Published in final edited form as:
Neurobiol Dis. 2012 January ; 45(1): 3–7. doi:10.1016/j.nbd.2011.06.001.

Gene expression profiling of the brain: pondering facts and fiction

Amanda C. Mitchell[1] and **Károly Mirnics**[1,2,*]

[1]Department of Psychiatry, Vanderbilt University, Nashville, USA

[2]Vanderbilt Kennedy Center for Research on Human Development, Vanderbilt University, Nashville, USA

Abstract

During the last decade brain transcriptome profiling by DNA microarrays has matured, developed sound experimental design standards, reporting practices, analytical procedures, and data sharing resources. It has become a powerful scientific tool in the exploratory research portfolio. Along this journey by trial and error, we encountered a number of intriguing questions and comments - pondering the value of hypothesis-driven research, appropriate sample size, the importance and interpretation of transcripts changes vis-à-vis protein changes, the role of statistical stringency, false discovery and magnitude of expression change, and many other interesting questions. Our field fully acknowledges and tries to address all of these challenges associated with high-throughput, data-driven transcriptomics. As a research field, we strongly advocate implementing the highest standards of our trade, and we deeply believe that transcriptome profiling studies will continue to be essential for deciphering the pathophysiological mechanisms leading to complex brain disorders.

Keywords

DNA microarray; transcriptome; brain; gene expression; data interpretation; experimental design; schizophrenia; psychiatric disorder

It has been over a decade since the publication of the first high-throughput gene expression profiling studies of the brain (Hakak et al., 2001; Lockhart and Barlow, 2001; Mimmack et al., 2002; Mirnics et al., 2000; Pasinetti, 2001). During the last decade the opinion and attitude of the scientific community has changed toward these technologies multiple times. The first phase, lasting about 3–4 years, was characterized by enthusiasm, excitement, and often unjustified optimism. Many viewed DNA microarray technology as a “magic bullet” that would fundamentally change our understanding of various brain disorders, and during this golden era of microarrays funding agencies were generous in supporting data-driven exploratory efforts. As a result, several important studies were generated, along with a large number of mediocre studies that resulted in long lists of gene dumpouts without meaningful interpretation of the findings. The backlash was predictable, and over a relatively short time

*Correspondence: **Karoly Mirnics**, Department of Psychiatry, Vanderbilt University, 8130A MRB III, 465 21st Avenue South, Nashville TN 37203, USA, karoly.mirnics@vanderbilt.edu, Office phone: 615-936-1074.

period the dominant scientific opinion transformed itself to one of skepticism toward anything that was DNA microarray-generated. In the meanwhile, proteomics became “hot”, followed by other novel “omics” technologies, and microarray expression studies fell in disgrace: microarray manuscripts started to be considered “descriptive studies”, and as such they routinely started to receive editorial rejections by the top journals in the field of neuroscience. The attitude of many journals and editors was best summarized by the boilerplate rejection letter of the Journal of Neuroscience from 2005, stating that “We tend to be circumspect when we receive an expression array paper.” (Journal-of-Neuroscience, 2005) - not caring about the quality of the presented science, but condemning the technology. Many microarray enthusiasts jumped on a bandwagon of new, hotter and better-funded “omic” technologies – while the real puritans of microarray technology went to work: they teamed up with expert teams of biostatisticians and bioinformaticians and started to generate standards of performing, reporting (Brazma et al., 2001), analyzing (Irizarry et al., 2003; Subramanian et al., 2005; Tusher et al., 2001) and sharing (Barrett et al., 2005) the experiments. To the classical “most changed gene” analyses novel pathway assessments were added (Curtis et al., 2005; Dennis et al., 2003; Langfelder and Horvath, 2008; Mirnics et al., 2001; Mirnics et al., 2000; Subramanian et al., 2005), and the best microarray experiments started to include both data verification and functional follow-up assays. The gene expression profiling field grew up, entered its current phase of “accepted method” and took its place as a powerful, yet not omnipotent, scientific tool in the exploratory research portfolio.

Being part of this evolution process was both fascinating and frustrating. Over the years we received many criticisms from our peers, editors and reviewers. Some were fully justified and pointed out our own ignorance, while others were clearly malicious and had no foundation in reality. However, a number of observations and statements made in this process constituted great starting points for interesting discussions. The latter ones are the topic of this manuscript, in a hope that public pondering of some of these issues will help achieve better experimental design, higher quality data, enhanced recognition of a good microarray experiment, and improved interpretation of findings.

“Without a testable hypothesis, there is no good science.”

We strongly disagree with this notion. “I believe in ignorance-based methods because humans have a lot of ignorance and we should play to our strong suit …you take raw ignorance and turn it into processed ignorance, and processed ignorance, well-defined ignorance, well-asked questions that we don't know the answer to, that's the root of experiment.” - Dr. Eric Lander, founding director of the Broad Institute, eloquently pointed out the prevailing philosophy of data-driven researchers (Lander, 2002). It should be noted that initial and unproven hypotheses *per se* carry little value, and are limited by our current perception of how a biological system might work (Horvath et al., 2010; Mirnics and Pevsner, 2004). Data-driven, hypothesis-free approaches allow simultaneous testing of thousands of unformulated hypotheses: comparing the whole genome transcription machinery allows us to find the truly unknown, the unexpected, and the counterintuitive. It allows us to generate novel hypotheses, which can, and should be followed up in detail, in a hypothesis-driven fashion. Thus, in our experiments the precise initial hypothesis becomes secondary to a stellar experimental design that maximizes the chance of finding meaningful and fundamentally novel data (Mirnics et al., 2006). What to analyze (e.g. which cell types or brain regions), which subjects to include (e.g. co-morbidity, endophenotypes, technical exclusion), how to analyze the dataset are critical, and often very challenging considerations, and the poorly performed expression profiling studies almost invariably fail at this.

“Sample size should be increased…”

In theory, this is true. I have never met a biostatistician who did not advocate an increased sample size. Yet, the reality is that in expression profiling studies sample size will always be limited, especially in postmortem brain studies. We simply cannot even approach the sample size of genome-wide association studies (GWAS) studies (Sullivan, 2010), where thousands of samples are required to perform a meaningful study – we can be considered fortunate if we have several dozens of high-quality postmortem samples to work with (Mirnics et al., 2006). So, due to the limited sample size, should we abandon this line of research? Absolutely not! There is a huge difference between the gene expression profiling and genome wide association studies, which is strongly in favor of transcriptome assessment: invariably the disease associated gene expression signature is much stronger than the genetic association signal! For example, genetics studies of schizophrenia (and also bipolar disorder, ADHD, major depression, and a host of other disorders) could only identify miniscule genetic signals for any putative susceptibility genes (Harrison and Weinberger, 2005), yet, gene expression studies, despite the small sample sizes, have successfully identified expression changes that are characteristic of >50% of the diseased cohort (Arion et al., 2007; Garbett et al., 2008; Hakak et al., 2001; Middleton et al., 2002; Mirnics et al., 2000). The explanation for this is functional convergence and the fundamental nature of the brain transcriptome itself – gene expression networks are interdependent (Horvath and Mirnics, 2009; Mirnics, 2008; Winden et al., 2009), and the various genetic and environmental insults converge at critical molecular pathways, resulting in common alteration of transcript levels of different origin (Mirnics et al., 2006). As a result, one might argue that while genetic predispositions speak of the *origin*s of the disturbance seen in psychiatric disorders, gene expression changes give us clues about the *mechanisms* by which the pathophysiology progresses and the disease symptoms arise.

“Transcript changes do not matter if the protein levels are not altered.”

This is an interesting reasoning, and at the first glance it makes a lot of sense – or does it? Let us follow this line of reasoning for a second. If the transcript alteration does not matter in the absence of an obvious protein change, than the protein level is also unimportant if the protein is not activated or transported to its place of action. Obviously, this trafficking also becomes irrelevant unless it affects the electrical conductivity and the responsiveness of the neuron. Similarly, any change in excitability of the single cell also becomes meaningless unless the activity of a whole neural network is altered, and neural network activity fluctuations mean virtually nothing if the behavior of an individual is not altered… This logic could be applied in such a manner to make any important discovery seem meaningless. We believe that such a line of reasoning is deeply flawed, and put serious limits on future discoveries.

Furthermore, the argument that the transcript changes are trivial in the absence of changing protein levels also demonstrates disrespect for the most fundamental principles of brain anatomy. The majority of neurons *project out* from the brain regions where the cell bodies reside, and they traffic proteins from the somata to thousands of synaptic terminals in remote brain regions. So, a proteomic profile of any brain region at any given moment is a sum of proteins that are produced by the cell somata which reside in the harvested area (intrinsic proteins) *and* proteins that are contained within the synaptic terminals that project into the area from other brain regions (extrinsic proteins). Importantly, many proteins (such as synaptic release machinery) have both an intrinsic and an extrinsic source, and bulk tissue proteomics cannot differentiate between them. In contrast, the mRNA expression profile is primarily made up by the somatodendritic transcripts of cells that reside in the harvested brain region (intrinsic transcripts), with a minimal contribution from extrinsic mRNA

species. Thus, transcriptomics and proteomics measure different anatomical substrates, and this is very important for the interpretation of both "omics" findings: a change in mRNA, without a corresponding protein change might mean that the intrinsic cell population is affected at both the mRNA and protein level, but the substantial extrinsic source of the same protein (which is unchanged), makes impossible to detect the intrinsic protein change (Pongrac et al., 2002).

"The magnitude of gene expression change is too small to be physiologically relevant"

This is clearly an assumption, yet it is surprising how little discussion it has attracted over the years. The "bigger is better" mentality, while characteristic of our society, is not clearly applicable to biological systems. Homeostatic systems are differently tuned, with various degrees of tolerances. For example, changing the blood pH by 20% is lethal, yet dynamic blood glucose concentration doubling after meals or significant hormonal oscillations are part of normal physiology. The same principle applies to gene expression changes – many knockout animals do not show any readily discernable phenotypes (McMahon et al., 1996; Schluter et al., 1999) and copy-number variations (CNVs) normally occur at a high rate in the disease-free human population (Vogler et al., 2010). Yet, modest decreases in mRNA expression of the 67-kilodalton isoform of glutamic acid decarboxylase (GAD67) appear to be a critical, functional component of schizophrenia pathophysiology (Hashimoto et al., 2008). Thus, it appears that the individual function of the gene, its place in the transcriptome network, other genes performing a similar function, associated regulatory and compensatory mechanisms, simultaneously occurring gene expression changes in the same molecular pathway and many other factors decide the functional consequence of a gene expression change. While the magnitude of the gene expression change is also certainly one part of this equation, by itself it is a poor predictor of functional consequences.

"The most significant expression changes are the most important"

This is clearly another assumption. While statistical assessment is essential, probability values in a transcriptome profiling experiments represent a continuum, and there is no way to predict if a gene expression change associated with a p=0.0001 is more functionally relevant than an expression alteration reporting a p=0.01.

Furthermore, if statistical assessments are not selected carefully, the results can be quite misleading. In the following example,,the frontal cortex of a genetically altered mouse was compared to the frontal cortex of its wild-type littermates using whole-genome expression profiling. The analysis of 5 transgenic and 5 control mice, after RMA normalization, revealed the following $\log_2$ expression values for *Gene X*: 6.0, 6.8, 7.0, 6.2, 6.5 for WT, and 7.2, 10.0, 15.0, 8.0, 7.8 for TG samples. Performing a standard, two-tailed groupwise, equal-variance Student t-test in Microsoft Excel will report a significance of p=0.0637. If not careful, such a result will not be even noticed, and discarded as "non-significant". In contrast, *Gene Y* in the same experiment might report the following values across the studies samples: 2.2, 2.2, 2.2, 2.3, 2.3 for WT, and 2.4, 2.4, 2.3, 2.3, 2.3 for the TG samples. A similar t-test analysis for *Gene Y* in Microsoft Excel will yield a "significant" p=0.0203, yet *Gene Y* is certainly less promising for follow up than *Gene X*. How did this happen? The variance for Gene X was big and for Gene Y was small (possibly a normalization artifact), and the t-test alone was quite inappropriate to analyze the transcriptome changes.

Involving a knowledgeable biostatistician in your experimental design and discussing the data mining strategies *before* performing the experiment is always a good idea. Determining "true expression changes" and subsequently identifying the "most promising expression

changes" requires pulling resources from both biological and statistical knowledge, requiring (an often painful) cross-field education for both the molecular biologist and biostatistician.

"The authors did not apply a Bonferroni correction…"

… because they did not want to throw out the baby with the bath water. False discovery assessment is very important in all transcriptome profiling experiments, but in most cases Bonferroni correction is ill-suited for this purpose. First, in a typical transcriptome profiling experiment the number of genes tested is greater than the number of samples by 3 orders of magnitude. Second, human brain samples show a molecular diversity similar to the genetic and symptom diversity of the disease, postulating that not all the affected brains will have an identical gene expression signature, and this weakens statistical significance measures. Third, in complex brain disorders typical gene expression differences are relatively modest, often in the range of 20–50%, and precise expression measurement can be challenging. As a result, gene expression changes almost never reach significance that that can withstand a whole-genome Bonferroni p-value correction. Should we have applied a Bonferroni correction to the schizophrenia microarray datasets over the last ten years, we would have not obtained any novel leads – and the immune (Arion et al., 2007), mitochondrial (Middleton et al., 2002), synaptic (Mirnics et al., 2000), oligodendroglial (Hakak et al., 2001) and GABAergic (Hashimoto et al., 2008) changes could have not been identified and consequently replicated.

So, how can we ensure that the gene expression changes we uncovered are "real", and not a result of experimental noise or cohort bias? There are multiple alternatives to extremely stringent statistical corrections that still ensure that true biological findings, and not experimental artifacts, are uncovered. First, one can use less stringent statistical corrections for multiple comparisons, such as the Benjamini-Hochberg procedure (Sibille et al., 2004). This method is very effective, especially if the gene expression dataset is *a priori* trimmed for non-expressed genes and genes with very low variance: proper "trimming" procedures can reduce the number of performed comparisons by up to 75%, thus reducing the stringency of statistical correction that is required for false discovery assessment. Second, assessment of false discovery by various kinds of permutation analyses has gained great popularity over the years (Gao, 2006; Sohn et al., 2009). At core of all these assessments is mixing the experimental and control samples randomly into two balanced groups, and performing the same analyses repeatedly – for both pathways and individual genes. If the disease effect is bigger than the random noise, the random assignment of microarrays into variously permutated groups will report less differentially expressed genes than the "pure" comparison of control and experimental samples (Unger et al., 2005). In the past such analyses have been challenging for laboratories without strong bioinformatics support, however, recently developed public domain software packages make these assessments (both at the individual gene and pathway level) straightforward (Gentleman et al., 2004; Kuehn et al., 2008). Third, a very elegant, biology-based assessment of false discovery includes defining the differentially expressed transcripts in the initial cohort, and then testing this pattern in a second, independent replication cohort (Lazarov et al., 2005). Although false discovery assessments should always be incorporated in the experimental design, the exact choice of the false discovery analysis method should depend on availability of samples, cohort size, experimental design and many other factors.

"The authors should have employed the analysis strategy by Doe et al, published in Nature"

There is no such a thing as a "universally good" microarray study design. Each experimental design and analysis strategy should be tailored to your own experiment. Applying an experimental design or analysis strategy only because it was published in a high-impact journal is a common mistake made most often by trainees who try to emulate successful studies. Rather, considerations should include sample size, type and diversity, employed technology, number of replicates, and many other factors. Establishing clear and carefully crafted experimental parameters *before* the start of the study goes a long way toward obtaining meaningful experimental data – "patching up" microarray experiments with changing inclusion-exclusion criteria and adding additional samples at a later time usually results in noisy (and often uninterpretable) experimental outcomes.

"In this transcriptome profiling study, genes X and Y and pathway Z were not changed"

This is true, but is potentially misleading: not *finding* gene expression changed and not *being* changed are fundamentally different statements. DNA microarrays studies notoriously carry a high percentage of type II errors – a true biological difference is often not detected in these experiments. The explanation to this is a technical limitation, and has three major sources. First, universal hybridization conditions are never ideal for all of the hundreds of thousands of probes on a single microarray. Second, some DNA microarray probes perform less then ideally, and cross-hybridize to other genes than their own target, and the specificity of the signal is lost. Third, genes that are expressed in only a small subpopulation of cells in the tissue, or are expressed at very low levels, are at the cusp of detection limit even on the highest quality microarrays. Thus, failing to find a gene expression change in a microarray experiment is not a definitive proof that a gene expression is absent between the compared samples (Hollingshead et al., 2005), and negative data must be interpreted with great care.

"Finding hundreds of diverse gene expression changes is uninterpretable"

Not so. Unfortunately, the human brain appears to like simple solutions, and we scientists are not immune to this. A list of several hundred gene expression changes between two conditions is overwhelming, and we would prefer to explain the main pathophysiological process by very few changes in mRNA level. More is less: a big panel of changed genes somehow became less informative than the alteration of only one mRNA. However, the vast majority of complex brain disorders cannot be explained by individual gene dysfunction and the transcriptome profiling results merely reflect this complexity of the pathophysiological process. Unfortunately, many microarray studies fall short of proper interpretation of data by simply discussing the role of several, usually, "most changed" genes. The ability to decipher and interpret the data, and subsequently build a falsifiable model related to the studied pathophysiological process is the trait that sets apart a great expression profiling manuscript from a mediocre one.

So, how can we avoid writing a mediocre manuscript? First, we must reduce the emphasis on single gene changes, and shift our attention to analysis of co-regulated transcript networks (Korade and Mirnics, 2011; Voineagu et al., 2011). This can be achieved by using a number of freely available tools such as weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath, 2008), WholePathwayScope (Yi et al., 2006), Database for Annotation, Visualization and Integrated Discovery (DAVID) (Dennis et al., 2003), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Aoki and Kanehisa, 2005; Arakawa et al., 2005), Gene Set Enrichment Analysis (GSEA) (Kuehn et al., 2008). Second,

we must start more actively comparing our datasets to those generated by other investigators, especially the ones that are deposited in the main microarray data repositories (Barrett et al., 2005; Parkinson et al., 2005). Third, beyond data verification, we should attempt to follow-up our findings with additional experiments, obtaining supporting readouts about the consequences of the observed gene expression changes (Horvath and Mirnics, 2009; Huffaker et al., 2009).

"Transcriptome profiling studies are descriptive"

Actually, gene expression profiling studies are not any more descriptive than anatomical, brain imaging, genetic association or any other "omic" studies. None of these studies can prove causality beyond doubt, yet they all offer critical information about the disease state. On the other hand, transgenic animal models, tissue culture experiments and certain electrophysiology studies test causal relationships, yet they all have serious limitations of a different kind: they will never be able to fully recapitulate complex brain disorders in their model systems: they study biological processes in isolation, and not the disease itself. Furthermore, "mechanistic" studies invariably obtain their leads from "descriptive" scientific discoveries. Is a genetic study reporting a higher proportion of CNVs in schizophrenia (Stefansson et al., 2008) or autism (Sebat et al., 2007) descriptive or causal? Neither of these two findings gave direct insight into the mechanism by which the CNVs might produce a disease, so, they must be considered "descriptive". Yet, they discovered a critical process by which these two devastating diseases might arise, and they suggest causality, so they must be considered "mechanistic" at the same time. Thus, separating studies into "descriptive" and "mechanistic" is highly artificial – the information value of the study is critical, and not the classification of the technology that was employed in the studies.

In summary, nowadays we are fully aware of the interpretational challenges associated with high-throughput expression profiling. We hope that we were able to convince you that a carefully designed, executed, and thoughtfully interpreted expression profiling experiment is a valuable source of scientific data, that our field has high standards, and that we are able to generate valuable leads for brain researchers. We expect that the "circumspect" attitude (Journal-of-Neuroscience, 2005) toward our experiments will fade into the sunset, and we deeply believe that transcriptome profiling studies will continue to be essential for deciphering the pathophysiological mechanisms leading to complex brain disorders.

Acknowledgments

We are helpful for Martin J. Schmidt for valuable comments of the manuscript. The work of the Mirnics laboratory is supported by NIH R01 MH067234 (KM) and MH079299 (KM). Training support for Amanda C. Mitchell was provided by NIH T32 MH064913-07.

REFERENCES

Aoki KF, Kanehisa M. Using the KEGG database resource. Curr Protoc Bioinformatics. 2005 Chapter 1, Unit 1 12.

Arakawa K, et al. KEGG-based pathway visualization tool for complex omics data. In Silico Biol. 2005; 5:419–23. [PubMed: 16268787]

Arion D, et al. Molecular evidence for increased expression of genes related to immune and chaperone function in the prefrontal cortex in schizophrenia. Biol Psychiatry. 2007; 62:711–21. [PubMed: 17568569]

Barrett T, et al. NCBI GEO: mining millions of expression profiles--database and tools. Nucleic Acids Res. 2005; 33:D562–6. [PubMed: 15608262]

Brazma A, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet. 2001; 29:365–71. [PubMed: 11726920]

Curtis RK, et al. Pathways to the analysis of microarray data. Trends Biotechnol. 2005; 23:429–35. [PubMed: 15950303]

Dennis G Jr. et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003; 4:P3. [PubMed: 12734009]

Gao X. Construction of null statistics in permutation-based multiple testing for multi-factorial microarray experiments. Bioinformatics. 2006; 22:1486–94. [PubMed: 16574697]

Garbett K, et al. Immune transcriptome alterations in the temporal cortex of subjects with autism. Neurobiol Dis. 2008; 30:303–11. [PubMed: 18378158]

Gentleman RC, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004; 5:R80. [PubMed: 15461798]

Hakak Y, et al. Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. Proc Natl Acad Sci U S A. 2001; 98:4746–51. [PubMed: 11296301]

Harrison PJ, Weinberger DR. Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. Mol Psychiatry. 2005; 10:40–68. image 5. [PubMed: 15263907]

Hashimoto T, et al. Alterations in GABA-related transcriptome in the dorsolateral prefrontal cortex of subjects with schizophrenia. Mol Psychiatry. 2008; 13:147–61. [PubMed: 17471287]

Hollingshead D, et al. Platform influence on DNA microarray data in postmortem brain research. Neurobiol Dis. 2005; 18:649–55. [PubMed: 15755690]

Horvath S, et al. Analyzing schizophrenia by DNA microarrays. Biol Psychiatry. 2010; 69:157–62. [PubMed: 20801428]

Horvath S, Mirnics K. Breaking the gene barrier in schizophrenia. Nat Med. 2009; 15:488–90. [PubMed: 19424205]

Huffaker SJ, et al. A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia. Nat Med. 2009; 15:509–18. [PubMed: 19412172]

Irizarry RA, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. 2003; 4:249–64. [PubMed: 12925520]

Journal-of-Neuroscience. Decision on Journal of Neuroscience, JN-RM-0759-05. In: Mirnics, K., editor. Society of Neuroscience. 2005.

Korade Z, Mirnics K. The autism disconnect. Nature. 2011 in press.

Kuehn H, et al. Using GenePattern for gene expression analysis. Curr Protoc Bioinformatics. 2008 Chapter 7, Unit 7 12.

Lander, E. “Winding Your Way through DNA” Symposium. 2002. http://www.accessexcellence.org/RC/CC/lander.php

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559. [PubMed: 19114008]

Lazarov O, et al. Environmental enrichment reduces Abeta levels and amyloid deposition in transgenic mice. Cell. 2005; 120:701–13. [PubMed: 15766532]

Lockhart DJ, Barlow C. Expressing what's on your mind: DNA arrays and the brain. Nat Rev Neurosci. 2001; 2:63–8. [PubMed: 11253360]

McMahon HT, et al. Synaptophysin, a major synaptic vesicle protein, is not essential for neurotransmitter release. Proc Natl Acad Sci U S A. 1996; 93:4760–4. [PubMed: 8643476]

Middleton FA, et al. Gene expression profiling reveals alterations of specific metabolic pathways in schizophrenia. J Neurosci. 2002; 22:2718–29. [PubMed: 11923437]

Mimmack ML, et al. Gene expression analysis in schizophrenia: reproducible up-regulation of several members of the apolipoprotein L family located in a high-susceptibility locus for schizophrenia on chromosome 22. Proc Natl Acad Sci U S A. 2002; 99:4680–5. [PubMed: 11930015]

Mirnics K. What is in the brain soup? Nat Neurosci. 2008; 11:1237–8. [PubMed: 18956006]

Mirnics K, et al. Critical appraisal of DNA microarrays in psychiatric genomics. Biol Psychiatry. 2006; 60:163–76. [PubMed: 16616896]

Mirnics K, et al. Analysis of complex brain disorders with gene expression microarrays: schizophrenia as a disease of the synapse. Trends Neurosci. 2001; 24:479–86. [PubMed: 11476888]

Mirnics K, et al. Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex. Neuron. 2000; 28:53–67. [PubMed: 11086983]

Mirnics K, Pevsner J. Progress in the use of microarray technology to study the neurobiology of disease. Nat Neurosci. 2004; 7:434–9. [PubMed: 15114354]

Parkinson H, et al. ArrayExpress--a public repository for microarray gene expression data at the EBI. Nucleic Acids Res. 2005; 33:D553–5. [PubMed: 15608260]

Pasinetti GM. Use of cDNA microarray in the search for molecular markers involved in the onset of Alzheimer's disease dementia. J Neurosci Res. 2001; 65:471–6. [PubMed: 11550214]

Pongrac J, et al. Gene expression profiling with DNA microarrays: advancing our understanding of psychiatric disorders. Neurochem Res. 2002; 27:1049–63. [PubMed: 12462404]

Schluter OM, et al. Rabphilin knock-out mice reveal that rabphilin is not required for rab3 function in regulating neurotransmitter release. J Neurosci. 1999; 19:5834–46. [PubMed: 10407024]

Sebat J, et al. Strong association of de novo copy number mutations with autism. Science. 2007; 316:445–9. [PubMed: 17363630]

Sibille E, et al. Gene expression profiling of depression and suicide in human prefrontal cortex. Neuropsychopharmacology. 2004; 29:351–61. [PubMed: 14603265]

Sohn I, et al. A permutation-based multiple testing method for time-course microarray experiments. BMC Bioinformatics. 2009; 10:336. [PubMed: 19832992]

Stefansson H, et al. Large recurrent microdeletions associated with schizophrenia. Nature. 2008; 455:232–6. [PubMed: 18668039]

Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–50. [PubMed: 16199517]

Sullivan PF. The psychiatric GWAS consortium: big science comes to psychiatry. Neuron. 2010; 68:182–6. [PubMed: 20955924]

Tusher VG, et al. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001; 98:5116–21. [PubMed: 11309499]

Unger T, et al. True and false discovery in DNA microarray experiments: transcriptome changes in the hippocampus of presenilin 1 mutant mice. Methods. 2005; 37:261–73. [PubMed: 16308155]

Vogler C, et al. Microarray-based maps of copy-number variant regions in European and sub-Saharan populations. PLoS One. 2010; 5:e15246. [PubMed: 21179565]

Voineagu I, et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. Nature. 2011 this issue.

Winden KD, et al. The organization of the transcriptional network in specific neuronal classes. Mol Syst Biol. 2009; 5:291. [PubMed: 19638972]

Yi M, et al. WholePathwayScope: a comprehensive pathway-based analysis tool for high-throughput data. BMC Bioinformatics. 2006; 7:30. [PubMed: 16423281]