Size does matter – distinguishing bacteriophages by genome length (and 'breadth')

Stephen T Abedon

Department of Microbiology
The Ohio State University
1760 University Drive
Mansfield, OH 44906, USA
Tel (1) 419-564-5538,
Fax (1) 419-755-4327,
Email abedon.1@osu.edu

Genome size and nucleic acid composition are basic characteristics of bacteriophages (phages). With larger genomes come increases in gene number, greater infection and virion sophistication, higher likelihood of gene acquisition via horizontal gene transfer (HGT), and additional genetic redundancy both within and between genomes (the latter referring to gene duplication and redundancy with host functions, respectively). Larger, that is, tailed phages in fact are among the most recombinogenic organisms on earth, often possessing numerous genes of mysterious function, many of which in at least some phages can be deleted without greatly impacting laboratory propagation. Larger phages, through the use of more sophisticated polymerases as well as proofreading and genome repair functions, also are less prone to replicative infidelities. As a consequence, their per-nucleotide mutation rates are substantially lower than those of smaller phages. Here, rather than macroevolutionary trends as typically considered from the perspective of comparative genomics, I emphasise instead microevolutionary processes, particularly the role of mutation as an immediate source of genetic variation within populations. Within this context I differentiate the viruses of domain *Bacteria* into four genome-size categories: very small (single-stranded RNA phages), small (single-stranded DNA phages), medium (lipid-containing, double-stranded DNA, tailless phages), and large, or at least larger (tailed phages).

Costs and utility of small genomes

Lower replication fidelity is particularly apparent in the smallest of phages, those with single-stranded RNA genomes, though small DNA phages, with their single-stranded genomes, also display comparatively high per-nucleotide mutation rates[1]. Mutation, consequently, can lead to rapid genetic diversification especially in populations of smaller phages, both per genome (single-stranded RNA phages) and per nucleotide (single-stranded RNA and single-stranded DNA phages). RNA viruses, in fact, are so mutation-prone that they have been loosely described as quasispecies[2] – massively exploring sequence space and thereby potentially adapting extremely rapidly to environmental changes (Figure 1). The utility of ever-higher organism mutation rates has a natural upper limit. Though the ultimate source of variation upon which natural selection can act, most mutations nonetheless are harmful. Rare, beneficial mutations thus will be paired with detrimental mutations too often if mutation rates too far exceed one per genome per replication event. Along with more general fitness concerns, genomes as a result can be only so large for a given per-nucleotide mutation rate. Phages with small genomes therefore can be somewhat trapped in their display of higher mutation rates: They may not have the ability, given high per-nucleotide mutation rates along with genome-packaging constraints, to display sufficiently large genome sizes to allow the encoding of fidelity-enhancing adaptations.

If genome lengths are short while genome 'breadth', the number of strands that must be replicated, also is "small" – that is, single-

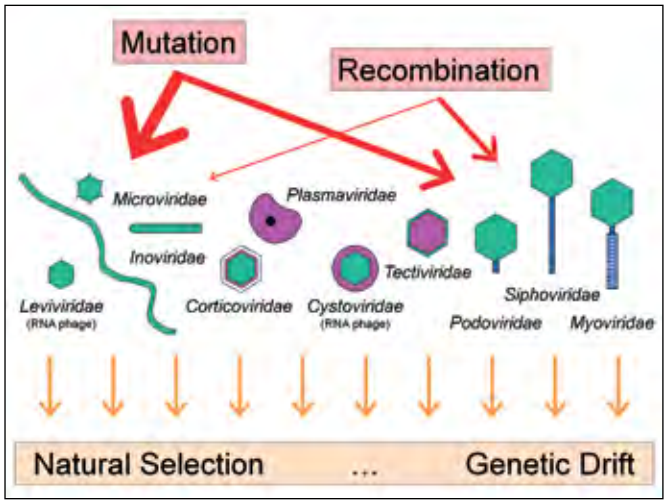

Figure 1. Schematic representation of the microevolutionary importance of mutation versus recombination (HGT) as sources of genetic variation in phages of different sizes. Mutations are less abundant in larger phages, on a per-nucleotide and therefore per-gene basis, whereas in smaller phages HGT is less abundant (thinner arrows). Smaller, single-stranded phages are found to the left. These are the families *Leviviridae* (RNA), *Microviridae* (DNA), and *Inoviridae* (DNA), the latter being filamentous and shown as two morphological types. Medium-sized, tailless phages, all of which have double-stranded genomes, are shown in the middle. These include the families *Corticoviridae* (DNA), *Plasmaviridae* (DNA), *Cystoviridae* (RNA), and *Tectiviridae* (DNA). Tailed phages, all of which have double-stranded DNA genomes, are shown to the right (order *Caudovirales* consisting of the families *Myoviridae, Siphoviridae,* and *Podoviridae).* Note that details of adsorption appendages, particularly tail fibres, are not shown. Virion illustrations are derived from Hyman and Abedon[9] and Ackermann[8] and phage families are shown in approximate increasing genome-size order (going from left to right). As presented here, magenta illustrates lipid content; capsid proteins, that is, as surround the phage nucleic acid, are green; and blue corresponds to tail structures.

stranded versus double-stranded – then a phage should be able to replicate that genome more easily. The number of proteins necessary to encapsidate a small-genomed phage also is tiny relative to that necessary to encapsidate a larger genomed phage. The result is a potential to make new phages faster, which can allow these phages to display impressive population growth rates and/or superlative yields through latent period shortening or burst size enhancement. In addition, as particle diffusion rates are functions of particle shape and size[3], having virions that are both small and simple (no tails) should enhance diffusion-limited rates of bacterial encounter. Smaller-genomed phages thus can be cheaper to generate, more mobile, and also can be produced in greater numbers per unit time than the virion particles of larger phages.

Phage genome size categories

Conceivably reflecting differences in how phage evolution and ecology might be optimised in terms of both phage genome length and strandedness, there seems to exist little overlap among four phage genome-size categories: very small, small, medium, and large. Very small phages, all single-stranded RNA and members of family *Leviviridae*, have genomes that range in size from about 3.5 to the low 4 kb range. Small phages, the smallest of which are members of family *Microviridae*, all have single-stranded DNA and have genomes that are slightly larger at their lower end than members of family *Leviviridae*. These range from about 4.5 kb up to approximately 9 kb. Medium-sized phages, all double-stranded, have genomes that range in size from 9 kb up to about 15 kb. Finally, larger-genomed phages, all double-stranded DNA and tailed, have genomes that are 16 kb and greater, though with one exception among sequenced, non-defective phages that have been morphologically characterised (see legend of Figure 2). Contrasting both smaller and larger phages, the virions of all medium-sized phages are noteworthy for their lipid content (Figure 1).

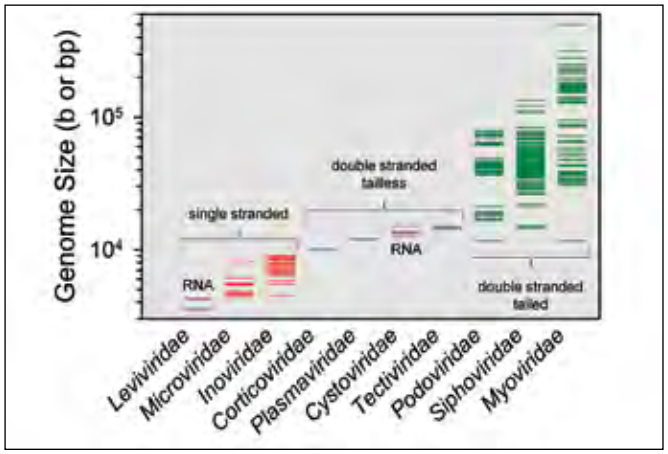

Figure 2. Phage morphologies and genome sizes. Except as indicated, these are based on the NCBI phage genome database (www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=10239&type=6&name=Phages) as well as the list provided in Abedon[4]. Where possible, morphologies were checked against the Bacteriophage Names 2000 database (www.phage.org/names.htm). In addition, I have included the *Bacillus* phage G, which has a reported genome size of 497.5 kb[7]. Tailed phages with genomes sizes below 16 kb include the 11.7 kb *Mycoplasma* phage P1 (family *Podoviridae);* the 11.6 kb *Myoviridae* defective/satellite phage, P4; and the 15.0, 14.5, and 15.2 kb (respectively) defective prophages biL310, biL311, and biL312 (host, *Lactococcus),* which according to the NCBI database have a *Siphoviridae* morphology, though this description is both controversial and something that I have been unable to confirm. At this time there are at least six additional sub-16 kb phages in the NCBI database that are not shown in the figure because they are of unknown morphology.

In Figure 2, I graph phage families against genome size, which range from the smallest members of the *Leviviridae*, at approximately 3.5 kb, to that of the *Bacillus* phage G, a member of family *Myoviridae*, which has a genome size of nearly 500 kb. In terms of total nucleotides present, with two times as many nucleotides per 'base' in double-stranded chromosomes, phage genome sizes span an over 250-fold range, or perhaps more if one considers the unclassified *Leuconostoc* phage, L5 (2.4 kb).

Further reading

For earlier consideration of the issues presented here, see Abedon[4]. For a primer on phage evolutionary biology, see Duffy and Turner[5]. Mutation rates are considered in both of these publications, but see also Duffy *et al.*[6]. The role of recombination especially among tailed phages is covered by Hendrix[7] while the taxonomy of phages is discussed in various places by Ackermann[8].

Acknowledgement

Thank you to Bob Blasdel for helpful discussion and to Rob Lavigne, Hans Ackermann, Cameron Thomas-Abedon, and an anonymous reviewer for helpful comments.

References

1. Cuevas, J.M. *et al.* (2009) Point mutation rate of bacteriophage ΦX174. *Genetics* 183, 747–749.
2. Holmes, E.C. (2010) The RNA virus quasispecies: fact or fiction? *J. Mol. Biol.* 400, 271–273.
3. Murray, A.G. and Jackson, G.A. (1992) Viral dynamics: A model of the effects of size, shape, motion and abundance of single-celled planktonic organisms and other particles. *Mar. Ecol. Prog. Ser.* 89, 103–116.
4. Abedon, S.T. (2009) Phage evolution and ecology. *Adv. Appl. Microbiol.* 67, 1–45.
5. Duffy, S. and Turner, P.E. (2008) Introduction to phage evolutionary biology. In *Bacteriophage Ecology* (Abedon, S.T., ed.), pp. 147–176, Cambridge University Press.
6. Duffy, S. *et al.* (2008) Rates of evolutionary change in viruses: patterns and determinants. *Nature Rev. Genet.* 9, 267–276.
7. Hendrix, R.W. (2009) Jumbo bacteriophages. *Curr Top. Microbiol. Immunol.* 328, 229–240.
8. Ackermann, H.-W. (2006) Classification of bacteriophages. In *The Bacteriophages* (Calendar, R. and Abedon, S.T., eds), pp. 8–16, Oxford University Press.
9. Hyman, P. and Abedon, S.T. (2009) Bacteriophage (overview). In *Encyclopedia of Microbiology* (Schaecter, M., ed.), pp. 322–338, Elsevier, Oxford.

Biography

Steve Abedon is an Associate Professor of Microbiology at The Ohio State University, USA. His single-authored monograph, *Bacteriophages and Biofilms: Ecology, Phage Therapy, Plaques*, was published early in 2011 (Nova Science Publishers). He also has edited or contributed to the editing of three volumes: a special issue on phage therapy published in the journal, *Current Pharmaceutical Biotechnology* (2010; www.bentham.org/cpb/contabs/cpb11-1.htm); the monograph, *Bacteriophage Ecology* (2008, Cambridge University Press); and the second edition of the monograph, *The Bacteriophages* (2006, Oxford University Press). In 1996 he founded the Bacteriophage Ecology Group (www.phage.org) and, since 1989, has published approximately 50 articles and chapters, most of which consider aspects of bacteriophage ecology and evolution.