Although SR-B1 has been shown to be an important lipoprotein receptor, it is believed to have other properties including the binding of phosphatidylserine, phosphatidylinositol (Rigotti et al. 1995; Fukasawa et al. 1996), and apoptotic thymocytes (Imachi et al. 2000). The closely related CD36 receptor has been demonstrated to bind *Plasmodium falciparum*-infected erythrocytes (Oquendo et al. 1989) and oxidatively altered erythrocytes appear to bind to an as yet unidentified scavenger receptor of liver cells (Terpstra & van Berkel 2000). Apoptotic or senescent erythrocytes are also believed to become enriched in phosphatidylserine in the outer leaflet of the plasma membrane leading to phagocytosis (Kiefer & Snyder 2000). The modest increases in β-globin and α-globin expression that we observed may reflect an increase in blood-derived cells that express high levels of globin accumulating in the liver. SR-BI may be sequestering these cells by virtue of its affinity for phosphatidylserine.

Previous studies have reported that plasma lipoproteins are diminished in the SR-B1 transgenic mice (Wang 1998; Veda 1999), which suggests increased uptake of cholesterol by the liver. It may therefore be expected that some genes involved in cholesterol homeostasis, such as the low density lipoprotein (LDL) receptor or HMG CoA reductase, would have decreased liver expression. There are several possibilities that could explain why changes in expression for these genes were not observed. First, if liver expression is low under normal conditions, any further decrease induced by increased cholesterol uptake, may result in decreased fluorescent intensities and greater ratio variability. Such data points have a greater chance of being excluded in the latter data-analysis stages as they would have relatively small *t* statistics and relatively large adjusted *P* values. We have used adjusted *P* values to decrease the likelihood of classifying false-positive genes as changed when in fact their large *t* statistics may arise by chance as in any study dealing with a large number of comparisons. One of the difficulties in looking at a measure of change in thousands of genes is that a true positive (i.e., a gene that is genuinely differentially expressed) may be lost (i.e., be a false negative) among the variability of scores or even hundreds of other genes that are not changing (i.e., true negatives). A second possibility why "expected" genes do not appear is increased output rates of cholesterol from the liver in the form of bile acids or de novo synthesized plasma lipoproteins. These processes may be sufficiently adaptable so as not to require dramatic

shifts in transcriptional response to the altered cholesterol flux induced by SR-BI overexpression and may also explain why a greater number of genes were not observed to be altered in these experimental models.

The decreased apoCIII expression seen in the apoAI-knockout mice has implications for studies using the knockout allele in a strain other than the 129 mouse strain in which the gene targeting was initially performed. Our results suggest that the decreased expression of apoCIII observed in the apoAI-knockout mice was caused by the 129-apoCIII allele in the C57Bl/6 background, rather than by the decrease in apoAI expression and plasma HDL cholesterol in these animals. This conclusion is supported by the fact that apoCIII and apoAI map to within 4 kb of each other and by our demonstration that 129 mice normally express apoCIII at approximately fivefold lower levels than C57Bl/6 mice, similar to that observed with the presence of the apoAI-knockout allele in the mostly C57Bl/6 background. A similar mechanism may explain the decreased expression of the sterol C5 desaturase gene in the apoAI-knockout mice. Although the chromosomal location of sterol C5 desaturase in the mouse is unknown at this time, in humans it has been mapped to the same chromosomal region as apoAI. The identification of polymorphisms at the gene-expression level in different mouse strains segregating with targeted loci suggests a novel application for microarrays and highlights the importance of genetic backgrounds when undertaking expression profiling studies.

The use of microarray expression profiling is gaining popularity through the general availability of ESTs and commercial availability of microarray printers and scanners. We have described a strategy that addresses the issues of detecting small but potentially significant changes in expression using microarrays and report a statistically simple approach to data analysis comparing two conditions. In our analysis we have also addressed the issue of multiple testing by calculating adjusted *P* values for comparisons of control and genetically modified mice. This analysis has yielded a number of genes with reproducible alterations in expression. These genes have drawn attention to cellular pathways and processes that warrant further investigation in studies of HDL metabolism.

METHODS

cDNAs and ESTs Selected for Printing

We selected 5600 ESTs from the I.M.A.G.E. consortium database (Lennon et al. 1996) including 257 ESTs that represent genes with possible and confirmed associations with lipid metabolism. One to two ESTs were selected for each gene and purchased from Research Genetics. All array clones identified in this report were resequenced. The vector inserts were PCR amplified using primers designed to the vector sequence surrounding the expressed sequence insert (provided by Research Genetics) and ethanol precipitated before resuspension in 15 µL of 3× SSC and 15 µL of Array-IT solution (Telechem).

Production of Microarrays

A printer was constructed based on design specifications from the Brown laboratory at Stanford University (Schena et al. 1995). The PCR-amplified products were printed from wells of a 384-well plate and deposited onto Poly-l-lysine coated slides using a 16-pin print head. Center-to-center spot spacings of ~200 microns were used.

Mice

All mice used in this study were males of 8–10 weeks of age and were killed by cervical dislocation. ApoAI-knockout mice that had been backcrossed to C57Bl/6 for 10 generations and C57Bl/6 control mice were obtained from Jackson Labs. SR-BI transgenic mice were a high-expressing line generated and maintained in the FVB inbred strain (Ueda et al. 1999). The livers were removed, snap frozen in liquid nitrogen, and the total RNA fraction was isolated using the Tri-zol reagent (Gibco BRL) following the manufacturers instructions. Cy3-dUTP or Cy5-dUTP (Amersham Pharmacia) was used to label cDNA synthesized from the oligo dT primed fraction of 60 µgrams of total RNA with Superscript II enzyme (Gibco BRL). Superscript II buffer, dithiothreitol, Cy3 or Cy5 dUTP, and the four unlabeled nucleotides were combined with the RNA and the reaction was allowed to proceed for 2–3 h before terminating by heat denaturation. In all cases the reference sample was generated by pooling equal amounts of RNA from each of the eight FVB control mice or the eight C57Bl/6 mice and the sample was then divided over eight reaction tubes for the labeling procedure with Cy3-dUTP and repooled before combining with the Cy5 labeled test sample. The test samples consisted of RNA labeled with the Cy5 dye from each individual transgenic, knockout or inbred control mouse.

The labeled cDNA mixture was concentrated on microcon-30 columns after the addition of mouse Cot1 DNA (Gibco BRL). After denaturation of the probe, the solution was added to the array and a coverslip was placed on top. The slide was then placed in a sealed, humidified hybridization chamber before being placed in a 65° C hybridization oven. After hybridization for 16 h, the slide was rinsed in 1× SSC, 0.05% SDS and then 0.2× SSC followed by a rinse in 0.1× SSC and spin dried.

Scanning of Microarrays

Arrays were scanned using a purpose-built dual-laser scanner. Data from the dual-laser scans was collected as graphics (.tiff format) files for each of the two lasers. The dual images were combined and processed using the ScanAlyze program (written by Michael Eisen, http://rana.Stanford.EDU/software/) and positions on the images were assigned grid coordinates for the matching of spot identity with average pixel intensity at that region. The background was determined from pixel intensity at the edge of the bounding box. Background corrected, log converted red/green-intensity ratios were determined for each element of the array. Variations in signal intensities of the fluorophores were normalized by subtracting the median value of log ratios from each log ratio determined for that array.

Quantitative PCR

Total RNA treated with DNAse I (Promega) was used for cDNA synthesis using random nanomers (New England Biolabs) and Superscript II enzyme according to the manufacturers instructions. Levels of cDNA were quantified using gene-specific primers and SYBR green detection in a Perkin Elmer 7700 sequence detector. Levels of 18s were quantified in the same samples using a primer/competimer pair purchased from Ambion Inc. and levels of cDNA were expressed relative to 18s content of the sample.

Statistical Analysis

For each experiment the expression log ratios were displayed in a 5600× 16 matrix with rows corresponding to array elements and columns corresponding to arrays. To test the null hypothesis H_j of equal mean expression for array element j in the control and transgenic or knockout mice, a two-sample t statistic was used

$$t_j = \frac{\bar{x}_{2j} - \bar{x}_{1j}}{\sqrt{\frac{s^2_{1j}}{n_1} + \frac{s^2_{2j}}{n_2}}}$$

where $\bar{x}_{1j}$ and $\bar{x}_{2j}$ denote the average log ratio of element j in the control and transgenic or knockout group and s^2_{1j} and s^2_{2j} denote the variances of element j's log ratios in the control and transgenic or knockout hybridizations, respectively. The number of arrays in the control and transgenic or knockout group is denoted by n_1 and n_2, respectively.

Two types of graphical presentations were used. In the normal q–q plot, the 5600 observed t statistics were plotted against standard normal quantiles (Venables and Ripley 1999). In addition, the t statistics for the 5600 cDNAs were plotted against a measure of average total intensity corresponding to each cDNA, this measure being the average over all 16 mice of the average of the logarithms (base 2) of the red and green intensities.

Assessing the strength of the evidence against the null hypotheses of equal expression in the control and transgenic or knockout mice is typically done by calculating P values for each hypothesis; that is, by calculating for each gene the chance of getting a t statistic as extreme, or more extreme, than the observed statistic under the null hypothesis. However, with a typical microarray data set comprising thousands of genes, an immediate concern is multiple testing because the probability that at least one null hypothesis is erroneously rejected (type-I error) can increase sharply with the number of hypotheses tested. To account for multiple testing we computed adjusted P values for each gene (Westfall & Young, 1993; Shaffer 1995).

The adjusted P value corresponding to the test of a null hypothesis H_j for a single element j can be defined as the level of the entire test procedure at which H_j would just be rejected, given the values of all test statistics involved. For our experiments, issues complicating P value calculations include an unknown null distribution of the test statistics and an unknown correlation structure between the elements. A suitable permutation distribution of the test statistics, as in Algorithm 4.1 of Westfall and Young (1993), was used to deal with these problems. In this algorithm, the permutation distribution of the t statistics was obtained by permuting the columns of the data matrix. Note that we are not assuming that the t statistics follow a t distribution or even a normal distribution, rather, we use a permutation distribution to estimate the null distribution of the t statistics.

ACKNOWLEDGMENTS

We thank Greg Barsh and David Kingsley for useful discussions and help in compiling the clone set. This work was supported by NIH project grant 1RO1 HL63897–01. Research was conducted at the E.O. Lawrence Berkeley National Laboratory (Department of Energy contract DE-AC0376SF00098), University of California. S.D. was supported by a postdoctoral fellowship from the Mathematical Sciences Research Institute. T.P.S. was supported by NIH project grant 5R01MH61665–02

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Acton, S., Rigotti, A., Landschulz, K.T., Xu, S.Z., Hobbs, H.H., and Krieger, M. 1996. Identification of scavenger receptor SR-BI as a high density lipoprotein receptor. *Science* **271:** 518–520.

Alon, U., Barkai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., and Levine, A.J. 1999. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc. Natl. Acad. Sci.* **96:** 6745–6750.

Chen, Z. and Lash, L.H. 1998. Evidence for mitochondrial uptake of glutathione by dicarboxylate and 2-oxoglutarate carriers. *J. Pharmacol. Exp. Ther.* **285:** 608–618.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O., and Herskowitz, I. 1998. The transcriptional program of sporulation in budding yeast. *Science* **282:** 699–705.

DeRisi, J.L., Iyer, V.R., and Brown, P.O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278:** 680–686.

DeRisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A., and Trent, J.M. 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nature Genetics* **14:** 457–460.

Friddle, C.J., Koga, T., Rubin, E.M., and Bristow, J. 2000. Expression profiling reveals distinct sets of genes altered during induction and regression of cardiac hypertrophy. *Proc. Natl. Acad. Sci.* **97:** 6745–6750.

Fukasawa, M., Adachi, H., Hirota, K., Tsujimoto, M., Arai, H., and Inoue, K. 1996. SRB1, a class B scavenger receptor, recognizes both negatively charged liposomes and apoptotic cells. *Exp. Cell. Res.* **222:** 246–250.

Honkakoski, P., Kojo, A., and Lang, M.A. 1992. Regulation of the mouse liver cytochrome P450 2B subfamily by sex hormones and phenobarbital. *Biochem. J.* **285:** 979–983.

Imachi, H., Murao, K., Hiramine, C., Sayo, Y., Sato, M., Hosokawa, H., Ishida, T., Kodama, T., Quehenberger, O., Steinberg, D., et al. 2000. Human scavenger receptor B1 is involved in recognition of apoptotic thymocytes by thymic nurse cells. *Lab. Invest.* **80:** 263–270.

Iyer, V.R., Eisen, M.B., Ross, D.T., Schuler, G., Moore, T., Lee, J.C.F., Trent, J.M., Staudt, L.M., Hudson, J., Jr., Boguski, M.S.,et al. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science* **283:** 83–87.

Ji, Y., Jian, B., Wang, N., Sun, Y., Moya, M.L., Phillips, M.C., Rothblat, G.H., Swaney, J.B., and Tall, A.R. 1997. Scavenger receptor BI promotes high density lipoprotein-mediated cellular cholesterol efflux. *J. Biol. Chem.* **272:** 20982–20985.

Khan, M.F., Srivastava, S.K., Singhal, S.S., Chaubey, M., Awasthi, S., Petersen, D.R., Ansari, G.A., and Awasthi, Y.C. 1995. Iron-induced lipid peroxidation in rat liver is accompanied by preferential induction of glutathione S-transferase 8–8 isozyme. *Toxicol. Appl. Pharmacol.* **131:** 63–72.

Kiefer, C.R. and Snyder, L.M. 2000. Oxidation and erythrocyte

senescence. *Curr. Opin. Hematol.* **7:** 113–116.
Kocarek, T.A., Kraniak, J.M., and Reddy, A.B. 1998. Regulation of rat hepatic cytochrome P450 expression by sterol biosynthesis inhibition: Inhibitors of squalene synthase are potent inducers of CYP2B expression in primary cultured rat hepatocytes and rat liver. *Mol. Pharmacol.* **54:** 474–484.
Kozarsky, K.F., Donahee, M.H., Rigotti, A., Iqbal, S.N., Edelman, E.R., and Krieger, M. 1997. Overexpression of the HDL receptor SR-BI alters plasma HDL and bile cholesterol levels. *Nature* **387:** 414–417.
Lashkari, D.A., DeRisi, J.L., McCusker, J.H., Namath, A.F., Gentile, C., Hwang, S.Y., Brown, P.O., and Davis, R.W. 1997. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc. Natl. Acad. Sci.* **94:** 13057–13062.
Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M.B. 1996. The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33:** 151–152.
McNeish, J., Aiello, R.J., Guyot, D., Turi, T., Gabel, C., Aldinger, C., Hoppe, K.L., Roach, M.L., Royer, L.J., de Wet, J., et al. 2000. High density lipoprotein deficiency and foam cell accumulation in mice with targeted disruption of ATP-binding cassette transporter. *Proc. Natl. Acad. Sci.* **97:** 4245–4250.
Oquendo, P., Hundt, E., Lawler, J., and Seed, B. 1989. CD36 directly mediates cytoadherence of Plasmodium falciparum parasitized erythrocytes. *Cell* **58:** 95–101.
Perou, C.M., Jeffrey, S.S., van de Rijn, M., Rees, C.A., Eisen, M.B., Ross, D.T., Pergamenschikov, A., Williams, C.F., Zhu, S.X., Lee, J.C., et al. 1999. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci.* **96:** 9212–9217.
Plump, A.S., Azrolan, N., Odaka, H., Wu, L., Jiang, X., Tall, A., Eisenberg, S., and Breslow, J.L. 1997. ApoA-I knockout mice: characterization of HDL metabolism in homozygotes and identification of a post-RNA mechanism of apoA-I up-regulation in heterozygotes. *J. Lipid Res.* **38:** 1033–1047.
Plump, A.S., Erickson, S.K., Weng, W., Partin, J.S., Breslow, J.L., and Williams, D.L. 1996. Apolipoprotein A-I is required for cholesteryl ester accumulation in steroidogenic cells and for normal adrenal steroid production. *J. Clin. Invest.* **97:** 2660–2671.
Rigotti, A., Acton, S.L., and Krieger, M. 1995. The class B scavenger receptors SR-BI and CD36 are receptors for anionic phospholipids. *J. Biol. Chem.* **270:** 16221–16224.
Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270:** 467–470.
Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P.O., and Davis, R.W. 1996. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci.* **93:** 10614–10619.
Shaffer, J.P. 1995. Multiple hypothesis testing. *Annu. Rev. Psychol.* **46:** 561–584.
Shalon, D., Smith, S.J., and Brown, P.O. 1996. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* **6:** 639–645.
Singhal, S.S., Saxena, M., Ahmad, H., Awasthi, S., Haque, A.K., and Awasthi, Y.C. 1992. Glutathione S-transferases of human lung: characterization and evaluation of the protective role of the alpha-class isozymes against lipid peroxidation. *Arch. Biochem. Biophys.* **299:** 232–241.
Sun, Y., Wang, N., and Tall, A.R. 1999. Regulation of adrenal scavenger receptor-BI expression by ACTH and cellular cholesterol pools. *J. Lipid Res.* **40:** 1799–1805.
Taton, M., Husselstein, T., Benveniste, P., and Rahier, A. 2000. Role of highly conserved residues in the reaction catalyzed by recombinant Δ7-sterol-C5(6)-desaturase studied by site-directed mutagenesis. *Biochemistry* **39:** 701–711.
Terpstra, V. and van Berkel, T.J. 2000. Scavenger receptors on liver Kupffer cells mediate the in vivo uptake of oxidatively damaged red blood cells in mice. *Blood* **95:** 2157–2163.
Tjalkens, R.B., Valerio, L.G., Jr., Awasthi, Y.C., and Petersen, D.R. 1998. Association of glutathione S-transferase isozyme-specific induction and lipid peroxidation in two inbred strains of mice subjected to chronic dietary iron overload. *Toxicol. Appl. Pharmacol.* **151:** 174–181.
Ueda, Y., Royer, L., Gong, E., Zhang, J., Cooper, P.N., Francone, O., and Rubin, E.M. 1999. Lower plasma levels and accelerated clearance of high density lipoprotein (HDL) and non-HDL cholesterol in scavenger receptor class B type I transgenic mice. *J. Biol. Chem.* **274:** 7165–7171.
Venables, W.N. and Ripley, B.D. 1999. *Modern applied statistics with S-PLUS*. 3rd ed. Statistics and computing, Springer, New York.
Wang, N., Arai, T., Ji, Y., Rinninger, F., and Tall, A.R. 1998. Liver-specific overexpression of scavenger receptor BI decreases levels of very low density lipoprotein ApoB, low density lipoprotein ApoB, and high density lipoprotein in transgenic mice. *J. Biol. Chem.* **273:** 32920–32926.
Westfall, P.H. and Young, S.S. 1993. *Resampling-based multiple testing: Examples and methods for* p*-value adjustment*. Wiley series in probability and mathematical statistics. Applied probability and statistics, Wiley, New York.
Williamson, R., Lee, D., Hagaman, J., and Maeda, N. 1992. Marked reduction of high density lipoprotein cholesterol in mice genetically modified to lack apolipoprotein A-I. *Proc. Natl. Acad. Sci.* **89:** 7134–7138.

Received May 9, 2000; accepted in revised form September 28, 2000.