Available online at www.sciencedirect.com
SCIENCE DIRECT
Methods 38 (2006) 227–234
METHODS
www.elsevier.com/locate/ymeth

Application of microarray technology in primate behavioral neuroscience research

Adriaan M. Karssen [a,1], Jun Z. Li [b,1], Song Her [a], Paresh D. Patel [c], Fan Meng [c], Simon J. Evans [c], Marquis P. Vawter [d], Hiroaki Tomita [d], Prabhakara V. Choudary [e], William E. Bunney Jr. [d], Edward G. Jones [e], Stanley J. Watson [c], Huda Akil [c], Richard M. Myers [b], Alan F. Schatzberg [a], David M. Lyons [a,*]

[a] *Department of Psychiatry and Behavioral Sciences, Stanford University, USA*
[b] *Stanford Human Genome Center, Department of Genetics, Stanford University, USA*
[c] *Molecular and Behavioral Neuroscience Institute, Department of Psychiatry, University of Michigan, USA*
[d] *Department of Psychiatry and Human Behavior, University of California, Irvine, USA*
[e] *Center for Neuroscience, University of California, Davis, USA*

Accepted 15 September 2005

Abstract

Gene expression profiling of brain tissue samples applied to DNA microarrays promises to provide novel insights into the neurobiological bases of primate behavior. The strength of the microarray technology lies in the ability to simultaneously measure the expression levels of all genes in defined brain regions that are known to mediate behavior. The application of microarrays presents, however, various limitations and challenges for primate neuroscience research. Low RNA abundance, modest changes in gene expression, heterogeneous distribution of mRNA among cell subpopulations, and individual differences in behavior all mandate great care in the collection, processing, and analysis of brain tissue. A unique problem for nonhuman primate research is the limited availability of species-specific arrays. Arrays designed for humans are often used, but expression level differences are inevitably confounded by gene sequence differences in all cross-species array applications. Tools to deal with this problem are currently being developed. Here we review these methodological issues, and provide examples from our experiences using human arrays to examine brain tissue samples from squirrel monkeys. Until species-specific microarrays become more widely available, great caution must be taken in the assessment and interpretation of microarray data from nonhuman primates. Nevertheless, the application of human microarrays in nonhuman primate neuroscience research recovers useful information from thousands of genes, and represents an important new strategy for understanding the molecular complexity of behavior and mental health.

Keywords: Primate; Brain; Behavior; Microarray; Gene expression; Prefrontal cortex; Hippocampus

1. Introduction

Studies of the neural basis of behavior in nonhuman primates continue to play a vital role in understanding the structure and function of brain circuits that mediate emotional, cognitive, and social aspects of behavior in humans. In part, this is due to practical limitations and ethical concerns that restrict opportunities to conduct controlled experiments in healthy humans or patients with behavioral disorders. Rodent models are of limited value because brain circuits that mediate emotion, cognition, and social behavior differ significantly between primates and rodents [1,2]. Despite the fact that certain key features of behavior in humans cannot be adequately modeled in nonhuman

* Corresponding author. Fax: +1 650 498 7761.
E-mail address: dmlyons@stanford.edu (D.M. Lyons).
[1] These authors contributed equally to this work.

1046-2023/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2005.09.017

primates [3], comparative studies of homologous brain regions in humans and nonhuman primates are required to advance our understanding of behavior and mental health.

Over the years, neuroscientists have addressed fundamental questions regarding brain structure and function by adopting an impressive assortment of techniques and methodologies. In this regard, microarray technology stands as a uniquely powerful new tool for the global characterization of novel molecular pathways and mechanisms underlying behavior. Microarray studies aim to measure the steady-state levels of all actively transcribed genes, with the implicit premise that observable features of behavior are often associated with recognizable patterns of gene expression that reflect structural, functional, and metabolic adaptations in relevant brain regions [4–6]. Since the introduction of microarray technology a decade ago [7,8], large-scale gene expression studies have grown exponentially in neuroscience research. Microarrays dramatically increase the capacity and efficiency of data collection, and facilitate a systems level approach [9,10]. At the same time, however, microarrays create unprecedented challenges in statistical analysis and biological interpretation in neuroscience research.

In this report, we broadly examine technical and analytical aspects of gene expression microarray methods relevant to the study of primate behavior. Specifically, we discuss the different platforms that are now available, the impact of various analysis decisions on the final results, and the challenges that arise when working with complex tissues in the brain. We also consider how gene sequence differences complicate the application of microarrays designed for humans to measure samples collected from monkeys or apes. To illustrate aspects of the analysis of microarray data from nonhuman primates, we present examples from our use of human arrays to examine the differences between hippocampus and dorsolateral prefrontal cortex in adult squirrel monkeys.

2. Microarray platforms

Microarrays measure gene expression by quantifying the amount of hybridization between the RNA (or cDNA) under study and DNA probes that are immobilized on a solid surface. The DNA probes on many recent, comprehensive platforms are designed to cover the entire transcriptome, i.e., the steady-state levels of all known transcribed genes. Typically, RNA is extracted from cell or tissue samples, labeled with a marker (usually a fluorescent dye), and hybridized to the arrays. The fluorescent intensity values at each probe location are then determined from the scanned optical image of the array, and these intensities reflect the abundance of the targeted RNA in the sample of interest. When each sample is analyzed on a different array, the relative expression levels of every mRNA can be compared across samples.

Currently, two kinds of microarray probes are prevalent: complementary DNA (cDNA) and oligonucleotides. Probes can be pre-synthesized and then robotically printed in a predefined matrix on microscope slides (i.e., spotted arrays), or lithographically synthesized directly on silicon chips (i.e., oligonucleotide arrays). The cDNA probes are generally spotted on standard glass slides. Oligonucleotides can be either printed (e.g., Agilent arrays) or directly synthesized (e.g., Affymetrix Genechips). Multiple probes may be designed to target the same mRNA. For example, Affymetrix Genechips use both perfectly matched (PM) oligonucleotide probes that correspond to a segment of the targeted transcript, and mismatched (MM) probes that are identical in sequence to the corresponding PM probe except for a single MM base at the central position [11]. MM probes are intended to provide a reference signal to control for nonspecific binding. Each PM probe paired with its corresponding MM probe forms a probe pair. Generally, 11–20 probe pairs form a probe set designed to measure a transcript of interest. Probes of different lengths have been used. Affymetrix uses 25-base PM and MM probes, whereas other manufacturers of microarrays use 50-, 60-, or 70-base PM probes without corresponding MM probes.

Researchers can obtain microarrays by purchasing them from commercial sources, or by printing them on their own [12,13]. Commercial arrays have the advantage of stability, quality control, standardized protocols, software for data processing, and a broad base of users for sharing expertise and for multi-study data integration. Self-printed arrays that are made at individual laboratories or academic facilities are usually less expensive and more flexible with regard to probe content, but require careful control of printing quality, and are generally less amenable to meta-analysis.

The various platforms are constantly improving in terms of sensitivity, specificity, ease of use, and coverage of all known genes. Recent studies indicate that when investigators carefully design and execute their experiments with standardized protocols and appropriate analyses, the most common microarray platforms are comparable in performance and consistent across platforms [14,15]. One of the major sources of variation between platforms is in probe design, as probes that interrogate different segments of a given gene transcript may produce different results due to alternative splicing of mRNA, or cross-hybridization with other transcripts. It is therefore important to apply multiple methods, including quantitative PCR and in situ hybridization histochemistry, to validate important conclusions.

3. Microarray data analysis

Like all other studies, a microarray-based study consists of an experimental design, data collection, statistical analysis, and interpretation of the results. Issues related to sample size, technical replication, and assignment of samples to arrays differ according to the type of study that the researcher plans to conduct. Generally, microarray experiments are designed to provide comparisons between groups of samples to generate lists of differentially expressed genes;

clustering and classification of genes or tissue samples; or class prediction of unknown samples [16,17].

After the collection of scanned optical images of each array and before the statistical analysis, a series of decisions must be made with respect to data preprocessing. This includes image analysis to extract the raw signal intensity values, quality control, background subtraction and normalization, gene filtering, and, in cases where several probes are designed to target each transcript (e.g., Affymetrix Genechips), summarization of the intensities of all individual probes in a given probe set. Due to space limitations, in the following sections we briefly review background correction, data normalization, probe set summary methods, probe annotation, statistical criteria, gene filtering, and aspects of gene ontology that are particularly relevant to primate research. The analysis of two-color cDNA microarray data must also deal with an additional factor known as "dye bias," i.e., some fluorescent molecules are incorporated more efficiently than others. We refer readers interested in this issue to several relevant reviews [18,19], and focus instead on single-color microarrays, particularly the Affymetrix Genechips.

3.1. Processing methods for Affymetrix data

For Affymetrix data, various software tools have been developed to perform background correction, normalization, and probe summary. These include Microarray Suite 5 (MAS5), the default software provided by Affymetrix, and third-party tools such as the Robust Multi-chip Average (RMA) method [20], DNA Chip Analyzer (dCHIP) [21], the Positional-Dependent-Nearest-Neighbor (PDNN) method [22], and RMA corrected for GC-content (GCRMA) [23]. All of these tools use the hybridization signals from all individual probes within a probe set to generate a single expression value for each probe set. They differ, however, in the algorithms applied; and this leads to significant differences in the outcome. One of the major differences is the way in which MM probe information is included in the analysis. MAS5 generates expression summary values based on PM–MM differences. In contrast, the default version of RMA uses only PM probes because it was reported that incorporating MM probe intensities appears to add noise with no obvious gain in sensitivity [24,25]. For this reason, many users of dCHIP also prefer the PM-only version to the PM–MM version. GCRMA is an extension of the RMA algorithm that incorporates some MM probe information by using physical models of nonspecific hybridization based on the GC-content of the probes [23].

The choice of which method to use affects the outcome of the downstream analysis, especially when the samples of interest do not have large gene expression differences [26–28]. In comparative studies of human samples, MAS5 often differs from the other methods, but none of the others has emerged as consistently better [20,27,29,30]. We have observed a similar pattern when the different methods are used to process human Affymetrix Genechips applied to samples of squirrel monkey hippocampus and dorsolateral prefrontal cortex. Of the 12,666 probe sets examined, MAS5 identified many more genes as being differentially expressed between brain regions compared to all other methods at a given fold-difference criterion (Table 1). However, a much smaller percentage of the differences identified by MAS5 were statistically significant, i.e., 26–33% for MAS5 compared to 71–98% for the other methods. Moreover, the correlations between MAS5 and each of the other methods were consistently less than the correlations among the other methods for the entire set of 12,666 between region fold-difference scores (Table 2). MAS5 appears to generate data for squirrel monkey brain tissue samples that differ from the other methods. This may reflect the effect of human–monkey sequence differences in PM probes compounded by additional background noise from sequence differences in the corresponding MM probes. It may be best to avoid the use of mismatch probe information on gene expression microarrays used in cross-species applications.

Table 1
Comparison of methods used to process human Affymetrix Genechips for a study of differences between brain regions in adult squirrel monkeys

Method	1.2-fold difference[a]	1.5-fold difference[b]	FDR < 5%[c]
MAS5	7250	3399	373
GCRMA	1720	580	680
RMA	1705	365	1069
dCHIP[d]	1945	331	1791

[a,b] Numbers of genes differentially expressed in dorsolateral prefrontal cortex compared to hippocampus at ±1.2- and ±1.5-fold difference thresholds.

[c] Also provided are the numbers of genes identified as differentially expressed at a False discovery rate (FDR) of 5%.

[d] Results from the PM-only version of dCHIP are presented.

3.2. A bioinformatic challenge: probe annotation

Appropriate interpretation of microarray results depends on the correct annotation of individual probes. The assignment of gene identity to probes can be problematic because the definitions of transcribed genes continue to evolve along with the ongoing progress in genome sequencing and annotation [31]. Currently, many models of known transcripts are available to serve as the basis for probe annotation, e.g., Unigene, Refseq, and ENSEMBL genes. Occasionally, different models may assign different gene identities to a given probe. This may result in different interpretations of the same array data. Therefore, it is important

Table 2
Correlations between fold-difference scores computed using four different methods to process human Affymetrix Genechips for a study of differences between squirrel monkey hippocampus and dorsolateral prefrontal cortex

	MAS5	GCRMA	RMA
GCRMA	0.44[a]		
RMA	0.54	0.86	
dCHIP[b]	0.41	0.71	0.81

[a] Each Pearson correlation coefficient is based on the between region fold-difference scores for 12,666 probe sets.

[b] Results from the PM-only version of dCHIP are presented.

to state in each published study the specific version of the gene model that is used for probe annotation.

Every few months, Affymetrix updates the mapping of the target GenBank sequences of each probe set to the most recent gene identity, but does not update the mapping of individual probes. When a probe from a given probe set matches the sequence of another unintended gene, the probe is not marked as unreliable, or reassigned to a new probe set that corresponds to the other gene. There is also much redundancy in the Affymetrix annotation system, as certain genes are targeted by more than one probe set. Microarray data are difficult to interpret when different probe sets for the same gene yield discordant results.

One way to improve annotation is to remap all individual probes to the most recent gene definitions for a variety of common gene models, and then reassemble the annotated probes into new probe sets. We have employed this approach for different versions of Affymetrix Genechips in our collaborative research effort to study gene expression changes in the human brain associated with psychiatric disorders. Every probe that can be uniquely assigned is periodically remapped to Refseq, DoTS transcripts, Unigene, ENSEMBL gene, ENSEMBL transcripts, and ENSEMBL exons. The annotated probes are then reassembled, and the resulting probe set definition files (i.e., CDF files) are made freely available for public use at http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genomic_curated_CDF.asp.

3.3. Statistical criteria

It is widely recognized that microarray experiments should include replicated samples [32]. When groups of replicated samples are compared, appropriate statistical methods are needed to assess the significance of observed differences in gene expression levels. Although fold-difference thresholds are often reported as the primary statistic in microarray studies, these fail to account for technical or biological variations, which usually differ from gene to gene [33]. To report the reliability of microarray results, statistics such as the Student t test and its associated P values are most often presented. As tens of thousands of genes are analyzed in a typical study, many false positives are expected by chance. For example, when testing 12,000 independent hypotheses (i.e., one for each probe set), approximately 600 genes will be misidentified as being statistically significant at a significance level of $P < 0.05$ due to chance alone. The overall level of statistical significance must therefore be adjusted to account for multiple tests.

The Bonferroni correction assumes that all genes vary independently, and is commonly regarded as too conservative for microarray data because many genes are in fact regulated in a correlated fashion. Permutation-based correction methods have been developed to determine the adjusted family-wise levels of significance for individual genes [34]. The calculation of false discovery rates (FDRs) [35–37] is a different method used to adjust for multiple tests. Instead of specifying the adjusted P value for each gene, FDR provides the estimated ratio of false positives among the entire set of significant genes at a certain cutoff. With an FDR cutoff of 5% applied to our squirrel monkey data, dCHIP and to a lesser extent RMA identify more genes compared to either MAS5 or GCRMA (Table 1). Species differences in GC-content may adversely influence the accuracy of data from GCRMA. The specific cutoffs applied to obtain a list of differentially expressed genes depend on the particular circumstances of a given study, and can strongly impact the downstream analysis [38]. In many situations cutoffs may be less important, however, than the ranking of genes by their strength of evidence. Rankings provide a prioritized list of candidate genes for follow-up validation, hypothesis formulation, or marker discovery.

3.4. Higher-level data interpretation using gene ontology

When a list of genes is obtained, the functions of a few key genes on the list may lead to valuable insights and important discoveries. Often, however, it is also informative to look for general biological "themes" represented by the list as a whole. One may, for example, examine the biological functions of all of the genes to determine whether particular themes are overrepresented by the list [13,26]. To accomplish this type of higher-level analysis, a bioinformatic infrastructure is required in which the functions of known gene products are systematically annotated to microarray probe sets in a form that is amenable to computation.

Gene ontology [39], KEGG [40], and Genmapp [41] are examples of gene ontology resources for managing the wealth of biological knowledge in a controlled and scalable fashion. Once each gene is linked to one or several different functional categories, researchers can begin to ask whether specific functional categories are enriched or depleted on a list of genes identified by applying certain cutoffs and filtering criteria to microarray data. Recently, numerous tools have been developed to identify patterns of annotation terms on lists of ranked or unranked genes (http://www.geneontology.org/GO.tools.shtml), and to evaluate the statistical significance of specific distribution patterns of annotation terms [42,43]. These tools are used to advance our understanding of the neurobiology of behavior beyond single gene-by-gene conclusions, and represent an important step toward fulfilling a central promise of the systems approach [9,10,44].

4. Challenges in studying brain tissue

Lifelong plasticity, individual variability, complexity, and heterogeneity of the brain, and the need for high quality tissue present problems for microarray studies in behavioral neuroscience research. Gene expression profiles of brain tissue samples likely reflect the cumulative effect of experiences gained over the entire life span. Differences in

early maternal availability in squirrel monkeys, for example, are known to alter numerous neurobiological outcomes measured much later in life [45]. Gene expression profiles in animals are often highly variable even in carefully controlled experiments [46]. Moreover, post-mortem brain tissue samples provide only a single snapshot of the transcriptome at one point in time. Gene expression profiles of brain tissue samples from humans are affected by conditions at death [47,48], that are generally unrelated to prior experiences or behavioral traits.

In terms of specific practical matters, the power to detect differential expression in brain tissue samples is influenced by sample size, effect size, variance, and the false positive rates that investigators are willing to accept. Sample size is limited by the availability of suitable subjects: some primate populations are larger than others, while human disorders may be common or rare. Because the great apes (i.e., chimpanzees, gorillas, orangutans, and gibbons) are extensively protected in the wild, and tissue samples from great apes in captivity are generally limited, Old-World and New-World monkeys have been most extensively studied. Of these primates, high quality brain tissue from Old-World monkeys (e.g., macaques) is most easily obtained because these animals are maintained in large federally funded breeding facilities for biomedical research.

Effect size varies from gene to gene, but gene expression differences in case-control comparisons rarely exceed a 2-fold threshold in microarray studies of the brain [49–51]. The brain is an unique complex organ in which many gene transcripts are expressed at low levels [52–54], or are restricted to specific subpopulations of cells [53,55]. Although new technologies will certainly increase the sensitivity of arrays, the current generation of microarrays most readily detects higher abundance gene transcripts [49,56]. The heterogeneous anatomical distribution creates problems because the levels of expression for certain transcripts and fold-difference scores in case-control comparisons may be "diluted" in microarray studies that rely on bulk brain tissue with limited anatomical resolution. Laser capture microdissection in combination with linear amplification may be used to isolate homogenous subsets of cells in situations where these can be readily discerned [50,57,58].

Small but real differences in gene expression can be obscured by extraneous sources of variation in microarray studies of the brain. Variations in gene expression levels may arise from numerous technical and biological factors, which, if not carefully controlled, will compromise the power and accuracy of microarray results. A sound experimental design must take all of these factors into consideration, and, whenever possible adopt extensive technical and biological replications. We have found, for example, that reagents from different lots, RNA labeling, hybridization, and optical scanning may all introduce systematic variation between different batches of hybridized arrays within a single experiment. Such technical variation is superimposed on naturally occurring interindividual differences between the biological samples of interest, and adversely affects the outcome of experiments when gene expression differences between cases and controls are small.

To assess all sources of variation, we routinely create color-coded correlation matrices to illustrate similarities and differences between all possible pairs of arrays within a particular study. These pair wise correlation "heatmaps" help to identify technical and biological outliers within a set of arrays (Fig. 1A). Poor-performing outlier arrays must be excluded from further analysis, or the affected samples need to be rerun. After the exclusion of outlier arrays, the pair wise color-coded correlation matrix is examined for

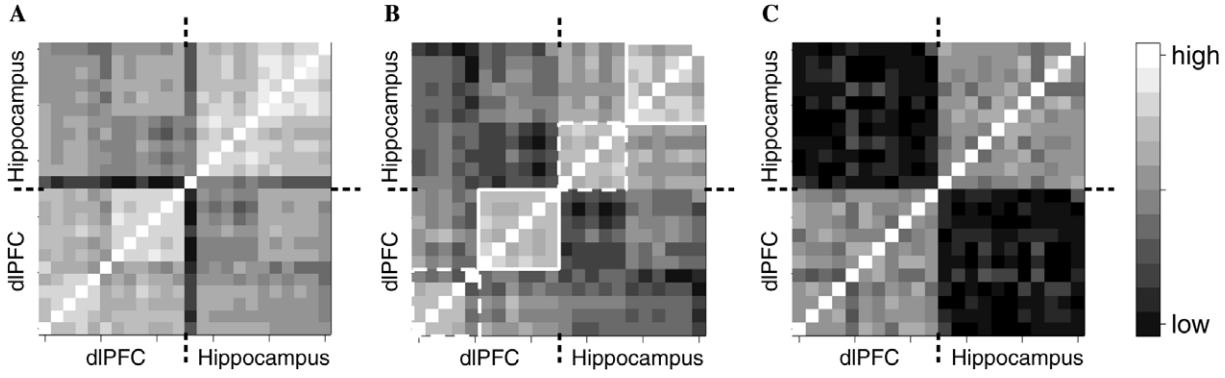

Fig. 1. Grayscale-coded correlation matrices illustrate similarities and differences between all possible pairs of arrays within a study of squirrel monkey hippocampus and dorsolateral prefrontal cortex (dlPFC). Within each matrix, individual arrays are listed in the same order from left to right and bottom to top. Lightly shaded cells indicate high Pearson correlations calculated from all probe set signal intensities for a given pair of arrays, and dark cells indicate low correlations. (A) Correlations between 12 dlPFC and 12 hippocampal samples collected from the same 12 monkeys show modest differences between the two regions: higher correlations are observed between pairs of arrays from the same region (lightly shaded cells) compared to arrays from different brain regions (darker cells). The hippocampal sample depicted by the vertical and horizontal stripe of dark cells is an aberrant outlier array. (B) After exclusion of the outlier array and the corresponding dlPFC array from the same monkey, the correlation matrix now reveals a strong batch effect, as shown by two distinct blocks of highly correlated arrays within each brain region. The batch effect is also evident in (A), but is less pronounced due the presence of the outlier array, which alters the overall grayscale-coding. (C) Median centering of the arrays within blocks adjusts for the batch effect, and highlights the brain region-related pattern with highly homogenous samples in each region.

evidence of sample batch effects (Fig. 1B). Technical factors related to processing subsets of samples together may alter an entire batch of microarrays with a constant effect. Once identified, batch differences can be adjusted by using fixed-effect models, or by subtracting from each probe set on each array the median intensity value of all samples in the batch. Median centering does not disturb the relative ranking of samples in the batch, and often provides a more homogeneous data set from which previously obscured patterns may emerge (Fig. 1C).

Because many sources of variation are impossible to control after the fact, microarray studies of primate brain tissue should adopt sound design features such as randomization, double-blind analysis, and/or balanced assignment of samples to different experimental conditions. After completion of the data analysis, other methods such as quantitative PCR, in situ hybridization histochemistry, and immunohistochemistry are often applied to the same set of samples to validate the microarray results. This type of validation is important, of course, but generally no amount of reanalysis of the same samples will be as convincing as confirmation with independent methods applied to new sample sets.

5. Species differences: a unique challenge

Currently, the application of microarray technology in primate research is hampered by the limited availability of species-specific arrays. This is primarily due to the lack of sufficient gene sequence information for most nonhuman primates. Even though the genomes of the chimpanzee (http://genome.ucsc.edu/cgi-bin/hgGateway?clade=vertebrate&org=Chimp) and rhesus macaque monkey (http://www.hgsc.bcm.tmc.edu/projects/rmacaque/) have been recently sequenced, human arrays provide the most broadly accessible option for the majority of primatologists. New microarray platforms designed for primates will gradually become available in the future. Affymetrix has announced that a macaque monkey Genechip will be released in 2005, and a program funded by the European Consortium (http://www.eupeah.org/) is now generating marmoset-specific cDNA arrays. Such efforts, however, are costly and time consuming. To develop spotted cDNA microarrays for a previously uncharacterized transcriptome, one needs to first obtain a large number of unique cDNA or expressed sequence tag (EST) sequences [59]. Known cDNA sequences are also required to design oligonucleotide probes. Without known sequences, primatologists will likely continue to rely on the use of microarrays designed for human transcripts [28,60,61]. This approach is based on the assumption that similar gene sequences in closely related species allow reasonably reliable detection of many orthologous genes.

Sequence divergence is, however, an important problem in all cross-species applications of microarrays. Even a 5% sequence difference means that, on average, each 25-base oligonucleotide probe will contain one mismatch. Because a contiguous stretch of 16 or more base pair matches has been considered sufficient for stable hybridization [62], longer probes like those on cDNA microarrays are thought to be less susceptible to sequence divergence-related problems. Recent evidence suggests, however, that even these probes can severely distort results obtained from between-species comparisons [63].

The problem of sequence divergence for cross-species microarray applications is especially troublesome when the research objective is to compare different species with respect to evolutionary questions [25,64–68]. Differences observed between species may reflect either genuine differences in gene expression levels, or methodological artifacts related to sequence differences that impair microarray hybridization. These two effects are entirely confounded, and hinder the interpretation of data obtained from human arrays applied to samples obtained from monkeys or apes. In cross-species comparisons, sequence divergence can also distort normalization of the signal intensities on each array, and may result in overestimation of the nonhuman primate expression levels of genes without sequence differences [69]. For studies that aim to compare groups of samples within a single species, sequence differences carry the same effect across all samples under study, and no longer represent a systematic bias in microarray studies. Yet even for these within-species applications, a detailed understanding of probe performance is valuable because poorly hybridized probes contribute to background noise, and adversely influence probe signal summarization for the affected probe sets.

Only recently has the impact of sequence divergence on microarray applications been systematically examined for nonhuman primates. For example, using a detection reliability tool provided by MAS5 software for human Affymetrix Genechips, Chismar et al. [70] reported that genes inconsistently called "Present" among pairs of technical replicates of frontal lobe tissue had 2-fold lower probe set signal intensities in rhesus macaque monkeys, and were more variable in monkeys compared to humans. This presumably reflects increased variability in probe signal intensities for probe sets affected by gene sequence mismatches. Although fewer genes were consistently called "Present" within pairs of technical replicates for monkey, a similar percentage of all genes (~8%) was found to switch from "Present" to "Absent," or vice versa, in both human and monkey. These findings suggest that sequence differences do not affect the consistency of calls (i.e., "Present" or "Absent"), but result in the loss of coverage of the monkey transcriptome. Other investigators have likewise reported that fewer genes are consistently called "Present" when nonhuman primate tissue is compared to human samples hybridized on human microarrays [65,67,71,72]. In 11 adult squirrel monkeys we found, for example, that 16% of 12,666 probe sets were consistently called "Present" in hippocampus and dorsolateral prefrontal cortex, and 43% of the probe sets were called "Absent" on all 22 microarrays. In 12 samples of hippocampus and 12 samples of dorsolateral prefrontal cortex obtained from 12 healthy adult humans,

these same assessments were, respectively, 24 and 36%. Similar results have been reported in studies of humans and rhesus macaque monkeys for different tissue types [67,71].

Several methods have been proposed to deal with the sequence divergence problem for studies of nonhuman primates. In primates with known gene sequences, the Affymetrix probes on current human platforms can be individually realigned to the nonhuman primate sequences to identify probes that perfectly match conserved segments of orthologous genes. Only these probes are then used to summarize probe signal intensities for the nonhuman primate samples [73]. Wang et al. [71] reported that the use of more highly homologous probe sets reduces the relative discordance between species with respect to signal intensities, but the sequence-based strategy is of limited utility in species for which most gene sequences are unknown. Wang et al. [71] attempt to define interspecies conserved probe sets on the Affymetrix HG-U133 Plus2.0 Genechip that covers nearly thirty thousand human Unigenes, successfully defined only 2704 macaque monkey genes and 1190 chimpanzee genes. These numbers will undoubtedly increase with the recent release of the draft genome sequence for rhesus macaque monkeys and chimpanzees. Recently, for example, we realigned the probes for different versions of human Affymetrix Genechips to putative chimpanzee transcripts defined according to human Unigenes. From the Affymetrix probes that can be uniquely assigned, we then created custom definition probe sets files corresponding to nearly 19,000 chimpanzee "Unigenes." These files are freely available for download at (http://brainarray.mhri.med.umich.edu/Brainarray/Database/CustomCDF/CDF_download.asp).

Given the limitations of sequence-based methods for the majority of nonhuman primates, a potentially fruitful approach is to exclude problematic probes based solely on probe hybridization characteristics obtained from actual microarray data [25,62,74]. Probes harboring between-species sequence differences are likely to show hybridization levels in nonhuman primates that are inconsistent with those in human samples compared to other probes in the same probe set, as long as the majority of the probes are showing comparable characteristics in humans and in nonhuman primates [25]. Masking out probes based on aberrant behavior may indirectly lead to the exclusion of probes containing sequence differences, even in the absence of actual gene sequence information for nonhuman primates. This approach has been used to identify a significant proportion of sequence differences in chimpanzee brain tissue [25,65,66], and effectively reduces the relative number and magnitude of gene expression differences between chimpanzee and humans [25]. Currently, the development of tools of this type is an active area research.

6. Summary and conclusions

Nonhuman primates will continue to play a vital role in research on the neural basis of emotional, cognitive, and social aspects of behavior. In recent years, the application of microarray technology has enabled behavioral neuroscientists to analyze global patterns of gene expression within defined brain regions. The use of microarrays to study brain tissue presents, however, various limitations and unique challenges. The primate brain is a complex organ, with a heterogeneous distribution of distinct subpopulations of cells, intricate signaling and regulatory circuits, and exquisite lifelong sensitivity to environmental variation. These factors result in high levels of interindividual variability in gene expression, and often subtle differences between the specific conditions under study. These inherent difficulties highlight the importance of a careful controlled experimental design, extensive replication, standard protocols, and sound analysis practices that encompass key considerations about normalization, probe summary and annotation, statistical criteria, and higher-level analysis of gene ontology. In particular, the lack of primate gene sequence information and, consequently, the limited availability of species-specific microarrays, is a major problem for the near future. The use of arrays designed for humans is currently the only option available to many primatologists, and great attention must be paid to the effect of sequence differences on the quality and interpretation of cross-species applications of microarrays. Finally, it is critical that the outcomes of microarray studies of nonhuman primates are validated and extended by the classic techniques (e.g., quantitative PCR, in situ hybridization, and immunohistochemistry) applied to new, independent sets of brain tissue samples.

Acknowledgments

The authors are members of the Pritzker Neuropsychiatric Disorders Research Consortium, which is supported by the Pritzker Neuropsychiatric Disorders Research Fund L.L.C. A shared intellectual property agreement exists between the Pritzker Neuropsychiatric Disorders Research Fund L.L.C and the University of Michigan, the University of California, and Stanford University to encourage the development of appropriate findings for research and clinical applications.

References

[1] L.J. Porrino, D. Lyons, Cereb. Cortex 10 (2000) 326–333.
[2] T.M. Preuss, J. Cogn. Neurosci. 7 (1995) 1–24.
[3] D.M. Lyons, in: A.F. Schatzberg, C.B. Nemeroff (Eds.), Textbook of Psychopharmacology, third ed., American Psychiatric Publishing, Washington, DC, USA, 2003, pp. 117–126.
[4] G.E. Robinson, Y. Ben-Shahar, Genes Brain Behav. 1 (2002) 197–203.
[5] D.P. Toma, K.P. White, J. Hirsch, R.J. Greenspan, Nat. Genet. 31 (2002) 349–353.
[6] C.W. Whitfield, A.M. Cziko, G.E. Robinson, Science 302 (2003) 296–299.
[7] D.J. Lockhart, H. Dong, M.C. Byrne, M.T. Follettie, M.V. Gallo, M.S. Chee, M. Mittmann, C. Wang, M. Kobayashi, H. Horton, E.L. Brown, Nat. Biotechnol. 14 (1996) 1675–1680.
[8] M. Schena, D. Shalon, R.W. Davis, P.O. Brown, Science 270 (1995) 467–470.
[9] J.D. Dougherty, D.H. Geschwind, Neuron 45 (2005) 183–185.

[10] H. Kitano, Science 295 (2002) 1662–1664.
[11] R.J. Lipshutz, S.P. Fodor, T.R. Gingeras, D.J. Lockhart, Nat. Genet. 21 (1999) 20–24.
[12] L.K. Nisenbaum, Genes Brain Behav. 1 (2002) 27–34.
[13] S.L. Karsten, L.C. Kudo, D.H. Geschwind, Int. Rev. Neurobiol. 60 (2004) 1–23.
[14] J.E. Larkin, B.C. Frank, H. Gavras, R. Sultana, J. Quackenbush, Nat. Methods 2 (2005) 337–344.
[15] R.A. Irizarry, D. Warren, F. Spencer, I.F. Kim, S. Biswal, B.C. Frank, E. Gabrielson, J.G. Garcia, J. Geoghegan, G. Germino, C. Griffin, S.C. Hilmer, E. Hoffman, A.E. Jedlicka, E. Kawasaki, F. Martinez-Murillo, L. Morsberger, H. Lee, D. Petersen, J. Quackenbush, A. Scott, M. Wilson, Y. Yang, S.Q. Ye, W. Yu, Nat. Methods 2 (2005) 345–350.
[16] G.A. Churchill, Nat. Genet. 32 (Suppl.) (2002) 490–495.
[17] Y.H. Yang, T. Speed, Nat. Rev. Genet. 3 (2002) 579–588.
[18] M.K. Kerr, M. Martin, G.A. Churchill, J. Comput. Biol. 7 (2000) 819–837.
[19] K. Dobbin, J.H. Shih, R. Simon, Bioinformatics 19 (2003) 803–810.
[20] R.A. Irizarry, B.M. Bolstad, F. Collin, L.M. Cope, B. Hobbs, T.P. Speed, Nucleic Acids Res. 31 (2003) e15.
[21] C. Li, W.H. Wong, Proc. Natl. Acad. Sci. USA 98 (2001) 31–36.
[22] L. Zhang, M.F. Miles, K.D. Aldape, Nat. Biotechnol. 21 (2003) 818–821.
[23] Z.J. Wu, R.A. Irizarry, R. Gentleman, F. Martinez-Murillo, F. Spencer, J. Am. Stat. Assoc. 99 (2004) 909–917.
[24] R.A. Irizarry, B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, T.P. Speed, Biostatistics 4 (2003) 249–264.
[25] W.P. Hsieh, T.M. Chu, R.D. Wolfinger, G. Gibson, Genetics 165 (2003) 747–757.
[26] C. Konradi, Brain Res. Rev. 50 (2005) 142–155.
[27] K. Shedden, W. Chen, R. Kuick, D. Ghosh, J. Macdonald, K.R. Cho, T.J. Giordano, S.B. Gruber, E.R. Fearon, J.M. Taylor, S. Hanash, BMC Bioinformatics 6 (2005) 26.
[28] P.E. Lachance, A. Chaudhuri, J. Neurochem. 88 (2004) 1455–1469.
[29] Y. Barash, E. Dehan, M. Krupsky, W. Franklin, M. Geraci, N. Friedman, N. Kaminski, Bioinformatics 20 (2004) 839–846.
[30] Z. Wu, R.A. Irizarry, Nat. Biotechnol. 22 (2004) 656–658.
[31] P. Wang, M. Dai, R. Thompson, S.J. Watson, F. Meng, in: M. He, G. Narasimhan, S. Petoukhov, (Eds.), Advances in bioinformatics and its applications: Proceedings of the International Conference, Fort Lauderdale, Florida edn, World Scientific Publishing Company, 2005.
[32] M.L. Lee, F.C. Kuo, G.A. Whitmore, J. Sklar, Proc. Natl. Acad. Sci. USA 97 (2000) 9834–9839.
[33] P. Pavlidis, W.S. Noble, Genome Biol. 2 (2001) 0042.1–0042.15.
[34] Y.C. Ge, S. Dudoit, T.P. Speed, Test 12 (2003) 1–77.
[35] Y. Benjamini, Y. Hochberg, J. R. Stat. Soc. Ser. B—Methodol. 57 (1995) 289–300.
[36] Y. Benjamini, D. Yekutieli, Ann. Stat. 29 (2001) 1165–1188.
[37] J.D. Storey, R. Tibshirani, Proc. Natl. Acad. Sci. USA 100 (2003) 9440–9445.
[38] K.H. Pan, C.J. Lih, S.N. Cohen, Proc. Natl. Acad. Sci. USA 102 (2005) 8961–8965.
[39] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, G. Sherlock, Nat. Genet. 25 (2000) 25–29.
[40] M. Kanehisa, S. Goto, S. Kawashima, Y. Okuno, M. Hattori, Nucleic Acids Res. 32 (2004) D277–D280.
[41] K.D. Dahlquist, N. Salomonis, K. Vranizan, S.C. Lawlor, B.R. Conklin, Nat. Genet. 31 (2002) 19–20.
[42] V.K. Mootha, C.M. Lindgren, K.F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstrale, E. Laurila, N. Houstis, M.J. Daly, N. Patterson, J.P. Mesirov, T.R. Golub, P. Tamayo, B. Spiegelman, E.S. Lander, J.N. Hirschhorn, D. Altshuler, L.C. Groop, Nat. Genet. 34 (2003) 267–273.
[43] D.A. Hosack, G. Dennis Jr., B.T. Sherman, H.C. Lane, R.A. Lempicki, Genome Biol. 4 (2003) R70.
[44] G.L. Henry, K. Zito, J. Dubnau, Curr. Opin. Neurobiol. 13 (2003) 570–576.
[45] D.M. Lyons, A.F. Schatzberg, Neurobiol. Learn. Mem. 80 (2003) 97–104.
[46] C.C. Pritchard, L. Hsu, J. Delrow, P.S. Nelson, Proc. Natl. Acad. Sci. USA 98 (2001) 13266–13271.
[47] J.Z. Li, M.P. Vawter, D.M. Walsh, H. Tomita, S.J. Evans, P.V. Choudary, J.F. Lopez, A. Avelar, V. Shokoohi, T. Chung, O. Mesarwi, E.G. Jones, S.J. Watson, H. Akil, W.E. Bunney Jr., R.M. Myers, Hum. Mol. Genet. 13 (2004) 609–616.
[48] H. Tomita, M.P. Vawter, D.M. Walsh, S.J. Evans, P.V. Choudary, J. Li, K.M. Overman, M.E. Atz, R.M. Myers, E.G. Jones, S.J. Watson, H. Akil, W.E. Bunney Jr., Biol. Psychiatry 55 (2004) 346–352.
[49] S.J. Evans, N.A. Datson, M. Kabbaj, R.C. Thompson, E. Vreugdenhil, E.R. De Kloet, S.J. Watson, H. Akil, Eur. J. Neurosci. 16 (2002) 409–413.
[50] N.A. Datson, L. Meijer, P.J. Steenbergen, M.C. Morsink, S. van der Laan, O.C. Meijer, E.R. de Kloet, Eur. J. Neurosci. 20 (2004) 2541–2554.
[51] K. Mirnics, Nat. Rev. Neurosci. 2 (2001) 444–447.
[52] H. Ozawa, E. Kushiya, Y. Takahashi, Neurosci. Lett. 18 (1980) 191–196.
[53] Y. Takahashi, Prog. Neurobiol. 38 (1992) 523–569.
[54] N.A. Datson, J. van der Perk, E.R. de Kloet, E. Vreugdenhil, Hippocampus 11 (2001) 430–444.
[55] A. Sawatari, E.M. Callaway, Neuron 25 (2000) 459–471.
[56] D.H. Geschwind, Proc. Natl. Acad. Sci. USA 97 (2000) 10676–10678.
[57] P.P. Sanna, A.R. King, L.D. van der Stap, V. Repunte-Canonigo, Brain Res. Brain Res. Protoc. 15 (2005) 66–74.
[58] V.A. Vincent, J.J. DeVoss, H.S. Ryan, G.M. Murphy Jr., J. Neurosci. Res. 69 (2002) 578–586.
[59] C.L. Magness, P.C. Fellin, M.J. Thomas, M.J. Korth, M.B. Agy, S.C. Proll, M. Fitzgibbon, C.A. Scherer, D.G. Miner, M.G. Katze, S.P. Iadonato, Genome Biol. 6 (2005) R60.
[60] D.E. Redmond Jr., J.L. Zhao, J.D. Randall, A.C. Eklund, L.O. Eusebi, R.H. Roth, S.R. Gullans, R.V. Jensen, Brain Res. Dev. Brain Res. 146 (2003) 99–106.
[61] G. Cheng, M.J. Mustari, S. Khanna, J.D. Porter, Invest. Ophthalmol. Vis. Sci. 44 (2003) 3842–3855.
[62] W. Ji, W. Zhou, K. Gregg, N. Yu, S. Davis, S. Davis, Nucleic Acids Res. 32 (2004) e93.
[63] Y. Gilad, S.A. Rifkin, P. Bertone, M. Gerstein, K.P. White, Genome Res. 15 (2005) 674–680.
[64] M.W. Karaman, M.L. Houck, L.G. Chemnick, S. Nagpal, D. Chawannakul, D. Sudano, B.L. Pike, V.V. Ho, O.A. Ryder, J.G. Hacia, Genome Res. 13 (2003) 1619–1630.
[65] M. Caceres, J. Lachuer, M.A. Zapala, J.C. Redmond, L. Kudo, D.H. Geschwind, D.J. Lockhart, T.M. Preuss, C. Barlow, Proc. Natl. Acad. Sci. USA 100 (2003) 13030–13035.
[66] P. Khaitovich, B. Muetzel, X. She, M. Lachmann, I. Hellmann, J. Dietzsch, S. Steigele, H.H. Do, G. Weiss, W. Enard, F. Heissig, T. Arendt, K. Nieselt-Struwe, E.E. Eichler, S. Paabo, Genome Res. 14 (2004) 1462–1473.
[67] M. Uddin, D.E. Wildman, G. Liu, W. Xu, R.M. Johnson, P.R. Hof, G. Kapatos, L.I. Grossman, M. Goodman, Proc. Natl. Acad. Sci. USA 101 (2004) 2957–2962.
[68] W. Enard, P. Khaitovich, J. Klose, S. Zollner, F. Heissig, P. Giavalisco, K. Nieselt-Struwe, E. Muchmore, A. Varki, R. Ravid, G.M. Doxiadis, R.E. Bontrop, S. Paabo, Science 296 (2002) 340–343.
[69] T.M. Preuss, M. Caceres, M.C. Oldham, D.H. Geschwind, Nat. Rev. Genet. 5 (2004) 850–860.
[70] J.D. Chismar, T. Mondala, H.S. Fox, E. Roberts, D. Langford, E. Masliah, D.R. Salomon, S.R. Head, Biotechniques 33 (2002) 516–524.
[71] Z. Wang, M.G. Lewis, M.E. Nau, A. Arnold, M.T. Vahey, BMC Bioinformatics 5 (2004) 165.
[72] M. Marvanova, J. Menager, E. Bezard, R.E. Bontrop, L. Pradier, G. Wong, FASEB J. 17 (2003) 929–931.
[73] P. Khaitovich, G. Weiss, M. Lachmann, I. Hellmann, W. Enard, B. Muetzel, U. Wirkner, W. Ansorge, S. Paabo, PLoS Biol. 2 (2004) E132.
[74] S. Nagpal, M.W. Karaman, M.M. Timmerman, V.V. Ho, B.L. Pike, J.G. Hacia, Nucleic Acids Res. 32 (2004) e51.