High parasite diversity maintained after an alga-virus coevolutionary arms race

Eva J.P. Lievens[1†]*, Samuel Kühn[1†], Elena L. Horas[1], Guénolé Le Pennec[2], Sarah Peter[1], Azade D. Petrosky[1], Sven Künzel[3], Philine G.D. Feulner[2,4], Lutz Becks[1]

[†]equal contribution

*corresponding author

[1]Aquatic Ecology and Evolution Group, Limnological Institute, University of Konstanz, Konstanz, Germany

[2]Department of Fish Ecology and Evolution, EAWAG, Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland

[3]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Biology, Plön, Germany

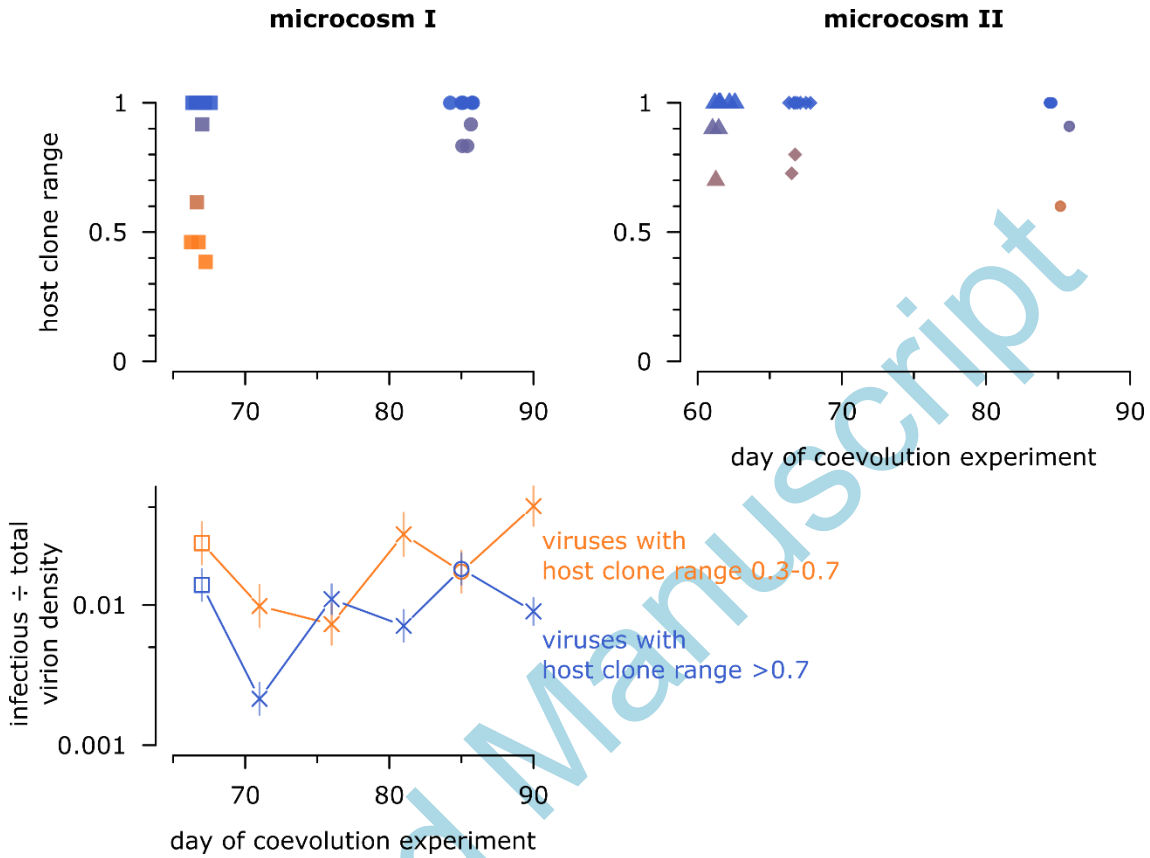
[4]Institute of Ecology and Evolution, University of Bern, Bern, Switzerland

CRediT: Conceptualization: EJPL, LB, ELH. Data curation: SaK, GLP, EJPL, PGDF. Formal analysis: GLP, EJPL, SaK. Funding acquisition: LB, PGDF. Investigation: SaK, ADP, SP, SvK, EJPL. Methodology: EJPL, SaK. Resources: ELH, LB. Supervision: EJPL, LB. Writing - original draft: EJPL. Writing - review & editing: EJPL, GLP, PGDF, LB, ELH, SaK.

Abstract: Arms race dynamics are a common outcome of host-parasite coevolution. While they can theoretically be maintained indefinitely, realistic arms races are expected to be finite. Once an arms race has ended, for example due to the evolution of a generalist resistant host, the system may transition into coevolutionary dynamics that favor long-term diversity. In microbial experiments, host-parasite arms races often transition into a stable coexistence of generalist resistant hosts, (semi-)susceptible hosts, and parasites. While long-term host diversity is implicit in these cases, parasite diversity is usually overlooked. In this study, we examined parasite diversity after the end of an experimental arms race between a unicellular alga (*Chlorella variabilis*) and its lytic virus (PBCV-1). First, we isolated virus genotypes from multiple time points from two replicate microcosms. A time-shift experiment confirmed that the virus isolates had escalating host ranges, i.e. that the arms races had occurred. We then examined the phenotypic and genetic diversity of virus isolates from the post-arms race phase. Post-arms race virus isolates had diverse host ranges, survival probabilities, and growth rates; they also clustered into distinct genetic groups. Importantly, host range diversity was maintained throughout the post-arms race phase, and the frequency of host range phenotypes fluctuated over time. We hypothesize that this dynamic polymorphism was maintained by a combination of fluctuating selection and demographic stochasticity. Together with previous work in prokaryotic systems, our results link experimental observations of arms races to natural observations of long-term host and parasite diversity.

Keywords: *Chlorella variabilis*, PBCV-1, host-parasite coevolution, arms race dynamics, fluctuating selection, diversity, time shift experiment, host range, resistance

Graphical abstract



Introduction

Host-parasite coevolution is regarded as an important driver of diversity, because it can generate and maintain polymorphisms in host resistance and parasite host range (Clarke 1979, Buckingham and Ashby 2022). Reciprocal selection between hosts and parasites can take various forms, including selection for static mono- or polymorphisms, fluctuating selection, and directional selection (reviewed in Agrawal and Lively 2002, Woolhouse et al. 2002, Best et al. 2017, Buckingham and Ashby 2022). Fluctuating selection acts on qualitative or quantitative differences in resistance or host range (e.g. matching alleles models, gene-for-gene & continuous range models with costs), producing dynamic polymorphisms that can maintain diversity indefinitely ('coevolutionary cycles', Best et al. 2017). Long-term diversity can also be maintained by selection for static polymorphisms of resistance and host range. In contrast, directional selection for escalating resistance and host range produces infinite arms race dynamics (e.g. gene-for-gene & continuous range models without costs). Infinite arms races preclude diversity, but are not ecologically realistic: if mutational limitations or costs of resistance/host range occur, arms race dynamics are expected to be transient (Lenski 1984, Brown and Tellier 2011, Best et al. 2017, Buckingham and Ashby 2022).

Arms race dynamics are a common outcome of experimental coevolution, particularly in microbial systems under permissive conditions (Chao et al. 1977, Lenski and Levin 1985, Brockhurst et al. 2003, Mizoguchi et al. 2003, Hall et al. 2011, Kashiwagi and Yomo 2011, Marston et al. 2012, Lopez Pascua et al. 2014, Perry et al. 2015, Gómez et al. 2015, Frickel et al. 2016, Larsen et al. 2019, Retel et al. 2019, Common et al. 2019). Microbial arms races are often finite, ending when a generalist resistant host phenotype evolves (a host phenotype that resists infection by parasites from all time points). The evolution of a generalist resistant host can drive the parasite extinct (Common et al. 2019). More commonly, however, the system transitions into a stable coexistence of parasites, generalist resistant hosts, and susceptible or 'semi-'susceptible hosts (host phenotypes that resist infection by parasites from respectively no or some time points)(Chao et al. 1977, Lenski and Levin 1985, Hall et al. 2011, Marston et al. 2012, Lopez Pascua et al. 2014, Frickel et al. 2016, Retel et al. 2019, Horas 2022). In other words, many microbial arms races transition into post-arms race dynamics that maintain host diversity (typically because resistance trades off with growth, Chao et al. 1977, Lenski and Levin 1985, Frickel et al. 2016, Retel et al. 2019). Whether these post-arms race dynamics also maintain parasite diversity is much less well-studied (Chao et al. 1977, Hall et al. 2011, Marston et al. 2012, Lopez Pascua et al. 2014). Since stable coexistence is likely to be more representative of natural host-parasite communities (Brown and Tellier 2011, Flores et al. 2011, Poisot et al. 2013), understanding when and how parasite diversity is maintained will help link laboratory experiments to natural observations.

In this study, we examined the maintenance of parasite diversity after a coevolutionary arms race between a unicellular alga, *Chlorella variabilis*, and its lytic virus, PBCV-1 (Van Etten et al. 2020). Coevolutionary dynamics in this system typically follow an arms race with three rounds of algal resistance evolution and two rounds of viral

host range evolution (Frickel et al. 2016, Retel et al. 2019, Horas 2022). The arms race ends with the evolution of a generalist resistant algal phenotype, but post-arms race algal polymorphism – and thus viral replication – is maintained by a trade-off between algal resistance and growth rate (Frickel et al. 2016, Retel et al. 2019, Horas 2022). As yet, nothing is known about the virus' phenotypic diversity during the post-arms race phase. To address this, we isolated virus genotypes from multiple time points of a previous coevolution experiment, and studied their phenotypic and genetic diversity.

Methods

Background

This study is based on an alga-virus coevolution experiment described in Horas (2022). Experimental microcosms were inoculated with single clones of *C. variabilis* and PBCV-1, and tracked for 90 days. Microcosms were shaken continuously to avoid environmental heterogeneity. Each day, 10 % of the volume was removed and replaced with fresh medium. The removed volume was used to measure the density of algae and viruses, to store a sample of the virus population (the free-living virus particles, stored at 4°C), and to store a sample of the algal clones (stored as colonies on agar plates). The stored samples were later used to perform time shift experiments: algal clones from ~18 time points were tested for their ability to resist infection by virus populations from the same time points. The time shift experiments provided measures of both algal resistance (the proportion of virus populations that an algal clone could resist) and viral host clone range (the proportion of algal clones that a virus population could infect). Three of four replicate microcosms demonstrated transient arms race dynamics.

We focused on two microcosms with similar dynamics (Fig. 1A&B). The arms race dynamics are visible as increases in algal resistance (detected on days 16, 43, and 67 in microcosm I; on days 16, 43, and 62 in microcosm II) and viral host clone range (detected on days 29 and 67 in microcosm I; on days 25 and 53 in microcosm II). The arms race ended once generalist-resistant algae evolved (detected on day 67 in microcosm I; on day 62 in microcosm II). The "post-arms race" phase, i.e. the phase characterized by the presence of generalist resistant algae, continued until the end of the experiment. Generalist resistance was the dominant algal phenotype during this phase, but viral replication continued (Fig. S1). We therefore assume that (semi-)susceptible algae occurred at low frequencies, as found in previous work (Frickel et al. 2016, Retel et al. 2019).

Virus isolation

In order to study viral diversity, we isolated single virus genotypes from the populations collected by Horas (2022). Hereafter, we use "virus isolate" to describe a virus suspension amplified from a single infectious virion. Virus isolates consist of a single genotype, which may or may not be shared with other virus isolates. "Virus" is used as a

general term; “virion” refers to free-living virus particles; “infectious virions” refers to virions that are capable of lysing a host cell and releasing progeny, i.e. capable of catalyzing the death of an algal culture (Van Etten et al. 1983b).

For each microcosm, we isolated viruses from the virus populations sampled from the coevolution experiment on days 34, 43, 53, 62, 67, and 85. To maximize the diversity of the virus isolates, we used different algal clones as isolation environments (the “isolation algae”, Fig. 1A&B). Each virus population could infect between 2 and 12 isolation algae; we collected one virus isolate per infectious combination (dots in Fig. 1A&B). Isolations were done using plaque or dilution methods. The isolation dilution method did not affect the phenotypes or genotypes of the virus isolates (data not shown). Plaque-based isolations were performed following Van Etten et al. (1983a): virus populations were diluted and plated onto a lawn of an isolation alga, and individual plaques were amplified in the same alga. Dilution-based isolations were done in two steps: first, we used ‘most probable number’ assays to measure the concentration of virions in a virus population that were infectious to an isolation alga (Zimmerman 2017); second, we aliquoted the virus population into 96 liquid culture wells of that alga such that each well received 0.05 infectious virions. After incubation, wells were identified as virus-positive (lysed) or virus-negative. Assuming 0.05 infectious virions per well, the probability that a virus-positive well was inoculated with a single infectious virion was 0.975; allowing for errors in the MPN measurement, the probability was >0.9. One virus-positive well was then amplified in the same isolation alga. For both the plaque and dilution methods, amplified virus isolates were filtered through 0.45 µm and stored at 4°C. Their virion concentration was measured by flow cytometry (on protocols.io: Lievens 2022a, Lievens et al. 2023).

Arms race confirmation

The time shift experiment done by Horas was performed at the virus population level (Fig. 1A&B). As in previous studies, it was assumed that broad host clone ranges at the population level reflected the evolution of broad host clone range genotypes, i.e. reflected an arms race (Frickel et al. 2016, Retel et al. 2019). However, it is also possible that virus populations contained multiple genotypes with complementary narrow host clone ranges. To confirm that arms race dynamics had occurred, we performed an isolate-level time shift experiment using the isolation algae and virus isolates (cf. Buckling and Rainey 2002, Hall et al. 2011, Lopez Pascua et al. 2014). For each isolation alga x virus isolate combination, we inoculated four liquid culture wells with $2x10^4$ virions/ml and $2x10^5$ algae/ml, to a final volume of 0.2 ml (multiplicity of particles 0.1). The cultures were incubated for 10 days, and the optical density at 680 nm was measured every second day. We identified wells as virus-positive (lysed) or virus-negative using longitudinal *k*-means clustering of the optical densities (package “kml”, Genolini et al. 2015) in R (R Core Team 2020). Isolation alga x virus isolate combinations were classed as susceptible if at least 3 of 4 replicates were virus-positive (gray in Fig. 1C&D), and as resistant if at least 3 of 4 replicates were virus-negative (green in Fig. 1C&D). Some combinations could not be completed due to the loss of algal clones over time (light gray or light green in Fig. 1C&D).

Diversity in the post-arms race phase

We assessed the phenotypic and genetic diversity of viruses in the post-arms race phase using virus isolates from the beginning and end of the phase. For microcosm I, these were isolates from respectively day 67 (*n* = 11 isolates) and day 85 (*n* = 8 isolates). For microcosm II, we used isolates from respectively days 62 and 67 (*n* = 8 and 9 isolates) and day 85 (*n* = 5 isolates) (Fig. 1C&D).

Phenotypic diversity

We measured the virus isolates' key life history phenotypes: host clone range, survival in the environment, and growth rate in the host. Host clone range was measured for all of the available isolates. Survival and growth rate were measured in a subset of the isolates (marked under "P" in Fig. 1C&D), after re-amplification in the ancestor alga. For re-amplification, ancestor algal cultures (10 ml, $1x10^6$ algae/ml) were inoculated with a modest number of viruses, incubated until lysis was visible (1-2 days), and filtered through 0.2 µm to remove algae and bacteria. Amplified isolates were stored at 4°C, and the concentration of virions was measured by flow cytometry (on protocols.io: Lievens 2022a, Lievens et al. 2023).

Host clone range was measured by the isolate-level time shift experiment described above, and expressed as the proportion of isolation algae that the virus isolate could infect (cf. Fig. 1C&D). A host clone range of 1.0 indicates that a virus could infect all of the isolation algae from its own microcosm. Note that the isolation algae did not include any generalist-resistant phenotypes, so a host clone range of 1.0 does not mean that a virus could infect generalist resistant algae.

To quantify survival, we measured the proportion of infectious virions that survived 7 days of exposure to the environmental conditions used in the coevolution experiment (constant white light at 50 µmol m^{-2} s^{-1}, 20°C), following the "modified survival assay" described in Lievens et al. (protocols.io: Lievens 2022b, Lievens et al. 2023). Briefly, we prepared viral suspensions with a 10,000-fold range of initial virion concentrations. Before and after exposure to the environment, aliquots of these suspensions were distributed across 16 liquid culture wells of the ancestor alga. The cultures were incubated for four days, after which the wells were classed as virus-positive (lysed) or virus-negative based on their optical density at 680 nm and 750 nm using Gaussian mixture modeling (package "mclust", Scrucca et al. 2016) in R (R Core Team 2020). (For experiments with a large number of wells and comparable data, i.e. a single algal clone at a single time point, Gaussian mixture models are more sensitive than the *k*-means clustering used above.) The assay was replicated twice. For each virus isolate, the numbers of virus-positive and -negative wells were used as binomial response variables in generalized non-linear mixed models, with replicate included as a random effect (in R, R Core Team 2020, package "repeated", Swihart and Lindsey 2022). The models estimated the initial concentration of infectious virions and the proportion of infectious virions that survived until day 7.

The virus isolates' growth rate was measured in the ancestor alga. We inoculated liquid culture wells with $2.5x10^4$ virions/ml and $1x10^6$ algae/ml, to a final volume of 0.2 ml (multiplicity of particles 0.025). The infections were incubated for 24 h. The wells were then centrifuged for 15 min at 2000 g to separate algae and viruses, and supernatants were stored at 4°C. The virion concentration of the supernatants was measured by flow cytometry (on protocols.io: Lievens 2022a, Lievens et al. 2023). We calculated the growth rate as *ln*(final virion concentration/initial virion concentration)/24 h. Each isolate was replicated twice.

Afterwards, we tested for pairwise correlations between the three phenotypes using Spearman's rank method. Correlations were tested within each microcosm and corrected for multiple testing using Holm's method (Holm 1979).

Genetic diversity

We investigated the genetic diversity of post-arms race viruses using a subset of the virus isolates (marked under "G" in Fig. 1C&D). These virus isolates were processed together with several viruses isolated from the arms race phase, whose results are not included in this study. To obtain enough viral DNA, we amplified each isolate in the ancestor alga. Ancestor algal cultures (100 ml, $1\text{-}4*10^6$ algae/ml) were inoculated with a modest number of viruses and incubated until lysis was visible (1-3 days). Algae and bacteria were removed from the lysed cultures by centrifuging for 30 min at 3000 g and filtering the supernatants through 0.2 µm. The virions were then collected onto a 0.1 µm filter, and DNA was extracted from the filter using a Qiagen DNeasy PowerWater kit (QIAGEN GmbH, Germany). Libraries were prepared analogously to Illumina Nextera Flex kits (Picelli et al. 2014) and sequenced on an Illumina MiSeq system (paired end 2x150 bp, Illumina, USA). The ancestor virus and virus isolate PBCV-1-H-79.2 were extracted and sequenced twice as a quality control measure. Reads were pre-processed in fastp v. 0.20.0 (Chen et al. 2018). This consisted of trimming Illumina adapter sequences and polyG tails (default settings), and merging forward and reverse reads in case they overlapped (with the settings overlap_len_re- quire to 20, overlap_diff_limit to 5 and overlap_diff_percent_limit to 5). We also trimmed 3' end tails of reads if the mean quality dropped below 15 and removed reads shorter than 70 bases. Reads of all virus isolates, including the ancestor virus, were mapped onto the PBCV-1 reference genome (NCBI GenBank, accession: NC_000852.5) using bwa v.0.7.17 (Li and Durbin 2009). The alignment files were then processed using tools of the GATK v. 4.2.2.0 (Van der Auwera and O'Connor 2020). We first applied FixMateInformation to all bam files, then merged paired-end and single-end bam files per sample with MergeSameFiles, and marked duplicate reads with MarkDuplicateSpark. The resulting bam files had an average depth of coverage of 63x (Table S1). Coverage was homogenous across the genome (standard deviation 23x), with only 100 sites with no coverage (0.03% of the genome) and 25964 sites with an average coverage below 10x (8% of the genome).

We processed the viral sequences following two independent pipelines. First, genetic variation was called separately on all bam files using GATK's HaplotypeCaller, with minimum mapping quality set to 20 and a minimum

base quality score of 10. The parameter -ploidy was set to 1 since isolates were expected to be single genotypes, and the option -ERC was set to output GVCF format. The resulting GVCFs were combined in a single cohort VCF with GATK's CombineGVCFs, and genotypes were called on this cohort with GATK's GenotypeGVCFs with --sample-ploidy equal to 1. The resulting VCF file contained 88 sites. In order to minimize sequencing artifacts and misalignments, we applied a stringent filtering strategy that removed all sites with ambiguous calls among technical replicates (ancestor virus and PBCV-1-H-79.2, see above). We used vcftools v. 0.1.15 (Danecek et al. 2011) to remove indels and sites with more than 2 alleles. The resulting VCF file contained 35 SNPs. SNP genotypes were extracted with a bash script, and loci with more than 20% missing genotypes were removed. This gave a final number of 24 SNPs compared to the PBCV-1 reference genome, of which 19 occurred in the ancestor or post-arms race viruses (Table S1). SNP coding effects were classified by SnpEff v5.1d (Cingolani et al. 2012, Table S2). Second, we looked for copy number variation along the genome. We used germline tools, which are well-suited to haploid data, in GATK-4.2.2.0 (Van der Auwera and O'Connor 2020). The PBCV-1 reference genome was divided into 1 kb intervals using PreprocessIntervals (default settings), and read counts were collected for every interval using CollectReadCounts (–interval-merging-rule set to overlapping only). We applied DetermineGermlineContigPloidy to the read counts to determine the baseline copy number for each virus isolate (contig ploidy priors: 97% for ploidy 1, 1% for ploidies 0, 2, and 3), and GermlineCNVCaller to obtain posterior probabilities and copy number variation calls. PostprocessGermlineCNVCalls was then used to generate genotype intervals and genotyped segments in VCF formats (--autosomal-ref-copy-number set to 1). This pipeline identified 5 copy number variations in the ancestor or post-arms race viruses compared to the PBCV-1 reference genome (Table S1). The patterns were consistent with two distinct deletions having occurred during the experimental coevolution: when compared to the ancestor virus, all post-arms race isolates had a large deletion between 15-41 kb or between 16-41 kb (each associated with the deletion of specific ancestral duplications, Table S1). There were no filtered SNPs in any of the duplicated or deleted regions.

After genotyping, we grouped the virus isolates into genetic clusters based on their differences from the ancestor virus. We coded the genotypes of the virus isolates as identical to (0) or different from (1) the ancestor virus at each of the 19 SNPs that occurred in these viruses, and recorded whether each isolate had (1) or did not have (0) one of the two large deletions. The distance between isolates was calculated using Gower's general dissimilarity coefficient for asymmetric binary variables, and clustering was done using a *k*-medoids approach (Gower 1971, Kaufman and Rousseeuw 1990). These analyses were done in R using the 'cluster' package; the optimal cluster number was found using the 'kneedle' algorithm (R Core Team 2020, Maechler et al. 2022, Tam 2022). Note that we assumed strict genetic linkage between the mutations identified in a single virus isolate, as the probability that it descended from a single infectious virion was >0.9 (see above).

After grouping the virus isolates into genetic clusters, we tested whether cluster had a significant effect on phenotype. We used generalized linear regressions with a beta error distribution (for survival & host clone range

(the binomial distribution performed poorly due to complete separation); package "betareg", Cribari-Neto and Zeileis 2010) or normal error distribution (for growth rate; package "lme4" Bates et al. 2015). Significance was tested using likelihood ratio tests and Tukey comparisons (package "multcomp", Hothorn et al. 2008). All analyses were done in R (R Core Team 2020).

Fluctuating frequencies in the post-arms race phase

In both microcosms, the population-level host clone range fluctuated during the post-arms race phase (gray lines in Fig. 2A). Since the population-level host clone range was determined using a fixed detection threshold (2000 virions, time shift experiment in Horas 2022) and virus isolates differed in their host clone range (points in Fig. 2A), we inferred that viral genotype frequencies fluctuated over time. For example, when the population-level host clone range was 1.0, genotypes with host clone range 1.0 must have been relatively common (cumulatively more than ~1 infectious virion/2000 virions). At time points with lower population-level host clone ranges, genotypes with host clone range 1.0 must have been rare (fewer than ~1 infectious virion/2000 virions).

Fortuitously, we were able to test the inference of fluctuating viral frequencies in microcosm I. In this microcosm, all isolates with host clone range >0.3 (hereafter "intermediate-HCR") could infect algal clone A_I_34_5, but only those with host clone range >0.7 (hereafter "high-HCR") could infect algal clone A_I_43_2 (Fig. 1C). We used 'most probable number' assays to measure the density of these phenotypes in the virus populations sampled from the coevolution experiment on days 67, 71, 76, 81, 85, and 90. Most probable number assays were done following Zimmerman (2017). For each virus population, 0.05 ml of a 10^{-2}, 10^{-3}, …, or 10^{-9}-diluted virus suspension was inoculated into 24 liquid culture wells containing A_I_34_5 or A_I_43_2, to a final volume of 0.2 ml and concentration of $2x10^5$ algae/ml. The cultures were incubated for 6 days, after which wells were visually identified as virus-positive (lysed) or virus-negative. If necessary, incubation was extended for 2-4 more days. The most probable numbers of infectious virions capable of infecting A_I_34_5 (intermediate- and high-HCR infectious virions) and A_I_43_2 (high-HCR infectious virions) were calculated in R using the 'MPN' package (Ferguson and Ihrie 2019, R Core Team 2020). The density of intermediate-HCR infectious virions was calculated by subtraction.

Experimental conditions

All assays described above were carried out in liquid medium in cell culture flasks (algal growth and virus amplification) or 96-well cell culture plates (all other assays). We used Bold's Basal Medium (BBM, Nichols and Bold 1965), with ammonium chloride substituted for sodium nitrate and double the concentration of trace element solution 4 (first used by Frickel et al. 2016). Flasks or culture plates were maintained at 20°C, under constant light (40-45 µmol s-1 m-2), and on an orbital shaker with diameter 10mm and frequency 120 rpm. Some culture plates were maintained without shaking in the final experiment (see above); this had no effect on the

results. Algal cultures were generally in late exponential to early stationary phase (1-4x10^6 algae/ml in this medium).

Results

Arms race confirmation

Using virus isolates obtained from the arms race and post-arms race phases, we confirmed that arms race dynamics had occurred. Virus isolates increased in host clone range over time (Fig. 1C&D), and many post-arms race isolates were able to infect the same set of isolation algae as their source populations (Fig. 2A, compare points to gray lines).

Diversity in the post-arms race phase

To assess viral diversity after the evolution of generalist-resistant algae, we focused on the viruses isolated from the beginning and end of the post-arms race phases (days 67 and 85 for microcosm I; days 62, 67, and 85 for microcosm II). We collected additional phenotypic and genotypic information for the majority of these isolates (marked under respectively "P" and "G", Fig. 1C&D).

In general, the virus isolates were phenotypically diverse throughout the post-arms race phase (Fig. 2). Isolates with host clone range 1.0 and host clone range <1.0 occurred at all time points, though in microcosm I the diversity was restricted on day 85 compared to day 67. Survival was diverse in microcosm I, but less so in microcosm II (particularly at time points 62 and 85). Although the growth rate in the ancestor alga was quantitatively diverse, qualitatively almost all isolates performed poorly (growth rate ~0). None of the phenotypes were significantly correlated ($p \geq 0.41$ after correction for multiple testing; Fig. S2).

As expected from their phenotypic diversity under standard conditions, the virus isolates were genetically diverse. The post-arms race isolates contained 19 loci with SNPs when compared to the ancestor virus (Table S1). They also contained deletions between the positions 15-41 kb or 16-41 kb (Table S1), a region which is regularly deleted in PBCV-1 (Landstein et al. 1995, Speciale et al. 2019). One isolate had low-quality sequences and was excluded from the analysis (marked under "G", Fig. 1C). The post-arms race isolates belonged to six genetic clusters, which were separated by mutations at loci 65130 (gene A122/A123R), 65158 (gene A122/A123R), 254136 (gene A533R), and 259116 (gene A540L), and by the two deletions (Fig. 3). Clusters 1-3 only occurred in microcosm I, with clusters 1 and 2 maintained throughout the post-arms race phase (Fig. 2D). Clusters 4-6 were typical of microcosm II. All three clusters were detected at the beginning of the post-arms race phase, but only cluster 5 was detected at the end (Fig. 2D).

The genetic diversity in the post-arms race phase isolates was associated with phenotypic variation in host clone range and survival (Fig. 4). Clusters differed in their host clone range ($\chi2(5) = 3.2$, $p < 0.001$), an effect driven by the significantly lower host clone range of clusters 2, 3, and 6 ($|z| \geq 4.1$, $p < 0.001$). Clusters also differed in survival ($\chi2(5) = 2.6$, $p < 0.001$). The overall pattern was of lower survival in microcosm II: cluster 4 had significantly lower survival than clusters 1 and 3 ($|z| \geq 3.1$, $p \leq 0.02$), and cluster 5 had significantly lower survival than clusters 1, 2, and 3 ($|z| \geq 3.0$, $p \leq 0.03$). There was no effect of genetic cluster on the growth rate in the ancestor alga ($\chi2(5) = -0.01$, $p = 0.20$).

Fluctuating frequencies in the post-arms race phase

Based on the fluctuating population-level host clone ranges (gray lines, Fig. 2A) and diverse isolate-level host clone ranges (points, Fig. 2A), we inferred that viral genotype frequencies fluctuated during the post-arms race phase (detailed explanation in Methods). We tested this inference in microcosm I by tracking the density of intermediate-HCR and high-HCR infectious viruses over the course of the post-arms race phase (host clone ranges 0.3-0.7 and >0.7, respectively). We found that the densities of infectious intermediate-HCR and high-HCR virions fluctuated dramatically over time (Fig. 5). The two phenotypes mostly replicated at different times: the high-HCR phenotype increased from day 71 to 76, the intermediate-HCR phenotype replicated from days 76 to 81 and 85 to 90, and both phenotypes increased from days 81 to 85. This is direct evidence of a dynamic polymorphism.

Discussion

We investigated the phenotypic and genetic diversity of the virus PBCV-1, after the end of its coevolutionary arms race with the unicellular alga *C. variabilis* - that is, after the evolution of generalist resistant algae. Our results confirmed that the coevolution followed a transient arms race dynamic, and found that parasite diversity was maintained throughout the post-arms race phase. The isolates had variable genotypes and phenotypes, and were particularly variable in their host clone range (Fig. 2). Moreover, we showed that host clone range followed a dynamic polymorphism in microcosm I (Fig. 5). This work deepens our understanding of coevolution in the alga-virus system and in microbial communities in general.

Post-arms race polymorphisms

Our findings align with the theoretical expectation that host-parasite arms races can end in long-term host range polymorphisms (Brown and Tellier 2011, Best et al. 2017, Buckingham and Ashby 2022). Experimental studies investigating this prediction are relatively scarce, as many coevolution experiments focus on the arms race phase or track only the parasite population. The available literature is dominated by bacteria-bacteriophage coevolution

experiments. Not all bacteria-bacteriophage experiments with arms race dynamics led to the evolution of generalist resistant hosts (Buckling and Rainey 2002, Brockhurst et al. 2003) or maintained phage populations after the end of the arms race (Common et al. 2019). Of those that did, several found that host range diversity was maintained after the evolution of a generalist resistant bacterium (Chao et al. 1977, Hall et al. 2011, Marston et al. 2012, Lopez Pascua et al. 2014), or that generalist resistant bacteria coexisted with diverse phage phenotypes at the experimental end points (Poullain et al. 2008, Wandro et al. 2019, Shaer Tamar and Kishony 2022 (with spatial structuring)). Combined with our results, it seems that the maintenance of post-arms race parasite diversity may be the norm in pro- and eukaryotic microbial systems. Since the evolution of generalist resistance and broad host range can be expected to be costly, this supports the overarching finding that coevolution is more likely to involve fluctuating selection dynamics as ecological "complications" increase (trade-offs, nutrient limitation, spatial heterogeneity; Chao et al. 1977, Gómez and Buckling 2011, Lopez Pascua et al. 2014, Gómez et al. 2015).

Although the post-arms race virus isolates were quite diverse (Fig. 2), the true diversity of the virus populations must be even higher. We found that the isolate-based measurements underestimated host clone range diversity in microcosm I: although intermediate-HCR phenotypes were present on day 85 (Fig. 5), these phenotypes were not represented among the virus isolates (Fig. 2A). This is likely to be a sampling artifact: as we only used three isolation algae that could be infected by intermediate-HCR viruses, they were easily missed by chance. In microcosm II, there was "missing" genetic diversity: isolates 6.1, 45.1, 53.1, 79.2, 17.2, and 94.1 had the same unique genotype (cluster 4, Fig. 3), but their host clone ranges varied between ~0.6 and 1.0 (Fig. 1D). It is possible that the mutations conferring these host clone range phenotypes were not detected due to our stringent filtering process (compare Retel et al. 2019, 2022).

Trade-offs associated with expanding host (clone) range are some of the most powerful arguments explaining parasite diversity in general, and the end of arms race dynamics in particular (Chao et al. 1977, Poullain et al. 2008, Brown and Tellier 2011). Although we did not find evidence for trade-offs in our post-arms race virus isolates, this may have been due to methodological limitations. In microcosm I, undersampling the intermediate-HCR isolates (see above) would have reduced our power to detect differences. To that end, it is interesting that the isolate with the highest survival in microcosm I had a narrow host clone range (Fig. 2A). We also measured the isolates' growth in the ancestor alga instead of their contemporary host clones. Although this did reveal that expanding host clone range is associated with reduced fitness in the ancestor alga, this cost would not have been relevant during the post-arms race phase.

Mechanisms maintaining viral diversity in the post-arms race phase

Theory shows that long-term parasite polymorphisms can be maintained by selection for an equilibrium frequency, producing a static polymorphism, or by fluctuating selection, causing coevolutionary cycling (Woolhouse et al. 2002, Best et al. 2017, Buckingham and Ashby 2022). The viral diversity at the beginning and end of the post-arms

race phase (Fig. 2) is compatible with either explanation, but the fluctuating population-level host clone ranges (Fig. 2A) and fluctuating densities of intermediate-HCR and high-HCR viruses (Fig. 5) indicate coevolutionary cycling.

It is clear that (semi-)susceptible algae were present during the post-arms race phase, since the intermediate-HCR and high-HCR viruses were actively growing (cf. Frickel et al. 2016, Retel et al. 2019). The population dynamics followed consumer-resource cycles, indicating that the density of (semi-)susceptible algae varied through time (Fig. S1; cycles cannot be caused by the unconsumable generalist resistant algae). The maintenance of (semi-)susceptible algae, and thus of virus replication, was driven by a trade-off between algal resistance and growth (Horas 2022).

Depending on the diversity of (semi-)susceptible algae, various selection mechanisms could have acted on the virus population. Two scenarios are possible given the data in Fig. 5 (more details in Fig. S3): A) The algal population contained a single (semi-)susceptible phenotype, susceptible to both intermediate- and high-HCR viruses, whose density fluctuated over time. If host clone range was not costly, neither intermediate-HCR nor high-HCR viruses would be selected for, and a static polymorphism would be maintained. If host clone range was costly, intermediate-HCR viruses would outcompete high-HCR viruses. B) The algal population contained two or more (semi-)susceptible phenotypes, whose densities fluctuated over time: a phenotype susceptible to intermediate- and high-HCR viruses, and a phenotype susceptible to high-HCR viruses. In a population without extinction where host clone range was not costly, fluctuations between the two algal phenotypes would be sufficient to maintain the viral polymorphism (Frickel et al. 2016). More realistically, high-HCR viruses would likely outcompete intermediate-HCR viruses, as they would be selected for when the second algal phenotype dominated. If host clone range traded off with growth or survival, high-HCR viruses would be selected for when the second algal phenotype was present, but not at other times. Thus fluctuating selection could maintain a dynamic viral polymorphism (coevolutionary cycling or “range FSD”; Hall et al. 2011, Lopez Pascua et al. 2014, Best et al. 2017).

In a deterministic world, therefore, the cyclical polymorphism in microcosm I (Fig. 5) would be proof of fluctuating selection for (costly) host clone range (scenario B with costs of host clone range). However, viruses are highly susceptible to demographic stochasticity (Gillespie 1998, Irwin et al. 2016). Viruses have highly skewed reproductive distributions (~1500 new virions per successful infection for PBCV-1, Lievens et al. 2023), and rely on chance to encounter host cells. When algal densities are low, host encounters are rare. For example, we estimate that ~10 infectious high-HCR virions would encounter an algal cell within 24 h on day 71 (adsorption constant conservatively estimated at $5x10^{-10}$, E.J.P. Lievens unpublished data, smoothed algal density ~1750 algae/ml, high-HCR density ~9000 infectious virions/ml, 1440 min; Hyman and Abedon 2009 eq. 18.2). Thus stochasticity in the order of host encounter, magnified by the large number of offspring, could obscure any fitness differences between intermediate-HCR and high-HCR viruses. This could lead to dynamic polymorphisms in the absence of fluctuating selection (more details in Fig. S3, cf. Lopez Pascua et al. 2014). Following the calculation above, we

expect that stochasticity played a larger role around days 71-76 (~10-50 infectious virions expected to encounter a host cell within 24 h), which could explain why only intermediate-HCR viruses replicated between days 76-81 (algae susceptible to intermediate-HCR viruses were also susceptible to high-HCR viruses). At time points with larger virus and algal populations, however, we expect fitness differences to be more important than stochasticity (e.g. days 67, 81, 85, 90: ~250-2500 infectious virions of each phenotype expected to encounter a host cell within 24 h).

Overall, we hypothesize that the (co)evolutionary cycling observed in microcosm I was caused by a combination of fluctuating selection for host clone range and demographic stochasticity. Microcosm II and the similar microcosms in Retel, Kowallik et al. (2019) also saw dramatic fluctuations in the size of algal and viral populations during the post-arms race phase (data not shown), so we expect the same mechanisms to have been at play. Theory predicts that stochasticity can reduce or increase polymorphism in host-parasite coevolution, depending on the model details (Kirby and Burdon 1997, Brown and Tellier 2011, Gokhale et al. 2013), so theoretical models that explicitly consider virus biology would be particularly welcome. Such models could test our intuitive expectations, explore the population densities at which stochasticity becomes important, and examine the persistence of stochasticity-driven polymorphisms.

Genetics of host clone range evolution

Roughly half of the SNPs we detected in the post-arms race phase were in genes that mutated in a previous *Chlorella*-virus coevolution experiment (Table S2; Retel et al. 2019, 2022). Gene A122/123R codes for a minor capsid 'fiber' protein (Shao et al. 2022), which may help attach the virion to the host cell (Van Etten et al. 1991). Gene A533R produces the 'penton' protein that surrounds the virion's spike structure, through which viral DNA enters the host cell (Milrot et al. 2017, Shao et al. 2022). Genes A246 and A540L are virion-associated proteins with unknown functions. However, both are members of "gene gangs" (genes in close proximity whose functions are expected to be related, Seitzer et al. 2018), which suggests that their functions are important and relatively conserved. Gene A540L is in the same gene gang as A533R, so its function may also be infection-related (Seitzer et al. 2018, Retel et al. 2022). The deletions between 15-41 or 16-41 kb, which occurred in all post-arms race isolates, also encompass genes associated with *Chlorella*-virus coevolution (Retel et al. 2019, 2022). Of particular interest are gene A025/027/029L (16.4-20.5 kb), which also encodes a minor capsid 'fiber' protein (Shao et al. 2022), and gene A064R, which glycosylates the major capsid protein (Speciale et al. 2019). Overall the mutated genes seem to be associated with the processes of attachment and DNA entry, consistent with the observation that *Chlorella* resistance usually manifests through attachment (Yamada et al. 2006, Quispe et al. 2017).

Interestingly, different mutations dominated in microcosms I and II (Fig. 3). Mutations in gene A122/123R were characteristic of microcosm I, but not detected in microcosm II. Conversely, mutations in A533R and A540L were characteristic of microcosm II (only two isolates in microcosm I had mutations in this gene, at a locus of A540L that

was absent in microcosm II). This finding could be explained by historical contingency - for example, if the SNP at locus 259116 was fixed early on in microcosm II -, but previous work found that PBCV-1 is not mutation-limited in microcosm experiments (Retel et al. 2022). In this case, it should have been possible for the virus populations to explore all allele combinations. The pattern is unlikely to be an artifact caused by repeated sampling of the same genotypes, as the lack of overlap is significant when considering only unique genotypes (n = 7, 8, and 2 genotypes with mutations in respectively A122/123R, A540L, and both; $\chi^2(1) = 7.6$, $p = 0.01$). Instead, potential explanations include: A) The algal resistance phenotypes in microcosms I and II required different counter-adaptations. This is somewhat supported by previous work, which found that algal populations from different microcosms did not share any high-frequency mutations (Frickel et al. 2018, Retel et al. 2019). However, this may not have affected the algal phenotypes: Frickel et al. (2016) also found that viruses from different microcosms were cross-infective, and this was also true of virus isolates from day 34 of microcosms I and II (S. Kühn, data not shown). B) Strong drift counteracted the diversifying effects of mutation supply, e.g. due to the virus' highly differential reproductive success (Lievens et al. 2023). C) Mutations at the two sites have negative epistatic interactions (Bedhomme et al. 2015, Zhao et al. 2019). Epistasis would need to be site-specific, as previous work found that mutations at different loci in genes A540L and A122/123R can co-occur (Retel et al. 2019). Negative epistasis may also occur between loci 65130 and 65158, since no isolates were found with both derived alleles in microcosm I ($\chi^2(1) = 4.3$, $p = 0.04$; Fig. 3).

Conclusion

Although experimental coevolution in microbial host-parasite systems often follows arms race dynamics, these arms races are usually finite. The post-arms race phase is likely to be more representative of natural conditions (Brown and Tellier 2011, Hall et al. 2011, Flores et al. 2011, Poisot et al. 2013), so understanding its development is a bridge to interpreting natural host-parasite systems. We have shown that in experimental alga-virus coevolutions, virus populations maintained considerable diversity after the end of the arms race. Viral host clone range phenotypes fluctuated in frequency, suggesting a role for fluctuating selection in the maintenance of diversity. Our results align with studies of experimental coevolution between bacteria and bacteriophage (Chao et al. 1977, Poullain et al. 2008, Hall et al. 2011, Marston et al. 2012, Lopez Pascua et al. 2014, Wandro et al. 2019, Shaer Tamar and Kishony 2022), and suggest that the maintenance of parasite diversity is common across microbial systems.

Acknowledgements

We thank the team at the Max Planck Institute sequencing center in Plön for carrying out the DNA sequencing, T. Häring for help with the time shift experiment, and M. Spagnuolo for help with the DNA extractions. We acknowledge the Flow Cytometry Centre (FlowKon) at the University of Konstanz for the use of flow cytometers and the expert support in instrument usage and data analysis. Finally, we thank the editor I. Bravo for his helpful comments on the bioinformatic analyses. This work was supported by the German Research Foundation (DFG) to L.B. (grants BE 4135/8-1, BE 4135/5-2), and by the Swiss National Science Foundation (SNSF) to P.G.D.F. (grant 310030E_179637); E.H. was supported by the International Max Planck Research School for Organismal Biology. Genomic data analysis was supported by collaboration with the Genetic Diversity Centre (GDC), ETH Zürich.

References

Agrawal, A., and C. M. Lively. 2002. Infection genetics: gene-for-gene versus matching-alleles models and all points in between. Evolutionary Ecology Research 4:79–90.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67:1–48.

Bedhomme, S., J. Hillung, and S. F. Elena. 2015. Emerging viruses: why they are not jacks of all trades? Current Opinion in Virology 10:1–6.

Best, A., B. Ashby, A. White, R. Bowers, A. Buckling, B. Koskella, and M. Boots. 2017. Host–parasite fluctuating selection in the absence of specificity. Proceedings of the Royal Society B: Biological Sciences 284:20171615.

Brockhurst, M. A., A. D. Morgan, P. B. Rainey, and A. Buckling. 2003. Population mixing accelerates coevolution. Ecology Letters 6:975–979.

Brown, J. K. M., and A. Tellier. 2011. Plant-parasite coevolution: bridging the gap between genetics and ecology. Annual Review of Phytopathology 49:345–367.

Buckingham, L. J., and B. Ashby. 2022. Coevolutionary theory of hosts and parasites. Journal of Evolutionary Biology 35:205–224.

Buckling, A., and P. B. Rainey. 2002. Antagonistic coevolution between a bacterium and a bacteriophage. Proceedings of the Royal Society of London, Series B 269:931–6.

Chao, L., B. R. Levin, and F. M. Stewart. 1977. A Complex Community in a Simple Habitat: An Experimental Study with Bacteria and Phage. Ecology 58:369–378.

Chen, S., Y. Zhou, Y. Chen, and J. Gu. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890.

Clarke, B. C. 1979. The evolution of genetic diversity. Proceedings of the Royal Society of London. Series B, Biological Sciences 205:453–474.

Common, J., D. Morley, E. R. Westra, and S. van Houte. 2019. CRISPR-Cas immunity leads to a coevolutionary arms race between *Streptococcus thermophilus* and lytic phage. Philosophical Transactions of the Royal Society B: Biological Sciences 374:20180098.

Cribari-Neto, F., and A. Zeileis. 2010. Beta regression in R. Journal Of Statistical Software 34:1–24.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Ferguson, M., and J. Ihrie. 2019. MPN: Most Probable Number and Other Microbial Enumeration Techniques.

Flores, C. O., J. R. Meyer, S. Valverde, L. Farr, and J. S. Weitz. 2011. Statistical structure of host–phage interactions. Proceedings of the National Academy of Sciences 108:E288–E297.

Frickel, J., P. G. D. Feulner, E. Karakoc, and L. Becks. 2018. Population size changes and selection drive patterns of parallel evolution in a host–virus system. Nature Communications 9:1706.

Frickel, J., M. Sieber, and L. Becks. 2016. Eco-evolutionary dynamics in a co-evolving host-virus system. Ecology

Letters 19:450–459.

Genolini, C., X. Alacoque, M. Sentenac, and C. Arnaud. 2015. kml and kml3d: R Packages to Cluster Longitudinal Data. Journal of Statistical Software 65:1–34.

Gillespie, J. H. 1998. Population Genetics: A Concise Guide. Second edition. John Hopkins University Press, Baltimore, MD.

Gokhale, C. S., A. Papkou, A. Traulsen, and H. Schulenburg. 2013. Lotka–Volterra dynamics kills the Red Queen: population size fluctuations and associated stochasticity dramatically change host-parasite coevolution. BMC Evolutionary Biology 13:254.

Gómez, P., B. Ashby, and A. Buckling. 2015. Population mixing promotes arms race host–parasite coevolution. Proceedings of the Royal Society B: Biological Sciences 282:20142297.

Gómez, P., and A. Buckling. 2011. Bacteria-phage antagonistic coevolution in soil. Science 332:106–9.

Gower, J. C. 1971. A general coefficient of similarity and some of its properties. Biometrics 27:857–874.

Hall, A. R., P. D. Scanlan, A. D. Morgan, and A. Buckling. 2011. Host-parasite coevolutionary arms races give way to fluctuating selection. Ecology Letters 14:635–642.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6:65–70.

Horas, E. L. 2022. The effects of symbiotic interactions on the ecology and evolution of the alga *Chlorella variabilis*. University of Konstanz, Kontanz, Germany.

Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous Inference in General Parametric Models. Biometrical Journal 50:346–363.

Hyman, P., and S. T. Abedon. 2009. Practical Methods for Determining Phage Growth Parameters. Pages 175–202 *in* M. R. J. Clokie and A. M. Kropinski, editors. Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions, vol. 401. Humana Press.

Irwin, K. K., S. Laurent, S. Matuszewski, S. Vuilleumier, L. Ormond, H. Shim, C. Bank, and J. D. Jensen. 2016. On the importance of skewed offspring distributions and background selection in virus population genetics. Heredity 117:393–399.

Kashiwagi, A., and T. Yomo. 2011. Ongoing Phenotypic and Genomic Changes in Experimental Coevolution of RNA Bacteriophage Qβ and Escherichia coli. PLOS Genetics 7:e1002188.

Kaufman, L., and P. J. Rousseeuw. 1990. Partitioning around medoids (program PAM). Page Finding groups in data: An introduction to cluster analysis. John Wiley & Sons.

Kirby, G. C., and J. J. Burdon. 1997. Effects of Mutation and Random Drift on Leonard's Gene-for-Gene Coevolution Model. Phytopathology 87:488–493.

Landstein, D., D. E. Burbank, J. W. Nietfeldt, and J. L. Van Etten. 1995. Large Deletions in Antigenic Variants of the Chlorella Virus PBCV-1. Virology 214:413–420.

Larsen, M. L., S. W. Wilhelm, and J. T. Lennon. 2019. Nutrient stoichiometry shapes microbial coevolution. Ecology Letters 22:1009–1018.

Lenski, R. E. 1984. Coevolution of bacteria and phage: are there endless cycles of bacterial defenses and phage counterdefenses? Journal of Theoretical Biology 108:319–325.

Lenski, R. E., and B. R. Levin. 1985. Constraints on the Coevolution of Bacteria and Virulent Phage: A Model, Some Experiments, and Predictions for Natural Communities. The American Naturalist 125:585–602.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760.

Lievens, E. J. P. 2022a. Virion quantification by flow cytometry, without fixation or freezing. dx.doi.org/10.17504/protocols.io.6qpvr6q93vmk/v1.

Lievens, E. J. P. 2022b. Modified survival (mS) assay. dx.doi.org/10.17504/protocols.io.81wgb6knylpk/v1.

Lievens, E. J. P., I. V. Agarkova, D. D. Dunigan, J. L. Van Etten, and L. Becks. 2023. Efficient assays to quantify the life history traits of algal viruses. Applied and Environmental Microbiology 89:e01659-23.

Lopez Pascua, L., A. R. Hall, A. Best, A. D. Morgan, M. Boots, and A. Buckling. 2014. Higher resources decrease fluctuating selection during host–parasite coevolution. Ecology Letters 17:1380–1388.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2022. cluster: Cluster Analysis Basics and Extensions.

Marston, M. F., F. J. Pierciey, A. Shepard, G. Gearin, J. Qi, C. Yandava, S. C. Schuster, M. R. Henn, and J. B. H. Martiny. 2012. Rapid diversification of coevolving marine *Synechococcus* and a virus. Proceedings of the National Academy of Sciences 109:4544–4549.

Milrot, E., E. Shimoni, T. Dadosh, K. Rechav, T. Unger, J. L. Van Etten, and A. Minsky. 2017. Structural studies demonstrating a bacteriophage-like replication cycle of the eukaryote-infecting *Paramecium bursaria* chlorella virus-1. PLOS Pathogens 13:e1006562.

Mizoguchi, K., M. Morita, C. R. Fischer, M. Yoichi, Y. Tanji, and H. Unno. 2003. Coevolution of Bacteriophage PP01 and *Escherichia coli* O157:H7 in Continuous Culture. Applied and Environmental Microbiology 69:170–176.

Nichols, H. W., and H. C. Bold. 1965. *Trichosarcina polymorpha* Gen. et Sp. Nov. Journal of Phycology 1:34–38.

Perry, E. B., J. E. Barrick, and B. J. M. Bohannan. 2015. The Molecular and Genetic Basis of Repeatable Coevolution between *Escherichia coli* and Bacteriophage T3 in a Laboratory Microcosm. PLOS ONE 10:e0130639.

Picelli, S., A. K. Björklund, B. Reinius, S. Sagasser, G. Winberg, and R. Sandberg. 2014. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Research 24:2033–2040.

Poisot, T., M. Lounnas, and M. E. Hochberg. 2013. The structure of natural microbial enemy-victim networks. Ecological Processes 2:13.

Poullain, V., S. Gandon, M. A. Brockhurst, A. Buckling, and M. E. Hochberg. 2008. The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. Evolution 62:1–11.

Quispe, C. F., A. Esmael, O. Sonderman, M. McQuinn, I. Agarkova, M. Battah, G. A. Duncan, D. D. Dunigan, P. L. Timothy, C. DeCastro, I. Speciale, F. Ma, and J. L. Van Etten. 2017. Characterization of a New Chlorovirus Type with Permissive and Non-permissive Features on Phylogenetically Related Algal Strains. Virology 500:103–113.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Retel, C., V. Kowallik, L. Becks, and P. G. D. Feulner. 2022. Strong selection and high mutation supply characterize experimental *Chlorovirus* evolution. Virus Evolution 8:veac003.

Retel, C., V. Kowallik, W. Huang, B. Werner, S. Künzel, L. Becks, and P. G. D. Feulner. 2019. The feedback between selection and demography shapes genomic diversity during coevolution. Science Advances 5:eaax0530.

Scrucca, L., M. Fop, T. Murphy Brendan, and A. Raftery E. 2016. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. The R Journal 8:289.

Seitzer, P., A. Jeanniard, F. Ma, J. L. Van Etten, M. T. Facciotti, and D. D. Dunigan. 2018. Gene Gangs of the Chloroviruses: Conserved Clusters of Collinear Monocistronic Genes. VIruses 10:576.

Shaer Tamar, E., and R. Kishony. 2022. Multistep diversification in spatiotemporal bacterial-phage coevolution. Nature Communications 13:7971.

Shao, Q., I. V. Agarkova, E. A. Noel, D. D. Dunigan, Y. Liu, A. Wang, M. Guo, L. Xie, X. Zhao, M. G. Rossmann, J. L. Van Etten, T. Klose, and Q. Fang. 2022. Near-atomic, non-icosahedrally averaged structure of giant virus *Paramecium bursaria* chlorella virus 1. Nature Communications 13:6476.

Speciale, I., G. A. Duncan, L. Unione, I. V. Agarkova, D. Garozzo, J. Jimenez-Barbero, S. Lin, T. L. Lowary, A. Molinaro, E. Noel, M. E. Laugieri, M. G. Tonetti, J. L. Van Etten, and C. De Castro. 2019. The N-glycan structures of the antigenic variants of chlorovirus PBCV-1 major capsid protein help to identify the virus-encoded glycosyltransferases. The Journal of Biological Chemistry 294:5688–5699.

Swihart, B., and J. Lindsey. 2022. repeated: Non-Normal Repeated Measurements Models.

Tam, E. 2022. Kneedle algorithm in R - detecting knees in graphs.

Van der Auwera, G. A., and B. D. O'Connor. 2020. Genomics in the Cloud. O'Reilly Media.

Van Etten, J. L., I. V. Agarkova, and D. D. Dunigan. 2020. Chloroviruses. Viruses 12:20.

Van Etten, J. L., D. E. Burbank, D. Kuczmarski, and R. H. Meints. 1983a. Virus infection of culturable *Chlorella*-like algae and development of a plaque assay. Science 219:994–996.

Van Etten, J. L., D. E. Burbank, Y. Xia, and R. H. Meints. 1983b. Growth cycle of a virus, PBCV-1, that infects *Chlorella*-like algae. Virology 126:117–125.

Van Etten, J. L., L. C. Lane, and R. H. Meints. 1991. Viruses and virus-like particles of eukaryotic algae. Microbiological Reviews 55:586–620.

Wandro, S., A. Oliver, T. Gallagher, C. Weihe, W. England, J. B. H. Martiny, and K. Whiteson. 2019. Predictable Molecular Adaptation of Coevolving *Enterococcus faecium* and Lytic Phage EfV12-phi1. Frontiers in Microbiology 9.

Woolhouse, M. E. J., J. P. Webster, E. Domingo, B. Charlesworth, and B. R. Levin. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. Nature Genetics 32:569–577.

Yamada, T., H. Onimatsu, and J. L. Van Etten. 2006. *Chlorella* viruses. Advances in virus research 66:293–336.

Zhao, L., M. Seth-Pasricha, D. Stemate, A. Crespo-Bellido, J. Gagnon, J. Draghi, and S. Duffy. 2019. Existing Host Range Mutations Constrain Further Emergence of RNA Viruses. Journal of Virology 93:e01385-18.

Zimmerman, A. 2017. MPN (Most Probable Number) assay for infectivity of algal viruses.

Figure captions

Figure 1. Population- and genotype-level arms race dynamics in alga-virus coevolution. A & B. *Infection matrices for the time shift experiment performed by Horas (2022) (reproduced with permission). Algal clones (columns) were tested for their ability to resist virus populations (rows), and classed as resistant (light green) or susceptible (light gray). A subset of the algal clones were used as environments to isolate virus genotypes ("isolation algae"); virus population x isolation alga combinations that were used to produce a virus isolate are marked with a black dot.* ***C & D.*** *Infection matrices for the time shift experiment confirming arms race dynamics. The isolation algae and ancestor alga (columns) were tested for their ability to resist virus isolates and the ancestor virus (rows), and classed as resistant (dark green) or susceptible (dark gray). If the combination could not be tested, the population-level result is shown (resistant: light green; susceptible: light gray). Selected virus isolates from the post-arms race phase were used to obtain additional phenotypic (✓ under "P") or genetic information (✓ under "G" if successful, × if not). Where relevant, the assignment of genetic cluster is appended (e.g. $✓_1$ for genetic cluster 1; see text). Names of the virus isolates are abbreviated from e.g. "PBCV-1-H-13.3" to "13.3".*

Figure 2. Virus diversity during the post-arms race phase. *Points represent virus isolates; x-axis values are jittered. Virus isolates are colored according to their host clone range; symbols group virus isolates from different populations. Note that not all virus isolates are represented in all panels (not all isolates were phenotyped and not all genotyping was successful, see Fig. 1C&D).* ***A-C.*** *Phenotypic diversity, expressed as the host clone range (**A**), proportion of infectious virions surviving after 7 days of exposure to 20°C and constant light (**B**), and growth rate in the ancestor alga (**C**). Error bars represent the SEs of survival and growth rate across two replicates (SEs for survival not shown for the two virus isolates with 0 survival). In **A**, the grey line represents the population-level host clone range when expressed on the same scale (i.e. proportion of the isolation algae + ancestor alga that could be infected).* ***D.*** *Genetic diversity, expressed as the genetic cluster to which a virus isolate was assigned.*

Figure 3. Composition of genetic clusters. *Virus isolates are organized by their genetic cluster in order to visualize the characteristic mutations. Mutations are identified by ORF and locus (SNPs) or by interval (deletions). The genotyped virus isolates were identical to (gray) or different from (black) the ancestor virus genotype. White squares are NA. Names of the virus isolates are abbreviated from e.g. "PBCV-1-H-13.3" to "13.3".*

Figure 4. Genotype-phenotype association. *Points represent virus isolates; x-axis values are jittered. Virus isolates are colored according to their host clone range; symbols group virus isolates from different populations. Phenotypes are expressed as the host clone range (**A**), proportion of infectious virions surviving after 7 days of exposure to 20°C and constant light (**B**), and growth rate in the ancestor alga (**C**). The SEs of survival and growth rate are not represented here, as they were not taken into account in the statistics.*

Figure 5. Density of virus phenotypes over the course of the post-arms race phase in microcosm I. *Intermediate-HCR viruses are shown in orange, high-HCR viruses in blue. Symbols following Fig. 2A. Both panels are on an* ln-*scale.* ***A.*** *Density of the two phenotypes at each time point. Error bars represent the SE of the most probable number.* ***B.*** *Frequencies of the two phenotypes (infectious virion density compared to the total virion density at that time point).*

Figure 1

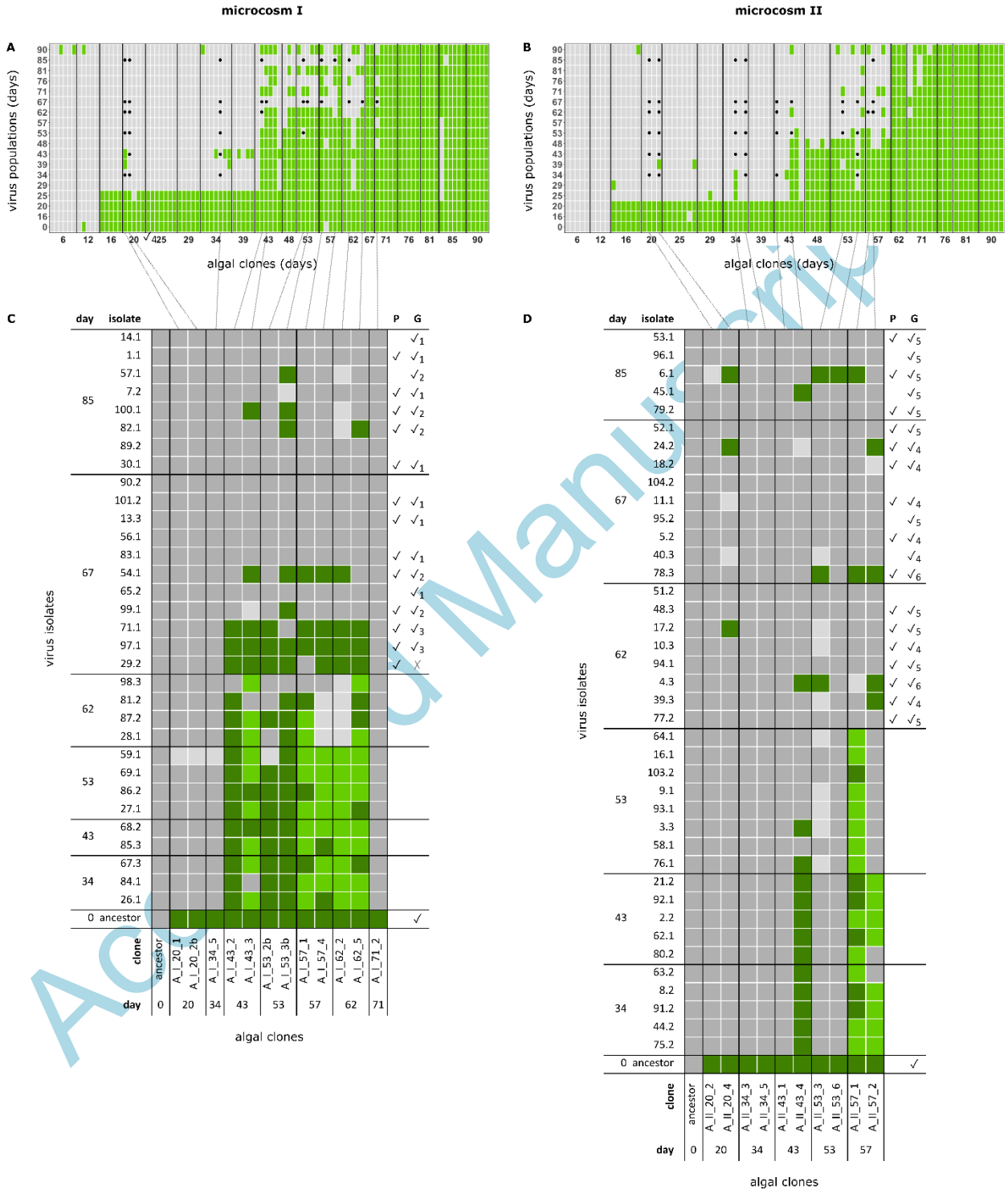


Figure 2

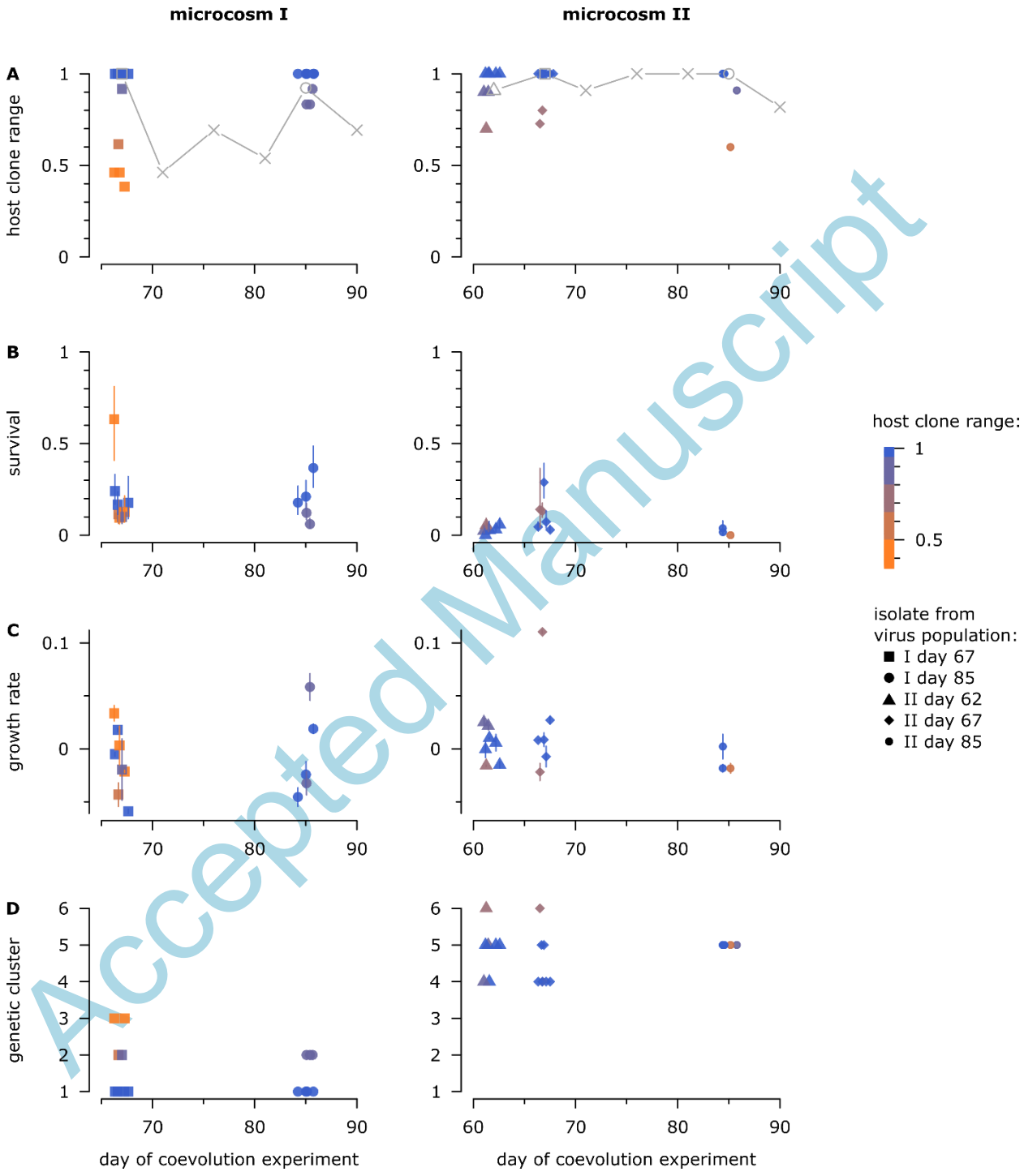


Figure 3

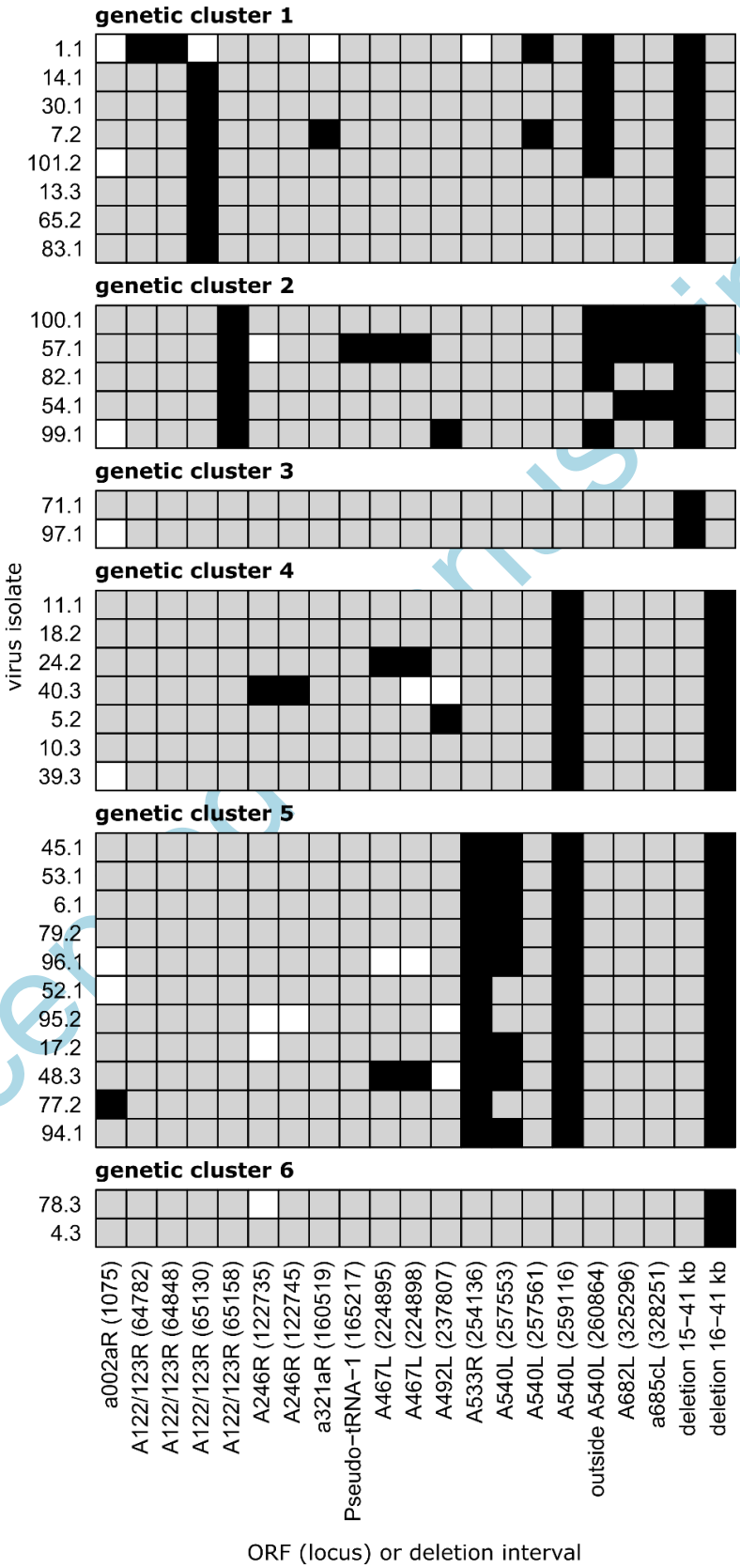


Figure 4

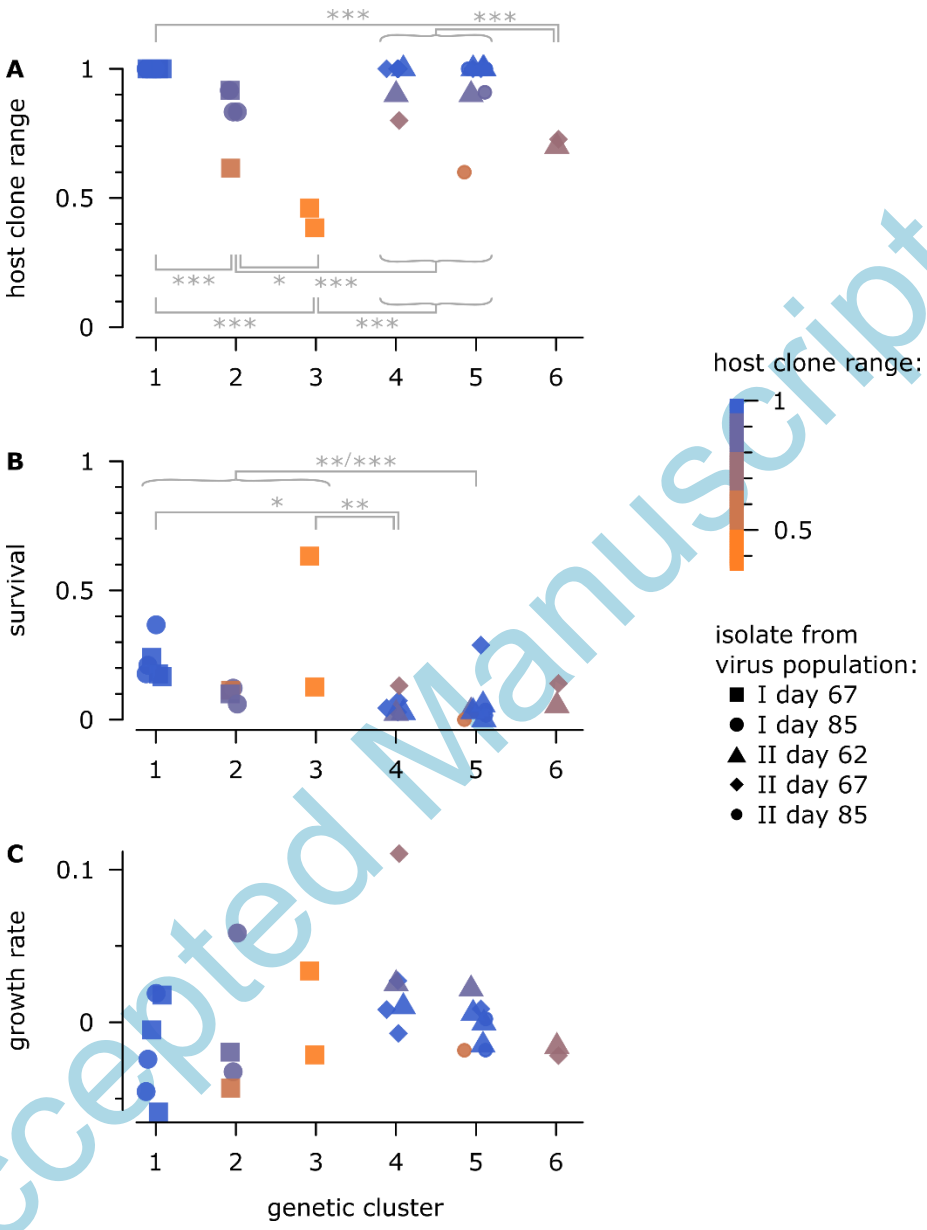


Figure 5