



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

# MICROARRAY ANALYSIS OF GENE EXPRESSION OF CANCER TO GUIDE THE USE OF CHEMOTHERAPEUTICS

Tzu-Hao Wang<sup>1,2\*</sup>, Angel Chao<sup>1</sup>

<sup>1</sup>Department of Obstetrics and Gynecology, Lin-Kou Medical Center, Chang Gung Memorial Hospital, Chang Gung University, and <sup>2</sup>Genomic Medicine Research Core Laboratory, Chang Gung Memorial Hospital, Tao-Yuan, Taiwan.

## SUMMARY

The beauty of microarray analysis of gene expression (MAGE) is that it can be used to discover some genes that were previously thought to be unrelated to a physiologic or pathologic event. During the period from 1999 to 2007, applications of MAGE in cancer investigation have shifted from molecular profiling, identifying previously undiscovered cancer types, predicting outcomes of cancer patients, revealing metastasis signatures of solid tumors, to guiding the use of therapeutics. The roles of cancer genomic signatures have evolved through three phases. In the first phase, genomic signatures were described in stored cancer specimens and dubbed as molecular portraits of cancer. When gene expression profiles were carefully correlated with sufficient clinical information of cancer patients, new subgroups of cancers with distinct outcomes were revealed. In studies of the second phase, validation of cancer signatures was emphasized and commonly performed with independent groups of cancer specimens or independent data set. In the third phase, cancer genomic signatures have been further expanded beyond depicting the molecular portrait of cancer to predicting patient outcomes and guiding the use of cancer therapeutics. Cancer genomic signatures have become an essential part of a new generation of cancer clinical trials. It is advocated that, in future clinical trials of cancer therapy, the cancer specimens of each participant should be tested for currently available predictor genomic signatures, so that the most effective treatment with the least adverse effects for each patient can be identified. Then, participants can be triaged to an appropriate study group. [*Taiwan J Obstet Gynecol* 2007;46(3):222-229]

**Key Words:** cancer therapeutics, clinical application, gene expression, genomic signatures, microarrays

## Introduction

A DNA microarray is an orderly arrangement of DNA on solid support, providing a medium for matching known and unknown DNA samples [1]. The types of DNA microarrays and relevant methodologies are reviewed by Chao et al [2]. In this article, we briefly review the current advances in microarray analysis of gene expression (MAGE), focusing on recent reports of the MicroArray Quality Control (MAQC) project and the

shift of MAGE usage from molecular cancer profiling to clinical cancer therapeutics.

## Applications of MAGE in Studying Physiologic and Pathologic Mechanisms

Gene expression profiling has generated and continues to generate extensive information on the molecular mechanisms of cellular function in particular tissues during physiologic or pathologic events. Microarray analysis technology is a high throughput platform for gene expression profiling. The beauty of MAGE is that we can usually discover some genes that were previously not linked to certain physiologic or pathologic events. For instance, we have gained insight into the host response to SARS infection [3], tumor biology of

\*Correspondence to: Dr Tzu-Hao Wang, Department of Obstetrics and Gynecology, Chang Gung Memorial Hospital, Chang Gung University, Tao-Yuan 333, Taiwan.  
E-mail: knoxtn@cgmh.org.tw  
Accepted: August 9, 2007

various cervical cancer types [4], molecular mechanisms in paclitaxel treatment of ovarian cancer [5], and the intrinsic difference among the mesenchymal stem cells derived from distinct origins [6].

## MAQC

When DNA microarrays are used to analyze similar tissues, gene expression profiles obtained from different studies used to be notoriously varied [7], sometimes even conflicting. Possible causes for the discrepancy include different assay platforms using different sequences to represent a particular gene, non-uniform coverage of gene sets, distinct data filtering strategies, various statistical stringency, as well as data complexity and variability [8–10]. The identification of differentially expressed genes in a studied condition with DNA microarray analysis is often determined by the criteria set by the investigators. Therefore, concerns have been raised regarding the reliability of microarray results due to the varied and often conflicting reports [11,12].

To address this concern, a collaborative effort led by the United States Food and Drug Administration that included 137 scientists from 51 organizations representing academia, industry, and the US government has completed the MAQC project [13–18]. In this project, identical specimens were aliquoted and assigned to participating laboratories to analyze using different microarray platforms, including those manufactured by Applied Biosystems, Affymetrix, Agilent, and GE Healthcare. To validate the quantitative capability of microarrays, microarray results were compared with real-time quantitative polymerase chain reaction (PCR). The correlation between Affymetrix gene expression results and TaqMan real-time quantitative PCR results has shown good linearity ( $r^2 = 0.95$ ) [15].

A fold-change ranking method with a  $p$ -value cutoff  $< 0.05$  has recently been shown to be reproducible in selecting the signature gene list from results using different microarray platforms [18]. These selection on criteria have been shown to more reproducible than  $t$ -test  $p$  value or significance analysis of microarrays [18]. We have applied this method in selecting the signature gene expression profiles with ease; after filtering using  $p < 0.05$ , we ranked genes by fold change and chose the top 25 genes that were upregulated in each group of mesenchymal stem cells derived from amniotic fluid, amniotic membrane, cord blood, and adult bone marrow [6].

Collectively, because of the remarkable improvement of microarray technology and the aforementioned critical evaluations, the majority of microarray researchers recognize the reliability and consistency of well-designed

and carefully conducted microarray results. Even in the 2 years before the publication of the MAQC project, the clinical and biologic findings derived from microarrays were regarded to be “remarkably robust, with a high level of quantitative precision” [19,20]. The recent MAQC results further demonstrate that microarray gene expression analysis itself is suitable as a stand-alone quantitative comparison [17]. Nevertheless, we should not ignore potential flaws. All the encouraging results of the MAQC project only establish that microarray technology is robust, but they do not imply that the technology is foolproof. Quoted from a commentary in the November 17, 2006 issue of *Cell*, “You can learn to do PCR well in a month. But with microarrays, it can take years.” [13].

## Advanced Microarray Data Analyses Beyond Detection of Differential Expression

To evaluate how MAGE can help to make a diagnosis or choose a therapy, researchers use one set of patients to identify a gene-expression pattern called a genetic signature that can correspond to a clinical issue, such as a 5-year survival rate, the response to a treatment, or the induction of side effects by a drug. The power of microarray technology is its ability to use changes in multiple genes as the pattern of gene expression rather than to choose thresholds of individual markers [19]. This genetic signature is then validated on other groups of patients [13].

During this trial period, it is critical that investigators understand how to minimize expression noise and bias through effective design. Expression noise can be defined as gene expression variation that does not correlate with the biology or behavior being studied and is introduced by both the technology itself and/or during tissue processing [21]. Bias is not inherent to microarray analysis but is easily introduced by faulty experimental design [21]. A series of sophisticated analytical strategies to address these problems have been discussed [19,22,23], as summarized in Table 1.

An unsupervised analysis does not use any a priori class definition, but it simply seeks to determine what structure is inherent in the data [19]. A commonly used example of unsupervised analysis is hierarchical clustering, i.e. letting the data define its own patterns by clustering genes that are most similar in expression profile [24]. A supervised analysis is more likely to reveal putative associations between genes and the cytogenetic class, but it may bias the outcome by forcing a model onto the data, i.e. the “overfitting” risk [19].

**Table 1.** Glossary of microarray analysis of gene expression

Analytic methods	Description	Reference
Unsupervised analysis	To simply seek to determine what structure is inherent in the data, without any a priori class definition	Eisen et al [24]
Supervised analysis	To reveal associations between genes and the different known disease classes	Golub et al [32]
Meta-analysis	To extract robust profiles from independent data sets	Rhodes et al [25]
Functional enrichment analysis	To use external databases for interpreting and summarizing large data sets	Rhodes and Chinnaiyan [22]
Transcriptional network analysis	To reduce a complex cancer signature to a small number of activated transcriptional programs	Rhodes and Chinnaiyan [22]
Analysis of expression modules	To transcend the investigation of cancer gene expression from individual genes to biologic processes using functional pathways as gene modules	Segal et al [23]

To extract robust profiles from multiple data sets, a meta-analysis has been done on 40 independent data sets derived from more than 3,700 array experiments, identifying 36 cancer signatures that were activated in cancer relative to the normal tissue from which the cancer arised [25]. A meta-analysis of these signatures further identified 67 genes that were activated in 12 or more signatures, suggesting a common transcriptional program pervading most types of cancer [22].

In functional enrichment analysis, a series of external functional information has been used to interpret and summarize large cancer signatures milestones [22]. Databases of external functional information include Gene Ontology ([www.geneontology.org](http://www.geneontology.org)) [26], Kyoto Encyclopedia of Genes and Genomes ([www.genome.jp/kegg/](http://www.genome.jp/kegg/)), Biocarta ([www.biocarta.com](http://www.biocarta.com)), and GenMAPP ([www.genmapp.org](http://www.genmapp.org)). Commercially maintained integrative databases and softwares include MetaCore by GeneGo ([www.genego.com/metacore.php](http://www.genego.com/metacore.php)) and Ingenuity Systems ([www.ingenuity.com](http://www.ingenuity.com)). We have recently used the MetaCore suite to analyze the signature profiles of mesenchymal stem cells of various origins and obtain insights into biologic processes of each group [6].

The goal of transcriptional network analysis is to simplify a complex cancer signature to a small number of activated transcriptional programs that may shed light on neoplastic mechanisms and further point to potential targets of therapeutic intervention [27]. In addition to the aforementioned functional enrichment analysis, in which many of the downstream effectors are transcription factors, chromatin immunoprecipitation coupled with promoter microarrays (ChIP-chip assays) allow for genome-wide identification of transcription factor-binding sites [28,29]. With hundreds of consensus binding sequences for transcription factors, which have been defined by sequence-based methods, it is feasible to perform a large-scale integrative

analysis of binding-site profiles and cancer signature expression profiles [22].

Analysis of expression modules, in which functional pathways (i.e. gene modules) are used as gene modules, was proposed to extend the investigation of cancer gene expression from individual genes to biologic processes [23]. When this concept of higher-level modules was applied to examine the joint behavior of differentially expressed genes in diabetic muscle, a significant change in the whole set of genes was noted, even though the expression of individual genes was not significantly different [30]. Segal et al used this module-level analysis to obtain a global view of the shared and unique molecule modules underlying human cancer [31]. They demonstrated that activation or repression of some modules (e.g. cell cycle) was shared across multiple tumor types and could be regarded as a general tumorigenesis, whereas others (e.g. growth-regulatory modules) were more specific to tissue origin or progression of particular tumors [23].

## MAGE in Clinical Cancer Investigation

Applications of MAGE in clinical cancer investigation have shifted from molecular profiling in the year 1999 [32,33], identifying previously undiscovered subgroups of particular type of cancer [34], predicting outcomes of cancer patients in 2002 [35,36], and revealing a metastasis signature of solid tumors [37], to guiding the use of therapeutics in 2006 [38], as summarized in Table 2. The use of MAGE as the guide of cancer therapeutics has also been compared in meta-analyses in large B-cell lymphoma [39].

In a leukemia data set of 38 bone marrow samples (27 acute lymphoblastic leukemia and 11 acute myeloblastic leukemia), Golub et al tested whether gene expression monitoring by DNA microarray could be

**Table 2.** Milestone studies of microarray analysis of gene expression in cancer research

Year	Reports on	References
1999	Molecular classification of cancer using supervised machine learning	Golub et al [32]
2000	Molecular profiling of breast cancer	Perou et al [33]
2000	Identification of subgroups of diffuse large B-cell lymphoma with different outcomes	Alizadeh et al [34]
2002	Prediction of clinical outcomes of breast cancer	van't Veer et al [35]
2003	Identification of metastasis signature that reflected both contributions of the tumor and the host environment	Ramaswamy et al [37]
2004	Identification of prognostic profiles of adult acute myeloid leukemia	Bullinger et al [45]; Valk et al [46]
2004	Using independent samples of lymphoma to test a meta-analysis derived signature profile for predicting overall survival in diffuse large-B-cell lymphoma	Lossos et al [39]
2006	Concordance among gene-expression based predictors for breast cancer	Fan et al [53]
2006	Development of a series of oncogenic pathway signatures in human cancers	Bild et al [54]
2006	Genomic signature to guide the use of cancer chemotherapeutics	Potti et al [38]

used to identify new cancer classes (class discovery) and to assign tumors to known classes (class prediction) [32]. Using a supervised learning classification algorithm, Golub et al first constructed class predictors and then evaluated them using cross-validation with the same collection of specimens with known outcomes. Their results suggested a general strategy for discovering and predicting cancer classes, which proved useful in predicting outcomes in patients with other tumor types [36].

As a proof of principle, Perou et al used cDNA microarrays to identify genes of differential expression between *in vitro* cultured human mammary epithelial cells and breast tumor specimens [40]. Their results supported the feasibility and usefulness of this systematic approach for studying variations in gene expression patterns in human cancers as a means to dissect and classify solid tumors [40]. Then, 64 surgical specimens of human breast tumors from 42 patients were analyzed for gene expression profiles [33]. They identified a set of co-expressed genes for which variation in mRNA levels could be related to specific features of physiologic variation. Molecular portraits of cancer with gene expression profiles were thus proposed [33].

Diffuse large B-cell lymphoma (DLBCL) is one disease in which further subgrouping by histology is difficult because of inter- and intra-pathologist irreproducibility [41]. Using hierarchical clustering for MAGE profiling, 40 DLBCL specimens could be divided into two distinct groups: 19 cases of germinal center B-like DLBCL, and 21 cases of activated B-like DLBCL [34]. Statistically, patients with germinal center B-like DLBCL had a better overall survival than those with activated B-like DLBCL. The molecular classification of tumors based on gene expression profiles has thus proved its ability to identify previously undetected and clinically significant subtypes of cancer [34].

Even in the same stage of disease, breast cancer is notorious for its unpredictable response to chemotherapy and variable overall outcome. Chemotherapy [42] or hormonal therapy [43] reduces the risk of distant metastasis by about a third. However, because of the lack of an accurate patient triage strategy to determine who should or should not undergo adjuvant therapy, many patients who might not develop cancer metastasis at all have unnecessarily undergone adjuvant therapy. To develop patient-tailored therapy strategies for breast cancer, van't Veer et al used supervised classification to analyze DNA microarray data on primary breast tumors of 117 young patients [35]. van't Veer et al identified a poor prognosis signature, by which they could predict a short time interval to distant metastasis. Therefore, these results provide a strategy to select patients who would benefit from adjuvant therapy.

In search of the molecular metastasis signature of cancer, Ramaswamy et al compared gene expression patterns between primary tumors and metastases [37]. They identified a gene expression signature that distinguished primary from metastatic adenocarcinomas. However, the authors found that a subset of primary tumors resembled metastatic tumors, and they further confirmed this finding by applying the expression signature to data on 279 primary solid tumors of diverse types. Notably, Ramaswamy et al analyzed whole tumor tissues including surrounding stromal cells, instead of pure cancer population that could be isolated using laser capture microdissection. In the 17-gene metastasis signature identified in that study, two collagen genes and one lamin gene were upregulated, suggesting that malignancy is the product of the tumor-host microenvironment [44].

Two articles about the prognostic usefulness of gene expression profiles in acute myeloid leukemia (AML)

were published back to back in the April 15, 2004 issue of *New England Journal of Medicine* [45,46]. Bullinger et al of Stanford University used cDNA microarrays to identify a 133-gene clinical outcome predictor of AML [45], while Valk et al of Erasmus University at Rotterdam used Affymetrix oligonucleotide microarrays to identify a unique cluster of genes in AML patients, which are thought to be associated with a poor outcome. Cancer signature gene profiles used to be criticized to be an highly unstable predictor of prognosis [47]. However, even using different microarray platforms, the results of the studies by Bullinger et al and Valk et al were surprisingly robust [19,20]. These studies also demonstrated the ability of microarray technology to use somewhat imprecise patterns of gene expression rather than exact thresholds of individual markers [19].

Currently, many prestigious journals, such as *Science* and *Nature* series, ask authors to deposit their microarray data in a Minimal Information About a Microarray Experiment (MIAME)-compliant form to one of two public repositories: Gene Expression Omnibus at National Center of Biotechnology and Informatics (<http://www.ncbi.nlm.nih.gov/geo/>) and Array Express at European Bioinformatics Institute (<http://www.ebi.ac.uk/array-express>) as a prerequisite of publication of microarray-derived research articles. Therefore, many large-scale microarray data sets became available for re-analyses by other researchers worldwide. For instance, multiple important papers of breast cancer gene expression profiles [35,48–50] have been derived from the same comprehensive microarray data set with sufficient clinical information [51].

Using several microarray data sets [34,36,52], Lossos et al did a meta-analysis of DBCL and selected 36 genes for further analysis with real-time quantitative PCR on independent samples of lymphoma from 66 patients [39]. Six genes that showed the strongest predicting power were *LMO2*, *BCL6*, *FN1*, *CCND2*, *SCYA3*, and *BCL2*. After testing these genes in two additional independent microarray data sets, Lossos et al concluded that measurement of the expression of these six genes to be sufficient in predicting overall survival in DBCL [39].

The use of microarray gene expression profiles in predicting outcomes of cancer patients has been validated in the aforementioned studies. However, even if the same group samples were analyzed, distinct prognostic profiles have been derived for outcome prediction [35,48–51]. To resolve this paradox, Fan et al compared the prediction powers of the gene sets for the same group of specimens by applying five gene-expression-based models: intrinsic subtypes, 70-gene profile, wound response, recurrence score, and the

two-gene ratio [53]. By performing Kaplan–Meier survival analyses of 295 patients with breast cancer, Fan et al found that four of the five models tested showed significant agreement in the outcome prediction for individual patients. The only exception was the model using two-gene ratio, which could not result in a reliable prediction [53]. The explanation for the surprising concordance among the four different models in predicting breast cancer outcomes is likely to be as follows: there was a large group of genes, which behaved differently and were related to biologic phenotypes of cancer, and ultimately, the patients' outcomes. In this large list of genes, each of the four models might have used only some of them to construct the signature profiles used for predicting outcomes.

To dissect oncogenic pathway signatures in human cancer, Bild et al used adenoviral vectors to express various oncogenic activities, such as Myc, Ras, E2F3, Src, and  $\beta$ -catenin in otherwise quiescent cells. Applying this strategy, they were able to specifically isolate the subsequent events as defined by the activation/deregulation of that single oncogenic pathway, and analyze these events with Affymetrix oligonucleotide microarrays [54]. In clinical samples of lung, breast and ovarian cancers, Bild et al combined signature-based predictions across several pathways and identified coordinated patterns of pathway deregulation that distinguish between specific cancer and tumor subtypes [54]. Using pathway-specific inhibitors such as Ras pathway inhibitors, either farnesyltransferase inhibitor (L-744832) or farnesylthiosalicylic acid, and the Src pathway inhibitor (SU6656), Bild et al could predict drug sensitivity of tested breast cancer cell lines according to their pathway deregulation. In summary, predictions of pathway deregulation in cancer cells can also predict the sensitivity to therapeutic agents that target components of the pathway. These results pave the path of using these oncogenic pathway signatures to guide the use of target therapeutics.

Because of the enormous complexity of cancer and a frequent inability to properly guide the use of available therapeutics, chemotherapy for solid tumors often results in marginal success. Most people with advanced solid tumors will relapse and die of their disease [55]. Furthermore, oncologists have always had to face the challenge of matching the right therapeutic regimen with the right individual, balancing relative benefit with risk to achieve the best outcome.

With the goal of using genomic signatures to guide the use of chemotherapeutics, Potti et al systematically extracted gene expression profiles from the following microarray data sets: NCI-60 cancer cell lines [56], additional 30 cancer cell lines [57], the authors' previously

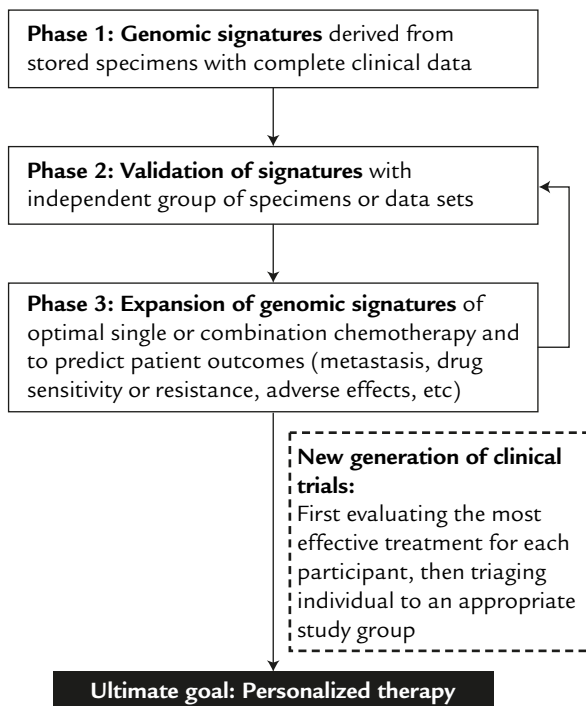
reported breast, ovarian and lung cancer specimens [54], and their newly analyzed 13 ovarian cancer cell lines and 119 advanced (FIGO stages III or IV) serous epithelial ovarian cancers [38]. Potti et al demonstrated patterns of predicted sensitivity of three human solid cancers (breast, lung, ovary) to seven common chemotherapeutic drugs (5-fluorouracil, paclitaxel, docetaxel, adriamycin, topotecan, cyclophosphamide, etoposide). To evaluate how individual signatures respond to a combination of drugs, Potti et al also analyzed 51 breast cancer patients who were in a breast neoadjuvant treatment study that used a combination of paclitaxel, 5-fluorouracil, adriamycin and cyclophosphamide (TFAC). The predicted response that was based on a combined probability of sensitivity built from the individual chemosensitivity predictions yielded a statistically significant distinction between responders and nonresponders [38].

### Evolving Roles of Genomic Signatures in a New Generation of Clinical Trials

As summarized in the Figure, the roles of cancer genomic signatures have evolved through three phases. In the first phase, genomic signatures were described in stored cancer specimens and dubbed as molecular portraits of cancer [33]. When gene expression profiles were

carefully correlated with sufficient clinical information of cancer patients, new subgroups of cancers with distinct outcomes were revealed [34]. In studies of the second phase, validation of cancer signatures was emphasized and commonly done with independent groups of cancer specimens or independent data sets [22,25,35,39]. In the third phase, cancer genomic signatures have been expanded beyond depicting the molecular portrait of cancer to predicting patient outcomes [45,46], including metastasis [37]. It has become a rule, in the third phase, that all of the prognostic genomic signatures be validated in additional data sets. Potti et al further demonstrated the role of cancer genomic signatures as a guide to the use of cancer therapeutics [38]. A new generation of cancer clinical trials was proposed, in which the cancer specimens of each participant should be tested for currently available predictor genomic signatures so that the most effective treatment and the least adverse effects for each patient could be identified. Then, participants can be triaged to an appropriate study group (Figure). In fact, successful examples in treating patients with early-stage non-small cell lung cancer were reported by Potti et al [59] and in treating those with advanced-stage ovarian cancer by Dressman et al [60].

It is commonly argued against microarray results by the fact that a transcriptome does not necessarily reflect the corresponding proteome, the collection of proteins that execute the majority of cellular functions. Indeed, mRNA expression is only a coarse surrogate for protein activation levels. For many genes, however, mRNA expression is a useful surrogate [23]. As documented in many studies, when one finds strong signals of differential expression, these are typically reflected later at the protein levels. The latter can be validated by protein assays such as enzyme-linked immunosorbent assays [5] or immunohistochemistry [4].



**Figure.** Three phases of genomic signatures in cancer therapy and a new generation of clinical trials. This schema is adapted from Herbst and Lippman [58] and Potti et al [38].

### Conclusion

At the current developmental pace of genomic technology, the clinical trend towards personalized medicine is almost certain (Figure). Many biomedical researchers and clinicians predict that microarray technology will be incorporated into clinical laboratories in hospitals in the near future. As extrapolated from the results of the studies discussed in this article, the use of a “focused array” to measure the expression of 50 to 100 genes in the signature profile for a selected disease would help clinicians predict the patient’s response to a drug, triage patients into a chemo-responsive versus chemo-resistant group, and evaluate a panel of risk

factors that may result in comorbidity of the patient. To achieve this practical feasibility, microarray technology would need to address a range of quality-control issues. Every aspect of the process should be so robust that it can be considered to be foolproof. It is predicted that within the next 10 years, microarray developers will meet these challenges [13].

## Acknowledgments

The authors like to thank Dr Yun-Shien Lee (Department of Biotechnology, Ming Chuan University) for his long-term contribution to the Genomic Medicine Research Core Laboratory of Chang Gung Memorial Hospital and Shih-Yee Mimi Wang (University of Illinois-Chicago College of Medicine) for editing the manuscript. This study was supported by grants NSC 93-2314-B-182A-102 (to TH Wang) and NMRPG450141 (to A Chao) from the National Science Council, and CMRPG 340462 (to TH Wang) and CMRPG36481 (to A Chao) from Chang Gung Memorial Hospital, Taiwan.

## References

1. Wang TH, Lee YS, Chen ES, et al. Establishment of cDNA microarray analysis at the Genomic Medicine Research Core Laboratory (GMRCL) of Chang Gung Memorial Hospital. *Chang Gung Med J* 2004;27:243–60.
2. Chao A, Wang TH, Lai CH. Overview of microarray analysis of gene expression and its applications to cervical cancer investigation. *Taiwan J Obstet Gynecol* 2007 (In press).
3. Lee YS, Chen CH, Chao A, et al. Molecular signature of clinical severity in recovering patients with severe acute respiratory syndrome coronavirus (SARS-CoV). *BMC Genomics* 2005;6:132.
4. Chao A, Wang TH, Lee YS, et al. Molecular characterization of adenocarcinoma and squamous carcinoma of the uterine cervix using microarray analysis of gene expression. *Int J Cancer* 2006;119:91–8.
5. Wang TH, Chan YH, Chen CW, et al. Paclitaxel (Taxol) upregulates expression of functional interleukin-6 in human ovarian cancer cells through multiple signaling pathways. *Oncogene* 2006;25:4857–66.
6. Tsai MS, Hwang SM, Chen KD, et al. Functional network analysis on the transcriptomes of mesenchymal stem cells derived from amniotic fluid, amniotic membrane, cord blood, and bone marrow. *Stem Cells* 2007 (In press).
7. Tan PK, Downey TJ, Spitznagel EL Jr, et al. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res* 2003;31:5676–84.
8. Yauk CL, Berndt ML, Williams A, Douglas GR. Comprehensive comparison of six microarray technologies. *Nucleic Acids Res* 2004;32:e124.
9. 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* 2005;6:107.
10. Wang H, He X, Band M, Wilson C, Liu L. A study of inter-lab and inter-platform agreement of DNA microarray data. *BMC Genomics* 2005;6:71.
11. Fortunel NO, Otu HH, Ng HH, et al. Comment on “‘Stemness’: transcriptional profiling of embryonic and adult stem cells” and “a stem cell molecular signature”. *Science* 2003;302:393; author reply 393.
12. Marshall E. Getting the noise out of gene arrays. *Science* 2004;306:630–1.
13. Strauss E. Arrays of hope. *Cell* 2006;127:657–9.
14. Tong W, Lucas AB, Shippy R, et al. Evaluation of external RNA controls for the assessment of microarray performance. *Nat Biotechnol* 2006;24:1132–9.
15. Canales RD, Luo Y, Willey JC, et al. Evaluation of DNA microarray results with quantitative gene expression platforms. *Nat Biotechnol* 2006;24:1115–22.
16. Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, et al. Performance comparison of one-color and two-color platforms within the MicroArray Quality Control (MAQC) project. *Nat Biotechnol* 2006;24:1140–50.
17. Shi L, Reid LH, Jones WD, et al, for MAQC Consortium. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 2006;24:1151–61.
18. Guo L, Lobenhofer EK, Wang C, et al. Rat toxicogenomic study reveals analytical consistency across microarray platforms. *Nat Biotechnol* 2006;24:1162–9.
19. Liu ET, Karuturi KR. Microarrays and clinical investigations. *N Engl J Med* 2004;350:1595–7.
20. Grimwade D, Haferlach T. Gene-expression profiling in acute myeloid leukemia. *N Engl J Med* 2004;350:1676–8.
21. Febbo PG, Kantoff PW. Noise and bias in microarray analysis of tumor specimens. *J Clin Oncol* 2006;24:3719–21.
22. Rhodes DR, Chinnaiyan AM. Integrative analysis of the cancer transcriptome. *Nat Genet* 2005;37(Suppl):S31–7.
23. Segal E, Friedman N, Kaminski N, Regev A, Koller D. From signatures to models: understanding cancer using microarrays. *Nat Genet* 2005;37(Suppl):S38–45.
24. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998;95:14863–8.
25. Rhodes DR, Yu J, Shanker K, et al. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci USA* 2004;101:9309–14.
26. Harris MA, Clark J, Ireland A, et al, for Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 2004;32(Database issue):D258–61.
27. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Barrette TR, Ghosh D, Chinnaiyan AM. Mining for regulatory programs in the cancer transcriptome. *Nat Genet* 2005;37:579–83.
28. Weinmann AS, Yan PS, Oberley MJ, Huang TH, Farnham PJ. Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev* 2002;16:235–44.



29. Odom DT, Zizlsperger N, Gordon DB, et al. Control of pancreas and liver gene expression by HNF transcription factors. *Science* 2004;303:1378–81.
30. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 2003;34:267–73.
31. Segal E, Friedman N, Koller D, Regev A. A module map showing conditional activity of expression modules in cancer. *Nat Genet* 2004;36:1090–8.
32. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531–7.
33. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
34. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503–11.
35. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530–6.
36. Shipp MA, Ross KN, Tamayo P, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* 2002;8:68–74.
37. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. *Nat Genet* 2003;33:49–54.
38. Potti A, Dressman HK, Bild A, et al. Genomic signatures to guide the use of chemotherapeutics. *Nat Med* 2006;12:1294–300.
39. Lossos IS, Czerwinski DK, Alizadeh AA, Wechsler MA, Tibshirani R, Botstein D, Levy R. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *N Engl J Med* 2004;350:1828–37.
40. Perou CM, Jeffrey SS, van de Rijn M, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci USA* 1999;96:9212–7.
41. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. *Blood* 1997;89:3909–18.
42. Polychemotherapy for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. *Lancet* 1998;352:930–42.
43. Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. *Lancet* 1998;351:1451–67.
44. Liotta LA, Kohn EC. Cancer's deadly signature. *Nat Genet* 2003;33:10–1.
45. Bullinger L, Dohner K, Bair E, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 2004;350:1605–16.
46. Valk PJ, Verhaak RG, Beijen MA, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* 2004;350:1617–28.
47. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;365:488–92.
48. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001;98:10869–74.
49. 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.
50. Chang HY, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci USA* 2005;102:3738–43.
51. Chang HY, Sneddon JB, Alizadeh AA, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol* 2004;2:E7.
52. Rosenwald A, Wright G, Chan WC, et al, for the Lymphoma/Leukemia Molecular Profiling Project. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 2002;346:1937–47.
53. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 2006;355:560–9.
54. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;439:353–7.
55. Breathnach OS, Freidlin B, Conley B, et al. Twenty-two years of phase III trials for patients with advanced non-small-cell lung cancer: sobering results. *J Clin Oncol* 2001;19:1734–42.
56. Staunton JE, Slonim DK, Collier HA, et al. Chemosensitivity prediction by transcriptional profiling. *Proc Natl Acad Sci USA* 2001;98:10787–92.
57. Györfy B, Surowiak P, Kiesslich O, Denkert C, Schafer R, Dietel M, Lage H. Gene expression profiling of 30 cancer cell lines predicts resistance towards 11 anticancer drugs at clinically achieved concentrations. *Int J Cancer* 2006;118:1699–712.
58. Herbst RS, Lippman SM. Molecular signatures of lung cancer—toward personalized therapy. *N Engl J Med* 2007;356:76–8.
59. Potti A, Mukherjee S, Petersen R, et al. A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N Engl J Med* 2006;355:570–80.
60. Dressman HK, Berchuck A, Chan G, et al. An integrated genomic-based approach to individualized treatment of patients with advanced-stage ovarian cancer. *J Clin Oncol* 2007;25:517–25.