

Genome-wide responses of a pathogenic bacterium to its host

David A. Relman

Departments of Microbiology & Immunology and Medicine, Stanford University School of Medicine, Stanford, California, USA and Veterans Affairs Palo Alto Health Care System, Palo Alto, California, USA

J. Clin. Invest. 110:1071–1073 (2002). doi:10.1172/JCI200216944.

Bacteria are remarkably sensitive and adaptable creatures. They are highly attuned to their environment, perceive a wide variety of cues, react using a complex set of mechanisms, and modify their composition and behavior in a sophisticated manner. Pathogenic bacteria are no exception. The benefits of a pathogenic lifestyle, e.g., replication within a host cell or persistence within the intravascular compartment as well as the costs of expressing virulence factors at inappropriate times or sites, has led to the development of redundant regulatory systems that facilitate the timely and coordinated control of these adaptation and virulence mechanisms (1). Some of the cues that are important to a human pathogen include those that indicate contact with a human host, arrival at a preferred niche, and transit through multiple potentially adverse microenvironments. In the laboratory, pathogens respond to signals such as temperature, low iron concentrations, oxygen and carbon dioxide partial pressures, and nutrient limitation. Because of the complexity and poor accessibility of host microenvironments, however, it has been difficult to determine which cues regulate bacterial responses *in vivo*.

Bacterial response mechanisms operate at multiple levels: (a) DNA structure and topology; (b) RNA

abundance and processing; and (c) protein expression, modification, and localization. Of these, a wealth of information bears on gene transcriptional regulation. Data from completed bacterial genome sequences suggest a direct relationship between the breadth of environmental conditions encountered by a bacterium and the number of genes in its genome that are predicted to encode transcriptional regulatory factors (2). *Escherichia coli* is particularly well endowed, presumably reflecting environmental transitions in and out of the intestinal tract (3). On closer examination of this species, we now recognize an immense degree of genomic heterogeneity among different serovars and pathovars.

Genetic and genomic analyses of pathogenic bacteria

Analyses of responses by bacterial pathogens to the host, or to conditions that mimic host environments relevant to the disease process, reveal fundamental aspects of virulence. Bacteria can be viewed as “microprobes” that report on the nature of a host microenvironment and provide unique insight into host physiology, but they are far from passive observers. Because pathogens survive on the basis of counter-defensive strategies, the information they provide on host environments also reflects the changes they bring to bear on the niche they occupy. A variety of powerful genetic approaches have been described for characterizing bacterial responses during host encounters. These approaches include *in vivo* expression technology (4), signature-tagged mutagenesis (5), and differential fluorescence induction (6). Despite the proven use-

fulness of these methods, selective pressures or colonization bottlenecks can create bias. The highly parallel, comprehensive assessment of microbial and host cell transcript abundance made possible by complete genome sequences and DNA microarrays provides an alternative strategy — one that has already dramatically deepened our appreciation of the complexities of host-microbe interactions (7–15).

In this issue of the *JCI*, Staudinger and colleagues describe genome-wide changes in *E. coli* transcript abundance early after bacterial ingestion by neutrophils (16). Their use of phagocyte oxidase-deficient neutrophils allows a more precise examination of bacterial responses that are attributable to the activity of this important anti-microbial host defense system and to the environment of the neutrophil phagosome. Their results, which are carefully controlled to exclude nonspecific effects of neutrophil contact and culture conditions, suggest that oxidative stress is a dominant signal perceived by *E. coli* in this environment. Many, but not all, of the phagocyte oxidase-dependent effects involve genes whose expression has previously been associated with the bacterial transcriptional regulator OxyR. Hydrogen peroxide also brings about OxyR-independent transcriptional events, as pointed out by these authors and by others previously (17). But as expected, an isogenic *oxyRS* mutant was much more susceptible to killing by hydrogen peroxide and by wild-type neutrophils than its *E. coli* parent.

Long-standing assumptions about the direct role of oxidative mechanisms in microbial killing within neutrophil phagocytic vacuoles have

Address correspondence to: David A. Relman, Veterans Affairs Palo Alto Health Care System 154T, Building 101, Room B4-185, 3801 Miranda Avenue, Palo Alto, California 94304, USA.
Phone: (650) 852-3308; Fax: (650) 852-3291; E-mail: relman@stanford.edu.

Conflict of interest declared: No conflict of interest has been declared.

recently been called into question. An alternative scenario holds that reactive oxygen species act indirectly by prompting an influx of potassium ions into the vacuole as well as pH changes, hypertonicity, and the release and activation of microbicidal granule proteases such as cathepsin G and elastase (18). In the setting of this recent discussion, the findings by Staudinger et al. seem to support a more direct role for reactive oxidative species and hydrogen peroxide. OxyR regulates expression of antioxidative defense factors in *E. coli* in response to hydrogen peroxide, as well as cellular thiol-disulfide redox status, and is activated by direct oxidation (19). If OxyR regulon activation is a reliable, direct, and specific indicator of oxidative conditions sufficient to cause significant bacterial stress, then these findings support a direct role for oxidative killing mechanisms. However, OxyR does respond to other signals, as illustrated by the enhanced susceptibility of the *oxyRS* mutant to phagocyte oxidase-deficient neutrophils (see ref. 16; Figure 6, right panel) (16). This observation reminds us that the mechanism of bacterial killing by phagocytic cells may be multifactorial and emphasizes the difficulties of interpreting simultaneous and sometimes competing or cacophonous signals in any complex real-world environment.

One clear advantage of microarray-based experiments in general is that they allow for the straightforward identification of genes whose transcripts decrease in abundance under a given condition. Genetic approaches that rely on reporter constructs with positive readouts and other traditional approaches tend to be biased against this half of regulatory systems. Staudinger et al. unexpectedly found decreased levels of transcripts for genes regulated by sigma factor σ^{54} and nitrogen starvation. As they point out, ATP depletion, resulting from disruption of proton-motive force in oxidatively stressed bacteria, may indirectly explain these observations. Disruption of proton-motive force and ATP generation is also consistent with the model proposed by Reeves et al. (18).

Caveats and challenges

Several limitations and confounding issues inherent to this study are illustrative of more general problems with

microarray-based analyses of genome-wide transcript patterns as they are currently conducted. DNA microarrays can only monitor those transcripts for which homologous probes are available. The recent release of complete genome sequences from pathogenic variants of earlier-sequenced laboratory type strains and the results of subtractive cloning are beginning to illustrate the breadth and abundance of genes that are unique to disease-causing bacteria. Thus, many of the currently-available bacterial gene microarrays ignore genes and transcripts that are germane to pathogenesis.

Staudinger et al. were aware of this problem with their study. They used a microarray designed from the genome sequence of the avirulent *E. coli* K-12 strain MG1655. In addition, they used the heavily-passaged *E. coli* type strain ATCC 11775 for their experiments. Perna et al. have shown that the genome of enterohemorrhagic *E. coli* O157:H7 contains 1.34 megabases of DNA, accounting for 26% of its genes, that are not present in *E. coli* K-12 (20). Hundreds of lineage-specific gene clusters are scattered through both genomes as groups of “islands” or as an “archipelago.” Many of these genome differences are related to virulence, and most appear to be the result of relatively recent horizontal transfer events. Uropathogenic *E. coli* also contain at least several hundred kilobases of unique DNA, some of which encodes hemolysins, iron acquisition and transport machinery, and other virulence factors (21, 22). It is possible that the virulence capabilities of a pathogenic *E. coli* strain would generate a different picture of the neutrophil phagosomal environment.

At face value, DNA microarrays provide a global view of changes in the relative abundance of gene transcripts. One should be cautious, however, in inferring mechanisms for complex phenotypes. Transcript abundance alone reflects more than just regulation of transcriptional initiation. A recently published study illustrates the wide variation in the half-life of different *E. coli* mRNAs (23). In this study, there was no correlation between half-life and a variety of transcript features, but there was a possible relationship between half-life

and the category of predicted gene function, raising the intriguing possibility that the neutrophil phagosomal environment causes nonuniform and selective degradation of bacterial transcripts. This possibility would have profound implications for interpreting the data presented by Staudinger et al. and other investigators of the host-pathogen interaction. It would be helpful to be able to turn to data that describe transcript half-life as a function of the environmental conditions that are relevant to niches occupied by pathogens. As an aside and as a separate issue, the magnitude of change in transcript abundance is not necessarily the most reliable indicator of relative biological importance. Relevant features of the response program may be revealed by the application of other analytical tools to microarray data, including the use of clustering methods for identifying coregulated sets of genes (24–26). And of course, RNA abundance levels do not necessarily correspond to protein levels or protein biological activity.

Future efforts to understand host-microbe interactions using genome-scale methodologies will benefit from progress in several areas. More complete gene inventories generated from relevant clinical and environmental isolates will enhance DNA microarray breadth of coverage. Standardized methods and reference pools of RNA and DNA will greatly facilitate the comparison of data generated by different research groups. A robust database with microbial response profiles from a wide variety of well-characterized environmental conditions will allow more insightful interpretation of transcription patterns recorded in novel host environments. Finally, the integration of host response data with bacterial response data will provide a more complete understanding of the two-way conversation between host and pathogen. Our capabilities as “eavesdroppers” are expanding on many fronts.

1. Finlay, B.B., and Falkow, S. 1997. Common themes in microbial pathogenicity revisited. *Microbiol. Mol. Biol. Rev.* **61**:136–169.
2. Stover, C.K., et al. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*. **406**:959–964.
3. Blattner, F.R., et al. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science*. **277**:1453–1474.

4. Slauch, J.M., and Camilli, A. 2000. IVET and RIVET: use of gene fusions to identify bacterial virulence factors specifically induced in host tissues. *Methods Enzymol.* **326**:73–96.
5. Chiang, S.L., Mekalanos, J.J., and Holden, D.W. 1999. In vivo genetic analysis of bacterial virulence. *Annu. Rev. Microbiol.* **53**:129–154.
6. Valdivia, R.H., and Falkow, S. 1997. Fluorescence-based isolation of bacterial genes expressed within host cells. *Science.* **277**:2007–2011.
7. Belcher, C.E., et al. 2000. The transcriptional responses of respiratory epithelial cells to *Bordetella pertussis* reveal host defensive and pathogen counter-defensive strategies. *Proc. Natl. Acad. Sci. USA.* **97**:13847–13852.
8. Cummings, C.A., and Relman, D.A. 2000. Using DNA microarrays to study host-microbe interactions. *Emerg. Infect. Dis.* **6**:513–525.
9. Derweiler, C.S., Cunanán, D.B., and Falkow, S. 2001. Host microarray analysis reveals a role for the Salmonella response regulator phoP in human macrophage cell death. *Proc. Natl. Acad. Sci. USA.* **98**:5850–5855.
10. Paustian, M.L., May, B.J., and Kapur, V. 2001. *Pasteurella multocida* gene expression in response to iron limitation. *Infect. Immun.* **69**:4109–4115.
11. Smoot, L.M., et al. 2001. Global differential gene expression in response to growth temperature alteration in group A Streptococcus. *Proc. Natl. Acad. Sci. USA.* **98**:10416–10421.
12. Whiteley, M., et al. 2001. Gene expression in *Pseudomonas aeruginosa* biofilms. *Nature.* **413**:860–864.
13. Sauvonnnet, N., Pradet-Balade, B., Garcia-Sanz, J.A., and Cornelis, G.R. 2002. Regulation of mRNA expression in macrophages after *Yersinia enterocolitica* infection. Role of different Yop effectors. *J. Biol. Chem.* **277**:25133–25142.
14. Boldrick, J.C., et al. 2002. Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc. Natl. Acad. Sci. USA.* **99**:972–977.
15. Schoolnik, G.K. 2002. Functional and comparative genomics of pathogenic bacteria. *Curr. Opin. Microbiol.* **5**:20–26.
16. Staudinger, B.J., Oberdoerster, M.A., Lewis, P., and Rosen, H. 2002. mRNA expression profiles for *Escherichia coli* ingested by normal and phagocyte oxidase-deficient human neutrophils. *J. Clin. Invest.* **110**:1151–1163. doi:10.1172/JCI200215268.
17. Zheng, M., et al. 2001. DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J. Bacteriol.* **183**:4562–4570.
18. Reeves, E.P., et al. 2002. Killing activity of neutrophils is mediated through activation of proteases by K⁺ flux. *Nature.* **416**:291–297.
19. Aslund, F., Zheng, M., Beckwith, J., and Storz, G. 1999. Regulation of the OxyR transcription factor by hydrogen peroxide and the cellular thiol-disulfide status. *Proc. Natl. Acad. Sci. USA.* **96**:6161–6165.
20. Perna, N.T., et al. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature.* **409**:529–533.
21. Janke, B., Dobrindt, U., Hacker, J., and Blum-Oehler, G. 2001. A subtractive hybridisation analysis of genomic differences between the uropathogenic *E. coli* strain 536 and the *E. coli* K-12 strain MG1655. *FEMS Microbiol. Lett.* **199**:61–66.
22. Melkerson-Watson, L.J., Rode, C.K., Zhang, L., Foxman, B., and Bloch, C.A. 2000. Integrated genomic map from uropathogenic *Escherichia coli* J96. *Infect. Immun.* **68**:5933–5942.
23. Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S., and Cohen, S.N. 2002. Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci. USA.* **99**:9697–9702.
24. Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA.* **95**:14863–14868.
25. Eisen, M.B., and Brown, P.O. 1999. DNA arrays for analysis of gene expression. *Methods Enzymol.* **303**:179–205.
26. Sherlock, G. 2001. Analysis of large-scale gene expression data. *Brief. Bioinform.* **2**:350–362.