

# Novel PhoH-encoding vibriophages with lytic activity against environmental *Vibrio* strains

Alice Nyambura Maina (✉ [nyamburagichuhi@gmail.com](mailto:nyamburagichuhi@gmail.com))

Technical University of Kenya <https://orcid.org/0000-0002-9425-8955>

Francis B. Mwaura

University of Nairobi College of Health Sciences

Miriam Jumba

University of Nairobi

Kristopher Kieft

University of Wisconsin-Madison

Hanzada T. Nour El-Din

Cairo University Faculty of Pharmacy

Ramy K. Aziz

Cairo University Faculty of Pharmacy

---

## Research Article

**Keywords:** *Vibrio cholerae*, vibriophages, myoviruses, phoH, biocontrol, genomics, genome annotation, phage therapy

**Posted Date:** July 26th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-185269/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Archives of Microbiology on August 11th, 2021. See the published version at <https://doi.org/10.1007/s00203-021-02511-3>.

# Abstract

Cholera is a devastating diarrheal disease that accounts for more than 10% of children's lives worldwide, but its treatment is hampered by a rise in antibiotic resistance. One promising alternative to antibiotic therapy is the use of bacteriophages to treat antibiotic-resistant cholera infections, and control *Vibrio cholera* in clinical cases and in the environment, respectively. Here, we report four novel, closely related environmental myoviruses, VP4, VP6, VP18, and VP24, which we isolated from two environmental toxigenic *Vibrio cholerae* strains from river Kuja and Usenge beach in Kenya. High-throughput sequencing followed by bioinformatics analysis indicated that the genomes of the four bacteriophages have closely related sequences, with sizes of 148,180 bp, 148,181 bp, 148,179 bp, and 148,179 bp, and a G + C content of 36.4%. The four genomes carry the *phoH* gene, which is overrepresented in marine cyanophages. The isolated phages displayed a lytic activity against 15 environmental, as well as one clinical, *Vibrio cholerae* strains. Thus, these novel lytic vibriophages represent potential biocontrol candidates for water decontamination against pathogenic *Vibrio cholerae* and ought to be considered for future studies of phage therapy.

## Introduction

Acute infectious diarrheal diseases remain among the most frequent causes of childhood deaths, accounting for 10–12% of the death toll in children under five years of age, and around 1.4–1.9 million fatalities worldwide (Levy et al. 2016). Cholera is a devastating watery diarrheal disease that causes severe dehydration and death if untreated. It is mainly caused by O1 and O139 toxigenic *Vibrio cholerae* serotypes. The disease is spread through the faecal-oral route and hence strongly associated with poverty, poor hygiene, clean water shortage, and lack of adequate sanitation facilities (Deen et al. 2020). The aquatic environment is the main reservoir for *V. cholerae*, specifically brackish, estuarine, and coastal waters (Almagro-Moreno and Taylor 2013).

Treatment of cholera is challenging. A combination of antibiotics with rehydration therapy relieves the symptoms of cholera and shortens the disease duration. Unfortunately, environmental drug-resistant *V. cholerae* strains have been recently reported, hampering the treatment option for cholera and urgently calling for adjunct or alternative approaches (Loo et al. 2020), such as bacteriophage therapy.

Bacteriophages, termed phages for short, are viruses that infect bacteria and exist in equilibrium with their bacterial hosts. The relationship between *V. cholerae* and bacteriophages dates to the 1920s, when Felix d'Herelle described the spread of phages in the environment after the onset of an outbreak (Jassim and Limoges 2014; Silva-Valenzuela and Camilli 2019). Back then, several studies proved the effectiveness of phage treatment against versatile diseases, including cholera, staphylococcal infections, typhoid fever, and bacterial dysentery (Sulakvelidze et al. 2001; El-Shibiny and El-Sahhar 2017)

Despite this early success, enthusiasm towards phage treatment and research immensely declined with the discovery of antibiotics (Abedon et al. 2011). Revisiting phage therapy ought to be taken into

consideration, given that human and environmental *V. cholerae* populations are naturally controlled by serogroup-specific bacteriophages (Faruque et al. 2005a). Since *V. cholerae* is a natural inhabitant of aquatic environments (Almagro-Moreno and Taylor 2013; Lutz et al. 2013), these environments are favorable for exploring candidate therapeutic bacteriophages. Indeed, several tailed bacteriophages, especially from family *Myoviridae*, were detected during environmental surveys in regions where outbreaks were reported, e.g., Peru (Talledo et al. 2003), Kolkata (Sen and Ghosh 2005) and Kenya (Maina et al. 2014). Of note, tailed bacteriophages are the most dominant viruses in the aquatic environment (Madhusudana Rao and Lalitha 2015; Letchumanan et al. 2016), which makes them a good initial candidate for screening, but will require elaborate efforts for their isolation.

Here we report the isolation, characterization and sequencing of four novel phages with contractile tails (typical myophage morphology). These lytic phages demonstrated a biocontrol potential against tested environmental and clinical *V. cholerae* strains. Hence, they should be considered as possible candidates for the highly needed phage therapy targeting the increasingly resistant *V. cholerae* and for water decontamination as well.

## Materials And Methods

### **Bacterial hosts used for phage isolation and propagation:**

Two environmental strains of *Vibrio cholerae*, previously isolated from different water sources, were used for bacteriophage isolation and propagation. The first of the two strains, Vc\_ke, isolated from river Kuja in Migori County and identified as toxigenic El Tor strain, was used for phages VP4, VP6, and VP18. The second strain, Vc\_Use, isolated from Usenge beach in Siaya, was used for VP24.

### **Phage isolation, propagation and DNA isolation:**

Like their bacterial hosts, phages VP4, VP6, and VP18 were isolated from river Kuja in Migori County, Kenya, while phage VP24 was isolated from Usenge beach in Siaya, Kenya.

The double agar layer method was used for isolation and purification of the phages, as described by van Twest and Kropinski (2009) but with a slight modification: The soft agar layer (0.6% agar) was made of Trypticase Soy agar instead of Luria-Bertani agar.

Genomic DNA was extracted from phage lysates by the standard phenol-chloroform protocol as described in Sambrook et al. (2001) and modified by Shah (2014).

### **Host range profiling:**

The host range for the isolated phages was profiled as previously described (Stenholm et al. 2008; Kutter 2009), with a minor modification (Trypticase Soy with 0.6% agar was used instead of Luria-Bertani agar). Each phage was tested for its ability to form plaques on lawns of each of the 15 used environmental *V. cholerae* strains isolated from the different Kenyan waters (Table 1), in addition to one clinical *V. cholerae*

isolate. The host range was estimated from the number of lysed strains out of the 16 tested strains above.

Table 1  
Sources of *Vibrio cholerae* Strains Used in Host range profiles

Code	Source	Region	Geographical site
Vc_ke*	Environmental	Migori	River Kuja
Vc_Kuj	Environmental	Migori	River Kuja
Vc_Kuj	Environmental	Migori	River Kuja
Vc_Kuj	Environmental	Migori	River Kuja
Vc_Kuj	Environmental	Migori	River Kuja
Vc_Koleche*	Environmental	Bondo	Koleche Pond
Vc-Kot	Environmental	Bondo	Kotonde Pond
Vc_Owi	Environmental	Bondo	Owira Pond
Vc_Use	Environmental	Rarieda	Usenge Beach
Vc_Osi	Environmental	Rarieda	Osieko Beach
Vc_Lau	Environmental	Rarieda	Lauda Beach
Vc_Nsongoni*	Environmental	Coast	River Nsongoni
Vc_Kiz	Environmental	Coast	River Kizulini
Vc_Kam	Environmental	Central	River Kamiti
Vc_Riv	Environmental	Central	Riverside River
VCM	Clinical	UoN	Medical Microbiology Laboratory
* Those bacteria were previously isolated and described (Maina et al. 2021).			

Among these isolates, three had been earlier confirmed by 16S rRNA gene sequencing. The 16S rRNA sequences were deposited in NCBI and given accession numbers MN467399.1 (Vc\_ke isolated from river Kuja, Migori), MN907464.1 (Vc\_Koleche, from Koleche pond, Siaya), and MN907465.1 (Vc\_Nsongoni, isolated from river Nsongoni, Mombasa).

Each isolate of the target strains was cultured on thiosulfate-citrate-bile salts-sucrose (TCBS) agar and incubated for 12h at 37°C. Thereafter, 10 ml of trypticase soy broth (TSB) was inoculated with a single

colony of each bacterial strain, and then incubated for 12h at 37°C. A 500 ul aliquot of the overnight culture was mixed with 4ml soft agar, and poured onto the surface of the TSA plates to make the host lawns. Subsequently, the plates were allowed to set, and 10µl of the phage lysate was spotted per lawn. After the spots were allowed to set, plates were incubated for 12h at 37°C. The plates were examined for zones of clearing/lysis wherever a phage had been spotted (Yu *et al.*, 2013). If a clear zone was observed, the isolate was declared sensitive to the phage. A control culture was set with sterile sodium chloride-magnesium sulfate (SM) buffer, instead of phage lysates, to verify the growth and purity of the culture.

The phages were also tested for their infectivity against three other Gram-negative bacteria isolated from the same environmental sources at the same geographical areas as the *Vibrio* isolates (Maina et al. 2021). These bacteria included two strains of *E. coli* (Ec\_Kuja, 16S rRNA sequence accession number MN907473.1, isolated from river Kuja, and EC\_ke, 16S rRNA sequence accession number MN467398.1, isolated from Nairobi River). The other two bacterial isolates, *Proteus mirabilis* and *Providencia sneebia*, were isolated from river Kuja, and their 16S rRNA sequences were determined and assigned accession numbers MN467400.1 and MN467401.1, respectively).

### **Morphological characterization by transmission electron microscopy (TEM):**

Purified bacteriophage samples (titer between  $10^9$  to  $10^{10}$  PFUs/ml) were prepared for TEM at the Wellcome Sanger Institute, operated by Genome Research Limited. Negative staining was performed with 5% of Uranyl acetate at a ratio of 5:1 for each phage sample and magnification of X 60000. Purified and washed phage samples were adhered to freshly glow-discharged carbon/Formvar grids, briefly stained with 5% uranyl acetate and then blotted and air dried. Grids were then viewed on a 120kV FEI Spirit Biotwin microscope (Thermo Fisher Scientific, Hillsboro, OR) and imaged by a Tietz 4.16 charge-coupled device (CCD) camera. Measurements were taken directly with the TVIPS EMTools (Germany).

### **Library construction and sequencing:**

A Covaris sonicator (Covaris, MA, USA) was used to randomly shred the genomic DNA, after its concentration was measured with a Qubit 3.0 fluorometer (Life Technologies, CA, USA) and adjusted. For library preparation, DNA fragments were end-repaired, A-tailed, ligated with adapters, purified, and amplified by the polymerase chain reaction (PCR).

After library construction, the library DNA was accurately quantified by Qubit 3.0 and diluted. Finally, its quality was assessed in an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Additional quality check was performed via quantitative PCR (qPCR) determination of the library DNA concentration. Once good quality of the libraries was confirmed, they were pooled and sequenced in an Illumina flow cell, according to the manufacturer's requirement of concentration and volume. The sequencing was performed on the high-throughput platform of Illumina NovaSeq 6000 (Illumina, CA, USA).

### **Annotation of phage genomes:**

The phages were annotated by both VIBRANT v1.2.1 (Kieft et al. 2020) and Rapid Annotation using Subsystems Technology (RAST) server (Aziz et al. 2008), after adjusting RAST settings to the optimized

phage pipeline (Aziz et al. 2018; McNair et al. 2018). Genes encoding tRNAs were detected by tRNAscan-SE (Lowe and Eddy 1997). Annotations of VP4 were visualized by Geneious Prime 2019.0.3 (Biomatters, Ltd., New Zealand) and color-coded according to the respective annotation category.

### **Phylogenetic analysis:**

The annotations from VIBRANT were used to identify ribonucleotide reductase (RnR) subunits alpha and beta, as well as the large terminase (LT) subunit, which were subsequently used for phylogenetic tree construction. RnR alpha, RnR beta and LT protein sequences from VP4 were used to query the NCBI nr database (accessed August 2020, taxid10239, e-value < 1e-5) for reference virus sequences. The top 100 results from each search were taken and individually dereplicated by CD-HIT v4.6, -aS 0.8 -c 0.9 (Huang et al. 2010).

Respective protein sequences for VP4, VP6, VP18, and VP24 were added to the reference sequences and aligned by MAFFT (v7.388, Katoh and Standley 2013), with default settings. Phylogenetic trees were constructed with RAxML (v8.2.4, Stamatakis 2014), with the command: raxmlHPC-PTHREADS -N 100 -f a -m PROTCATLG, and the resulting trees, representing 100 bootstrap iterations and rooted by outgroup, were visualized by FigTree v1.4.4 (Rambaut 2007).

### **Genome alignment:**

The genomes of two closely related phages (MG592609.1 and NC\_048709.1), chosen from the phylogenetic trees, were aligned to VP4 genome by tBLASTX (Altschul et al. 1990) and visualized by use of EasyFig (v2.2.2) software (Sullivan et al. 2011) with minimum BLAST alignment length = 100, maximum e-value = 1e-5, and minimum identity = 50%. Tools of the Pathosystems Resource Integration Center (PATRIC) portal (Davis et al. 2020) were used for comparing predicted proteins of the four phages and for generating a CIRCOS circular map with the GC% and GC skew of the phage genome (Krzywinski et al. 2009).

### **Nucleotide sequence accession number:**

The full genome sequence of phage VP4 (systematically named vB\_vcM\_Kuja) is currently available in the NCBI Nucleotide database under the accession number MN718199.1 The corresponding RefSeq record is NC\_048827.1

## **Results And Discussion**

### **Morphology of vibriophages VP4, VP6, VP18, and VP24:**

Morphological characterization, based on the TEM images, of the four studied phages VP4, VP6, VP18, and VP24, showed an icosahedral capsid head, ranging in size from 78nm to 85nm and a contractile tail, ranging in length from 95nm to 103nm (Fig. 1). The virion morphology suggested that the phages belong to the family *Myoviridae*, order *Caudovirales*, as previously described (Maina et al. 2014).

### **Genomic features and primary classification of VP4, VP6, VP18, and VP24:**

The whole genomes of the four phages were sequenced. The linear double-stranded DNA size ranged between 148,179 bp and 148,181 bp, and its molecular weight was 91.54 MDa, which suggests relatively large genomes compared to those of other myoviruses. The GC content was 36.4%. The annotation suggested 186 protein-coding sequences (CDSs) for phages VP4 and VP18, and 185 CDSs for VP6 and VP24.

Among the well-studied bacteriophages that infect *V. cholerae* O1, the myophage ICP1, a double-stranded DNA virus (previously known as JSF1), has a genome size of 125,956 bp, G + C content of 37% and 230 coding sequences. A similar phage of the same series, JSF7, propagated on strain *V. cholerae* O1, has a double-stranded DNA genome with G + C content of 48.42% has a genome size of 46kb (Naser et al. 2017). The genome size of JSF7 is much smaller than the four phages in this study.

The best-characterized phage of the *Myoviridae* family is phage T4. T4 has a genome size of 168,903 bp, an elongated head (110x80 nm), a contractile tail that ends with a complex base plate with six long fibers radiating from it. The T4 genome (NCBI accession number NC\_000866) has a G + C content of 35% and eight tRNA genes. The propagating strain is *Escherichia coli*. More than 200 similar phages have been described that share common virion morphology and related features (Comeau and Krisch 2008). The genomes of the four closely related phages isolated in this study can possibly be placed with the class of T4-like phages.

About 90% of known T4-like phages grow on *E. coli* or other enterobacteria, but 10% grow on phylogenetically more distant bacteria, such as *Aeromonas*, *Vibrio*, *Cyanobacteria*, among others, and they significantly vary in virion morphology (Comeau and Krisch 2008). Phage JS98 (NCBI accession number NC\_010105), propagated on *E. coli*, with a genome size of 170,523, three tRNAs, G + C content of 39%, can be considered to be closely related to the four phage genomes in this study.

A well-studied vibriophage, KVP40 has a giant genome of 244,834 bp (NCBI accession number NC\_005083), host strain *Vibrio parahemolyticus*, G + C content of 42.6%, and 30 tRNAs. The phage, isolated from polluted sea water in Japan, belongs to the *Myoviridae* family and is classified as a T4-like phage with a broad host range (Miller et al. 2003).

The four phages are closely related to two non-*Vibrio* phages: S-PM2 (NCBI genome accession NC\_006820) and S-RSM4 (NCBI genome accession number NC\_013085), which infect *Synechococcus*. Their G + C content is 37% and 41%, respectively, and their genomes sizes are 196,280 bp and 194,454 bp (Clokic et al. 2010).

Two other phages are closely related in percent G + C content. These are P-SSM2 and P-SSM4, known to infect *Prochlorococcus*. The genomes of these two phages (NCBI accession numbers NC\_006883 and NC\_006884, respectively) have G + C content of 35.5% and 36.7%, respectively, and genome sizes 252,401 bp and 178,249 bp, respectively. P-SSM2 only encodes for one tRNA while P-SSM4 lacks tRNA genes. These two bacteria, *Prochlorococcus* and *Synechococcus*, are globally ubiquitous marine cyanobacteria, and their phages are among the most abundant in the world's oceans (Clokic et al. 2010; Aziz et al. 2015).

Because the four phages isolated in this study have relatively large genomes, they are potentially interesting subjects of further studies.

### **Genome annotation:**

As mentioned above, 186 CDSs were defined for phages VP4 and VP18, and 185 CDSs for phages VP6 and VP24. Out of these, 103 genes (55%) were annotated as 'hypothetical' or 'unknown' proteins. High frequency of proteins with unassigned functions is typical for phage genomes from previously unsampled geographical sites, and calls for further studies to identify the potential functions of those proteins.

On the other hand, the annotation process identified 82 genes (45% of all CDSs) associated with two main categories: (1) functional subsystems/modules, including nucleotide replication, repair, recombination, and metabolism, and (2) phage structural and hallmark genes (Fig. 2).

The four phages had near identical set of predicted protein, with marginal differences in 2–3 proteins per genome (Fig. 3 and Supplementary data table S1). Other than structural and functional phage domains, detailed for phage VP4 (Fig. 2), no unusual genes were detected. Specifically, both VIBRANT (Kieft et al. 2020) and the PATRIC (Davis et al. 2020) database indicated no known resistance or virulence genes, and no evidence of integrases was found.

### **Detection and significance of *phoH*:**

A notable exception to the above categories was the *phoH* gene, the expression of which was linked to phosphate starvation conditions. The *phoH* is a host-derived auxiliary metabolic gene (AMG), sometimes known as a moron (Hendrix et al. 2000; Hendrix et al. 2003), and is commonly carried by some phages. It belongs to the phosphate regulon that regulates phosphate uptake and metabolism under conditions of low phosphate and phosphate limitation. The *phoH* gene homologs were detected in phages with various morphological types, e.g., siphophages, myophages, and podophages, and with a wide bacterial host range (including autotrophic and heterotrophic bacteria). They were even detected in viruses of autotrophic eukaryotes (Goldsmith et al. 2011).

The *phoH* gene is not restricted to a certain morphological type of phage, which suggests that it could be a powerful biomarker gene for studying phage diversity. Goldsmith et al. (2011) found out that nearly 40% of marine phages contained *phoH*, compared to only 4% of nonmarine phages (Goldsmith et al. 2011).

In a study by Wang et al. (2016), more than 400 phage-harbored *phoH* sequences were obtained from several paddy floodwaters in northeast China. Precisely, four specific groups and seven subgroups of this gene family were detected in phages from paddy waters (Wang et al. 2016). The study demonstrated that *phoH* was present in phage genomes of terrestrial environments and that this gene was useful for studying phage ecology in paddy ecosystems. These findings support the evidence that this biomarker

gene can be used to investigate the diversity of phages in both marine and terrestrial environments (Li et al. 2019).

Phosphorus is a major element, necessary for nucleotide biosynthesis and DNA replication, but is extremely scarce in oligotrophic waters and is consequently thought to be one of the limiting factors for cyanobacterial growth (Martiny et al. 2006; Tetu et al. 2009; Kelly et al. 2013). Thus, it is not surprising that some phosphorus-acquisition genes, such as the phosphate-inducible genes, *pstS* and *phoH*, and the alkaline phosphatase gene *phoA*, which are regulated by the PhoR/PhoB two-component regulatory system to sense phosphorus availability, were found in the genomes of cyanophages infecting cyanobacteria (Martiny et al. 2006; Sullivan et al. 2010; Zeng and Chisholm 2012). These genes could be upregulated in response to phosphate starvation in host cells, and their products could play an important role in regulating phosphorus absorption and transportation of host cells under low-phosphorus content or phosphorus-deprived conditions (Gao et al. 2016). It was also proposed that cyanophages maintain *phoH* to allow their host increased phosphate uptake during infection; however, the mechanism of how it occurs is not well known (Clokic et al. 2010), and *phoH* expression in phosphate-limited conditions appears to vary between hosts (Lindell et al. 2007; Tetu et al. 2009).

Based on the above observations, *phoH* has been proposed as a novel signature gene to assess the genetic diversity of viruses in multiple families of double-stranded DNA tailed phages (Goldsmith et al. 2011). *phoH* was commonly discovered in cyanophages, such as marine cyanophages P-SSM2, P-SSM4, and Syn9 (Sullivan et al. 2005; Weigele et al. 2007), and freshwater cyanophages Ma-LMM01 (Yoshida et al. 2008) and MaMV-DC (Ou et al. 2015). It was also found in non-marine phages such as coliphage T5 (Wang et al. 2005) and *Staphylococcus* phages K, G1, and Twort (O'Flaherty et al. 2004; Kwan et al. 2005). The frequent detection of *phoH* genes in phage genomes suggests that their products play a role in the phosphate metabolism of the phage-infected cell (Sullivan et al. 2005). Based on bioinformatic analyses, *phoH* genes were suggested to be part of a multi-gene family with divergent functions from phospholipid metabolism and RNA modification to fatty acid beta-oxidation (Kazakov et al. 2003).

*phoH* was also reported in vibriophages, represented by phage KVP40 (Miller et al. 2003) the well-studied T4-like phage isolated from polluted coastal sea water in Japan. The propagating host bacterium of phage KVP40 was *Vibrio parahaemolyticus*. Here, *phoH* was found in all the four phage genome sequences reported here. *phoH* has already been reported in phages isolated from versatile geographic locations. Although most phages harboring the *phoH* gene originated from marine environments, some belonged to other habitats, such as soil, sewage and stool (Adriaenssens and Cowan 2014). Taken together, the presence of genes involved in phosphorous acquisition demonstrate how phages might have developed adaptation to life in oligotrophic environment (Baudoux et al. 2012).

#### **Phylogenetic context of VP4, VP6, VP18, and VP24:**

Phylogenetic analysis of RnR alpha, RnR beta and LT protein sequences of the four phages, compared to those from reference phages, displayed separate branching of the four phages (Fig. 4). Although the four phages were phylogenetically distinct from reference phages, they were most closely related to other

*Vibrio*-infecting phages. Based on the trees for RnR alpha and RnR beta, the four phages are most closely related to *Vibrio* phage 1.244.A\_10N.261.54.C3 (NCBI accession number MG592609.1, 159 kb length), from seawater, which has a partial genome sequence and is estimated to belong to the family *Ackermannviridae* (Kauffman et al. 2018).

As for the LT-based phylogeny, LT from the four phages are most closely related to *Vibrio* phage YC (RefSeq NC\_048709.1, 147 kb length), which belongs to *Ackermannviridae* and infects *Vibrio coralliilyticus* (Cohen et al. 2013). Other phages that are related based on the LT tree include *Vibrio* phage VP-1 (NCBI MH363700.1, 150 kb length) (Mateus et al. 2014) and *Vibrio* phage VAP7 (RefSeq NC\_048765.1, 144 kb length), both of which belong to the family *Ackermannviridae* as well (Gao et al. 2020), but were found to infect *Vibrio parahaemolyticus* and *Vibrio alginolyticus*, respectively.

Although the four phages described here seem to belong to family *Ackermannviridae* and share a similar genome length (~ 150kb) to those other phages, their separate phylogenetic branching and different host range suggest that they are evolutionarily distinct from other known phages.

Alignment of phage VP4 with two closely related *Vibrio*-infecting phages shows that VP4 shares a few genes with known *Vibrio* phages (Fig. 5). The genes with the highest similarity are those predicted to encode terminases and major capsid proteins. These similarities, along with the phylogenetic trees, confirm the phylogenetic relatedness, yet distinction, of the four novel phages relative to known *Vibrio* phages.

#### **Potential applications and therapeutic value:**

Cholera epidemics are known to be self-limiting in nature, since the epidemics subside after reaching a peak, even without any active human intervention (Hoque et al. 2016). Among other factors, lytic phages that kill *V. cholerae* have been shown to play a significant role in modulating the course of epidemics, presumably through their inherent bactericidal activity. Studies suggested that seasonal cholera epidemics may end as a result of phage predation of the causative epidemic *V. cholerae* strains (Faruque et al. 2005a; Faruque et al. 2005b; Nelson et al. 2009). This natural predatory role of phages make them appealing tools for the biocontrol of epidemics before they claim human lives; hence, the value of lytic phages is not only therapeutic, but also preventive, in some sense.

One of the pivotal phage features, which affects their therapeutic value, is their host spectrum. The four phages isolated here were infective against the 15 different tested environmental *V. cholerae* strains, in addition to a clinical strain. However, the phages were not infective against bacteria representing three other species: *E. coli*, *Proteus mirabilis* and *Providencia sneebia* so they remain of limited spectrum. A cocktail, composed of three different phages isolated from surface waters in Bangladesh and designated as JSF7, JSF4, and JSF3, could significantly influence the distribution and concentration of the active planktonic form and biofilm-associated form of toxigenic *V. cholerae* in water (Naser et al. 2017). Therefore, the four phages in this study are potential candidates to be added to cocktails for water decontamination and control of *V. cholerae* in environmental waters, used by poor communities for domestic purposes in Kenya.

## Conclusion and outlook.

Here, we report the initial characterization and whole genome sequencing of four novel vibriophages that could be primarily classified in the subfamily *Ackermannviridae*. Our analysis showed they possess relatively large genomes, with over half of their genes encoding for unknown proteins that might be involved in manipulating the *Vibrio* host during infection.

Phage therapy is a potential life saver during cholera outbreaks in underprivileged countries, owing to the relative ease and speed of phage preparation with basic inexpensive laboratory equipment (Bhandare et al. 2019). To further circumvent antibiotic resistance, more complex and stable phage formulation methods are being explored (e.g., lyophilization, spray drying, emulsification, and microencapsulation (Malik et al. 2017). Future directions should target further characterization of the presented vibriophages and aim to test their control over pathogenic *V. cholerae* for water decontamination and phage therapy.

## Post scriptum:

It is important to note that, while this manuscript was being written and was going through several rounds of revision, the Bacterial Viruses Subcommittee of International Committee on Taxonomy of Viruses (ICTV, URL: <https://talk.ictvonline.org/>) has set out to make major changes in nomenclature and taxonomy, with a re-assignment of order *Caudovirales* into class *Caudoviricetes*, and an eventual abolishment of families *Myoviridae*, *Podoviridae* and *Siphoviridae* (Turner et al. 2021). However, we opted to keep the current description of phages with contractile tails as “myophages” or “myoviruses” to allow comparison with literature, until a genome-based taxonomic system is fully established for the members of those three families.

## Declarations

**Acknowledgements:** We wish to acknowledge David Goulding for help with TEM imaging. We are grateful for the generous financial assistance provided by Professor Elizabeth Kutter (The Evergreen State College, Olympia, WA, USA) and Tobi Nagel (Phages for Global Health) to phage researchers from Africa. The clinical *Vibrio* isolate was a kindly provided by the Department of Microbiology, University of Nairobi, Kenya.

**Funding:** RKA is funded by the Egyptian Academy for Scientific Research and Technology (ASRT) JESOR program (Project #3046). The funder has no interference with the scientific content of the article.

**Conflict of interest:** None of the authors has a financial or personal conflict of interest to declare.

**Supplementary Material:** A data table (Table S1, in Excel .xlsx format) with all defined and annotated genes

## References

- Abedon ST, Kuhl SJ, Blasdel BG, Kutter EM (2011) Phage treatment of human infections. *Bacteriophage* 1:66-85 doi: 10.4161/bact.1.2.15845
- Adriaenssens EM, Cowan DA (2014) Using signature genes as tools to assess environmental viral ecology and diversity. *Appl Environ Microbiol* 80:4470-4480 doi: 10.1128/aem.00878-14
- Almagro-Moreno S, Taylor RK (2013) Cholera: environmental reservoirs and impact on disease transmission. *Microbiol Spectr* 1 doi: 10.1128/microbiolspec.OH-0003-2012
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403-410 doi: 10.1016/s0022-2836(05)80360-2
- Aziz RK, Ackermann HW, Petty NK, Kropinski AM (2018) Essential steps in characterizing bacteriophages: biology, taxonomy, and genome analysis. *Methods Mol Biol* 1681:197-215 doi: 10.1007/978-1-4939-7343-9\_15
- Aziz RK, Bartels D, Best AA et al. (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75 doi: 10.1186/1471-2164-9-75
- Aziz RK, Dwivedi B, Akhter S, Breitbart M, Edwards RA (2015) Multidimensional metrics for estimating phage abundance, distribution, gene density, and sequence coverage in metagenomes. *Front Microbiol* 6:381 doi: 10.3389/fmicb.2015.00381
- Baudoux AC, Hendrix RW, Lander GC et al. (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 doi: 10.1111/j.1462-2920.2011.02685.x
- Bhandare S, Colom J, Baig A et al. (2019) Reviving phage therapy for the treatment of cholera. *J Infect Dis* 219:786-794 doi: 10.1093/infdis/jiy563
- Clokier MR, Millard AD, Mann NH (2010) T4 genes in the marine ecosystem: studies of the T4-like cyanophages and their role in marine ecology. *Viol J* 7:291 doi: 10.1186/1743-422x-7-291
- Cohen Y, Joseph Pollock F, Rosenberg E, Bourne DG (2013) Phage therapy treatment of the coral pathogen *Vibrio coralliilyticus*. *Microbiology Open* 2:64-74 doi: 10.1002/mbo3.52
- Comeau AM, Krisch HM (2008) The capsid of the T4 phage superfamily: the evolution, diversity, and structure of some of the most prevalent proteins in the biosphere. *Mol Biol Evol* 25:1321-1332 doi: 10.1093/molbev/msn080
- Davis JJ, Wattam AR, Aziz RK et al. (2020) The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res.* 48:D606-d612 doi: 10.1093/nar/gkz943

- Deen J, Mengel MA, Clemens JD (2020) Epidemiology of cholera. *Vaccine* 38 Suppl 1:A31-a40 doi: 10.1016/j.vaccine.2019.07.078
- El-Shibiny A, El-Sahhar S (2017) Bacteriophages: the possible solution to treat infections caused by pathogenic bacteria. *Can J Microbiol* 63:865-879 doi: 10.1139/cjm-2017-0030
- Faruque SM, Islam MJ, Ahmad QS et al. (2005a) Self-limiting nature of seasonal cholera epidemics: role of host-mediated amplification of phage. *Proc Natl Acad Sci USA* 102:6119-6124 doi: 10.1073/pnas.0502069102
- Faruque SM, Naser IB, Islam MJ et al. (2005b) Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. *Proc Natl Acad Sci USA* 102:1702-1707 doi: 10.1073/pnas.0408992102
- Gao EB, Huang Y, Ning D (2016) Metabolic genes within cyanophage genomes: implications for diversity and evolution. *Genes (Basel)* 7 doi: 10.3390/genes7100080
- Gao M, Qin Y, Fan H et al. (2020) Characteristics and complete genome sequence of the virulent *Vibrio alginolyticus* phage VAP7, isolated in Hainan, China. *Arch Virol* 165:947-953 doi: 10.1007/s00705-020-04535-4
- Goldsmith DB, Crosti G, Dwivedi B et al. (2011) Development of *phoH* as a novel signature gene for assessing marine phage diversity. *Appl Environ Microbiol* 77:7730-7739 doi: 10.1128/aem.05531-11
- Hendrix RW, Hatfull GF, Smith MC (2003) Bacteriophages with tails: chasing their origins and evolution. *Res Microbiol* 154:253-257 doi: 10.1016/s0923-2508(03)00068-8
- Hendrix RW, Lawrence JG, Hatfull GF, Casjens S (2000) The origins and ongoing evolution of viruses. *Trends Microbiol* 8:504-508 doi: 10.1016/s0966-842x(00)01863-1
- Hoque MM, Naser IB, Bari SM, Zhu J, Mekalanos JJ, Faruque SM (2016) Quorum regulated resistance of *Vibrio cholerae* against environmental bacteriophages. *Sci Rep* 6:37956 doi: 10.1038/srep37956
- Huang Y, Niu B, Gao Y, Fu L, Li W (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* 26:680-682 doi: 10.1093/bioinformatics/btq003
- Jassim SA, Limoges RG (2014) Natural solution to antibiotic resistance: bacteriophages 'The Living Drugs'. *World J Microbiol Biotechnol* 30:2153-2170 doi: 10.1007/s11274-014-1655-7
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772-780 doi: 10.1093/molbev/mst010
- Kauffman KM, Hussain FA, Yang J et al. (2018) A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria. *Nature* 554:118-122 doi: 10.1038/nature25474

- 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
- Kelly L, Ding H, Huang KH, Osburne MS, Chisholm SW (2013) Genetic diversity in cultured and wild marine cyanomyoviruses reveals phosphorus stress as a strong selective agent. *Isme j* 7:1827-1841 doi: 10.1038/ismej.2013.58
- Kieft K, Zhou Z, Anantharaman K (2020) VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* 8:90 doi: 10.1186/s40168-020-00867-0
- Krzywinski M, Schein J, Birol I et al. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.* 19:1639-1645 doi: 10.1101/gr.092759.109
- Kutter E (2009) Phage host range and efficiency of plating. *Methods Mol Biol* 501:141-149 doi: 10.1007/978-1-60327-164-6\_14
- Kwan T, Liu J, DuBow M, Gros P, Pelletier J (2005) The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. *Proc Natl Acad Sci USA* 102:5174-5179 doi: 10.1073/pnas.0501140102
- Letchumanan V, Chan KG, Pusparajah P et al. (2016) Insights into bacteriophage application in controlling *Vibrio* species. *Front. Microbiol.* 7:1114 doi: 10.3389/fmicb.2016.01114
- Levy K, Woster AP, Goldstein RS, Carlton EJ (2016) Untangling the impacts of climate change on waterborne diseases: a systematic review of relationships between diarrheal diseases and temperature, rainfall, flooding, and drought. *Environ Sci Technol* 50:4905-4922 doi: 10.1021/acs.est.5b06186
- Li X, Sun Y, Liu J, Yao Q, Wang G (2019) Survey of the bacteriophage *phoH* gene in wetland sediments in northeast China. *Sci Rep* 9:911 doi: 10.1038/s41598-018-37508-4
- Lindell D, Jaffe JD, Coleman ML et al. (2007) Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* 449:83-86 doi: 10.1038/nature06130
- Loo K-Y, Letchumanan V, Law JW-F et al. (2020) Incidence of antibiotic resistance in *Vibrio* spp. *Rev Aquacult* 12:2590-2608 doi: <https://doi.org/10.1111/raq.12460>
- Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955-964 doi: 10.1093/nar/25.5.955
- Lutz C, Erken M, Noorian P, Sun S, McDougald D (2013) Environmental reservoirs and mechanisms of persistence of *Vibrio cholerae*. *Front. Microbiol.* 4:375 doi: 10.3389/fmicb.2013.00375

- Madhusudana Rao B, Lalitha KV (2015) Bacteriophages for aquaculture: Are they beneficial or inimical. *Aquaculture* 437:146-154 doi: <https://doi.org/10.1016/j.aquaculture.2014.11.039>
- Maina A, Mwaura F, Jumba M, Kieft K (2021) Isolation, characterisation of four Gram-negative environmental bacteria and bacteriophage plaques formation on their lawns. *Int. J.I Pharma and bio sci.* 11:121-129
- Maina AN, Mwaura FB, Oyugi J, Goulding D, Toribio AL, Kariuki S (2014) Characterization of *Vibrio cholerae* bacteriophages isolated from the environmental waters of the Lake Victoria region of Kenya. *Curr Microbiol* 68:64-70 doi: 10.1007/s00284-013-0447-x
- Malik DJ, Sokolov IJ, Vinner GK et al. (2017) Formulation, stabilisation and encapsulation of bacteriophage for phage therapy. *Adv Colloid Interface Sci* 249:100-133 doi: 10.1016/j.cis.2017.05.014
- Martiny AC, Coleman ML, Chisholm SW (2006) Phosphate acquisition genes in *Prochlorococcus* ecotypes: evidence for genome-wide adaptation. *Proc Natl Acad Sci USA* 103:12552-12557 doi: 10.1073/pnas.0601301103
- Mateus L, Costa L, Silva Y, Pereira C, Cunha A, Almeida A (2014) Efficiency of phage cocktails in the inactivation of *Vibrio* in aquaculture. *Aquaculture* 424:167-173 doi: <https://doi.org/10.1016/j.aquaculture.2014.01.001>
- McNair K, Aziz RK, Pusch GD, Overbeek R, Dutilh BE, Edwards R (2018) Phage genome annotation using the RAST pipeline. *Methods Mol Biol* 1681:231-238 doi: 10.1007/978-1-4939-7343-9\_17
- Miller ES, Heidelberg JF, Eisen JA et al. (2003) Complete genome sequence of the broad-host-range vibriophage KVP40: comparative genomics of a T4-related bacteriophage. *J Bacteriol* 185:5220-5233 doi: 10.1128/jb.185.17.5220-5233.2003
- Naser IB, Hoque MM, Abdullah A, Bari SMN, Ghosh AN, Faruque SM (2017) Environmental bacteriophages active on biofilms and planktonic forms of toxigenic *Vibrio cholerae*: potential relevance in cholera epidemiology. *PLoS One* 12:e0180838 doi: 10.1371/journal.pone.0180838
- Nelson EJ, Harris JB, Morris JG, Jr., Calderwood SB, Camilli A (2009) Cholera transmission: the host, pathogen and bacteriophage dynamic. *Nat Rev Microbiol* 7:693-702 doi: 10.1038/nrmicro2204
- O'Flaherty S, Coffey A, Edwards R, Meaney W, Fitzgerald GF, Ross RP (2004) Genome of staphylococcal phage K: a new lineage of Myoviridae infecting Gram-positive bacteria with a low G+C content. *J Bacteriol* 186:2862-2871 doi: 10.1128/jb.186.9.2862-2871.2004
- Ou T, Gao XC, Li SH, Zhang QY (2015) Genome analysis and gene *nblA* identification of *Microcystis aeruginosa* myovirus (MaMV-DC) reveal the evidence for horizontal gene transfer events between cyanomyovirus and host. *J Gen Virol* 96:3681-3697 doi: 10.1099/jgv.0.000290

- Rambaut A (2007) FigTree, a graphical viewer of phylogenetic trees. Edinburgh, UK: Institute of Evolutionary Biology University of Edinburgh. Available from: <http://tree.bio.ed.ac.uk/software/figtree/>
- Sambrook J, Russell DW, Russell DW (2001) Molecular Cloning: a Laboratory Manual (3-volume set), Third edn. Cold Spring Harbor Laboratory Press, New York
- Sen A, Ghosh AN (2005) New *Vibrio cholerae* O1 biotype ElTor bacteriophages. Virol. J. 2:28 doi: 10.1186/1743-422x-2-28
- Shah M (2014) An application of sequencing batch reactors in the identification of microbial community structure from an activated sludge. J Appl Env Microbiol 2:176-184 doi: 10.12691/jaem-2-4-11
- Silva-Valenzuela CA, Camilli A (2019) Niche adaptation limits bacteriophage predation of *Vibrio cholerae* in a nutrient-poor aquatic environment. Proc Natl Acad Sci USA 116:1627-1632 doi: 10.1073/pnas.1810138116
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313 doi: 10.1093/bioinformatics/btu033
- Stenholm AR, Dalsgaard I, Middelboe M (2008) Isolation and characterization of bacteriophages infecting the fish pathogen *Flavobacterium psychrophilum*. Appl Environ Microbiol 74:4070-4078 doi: 10.1128/aem.00428-08
- Sulakvelidze A, Alavidze Z, Morris JG, Jr. (2001) Bacteriophage therapy. Antimicrob Agents Chemother 45:649-659 doi: 10.1128/aac.45.3.649-659.2001
- Sullivan MB, Coleman ML, Weigle P, Rohwer F, Chisholm SW (2005) Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. PLoS Biol 3:e144 doi: 10.1371/journal.pbio.0030144
- Sullivan MB, Huang KH, Ignacio-Espinoza JC et al. (2010) Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. Environ Microbiol 12:3035-3056 doi: 10.1111/j.1462-2920.2010.02280.x
- Sullivan MJ, Petty NK, Beatson SA (2011) Easyfig: a genome comparison visualizer. Bioinformatics 27:1009-1010 doi: 10.1093/bioinformatics/btr039
- Talledo M, Rivera IN, Lipp EK et al. (2003) Characterization of a *Vibrio cholerae* phage isolated from the coastal water of Peru. Environ. Microbiol. 5:350-354 doi: 10.1046/j.1462-2920.2003.00411.x
- Tetu SG, Brahmsha B, Johnson DA et al. (2009) Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp. WH8102. Isme j 3:835-849 doi: 10.1038/ismej.2009.31

Turner D, Kropinski AM, Adriaenssens EM (2021) A roadmap for genome-based phage taxonomy. *Viruses* 13 doi: 10.3390/v13030506

van Twest R, Kropinski AM (2009) Bacteriophage enrichment from water and soil. *Methods Mol Biol* 501:15-21 doi: 10.1007/978-1-60327-164-6\_2

Wang J, Jiang Y, Vincent M et al. (2005) Complete genome sequence of bacteriophage T5. *Virology* 332:45-65 doi: 10.1016/j.virol.2004.10.049

Wang X, Liu J, Yu Z, Jin J, Liu X, Wang G (2016) Novel groups and unique distribution of phage *phoH* genes in paddy waters in northeast China. *Sci Rep* 6:38428 doi: 10.1038/srep38428

Weigele PR, Pope WH, Pedulla ML et al. (2007) Genomic and structural analysis of Syn9, a cyanophage infecting marine *Prochlorococcus* and *Synechococcus*. *Environ Microbiol* 9:1675-1695 doi: 10.1111/j.1462-2920.2007.01285.x

Yoshida T, Nagasaki K, Takashima Y et al. (2008) Ma-LMM01 infecting toxic *Microcystis aeruginosa* illuminates diverse cyanophage genome strategies. *J Bacteriol* 190:1762-1772 doi: 10.1128/jb.01534-07

Zeng Q, Chisholm SW (2012) Marine viruses exploit their host's two-component regulatory system in response to resource limitation. *Curr Biol* 22:124-128 doi: 10.1016/j.cub.2011.11.055

## Tables

**Table 1:** Sources of *Vibrio cholerae* Strains Used in Host range profiles

Code	Source	Region	Geographical site
Vc_ke*	Environmental	Migori	River Kuja
Vc_Kuj	Environmental	Migori	River Kuja
Vc_Kuj	Environmental	Migori	River Kuja
Vc_Kuj	Environmental	Migori	River Kuja
Vc_Kuj	Environmental	Migori	River Kuja
Vc_Koleche*	Environmental	Bondo	Koleche Pond
Vc-Kot	Environmental	Bondo	Kotonde Pond
Vc_Owi	Environmental	Bondo	Owira Pond
Vc_Use	Environmental	Rarieda	Usenge Beach
Vc_Osi	Environmental	Rarieda	Osieko Beach
Vc_Lau	Environmental	Rarieda	Lauda Beach
Vc_Nsongoni*	Environmental	Coast	River Nsongoni
Vc_Kiz	Environmental	Coast	River Kizulini
Vc_Kam	Environmental	Central	River Kamiti
Vc_Riv	Environmental	Central	Riverside River
VCM	Clinical	UoN	Medical Microbiology Laboratory

\* Those bacteria were previously isolated and described (Maina et al. 2021).

## Figures

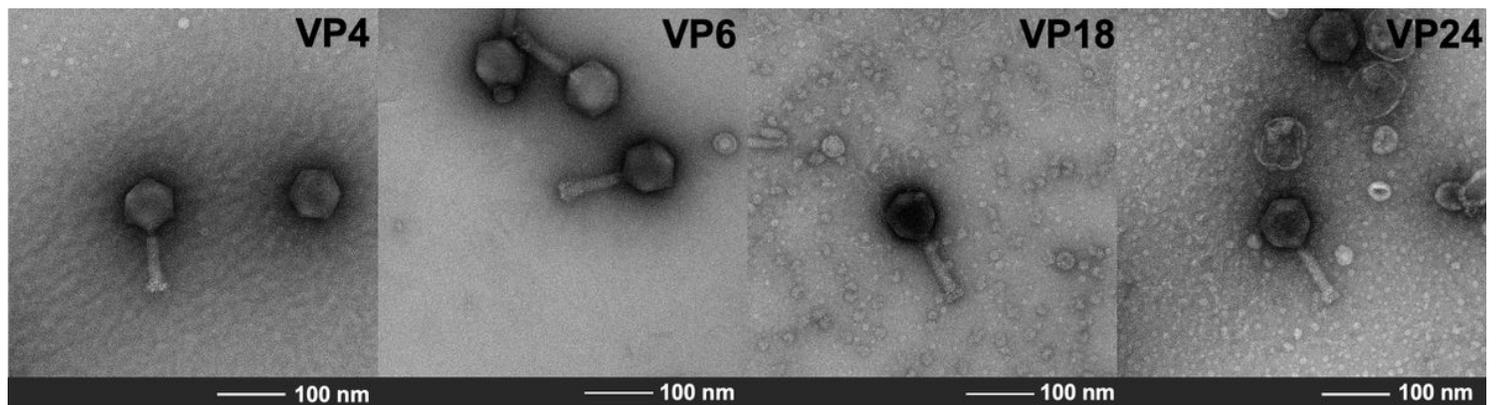
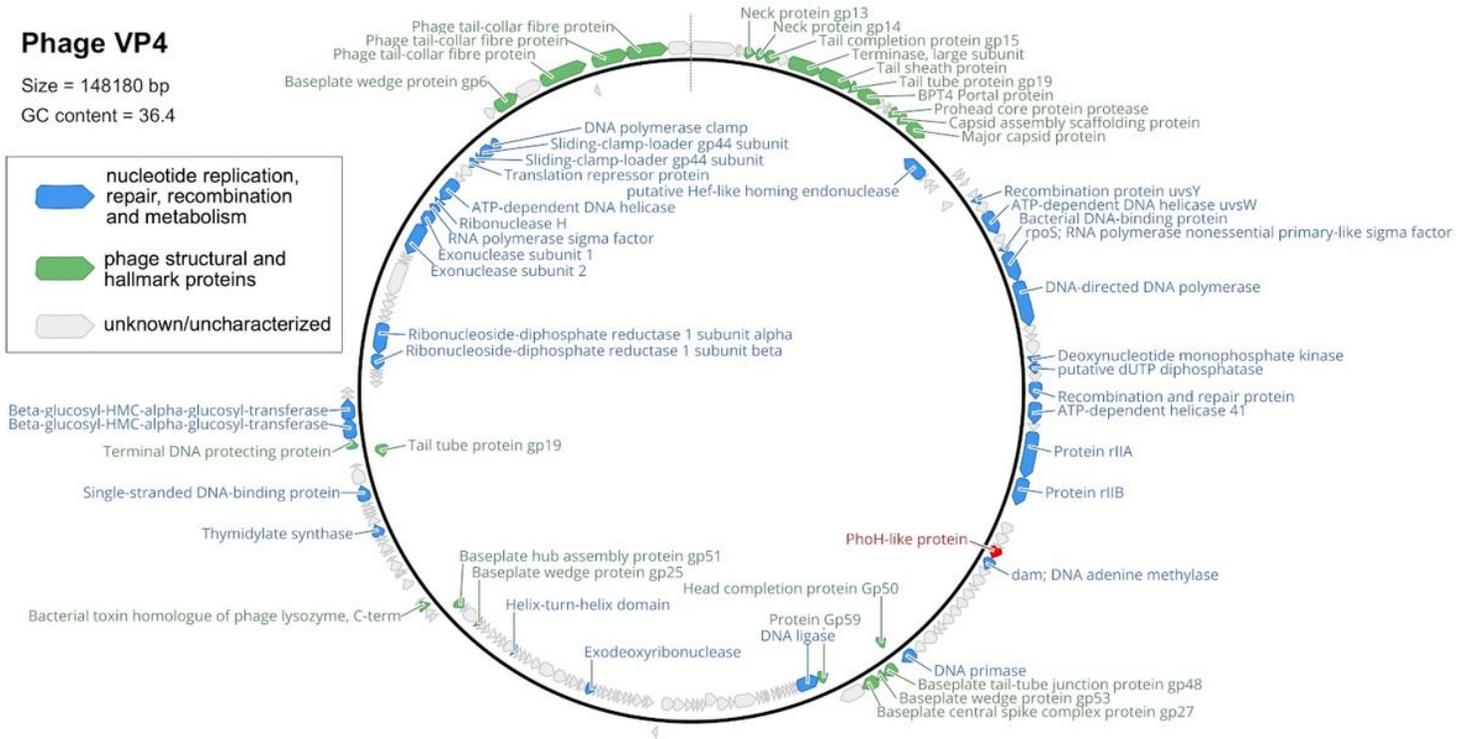


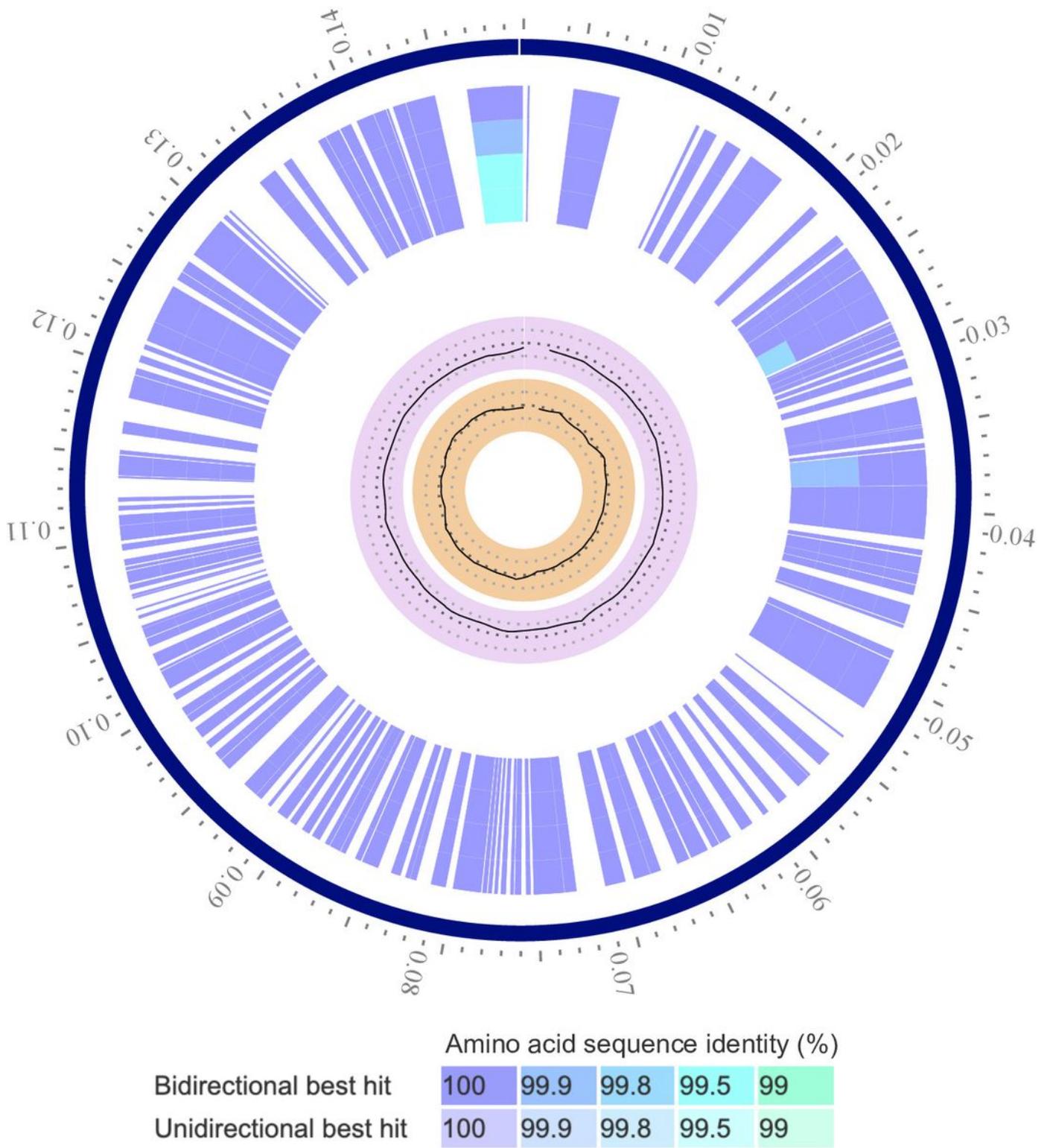
Figure 1

Transmission electron micrographs of VP4, VP6, VP18, and VP24 phages after negative staining. The four vibriophages have icosahedral heads and contractile tails. Bar = 100nm.



**Figure 2**

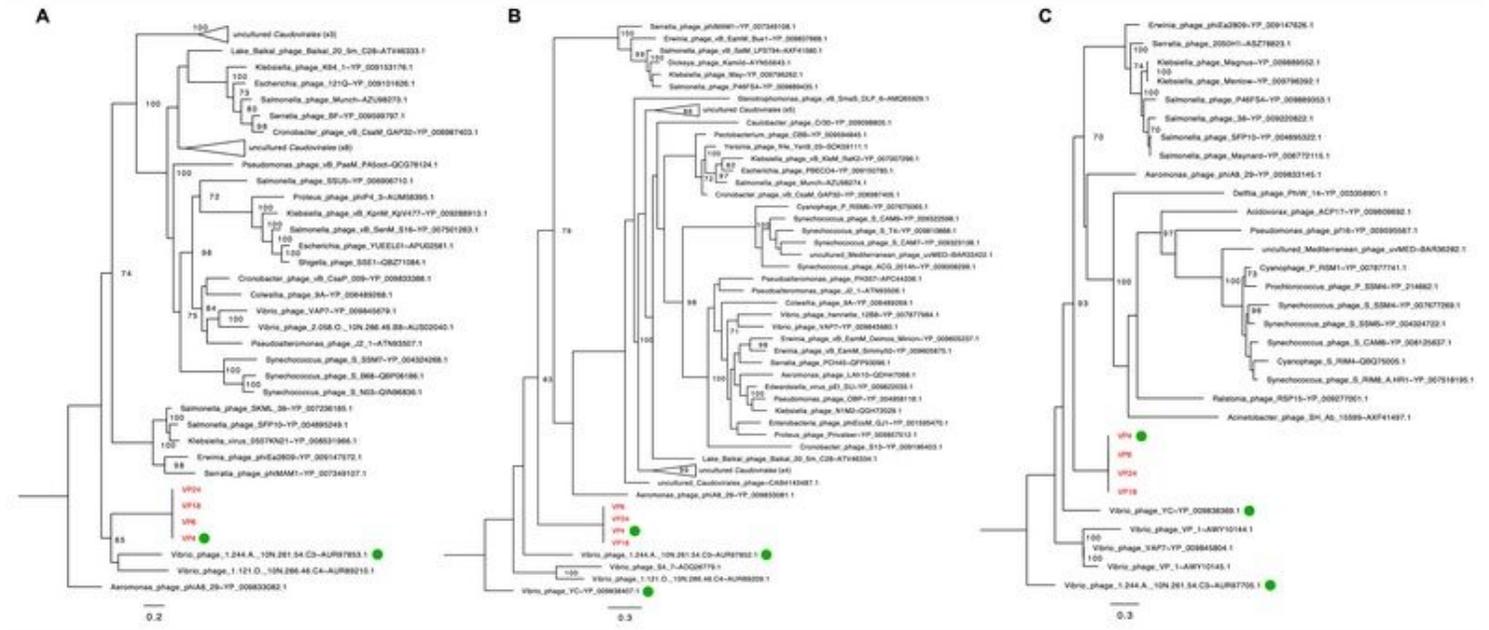
Circular diagram of annotated phage VP4 genome. Annotations from VIBRANT v1.2.1 are color-coded as follows: blue for proteins involved in nucleotide replication, repair, recombination, and metabolism; green for phage structural and hallmark proteins; and gray for proteins of unknown or uncategorized functions. *phoH* is shown in red. The dotted line indicates the site of circularization of the linear sequence. Geneious Prime 2019.0.3 was used to generate the diagram.



**Figure 3**

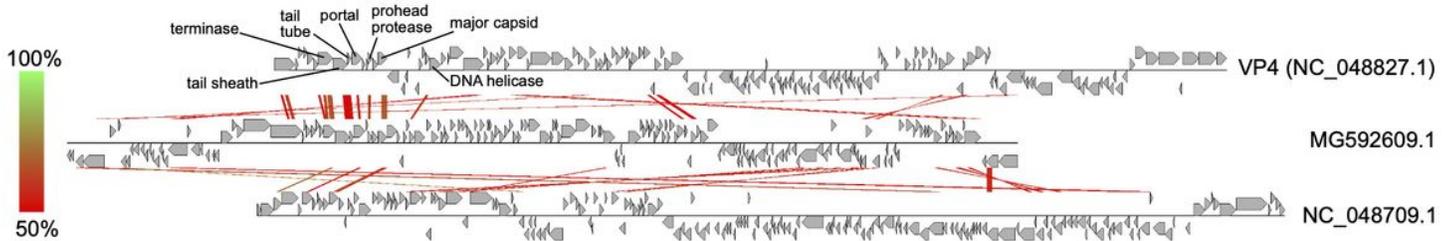
Proteome-based comparative diagram of protein-coding genes of phages VP4, VP6, VP18, and VP24. All protein-coding genes of the four phages VP4, VP6, VP18, and VP24 were compared (at the amino-acid level), and the comparison is represented in a circular diagram, with color-coded percent amino acid identity (colors and shades are indicated in the bottom). Pairwise bidirectional similarity suggests orthology. Tracks, from outside to inside: nucleotide position in millions of bases, genome scaffold

(continuous blue line), phage VP4 open reading frames (ORFs), phage VP18 ORFs, phage VP6 ORFs, phage VP24 ORFs, GC% (with window size of 2000 nucleotides), and finally GC skew (with window size of 2000 nucleotides, and a median dotted line equivalent to zero). The figure is a merger of two figures created in the PATRIC platform (Davis et al. 2020).



**Figure 4**

Marker gene-based phylogenetic trees of phages VP4, VP6, VP18, and VP24 with reference viruses. The phylogenetic trees were constructed based on compared protein sequences of (A) ribonucleotide reductase subunit alpha (RnR alpha), (B) ribonucleotide reductase subunit beta (RnR beta), and (C) large terminase (LT) subunit. Clades representing phages VP4, VP6, VP18, and VP24 are labeled in red. Phages used for genome alignments are highlighted with green dots. Bootstrap values greater than 70 are shown, and trees are rooted by outgroups. Trees were graphically generated by FigTree (v1.4.4).



**Figure 5**

Genome alignments of phage VP4 against reference phages. Genomes maps are connected by lines according to high (green) or low (red) tBLASTX identity. Open-reading frames are depicted by gray arrows, and genes with high identity between VP4 and the reference phages are annotated. Accession numbers

for reference phages are MG592609.1 (Vibrio phage 1.244.A\_10N.261.54.C3) and NC\_048709.1 (Vibrio phage YC).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.xlsx](#)