Toolbox

Editor's Note: Toolboxes are a new, occasional feature in the *Journal* designed to briefly highlight a new method or a resource of general use in neuroscience or to critically analyze existing approaches or methods. For more information, see <http://www.jneurosci.org/misc/itoa.shtml>.

Brain Microarray: Finding Needles in Molecular Haystacks

Nicole M. Lewandowski and Scott A. Small

The Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Department of Neurology, and Center for Neurobiology and Behavior, Columbia University College of Physicians and Surgeons, New York, New York 10032

Key words: entorhinal; gene; mRNA; dentate gyrus; fMRI; Alzheimer's disease

Introduction

When technical advances generated novel genetic markers in the 1980s, they were combined with linkage analysis allowing thousands of genes to be profiled simultaneously (Botstein et al., 1980; Weber and May, 1989). This capability has been used to isolate the molecular causes of most monogenetic disorders of the brain. With this success as a backdrop, the persistent difficulty in isolating the molecular defects that underlie complex brain disorders is particularly glaring. Complex disorders emerge from an interplay between genes and the environment and constitute by far the vast majority of diseases. It is no surprise that when microarray, a technique that simultaneously profiles the levels of thousands of mRNA transcripts, was introduced in the late 1990s (Lockhart et al., 1996; DeRisi et al., 1997), this technical advance was enthusiastically met by clinical neuroscientists. Because mRNA expression profiles are anatomically specific, and because expression levels are influenced by both genetic polymorphisms and environmental input, microarray held great promise for uncovering the pathogenic molecules underlying complex disorders.

Despite a number of interesting and potentially important findings (Mirnics et

al., 2001; Geschwind, 2003), this promise has not yet been fully realized. Molecular heterogeneity can be invoked to account for this lack of success, in which defects in many separate molecular pathways produce overlapping disease phenotypes. Nevertheless, we believe that molecular parsimony should still be assumed (i.e., that each disease is driven primarily by relatively few molecular pathways) and that the difficulty in pinpointing pathogenic molecules with microarray is more likely a reflection of its analytical challenges and technical limitations.

Here we will review the analytical challenges presented by microarray and show how combining microarray with other technologies, such as *in vivo* brain imaging, can be used to: (1) maximize "signal amplitude" in a microarray experiment; (2) constrain sources of noise; and (3) address the high rate of false-positivity that naturally occurs with multiple comparisons. Next, we discuss the technical limitations presented by microarray and the importance of confirming microarray findings at the protein level. We conclude by considering how a microarray finding can be validated, testing whether the identified molecule is causally related to the disease under investigation.

We use our recent experience applying microarray to late-onset Alzheimer's disease (Small et al., 2005) to illustrate these points. Indeed, with both an early-onset and a late-onset form, Alzheimer's disease provides an informative example. Because the early-onset form is monogenetic, linkage analysis has successfully isolated most of the molecular defects that cause this extremely rare form of disease (St George-

Hyslop, 1999). In contrast, the late-onset form, accounting for >95% of all cases, is a complex disorder, and its molecular causes remain elusive.

Maximizing signal amplitude

Comparing the expression levels of a hypothetical molecule underlying Alzheimer's disease with the expression levels of a hypothetical melanoma-causing molecule is useful in illustrating the analytical challenges inherent to microarray (Fig. 1). In melanoma, as in all neoplastic disorders, an unregulated increase in transcription is a defining feature of the disease. The expression levels of cases compared with controls (the signal amplitude in a microarray experiment) are therefore expected to differ by orders of magnitude (Fig. 1A). In contrast, low signal amplitude must be assumed for many brain disorders, including Alzheimer's disease, because of molecular mechanisms underlying neuronal dysfunction (Fig. 1B). Although protein aggregates and cell death ultimately mark Alzheimer's disease and other neurodegenerative disorders, "cell sickness" is the key feature of the earliest and most informative stage of disease (e.g., synaptic dysfunction in a relatively intact neuron) (Selkoe, 2002). Alzheimer's and other neurodegenerative diseases are therefore part of a larger category of brain disorders characterized by physiological rather than structural lesions, a category that includes most psychiatric and many developmental disorders, as well as age-related cognitive decline. Because synaptic dysfunction can occur with relatively subtle changes in molecular expression, low single amplitude is the first challenge presented by microarray when

Received Sept. 21, 2005; revised Oct. 5, 2005; accepted Oct. 5, 2005.

This work was supported in part by National Institutes of Health Grants AG025161 and AG08702, the Beeson Faculty Scholar Award from the American Federation of Aging, the McKnight Neuroscience of Brain Disorders Award, and the James S. McDonnell Foundation. We thank Joseph Lee for his helpful comments.

Correspondence should be addressed to Scott A. Small, Center for Neurobiology and Behavior, Columbia University, 630 West 168th Street, New York, NY 10032. E-mail: sas68@columbia.edu.

DOI:10.1523/JNEUROSCI.4006-05.2005

Copyright © 2005 Society for Neuroscience 0270-6474/05/2510341-06\$15.00/0

applied to these intractable disorders of the brain.

Almost all brain disorders are regionally selective, and it would seem self-evident that microarray performed on brain tissue most affected by the disease process should maximize signal amplitude (Fig. 1C). Microarray performed on large swatches of brain, containing a mixture of affected, less affected, and unaffected tissue, is likely to dilute an already low but meaningful abnormality in expression (Fig. 1B). Although obvious from an analytical perspective, it is not always clear how to identify the one brain region most vulnerable and affected by disease. This issue is particularly problematic for physiological rather than structural disorders, where telltale histological or structural lesions cannot be used as landmarks. Even in neurodegenerative diseases, such as Alzheimer's disease (Selkoe, 2002), Parkinson's disease (Dauer and Przedborski, 2003), and Huntington's disease (Arrasate et al., 2004), it is generally agreed that the histological distribution of protein aggregates does not necessarily mark the sites of greatest physiologic dysfunction. Furthermore, although cells do eventually degenerate after many years of sublethal injury, patterns of cell death might reflect differential sensitivity to apoptosis rather than to synaptic dysfunction. Finally, because these disorders have long protracted courses, postmortem studies might bias against the earliest and most informative sites of disease. In Alzheimer's disease, for example, which begins by causing hippocampal-dependent memory loss, postmortem maps of cell death have implicated the CA1 subfield (West et al., 1994), the entorhinal cortex (Shoghi-Jadid et al., 2002), or both (Price et al., 2001). Indeed, relying on postmortem findings, many microarray studies of Alzheimer's disease have focused on the CA1 subfield as the targeted region of investigation (Ginsberg et al., 2000).

With the improvement in functional imaging techniques, the physiological integrity of small and discrete subregions of the brain can now be assessed (Small et al., 2000, 2002, 2004), and these techniques are well suited for pinpointing sites most affected by physiological disorders. By imaging the functional integrity of multiple hippocampal subregions in living subjects (Small et al., 1999, 2002), imaging studies have clearly shown that the entorhinal cortex, not the CA1 subfield, is the first and most profoundly affected site in Alzheimer's disease (Fig. 1C). This pattern of physiological dysfunction matches the an-

Figure 1. Enhancing signal-to-noise in a microarray experiment. Signal amplitude in a microarray experiment is defined as the difference in expression levels between tissue isolated from patients and controls. Signal noise is defined as the inter-individual variance within a group. **A**, For a hypothetical melanoma molecule isolated from a melanoma tissue sample, signal amplitude is predicted to be relatively large. **B**, In contrast, for a hypothetical Alzheimer molecule isolated from the entire hippocampal formation, signal amplitude is predicted to be relatively small, a problem inherent to all physiological disorders of the brain. Furthermore, the noise in Alzheimer's disease is expected to be relatively large because tissue, as in most brain disorders, is harvested from postmortem brains. **C**, Profiling expression levels from a select brain region most affected by a disease process, the entorhinal cortex in the case of Alzheimer's disease, will maximize signal amplitude. **D**, Expression levels of a neighboring brain region unaffected by a disease, the dentate gyrus in the case of Alzheimer's disease, can be used to constrain sources of noise.

atomical distribution of intracellular neurofibrillary tangles (Braak and Braak, 1991). Guided by these functional imaging findings, we decided to focus on the entorhinal cortex rather than the CA1 subfield as the target site in our microarray analysis of Alzheimer's disease (Small et al., 2005).

Constraining sources of noise

The hypothetical example comparing Alzheimer's disease with melanoma highlights a second analytical challenge inherent to microarray studies of the brain. Compared with melanoma, an Alzheimer's-related molecule is expected to manifest greater inter-individual variance, factors that influence mRNA expression independent of the disease. A main reason for the increase in "signal noise" (Fig. 1B) is that tissue samples used to investigate brain disorders are typically harvested from postmortem brains. Thus, in contrast to neoplasms in which tissue is typically harvested from biopsy material, sources of expression–variance extend beyond differences in genetic heritage and environment, to include differences that occur during the dying process, a particularly powerful and vexing source of noise (Li et al., 2004). For example, a slow ago-

nal death, typically accompanied by hypoperfusion and blood-borne toxins or drugs, will dramatically affect mRNA levels compared with death caused by acute trauma.

Identifying a brain region relatively unaffected by a disease can, in principle, be used to constrain inter-individual sources of variance (Fig. 1D). Global sources of noise, such as vascular-related changes caused by blood-borne drugs or hypoperfusion, can be assumed to influence expression levels of a particular gene within both the affected and unaffected brain regions. Thus, expression levels of the unaffected region can be used to statistically normalize against these global sources of noise. Of course, this approach will not constrain variance caused by noise factors that are region specific. Many brain regions are typically resistant to a disease process, and so careful consideration should be given in selecting an unaffected brain region as a noise reducer in a microarray study. We advocate the use of an unaffected region that neighbors the affected region, because neighboring regions are more likely to share vascular supply (Small et al., 2005).

Identifying a neighboring unaffected region is relatively trivial when studying

structural diseases, such as brain tumors, stroke, or multiple sclerosis. Isolating an unaffected region is more challenging for physiological disorders, and here again functional imaging is well suited for the task. For example, functional imaging studies have extended on postmortem observations showing that the physiological integrity of the dentate gyrus is relatively preserved even in late stages of Alzheimer's disease (Small et al., 1999).

To summarize, establishing the spatial profile of a disease can be used to perform an analytical "double subtraction," comparing the expression levels from the affected and unaffected regions and comparing this difference between cases and controls. In statistical terms, the spatial profile can be used to convert a microarray analysis from a simple *t* test design to a more sophisticated factorial ANOVA, including both within- and between-group factors, a design that has proven to be effective in addressing analytical challenges of microarray (Kerr et al., 2000). Accordingly, in our Alzheimer's disease study, we applied a factorial design to microarray data generated from the entorhinal cortex and the dentate gyrus, harvested from each brain with and without disease. This analysis yielded 33 molecules that conformed to the spatial profile of Alzheimer's disease, molecules that are differentially expressed in the entorhinal cortex compared with the dentate gyrus, between cases and controls (Small et al., 2005). Nevertheless, because of reasons discussed below, we were forced to acknowledge that many of these molecules were not necessarily linked to disease pathogenesis.

Distinguishing true from false-positive findings

With the improvement in gene-chip technology, the expression levels of >20,000 genes can be assessed. A statistical analysis comparing expression profiles between disease and control tissue must therefore contend with the high false-positive rate (the type I error) that naturally occurs when thousands of comparisons are performed (Slonim, 2002). By applying a statistical cutoff, a typical microarray study yields a laundry list of molecules, and only some of these molecules are truly relevant to the disease under investigation, whereas many others are simply false-positive findings.

How can the type I error be effectively addressed? Acquiring data from thousands of tissue samples is considered impractical, and because expression levels are not independent events, applying sim-

Figure 2. Separating true from false-positive findings. As in all multivariate analyses, a microarray experiment must address the high false-positivity that naturally occurs with multiple comparisons. **A**, As evidenced in microarray experiments applied to simple organisms such as yeast [DeRisi et al. (1997), their Figure 5A, reprinted with permission], the temporal profile of the behavior under investigation can be used as an analytical filter against false-positive findings. **B**, Functional imaging is well suited for extracting temporal profiles of brain dysfunction, particularly in physiological disorders of the brain. For example, imaging studies have suggested that once entorhinal dysfunction begins in Alzheimer's disease, dysfunction does not worsen across age or over time (control data points are shown in blue; Alzheimer's disease is shown in red). **C**, In contrast, imaging studies have shown that brain regions implicated in cognitive aging decline linearly across the lifespan. As in simpler organisms, these temporal profiles of brain dysfunction can be used as analytical filters against false-positive findings when applied to specifically tailored microarray data.

ple statistical corrections is considered inappropriate. Using more sophisticated statistical methods is one general approach for dealing with the problem of multiple comparisons. For example, statistical techniques such as principle components or cluster analysis can be used, essentially looking for covariate patterns of expression levels among groups of molecules. However, given the potential discordance between mRNA and protein, as discussed below, it is not clear how the covariate pattern observed among levels of different mRNAs easily translates to a meaningful relationship at the protein level. Furthermore, attempting to extract a covariate pattern among a group of molecules might, in principle, make it difficult to pinpoint a single molecule or molecular pathway that plays a primary role in disease pathogenesis.

Relying on a more sophisticated understanding of the temporal profile of the disease under investigation is an alternative approach that has proven effective for addressing the type I error. For example, in microarray studies applied to simpler organisms, such as yeast (DeRisi et al., 1997), worms (Kim et al., 2001), and flies (White et al., 1999), the most relevant molecules have been successfully isolated by requiring that expression patterns match the temporal profile of a phenotype under investigation (Fig. 2A). In principle, the same logic can be applied to brain disorders, by assuming that the most pathogenic molecules will match the temporal profile of regional dysfunction. In Alzheimer's disease, for example, functional imaging has mapped the temporal profile of dysfunction, showing that compared with controls entorhinal dysfunction is age and time independent (de Leon

et al., 2001; Small et al., 2002) (Fig. 2B). The implication is that once dysfunction occurs, it does not change with time, an interpretation supported by theoretical studies of neurodegeneration (Clarke et al., 2000). In contrast, and as another example, functional imaging studies have shown that in cognitive aging, regional dysfunction progresses linearly across the lifespan (Small et al., 2002, 2004) (Fig. 2C).

As with simpler organisms, this temporal information can be used to generate an a priori model predicting how a molecule most relevant to a particular disorder should temporally behave (Fig. 2). Then this model can be forward-applied onto a specifically tailored microarray dataset, identifying only those molecules with a temporal profile that best conforms to the model. Thus, by forcing a microarray experiment to be hypothesis driven, a temporal profile of dysfunction can be used as an analytical filter against false-positive findings.

Accordingly, in our Alzheimer's disease study, in which we anticipated the problem of false-positivity, we purposefully harvested samples from brains that covered a broad age-span. This allowed us to perform a secondary analysis on the 33 molecules originally derived from the spatial profile of Alzheimer's disease and to ask which of these molecules also conformed to its temporal profile (Fig. 2B). Only five molecules survived this analytical filter, and among these, VPS35, the core of the retromer trafficking complex (Seaman, 2005), best conformed to the spatiotemporal profile of disease (Small et al., 2005). Nevertheless, despite plausible mechanisms for Alzheimer's pathogenesis suggested by the trafficking itinerary of the retromer (Seaman, 2005), microarray

findings by themselves do not support mechanistic interpretations, as discussed below.

Confirming a microarray finding

Two molecular confirmations are required of any microarray finding. The first is to confirm the basic mRNA finding using reverse transcription (RT)-PCR, the gold-standard method for quantifying mRNA. In truth, rarely have RT-PCR and microarray findings diverged, and so the requirement for RT-PCR confirmation has become more relaxed.

The second confirmation, testing the protein levels of an associated mRNA finding, is the one that has emerged as more important. Protein, not mRNA, is the meaningful end product of gene expression, and most microarray studies have implicitly assumed a simple relationship between transcription and translation (i.e., that more mRNA necessarily reflects more protein). By systematically quantifying mRNA and protein levels from the same tissue sample, a number of studies (Ideker et al., 2001; Chen et al., 2002; Greenbaum et al., 2003; Lee et al., 2003; Mehra et al., 2003; Beyer et al., 2004; Tian et al., 2004) have shown that this simple relationship does not hold true (Fig. 3A). Although in many cases there is a positive correlation between levels of mRNA and protein, often there is no correlation, and frequently a negative correlation is observed.

In retrospect, based on a more realistic model of gene expression, the complex relationship between mRNA and protein should not be surprising (Fig. 3B). Not only are transcription and translation governed by independent mechanisms and separate time constants, but mRNA and protein are likewise degraded by different pathways. Moreover, feedback mechanisms exist such that a protein that undergoes accelerated degradation might lead to increased translation that could account for a negative correlation between mRNA and protein. Thus, at best, a disease-related abnormality in levels of mRNA can indicate a protein abnormality, but by itself does not inform on the directionality of the defect. At worst, an abnormality in mRNA levels may not be accompanied by a coexistent protein defect.

The potential discordance between mRNA and protein levels can be addressed with relative ease: Any mRNA finding detected with microarray can, and in fact must, be tested at the protein level, for example using Western blot analysis.

Figure 3. Clarifying the complex mRNA–protein relationship. Although often assumed, the relationship between the levels of mRNA and protein is not simple. **A**, By systematically measuring mRNA and protein levels from the same tissue sample, mRNA–protein correlation may be positive, negative, or absent. The graph shows data acquired from one study, in which the correlation between mRNA and protein level is plotted against the 66 different molecules assessed [adapted from Chen et al. (2002)]. **B**, Gene expression is governed by multiple independent components, accounting for the complex and often discordant mRNA–protein relationship; thus, protein levels must be assessed as follow-up to any microarray finding.

Figure 4. Establishing causality. Although an isolated molecule might be linked to a disease, concluding that the molecule plays a causal role in disease pathogenesis is prohibited by the correlational nature of microarray. The results of three experimental approaches can lend support to claims of molecular causality: **A**, First, the candidate molecule, or the molecular pathway to which it belongs, can be systematically manipulated in cell culture, measuring a meaningful biochemical readout. For example, manipulating a molecule that is thought to be pathogenic in Alzheimer's disease should affect the levels of A β peptide. **B**, Second, the candidate molecule, or the molecular pathway to which it belongs, can be manipulated in animal models using transgenic technologies. In the case of Alzheimer's disease, this manipulation should phenocopy the behavioral, electrophysiological, and biochemical phenotype of Alzheimer's disease. **C**, Finally, genomic screens can be performed in humans, testing whether polymorphisms in the molecular pathway increase the risk of Alzheimer's disease.

Validating a microarray finding and establishing causality

A microarray experiment on human tissue is fundamentally correlative. No matter how sophisticated the experimental design or how significant the statistics, it is impossible to conclude that a molecule isolated by microarray plays a primary role in disease pathogenesis. At best, a well designed microarray experiment increases the odds of filtering out false-positive findings. However, given the analytical challenges, no microarray finding can ever be free of this shadow of doubt. Even if the finding is true (namely, that the isolated molecule is linked to the disease under investigation), the possibility always exists that a shift in expression is a secondary response to a protracted illness and is not by itself an “upstream” cause of neuronal dysfunction.

In general, three experimental approaches can be used to validate that an identified molecule, or the molecular pathway to which it belongs, plays a causal role in disease pathogenesis (Fig. 4): The first is cell-culture experiments that can be used for disorders for which there is a meaningful molecular or cellular readout (Fig. 4A). In Alzheimer's disease, for example, β -amyloid (A β) peptide is the biochemical “smoking gun” of the disease, and so manipulating a molecule relevant to disease pathogenesis should affect the levels of this peptide. Accordingly, as validation of our microarray study, we systematically increased or decreased the expression of retromer proteins, using expression vectors or small interfering RNA, and we did indeed observe the predicted affect on A β (Small et al., 2005).

This finding suggests that retromer dysfunction can contribute to disease pathogenesis, but also acts to validate the spatiotemporal assumptions that went into the design and analysis of the microarray study.

Transgenic or viral vector manipulations in animals are a second approach to test for causality (Fig. 4B). In our case, we are investigating a transgenic mouse with retromer dysfunction (Lee et al., 1992) and testing whether these mice phenocopy Alzheimer's disease: at the behavioral level, testing for hippocampal-dependent memory deficits; at the electrophysiological level, testing for synaptic dysfunction; and, at the biochemical level, testing for A β and other indicators of disease. Manipulating the expression of a molecule of interest in a living brain has, of course, many advantages. Nevertheless, many human disorders are species-specific, suggesting that in some cases humans possess unique molecular machinery for disease pathogenesis (Hill and Walsh, 2005). Thus, animal models may not always be expected to replicate the disease process.

Returning to humans, screening for polymorphisms in the identified molecule pathway is another approach for testing causality (Fig. 4C). We have recently approached our colleagues who have genotyped a large cohort of families with late-onset Alzheimer's disease, and they are testing whether polymorphisms in any of the molecules that make up the retromer trafficking pathway increase disease risk. If so, this would provide the strongest evidence for causality. Nevertheless, this approach will only work for diseases that have a strong genetic component. In a disease driven primarily by the environment, a microarray finding might still reflect a causal process even in the absence of an identified genetic polymorphism.

Conclusions

After a first blush of excitement, microarray must now contend with a certain degree of backlash. Frustration is a common response to the first-generation studies, in which the most meaningful findings are typically buried in a laundry list of differentially expressed molecules. Although legitimate, this frustration does not indict the fundamental utility of microarray as a technique that can pinpoint pathogenic molecules of the brain. In fact, microarray suffers from its technical precision, sensitive enough to detect subtle mRNA differences that cause a disease but also to detect disease-independent differences in ge-

netic heritage, environment, and regional anatomy.

In an influential review on gene mapping, Terwilliger and Goring (2000) make the compelling point that despite technical breakthroughs that have generated large amounts of genetic information and the availability of powerful computational statistics, the most important development for future discovery is "study design." Namely, the analytical challenges and limitations inherent to genetic screening are best overcome by first establishing a more sophisticated understanding of the disease phenotype and then using this understanding to guide data acquisition and data analysis.

This conclusion also applies to gene-expression profiling. We have the technology, with microarray soon being able to measure all of the mRNA expressed in a given sample, and we have the computational power. Now, by using complementary techniques such as *in vivo* brain imaging, we can better "phenotype" a brain disorder: anatomically, by identifying brain regions vulnerable and resistant to a disease, and temporally, by mapping regional dysfunction over time or across age groups. Just as with Alzheimer's disease, high-resolution brain imaging can now map the spatiotemporal phenotype of many other complex disorders, by imaging the subregions of the hippocampal formation or prefrontal cortex in schizophrenia or cognitive aging, the subregions of the amygdala in autism or anxiety disorders, or the subregions of the brainstem or basal ganglia in Parkinson's disease and other movement disorders. Once established, a spatiotemporal phenotype of dysfunction can be used to improve the design of a microarray experiment, harnessing its precision and allowing microarray to fully realize its promise for isolating pathogenic molecules of the brain.

References

- Arrasate M, Mitra S, Schweitzer ES, Segal MR, Finkbeiner S (2004) Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431:805–810.
- Beyer A, Hollunder J, Nasheuer HP, Wilhelm T (2004) Post-transcriptional expression regulation in the yeast *Saccharomyces cerevisiae* on a genomic scale. *Mol Cell Proteomics* 3:1083–1092.
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314–331.
- Braak H, Braak E (1991) Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol (Berl)* 82:239–259.
- Chen G, Gharib TG, Huang CC, Taylor JM, Misek DE, Kardias SL, Giordano TJ, Iannettoni MD, Orringer MB, Hanash SM, Beer DG (2002) Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics* 1:304–313.
- Clarke G, Collins RA, Leavitt BR, Andrews DF, Hayden MR, Lumsden CJ, McInnes RR (2000) A one-hit model of cell death in inherited neuronal degenerations. *Nature* 406:195–199.
- Dauer W, Przedborski S (2003) Parkinson's disease: mechanisms and models. *Neuron* 39:889–909.
- de Leon MJ, Convit A, Wolf OT, Tarshish CY, DeSanti S, Rusinek H, Tsui W, Kandil E, Scherer AJ, Roche A, Imossi A, Thorn E, Bobinski M, Caras C, Lesbre P, Schlyer D, Poirier J, Reisberg B, Fowler J (2001) Prediction of cognitive decline in normal elderly subjects with 2-[(18)F]fluoro-2-deoxy-D-glucose/positron-emission tomography (FDG/PET). *Proc Natl Acad Sci USA* 98:10966–10971.
- DeRisi JL, Iyer VR, Brown PO (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680–686.
- Geschwind DH (2003) DNA microarrays: translation of the genome from laboratory to clinic. *Lancet Neurol* 2:275–282.
- Ginsberg SD, Hemby SE, Lee VM, Eberwine JH, Trojanowski JQ (2000) Expression profile of transcripts in Alzheimer's disease tangle-bearing CA1 neurons. *Ann Neurol* 48:77–87.
- Greenbaum D, Colangelo C, Williams K, Gerstein M (2003) Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol* 4:117.
- Hill RS, Walsh CA (2005) Molecular insights into human brain evolution. *Nature* 437:64–67.
- Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292:929–934.
- Kerr MK, Martin M, Churchill GA (2000) Analysis of variance for gene expression microarray data. *J Comput Biol* 7:819–837.
- Kim SK, Lund J, Kiraly M, Duke K, Jiang M, Stuart JM, Eizinger A, Wylie BN, Davidson GS (2001) A gene expression map for *Caenorhabditis elegans*. *Science* 293:2087–2092.
- Lee JJ, Radice G, Perkins CP, Costantini F (1992) Identification and characterization of a novel, evolutionarily conserved gene disrupted by the murine H beta 58 embryonic lethal transgene insertion. *Development* 115:277–288.
- Lee PS, Shaw LB, Choe LH, Mehra A, Hatzimanikatis V, Lee KH (2003) Insights into the relation between mRNA and protein expression patterns. II. Experimental observations in *Escherichia coli*. *Biotechnol Bioeng* 84:834–841.
- Li JZ, Vawter MP, Walsh DM, Tomita H, Evans SJ, Choudhary PV, Lopez JF, Avelar A, Shokoochi V, Chung T, Mesarwi O, Jones EG, Watson SJ, Akil H, Bunney Jr WE, Myers RM (2004) Systematic changes in gene expression in postmortem human brains associated with

- tissue pH and terminal medical conditions. *Hum Mol Genet* 13:609–616.
- Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14:1675–1680.
- Mehra A, Lee KH, Hatzimanikatis V (2003) Insights into the relation between mRNA and protein expression patterns. I. Theoretical considerations. *Biotechnol Bioeng* 84:822–833.
- Mirnics K, Middleton FA, Lewis DA, Levitt P (2001) Analysis of complex brain disorders with gene expression microarrays: schizophrenia as a disease of the synapse. *Trends Neurosci* 24:479–486.
- Price JL, Ko AI, Wade MJ, Tsou SK, McKeel DW, Morris JC (2001) Neuron number in the entorhinal cortex and CA1 in preclinical Alzheimer disease. *Arch Neurol* 58:1395–1402.
- Seaman MN (2005) Recycle your receptors with retromer. *Trends Cell Biol* 15:68–75.
- Selkoe DJ (2002) Alzheimer's disease is a synaptic failure. *Science* 298:789–791.
- Shoghi-Jadid K, Small GW, Agdeppa ED, Kepe V, Ercoli LM, Siddarth P, Read S, Satyamurthy N, Petric A, Huang SC, Barrio JR (2002) Localization of neurofibrillary tangles and beta-amyloid plaques in the brains of living patients with Alzheimer disease. *Am J Geriatr Psychiatry* 10:24–35.
- Slonim DK (2002) From patterns to pathways: gene expression data analysis comes of age. *Nat Genet [Suppl]* 32:502–508.
- Small S, Wu E, Bartsch D, Lacefield C, DeLaPaz R, Mayeux R, Stern Y, Kandel E (2000) Imaging physiologic dysfunction of individual hippocampal subregions in humans and genetically modified mice. *Neuron* 28:653–664.
- Small S, Kent K, Pierce A, Leung C, Kang M, Okada H, Honig L, Vonsattel J, Kim T (2005) Model-guided microarray implicates the retromer complex in Alzheimer's disease. *Ann Neurol*, in press.
- Small SA, Perera GM, DeLaPaz R, Mayeux R, Stern Y (1999) Differential regional dysfunction of the hippocampal formation among elderly with memory decline and Alzheimer's disease. *Ann Neurol* 45:466–472.
- Small SA, Tsai WY, DeLaPaz R, Mayeux R, Stern Y (2002) Imaging hippocampal function across the human life span: is memory decline normal or not? *Ann Neurol* 51:290–295.
- Small SA, Chawla MK, Buonocore M, Rapp PR, Barnes CA (2004) Imaging correlates of brain function in monkeys and rats isolates a hippocampal subregion differentially vulnerable to aging. *Proc Natl Acad Sci USA* 101:7181–7186.
- St George-Hyslop PH (1999) Molecular genetics of Alzheimer disease. *Semin Neurol* 19:371–383.
- Terwilliger JD, Goring HH (2000) Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. *Hum Biol* 72:63–132.
- Tian Q, Stepaniants SB, Mao M, Weng L, Feetham MC, Doyle MJ, Yi EC, Dai H, Thorsson V, Eng J, Goodlett D, Berger JP, Gunter B, Linseley PS, Stoughton RB, Aebersold R, Collins SJ, Hanlon WA, Hood LE (2004) Integrated genomic and proteomic analyses of gene expression in mammalian cells. *Mol Cell Proteomics* 3:960–969.
- Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44:388–396.
- West MJ, Coleman PD, Flood DG, Troncoso JC (1994) Differences in the pattern of hippocampal neuronal loss in normal ageing and Alzheimer's disease. *Lancet* 344:769–772.
- White KP, Rifkin SA, Hurban P, Hogness DS (1999) Microarray analysis of *Drosophila* development during metamorphosis. *Science* 286:2179–2184.