DOI: 10.1038/S41564-022-01157-1; https://www.nature.com/articles/s41564-022-01157-1

Phage–host coevolution in natural populations

Damien Piel[1,2], Maxime Bruto[2], Yannick Labreuche[1,2], François Blanquart[3,4], David Goudenège[1,2], Rubén Barcia-Cruz[2,5], Sabine Chenivesse[2], Sophie Le Panse[2], Adèle James[1], Javier Dubert[2,5,6], Bruno Petton[1], Erica Lieberman[7], K. Mathias Wegner[8], Fatima A. Hussain[6], Kathryn M. Kauffman[6], Martin F. Polz[6,9], David Bikard[10], Sylvain Gandon[11], Eduardo P. C. Rocha[12] and Frédérique Le Roux[1,2] ✉

Coevolution between bacteriophages (phages) and their bacterial hosts occurs through changes in resistance and counter-resistance mechanisms. To assess phage–host evolution in wild populations, we isolated 195 *Vibrio crassostreae* strains and 243 vibriophages during a 5-month time series from an oyster farm and combined these isolates with existing *V. crassostreae* and phage isolates. Cross-infection studies of 81,926 host–phage pairs delineated a modular network where phages are best at infecting co-occurring hosts, indicating local adaptation. Successful propagation of phage is restricted by the ability to adsorb to closely related bacteria and further constrained by strain-specific defence systems. These defences are highly diverse and predominantly located on mobile genetic elements, and multiple defences are active within a single genome. We further show that epigenetic and genomic modifications enable phage to adapt to bacterial defences and alter host range. Our findings reveal that the evolution of bacterial defences and phage counter-defences is underpinned by frequent genetic exchanges with, and between, mobile genetic elements.

The ongoing battle between bacteriophages (phages) and their bacterial hosts is probably billions of years old and involves a large set of defence and counter-defence systems[1,2]. Coevolutionary dynamics of host–phage interactions involve different steps of the infection process[3–5]. Infection requires phages to adsorb to specific receptors at the bacterial cell surface and then bypass intracellular host defences[6]. Recent work on marine *Vibrionaceae* (hereafter vibrios) highlighted that most phage defence genes are encoded on mobile genetic elements (MGEs) that are gained and lost at high rates, differentiating otherwise genomically identical isolates[7]. At the same time, gene-specific selective sweeps can result in emergence of diverse species with monomorphic phage receptors[7]. Hence, phage can infect more diverse hosts than they can kill, owing to the presence of highly variable anti-phage cell defence systems[7,8]. Bacterial genomes encode numerous defence systems, either because additional immune barriers provide better protection, or because different phages are inhibited by different defence systems[7,9,10]. A single bacterial strain can acquire immune defence mechanisms by horizontal gene transfer to form what has been called a pan-immune system[11].

As phages are thought to rapidly evolve counter-defences to thrive, it is expected that the diversity of escape mechanisms within closely related phages will mirror the hosts' pan-immune system[11]. While the known repertoire of defence systems has recently expanded, counter-defence strategies of phages are much less well characterized. Bacteria can considerably expand their genome size and incorporate multiple defence systems[11]. In contrast, the size of capsids constrains phage genome size to a range of variation of 5–10%, limiting the number of escape mechanisms a phage can encode[12]. Phages are, however, able to exchange genes[13], so an open question is how phage populations counter the numerous defence systems in their hosts and how this affects phage specificity in natural populations.

We address this question by analysing the dynamics of phage–bacteria coevolution in natural populations of the oyster pathogen, *Vibrio crassostreae* and its phages. We previously reported that juveniles affected by Pacific oyster mortality syndrome are infected by diverse virulent genotypes of the *V. crassostreae* species[14,15]. We also showed that vibrio populations are dynamic[14]. Here we combine cultivation, comparative genomics and molecular genetics to analyse a large collection of bacterial isolates and their phages collected at two separate marine sites.

Isolation of oyster-pathogenic vibrios and phages

To study the dynamics of *V. crassostreae* at fine temporal resolution, we sampled vibrios from juvenile oysters in an oyster farm (Bay of Brest, France) and from the surrounding seawater on 57 dates over 5 months, including a period with mortality events (Methods). Vibrios were isolated on selective media and 195 isolates (Supplementary Table 1) were further assigned to the *V. crassostreae* species by the sequencing and analysis of the gene encoding the DNA gyrase subunit B. *V. crassostreae* was only isolated during the disease outbreak,

[1]Ifremer, Unité Physiologie Fonctionnelle des Organismes Marins, ZI de la Pointe du Diable, Plouzané, France. [2]Sorbonne Université, UPMC Paris 06, CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, Roscoff cedex, France. [3]Centre for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS, INSERM, PSL Research University, Paris, France. [4]Infection Antimicrobials Modelling Evolution, UMR 1137, INSERM, Université de Paris, Paris, France. [5]Department of Microbiology and Parasitology, CIBUS-Faculty of Biology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain. [6]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [7]Eligo Bioscience, Paris, France. [8]AWI - Alfred Wegener Institut - Helmholtz-Zentrum für Polar- und Meeresforschung, Coastal Ecology, Waddensea Station Sylt, List, Germany. [9]Division of Microbial Ecology, Centre for Microbiology and Environmental Systems Science, University of Vienna, Vienna, Austria. [10]Synthetic Biology Group, Department of Microbiology, Institut Pasteur, Paris, France. [11]CEFE (UMR 5175 CNRS - Université de Montpellier - Université Paul Valéry - EPHE), Montpellier, France. [12]Institut Pasteur, Université de Paris, CNRS UMR3525, Microbial Evolutionary Genomics, Paris, France. ✉e-mail: frederique.le-roux@sb-roscoff.fr

with frequencies varying from 0–16% for seawater and 0–58% for oysters depending on the date of sampling (Extended Data Fig. 1), consistent with the previously determined increased prevalence of this species in diseased oysters[14]. We used the 195 *V. crassostreae* isolates as 'bait' to isolate 243 phages from viral concentrates or oyster tissues collected at the same location (Methods, Supplementary Fig. 1 and Table 2). This collection was complemented with 51 bacteria and 31 phages isolated in 2016 in Sylt (Germany), where the oyster beds have not suffered *V. crassostreae*-related disease outbreaks. Accordingly, the *V. crassostreae*-specific virulence plasmid was absent from the Sylt collection[16].

Genomic analyses of phage–bacteria collection

To investigate phage host-range, our set of 274 phages was tested against 117 *V. crassostreae* isolates from the Brest time series, 34 strains previously sampled in Brest, 51 strains from Sylt and 97 representative members of other vibrio species (Methods, Supplementary Table 1 and Supplementary Table 2). We found that 1,861 of the 81,926 tested host–phage interactions resulted in infection (positive) (2.2%, Extended Data Fig. 2). Analyses of the infection outcomes revealed highly specific interactions between phage and bacteria. We found that 14/274 phages were able to infect members of other vibrio species and only one phage with a broad host range (Extended Data Fig. 2). These results are consistent with a previous report of a sparse infection network using a diverse collection of oceanic vibrios[17].

Next we sequenced and assembled the genomes of selected *V. crassostreae* strains (Supplementary Table 3). Representative strains (n =157) were chosen by clustering the *V. crassostreae* isolates with identical *gyrB* sequences and identical patterns of susceptibility in the cross-infection assays. Overall, genomes varied in size between 4.9 Mb to 6.4 Mb (Supplementary Table 3), confirming the existence of extensive genomic diversity among very closely related bacterial isolates. We analysed the core genome phylogeny of *V. crassostreae* and delineated 8 clades (named V1 to V8) with different phylogenetic depth (Fig. 1). The most closely related genomes were separated by just 16 single nucleotide polymorphisms (SNPs) in the core genome. Even these closely related genomes harbour dozens of unique genes.

We grouped phages according to host, date of isolation and patterns of infectivity. We selected 75 representatives and characterized their morphology, genome sequence and taxonomy (Supplementary Table 4). Electron microscopy revealed that all phages belong to the double-stranded DNA *Caudovirales*, including podoviruses (n = 31), myoviruses (n = 15) and siphoviruses (n = 29) (Supplementary Figs. 2–4). Genome sequencing revealed extensive variability in genome length (from 36.8 to 253.3 kbp) and gene content (39–444 predicted coding sequences) (Supplementary Table 4). We could not find any genomic evidence that the phages are temperate (Methods). To evaluate phage phylogeny, we used the Virus Intergenomic Distance Calculator (VIRIDIC[18]), which grouped phages into 28 clusters (VIRIDIC genus rank). Intra-genus genomic comparisons revealed high conservation of phage genome size, high nucleotide identity in core genes (average nucleotide identity (ANI) from 95.5 to 99.9%; >99% for 10 out of 14 clusters) but variable gene repertoires (Supplementary Table 4). Recombinases that allow exchanges of divergent sequences underpin variation in genome repertoires of phages[13,17,19] and we detected putative recombinase(s) in 46 of the 75 phage genomes (Supplementary Table 4).

Building a modular host–phage interaction network

We analysed host–phage interactions, taking the vibrio core genome phylogeny into consideration (Fig. 1 and Supplementary Table 5) and found that phages from each cluster kill bacteria from a single clade of *V. crassostreae*, with few exceptions. This enabled us to build a modular interaction network. We named phage–host clusters Px_y, where P indicates phage, x the vibrio clade they infect (1 to 8 or N if a genome was not associated with a larger clade) and y is the VIRIDIC genus number (1 to 28).

The likelihood of success of phage infection is often attributed to the presence of cell-surface phage receptors, but it has also been proposed that intracellular defences determine host range[7]. To distinguish these mechanisms, we estimated whether representative phages from 16 clusters were able to adsorb to receptors on diverse hosts. We found that phages only adsorb to bacteria from their respective vibrio clade, except for one phage (P5_16) that adsorbs to bacteria from clades V5 and V6 (Fig. 2 and Supplementary Fig. 5). Within the same vibrio clade, phages adsorb to all tested isolates, regardless of the production of progeny and cell lysis. We propose that host-range results from two successive mechanisms. Specificity is determined by a matching clade-specific bacterial receptor(s) and cluster-specific phage receptor-binding protein(s). Once the phage has adsorbed to the cell, intracellular defence systems further narrow the range of strains that are successfully infected and killed. To test this hypothesis, we modelled phage–host dynamics over time.

Temporal and spatial variations in the patterns of phage infection

Our cross-infection experiments between phages and bacteria originating from different locations (Brest or Sylt) show that the proportion of infected hosts is higher on sympatric than on allopatric combinations (Fig. 3a). This spatial pattern results from the non-overlapping distribution of vibrio clades and phage clusters across locations (Extended Data Fig. 3) and suggests that phages sampled from a given location are rarely capable of infecting bacteria from the other location. In addition, our cross-infection experiments among phages and bacteria sampled during the Brest time series allowed us to detect temporal variations in the patterns of infection[8,20]. The mean phage infectivity (that is, the average ability of phages isolated from a sample to infect and kill bacteria) peaked for contemporary combinations (that is, phages and bacteria sampled at the same time point) and declined as phages were inoculated on bacterial strains collected from more distant sampling dates in the past or in the future (Fig. 3b and Supplementary Fig. 6). We modelled the number of isolated phages that were found to belong to a given cluster as a function of the frequency of matching-clade and non-matching-clade bacteria (Methods). This analysis revealed that the presence of phages of a given cluster was significantly associated with the presence of the corresponding *V. crassostreae* clade at that time (likelihood ratio test, $P < 0.001$).

These patterns of phage adaptation are consistent with predictions of models of fluctuating selection driven by coevolutionary dynamics, where pathogen populations are maximally adapted to their sympatric and contemporary host populations[21,22]. Bacteria were found to be more resistant to contemporary phages despite phage local adaptation[23]. Even if other biotic and abiotic factors are involved in the fluctuations of vibrio populations in the wild, our results suggest that coevolutionary dynamics between phages and bacteria are driving the observed patterns of adaptation. Next we set out to identify the genetic basis of the specificity of these antagonistic interactions.

Genetic basis of the specificity of host–phage interactions

The existence of genetic diversity in phages and hosts and the modular structure of the interaction matrix open the possibility of understanding the genetic basis of the multiple mechanisms regulating phage–host interactions. To identify the putative mechanisms operating in the bacteria to block phage infection, we used Defense-Finder[24]. This revealed a wealth of different known systems, such as the restriction modification systems[6] R-M type I (200), type IV (51), the DNA phosphorothioation-based defence system Dnd[25] (51) and diverse systems involved in abortive infection[26] (Fig. 4a).

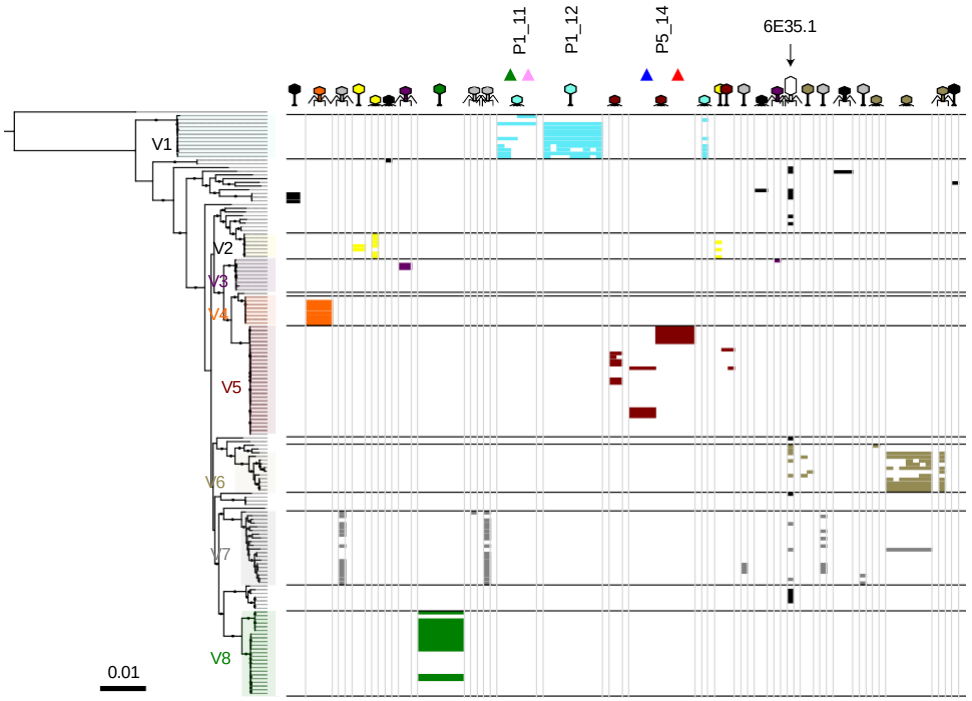


Fig. 1 | A modular phage–vibrio infection network. Rows represent sequenced vibrio strains ordered as a function of a maximum-likelihood core genome phylogeny of 157 *V. crassostreae* isolates and using *V. gigantis* strain 43_P_281 as an out-group (2,498 genes). Bacterial clades (V1 to V8 containing at least 6 strains) are labelled with different colours. Columns represent sequenced phages ($n = 75$) ordered by VIRIDIC genus (from 1 to 28). Coloured squares indicate vibrio killing by the phage (see the host-range matrix in Supplementary Table 5). Black squares indicate killing of vibrios that are outside clades V1 to V8. Phage morphotypes are indicated by specific icons for siphoviruses (long tail), myoviruses (medium tail) and podoviruses (short tail), and the colour of each phage genome refers to the clade assignment of the host they kill (black when not in clade). The myovirus 6E35.1 has a broader host range, a comparatively larger capsid and genome size (253 kb) (Supplementary Table 4). The names of phage clusters that have been subjected to mechanistic analyses are indicated as Px_y (P is for phage, x for the vibrio clade they infect and y is the VIRIDIC genus number). Coloured triangles distinguish groups of phages in the P1_11 (pink, green) and P5_14 (blue, red) genera on the basis of their infectivity.

Overall, the number of defence systems is strongly correlated with bacterial genome size (Spearman rho = 0.769, $P < 0.001$) and both measures correlate negatively with the number of different infecting phages (Spearman rho = −0.281, $P < 0.001$ for genome size and rho = −0.260, $P = 0.001$ for number of defence elements). When correcting for phylogenetic relatedness of the strains, only the correlation of defence elements and number of infecting phages remained significant (phylogenetic independent contrasts, Spearman rho = −0.175, $P = 0.029$) (Fig. 4b), pinpointing the key role of the size of defence system repertoires in protecting the cell.

Defence systems vary widely across isolates, consistent with the high specificity of phage–host interactions. To identify the mechanisms of genetic mobility of defence systems, we identified regions of genomic plasticity (RGPs) in the species pangenome graph using PanRGP[27]. These islands vary in size from 3 kbp to more than 100 kbp. They account for 19% of the pangenome genes but include 92.4% of all defence systems (Fisher test, odds-ratio of 53.02, $P < 0.001$). Genomic islands usually result from horizontal gene transfer driven by MGEs such as prophages or conjugative elements. To test this hypothesis, we analysed the gene repertoires of the RGPs in relation to MGE functions (conjugation-related genes, phage functions and integrases). While we cannot ascertain that all these RGPs are still functional MGEs, inactivation being frequent in the chromosome[28], many of them include the hallmark functions of MGEs. We could classify 47% of the RGPs into putative MGEs: prophages (19% of the RGP genes), plasmids (19%), and among the remaining RGPs we selected those encoding an integrase (9%). Some of the latter may be related to MGE families that were recently named 'phage defence elements' (PDE)[7] and whose only known function is phage defence. These three types of MGEs account for less than 10% of the genes but contain 72% of all defence systems: 29% in prophage, 14% in plasmid and 28% in the 'others' (Fig. 4c). This is significantly more than the RGPs that could not be classified as MGEs (Fisher test, odds-ratio of 2.32, $P < 0.001$) and much more than the rest of the pangenome. We conclude that MGEs play a major role in the emergence of phage resistance since they account for many differences in gene repertoires among closely related strains and encode most phage defence genes.

Phage defence in vibrios with the smallest genomes

The number and diversity of known defence systems carried by closely related strains complicate the prediction of resistance genotypes from the presence of specific systems (Fig. 4a and Supplementary Table 6). In addition, we can expect that many defence systems remain to be identified. We hypothesized that the genetic characterization of immunity mechanisms should be simpler in the vibrios from clade V1 because they have a smaller genome size (average 5.01 Mbp), a lower number of known defence systems (average 5.6) and a higher number of infecting phages (average 8.6) (Fig. 4b). One strain (7F1_18) was resistant to 14 out of 15 phages (Supplementary Table 5). The abovementioned cross-infection tests were done using a constant, relatively low number of phage particles (10^3 p.f.u.) and we first sought to refine these estimates of host susceptibility using increasing quantities of phages (Methods). We classified the strains as 'sensitive' or 'partially sensitive' if a clearing and viable phage production were obtained using a low (1–10 p.f.u.) or higher quantity of phages (10^4–10^5 p.f.u.), respectively. The strains

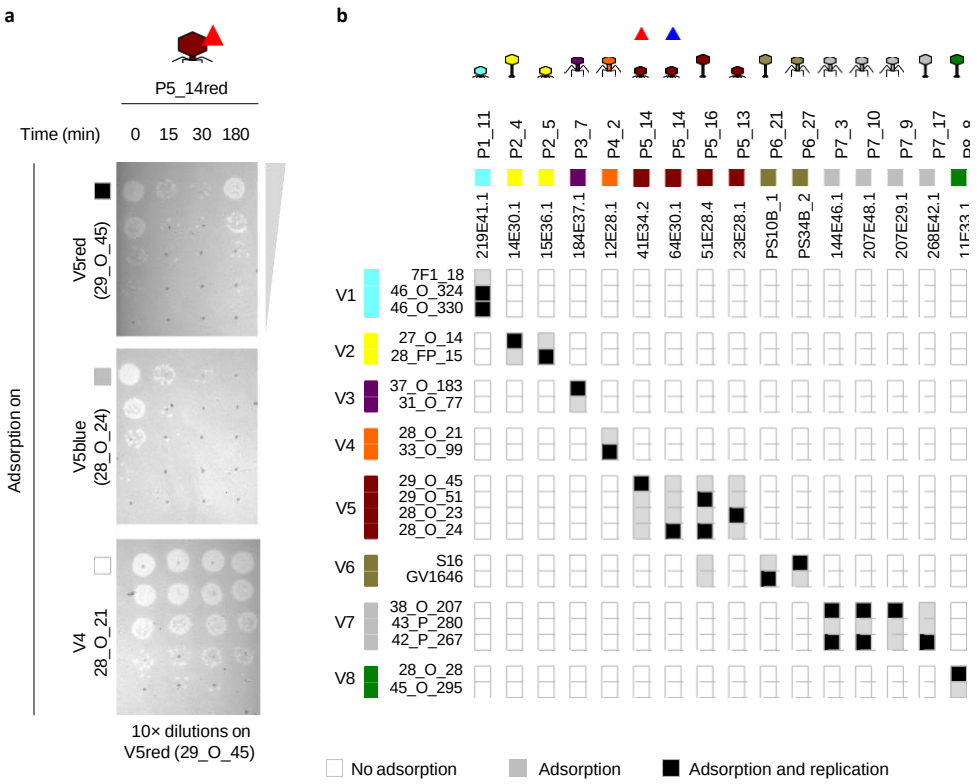


Fig. 2 | Phage adsorption and cell death are not the same. After allowing a fixed concentration of phages (MOI 0.01) to adsorb to a *V. crassostreae* strain, free phages that remained unattached were serially diluted and plated with sensitive hosts. In this assay, a drop in the number of infectious particles at 15 or 30 min indicates phage adsorption. **a**, Example of the phage 41E34.2 from cluster P5_14 tested for adsorption on vibrio 29_O_45 and 28_O_24 from clade V5, and vibrio 28_O_21 from clade V4. Serial dilutions were plated on the sensitive vibrio 29_O_45. Adsorption was complete after 30 min using the two V5 hosts. After 180 min, production of phages was observed only when using the strain 29_O_45 as host. On the basis of killing, V5 strains were divided into two groups, V5red (29_O_45) and V5blue (28_O_24). **b**, Graphic summary of adsorption and killing assays. Bacteria and phage belonging to a specific clade and cluster are arrayed in rows and columns, respectively. For phage cluster P5_14, two phages (indicated by red and blue triangles) that differ in their ability to kill the host (respectively V5red and V5blue) are shown. All 320 adsorption assays are shown in Supplementary Fig. 5.

were classified as 'resistant but impaired' if we observed a clearing zone but no production of viable phages when using high titers.

Our exploration of host susceptibility at a finer resolution led us to classify strain 7F1_18 as resistant but impaired by 8 siphoviruses from cluster P1_12 (Supplementary Fig. 7). We searched for genes present in the resistant strain (7F1_18) and absent from 7 suscep- tible strains to identify putative host genes involved in resistance against P1_12 siphoviruses. A total of 35 genes were found to be specific to the resistant strain (Supplementary Table 7), 31 of which were localized in two distinct regions (R1 and R2) that are probably MGEs (Fig. 5a). The region R1 encodes an integrase and a recom- binase. The region R2 encodes a recombinase and a relaxase, sug- gesting that it is an integrative mobilizable element. We used genetic knockouts to assess the role of these regions in phage resistance. The deletion of R1 (mutant ΔR1) was sufficient to restore full sensitiv- ity to all P1_12 siphoviruses (Fig. 5b and Supplementary Fig. 7). Single-gene inactivation further showed that only two genes encod- ing a putative NTPase (GV2333_v1_250051) and an adenosine deaminase (GV2333_v1_250052) are involved in defence against the siphoviruses. A two-gene cassette consisting of an adenosine triphosphatase and a divergent adenosine deaminase has been previously described as a phage defence system named RADAR, causing editing-dependent abortive infection in the presence of the phages[10]. However, Defence-Finder did not detect a RADAR sys- tem in the strain 7F1_18 (Supplementary Table 6) and we did not observe abortive infection when infecting this strain with the P1_12 phages (Supplementary Fig. 8). Hence, this system seems different from described RADAR systems.

Epigenetic escape of phage from host resistance

Strain 7F1_18 was 'resistant but impaired' to 3 out of 6 podoviruses from cluster P1_11 (here termed green phages), while being partially sensitive to the 3 others (termed pink phages) (Fig. 5b and Supplementary Fig. 9). The pink phages were detected after the green phages during the time series, raising the possibility that they coevolved with the hosts. To investigate this hypothesis, we first explored the role of region R1 in vibrio resistance against all P1_11 podoviruses. The deletion of R1 resulted in a full sensitivity to the pink phages and to partial sensitivity to the green phages (previously resistant). The inactivation of a gene encoding a DNA methylase (GV2333_v1_250053) or a gene encoding an ATP-dependent helicase (GV2333_v1_250054) was sufficient to alter the resistance to P1_11 phages (Fig. 5b and Supplementary Fig. 9). In contrast, the inactivation of the putative NTPase and an adenosine deaminase encoding genes had no effect, demonstrating that distinct mechanisms are involved in host defence against P1_12 and P1_11 phages. Subsequent deletion of the second region (mutant ΔR1ΔR2), and more specifically the inactivation of an R-M III system within region R2, dramatically increased sensitivity to the green phages (Fig. 5b and Supplementary Fig. 9). Hence, the green phages are neutralized by the R-M III defence system, whereas the pink ones manage to escape it.

Phage can escape from R-M by genetic or epigenetic modifications[6]. The investigation of genetic differences between pink and green phages highlighted the presence of two genes of unknown function that were exclusive to pink phages (Extended Data Fig. 4). However, these genes did not disable the R-M III when expressed

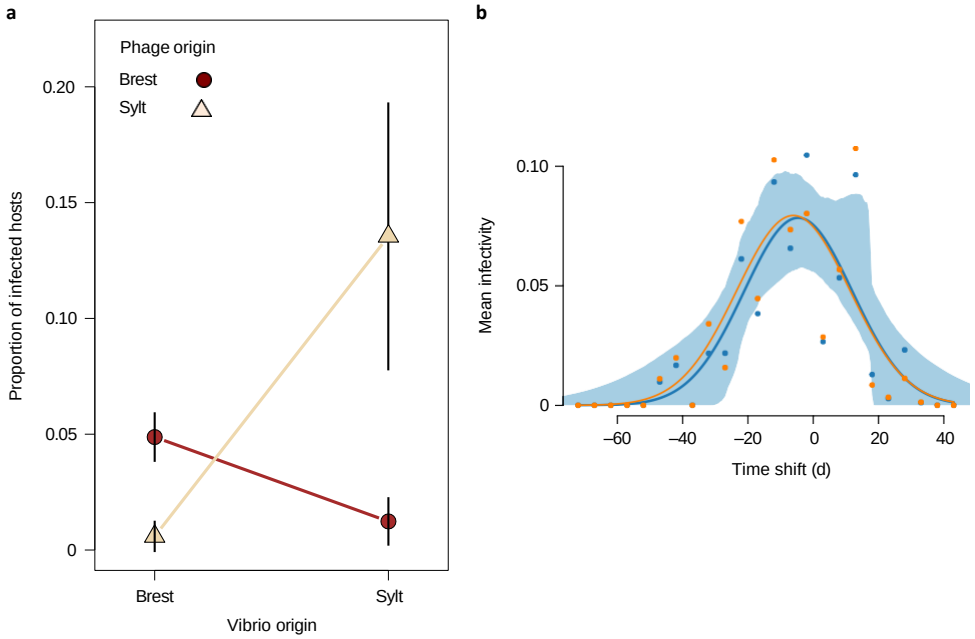


Fig. 3 | Patterns of phage adaptation across space and time. **a**, The phage is locally adapted to its vibrio host. The proportion of infected hosts is higher on sympatric than on allopatric combinations (Brest, 56 phages, 120 vibrio isolates; Sylt, 17 phages, 33 vibrio isolates) with the vibrio origin × phage origin explaining most of the variation in the data (GLMM, analysis of deviance based on Type III Wald χ^2 tests, vibrio origin $\chi^2 = 18.119$, $P = 2.1 \times 10^{-5}$, phage origin $\chi^2 = 39.552$, $P = 3.2 \times 10^{-10}$, vibrio × phage origin $\chi^2 = 161.063$, $P < 2.2 \times 10^{-16}$). The error bars represent 4 standard errors. **b**, The phage is adapted to its contemporary vibrio host in Brest. Mean infectivity (that is, the average ability of phages isolated from a sample to infect and kill bacteria) as a function of the time shift (in days) between phages and vibrios, using 1 to 15 *V. crassostreae* colonies per sampling event (median 3 colonies), and 1 to 47 phage strains (median 8 strains) in a total of 25,794 cross-inoculations. The mean infectivity for each time shift is calculated after grouping the different combinations of phage–bacteria in time-shift bins of 5 d. The blue points show the mean infectivity, with the blue curve and confidence interval (CI) (shaded area) showing the least-square fit of a smooth unimodal function (Methods). The orange points and curve refer to similar quantities calculated after excluding the phages isolated on the vibrio strains that are used for this analysis. Mean phage infectivity was found to be significantly higher on contemporary bacteria (Methods).

from a plasmid in the bacterial host (Supplementary Fig. 10). Instead, neutralization of R-M III might result from epigenetic modifications protecting the pink phage DNA from degradation by the endonuclease of the R-M system. All pink phages were produced from the same bacterial host (strain 48_O_431) and when they were propagated on another susceptible host (46_O_330), they lacked the ability to neutralize the R-M III (Fig. 5c and Extended Data Fig. 5). We conclude that the pink phages were protected from the R-M III system by DNA methylation or other epigenetic modifications. In summary, we identified three defence systems targeting distinct phage clusters (P1_11 and P1_12) among which those in RGP regions 1 and 2 (potentially corresponding to PDEs[7]) enhance the protection against P1_11 phages. Hence, immunity requires multiple defence systems encoded in MGEs that tackle different groups of phages, but it can be partly circumvented by epigenetic modifications in the phage that adapts to host defence.

Modular pattern of interaction within clade V5

To study vibrio genomes with greater ability to defend against phages, we focused on clade V5 whose strains have large genomes (average 5.7 Mbp) and more known defence systems than clade V1 (average 8.9 vs 5.6, respectively) (Fig. 4b). Accordingly, while all V1 strains from our collection were susceptible to at least one phage, 14 out of 30 strains from V5 were resistant to all tested phages (Fig. 1 and Supplementary Table 5). The interactions between vibrios from clade V5 and podoviruses from cluster P5_14 also revealed modularity: a subset of P5_14 phages (differentiated as group 'red') is only produced by a subset of V5 vibrios (thus named V5red), and another very closely related subset (group 'blue') is produced by another non-overlapping set of vibrios (V5blue) (Fig. 6a). Representatives of both phage groups adsorb to all tested V5 strains (Fig. 2), suggesting that the specificity of killing depends on the interplay between bacterial defences and phage anti-defences.

Synergistic effects between defence systems

To characterize the host mechanisms involved in defence from P5_14 phages, we focused on the V5red strains, that is, the strains from V5 clade that are killed by the red phages. We found 104 genes in 5 V5red strains that were absent from 4 V5blue strains (that is, killed by the blue phages). Among these genes, 103 clustered in 6 genomic regions (R1 to R6) (Supplementary Table 8). Deleting region R5 (Fig. 6b) in strain 29_O_45 resulted in sensitivity to all blue phages (Fig. 6d and Extended Data Fig. 6). This region encodes homologues of a two-gene phage resistance system from the retron family Ec48 (REV045_v1_590001 and 590002 in 29_O_45)[29]. However, while the production of blue phage was 10^{10} p.f.u. ml^{-1} in the sensitive V5blue host, only 10^4 p.f.u. ml^{-1} were produced using the V5red derivative lacking the retron (ΔRetron). When investigating the cause of additional resistance mechanisms explaining the differences in phage production, we found that the additional deletion of region R2 (Fig. 6b) encoding a Dnd defence system (REV045_v1_90119 to 21) increased the production of blue phage to 5.10^8 p.f.u. ml^{-1} in the V5red derivative (ΔDnd ΔRetron) (Fig. 6d and Extended Data Fig. 6). The greater effect (>10^4 more phage titre) of the double deletion than single deletion indicates that the Dnd and retron systems are synergistic. Blue phages were produced by the mutant lacking the retron and carrying the Dnd system (ΔRetron) but not by the derivative lacking Dnd and carrying the

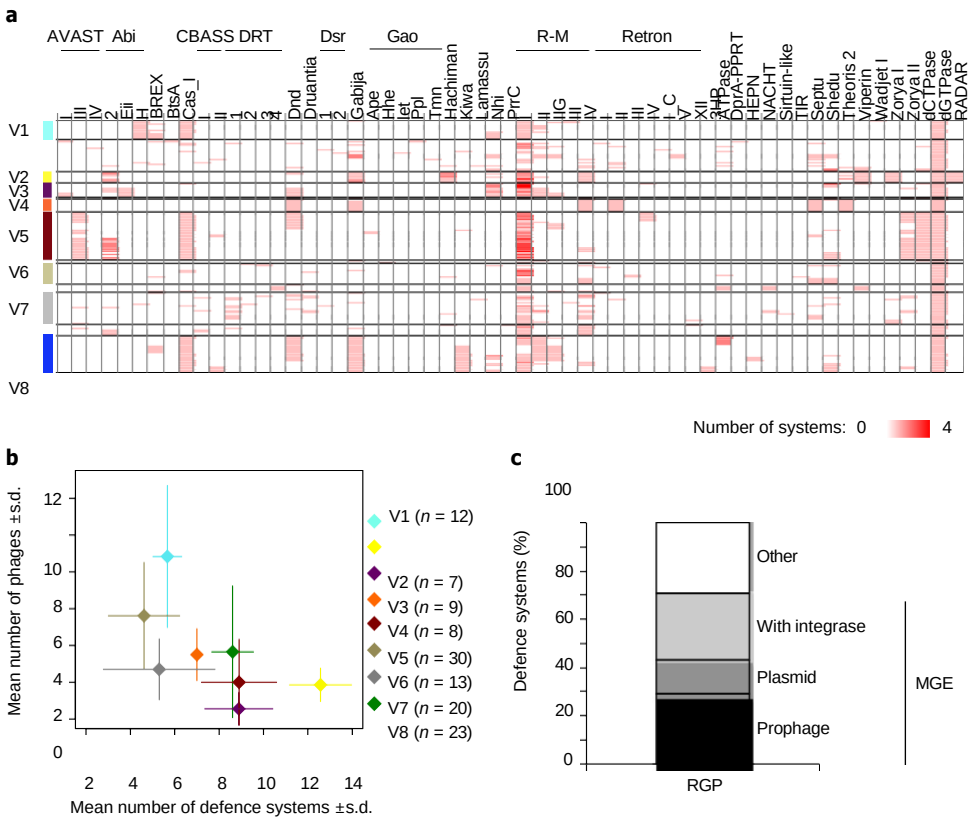


Fig. 4 | Distribution of phage-defence systems in *V. crassostreae* genomes. **a**, Heat map showing the distribution of all defence systems found by Defence-finder[24] in *V. crassostreae*. V1 to V8 indicate the bacterial phylogenetic clades as described in Fig. 1. **b**, Mean number of infecting phages versus the mean number of defence elements per strain. The means (±1 s.d.) per bacterial clade are shown as coloured diamonds and numbers next to the symbol give the sample size for each clade. **c**, Percentage of defence systems found in annotated RGPs. Putative mobile genetic elements (MGE) are classified as phage, plasmid and RGP with at least one gene encoding an integrase.

retron system (ΔDnd). This shows that the retron is more effective than the Dnd defence system in preventing the production of blue phages.

A retron from the Ec48 family has been shown to confer resistance to phage via abortive infection[29]. We thus compared the infection dynamics of the V5red lacking Dnd but carrying the retron (ΔDnd) or lacking both the Dnd and the Ec48 retron systems (ΔDnd ΔRetron) with the blue phage. At an multiplicity of infection (MOI) of 0.2 and 2, ΔDnd ΔRetron cell cultures collapsed due to phage lytic infection after 4 and 1.5 h, respectively (Extended Data Fig. 7). ΔDnd-containing cultures were not affected by the phage at an MOI of 0.2, while they entered a state of growth stasis at an MOI of 2, a hallmark of abortive infection. Such a pattern might result from a toxic effect of the retron system leading to abortive infection. Indeed, at an MOI of 0.2, only a fraction of the bacteria is exposed to phages and enters into abortive infection. This has a limited effect on population growth because after the suicide of infected cells, the bacteria are free of phages. At an MOI of 2, all bacteria are exposed to phages and undergo abortive infection, which precludes bacterial growth. In summary, among 6 genomic regions specific to V5red strains, we identified 2 synergistic systems in 2 RGP regions, flanked by integrases, providing resistance to blue phages—the Dnd defence system targeting the viral DNA and the retron leading to host suicide.

Phage counter-defences result in host shift

Considering the two systems in V5red that provide resistance to blue phages and the success of red phages in infecting these strains, we postulated that the red phages evolved to counter these systems. To gain insight into the genetic means of Dnd escape, we compared the genomes of the viruses from cluster P5_14. The gene content

and order of blue and red phages are almost identical, with the exception of a specific region where red phages carry two genes: the first one (p0021) encodes a protein of unknown function and the second one (p0020), a protein with an N-terminal (aa 4–175) phosphoadenosine phosphosulphate reductase (PAPS) domain and a C-terminal (aa 299–470) DNA *N*-6-adenine-methyltransferase (Dam) domain (Fig. 6c and Extended Data Fig. 8). A PAPS domain (aa 46–228) is also present in the sulphotransferase encoded by *dndC* in the vibrio Dnd defence system. We thus hypothesized that these two genes provide resistance to the Dnd system. To test this hypothesis, we engineered a blue phage by exchanging the four blue-specific genes found at that location by the red-specific genes (Methods). All isolated recombinants (designated Blue-PAPS) were able to infect the V5red derivatives lacking the retron, with a production of 10^{11} p.f.u. ml^{-1} (Fig. 6d and Supplementary Fig. 11). Thus, the red phage-specific genes encode an anti-Dnd system. As Blue-PAPS phage was unable to infect V5blue (Fig. 6d and Supplementary Fig. 11), the loss of four blue genes might result in maladaptation to V5 blue. This is probably not caused by a putative loss of epigenetic modifications when propagated on the V5red derivative for the genetic engineering, because blue phages propagated in V5red derivatives were still able to kill V5blue (Supplementary Fig. 12). Hence, exchanges of anti-defence genes in phages from cluster P5_14 allow them to switch from one to another set of hosts.

Next we tried to understand how red phages could overcome the retron defence system. We noticed that infection by the recombinant phage Blue-PAPS resulted in the production of phages that escape retron immunity (Fig. 6d). The production was 10^{6} p.f.u. ml^{-1} by the wild-type V5red host and 10^{8} p.f.u. ml^{-1} by the derivative lacking the Dnd but carrying the retron system (ΔDnd), confirming the synergy of Dnd and the retron systems. Importantly, escape

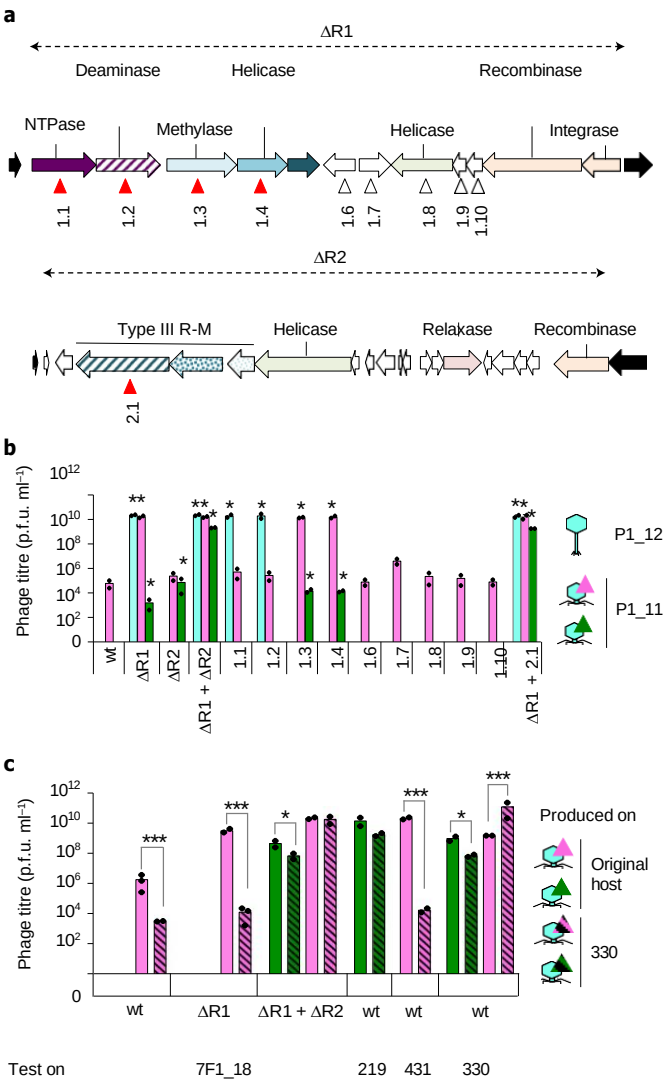


Fig. 5 | Anti-phage systems identified in *V. crassostreae* clade V1. **a**, Gene diagrams of the two defence regions (R1 and R2) specific to 7F1_18. Genes in white encoded for an unknown function. Genes in black indicate the region boundaries that have been targeted for the deletion of the entire region (Methods), leading to ΔR1, ΔR2 and the double mutant ΔR1ΔR2. Triangles indicate the integration site of a suicide plasmid by single recombination in single genes (red, modification of sensitivity; white, no phenotype). **b**, Phage titer obtained using representative phages from siphoviruses P1_12 (70E37.6), podoviruses P1_11pink (pink triangle, 431E48.2) and P1_11green (green triangle, 234P8) to infect the wild-type (wt) 7F1_18 host and its derivatives (for all phages, see Supplementary Figs. 7 and 9). All experiments were performed twice (individual dots) and showed that phage reproduction strongly depended on the specific combination of phage and gene knockout (ANOVA, two-way interaction $F_{26,42} = 205.20$, all two-sided $P < 0.001$, asterisks above bars show significant differences in separate linear models calculated for each phage in comparison with the respective wt strain). **c**, Phage titer obtained using a green and a pink phage produced in their original host (the strain used to isolate phage from seawater flocculate) or in the strain 46_O_330 (330) that is sensitive to all P1_11 phages (for all phages, see Extended Data Fig. 5). The green phage 219E41.2 originated from 38_P_219 (219) and the pink phage 431E48.2 originated from 48_O_431 (431). All experiments were performed twice (individual dots). The resulting virus titers depended on the specific combination of the strain used to produce the phage, the strain used in the infection and the phage used (three-way interaction, ANOVA $F_{1,8} = 160.263$, $P < 0.001$). The host used to produce a progeny also played a significant role when infecting the wt strain 7F1_18 and its derivatives carrying the RM-III system (ΔR1) or not (ΔR1+ΔR2) (three-way interaction production host × knockout × phage, ANOVA $F_{2,12} = 130.56$, $P < 0.001$). Asterisks above bars indicate significant pairwise differences between phages produced on original hosts and the strain 330 (*$P < 0.05$, ***$P < 0.001$, two-sided).

was not observed when the wild-type blue phage was propagated on a host lacking the Dnd system but encoding the retron system, suggesting that red phage-specific genes potentiate escape from the retron system. When isolated and further propagated on V5red, these escapees (named Blue-PAPS-RetronR) showed the same infectivity as the red phage (Fig. 6d and Supplementary Fig. 11). Spontaneous coliphage mutants that overcome the defence conferred by Ec48 retron have previously been isolated in *E. coli*[29] and they all abolished the function of RecBCD phage-encoded inhibitors (Gam in λ, gp5.9 in T7), a host complex involved in DNA repair and anti-phage activity. It was proposed that Ec48 'guards' the exonuclease RecBCD and that Ec48 activity is triggered by phage-mediated RecBCD inhibition[29]. No homologues of these RecBCD inhibitors were identified in the red phage genomes, suggesting that the red phage neutralizes RecBCD in another way. Alternatively, the retron from V5red may not guard RecBCD. To gain further insight into genetic determinism of overcoming retron in P5_14 podoviruses, we sequenced the genome of eight laboratory-generated Blue-PAPS-RetronR phages. This sequencing allowed us to identify several mutations (Extended Data Fig. 9), all of which were localized in an exonuclease (p0030 in 66E30.1, 185 aa). Half of the escapees gained a stop mutation leading to truncated proteins at position 43, 49, 113 and 142. The other half of the escapees acquired missense mutations at position 21, 27 and 138 (Supplementary Fig. 13). The comparison of wild-type blue and red phages also reveals divergence in this exonuclease (5 amino acid substitutions along the protein and 4 additional amino acids in the C terminal of blue phages, Supplementary Fig. 13), which may explain their different sensitivity to the retron. Collectively, these data suggest that modifying an exonuclease enabled a natural population of phages to neutralize the retron defence by a mechanism that remains to be elucidated.

Discussion

In this study, we determined the genetic mechanisms driving specificity of phage for *V. crassostreae* at different initial stages of phage

infection (Extended Data Fig. 10). We identified two-step protection from phage infection, involving receptor adaptations and host defence systems. The patterns of phage adsorption matched the bacterial clades within the *V. crassostreae* species, suggesting the existence of clade-specific receptor(s) and cluster-specific receptor-binding protein(s).

In the future, identification of the receptor(s) for each phage cluster will enable an understanding of evolutionary selection processes. In wild vibrio populations, it was recently reported that receptor alleles swept through host populations[7]. This is consistent with the expectation that evolution of surface receptors results from a trade-off between physiological relevance for bacteria in the natural environment (including virulence for oysters) and the cost they induce under phage predation[30,31].

After adsorption and entry, many phages cannot complete the lytic cycle owing to cellular defence systems. Our results are consistent with a previously described major role for MGEs in the emergence of bacteria resistant to phages[7,8]. Notably, differences in phage defence gene repertoires among closely related bacterial strains are found in regions of genome plasticity, mainly in MGEs. This may explain the positive correlation between the number of defence genes and genome size, since genome size increases upon acquisition of MGEs. The number of defence systems is also correlated with resistance to phages and might explain why genomes with large size differences can coexist within natural bacterial populations in the wild[32]. The presence of defence systems in MGEs facilitates rapid acquisition of defence systems to provide resistance against multiple phages. Further work will be needed to understand the cost of these defence systems, for example, on cell growth and in precluding the acquisition of novel function by horizontal gene transfer[33], which could in certain conditions select for loss of defence

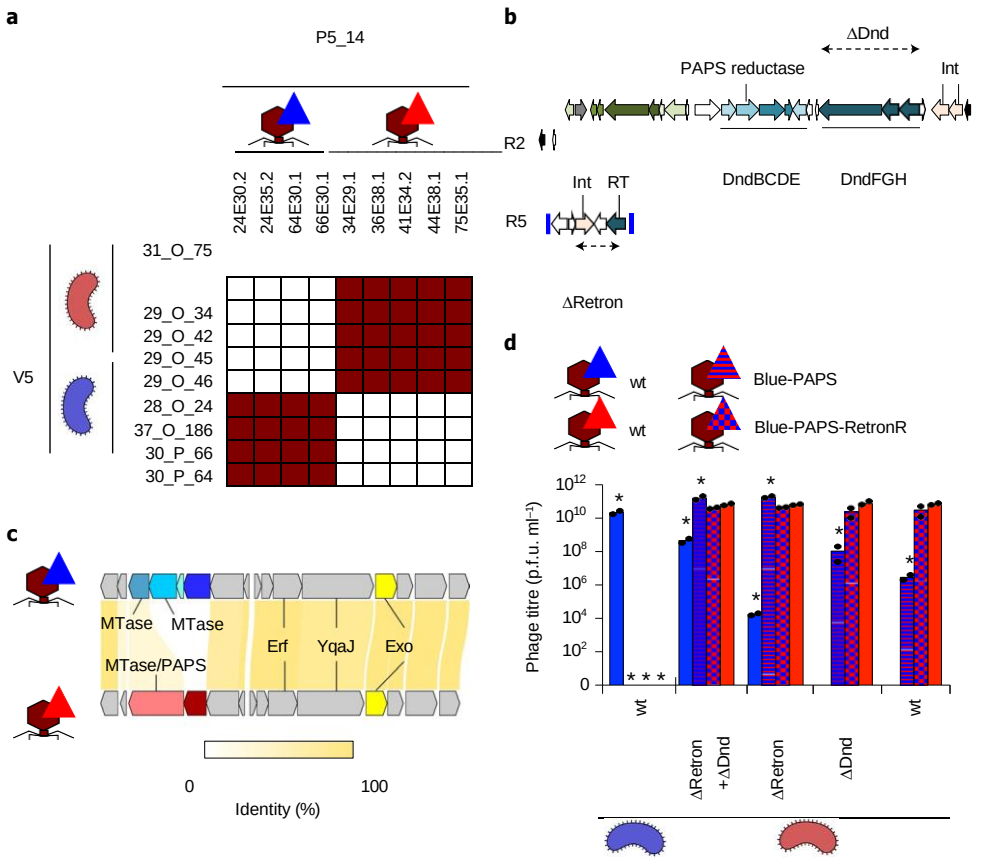


Fig. 6 | Bacterial defence and phage counter-defence. **a**, Modularity of the interactions between vibrios from clade V5 and phages from cluster P5_14. Brown squares indicate vibrio sensitivity to killing by the phage. A subset of P5_14 phages (red triangle) only reproduces in a subset of V5 vibrios (red bacteria), and another very closely related subset (blue triangle) in another non-overlapping set of vibrios (blue bacteria). **b**, Gene diagrams of the two regions specific to the vibrio V5red strain 29_0_45 and demonstrated to be involved in resistance to phage P5_14 blue. Blue lines indicate the end of a contig. Deletions include DndFGH (ΔDnd) and the complete retron Ec48 (ΔRetron) systems for regions R2 and R5, respectively. RT, Reverse Transcriptase; Int, Integrase. **c**, Only one syntenic region consisting of two (in red, p0020 and p0021 in phage 44E38.1) and four (in blue, p0020 to p0023 in phage 66E30.1) genes differed between P5_14 red and blue, while the flanking genes encoded conserved core proteins (grey, see also Extended Data Fig. 8). Genes encoding putative recombinases (Erf, YqaJ) and methylase (MTase) were identified using dedicated Hidden Markov Model (HMM) profiles. The exonuclease involved in retron escape is indicated in yellow. **d**, Phage titres obtained using wt and mutant P5_14 phages to infect wt V5 strains and V5red derivatives. The Blue-PAPS phage mutant is a blue phage in which we exchanged the four blue-specific genes with the two red-specific genes (including the PAPS-domain protein). Blue-PAPS-RetronR phage is a retron escapee obtained when we propagated Blue-PAPS phage on V5red (for all phages, see Supplementary Fig. 11). All experiments were performed twice (individual dots) and showed that phage reproduction depended on the combination of phage and host derivates (ANOVA, phage × gene knockout interaction $F_{12,20} = 801.49$, $P < 0.001$). Asterisks show significant differences in separate linear models calculated for each phage derivate in comparison to the respective wt V5red (all two-sided $P < 0.001$, except for ΔDnd × Blue-PAPS where $P = 0.007$).

systems. The smallest genomes in our datasets have fewer defence systems and are more sensitive to phage. We speculate that if phage abundance is low, those bacteria with smaller genomes may have a fitness advantage.

Our results reveal striking similarities in the evolution of bacterial defences and phage counter-defences. First, bacteria make extensive use of epigenetic modifications for protection, as revealed by the abundance of R-M and Bacteriophage Exclusion (BREX) systems. Phages overcome host defences through epigenetic escape from restriction modification and exchange of genes encoding anti-restriction proteins. Where patterns differ is that changes in genome size are tolerated by bacterial genomes, which can increase by hundreds of genes in single events of horizontal gene transfer (HGT), whereas genome size is constrained in phages by the capsid, probably limiting the number of escape genes a phage can encode. As a result, changes in the patterns of phage counter-defence among closely related phages might be caused by gene replacements instead of gene additions. As the exchange of anti-defence genes can lead to mismatches with host defence and a switch in phage host, the outcome of the interplay between defence and counter-defence is

a key determinant of phage host-range. The study of other interaction modules, for example understanding why red phages are not capable of infecting V5blue and how blue phages counter the V5blue defence systems, will allow us to link these molecular mechanisms with coevolutionary dynamics in nature. Second, the analysis of phage encoding counter-defences suggests that evolution is mediated by genetic exchanges between phages, and possibly other MGEs, by recombination. Accordingly, specific recombination systems have been identified in numerous phage genomes[19] and could play a role in overcoming bacterial anti-phage defences[17]. This suggests that gene variation in phages mirrors the turnover of MGEs encoding for resistance in the hosts. Although bacteria acquire novel systems from a vast pool of defence systems present in other bacteria[11], phages use recombination to exchange anti-defence mechanisms with other phages, and possibly other MGEs, that co-infect the same bacterial host.

In conclusion, we show that phages contain a pangenome of counter-defence systems that they frequently exchange. This results in variations matching those of bacterial defences in a continuous process of antagonistic coevolution.

Methods

Sampling. Samples were collected from an oyster farm located at the Bay of Brest (Pointe du Château, 48° 20′ 6.19′′ N, 4° 19′ 6.37′′ W), every Monday, Wednesday and Friday from 3 May to 11 September 2017. Specific pathogen-free juvenile oysters[34] were deployed in the field in batches of 100 animals. When the first mortalities were observed in the first batch, another batch of specific pathogen-free animals was placed in the field, leading to the consecutive deployment of seven batches from 26 April to 11 September. Oyster mortalities were recorded on each sampling day. Seawater temperature reached 16 °C on 22 May, a previously observed threshold for oyster mortalities[35]. Mortalities began on 29 May and persisted until 25 August. On each sampling date, five living oysters were collected from a batch showing less than 50% mortality. We previously showed that the relative abundance of *V. crassostreae* (number of *V. crassostreae* isolates out of total vibrios) is similar among diseased animals[14]. The animals were cleaned, shucked, weighed and 2216 Marine Broth (MB) was added (10 mg ml^{-1}) for homogenization using an ultra-turrax. A volume of 100 µl homogenate was used for vibrio isolation, the remaining volume was centrifuged (10 min,17,000 *g*), the supernatant filtered through a 0.2 µm filter and stored at 4 °C until the phage isolation stage. Two litres of seawater were collected and size fractionated as previously described[14]. 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 1,000-fold concentrated viral samples from 2 l passaged through a 0.2 µm filter, following a previously described protocol[36]. Virus flocculates were suspended in 2 ml 0.1 M EDTA, 0.2 M $MgCl_2$, 0.2 M oxalate buffer at pH 6 and stored at 4 °C until the phage isolation stage.

***Vibrio crassostreae* isolation, identification and genome analysis.** *Isolation and identifications.* 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. *V. crassostreae* isolates were first identified by PCR using a primer set (Eurofins Genomics) targeting the *r5.2* gene (previously identified as population-specific marker[15]) (Supplementary Table 9) and colonies as template. PCR-positive isolates were grown in MB and stored at −80 °C in 10% dimethyl sulfoxide (DMSO). Taxonomic assignment was further refined by *gyrB* gene sequencing[15]. Bacteria were grown overnight in MB and DNA extracted using an extraction kit (Wizard, Promega) according to the manufacturer's instructions. The partial *gyrB* gene was amplified using degenerate primers (Supplementary Table 9), Sanger sequenced (Macrogen) and manually corrected with the chromatogram. Sequences were aligned with Muscle and phylogenetic reconstruction was done with the RAxML v8 GTR model of evolution, a gamma model and default parameters[37]. Sequencing of the *gyrB* gene confirmed that 195 isolates were *V. crassostreae* (Supplementary Table 1).

Genome sequencing, assembly and annotation. DNAs were sequenced by the Joint Genome Institute using 300 bp library and HiSeq2000 Illumina sequencing technology or at the BioMicro Center at MIT using Nextera FLEX for library and NextSeq 80PE for sequencing. De novo reads assembly was performed using Spades v3.11[38] in careful mode and with three kmer sizes of 33;55;77. Computational prediction of coding sequences and functional assignments were performed using the automated annotation pipeline implemented in the MicroScope platform[39]. Annotation of phage defence genes was performed with Defence-Finder v0.0.11[24] using default options. The databases used for the search were downloaded on 9 August 2021.

Core genome phylogeny. The proteome of each isolate was compared by performing a Blastp all-vs-all. Silix v1.2.9[40] was used to reconstruct protein families on the basis of 80% of reciprocal length of alignment and an identity of at least 80% for *V. crassostreae*. Protein sequences of each family were first aligned with Muscle v3.8.31 using default parameters, filtered using Gblocks v0.91b with relaxed parameters[41] (maximum number of contiguous non-conserved positions, 10; minimum length of a block, 5; allowed gap positions, half) and concatenated. Phylogenetic reconstruction was done using RAxML v8[37] on this concatemer using an LG model of evolution, a gamma model and default parameters.

Comparative genomics. The ANI value of genomes was determined using pyani v0.2.11 (https://github.com/widdowquinn/pyani). The phylogenetic profile method implemented in the MicroScope platform[39] was used to identify putative phage resistance genes and regions. For this, we searched for genes present in all strains resistant to phage (80% identities on 80% coverage) and absent from sensitive strains. The same approach was used to estimate specific genes of vibrios in pairwise genome comparisons.

Genomic islands identification. RGPs were identified with PPanGGOLIN v1.1.136[42] using the module panRGP[27]. Default options were used to identify RGPs with a minimum of 3 kbp.

Plasmid identification. Identification of plasmids was first based on the search for conjugative systems with macsyfinder v2.0[43] using the CONJScan module[44]. Complete and incomplete systems were searched with default options on each RGP. In complement, we used RFPlasmid v0.0.18[45] on total genomic sequences to classify each contig as plasmid- or chromosome-like on the basis of hallmark genes and pentamer frequencies. A random forest classifier attributes to each contig votes ranging from 0–1, allowing discrimination between plasmid and chromosomal contigs. To conservatively identify plasmids, we defined rules that used both macsyfinder and RFPlasmid analyses. RGPs were assigned as plasmid if (1) the RFPlasmid votes for plasmid were above 0.7; (2) the RFPlasmid votes for plasmid were above 0.5 and contained either complete or incomplete conjugative system; and (3) the entire contig was defined as an RGP to discard potential integrative and conjugative elements.

Prophage identification and characterization. Identification of prophages from RGPs in vibrio genomes was done using CheckV v0.8.1[46] and the following criteria. First, RGPs assigned as 'low', 'medium' or 'high' quality on the basis of their completeness were selected as candidates. To filter out false positives, we discarded RGPs with no viral genes and more than one host gene called, as advised by the designers of CheckV. We observed that plasmids are sometimes assigned as prophages by CheckV, so candidate RGPs with overly large size (>50 kbp) were checked for the presence of potential plasmid hallmark genes on the basis of the MicroScope platform functional annotation (*repB*, *parA*, *parB*) and conjugative systems on the basis of CONJScan results. If a conjugative system or at least two genes implicated in replication and/or partition of plasmids were found, the candidate RGP was not considered as a phage.

Phage isolation, identification and genome analysis. *Sympatric and allopatric sampling strategy*. First, we used the 195 *V. crassostreae* strains as 'bait' to isolate phages from 20 ml seawater equivalents of viral concentrate (1,000×) or oyster tissues (0.1 mg) collected on the same day (Supplementary Fig. 1). Phage infection was assessed by plaque formation in soft agar overlays of host lawns mixed with a viral source. This approach yielded 45 phages from 18 of 195 tested hosts (9.2%). Second, 90 *V. crassostreae* isolated from 12 June to 28 July were screened for phages using: (1) 10 pooled seawater viral concentrates from 5 consecutive dates, and (2) 20 time-shift combinations of single seawater viral concentrates. Each approach resulted in the isolation of 21 and 177 additional phages from 5/90 (5%) and 38/90 (42%) plaque-positive hosts, respectively, in total resulting in a collection of 243 phages from Brest (Supplementary Table 2). This collection was complemented with 51 bacteria and 31 phages previously isolated in Sylt (Germany) during a 2 d sampling campaign.

Isolation and generation of high-titre stocks. We used previously described methods[36] to isolate and generate high-titre stocks. Briefly, isolation of phages was performed by directly plating on a bottom agar plate (1.5% agar, in MB) 100 µl of an overnight bacterial culture, 20 µl of seawater flocculate (equivalent to 20 ml of seawater containing viruses) or 20 µl of oyster homogenate and 3.5 ml molten top agar (55 °C, 0.4% agar, in MB) to form host lawns in overlay and allow for plaque formation. After incubation for 48 h at room temperature (r.t.), a maximum of six plaques per morphotype was archived. Plaque plugs were eluted in 500 µl MB for 24 h at 4 °C, 0.2 µm filtered to remove bacteria, re-isolated three times on the sensitive host, stored at 4°C, and after supplementation with 25% glycerol at −80 °C and stored at 4 °C. High-titre stocks (>10^9 p.f.u. ml^{-1}) were generated by confluent lysis in agar overlays.

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). Imaging was performed using a Jeol JEM-1400 transmission electron microscope equipped with an Orious Gatan camera at the platform MERIMAGE (Station Biologique).

DNA extraction, sequencing, assembly and annotation. Phage suspensions (10 ml, >10^{10} p.f.u. ml^{-1}) were concentrated to approximately 500 µl 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 2.5 µl RNAse (Macherey-Nagel) at 1,000 units and 3.5 mg ml^{-1}, respectively. The nucleases were inactivated by adding EDTA (20 mM, pH 8). DNA extractions were performed using the MasterPure Complete DNA and RNA purification kit (Epicentre) according to the manufacturer's instructions. Alternatively, DNA was extracted following a previously described protocol[36]. Briefly, DNA extraction encompassed a first step of protein lysis (0.02 M EDTA pH 8.0, 0.5 mg ml^{-1} proteinase K, 0.5% sodium dodecyl sulfate) for 30 min incubation at 55 °C, a phenol chloroform extraction and an ethanol precipitation. DNA was visualized by agarose gel electrophoresis (0.7% agarose, 50 V, overnight at 4 °C) and quantified using QuBit.

Phages were sequenced by the Biomics platform at the Pasteur Institute. The sequencing library was prepared using a TruSeq Illumina kit, 2 × 150 bp paired end and sequenced on a NextSeq550 Illumina sequencer. After reads trimming with Trimmomatic v0.39 (LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:36)[47], de novo assembly was performed using SPAdes v3.15.2 (–careful–cov-cutoff auto -k 21,33,55,77 -m 10)[48]. Vector contamination contigs were excluded using the UniVec Database (build 10.0) and one-contig phage genome

was manually decircularized. The large terminase subunit was identified using DIAMOND v2.0.8.146 (blastp, default parameters)[49] on previous annotated vibrio large terminase and using HMMSCAN (HMMER v3.3.2, default parameters with *e*-value ≤ 10^{-3}) with PFAM profiles (PF03354.17, PF04466.15, PF05876.14, PF06056.14)[50]. The phage genome was manually ordered starting from the gene for the large terminase subunit.

Syntaxic annotation was done with PHANOTATE v1.5.0 (https://github.com/deprekate/PHANOTATE). Transfer RNAs were identified with tRNAscan-SE v2.0.9[51]. The functional annotation used multiple approaches. First, we used HMMSCAN (*e*-value ≤ 10^{-3}) with recombinase (PF00154.23, PF00239.23, PF00589.24, PF02920.17, PF03930.16, PF05772.14, PF09588.12, PF10136.11), methylase (PF00145.19, PF01555.20, PF02086.17, PF02384.18, PF05175.16, PF05869.13, PF10237.11), phosphoadenosine phosphosulfate reductase (PAPS, PF01507.21), VirFam[52](08/05/21), PVOGS profiles[53](09/22/16) and Gam (PF06064.13, PF07352.14). We also used DIAMOND blastP against the Nahant collection genomes (30% identities, 50% coverage)[36] combined with InterProScan v5.52-86.0 (*e*-value ≤ 10^{-3}) and eggNOG v5.0.2 (30% identities, 50% coverage) default results[54,55].

We clustered phages using VIRIDIC v1.0r3.6 (default parameters)[18]. Intergenomic similarities were identified using BLASTN pairwise comparisons. Virus assignment into genera (≥70% similarities) and species (≥95% similarities) ranks follows the International Committee on Taxonomy of Viruses (ICTV) genome identity thresholds. For each phage clusters, the ANI that corresponds to the orthologous gene pairs mean identity was determined using fastANI v1.32 (–fragLen=100)[56].

Comparative genomics used MMseqs2 v13-45111 for protein clustering[57] and IQ-TREE v2.1.2 for phylogenetic tree construction[58]. We used Snippy v4.6.0 to identify retron-escaping mutations from assembled contigs (default parameters) [https://github.com/tseemann/snippy].

Identification of temperate phages. Phage integration into the bacterial genome was searched using BLASTn v2.9.0 between the phage and vibrio genomes in our collection. No phage integration was found (*e*-value ≤ 0.01). We also searched for genes associated with lysogeny, that is, integrase (IPR002104) and replicative transposase (IPR004189), with InterProScan v5.52-86.0 and no such gene was found for any phage (*e*-value ≤ 0.01)[7].

Host-range determination. *Single-phage-by-single-host host-range infection assay*. Host-range assays were carried out using a robot hosted at Eligo Bioscience (Paris, France) or manually using an electronic multichannel pipette by spotting 5 µl of the phage suspension normalized at 2×10^5 p.f.u. ml^{-1} (10^3 p.f.u. per spot) on the agar overlay inoculated with the tested host. Plates were incubated overnight at r.t. and plaque formation was observed after 24 h. Spot assays were performed in duplicate and positive interactions (observation of a clearing zone) were confirmed in a third experiment.

Classification of host sensitivity. To explore the sensitivity of bacteria, 10-fold serial dilutions of phages from a high-titre stock (>10^9 p.f.u. ml^{-1}) were prepared and 5 µl drop spots of each dilution were pipetted onto bacterial host lawns. For some spot tests, clearing zones (sometimes turbid) were observed for the highest concentrations of phage lysates. To determine whether the bacterial host was sensitive, partially sensitive or insensitive but impaired, we explored the titre of the phage on a given bacteria compared to the maximum titre observed (that is, with the host used to produce the phage). A total of 5 µl of serial phage dilutions was mixed with 100 µl of an overnight host culture and 2.5 ml top agar to form host lawns in overlay and plaques were counted after 24 h. In sensitive and partially sensitive hosts, plaques were obtained using 1–10 and 10^4–10^5 p.f.u., respectively. In resistant but impaired hosts, no plaque was observed using up to 10^7 p.f.u. To explore whether the 'resistant but impaired' phenotype results from contaminating temperate phage(s) that lysogenize the test strain, the vibrios used to produce the phages were grown overnight in MB. After centrifugation at 3,894 *g* for 10 min, the supernatant was filtered on 0.2 µm and 10-fold dilutions of the supernatant (or the phages as control) were plated on the strain defined as 'resistant but impaired'. We never observed any clearing or turbid zone using the supernatant.

Adsorption estimation. Phage adsorption experiments were performed as previously described[59]. Phages were mixed with exponentially growing cells (optical density (OD) 0.3; 10^7 c.f.u. ml^{-1}) at an MOI of 0.01 and incubated at r.t. 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 17,000 *g* 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 (Fig. 2a) indicates bacteriophage adsorption.

Temporal adaptation analysis. To characterize coevolutionary dynamics between *V. crassostreae* and its phages, we examined how phage infectivity varied with the time shift between *V. crassostreae* and phage isolates[20,21]. This was done for 1 to 15 *V. crassostreae* colonies per sampling day (median 3 colonies), and 1 to 47 phage strains (median 8 strains) for a total of 25,974 cross-inoculations. Infectivity was coded as a binary variable associated with the phage ability to infect (or not) the bacterial isolate. The mean infectivity was averaged over all combinations of phages sampled at time t and bacteria sampled at time $t + \tau$, where τ measures the time shift in days (that is, the time difference between the sampling time of the bacteria and the sampling time of the virus). The mean infectivity of the phage peaked for contemporary associations and declined as phages were inoculated on bacterial strains from the past or from the future (Fig. 3b and Supplementary Fig. 6). This pattern is characteristic of fluctuating selection dynamics, whereby phage populations are maximally adapted to bacterial populations from the recent past or from contemporary time points. To better visualize the pattern, we fitted a smooth unimodal function (of the form $c\left(1 + \mathrm{erf}\left[\frac{\sqrt{2}\sigma}{\alpha(\tau-\mu)}\right]\right)\exp\left[\frac{2\sigma^2}{-(\tau-\mu)^2}\right]$) using the

least-squares method (nls function of the R software v4.2.1), where erf is the error function, τ is time shift and c, α, σ, μ are the four parameters of the function. The pattern was unchanged when removing the phages isolated on bacterial strains used in this analysis (Fig. 3b). As an attempt to test the significance of this pattern, we compared the difference in infectivity of contemporary vs non-contemporary combinations. In this test, 'contemporary combinations' were defined as those where the phages were isolated 10 d before to 10 d after the bacteria. Phages could infect contemporary bacteria in 789/12,045 combinations (frequency: 0.0655) and non-contemporary bacteria in 610/13,929 combinations (frequency: 0.0437). The difference was significant according to a test based on binomial probabilities ($P < 0.001$). This pattern remained significant after removing phages isolated after amplification on bacteria used in this analysis: infection occurred in 595/9,801 (frequency 0.0607) contemporaneous combinations and 540/11,259 (frequency: 0.0479) non-contemporaneous combinations ($P < 0.001$).

We used another complementary approach to explore the correspondence in the temporal fluctuations of the composition of the phage and vibrio populations. We modelled the number of isolated phages that were found to belong to cluster *i* at time *t* as a Poisson distributed number with a rate given by a linear function of the frequency of matching-clade bacteria (that is, clade *i*) and frequency of non-matching-clade bacteria (that is, clade $j \neq i$). The number of phages is the number of plaque-forming units identified on contemporary vibrios. The frequency of bacterial clades was estimated from the 48 vibrio colonies isolated at each time point. The maximum-likelihood model indicates that a 1% increase in frequency of a *V. crassostreae* clade translated to 0.27 additional phages of the corresponding cluster, compared with only 0.03 additional phages of a different cluster; this model was highly supported compared with a model with no distinction between clades (likelihood ratio test, $P < 0.001$), showing a significant temporal association between *V. crassostreae* clades and corresponding phage clusters.

Overall, these two approaches concur that the temporal fluctuations in the composition of the phage population are driven by the fluctuations in the composition of the vibrio population. Whether these fluctuations in vibrios are, in turn, solely driven by coevolution (that is, negative-frequency-dependent selection) or by another fluctuating environmental factor (for example, competition with other species, other predators, variations in the abiotic environmental factors and so on) warrants further studies.

Statistical analyses. Spatial specificity of phage–vibrio infection success (Fig. 3a) was investigated using a binomial generalized linear mixed model (GLMM), fitting the origin of phages and vibrio as main effects and phage and vibrio strain identity as random effect to account for non-independence between single assays. Analysis of deviance based on Type III Wald χ^2 tests was used to summarize the results. The correlation between the number of infecting phages, genome size and the number of defence elements was calculated using non-parametric Spearman's rank correlation. To remove phylogenetic dependence between closely related strains, we calculated phylogenetic independent contrasts on the basis of the maximum-likelihood tree obtained from the core phylogeny using the *ade4* package (Fig. 4b). Effects of genetic knockout (Figs. 5 and 6) were analysed as linear models, fitting the phage and gene knockout plus their interaction as factors, and results were summarized by analysis of variance (ANOVA). To test for the effect of the host used to produce the phages, we fitted two separate models.
(1) To test differences between phage production on the original host versus the tester host (strain 46_O_330, or 330) in the original and the tester strain, we fitted a linear model with log-transformed virus titre containing production host (original vs 330), test host (original vs 330) and phage (green vs pink) with all possible interactions and analysed the model by ANOVA. (2) To test differences between phage production on the original host vs the tester host in the 7F1_18 knockouts, we fitted a linear model containing production host (original vs 330), gene knockout (wt vs ΔR1, ΔR1+ΔR2) and phage (green vs pink) with all possible interactions and analysed the model by ANOVA. Full models were complemented by Tukey posthoc analyses focusing on the pairwise differences between production hosts (original vs tester) to show epigenetic effects induced by the production host.

Molecular microbiology. *Strains and plasmids*. All plasmids and strains used or constructed in the present study are described in Supplementary Table 10 and Fig. 11. *V. crassostreae* isolates were grown in Luria-Bertani (LB) or LB-agar + 0.5 M NaCl at r.t. *Escherichia coli* strains were grown in LB or on LB-agar at 37 °C. Chloramphenicol (Cm; 5 or 25 $\mu g\,ml^{-1}$ for *V. crassostreae* and *E. coli*, respectively), thymidine (0.3 mM) and diaminopimelate (0.3 mM) were added as supplements

when necessary. Induction of the arabinose-inducible and glucose-repressible promoter, P_{BAD}, was achieved by the addition of 0.2% L-arabinose to the growth media, while repression was achieved by the addition of 1% D-glucose. Conjugation between *E. coli* and vibrios was performed at 30 °C as described previously[60], with the modification of using TSB-2 (tryptic soy broth supplemented with 1.5% NaCl) instead of LB for mating and selection.

Clonings. All clonings in pSW7848T were 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. Before cloning in pSW23T, a PCR fragment was amplified using Go*Taq* DNA polymerase (Promega) and subcloned in a TOPO cloning vector (Invitrogen). The plasmid miniprep (Wizard Plus SV Minipreps DNA purification system, Promega) was digested with EcoR1 (NEB) and the insert was cloned in pSW23T. The phage-specific region was amplified using the Herculase II, digested with Apa1 and Xba1, and cloned in pSU18T-P_{BAD} instead of the *gfp* gene. All clonings were first confirmed by digesting plasmid minipreps with specific restriction enzymes and then by sequencing the insert (Macrogen).

Vibrio mutagenesis. Gene inactivation was performed by cloning an internal region of the target gene in the suicide plasmid pSW23T[61]. After conjugative transfer, selection of the plasmid-borne drug marker (Cm^R) resulted from integration of pSW23T in the target gene by a single crossing-over. Region deletion was performed by cloning 500 bp fragments flanking the region in the pSW7848T suicide plasmid[62]. This pSW23T derivative vector encodes the *ccdB* toxin gene under the control of the P_{BAD} promoter[60]. 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. P5_14blue phage was engineered using double crossing-over with a plasmid carrying regions of homology (438 and 156 bp) to the phage genome P5_14red. A 3,745 bp region of the phage P5_14red (44E38.1) was amplified by PCR and cloned in a replicative plasmid pSU18T-P_{BAD} (P15A *oriV;* Cm^R) under the control of the conditional P_{BAD} promoter. Selection of the transformants on Cm + glucose (1%) prevented the expression of toxic phage genes. This plasmid was transferred by conjugation to a V5red strain (29_O_45) derivative lacking the retron (Δretron). Plate lysates were generated by mixing 500 µl of an overnight culture of the transconjugant with the P5_14blue phage (66E30.1) and plating in 7.5 ml agar overlay. After the development of a confluent lysis of lawns, the lysate was collected by addition of 10 ml MB, shredding of the agar overlay and storing overnight at 4 °C for diffusion of phage particles. The lysates were next centrifuged at 3,894 *g* for 10 min, the supernatant filtered through 0.2 µm filter and stored at 4 °C. Recombinant phages were enriched by infecting the Δretron mutant in agar overlays. Recombinant phages were screened by PCR using a primer set targeting the P5_14red-specific gene (PODOV008_V2_p0019 in 44E30.1) and single plaque as template. The recombination was further confirmed by sequencing genes that are polymorphic between P5_14red and P5_14blue phages.

To isolate mutant phages that escape Ec48 defence, P5_14red or blue-PAPS phages were plated on wild-type V5red or derivative ΔDnd using the double-layer plaque assay. Plaques were obtained using only the blue-PAPS as viral source and 8 single plaques were picked for re-isolation. These phages were produced at high titre, DNA extracted and genomes sequenced. Reads were aligned to the ancestor genome.

Abortive infection assay. Abortive infection experiments were performed as previously described[29]. Overnight cultures of bacteria were diluted 1:100 in LB NaCl or MB and incubated at room temperature with agitation. Bacterial cultures (90 µl) were transferred into wells in a 96-well plate (Cellstar 655 180) containing 10 µl of phage for diverse MOIs or 10 µl MB (negative control). Plates were incubated at room temperature with shaking in a FLUOstar Omega plate reader and OD_{600} was measured every 30 min. Infections were performed in triplicate.

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

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequenced genomes have been deposited under the NCBI BioProject, with accession numbers PRJEB5876 to PRJEB5880, PRJEB5882 to PRJEB5885, PRJNA499864 to PRJNA499870, PRJNA500024 to PRJNA500069, PRJNA538125 to PRJNA538127, PRJNA548064 and PRJNA712984 for vibrios; MW824369 to MW824434, MW865291 to MW865292 and MW865297 to MW865379 for phages. Source data are provided with this paper.

References

1. Bernheim, A. & Sorek, R. Viruses cooperate to defeat bacteria. *Nature* **559**, 482–484 (2018).
2. Koonin, E. V., Senkevich, T. G. & Dolja, V. V. The ancient Virus World and evolution of cells. *Biol. Direct* **1**, 29 (2006).
3. Breitbart, M., Bonnain, C., Malki, K. & Sawaya, N. A. Phage puppet masters of the marine microbial realm. *Nat. Microbiol.* **3**, 754–766 (2018).
4. Chevallereau, A., Pons, B. J., van Houte, S. & Westra, E. R. Interactions between bacterial and phage communities in natural environments. *Nat. Rev. Microbiol*. https://doi.org/10.1038/s41579-021-00602-y (2021).
5. Koskella, B. & Brockhurst, M. A. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **38**, 916–931 (2014).
6. Hampton, H. G., Watson, B. N. J. & Fineran, P. C. The arms race between bacteria and their phage foes. *Nature* **577**, 327–336 (2020).
7. Hussain, F. A. et al. Rapid evolutionary turnover of mobile genetic elements drives bacterial resistance to phages. *Science* **374**, 488–492 (2021).
8. LeGault, K. N. et al. Temporal shifts in antibiotic resistance elements govern phage-pathogen conflicts. *Science* https://doi.org/10.1126/science.abg2166 (2021).
9. Doron, S. et al. Systematic discovery of antiphage defence systems in the microbial pangenome. *Science* https://doi.org/10.1126/science.aar4120 (2018).
10. Gao, L. et al. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077–1084 (2020).
11. Bernheim, A. & Sorek, R. The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.* **18**, 113–119 (2020).
12. Nurmemmedov, E., Castelnovo, M., Medina, E., Catalano, C. E. & Evilevitch, A. Challenging packaging limits and infectivity of phage lambda. *J. Mol. Biol.* **415**, 263–273 (2012).
13. Hendrix, R. W. Hot new virus, deep connections. *Proc. Natl Acad. Sci. USA* **101**, 7495–7496 (2004).
14. Bruto, M. et al. *Vibrio crassostreae*, a benign oyster colonizer turned into a pathogen after plasmid acquisition. *Isme J.* **11**, 1043–1052 (2017).
15. Lemire, A. et al. Populations, not clones, are the unit of vibrio pathogenesis in naturally infected oysters. *Isme J.* **9**, 1523–1531 (2014).
16. Piel, D. et al. Selection of *Vibrio crassostreae* relies on a plasmid expressing a type 6 secretion system cytotoxic for host immune cells. *Environ. Microbiol.* **22**, 4198–4211 (2020).
17. Kauffman, K. M. et al. Resolving the structure of phage-bacteria interactions in the context of natural diversity. *Nat. Commun.* **13**, 372 (2022).
18. Moraru, C., Varsani, A. & Kropinski, A. M. VIRIDIC—a novel tool to calculate the intergenomic similarities of prokaryote-infecting viruses. *Viruses* https://doi.org/10.3390/v12111268 (2020).
19. Moura de Sousa, J. A., Pfeifer, E., Touchon, M. & Rocha, E. P. C. Causes and consequences of bacteriophage diversification via genetic exchanges across lifestyles and bacterial taxa. *Mol. Biol. Evol.* **38**, 2497–2512 (2021).
20. Gandon, S., Buckling, A., Decaestecker, E. & Day, T. Host-parasite coevolution and patterns of adaptation across time and space. *J. Evol. Biol.* **21**, 1861–1866 (2008).
21. Blanquart, F. & Gandon, S. Time-shift experiments and patterns of adaptation across time and space. *Ecol. Lett.* **16**, 31–38 (2013).
22. Brockhurst, M. A. & Koskella, B. Experimental coevolution of species interactions. *Trends Ecol. Evol.* **28**, 367–375 (2013).
23. Gomez, P. & Buckling, A. Bacteria-phage antagonistic coevolution in soil. *Science* **332**, 106–109 (2011).
24. Tesson, F. et al. Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat. Commun.* **13**, 2561 (2022).
25. Wang, L. et al. DNA phosphorothioation is widespread and quantized in bacterial genomes. *Proc. Natl Acad. Sci. USA* **108**, 2963–2968 (2011).
26. Lopatina, A., Tal, N. & Sorek, R. Abortive infection: bacterial suicide as an antiviral immune strategy. *Annu. Rev. Virol.* **7**, 371–384 (2020).
27. Bazin, A., Gautreau, G., Medigue, C., Vallenet, D. & Calteau, A. panRGP: a pangenome-based method to predict genomic islands and explore their diversity. *Bioinformatics* **36**, i651–i658 (2020).
28. Touchon, M., Bobay, L. M. & Rocha, E. P. The chromosomal accommodation and domestication of mobile genetic elements. *Curr. Opin. Microbiol.* **22**, 22–29 (2014).
29. Millman, A. et al. Bacterial retrons function in anti-phage defense. *Cell* **183**, 1551–1561.e12 (2020).
30. Alseth, E. O. et al. Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. *Nature* **574**, 549–552 (2019).
31. Seed, K. D. et al. Evolutionary consequences of intra-patient phage predation on microbial populations. *eLife* **3**, e03497 (2014).
32. Thompson, J. R. et al. Genotypic diversity within a natural coastal bacterioplankton population. *Science* **307**, 1311–1313 (2005).
33. Rocha, E. P. C. & Bikard, D. Microbial defenses against mobile genetic elements and viruses: who defends whom from what? *PLoS Biol.* **20**, e3001514 (2022).

34. Petton, B., Boudry, P., Alunno-Bruscia, M. & Pernet, F. Factors influencing disease-induced mortality of Pacific oysters *Crassostreae gigas*. *Aquac. Environ. Interact.* **6**, 205–222 (2015).
35. de Lorgeril, J. et al. Immune-suppression by OsHV-1 viral infection causes fatal bacteraemia in Pacific oysters. *Nat. Commun.* **9**, 4215 (2018).
36. Kauffman, K. M. et al. Viruses of the Nahant Collection, characterization of 251 marine Vibrionaceae viruses. *Sci. Data* **5**, 180114 (2018).
37. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).
38. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
39. Vallenet, D. et al. MicroScope–an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res.* **41**, D636–D647 (2013).
40. Miele, V., Penel, S. & Duret, L. Ultra-fast sequence clustering from similarity networks with SiLiX. *BMC Bioinformatics* **12**, 116 (2011).
41. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).
42. Gautreau, G. et al. PPanGGOLiN: depicting microbial diversity via a partitioned pangenome graph. *PLoS Comput. Biol.* **16**, e1007732 (2020).
43. Abby, S. S., Neron, B., Menager, H., Touchon, M. & Rocha, E. P. MacSyFinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PLoS ONE* **9**, e110726 (2014).
44. Cury, J., Abby, S. S., Doppelt-Azeroual, O., Neron, B. & Rocha, E. P. C. Identifying conjugative plasmids and integrative conjugative elements with CONJscan. *Methods Mol. Biol.* **2075**, 265–283 (2020).
45. van Bloois, L. V. D. G., Wagenaar, J. A. & Zomer, A. L. RFPlasmid: predicting plasmid sequences from short read assembly data using machine learning. *Microb. Genom.* **7**, 000683 (2021).
46. Nayfach, S. et al. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat. Biotechnol.* **39**, 578–585 (2021).
47. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
48. Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes de novo assembler. *Curr. Protoc. Bioinformatics* **70**, e102 (2020).
49. Buchfink, B., Reuter, K. & Drost, H. G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
50. Mistry, J. et al. Pfam: the protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
51. Chan, P. P. & Lowe, T. M. tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* **1962**, 1–14 (2019).
52. Lopes, A., Amarir-Bouhram, J., Faure, G., Petit, M. A. & Guerois, R. Detection of novel recombinases in bacteriophage genomes unveils Rad52, Rad51 and Gp2.5 remote homologs. *Nucleic Acids Res.* **38**, 3952–3962 (2010).
53. Grazziotin, A. L., Koonin, E. V. & Kristensen, D. M. Prokaryotic Virus Orthologous Groups (pVOGs): a resource for comparative genomics and protein family annotation. *Nucleic Acids Res.* **45**, D491–D498 (2017).
54. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
55. Jones, P. et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
56. Jain, C., Rodriguez, R. L., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
57. Steinegger, M. & Soding, J. Clustering huge protein sequence sets in linear time. *Nat. Commun.* **9**, 2542 (2018).
58. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
59. Hyman, P. & Abedon, S. T. Practical methods for determining phage growth parameters. *Methods Mol. Biol.* **501**, 175–202 (2009).
60. Le Roux, F., Binesse, J., Saulnier, D. & Mazel, D. Construction of a *Vibrio splendidus* mutant lacking the metalloprotease gene vsm by use of a novel counterselectable suicide vector. *Appl. Environ. Microbiol.* **73**, 777–784 (2007).
61. Demarre, G. et al. A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries and their cognate *Escherichia coli* host strains. *Res. Microbiol.* **156**, 245–255 (2005).
62. Val, M. E., Skovgaard, O., Ducos-Galand, M., Bland, M. J. & Mazel, D. Genome engineering in *Vibrio cholerae*: a feasible approach to address biological issues. *PLoS Genet.* **8**, e1002472 (2012).

Acknowledgements

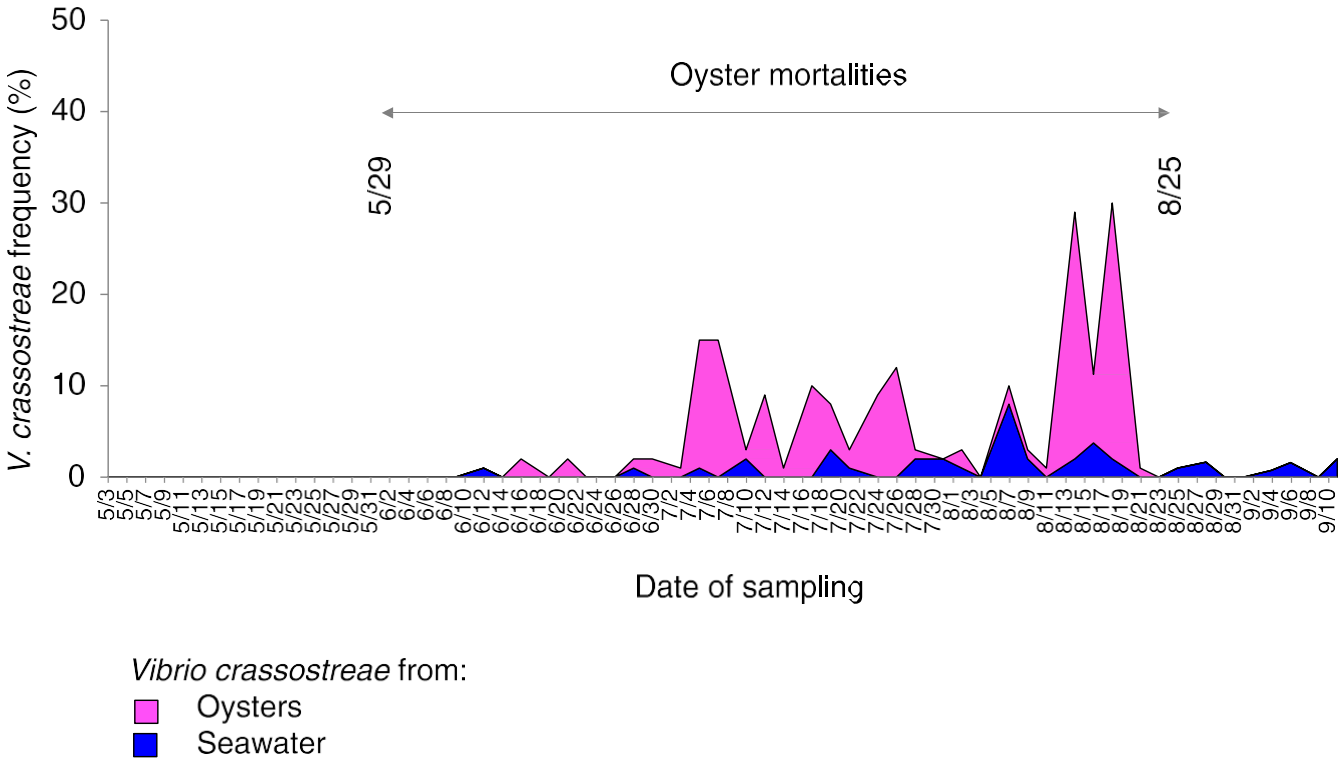
We thank M. A. Petit, M. Blokesch, M. Sullivan and A. Bernheim for valuable suggestions; M. Touchon and A. Bernheim for assistance with vibrio genome annotation; Z. Chaplain for the illustrations and help during the time series sampling; the staff of the station Ifremer Argenton and Bouin, the ABIMS (Roscoff) and LABGeM (Evry) platforms for technical assistance; Z. Allouche, Biomics Platform, C2RT, Institut Pasteur, Paris, France, supported by France Génomique (ANR-10-INBS-09-09) and IBISA; and G. Riddihough from Life Science Editors for help with the manuscript. This work was supported by funding from the Agence Nationale de la Recherche (ANR-16-CE32-0008-01, REVENGE; ANR-20-CE35-0014, RESISTE), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 884988, Advanced ERC Dynamic) to F.L.R. and Ifremer to D.P. The work was further supported by a grant from the Simons Foundation (LIFE ID 572792) to M.F.P. Part of the *Vibrio crassostreae* genome sequencing was conducted by the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, and is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

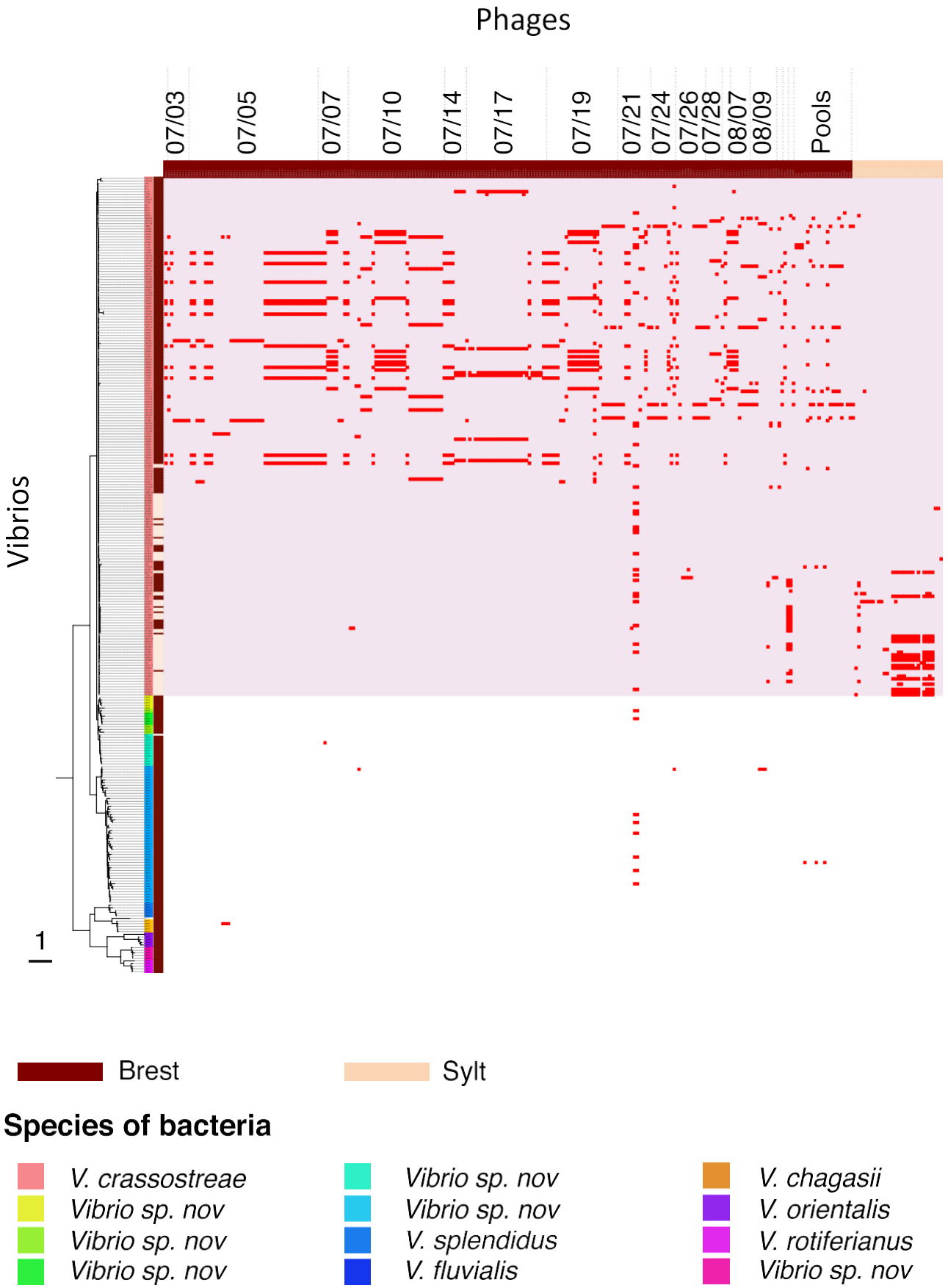
F.L.R. and D.B. conceived the project. F.L.R. wrote the paper with contributions from D.P., M.B., Y.L., F.B., K.M.W., F.A.H., K.M.K., M.F.P., D.B., S.G. and E.P.C.R. D.P. performed phage–vibrio interaction experiments with assistance from S.C., R.B.-C. and E.L. M.B. and D.G. performed the in silico analyses with assistance from K.M.K. and supervision by E.P.C.R. Y.L., F.L.R. and D.P. performed the genetics, and R.B.-C. the epigenetic experiments. S.L.P. performed the electronic microscopy analyses. D.P., Y.L., S.C., A.J., B.P. and F.L.R. established the time-series sampling. K.M.W., F.L.R. and J.D. isolated the phage and vibrio collections from Sylt. F.A.H. and M.F.P. performed and funded part of the vibrio sequencing. F.B. and S.G. designed, and F.B. performed the time-shift analysis. F.L.R. supervised the project and secured funding.

Competing interests

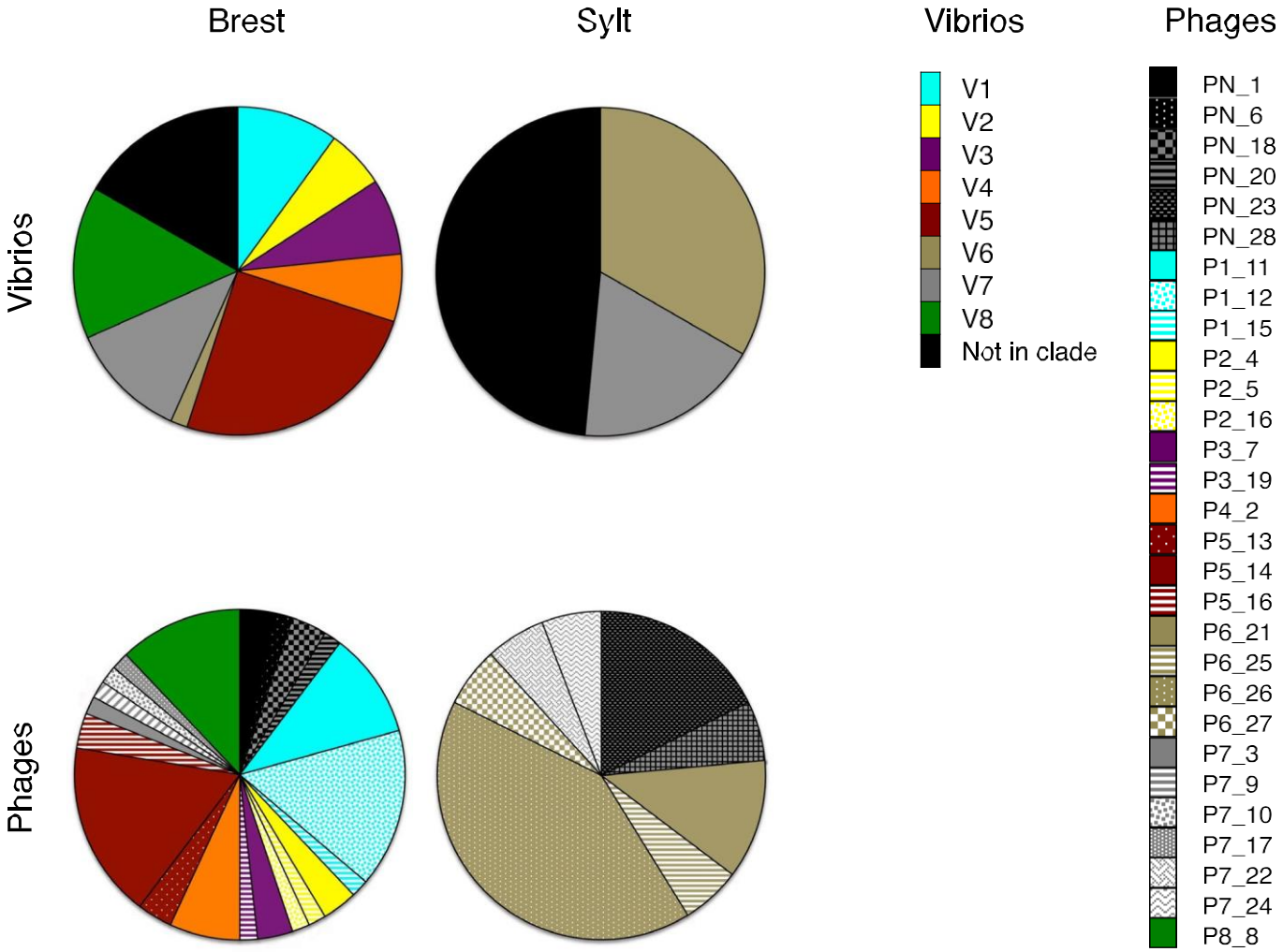
Authors declare no competing interests.



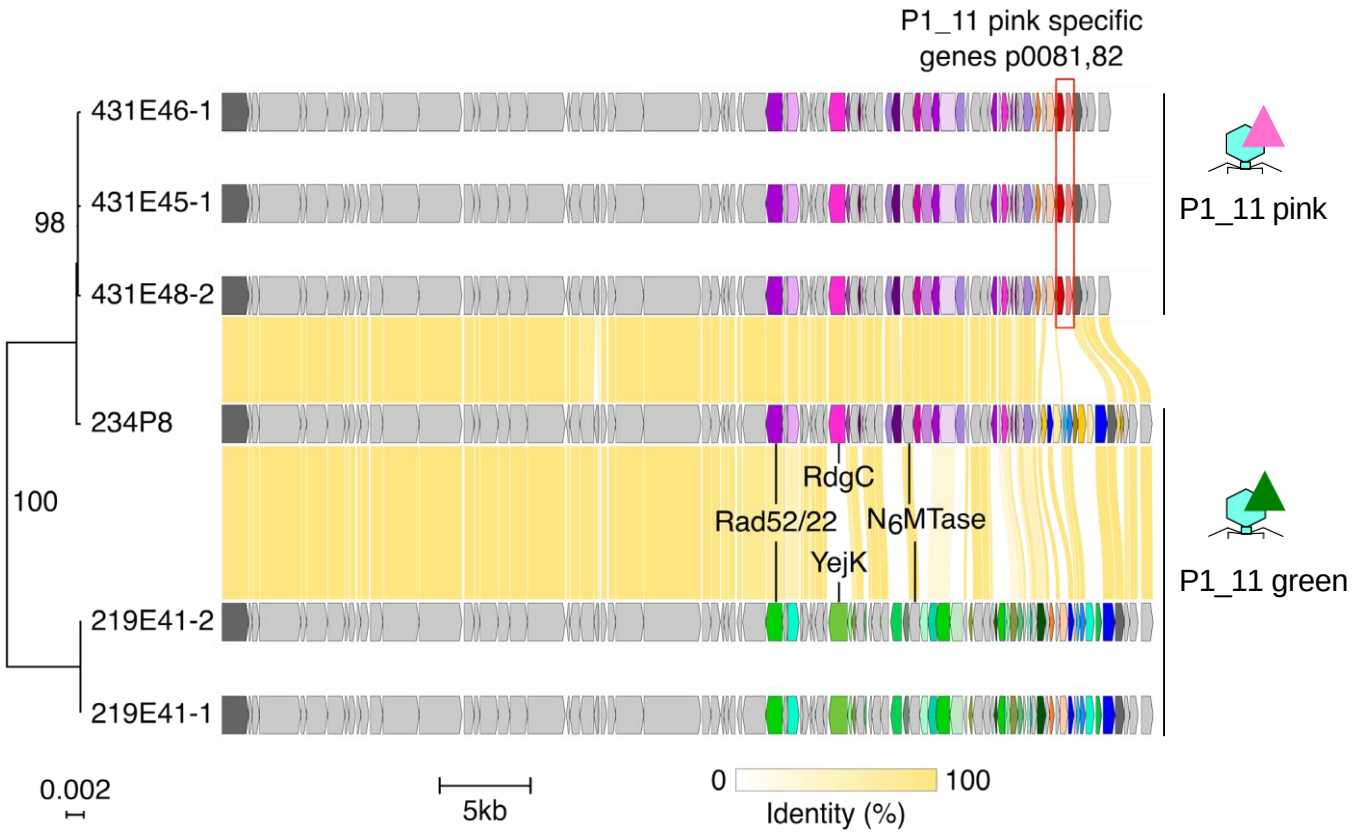
Extended Data Fig. 1 | Dynamics of *Vibrio crassostreae*. On each sampling date (May 3rd- September 11th 2017, x-axis), vibrios from seawater (size fraction 1-0.2µm, blue) or five oyster tissues (pink) were selected on TCBS and genotyped to identify *V. crassostreae* isolates. The y-axis indicates the frequency of *V. crassostreae* (number of positive isolates out of the randomly picked 48 colonies*100). The arrow indicates the period of oyster mortalities, that is May 29th- August 25th.



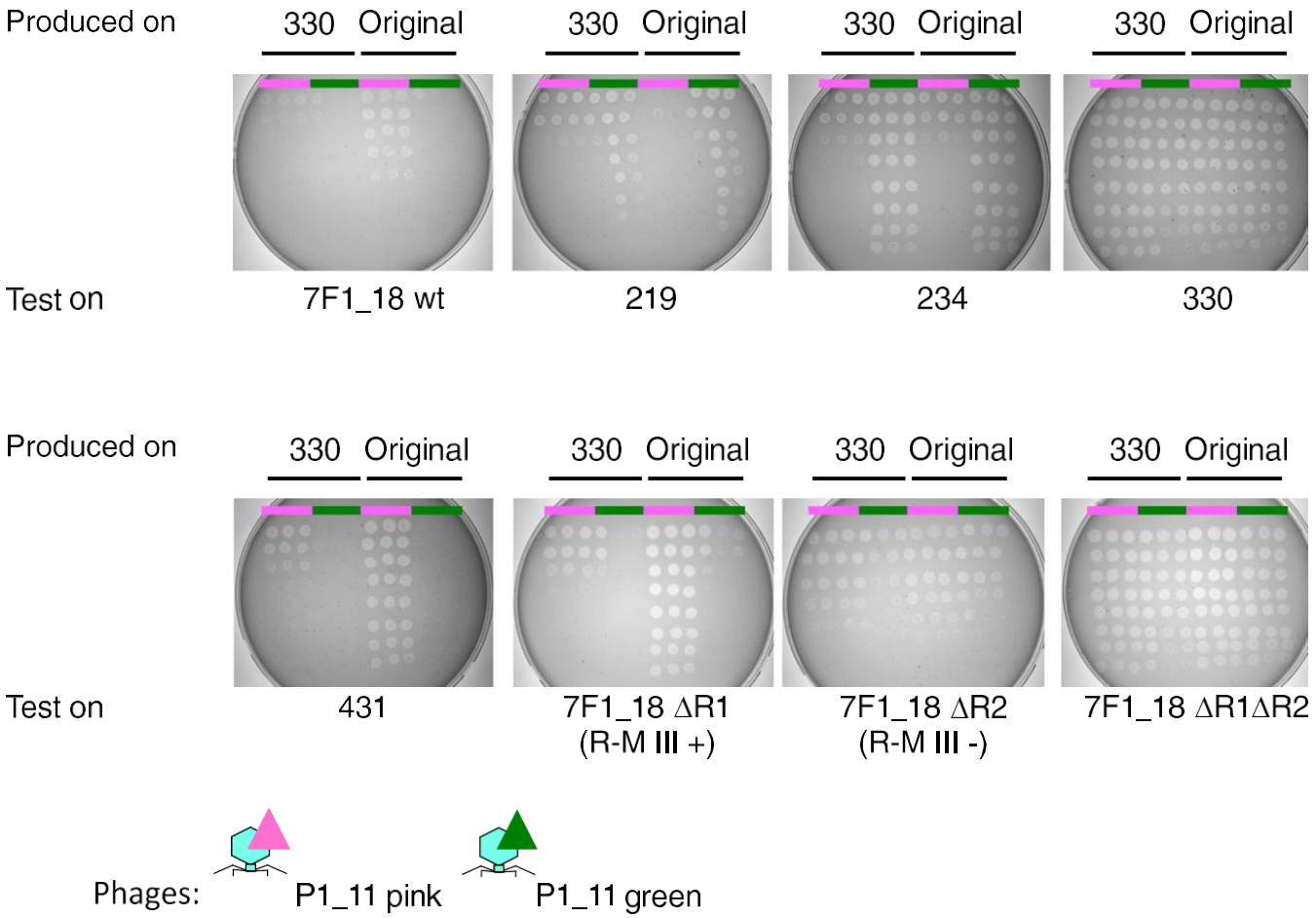
Extended Data Fig. 2 | Host range matrix for assay of phages on *gyrB*-sequenced hosts. Rows represent vibrio strains (n = 299), columns represent phages (n = 243 phages isolated in Brest and ordered by date or "Pools" when phages were isolated from a mix of viruses from 5 consecutive dates; n = 31 phages isolated in Sylt), red marks indicate infection of host. Pink shade discriminate *V. crassostreae* isolates from representatives of other species.



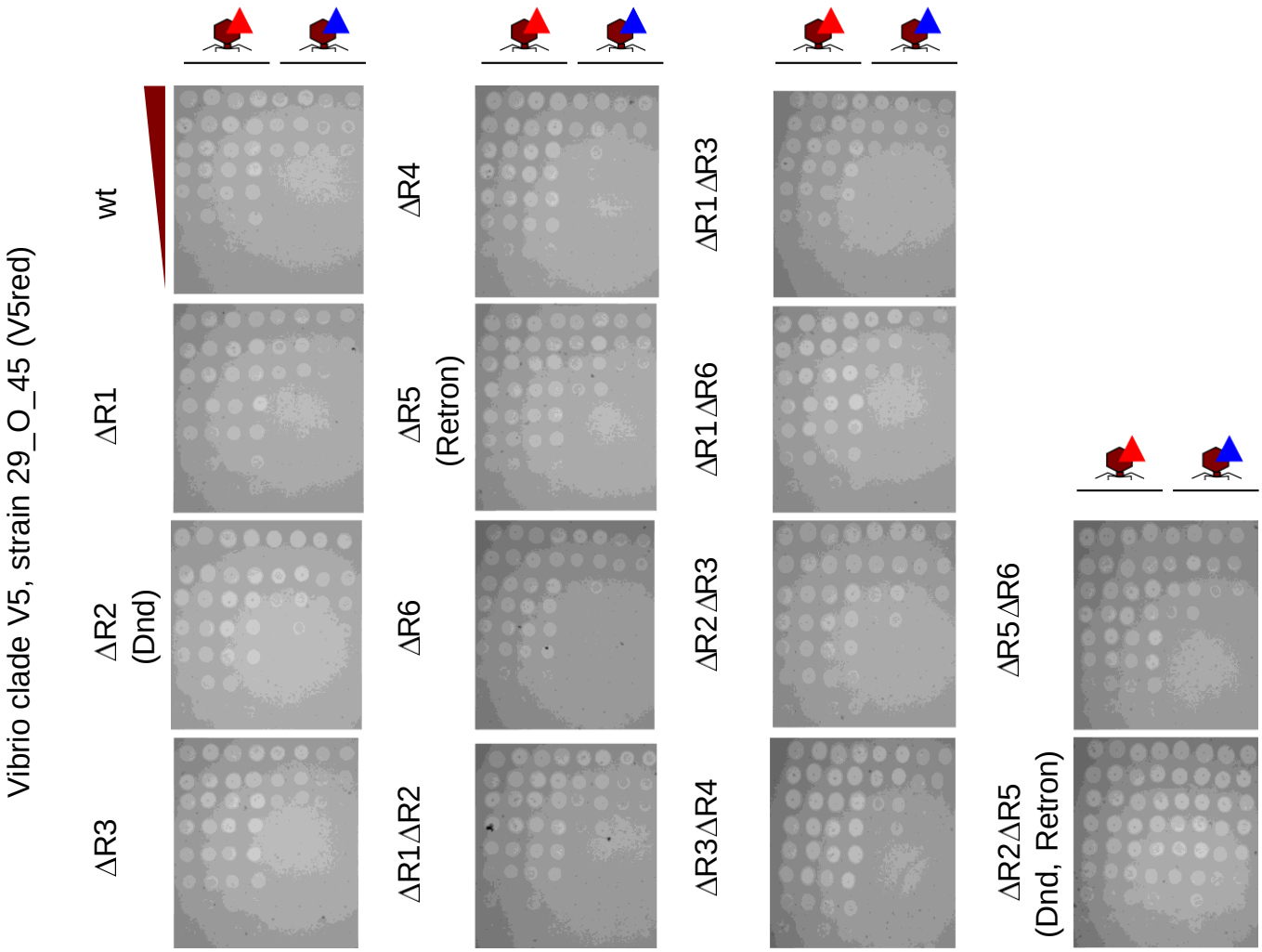
Extended Data Fig. 3 | Non-overlapping distribution of vibrio clades and phage genus across locations. The all-by-all host range infection assay (Fig.1) reveals that sympatric killing of vibrio by phages is more frequent than allopatric killing (Fig. 3a). The presence of phage of a given cluster is associated with the presence of the corresponding *V. crassostreae* clade at each location. Upper pie charts show the number of isolates per clade (V1 to V8) or not in clade among 120 and 33 vibrios from Brest and Sylt respectively. Lower pie charts show the number of isolates per phage cluster among 58 and 17 phages from Brest and Sylt respectively. Phage clusters are indicated as Px_y, where P indicates phage, x the vibrio clade they infect (1 to 8 and N when not in clade), and y is the VIRIDIC genus number (1 to 28).



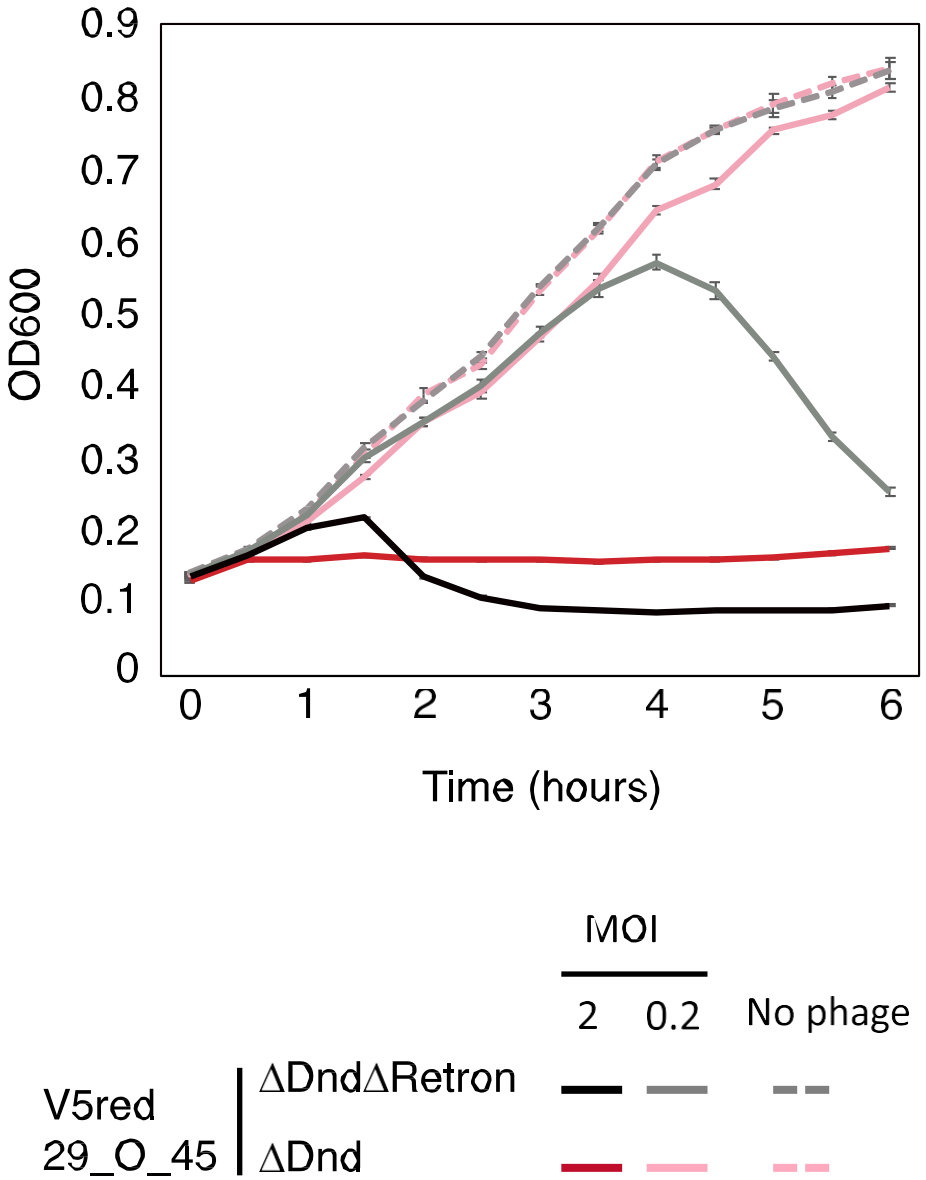
Extended Data Fig. 4 | Core phylogenetic tree and genes synteny for phages from the cluster P1_11. The three pink (but not the three green) phages are able to neutralize the R-M III defence. Core proteins are indicated in grey (with 30% identity and 80% coverage thresholds), large and small terminase subunit in dark grey and accessory proteins in colors. The tree was built using IQ-TREE2 with 1000 bootstraps and the GTR model. Genes encoding putative recombinases (Rad52/22, RdgC, YejK) and methylases (MTases) were identified using dedicated HMM profiles.



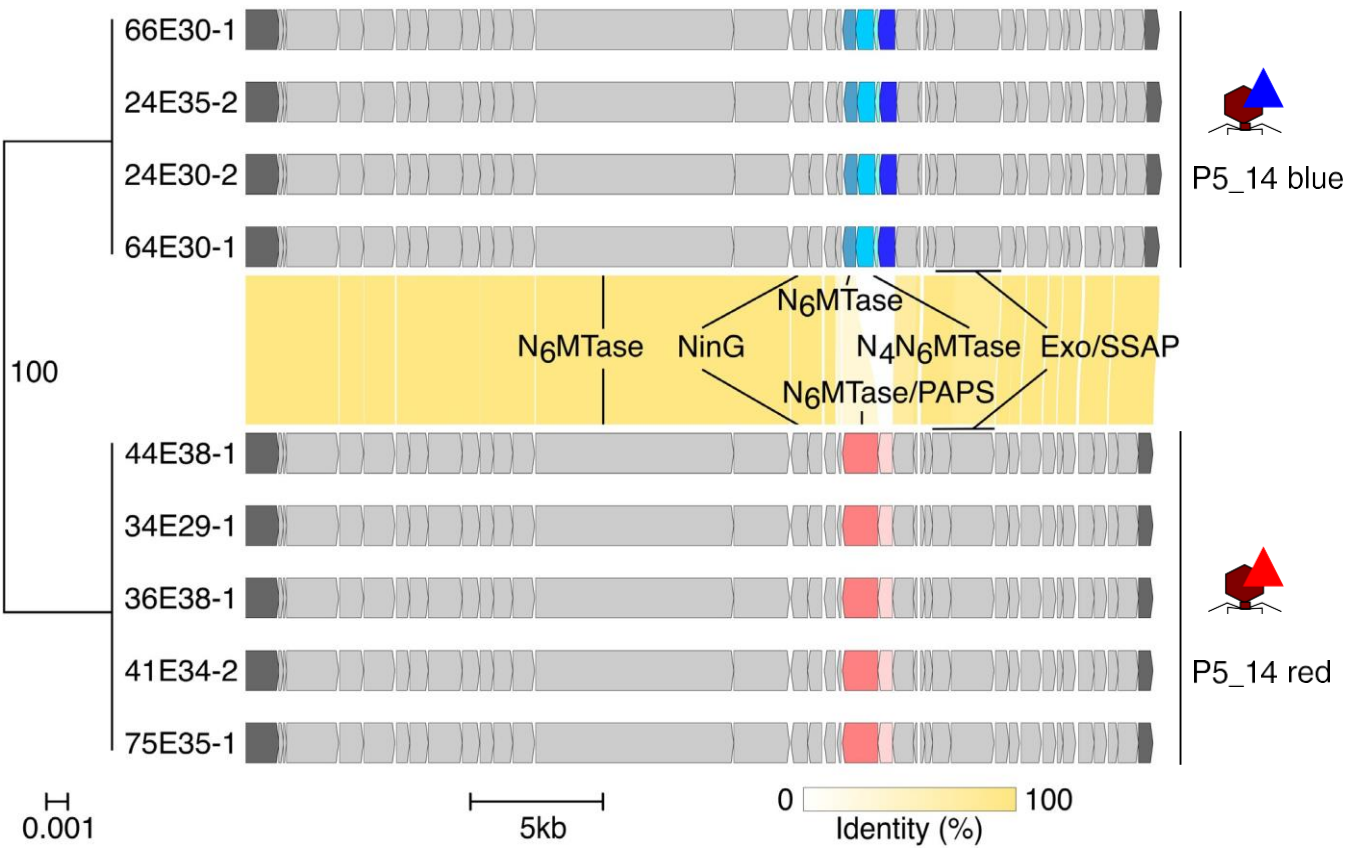
Extended Data Fig. 5 | Host-dependent epigenetic modification allows P1_11 podoviruses to neutralize a R-M III system. Three green and three pink phages were produced in their original host (the strain used to isolate phage from seawater flocculate) or in the strain 46_O_330 (name here 330) that is sensitive to all P1_11 phages. The green phages were 234P8, 219E41.1, 219E41.2 and original hosts were vibrio V1 strains 40_O_234 (234) and 38_P_219 (219). The pink phages were 431E45.1, 431E46.1, 431E48.2 all isolated from 48_O_431 (431). Tenfold dilutions of each progeny were spotted on the strain 7F1_18 wild type and derivatives DR1, DR2 and DR1DR2.



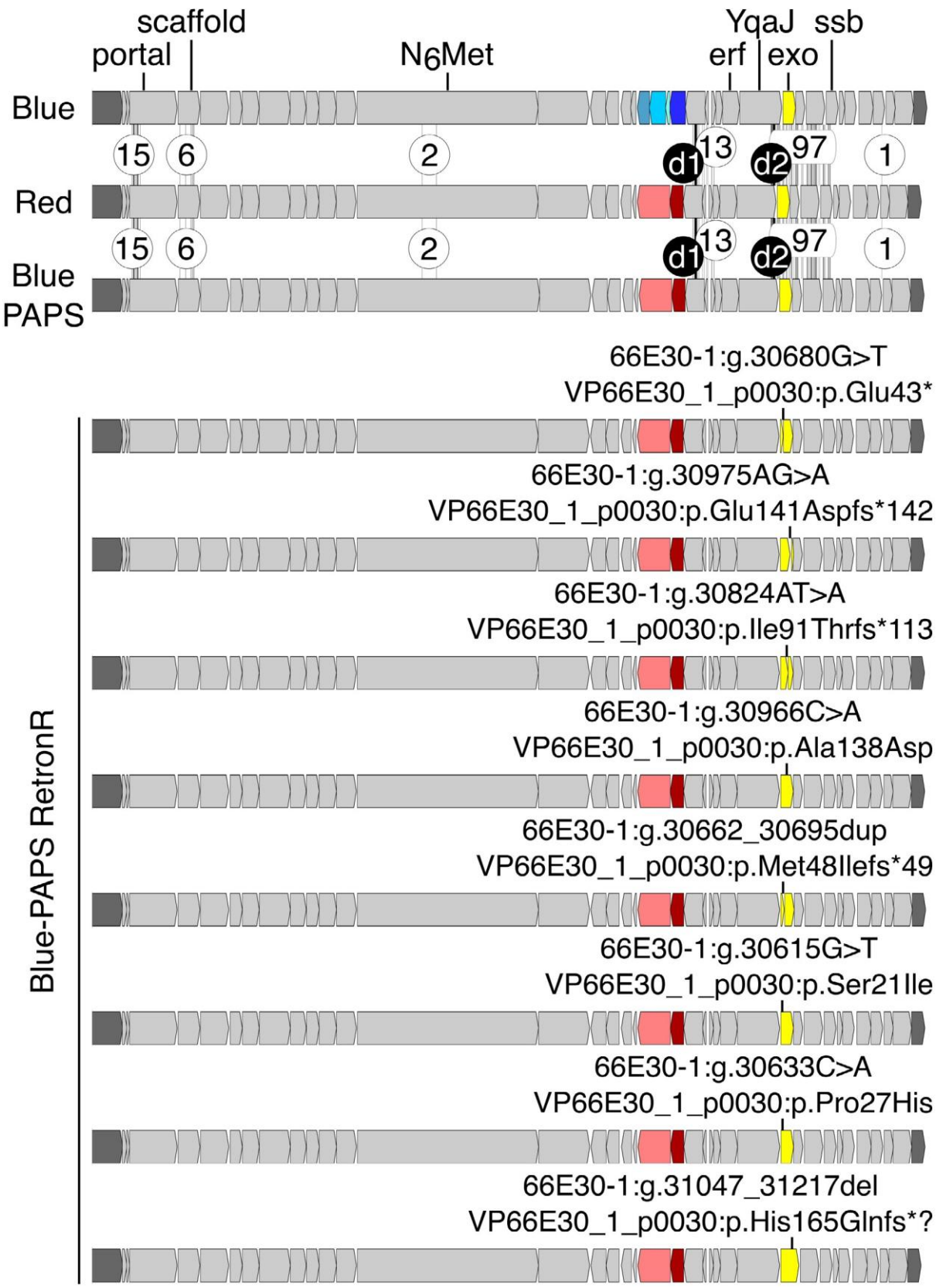
Extended Data Fig. 6 | Changes in susceptibility to phage killing observed for V5red specific region deletions. Single and double deletions of the six regions found in all V5red and absent in all V5blue strains were performed in one strain (29_O_45) of vibrio clade V5. Lawns of bacterial hosts with drop spots of a 1:10 dilution series of phages from cluster P5_14, red (red triangle; phage) and blue (blue triangle). The red phages were 36E38.1, 41E34.2, 44E38.1 and 44E38.2. The blue phages were 24E30.2, 24E35.2, 64E30.1 and 66E30.1.



Extended Data Fig. 7 | The Ec48 retron system from V5red strain causes abortive Infection. Infection dynamics in liquid of the V5red (29_O_45) lacking Dnd (ΔDnd) or both Dnd and the Ec48 retron systems (ΔDndΔRetron) infected by blue P5_14 phage (66E30.1) at MOI 0.2 and 2. Infections were performed in technical triplicate (± SD from three replicates). Data are representative of two independent experiments.

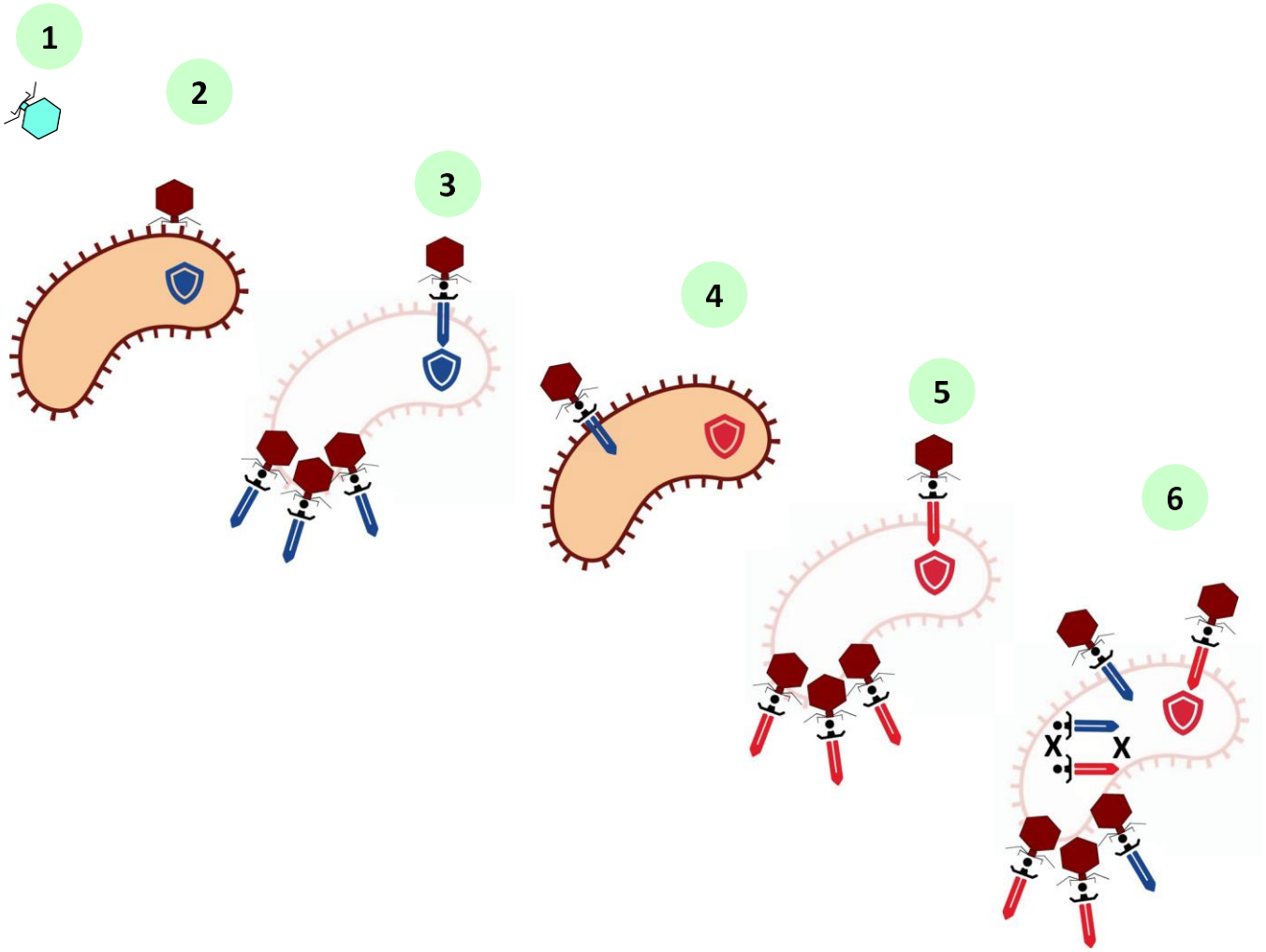


Extended Data Fig. 8 | Core genome phylogenetic tree and gene synteny representation for phage from the cluster P5_14. Core proteins are indicated in grey (blue/red pairwise identities indicated by a yellow gradiant) large and small terminase subunit in dark grey. P5_14 red (p0020 and p0021 in phage 44E38.1) and blue (p0020 to 23 in phage 66E30.1) specific genes are indicated by red and blue gradient respectively. The tree was built using IQ-TREE2 with 1000 bootstraps and the GTR model. Genes encoding putative recombinases (NinG, Exo/SSPAP) and methylases (MTases) were identified using dedicated HMM profiles.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Gene synteny of P5_14 phages and derivative Blue-PAPS-retron escapers. Core proteins are indicated in grey, large and small terminase subunit in dark grey, specific P5_14 red proteins in red gradient and specific P5_14 blue proteins in blue gradient and retron-escaping exonuclease in yellow. The number of SNPs and presence of deletion between phages are indicated in white and black circles respectively. SNPs were labeled using HGVS nomenclature.



Extended Data Fig. 10 | Genetic mechanisms driving the specificity of the interaction between a natural population of vibrio and their viral predators. A successful infection requires the phage to adsorb on specific receptor(s) at the surface of the bacterial cell and then bypass the host's intracellular defences. In *Vibrio crassostreae*, adsorption governs a first level of specificity of the interaction between a phage cluster and a clade within this bacterial species (1). Defence systems, mostly transmitted by mobile genetic elements, constitute a second barrier to infection (2). The phage can escape these defences by epigenetic or genetic modifications (3). Facing this new threat, a bacterium that has acquired a new defence system can be selected (4) and in the same way, the phage can adapt by acquiring new counter-defences (5). Recombination can allow the exchange of counter-defence systems and generate a genetically diverse progeny (6).