1   <u>Title:</u> **Host traits drive viral life histories across phytoplankton viruses**

2

3   Kyle F. Edwards[1]*, Grieg F. Steward[2]

4

5   [1]Department of Oceanography

6   University of Hawai'i at Mānoa

7   kfe@hawaii.edu

8

9   [2]Department of Oceanography

10   Center for Microbial Oceanography: Research and Education

11   University of Hawai'i at Mānoa

12   grieg@hawaii.edu

13

14   *Corresponding author:

15   Department of Oceanography

16   University of Hawai'i at Mānoa

17   1000 Pope Road

18   Honolulu, HI 96822

19   808-956-6192

20   kfe@hawaii.edu

21

22   <u>Keywords</u>: burst size, latent period, genome size, evolution, adaptive dynamics, life

23   history

24

25   <u>Manuscript type:</u> Article

26   <u>Elements of expanded online edition</u>: Appendix Table S1, Supplementary Figures S1-S7

27

28

29

30

31

32

33

34

35

Abstract

Viruses are integral to ecological and evolutionary processes, but we have a poor understanding of what drives variation in key traits across diverse viruses. For lytic viruses, burst size, latent period, and genome size are primary characteristics controlling host-virus dynamics. Burst size and latent period are analogous to organismal traits of fecundity and generation time, and genome size affects the size of the virion as well as viral control of host metabolism. Here we synthesize data on these traits for 75 strains of phytoplankton viruses, which play an important role in global biogeochemistry. We find that primary traits of the host (genome size, growth rate) are major ecological drivers, explaining 40-50% of variation in burst size and latent period. We analyze an eco-evolutionary model to explore mechanisms underlying these patterns. We find that burst size may be set by the host genomic resources available for viral construction, while latent period evolves to permit this maximal burst size, modulated by host metabolic rate. These results suggest that general mechanisms may underlie the evolution of diverse viruses, which will facilitate our understanding of viral community processes, ecosystem impacts, and coevolutionary dynamics.

Introduction

59

60   Viruses are integral to the ecology and evolution of all cellular life (Villarreal and

61   Witzany 2010, Koonin and Dolja 2013). By selectively lysing or altering the physiology

62   of the cells they infect, and mediating horizontal gene transfer among genetically distinct

63   lineages, they can generate and maintain host diversity and influence biogeochemical

64   cycles (Suttle 2007, Breitbart 2012). Viruses are highly host-specific and exhibit great

65   taxonomic and genetic diversity, with many forms still being discovered and

66   characterized (Lang et al. 2009, Rosario et al. 2012, Fischer 2016). Although viruses

67   were discovered over a century ago, and some model systems have been studied in great

68   detail, we have a poor understanding of what drives variation in key traits across diverse

69   viruses. Research on functional trait diversity has been fruitful in linking physiology,

70   community structure, and ecosystem dynamics, particularly in terrestrial plants (Westoby

71   and Wright 2006) and phytoplankton (Litchman and Klausmeier 2008). If broad patterns

72   and mechanisms underlying how viruses function can be discerned, this will promote a

73   general framework for predicting viral ecology and the effects of viral interactions on

74   host populations, communities, and ecosystems (Gudelj et al. 2010). In this study we

75   examine viruses that infect phytoplankton, a group of viruses that have a direct influence

76   one of the most important biogeochemical transformations on the planet. Phytoplankton

77   account for nearly half of global primary production and participate in multiple elemental

78   cycles (Falkowski et al. 2004). Viral effects on mortality, element cycling, and

79   community structure of phytoplankton potentially have global consequences (Brussaard

80   2004, Weitz et al. 2015).

81   We will focus on several traits of lytic viruses, which always produce virions and

82   lyse the host cell, provided the cell has sufficient resources and integrity to support the

83   infection. Viruses replicating via the lytic cycle can have particularly dramatic effects on

84   host populations, because of their rapid rate of reproduction and corresponding

85   destruction of the host population. Among the key traits that define the function of lytic

86   viruses are burst size (new virions produced per infected host), latent period (time elapsed

87   between infection and lysis), and viral genome size. Burst size and latent period are

88   analogous to the organismal life history traits of fecundity and generation time, and are

89   key parameters for host-virus population dynamics. The burst size and latent period of a

90   virus can vary with environmental conditions or host genotype (Wilson et al. 1996, Maat

91   and Brussaard 2016), but these parameters also vary greatly among different viral strains

92   measured under similar conditions (typically, resource-replete exponential growth of the

93   host). For simplicity we will refer to differences across isolates as variation in 'viral

94   traits', although such differences are likely driven by both viral genotype and host

95   genotype. Viral genome size is a trait that will influence the ability of the virus to control

96   host metabolism, while also determining the metabolic cost of synthesizing new virions

97   (Bragg and Chisholm 2008, Thomas et al. 2011), as well as the size of the virion

98   (Steward et al. 2013), which will affect diffusivity and other processes (Murray and

99   Jackson 1992). Here we analyze variation in burst size, latent period, and genome size

100   across viruses, because these traits have been quantified for numerous isolates that infect

101   phytoplankton and other unicellular algae.

102   Guided by previous work, we can make predictions about drivers of variation in

103   these traits: conditions that may select for particular trait values, and constraints that may

104   cause the traits to covary. Viruses use their host's molecular machinery to reproduce, and

105   therefore host structure and physiology are primary selective forces for viral trait

106   evolution. Nucleotides for viral genome synthesis come from host nucleotide pools and

107   degradation of the host genome, with an uncertain and variable contribution of *de novo*

108   nucleotide synthesis during infection (Van Etten et al. 1984, Wikner et al. 1993, Brown et

109   al. 2007, Thompson et al. 2011). Host genome size may thus influence the rate of viral

110   production or burst size, and a prior synthesis of 15 host-virus pairs found that viral

111   nucleotide production correlates with host genome size (Brown et al. 2006). The growth

112   rate of the host will likely correlate with the concentration of ribosomes and enzymes

113   required to synthesize host proteins, RNA, and DNA, which are also used to construct

114   new virus particles (You et al. 2002, Daines et al. 2014). Therefore, host growth rate may

115   also affect the rate of viral production.

116   Within the context set by host conditions, viral traits should evolve to maximize

117   fitness. Burst size and latent period are intrinsically related, because lysing the host

118   sooner will reduce burst size, all else equal. Theory and experiments with *E. coli* show

119   that the optimal latent period (and burst size) depends on host density and quality, such

120   that a shorter latent period is selected for when hosts are more dense or of higher quality

121   (Wang et al. 1996, Abedon et al. 2003). However, these models and experiments assume

122   that host density is relatively constant, rather than being regulated by the virus, and so it

123   is less clear how latent period will evolve when host and virus populations are dynamic.

124   Furthermore, the benefit of a longer latent period may be diminished if viral reproduction

125   exhausts nucleotide resources within the host. Models that include bacteriophage-host

126   dynamics and latent period evolution have been developed recently (Bonachela and

5

127   Levin 2014). By parameterizing an eco-evolutionary model for phytoplankton-virus

128   interactions, and manipulating host traits relevant to our hypotheses, we can test whether

129   empirical patterns are consistent with theoretical expectations for particular mechanisms.

130   In sum, we can predict that burst size and latent period will tend to be correlated

131   across viral strains, but that both of these traits may be modulated by the genome size of

132   the host (larger hosts may permit a larger optimal burst size), the growth rate of the host

133   (faster growth may decrease optimal latent period), and the genome size of the virus

134   (larger viruses require more resources per virion, which may reduce optimal burst size).

135   To test these predictions we compile published data on burst size, latent period, and

136   genome size for lytic viruses isolated from phytoplankton/microalgae. We explore

137   hypothesized relationships among virus traits, and between virus and host traits. We also

138   analyze a model of viral trait evolution to explore potential mechanisms for the empirical

139   patterns of trait variation.

140

141   **Methods**

142   <u>Virus trait compilation</u>. The literature was searched for studies that measured burst size

143   and latent period of viruses isolated on phytoplankton or other microalgal hosts

144   (Appendix Table S1). We only included experiments where the host was grown under

145   nutrient-replete conditions, so that we could quantify functional variation across viruses

146   cultured under similar conditions, as opposed to plastic responses of individual strains.

147   We recorded the name of the virus strain, virus genome type

148   (dsDNA/dsRNA/ssDNA/ssRNA), virus source location, host species, host taxon

149   (chlorophyte/cryptophyte/cyanobacterium/diatom/dinoflagellate/haptophyte

6

150 /pelagophyte/raphidophyte), and environment (marine/freshwater). We also recorded

151 whether burst size was estimated by counting infectious units or free virions. Virus capsid

152 size (diameter) and genome size estimates were taken from the same study, or other

153 studies on the same isolate. Virus genome size correlates strongly with capsid diameter,

154 and thus we only use genome size to represent virus size in all analyses (Fig. S1). For

155 some analyses we quantify total viral nucleotide output at lysis (burst size * virus genome

156 size); for these calculations we divided the genome size of the single-stranded viruses by

157 a factor of 2. Host genome size and cell volume estimates were taken from the literature,

158 and if an estimate for the host species was not available, an estimate from a congener was

159 used if available. It is noteworthy that 10 of the 13 single-stranded viruses have been

160 isolated from *Chaetoceros* species, and, in the absence of published information on

161 genome sizes for most of the hosts, we assigned all of the host species the same genome

162 size (measured for *C. muelleri*). For the double-stranded viruses, genome size estimates

163 were available for nearly all host species. When possible, host exponential growth rate

164 was estimated using DataThief (Tummers 2006) to extract growth curves measured on

165 uninfected hosts, or hosts growing prior to infection. Temperature and irradiance under

166 which the hosts were cultured during one-step growth experiments were also recorded.

167

168 <u>Statistical methods.</u> Relationships between viral traits, or viral traits and host traits, were

169 analyzed using mixed models (R package lme4; Bates et al. 2015). We used random

170 effects to appropriately account for non-independence in the data resulting from multiple

171 viruses infecting similar hosts (host genus), multiple viruses measured in the same study

172 (study ID), host taxonomy (diatom/cyanobacteria/haptophyte/etc.), or similarities among

173   virus type (dsDNA/dsRNA/ssDNA/ssRNA). These random effects were included in all

174   models, and different fixed effects and response variables were used to test different

175   relationships (e.g., host genome size as a predictor of burst size). Significance of fixed

176   effects was tested using approximate F-tests (R package lmerTest; Kuznetsova et al.

177   2016), and variation explained by fixed effects was quantified as marginal R^2_{GLMM} (R

178   package MuMIn; Bartón 2016). Preliminary results showed that patterns did not vary

179   between marine and freshwater strains, or between estimates of burst size using

180   infectious units vs. direct counts; therefore these factors were excluded from the analyses

181   for simplicity.

182

183   <u>Model of evolution of latent period / burst size</u>. As shown below, there is evidence that

184   viral traits are affected by the genome size and growth rate of the host. To ask how viral

185   strategies are expected to evolve across phytoplankton hosts that vary in growth rate and

186   genome size, we analyzed a model of viral trait evolution. The model is adapted from

187   previous work on bacteriophage (Levin et al. 1977, Bonachela and Levin 2014), and the

188   contribution of our analysis is to test the effect of parameters representing host growth

189   and genome size. The model can be written as follows:

190

191    $\frac{dP}{dt} = \frac{\mu_{max} P N}{h+N} - \sum_i kPV_i - d_P P + m(P_{in} - P)$ \hfill (1)

192

193    $\frac{dV_i}{dt} = b_i k V_i^{t-L_i} P^{t-L_i} e^{-d_P L_i} - kV_i P - d_V V_i - mV_i$ \hfill (2)

194

195    $\frac{dI}{dt} = \sum_i kPV_i - \sum_i kV_i^{t-L_i} P^{t-L_i} e^{-d_P L_i} - d_P I - mI$ \hfill (3)

8

196

197 $$\frac{dN}{dt} = m(N_{\text{in}} - N) - q\frac{\mu_{\text{max}}PN}{h+N} \qquad (4)$$

198

199   The model explicitly represents latent period as a delay between infection and production

200   of new virions (Bonachela and Levin 2014). In eqn. (1), susceptible phytoplankton (P)

201   grow, limited by nutrient (N), with maximum growth rate μ_{max} and half-saturation

202   constant h. Susceptible hosts are lost due to infection from each viral strain i (V_i), with

203   adsorption rate k. There is loss due to other mortality d_P, which is meant to primarily

204   represent grazing. There is also a slow rate of mixing with adjacent waters (m), which

205   causes susceptible hosts from elsewhere (P_{in}) to enter the system. In eqn. (2), viral strain i

206   has burst size b_i, produced from hosts that were infected at time (t-L_i), where L_i is the

207   latent period. The term $e^{-d_P L_i}$ represents the fraction of infected hosts that have not died

208   by the end of the latent period. Free virions are lost as a result of adsorption to new hosts,

209   decay at rate d_V, or mixing at rate m. In eqn. (3), infected hosts are created through

210   adsorption, are lost to viral lysis, other sources of mortality, or mixing. In eqn. (4),

211   nutrients from elsewhere (N_{in}) enter the system by mixing, and are taken up during

212   phytoplankton growth, with constant cellular quota q. This model does not include

213   recycling of phytoplankton nutrients from lysis or other mortality for computational

214   simplicity; we have checked to ensure this does not affect the trait evolution results.

215   To explore the influence of host characteristics, viral burst size and latent period

216   are related to each other and to host genome size and growth rate:

217

218     $b_i = \text{Min}(r(L_i - E), b_{\text{max}})$ (5)

219     $$r = r_{\max} * \frac{\mu_{\max}N}{h+N}$$ (6)

220     $$E = E_{\min} \frac{h+N}{\mu_{\max}N}$$ (7)

221

222   Where r is the linear rise rate for viral particle production and E is the eclipse period. We

223   assume r is proportional to host growth rate (with maximum $r_{\max}*\mu_{\max}$), and E is

224   inversely proportional to host growth rate (with minimum $E_{\min}/\mu_{\max}$). The functional form

225   of eqns. (5-7) are based on the infection cycle of phage T7 infecting *E. coli* (You et al.

226   2002). To represent limitation of particle production by host genomic resources, burst

227   size is also given an upper limit $b_{\max}$, which is meant to represent the ratio of host

228   genome size : virus genome size. The magnitude of this upper limit is varied to test the

229   effect of host genomic resources on viral life history strategy. The magnitude of $\mu_{\max}$ is

230   varied to test the effect of host growth rate.

231   To simulate evolution under this model, 100 virus genotypes were initialized at

232   equal density, with latent periods ranging from 3.5 to 96 hrs, and corresponding burst

233   sizes calculated from the above equations. All virus genotypes compete for a single host,

234   and the model was run until one virus genotype competitively excluded all others. The

235   delay differential equations were solved with LSODA in the R package deSolve (Soetaert

236   et al. 2010). External inputs to the model are constant (it is a nutrient-limited chemostat /

237   mixed layer model), but the host and viruses oscillate substantially in abundance (Fig.

238   S2). The amplitude of the oscillations is somewhat reduced by the diffusion of

239   susceptible hosts into the system, which aids in computational tractability. Parameter

240   values were chosen to represent a typical phytoplankton-virus system (Table 1).

241

10

242 **Results**

243 Compilation. The literature compilation yielded data on 75 unique virus strains, including

244 51 dsDNA, 1 dsRNA, 7 ssDNA, and 6 ssRNA viruses, and 12 of unknown type

245 (Appendix Table S1). The viruses were isolated from 26 phytoplankton genera, with 38

246 strains isolated from cyanobacteria, 15 from diatoms, 10 from haptophytes, 7 from

247 chlorophytes, 4 from dinoflagellates, and 1 each from a cryptophyte, raphidophyte, and

248 pelagophyte. The majority was isolated from marine systems (58 vs. 19 from fresh

249 waters). The genome size of the isolates ranges from 4.4 to 560 kb, and capsid diameter

250 ranges from 22 to 310 nm. In our analyses, we consider relationships across all viruses,

251 and also relationships within the dsDNA viruses, which are the most numerous and

252 sometimes show distinct patterns compared to single-stranded viruses.

253

254 Empirical trait relationships. We tested a variety of hypothesized correlations between

255 viral traits, and between viral and host traits. In brief, burst size is most strongly related to

256 the ratio $\frac{host\ genome\ size}{viral\ genome\ size}$, which we refer to as the genome size ratio. Latent period is

257 most strongly related to a combination of host growth rate and the genome size ratio.

258 Figure 1A shows that burst size ranges over four orders of magnitude, and that a larger

259 host genome size, relative to viral genome size, is correlated with a greater burst size. The

260 genome size ratio can explain half of the variation in burst size ($R^2 = 0.49$, $F_{1,44} = 33$, $p <$

261 0.001), and the relationship is strongest for dsDNA viruses, with greater variability for

262 the single-stranded viruses that infect diatoms. A similar pattern is found when

263 comparing total viral nucleotide output to host genome size (Fig. S3; $R^2 = 0.54$, $F_{1,5} = 16$,

264 $p = 0.01$); for dsDNA viruses these quantities tend to be directly proportional, which

11

265   means that the number of nucleotides in released virions is similar to the number in the

266   host genome.

267   Latent period also increases with the genome size ratio, though less steeply than

268   burst size (Fig. 1B; $R^2 = 0.30$, $F_{1,14} = 9.3$, p = 0.008). For dsDNA viruses there is a

269   stronger relationship between host growth rate and latent period, with latent period

270   declining about 10-fold for a 10-fold increase in host growth rate, and latent period

271   roughly equal to half of the host doubling time (Fig. 1C; $R^2 = 0.47$, $F_{1,9} = 29$, p < 0.001).

272   This pattern is weaker when single-stranded viruses are included ($R^2 = 0.14$, $F_{1,20} = 6.6$, p

273   = 0.02). In a multivariate model, both genome size ratio and host growth rate are

274   significant predictors of latent period, and they jointly explain 38% of variation in latent

275   period across all viruses, and 57% of variation for dsDNA viruses. In contrast, burst size

276   is unrelated to host growth rate (Fig. S3).

277   Several hypothesized trait correlations are weaker but statistically significant.

278   Burst size and latent period are positively correlated, but only within the dsDNA strains;

279   this corresponds to a 5-fold increase in burst size across a 50-fold increase in latent

280   period (Fig. 1D; $R^2 = 0.05$, $F_{1,22} = 4.6$, p = 0.043). Larger viruses tend to have a smaller

281   burst size when comparing all strains, corresponding to a ~100-fold decrease in burst size

282   over a 100-fold increase in genome size (Fig. 1E; $R^2 = 0.2$, $F_{1,20} = 11.7$, p = 0.003). It is

283   also noteworthy that within the dsDNA viruses there is a trend for larger viruses to infect

284   larger hosts ($R^2 = 0.27$, $F_{1,5.2} = 5.9$, p = 0.058; Fig. 2F), while the smallest viruses, which

285   are single-stranded, have only been isolated from relatively large eukaryotes so far (Fig.

286   1F). There is no relationship between latent period and viral genome size (Fig. S3).

287

12

288 <u>Model of viral trait evolution vs. host genome size</u>. The empirical results suggest that

289 relative genome size and host growth rate are important drivers of viral trait evolution,

290 while burst size and latent period are themselves not strongly correlated across viral

291 strains. We used a model to ask whether these patterns can be explained by mechanistic

292 theory for the evolution of latent period. To ask how host genomic resources might

293 influence viral traits, we used a model in which latent period and burst size can evolve,

294 but burst size is constrained by an upper limit. This limit could have various causes, but

295 here we will imagine that the total number of nucleotides available for viral genome

296 synthesis is approximately equal to the size of the host genome, and therefore the

297 maximum burst size is given by the host:virus genome size ratio (Fig. 1A).

298   The model predicts a nonlinear relationship between latent period / burst size and

299 the genome size ratio. At low size ratios, selection leads to a strategy that maximizes

300 burst size, i.e. latent period increases until the maximal burst size is reached. This means

301 that both latent period and burst size increase steeply as the genome size ratio increases,

302 until a threshold is reached at a size ratio of ~1000 (Fig. 2). Above this size ratio, burst

303 size is not maximized and is unrelated to the size ratio. The value of other model

304 parameters affects the asymptotic strategy. For example, a lower virion decay rate selects

305 for longer latent period and greater burst size (Fig. 2). The mortality rate of the host and

306 the concentration of nutrient input have modest effects; reduced host mortality slightly

307 increases asymptotic latent period and burst size, and increased nutrient input slightly

308 reduces latent period (Fig. S4, S5).

309   These results are consistent with the data showing burst size and latent period

310 tend to increase with the genome size ratio (Fig. 1A-B). They are also consistent with the

13

311   fact that latent period seems to be capped at about 50 hrs, and may be consistent with the

312   fact that burst size for single-stranded viruses, which infect very large hosts, is highly

313   variable.

314

315   <u>Modeled effects of host growth rate.</u> The model predicts that latent period is inversely

316   proportional to host growth rate, while burst size is insensitive to host growth rate (Fig. 3,

317   Fig. S6). In other words, a higher host growth rate allows the optimal burst size to be

318   achieved with a shorter latent period. We also varied r_{max}, which accounts for factors

319   beyond host growth rate that influence the rate of virion production. When burst size is

320   maximized, a higher r_{max} reduces the latent period (Fig. 5); when burst size is not

321   maximized, as observed at high genome size ratios, a higher r_{max} increases the burst size

322   instead (Fig. S6). Overall, these results are consistent with the empirical patterns showing

323   a negative correlation between latent period and host growth rate, but no relationship

324   between burst size and growth rate (Fig. 1C, Fig. S6). The results are also consistent with

325   the fact that there is only a weak correlation between burst size and latent period across

326   viruses (Fig. 1D). This is because variation across host species in growth rate or variation

327   across viruses in r_{max} will decouple burst size and latent period across viral strains, even

328   though there is a mechanistic connection between the traits.

329

330   **Discussion**

331   Comparative analyses inevitably deal with noisy patterns, due to the variety of

332   researchers and methods involved. Nonetheless, we find that traits of the phytoplankton

333   host (genome size and growth rate), in combination with virus genome size, can

334   collectively explain ~40-50% of variation in burst size and latent period across

335   phytoplankton viruses characterized thus far. We also find that a model of viral trait

336   evolution, parameterized with realistic values, produces patterns similar to the empirical

337   results, which lends support to the hypothesized underlying mechanisms. Our

338   interpretation of these results is that phytoplankton cells are a sparse resource in a world

339   that is relatively hazardous. This leads to selection for latent periods that are long enough

340   to exhaust host resources before lysis, for viruses that are not extremely small compared

341   to their host. This results in a correlation between total nucleotide output and host

342   genome size, or between burst size and the host:virus genome size ratio. In addition,

343   latent period is jointly influenced by the genome size ratio and the host growth rate,

344   because the physiology of more rapid host growth allows for more rapid virion

345   production. When the genome size ratio is very large (>~1000), it may be that latent

346   period would have to be several days long in order to exhaust host resources, which

347   increases the likelihood of host mortality during infection. For these cases, other factors

348   may determine the evolution of latent period, such as mortality rates of the host and free

349   virions. In total, our results argue that a trait-based approach to viral ecology is

350   promising, and important aspects of community dynamics and evolution may be

351   predictable from a relatively simple set of underlying principles.

352   The patterns across viral strains can be compared with short-term experiments on

353   a single strain, where host growth is manipulated and virus reproduction is characterized.

354   Experiments with *E. coli* typically show that increased host growth rate decreases latent

355   period and increases burst size (You et al. 2002). In an experiment with *Synechococcus*,

356   an increase in growth due to stirring did not alter latent period (Wilson et al. 1996), while

15

357 experiments with *Micromonas* and *Phaeocystis* found that both nitrogen- and

358 phosphorus-limited growth increased latent period and reduced burst size (Maat and

359 Brussaard 2016). Observed correlations between burst size and host growth rate differ

360 from our empirical and model results (Fig. 3, Fig. S3). In this context, it is important to

361 note that short-term plastic responses of a strain may differ from the evolved strategies

362 that vary across strains and hosts. Indeed, in our model the immediate effect of faster host

363 growth is to increase the rate of viral production, which will lead to a greater burst size

364 for a particular latent period. However, over the long term the optimal genotype is one

365 that reduces the latent period while maintaining the same burst size.

366   Our compilation is relevant to some additional questions not addressed by our

367 model, which focused on the evolution of burst size and latent period. Viruses of

368 phytoplankton vary greatly in size, and the maintenance of this size variation remains to

369 be explained. All else equal, larger virions should take longer to synthesize/assemble, and

370 if host resources are limiting then fewer virions can be produced. Both considerations

371 should reduce burst size for larger viruses, but this is only modestly evident in the data

372 (Fig. 1E). A similar pattern is found when (burst size)/(latent period) is used to

373 approximate the rate of virion production (Fig. S7). There is evidence that larger viral

374 genome size allows for better control of host metabolism, which could increase the rate

375 of viral production and/or the contribution of *de novo* nucleotide synthesis (Lindell et al.

376 2005, Hurwitz et al. 2013). Modeling a general mechanistic relationship between genome

377 size and viral reproduction will require a complex model of host-viral metabolism and its

378 evolution (Bragg and Chisholm 2008, Birch et al. 2012).

379   The patterns in our compilation are most evident for dsDNA viruses, which have

380   been studied in much greater detail than other types. It is possible that single-stranded

381   viruses are under distinct selective pressures as a consequence of their very small size or

382   different type of interaction with host systems of transcription/translation/replication. In

383   addition, at this point single-stranded viruses of phytoplankton have been isolated

384   primarily from *Chaetoceros* spp. (e.g., Nagasaki et al. 2005, Tomaru et al. 2008, Kimura

385   and Tomaru 2015). The authors of those studies further note that burst size is hard to

386   define, because it appears that new viral particles are released prior to lysis. Therefore, a

387   somewhat different mode of infection may be part of the reason the strains show different

388   patterns in our analyses. An additional factor that could add noise or bias to the data is the

389   fact that viral strains maintained in culture could evolve in response to propagation,

390   during which they are periodically exposed to relatively high host density, which could

391   select for a shorter latent period.

392   The results of this study can contribute to models of virus-microbe interactions,

393   community structure, and ecosystem consequences. Recent models have explored how

394   viruses can promote host diversity (Thingstad et al. 2014), how hosts and viruses with

395   overlapping host ranges can coexist (Jover et al. 2013), and how viruses alter the

396   dynamics of standard ecosystem models (Weitz et al. 2015), among other topics. The

397   relationships between virus and host traits described here could be used to constrain

398   model parameters, allowing host and virus community structure to self-organize across

399   environmental gradients (Follows and Dutkiewicz 2011). Cell size is a 'master trait' that

400   influences many aspects of plankton ecology (Finkel et al. 2009, Edwards et al. 2012),

401   and size has been used to structure trait-based models via allometric scaling relationships

17

402   (Dutkiewicz et al. 2011). Genome size is well-correlated with cell size in phytoplankton

403   and other organisms (Veldhuis et al. 1997), and therefore the results of the current study

404   could be used to derive scaling relationships and incorporate viral dynamics into size-

405   structured models of plankton ecosystems. Successful development of virus community

406   models will also require considerable empirical progress on how host range and host

407   resistance coevolve, and how infection network structure is related to both viral traits

408   (burst size, adsorption rate, etc.) and host traits (resistance and its physiological costs).

409

References

411   Abedon, S. T., P. Hyman, and C. Thomas. 2003. Experimental Examination of Bacteriophage Latent-

412   Period Evolution as a Response to Bacterial Availability Experimental Examination of Bacteriophage

413   Latent-Period Evolution as a Response to Bacterial Availability. Applied and Environmental

414   Microbiology 69:7499–7506.

415   Bartón, K. 2016. MuMIn: Multi-Model Inference. R package version 1.15.6.

416   Bates, D., Maechler, M., Bolker, B., and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using

417   lme4. Journal of Statistical Software 67:1-48.

418   Birch, E. W., N. A. Ruggero, and M. W. Covert. 2012. Determining Host Metabolic Limitations on Viral

419   Replication via Integrated Modeling and Experimental Perturbation. PLoS Computational Biology 8.

420   Bonachela, J. A., and S. A. Levin. 2014. Evolutionary comparison between viral lysis rate and latent

421   period. Journal of Theoretical Biology 345:32–42.

422   Bragg, J. G., and S. W. Chisholm. 2008. Modeling the fitness consequences of a cyanophage-encoded

423   photosynthesis gene. PLoS ONE 3:1–9.

424   Breitbart, M. 2012. Marine Viruses: Truth or Dare. Annual Review of Marine Science 4:425–448.

425   Brown, C. M., J. E. Lawrence, and D. A. Campbell. 2006. Are phytoplankton population density maxima
426   predictable through analysis of host and viral genomic DNA content? Journal of the Marine
427   Biological Association of the UK 86:491.

428   Brown, C. M., D. A. Campbell, and J. E. Lawrence. 2007. Resource dynamics during infection of
429   *Micromonas pusilla* by virus MpV-Sp1. Environmental Microbiology 9:2720–2727.

430   Brussaard, C. P. D. 2004. Viral control of phytoplankton populations: a review. The Journal of Eukaryotic
431   Microbiology 51:125–138.

432   Daines, S. J., J. R. Clark, and T. M. Lenton. 2014. Multiple environmental controls on phytoplankton
433   growth strategies determine adaptive responses of the N:P ratio. Ecology Letters 17:414–425.

434   Edwards, K. F., M. K. Thomas, C. A. Klausmeier, and E. Litchman. 2012. Allometric scaling and
435   taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton.
436   Limnology and Oceanography 57:554–566.

437   Finkel, Z. V., J. Beardall, K. J. Flynn, A. Quigg, T. A. V Rees, and J. A. Raven. 2009. Phytoplankton in a
438   changing world: Cell size and elemental stoichiometry. Journal of Plankton Research 32:119–137.

439   Fischer, M. G. 2016. ScienceDirect Giant viruses come of age. Curr Opin Microbiol 31:50–57.

440   Gudelj, I., J. S. Weitz, T. Ferenci, M. Claire Horner-Devine, C. J. Marx, J. R. Meyer, and S. E. Forde.
441   2010. An integrative approach to understanding microbial diversity: from intracellular mechanisms to
442   community structure. Ecology Letters 13:1073–84.

443   Hurwitz, B. L., S. J. Hallam, and M. B. Sullivan. 2013. Metabolic reprogramming by viruses in the sunlit
444   and dark ocean. Genome biology 14:R123.

19

445   Kimura, K., and Y. Tomaru. 2015. Discovery of two novel viruses expands the diversity of single-stranded

446   DNA and single-stranded RNA viruses infecting a cosmopolitan marine diatom. Applied and

447   Environmental Microbiology 81:1120–1131.

448   Koonin E. V., and V. V. Dolja. 2013. A virocentric perspective on the evolution of life. Current Opinion in

449   Virology 3:546–557.

450   Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2016. lmerTest: Tests in linear mixed effects

451   models. R package version 2.0-32.

452   Lang A. S., M. L. Rise, A. I. Culley, and G. Steward. 2009. RNA viruses in the sea. FEMS Microbiol Rev

453   33:295–323.

454   Levin, B., F. Stewart, and L. Chao. 1977. Resource-limited growth, competition, and predation: a model

455   and experimental studies with bacteria and bacteriophage. American Naturalist 111:3–24.

456   Lindell, D., J. D. Jaffe, Z. I. Johnson, G. M. Church, and S. W. Chisholm. 2005. Photosynthesis genes in

457   marine viruses yield proteins during host infection. Nature 438:86–9.

458   Litchman, E., and C. A. Klausmeier. 2008. Trait-Based Community Ecology of Phytoplankton. Annual

459   Review of Ecology, Evolution, and Systematics 39:615–639.

460   Maat, D. S., and C. P. D. Brussaard. 2016. Both phosphorus- and nitrogen limitation constrain viral

461   proliferation in marine phytoplankton. Aquatic Microbial Ecology 77:87–97.

462   Murray, A. G., and G. A. Jackson. 1992. Viral Dynamics A Model of the Effects of Size Shape Motion and

463   Abundance of Single-Celled Planktonic Organisms and Other Particles. Marine ecology progress

464   series 89:103–116.

465   Nagasaki, K., Y. Tomaru, Y. Takao, K. Nishida, Y. Shirai, H. Suzuki, et al. 2005. Previously unknown

466   virus infects marine diatom. Applied and Environmental Microbiology 71:3528–3535.

467   Rosario, K., S. Duffy, and M. Breitbart. 2012. A field guide to eukaryotic circular single-stranded DNA

468   viruses: insights gained from metagenomics. Arch Virol 157:1851–1871.

469   Soetaert, K. T. Petzoldt, and R. W. Setzer. 2010. Solving Differential Equations in R: Package deSolve.

470   Journal of Statistical Software, 33:1-25.

471   Steward, G. F., A. I. Culley, and E. M. Wood-Charlson. 2013. Marine Viruses. Pages 127–144 In: Levin,

472   S.A. (ed.) Encyclopedia of Biodiversity. Second Edition.

473   Suttle, C. A. 2007. Marine viruses--major players in the global ecosystem. Nature Reviews Microbiology

474   5:801–12.

475   Thomas, R., N. Grimsley, M.-L. Escande, L. Subirana, E. Derelle, and H. Moreau. 2011. Acquisition and

476   maintenance of resistance to viruses in eukaryotic phytoplankton populations. Environmental

477   microbiology 13:1412–20.

478   Tomaru, Y., Y. Shirai, H. Suzuki, T. Nagumo, and K. Nagasaki. 2008. Isolation and characterization of a

479   new single-stranded DNA virus infecting the cosmopolitan marine diatom *Chaetoceros debilis*.

480   Aquatic Microbial Ecology 50:103–112.

481   Tummers, B. DataThief III. 2006 <http://datathief.org/>

482   Van Etten, J. L., D. E. Burbank, J. Joshi, and R. H. Meints. 1984. DNA synthesis in a *Chlorella*-like alga

483   following infection with the virus PBCV-1. Virology 134:443–449.

484   Veldhuis, M. J. W., T. L. Cucci, and M. E. Sieracki. 1997. Cellular DNA content of marine phytoplankton

485   using two new fluorochromes: taxanomic and ecological implications. Journal of Phycology 33:527–

486   541.

487   Villarreal L. P., and G. Witzany. 2010. Viruses are essential agents within the roots and stem of the tree of

488   life. Journal of Theoretical Biology 262:698–710.

489 Wang, I.-N., D. E. Dykhuizen, and L. B. Slobodkin. 1996. The evolution of phage lysis timing.

490 Evolutionary Ecology 10:545–558.

491 Weitz, J. S., C. A. Stock, S. W. Wilhelm, L. Bourouiba, M. L. Coleman, A. Buchan, et al. 2015. A

492 multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem

493 processes. The ISME Journal 9:1352–1364.

494 Westoby, M., and I. J. Wright. 2006. Land-plant ecology on the basis of functional traits. Trends in

495 Ecology & Evolution 21:261–268.

496 Wikner, J., J. J. Vallino, G. F. Steward, D. C. Smith, and F. Azam. 1993. Nucleic-Acids from the Host

497 Bacterium as a Major Source of Nucleotides for 3 Marine Bacteriophages. Fems Microbiology

498 Ecology 12:237–248.

499 Wilson, W. H., N. G. Carr, and N. H. Mann. 1996. The effect of phosphate status on the kinetics of

500 cyanophage infection in the oceanic cyanobacterium *Synechococcus* sp. WH7803. Journal of

501 Phycology 32:506–516.

502 You, L., P. F. Suthers, and J. Yin. 2002. Effects of *Escherichia coli* physiology on growth of phage T7 in

503 vivo and in silico. Journal of Bacteriology 184:1497–1500.

504

505

506

507

508

509

510

511

Figures

Figure 1. Empirical patterns of trait variation across phytoplankton viruses. (A) Burst size vs. the genome size ratio ($\frac{host\ genome\ size}{viral\ genome\ size}$). (B) Latent period vs. the genome size ratio. (C) Latent period vs. host growth rate. (D) Burst size vs. latent period. (E) Burst size vs. viral genome size. (F) Virus genome size vs. host genome size. Dashed lines show fitted relationships for all strains, dotted lines show fitted relationships for dsDNA strains. Fitted relationships are from mixed models in which host genus, host taxon, virus type, and publication were included as random effects.
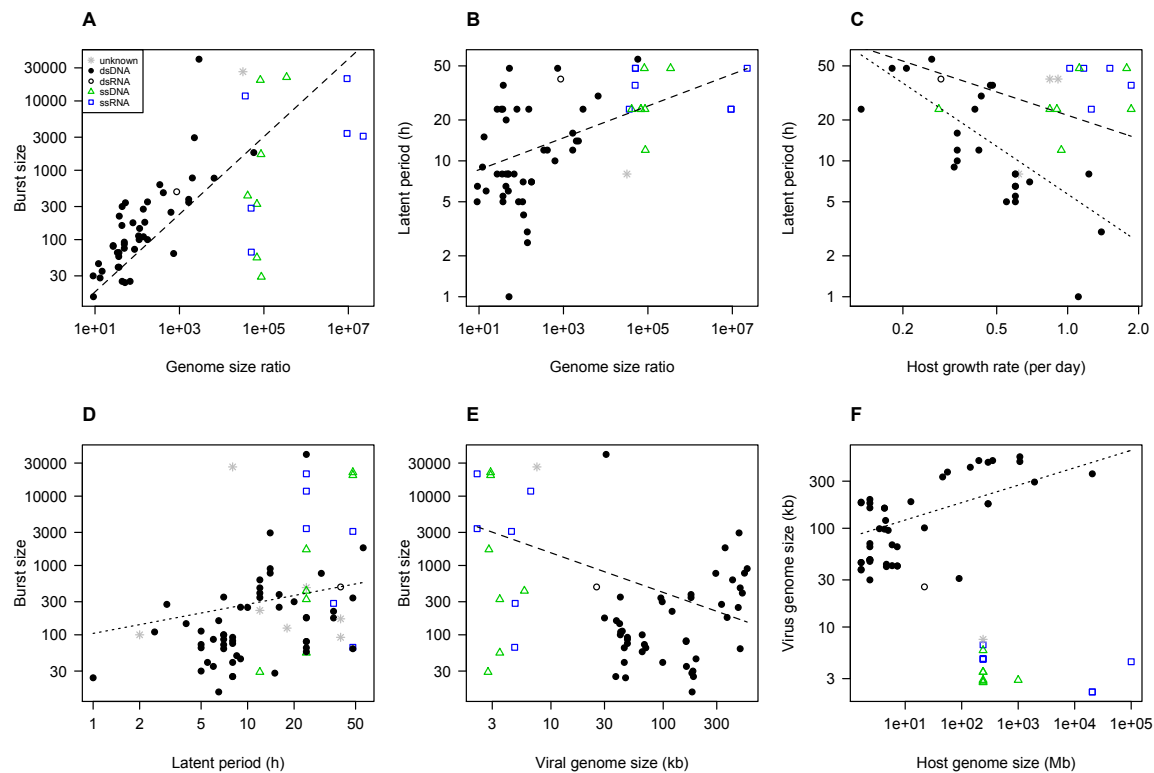
Figure 2. Evolved latent period (A) and burst size (B) as a function of the host:virus genome size ratio. In both panels, the x-axis shows the maximum permitted burst size, which is interpreted here as the ratio of host:virus genome sizes. The y-axis shows the traits of the genotypes that persist in the model. In the model, burst size is calculated instantaneously from the host growth rate (eqns. 5-7); panel (B) shows the burst size for each genotype at maximum host growth, analogous to nutrient-sufficient culture conditions. Results are shown for three virion decay rates (d_v, d^{-1}). $\mu_{max} = 1$ d^{-1}, $r_{max} = 480$ virions d^{-1}.
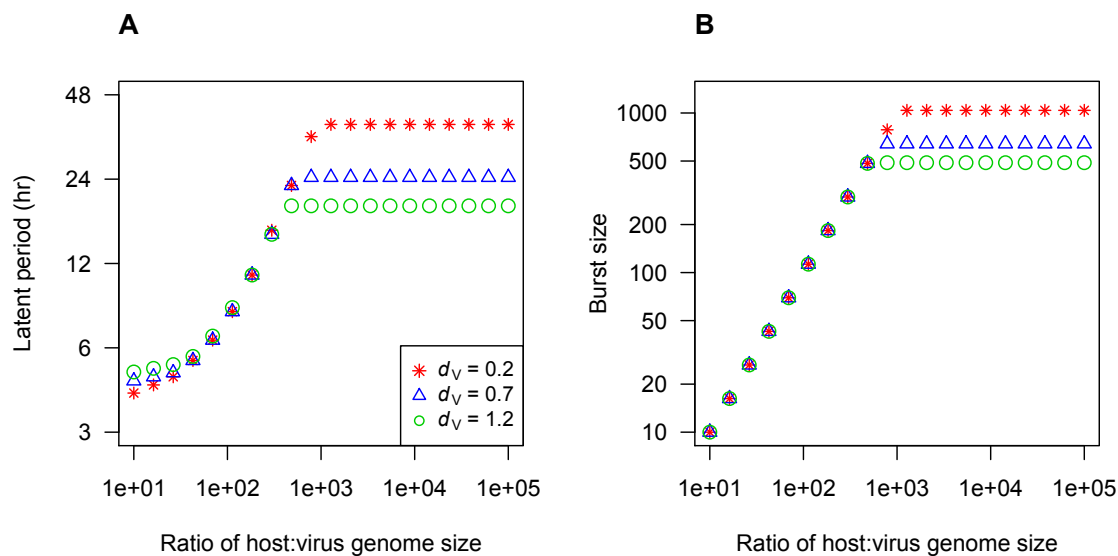
Figure 3. Evolved latent period (A) and burst size (B) as a function of host growth rate. In both panels, the x-axis shows the growth rate of the host, which directly affects the eclipse period and rise rate in the model. The y-axis shows the traits of the genotypes that persist in the model. The maximum burst size (host:virus genome size ratio) is 100 for all cases. Results are shown for three values of r_{max} (Eqn. 6), which represents factors beyond host growth rate that affect the virion production rate. $d_V = 0.7$ d^{-1}.