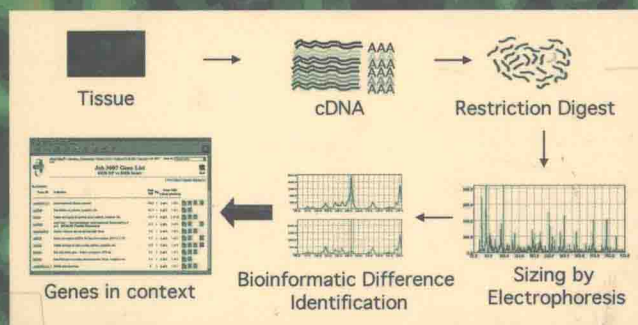


Gene Expression Profiling

Methods and Protocols

Edited by

Richard A. Shimkets



Gene Expression Profiling

Methods and Protocols

Edited by

Richard A. Shimkets

CuraGen Corporation, New Haven, CT


HUMANA PRESS  TOTOWA, NEW JERSEY

© 2004 Humana Press Inc.
999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

www.humanapress.com

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise without written permission from the Publisher. Methods in Molecular Biology™ is a trademark of The Humana Press Inc.

All papers, comments, opinions, conclusions, or recommendations are those of the author(s), and do not necessarily reflect the views of the publisher.

This publication is printed on acid-free paper. 
ANSI Z39.48-1984 (American Standards Institute)

Permanence of Paper for Printed Library Materials.

Cover illustration: From Fig. 1, in Chapter 4, "GeneCalling: Transcript Profiling Coupled to a Gene Database Query," by Richard A. Shimkets.

Production Editor: Robin B. Weisberg

Cover design by Patricia F. Cleary.

For additional copies, pricing for bulk purchases, and/or information about other Humana titles, contact Humana at the above address or at any of the following numbers: Tel.: 973-256-1699; Fax: 973-256-8341; E-mail: humana@humanapress.com; or visit our Website: www.humanapress.com

Photocopy Authorization Policy:

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Humana Press Inc., provided that the base fee of US \$25.00 per copy is paid directly to the Copyright Clearance Center at 222 Rosewood Drive, Danvers, MA 01923. For those organizations that have been granted a photocopy license from the CCC, a separate system of payment has been arranged and is acceptable to Humana Press Inc. The fee code for users of the Transactional Reporting Service is: [1-58829-220-7/04 \$25.00].

Printed in the United States of America. 10 9 8 7 6 5 4 3 2 1

E-ISBN 1-59259-751-3

Library of Congress Cataloging-in-Publication Data

Gene expression profiling : methods and protocols / edited by Richard A. Shimkets.

p. ; cm. -- (Methods in molecular biology, ISSN 1064-3745 ; v. 258)

Includes bibliographical references and index.

ISBN 1-58829-220-7 (alk. paper)

1. Gene expression--Laboratory manuals.

[DNLM: 1. Gene Expression Profiling--methods. QZ 52 G326 2004] I. Shimkets, Richard A.

II. Series: Methods in molecular biology (Totowa, N.J.) v. 258.

QH450.G46255 2004

572.8'65--dc22

2003022875

Preface

Why Quantitate Gene Expression?

The central dogma of molecular biology is the concept of the transcription of messenger RNA from a DNA template and translation of that RNA into protein. Since the transcription of RNA is a key regulatory point that may eventually signal many other cascades of events, the study of RNA levels in a cell or organ can help the understanding of a wide variety of biological systems.

It is assumed that readers of *Gene Expression Profiling: Methods and Protocols* have an appreciation for this fact, so that the present volume focuses on the practical and technical considerations that guide the choice of methodology in this area. Thoroughly thinking through the specific scientific questions at hand will enable the choice of the best technology for that application and examples of which technologies fit which applications best will be discussed.

Richard A. Shimkets

Contributors

- LUANNE CHEHAK • *Third Wave Technologies Inc., Madison, WI*
ZHUIAN J. CHEN • *Department of Pharmacology, Fox Chase Cancer Center, Philadelphia, PA*
MARINA E. CHICUREL • *Affymetrix Inc., Santa Clara, CA*
ERIN L. CRAWFORD • *Departments of Medicine and Pathology, Medical College of Ohio, Toledo, OH*
MICHELLE L. CURTIS • *Third Wave Technologies Inc., Madison, WI*
DENNISE D. DALMA-WEISZHAUSZ • *Affymetrix Inc., Santa Clara, CA*
SEJAL M. DESAI • *BD Biosciences Clontech, Palo Alto, CA*
TIMOTHY G. GRAVES • *Departments of Medicine and Pathology, Medical College of Ohio, Toledo, OH*
ELIZABETH A. HERNES • *Gene Express Inc., Toledo, OH*
CHARLES R. KNIGHT • *Departments of Medicine and Pathology, Medical College of Ohio, Toledo, OH*
ROBERT W. KWIATKOWSKI • *Third Wave Technologies Inc., Madison, WI*
SERGEY A. LUKYANOV • *Evrogen JSC and Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*
MARTIN D. LEACH • *Bioinformatics, CuraGen Corporation, New Haven, CT*
JENNIFER LESCALLETT • *Affymetrix, Inc., Santa Clara, CA*
ROBERT LIPSHUTZ • *Affymetrix Inc., Santa Clara, CA*
CHERYL A. MOTTEN • *Departments of Medicine and Pathology, Medical College of Ohio, Toledo, OH*
MARILYN C. OLSON • *Third Wave Technologies Inc., Madison, WI*
SARAH M. OLSON • *Third Wave Technologies Inc., Madison, WI*
DENIS V. REBRIKOV • *Evrogen JSC and Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*
RICHARD A. SHIMKETS • *Drug Discovery, CuraGen Corporation, New Haven, CT*
PAUL D. SIEBERT • *BD Biosciences Clontech, Palo Alto, CA*
TSETSKA TAKOVA • *Third Wave Technologies Inc., Madison, WI*
KENNETH D. TEW • *Department of Pharmacology, Fox Chase Cancer Center, Philadelphia, PA*
GILBERT VASSART • *Institute of Interdisciplinary Research (IRIBHM) Université Libre de Bruxelles, Brussels, Belgium and Department of Medical Genetics of the Erasme Hospital, Brussels, Belgium*

CATHELINE VILAIN • *Institute of Interdisciplinary Research (IRIBHM)*
Université Libre de Bruxelles, Brussels, Belgium

K. A. WARNER • *Departments of Medicine and Pathology, Medical College*
of Ohio, Toledo, OH

JAMES C. WILLEY • *Departments of Medicine and Pathology, Medical*
College of Ohio, Toledo, OH

ROBERT J. ZAHORCHAK • *Gene Express Inc., Toledo, OH*

Contents

Preface	v
Contributors	xi
1 Technical Considerations in Quantitating Gene Expression <i>Richard A. Shimkets</i>	1
2 Gene Expression Quantitation Technology Summary <i>Richard A. Shimkets</i>	7
3 Standardized RT-PCR and the Standardized Expression Measurement Center <i>James C. Willey, Erin L. Crawford, Charles R. Knight, K. A. Warner, Cheryl A. Motten, Elizabeth A. Herness, Robert J. Zahorchak, and Timothy G. Graves</i>	13
4 GeneCalling: Transcript Profiling Coupled to a Gene Database Query <i>Richard A. Shimkets</i>	43
5 Invader Assay for RNA Quantitation <i>Marilyn C. Olson, Tsetska Takova, LuAnne Chehak, Michelle L. Curtis, Sarah M. Olson, and Robert W. Kwiattkowski</i>	53
6 Monitoring Eukaryotic Gene Expression Using Oligonucleotide Microarrays <i>Jennifer Lescallett, Marina E. Chicurel, Robert Lipshutz, and Dennise D. Dalma-Weiszhausz</i>	71
7 Amplified Differential Gene Expression Microarray <i>Zhijian J. Chen and Kenneth D. Tew</i>	95
8 Suppression Subtractive Hybridization <i>Denis V. Rebrikov, Sejal M. Desai, Paul D. Siebart, and Sergey A. Lukyanov</i>	107
9 Small Amplified RNA-SAGE <i>Catheline Vilain and Gilbert Vassart</i>	135
10 Gene Expression Informatics <i>Martin D. Leach</i>	153
Index	167

Technical Considerations in Quantitating Gene Expression

Richard A. Shimkets

1. Introduction

Scientists routinely lecture and write about gene expression and the abundance of transcripts, but in reality, they extrapolate this information from a variety of measurements that different technologies may provide. Indeed, there are many reasons that applying different technologies to transcript abundance may give different results. This may result from an incomplete understanding of the gene in question or from shortcomings in the applications of the technologies.

The first key factor to appreciate in measuring gene expression is the way that genes are organized and how this influences the transcripts in a cell. **Figure 1** depicts some of the scenarios that have been determined from sequence analyses of the human genome. Most genes are composed of multiple exons transcribed with intron sequences and then spliced together. Some genes exist entirely between the exons of other genes, either in the forward or reverse orientation. This poses a problem because it is possible to recover a fragment or clone that could belong to multiple genes, be derived from an unspliced transcript, or be the result of genomic DNA contaminating the RNA preparation. All of these events can create confusing and confounding results. Additionally, the gene duplication events that have occurred in organisms that are more complex have led to the existence of closely related gene families that coincidentally may lie near each other in the genome. In addition, although there are probably less than 50,000 human genes, the exons within those genes can be spliced together in a variety of ways, with some genes documented to produce more than 100 different transcripts (*1*).

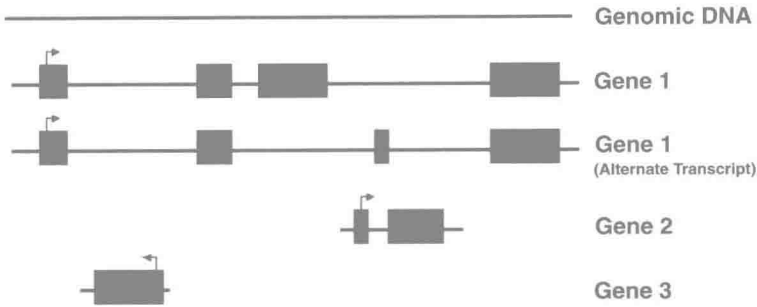


Fig. 1. Typical gene exon structure.

Therefore, there may be several hundred thousand distinct transcripts, with potentially many common sequences. Gene biology is even more interesting and complex, however, in that genetic variations in the form of single nucleotide polymorphisms (SNPs) frequently cause humans and diploid or polyploid model systems to have two (or more) distinct versions of the same transcript.

This set of facts negates the possibility that a single, simple technology can accurately measure the abundance of a specific transcript. Most technologies probe for the presence of pieces of a transcript that can be confounded by closely related genes, overlapping genes, incomplete splicing, alternative splicing, genomic DNA contamination, and genetic polymorphisms. Thus, independent methods that verify the results in different ways to the exclusion of confounding variables are necessary, but frequently not employed, to gain a clear understanding of the expression data. The specific means to work around these confounding variables are mentioned here, but a blend of techniques will be necessary to achieve success.

2. Methods and Considerations

There are nine basic considerations for choosing a technology for quantitating gene expression: architecture, specificity, sensitivity, sample requirement, coverage, throughput, cost, reproducibility, and data management.

2.1. Architecture

We define the architecture of a gene-expression analysis system as either an open system, in which it is possible to discover novel genes, or a closed system in which only known gene or genes are queried. Depending on the application, there are numerous advantages to open systems. For example, an open system may detect a relevant biological event that affects splicing or genetic variation. In addition, the most innovative biological discovery processes have involved the

discovery of novel genes. However, in an era where multiple genome sequences have been identified, this may not be the case. The genomic sequence of an organism, however, has not proven sufficient for the determination of all of the transcripts encoded by that genome, and thus there remain prospects for novelty regardless of the biological system. In model systems that are relatively uncharacterized at the genomic or transcript level, entire technology platforms may be excluded as possibilities. For example, if one is studying transcript levels in a rabbit, one cannot comprehensively apply a hybridization technology because there are not enough transcripts known for this to be of value. If one simply wants to know the levels of a set of known genes in an organism, a hybridization technology may be the most cost-effective, if the number of genes is sufficient to warrant the cost of producing a gene array.

2.2. Specificity

The evolution of genomes through gene or chromosomal fragment duplications and the subsequent selection for their retention, has resulted in many gene families, some of which share substantial conservation at the protein and nucleotide level. The ability for a technology to discriminate between closely related gene sequences must be evaluated in this context in order to determine whether one is measuring the level of a single transcript, or the combined, added levels of multiple transcripts detected by the same probing means. This is a double-edged sword because technologies with high specificity, may fail to identify one allele, or may do so to a different degree than another allele when confronted with a genetic polymorphism. This can lead to the false positive of an expression differential, or the false negative of any expression at all. This is addressed in many methods by surveying multiple samples of the same class, and probing multiple points on the same gene. Methods that do this effectively are preferred to those that do not.

2.3. Sensitivity

The ability to detect low-abundance transcripts is an integral part of gene discovery programs. Low-abundance transcripts, in principle, have properties that are of particular importance to the study of complex organisms. Rare transcripts frequently encode for proteins of low physiologic concentrations that in many cases make them potent by their very nature. Erythropoietin is a classic example of such a rare transcript. Amgen scientists functionally cloned erythropoietin long before it appeared in the public expressed sequence tag (EST) database. Genes are frequently discovered in the order of transcript abundance, and a simple analysis of EST databases correctly reveals high, medium, and low abundance transcripts by a direct correlation of the number of occurrences in that

database (data not shown). Thus, using a technology that is more sensitive has the potential to identify novel transcripts even in a well-studied system.

Sensitivity values are quoted in publications for available technologies at concentrations of 1 part in 50,000 to 1 part in 500,000. The interpretation of these data, however, should be made cautiously both upon examination of the method in which the sensitivity was determined, as well as the sensitivity needed for the intended use. For example, if one intends to study appetite-signaling factors and uses an entire rat brain for expression analysis, the dilution of the target cells of anywhere from 1 part in 10,000 to 1 part in 100,000 allows for only the most abundant transcripts in the rare cells to be measured, even with the most sensitive technology available. Reliance on cell models to do the same type of analysis, where possible, suffers the confounding variable that isolated cells or cell lines may respond differently in culture at the level of gene expression. An ideal scenario would be to carefully micro dissect or sort the cells of interest and study them directly, provided enough samples can be obtained.

In addition to the ability of a technology to measure rare transcripts, the sensitivity to discern small differentials between transcripts must be considered. The differential sensitivity limit has been reported for a variety of techniques ranging from 1.5-fold to 5-fold, so the user must determine how important small modulations are to the overall project and choose the technology while taking this property into account as well.

2.4. Sample Requirement

The requirement for studying transcript abundance levels is a cell or tissue substrate, and the amount of such material needed for analysis can be prohibitively high with many technologies in many model systems. To use the above example, dozens of dissected rat hypothalami may be required to perform a global gene expression study, depending on the quantitating technology chosen. Samples procured by laser-capture microdissection can only be used in the measuring of a small number of transcripts and only with some technologies, or must be subjected to amplification technologies, which risk artificially altering transcript ratios.

2.5. Coverage

For open architecture systems where the objective is to profile as many transcripts as possible and identify new genes, the number of independent transcripts being measured is an important metric. However, this is one of the most difficult parameters to measure, because determining what fraction of unknown transcripts is missing is not possible. Despite this difficulty, predictive models can be made to suggest coverage, and the intuitive understanding of the technology is a good gage for the relevance and accuracy of the predictive model.

The problem of incomplete coverage is perhaps one of the most embarrassing examples of why hundreds of scientific publications were produced in the 1970's and 1980's having relatively little value. Many of these papers reported the identification of a single differentially expressed gene in some model system and expounded upon the overwhelmingly important new biological pathway uncovered. Modern analysis has demonstrated that even in the most similar biological systems or states, finding 1% of transcripts with differences is common, with this number increasing to 20% of transcripts or more for systems when major changes in growth or activation state are signaled. In fact, the activation of a single transcription factor can induce the expression of hundreds of genes. Any given abundantly altered transcript without an understanding of what other transcripts are altered, is similar to independent observers describing the small part of an elephant that they can see. The person looking at the trunk describes the elephant as long and thin, the person observing an ear believes it to be flat, soft and furry, and the observer examining a foot describes the elephant as hard and wrinkly. Seeing the list of the majority of transcripts that are altered in a system is like looking at the entire elephant, and only then can it be accurately described. Separating the key regulatory genes on a gene list from the irrelevant changes remains one of the biggest challenges in the use of transcript profiling.

2.6. Throughput

The throughput of the technology, as defined by the number of transcript samples measured per unit time, is an important consideration for some projects. When quick turnaround is desired, it is impractical to print microarrays, but where large numbers of data points need to be generated, techniques where individual reactions are required are impractical. Where large experiments on new models generate significant expense, it may be practical to perform a higher throughput, lower quality assay as a control prior to a large investment. For example, prior to conducting a comprehensive gene profiling experiment in a drug dose-response model, it might be practical to first use a low throughput technique to determine the relevance of the samples prior to making the investment with the more comprehensive analysis.

2.7. Cost

Cost can be an important driver in the decision of which technologies to employ. For some methods, substantial capital investment is required to obtain the equipment needed to generate the data. Thus, one must determine whether a microarray scanner or a capillary electrophoresis machine is obtainable, or if X-ray film and a developer need to suffice. It should be noted that as large companies change platforms, used equipment becomes available at prices dramati-

cally less than those for brand new models. In some cases, homemade equipment can serve the purpose as well as commercial apparatuses at a fraction of the price.

2.8. Reproducibility

It is desired to produce consistent data that can be trusted, but there is more value to highly reproducible data than merely the ability to feel confident about the conclusions one draws from them. The ability to forward-integrate the findings of a project and to compare results achieved today with results achieved next year and last year, without having to repeat the experiments, is key to managing large projects successfully. Changing transcript-profiling technologies often results in datasets that are not directly comparable, so deciding upon and persevering with a particular technology has great value to the analysis of data in aggregate. An excellent example of this is with the serial analysis of gene expression (SAGE) technique, where directly comparable data have been generated by many investigators over the course of decades and are available online (<http://www.ncbi.nlm.nih.gov>).

2.9. Data Management

Management and analysis of data is the natural continuation to the discussion of reproducibility and integration. Some techniques, like differential display, produce complex data sets that are neither reproducible enough for subsequent comparisons, nor easily digitized. Microarray and GeneCalling data, however, can be obtained with software packages that determine the statistical significance of the findings and even can organize the findings by molecular function or biochemical pathways. Such tools offer a substantial advance in the generation of accretive data. The field of bioinformatics is flourishing as the number of data points generated by high throughput technologies has rapidly exceeded the number of biologists to analyze the data.

Reference

1. Ushkaryov, Y. A. and Sudhof, T. C. (1993) Neurexin III α : extensive alternative splicing generates membrane-bound and soluble forms. *Proc. Natl. Acad. Sci. USA* **90**, 6410–6414.

Gene Expression Quantitation Technology Summary

Richard A. Shimkets

Summary

Scientists routinely talk and write about gene expression and the abundance of transcripts, but in reality they extrapolate this information from the various measurements that a variety of different technologies provide. Indeed, there are many reasons why applying different technologies to the problem of transcript abundance may give different results, owing to an incomplete understanding of the gene in question or from shortcomings in the applications of the technologies. There are nine basic considerations for making a technology choice for quantitating gene expression that will impact the overall outcome: architecture, specificity, sensitivity, sample requirement, coverage, throughput, cost, reproducibility, and data management. These considerations will be discussed in the context of available technologies.

Key Words: Architecture, bioinformatics, coverage, quantitative, reproducibility, sensitivity, specificity, throughput

1. Introduction

Owing to the intense interest of many groups in determining transcript levels in a variety of biological systems, there are a large number of methods that have been described for gene-expression profiling. Although the actual catalog of all techniques developed is quite extensive, there are many variations on similar themes, and thus we have reduced what we present here to those techniques that represent a distinct technical concept. Within these groups, we discovered that there are methods that are no longer applied in the scientific community, not even in the inventor's laboratory. Thus, we have chosen to focus the methods chapters of this volume on techniques that are in common use in the community

at the time of this writing. This work also introduces two novel technologies, SEM-PCR and the Invader Assay, that have not been described previously. Although these methods have not yet been formally peer-reviewed by the scientific community, we feel these approaches merit serious consideration.

In general, methods for determining transcript levels can be based on transcript visualization, transcript hybridization, or transcript sequencing (**Table 1**).

The principle of transcript visualization methods is to generate transcripts with some visible label, such as radioactivity or fluorescent dyes, to separate the different transcripts present, and then to quantify by virtue of the label the relative amount of each transcript present. Real-time methods for measuring label while a transcript is in the process of being linearly amplified offer an advantage in some cases over methods where a single time-point is measured. Many of these methods employ the polymerase chain reaction (PCR), which is an effective way of increasing copies of rare transcripts and thus making the techniques more sensitive than those without amplification steps. The risk to any amplification step, however, is the introduction of amplification biases that occur when different primer sets are used or when different sequences are amplified. For example, two different genes amplified with gene-specific primer sets in adjacent reactions may be at the same abundance level, but because of a thermodynamic advantage of one primer set over the other, one of the genes might give a more robust signal. This property is a challenge to control, except by multiple independent measurements of the same gene. In addition, two allelic variants of the same gene may amplify differently if the polymorphism affects the secondary structure of the amplified fragment, and thus an incorrect result may be achieved by the genetic variation in the system. As one can imagine, transcript visualization methods do not provide an absolute quantity of transcripts per cell, but are most useful in comparing transcript abundance among multiple states.

Transcript hybridization methods have a different set of advantages and disadvantages. Most hybridization methods utilize a solid substrate, such as a microarray, on which DNA sequences are immobilized and then labeled. Test DNA or RNA is annealed to the solid support and the locations and intensities on the solid support are measured. In another embodiment, transcripts present in two samples at the same levels are removed in solution, and only those present at differential levels are recovered. This suppression subtractive hybridization method can identify novel genes, unlike hybridizing to a solid support where information generated is limited to the gene sequences placed on the array. Limitations to hybridization are those of specificity and sensitivity. In addition, the position of the probe sequence, typically 20–60 nucleotides in length, is critical to the detection of a single or multiple splice variants. Hybridization methods employing cDNA libraries instead of synthetic oligonucleotides give

inconsistent results, such as variations in splicing and not allowing for the testing of the levels of putative transcripts predicted from genomic DNA sequence.

Hybridization specificity can be addressed directly when the genome sequence of the organism is known, because oligonucleotides can be designed specifically to detect a single gene and to exclude the detection of related genes. In the absence of this information, the oligonucleotides cannot be designed to assure specificity, but there are some guidelines that lead to success. Protein-coding regions are more conserved at the nucleotide level than untranslated regions, so avoiding translated regions in favor of regions less likely to be conserved is useful. However, a substantial amount of alternative splicing occurs immediately distal to the 3' untranslated region and thus designing in proximity to regions following the termination codon may be ideal in many cases. Regions containing repetitive elements, which may occur in the untranslated regions of transcripts, should be avoided.

Several issues make the measurement of transcript levels by hybridization a relative measurement and not an absolute measurement. Those experienced with hybridization reactions recognize the different properties of sequences annealing to their complementary sequences, and thus empirical optimization of temperatures and wash conditions have been integrated into these methods.

Principle disadvantages to hybridization methods, in addition to those of any closed system, center around the analysis of what is actually being measured. Typically, small regions are probed and if an oligonucleotide is designed to a region that is common to multiple transcripts or splice variants, the resulting intensity values may be misleading. If the oligonucleotide is designed to an exon that is not used in one sample of a comparison, the results will indicate lack of expression, which is incorrect. In addition, hybridization methods may be less sensitive and may yield a negative result when a positive result is clearly present through visualization.

The final class of technologies that measure transcript levels, transcript sequencing, and counting methods can provide absolute levels of a transcript in a cell. These methods involve capturing the identical piece of all genes of interest, typically the 3' end of the transcript, and sequencing a small piece. The number of times each piece was sequenced can be a direct measurement of the abundance of that transcript in that sample. In addition to absolute measurement, other principle advantages of this method include the simplicity of data integration and analysis and a general lack of problems with similar or overlapping transcripts. Principle disadvantages include time and cost, as well as the fact that determining the identity of a novel gene by only the 10-nucleotide tag is not trivial.

We would like to mention two additional considerations before providing detailed descriptions of the most popular techniques. The first is contamination

Table 1
Common Gene Expression Profiling Methods

Technique	Class	Architecture	Kits Available	Service Available	Detect Alt. Splicing	Detect SNPs
5'-nuclease assay/real-time RT-PCR	Visualization	Open	Yes	No	No	No
AFLP (amplified-fragment length polymorphism fingerprinting)	Visualization	Open	No	No	No	Yes
Antisense display	Visualization	Open	No	No	No	No
DDRT-PCR	Visualization	Open	Yes	No	No	No
(differential display RT-PCR)						
DEPD (digital expression pattern display)	Visualization	Open	No	No	Yes	No
Differential hybridization (differential cDNA library screening)	Hybridization	Open	No	No	No	No
DSC (differential subtraction chain)	Hybridization	Open	No	No	No	No
GeneCalling	Visualization	Open	No	Yes	Yes	Yes
In situ Hybridization	Hybridization	Closed	Yes	No	No	No
Invader Assay	Visualization	Closed	Yes	Yes	No	Yes
Microarray hybridization	Hybridization	Closed	Yes	Yes	No	No
Molecular indexing (and computational methods)	Visualization	Open	No	No	No	No
MPSS (massively parallel signature sequencing)	Sequencing	Open	No	No	No	No
Northern-Blotting (Dot-/Slot-Blotting)	Hybridization	Closed	Yes	No	No	No
Nuclear run on assay/nuclease S1 analysis	Visualization	Closed	Yes	No	No	No
ODD (ordered differential display)	Visualization	Open	No	No	No	No
Quantitative RT-PCR	Visualization	Closed	Yes	Yes	No	No