

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

U.S. Department of Veterans Affairs Staff
Publications

U.S. Department of Veterans Affairs

4-2000

How the host 'sees' pathogens: global gene expression responses to infection

Ian D. Manger

Stanford University School of Medicine

David A. Relman

Stanford University School of Medicine, relman@stanford.edu

Follow this and additional works at: <http://digitalcommons.unl.edu/veterans>

Manger, Ian D. and Relman, David A., "How the host 'sees' pathogens: global gene expression responses to infection" (2000). *U.S. Department of Veterans Affairs Staff Publications*. 6.

<http://digitalcommons.unl.edu/veterans/6>

This Article is brought to you for free and open access by the U.S. Department of Veterans Affairs at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in U.S. Department of Veterans Affairs Staff Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

How the host 'sees' pathogens: global gene expression responses to infection

Ian D Manger and David A Relman

Innate immune responses to pathogens are believed to be patterned and stereotyped. Adaptive responses display variety but in relatively few types of products and with limited numbers of mechanisms. Is this apparent disparity between microbial pathogen diversity and a restricted set of host responses an accurate picture of infection or is it the result of a limited collection of analytic tools? DNA microarray technology permits one to address simple descriptive questions about global gene expression inside cells. In particular, it offers an opportunity to examine the relationship between host and pathogen in much greater detail than has been possible previously. One can now ask, firstly, how a host cell or organism 'sees' a microbial pathogen from the viewpoint of gene expression responses and, secondly, at what level it is able to discriminate between different agents. Other potential insights to be reaped include the identification of microbial determinants of the host response, the temporal features of the 'conversation' between host and pathogen, novel strategies for therapeutic and prophylactic intervention and prognostic markers of outcome.

Addresses

Departments of Medicine, and Microbiology and Immunology, Stanford University, 299 Campus Drive, Room D300, Stanford, CA 94305-5124, USA; Veterans Affairs Palo Alto Health Care System, 3801 Miranda Avenue, 154T, Palo Alto, CA 94304, USA
Correspondence: David A Relman; e-mail: relman@cmgm.stanford.edu

Current Opinion in Immunology 2000, **12**:215-218

0952-7915/00/\$ – see front matter. Published by Elsevier Science Ltd.

Abbreviations

CMV cytomegalovirus
LPS lipopolysaccharide

Introduction

The interactions between a host and microbial pathogens are diverse, choreographed and regulated. At one level, the features and outcomes of these interactions can be assigned to broad, generic categories, based on the predominant type of locally recruited leukocyte or T lymphocyte subset or on the profile of elicited cytokines. The frequently described 'acute phase response' and the recent focus on the recognition of pathogen-associated molecular patterns (e.g. lipopolysaccharide [LPS] lipid A, lipoteichoic acid, peptidoglycan and bacterial lipopeptides) by the innate immune system underscore the commonalities of the host response to microbial pathogens. The molecular description of microbial pathogenesis likewise reveals a number of common themes that involve families of structurally and functionally related adherence factors, toxins, secretion systems and regulators of microbial gene expression. On the other hand, detailed analysis of any specific interactions between pathogen and host uncovers unique mechanisms and molecules. Each pathogen evolves its own particular

strategy for manipulation of the host that optimizes microbial survival and transmission. Those that find themselves in a new host without an opportunity for co-adaptation are more likely to elicit a more dramatic set of host responses. Thus, in the context of this discussion, to what degree is the response of the host to any microbial pathogen stereotyped and limited to certain patterned scripts? Does one find a greater degree of discrimination by the host if one looks at these responses in more detail? One can certainly argue that the selection of time and place during the course of the interaction affects the picture so obtained and might reflect in part differences between innate and adaptive immune responses.

High density DNA microarrays have the potential to offer an unparalleled view of the transcriptional events that underlie the host response to microbial pathogens. This view is but one of several perspectives on this process; other perspectives would include those that focus upon global host cellular protein expression [1], the global set of interactions between cellular proteins, a profile of protein activation states and spatial arrangement of structural and signaling molecules within a cell. Currently available microarray technology can provide a quantitative description of the behavior of tens of thousands of genes and promises within the near future to encompass the complete transcriptome of humans and model organisms. To date, however, few reports of the application of this technology to the exploration of interactions between host and pathogen have been published. These reports are largely limited to 'proof-of-concept' experiments that catalog genes whose transcriptional activities vary most widely under defined conditions. As the technology becomes more widely available, exploration of the relationships between individual pathogens and their target cells, of the development of the proinflammatory response and of cell recruitment, activation and maturation will enable development of a detailed picture of the transcriptional events that accompany the innate and adaptive responses.

Computational tools offer new ways to visualize microarray-based data and to reveal possible gene functions and interactions. They currently provide the capacity to identify gene expression patterns that reflect common themes as well as unique features of important human diseases. As these tools are improved and developed further, we anticipate that the detailed description of the interaction of infectious agents and a variety of 'target' host cells, including immune cells, will have a considerable impact on the diagnosis and therapy of infectious diseases. In the following brief review, we restrict our comments to the study of host response to infection using DNA microarrays and not to the use of microarrays to study global microbial responses.

The latter field is also evolving quickly through its infancy, as complete genome sequences from a rapidly increasing list of agents become widely available (<http://www.tigr.org/tdb/mdb/mdb.html>) and methods for efficient generation and comprehensive display of all microbial ORFs (open reading frames) are quickly developed [2]. A few early papers suggest the significant potential power of this approach [3,4]. Microbial genome structural (polymorphism) analysis is another potential application of microbial microarray-based approaches [5,6].

Microarray techniques

The techniques for fabricating materials and reagents, and probing and gathering data from oligonucleotide or cDNA microarrays have been reviewed extensively in the past year and discussion of these topics lies outside of the scope of this article. The interested reader should refer to [7–9,10*,11,12] and references therein for further details. At this early stage in the evolution of this technology it is difficult to predict whether one particular format or approach will provide consistent advantages over others. However, the abilities to examine responses from a large number of genes — including those without a known role in immune response — and to compare multiple methods for pattern recognition may be especially important. A disproportionate emphasis has been placed so far on a small number of genes that are strongly induced in response to many different inflammatory stimuli, for example genes that encode the known cytokines.

What happens upon stimulation/infection of host cells?

At the present time, the majority of examples of DNA microarray application to the study of the immune responses are limited to manipulated immune cell activation with a variety of generic stimulants and the exploration of subsequent signaling events. For example, Der *et al.* [13] stimulated HT1080 fibrosarcoma cells with interferon- α , - β and - γ for six hours and identified more than 1000 genes (out of more than 6800 studied) whose mRNA expression levels varied after treatment with these agents. Application of simple filters to the data enabled the authors to obtain a list of 122 genes whose expression varied most widely. Of these, some 40 genes had previously been identified as interferon-stimulated genes (ISGs) whereas the remainder represented novel ISGs. Similar experiments examining responses to IL-4 and IL-2 [14] and T cell responses to heat shock or phorbol esters [15] have also been reported. It is important, however, to remember that the magnitude of the expression response may not be the most useful indicator of biological significance.

Zhu *et al.* [16*] have reported one of the first analyses of the interaction of a human pathogen with a host cell using microarrays. Molecular dissection of the genome of human cytomegalovirus (CMV) had shown that the virus encodes a number of putative products that would be predicted to alter gene expression in primary human foreskin fibroblasts (HFF) and that the virus is known to perturb the cell cycle. These investigators catalogued 258 human genes (of a total

of approximately 6600 studied) whose expression levels varied 4-fold (up or down) early in the virus replication cycle and for a number of cases verified the measured levels by northern blotting. Only a limited number of samplings were made (40 minutes, 8 hours and 24 hours post infection). Nonetheless, the data enabled the authors to speculate on a number of interesting and previously unrecognized features of CMV infection, such as the upregulation of the gene encoding the nonclassical class I molecule, HLA-E, as a means to avoid natural killer (NK) cell surveillance and the increased expression of cytosolic phospholipase A2 and cyclooxygenase-2 (COX2) as a means to increase synthesis of the proinflammatory mediator prostaglandin E2. The latter effect may benefit both host and virus, since inflammation would both stimulate a cell-mediated immune response and at the same time attract a reservoir of monocytic cells that might serve to disseminate the infection. A number of herpesvirus family members (mouse and human CMVs, as well as the Kaposi's-sarcoma-associated herpes virus, HHV-8) have recently been shown to encode chemokine-like functions that also appear to promote virus dissemination [17,18].

A poster describing the early events that precede the wanton destruction caused by Ebola virus in human monocytes was presented by K Anderson *et al.* (abstract VW8.01, 11th International Congress of Virology, Sydney, August 9–13, 1999). This virus induces strong and sustained release of a number of cytokines and chemokines as well as the cellular inhibitors of apoptosis, cIAP1 and XIAP, and the apoptosis inducer, TRAIL. Recent evidence suggests that survival of Ebola infection by both guinea pigs and by humans is dependent upon the development of a sustained T cell response that precedes effective humoral responses. The virus itself appears to interfere with the ability to maintain T cell responses, inducing massive apoptosis in peripheral blood leukocytes in nonsurvivors [19,20]. In a related study reported by Xiang *et al.* [21], the transcriptional responses induced upon infection of primary human monocytes with the Zaire and Reston strains of Ebola were analyzed; striking differences between responses to the two strains that might underlie the differences in their pathogenicity in humans were reported. Wallace *et al.* [22] reported on experiments designed to investigate the effects of *Helicobacter pylori* infection on gastric epithelial cells and cataloged changes in the transcriptional activity of 260 genes of known and unknown function.

Making sense of the data

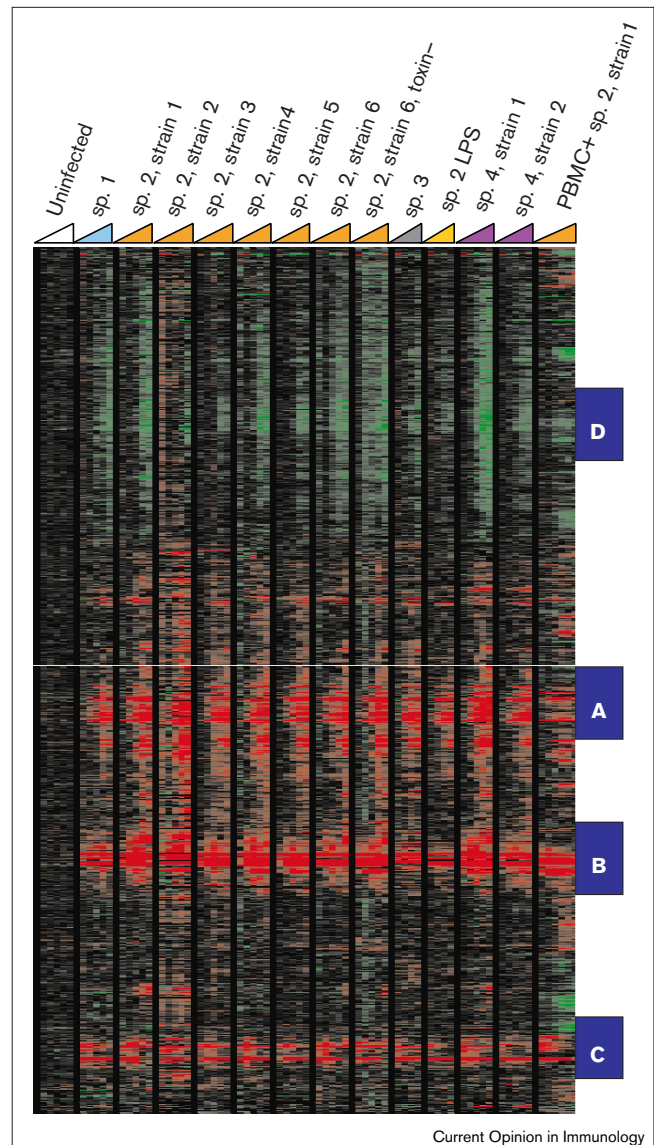
These 'proof-of-concept' experiments demonstrate many of the fundamental strengths of the microarray approach and the some of the extraordinary conceptual challenges posed by data gathering at this scale. Sherlock (this issue, pp 201–205) has provided a description of some of the tools already available for storing and analyzing microarray data. Many of the presentations at The Microarray Meeting (Scottsdale, Arizona, September 22–25, 1999) suggest that a formidable arsenal of mathematical and statistical tools are

under development. Additional issues already tabled [23] are the establishment of a public database for warehousing and integrating array data and the adoption of a minimal standard for submission that will allow data from many laboratories to be integrated and incorporated into larger analyses. Bassett *et al.* [23] emphasize one of the fundamental assumptions implicit in array-based expression analysis: “knowledge of where and when a gene is expressed carries important information about what the gene does”. It follows from this assumption that a first step in the analysis of array-based data would be to group together genes whose expression patterns are similar. Cluster analysis methods [24*,25*] provide the ability to reveal structure within a mass of array data, grouping both known and unknown genes together in ways that may suggest coordinate control mechanisms and/or function in common pathways. These sorts of methods have led to the classification of cell and tissue types, and the inference of functional categories [26,27*]. Golub *et al.* [28*] and Alizadeh *et al.* [29*] have recently demonstrated the practical utility of microarray expression analysis in cancer diagnostics. The first group identified gene ‘classes’ in acute lymphoblastic and myeloblastic leukemias that serve as predictors when screening unknown samples. Further refinement of these classes enabled discrimination of B cell and T cell leukemias as well as prediction of new classes. The second group were able to define and distinguish two types of diffuse large-B-cell lymphoma with different gene expression features, as well as different clinical outcomes. It is easy to imagine the utility of gene classes in predicting disease mechanism or responsiveness to therapy. Many of these same approaches and principles can be applied to infectious disease and to the identification of diagnostic signatures, predictors of clinical course of infection and targets for novel intervention.

The potential power of microarray analysis to reveal mechanisms of microbial pathogenesis through host expression profiling has been barely touched. A particularly effective approach would compare the responses induced by sets of microorganisms that differ in a well-defined manner. For instance, comparisons of a wild-type strain with isogenic derivatives, each with well-characterized defects in genes encoding virulence factors, are likely to generate data that reveal possible mechanisms of action and are more useful than comparisons between unrelated strains. Purified virulence factors, microbial cell wall components and genetically inactivated variants of these are also useful reagents and would serve to help answer questions regarding sufficiency, rather than essentiality. The question raised at the beginning of this article concerning possible host discrimination between pathogens and the identification of diagnostic signatures among host response profiles will require the examination of a wide variety of naturally occurring strains and species, and application of two-way clustering across all genes as well as microbial stimuli. We have begun this kind of investigation, in collaboration with Pat Brown (Stanford University, Stanford, CA) and Lou Staudt (National Cancer Institute, Bethesda, MD), by collecting and analyzing response profiles obtained from human cells exposed to

various infectious agents *in vitro* [30–32]. Figure 1 presents the type of data obtained from these experiments. These particular data illustrate the hierarchical clustering of responsive genes in U937 monocytic cells after exposure to several members of the same Gram-negative bacterial genus, multiple strains of the same species and some

Figure 1



Cluster diagram of U937 monocytic cell expression responses to a variety of bacterial species and strains. *Bordetella avium*, *B. parapertussis*, two strains of *B. bronchiseptica*, multiple strains of *B. pertussis* – including an isogenic pertussis toxin mutant – and *B. pertussis* LPS were studied. An 18,432 element human cDNA microarray was employed. Each column represents a single experiment and each row a single gene. U937 responses to each stimulus were studied at six different time points (represented as six columns grouped under each colored triangle in temporal order). Large groups of genes (A–D) respond in a similar fashion (induced, red; repressed, green) to many or all stimuli. Other genes respond in a more discriminating manner. The left-hand column shows uninfected cells; the right-hand column shows the response of peripheral blood mononuclear cells (PBMC) to *B. pertussis*.

purified exo- and endo-toxins (LPS). A much larger effort will be necessary before any conclusions can be reached about the presence of diagnostic signatures *in vivo* (e.g. in peripheral blood mononuclear cells), the relative importance of host-to-host variability and the possible basis for a novel means of classifying infectious agents of disease.

Conclusions

High density DNA microarray analysis of host gene expression provides a powerful method of examining microbial pathogens from a novel perspective. The ability to survey the responses of a large subset of the host genome, and to find patterns among the profiles from many different microorganisms and hosts, allows fundamental questions to be addressed (or re-addressed) about the basis of pathogen recognition, the features of the interaction between host and pathogen and the mechanisms of host defense and microbial virulence.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Lueking A, Horn M, Eickhoff H, Bussow K, Lehrach H, Walter G: **Protein microarrays for gene expression and antibody screening.** *Anal Biochem* 1999, **270**:103-111.
2. Saunders NJ, Moxon ER: **Implications of sequencing bacterial genomes for pathogenesis and vaccine development.** *Curr Opin Biotechnol* 1998, **9**:618-623.
3. de Saizieu A, Certa U, Warrington J, Gray C, Keck W, Mous J: **Bacterial transcript imaging by hybridization of total RNA to oligonucleotide arrays.** *Nat Biotechnol* 1998, **16**:45-48.
4. Wilson M, DeRisi J, Kristensen HH, Imboden P, Rane S, Brown PO, Schoolnik GK: **Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization.** *Proc Natl Acad Sci USA* 1999, **96**:12833-12838.
5. Gingeras TR, Ghandour G, Wang E, Bero A, Small PM, Drobniewski F, Alland D, Desmond E, Holodniy M, Drenkow J: **Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays.** *Genome Res* 1998, **8**:435-448.
6. Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM: **Comparative genomics of BCG vaccines by whole-genome DNA microarray.** *Science* 1999, **284**:1520-1523.
7. Bowtell DD: **Options available – from start to finish – for obtaining expression data by microarray.** *Nat Genet* 1999, **21**:25-32. [Published erratum appears in *Nat Genet* 1999, **21**:241.]
8. Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapati R, Childs G: **Making and reading microarrays.** *Nat Genet* 1999, **21**:15-19.
9. Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM: **Expression profiling using cDNA microarrays.** *Nat Genet* 1999, **21**:10-14.
10. Eisen MB, Brown PO: **DNA arrays for analysis of gene expression.** • *Methods Enzymol* 1999, **303**:179-205. A useful set of procedures for widespread use.
11. Gerhold D, Rushmore T, Caskey CT: **DNA chips: promising toys have become powerful tools.** *Trends Biochem Sci* 1999, **24**:168-173.
12. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ: **High density synthetic oligonucleotide arrays.** *Nat Genet* 1999, **21**:20-24.
13. Der SD, Zhou A, Williams BR, Silverman RH: **Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays.** *Proc Natl Acad Sci USA* 1998, **95**:15623-15628.
14. Alizadeh A, Eisen M, Botstein D, Brown PO, Staudt LM: **Probing lymphocyte biology by genomic-scale gene expression analysis.** *J Clin Immunol* 1998, **18**:373-379.
15. Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW: **Parallel human genome analysis: microarray-based expression monitoring of 1000 genes.** *Proc Natl Acad Sci USA* 1996, **93**:10614-10619.
16. Zhu H, Cong JP, Mamtara G, Gingeras T, Shenk T: **Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays.** *Proc Natl Acad Sci USA* 1998, **95**:14470-14475. This analysis reveals several insights into CMV pathogenesis.
17. Saederup N, Lin YC, Dairaghi DJ, Schall TJ, Mocarski ES: **Cytomegalovirus-encoded beta chemokine promotes monocyte-associated viremia in the host.** *Proc Natl Acad Sci USA* 1999, **96**:10881-10886.
18. Penfold ME, Dairaghi DJ, Duke GM, Saederup N, Mocarski ES, Kemble GW, Schall TJ: **Cytomegalovirus encodes a potent alpha chemokine.** *Proc Natl Acad Sci USA* 1999, **96**:9839-9844.
19. Baize S, Leroy EM, Georges-Courbot MC, Capron M, Lansoud-Soukate J, Debre P, Fisher-Hoch SP, McCormick JB, Georges AJ: **Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients.** *Nat Med* 1999, **5**:423-426.
20. Nabel GJ: **Surviving Ebola virus infection.** *Nat Med* 1999, **5**:373-374.
21. Xiang C, Young H, Alterson H, Reynolds D, Bittner M, Chen Y, Gooden G, Jiang Y, Meltzer P, Trent J *et al.*: **Comparison of cellular gene expression in Ebola-Zaire and Ebola-Reston virus-infected primary human monocytes.** *Nat Genet* 1999, **23**:82.
22. Wallace DM, Clayton CL, Cox JM, Falciani F, Trower M, Robinson PA, Crabtree JE: **Identification of *Helicobacter pylori* epithelial cell response genes by screening high-density cDNA arrays.** *Nat Genet* 1999, **23**:80.
23. Bassett DE Jr, Eisen MB, Boguski MS: **Gene expression informatics – it's all in your mine.** *Nat Genet* 1999, **21**:51-55.
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**:14863-14868. Hierarchical clustering provides a means for organizing genome-wide data in a useful visual format and tends to group genes according to similar function.
25. Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander ES, Golub TR: **Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation.** *Proc Natl Acad Sci USA* 1999, **96**:2907-2912. An alternative method for gene clustering (see [22]) that might be combined with a hierarchical method for optimal association of genes with like function.
26. Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, Levine AJ: **Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays.** *Proc Natl Acad Sci USA* 1999, **96**:6745-6750.
27. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC *et al.*: **Distinctive gene expression patterns in human mammary epithelial cells and breast cancers.** *Proc Natl Acad Sci USA* 1999, **96**:9212-9217. Breast cancer can be distinguished from other disturbances in mammary epithelial cells using expression microarray analysis and the relative contribution of stromal and inflammatory cells to overall breast cancer expression profiles can be deduced.
28. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA *et al.*: **Molecular classification of cancer: class discovery and class prediction by gene expression monitoring.** *Science* 1999, **286**:531-537. Acute leukemia subtypes are accurately classified and novel subtypes predicted using the expression response profiles from classes of host genes.
29. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X *et al.*: **Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling.** *Nature* 2000, **403**:503-511. Expression profile analysis reveals two distinct, previously unrecognized types of diffuse large-B-cell lymphoma with different clinical courses.
30. Brown PO, Hartwell L: **Genomics and human disease – variations on variation.** *Nat Genet* 1998, **18**:91-93.
31. Relman DA: **Detection and identification of previously unrecognized microbial pathogens.** *Emerg Infect Dis* 1998, **4**:382-389.
32. Relman DA: **The search for unrecognized pathogens.** *Science* 1999, **284**:1308-1310.