

SCIENTIFIC REPORTS



OPEN

A novel roseobacter phage possesses features of podoviruses, siphoviruses, prophages and gene transfer agents

Received: 29 February 2016

Accepted: 30 June 2016

Published: 27 July 2016

Yuanchao Zhan¹, Sijun Huang², Sonja Voget³, Meinhard Simon⁴ & Feng Chen¹

Bacteria in the Roseobacter lineage have been studied extensively due to their significant biogeochemical roles in the marine ecosystem. However, our knowledge on bacteriophage which infects the Roseobacter clade is still very limited. Here, we report a new bacteriophage, phage DSS3Φ8, which infects marine roseobacter *Ruegeria pomeroyi* DSS-3. DSS3Φ8 is a lytic siphovirus. Genomic analysis showed that DSS3Φ8 is most closely related to a group of siphoviruses, CbK-like phages, which infect freshwater bacterium *Caulobacter crescentus*. DSS3Φ8 contains a smaller capsid and has a reduced genome size (146 kb) compared to the CbK-like phages (205–279 kb). DSS3Φ8 contains the DNA polymerase gene which is closely related to T7-like podoviruses. DSS3Φ8 also contains the integrase and repressor genes, indicating its potential to involve in lysogenic cycle. In addition, four GTA (gene transfer agent) genes were identified in the DSS3Φ8 genome. Genomic analysis suggests that DSS3Φ8 is a highly mosaic phage that inherits the genetic features from siphoviruses, podoviruses, prophages and GTAs. This is the first report of CbK-like phages infecting marine bacteria. We believe phage isolation is still a powerful tool that can lead to discovery of new phages and help interpret the overwhelming unknown sequences in the viral metagenomics.

Bacteria in the Roseobacter clade are abundant and widely distributed in the marine environment^{1,2}. The broad distribution of the Roseobacter clade is in nature largely attributed to their diverse metabolic capabilities^{3–6}. Currently, more than 140 genomes of different roseobacter species have been sequenced, and analyses of representative genome sequences have generated new insight into ecological adaptation and evolution of the Roseobacter clade⁷.

Compared to the knowledge we gained on the Roseobacter clade, much less is known about bacteriophages which infect marine roseobacters (roseophage). Marine viruses, as the most abundant biological entities in the biosphere, play important roles in shaping host population structures, mediating genetic exchange between hosts, and modulating trophic transfer in marine food webs^{8–11}. Viral metagenomic studies have greatly expanded our knowledge on viral diversity in nature^{12–14}, but the interpretation of data is often hampered due to the limitation of our knowledge on isolated viruses. Recent isolations of phages infecting two abundant marine bacteria are the great examples for linking the unknown phages in the metagenomic database with known phages^{15,16}. In addition, for the isolated phages, a wealth of biological information such as morphology, lytic activity, host specificity, genomics, and evolution can be retrieved.

Currently, only a handful of phages which infect roseobacters have been isolated and characterized. Roseophage SIO1 was the first described phage which infects *Roseobacter* SIO67¹⁷, and similar SIO1-like phages were found at the same site after 10 years¹⁸. Two N4-like phages infecting two different roseobacter strains (*Silicibacter pomeroyi* DSS-3 and *Sulfitobacter* sp. EE-36) were discovered, and they represent the first report of N4-like phages infecting marine bacteria¹⁹. Recently, several N4-like phages were isolated from different strains of the Roseobacter clade^{20–22}, suggesting that N4-like phages may be a common type of phage infecting

¹Institution of Marine and Environmental Technology, University of Maryland Center for Environmental Science, USA. ²CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China. ³Institute for Microbiology and Genetics, University of Göttingen, Germany. ⁴Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Germany. Correspondence and requests for materials should be addressed to F.C. (email: chenf@umces.edu)

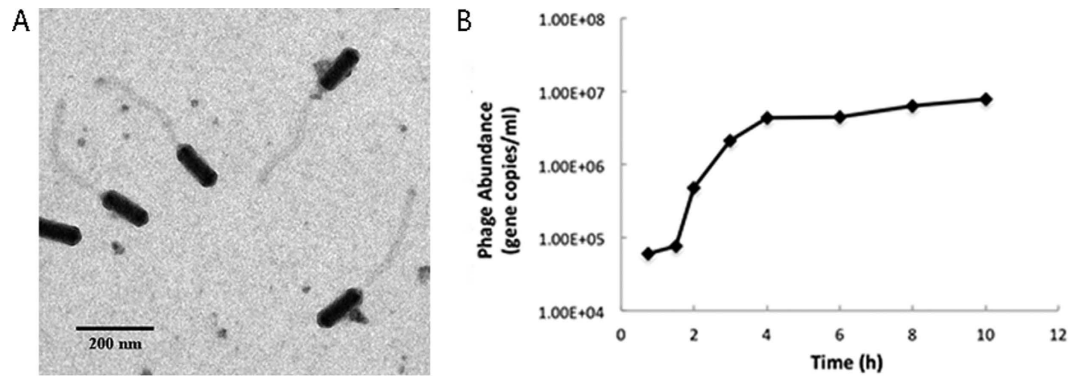


Figure 1. (A) Transmission electron microscopy image of roseophage DSS3Φ8. Scale bar = 200 nm. (B) One-step growth curve of roseophage DSS3Φ8.

roseobacters. One siphovirus, phage RDJLΦ1, which infects *Roseobacter denitrificans* OCh114 was isolated from the South China Sea²³. Most recently, two temperate podoviruses infecting *Sulfitobacter* sp. strain 2047 were also described²⁴. These two temperate phages are almost identical at nucleotide level and were isolated from a mesocosm study in Raunefjorden.

The *Roseobacter* clade also contains inducible prophages. Prophages have been induced from two *roseobacter* strains, *Roseobacter* sp. TM1040²⁵ and *Roseovarius nubinihibens* ISM²⁶. Some bacterial strains (i.e. TM1040) may contain multiple prophages in their genomes²⁷. Another phage-like structure, the gene transfer agent (GTA), has been found in nearly all the *roseobacter* genomes^{7,28,29}. GTA is a phage-like entity that transfers random pieces of genome from donor cells to recipient cells³⁰. Interestingly, roseophage RDJLΦ1 contains four GTA-like genes in its genome³¹. The *Roseobacter* clade are a group of highly diverse and adaptable bacteria. It is believed that phages (both lytic and temperate), prophages and GTA are the driving force for genomic diversification of the *Roseobacter* clade.

In this study, we report a new phage, phage DSS3Φ8, which infects a marine *roseobacter*, *R. pomeroyi* DSS-3. Phage DSS3Φ8 is a highly mosaic phage because it contains genetic elements of podoviruses, siphoviruses, prophages, and GTAs. To our knowledge, no phage similar to the very unusual DSS3Φ8 has been found in marine bacteria prior to this study.

Results and Discussion

Morphology and biological features. Roseophage DSS3Φ8 was isolated from Baltimore Inner Harbor, Baltimore, USA in January 2012. DSS3Φ8 has a prolate cylindrical head and a long, flexible, non-contractile tail. The capsid size of DSS3Φ8 is ca. 130 nm long and 40 nm wide, and the tail length is ca. 300 nm (Fig. 1A). Based on its morphology, DSS3Φ8 is a member of B3 morphotype of *Siphoviridae*, which comprises approximately 1% of all known phages³². Genome sequence showed that DSS3Φ8 belongs to the phiCbK genus under *Siphoviridae*, which contains five B3 morphotype bacteriophages infecting the freshwater bacterium *Caulobacter crescentus*³³. Phage CbK was first isolated in 1970, and has been an important genetic and cytological tool for studying cell cycle regulation, because it specifically infects “swarmer” host cells³⁴. Roseophage DSS3Φ8 is much smaller than the five CbK-like phages in terms of capsid size and tail length.

The cross infectivity test showed that DSS3Φ8 did not infect the other four marine *roseobacter* strains, even though *Silicibacter* sp. TM1040 is in the same genus with DSS3. The latent period of DSS3Φ8 was about 2 hours, and it reached the growth plateau in 6 hours. The burst size of DSS3Φ8 is ca. 120 (Fig. 1B).

General genome features. DSS3Φ8 is a circular, double-stranded DNA virus. The genome size of DSS3Φ8 is 146,135 bp. In general, the genome of DSS3Φ8 is homologous to that of the CbK-like phages (see the later section), but the genome size of DSS3Φ8 is much smaller compared to the known CbK-like phages (205–279 kb)³³. The G + C content of the phage genome is 56%, lower than that of the CbK-like phages (62–66%). The G + C content is 64 and 67% for *R. pomeroyi* DSS3 and *Caulobacter crescentus*, respectively.

A total of 229 predicted open reading frames (ORFs) was identified in the DSS3Φ8 genome, of which 106 ORFs had recognizable homologues in GenBank (Fig. 2). Among the recognizable proteins, only 50 have described function, consistent with the idea that marine siphoviruses encode proteins that are under-represented in the database³⁵.

DSS3Φ8 contains 24 tRNA genes coding 14 different amino acids. Most of tRNA genes found in phage DSS3Φ8 overlapped with those found in the host *R. pomeroyi* DSS3. The parallel possession of tRNA genes between DSS3Φ8 and host DSS3 is considered as a result of co-evolution of phage and host^{36,37}. On the other hand, two tRNAs, tRNA^{Lys/CTT} and tRNA^{Gln/CTG}, were only found in DSS3Φ8. The anticodon CTT and CTG account for more than 90% codon usage for Lys and Gln respectively for phages, however, these anticodon are the rarest codons in the host³⁸. This finding supported previous findings that phages tend to carry tRNAs corresponding to codons that are used frequently by the phage genes while rare in the host genome³⁹. Also, it is possible that tRNA genes present in phages facilitate the expression of the late phage gene encoding structural proteins as proposed from cyanophage Syn9⁴⁰.

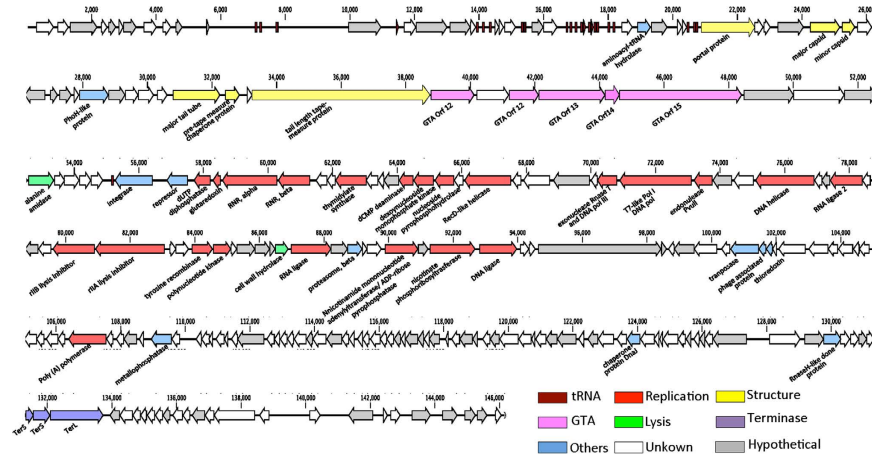


Figure 2. Genome map of roseophage DSS3Φ8. Gene features (unknown, hypothetical, tRNA genes, GTA and others) and genome modules (structure, lysis and DNA replication) are color-coded according to the legend below the figure.

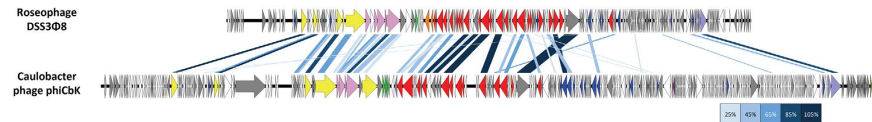


Figure 3. Genome wide comparison between roseophage DSS3Φ8 and *Caulobacter* phage phiCbK. White arrows represent ORFs for which no putative function can be attributed. Yellow, green, red and purple arrows represent structure, lysis, replication and packaging ORFs, respectively. Pink arrows stand for GTA structure, while orange arrows represent prophage-related ORFs. Grey color stands for hypothetical proteins. Related genes of these two phages are connected by blue shading. The color box corresponds to different amino acid identities.

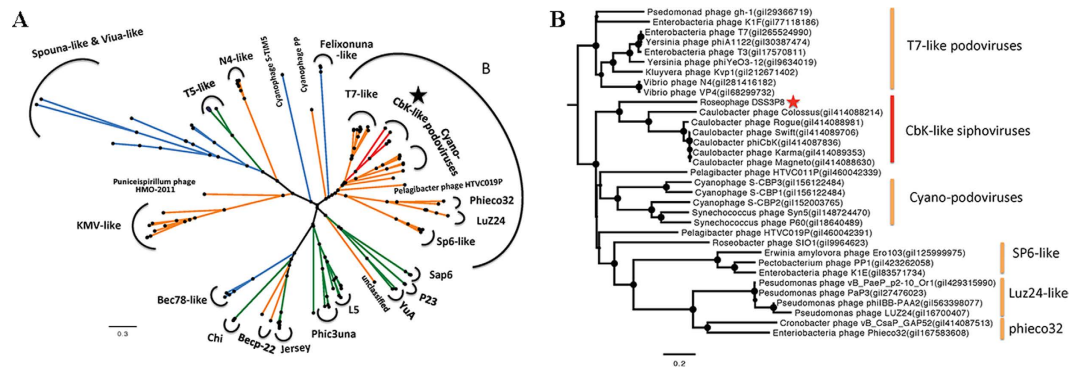


Figure 4. (A) The maximum likelihood phylogenetic tree of DNA polymerase I of bacteriophages. Orange, blue and green colors represent *Podoviridae*, *Myoviridae* and *Siphoviridae*, respectively. Red color represents the CbK-like phages. Bootstrap = 500. **(B)** Zoom in of region B in Fig. 4A.

Homology between DSS3Φ8 and CbK-like phages. Genome wide comparison between DSS3Φ8 and the CbK-like phages reveals that they all share similar modules (Fig. 3). Of the 106 characterized ORFs with homologs in the DSS3Φ8 genome, 24 were related to the CbK-like phages, including the genes involved in DNA metabolism, replication, structure and genes with unknown function (Supplementary Table S1). The phylogenetic analyses based on the DNA polymerase gene (Fig. 4), major capsid gene (Supplementary Fig. S1) and terminase large subunit gene (Supplementary Fig. S2) all supported that DSS3Φ8 is more closely related to the CbK-like phages than to other phages. Also DSS3Φ8 forms a deep branch of its own among the CbK-like phages.

Although siphoviruses are usually highly mosaic in terms of their genomic evolution, siphoviruses with prolate capsid appear to maintain a conserved genomic architecture among them even though their genome sizes vary dramatically. For example, phage S-CBS2⁴¹, a siphovirus with prolate capsid, infecting a marine *Synechococcus* strain, exhibits a genomic arrangement similar to P-SS2³⁵, a siphovirus with prolate head, infecting a marine *Prochlorococcus* strain. The genome size of phage S-CBS2, however, is only two thirds of that of phage P-SS2⁴¹.

In our case, DSS3Φ8 infecting a marine alphaproteobacterium, *R. pomeroyi* DSS3, shared the highest genome homology with six siphoviruses CbK-like phages infecting a freshwater alphaproteobacterium, *Caulobacter crescentus*, despite that the genome size of DSS3Φ8 is about two-third of the genome sizes of CbK-like phages. The high genomic homology between DSS3Φ8 and CbK-like phages is remarkable because they infect bacteria inhabiting two different environments (marine vs. freshwater). A recent survey of the global oceans showed that siphoviruses with prolate capsids can make up 48% of siphoviruses observed in the marine environment¹⁴, suggesting that the genetic diversity of marine siphoviruses with prolate capsids remains largely unexplored.

Difference between DSS3Φ8 and CbK-like phages. Despite the conserved genome arrangement between DSS3Φ8 and the CbK-like phages, an obvious genome reduction of DSS3Φ8 was observed. DSS3Φ8 contains 229 ORFs, while CbK-like phages have 318 to 448 ORFs³³. The majority of missing genes are those without any matches in the NCBI database. The structure genes of DSS3Φ8 are much smaller than those of phiCbK, especially for the tail related genes. For instance, the length of pre-tape measure gene of DSS3Φ8 is half of phiCbK. The reduction of structure genes in DSS3Φ8 is consistent with the smaller head and shorter tail of DSS3Φ8 in comparison to the CbK-like phages. It appears that the loss of genes and the reduction of gene size both contribute to the reduction of the DSS3Φ8 genome.

For CbK-like phages, a lysis gene cassette coding for endolysin, holin and spanin was found between structure and replication gene cassettes. However, at the same region, DSS3Φ8 only contains one conserved lysis gene - peptidoglycan amidohydrolase (Fig. 3). On the other hand, DSS3Φ8 contains two hypothetical genes upstream to the endolysin gene, which we predict are involved in the lysis of the host cell based on their location.

Replication System. DSS3Φ8 contains a conserved replication system, which includes genes coding for DNA polymerase, helicase, ligase, recombinase and nucleotide metabolism (Supplementary Table S1). DSS3Φ8 and the CbK-like phages are siphoviruses, but they likely hijacked the DNA polymerase gene from T7-like podoviruses. Such a genetic recombination between siphoviruses and podoviruses was first reported in the CbK like phages³³. Phylogenetic analysis based on *polA* showed that DSS3Φ8 and the CbK-like phages form their own cluster within the *Podoviridae* clade, and closed related with T7-like *polA* cluster (Fig. 4).

T7-like podoviruses typically contain a simple and conserved replication system, which includes the genes coding for DNA polymerase, helicase-primase, a single-stranded DNA binding protein, and thioredoxin⁴². DSS3Φ8 and the CbK-like phages possess the helicase gene, but lack the primase gene and the single strand binding protein gene. Possessing the thioredoxin gene may help proliferating phage DSS3Φ8 in marine environments since thioredoxin is known to promote DNA replication rate of T7-like DNA polymerase gene⁴³. The thioredoxin gene has been found in many podoviruses infecting marine bacteria^{44,45}.

Prophage. DSS3Φ8 encodes an integrase and transcriptional regulator between the structure and replication modules, implying the lysogenic potential of this phage. The closest match (E-value < 2.5E-51, 32% identity) with the DSS3Φ8 integrase gene is in roseobacter *Ruegeria* sp. 6PALISEP08, which is 99% identical to the host strain *R. pomeroyi* DSS-3 based on the 16S rRNA sequences. The transcriptional regulator gene contains Cro/C1-type HTH and peptidase s24 domains, therefore was identified as a repressor. The genome of *R. pomeroyi* DSS-3 does not contain a prophage³⁸. We searched for the site-specific attachment sites (*attP* and *attB*) in the genomes of phage DSS3Φ8 and host DSS-3, and found one possible attachment site, a 63bp sequence near the tRNA upstream of the integrase gene, which is 81% identical to the tRNA-Met4 gene in the host genome. Within the 63bp region, a 18bp segment was 100% identical between the phage and host. The *attP-attB* sites have been found in many bacteriophages including siphoviruses infecting marine picocyanobacteria^{35,41}, and have been used to link phages with their potential hosts⁴⁶. Because of the conserved nature of tRNAs, they can become the integration sites for temperate phages⁴⁷. Whether this 63bp region is the potential integration site for DSS3Φ8 remains to be tested.

The GTA genes. Several hallmark genes of DSS3Φ8 are homologous to genes 12, 13, 14 and 15 in RcGTA, the gene transfer agent found in *Rhodobacter capsulatus*. These four GTA related genes have been found in the CbK-like phages³³, phage ΦJL001 infecting marine *Rhizobiales* str. JL001⁴⁸, and roseophage RDJLΦ1 infecting *Roseobacter denitrificans* Och114³¹. It is interesting that DSS3Φ8 carries two copies and larger size of RcGTA gene 12, which separated by an unknown protein.

GTA is a well-preserved genetic structure in alpha-proteobacteria, especially in the Roseobacter lineage. A phylogenetic analysis based on concatenated translated sequences of these four GTA genes showed that all members of roseobacters fell into their own clades (Fig. 5), which are consistent with the phylogeny based on their whole genome sequences²⁹. Interestingly, phage-derived four GTA genes belonged to three different clades, which were not closely related to those of their potential hosts. This result suggests a host-independent evolution of phage-encoded GTA genes. Despite the fact that DSS3Φ8 and the CbK-like phages infect different bacterial species, the GTA gene phylogeny clusters these phages together, suggesting that DSS3Φ8 and the CbK-like phages acquired these four GTA genes from a common ancestor (Fig. 5).

It is noteworthy that these four GTA genes are not found in several podoviruses including SIO-like phages and N4-like phages that infect different marine roseobacter strains^{18,19}. It is possible that these four GTA genes are only common in siphoviruses infecting different roseobacter strains including phages RDJLΦ1, and DSS3Φ8. These four GTA genes were also found in a few unpublished siphoviruses infecting roseobacters (e.g. KT253150, NC_026608 and KT266805). The finding of roseophages carrying the four GTA genes indicates that there is a long and tangled evolutionary relationship between GTAs, phages and even prophages⁵⁰. Isolation and genome sequencing of more roseophages will provide new insight into the evolutionary relationship between roseophages, GTAs and the Roseobacter clade.

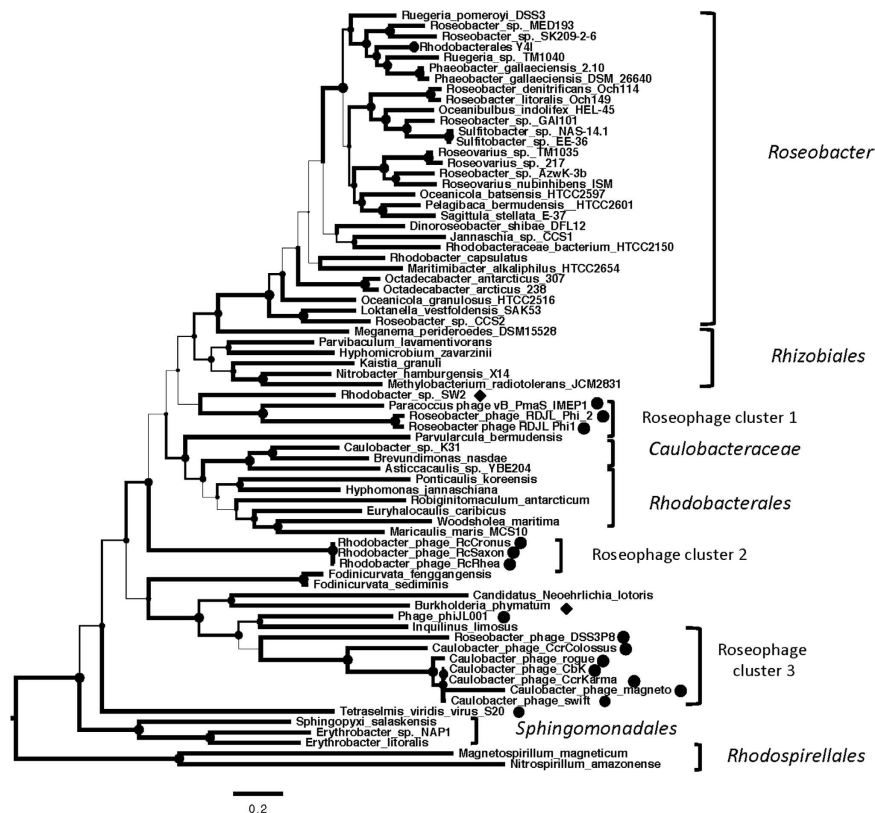


Figure 5. The maximum likelihood phylogenetic tree of concatenated protein sequences of RcGTA-like genes 12–15 from bacteriophages and bacteria. The circle represents the phage-derived four GTA genes, while the diamond indicates the GTA gene cluster came from the bacteria. Bootstrap = 500. The width of the branch corresponds to the bootstrap value.

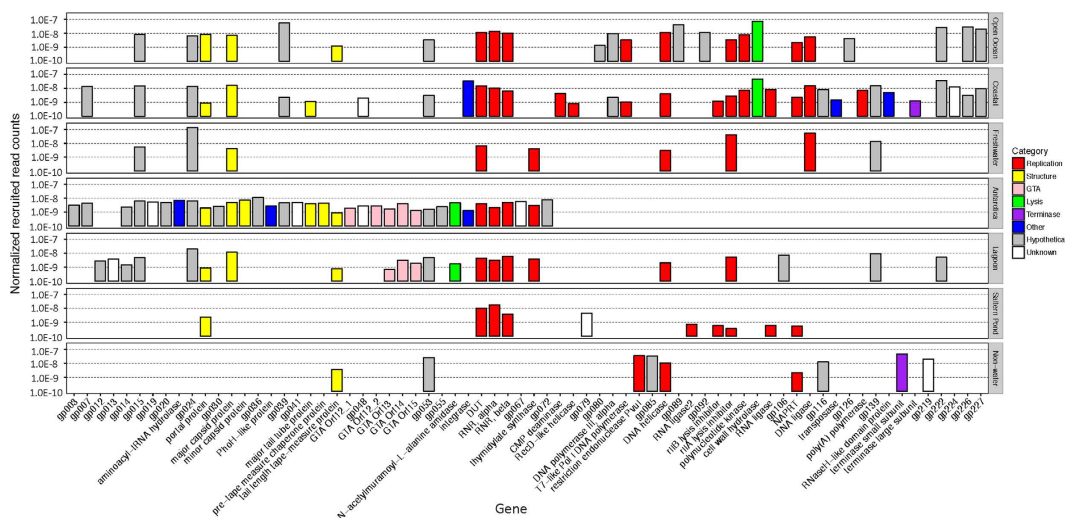


Figure 6. Rates of occurrences of DSS3Φ8 genes in various metagenomic databases. The colors of the bars indicate the categories of genes. The absolute counts of retrieved reads were normalized against the data sizes of metagenomes and the average gene size.

Environmental distribution of DSS3Φ8. Metagenomic recruitments show that DSS3Φ8 reads were detected at the rates ranging from 10^{-10} to 10^{-7} per base pair in the metagenomic databases. The DSS3Φ8 homologs can be found in diverse habitats, ranging from freshwater to open ocean (Fig. 6). The highest recruitment rate came from the samples from Scripps Pier of Pacific Ocean Virome, where many SIO-like roseophages were isolated^{17,18}. DSS3Φ8 was present in Antarctica, especially Organic Lake and Ace Lake, where abundant N4-like phages (Roseo N4-like phages) were observed^{21,49}. DSS3Φ8 homologs were also recruited from other

environments, such as chimney biofilm, stromatolite and whalebone. It appears that DSS3phi8-like phages are widespread in aquatic and marine environments.

Conclusion

Lytic phage DSS3Φ8 is a new member of marine roseobacter phages, and represents the first CbK-like phage found in the marine environment. The genome sequence of DSS3Φ8 differs largely from all the known roseophages. DSS3Φ8 contains many features related to *Podoviridae*, *Siphoviridae*, prophage and GTA. Acquisition of T7-like DNA pol gene provides an efficient DNA replication mechanism to phage DSS3Φ8. On the other hand, carrying the integrase and repressor genes implies the lysogenic potential of DSS3Φ8. With its highly mosaic genome, DSS3Φ8 could serve as an ideal phage system to study the genomic evolution of phages and how the acquisition of different genetic elements (i.e. *pol*, *int*, or GTA genes) affect the phenotypic characteristics of phages.

Methods and Materials

Isolation of phage. Host strain *R. pomeroyi* DSS-3 was grown in ½ YTSS medium (4 g yeast extract, 2.5 g tryptone and 20 g Crystal Sea per liter) at 28 °C. Water samples, collected from Baltimore Inner Harbor Pier V in January 2012, were filtered through 0.22 μm polycarbonate membrane filters (Millipore, Bedford, MA, USA). Filtrate of 10 ml was added into 50 ml of exponentially growing bacterial cultures and incubated for two days. The DSS-3 culture was filtered through 0.22 μm polycarbonate membrane filter (Millipore, Bedford, MA, USA) to remove bacteria. 100 μl of this cell-free lysate was added to 500 μl of exponentially growing DSS-3 culture, and plated using plaque assay. Phage isolates were purified three times by plaque assay.

Transmission electron microscopy (TEM). For TEM image, one drop of purified phage lysate was adsorbed to the 200-mesh Formvar/carbon-coated copper grid and stained with 1% phosphotungstic acid (PTA) for one minute⁵¹. Samples were examined with a JEOL JEM2100F transmission electron microscope, at the University of Oldenburg.

Cross infection. Cross-infectivity of roseophage DSS3Φ8 was tested against other five marine roseobacter stains, including *Roseovarius nubinhibens* ISM, *Silicibacter* sp. TM1040, *Sulfitobacter* sp. EE-36, *Dinoroseobacter shibae* DFL-12 and *Roseobacter denitrificans* OCh 114. For each roseobacter strain, 500 μl of exponentially growing cells was added to 5 ml top agar and poured on the plates. After the agar was solidified, one drop of purified phage lysate was spotted onto each plate, and incubated for 2–3 days at 28 °C¹⁸. The formation of plaques was examined for possible cross-infectivity.

One-step growth curve. One-step growth curve was determined using a method described elsewhere⁵². Generally, exponentially growing cultures of *R. pomeroyi* DSS-3 were incubated on ice for 20 min to synchronize the growth of host. After the DSS-3 culture was recovered to room temperature, roseophage DSS3Φ8 was added into the culture at a multiplicity of infection (moi) of 0.5. After 20 min attachment, cells were pelleted, re-suspended and diluted 100 times to avoid possible secondary infection. An aliquot of the cell suspension was collected at different time points between 30 min and 12 h.

The number of phages was enumerated by using qPCR^{53–55}. Collected samples were filtered through 0.22 μm polycarbonate membrane filters (Millipore, Bedford, MA, USA) and DNA was extracted by WIZARD PCR Preps DNA Purification System (Promega Corporation, WI, USA). A set of PCR primers was designed based on the sequence of the phage DNA polymerase (Forward: ATGCTGCTCCGAACGTATCT, Reverse: ACTCGCCCTTCTTTTCCTTC). qPCR was conducted by APPLIED BIOSYSTEMS, 7500 Fast Real-time PCR system (ThermoFisher Scientific, New York, USA). qPCR reactions were performed in a 25 μl volume with 12.5 μl PERFECTA SYBR Green SuperMix (Quanta Biosciences, Gaithersburg, MD, USA), 400 nM forward and reverse primers and 5 μl of template. The amplification program was set as: 95 °C for 10 min, 40 cycle of 95 °C for 30 s, 60 °C for 30 s, 72 °C for 30 s. Fluorescence measurements were conducted at 60 °C of each cycle. Standards were developed using DNA extracted from phage lysate with known titer and 10-fold serial dilutions of the DNA were used in the reactions. Standard curves were determined as the correlation between the log of gene copy numbers and the Ct ($R^2 = 0.99$).

DNA extraction. One liter phage lysate was treated with DNase and RNase at a final concentration of 2 μg/ml (for both enzymes) at 4 °C to remove free DNA and RNA. Phage particles in the supernatant were precipitated with polyethylene glycol 8000 (final 10% w/v), and centrifuged in an OPTIPREP Density Gradient Medium for 12 h at 41,000 g. The visible viral band was extracted and dialyzed against TM buffer overnight at 4 °C. The purified phage particles were treated with a combination of sodium dodecyl sulfate (SDS, final concentration 1% w/v), proteinase K (final concentration 30 μg/ml) and EDTA (final concentration 5 μM) at 55 °C for 3 h. Phage DNA was then extracted using phenol/chloroform/isoamyl alcohol (25:24:1).

Sequencing and annotation. The genome of roseophage DSS3Φ8 was sequenced by Illumina GAIIX sequencer. A 112 bp paired-end run resulted in 157,500 reads. The high quality reads were assembled by MIRA assembler, resulting in three contigs (139,090 bp, 2,108 bp and 4,767 bp in size). Primer walking with the Sanger sequencing method was followed to close the phage genome. Sanger sequencing was performed using ABI 3100 genetic analyzer (PE Applied Biosystems, Foster City, CA, USA) at the Institute of Marine and Environmental Technology, University of Maryland Center for Environmental Science. The circular feature of DSS3Φ8 was evident as we closed the genome sequences by primer walking from opposite directions.

The open reading frames (ORFs) were predicted using GeneMarkS⁵⁶ and tRNAscan-SE⁵⁷ was used to identify tRNA sequence. Translated ORFs were compared with known protein sequences using BLASTP (E-value < 10⁻⁵ and identity > 30%). The complete genome sequence was submitted to GenBank with the accession number of KT870145.

Sequence alignment and phylogenetic analysis for *polA*, major capsid protein, terminase large subunit and concatenated four GTA genes were performed using the CLC Genomic 7 program. A maximum-likelihood method was used to construct the phylogenetic tree, with bootstrap value of 500.

Metagenomic recruitment. To explore the geographic distribution of DSS3Φ8, the phage genes were used as queries to search against all the available metagenomic databases in iMicrobe Database (<http://data.imicrobe.us/>) by June 2015. A reciprocal Blast analysis was conducted. Each of the recruited metagenomic reads was compared against the NCBI RefSeq database, which contains all complete viral and bacterial genomes by June 2015. The metagenomic reads were considered as DSS3Φ8 origin only if the DSS3Φ8 genes were the best hits. The relative abundance of each gene in every metagenomic database was calculated following the methods described elsewhere¹⁵. The count for each recruited read was divided by the number of total reads in the database, and further divided by the size of the gene product. The metagenomic databases were then categorized by their metadata. Samples were scaled using the mean of all samples.

References

- González, J. M. & Moran, M. A. Numerical dominance of a group of marine bacteria in the alpha-subclass of the class *Proteobacteria* in coastal seawater. *Appl. Environ. Microbiol.* **63**, 4237–4242 (1997).
- Moran, M. A. *et al.* Ecological genomics of marine Roseobacters. *Appl. Environ. Microbiol.* **73**, 4559–4569 (2007).
- Brinkhoff, T., Giebel, H. A. & Simon, M. Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch. Microbiol.* **189**, 531–539 (2008).
- Buchan, A., González, J. M. & Moran, M. A. Overview of the marine *Roseobacter* lineage. *Appl. Environ. Microbiol.* **71**, 5665–5677 (2005).
- Moran, M. A., Reisch, C. R., Kiene, R. P. & Whitman, W. B. Genomic insights into bacterial DMSP transformations. *Ann. Rev. Mar. Sci.* **4**, 523–542 (2012).
- Wagner-Döbler, I. & Biebl, H. Environmental biology of the marine *Roseobacter* lineage. *Annu. Rev. Microbiol.* **60**, 255–280 (2006).
- Luo, H. & Moran, M. A. How do divergent ecological strategies emerge among marine bacterioplankton lineages? *Trends Microbiol.* **23**, 577–584 (2015).
- Fuhrman, J. A. Marine viruses and their biogeochemical and ecological effects. *Nature* **399**, 541–548 (1999).
- Suttle, C. A. Marine viruses - major players in the global ecosystem. *Nature Rev. Microbiol.* **5**, 801–812 (2007).
- Suttle, C. A. Viruses in the sea. *Nature* **437**, 356–361 (2005).
- Wommack, K. E. & Colwell, R. R. Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* **64**, 69–114 (2000).
- Angly, F. E. *et al.* The marine viromes of four oceanic regions. *PLoS Biol.* **4**, 2121–2131 (2006).
- Hurwitz, B. L. & Sullivan, M. B. The Pacific Ocean Virome (POV): A marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. *PLoS One* **8**, 1–12 (2013).
- Brum, J. R., Ignacio-Espinoza, J. C., Roux, S. & Doucier, G. Patterns and ecological drivers of ocean viral communities. *Science* **348**, 1–10 (2015).
- Zhao, Y. *et al.* Abundant SAR11 viruses in the ocean. *Nature* **494**, 357–360 (2013).
- Kang, I., Oh, H.-M., Kang, D. & Cho, J.-C. Genome of a SAR116 bacteriophage shows the prevalence of this phage type in the oceans. *Proc. Natl. Acad. Sci. USA* **110**, (2013).
- Rohwer, F. *et al.* The complete genomic sequence of the marine phage Roseophage SIOI shares homology with nonmarine phages. *Limnol. Oceanogr.* **45**, 408–418 (2000).
- Angly, F. *et al.* Genomic analysis of multiple Roseophage SIO1 strains. *Environ. Microbiol.* **11**, 2863–2873 (2009).
- Zhao, Y., Wang, K., Jiao, N. & Chan, F. Genome sequences of two novel phages infecting marine roseobacters. *Environ. Microbiol.* **11**, 2055–2064 (2009).
- Ankrah, N. Y., Budinoff, C. R., Wilson, W. H., Wilhelm, S. W. & Buchan, A. Genome sequence of the *Sulfitobacter* sp. strain 2047-infecting lytic phage CB2047-B. *Genome Announc* **2**, 1–2 (2014).
- Chan, J. Z. M., Millard, A. D., Mann, N. H. & Schäfer, H. Comparative genomics defines the core genome of the growing N4-like phage genus and identifies N4-like Roseophage specific genes. *Front. Microbiol.* **5**, 1–14 (2014).
- Ji, J., Zhang, R. & Jiao, N. Complete genome sequence of Roseophage vB_DshP-R1, which infects *Dinoroseobacter shibae* DFL12. *Stand. Genomic Sci.* **10**, 1–8 (2015).
- Zhang, Y. & Jiao, N. Roseophage RDJLΦ1, Infecting the aerobic anoxygenic phototrophic bacterium *Roseobacter denitrificans* OCh114. *Appl. Environ. Microbiol.* **75**, 1745–1749 (2009).
- Ankrah, N. Y. D., Budinoff, C. R., Wilson, W. H., Wilhelm, S. W. & Buchan, A. Genome sequences of two temperate phages, CB2047-A and CB2047-C, infecting *Sulfitobacter* sp. strain 2047. *Genome Announc* **2**, 1–2 (2014).
- Chen, F., Wang, K., Stewart, J. & Belas, R. Induction of multiple prophages from a marine bacterium: a genomic approach. *Appl. Environ. Microbiol.* **72**, 4995–5001 (2006).
- Zhao, Y. *et al.* Searching for a “hidden” prophage in a marine bacterium. *Appl. Environ. Microbiol.* **76**, 589–595 (2010).
- Paul, J. H. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J* **2**, 579–589 (2008).
- Zhao, Y. *et al.* Gene transfer agent (GTA) genes reveal diverse and dynamic *Roseobacter* and *Rhodobacter* populations in the Chesapeake Bay. *ISME J* **3**, 364–373 (2008).
- Newton, R. J. *et al.* Genome characteristics of a generalist marine bacterial lineage. *ISME J* **4**, 784–798 (2010).
- Lang, A. S. & Beatty, J. T. Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Natl. Acad. Sci. USA* **97**, 859–864 (2000).
- Huang, S., Zhang, Y., Chen, F. & Jiao, N. Complete genome sequence of a marine roseophage provides evidence into the evolution of gene transfer agents in alphaproteobacteria. *Viol. J.* **8**, 1–6 (2011).
- Ackermann, H. W. Frequency of morphological phage descriptions in the year 2000. *Arch. Virol.* **146**, 843–857 (2001).
- Gill, J. J. *et al.* The *Caulobacter crescentus* phage phiCbK: genomics of a canonical phage. *BMC Genom.* **13**, 1–20 (2012).
- Agabian-Keshishian, N. & Shapiro, L. Stalked bacteria: properties of deoxyribonucleic acid bacteriophage φCbK. *J. Virol.* **5**, 795–800 (1970).
- Sullivan, M. B. *et al.* The genome and structural proteome of an ocean siphovirus: a new window into the cyanobacterial ‘mobilome’. *Environ. Microbiol.* **11**, 2935–2951 (2009).
- Bahir, I., Fromer, M., Prat, Y. & Linial, M. Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. *Mol. Syst. Biol.* **5**, 1–14, 311 (2009).

37. Krakauer, D. C. & Jansen, V. A. A. Red queen dynamics of protein translation. *J. Theor. Biol.* **218**, 97–109 (2002).
38. Moran, M. A. *et al.* Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* **432**, 910–913 (2004).
39. Bailly-Bechet, M., Vergassola, M. & Rocha, E. Causes for the intriguing presence of tRNAs in phages. *Genome Res.* **17**, 1486–1495 (2007).
40. Weigle, P. R. *et al.* Genomic and structural analysis of Syn9, a cyanophage infecting marine *Prochlorococcus* and *Synechococcus*. *Environ. Microbiol.* **9**, 1675–1695 (2007).
41. Huang, S., Wang, K., Jiao, N. & Chen, F. Genome sequences of siphoviruses infecting marine *Synechococcus* unveil a diverse cyanophage group and extensive phage–host genetic exchanges. *Environ. Microbiol.* **14**, 540–558 (2012).
42. Lee, S. J. & Richardson, C. C. Choreography of bacteriophage T7 DNA replication. *Curr. Opin. Chem. Biol.* **15**, 580–586 (2011).
43. Tabor, S., Huber, H. E. & Richardson, C. C. *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem.* **262**, 16212–16223 (1987).
44. Chen, F. & Lu, J. Genomic sequence and evolution of marine cyanophage P60: a new insight on lytic and lysogenic phages. *Appl. Environ. Microbiol.* **68**, 2589–2594 (2002).
45. Pope, W. H. *et al.* Genome sequence, structural proteins, and capsid organization of the cyanophage Syn5: a “horned” bacteriophage of marine *Synechococcus*. *J. Mol. Biol.* **368**, 966–981 (2007).
46. Mizuno, C. M., Rodriguez-Valera, F., Kimes, N. E. & Ghai, R. Expanding the marine virosphere using metagenomics. *PLoS Genet* **9**, 1–13 (2013).
47. Campbell, A. Prophage insertion sites. *Res Microbiol* **154**, 277–282 (2003).
48. Lohr, J. E., Chen, F. & Hill, R. T. Genomic analysis of bacteriophage ΦJL001: insights into its interaction with a sponge-associated alpha-proteobacterium. *Appl. Environ. Microbiol.* **71**, 1598–1609 (2005).
49. Zhan, Y., Buchan, A. & Chen, F. Novel N4 bacteriophages prevail in the cold biosphere. *Appl. Environ. Microbiol.* **81**, 5196–5202 (2015).
50. Lang, A. S., Zhaxybayeva, O. & Beatty, J. T. Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* **10**, 472–482 (2012).
51. Brenner, S. & Horne, R. W. A negative staining method for high resolution electron microscopy of viruses. *Biophys. Acta* **34**, 103–110 (1959).
52. John, M. W., Martha, R. J. C., Andrew, M. K., Paul, H. & Stephen, T. A. Practical methods for determining phage growth parameters. *Methods in Molecular Biology*, 175–202 (2009).
53. Nana Yaw, D. A. *et al.* Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. *ISME J* **8**, 1089–1100 (2014).
54. Anderson, B. *et al.* Enumeration of bacteriophage particles: Comparative analysis of the traditional plaque assay and real-time QPCR-and nanosight-based assays. *Bacteriophage* **1**, 86–93 (2011).
55. Edelman, D. C. & Barletta, J. Real-time PCR provides improved detection and titer determination of bacteriophage. *Biotechniques* **35**, 368–375 (2003).
56. Besemer, J., Lomsadze, A. & Borodovsky, M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* **29**, 2607–2618 (2001).
57. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).

Acknowledgements

We thank E. Rhiel for help with electron microscopic analysis. Funding by Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center *Roseobacter* (TRR51) to MS, SV and SH and by fellowships of the Hanse Institute for Advanced Study Delmenhorst (HWK) to FC and of the Chinese Scholar Council to YCZ are gratefully acknowledged.

Author Contributions

Y.Z. performed experiments, analyzed data and wrote the manuscripts. F.C. supervised the project and revised manuscript. M.S. participated in manuscript preparation and provided funding for sequencing phage genome. S.H. prepared the TEM images and S.V. sequenced the phage genome.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Zhan, Y. *et al.* A novel roseobacter phage possesses features of podoviruses, siphoviruses, prophages and gene transfer agents. *Sci. Rep.* **6**, 30372; doi: 10.1038/srep30372 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016