DNA microarrays: a powerful new tool for analysis of the virus–host interaction

Bret A. Wing, Edward P. Browne and Thomas Shenk

DNA microarray technology enables the expression profiles of thousands of genes to be assayed in parallel. Recently, DNA microarrays have been utilized to study the modulation of host-cell gene expression during virus infection. These experiments have led to the identification of a number of hitherto uncharacterized changes in host-cell gene expression that occurs in the presence of virus. Understanding how these changes impact both virus and host should lead to an increased understanding of the mechanisms of virus replication, host anti-viral responses and viral pathogenesis.

Bret A. Wing, Edward P. Browne and Thomas Shenk*
Dept of Molecular Biology
Princeton University
Princeton
NJ 08544, USA
*tel: +1 609 258 5992
fax: +1 609 258 1704
e-mail: tshenk@princeton.edu

▼ The physiological consequences of viral infections are complex and variable. Some viruses, such as influenza, are highly pathogenic, rapidly produce disease and sometimes kill their host. Others, such as human immunodeficiency virus (HIV), are more subtle pathogens and cause a slow progressing pathology. However, other viruses, such as cytomegalovirus, can reside within a healthy host for many years without resulting in overt disease.

This range of viral behaviors is also evident at the cellular level. Some viruses rapidly shut down the expression of most host-cell genes, ultimately killing their host, whereas others reside within an infected cell on a long-term basis. Host organisms and cells often respond to viral infection by instituting anti-viral measures. Cytokines and the immune system impact dramatically on infected cells, and individual cells can initiate anti-viral responses, including the induction of apotosis.

DNA array technology provides an opportunity to identify and dissect aspects of the multifaceted and dynamic interaction between host and pathogen. The identification of cellular genes whose mRNA levels change during infection, will inevitably provide new insights into not only viral replication and spread strategies but also pathophysiological changes and host anti-viral responses.

Types and relative advantages of different microarray technologies

There are three basic types of DNA array technology available: Affymetrix gene chips, glass slide microarrays and membrane-based arrays. All three technologies employ modified RNA:DNA or DNA:DNA hybridization protocols to measure expression levels of large numbers of genes in parallel. Affymetrix gene chips, and glass slide cDNA/oligonucleotide arrays utilize fluorescent probes and fluorescent scanning microscopy to quantify hybridization signals. By contrast, membrane-based arrays utilize probes labeled with a radioisotope and use autoradiography to measure hybridization signals. Whereas Affymetrix and glass slide arrays typically contain oligonucleotides or cDNAs corresponding to 12,000 or more genes, membrane-based arrays are less dense and generally contain DNAs corresponding to significantly fewer genes.

Each type of array offers specific advantages and disadvantages depending on the nature of the experiment. The high-density DNAs on Affymetix and glass slide arrays make these technologies particularly useful for more comprehensive studies, or where maximum potential for discovery of novel genes is important. For experiments requiring analysis of genes not found on commercial

1359-6446/01/$ – see front matter PII: S1471-1931(01)00171-9

gene chips, custom glass slide microarrays can be synthesized using commercially available or home-built microarray printers. For example, we have made glass slide arrays containing probes for the 178 predicted open reading frames (ORF) encoded by mouse cytomegalovirus for our virus–host interaction studies (Bret Wing and Thomas Shenk, unpublished results). Glass slide microarrays also offer the advantage of two-color fluorescent analysis. Probes are produced from experimental and baseline RNA samples by labeling with different fluorescent dyes and then hybridized to the same gene chip, with the results for each gene being expressed as a ratio of the fluorescent signals. By contrast, Affymetrix and membrane-based microarrays both require hybridization of experimental and control samples to different arrays, increasing the number of arrays required for each experiment.

Affymetrix arrays are expensive in themselves, and both Affymetrix and glass slide technologies require expensive equipment for array printing and/or post-hybridization array processing, and fluorescent signal acquisition. If the focus of the experiment is narrower, or specifically involves the study of genes that are known to play a role in well-studied biological processes (e.g. cell cycle regulation and apoptosis), membrane-based arrays can provide a more cost-effective alternative, as most laboratories are already set up to perform and analyze radioisotope-based assays.

A final concern, which applies to all microarray technologies, is the question of reproducibility and cross comparison. Generally, microarray experiments need to be repeated at least 2–3 times to obtain a reliable dataset, and even then their reliability is primarily limited to whether or not the expression of a gene is increased or decreased in the experimental sample compared with the control. Absolute levels of transcripts tend to be underestimated when compared with results obtained using standard techniques such as northern blot or RNase protection assays. In addition, it is also difficult to make cross comparisons of results obtained using different microarray technologies, owing to the inherent differences in sensitivity of each system. Thus, one challenge facing the microarray community is the need for a set of standards that enable meaningful cross comparison of data. To this end, the establishment of a gene expression database for comparison of results obtained in different systems would appear to be an essential first step.

Proof of concept: DNA array results in different virus systems

Historically, DNA arrays have primarily been used to generate catalogs of cellular genes whose steady-state mRNA levels change in response to infection. Such changes have been documented in human cytomegalovirus (HCMV)-infected fibroblasts[1], HIV type 1-infected T cells[2], poliovirus-infected HeLa cells[3], echovirus type 1-infected human osteogenic sarcoma cells[4], human foamy virus-infected embryonic lung cells[5], human T cell leukemia type 1 virus-infected T cells[6], herpes simplex virus (HSV) type 1-infected human embryonic lung cells[7], and Merek's disease virus-infected chicken embryo fibroblasts[8]. Similarly, differences in gene expression were observed for human papillomavirus type 31-immortalized cells as compared with normal keratinocytes[9]. These studies clearly demonstrate changes in the expression of host cell genes, subsequent to virus infection or immortalization, as reflected by altered mRNA levels. However, they fail to distinguish between primary effects of viral proteins and secondary effects mediated by altered cellular genes. It is possible that after a cellular gene is altered, it subsequently affects the expression or stability of other cellular mRNAs.

It is difficult to infer generalizations from these early studies because different viruses were used to infect a variety of different cultured cells, some of which were actively growing and some of which were residing in the G_0 compartment of the cell cycle. Furthermore, some experiments employed purified virions as a source of virus whereas others used crude virus stocks (i.e. infected-cell lysates). Infection with cell lysates could generate misleading results as the lysate can contain non-viral constituents such as cytokines, which dramatically modulate cellular gene expression.

The effects of several individual viral gene products on cellular mRNA levels in cultured cells have also been monitored using DNA arrays. The hepatitis B virus X protein[10] and the hepatitis C non-structural proteins[11] have been studied in HepG2 cells, and the human herpes virus 8 latency-associated nuclear antigen (LANA)[12] has been studied in BJAB cells (a Burkitt's lymphoma cell line). Each of these experiments identified cellular genes that were induced in the cells expressing viral genes. However, as the authors of both papers note, a weakness in the experimental design for analysis of the X and LANA proteins, is that these proteins are constitutively expressed in cells, forcing the comparison of different cell clones. Consequently, the changes observed might not be caused by the viral proteins. The experiment with hepatitis C non-structural proteins benefited from the use of a promoter that was conditionally regulated by tetracycline[11]. This enabled the comparison of the same cell line before and after induction, increasing the likelihood that changes resulted from the action of the viral proteins.

Finally, cellular mRNA expression in both coxsackievirus B3-infected murine myocardium[13], and human liver tissue chronically infected with hepatitis B and C, was compared with that in normal tissues[14]. A variety of changes were identified in each case. These analyses are, inherently, more biologically

relevant than experiments in cultured cells, but they are also more complex and difficult to interpret. If all cells are not infected in a tissue sample, it is very difficult to identify mRNAs whose level is decreased in response to infection. In addition, the samples contain multiple cell types, which might respond to infection differently. Furthermore, the cells can be infiltrated with inflammatory cells, whose gene expression profiles might confound interpretation of the results. It is important to dissect and identify cell types and determine their infection status to more effectively evaluate the *in vivo* consequences of infections. However, early experiments already demonstrate that mRNA profiles will significantly advance our understanding of viral pathogenesis.

Modulation of interferon-responsive gene expression: a common theme

Expression of cellular genes could change in response to infection for one of three reasons. First, the level of a cellular mRNA might increase or decrease as part of an active cellular anti-viral response. Second, the level of a cellular mRNA might be induced or repressed by a virus in order to facilitate its replication or spread. Finally, changes can result from bystander effects. For example, some cellular mRNAs could change as part of a set of coordinated, regulated genes that are turned on or off, but the change in levels of many or most mRNAs in the set might have no effect on either viral replication or the survival of the cell or host organism.

The DNA array experiments with human cytomegalovirus[1] and HSV-1 (Ref. 7) (two distantly related members of the herpes virus family), provide an example of the first category of change. Although cytomegalovirus infection induced changes in the expression profile of a much larger number of mRNAs in human fibroblasts than HSV-1 did in human embryonic lung cells, both viruses induced the accumulation of numerous mRNAs encoded by interferon-responsive genes. The induction of interferon-responsive genes by cytomegalovirus was first noted by Zhu *et al.*[15] in a differential display analysis, and the induction appears to result, at least in part, from the interaction of a virion glycoprotein with an unknown cell surface receptor[16]. Mossman *et al.*[7] showed that infection with HSV activates the expression of interferon-responsive genes, but a virus-coded product rapidly blocks the accumulation of most of these mRNAs. We have recently found that the same appears to be true for cytomegalovirus, although some interferon-responsive mRNAs continue to be present at induced levels (Edward Browne, Bret Wing and Thomas Shenk, unpublished results). These DNA array analyses have therefore highlighted a cellular anti-viral measure (induction of interferon-responsive genes) and a herpes virus counter-measure (inhibition of continued expression). It will be of great interest to identify the mechanism underlying the subversion of the interferon response by the two herpes viruses.

Given the anti-viral effects of well-characterized interferon-responsive genes, it is somewhat surprising that a subset of these genes continue to be expressed during cytomegalovirus infection (Edward Browne, Bret Wing and Thomas Shenk, unpublished results). It is not clear to what extent the sustained expression of some interferon-responsive genes represent a host response to infection that is thwarted by the virus using a post-transcriptional mechanism, or whether the response is somehow beneficial to, and perhaps promoted by, the virus. Interestingly, the interferon-inducible *p204* gene, a member of the interferon-inducible (*Ifi*) 200 gene family (which plays a role in the anti-proliferative effects of interferon-α), shows increased expression in DNA microarray analysis of murine cytomegalovirus-infected cells (Bret Wing, Edward Browne and Thomas Shenk, unpublished results). It was recently reported that inactivation of p204 inhibits replication of the virus[17], suggesting that *p204* induction might be beneficial to the virus. In a similar vein, interferon-γ-responsive human leukocyte antigen E (HLA-E) mRNA is induced by human cytomegalovirus[1], and this is expected to protect infected cells from natural killer (NK) cell-mediated lysis. NK cells recognize and spare target cells expressing this non-classical class I molecule on their surface[18]. Thus, it would appear, paradoxically, that at least some of the interferon-responsive genes induced during cytomegalovirus infection actually facilitate viral replication.

Other viruses have developed mechanisms to limit the expression of interferon-inducible genes. In the case of human papillomavirus, interferon-responsive mRNAs were among the most significantly decreased[9]. Induction of these genes in human papillomavirus type 31-immortalized cells was also resistant to the addition of exogenous interferon, indicating that the virus has developed a highly effective mechanism for preventing the induction of interferon-responsive genes.

Prostaglandin E_2: a cellular mediator of viral replication

The cytomegalovirus array analysis[1] also provides an example of a change in cellular gene expression that appears to facilitate viral replication. Multiple constituents of the pathway that produces prostaglandin E_2 are modulated in human fibroblasts after infection with cytomegalovirus[1]. Cyclooxygenase-2 (COX-2) and cytosolic phospholipase A_2 ($cPLA_2$) mRNAs are significantly elevated, and lipocortin-1 mRNA is significantly decreased after infection. When $cPLA_2$ is activated by phosphorylation, it translocates to the membranes where it selectively cleaves and releases arachidonic acid. COX-2 converts arachidonic acid to prostaglandin H_2, which is rapidly converted to prostaglandin E_2 in fibroblasts. Lipocortin-1 inhibits

the activation of $cPLA_2$ (Ref. 19). Thus, mRNA levels show that cytomegalovirus infection induces the synthesis of prostaglandin E_2 by elevating $cPLA_2$ and COX-2 levels, and reducing the amount of the negative regulator lipocortin-1. Furthermore, HCMV infection has been shown to activate latent $cPLA_2$ by inducing its phosphorylation[20].

Consistent with the altered mRNA levels, we have observed that the level of prostaglandin E_2 is dramatically induced in cytomegalovirus-infected human fibroblasts, and that this induction can be blocked by specific COX-2 inhibitors (Hua Zhu, Jian-Ping Cong, Wade Bresnahan and Thomas Shenk, unpublished results). Concentrations of these drugs that completely block the accumulation of prostaglandin E_2 in response to cytomegalovirus infection, also substantially inhibit the replication of cytomegalovirus in human fibroblasts. Importantly, addition of exogenous prostaglandin E_2 to drug-treated cultures substantially reduces the block on viral replication. Consequently, it appears that the induction of prostaglandin E_2 accumulation is required for the efficient replication of cytomegalovirus in fibroblasts.

Prostaglandins serve as second messengers to provoke a variety of responses, including inflammation. Perhaps the activation of this pathway is a cellular reaction to cytomegalovirus infection, which is intended to induce a cell-mediated response that will kill the infected cell, thereby inhibiting spread of the infection. If this is the case, it appears that the virus weakens this response. Prostaglandins can induce transcription, and it is possible that cytomegalovirus uses this pathway to aid in the activation of its own genes. Experiments are in progress to test this hypothesis.

Implications for drug development

The anti-viral drugs that are currently available inhibit virus-coded products, such as polymerases or proteases. Viral proteins have been chosen as targets for drug development, in part, because few cell-coded candidates have been identified, but also because inhibition of virus-coded proteins should provide specificity, reducing toxicity to the host. However, therapeutics that target viral proteins often suffer from the rapid emergence of drug-resistant variants. The large size of viral populations and a lack of proofreading function in some viral polymerases favor the accumulation of variants with mutations that make them less sensitive to inhibition by drugs. Cellular genes, whose products facilitate viral pathogenesis, are potential targets for anti-viral therapeutics. Compounds that inhibit cellular functions are, of course, more likely to be toxic than drugs blocking virus-coded activities. However, compounds that target host-cell functions probably share an important advantage in that viral populations will less readily develop resistance.

Concluding remarks

With the availability of complete genome sequences, microarrays will soon become available that represent nearly-complete transcriptomes. Consequently, it will be possible to monitor the expression of a full set of host genes as they change in response to infection. This will inevitably provide not only important new insights into the virus–host cell interaction but also advance our understanding of viral pathogenesis and identify new targets for anti-viral drug development.

One of the weaknesses of the early efforts to produce transcription profiles for infected cells and organisms is the lack of a facile means for cross comparison of datasets. Comparative analyses will help to distinguish the strategies for replication and spread that is employed by different viruses, and will help to elucidate the anti-viral countermeasures employed by infected host cells and organisms, thereby distinguishing generalized from pathogen-specific strategies. A specialized database, such as that established for the yeast genome[21,22], could markedly improve the comparability of microarray data in infectious disease by providing clear descriptions of experimental designs and including raw data that can be re-analyzed for comparative purposes.

References

1 Zhu, H. *et al.* (1998) Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14470–14475

2 Geiss, G.K. *et al.* (2000) Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology* 266, 8–16

3 Johannes, G. *et al.* (1999) Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microassay. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13118–13123

4 Pietiäinen, V. *et al.* (2000) Effects of echovirus 1 infection on cellular gene expression. *Virology* 276, 243–250

5 Wagner, A.A. *et al.* (2000) Induction of cellular genes is mediated by the *Bel1* transactivator in foamy virus-infected cells. *J. Virol.* 74, 4441–4447

6 de La Fuente, C. *et al.* (2000) Gene expression array of HTLV type 1-infected T cells: upregulation of transcription factors and cell cycle genes. *AIDS Res. Hum. Retroviruses* 16, 1695–1700

7 Mossman, K.L. *et al.* (2001) Herpes simplex virus triggers and then disarms a host antiviral response. *J. Virol.* 75, 750–758

8 Morgan, R.W. *et al.* (2001) Induction of host gene expression following infection of chicken embryo fibroblasts with oncogenic Merek's disease virus. *J. Virol.* 75, 533–539

9 Chang, Y.E. and Laimins, L.A. (2000) Microarray analysis identifies interferon-inducible genes and *Stat-1* as major transcriptional targets of human papillomavirus type 31. *J. Virol.* 74, 4174–4182

10 Han, J. *et al*. (2000) Selective transcriptional regulation in the human liver cell by hepatitis B viral X protein. *Biochem. Biophys. Res. Commun.* 272, 525–530
11 Huang, Y. *et al*. (2001) A human hepatoma cell line expressing hepatitis C virus nonstructural proteins tightly regulated by tetracycline. *Biochem. Biophys. Res. Commun*. 281, 732–740
12 Renne, R. *et al*. (2001) Modulation of cellular and viral gene expression by the latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus. *J. Virol.* 75, 458–468
13 Taylor, L.A. *et al*. (2000) Host gene regulation during coxsackievirus B3 infection in mice: assessment by microarrays. *Circ. Res.* 87, 328–334
14 Honda, M. *et al*. (2001) Differential gene expression between chronic hepatitis B and C hepatic lesions. *Gastroenterology* 120, 955–966
15 Zhu, H. *et al*. (1997) Use of differential display analysis to assess the effect of human cytomegalovirus infection on the accumulation of cellular mRNAs: induction of interferon-responsive RNAs. *Proc. Natl. Acad. Sci. U. S. A.* 94, 13985–13990
16 Boyle, K.A. *et al*. (1999) Engagement of the cellular receptor for glycoprotein B of human cytomegalovirus activates the interferon-responsive pathway. *Mol. Cell. Biol.* 19, 3607–3613
17 Hertel, L. *et al*. (1999) The interferon-inducible 204 gene, a member of the Ifi 200 family, is not involved in the antiviral state induction by IFN-alpha, but is required by the mouse cytomegalovirus for its replication. *Virology* 262, 1–8
18 Borrego, F. *et al*. (1998) Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell mediated lysis. *J. Exp. Med*. 187, 813–818
19 Croxtall, J.D. *et al*. (1996) Lipocortin 1 and the control of cPLA2 activity in A549 cells: glucocorticoids block EGF stimulation of cPLA2 phosphorylation. *Biochem. Pharmacol*. 52, 351–356
20 Shibutani, T. *et al*. (1997) Pertussis toxin-sensitive G proteins as mediators of the signal transduction pathways activated by cytomegalovirus infection of smooth muscle cells. *J. Clin. Invest*. 100, 2054–2061
21 Aach, J. *et al*. (2000) Systematic management and analysis of yeast genome expression data. *Genome Res.* 10, 431–445
22 Ball, C.A. *et al*. (2001) *Saccharomyces* genome database provides tools to survey gene expression and functional analysis data. *Nucl. Acids Res*. 29, 80–81