**Phi8**

Zachary Romer1,2, Siobhan Watkins3, and Catherine Putonti1,2,3\*

1 Bioinformatics Program, Loyola University Chicago, Chicago, IL 60660

2 Department of Computer Science, Loyola University Chicago, Chicago, IL 60611

3 Department of Biology, Loyola University Chicago, Chicago, IL 60660

\* Corresponding author: cputonti@luc.edu

**Abstract**

Strong correlations between codon biases in bacteriophage and the highly expressed genes of their hosts, a likely result of selection for translational efficiency and/or accuracy, have been observed for a number of species. These correlations are most apparent for phage species with narrow host ranges. The codon usage of phage with an expanded host-range, however, can be the result of competing forces if, for example, the phage infects host species with different codon biases or strengths of these biases. We have conducted bioinformatic and experimental examination into the dsRNA *Cystoviridae* family which includes species with both the specialist and generalist phenotype. While the native host of these phage is *Pseudomonads* having little or no biases in their codon usage, three species are capable of also infecting the translationally biased *Salmonella typhimurium*. SOME CONCLUDING STATEMENT

**Keywords:**

**Introduction**

Biases in nucleotide usage and preferences in the usage of synonymous codons have long been known to occur within the genomes of species across all branches of the tree of life. This bias is often strongest within highly expressed genes. Within bacterial species, the strength of this bias correlates to the organism’s doubling time or generation time as the time spent in translation can be a limiting factor in cell division (Higgs and Ran 2008). Correspondence in virus-host codon usage has been observed for a wide variety of bacteriophages (Kunisawa *et al*. 1998; Lucks *et al*. 2008; Carbone 2008). Given the fact that viruses make substantial use of host biosynthesis, it stands to reason that these viruses would be under selection to utilize codons that confer a translational advantage and thus increased fitness within their host. This correspondence is most prominent for those phage species which infect translationally biased bacterial hosts. While bioinformatic analyses revealing these trends have primarily focused on phage with a limited host range (specialists) that are well adapted to laboratory conditions, phage isolated from the environment have repeatedly revealed the ability to infect bacterial species belonging to different taxa (some reference). Elucidating the effect of selection shaping codon usage within these generalists presents new challenges, particularly when the host species vary in their codon preferences and/or the strength of their codon biases.

The members of the 3-segmented dsDNA *Cystoviridae* family of phage provide an opportunity to explore the effects of translational selection on viral codon usage as this family includes species with different hosts and host-ranges. The five members within this family – Φ6, Φ8, Φ12, Φ13, and Φ2954 – were all isolated from plant leaves and infect pseudomonads, particularly the strain *Pseudomonas syringae* pv. phaseolicola (Vidaver et al. 1973; Mindich et al. 1999) although some of these species have been found to be able of infecting other *Pseudomonas* species and *P. syringae* strains (Mindich et al. 1976, Guyader and Burch 2008; Bono et al. 2013). While Φ6 and Φ2954 are restricted to infection via the pilus (Mindich et al. 1999; Qiao et al. 2010), Φ8, Φ12 and Φ13 attach to their host cells directly via the rough LPS and thus can infect bacterial species with a similar outer membrane including *Salmonella typhimurium* (Mindich et al. 1999)*.* This expanded host-range presents the phage with two rather different environments. While *S. typhimurium* is a translationally biased host, pseudomonads have little or no selective pressures shaping their codon usage within their set of highly expressed genes; the strength of selected codon usage bias of *P. syringae* is half that of *S. typhimurium* (Sharp et al. 2005). This is further reflected in their difference in doubling time; the generation of *Salmonella* is half that of *Pseudomonas* (Cooper & Ruettinger; Young et al; Rocha 2004). Because the cystoviruses are completely dependent upon their host(s) for biosynthesis, the availability of host resources have likely left their mark on the genome.

In an effort to begin to tease apart the potentially confounding pressures, we have conducted bioinformatic and experimental analysis into the codon usage and fitness of the five cystoviruses*.*

**Results**

***Codon usage of generalist Cystoviruses***

The codon usage of each gene within the five *Cystoviridae* genomes was calculated and compared to the codon usage preferences within the highly expressed genes of *P. syringae*. While the average codon preferences of the specialist Φ6 were more similar to *P. syringae* than Φ8’s, the average codon preferences of Φ12, Φ13 and Φ2954 were found to be more similar to *P. syringae* than Φ6 and Φ8. While one would expect the phage codon usage to be fine-tuned to the host tRNA abundances, one must recall that *P. syringae* is a weakly biased species. Moreover, both bacterial species have relatively similar codon usage profiles within their set of highly expressed genes despite their difference in the strength of selected codon usage bias (*R*2=0.8489). There are, however, different biases present for several amino acids, namely glycine, lysine, leucine, proline, serine and valine. Further investigation into these key amino acids does not reveal a statistically significant difference between the usage preferences of the generalist and specialist species relative to the two hosts’ usage profiles.

***Fitness of Generalist Φ8 in Different Hosts***

The fitness of the generalist Φ8 was examined in each of its hosts, *P. syringae* and *S. typhimurium* (see Methods). The growth rate, generation time, and latent period of Φ8 was longer when infecting *P. syringae* (Table 1). Moreover, the phage was more quickly absorbed into the *S. typhimurium* host cells. While Φ8 plaques were found to be turbid when plated on both hosts, the plaques on the *S. typhimurium* lawns were found to be slightly larger. Despite the shorter generation time of *S. typhimurium*, Φ8 is able to take advantage of the host’s biosynthesis machinery and lyse quickly the cell. CAN WE PUT IN SOME GROWTH CURVES? I THINK THAT’D MAKE IT COMPLETE

***P. syringae***

***S. typhimurium***

Phage generation time

120 min

20 min

Latent period

40 min

9 min

Growth rate

0.0247 phage/μL min

0.0113 phage/μL min

Attachment rate

1.239 10-11

9.1551 10-11

**Table 2.** Comparison of growth parameters for Φ8 in each host.

**Concluding Remarks**

In sharp contrast to phage which can only infect a single and strongly translationally biased host, such as ΦX174 and its host *E. coli* (Carbone 2008), the effect of translational selection of the *Cystoviridae* generalist species is not clear. It is important to note, however, that the sequenced strains used for this analysis were isolated from growth in the laboratory with *P. syringae*. Thus, long-term passage of the generalist cystoviruses in *S. typhimurium* may result in the rise of a more *Salmonella*-like codon usage. This, however, may come at a cost; the *Salmonella* adapted strain could become specialized or possibly lead to reduced fitness in *P. syringae*. The assays conducted here for Φ8 provide insight into the fitness of this phage within each of its hosts. The fact that Φ8 does not display a strong correspondence to the codon preferences of either host may in fact provide it with the opportunity to successfully infect both pseudomonads and *S. typhimurium* and thus maintain its generalist phenotype.

The explicit quantification of the infectivity of Φ8 within both *P. syringae* and *S. typhimurium* provides a foundation from which future studies into the cost of expanded host-range can be launched. In contrast to existing “generalist” laboratory phage strains, Φ8 infects two species belonging to different taxonomical orders which are both well-studied, well-characterized, and lab-adapted bacterial species. In addition to providing a means to investigate the cost of host-range at the phenotypic level, the Φ8-*Pseudomonas*-*Salmonella* system can be used as a model for unraveling the effects of translational selection as these two host species present very different pressures.

**Methods and Materials**

**Calculating Codon Usages and Biases**

The annotations and sequences for each of segments of the five *Cystoviridae* sequences were obtained from NCBI (Table 2). The codon usage for each gene was calculated using software previously developed by the authors (Hilterbrand et al. 2012). Furthermore, the codon usage of each coding region in the viral genomes was compared to the codon usage of the highly expressed genes within the *Pseudomonas syringae* pv. phaseolicola (NC\_005773) and *Salmonella enterica subsp. enterica* serovar Typhimurium str. LT2 genome (NC\_003197). Variations in codon biases for these two bacterial hosts can be seen via the resource CBDB (Hilterbrand et al. 2012).

**Species**

**Means of Host Attachment**

**Segment**

**S**

**M**

**L**

Φ6

pilus

NC\_003714

NC\_003716

NC\_003715

Φ8

LPS

NC\_003301

NC\_003300

NC\_003299

Φ12

LPS

NC\_004174

NC\_004175

NC\_004173

Φ13

LPS

NC\_004170

NC\_004171

NC\_004172

Φ2954

pilus

NC\_012093

NC\_012092

NC\_012091

**Table 2.** Genome Sequences for annotated *Cystoviridae* species.

**Fitness Assays of Φ8 in Different Hosts**

Both *S. typhimurium* (#, obtained from ATCC) and *P. syringae* var. phaseolicola (HER #1419, obtained from the Félix d’Hérelle Reference Center for Bacterial Viruses) were maintained in liquid culture with LB media. 6 mL of *S. typhimurium* culture at log phase was inoculated with 50 μL Φ8 (HER #419, obtained from the Félix d’Hérelle Reference Center for Bacterial Viruses) diluted in 950 μL saline and placed into a shaking incubator at 37°C; 0.5 mL samples were taken from the inoculate every 10 minutes for 2 hours and immediately centrifuged at 13000 rpm for 10 minutes. 9 mL of *P. syringae* culture at log phase was inoculated with 50 μL Φ8 diluted in 950 μL saline and placed into a shaking incubator at 27°C; 0.5 mL samples were taken from the inoculate every 10 minutes for 3 hours and centrifuged. The samples from each collection were then diluted in an 8-fold 10x dilution series and spotted on lawns composed of 1 mL of naïve host in 0.7% LB soft agar. (Due to the “sticky” nature of *P. syringae*, the 1 mL used for plating was vortexed for 20 seconds prior to plating the lawn.) A dilution series was also conducted for the initial 50 μL Φ8 diluted in 950 μL saline and spotted on a *S. typhimurium* lawn, serving as a baseline for the titer of the inoculate. Each of these assays was conducted five times.

The growth rate of Φ8 in each of these hosts was calculated using the natural log concentration of each sample was plotted against time, following the protocols previously used by Dennehy & Turner (2) for Φ6; the slope of the regression on this data provides the growth rate of the phage. The phage latent period was calculated using the equation *tG = L + (kN)-1*, employed previously by Abedon & Stopar (3), where tG is the phage generation time as a function of the latent period *L*, the inverse of the product of the phage absorption constant *k* and host cell density *N*. The phage absorption constant was calculated as -*ln(NU/N0)/(C\*t)*, with *NU*and *N0*being the final and initial phage concentrations, respectively, *C* being the host cell density and *t* representing time.