Published in final edited form as:

J Clin Epidemiol. 2010 September ; 63(9): 950–959. doi:10.1016/j.jclinepi.2010.02.018.

EXPECTATIONS, VALIDITY, AND REALITY IN GENE EXPRESSION PROFILING

Kyoungmi Kim^{1,2}, Stanislav O. Zakharkin^{3,*}, and David B Allison⁴

¹Division of Biostatistics, Department of Public Health Sciences, University of California, Davis, CA, USA

²Rowe Program in Human Genetics, University of California, Davis, CA, USA

³Technology and Innovation, Solae, St Louis, MO, USA

⁴Section on Statistical Genetics, Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA

Abstract

Objective: To provide a critical overview of gene expression profiling methodology and discuss areas of future development.

Results: Gene expression profiling has been used extensively in biological research and has resulted in significant advances in the understanding of the molecular mechanisms of complex disorders, including cancer, heart disease, and metabolic disorders. However, translating this technology into genomic medicine for use in diagnosis and prognosis faces many challenges. In addition, gene expression profile analysis is frequently controversial, because its conclusions often lack reproducibility and claims of effective dissemination into translational medicine have, in some cases, been remarkably unjustified. In the last decade, a large number of methodological and technical solutions have been offered to overcome the challenges.

Study Design and Setting: We consider the strengths, limitations, and appropriate applications of gene expression profiling techniques, with particular reference to the clinical relevance.

Conclusion: Some studies have demonstrated the ability and clinical utility of gene expression profiling for use as diagnostic, prognostic, and predictive molecular markers. The challenges of gene expression profiling lie with the standardization of analytic approaches and the evaluation of the clinical merit in broader heterogeneous populations by prospective clinical trials.

Keywords

microarray; gene expression profiling; classification; experimental design; technical challenges; statistical issues

© 2010 Elsevier Inc. All rights reserved.

Correspondence to: Kyoungmi Kim, PhD Division of Biostatistics Department of Public Health Sciences University of California, Davis, CA 95616, USA kmkim@ucdavis.edu.

*Currently at Science and Technology, Cadbury, Whippany, NJ, USA

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

K. Kim and S. O. Zakharkin contributed equally.

INTRODUCTION

Microarray technology has become a widely used tool for genome-wide gene expression profiling where expression levels of thousands of genes are measured at once. The hope is that by analyzing patterns of gene expression (e.g. profiling) scientists will be able to better understand the molecular etiology of multi-factorial disorders such as obesity, diabetes, heart disease, or cancer. Microarray technology offers an opportunity to pinpoint a few genes that may be the “key players” in the observable biological phenomena as well as to view a “big picture” and reveal important multi-gene interactions and understand changes at the level of molecular pathways and networks.

However, major challenges exist. Successful microarray experiment requires proper planning and sound experimental design that accounts for various sources of variability; careful sampling and preparation of biological material; thorough array processing, hybridization, scanning, and image analysis. The application of different analytic approaches to these massive datasets can result in different outcomes. Additionally, comparison of results obtained using different microarray platforms remains a challenge due to various informatics issues. Thus the success of microarray studies depends not only on the quality of experimental design and data but also on the statistical and bioinformatic methods of analysis. Herein we discuss advances in gene expression profiling studies over the last decade and identify areas that need further research. There are many related fields where transcriptional data are used that we do not cover, such as eQTL studies, systems biology or genome-wide association studies (1-3).

EXPECTATIONS OF GENE EXPRESSION PROFILING STUDIES

DNA microarray experiments are typically designed to achieve one or several of the following objectives: (1) to identify individual genes (transcripts) whose expression is correlated with a phenotypic trait such as response to treatment, (2) to identify multiple genes that are interactively involved in regulatory networks and mediating biological phenomena or disease pathogenesis, (3) to discover potential molecular targets for drug development, and (4) to identify molecular markers that can be used as tools for disease diagnosis and prognosis or as predictors of clinical outcomes. Regardless of the purpose of a microarray study, there is one common underlying assumption—the primary end point of the study is associated with the expression levels of a number of genes. For example, Kim et al profiled genes from perioperative adipose tissue whose expression was associated with weight loss outcome following Roux-en-Y gastric bypass (RYGB) surgery (4). Their results showed that genes for glycerolipid syntheses that are directly involved in regulating adipose tissue fat stores were significantly correlated with postoperative body weight loss. More interestingly, the network analysis demonstrated the unique impact of the regulation of genes in lipid metabolism on weight loss. Their gene expression profiling study has limited “generalizability” but is perfectly satisfactory to establish a ‘proof of principle’ and to provide the basis for later confirmatory (usually larger prospective) studies to assess the validity of gene expression profiles in obesity intervention. Clearly, it is possible for microarray studies to yield clinically significant results, but it is also important to recognize that different questions are addressed in different phases of research (feasibility, validation, confirmation, and proof of clinical merit) and that the appropriate design and analysis strategy of a gene expression profiling study should be driven by a specific objective in the planning stage of each phase of research.

CLINICAL AND THERAPEUTIC MERITS OF GENE EXPRESSION PROFILING

Gene expression profiling is changing the approach to discovery of biomarkers in clinical research. Gene expression profiling of disease suggests reliance on the characteristic

genomic ‘signatures’ (groups of genes that can discriminate disease samples from healthy samples) with prognostic and predictive implications in clinical settings rather than on traditional clinical prognostic assessment. For example, investigators from the Netherlands Cancer Institute in Amsterdam constructed a gene expression profile through the analysis of 78 breast tumors, obtaining a ‘70-gene prognostic signature’ whose differential expression had prognostic value in patients with node-negative (N0) breast cancer. The 70-gene signature was capable to classify N0 patients into two groups: good prognosis (without recurrence during a 5-year follow-up) and poor prognosis (recurrence/metastasis in a 5-year follow-up) with high sensitivity and specificity. (5). Subsequently, the investigators confirmed the results with a larger cohort of 234 breast tumors, including stage I-II and N0-N1 disease (6), which established the basis of an ongoing European clinical trial called MINDACT (Microarray In Node negative Disease may Avoid ChemoTherapy) (Buyse M). 307 Patients from five European centers were stratified into high- and low-risk groups based on two prognostic models: the 70-gene signature and traditional clinical prognosis assessment. The two prognostic models were compared in terms of time to distant metastases, disease-free survival, and overall survival in high- versus low-risk groups. These studies suggest that there can indeed be clinical utility of gene expression profiling as a prognostic marker. Other studies have been carried out to evaluate gene expression profiling as a predictive treatment decision model. Gene expression profile of 112 ER-positive breast cancer patients treated with tamoxifen could identify a 44-gene predictive signature whose expression differentiates between hormone-responsive and –resistant carcinomas, being superior to traditional clinical predictive factors (7). Another research concerning the responsiveness to preoperative docetaxel chemotherapy regimen has been carried out analyzing 44 breast cancers. Genes differentially expressed between responders and resistant carcinomas have constructed a 85-gene signature with accuracy >80% in classifying responders and non-responders (8). These studies are further examples suggesting the value of gene expression profiling as prognostic and predictive markers that may help clinical decision making

CHALLENGES OF GENE EXPRESSION PROFILING STUDIES

Gene expression profiling seems to have value in the discovery of molecular markers for potential use in diagnosis or as therapeutic targets. However, translating this technology into genomic medicine is still a work in progress. To better understand strengths and limitations of gene expression profiling techniques, we need to understand biological, technological, statistical, and informatics challenges and caveats.

Biological challenges

Gene expression is dynamic—Biological systems are dynamic and constantly change. A microarray experiment presents a snapshot of a biological system at a given time point. The mRNAs are being synthesized and degraded each at its own rate and the measured levels are the steady state levels. The rates of synthesis and degradation are different; some mRNAs are synthesized more slowly than others but are more stable and accumulate at higher levels. Thus, the presence of mRNA does not explicitly mean that it was just synthesized. Likewise, the inability to detect an unstable transcript may be due to its high degradation rate (9). The expression of some genes (“housekeeping genes”) is thought to be more stable, and these genes are often used as controls for the normalization of expression levels of other genes. However, the expression of such traditionally used controls as ribosomal RNA genes also changes across different tissues and experimental conditions making it difficult to select “gold standards” (10).

Sampling issues—In some cases (e.g. fat tissue obtained during liposuction), the source of biological material is abundant, but in other cases (e.g. acquiring the brain tissue of patients with brain tumor), obtaining the biological material can be quite challenging. The choice of a biopsy method can significantly affect the results as certain fractions of the biological material may be over- or under-represented (11). Often, a few cells from the neighboring tissue may contaminate the sample and seriously bias the results for certain genes. This situation can be further complicated if the organs or tissues are heterogeneous (e.g. liver or brain) and slight variations in sampling locations may result in different expression profiles (12,13). Recent methods like laser microdissection and laser pressure catapulting enable pure and homogeneous sample preparation and may significantly improve the situation (14-16).

RNA quality issues—RNA quality is critical in genome-wide analysis of gene expression. RNA is less stable than DNA and care should be taken and adequate protocols followed to preserve the quality of biological material. This is particularly important in clinical setting. Preservation of frozen tissue is fairly common in many pathology laboratories when molecular analysis is required. In most oncology contexts, labs rely on the formalin-fixed, paraffin-embedded technique (FFPE) as a standard method of preservation (17). But several studies have shown that freezing in liquid nitrogen preserve quality better than FFPE (18-20). Interestingly, RT-PCR techniques were shown to be capable of amplifying mRNA even from the degraded FFPE tissues (21,22) though practical considerations have limited these assays to a few selected genes (23).

Microarrays are just a part of the picture—One of the goals of a typical microarray experiment is to get a better understanding of molecular mechanisms involved in the biological phenomenon or in the development of a disease. But, while changes in transcriptional profiles are important, they are just a part of the big picture. Many biologically important changes do not necessarily manifest themselves in alterations of the RNA levels. Most of the cellular functions are performed by proteins and physiological changes can be modulated by not only changes in protein levels but also by protein modifications such as glycosylation, methylation, acetylation, and phosphorylation. These modifications could change protein conformation and lead to changes in activity. The classic example of the importance of protein conformation is misfolding of prion proteins. Changes in conformation from the alpha helixes to beta-sheets in proteins with exactly the same sequence lead to extracellular protein aggregation and forming plaques resulting in brain damage and death (24). The process replicates itself without involvement of nucleic acids.

Technological challenges

Microarray platforms—As the quality of microarray chips continues to improve, costs continue to drop, and methods of analysis and experimental designs continue to standardize, microarrays may well have a transforming effect on biomedical research. Several microarray platforms available on the market today employ different hybridization strategies: synthesized oligonucleotide arrays (Affymetrix), spotted oligonucleotides (Agilent), beads-based (Illumina), and custom spotted arrays. Typically, oligonucleotide arrays are single-channeled, meaning that only one dye is used. In oligonucleotide arrays, an individual RNA sample hybridizes to a single array. Traditionally, oligonucleotides close to 3' end were used, from one to several per gene. The cDNA arrays are typically, dual-channeled, meaning that cDNA specimens from two different samples—each representing a different condition or treatment—are labeled with two different fluorescent dyes such as Cy5 (red) and Cy3 (green). The cDNA labeled with Cy5 and Cy3 are mixed together and cohybridized against the same array. A ratio rather than the two individual fluorescent intensities at each spot is used to profile gene expression patterns. Despite the high variability in gene expression

attributed to differences in microarray platforms, studies have demonstrated that reproducibility across platforms can be dramatically improved when standardized protocols are implemented for RNA labeling, hybridization, data processing, data acquisition, and data normalization. When these technical variables are standardized, different microarray platforms can produce comparable outcomes (25,26). Nevertheless, the results from comparison across different platforms can be misleading and should be interpreted with great caution (27).

Binding efficiency—Gene expression level is determined by measuring the amount of labeled target bound to the respective oligonucleotide probe. Binding affinity depends on the oligonucleotide length and composition as well as hybridization conditions such as temperature, concentration of ions, oligonucleotide conformations, etc. (28,29). The selection of probes for different genes in a way to ensure that all probes have similar binding properties is quite challenging (30,31). Consequently, assessments of changes in transcription level rather than of absolute measures of probe are more meaningful.

Sources of technical variation—Gene expression profiles can be affected by many different technical factors during experiment. Technical variation reflects changes in experimental conditions during conduct and data processing which can significantly impact the quality of data and introduce bias to the results. Identification of sources of variation and assessment of their potential influence are important to ensure highly reproducible microarray data. Thus technical variability should be minimized in the planning of experiments by controlling the quality of the RNA samples and by efficient and uniform data processing procedures (32). With regards to prospective experiments, the uniformity of experimental conduct will help to minimize potential bias and thus improve the validity of a study. Additionally, the choice of an image processing algorithm, such as dChip (33), MAS 5.0 (34), RMA (35), etc, can result in different variance components estimates and thus impact on the results of microarray data analysis (36). Currently, there is no consensus about what is the best way of processing images to quantify levels of gene expression.

Statistical and bioinformatic challenges

Design of microarray studies—When a microarray study is proposed, it should have a clear goal and a specific hypothesis to test. In the design of a microarray experiment, all potential sources of variation should be taken into account to avoid any systematic bias. When designing a microarray study, researchers should adhere to the sound principles of study and match the experimental variables of cases and controls to the fullest extent possible, selecting biologically homogenous sample populations, balancing a design with respect to all factors that can confound results among the comparison groups, and handling samples uniformly through the course of the entire experiment. Randomization will assure baseline equality between the groups being compared. Violation of these principles will lead to biased results and can cause a loss in power. It should be pointed out that statistical analysis of data cannot solve fundamental problems of study design. Significantly, the validity of gene expression profiles depends on the characteristics of samples and selection bias, regarding all eligible criteria of participation and other confounding factors.

As with all such studies, an adequate sample size is necessary to achieve sufficient power to demonstrate significance of findings, and this is particularly the case with microarray studies where thousands of genes are tested simultaneously. Determination of an adequate sample size is affected by many different factors of variability which are attributed to biological causes as well as technical sources as previously described (37). Biological variation results from heterogeneity among samples due to disease, sex, age, race, gene, gene-environment interaction and other confounding factors. For example, when inter-individual variability is

the major contributor to variation in data, it is necessary to increase the number of independent biological samples rather than performing technical replicate arrays using specimens from a small number of biological samples. By contrast, if inter-individual variability is relatively small for experiments involving inbred strains of model species which have less inherent biological variation than outbred populations, then technical variation during data processing will likely have a bigger impact on data quality. In this case, replicate arrays of the same RNA sample may help to measure and improve the sensitivity of the study. Also, the sample size required for a study depends in part on methods of analysis. For example, when microarrays are used for the identification of classifiers (e.g., transcripts or genes) that predict a sample's class membership, the prediction accuracy for internal validity of the study needs to be evaluated against an independent separate test set that was not used in the model construction. Because a training and test set have different goals, the sample size for each set should differ. The size of the test set increases as the number of selected classifiers that need to be evaluated increases; this principle urges researchers to select a small number of classifiers that are most informative in classification.

Normalization and transformation—Appropriate preprocessing of microarray data prior to analysis is critical for identifying differentially expressed genes. Normalization attempts to remove variability among chips and other systematic biases that are unrelated to biological variation so that a meaningful biological comparison can be made. Transformation is used for multiple purposes, including stabilizing variance in data so that underlying assumptions required for the statistical analysis method are met. Approaches to normalizing expression levels range from simpler methods such as global median normalization (38), to complex methods including linear or non-linear intensity dependent normalization (39,40), to rank invariant methods (41), and others. Similarly, a wide range of transformation methods can be used to stabilize variance. Because variance in microarray data tends to rise with intensities, the most common transformation is the logarithm that reduces the effect of extreme values. Other sophisticated methods have been developed (42-45). For example, Durbin and Rocke presented three variance-stabilizing transformations for microarray data, including the generalized-log, started-log and log-linear hybrid transformation (45). They showed that each of these transformations appears to stabilize the variance of transformed data and provides better variance stabilization than the log transformation. Although it is expected that the choice of a preprocessing procedure does not affect the core results of microarray data, different normalization and/or transformation methods may result in different outcomes (46).

Application of appropriate analytic methods—Various statistical data analysis methods for gene expression profiling have been developed. Commonly used statistical methods can be classified as differential analysis (such as t-test, analysis of variance (ANOVA)), supervised classification methods (linear discriminant analysis), dimension reduction (principal component analysis), and unsupervised cluster analysis (hierarchical clustering, K-means clustering). Differential analysis methods are statistical tests for identifying differentially expressed genes by determining differences in mean values between different groups or between groups in different conditions (refer to as group comparison (47)). Researchers often rely on an alternative measure, fold change, for the selection of genes differentially expressed between the two groups. Fold change is the ratio of average expression levels of two groups, assuming no variability exists across measured gene expression levels. However, this assumption is unrealistic because the ability of a gene (probe) to bind to a unique target sequence varies and results in a signal (measured expression level) even though all genes are quantified under the uniform condition. Hence, fold change may be a useful tool for ranking genes on the basis of effect sizes, but it is not a

valid statistical test for significance (48). Valid measures for testing include, for example, the t-statistic (or F-statistic of ANOVA) that considers the underlying variability in data.

Classification and cluster analysis are analytical approaches that attempt to divide data into classes. As a supervised method of class prediction (47), classification analysis differs from cluster analysis because it uses information on predefined classes. In studies of class prediction which distinguish samples in one group from samples in the other group, a perfect discriminatory classifier can be found by chance, especially when assessing a large number of possible predictors—this problem is referred to ‘overfitting.’ Overfitting occurs when a classifier is made to perfectly fit a set of data that was used in the model development, but has no discriminatory power so that the results cannot be reproduced in a set of completely independent samples. By contrast, cluster analysis is referred to as a class discovery method because no classes are predefined, and the procedure groups genes or samples into arbitrary classes based on their similarity/dissimilarity in expression regardless of their true class membership. Although clusters are rarely meaningful in themselves, cluster analysis is useful for a general description of gene expression patterns across the genome and may help investigators in formulating specific questions for future research. It should be mentioned that while there is no “best” clustering algorithm and different approaches have their merits, some methods produce more stable results. Clustering results may be quite sensitive to the choice of linkage distance and an algorithm as well as to noise in the data and sample size (49). Some evidence suggests that divisive hierarchical clustering and methods involving resampling and bootstrapping tend to produce more stable clusters (50,51).

Multiple testing—Because analyses of genome-wide gene expression profiles involve multiple comparisons, they are typically accompanied by a multiple testing correction such as Benjamini-Hochberg’s false discovery rate (52), Westfall-Young step-down permutation correction (53), or the Bonferroni correction. The goal of multiple testing correction is to minimize the chance of false positives. The general consensus is that the selection of differentially expressed genes should be based on the false discovery rate (48).

Validation of results—The most important test for developing a predictor or classifier is that the proposed classifier be able to predict class membership for new samples whose class membership is not known. Hence it is important to estimate the accuracy of class prediction for future blind samples. The most commonly used approach for properly estimating the accuracy of a class predictor is to develop models using a training dataset and then to test the resulting model on an independent test set that was not used in the model construction. The independent test set should not be used in any way for the development of the prediction model, for estimating the parameters of the model, or, importantly, for selecting the specific predictors. In other words, the test set should be used *only* to ascertain the prediction accuracy of the model which was developed using *only* the training set. Feature (gene) ranking, feature selection, together with any parameter estimation are all part of the modeling and should be tested within validation. If the predicted class of a sample in the test set does not agree with the true class of that sample, a prediction error has occurred.

In cases where data sets are too small or independent test sets are not available, cross-validation is an alternative approach. Cross-validation systematically splits the given samples into a training set and a test set. The test set is the set of future samples for which class labels are to be determined. It is set aside until a specified predictor has been developed using only the training set. The process is *repeated* a number of times—each time creates a different training set and its complementary test set—and the performance scores are averaged over all splits. It is important to cross-validate all steps of predictor construction in estimating the prediction error. Incorrect use of cross-validation can result in a seriously biased underestimate of the prediction error for microarray data when there are

thousands of potential predictors (54). A second useful cross-validation technique is multiple random validation in which the samples are randomly split into a training and test set (55). Michiels et al demonstrated that the feature selection strongly depended on the selection of samples in the training set because every training set of samples led to different outcomes (e.g., a different set of genes), suggesting that selection bias can be problematic (55). Nonetheless, they showed that the misclassification rate could be improved with increased sample sizes of training set. An adequate sample size is essential for any cross-validation technique to be effective.

Gene regulatory networks—Genes do not act in isolation; changes in their activity are tightly regulated and influenced by other genes. One way to summarize such overwhelming information is to use a network, consisting of the nodes (e.g. genes) connected by the edges (lines or arrows) where edges indicate dependencies and possibly functional relations among nodes (56). One significant challenge for researchers is to reconstruct network structure from available expression data. Many different methods for network inference have been proposed (57). A common problem of such models is exponential complexity: the number of parameters increases exponentially with the number of variables. Thus, many alternative and equally probable network structures may be constructed from a given dataset. Successful reconstruction of biological networks requires a systematic source of perturbation such as genetic variations or environmental changes. Integration of genetic and gene expression data has been successfully applied to identify a number of novel disease-related genes (58).

Informatics challenges

Data management is a critical issue in any large scale experimental project, no matter if it is a multicenter clinical trial or a metabolomic, proteomic or transcriptomic experiment (59). Microarray experiments share several issues similar to those in other “omics” fields, and they also present their own unique challenges. Some of the challenges that seemed almost insurmountable at the dawn of microarray era are now being resolved while others remain critical (60).

Different gene nomenclatures—One of the critical issues in microarray data analysis has been the reproducibility of results. There are multiple objective biological, technical, and analytical reasons for this problem, but adding to the challenge is the fact that studies can be difficult to compare because they report results using different gene nomenclatures, such as Genbank, Locuslink, EMBL, RefSeq and Affymetrix gene IDs. Significant efforts have been put forward by the bioinformatics community to standardize nomenclature and to develop translational tools. Currently, several software packages and websites (e.g. DAVID (61) and GoMiner (62)) are available for effective high-throughput cross-conversion of gene ID types.

Different probes for the same gene—Different microarray platforms may use different probes to interrogate the same transcript because different bioinformatics algorithms are used for probe selection (copyright restrictions further complicate this problem) (63-66). As a result, hybridization efficiency may be different and may lead to observed differences in expression levels even if the actual mRNA levels were identical.

Common quality standard—To prevent common procedural failures and to establish quality tools for the microarray community, several institutions—including six FDA centers, commercial array and reagent providers, NIH, EPA, NIST, and several academic laboratories—established the Microarray Quality Control (MAQC) project in 2005. The project encompasses all the major microarray platforms and generates data from over 100 microarrays and quantitative PCR validation reactions. The first phase of the project,

MAQC-I, aimed to develop procedural guidelines and quality control metrics (67). The whole September 2006 issue of *Nature Biotechnology* describes the initiative and shows examples. The second phase, MAQC-II, involves thirty six teams and aims to evaluate various data analysis methods and predictive models.

Common data format—One of the serious problems historically has been a wide diversity of data formats used in microarray experiments. As a result, the Microarray Gene Expression Database Society (MGED) was created in 1999 to develop a common standard for data input and reporting that could be shared among scientists in the microarray field. In 2001 the MGED in turn created the Minimum Information About a Microarray Experiment (MIAME) guidelines which serve as a template for researchers to report an adequate description of how microarray data were obtained. The MIAME has six essential elements (68): (1) Experimental design; (2) Type of array and description of each element on the array; (3) Description of samples and their preparation and labeling procedures; (4) Hybridization procedures and parameters; (5) Images and description of measurement specifications; (6) Description of normalization procedures. Currently, many journals require MIAME compliance as a part of the manuscript submitting procedure.

Repositories of microarray data—The largest publicly available repository of microarray and other high-throughput functional genomic data is the Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI) (69). GEO is MIAME-compliant and offers tools to effectively explore, analyze, and download expression data from both gene-centric and experiment-centric perspectives. Among other repositories are Stanford Microarray Database (70), ArrayExpress (71), CIBEX (72), and many other specialized databases (73-75).

FUTURE DIRECTIONS

What is realistic?

Integration of microarray studies—Genomic scale profiling of gene expression has become a routine method in various areas of biological research. Of major concern is the reproducibility and validity of published results that complicate comparisons of results from multiple studies (76). Validating results is difficult or almost impossible—related studies often lead to contradictory results, even if they investigate the same biological phenomena. Inter-experiment variation can also affect the fundamental comparison of results. There are a number of sources of variation that affect results. Technological differences in platforms, including differences in probe sizes and composition, the number of probe sets, and the total number of probes per array, may bias results. In addition, biological, experimental, and technical variations between studies create even more substantial differences in outcomes. More importantly, it has been demonstrated that different methods of analysis can result in very different outcomes from the same data (77,78).

Despite the common discordance of microarray study results, the integration of somewhat contradictory findings can provide important information. There are two general approaches to integrating multiple microarray studies: meta-analysis of merged raw data or meta-review analysis of the published results (i.e., the published gene lists which resulted from the analysis of the raw data). Although meta-analysis of merged raw datasets from similar microarray studies is most desirable, there are limitations to the approach. Meaningful raw data that can be used in meta-analysis are simply not accessible for most microarray studies. An alternative approach for ranking genes is a meta-review method that employs a vote-counting strategy in an analysis of published evidence (79). The meta-review approach is built on the assumption that meta-genes or aggregate patterns of gene expression that are consistently expressed in multiple independent studies are more likely to share a common

biological process, when experimental-specific spurious genes are not overlapped between the different studies and their biological relevance is minimal. The potential of meta-analysis can improve when all prospective microarray studies are considered in a framework that integrates diverse approaches taken from different studies.

Integration of different sources of biological information—During the last decade, bioinformatics community has witnessed explosive growth of the information and available tools. The January 2009 issue of *Nucleic Acids Research* includes descriptions of 179 biological databases, of which 95 are new. These databases (along with several molecular biology databases described in other journals) have been included in the *Nucleic Acids Research* online Molecular Biology Database Collection, bringing the total number of databases in the collection to 1170 (80). The question is: how can we handle and make use of all this overwhelming information? Certainly, this is a big challenge and variety of approaches will be used to find a solution. Researchers at the Institute for Systems Biology have developed a system to integrate data seamlessly from different sources (81,82). In another example, SIGMA2 software allows identification of genes that have simultaneous changes in number of copies, loss of heterozygosity, DNA methylation and altered expression profiles (83). The SMART extension to MIAME, Standard MicroArray Reporting Template, is based on a semantically consistent markup language for communicating MIAME data, called MAGE-ML (84). It includes extensions for adequate recording of gene identifiers relevant for microarray studies allowing different databases to exchange information with minimal reformatting. With the significant advances and promises offered by the transcriptomic approach to biological science and medicine, today systematic and integrated approaches are becoming needed for the integration of heterogeneous data drawn from emerging studies in genetics, genomics, proteomics, and metabolomics.

What is not realistic or too optimistic in the near future?

Widespread replacement of existing clinical tests as predictors of clinical outcomes—Traditional tests in clinical practice are not always adequate. For example, brain tumors are one of the most difficult to cure. Gliomas have several grades that have different survival prognoses but that are difficult to distinguish morphologically (85,86). Gene expression profiling studies would substantially help researchers to elucidate molecular pathways implicated in the development of these devastating tumors. However, despite the progress in genomic methods in gene signature discovery and characterization, criticism has emerged against predictive models based on gene expression data. Researchers point out the complex relationships among genetic variation, the environment, and disease, and the limited clinical validity and utility of genetic risk prediction in translational personalized and genomic medicine (87). Quite often, claims are prematurely optimistic and may lead to unrealistic expectations. In some cases, simple errors in data processing resulted in erroneous models (88,89). Ein-Dor et al used a theoretical analysis to show that thousands of patients may be needed to obtain reliable gene lists (90). Not only genetics but environment and lifestyle may have significant contribution to many complex diseases. The integration of genomics with clinical and environmental data, joining “nature and nurture,” (91) will hopefully improve the robustness and practical impact of gene expression profiling studies.

What is the take-home message?

Many studies have suggested the ability and clinical utility of gene expression profiling for use as diagnostic, prognostic, and predictive molecular markers. The challenges of gene expression profiling lie with the standardization of analytic approaches, replication, attaining

adequate sample size, and the evaluation of the clinical utility in broader heterogeneous populations by prospective clinical trials.

Acknowledgments

This work was supported in part by NIH/NIA P01AG025532 (KK) and by NIH/NIDDK P30DK056336 (DBA).

References

1. Baliga NS. Systems biology. The scale of prediction. *Science* Jun 6;2008 320(5881):1297–8. [PubMed: 18535232]
2. Gilad Y, Rifkin SA, Pritchard JK. Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet* Aug;2008 24(8):408–15. [PubMed: 18597885]
3. Manolio TA, Collins FS. The HapMap and genome-wide association studies in diagnosis and therapy. *Annu Rev Med* 2009;60:443–56. [PubMed: 19630580]
4. Kim K, Perroud B, Espinal G, Kachinskas D, Austrheim-Smith I, Wolfe BM, Warden CH. Genes and networks expressed in perioperative omental adipose tissue are correlated with weight loss from Roux-en-Y gastric bypass. *Int J Obes (Lond)* 2008;32(9):1395–406. [PubMed: 18645576]
5. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530–536. [PubMed: 11823860]
6. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999–2009.
7. Jansen MP, Foekens JA, van Staveren IL, et al. Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. *J Clin Oncol* 2005;23:732–740. [PubMed: 15681518]
8. Iwao-Koizumi K, Matoba R, Ueno N, et al. Prediction of doctaxel response in human breast cancer by gene expression profiling. *J Clin Oncol* 2005;23:422–431. [PubMed: 15659489]
9. Cheadle C, Fan J, Cho-Chung YS, Werner T, Ray J, Do L, Gorospe M, Becker KG. Stability regulation of mRNA and the control of gene expression. *Ann N Y Acad Sci Nov;2005* 1058:196–204. [PubMed: 16394137]
10. Thorrez L, Van Deun K, Tranchevent LC, Van Lommel L, Engelen K, Marchal K, Moreau Y, Van Mechelen I, Schuit F. Using ribosomal protein genes as reference: a tale of caution. *PLoS ONE* Mar 26;2008 3(3):e1854. [PubMed: 18365009]
11. Mutch DM, Tordjman J, Pelloux V, Hanczar B, Henegar C, Poitou C, Veyrie N, Zucker JD, Clément K. Needle and surgical biopsy techniques differentially affect adipose tissue gene expression profiles. *Am J Clin Nutr* Jan;2009 89(1):51–7. [PubMed: 19056587]
12. Karsten SL, Kudo LC, Geschwind DH. Gene expression analysis of neural cells and tissues using DNA microarrays. *Curr Protoc Neurosci*. Oct;2008 Chapter 4:Unit 4.28.
13. Malarkey DE, Johnson K, Ryan L, Boorman G, Maronpot RR. New insights into functional aspects of liver morphology. *Toxicol Pathol* 2005;33(1):27–34. [PubMed: 15805053]
14. Niyaz Y, Stich M, Sägmüller B, Burgemeister R, Friedemann G, Sauer U, Gangnus R, Schütze K. Noncontact laser microdissection and pressure catapulting: sample preparation for genomic, transcriptomic, and proteomic analysis. *Methods Mol Med* 2005;114:1–24. [PubMed: 16156095]
15. Bernard R, Kerman IA, Meng F, Evans SJ, Amrein I, Jones EG, Bunney WE, Akil H, Watson SJ, Thompson RC. Gene expression profiling of neurochemically defined regions of the human brain by in situ hybridization-guided laser capture microdissection. *J Neurosci Methods* Mar 30;2009 178(1):46–54. [PubMed: 19070632]
16. Ma LJ, Li W, Zhang X, Huang DH, Zhang H, Xiao JY, Tian YQ. Differential gene expression profiling of laryngeal squamous cell carcinoma by laser capture microdissection and complementary DNA microarrays. *Arch Med Res* Feb;2009 40(2):114–23. [PubMed: 19237021]
17. Medeiros F, Rigl CT, Anderson GG, Becker SH, Halling KC. Tissue handling for genome-wide expression analysis: a review of the issues, evidence, and opportunities. *Arch Pathol Lab Med* Dec;2007 131(12):1805–16. [PubMed: 18081440]

18. Coudry RA, Meireles SI, Stoyanova R, et al. Successful application of microarray technology to microdissected formalin-fixed, paraffin-embedded tissue. *J Mol Diagn* 2007;9:70–79. [PubMed: 17251338]
19. D'Orazio D, Stumm M, Sieber C. Accurate gene expression measurement in formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol* 2002;160:383–384. [PubMed: 11786432]
20. Penland SK, Keku TO, Torrice C, et al. RNA expression analysis of formalin-fixed paraffin-embedded tumors. *Lab Invest* 2007;87:383–391. [PubMed: 17297435]
21. Paik S, Kim CY, Song YK, Kim WS. Technology insight: application of molecular techniques to formalin-fixed paraffin-embedded tissues from breast cancer. *Nat Clin Pract Oncol* 2005;2:246–254. [PubMed: 16264960]
22. Specht K, Richter T, Muller U, Walch A, Werner M, Hofler H. Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol* 2001;158:419–429. [PubMed: 11159180]
23. Popovici V, Goldstein DR, Antonov J, Jaggi R, Delorenzi M, Wirapati P. Selecting control genes for RT-QPCR using public microarray data. *BMC Bioinformatics* 2009;10:42. [PubMed: 19187545]
24. Moor RA, Taubner LM, Priola SA. Prion protein misfolding and disease. *Curr Opin Struct Biol* Feb;2009 19(1):14–22. [PubMed: 19157856]
25. Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. *Nat Methods* 2005;2(5):337–344. [PubMed: 15846360]
26. Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, Frank BC, Gabrielson E, Garcia JG, Geoghegan J, Germino G, Griffin C, Hilmer SC, Hoffman E, Jedlicka AE, Kawasaki E, et al. Multiple-laboratory comparison of microarray platforms. *Nat Methods* 2005;2(5):345–350. [PubMed: 15846361]
27. Toxicogenomics Research Consortium. Standardizing global gene expression analysis between laboratories and across platforms. *Nat Methods* 2005;2(5):351–356.
28. Bishop J, Blair S, Chagovetz AM. A competitive kinetic model of nucleic acid surface hybridization in the presence of point mutants. *Biophys J* Feb 1;2006 90(3):831–40. [PubMed: 16284267]
29. Singh R, Nitsche J, Andreadis ST. An integrated reaction-transport model for DNA surface hybridization: implications for DNA microarrays. *Ann Biomed Eng* Jan;2009 37(1):255–69. [PubMed: 18941894]
30. Li X, He Z, Zhou J. Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation. *Nucleic Acids Res* Oct 24;2005 33(19):6114–23. [PubMed: 16246912]
31. Wernersson R, Juncker AS, Nielsen HB. Probe selection for DNA microarrays using OligoWiz. *Nat Protoc* 2007;2(11):2677–91. [PubMed: 18007603]
32. Dumur CI, Nasim S, Best AM, Archer KJ, Ladd AC, Mas VR, Wilkinson DS, Garrett CT, Ferreira-Gonzalez A. Evaluation of quality-control criteria for microarray gene expression analysis. *Clin Chem* 2004;50(11):1994–2002. [PubMed: 15364885]
33. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci* 2001;98:31–36. [PubMed: 11134512]
34. Hubbell E, Liu WM, Mei R. Robust estimators for expression analysis. *Bioinformatics* 2002;28:2585–92.
35. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4(2):249–264. [PubMed: 12925520]
36. Shedden K, Chen W, Kuick R, Ghosh D, Macdonald J, Cho KR, Giordano TJ, Gruber SB, Fearon ER, Taylor JM, Hanash S. Comparison of seven methods for producing Affymetrix expression scores based on False Discovery Rates in disease profiling data. *BMC Bioinformatics* 2005;6:26. [PubMed: 15705192]
37. Zakharkin SO, Kim K, Mehta T, Chen L, Barnes S, Scheirer KE, Parrish RS, Allison DB, Page GP. Sources of variation in Affymetrix microarray experiments. *BMC Bioinformatics* 2005;6:214. [PubMed: 16124883]

38. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* Feb 15;2002 30(4):e1. [PubMed: 11788727]
39. Kepler TB, Crosby L, Morgan KT. Normalization and analysis of DNA microarray data by self-consistency and local regression. *Genome Biol* Jun 28;2002 3(7) RESEARCH0037. Epub 2002 Jun 28.
40. Edwards D. Non-linear normalization and background correction in one-channel cDNA microarray studies. *Bioinformatics* May 1;2003 19(7):825–33. [PubMed: 12724292]
41. Tseng GC, Oh MK, Rohlin L, Liao JC, Wong WH. Issues in cDNA micarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res* 2001;29:2549–2557. [PubMed: 11410663]
42. Rocke DM, Durbin B. A model for measurement error for gene expression arrays. *J. Comp. Biol* 2001;8:557–569.
43. Huber W, Von Heydebrek A, Sültmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 2002;18:S96–104. [PubMed: 12169536]
44. Geller SC, Gregg JP, Hagerman P, Rocke D. Transformation and normalization of oligonucleotide microarray data. *Bioinformatics* 2003;19:1817–1823. [PubMed: 14512353]
45. Durbin B, Rocke D. Variance-stabilizing transformations for two-color microarrays. *Bioinformatics* 2004;20:660–667. [PubMed: 15033873]
46. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* Jan 22;2003 19(2): 185–93. [PubMed: 12538238]
47. Simon R. Diagnostic and prognostic prediction using gene expression profiles in high-dimensional microarray data. *British Journal of Cancer* 2003;89:1599–1604. [PubMed: 14583755]
48. Allison DB, Cui C, Page PG, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genetics* 2005;7:55–65. [PubMed: 16369572]
49. Garge NR, Page GP, Sprague AP, Gorman BS, Allison DB. Reproducible clusters from microarray research: whither? *BMC Bioinformatics*. 2005
50. Datta S, Datta S. Comparisons and validation of statistical clustering techniques for microarray gene expression data. *Bioinformatics*. 2003
51. Suzuki R, Shimodaira H. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* Jun 15;2006 22(12):1540–2. [PubMed: 16595560]
52. Benjamini Y, Hochberg Y. Controlling the false discovery data- a practical and powerful approach to multiple tests. *J. R. Stat. Soc. Ser. B* 1995;57:289–300.
53. Westfall, PH.; Young, SS. Resampling-based multiple testing: examples and methods for multiple p-value adjustment. John Wiley & Sons; New York: 1993.
54. Ambrose C, McLachlan GJ. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc Natl Acad Sci U S A* May 14;2002 99(10):6562–6. Epub 2002 Apr 30. [PubMed: 11983868]
55. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;365:488–92. [PubMed: 15705458]
56. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet* Feb;2004 5(2):101–13. [PubMed: 14735121]
57. Lee WP, Tzou WS. Computational methods for discovering gene networks from expression data. *Brief Bioinform* Jul;2009 10(4):408–23. [PubMed: 19505889]
58. Schadt EE, Lamb J, Yang X, Zhu J, Edwards S, Guhathakurta D, Sieberts SK, Monks S, Reitman M, Zhang C, Lum PY, Leonardson A, Thieringer R, Metzger JM, Yang L, Castle J, Zhu H, Kash SF, Drake TA, Sachs A, Lusk AJ. An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genet* Jul;2005 37(7):710–7. [PubMed: 15965475]
59. Haquin S, Oeuillet E, Pajon A, Harris M, Jones AT, van Tilbeurgh H, Markley JL, Zolnai Z, Poupon A. Data management in structural genomics: an overview. *Methods Mol Biol* 2008;426:49–79. [PubMed: 18542857]

60. Kawasaki ES. The end of the microarray Tower of Babel: will universal standards lead the way? *J Biomol Tech* 2006;17(3):200–6. [PubMed: 16870711]
61. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003;4(5):P3. [PubMed: 12734009]
62. Zeeberg BR, Feng W, Wang G, Wang MD, Fojo AT, Sunshine M, Narasimhan S, Kane DW, Reinhold WC, Lababidi S, Bussey KJ, Riss J, Barrett JC, Weinstein JN. GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol* 2003;4(4):R28. [PubMed: 12702209]
63. Mecham BH, Klus GT, Strovel J, Augustus M, Byrne D, Bozso P, Wetmore DZ, Mariani TJ, Kohane IS, Szallasi Z. Sequence-matched probes produce increased cross-platform consistency and more reproducible biological results in microarray-based gene expression measurements. *Nucleic Acids Res* May 25;2004 32(9):e74. [PubMed: 15161944]
64. Verdugo RA, Medrano JF. Comparison of gene coverage of mouse oligonucleotide microarray platforms. *BMC Genomics* Mar 21;2006 7:58. [PubMed: 16551360]
65. Carter SL, Eklund AC, Mecham BH, Kohane IS, Szallasi Z. Redefinition of Affymetrix probe sets by sequence overlap with cDNA microarray probes reduces cross-platform inconsistencies in cancer-associated gene expression measurements. *BMC Bioinformatics* Apr 25;2005 6:107. [PubMed: 15850491]
66. Rouse R, Hardiman G. Microarray technology--an intellectual property retrospective. *Pharmacogenomics* Sep;2003 4(5):623–32. [PubMed: 12943468]
67. MAQC Consortium. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* Sep;2006 24(9):1151–61. [PubMed: 16964229]
68. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* Dec;2001 29(4):365–71. [PubMed: 11726920]
69. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muetter RN, Edgar R. NCBI GEO: archive for high-throughput functional genomic data. *Nucleic Acids Res*. Oct 21;2008 Epub ahead of print.
70. Demeter J, Beauheim C, Gollub J, Hernandez-Boussard T, Jin H, Maier D, Matese JC, Nitzberg M, Wymore F, Zachariah ZK, Brown PO, Sherlock G, Ball CA. The Stanford Microarray Database: implementation of new analysis tools and open source release of software. *Nucleic Acids Res* Jan 1;2007 35:D766–770. [PubMed: 17182626]
71. Parkinson H, Kapushesky M, Shojatalab M, Abeygunawardena N, Coulson R, Farne A, Holloway E, Kolesnykov N, Lilja P, Lukk M, Mani R, Rayner T, Sharma A, William E, Sarkans U, Brazma A. ArrayExpress—a public database of microarray experiments and gene expression profiles. *Nucl. Acids Res* 2007;35:D747–D750. [PubMed: 17132828]
72. Ikeo K, Ishi-i J, Tamura T, Gojobori T, Tateno Y. CIBEX: Center for information biology gene expression database. *C R Biol* 2003;326:1079–1082. [PubMed: 14744116]
73. Reddy TB, Riley R, Wymore F, Montgomery P, DeCaprio D, Engels R, Gellesch M, Hubble J, Jen D, Jin H, Koehrsen M, Larson L, Mao M, Nitzberg M, Sisk P, Stolte C, Weiner B, White J, Zachariah ZK, Sherlock G, Galagan JE, Ball CA, Schoolnik GK. TB database: an integrated platform for tuberculosis research. *Nucleic Acids Res* Jan;2009 37(Database issue):D499–508. [PubMed: 18835847]
74. Smith CM, Finger JH, Hayamizu TF, McCright IJ, Eppig JT, Kadin JA, Richardson JE, Ringwald M. The mouse Gene Expression Database (GXD): 2007 update. *Nucleic Acids Res* 2007;35:D618–D623. [PubMed: 17130151]
75. Cheng KC, Strömvik MV. SoyXpress: a database for exploring the soybean transcriptome. *BMC Genomics* Aug 1;2008 9:368. [PubMed: 18671881]

76. Ioannidis JP, Allison DB, Ball CA, Coulibaly I, Cui X, et al. Repeatability of published microarray gene expression analyses. *Nat Genet* Feb;2009 41(2):149–55. [PubMed: 19174838]
77. Harr B, Schlotterer C. Comparison of algorithms for the analysis of Affymetrix microarray data as evaluated by co-expression of genes in known operons. *Nucleic Acids Res* 2006;34:e8. [PubMed: 16432259]
78. Song S, Black MA. Microarray-based gene set analysis: a comparison of current methods. *BMC Bioinformatics* 2008;9:502. [PubMed: 19038052]
79. Griffith OL, Melck A, Jones SJ, Wiseman SM. Meta-analysis and meta-review of thyroid cancer gene expression profiling studies identifies important diagnostic biomarkers. *J Clin Oncol* 2006;24(31):5043–51. [PubMed: 17075124]
80. Galperin MY, Cochrane GR. Nucleic Acids Research annual Database Issue and the NAR online Molecular Biology Database Collection in 2009. *Nucleic Acids Research* 2009;37(Database issue):D1–D4. [PubMed: 19033364]
81. Nykter M, Lähdesmäki H, Rust A, Thorsson V, Shmulevich I. A data integration framework for prediction of transcription factor targets. *Ann N Y Acad Sci* Mar;2009 1158:205–14. [PubMed: 19348642]
82. Boyle J, Rovira H, Cavnor C, Burdick D, Killcoyne S, Shmulevich I. Adaptable data management for systems biology investigations. *BMC Bioinformatics* Mar 6;2009 10(1):79. [PubMed: 19265554]
83. Chari R, Coe BP, Wedseltoft C, Benetti M, Wilson IM, Vucic EA, MacAulay C, Ng RT, Lam WL. SIGMA2: a system for the integrative genomic multi-dimensional analysis of cancer genomes, epigenomes, and transcriptomes. *BMC Bioinformatics* Oct 7;2008 9:422. [PubMed: 18840289]
84. Ball CA, Brazma A. MGED standards: work in progress. *OMICS* 2006;10(2):138–44. [PubMed: 16901218]
85. Pope WB, Chen JH, Dong J, Carlson MR, Perlina A, Cloughesy TF, Liao LM, Mischel PS, Nghiemphu P, Lai A, Nelson SF. Relationship between gene expression and enhancement in glioblastoma multiforme: exploratory DNA microarray analysis. *Radiology* Oct;2008 249(1):268–77. [PubMed: 18796682]
86. Belda-Iniesta C, de Castro Carpeño J, Casado Sáenz E, Cejas Guerrero P, Perona R, González Barón M. Molecular biology of malignant gliomas. *Clin Transl Oncol* Sep;2006 8(9):635–41. [PubMed: 17005465]
87. Bellazzi R, Zupan B. Predictive data mining in clinical medicine: current issues and guidelines. *Int J Med Inform* 2008;77(2):81–97. [PubMed: 17188928]
88. Baggerly K, Coombes K. Deriving Chemosensitivity from Cell Lines: Forensic Bioinformatics and Reproducible Research in High-Throughput Biology. Submitted to the *Annals of Applied Statistics*. Available at http://www.imstat.org/aoas/next_issue.html.
89. Morris J. Fatal flaws in cancer research: AOAS paper. *IMS Bulletin* 2010;39(1):5.
90. Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc Natl Acad Sci U S A* Apr 11;2006 103(15):5923–8. [PubMed: 16585533]
91. Ridley, M. *Nature Via Nurture: Genes, Experience, and What Makes Us Human*. HarperCollins; USA: 2003. p. 336