JOURNAL OF BACTERIOLOGY, Nov. 2004, p. 7029–7031 Vol. 186, No. 21
0021-9193/04/$08.00+0 DOI: 10.1128/JB.186.21.7029–7031.2004

Phage Taxonomy: We Agree To Disagree

Daniel Nelson*

Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, New York, New York

Bacteriophage (phage) researchers are in universal agreement that phage biology has undergone a renaissance in recent years. No longer just tools of molecular biology, phage are now recognized to play critical roles in bacterial pathogenesis (3, 29) and bacterial population dynamics (5, 10). Likewise, phage therapy is enjoying renewed interest in Western medicine (24). Phage-derived proteins are currently being used as molecular machines (23), diagnostic agents (22), and therapeutic agents (15, 17, 22) and for drug discovery (14). However, the resurgence in phage popularity has only fanned the flames of an ongoing debate, namely, that of phage taxonomy. Many phage biologists feel that the current taxonomic classification scheme, devised over 3 decades ago by the International Committee on Taxonomy of Viruses (ICTV), is outdated and in need of revision in light of the postgenomic age. However, choosing one methodology on the basis of which to classify all known and yet to be discovered phage is an area where phage researchers agree to disagree. In this issue of the *Journal of Bacteriology*, Chibani-Chennoufi et al. provide support for a comparative genomics-based alternative classification scheme (6).

EPOCHS OF PHAGE TAXONOMY

Since the discovery of phage independently by Frederick Twort in 1915 (26) and Felix d'Herelle in 1917 (7), taxonomic classification of phage has gone through several epochs. In the 1920s and 1930s it was observed that different phage types, or "races," displayed bacterial host specificity; thus, the field of phage typing was founded (8) and an initial classification system for phage was established. The advent of the electron microscope allowed phage biologists to measure the physical size of phage and the length of the tail fibers and to determine the symmetry of the capsid (16), giving rise to a taxonomy based on morphotypes in the 1940s and 1950s. During the 1960s, advances in biochemical methods to isolate nucleic acids from phage and resolve both the genome size and type (single-stranded DNA, double-stranded DNA, single-stranded RNA, double-stranded RNA) added further information to augment phage taxonomic schemes (25).

The first global attempt to systematically classify viruses took place at the International Congress of Microbiology held in Moscow in 1966. This meeting established the ICTV, whose mission was to develop a universal taxonomic system for all viruses infecting animals, plants, fungi, bacteria, and later, archaea. Beginning with its first report in 1971, the ICTV has met regularly to update virus definitions and taxonomy guidelines, with the seventh and most recent report published in 2000 (27). The ICTV conventions are based on a hierarchical system originally devised by Linnaeus in 1758 to classify plants and animals (13). On this premise, viruses are grouped together by shared characteristics, with subgroups having smaller clusters of shared attributes. For example, the tailed order of phage (*Caudovirales*) is broken down into three families: phage with long, contractile tails (*Myoviridae*), phage with long, noncontractile tails (*Siphoviridae*), and phage with short tails (*Podoviridae*). The families are further broken down into genus and subgenus by criteria such as genome configuration (linear, circular, supercoiled), host range, and genome size.

However, there are obvious shortcomings of the ICTV paradigm that critics are justifiably quick to point out. First, this model largely overlooks genomic and proteomic information. Since it is now possible to sequence whole genomes in a matter of weeks, and the entire proteome for a particular phage can be run on a single two-dimensional gel, one must question the wisdom of disregarding this wealth of information. For instance, despite a complete lack of DNA sequence relatedness, the ICTV lists both the *Salmonella* P22 phage and T7 as belonging to the *Podoviridae* family based on the shared presence of short tails. Yet it has been known for 30 years that P22 is in fact so genetically related to the long-tailed lambda phage that recombination between the two genomes forms functional hybrids (2). Thus, it would make more sense for P22 to be clustered with the lambdoid-like phage of the *Siphoviridae* family in spite of its short tails. Second, the ICTV system is based on vertical transmission of genetic characteristics, whereas phage genomes are known to be highly mosaic owing to horizontal exchange among common genetic pools (11, 18). Third, the model is dependent on electron microscopy for assignment to a given taxonomic family. This is problematic for a multitude of reasons. Whole genomes are being sequenced with such speed that some authors are not even taking the time to obtain electron micrographs. Alternately, it may be impossible to isolate corresponding phage particles when the genome data are derived from a whole-community sequencing project (4). In other instances, the phage may exist as a lysogenic prophage that does not produce mature virions. Recent genome sequencing of the bacterium *Streptococcus pyogenes* revealed the presence of six prophage or prophage-like elements, none of which are recognized by the ICTV (9).

The inadequacy of the ICTV classification is evident when one looks closely at the numbers. Of the completed phage genomes currently deposited in GenBank (1), 40% (92 of 228) are unclassified beyond the level of family according to ICTV conventions. Furthermore, an additional 10% (23 of 228) are not even assigned an order and are simply listed as "unclassi-

* Mailing address: Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, 1230 York Ave., New York, NY 10021. Phone: (212) 327-8167. Fax: (212) 327-7584. E-mail: nelsond@rockefeller.edu.

fied bacteriophage." Thus, the current taxonomic system fails to classify half of all the phage for which the entire genomic sequence is known!

A NEW SYSTEM OF PHAGE TAXONOMY?

Complaining about the current phage classification scheme does not accomplish anything unless one is prepared to propose a solution. It is generally agreed that future phage classifications must reflect genomic data as a primary component. In bacteria, this is easily accomplished by examining the conserved 16S ribosomal genes. However, phage lack ribosomal DNA and there are no conserved gene or protein sequences common to all phage on which to base a classification (21). Furthermore, any immediate attempt to reclassify phage based exclusively on comparative genomics may be biased toward the lambdoid phages or those involved in industrial fermentation, since these phages dominate the current assemblage of sequenced genomes. Nonetheless, in the fall of 2002, three separate phage research groups proposed alternative classification schemes as outlined below.

An option proposed by Rohwer and Edwards is based on a "phage proteome tree," which is constructed by grouping phage both relative to their near neighbors and in the context of all other phage (21). This method analyzed the entire predicted proteome for a given phage by using either the BLASTP or the PROTDIST program, and the results were then transformed into a distance matrix. By parsing the data according to cutoff E-values for BLASTP or penalty scores for PROTDIST, a tree was constructed based on relationships between phage proteins. The authors list several anomalies in the ICTV system that could be resolved by the proteome tree, including a reclassification of the P22 phage (Podoviridae) mentioned above into the lamboid-like family *Siphoviridae*. Additionally, the proteome tree moves the enteric PRD1 phage, which the ICTV classifies as *Tectiviridae* due to the presence of a lipid membrane below the capsid, to a subgenus of the family *Podoviridae* that shares a similar protein-primed DNA replication machinery.

A second approach, put forth by a group at the Pittsburgh Bacteriophage Institute, suggests that it may be impossible to have a strictly hierarchical taxonomic system given the genetic mosaicism arising from horizontal gene transfer among phage (12). In this paradigm, the top levels of taxonomy would still follow the hierarchical Linnean approach, i.e., viruses would be divided into "domains" according to genome type (double-stranded DNA, single-stranded DNA, single-stranded RNA, and double-stranded RNA), with a further partition known as "divisions" to separate defining characteristics such as tailed phage from filamentous phage. However, from this point on, three basic tenets would guide the remainder of the classification. First, members of a group should exhibit similarity in one or more loosely defined cohesion mechanisms. Second, significant sequence data, preferably from whole genomes, should be available for evolutionary assignment to a taxonomic cluster. Third, the groups may be reticulate, i.e., phage may simultaneously belong to several groups based on multiple and/or differing criteria from the first two tenets. This web-like design has an inherent flexibility that is not afforded by either phenetic or genetic hierarchical approaches. Nonetheless, one can envision how this system could become unwieldy as the complexity multiplies with every additional reticulate classification.

Finally, a third classification scheme, based on comparative genomics of a structural gene module, has been proposed (19). Since it is believed that the structural genes are the oldest and most conserved module in dairy phage, dot plots of temperate lactococcal phage were used to observe graded relatedness between DNA and protein sequences as well as similarity in the organization of the structural genes in the absence of sequence relatedness. Whereas comparative genomics of nonstructural genes tended to lump all of the lactococcal phages studied into one species, comparative genomics of the structural genes delineated four phage species and distinguished two genera based on head morphology. Using these criteria, the authors proposed a reclassification of the lactococcal phages into five distinct genera.

In this issue of the *Journal of Bacteriology*, Chibani-Chennoufi and colleagues at the Nestlé Research Center offer further support for the comparative genomics approach based on structural genes (6). They begin by providing the complete genome sequence of bacteriophage LP65, which infects *Lactobacillus plantarum*, an organism important to the fermentation industry. Although the detailed characterization of the phage, including spectacular electron micrographs of ultrastructural features, is superb, the most significant finding comes from comparative genomics. The LP65 structural genes share extensive protein sequence identity with those of *Bacillus* phage SPO1, *Listeria* phage A511, and *Staphylococcus* phage K, all members of the SPO1-like genus of the family *Myoviridae*. Significantly, it was also determined that the structural module of the SPO1-like genus shares a closer relationship with the lambda superfamily of *Siphoviridae* than it does with other genera of the *Myoviridae*. In addition, it was noted that after contraction, reorganization of the baseplate for both SPO1 and LP65 is markedly different from that for the T4-like *Myoviridae*, implying that contractile tails may have evolved in *Myoviridae* through convergent evolution. Furthermore, comparative genomics data and morphological data, taken together, suggest that the genus of SPO1-like *Myoviridae* forms a bridge between the tail-contracting *Myoviridae* and the noncontractile *Siphoviridae*.

THE OUTLOOK FOR PHAGE TAXONOMY

In the end, it may not be such a bad idea that we agree to disagree. With the exception of a few remarkable outliers, some of which are noted above, all three approaches would ultimately cluster the majority of bacteriophage within the same groups in which they are currently classified under the ICTV conventions. Consequently, selecting one method over another as a universal taxonomic scheme is unlikely to have any immediate significant impact. Chibani-Chennoufi et al. do provide good evidence that comparative genomics works well for the SPO1-like genus; however, given that the current knowledge of phage genomics comprises a relatively narrow spectrum of closely related phage groups (coliphage, dairy phage, etc.), it may be premature to establish the criteria for a universal system. We are just now beginning to sample whole-genome communities of phage and are discovering that phage diversity is even greater than we had previously postulated.

Sequencing expeditions similar to those recently carried out at Scripps Pier and Mission Bay (4) and the Sargasso Sea (28) have revealed that the majority of all new phage sequences are completely unrelated to anything contained in GenBank. Indeed, it has been estimated that we have sampled less than 0.0002% of the global phage metagenome (20). Given our rather myopic view of phage diversity, perhaps we will have to wait until larger sequence collections of unrelated bacteriophage genomes become available before we attempt to reconstruct the phage phylogenetic tree. In the meantime, the debate continues. This new report by Chibani-Chennoufi et al. maintains the open dialogue between phage taxonomists and further validates the use of structural genes as one alternative approach to classification.

ACKNOWLEDGMENTS

I thank Vincent A. Fischetti, Head of the Laboratory of Bacterial Pathogenesis and Immunology, for comments on the manuscript and general mentoring in his laboratory. Additionally, I thank Raymond Schuch, Mattias Collin, and John Mayo for critical reading of the manuscript.

Research in the laboratory is supported by grants from the James D. Watson Investigator Program of the New York State Office of Science, Technology, and Academic Research (NYSTAR) to D.N. and from the Defense Advanced Research Projects Agency (DARPA) to V.A.F.

REFERENCES

1. **Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler.** 2004. GenBank: update. Nucleic Acids Res. **32:**D23–D26.
2. **Botstein, D., and I. Herskowitz.** 1974. Properties of hybrids between *Salmonella* phage P22 and coliphage lambda. Nature **251:**585–589.
3. **Boyd, E. F., B. M. Davis, and B. Hochhut.** 2001. Bacteriophage-bacteriophage interactions in the evolution of pathogenic bacteria. Trends Microbiol. **9:**137–144.
4. **Breitbart, M., P. Salamon, B. Andresen, J. M. Mahaffy, A. M. Segall, D. Mead, F. Azam, and F. Rohwer.** 2002. Genomic analysis of uncultured marine viral communities. Proc. Natl. Acad. Sci. USA **99:**14250–14255.
5. **Burroughs, N. J., P. Marsh, and E. M. H. Wellington.** 2000. Mathematical analysis of growth and interaction dynamics of streptomycetes and a bacteriophage in soil. Appl. Environ. Microbiol. **66:**3868–3877.
6. **Chibani-Chennoufi, S., M.-L. Dillmann, L. Marvin, S. Rami-Shojaei, and H. Brussow.** 2004. *Lactobacillus plantarum* bacteriophage LP65: a new member of the SPO1-like genus of the family *Myoviridae.* J. Bacteriol. **186:**7069–7083.
7. **d'Herelle, F. H.** 1917. Sur un microbe invisible antagoniste des bacilles dysenteriques. C. R. Acad. Sci. **165:**373–375.
8. **Evans, A. C.** 1934. *Streptococcus* bacteriophage: a study of four serological types. Public Health Rep. **49:**1386–1401.
9. **Ferretti, J. J., W. M. McShan, D. Ajdic, D. J. Savic, G. Savic, K. Lyon, C. Primeaux, S. Sezate, A. N. Suvorov, S. Kenton, H. S. Lai, S. P. Lin, Y. Qian, H. G. Jia, F. Z. Najar, Q. Ren, H. Zhu, L. Song, J. White, X. Yuan, S. W. Clifton, B. A. Roe, and R. McLaughlin.** 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes.* Proc. Natl. Acad. Sci. USA **98:**4658–4663.
10. **Hendrix, R. W.** 2003. Bacteriophage genomics. Curr. Opin. Microbiol. **6:**506–511.
11. **Hendrix, R. W., M. C. M. Smith, R. N. Burns, M. E. Ford, and G. F. Hatfull.** 1999. Evolutionary relationships among bacteriophages and prophages: all the world's a phage. Proc. Natl. Acad. Sci. USA **96:**2192–2197.
12. **Lawrence, J. G., G. F. Hatfull, and R. W. Hendrix.** 2002. Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. J. Bacteriol. **184:**4891–4905.
13. **Linnaeus, C.** 1758. Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis, 10th ed., vol. 1. Stockholm, Sweden.
14. **Liu, J., M. Dehbi, G. Moeck, F. Arhin, P. Bauda, D. Bergeron, M. Callejo, V. Ferretti, N. Ha, T. Kwan, J. McCarty, R. Srikumar, D. Williams, J. J. Wu, P. Gros, J. Pelletier, and M. DuBow.** 2004. Antimicrobial drug discovery through bacteriophage genomics. Nat. Biotechnol. **22:**185–191.
15. **Loeffler, J. M., D. Nelson, and V. A. Fischetti.** 2001. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. Science **294:**2170–2172.
16. **Luria, S. E., M. Delbruck, and T. F. Anderson.** 1943. Electron microscope studies of bacterial viruses. J. Bacteriol. **46:**57–67.
17. **Nelson, D., L. Loomis, and V. A. Fischetti.** 2001. Prevention and elimination of upper respiratory colonization of mice by group A streptococci using a bacteriophage lytic enzyme. Proc. Natl. Acad. Sci. USA. **98:**4107–4112.
18. **Pedulla, M. L., M. E. Ford, J. M. Houtz, T. Karthikeyan, C. Wadsworth, J. A. Lewis, D. Jacobs-Sera, J. Falbo, J. Gross, N. R. Pannunzio, W. Brucker, V. Kumar, J. Kandasamy, L. Keenan, S. Bardarov, J. Kriakov, J. G. Lawrence, W. R. Jacobs, Jr., R. W. Hendrix, and G. F. Hatfull.** 2003. Origins of highly mosaic mycobacteriophage genomes. Cell **113:**171–182.
19. **Proux, C., D. van Sinderen, J. Suarez, P. Garcia, V. Ladero, G. F. Fitzgerald, F. Desiere, and H. Brussow.** 2002. The dilemma of phage taxonomy illustrated by comparative genomics of Sfi21-like *Siphoviridae* in lactic acid bacteria. J. Bacteriol. **184:**6026–6036.
20. **Rohwer, F.** 2003. Global phage diversity. Cell **113:**141.
21. **Rohwer, F., and R. Edwards.** 2002. The Phage Proteomic Tree: a genome-based taxonomy for phage. J. Bacteriol. **184:**4529–4535.
22. **Schuch, R., D. Nelson, and V. A. Fischetti.** 2002. A bacteriolytic agent that detects and kills *Bacillus anthracis.* Nature **418:**884–889.
23. **Smith, D. E., S. J. Tans, S. B. Smith, S. Grimes, D. L. Anderson, and C. Bustamante.** 2001. The bacteriophage ϕ29 portal motor can package DNA against a large internal force. Nature **413:**748–752.
24. **Thacker, P. D.** 2003. Set a microbe to kill a microbe: drug resistance renews interest in phage therapy. JAMA **290:**3183–3185.
25. **Thomas, C. A., Jr., and J. Abelson.** 1966. The isolation and characterization of DNA from bacteriophages. Prog. Nucleic Acid Res. **1:**553–561.
26. **Twort, F. W.** 1915. An investigation on the nature of ultra-microscopic viruses. Lancet **ii:**1241–1246.
27. **van Regenmortel, M. H. V., C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wickner.** 2000. Virus taxonomy: classification and nomenclature of viruses. Seventh report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, Calif.
28. **Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. T. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y.-H. Rogers, and H. O. Smith.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science **304:**66–74.
29. **Wagner, P. L., and M. K. Waldor.** 2002. Bacteriophage control of bacterial virulence. Infect. Immun. **70:**3985–3993.

The views expressed in this Commentary do not necessarily reflect the views of the journal or of ASM.