



Genomic and Seasonal Variations among Aquatic Phages Infecting the Baltic Sea Gammaproteobacterium *Rheinheimera* sp. Strain BAL341

E. Nilsson,^a K. Li,^a J. Fridlund,^a S. Šulčius,^{a*} C. Bunse,^{a*} C. M. G. Karlsson,^a M. Lindh,^{a*} D. Lundin,^a  J. Pinhassi,^a K. Holmfeldt^a

^aFaculty of Health and Life Sciences, Department of Biology and Environmental Science, Centre for Ecology and Evolution in Microbial Model Systems, Linnaeus University, Kalmar, Sweden

ABSTRACT Knowledge in aquatic virology has been greatly improved by culture-independent methods, yet there is still a critical need for isolating novel phages to identify the large proportion of “unknowns” that dominate metagenomes and for detailed analyses of phage-host interactions. Here, 54 phages infecting *Rheinheimera* sp. strain BAL341 (*Gammaproteobacteria*) were isolated from Baltic Sea seawater and characterized through genome content analysis and comparative genomics. The phages showed a myovirus-like morphology and belonged to a novel genus, for which we propose the name *Barbavirus*. All phages had similar genome sizes and numbers of genes (80 to 84 kb; 134 to 145 genes), and based on average nucleotide identity and genome BLAST distance phylogeny, the phages were divided into five species. The phages possessed several genes involved in metabolic processes and host signaling, such as genes encoding ribonucleotide reductase and thymidylate synthase, *phoH*, and *mazG*. One species had additional metabolic genes involved in pyridine nucleotide salvage, possibly providing a fitness advantage by further increasing the phages’ replication efficiency. Recruitment of viral metagenomic reads (25 Baltic Sea viral metagenomes from 2012 to 2015) to the phage genomes showed pronounced seasonal variations, with increased relative abundances of barba phages in August and September synchronized with peaks in host abundances, as shown by 16S rRNA gene amplicon sequencing. Overall, this study provides detailed information regarding genetic diversity, phage-host interactions, and temporal dynamics of an ecologically important aquatic phage-host system.

IMPORTANCE Phages are important in aquatic ecosystems as they influence their microbial hosts through lysis, gene transfer, transcriptional regulation, and expression of phage metabolic genes. Still, there is limited knowledge of how phages interact with their hosts, especially at fine scales. Here, a *Rheinheimera* phage-host system constituting highly similar phages infecting one host strain is presented. This relatively limited diversity has previously been seen only when smaller numbers of phages have been isolated and points toward ecological constraints affecting the *Rheinheimera* phage diversity. The variation of metabolic genes among the species points toward various fitness advantages, opening up possibilities for future hypothesis testing. Phage-host dynamics monitored over several years point toward recurring “kill-the-winner” oscillations and an ecological niche fulfilled by this system in the Baltic Sea. Identifying and quantifying ecological dynamics of such phage-host model systems *in situ* allow us to understand and study the influence of phages on aquatic ecosystems.

KEYWORDS Baltic Sea, bacteriophage, genomics, temporal variation

Viruses are the most abundant biological entities on the planet (1). With an approximate abundance of 10^7 viral particles ml^{-1} in surface waters of the global oceans, they outnumber their microbial hosts approximately 10-fold (2). Through

Citation Nilsson E, Li K, Fridlund J, Šulčius S, Bunse C, Karlsson CMG, Lindh M, Lundin D, Pinhassi J, Holmfeldt K. 2019. Genomic and seasonal variations among aquatic phages infecting the Baltic Sea gammaproteobacterium *Rheinheimera* sp. strain BAL341. *Appl Environ Microbiol* 85:e01003-19. <https://doi.org/10.1128/AEM.01003-19>.

Editor Eric V. Stabb, University of Illinois at Chicago

Copyright © 2019 Nilsson et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to K. Holmfeldt, karin.holmfeldt@lnu.se.

* Present address: S. Šulčius, Nature Research Centre, Laboratory of Algology and Microbial Ecology, Vilnius, Lithuania; C. Bunse, Alfred-Wegener-Institut Helmholtz-Zentrum Für Polar- Und Meeresforschung, Bremerhaven, Germany, and Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg, Oldenburg, Germany; M. Lindh, Swedish Meteorological and Hydrological Institute, Göteborg, Sweden.

Received 2 May 2019

Accepted 10 July 2019

Accepted manuscript posted online 19 July 2019

Published 29 August 2019

infection and lysis of 20% of the bacterial community on a daily basis, viruses that infect bacteria (bacteriophages, or phages for short) are important factors influencing bacterial mortality and genetic diversity (3). Lately, high-throughput sequencing techniques have improved our knowledge regarding the viral communities, including viruses that infect so-far uncultured hosts. Metagenomic analyses have unraveled large-scale spatial variations within viral populations (4, 5), shed light on virus-host dynamics (6, 7), and predicted the functional gene potential of viral communities (8, 9). Moreover, single-cell genomics (10) of phage-host pairs from natural communities and viral tagging (11), where host-specific metagenomes are analyzed, have uncovered viral diversity and host interactions on finer scales. Since the discovery of auxiliary metabolic genes (AMGs) (12–14), which typically are metabolic genes transferred from the host's genome to the phage genome that enable continued or increased phage replication during infection (15), the functional capacity of viral communities and the potential of these genes to influence biogeochemical cycles have gained increased interest.

The isolation of phage-host pairs is essential to elucidate fine-scale patterns of phage-host interactions. For example, different phages infecting the same bacterial species can display a wide range of infection abilities, both on a qualitative (whom they infect) and quantitative (how well they infect) scale (16). In addition, microdiversity within viral species has the ability to affect not only host range but also replication efficiency, seen, e.g., as burst size (17). Further, a viral strain can have different replication efficiencies on different host strains (17, 18) due to differences in the host transcriptional (19) and translational (20) responses. Thus, the knowledge gained from phage isolation and experimental studies is imperative to fully understand phage-host systems and hypotheses that derive from sequencing methods such as metagenomics, particularly with regard to viral diversity and ecological importance. This is especially important for nonmodel systems since we do not know how they differ from model systems before they have been isolated.

Clinically important phages infecting different gammaproteobacterial species, including, e.g., *Escherichia* spp. and *Pseudomonas* spp., are among the most well-studied phage-host systems in the world (21). However, in marine environments, phages infecting *Gammaproteobacteria* have received less attention, and research has been focused mainly on phages infecting *Pseudoalteromonas* (22–24) or *Vibrio* (13, 25, 26). These isolated *Pseudoalteromonas* (27, 28) and *Vibrio* (26) phages display a large diversity, with phages belonging to a variety of genera and families, which is similar to what has been seen for *Escherichia* (29) and *Pseudomonas* (30) phages. In addition, a podovirus and a myovirus, respectively, have been identified infecting the uncultured gammaproteobacterial SAR92 and SAR86 clades using single amplified genomics (10). However, information regarding phages infecting other *Gammaproteobacteria* species is sparse.

Members of the bacterial genus *Rheinheimera* exist in various environments, and they have been found in soils (31) and aquatic habitats, including both freshwater (32) and marine (33) environments. They have also been shown to proliferate in more extreme environments, for example, in iron backwash sludge (34), and are able to degrade hydrocarbons such as *n*-alkanes (35). Bacteria within this genus have been detected in the brackish Baltic Sea (36), and transplant experiments with shifting salinities have noted stabilizing or increasing abundances as adjustment effects (37). Even though species in the *Rheinheimera* genus are widespread, to our knowledge, no phages infecting these species have been isolated.

The aim of this study was to isolate and characterize aquatic phages infecting environmentally important bacteria and to increase our knowledge about phage diversity and the ecological relevance of phages within aquatic ecosystems. We present 54 genetically similar phages that are distinct from previously described phages and infect the Baltic Sea isolate *Rheinheimera* sp. strain BAL341. We propose that these phages are assigned to a novel viral genus, *Barbavirus*, consisting of five species which contain genes potentially involved in various metabolic pathways. The phages' preva-

lence is evident by metagenomic recruitment, and their temporal abundances coincide with the host's abundance in late summer.

RESULTS

Genomic characteristics. Fifty-four phage isolates, originating from individual plaques, infecting *Rheinheimera* sp. strain BAL341 were isolated from the long-term sampling station Linnaeus Microbial Observatory (LMO) in the Baltic Sea Proper (38, 39); 31 isolates were obtained in August 2015, and 23 were obtained in September 2015. The proposed names of the isolates, following standards of the International Committee on the Taxonomy of Viruses (ICTV) (40), take the form *Rheinheimera* phage vB_RspM-barba, followed by numbers indicating the order of isolation and A or S, representing the time of sampling (August or September), e.g., *Rheinheimera* phage vB_RspM-barba18A (see Table S1 in the supplemental material). All 54 phage isolates were whole-genome sequenced, and the genomes were assembled with a coverage ranging from 193× to 4,789× (Table S1). Identical *k*-mers at both ends of the genomes indicated that they were circular. The phage genome sizes varied between 80 and 84 kb, with between 134 and 145 predicted genes in each genome. All-versus-all genome comparison showed that the barba phages shared more than 70% average nucleotide identity (ANI) (Fig. 1; Table S2) across their entire genomes, indicating that they belong to the same genus (>50% ANI) (41). Phylogenetic analysis using VICTOR (42) showed that the isolated phages formed a separate clade that was distinct from previously sequenced and characterized phages (Fig. 2; Table S3). Whereas the low similarity to the reference genomes prevents us from drawing firm conclusions between the novel phages and the reference genomes, the short branches within the clade containing the novel phages let us propose that these phages should be assigned to a novel genus, for which the name *Barbavirus* (Baltic Sea *Rheinheimera* strain BAL341) is suggested. Within this novel genus, 48 of the phage genomes shared more than 95% ANI and therefore belong to one species (Fig. 1; Table S2). Using this species cutoff, four additional species were distinguished, with two species represented by two isolates and two species represented by only a single isolate each (Fig. 1; Table S2). This agreed with the VICTOR analysis for which five species within this one genus were identified (Fig. 2; Table S3). The five species were named based on their suggested type phages: (i) *Rheinheimera* virus Barba18A, (ii) *Rheinheimera* virus Barba21A, (iii) *Rheinheimera* virus Barba5S, (iv) *Rheinheimera* virus Barba8S, and (v) *Rheinheimera* virus Barba19A (Fig. 1). Hereafter, species are indicated with full species name, i.e., *Rheinheimera* virus Barba18A, while isolates are indicated with their unique identifier, e.g., barba18A; also, we use the term barba phages to denote all isolates collectively while *Barbavirus* indicates the species within the genus.

Morphology and replication characteristics. All phages displayed similar plaque morphologies, with round, clear plaques up to 4 mm in diameter that appeared within 24 to 48 h. Transmission electron microscopy (TEM) of barba18A, the type phage of the genus, showed a myovirus morphology (Fig. 3a), with a capsid diameter of 72.1 nm (standard deviation, ±2.7 nm), a tail length of 88.7 (±2.2) nm, and a tail width of 18.8 (±1.5) nm ($n = 50$). When grown in Zobell medium and replicating on *Rheinheimera* sp. strain BAL341, barba18A had a burst size of 162 (±20) phages and a latent period of 50 (±8.7) min ($n = 3$) (Fig. 3b). barba19A, the type phage of the most divergent species (Fig. 1 and 2), had a significantly smaller burst size of 101 phages (±17 phages; $n = 3$; Welch two-sample *t* test, $df = 3.88$ and $P = 0.017$) and longer latent period of 80 min (±0 min; $n = 3$; Welch two-sample *t* test, $df = 2$ and $P = 0.027$) (Fig. 3b).

We found three bacterial strains in our in-house bacterial collection (namely *Rheinheimera* sp. strain BAL331, *Rheinheimera* sp. strain BAL335, and *Rheinheimera* sp. strain BAL336) that had 97 to 99% 16S rRNA gene sequence identity to other cultured *Rheinheimera* bacterial strains in NCBI (Table S4). Compared to the model bacterium *Rheinheimera* sp. strain BAL341, they had 95 to 98% sequence identity on the sequenced parts of the 16S rRNA gene (Table S4). Neither barba18A nor barba19A could infect any of these other *Rheinheimera* strains.

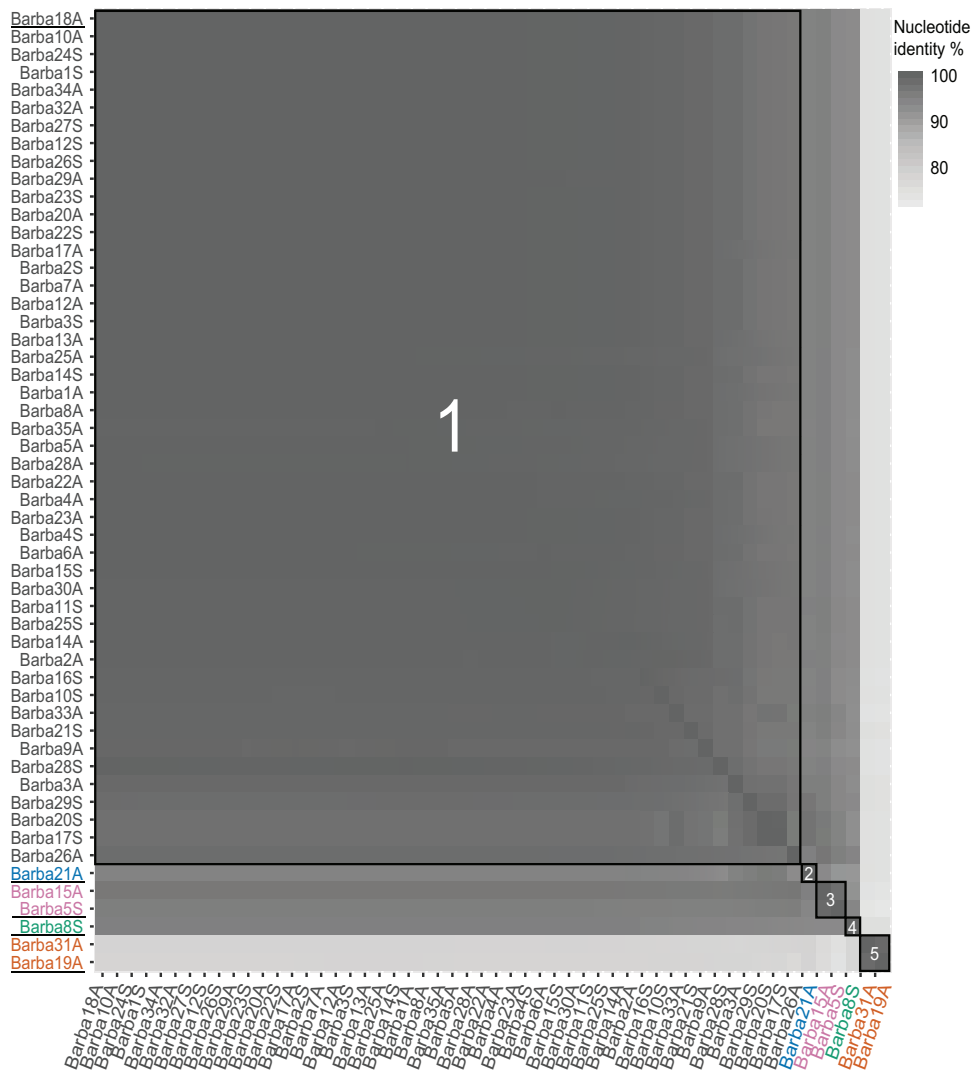


FIG 1 Pairwise blastn comparisons of average nucleotide identities across entire genomes were calculated using gegenees (version 3.0.0) (settings: fragment size of 500 and step size of 500) (103). The isolates belonging to the different species are boxed, and the names are color coded as follows: black, *Rheinheimera* virus Barba18A; blue, *Rheinheimera* virus Barba21A; pink, *Rheinheimera* virus Barba5S; green, *Rheinheimera* virus Barba8S; orange, *Rheinheimera* virus Barba19A. The type phages are underlined on the left axis.

Gene functionality. The core genome shared among the 54 novel phage genomes consisted of 97 genes with at least 70% amino acid identity, while the flexible genome consisted of 99 additional genes, most of which were found in *Rheinheimera* virus Barba19A. Thus, the pangenome of these phages was made up of a total of 196 genes (Table S5). Of the genes in the pangenome, 81 genes had significant (E value of <0.001) sequence identity to proteins in the NCBI nonredundant (nr) database, where 81% of the matches were to either gammaproteobacteria or phages infecting bacteria within this class (Table S5). Combined with matches to the pfam database (adding function to 16 additional genes), 43 genes could be assigned to a predicted function (Table 1; Table S5). No tRNAs were detected in the genomes.

The genomes of the five species were similarly arranged and consisted of two larger modules, one for structural genes and one for host interaction genes (Fig. 4). Within the structural module, genes involved in the structure of the phage (e.g., capsid proteins and tail elements), packaging of DNA (e.g., large terminase subunit), and peptidases were detected (Table 1; Table S5). The host interaction module consisted of genes involved in, for example, DNA processing, replication, and recombination (e.g., DNA



FIG 2 Genome BLAST distance phylogeny (GBDP) tree created with FastME (SPR branch swapping), as a part of VICTOR (42). The numbers above branches are GBDP pseudobootstrap support values from 100 replications (values of >50 are reported), and the branch lengths are scaled in terms of the D0 distance formula (settings: word length, 11; E value filter, 1.0; algorithm, greedy-with-trimming). Isolates belonging to *Barbavirus* are colored based on the species to which they belong as described in the legend of Fig. 1.

polymerase and primase), host signaling (nucleoside triphosphate pyrophosphohydrolyase [*mazG*]), and various metabolic processes (e.g., ribonucleotide reductase [RNR], thymidylate synthase [*thy1*], and protein PhoH [*phoH*]) (Table 1; Table S5). Of the metabolic genes, two were part of the core genome, the RNR genes *nrdA* and *nrdB*. These genes had 99.9 to 100% amino acid sequence identity in all barba phage isolates. The other metabolic genes found in the barba phages, glutaredoxin (*glrx*) and *thy1*, involved in nucleotide metabolism, and *phoH*, potentially involved in phosphate metabolism, were part of the flexible genome. For these genes, two different variants were detected among the barba phages (Table 2), with the common version shared among all phages except the phages within one particular species, either *Rheinheimera* virus Barba19A or *Rheinheimera* virus Barba5S (Table 2). The genes annotated as *glrx* in barba18A (the common one) and in barba19A (the rare one) shared 43% amino acid

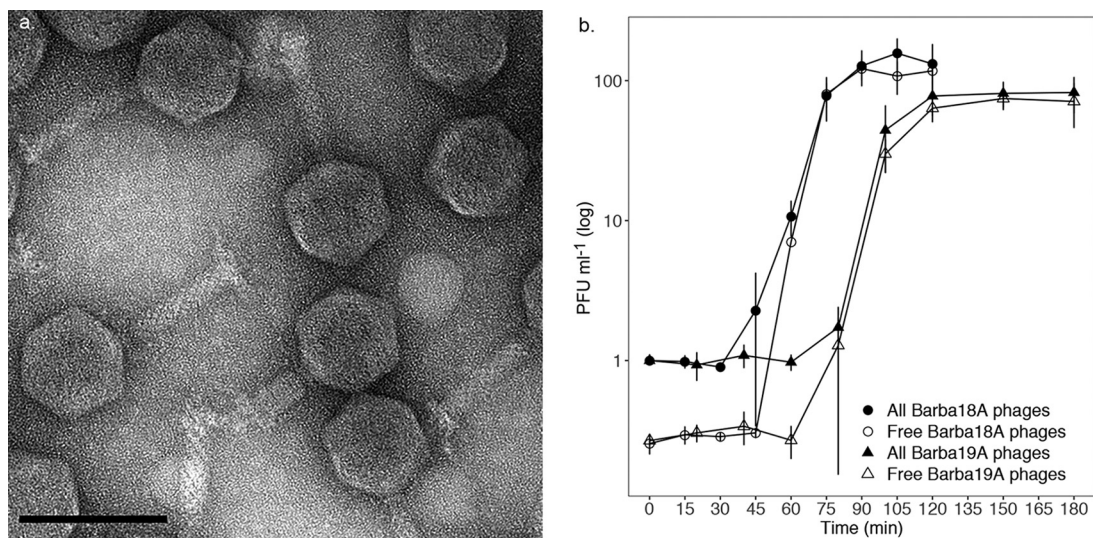


FIG 3 (a) TEM of barba18A shows a myovirus morphology with a capsid that was measured to 72.1 nm (standard deviation, ± 2.7 nm), tail length of 88.7 nm (± 2.2 nm), and width of 18.8 nm (± 1.5 nm). Scale bar, 100 nm. (b) One-step growth curve of barba18A- and barba19A-infected *Rheinheimera* sp. strain BAL341. All phages, phages within and attached to cells as well as unattached phages; free phages, only unattached phages. The values are normalized based on the concentration of all phages at time zero for each replicate. Error bars indicate standard deviations ($n = 3$) but are not shown for barba18A at 60 min for free phages as the deviation was too large (± 7.18 PFU/ml).

identity across 92% of the genes (Table 2). Similarly, the common and rare (in barba5S) versions of the genes annotated as *phoH* shared 31% amino acid similarity across 94% of the genes. However, the two versions of the thymidylate synthase gene did not show significant sequence identity (by blastp, no E value of < 0.001) between the two different phage variants (Table 2). In addition, the *Rheinheimera* virus Barba19A isolates contained two metabolic genes that the other isolates lacked, a ribose-phosphate pyrophosphokinase (*prs*) and a nicotinate phosphoribosyltransferase (*pncB*) that are involved in pyridine nucleotide salvage. These phage genes, potentially involved in different metabolic processes, either showed low sequence similarity matches ($< 40\%$ amino acid identity at an E value of < 0.001) or did not provide any significant matches (no E value of < 0.001) to genes with similar functions within the host genome (Table 2).

Temporal variations. The barba phages were isolated at two different time points, 19 August 2015 and 16 September 2015. The majority of the isolates belonged to *Rheinheimera* virus Barba18A and were isolated from both time points (Fig. 1). In addition, the two *Rheinheimera* virus Barba5S phages were isolated in both August and September (Fig. 1). Phages belonging to the other species were isolated in either August (*Rheinheimera* virus Barba21A and *Rheinheimera* virus Barba19A, one and two isolates, respectively) or September (*Rheinheimera* virus Barba8S, one isolate).

The seasonal dynamics of *Barbavirus* were investigated by competitive recruitment of viral metagenomic reads from the LMO station against the genomes of the five representative phages of each species (barba18A, barba21A, barba5S, barba8S, and barba19A) (Table S6). In order for a phage to be counted as present in a metagenomic data set, at least 75% of a phage genome should be covered by viral metagenomic reads with 90% nucleotide identity (5, 43). According to this definition and using the competitive recruitment, barba19A was detected at one time point while the other four species did not have enough coverage at any time point (Table S6). Due to the high sequence similarity among the barba phages, competitive recruitment was unable to distinguish the true source of reads that belong to the identical parts of the genomes and thus was unable to convey the prevalence of the phages *in situ*. Therefore, barba19A was selected to represent the barba phages and was used for individual recruitment analysis, which resulted in detection of barba19A ($< 75\%$ coverage) in

TABLE 1 Predicted genes with their functions and the functional group to which they belong

Gene name	Length of protein (aa) ^a	Functional group ^b	Function
Barba18A_gp001	334	Nucleotide metabolism and recycling	Ribonucleotide reductase
Barba18A_gp002	951	Nucleotide metabolism and recycling	Ribonucleotide reductase
Barba18A_gp004	223	Nucleotide metabolism and recycling	Thymidylate synthase
Barba18A_gp005	85	Nucleotide metabolism and recycling	Glutaredoxin
Barba18A_gp006	119	Host signaling	<i>mazG</i>
Barba18A_gp008	404	DNA processing, replication, and recombination	Metallo-dependent phosphatase
Barba18A_gp012	202	DNA processing, replication, and recombination	Homing endonuclease HNH
Barba18A_gp013	350	DNA processing, replication, and recombination	Putative exodeoxyribonuclease
Barba18A_gp022	730	DNA processing, replication, and recombination	DNA polymerase
Barba18A_gp023	631	DNA processing, replication, and recombination	DNA primase
Barba18A_gp030	164	Host signaling	macro domain protein
Barba18A_gp071	271	DNA processing, replication, and recombination	DNA adenine methylase
Barba18A_gp072	228	Peptidase	CLP ₂ protease
Barba18A_gp073	262	Phosphate metabolism	<i>phoH</i>
Barba18A_gp082	102	Structural and packaging	DNA-binding motif containing protein
Barba18A_gp090	205	Structural and packaging	Putative DNA binding protein
Barba18A_gp095	455	Structural and packaging	Terminase large subunit
Barba18A_gp096	491	Structural and packaging	Putative portal protein
Barba18A_gp097	469	Peptidase	Peptidase
Barba18A_gp099	352	Structural and packaging	Major capsid protein
Barba18A_gp104	359	Structural and packaging	Tail sheath protein
Barba18A_gp105	154	Structural and packaging	Putative structural protein
Barba18A_gp108	213	Structural and packaging	Putative DNA binding protein
Barba18A_gp109	566	Structural and packaging	Tail length tape measure protein
Barba18A_gp113	222	Structural and packaging	Putative baseplate assembly protein
Barba18A_gp116	504	Structural and packaging	Baseplate
Barba18A_gp118	220	Structural and packaging	Tail fiber
Barba18A_gp119	414	Structural and packaging	Tail fiber
Barba18A_gp123	121	Peptidase	Peptidase M15
Barba19A_gp006	82	Nucleotide metabolism and recycling	Glutaredoxin
Barba19A_gp007	190	Host signaling	<i>mazG</i>
Barba19A_gp009	209	DNA processing, replication, and recombination	Putative DNA binding protein
Barba19A_gp010	338	DNA processing, replication, and recombination	Metallo-dependent phosphatase
Barba19A_gp033	496	Pyridine nucleotide salvage	Nicotinate phosphoribosyltransferase
Barba19A_gp034	268	Pyridine nucleotide salvage	Ribose-phosphate pyrophosphokinase
Barba19A_gp035	167	Host signaling	Macrodomain protein
Barba19A_gp061	322	DNA processing, replication, and recombination	Nucleotidyltransferase
Barba19A_gp063	419	DNA processing, replication, and recombination	tRNA nucleotidyltransferase
Barba19A_gp098	219	Structural and packaging	Putative DNA binding protein
Barba19A_gp103	427	Structural and packaging	Terminase large subunit
Barba23A_gp115	374	DNA processing, replication and recombination	ISAs1 family transposase
Barba55_gp006	303	Nucleotide metabolism and recycling	Thymidylate synthase
Barba55_gp080	254	Phosphate metabolism	<i>phoH</i>

^aaa, amino acids.^bGenes in the "Structural and packaging" group belong to the structural module, and the other functional groups belong to the host interaction module. For details, see Table S5 in the supplemental material.

August and September samples and in one July sample (Fig. 5). The seasonal dynamics of host abundance in the Baltic Sea was determined by quantification of a 16S rRNA gene amplicon sequence variant (ASV) with 100% sequence identity to *Rheinheimera* sp. strain BAL341. These results showed increasing relative host abundances in June to August between 2012 and 2015, which largely mirrored or slightly preceded the phage dynamics (Fig. 5).

DISCUSSION

In this study, we describe a previously unknown phage genus, all of whose members infected the genome-sequenced bacterial isolate *Rheinheimera* sp. strain BAL341. Genetically, the phage genes showed highest similarity to different gammaproteobacteria or phages infecting gammaproteobacteria. Several of the best viral matches, in particular among genes within the structural module, were to myoviruses (see Table S5 in the supplemental material), which, together with the morphological characterization, placed this genus within the family of *Myoviridae* (Fig. 3a). The head and tail width were

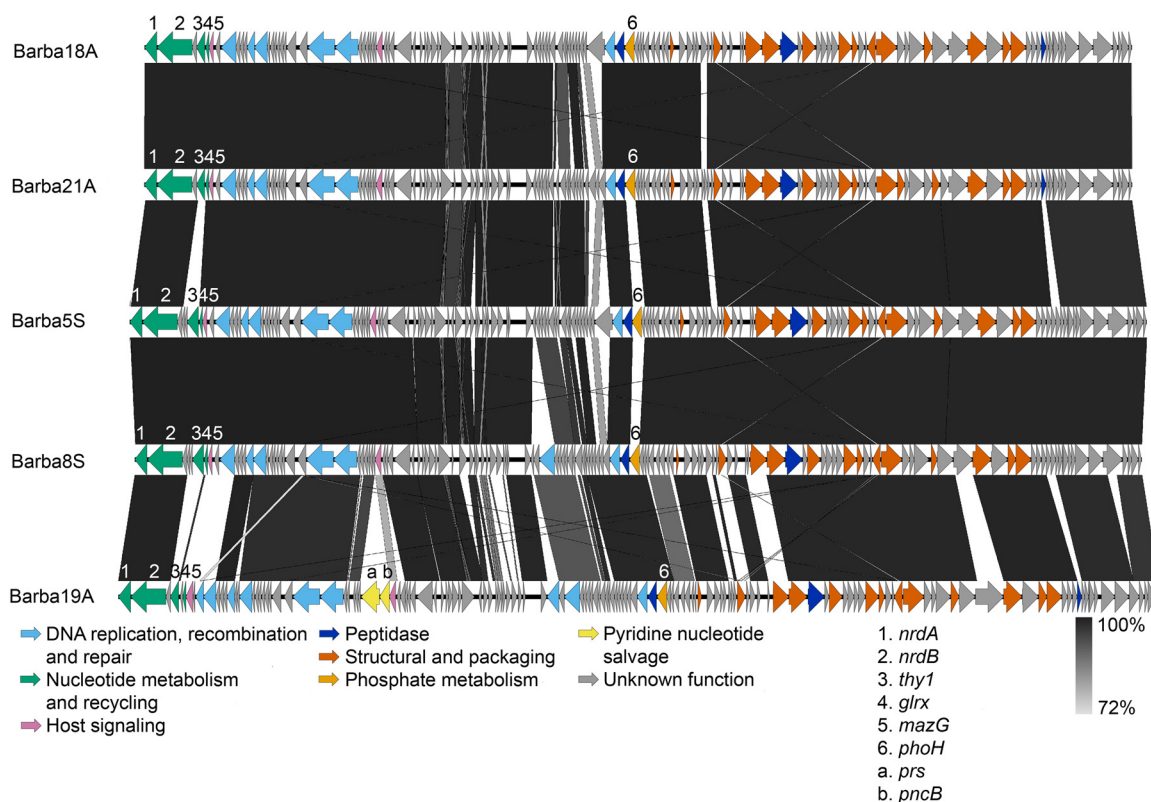


FIG 4 Comparison of genomes from the type phages of the five species within the novel genus *Barbavirus*. Arrows indicate predicted genes, colored by functionality, and the shade of gray between the genomes indicates percent nucleotide identity (blastn), according to the legend on the figure. The metabolic genes and *mazG* discussed in the text are indicated with numbers (1 to 6) or letters (a and b) if they exist in all species or only one, respectively. *nrdA*, ribonucleotide reductase alpha subunit; *nrdB*, ribonucleotide reductase beta subunit; *thy1*, thymidylate synthase; *glrx*, glutaredoxin; *mazG*, nucleoside triphosphate pyrophosphohydrolase; *phoH*, protein PhoH; *prs*, ribose-phosphate pyrophosphokinase; *pncB*, nicotinate phosphoribosyltransferase. The figure was created with EasyFig (106).

of average size compared to the size of other myoviruses, whereas the tail length was relatively short (44, 45). Replication studies of the type phage of the genus, barba18A, and the type phage of the most divergent species, barba19A, showed a burst size (162 and 101 phages, respectively) and latent period (50 and 80 min, respectively) within ranges that are commonly seen among marine phages (17, 25, 46–49). These myoviruses do not fall into any previously known genera and do not share enough sequence similarity to the most similar phages to be able to determine which are their closest evolutionary relatives (Fig. 2). Therefore, we propose that this novel phage genus should be named *Barbavirus*.

All 54 barba phages belonged to the same genus, and thus their diversity appeared limited compared to the genetic diversity seen among phages infecting, e.g., human-associated *Escherichia coli* and *Pseudomonas aeruginosa* as well as the Baltic Sea bacterium *Cellulophaga baltica*, where phages belonging to multiple genera and even different families have been isolated (50). Potentially, the limited diversity within the barba phages could be due to the use of a single host strain for isolation compared to, e.g., 20 host strains used for the *C. baltica* system (51). This limited diversity is also known from other phage-host systems where single host strains have been used for isolation of phages, e.g., among phages infecting *Roseobacter* strain SIO67 (52) and *Bacillus thuringiensis* (53). However, it should be taken into consideration that a smaller number of phages were isolated for these systems than of barba phages for the *Rheinheimeria* model system. On the other hand, multiple phage genera and even phages of different families have been isolated for other model systems using an individual host strain (47, 54–56) also when few phages were isolated (47). Thus, the use of one individual host strain for isolation cannot be the sole explanation for the limited

TABLE 2 Alignment of genes with metabolic functions within the pangenome of the barba phages^a

Function	Common gene			Rare gene			Best matching host gene against common gene ^d			Best matching host gene against rare gene ^d					
	No. of isolates	Gene name	No. of isolates	Gene name	E value	% identity	% coverage	Gene name	E value	% identity	% coverage	Gene name	E value	% identity	% coverage
Thymidylate synthase	52	Barba18A_gp004	2	Barba55_gp006	5.00E-20	43	92	fig 67575.7.peg.3030	8.00E-10	23	29	fig 67575.7.peg.2541	4.00E-05	25	96
Glutaredoxin	52	Barba18A_gp005	2	Barba19A_gp006	1.00E-42	31	94	fig 67575.7.peg.2555	6.00E-04	27	88	fig 67575.7.peg.1895	4.00E-31	35	80
Protein PhoH	52	Barba18A_gp073	2	Barba55_gp080	5.00E-26	28	83	fig 67575.7.peg.2145	5.00E-26	28	83	fig 67575.7.peg.2145	4.00E-31	35	80
Ribonucleotide reductase <i>nrdB</i>	54	Barba18A_gp001	— ^b		4.00E-06	22	54	fig 67575.7.peg.2913	4.00E-06	22	54				
Ribonucleotide reductase <i>nrdA</i>	54	Barba18A_gp002	— ^b		8.00E-10	23	29	fig 67575.7.peg.2912	8.00E-10	23	29				
Nicotinate phosphoribosyltransferase	2	Barba19A_gp033	— ^c					fig 67575.7.peg.3024							
Ribose-phosphate pyrophosphokinase	2	Barba19A_gp034	— ^c		2.00E-08	27	79	fig 67575.7.peg.1160	2.00E-08	27	79				

^aIsolate barba18A was used as the subject in phage-phage alignments where the rare gene variant was compared to the common gene, while the host was used as the query for the phage-host alignment. Only significant alignments (E value of <0.001) are included. Alignment was performed using blastp.

^bAll phages have the same gene.

^cNo other phages had this gene.

^dDetermined by blastp.

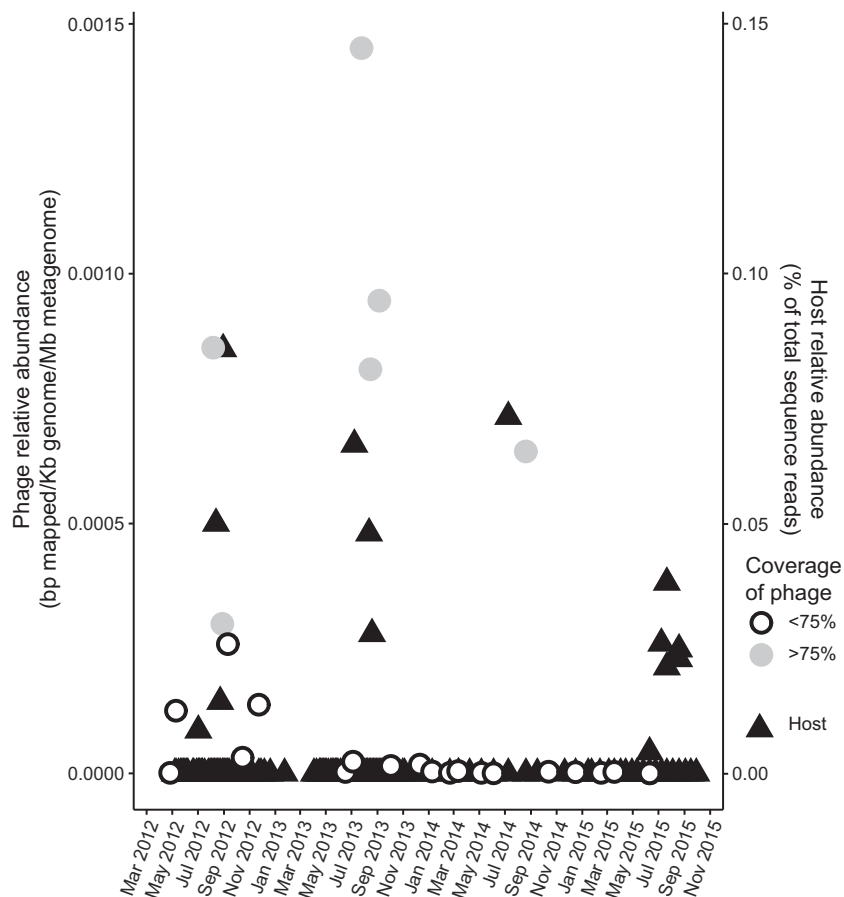


FIG 5 Temporal variation of barba19A and relative abundances of 16S rRNA gene amplicon sequences similar to those of the bacterial host *Rheinheimeria* sp. strain BAL341. Circles indicate phage coverage within the sample.

diversity noted among the barba phages. Instead, this isolation of only highly similar phages might suggest that the barba phages are the dominant viral predators for *Rheinheimeria* sp. strain BAL341, as observed during late summer 2015. Further, the high nucleotide similarity among the barba phages might be a result of ecological constraints, such as the reliance on host genes, imposed on the phages while infecting *Rheinheimeria* sp. strain BAL341.

The majority (78%) of the pangome of the barba phages could not be given a predicted function, but the genes for which a functional annotation could be provided showed a genome organization within the barba phages that is typical among phages (22). One module contained genes involved in host interactions, and one module contained genes involved in phage structure and packaging. The host interaction module contained a few genes involved in DNA replication, for example, DNA primase, DNA polymerase, and DNA binding protein. Such genes are commonly found in phages (52, 55) and are involved in the replication of the phage DNA. However, several functions needed for DNA replication were not detected through amino acid similarity searches; potentially, the phage could use host genes for those functions (57, 58), or these functions are among the large number of genes without a predicted function.

The barba phage genomes contained several genes potentially involved in metabolic processes, which are of interest since they can increase the amount of phage progeny (59). Genes involved in nucleotide metabolism, e.g., RNR and thymidylate synthase, are commonly detected within phage genomes and viral metagenomes (60–62). This includes aquatic myoviruses, for example, T4-like cyanophages (61, 63) and *Vibrio* phage KVP40 (13). All barba phages have a complete class I RNR, consisting

of the *nrdA* and *nrdB* genes, a glutaredoxin gene, and thymidylate synthase, which are all clustered close together in the genome (Fig. 4). The proteins encoded by the RNR genes *nrdA* and *nrdB* form an enzyme which reduces ribonucleotides to deoxyribonucleotides (64). The class I enzyme is thereafter reduced to its active form by either glutaredoxin or thioredoxin (64), which for the barba phages would suggest the use of the phage-encoded glutaredoxin. Another key part of nucleotide metabolism is thymidylate synthase, the enzyme that converts deoxyuridine 5'-monophosphate to deoxythymidine 5'-monophosphate (65) and is thus an important function that allows phages to be self-sustaining on the nucleotide dTTP. The barba phage gene is most similar to *thy1* (*thyX*), which has the greatest efficiency in the presence of FAD (65). Thymidylate synthase is commonly found among other marine phages, in particular, in relatively large (>70 kb) T4-like cyanomyophages (14), N4-like *Roseobacter* phages (66), and large siphon- and podophages infecting *C. baltica* (50). These genes, thymidylate synthase, RNRs, and glutaredoxin, all potentially involved in different parts of the nucleotide metabolism, would provide the barba phages with the potential to utilize their own genes to acquire the metabolites they need for DNA replication if the host nucleotide metabolism is not sufficient or is downregulated during phage infection.

The *Rheinheimera* virus Barba19A isolates also contained two additional genes potentially involved in pyridine nucleotide salvage, encoding a ribose-phosphate pyrophosphokinase and a nicotinate phosphoribosyltransferase (NAPRTase) (67). Ribose-phosphate pyrophosphokinase is an enzyme that transforms ribose 5-phosphate and ATP to phosphoribosyl pyrophosphate (PRPP) and AMP (68). PRPP is a substrate in several enzymatic processes in bacteria (69) and is a precursor of purine and pyrimidine nucleotides, as well as the pyridine nucleotide NAD⁺ (70). Purine and pyrimidine are used by RNR to produce the nucleotides that are needed during the replication of the phage, whereas within the pyridine nucleotide salvage pathway, PRPP and nicotinic acid are converted by NAPRTase to nicotinate mononucleotide and pyrophosphate, most efficiently by hydrolysis of ATP (67). Nicotinate mononucleotide can then be converted to NAD⁺, which is involved in amino acid catabolism and DNA ligase reactions and is a key coenzyme in redox reactions where its reduced form (NADH) transfers electrons to the electron transport chain (71). However, the function within phages has not been clarified. A complete pyridine nucleotide salvage pathway, with *nadV* and *natV*, is functional in *Vibrio* phage KVP40, where *nadV* has functional similarity to NAPRTase (72). The *nadV* gene has also been found in other phages, including another *Vibrio* phage (73) and a marine *Pseudoalteromonas* phage (74). The reaction of *nadV* is followed by the conversion of nicotinamide mononucleotide to NAD⁺ by *natV* in the presence of ATP (72), for which a homologue is missing in *Rheinheimera* virus Barba19A. While the particular functionalities of the ribose-phosphate pyrophosphokinase and the NAPRTase in the *Rheinheimera* virus Barba19A phages have not been verified, these extra metabolic gene products could be hypothesized to provide this species a fitness advantage over the other barba species by increasing the efficiency of the RNRs through redox reactions. However, this is not evident for replication on *Rheinheimera* sp. strain BAL341, where barba19A had a significantly smaller burst size and longer latent period than barba18A. It should be taken into consideration that during the replication experiment, the nutrient supply was high and did not correspond to *in situ* conditions, rendering it hard to detect the increased metabolic potential of barba19A that might be evident as a fitness advantage in the Baltic Sea.

All barba phage genomes contained a *mazG* homologue, even though the *Rheinheimera* virus Barba19A gene showed no sequence similarity to the gene in the other phages (Tables 1 and 2). MazG is predicted to halt self-programmed cell death (75) and has been hypothesized to be used by the phage to keep the host alive during phage propagation (76). A functional MazG acts by decreasing the cellular pool of guanosine 3',5'-bispyrophosphate (ppGpp), which changes the metabolism of the cell as a stress response, that in turn will halt the toxic effects of MazF by resynthesizing MazE (75). Bioinformatic analysis has suggested that MazG is overrepresented among marine phages (22), such as the marine *Pseudoalteromonas* phage H105/1 (22), Roseophage

SIO1 (52), and T4-like cyanophages (77). Therefore, MazG has been suggested to have an important role in marine phage systems (22). Also, *phoH* was detected in all barba phages even though the *phoH* gene in *Rheinheimera* virus Barba5S differed from the version found in the other barba phages. In *E. coli*, *phoH* is part of the Pho regulon, where it is expressed during phosphate limitation, potentially to transport or use phosphate (78). However, a bioinformatic approach to investigate the function of PhoH suggests that the protein could also be involved in phospholipid metabolism and RNA modification or fatty acid beta-oxidation (79). *phoH* has been shown to be upregulated during late infection of the cyanobacterium *Prochlorococcus* strain MED4 by cyanophage P-SSP7 (80). This upregulation is speculated to be a reaction to stress in the host due to phage infection (80) or a consequence of P limitation caused by phage production (27), where an increased uptake of phosphorus would increase phage replication success. Even though the exact mechanism of the *phoH* gene is not fully determined, the gene appears to be more important for marine phages than for phages from other environments; the *phoH* gene is present in 40% of sequenced marine phages in contrast to 4% of phages from other environments (81). The presence of both *mazG* and *phoH* in the phage genome might be indicative of adaptations to counter the host response to stressors caused by the phage infection and might increase the quantity of newly created phages.

Phage metabolic genes can be used to increase replication success and might originate from the host's genome. However, with respect to the barba phages, their metabolic genes were distantly related to genes of similar function occurring in the host *Rheinheimera* sp. strain BAL341 (Table 2). For example, the RNR genes within the barba phages shared low sequence similarity to the host's RNR genes (23% for *nrpA* and 22% for *nrpB*) (Table 2). In general, phages have the same RNRs as their host, but the opposite has also been noted in several instances (60, 82) and has been suggested to be a sign of either horizontal gene transfer within phages with extended host range or of a shifted host range in phages over longer time periods (83), which potentially could be the case for these genes in the barba phages.

The phages isolated in August and September were highly similar at both the nucleotide and amino acid levels, and there were no obvious indications that this phage population had changed genetically between August and September. Within the temporal viral metagenomes from 25 different samplings between 2012 and 2015, the barba phages, represented by barba19A, were detected from July until September, when their host was either abundant or had just declined (Fig. 5). The relative abundance of barba19A within this viral metagenomic data set was low compared to that of investigated uncultured phages within marine metagenomes (5), but the recurring presence of barba19A during the entire investigated time frame indicates that the barba phages form a stable part of the Baltic Sea phage community in late summer (Fig. 5).

While barba19A was the species that was covered by the largest number of viral metagenomic reads during a competitive recruitment (Table S6) and therefore represented the abundance of the barba phages in the temporal data set (Fig. 5), *Rheinheimera* virus Barba18A had the most associated isolates (Fig. 1). This could be due to a shift in the phage community whereby *Rheinheimera* virus Barba19A dominated the barba phage population during the earlier years, as seen by coverage of viral metagenome reads, while *Rheinheimera* virus Barba18A was the dominant species at the time points of isolation and, hence, had a large number of associated isolates. On the other hand, *Rheinheimera* sp. strain BAL341 may not be the optimal or preferred host for the wild-type *Rheinheimera* virus Barba19A phages, and therefore not as many plaques were retrieved as for *Rheinheimera* virus Barba18A. This could explain the reduced replication success for barba19A when it replicates on BAL341 compared to that of barba18A (Fig. 3b). Similar behavior is known from aquatic *Cellulophaga* phages replicating on nonoptimal hosts (16, 17). Efficiency of infection can be another indication of host suitability, but based on the one-step growth curves (Fig. 3b, difference between total and free phages before the burst), efficiency levels were similar for

barba18A and barba19A. Yet efficiency of infection is highly linked to the host bacterium that was latest infected (16), and thus a high efficiency of infection of our isolated barba19A, which has been passed through *Rheinheimera* sp. strain BAL341 multiple times, might not be representative of the efficiency of infection of the wild-type barba19A. However, no alternative host could be detected for barba18A and barba19A within our bacterial culture collection, but it should be noted that the range of suitable alternative hosts was limited. Still, *Rheinheimera* virus Barba19A is the species that could be detected in the Baltic Sea during several consecutive years (Fig. 5; Table S6), and potentially their additional metabolic genes provide them with the fitness advantage to make them the dominant member of the *Barbavirus* population.

Whether or not barba18A or barba19A was the dominant barba phage species during 2015, these replication differences and the temporal dynamics of phages and their host seen over multiple years are indicative of a within-genus phage variation, where different phage species will have varied impacts on their host strains. The pattern of increased phage abundance that coincides with the increase of host abundance, measured as relative abundance of 16S rRNA gene amplicon reads (Fig. 5), is representative of the Lotka-Volterra-like phage-host oscillations suggested in the “kill-the-winner” hypothesis (84). The yearly, reoccurring increases of *Rheinheimera* sp. strain BAL341 and its phages after the summer cyanobacteria bloom (38) point toward the ecological importance of this phage-host system connected to bloom degradation or the exudates excreted during bloom declines. Given its ecological relevance and the high number of metabolic genes, the *Rheinheimera* phage-host system represents an important novel model system for hypothesis-driven research, such as describing different genes of unknown function, how the expression of metabolic genes during infection under various growth conditions influences phage progeny, and how the phage replication success is affected by different environmental factors, such as temperature or nutrient concentrations. These are important aspects to consider for an increased understanding of phage-host interactions and the ecological implications of these microbial players in aquatic environments.

MATERIALS AND METHODS

Bacterial isolation. Water for bacterial isolation was collected at a 2-m depth at LMO (56°55.8540'N, 17°3.6420'E), situated 10 km east off the coast of Öland, Sweden, with a Ruttner sampler on 12 July 2012. Bacterial colonies were grown on Zobell agar plates (1 g of yeast extract [Becton, Dickinson and Company (BD)], 5 g of Bacto peptone [BD], and 15 g of Bacto agar [BD] in 800 ml of filtered Baltic Sea water and 200 ml of Milli-Q water) at room temperature for 3 to 4 days, and each isolate was clean streaked three times. Bacterial DNA for sequencing of the 16S rRNA gene (BAL331, BAL335, BAL336, and BAL341) and for whole-genome sequencing (BAL341) was isolated using an EZNA tissue DNA kit (Omega Bio-tek) according to the manufacturer's instructions. Briefly, cells were lysed, protease treated, and bound to a HiBind DNA Mini column. While bound to the column, the DNA was washed with DNA wash buffer diluted with 100% ethanol and then eluted in 10 mM Tris. The 16S rRNA gene was PCR amplified with the primers 27F and 1492R and sequenced using Sanger dideoxy sequencing at MacroGen Europe, Amsterdam, Netherlands. The whole-genome DNA was verified with regard to quality and quantity with Nanodrop (Thermo Scientific) and Qubit (high-sensitivity DNA kit; Invitrogen, Life Technologies), respectively, and sequenced with an Illumina HiSeq system at the Science for Life Laboratory, the National Genomics Infrastructure (SciLife/NGI) (see the paragraph “Sequencing of whole genomes” below).

Viral isolation. Water for viral isolation was collected on 19 August 2015 and 16 September 2015 at LMO. At each time point, 10 liters of water was prefiltered through 0.22- μ m-pore-size Sterivex filters (Millipore), and the water was concentrated using a 30-kDa tangential flow filter (Millipore). The resulting 100 to 150 ml was further concentrated through 50-kDa Amicon Ultra centrifugal filters (Millipore) to 13 to 30 ml.

In order to isolate phages, plaque assays (85) with the two viral concentrates were conducted using *Rheinheimera* sp. strain BAL341 as the host. Here, bacteria grown overnight in liquid Zobell medium (1 g of yeast extract [BD] and 5 g of Bacto peptone [BD] in 800 ml of filtered Baltic Sea water and 200 ml of Milli-Q water) were mixed with the viral concentrates and molten (32°C) top agar (marine sodium magnesium [MSM] buffer consisting of 450 mM NaCl [Sigma], 50 mM MgSO₄·7H₂O [Fisher], and 50 mM Trizma base [Sigma], pH 8, with 0.5% low-melting-point agarose [Thermo Fisher Scientific]). The mixture of bacteria, viral concentrate, and molten agar was spread evenly onto Zobell agar plates. The plates were incubated at room temperature (RT) for 2 days, and plaque formation was monitored on a daily basis. To obtain pure phage isolates, individual plaques (54 in total) were picked with a sterile 100- μ l pipette tip, dispersed in MSM buffer, and replated three times as described above, but now with phages suspended in MSM buffer and not viral concentrates. Pure isolates were harvested by adding 5 ml of MSM buffer to fully lysed plates, the top agar layer was shredded with an inoculation loop, and the plates

were incubated on a shaking table (40 rpm) for at least 30 min. The MSM buffer with suspended phages was collected into a Falcon tube, centrifuged for 10 min at $10,000 \times g$, filtered through a $0.2\text{-}\mu\text{m}$ -pore-size syringe filter (BD), and stored at 4°C .

TEM. Phage particles were purified by cesium chloride (CsCl_2) density gradient centrifugation (described by Carlson and Miller [86]), with modifications as described below, and analyzed by transmission electron microscopy (TEM). Briefly, a high titer of phage lysate ($\sim 10^{11} \text{ ml}^{-1}$) was applied to a CsCl_2 step gradient (densities of 1.1, 0.9, 0.7, and 0.5 g ml^{-1}) and centrifuged in a Spinco SW39 rotor for 3 h at $24,000 \text{ rpm}$ ($280,000 \times g$), 4°C . The band with the highest opalescence was collected with a syringe and dialyzed three times with SM buffer (100 mM NaCl, 8 mM MgSO_4 , 50 mM Tris-HCl, pH 7.5) in dialysis tubes (0.020 mm ; Viskase, USA) for 24 h at 4°C . The suspension was applied to carbon-coated nitrocellulose grids (Agar Scientific Elektron Technology, United Kingdom), stained with two successive drops of 2% uranyl acetate (pH 4.5), dried, and examined using a Morgagni 268(D) transmission electron microscope (FEI, USA). In total, 50 phage particles were measured for determination of capsid size and tail length using the Morgagni integrated image acquisition and analysis software.

Host range. Host range analysis based on efficiency of plating was performed with two species type phages from this study, barba18A and barba19A, on the in-house *Rheinheimera* bacterial strains, *Rheinheimera* sp. BAL341, BAL331, BAL335, and BAL336, isolated from LMO. Plaque assays as described previously were performed in duplicate for each phage-host pair with phage dilutions of 10^{-2} to 10^{-4} -fold, except for *Rheinheimera* sp. strain BAL341, for which the phages were diluted from 10^{-2} to 10^{-9} , and plates were monitored daily for 5 days to detect plaque formation.

One-step growth curve. To determine the exponential growth phase of *Rheinheimera* sp. strain BAL341, the bacterium was grown in 10 ml of Zobell medium in glass tubes at RT in triplicates, and optical density at 600 nm with a CO8000 cell density meter (WPA, Cambridge, United Kingdom) was measured every hour for 14 h. One-step growth curves, providing phage replication characteristics, were performed as described in Holmfeldt et al. (16), with minor modifications. Briefly, exponentially growing bacteria were infected with barba18A or barba19A at a multiplicity of infection (MOI) of 0.1 for 5 min, after which the solution was diluted 1,000-fold in order to prevent new infections. Free and total numbers of phages were enumerated using a plaque assay (described above) every 15 min for 2 h for barba18A and every 20 min for 2 h and then every 30 min for one additional hour for barba19A. To enumerate free (extracellular) phages, subsamples of the diluted infection were filtered through a $0.2\text{-}\mu\text{m}$ -pore-size syringe filter (BD), and the flowthrough was used for the plaque assay. For total phages, including both extracellular phages and phages attached to or replicating inside bacteria, subsamples were retrieved directly from the incubation and used for plaque assay. Plaque formation was examined within 48 h. Statistical analysis and plotting were done with R (version 3.5.1) (87) through RStudio (version 1.1.383) (88) using the packages plyr (version 1.8.4) (89) and ggplot2 (version 3.1.0) (90).

Viral DNA extraction. DNA extraction from the 54 isolated phages was performed with Wizard PCR DNA purification resin and minicolumns (Promega). First, 1 ml of resin was added to 1 ml of the newly harvested, high-titer viral sample, thoroughly mixed, and then pushed through a minicolumn. The column was washed twice using 1 ml of 80% isopropanol, and residues of isopropanol were removed by centrifuging the column for 2 min at $10,000 \times g$. The DNA was then eluted with preheated (80°C) 10 mM Tris (Trizma base [Sigma], pH 8) by first vortexing the column and then centrifuging it at $10,000 \times g$ for 30 s. Quality and quantity of the DNA were verified with Nanodrop (Thermo Scientific) and Qubit (high-sensitivity DNA kit; Invitrogen, Life Technologies), respectively.

Sequencing of whole genomes. Library preparation of DNA from whole-genome extraction of *Rheinheimera* sp. BAL341 and the 54 viral isolates was done with Nextera XT (Illumina, Inc., San Diego, CA, USA), and sequencing was performed by SciLife/NGI (Solna, Sweden) on a HiSeq 2500 system (Illumina, Inc.), which resulted in paired-end 125-bp reads (average of 1.2 million reads per sample).

Bioinformatics analysis. Reads in fastq format were trimmed to remove adapters and poor-quality reads (trimmomatic version 0.30) (settings: -PE -threads 2 -phred33 ILLUMINACLIP:nextera_linkers.txt:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:30) (91), and quality was then evaluated with FastQC (92). Reads were assembled with Spades (version 3.6.0) (settings: -careful -t 8 -pe1-1 -pe1-2 -o) (93) and Abyss (version 1.3.5) (settings: -np 16 -k61 -q3 -coverage-hist -s -o) (94), and the assembler providing the longest contig was chosen. The bacterial genome was annotated with the RAST online service (95, 96). The phage genomes were manually inspected, and gene prediction was performed with GeneMark (version 3.26; gene code 11, heuristic parameters as 2010) (97) and manually curated on completed genomes. Predicted genes were then annotated with Diamond (version 0.8.26) (settings: blastp -p 16 -q -k 500 -min-score 30 -sensitive -tmpdir -f 6 -o) (98) against the NCBI nonredundant database (February 2018) and hmmsearch (version 3.1b2) (settings: -tblout -E 1e-3 -cpu 4) (99) against the Pfam database (100) (February 2018, version 31.0), and an alignment was considered significant if the E value was less than 0.001. Also, to identify the phages with highest similarity to the barba phages, a Diamond search with the same parameters as described above was run against a custom viral database (viraldb) consisting of viral genomes from RefSeq (downloaded 14 September 2018). The results from the search against the viraldb were used to define reference genomes for the VICTOR analyses described below. Searches for tRNA were performed with tRNAscan-SE (settings: -qQ -detail -o# -m# -f# -l# -c tRNAscan-SE.conf -s# -B) (101). The core genome and pangenome were calculated with Roary (102), with settings not to split paralogs and to group genes with more than 70% amino acid similarity (-s -i 70). Whole-genome comparisons on the nucleotide level were done with gegenees (version 3.0.0) (settings: fragment size of 500 and step size of 500; blastn) (103) that fragments the genomes and performs all-against-all blastn alignments before calculation of the average nucleotide identity across the entire genomes. The gegenees results were plotted with reshape2 (version 1.4.3) (104) and Tidyverse (version

1.2.1) (105) in R (version 3.5.1) through RStudio (version 1.1.383). Comparison of gene distributions between a select set of genomes was conducted using EasyFig (106).

Phylogenetic analysis. To distinguish taxonomic classification of the barba phages from that of other viruses, all 54 barba phage isolates and selected reference genomes, based on alignments to the NCBI nr database and viraldb, were analyzed with VICTOR, a genome-to-genome distance calculator (42). Pairwise comparisons of nucleotide sequences were done by using the genome-BLAST distance phylogeny (GBDP) method (107) under settings for the D0 formula (settings: word length, 11; E value filter, 1.0; algorithm, greedy-with-trimming) recommended for prokaryotic viruses (42). The distances were then used to infer a balanced minimum evolution tree with branch support (subtree pruning and regrafting [SPR] branch swapping, 100 pseudobootstrap replicates) via FastME (108). Also, the genomes were clustered at species, genus, and family levels with the OPTSIL program (109) with recommended thresholds (42) and an *F* value (fraction of links required for cluster fusion) of 0.5 (110).

Viral metagenome sampling and recruitment to the phage genomes. At 25 time points between 2012 and 2015 (for dates, see Table S7 in the supplemental material), water was collected at the LMO in the same manner as described above. The water was prefiltered through an 0.22- μm -pore-size filter (Sterivex cartridge filter; Millipore) to remove larger organisms. Viruses were precipitated using the iron chloride method (111), where the flowthrough was treated with FeCl_3 (final concentration, 1 mg liter⁻¹; Sigma) to aggregate the viral particles, which were collected on a 1.0- μm -pore-size filter (polycarbonate; Maine Manufacturing LLC) and stored in falcon tubes at +4°C until DNA extraction. For DNA extraction, the viral particles were resuspended in ascorbate solution (20 ml of 0.5 M EDTA, pH 8, 12.5 ml of 1 M Tris, pH 8, 1.9 g of MgCl_2 , 3.52 g of ascorbic acid, ~4 ml of NaOH, and Milli-Q water up to 100 ml) within 1 to 18 months after sampling and concentrated using Amicon spin filters (Millipore). The viral concentrates were thereafter treated with DNase I (Invitrogen) according to the manufacturer's recommendations and inactivated with EDTA (100 mM; Sigma) by incubation at 65°C for 15 min. The viral particles were treated with preheated (37°C for 30 min) proteinase K (20 mg/ml; Fisher) at 37°C for 12 h, and the DNA was then extracted with a Wizard PCR DNA purification kit as described above. The quality and quantity of DNA were assessed with Qubit (Invitrogen) and Nanodrop (Thermo Scientific), respectively, and sequenced at SciLife/NGI (Solna, Sweden). Nine of the viral metagenomes were sequenced during 2014, and the last 16 samples were sequenced during 2016. Libraries for the first set were prepared using Illumina TruSeq DNA, clustered with cbot, and sequenced with a 2-by-101 setup. For the second set, the libraries were prepared using Illumina TruSeq Nano, clustered with cbot, and sequenced with a 2-by-126 setup (Table S7). Both sets were sequenced on an Illumina HiSeq 2500 system (Illumina, Inc., San Diego, CA, USA). Reads from the 25 viral metagenomes were recruited against barba18A, barba21A, barba5S, barba8S, and barba19A competitively and against barba19A exclusively using Bowtie2 (version 2.3.3.1, with default settings) (112). To calculate coverage and depth, samtools (version 1.9; setting -F4) (113) was used to filter out alignments with low quality (<90% identity to the reference), and bedtools (version 2.27.1) (114) was used to perform calculations of coverage and depth. Read depth was normalized based on the length of the genome in kilobases and the size of the viral metagenome in megabases as previously done for other viral metagenome analyses (5).

Amplicon sequencing and host identification. Bacterial 16S rRNA gene amplicon sequence data for monitoring host abundance were obtained according to Boström et al. (115), as modified by Bunse et al. (116). Briefly, water samples from LMO (2012 to 2015) were collected and either sequentially filtered through 3.0- μm -pore-size, 48-mm-diameter, polycarbonate filters (Pall Life Sciences) and then through 0.2- μm -pore-size Sterivex cartridge filters (Millipore) or directly through Sterivex cartridge filters (Millipore). DNA was extracted using the phenol-chloroform method (117) from all filters, and the V3-V4 region of the 16S rRNA gene was amplified in PCR with the primer pair 341f-805r (118, 119). Illumina adapters were attached during a second PCR, and the samples were sequenced (300 bp; pair end) on the Illumina MiSeq platform (Illumina, Inc., USA) at SciLife/NGI (Solna, Sweden). Raw reads were subsequently processed with the DADA2 pipeline (120) to produce amplicon sequencing variants (ASVs) (filterAndTrim: -trimLeft = 8,8 -truncLen = 290,210, learnErrors: -randomize=TRUE, -nsamples = 24, mergePairs: -minOverlap = 10, isBimeraDenovo: minParentAbundance = 8, minFoldParentAbundance = 4). Using RStudio (version 1.1.383) (88) and the Tidyverse package (version 1.2.1) (105), the ASV sequence matching *Rheinheimera* sp. strain BAL341 (100% nucleotide identity, by blastn) was extracted from the data set from all three filter fractions (3.0 μm , 0.2 μm , and 0.2 μm after 3.0 filtration) and plotted as relative abundances (proportions) over time together with the phage abundances using ggplot2 (90).

Data availability. 16S rRNA gene accession numbers for *Rheinheimera* sp. BAL341, BAL331, BAL335, and BAL336 were deposited in GenBank under accession numbers [KM586890](#), [KM586876](#), [KM586880](#), and [KM586917](#), respectively. Sequences for the barba phages are deposited under accession numbers [MK719701](#) to [MK719754](#).

Raw reads for *Rheinheimera* sp. strain BAL341 were deposited in NCBI under BioProject accession number [PRJEB29737](#). Raw sequence reads for the viral metagenome were deposited under BioProject accession number [PRJNA474405](#). Whole-genome and annotation information for *Rheinheimera* sp. strain BAL341 is available in the EMBL/ENA database under accession number [CAAJGR010000000](#).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/AEM.01003-19>.

SUPPLEMENTAL FILE 1, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

We acknowledge Anders Månsson, Kristofer Bergström, Emil Fridolfsson, and the others in the LMO sampling crew and Ecochange; we also thank Sabina Arnautovic for help with bacterial isolation, DNA extraction, and sequencing, Stina Israelsson for help with viral metagenome sampling, and Diego Brambilla for help with genome submission. We also acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure, and the Swedish National Infrastructure for Computing at Uppmax for helping in massive parallel sequencing and resources for computation.

The project was funded by a grant from the Swedish Research Council (2013-4554) and the Crafoord Foundation (20140681) to K.H.

K.H. and J.P. conceived the study. E.N., K.L., and S.S. isolated and processed viral isolates and carried out viral experiments. E.N. and K.H. prepared and processed viral metagenome samples. M.L. isolated bacteria, and C.B. and M.L. undertook LMO sampling efforts and processed 16S sequencing data. E.N., J.F., C.B., M.L., and D.L. performed bioinformatic analyses. E.N., K.H., C.B., C.M.G.K., D.L., and J.P. analyzed and interpreted the data. E.N. and K.H. wrote the initial draft of the manuscript with the help of S.S., C.M.G.K., and C.B. All authors commented and approved the final manuscript.

We declare that we have no conflicts of interest.

REFERENCES

- Suttle CA. 2007. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol* 5:801–812. <https://doi.org/10.1038/nrmicro1750>.
- Bergh O, Borsheim KY, Bratbak G, Heldal M. 1989. High abundance of viruses found in aquatic environments. *Nature* 340:467–468. <https://doi.org/10.1038/340467a0>.
- Suttle CA. 1994. The significance of viruses to mortality in aquatic microbial communities. *Microb Ecol* 28:237–243. <https://doi.org/10.1007/BF00166813>.
- Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, Carlson C, Chan AM, Haynes M, Kelley S, Liu H, Mahaffy JM, Mueller JE, Nulton J, Olson R, Parsons R, Rayhawk S, Suttle CA, Rohwer F. 2006. The marine viromes of four oceanic regions. *PLoS Biol* 4:e368. <https://doi.org/10.1371/journal.pbio.0040368>.
- Brum JR, Ignacio-Espinoza JC, Roux S, Doucier G, Acinas SG, Alberti A, Chaffron S, Cruaud C, de Vargas C, Gasol JM, Gorsky G, Gregory AC, Guidi L, Hingamp P, Iudicone D, Not F, Ogata H, Pesant S, Poulos BT, Schwenck SM, Speich S, Dimier C, Kandel-Lewis S, Picheral M, Searson S, Tara Oceans Coordinators, Bork P, Bowler C, Sunagawa S, Wincker P, Karsenti E, Sullivan MB. 2015. Ocean plankton. Patterns and ecological drivers of ocean viral communities. *Science* 348:1261498. <https://doi.org/10.1126/science.1261498>.
- Roux S, Hallam SJ, Woyke T, Sullivan MB. 2015. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *Elife* 4:e08490. <https://doi.org/10.7554/eLife.08490>.
- Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. 2016. Computational approaches to predict bacteriophage-host relationships. *FEMS Microbiol Rev* 40:258–272. <https://doi.org/10.1093/femsre/fuv048>.
- Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L, McDaniel L, Moran MA, Nelson KE, Nilsson C, Olson R, Paul J, Brito BR, Ruan Y, Swan BK, Stevens R, Valentine DL, Thurber RV, Wegley L, White BA, Rohwer F. 2008. Functional metagenomic profiling of nine biomes. *Nature* 452:629–632. <https://doi.org/10.1038/nature06810>.
- Hurwitz BL, Hallam SJ, Sullivan MB. 2013. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol* 14:R123. <https://doi.org/10.1186/gb-2013-14-11-r123>.
- Labonté JM, Swan BK, Poulos B, Luo H, Koren S, Hallam SJ, Sullivan MB, Woyke T, Wommack KE, Stepanauskas R. 2015. Single-cell genomics-based analysis of virus-host interactions in marine surface bacterioplankton. *ISME J* 9:2386–2399. <https://doi.org/10.1038/ismej.2015.48>.
- Deng L, Ignacio-Espinoza JC, Gregory AC, Poulos BT, Weitz JS, Hugenholtz P, Sullivan MB. 2014. Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space. *Nature* 513:242–245. <https://doi.org/10.1038/nature13459>.
- Mann NH, Cook A, Millard A, Bailey S, Clokie M. 2003. Marine ecosystems: bacterial photosynthesis genes in a virus. *Nature* 424:741. <https://doi.org/10.1038/424741a>.
- Miller ES, Heidelberg JF, Eisen JA, Nelson WC, Durkin AS, Ciecko A, Feldblyum TV, White O, Paulsen IT, Nierman WC, Lee J, Szczepinski B, Fraser CM. 2003. Complete genome sequence of the broad-host-range vibriophage KVP40: comparative genomics of a T4-related bacteriophage. *J Bacteriol* 185:5220–5233. <https://doi.org/10.1128/jb.185.17.5220-5233.2003>.
- Sullivan MB, Coleman ML, Weigle P, Rohwer F, Chisholm SW. 2005. Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biol* 3:e144. <https://doi.org/10.1371/journal.pbio.0030144>.
- Breitbart M, Thompson LR, Suttle CA, Sullivan MB. 2007. Exploring the vast diversity of marine viruses. *Oceanography* 20:135–139. <https://doi.org/10.5670/oceanog.2007.58>.
- Holmfeldt K, Solonenko N, Howard-Varona C, Moreno M, Malmstrom RR, Blow MJ, Sullivan MB. 2016. Large-scale maps of variable infection efficiencies in aquatic *Bacteroidetes* phage-host model systems. *Environ Microbiol* 18:3949–3961. <https://doi.org/10.1111/1462-2920.13392>.
- Holmfeldt K, Howard-Varona C, Solonenko N, Sullivan MB. 2014. Contrasting genomic patterns and infection strategies of two co-existing *Bacteroidetes* podovirus genera. *Environ Microbiol* 16:2501–2513. <https://doi.org/10.1111/1462-2920.12391>.
- Dang VT, Howard-Varona C, Schwenck S, Sullivan MB. 2015. Variably lytic infection dynamics of large *Bacteroidetes* podovirus phi38:1 against two *Cellulophaga baltica* host strains. *Environ Microbiol* 17:4659–4671. <https://doi.org/10.1111/1462-2920.13009>.
- Howard-Varona C, Roux S, Dore H, Solonenko NE, Holmfeldt K, Markillie LM, Orr G, Sullivan MB. 2017. Regulation of infection efficiency in a globally abundant marine *Bacteroidetes* virus. *ISME J* 11:284–295. <https://doi.org/10.1038/ismej.2016.81>.
- Howard-Varona C, Hargreaves KR, Solonenko NE, Markillie LM, White RA, Brewer HM, Ansong C, Orr G, Adkins JN, Sullivan MB. 2018. Multiple mechanisms drive phage infection efficiency in nearly identical hosts. *ISME J* 12:1605–1618. <https://doi.org/10.1038/s41396-018-0099-8>.
- Rodrigues RAL, Andrade A, Boratto PVM, Trindade GS, Kroon EG, Abrahao JS. 2017. An anthropocentric view of the virosphere-host relationship. *Front Microbiol* 8:1673. <https://doi.org/10.3389/fmicb.2017.01673>.
- Duhaime MB, Wichels A, Waldmann J, Teeling H, Glöckner FO. 2011. Ecogenomics and genome landscapes of marine *Pseudoalteromonas* phage H105/1. *ISME J* 5:107–121. <https://doi.org/10.1038/ismej.2010.94>.
- Lara E, Holmfeldt K, Solonenko N, Sà EL, Ignacio-Espinoza JC, Cornejo-Castillo FM, Verberkmoes NC, Vaqué D, Sullivan MB, Acinas SG. 2015. Life-style and genome structure of marine *Pseudoalteromonas* siphovirus B8b isolated from the northwestern Mediterranean Sea. *PLoS One* 10:e0114829. <https://doi.org/10.1371/journal.pone.0114829>.

24. Wichels A, Gerdt G, Schütt C. 2002. *Pseudoalteromonas* spp. phages, a significant group of marine bacteriophages in the North Sea. *Aquat Microb Ecol* 27:233–239. <https://doi.org/10.3354/ame027233>.
25. Baudoux AC, Hendrix RW, Lander GC, Bailly X, Podell S, Paillard C, Johnson JE, Potter CS, Carragher B, Azam F. 2012. Genomic and functional analysis of *Vibrio* phage SIO-2 reveals novel insights into ecology and evolution of marine siphoviruses. *Environ Microbiol* 14:2071–2086. <https://doi.org/10.1111/j.1462-2920.2011.02685.x>.
26. Comeau AM, Chan AM, Suttle CA. 2006. Genetic richness of vibriophages isolated in a coastal environment. *Environ Microbiol* 8:1164–1176. <https://doi.org/10.1111/j.1462-2920.2006.01006.x>.
27. Duhaime MB, Solonenko N, Roux S, Verberkmoes NC, Wichels A, Sullivan MB. 2017. Comparative omics and trait analyses of marine *Pseudoalteromonas* phages advance the phage OTU concept. *Front Microbiol* 8:1241. <https://doi.org/10.3389/fmicb.2017.01241>.
28. Yu ZC, Chen XL, Shen QT, Zhao DL, Tang BL, Su HN, Wu ZY, Qin QL, Xie BB, Zhang XY, Yu Y, Zhou BC, Chen B, Zhang YZ. 2015. Filamentous phages prevalent in *Pseudoalteromonas* spp. confer properties advantageous to host survival in Arctic sea ice. *ISME J* 9:871–881. <https://doi.org/10.1038/ismej.2014.185>.
29. Grose JH, Casjens SR. 2014. Understanding the enormous diversity of bacteriophages: the tailed phages that infect the bacterial family *Enterobacteriaceae*. *Virology* 468–470:421–443. <https://doi.org/10.1016/j.virol.2014.08.024>.
30. Kwan T, Liu J, Dubow M, Gros P, Pelletier J. 2006. Comparative genomic analysis of 18 *Pseudomonas aeruginosa* bacteriophages. *J Bacteriol* 188:1184–1187. <https://doi.org/10.1128/JB.188.3.1184-1187.2006>.
31. Ryu SH, Chung BS, Park M, Lee SS, Lee SS, Jeon CO. 2008. *Rheinheimera soli* sp. nov., a gammaproteobacterium isolated from soil in Korea. *Int J Syst Evol Microbiol* 58:2271–2274. <https://doi.org/10.1099/ijs.0.65489-0>.
32. Chen WM, Lin CY, Young CC, Sheu SY. 2010. *Rheinheimera aquatica* sp. nov., antimicrobial activity-producing bacterium isolated from freshwater culture pond. *J Microbiol Biotechnol* 20:1386–1392. <https://doi.org/10.4014/jmb.1004.04048>.
33. Yoon J-H, Bae SE, Park SE, Kang S-J, Oh T-K. 2007. *Rheinheimera aquimaris* sp. nov., isolated from seawater of the East Sea in Korea. *Int J Syst Evol Microbiol* 57:1386–1390. <https://doi.org/10.1099/ijs.0.64898-0>.
34. Schröder J, Braun B, Liere K, Szewzyk U. 2016. Draft genome sequence of *Rheinheimera* sp. strain SA_1 isolated from iron backwash sludge in Germany. *Genome Announc* 4:e00853-16. <https://doi.org/10.1128/genomeA.00853-16>.
35. Karlsson CM, Cerro-Gálvez E, Lundin D, Karlsson C, Vila-Costa M, Pinhassi J. 4 July 2019. Direct effects of organic pollutants on the growth and gene expression of the Baltic Sea model bacterium *Rheinheimera* sp. BAL341. *Microb Biotechnol* <https://doi.org/10.1111/1751-7915.13441>.
36. Brettar I, Christen R, Höfle MG. 2002. *Rheinheimera baltica* gen. nov., sp. nov., a blue-coloured bacterium isolated from the central Baltic Sea. *Int J Syst Evol Microbiol* 52:1851–1857. <https://doi.org/10.1099/00207713-52-5-1851>.
37. Lindh MV, Figueroa D, Sjöstedt J, Baltar F, Lundin D, Andersson A, Legrand C, Pinhassi J. 2015. Transplant experiments uncover Baltic Sea basin-specific responses in bacterioplankton community composition and metabolic activities. *Front Microbiol* 6:223. <https://doi.org/10.3389/fmicb.2015.00223>.
38. Bunse C, Israelsson S, Baltar F, Bertos-Fortis M, Fridolfsson E, Legrand C, Lindehoff E, Lindh MV, Martínez-García S, Pinhassi J. 2018. High frequency multi-year variability in Baltic Sea microbial plankton stocks and activities. *Front Microbiol* 9:3296. <https://doi.org/10.3389/fmicb.2018.03296>.
39. Lindh MV, Sjöstedt J, Andersson AF, Baltar F, Hugert LW, Lundin D, Muthusamy S, Legrand C, Pinhassi J. 2015. Disentangling seasonal bacterioplankton population dynamics by high-frequency sampling. *Environ Microbiol* 17:2459–2476. <https://doi.org/10.1111/1462-2920.12720>.
40. Kropinski AM, Prangishvili D, Lavigne R. 2009. Position paper: the creation of a rational scheme for the nomenclature of viruses of *Bacteria* and *Archaea*. *Environ Microbiol* 11:2775–2777. <https://doi.org/10.1111/j.1462-2920.2009.01970.x>.
41. Adriaenssens EM, Brister JR. 2017. How to name and classify your phage: an informal guide. *Viruses* 9:E70. <https://doi.org/10.3390/v904070>.
42. Meier-Kolthoff JP, Göker M. 2017. VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* 33:3396–3404. <https://doi.org/10.1093/bioinformatics/btx440>.
43. Roux S, Emerson JB, Eloe-Fadrosh EA, Sullivan MB. 2017. Benchmarking viromics: an in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. *PeerJ* 5:e3817. <https://doi.org/10.7717/peerj.3817>.
44. King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed). 2011. *Virus taxonomy. Classification and nomenclature of viruses. Ninth report of the International Committee on Taxonomy of Viruses*. Elsevier Academic Press, San Diego, CA.
45. Brum JR, Schenck RO, Sullivan MB. 2013. Global morphological analysis of marine viruses shows minimal regional variation and dominance of non-tailed viruses. *ISME J* 7:1738–1751. <https://doi.org/10.1038/ismej.2013.67>.
46. Jiang SC, Kellogg CA, Paul JH. 1998. Characterization of marine temperate phage-host systems isolated from Mamala Bay, Oahu, Hawaii. *Appl Environ Microbiol* 64:535–542.
47. Zhao Y, Temperton B, Thrash JC, Schwalbach MS, Vergin KL, Landry ZC, Ellisman M, Deerinck T, Sullivan MB, Giovannoni SJ. 2013. Abundant SAR11 viruses in the ocean. *Nature* 494:357–360. <https://doi.org/10.1038/nature11921>.
48. Suttle CA, Chan AM. 1993. Marine cyanophages infecting oceanic and coastal strains of *Synechococcus*: abundance, morphology, cross-infectivity and growth characteristics. *Mar Ecol Prog Ser* 92:99–109. <https://doi.org/10.3354/meps092099>.
49. Hadas H, Einav M, Fishov I, Zaritsky A. 1997. Bacteriophage T4 development depends on the physiology of its host *Escherichia coli*. *Microbiology* 143:179–185. <https://doi.org/10.1099/00221287-143-1-179>.
50. Holmfeldt K, Solonenko N, Shah M, Corrier K, Riemann L, Verberkmoes NC, Sullivan MB. 2013. Twelve previously unknown phage genera are ubiquitous in global oceans. *Proc Natl Acad Sci U S A* 110:12798–12803. <https://doi.org/10.1073/pnas.1305956110>.
51. Holmfeldt K, Middelboe M, Nybroe O, Riemann L. 2007. Large variabilities in host strain susceptibility and phage host range govern interactions between lytic marine phages and their *Flavobacterium* hosts. *Appl Environ Microbiol* 73:6730–6739. <https://doi.org/10.1128/AEM.01399-07>.
52. Angly F, Youle M, Nosrat B, Srinagesh S, Rodriguez-Brito B, McNairnie P, Deyanat-Yazdi G, Breitbart M, Rohwer F. 2009. Genomic analysis of multiple Roseophage SIO1 strains. *Environ Microbiol* 11:2863–2873. <https://doi.org/10.1111/j.1462-2920.2009.02021.x>.
53. Sauder AB, Quinn MR, Brouillette A, Caruso S, Cresawn S, Erill I, Lewis L, Loesser-Casey K, Pate M, Scott C, Stockwell S, Temple L. 2016. Genomic characterization and comparison of seven *Myoviridae* bacteriophage infecting *Bacillus thuringiensis*. *Virology* 489:243–251. <https://doi.org/10.1016/j.virol.2015.12.012>.
54. Gregory AC, Solonenko SA, Ignacio-Espinoza JC, LaButti K, Copeland A, Sudek S, Maitland A, Chittick L, Dos Santos F, Weitz JS, Worden AZ, Woyke T, Sullivan MB. 2016. Genomic differentiation among wild cyanophages despite widespread horizontal gene transfer. *BMC Genomics* 17:930. <https://doi.org/10.1186/s12864-016-3286-x>.
55. Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, Weber RJ, Patel MC, Germane KL, Edgar RH, Hoyte NN, Bowman CA, Tantoco AT, Paladin EC, Myers MS, Smith AL, Grace MS, Pham TT, O'Brien MB, Vogelsberger AM, Hryckowian AJ, Wynalek JL, Donis-Keller H, Bogel MW, Peebles CL, Cresawn SG, Hendrix RW. 2010. Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. *J Mol Biol* 397:119–143. <https://doi.org/10.1016/j.jmb.2010.01.011>.
56. Pope WH, Jacobs-Sera D, Russell DA, Peebles CL, Al-Atrache Z, Alcoser TA, Alexander LM, Alfano MB, Alford ST, Amy NE, Anderson MD, Anderson AG, Ang AAS, Ares M, Barber AJ, Barker LP, Barrett JM, Barshop WD, Bauerle CM, Bayles IM, Belfield KL, Best AA, Borjon A, Bowman CA, Boyer CA, Bradley KW, Bradley VA, Broadway LN, Budwal K, Busby KN, Campbell IW, Campbell AM, Carey A, Caruso SM, Chew RD, Cockburn CL, Cohen LB, Corajod JM, Cresawn SG, Davis KR, Deng L, Denver DR, Dixon BR, Ekram S, Elgin SCR, Engelsen AE, English BEV, Erb ML, Estrada C, Filliger LZ, et al. 2011. Expanding the diversity of mycobacteriophages: insights into genome architecture and evolution. *PLoS One* 6:e16329. <https://doi.org/10.1371/journal.pone.0016329>.
57. Weigel C, Seitz H. 2006. Bacteriophage replication modules. *FEMS Microbiol Rev* 30:321–381. <https://doi.org/10.1111/j.1574-6976.2006.00015.x>.

58. Seco EM, Zinder JC, Manhart CM, Lo Piano A, McHenry CS, Ayora S. 2013. Bacteriophage SPP1 DNA replication strategies promote viral and disable host replication in vitro. *Nucleic Acids Res* 41:1711–1721. <https://doi.org/10.1093/nar/gks1290>.
59. Breitbart M, Middelboe M, Rohwer F. 2008. Marine viruses: community dynamics, diversity and impact on microbial processes, p 443–479. In Kirchman DL (ed), *Microbial ecology of the oceans*, 2nd ed. John Wiley & Sons, Hoboken, NJ.
60. Dwivedi B, Xue B, Lundin D, Edwards RA, Breitbart M. 2013. A bioinformatic analysis of ribonucleotide reductase genes in phage genomes and metagenomes. *BMC Evol Biol* 13:33. <https://doi.org/10.1186/1471-2148-13-33>.
61. Ignacio-Espinoza JC, Sullivan MB. 2012. Phylogenomics of T4 cyanophages: lateral gene transfer in the 'core' and origins of host genes. *Environ Microbiol* 14:2113–2126. <https://doi.org/10.1111/j.1462-2920.2012.02704.x>.
62. Roux S, Brum JR, Dutilh BE, Sunagawa S, Duhaime MB, Loy A, Poulos BT, Solonenko N, Lara E, Poulain J, Pesant S, Kandels-Lewis S, Dimier C, Picheral M, Searson S, Cruaud C, Alberti A, Duarte CM, Gasol JM, Vaqué D, Bork P, Acinas SG, Wincker P, Sullivan MB. 2016. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* 537:689–693. <https://doi.org/10.1038/nature19366>.
63. Mann NH, Clokie MR, Millard A, Cook A, Wilson WH, Wheatley PJ, Letarov A, Krusch HM. 2005. The genome of S-PM2, a "photosynthetic" T4-type bacteriophage that infects marine *Synechococcus* strains. *J Bacteriol* 187:3188–3200. <https://doi.org/10.1128/JB.187.9.3188-3200.2005>.
64. Nordlund P, Reichard P. 2006. Ribonucleotide reductases. *Annu Rev Biochem* 75:681–706. <https://doi.org/10.1146/annurev.biochem.75.103004.142443>.
65. Myllykallio H, Lipowski G, Leduc D, Filee J, Forterre P, Liebl U. 2002. An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* 297:105–107. <https://doi.org/10.1126/science.1072113>.
66. Zhao Y, Wang K, Jiao N, Chen F. 2009. Genome sequences of two novel phages infecting marine roseobacters. *Environ Microbiol* 11:2055–2064. <https://doi.org/10.1111/j.1462-2920.2009.01927.x>.
67. Gross JW, Rajavel M, Grubmeyer C. 1998. Kinetic mechanism of nicotinic acid phosphoribosyltransferase: implications for energy coupling. *Biochem* 37:4189–4199. <https://doi.org/10.1021/bi972014w>.
68. Miller GA, Jr, Rosenzweig S, Switzer RL. 1975. Oxygen-18 studies of the mechanism of pyrophosphoryl group transfer catalyzed by phosphoribosylpyrophosphate synthetase. *Arch Biochem Biophys* 171:732–736. [https://doi.org/10.1016/0003-9861\(75\)90086-7](https://doi.org/10.1016/0003-9861(75)90086-7).
69. Hove-Jensen B. 1988. Mutation in the phosphoribosylpyrophosphate synthetase gene (*prs*) that results in simultaneous requirements for purine and pyrimidine nucleosides, nicotinamide nucleotide, histidine, and tryptophan in *Escherichia coli*. *J Bacteriol* 170:1148–1152. <https://doi.org/10.1128/jb.170.3.1148-1152.1988>.
70. Jensen KF. 1983. Metabolism of 5-phosphoribosyl 1-pyrophosphate (PRPP) in *Escherichia coli* and *Salmonella typhimurium*, p 1–25. In Munch-Petersen A (ed), *Metabolism of nucleotides, nucleosides and nucleobases in microorganisms*. Academic Press, London, United Kingdom.
71. Horton H, Moran L, Ochs R, Rawn D, Scrimgeour K. 1996. Principles of biochemistry, 2nd ed. Prentice Hall, Upper Saddle River, NJ.
72. Lee JY, Li ZQ, Miller ES. 2017. Vibrio phage KVP40 encodes a functional NAD⁺ salvage pathway. *J Bacteriol* 199:e00855-16. <https://doi.org/10.1128/JB.00855-16>.
73. Seed KD, Bodi KL, Kropinski AM, Ackermann HW, Calderwood SB, Qadri F, Camilli A. 2011. Evidence of a dominant lineage of *Vibrio cholerae*-specific lytic bacteriophages shed by cholera patients over a 10-year period in Dhaka, Bangladesh. *mBio* 2:e00334-10. <https://doi.org/10.1128/mBio.00334-10>.
74. Wang DB, Sun MQ, Shao HB, Li Y, Meng X, Liu ZY, Wang M. 2015. Characterization and genome sequencing of a novel bacteriophage PH101 infecting *Pseudoalteromonas marina* BH101 from the Yellow Sea of China. *Curr Microbiol* 71:594–600. <https://doi.org/10.1007/s00284-015-0896-5>.
75. Gross M, Marianovsky I, Glaser G. 2006. MazG - a regulator of programmed cell death in *Escherichia coli*. *Mol Microbiol* 59:590–601. <https://doi.org/10.1111/j.1365-2958.2005.04956.x>.
76. Clokie MR, Mann NH. 2006. Marine cyanophages and light. *Environ Microbiol* 8:2074–2082. <https://doi.org/10.1111/j.1462-2920.2006.01171.x>.
77. Sullivan MB, Huang KH, Ignacio-Espinoza JC, Berlin AM, Kelly L, Weigle PR, DeFrancesco AS, Kern SE, Thompson LR, Young S, Yandava C, Fu R, Krastins B, Chase M, Sarracino D, Osburne MS, Henn MR, Chisholm SW. 2010. Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environ Microbiol* 12:3035–3056. <https://doi.org/10.1111/j.1462-2920.2010.02280.x>.
78. Kim SK, Makino K, Amemura M, Shinagawa H, Nakata A. 1993. Molecular analysis of the *phoH* gene, belonging to the phosphate regulon in *Escherichia coli*. *J Bacteriol* 175:1316–1324. <https://doi.org/10.1128/jb.175.5.1316-1324.1993>.
79. Kazakov AE, Vassieva O, Gelfand MS, Osterman A, Overbeek R. 2003. Bioinformatics classification and functional analysis of PhoH homologs. *In Silico Biol* 3:3–15.
80. Lindell D, Jaffe JD, Coleman ML, Futschik ME, Axmann IM, Rector T, Kettler G, Sullivan MB, Steen R, Hess WR, Church GM, Chisholm SW. 2007. Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* 449:83–86. <https://doi.org/10.1038/nature06130>.
81. Goldsmith DB, Crosti G, Dwivedi B, McDaniel LD, Varsani A, Suttle CA, Weinbauer MG, Sandaa RA, Breitbart M. 2011. Development of *phoH* as a novel signature gene for assessing marine phage diversity. *Appl Environ Microbiol* 77:7730–7739. <https://doi.org/10.1128/AEM.05531-11>.
82. Sakowski EG, Munsell EV, Hyatt M, Kress W, Williamson SJ, Nasko DJ, Polson SW, Wommack KE. 2014. Ribonucleotide reductases reveal novel viral diversity and predict biological and ecological features of unknown marine viruses. *Proc Natl Acad Sci U S A* 111:15786–15791. <https://doi.org/10.1073/pnas.1401322111>.
83. Loderer C, Holmfeldt K, Lundin D. 2019. Non-host class II ribonucleotide reductase in Thermus viruses: sequence adaptation and host interaction. *PeerJ* 7:e6700. <https://doi.org/10.7717/peerj.6700>.
84. Thingstad TF. 2000. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr* 45:1320–1328. <https://doi.org/10.4319/lo.2000.45.6.1320>.
85. Sambrook J, Russel DW. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
86. Carlson K, Miller E. 1994. Experiments in T4 genetics, p 421–483. In Karam JD, Drake JW, Kreuzer KN, Mosig G, Hall DH, Eiserling FA, Black LW, Spicer EK, Dutter E, Carlson K, Miller ES (ed), *Molecular biology of bacteriophage T4*. ASM Press, Washington, DC.
87. R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
88. RStudio Team. 2015. RStudio: integrated development for R. RStudio, Inc., Boston, MA.
89. Wickham H. 2011. The split-apply-combine strategy for data analysis. *J Stat Softw* 40:1–29.
90. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer, New York, NY.
91. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
92. Andrews S. 2010. FastQC: a quality control tool for high throughput sequencing data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
93. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
94. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Res* 19:1117–1123. <https://doi.org/10.1101/gr.089532.108>.
95. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid annotations using subsystems technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
96. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia FF, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.

97. Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* 33: W451–W454. <https://doi.org/10.1093/nar/gki487>.
98. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.
99. Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol* 7:e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
100. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285. <https://doi.org/10.1093/nar/gkv1344>.
101. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* 44: W54–W57. <https://doi.org/10.1093/nar/gkw413>.
102. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>.
103. Ågren J, Sundström A, Håfström T, Segerman B. 2012. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. *PLoS One* 7:e39107. <https://doi.org/10.1371/journal.pone.0039107>.
104. Wickham H. 2007. Reshaping data with the reshape package. *J Stat Softw* 21:1–20.
105. Wickham H. 2017. Tidyverse: easily install and load “tidyverse” packages. <https://tidyverse.tidyverse.org/>.
106. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. *Bioinformatics* 27:1009–1010. <https://doi.org/10.1093/bioinformatics/btr039>.
107. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. <https://doi.org/10.1186/1471-2105-14-60>.
108. Lefort V, Desper R, Gascuel O. 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol* 32:2798–2800. <https://doi.org/10.1093/molbev/msv150>.
109. Göker M, García-Blázquez G, Voglmayr H, Tellería MT, Martín MP. 2009. Molecular taxonomy of phytopathogenic fungi: a case study in *Peronospora*. *PLoS One* 4:e6319. <https://doi.org/10.1371/journal.pone.0006319>.
110. Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, Fiebig A, Rohde C, Rohde M, Fartmann B, Goodwin LA, Chertkov O, Reddy TBK, Pati A, Ivanova NN, Markowitz V, Kyrpides NC, Woynke T, Göker M, Klenk HP. 2014. Complete genome sequence of DSM 30083^T, the type strain (U5/41^T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci* 9:2. <https://doi.org/10.1186/1944-3277-9-2>.
111. John SG, Mendez CB, Deng L, Poulos B, Kauffman AK, Kern S, Brum J, Polz MF, Boyle EA, Sullivan MB. 2011. A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ Microbiol Rep* 3:195–202. <https://doi.org/10.1111/j.1758-2229.2010.00208.x>.
112. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
113. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
114. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
115. Boström KH, Simu K, Hagström Å, Riemann L. 2004. Optimization of DNA extraction for quantitative marine bacterioplankton community analysis. *Limnol Oceanogr Methods* 2:365–373. <https://doi.org/10.4319/lom.2004.2.365>.
116. Bunse C, Bertos-Fortis M, Sassenhagen I, Sildever S, Sjöqvist C, Godhe A, Gross S, Kremp A, Lips I, Lundholm N, Rengefors K, Seftom J, Pinhassi J, Legrand C. 2016. Spatio-temporal interdependence of bacteria and phytoplankton during a Baltic Sea spring bloom. *Front Microbiol* 7:517. <https://doi.org/10.3389/fmicb.2016.00517>.
117. Riemann L, Steward GF, Azam F. 2000. Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. *Appl Environ Microbiol* 66:578–587. <https://doi.org/10.1128/aem.66.2.578-587.2000>.
118. Herlemann DP, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* 5:1571–1579. <https://doi.org/10.1038/ismej.2011.41>.
119. Hugerth LW, Wefer HA, Lundin S, Jakobsson HE, Lindberg M, Rodin S, Engstrand L, Andersson AF. 2014. DegePrime, a program for degenerate primer design for broad-taxonomic-range PCR in microbial ecology studies. *Appl Environ Microbiol* 80:5116–5123. <https://doi.org/10.1128/AEM.01403-14>.
120. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583. <https://doi.org/10.1038/nmeth.3869>.