Minireview

The complete genome sequence of *Vibrio cholerae*: a tale of two chromosomes and of two lifestyles

Gary K Schoolnik and Fitnat H Yildiz

Address: Departments of Medicine (Infectious Diseases and Geographic Medicine) and Microbiology and Immunology, Stanford University Medical School, Stanford, CA 94305, USA.

Correspondence: Gary K Schoolnik. E-mail: schoolni@cmgm.stanford.edu

Published: 1 September 2000

Genome ***Biology*** 2000, **1(3)**:reviews1016.1–1016.3

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2000/1/3/reviews/1016

Abstract

Vibrio cholerae O1 has figured prominently in the history of infectious diseases as a cause of periodic global epidemics, an affliction of refugees in areas of social strife and as the disease first subjected to modern epidemiological analysis during the classic investigations of John Snow in mid-19th century London [1]. Thus, publication of the entire genome sequence of *V. cholerae* O1 (biotype El Tor) in *Nature* [2] by a consortium of investigators from The Institute for Genomic Research, the University of Maryland and Harvard Medical School is properly regarded as an historic event that will trigger a paradigm shift in the study of this organism.

Vibrio cholerae O1 causes Asiatic cholera, a purging diarrheal illness that can kill an adult within 24 hours. Its most intriguing and least understood feature, however, comes from the study of its annual epidemic profile in the Bengal region of Bangladesh and India. There, nearly all cases each year occur in a synchronized, massive outbreak in the months of October and November, just as the monsoon rains decline [3]. During most other months of the year, cholera cases occur sporadically or not at all. This epidemic profile and its correlation with major transitions of climate point to the following: *V. cholerae* O1 resides in a stable environmental reservoir; then, seasonally determined changes in rainfall and sunlight trigger its periodic and transient emergence as a human pathogen [4].

Between epidemics, *V. cholerae* O1 lives in natural aquatic habitats formed by the confluence of the Ganges and the Brahmaputra rivers. All the physiochemical and ecological features of this system are dramatically influenced by the monsoon climate. Within this ecosystem, five distinctive stages have been proposed to comprise the *V. cholerae* environmental life cycle: an independent, free-swimming form; a symbiont of phytoplankton [5]; a commensal of zooplankton [6]; a viable, but not culturable state [7]; and a biofilm community attached to abiotic or chitinous surfaces (Figure 1) [8-12]. Thus the functional repertoire of the *V. cholerae* O1 genome must be unusually broad as it accommodates two, quite distinctive, lifestyles: the milieu of the human intestine and long term residence in aquatic habitats that are subject to climate-determined changes of the microenvironment [13].

The functional annotation of the *V. cholerae* O1 sequence [2] sheds light on this remarkable capacity. In particular, the distribution of genes of known function between the large and small chromosomes of the organism provides tantalizing clues about how the two-chromosome configuration of the *Vibrionaceae* might confer an evolutionary advantage in habitats that vary with climate change. Although the large chromosome (chrI) contains most of the genes that are required for growth and pathogenicity, some of the components of several essential metabolic and regulatory pathways reside on the small chromosome (chrII). Thus, retention of chrII is presumably required for survival of the organism, at least under some conditions of growth. By contrast, a larger

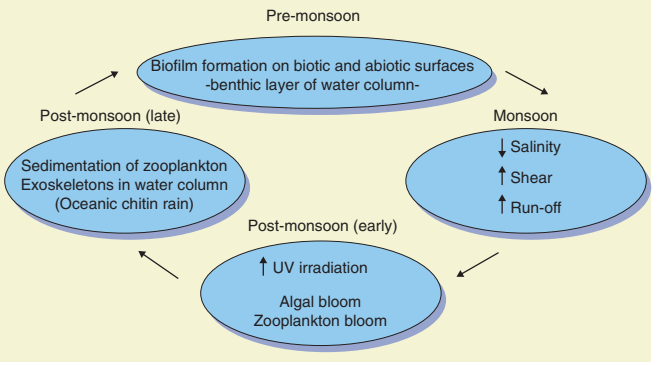


Figure 1
The environmental life cycle of *Vibrio cholerae* O1.

percentage of hypothetical genes and genes of unknown function reside on chrII. This asymmetrical distribution of genes suggests that under certain conditions differences in the copy number of chrI and chrII might occur, potentially increasing the effective level of expression of genes on the more numerous chromosome to the organism's advantage [2]. A related hypothesis, which can now be tested, predicts that chrI genes mainly adapt the organism for growth in the intestine whereas chrII genes are essential within environmental niches. The extreme in chromosome copy number asymmetry is the loss of one chromosome and Heidelberg *et al*. suggest that non-replicating, single chromosomal cells, within a nutrient-stressed population of normal two-chromosomal cells, might contribute to the survival of the community by continuing to secrete enzymes that break down molecules in the surrounding microenvironment. They propose that the resulting nutrients might be used by normal members of a community at risk, thereby promoting survival of the species.

The organism's capacity to access nutrients within diverse environmental habitats is also enhanced by the extensive duplication of genes involved in nutrient scavenging. A striking example of this is the apparent duplication of genes coding for chitinase, which together with the corresponding phosphoenolpyruvate phosphotransferase system, produce and transport chitin-derived disaccharides released from the exoskeletons of zooplankton to which *V. cholerae* attach [2,6].

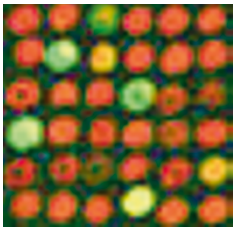
While the identification of genes of known function is an important achievement, it must be remembered that only 54% of the identified open reading frames (ORFs) are predicted to encode proteins of known function [2] and, even for these, differences in sequence could confer unexpected functions. Thus, although completion of the *V. cholerae* sequence is an important first step, now it is important to use this information to launch functional genomic studies by mutational approaches, comparative studies of related organisms and expression profiling using DNA microarrays (see Box 1).

Bioinformatic analysis of microarray-derived expression profiles will reveal sets of co-regulated genes that mediate the adaptive response of the organism to different conditions of growth [14,15]. Further analysis of co-regulated gene clusters will disclose components of metabolic and biosynthetic pathways [16]. Initially employed for eukaryotic expression analysis, this method has recently been used to study bacterial transcriptional responses as well [17] and methods to stabilize

Box 1

A microarray is a surface that contains representations of each open reading frame (ORF) of a sequenced and annotated genome. Of the several available formats, the most commonly used in academic labs (developed by Patrick Brown and colleagues at Stanford University) consists of a microscope slide whose surface displays a matrix of printed spots, each spot containing a PCR-derived amplicon that corresponds to all or part of an ORF of the sequenced genome. Thus, each ORF of the genome is represented on the array as a separate spot, its location designated by its matrix coordinates.

One principal innovation in gene-expression profiling involved the introduction of two-color hybridization. This method employs two populations of cDNAs that have been differentially labeled with two different fluorochromes – the cDNAs usually having been derived from RNA prepared from the same organism cultivated under or exposed to, two contrasting conditions. Equal masses of the two differentially labeled populations of cDNAs are combined, applied to the array surface and allowed to hybridize to the corresponding ORF-specific targets. The array is then scanned and the intensity of each label for each ORF-specific spot is quantitated. These values are compared, yielding ratios that serve as a measure of the relative expression levels of each ORF for the two tested conditions.



Other microarray systems and methods, such as those developed by Affymetrix, really differ only in detail (for example, the oligonucleotides representing each ORF are synthesized *in situ* on the solid support, rather than being spotted onto glass slides or nylon membranes), but the underlying principles of experimental design remain the same.

bacterial RNA have been developed [18] and modified for this application. To date, most such studies have explored physiological responses, changes during development, and the mode of action of antibiotics. Now, with the identification of the 3,885 ORFs of the *V. cholerae* O1 genome sequence [2], fabrication and use of the corresponding DNA microarray will surely follow.

What ecological and pathogenic questions might whole-genome expression profiling of *V. cholerae* O1 address? Investigators will want to learn more about how this species adapts to the many niches it so successfully occupies. To this end, microarray experiments will be conducted to identify genes that are differentially expressed in biofilms formed on plastic, glass and chitin membranes; during co-cultivation with algae and zooplankton; after exposure to different concentrations of saline; and within the human gastrointestinal tract. Because the annotated genome sequence highlights a number of transcription factors [2], expression profiling will also be used to identify which genes these factors control. For example, the regulatory network governed by *rpoS*, which encodes an alternative σ factor postulated to be required for the environmental survival of *V. cholerae* O1 [19] - could be dissected by identifying which genes are not expressed by a *rpoS* mutant when compared to the wild-type parent.

Finally, microarray DNA-DNA hybridization experiments will be performed with labeled DNA to compare the genomes of closely related, but non-identical, strains and species [20]. The most compelling study of this kind will seek to identify genes present in the sequenced El Tor biotype of *V. cholerae* O1, but missing from the yet-to-be-sequenced Classical biotype. The Classical biotype caused the fifth (1881-1896) and sixth (1899-1923) cholera pandemics, but was evicted from most endemic areas by the El Tor biotype during the on-going seventh pandemic [21]. This has led to the idea that the El Tor biotype is somehow more environmentally fit than its close, Classical cousin [22]. Accordingly, within the set of El Tor-specific genes may be those that confer a selective advantage in the estuaries and rivers of South Asia.

V. cholerae O1 has become the model organism to study the effect of climate on the evolution, emergence and epidemiology of an infectious agent [4]. The publication of the entire genome sequence of this species, particularly when combined with array-based functional genomics and on-site and remote imaging of the Bengal river delta [13], will probably provide the first opportunity to study how climate change affects an organism's transcriptional activity.

References

1. Snow J: *On the mode of communication of cholera*. London: John Churchill, 1849.
2. Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, *et al.*: **DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.** *Nature* 2000, **406:**477-483
3. Glass RI, Claeson M, Blake PA, Waldman RJ, Pierce NF: **Cholera in Africa: lessons on transmission and control for Latin America.** *Lancet* 1991, **338:**791-795.
4. Colwell RR: **Global climate change and infectious diseases: the cholera paradigm.** *Science* 1996 **274:**2025-2031.
5. Islam MS, Drasar BS, Sack RB: **The aquatic flora and fauna as reservoirs of *Vibrio cholerae*: a review.** *J Diarrhoeal Dis Res* 1994, **12:**87-96
6. Huq A, Small EB, West PA, Huq MI, Rahman R, Colwell RR: **Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepodes.** *Appl Environ Microbiol* 1983, **45:**275-83
7. Xu H-S, Roberts N, Singleton FL, Attwell RW, Grimes DJ, Colwell RR: **Survival and viability of non-culturable *Escherichia coli* and *Vibrio cholerae* in the estuarine and marine environment.** *Microb Ecol,* 1982, **8:**313-323.
8. Wai SN, Mizunoe A, Takade A, Kawabata SI, Yoshida SI: ***Vibrio cholerae* O1 strain TSI-4 produces the exopolysaccharide materials that determine colony morphology, stress resistance, and biofilm formation.** *Appl Environ Microbiol* 1998, **64:**3648-3655.
9. Yildiz FH, Schoolnik GK: ***Vibrio cholerae* O1 El Tor: identification of a gene cluster required for the rugose colony type, exopolysaccharide production, chlorine resistance and biofilm formation.** *Proc Natl Acad Sci USA* 1999, **96:**4028-4033.
10. Watnick PI, Fullner KJ, Kolter R: **A role for the mannose-sensitive hemagglutinin in biofilm formation by *Vibrio cholerae* El Tor.** *J Bacteriol* 1999, **181:**3606-3609.
11. Watnick PI, Kolter R: **Steps in the development of a *Vibrio cholerae* biofilm.** *Mol Microbiol* 1999, **34:**586-595.
12. Watnick PI, Kolter R: **Biofilm, city of microbes.** *J Bacteriol* 2000, **182:**2675-2679.
13. Lobitz B, Beck L, Huq A, Wood B, Fuchs G, Faruque ASG, Cowell R: **Climate and infectious disease: use of remote sensing for detection of *Vibrio cholerae* by indirect measurement.** *Proc Natl Acad Sci USA* 2000, **97:**1438-1443.
14. 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.
15. 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.
16. Karp PD, Krummenacker M, Paley S, Wagg J: **Integrated pathway-genome databases and their role in drug discovery.** *Trends Biotech* 1999, **17:**275-281.
17. Wilson M, DeRisi J, Kristensen H-H, 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.
18. Mangan JA, Butcher PD: **Analysis of mycobacterial differential gene expression by RAP-PCR.** *Methods Mol Biol* 1998, **101**:307-22.
19. Yildiz FH, Schoolnik GK: **Role of rpoS in stress survival and virulence of *Vibrio cholera*.** *J Bacteriol* 1998, **180:**773-784.
20. Behr MA, Wilson MA, Gill WP, Salamon, H, Schoolnik GK, Rane S, Small PA: **Comparative genomics of BCG vaccines by whole-genome DNA microarray.** *Science* 1999, **284:**1520-1523
21. Huq MI, Glass RI, Stoll BJ: **Epidemiology of cholera.** *N Eng J Med* 1980, **303:**643-644.
22. Neogy KN: **Viability of *V. cholerae* and *V. eltor* in food and water.** *Bull Calcutta School Trop Med* 1965, **13:**10-11.