Environmental Microbiology
August 2023, Volume 25, Issue 8, Pages 1424-1438
https://doi.org/10.1111/1462-2920.16366 https://archimer.ifremer.fr/doc/00825/93671/

Environmental vibrio phage- bacteria interaction networks reflect the genetic structure of host populations

Cahier Karine [1, 2], Piel Damien [1, 2], Barcia-cruz Rubén [2, 3], Goudenège David [1, 2], Wegner K. Mathias [4], Monot Marc [5], Romalde Jesús L. [3, 6], Le Roux Frédérique [1, 2, *]

[1] Ifremer, Unité Physiologie Fonctionnelle des Organismes Marins, ZI de la Pointe du Diable, CS 10070, F-29280 Plouzané ,France
[2] Sorbonne Université, CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, F-29688 Roscoff cedex ,France
[3] Department of Microbiology and Parasitology CIBUS-Faculty of Biology, Universidade de Santiago de Compostela Santiago de Compostela ,Spain
[4] AWI - Alfred Wegener Institut - Helmholtz-Zentrum für Polar- und Meeresforschung, Coastal Ecology, Waddensea Station Sylt, Hafenstrasse 43, 25992 List ,Germany
[5] Institut Pasteur, Université Paris Cité, Plate-forme Technologique Biomics Paris ,France
[6] Cross-disciplinary Research Center in Environmental Technologies (CRETUS), Universidade de Santiago de Compostela Santiago de Compostela ,Spain

* Corresponding author : Frédérique Le Roux, email addresses : frederique.le-roux@sb-roscoff.fr ; fleroux@sb-roscoff.fr

Abstract :

Phages depend on their bacterial host to replicate. The habitat, density and genetic diversity of host populations are therefore key factors in phage ecology, but our ability to explore their biology depends on the isolation of a diverse and representative collection of phages from different sources. Here, we compared two populations of marine bacterial hosts and their phages collected during a time series sampling program in an oyster farm. The population of Vibrio crassostreae, a species associated specifically to oysters, was genetically structured into clades of near clonal strains, leading to the isolation of closely related phages forming large modules in phage-bacterial infection networks. For Vibrio chagasii, which blooms in the water column, a lower number of closely related hosts and a higher diversity of isolated phages resulted in small modules in the phage-bacterial infection network. Over time, phage load was correlated with V. chagasii abundance, indicating a role of host blooms in driving phage abundance. Genetic experiments further demonstrated that these phage blooms can generate epigenetic and genetic variability that can counteract host defense systems. These results highlight the importance of considering both the environmental dynamics as well as the genetic structure of the host when interpreting phage-bacteria networks.

INTRODUCTION

Phages, the viral predators of bacteria, constitute the most abundant and diverse biological entities in the ocean and are thought to play a significant role in controlling the composition of microbial communities (Brum and Sullivan 2015, Breitbart, Bonnain et al. 2018). As obligate parasites, phages depend on their bacterial hosts to thrive in the environment and therefore the habitat, density and genetic diversity of host populations are expected to be key determinants of phage ecology. Phage ecology encompasses the breadth of host strains a phage is capable of infecting (host range), as well as adsorption, production and dispersion efficiency in the ecosystem (phage-predator load). The ecology of phages in marine environments remains poorly understood, but such knowledge is important for understanding how they control the expansion of bacterial populations in healthy natural ecosystems and for devising phage applications in disease ridden aquaculture settings (Nobrega, Costa et al. 2015, Gordillo Altamirano and Barr 2019, Kortright, Chan et al. 2019).

The farming of Pacific oysters is a prime example of an aquaculture setting susceptible to disease, with the *Vibrionaceae* (hereafter referred to as "vibrios") being the best described bacteria associated with oyster diseases (Le Roux, Wegner et al. 2015). The ecology, diversity, and population structure of vibrios have been extensively studied (reviewed in (Cordero and Polz 2014, Le Roux, Wegner et al. 2016). Despite their very broad genomic diversity, vibrios fall into well-defined phylogenetic clusters that are specifically associated to organic particles or host organisms, or correspond to free-living forms in the water column (Hunt, David et al. 2008, Szabo, Preheim et al. 2012). These phylogenetic clusters represent distinct populations with respect to gene flow (Shapiro, Friedman et al. 2012, Hehemann, Arevalo et al. 2016) and interactions with other organisms (social/behavioural networks) (Cordero, Ventouras et al. 2012, Cordero, Wildschutte et al. 2012, Yawata, Cordero et al. 2014, Hehemann, Arevalo et al. 2016) and are predominantly congruent with described *Vibrio* species (Preheim, Timberlake et al. 2011), although with finer evolutionary divergence (Shapiro, Friedman et al. 2012). Following this finer distinction, here we define "populations" as corresponding to members of a vibrio species isolated from the same location.

By investigating the disease ecology of vibrios in an oyster farming area, we showed that the onset of disease in oysters is associated with progressive replacement of diverse benign colonizers by members of a virulent population, assigned to *V. crassostreae* (Lemire, Goudenège et al. 2015). Although genetically diverse, most members of this population can

cause the disease. *V. crassostreae* was abundant in diseased animals and nearly absent in the surrounding seawater, suggesting that its primary habitat is the oyster (Bruto, James et al. 2017). The distribution of *V. crassostreae* in the tissues and circulatory system of oysters revealed a positive association with the hemolymph (blood), supporting the conclusion that this population is adapted to overcoming the immune system of the oysters. This was later confirmed by the identification of virulence factors involved in hemocyte (immune cell) cytotoxicity (Rubio, Oyanedel et al. 2019, Piel, Bruto et al. 2020).

An important consequence of the population framework outlined above is that it allows analysis of environmental dynamics of populations, e.g. blooms and busts triggered by specific ecological conditions. We observed that vibrio populations are highly dynamic in the natural ecosystem (Bruto, James et al. 2017), presumably in part due to predation by phages (Piel, Bruto et al. 2022). We isolated a large collection of *V. crassostreae* and vibriophage strains from an oyster farm during a five month time-series sampling program. Cross-infection experiments of roughly 82,000 host-phage pairs revealed a low rate of lytic interaction (2.2%), with interconnected groups of phages and hosts (or "modules"). These modules involved genomic clusters of phages (taxonomically assigned to the genus rank) that adsorb on distinct phylogenetic clades within the *V. crassostreae* species, suggesting the existence of clade-specific receptor(s) and genus-specific receptor binding protein(s).

It is unclear, however, whether the structure of phage-bacteria infection networks in large modules is a specific characteristic of *V. crassostreae* or a general feature of phage-vibrio interactions. A study exploring diverse populations of vibrios and their phages revealed that singleton modules of interaction (one phage infecting only one strain) were prevalent, highlighting a remarkably narrow host range of phages with respect to their local hosts in the environment (Kauffman, Chang et al. 2022). Only a few larger modules were described involving 1) broader host range phages that infect strains from diverse populations, and 2) diverse phages that exclusively infect *V. breoganii,* an algae specialist (Corzett, Elsherbini et al. 2018). This suggested that divergence in bacterial ecology, for example the divergence between host-associated and free-living lifestyles, can be reflected in the topography of the interaction network and host range of the phages (Kauffman, Chang et al. 2022).

To further investigate the connection between ecology, genetic diversity and infection networks, here we compare the structure of phage-bacteria interactions involving *V. crassostreae* and *V. chagasii*. *V. chagasii* was chosen because previous data suggest that its

habitat and lifestyle diverge from those of *V. crassostreae*, although the two populations can co-infect the same animal (Bruto, James et al. 2017). Firstly, the relative abundance of *V. chagasii* in oysters has been reported to resemble that in surrounding seawater. Secondly, this population preferentially occurred in the free-living (1- 0.2 μm) fraction of seawater and oyster digestive glands, where food particles are concentrated in filter feeding oysters (Froelich and Noble 2014). The presence of *V. chagasii* in the digestive gland is thus likely due to passive transfer by filter-feeding and this population might rather bloom in seawater. To directly compare *V. crassostreae* and *V. chagasi*, we analysed the dynamics of *V. chagasii* and their vibriophages in the same time series previously used to describe the *V. crassostreae*-phage infection network (Piel, Bruto et al. 2022). We combined cultivation, comparative genomics and molecular genetics to analyze a large collection of bacterial isolates and their phages. This allowed determination of crucial parameters of phage ecology by addressing how different host ecologies can feed back on genetic diversity and the structure of the resulting host-phage interaction networks.

RESULTS AND DISCUSSION

Phage load correlates with *V. chagasii* frequency

We sampled vibrios from juvenile oysters deployed in an oyster farm (Bay of Brest, France) and from the surrounding seawater on 57 dates over five months, including a period with mortality events (Piel, Bruto et al. 2022). On each sampling date, total vibrios from oysters and seawater were isolated on selective media. Of the 5268 randomly picked colonies, 252 isolates were assigned to *V. chagasii* by sequencing and analysis of the gene encoding the DNA gyrase subunit B. A total of 144 *V. chagasii* isolates were obtained from the 1-0.2 μm (free-living) seawater fraction, while the remaining 108 isolates were obtained from oyster tissues (Table 1). *V. chagasii* was detected later than *V. crassostreae* and persisted after the disease outbreak, suggesting that a role in oyster disease is unlikely (Fig. 1a).

Each of the 252 *V. chagasii* isolates were used as "bait" to detect co-occurring lytic phages from both 1) seawater viral concentrate, and 2) oyster tissues (504 combinations tested). Phage infection was assessed by plaque formation in soft agar overlays of host lawns mixed with the viral source (see Methods). Of the 252 strains, 113 (44.8%) were plaque positive, 59 of which were isolated from seawater and 54 from oysters (Table 1). The vast majority of plaque-forming units (PFUs) were obtained using seawater concentrate as the viral source (2023 PFUs from

seawater vs 35 PFUs from oysters) which could be expected considering that viruses from seawater were concentrated 1000X. The number of phages infecting *V. chagasii* in seawater was estimated to range from 50 L^{-1} (1 PFU observed on the plate using 20mL-seawater equivalents of viral concentrate 1000X) to more than 5.10^3 L^{-1} depending on the strain and date. Phage load was significantly correlated with *V. chagasii* frequency in both isolation fractions (Fig. 1b), consistent with a role of host blooms in driving phage abundance (Kauffman, Chang et al. 2022). Twice as many PFUs were obtained using vibrios isolated from oysters as prey (1347 vs 676 when using vibrio from seawater, Table 1), but this difference was driven by the high proportion of *V. chagasii* isolated from oysters on August 25th (30/108; 30%), with 29 out of every 30 isolates being plaque positive leading to a total of 1088 PFUs. By comparison, the ratio of plaque positive to total number of hosts (18/194, 9.2%) and the number of plaques per host were significantly lower for *V. crassostreae* in both isolation fractions (Table 1). Detection of only 47 PFUs from seawater viral concentrates vs 2023 PFUs for *V. chagasii* confirmed the scarcity of *V. crassostreae* in the seawater. It is noteworthy that during oyster disease at least 50% of oysters die rapidly and massive numbers of *V. crassostreae* will be released into the water column. The low number of PFUs in seawater suggests that vibrios from this population rapidly colonize new animals rather than remaining in the water column.

Structure of phage-bacteria interactions reflects genomic diversity and phylogeny

To investigate the host range of phages isolated using *V. chagasii,* we purified one phage from each plaque positive host to obtain a final set of 95 phage isolates. These phages were tested against 252 *V. chagasii* isolates from the time series and 106 representative members of other vibrio populations previously isolated in the same location (Bruto, James et al. 2017). This cross-infection assay revealed highly specific interactions between phages and bacteria since only 1047 of the 34010 tested interactions (3%) were positive (Fig.S1), with no phages infecting members of other populations.

We next sequenced and assembled the genomes of selected host strains (Table S1). A total of 136 representatives were chosen by excluding the *V. chagasii* isolates with identical *gyrB* sequences and identical patterns of susceptibility in the cross-infection assays. The core genome phylogeny and pairwise ANI values revealed striking differences between hosts (Fig. 2a). In *V. crassostreae* the genetic distance was low within clades and large between clades (Fig. 2b). In *V. chagasii,* genetic distances within the population were more homogenous, being lower overall but larger than the within-clade distance in *V. crassostreae*.

Genomes of *V. chagasii* varied in size between 5.3 to 6.5 Mb (Table S1), which is indicative of extensive flexibility of genome diversity among isolates. A total of 191 genes were exclusively found in all *V. chagasii* isolates compared to only 53 specific genes in *V. crassostreae* isolates (Table S2). Among the *V. chagasii* specific genes we found clusters presumably involved in bacterial attachment (tight adherence and curling) and motility (flagellar). On the other hand, none of the *V. chagasii* strains carried the virulence plasmid pGV1512 that was present in 72% of *V. crassostreae* isolates and encodes a type 6 secretion system cytotoxic for oyster immune cells (Piel, Bruto et al. 2020).

To analyze the diversity of phages isolated from *V. chagasii* we grouped them according to their host, date of isolation and patterns of infectivity, resulting in 49 clusters for which we investigated phage morphology, genome sequence, and taxonomy. Electron microscopy revealed that all phages belong to the *Caudovirales*, including morphotypes of myoviruses (n=15), podoviruses (n=22) and siphoviruses (n=11) (Fig.S2). Genome sequencing of these double-stranded DNA viruses revealed extensive variability in genome length (from 31,648 to 161,442 bp) and gene content (45 to 305 predicted coding sequences) (Table S3). To explore the genomic diversity of these phages we used the Virus Intergenomic Distance Calculator (VIRIDIC, (Moraru, Varsani et al. 2020)) (Fig. 3a) which led us to group the 49 phages infecting *V. chagasii* into 25 VIRIDIC genera (ANI>70%) and 41 species (ANI>95%). By comparison, the higher number of 57 phages isolated using *V. crassostreae* were assigned to only 19 genera and 23 species (Fig. 3b), indicative of a higher taxonomic diversity in our collection of *V. chagasii* phages.

Only one genus (Genus 1, Podoviruses, Table S3) contained phages predicted as temperate phages by BACPHILIP (score 0.975). A gene encoding a putative integrase (tyrosine recombinase) was detected in the genomes of all 11 phages in this genus. However, this prediction was contradicted by the absence of integration of the phages in *V. chagasii* genomes. Tyrosine recombinase constitutes a large protein family involved in a wide variety of biological processes (post-replicative segregation, genetic switches and movement of mobile genetic elements including phages) (Smyshlyaev, Bateman et al. 2021) and functional annotation of these proteins is still limited by their diversity and lack of experimental data. Therefore, the classification of these phages as lysogens will have to be validated experimentally.

We identified a single genus (Genus 10, Table S3) that contains phages infecting both *Vibrio* species (Fig. 3a). This genus splits into two VIRIDIC species (identities >95%), the first

(species 38) containing a single phage isolated from *V. chagasii* and the second (species 15) containing seven phages isolated from *V. crassostreae.* Among genes specific to VIRIDIC species 15, we found a gene encoding a protein with a N-terminal phosphoadenosine phosphosulphate reductase (PAPS) domain and a C-terminal DNA N-6-adenine-methyltransferase (Dam) domain (Fig. S3). We previously demonstrated that this gene (called PAPS-Dam) is involved in phage counter-defense against the DNA phosphorothioation-based defense system Dnd (Piel, Bruto et al. 2022). We used Defense-Finder (Tesson, Hervé et al. 2021) to identify known defense systems in the hosts used to isolate the phages from Genus 10 and found the Dnd_ABCDEFGH system (Wang, Chen et al. 2011) in all *V. crassostreae*, but not in *V. chagasii* hosts. This suggested that the PAPS-Dam phage anti-defense system is involved in the adaptation of phages from species 15 to their *V. crassostreae* host.

Some genera of phages isolated from *V. crassostreae* or *V. chagasii* could be clustered in VIRIDIC families (identities >50%). A core proteome phylogeny of the best represented family indicated that phages differentiated from a common ancestor into genus and species in their respective vibrio host species (Fig. 3a, Fig. S4). Neither *V. crassostreae*-infecting phages nor *V. chagasii*-infecting phages form monophyletic clades, which indicates that phages have jumped several times from one host species to another. To investigate evidence for past recombination among phages, we compared the phylogenetic relationships of each core protein (Fig. S5). Only 3 out of 38 proteins, a large terminase, a DNA binding protein and a putative primase, cluster all *V. crassostreae*-infecting phages in a clade with 56, 100 and 85% bootstrap values, respectively. Incongruences in other protein tree topologies suggest that recombination has shuffled genes between phages. We speculate that vibrio diversity and coinfection of oysters could enhance genomic diversification of the viral population, promoting phage host jumps, as previously described for gut microbiota (De Sordi, Khanna et al. 2017).

We analyzed host-phage interactions taking the vibrio core genome phylogeny and phage clustering into consideration (Fig. 4a). In *V. chagasii,* the size of modules was significantly smaller than in *V. crassostreae* (negative binomial GLM, estimate = 1.234 ± 0.241, z = 5.121, $p < 0.001$). Phages isolated using *V. chagasii* were more diverse than those from *V. crassostreae*, which might at least partially explain the difference in module size. Smaller modules might also indicate higher specialization of phages, i.e. a lower number of susceptible hosts. Comparing pairwise number of shared phages and phylogenetic distance, we found that the likelihood of sharing phages was highest at the smallest genetic distances (Fig. 4b). Both species exhibited a multimodal distribution of sharing probabilities reflecting the modular

structure of the infection network. However, the distribution of these modes differed substantially between the two species. While *V. chagasii* strains showed a comparatively high sharing probability (~20%) at moderate genetic distances of 0.02, the largest genetic distance with compatible phages was observed for *V. crassostreae* strains at genetic distances larger than 0.04. This suggests a higher degree of specialization of *V. chagasii* phages. In fact, only one phage (382E49-1) out of 49 was able to infect strains with larger genetic distances (>0.024). We therefore conclude that smaller modules in *V. chagasii* are a consequence of the lower number of genetically similar hosts, as well as a higher diversity of the phages coupled to a higher degree of specialization.

With such a narrow host range, phages have to find strategies that increase encounter rates with their host. On the one hand, *V. chagasii* inhabits an open system (seawater), which might require a high viral load to promote phage-bacteria encounters. Individual strains of the most abundant *Vibrio* species have previously been estimated to occur at concentrations of on average 10^3 L^{-1} (Kauffman, Chang et al. 2022) and in the present study we showed that the number of phages infecting a *V. chagasii* strain could reach 5 10^3 L^{-1} in seawater. On the other hand, the oyster specialist *V. crassostreae* reaches a higher density in the animal host, which might favor physical contact and promote phage infection. Furthermore, as the *V. crassostreae* population is genetically structured by distinct clades of nearly clonal strains, the same phage can potentially infect more hosts. Exploring to what extent clades bloom in distinct oysters, co-occur in the same animal, or bloom sequentialy, might be necessary to decipher whether the loss of a specific *V. crassostreae* clade is frequently accompanied by the rise of a population of phages which would be a clear indication that the latter moderate the population of the former.

Large numbers of phages during blooms generate epigenetic and genetic variability

Blooms of specific strains can dramatically increase the abundances of specific phages as observed in the time series on the 25th August (Fig. 1a, Fig. S1). The only large module observed in the *V. chagasii* phage-bacteria network connecting hosts and phages was isolated from samples taken on this date (Fig. 4a). On the host side, the phylogenetic relationship based on the core genome revealed the grouping of eight strains in a clade with very little intra-clade diversity (2 to 20 SNPs) (Fig. S6). Seven of these strains were isolated from oysters sampled on the 25th August, and the absence of strain specific genes further confirmed their clonality. The most closely related clade contains a group of four strains with higher diversity (5 to 823 SNPs, 40 genes specific to strain 52_P_461 and 10 genes specific to strain 37_P_203) that were

isolated from seawater collected before and after the 25th August. Similarly, among the nine phages with sequenced genomes isolated from samples taken on the 25th August, five belong to the same species (species 26, genus 18 Table S3) and differed by a unique single nucleotide polymorphism (SNP) (Fig. S7). We compared the virulence of these phages for all hosts from the two previously described sister clades using efficiency of plating (EOP), i.e. the titer of the phage on a given bacteria compared to the titer on the strain used to isolate the phage (original host) (Fig. 5a). This revealed that only phage 409E50-1 showed noticeable variation separating the four strains from the more diverse clade with low susceptibility. These four low susceptibility (LS) strains showed significantly lower susceptibility than the eight remaining clonal high susceptibility strains (HS, LM, estimate = -3.322 ± 0.081, t = -40.609, $p < 0.001$, Fig.5b). Lower infection could not be attributed to difference in cell-surface phage receptors, as phage 409E50-1 adsorbed to all tested isolates regardless of the production of progeny (Fig. S8). We thus concluded that an intracellular defense system found in LS strains controls the efficiency of phage 409E50-1 infection, while other phages from the same species are able to counteract this defense.

Genome comparisons revealed 276 genes that were present in all LS but absent in all HS strains (Table S4). Among these we found two known phage defense systems, the phosphorothioation-sensing bacterial defense system (sspBCDE) targeting viral DNA (Xiong, Wu et al. 2020) and the cyclic-oligonucleotide-based anti-phage signalling systems (CBASS_I) leading to host suicide (Millman, Melamed et al. 2020). However the double deletion of these systems in 36_P_168 did not change the susceptibility of the host (not shown). The defense mechanism(s) of LS acting on 409E50-1 therefore remain(s) to be identified.

We also sought to understand how the other phages of this species escape the LS defense. A phage can escape from host defenses by epigenetic or genetic modifications. Phages 409E50-1 and 521E56-1 were isolated from different strains (respectively 50_O_409 and 56_P_521) and diverge by a single SNP in their genome. To explore the hypothesis of epigenetic modification, we produced these phages in the two host strains (Fig. 5c). When phage 409E50-1 was propagated on vibrio 56_P_521, its infectivity for LS strains was strongly increased compared to the progeny obtained from the original host (LMM, estimate = 2.016 ± 0.296, t = 6.821, $p = 0.002$). The infectivity of phage 521E56-1 was slightly increased when produced on 50_O_409 compared to its original host (Fig. 5c, LMM estimate = 0.656 ± 0.062, t = 10.62, $p < 0.001$). However, the effect of 409E50-1 was three times greater than that of 521E56-1, indicating that the latter was less affected by the host used to produce progenies. This suggests that vibrio

strain 56_P_521 confers an epigenetic advantage to phage 409E50-1 and we hypothesize that phage 521E56-1 does not require host-mediated modification because of genetic divergence.

Phage 409E50-1 differs from 521E56-1 by a single SNP localized in a gene encoding a protein of unknown function (label VP409E501_p0076 in phage 409E50-1; 125 amino acid) (Fig. S7). This SNP results in the change of a threonine (409E50-1) to a proline (521E56-1) at amino acid position 8 (Fig. S9). We exchanged the allele T>G in phage 409E50-1 by recombination (see Methods and Fig. 5d) and subsequently observed a significant decrease in the infectivity of the control phages for LS strains (EOP of 10^{-6} instead of 10^{-4} for the ancestral phage, Fig. 5b and 5c). Irrespective of this difference, control phage infections showed significantly lower EOP on LS strains compared to the original host, 50_0_409 (LM, estimate = -5.843 ± 0.123, t = -47.571, $p < 0.001$). The recombinant phages (T>G) were on average 10^6 fold more infectious than control phages on LS strains and did not differ significantly from the original host, 50_0_409 (Fig. 5e, LM, estimate = -0.160 ± 0.120, t = -1.327, p = 0.196). We therefore show that a single SNP provides protection from the unknown phage defense system present in LS strains. We speculate that this locus could also be the target for epigenetic protection conferred by the host 56_P_521 to phage 409E50-1.

With this mechanistic understanding of the bacteria-phage infection process, we can hypothesize as to how defense elements could spread in the time series samples. The bloom-like proliferation observed on the 25th August most likely resulted from the clonal expansion of a HS strain. The expansion of a *V. chagasii* clone in oysters coincided with the disappearance of *V. crassostreae* in the animals. This could be explained by biotic or abiotic factors that are no longer conducive to the replication of *V. crassostreae* in the oyster, leaving this niche accessible to the expansion of another population (population shift). Finally, the clonal expansion presumably drove the overproportional increase of phages assigned to species 26. It is conceivable that this large phage population generated genetic diversity by mutation, including the previously identified SNP that led to adaptation to the host defense system. This adaptation enabled the predominant phages to reproduce in LS strains after the HS bloom and persist for prolonged periods also outside of the oyster environment.

CONCLUSION

In this study we explored how different genetic structures of hosts can feed back on host-phage interaction networks and genetic diversity of both hosts and phages. The different genetic

structures also reflect differences in host ecology and may therefore also influence phage ecology (host range and population dynamics). Between host species, contrasting patterns of genetic diversity for both host and phage resulted in different structures of the respective phage-bacteria interaction networks. Smaller modules in *V. chagasii* were a consequence of a lower number of clonal hosts. More diverse hosts used as prey will lead to the isolation of more diverse phages. A higher specialization of the phages infecting *V. chagasii* was, however, supported by the highest likelihood to share phages when genetic distances were small. Host blooms result in a high number of phages and we showed that this can generate adaptive genetic variability. Selection can work on this increased variability and may favor variants that can also infect hosts outside the bloom, that were less susceptible to the original phage. The free living lifestyle of *V. chagasii* with blooms inside and outside oyster hosts may thus increase genetic variability of phages, but prevent the formation of modules in this host with low population structure.

MATERIAL AND METHODS

Sampling

Sampling was performed at the same time as for Piel et al. (2022). Briefly, samples were collected from an oyster farm located in the Bay of Brest (Pointe du Château, 48◦ 20′ 06.19″ N, 4◦ 19′ 06.37″ W) every Monday, Wednesday and Friday from the 3rd May to the 11th September 2017. Specific Pathogen Free (SPF) juvenile oysters (Petton, Bruto et al. 2015) were deployed in the field in batches of 100 animals. When the first mortalities were observed in the first batch, another batch of SPF animals was placed in the field, leading to the consecutive deployment of seven batches from the 26th April to the 11th September. Oyster mortalities were recorded on each sampling day. Seawater temperature reached 16°C on the 22nd May, a previously observed threshold for oyster mortalities (de Lorgeril, Lucasson et al. 2018). Mortalities began on the 29th May and persisted until the 25th August. On each sampling date, five living oysters were collected from a batch showing less than 50% mortality. The animals were cleaned, shucked, weighed and 2216 Marine Broth (MB) was added (10 mg/ml) for homogenization using an ultra-turrax. 100 µL of the homogenate was used for vibrio isolation, with the remaining volume centrifuged (10 min,17,000 g) and the supernatant filtered through a 0.2 µm filter and stored at 4°C until the phage isolation stage. Two liters of seawater were collected and size fractionated as previously described (Bruto, James et al. 2017). Bacterial cells from 0.2 µm filters were suspended in 2 mL MB and 100 µL of this suspension was used for vibrio isolation. The iron

chloride flocculation method was used to generate 1000-fold concentrated viral samples from 2 liters passed through a 0.2 μm filter, following the previously described protocol (Kauffman, Brown et al. 2018). Virus-flocculates were suspended in 2mL 0.1M EDTA, 0.2M $MgCl_2$, 0.2M oxalate buffer at pH6 and stored at 4°C until the phage isolation stage.

Vibrio isolation, identification and genome analysis

Isolation and identification. Vibrios from seawater or oyster tissues were selected on thiosulfate-citrate-bile salts-sucrose agar (TCBS). Roughly 48 colonies were randomly picked from each plate and re-isolated once on TCBS, then on 2216 Marine agar (MA). Colonies were screened by PCR targeting the *r5.2* gene encoding for a regulator (Lemire, Goudenege et al. 2014). PCR positive isolates were grown in MB and stored at -80°C in 10% DMSO. Bacteria were grown overnight in 5mL MB and DNA extracted using an extraction kit (Wizard, Promega) according to the manufacturer's instructions. Taxonomic assignment was further refined by *gyrB* gene sequencing to identify *V. crassostreae* and *V. chagasii* isolates. The partial *gyrB* gene was amplified using degenerate primers (Table S5), Sanger sequenced (Macrogen), with sequences manually corrected by visualization of chromatograms. Sequences were aligned with Muscle and phylogenetic reconstruction was performed with RAxML version 8 GTR model of evolution, a gamma model and default parameters (Stamatakis 2006).

Genome sequencing, assembly and annotation. V. chagasii libraries were prepared from 500 ng of genomic DNA using MGIEasy Universal DNA Library Prep Set following the manufacturer's instruction. The 100bp paired-end library pools were circularised and sequenced using DNBSEQ-G400 (BGI) by the Biomics platform at the Pasteur Institute (Paris, France). Reads were trimmed using Trimmomatic v0.39 (LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:36) (Bolger, Lohse et al. 2014). *De novo* read assembly was performed using Spades version 3.15.2 (--careful --cov-cutoff auto -k 21,33,55,77 -m 10). In order to confirm BGI sequencing quality, we compared these assembly results with those obtained for 20 of the same *V. chagasii* strains sequenced by Illumina HiSeq using QUAST v5.0.2 (Gurevich, Saveliev et al. 2013) (Table S6). We confirmed the previous benchmark studies showing similar sequencing and assembly qualities across both platforms (Foox, Tighe et al. 2021). Computational prediction of coding sequences and functional assignments were performed using the automated annotation pipeline implemented in the MicroScope platform (Vallenet, Belda et al. 2013). Phage defense gene annotation was performed using Defense-Finder version 1.0.8 (Tesson, Herve et al. 2022) with MacSyFinder models version 1.1.0 and

default options. Persistent genome phylogeny was undertaken using the PanACoTA workflow (Perrin and Rocha 2021) version 1.3.1-dev2 with mash (version 2.3) distance filtering from 0 to 0.3, prodigal version 2.6.3 for syntaxic annotation, mmseqs version 2.11.e1a1c for gene clustering, mafft version 7.407 for persistent genome alignment defined as 90.0% of all genomes in each family and IQ-TREE (Nguyen, Schmidt et al. 2015)version 2.0.3 based on GTR model and 1000 bootstrap for tree construction.

Comparative genomics. The average nucleotide identity (ANI) value of genomes was determined using fastANI version 1.32 (fragment length 500) (Jain, Rodriguez et al. 2018). *V. crassostreae* and *V. chagasii* core, variable, accessory and specific genomes were established using mmseqs2 reciprocal best hit version 13-45111 with 80% identities on 80% coverage thresholds (Steinegger and Soding 2018). Synteny plots were generated using dedicated python scripts based on the "DNA Features Viewer" library (https://github.com/Edinburgh-Genome-Foundry/DnaFeaturesViewer) and clinker for conservation representation (https://github.com/gamcil/clinker).

Phage isolation, identification and genome analysis

Phage isolation. We used the 194 and 252 isolates of *V. crassostreae* and *V. chagasii* respectively as 'bait' to isolate phages from the same date of sampling from 20mL-seawater equivalents of viral concentrate (1,000X) or oyster tissues (0.1mg). Phage infection was assessed by plaque formation in soft agar overlays of host lawns mixed with both viral sources. For *V. crassostreae,* due to the limited number of co-occuring phages, the collection was completed by isolating viruses from seawater concentrates sampled on several dates (Piel, Bruto et al. 2022). For *V. chagasii*, we purified one phage from each plaque positive host, representing a final set of 95 phage isolates. To generate high titer stocks, plaque plugs were first eluted in 500 µl of MB for 24 hours at 4°C, 0.2-µm filtered to remove bacteria, and re-isolated three times on the sensitive host for purification before storage at 4°C and, after addition of 25% glycerol, at -80°C. High titer stocks (>10^9 PFU/ml) were generated by confluent lysis in agar overlays and phage concentration was determined by dropping spots of ten-fold serial dilutions onto bacterial host lawns.

Electron microscopy. Following concentration on centrifugal filtration devices (Millipore, amicon Ultra centrifugal filter, Ultracel 30K, UFC903024), 20 µl of the phage concentrate were adsorbed for 10 min to a formvar film on a carbon-coated 300 mesh copper

grid (FF-300 Cu formvar square mesh Cu, delta microscopy). The adsorbed samples were negatively contrasted with 2% uranyl acetate (EMS, Hatfield, PA, USA). Imaging was performed using a Jeol JEM-1400 Transmission Electron Microscope equipped with an Orious Gatan camera at the MERIMAGE platform (Station Biologique, Roscoff, France).

DNA extraction. Phage suspensions (10 mL, $>10^{10}$ PFU/mL) were concentrated to approximately 500µL on centrifugal filtration devices (30 kDa Millipore Ultra Centrifugal Filter, Ultracel UFC903024) and washed with 1/100 MB to decrease salt concentration. The concentrated phages were next treated for 30 min at 37°C with 10 µL of DNAse (Promega) and 2.5µL of RNAse (Macherey-Nagel) at 1000 units and 3.5 mg/mL, respectively. Nucleases were inactivated by adding EDTA (20 mM, pH8). DNA extractions involved a first protein lysis step (0.02 M EDTA pH 8.0, 0.5 mg/mL proteinase K, 0.5% SDS) for 30 min incubation at 55°C, followed by phenol chloroform extraction and ethanol precipitation. DNA was visualized by agarose gel electrophoresis (0.7% agarose, 50V, overnight at 4°C) and quantified using QuBit.

Genome sequencing, assembly. Phages were sequenced by the Biomics platform at the Pasteur Institute (Paris, France). The sequencing library (2 × 75 bp paired end) was prepared using a TruSeq Illumina kit and sequenced on a NextSeq550 Illumina sequencer. After read trimming conducted with Trimmomatic v0.39 (LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:36) (Bolger, Lohse et al. 2014), de novo assembly was perfomed using SPAdes v3.15.2 (--careful --cov-cutoff auto -k 21,33,55,77 -m 10). Vector contamination contigs were excluded using the UniVec Database and the one-contig phage genome was manually decircularized. Syntaxic annotation was performed with PHANOTATE v1.5.0 (McNair, Zhou et al. 2019). Large terminase subunit was identified using DIAMOND v2.0.8.146 (blastp, default parameters) (Buchfink, Reuter et al. 2021) on previously annotated vibrio large terminase and using HMMSCAN (HMMER v3.3.2, default parameters with evalue$\leq 10^{-3}$) with PFAM profiles (PF03354.17, PF04466.15, PF05876.14, PF06056.14) (Mistry, Chuguransky et al. 2021). The phage genome was manually ordered starting from the gene for the large terminase subunit.

*Genome annotation.*The tRNA were identified with tRNAscan-SE v2.0.9 (Chan and Lowe 2019). Functional annotation used multiple approaches. First, we used HMMSCAN (evalue$\leq 10^{-3}$) with PVOGS profiles (Grazziotin, Koonin et al. 2017). We also used DIAMOND blastP against first the Nahant collection genomes and then against other bacterial RefSeq (version 212) and phage GenBank genomes (2022-24-06, with 30% identities and 50%

coverage) (Kauffman, Brown et al. 2018, Buchfink, Reuter et al. 2021). Finally, we added InterProScan v5.52-86.0 (evalue≤10^{-3}) and eggNOG v5.0.2 annotation results (Jones, Binns et al. 2014, Huerta-Cepas, Szklarczyk et al. 2019).

Clustering, lifestyle and comparative genomics. We clustered phages using VIRIDIC v1.0r3.6 (default parameters) (Moraru, Varsani et al. 2020). Intergenomic similarities were identified using BLASTN pairwise comparisons. Virus assignment into family (≥50% similarities), genera (≥70% similarities) and species (≥95% similarities) ranks followed the International Committee on Taxonomy of Viruses (ICTV) genome identity thresholds. Phage lifestyle was predicted using BACPHILIP version 0.9.3-alpha and lysogeny associated genes were searched using InterProScan (v.5.52-86.0) based on integrase (IPR002104), resolvase (IPR006119), replicative transposases (IPR004189), and transcriptional repressor domains (IPR010982, IPR010744, IPR001387, IPR032499). Absence of prophage integration was also checked using BLASTn searches between both phage and bacterial genomes (evalue ≤0.01) (Hussain, Dubert et al. 2021). All comparative genomic analyses were conducted using MMseqs2 version 13-45111 for protein clustering (Steinegger and Soding 2018), FAMSA version 1.6.2 for gene/protein alignments (Deorowicz, Debudaj-Grabysz et al. 2016) and IQ-TREE version 2.1.2 for phylogenetic tree construction (Nguyen, Schmidt et al. 2015).

Host range determination

Single-phage-by-single-host host range infection assay. Host range assays were carried out using an electronic multichannel pipette by spotting 5 µL of the phage suspension normalized at $2x10^5$ PFU/ml (10^3 PFU/spot) on the agar overlay inoculated with the tested host. Plates were incubated overnight at RT and plaque formation was observed after 24 hours. Spot assays were performed for *V. chagasii*, first by infecting 358 vibrio strains (252 *V. chagasii* + 106 strains belonging to other vibrio populations) with 95 phages in duplicate. Then, the host range of 49 selected phages was confirmed on 136 *V. chagasii* hosts. Concerning *V. crassostreae*, results were obtained in (Piel, Bruto et al. 2022).

Efficiency of plating (EOP). Ten-fold serial dilutions of phages from a high titer stock (>10^9 PFU/mL) were prepared and 10 µL of each dilution were pipetted onto bacterial host lawns. The EOP was calculated as the ratio between the titer of a phage on a given strain compared to the titer of the same phage on its reference strain (host used to isolate and produce the phage).

Adsorption estimation. Phage adsorption experiments were performed as previously described (Hyman and Abedon 2009). Phages were mixed with exponentially growing cells (OD 0.3; 10^7 CFU/mL) at a MOI of 0.01 and incubated at RT without agitation. At different time points, 250 µL of the culture was transferred into a 1.5 mL tube containing 50 µL of chloroform and centrifuged at 17,000g for 5 min. The supernatant was 10-fold serially diluted and drop spotted onto a fresh lawn of a sensitive host to quantify the remaining free phage particles. In this assay, a drop in the number of infectious particles at 15 or 30 min indicates bacteriophage adsorption.

Statistical analyses

Statistical analyses were based on linear models (LM) when data were normally distributed or on negative bionmial generalized linear models (negbin GLM) when data were left skewed. Correlations were based on non-parametric Spearman's rank correlations. Paired tests were analysed as linear mixed models adding the paired factor as a random intercept. All analyses were performed using the R statistical environment (v.4.0.5).

Molecular microbiology

Strains and plasmids. All plasmids and strains used or constructed in the present study are described in Table S7 and S8. *V. chagasii* isolates were grown in Marine broth (MA) or LB+0.5 M NaCl at RT. *Escherichia coli* strains were grown in LB at 37°C. Chloramphenicol (Cm; 5 or 25µg/mL for *V. chagasii* and *E. coli*, respectively), thymidine (0.3 mM) and diaminopimelate (0.3 mM) were added as supplements when necessary. The P_{BAD} promoter was induced by the addition of 0.2% L-arabinose to the growth media and repressed by the addition of 1% D-glucose. Conjugation between *E. coli* and *V. chagasii* were performed at 30°C as described previously (Le Roux, Binesse et al. 2007).

Cloning. For vibrio knock outs, all cloning in the suicide vector pSW7848T was performed using Herculase II fusion DNA polymerase (Agilent) for PCR amplification and the Gibson Assembly Master Mix (New England Biolabs, NEB) according to the manufacturer's instructions. For recombination of the 409E50-1 phage, a 227 bp region flanking the SNP (T in 409E50-1, G in other phages from species 26) was amplified using the Herculase II and DNA from the 521E56-1 phage and cloned by Gibson Assembly in pMRB (Le Roux, Davis et al. 2011) instead of the *gfp* gene. All cloning was firstly confirmed by digesting plasmid minipreps with specific restriction enzymes and secondly by sequencing the insert (Eurofins).

Vibrio mutagenesis. Knock out was performed by cloning 500bp fragments flanking the pSW7848T region (Le Roux, Binesse et al. 2007). This suicide vector encodes the *ccdB* toxin gene under the control of an arabinose-inducible and glucose-repressible promoter, P_{BAD}. Selection of the plasmid-borne drug marker on Cm and glucose resulted from integration of pSW7848T in the genome. The second recombination leading to pSW7848T elimination was selected on arabinose-containing media.

Phage mutagenesis. Recombinant phage (T>G in p0076 of phage 409E50-1) was engineered using double crossing over with a plasmid carrying a 227 region of homology to the phage genome 521E56-1 (Fig. 5d). This plasmid, or as control the original pMRB-*gfp*, was transferred by conjugation to strain 50_O_409. Plate lysates were generated by mixing 100 µL of an overnight culture of the transconjugant with the 409E50-1 phage and plating in 5 mL agar overlay. After the development of a confluent lysis of lawns, the lysate was harvested by addition of 7 mL of MB and shredding of the agar overlay and stored ON at 4°C for diffusion of phage particles. The lysates were next centrifuged, the supernatant filtered through 0.2 µm filter and stored at 4°C. We could not enrich the recombinant phages by infecting LS strains because these strains confer epigenetic protection to the 409E50-1 phage. We therefore performed several successive passages of the phage on the strain carrying the plasmid, testing the infectivity of the phages on the original strain and the LS strain at each passage. After three passages, a different titer between the putative recombinant and controls was observed. The p0076 gene was PCR amplified using single plaque as template (eight recombinant or eight control phages) and primers flanking the 227 region in the phage genome (external primers). The SNP was confirmed by sequencing. Three control or recombinant phages were tested by EOP on four LS strains in addition to the source 50_O_409.

Data availability

Sequenced genomes have been deposited under the ENA Project with accession numbers PRJEB53320 for *V. chagasii* and PRJEB53960 for phages.

Materials availability

All vibrio strains, phage strains and plasmids are available upon request.

ACKNOWLEDGEMENTS

We warmly thank Sylvain Gandon, François Blanquart and Martin Polz for fruitful discussions on the manuscript. We are grateful to Ian Probert for proofreading the manuscript. We thank Sabine Chenivesse, Sophie Le Panse, the staff of the Ifremer Argenton and Bouin stations and of the ABIMS (Roscoff) and LABGeM (Evry) platforms for technical assistance. We thank all members of the "GV team" for support with field sampling. We thank Zachary Allouche, Biomics Platform, C2RT, Institut Pasteur (Paris, France) supported by France Génomique (ANR-10-INBS-09) and IBISA. This work was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 884988, Advanced ERC Dynamic) and the Agence Nationale de la Recherche (ANR-20-CE35-0014 « RESISTE ») to FLR. R.B.-C. acknowledges the Spanish Ministerio de Ciencia e Innovación for his FPI predoctoral contract (BES-2017-079730).

Competing interests

Authors declare no competing interests.

REFERENCES

Bolger, A. M., M. Lohse and B. Usadel (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics **30**(15): 2114-2120.

Breitbart, M., C. Bonnain, K. Malki and N. A. Sawaya (2018). "Phage puppet masters of the marine microbial realm." Nat Microbiol **3**(7): 754-766.

Brum, J. R. and M. B. Sullivan (2015). "Rising to the challenge: accelerated pace of discovery transforms marine virology." Nat Rev Microbiol **13**(3): 147-159.

Bruto, M., A. James, B. Petton, Y. Labreuche, S. Chenivesse, M. Alunno-Bruscia, M. F. Polz and F. Le Roux (2017). "Vibrio crassostreae, a benign oyster colonizer turned into a pathogen after plasmid acquisition." ISME J **11**(4): 1043-1052.

Buchfink, B., K. Reuter and H. G. Drost (2021). "Sensitive protein alignments at tree-of-life scale using DIAMOND." Nat Methods **18**(4): 366-368.

Chan, P. P. and T. M. Lowe (2019). "tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences." Methods Mol Biol **1962**: 1-14.

Cordero, O. X. and M. F. Polz (2014). "Explaining microbial genomic diversity in light of evolutionary ecology." Nat Rev Microbiol **12**(4): 263-273.

Cordero, O. X., L. A. Ventouras, E. F. Delong and M. F. Polz (2012). "Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations."

Proceedings of the National Academy of Sciences of the United States of America **109**(49): 20059-20064.

Cordero, O. X., H. Wildschutte, B. Kirkup, S. Proehl, L. Ngo, F. Hussain, F. Le Roux, T. Mincer and M. F. Polz (2012). "Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance." Science **337**(6099): 1228-1231.

Corzett, C. H., J. Elsherbini, D. M. Chien, J. H. Hehemann, A. Henschel, S. P. Preheim, X. Yu, E. J. Alm and M. F. Polz (2018). "Evolution of a Vegetarian Vibrio: Metabolic Specialization of Vibrio breoganii to Macroalgal Substrates." J Bacteriol **200**(15).

de Lorgeril, J., A. Lucasson, B. Petton, E. Toulza, C. Montagnani, C. Clerissi, J. Vidal-Dupiol, C. Chaparro, R. Galinier, J. M. Escoubas, P. Haffner, L. Degremont, G. M. Charriere, M. Lafont, A. Delort, A. Vergnes, M. Chiarello, N. Faury, T. Rubio, M. A. Leroy, A. Perignon, D. Regler, B. Morga, M. Alunno-Bruscia, P. Boudry, F. Le Roux, D. Destoumieux-Garzomicronn, Y. Gueguen and G. Mitta (2018). "Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters." Nat Commun **9**(1): 4215.

De Sordi, L., V. Khanna and L. Debarbieux (2017). "The Gut Microbiota Facilitates Drifts in the Genetic Diversity and Infectivity of Bacterial Viruses." Cell Host Microbe **22**(6): 801-808 e803.

Deorowicz, S., A. Debudaj-Grabysz and A. Gudys (2016). "FAMSA: Fast and accurate multiple sequence alignment of huge protein families." Sci Rep **6**: 33964.

Foox, J., S. W. Tighe, C. M. Nicolet, J. M. Zook, M. Byrska-Bishop, W. E. Clarke, M. M. Khayat, M. Mahmoud, P. K. Laaguiby, Z. T. Herbert, D. Warner, G. S. Grills, J. Jen, S. Levy, J. Xiang, A. Alonso, X. Zhao, W. Zhang, F. Teng, Y. Zhao, H. Lu, G. P. Schroth, G. Narzisi, W. Farmerie, F. J. Sedlazeck, D. A. Baldwin and C. E. Mason (2021). "Performance assessment of DNA sequencing platforms in the ABRF Next-Generation Sequencing Study." Nat Biotechnol **39**(9): 1129-1140.

Froelich, B. A. and R. T. Noble (2014). "Factors affecting the uptake and retention of Vibrio vulnificus in oysters." Appl Environ Microbiol **80**(24): 7454-7459.

Gordillo Altamirano, F. L. and J. J. Barr (2019). "Phage Therapy in the Postantibiotic Era." Clin Microbiol Rev **32**(2).

Grazziotin, A. L., E. V. Koonin and D. M. Kristensen (2017). "Prokaryotic Virus Orthologous Groups (pVOGs): a resource for comparative genomics and protein family annotation." Nucleic Acids Res **45**(D1): D491-D498.

Gurevich, A., V. Saveliev, N. Vyahhi and G. Tesler (2013). "QUAST: quality assessment tool for genome assemblies." Bioinformatics **29**(8): 1072-1075.

Hehemann, J. H., P. Arevalo, M. S. Datta, X. Yu, C. H. Corzett, A. Henschel, S. P. Preheim, S. Timberlake, E. J. Alm and M. F. Polz (2016). "Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes." Nat Commun **7**: 12860.

Huerta-Cepas, J., D. Szklarczyk, D. Heller, A. Hernandez-Plaza, S. K. Forslund, H. Cook, D. R. Mende, I. Letunic, T. Rattei, L. J. Jensen, C. von Mering and P. Bork (2019). "eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses." Nucleic Acids Res **47**(D1): D309-D314.

Hunt, D. E., L. A. David, D. Gevers, S. P. Preheim, E. J. Alm and M. F. Polz (2008). "Resource partitioning and sympatric differentiation among closely related bacterioplankton." Science **320**(5879): 1081-1085.

Hussain, F. A., J. Dubert, J. Elsherbini, M. Murphy, D. VanInsberghe, P. Arevalo, K. Kauffman, B. K. Rodino-Janeiro, H. Gavin, A. Gomez, A. Lopatina, F. Le Roux and M. F. Polz (2021). "Rapid evolutionary turnover of mobile genetic elements drives bacterial resistance to phages." Science **374**(6566): 488-492.

Hyman, P. and S. T. Abedon (2009). "Practical methods for determining phage growth parameters." Methods Mol Biol **501**: 175-202.

Jain, C., R. L. Rodriguez, A. M. Phillippy, K. T. Konstantinidis and S. Aluru (2018). "High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries." Nat Commun **9**(1): 5114.

Jones, P., D. Binns, H. Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S. Y. Yong, R. Lopez and S. Hunter (2014). "InterProScan 5: genome-scale protein function classification." Bioinformatics **30**(9): 1236-1240.

Kauffman, K. M., J. M. Brown, R. S. Sharma, D. VanInsberghe, J. Elsherbini, M. Polz and L. Kelly (2018). "Viruses of the Nahant Collection, characterization of 251 marine Vibrionaceae viruses." Scientific Data **5**(1): 180114.

Kauffman, K. M., W. K. Chang, J. M. Brown, F. A. Hussain, J. Yang, M. F. Polz and L. Kelly (2022). "Resolving the structure of phage-bacteria interactions in the context of natural diversity." Nat Commun **13**(1): 372.

Kortright, K. E., B. K. Chan, J. L. Koff and P. E. Turner (2019). "Phage Therapy: A Renewed Approach to Combat Antibiotic-Resistant Bacteria." Cell Host Microbe **25**(2): 219-232.

Le Roux, F., J. Binesse, D. Saulnier and D. Mazel (2007). "Construction of a Vibrio splendidus mutant lacking the metalloprotease gene vsm by use of a novel counterselectable suicide vector." Appl Environ Microbiol **73**(3): 777-784.

Le Roux, F., B. M. Davis and M. K. Waldor (2011). "Conserved small RNAs govern replication and incompatibility of a diverse new plasmid family from marine bacteria." Nucleic Acids Res **39**(3): 1004-1013.

Le Roux, F., K. M. Wegner, C. Baker-Austin, L. Vezzulli, C. R. Osorio, C. Amaro, J. M. Ritchie, T. Defoirdt, D. Destoumieux-Garzon, M. Blokesch, D. Mazel, A. Jacq, F. Cava, L. Gram, C. C. Wendling, E. Strauch, A. Kirschner and S. Huehn (2015). "The emergence of Vibrio pathogens in Europe: ecology, evolution, and pathogenesis (Paris, 11-12th March 2015)." Front Microbiol **6**: 830.

Le Roux, F., K. M. Wegner and M. F. Polz (2016). "Oysters and Vibrios as a Model for Disease Dynamics in Wild Animals." Trends Microbiol.

Lemire, A., D. Goudenege, T. Versigny, B. Petton, A. Calteau, Y. Labreuche and F. Le Roux (2014). "Populations, not clones, are the unit of vibrio pathogenesis in naturally infected oysters." ISME J **9**(7): 1523-1531.

Lemire, A., D. Goudenège, T. Versigny, B. Petton, A. Calteau, Y. Labreuche and F. Le Roux (2015). "Populations, not clones, are the unit of vibrio pathogenesis in naturally infected oysters." The ISME journal **9**(7): 1523-1531.

McNair, K., C. Zhou, E. A. Dinsdale, B. Souza and R. A. Edwards (2019). "PHANOTATE: a novel approach to gene identification in phage genomes." Bioinformatics **35**(22): 4537-4542.

Millman, A., S. Melamed, G. Amitai and R. Sorek (2020). "Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems." Nat Microbiol **5**(12): 1608-1615.

Mistry, J., S. Chuguransky, L. Williams, M. Qureshi, G. A. Salazar, E. L. L. Sonnhammer, S. C. E. Tosatto, L. Paladin, S. Raj, L. J. Richardson, R. D. Finn and A. Bateman (2021). "Pfam: The protein families database in 2021." Nucleic Acids Res **49**(D1): D412-D419.

Moraru, C., A. Varsani and A. M. Kropinski (2020). "VIRIDIC-A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses." Viruses **12**(11).

Nguyen, L. T., H. A. Schmidt, A. von Haeseler and B. Q. Minh (2015). "IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies." Mol Biol Evol **32**(1): 268-274.

Nobrega, F. L., A. R. Costa, L. D. Kluskens and J. Azeredo (2015). "Revisiting phage therapy: new applications for old resources." Trends Microbiol **23**(4): 185-191.

Perrin, A. and E. P. C. Rocha (2021). "PanACoTA: a modular tool for massive microbial comparative genomics." NAR Genom Bioinform **3**(1): lqaa106.

Petton, B., M. Bruto, A. James, Y. Labreuche, M. Alunno-Bruscia and F. Le Roux (2015). "Crassostrea gigas mortality in France: the usual suspect, a herpes virus, may not be the killer in this polymicrobial opportunistic disease." Front Microbiol **6**: 686.

Piel, D., M. Bruto, A. James, Y. Labreuche, C. Lambert, A. Janicot, S. Chenivesse, B. Petton, K. M. Wegner, C. Stoudmann, M. Blokesch and F. Le Roux (2020). "Selection of Vibrio crassostreae relies on a plasmid expressing a type 6 secretion system cytotoxic for host immune cells." Environ Microbiol **22**(10): 4198-4211.

Piel, D., M. Bruto, Y. Labreuche, F. Blanquart, D. Goudenege, R. Barcia-Cruz, S. Chenivesse, S. Le Panse, A. James, J. Dubert, B. Petton, E. Lieberman, K. M. Wegner, F. A. Hussain, K. M. Kauffman, M. F. Polz, D. Bikard, S. Gandon, E. P. C. Rocha and F. Le Roux (2022). "Phage-host coevolution in natural populations." Nat Microbiol.

Preheim, S. P., S. Timberlake and M. F. Polz (2011). "Merging taxonomy with ecological population prediction in a case study of Vibrionaceae." Appl Environ Microbiol **77**(20): 7195-7206.

Rubio, T., D. Oyanedel, Y. Labreuche, E. Toulza, X. Luo, M. Bruto, C. Chaparro, M. Torres, J. de Lorgeril, P. Haffner, J. Vidal-Dupiol, A. Lagorce, B. Petton, G. Mitta, A. Jacq, F. Le Roux, G. M. Charriere and D. Destoumieux-Garzon (2019). "Species-specific mechanisms of cytotoxicity toward immune cells determine the successful outcome of Vibrio infections." Proc Natl Acad Sci U S A **116**(28): 14238-14247.

Shapiro, B. J., J. Friedman, O. X. Cordero, S. P. Preheim, S. C. Timberlake, G. Szabo, M. F. Polz and E. J. Alm (2012). "Population genomics of early events in the ecological differentiation of bacteria." Science **336**(6077): 48-51.

Smyshlyaev, G., A. Bateman and O. Barabas (2021). "Sequence analysis of tyrosine recombinases allows annotation of mobile genetic elements in prokaryotic genomes." Mol Syst Biol **17**(5): e9880.

Stamatakis, A. (2006). "RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models." Bioinformatics **22**(21): 2688-2690.

Steinegger, M. and J. Soding (2018). "Clustering huge protein sequence sets in linear time." Nat Commun **9**(1): 2542.

Szabo, G., S. P. Preheim, K. M. Kauffman, L. A. David, J. Shapiro, E. J. Alm and M. F. Polz (2012). "Reproducibility of Vibrionaceae population structure in coastal bacterioplankton." ISME J **7**(3): 509-519.

Tesson, F., A. Herve, E. Mordret, M. Touchon, C. d'Humieres, J. Cury and A. Bernheim (2022). "Systematic and quantitative view of the antiviral arsenal of prokaryotes." Nat Commun **13**(1): 2561.

Tesson, F., A. Hervé, M. Touchon, C. d'Humières, J. Cury and A. Bernheim (2021). "Systematic and quantitative view of the antiviral arsenal of prokaryotes." bioRxiv: 2021.2009.2002.458658.

Vallenet, D., E. Belda, A. Calteau, S. Cruveiller, S. Engelen, A. Lajus, F. Le Fevre, C. Longin, D. Mornico, D. Roche, Z. Rouy, G. Salvignol, C. Scarpelli, A. A. Thil Smith, M. Weiman and C. Medigue (2013). "MicroScope--an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data." Nucleic Acids Res **41**(Database issue): D636-647.

Wang, L., S. Chen, K. L. Vergin, S. J. Giovannoni, S. W. Chan, M. S. DeMott, K. Taghizadeh, O. X. Cordero, M. Cutler, S. Timberlake, E. J. Alm, M. F. Polz, J. Pinhassi, Z. Deng and P. C. Dedon (2011). "DNA phosphorothioation is widespread and quantized in bacterial genomes." Proc Natl Acad Sci U S A **108**(7): 2963-2968.

Xiong, X., G. Wu, Y. Wei, L. Liu, Y. Zhang, R. Su, X. Jiang, M. Li, H. Gao, X. Tian, Y. Zhang, L. Hu, S. Chen, Y. Tang, S. Jiang, R. Huang, Z. Li, Y. Wang, Z. Deng, J. Wang, P. C. Dedon, S. Chen and L. Wang (2020). "SspABCD-SspE is a phosphorothioation-sensing bacterial defence system with broad anti-phage activities." Nat Microbiol **5**(7): 917-928.

Yawata, Y., O. X. Cordero, F. Menolascina, J. H. Hehemann, M. F. Polz and R. Stocker (2014). "Competition-dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations." Proc Natl Acad Sci U S A **111**(15): 5622-5627.

TITLES AND LEGENDS TO FIGURES

Figure 1. Time-series sampling of *V. chagasii* and vibriophages

a, On each sampling date, vibrios from seawater (1-0.2µm size fraction) or tissue from five oysters were selected on TCBS and genotyped to identify *V. chagasii* isolates. Oyster mortalities occurred between May 29th-August 25th (black bar) and *V. crassostreae* was detected in oysters from June 16th-August 18th (pink bar). Green lines show the frequency of *V. chagasii* (number of positive isolates out of 48 randomly picked colonies*100) in seawater (solid) or oyster tissue (dashed). Each of the *V. chagasii* isolates was used as "bait" to detect co-occurring lytic phages by plaque formation in soft agar overlays of host lawns mixed with viral concentrate from 1) seawater and 2) oyster tissues. Black circles and triangles indicate the mean number ± se of plaque forming units (PFU) obtained for each plaque positive host using viruses collected from seawater (circles) or oyster tissue (triangles) collected on the same day. The red star indicates the date of the *V. chagasii* bloom (25th August). **b,** Correlation between mean number of PFUs and *V. chagasii* frequency for viral concentrates from seawater (upper panel) or oysters (lower panel). Regression lines show similar correlation estimates for both isolation fractions (seawater: estimate = 0.627 ± 0.193, t = 3.252, p = 0.004; oyster: estimate = 0.849 ± 0.342, t = 2.487, p = 0.022).

Figure 2. Genome diversity of *V. crassostreae.* and *V. chagasii*. a, Core genome phylogeny based on 2689 gene families of 88 *V. crassostreae* and 136 *V. chagasii* isolates from the time series, with pairwise ANI values revealing highly structured clades in *V. crassostreae*, but not in *V. chagasii*. Clade numbers refer to (Piel, Bruto et al. 2022). **b,** The resulting distribution of pairwise genetic distance within *V. chagasii* (green) and *V. crassostreae* (pink) indicates striking differences between these species (Two-sample Kolmogorov-Smirnov test, D = 0.762, $p < 0.001$): *V. chagasii* exhibits a unimodal distribution with few clonal strains and a homogenous pairwise genetic distances with a maximum differentiation of 0.024, whereas *V. crassostreae* exhibits a multimodal distribution corresponding to the distinct clades with low within-clade and large between clade genetic distances.

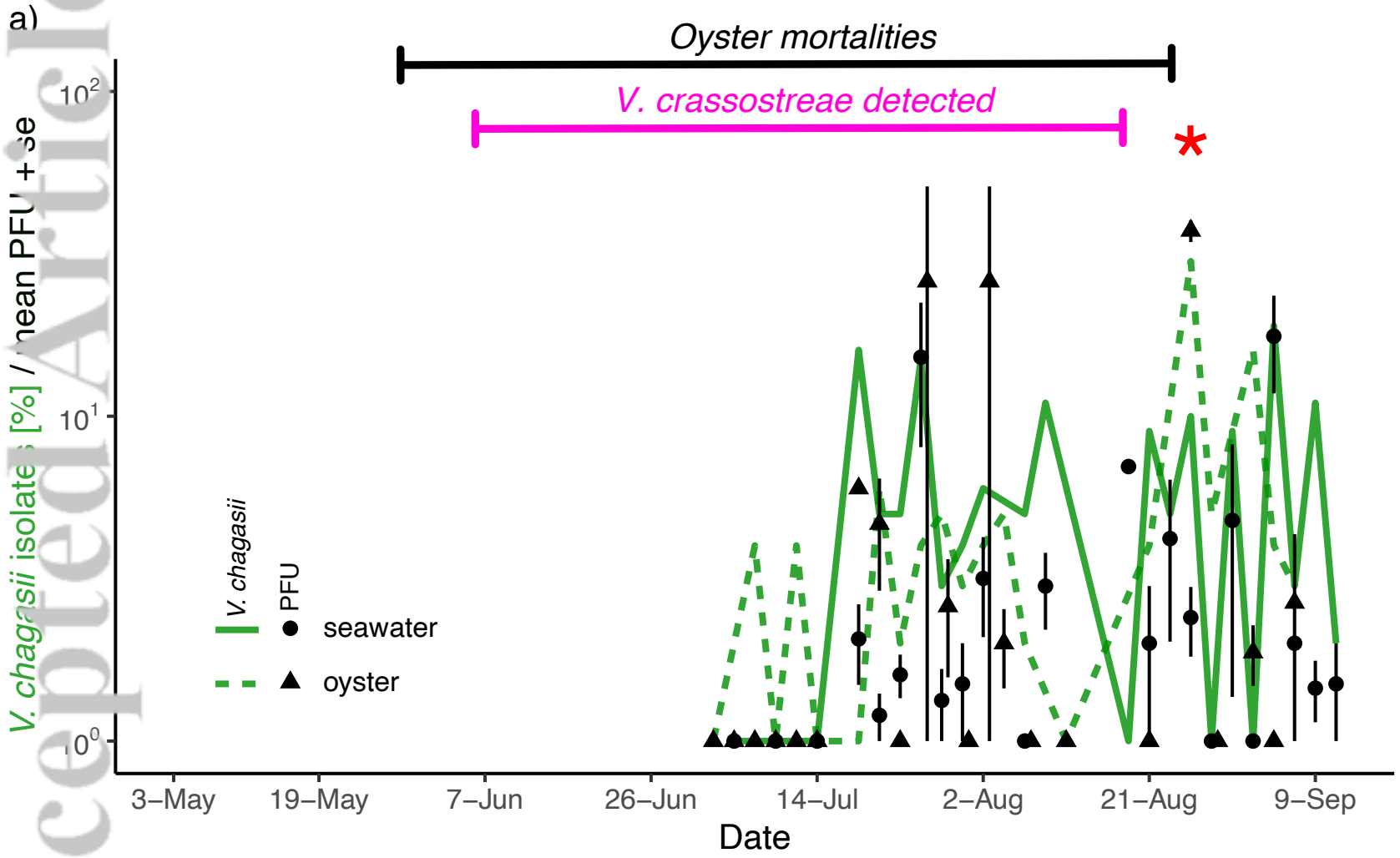
Figure 3. Diversity of phages isolated using *V. chagasii* and *V. crassostreae* as hosts. a, VIRIDIC intergenomic similarity between phage genomes. The 106 phages included in the study were grouped into 43 clusters assigned to VIRIDIC genus rank (>70% identities, indicated with a plain line). Only one genus contains phages isolated using *V. crassostreae* or

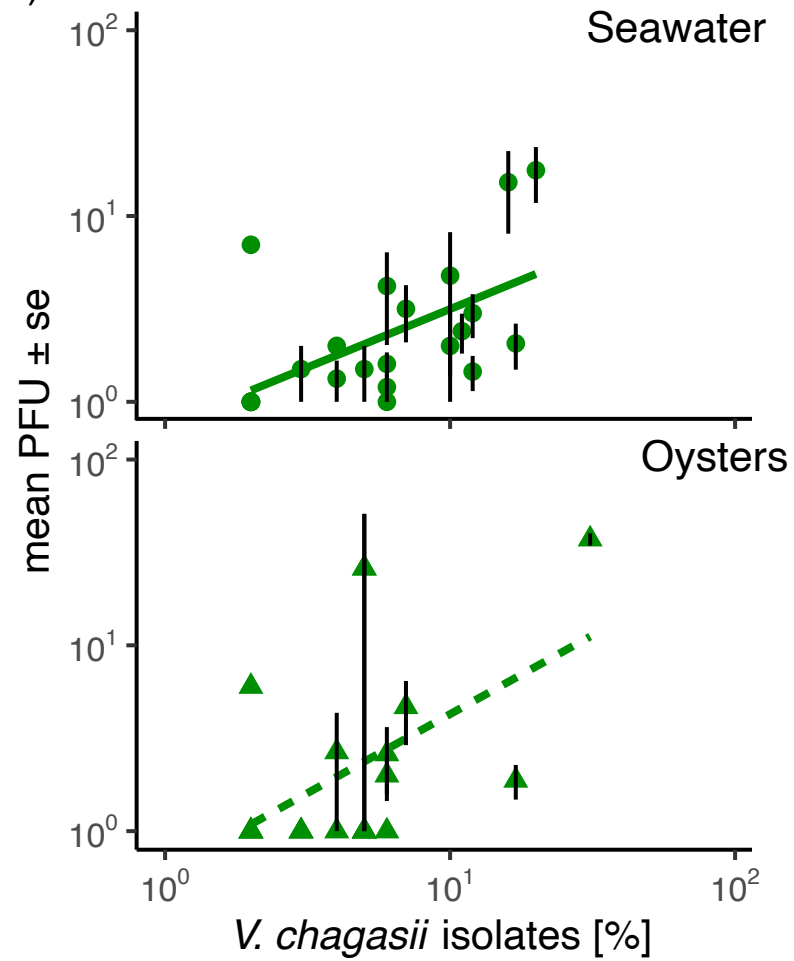
V. chagasii. The phylogenic tree based on 41 core proteins (25% identities and 80% coverage) shows that this genus splits into two clades, one represented by a single phage isolated from *V. chagasii* and the second containing seven phages all isolated from *V. crassostreae.* Some genera were grouped into clusters assigned to VIRIDIC family rank (>50%). For the larger family, the phylogenic tree based on 38 core proteins (25% identities and 80% coverage) delineates the phages isolated from *V. crassostreae* and *V. chagasii* populations.
b, The 49 and 57 phages infecting *V. chagasii* and *V. crassostreae* were assigned to 25 and 18 genera and 41 and 23 species, respectively.

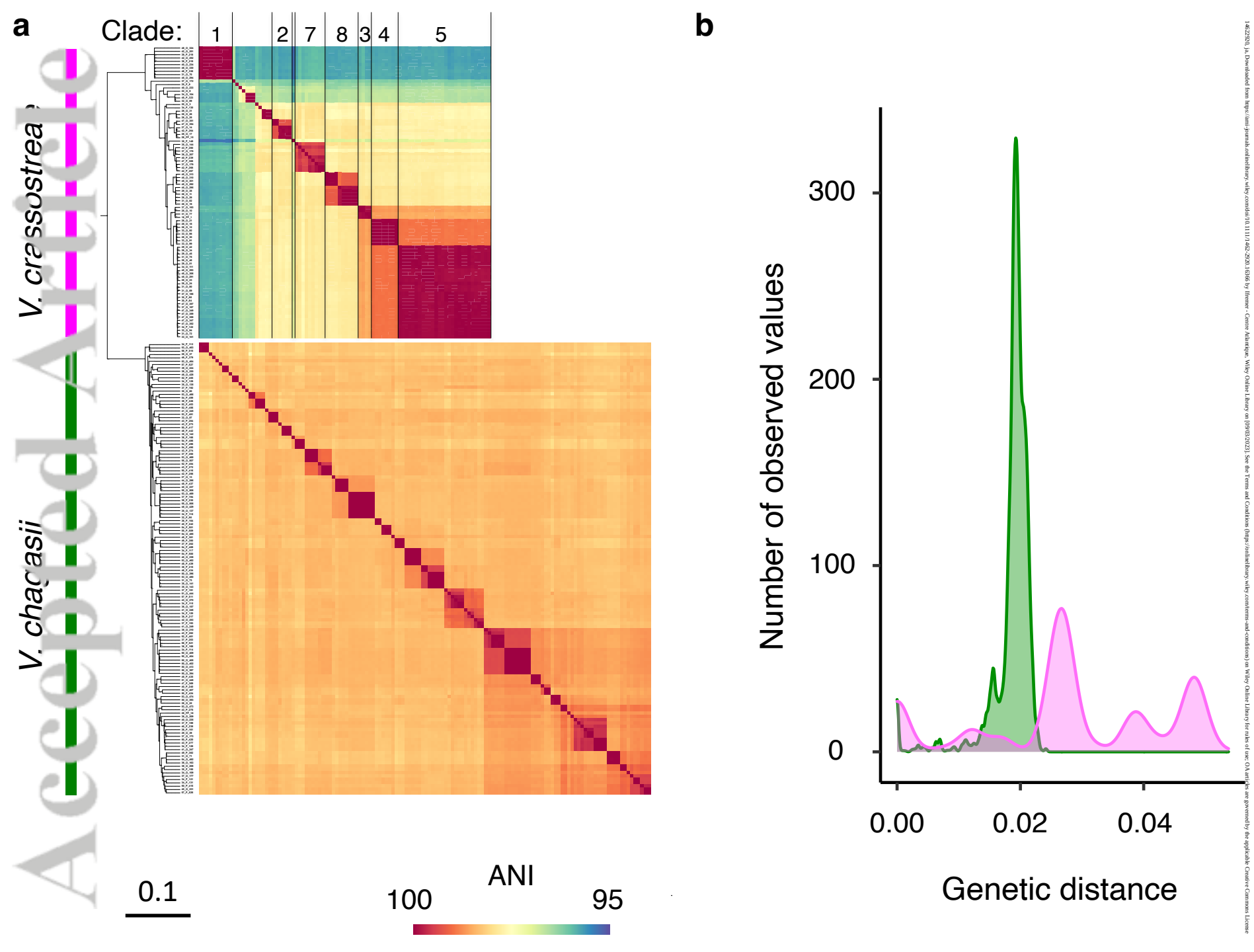
Figure 4. Modularity is driven by phylogenetic distances within the host population. a, Comparisons of the phage-bacteria infection matrix for *V. crassostreae* and *V. chagasii* as hosts. Rows represent sequenced *Vibrio* strains ordered according to the Maximum Likelihood core genome phylogeny of *V. chagasii* (n=136 green matrix) and *V. crassostreae* strains (n=88; pink matrix). Columns represent sequenced phages (49 and 57 isolated from *V. chagasii* and *V. crassostreae*) ordered by VIRIDIC genus (see dendrogram in Fig. 3a). Colored squares indicate killing of the respective vibrio isolate by the phage. Vertical lines represent VIRIDIC genus (solid) and species (dashed). **b,** Sharing probability (i.e. proportion of phages shared in all pairs within a window of genetic distances) weighed by the frequency of pairwise comparisons (i.e. number of pairwise comparisons within each window divided by the total number of pairwise comparisons) as a function of host genetic distance. Sharing probability was assessed in sliding windows with an interval size of 0.003 and a step size of 0.001 along observed genetic distances.

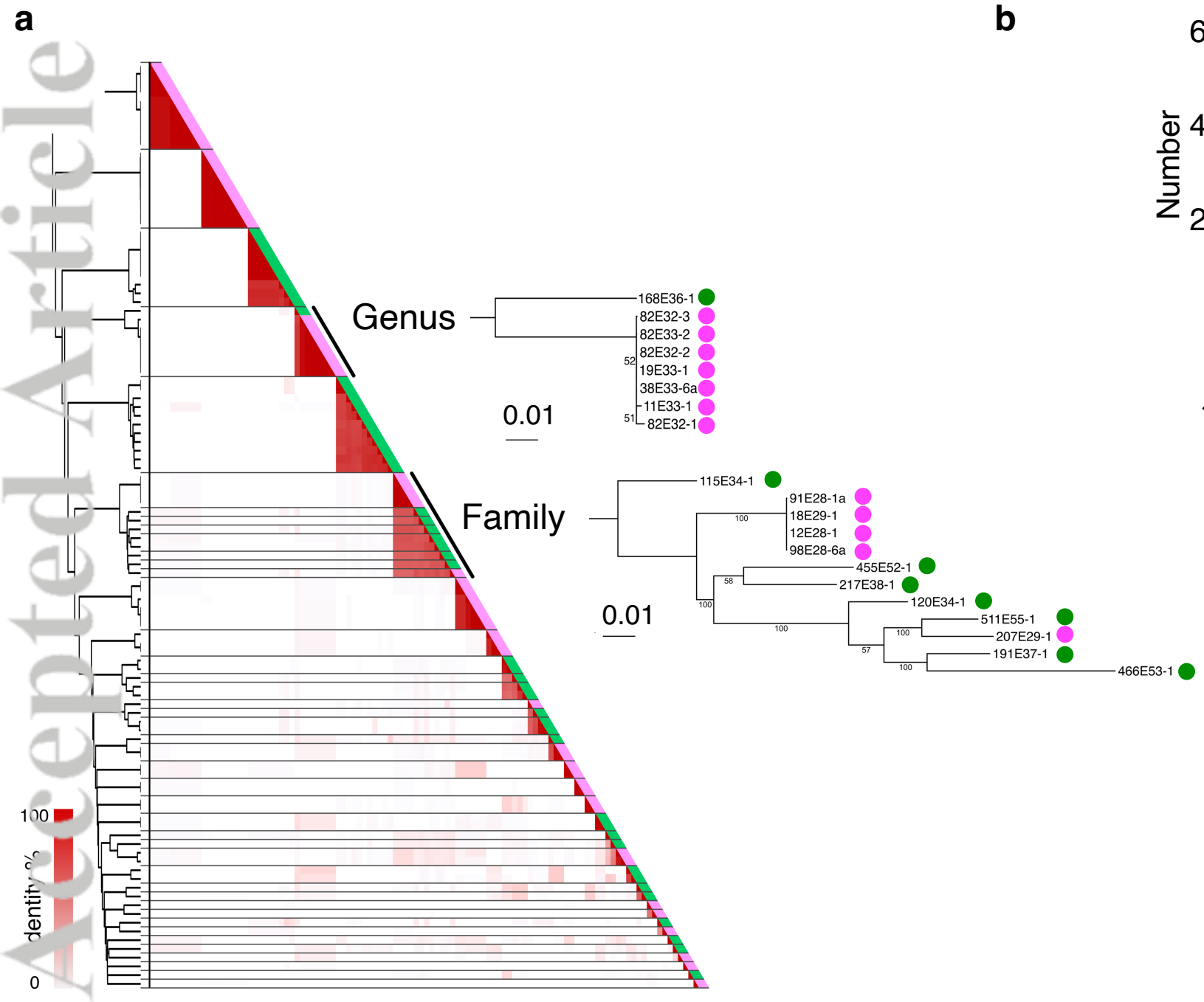
Figure 5: Epigenetic and genetic modification are involved in phage adaptation. a, Efficiency of plating (EOP, the titer of the phage on a given bacteria divided by the titer on the strain used to isolate the phage or 'original host') of six phages from VIRIDIC species 26 (Fig. S6) on 12 *V. chagasii* strains. Individual dots correspond to the mean from three independent experiments. **b,** EOP of the phage 409E50-1 infecting *V. chagasii* strains with low susceptibility (LS) or high susceptibility (HS). Bar charts show the mean +/- s.d. from three independent experiments (individual dots). **c,** EOP obtained using *V. chagasii* strains 50_O_409 (HS) or 56_P_521 (LS) to produce the phages 409E50-1 or 521E56-1. Results were standardized by experiment, i.e. the production and testing on the original host was normalized to 1 for each experiment. **d,** Graphic summary of the strategy used to exchange the nucleotide T to G and codon Thr to Pro in phage 409E50-1. **e**, EOP of the recombinant phages or control infecting the

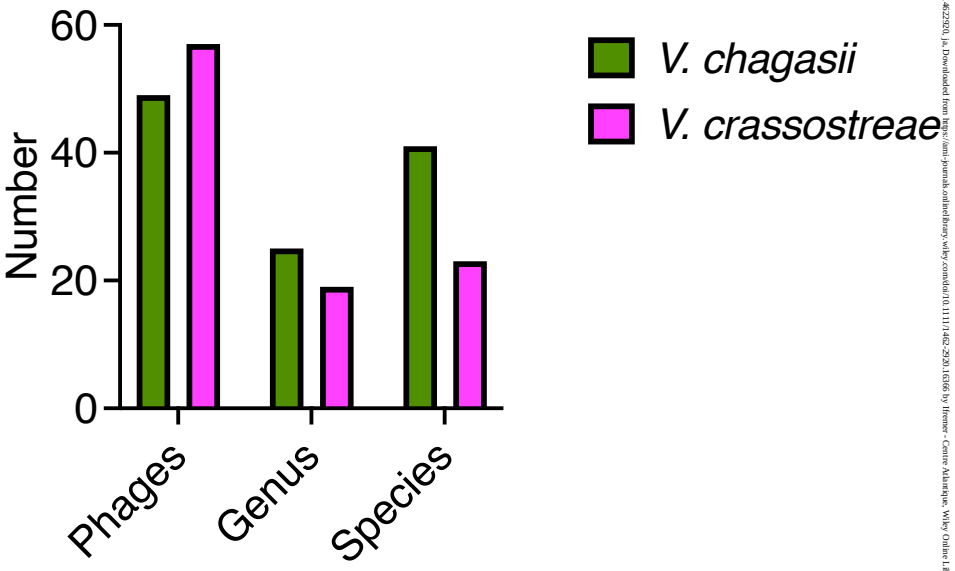
four *V. chagasii* strains with low susceptibility (LS). Results were standardized by experiment, i.e. the production in original host (50_0_409) was normalized to 1 for each experiment.

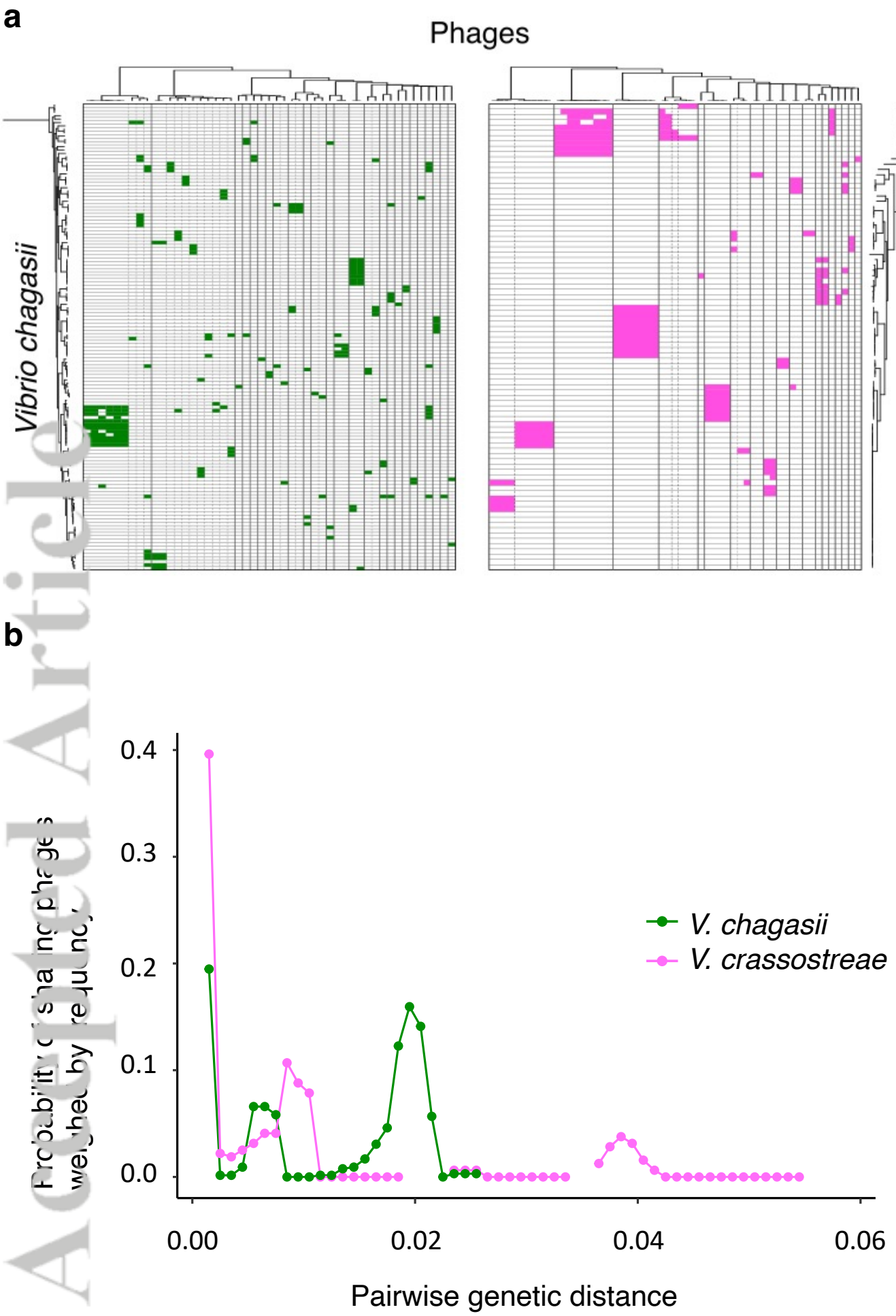












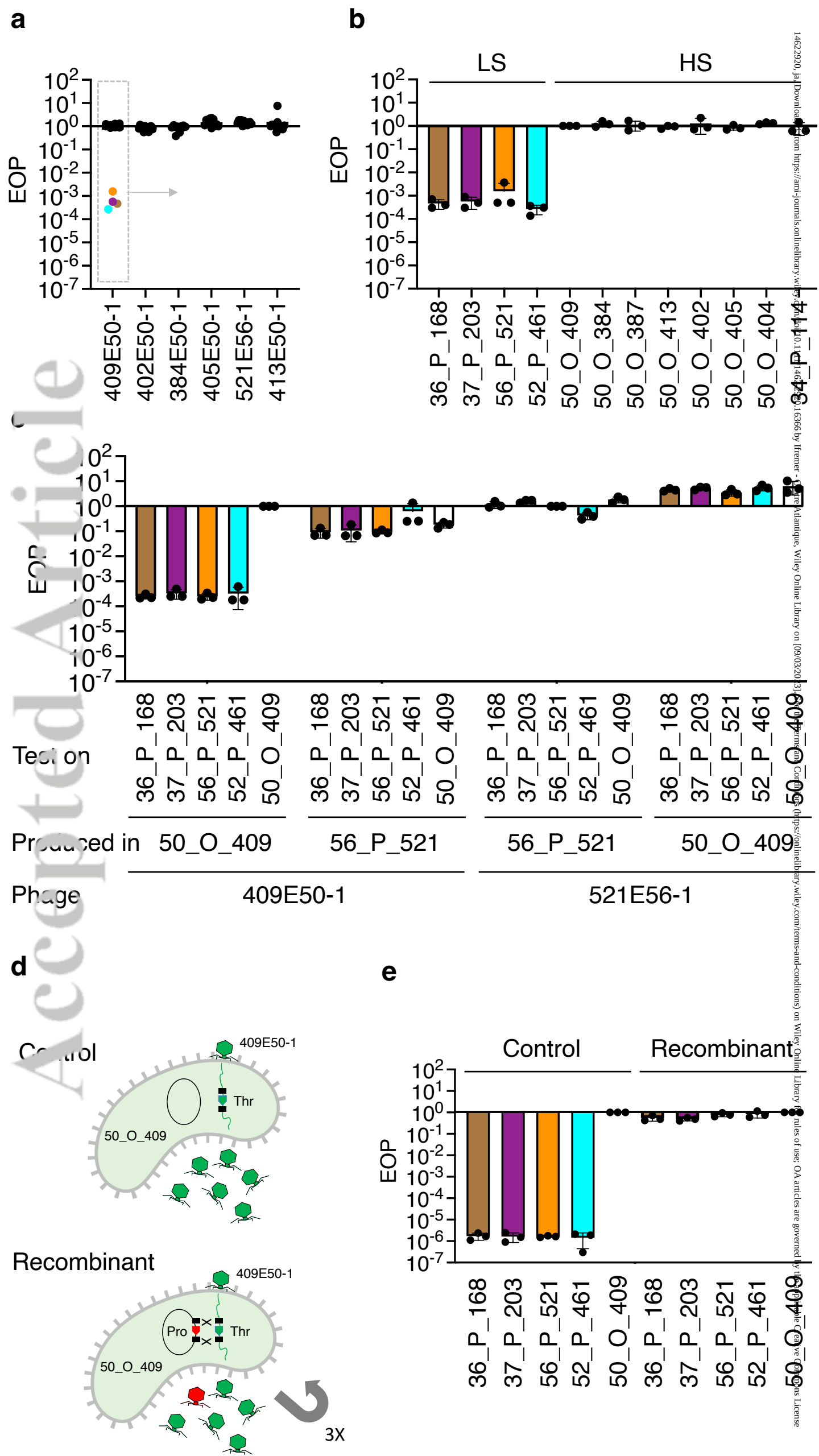


Table 1. Summary of the time series sampling of vibrios and their infecting vibriophages.

Vibrio		Number of isolated hosts	Number of plaque sensitive hosts using viral source				Number of PFU using viral source			
			From seawater	Ratio	From oysters	Ratio	From seawater	Ratio	From oysters	Ratio
V. chagasii	From seawater	144	54	54/144=0.375	7		676	676/144=4.69	16	
	From oysters	108	47		12	12/108=0.111	1347		19	19/108=0.175
V. crassostreae	From seawater	39	4	4/39=0.102	1		17	17/39=0.43	6	
	From oysters	155	14		3	3/155=0.019	30		3	3/155=0.019
				$\chi2$ = 5.298, p = 0.021		$\chi2$ = 7.25 p = 0.007		$\chi2$ = 82.916, p < 0.001		$\chi2$ = 15.079, p < 0.001