



HAL
open science

Genome analysis of *Marinobacter* phage AS1 suggests its close interactions with host *Marinobacter* sp

Sreekumar Aparna, Ammini Parvathi, Angia Sriram Pradeep Ram,
Télesphore Sime-Ngando

► To cite this version:

Sreekumar Aparna, Ammini Parvathi, Angia Sriram Pradeep Ram, Télesphore Sime-Ngando. Genome analysis of *Marinobacter* phage AS1 suggests its close interactions with host *Marinobacter* sp. *Aquatic Microbial Ecology*, 2019. hal-03043263

HAL Id: hal-03043263

<https://hal.science/hal-03043263>

Submitted on 7 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Genome analysis of Marinobacter phage AS1 suggests its close interactions with host**

2 ***Marinobacter* sp.**

3 Aparna S.¹, Parvathi A.^{1*}, Pradeep Ram A. S.² and Sime-Ngando T.²

4
5 ¹ CSIR-National Institute of Oceanography, Regional Centre (CSIR), Kochi-682 018, India

6 ² Laboratoire Microorganismes : Génome et Environnement, UMR CNRS 6023, Université Clermont-
7 Auvergne, 1 Impasse Amélie Murat, 63178 Aubière Cedex, France

8
9
10
11 * **Corresponding author**

12 Dr. Parvathi A.

13 CSIR-National Institute of Oceanography

14 Regional Centre, Dr. Salim Ali Road,

15 Post Box No. 1913, Kochi-682 018, India

16 Ph: 91-(0) 484-2390814, Fax: 91-(0) 484-2390618

17 Email: parvathi@nio.org

18
19
20 **Running title: Whole genome characterization of Marinobacter phage AS1**

1 **Abstract**

2 *Marinobacter* comprises a widespread and ecologically significant genus of marine bacteria
3 with vast metabolic capabilities. However, very few reports on phages infecting this genus exist. In
4 this study, we isolated and characterized a linear dsDNA genome of *Marinobacter* phage AS1,
5 infecting *Marinobacter sp.* strain, D1S9. The phage is a member of Podoviridae family having a
6 genome size of ~37 kb and a 57 % GC content with no significant overall similarity to any of the
7 available viral whole genome sequences in the database. The genome which encodes for 52 protein
8 coding genes with no tRNA genes, contains 21 protein coding genes with assigned putative functions.
9 High degree of recombination events with other members of the family *Caudovirales* is evident from
10 the mosaic pattern of inheritance of genes such as major capsid protein, terminase large sub unit,
11 portal protein and integrase. Presence of integrase gene, DNA methyl transferases and the occurrence
12 of AS1 sequences within the whole genome sequences of some members of the host taxon were
13 suggestive of its intense association with the host.

14

15 **Keywords:** *Marinobacter*, phage, complete genome sequence, Podovirus, Arabian Sea

16

17

18

19

20

21

22

23

1 **1. Introduction**

2 Viruses, primarily bacteriophages are the most prolific and diverse obligate parasites in the
3 marine biosphere where they outnumber bacteria on an average by one order of magnitude (Suttle,
4 2007, Sime-Ngando 2014). Bacteriophages have indispensable involvement in the ecological and
5 biogeochemical processes by regulating bacterial abundance, bacterial diversity, nutrient fluxes,
6 structuring the microbial food web dynamics and to a great extent in influencing host diversification
7 and speciation through horizontal gene transfer (Middelboe and Lyck, 2002, Bouvier and Del Giorgio
8 2007, McDaniel et al. 2010, Winget et al. 2011). A considerable fraction of bacterial genetic diversity
9 is acquired through horizontal gene transfer, imparting ecological and pathogenic attributes like
10 antibiotic resistance, resistance to phage infection, toxins and heavy metals (Stern and Sorek, 2010,
11 Xia and Wolz, 2014). Phage genomes are also coevolving concomitantly with their hosts. Compared
12 to other host-parasite systems, bacterial-phage coevolution is faster owing to their large population
13 sizes and short generation times (Buckling and Brockrust 2012). Studies with coevolving populations
14 of hosts and their infective phages revealed two modes of coevolution: arms race dynamics (ARD)
15 and fluctuating selection dynamics (FSD). These two were found to operate at different times in the
16 same population with arms race dynamics giving way to fluctuating selection dynamics in the long
17 run (Hall et al. 2011, Gokhale et al. 2013).

18 Most of the dsDNA phages share a large pool of genetic elements (Hendrix et al. 1999).
19 Phage genomes exhibit prevalent mosaic architecture as a result of extensive horizontal gene transfer
20 and in fact different genomic segments may have different ancestral origins. So, it is challenging to
21 deduce whole genome similarities and to assign phylogenetic relationships to newly sequenced
22 phages (Hatfull and Hendrix 2011). Isolation and genome characterization of additional phages would
23 help to resolve this problem to some extent.

1 Despite of the profound knowledge on abundance and role of marine viruses in
2 biogeochemical cycles and host population dynamics, genome level information on phage-host
3 interactions of ecologically significant bacteria are very limited. Sequencing and whole genome
4 characterization of phages infecting these bacteria might provide insights into the phage-host
5 interactions and how they shape the ecological roles of the hosts. *Marinobacter*, an ecologically
6 significant genus of bacteria, and its first representative, *M. hydrocarbonoclasticus*, was proposed by
7 Gauthier in 1992. Since then, new members belonging to this genus have been isolated from diverse
8 marine habitats ranging from oil producing wells to deep-sea benthic sediments with fifty-eight
9 species reported so far (Martin et al. 2003, Yoon et al. 2007, Xu et al. 2008, Zhang et al. 2008, Wang
10 et al. 2009, Wang et al. 2012, Chua et al. 2018). Several species have been associated with marine
11 sponges and microalgae (Green et al. 2006, Kaepfel et al. 2012, Lee et al. 2012, Lupette et al. 2016,
12 Sandhya et al. 2017). Members of this genus are known to have substantial impact on marine snow
13 formation and on various biogeochemical cycles with quite a few strains possessing a repertoire of
14 activities like degradation of various hydrocarbon compounds, denitrification, Fe (II) oxidation, Mn
15 (II) oxidation, Fe (III) reduction and redox cycling of arsenic and fumarate (Kaye et al. 2010, Kaepfel
16 et al 2012, Wang et al. 2012, Handley and Lloyd 2013, Bonis and Gralnick 2015). Although their
17 versatile functional capabilities enable them to survive in diverse environmental conditions and to
18 exploit variable resources (Singer et al. 2011), viruses infecting this genus are reported rarely (Zhu et
19 al. 2018), thus hindering an integrated understanding of the life cycle of these microbes in the marine
20 environments. In this study, we report the isolation of a bacteriophage infecting the genus
21 *Marinobacter* from surface waters of the Arabian Sea (southwest coast of India) and provide first-
22 hand information on the sequencing and whole genome analysis of it.

23

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

2. Materials and Methods

2.1. Isolation and purification of Marinobacter phage AS1 and its host

Marinobacter phage AS1 was isolated from the surface water of the tropical Arabian Sea, off the coast of Kochi, southwest coast of India (9.9586° N and 76.0825° E, Fig. 1A) in December 2014 as described by Middelboe et al. (2010). Briefly, 25 mL water sample was amended with 1 mL 10× Zobell marine broth and incubated at room temperature for 2 days in order to propagate potential bacterial hosts and thereby the specific viruses in the sample (Himedia, Mumbai, India). The viral enriched seawater was screened against lawns of bacterial isolates obtained from the same water sample by means of soft agar overlay technique. Culturing of bacteria was done by spread plating water samples on Zobell Marine Agar. The plaques formed on the bacterial lawn after overnight incubation (37 °C) were picked and transferred to phage buffer (SM buffer: 450 mM NaCl, 50 mM MgSO₄, 50 mM Tris, 0.01 % gelatin, pH 8), vortexed well and kept at 4°C. Centrifugation (10,000 ×g, 10 min) was done the next day to remove the soft agar and host cell remnants which was followed by filtration through 0.2µm pore size Acrodisc syringe filters (Sigma-Aldrich, India). The filtrate was then serially diluted and plated by soft agar overlay method. Single plaque from the plate was isolated and the purification was repeated for five consecutive times. The phage stock was stored at 4°C until use.

The bacterial isolate for which the phage was isolated was identified using 16S rRNA gene polymerase chain reaction (Amann et al. 1995). The PCR products were sequenced and identified by BLAST search against non-redundant database of NCBI.

2.2. Transmission Electron microscopy (TEM)

1 For morphometric characterization of phage by TEM analysis, the purified phage particles
2 were collected onto 400-mesh carbon coated formvar grids by centrifugation (70,000 ×g, 20 minutes
3 at 4°C) using a SW 40Ti rotor and stained at room temperature for 30 s with uranyl acetate stain (2 %
4 w/v) (Pradeep Ram et al. 2010). Excess stain on grids was removed by rinsing twice with 0.02 µm-
5 filtered distilled water and dried on filter paper. Grids were examined using a JEOL 1200Ex
6 transmission electron microscope (JEOL, Akishima, Japan) operated at 80 kV at a magnification
7 between × 60,000 and × 100,000.

8 **2.3. Phage DNA isolation**

9 Phage particles were precipitated with 8% polyethylene glycol (PEG) 8000 containing 1M
10 NaCl after overnight incubation at 4°C (Lawrence and Steward 2010). The mixture was centrifuged at
11 10000×g for 20 min at 4°C and the pellet was suspended in SM buffer to disaggregate overnight at
12 4°C. To remove the PEG, KCl was added to the viral suspension to a final concentration of 1M and
13 kept on ice for 30 min. PEG was pelleted by centrifugation at 12000×g for 10 min at 4°C and the
14 supernatant containing viral particles were transferred to a fresh tube. The viral suspension was
15 concentrated using Amicon Ultra-15 centrifugal filter unit with a molecular weight cut off of 30 kDa
16 (Millipore, India). The concentrated phage preparation was treated with nucleases (DNase I and
17 RNase A at final concentrations of 1 mg mL⁻¹) to remove any host DNA and RNA. Phage DNA was
18 extracted using the phenol:chloroform (24:1) extraction method as described previously (Yang et al.
19 2017). Briefly, the nuclease treated phage preparation was treated with proteinase K (100 mg mL⁻¹),
20 10% (w/v), SDS and 0.5 M EDTA (pH 8) and incubated in water bath at 55 °C for 1 hr. This was
21 followed by two rounds of phenol:chloroform:isoamyl alcohol (25:24:1) extraction at 12,000 rpm at
22 4°C for 10 min . Again, the aqueous phase was purified by chloroform:isoamyl alcohol (24:1)
23 extraction at 12,000 rpm at 4°C for 10 min. The supernatant was mixed with isoamyl alcohol and

1 incubated overnight at -20 °C. The pellet was washed with 70% ethanol and air dried. DNA was
2 resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and stored at 4 °C.

3

4 **2.4. Genome sequencing and bioinformatic analyses**

5 Whole genome sequencing library was prepared with Illumina-compatible NEXTFlex rapid
6 DNA sequencing kit (BIOO Scientific, Inc., USA) at Genotypic Technology Pvt. Ltd., Bangalore,
7 India. 100 ng of Qubit quantified DNA was sheared using Covaris S220 sonicator (Covaris, Inc.,
8 USA) to generate specific fragments with size range of 300-400 bp. The fragment size distribution
9 was verified on Agilent 2200 TapeStation and subsequently purified using HighPrep magnetic beads
10 (MagBio Genomics, Inc., USA). Purified fragments were end-repaired, adenylated and ligated to
11 Illumina multiplex barcode adapters as per NEXTFlex rapid DNA sequencing kit protocol.

12 Illumina universal adapters used in the study are:

13 5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-
14 3' and Index Adapter: 5'-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC [INDEX]
15 ATCTCGTATGCCGTCTTCTGCTTG-3'.

16 Adapter-ligated DNA was purified using HighPrep beads. Resultant fragments were amplified
17 for 5 cycles of PCR using Illumina-compatible primers provided in the NEXTFlex rapid DNA
18 sequencing kit. Final PCR product (sequencing library) was purified with HighPrep beads, followed
19 by library quality control. Sequencing library was quantified by Qubit fluorometer (Thermo Fisher
20 Scientific, MA, USA) and its fragment size distribution was analyzed on Agilent 2200 TapeStation.

21 The library was sequenced using Illumina MiSeq (275×2 chemistry) platform. The Illumina
22 paired-end raw reads were quality checked using FastQC (Andrews 2010). The high-quality reads
23 were then assembled using SPAdes assembler (Bankevich et al. 2012). Gene calling and primary

1 annotation of the assembled genome was done using the RAST server (Aziz et al. 2008). The proteins
2 thus predicted were similarity searched against Uniprot Virus and bacterial protein databases with an
3 E-value cut off of $<10^{-5}$. The conserved domains on proteins were identified using NCBI conserved
4 domain database (Marchler-Bauer et al. 2016). The tRNAs were predicted in the assembled genome
5 using tRNAscan-SE (Lowe and Chan 2016). ARNold12 was used for the identification of rho-
6 independent transcriptional terminators (Naville et al. 2011). The search procedure used two
7 complementary programs Erpin and RNAmotif. Potential promoter sequences and the trans-
8 membrane helices in the proteins were predicted using Martin Reese's Neural Network Promoter
9 Prediction (Reese 2001) and TMHMM2.014 server (<http://www.cbs.dtu.dk/services/TMHMM/>),
10 respectively. The SignalP 4.1 server (<http://www.cbs.dtu.dk/services/SignalP/>) was used to detect
11 presence and location of signal peptide in amino acid sequences. Phylogenetic trees were constructed
12 using MEGA version 7 (Kumar et al. 2016) by maximum likelihood method based on the LG+G+I
13 model with ClustalW aligned amino acid sequences of major capsid protein, phage portal protein,
14 terminase large subunit, major tail protein and integrase. Bootstrap values were based on 100
15 replications. Genome sequence of Marinobacter phage AS1 was deposited in GenBank database
16 under accession number, MK088078.

17 **3. Results and Discussion**

18 **3.1. Growth and morphometric characteristics of Marinobacter phage AS1**

19 The host, *Marinobacter sp.* strain D1S9 is a Gram negative, rod shaped bacterium (Fig 1B)
20 isolated from the surface waters of the Arabian Sea. The host present in the water sample was
21 enriched by the addition of nutrients which in turn amplified the specific viruses for the host. The
22 isolated phages formed clear, round plaques of 1-2 mm diameter with regular edges on lawns of the
23 host after 10-12 h of incubation (Fig 1C). Morphometric characteristics of phage AS1 examined by

1 transmission electron microscopy analysis indicated it belonged to family Podoviridae of the order
2 Caudovirales. The phage resembled morphotype C1 with an icosahedral head (isometric) having a
3 capsid diameter of ~50nm (Bradley, 1967). With a presence of short, stubby, non-contractile tail,
4 characteristic of this group (Fig. 2). Host range experiments with 9 laboratory isolates of
5 *Marinobacter* did not result any infection, indicating its narrow host range.

6 **3.2. Genome features and annotation**

7 Illumina sequencing and assembly of AS1 genome generated a single contig with a coverage
8 of 92.5%. The assembled genome of *Marinobacter* phage AS1 was linear with a sequence length of
9 37 kb with a GC content of 57% (Fig. 3). The general features of AS1 genome are listed in Table 1.
10 An initial whole genome similarity search against the NCBI non-redundant database using basic local
11 alignment search tool (BLAST) showed no close relatives to *Marinobacter* phage AS1. AS1 genome
12 showed no significant similarity to any of the three already reported *Marinobacter* phages, PS3
13 (GenBank accession, MF959999), PS6 (GenBank accession, MF959998) or B23 (Zhu et al. 2018;
14 GenBank accession, KY939598). However, the AS1 phage represents the first *Marinobacter* phage
15 belonging to the family Podoviridae. When phage AS1 genome was searched against the whole
16 genome shotgun sequences of the taxon *Marinobacter* (taxid: 2742) using NCBI- BLAST, it showed
17 an identity of 96 % with *Marinobacter manganoxydans* isolate UBA5690_contig_21412
18 (DIHS01000048) and 94 % with *Marinobacter* sp. N4 KEHDKFFH_1 (PSSX01000001) with query
19 coverages of 85% and 73 % respectively. This observation indicates the existence of fragments of
20 AS1 genome within *Marinobacter* population and its possible interactions with the hosts. A total of
21 52 protein coding genes were predicted in the genome, of which 21 have assigned putative functions
22 (Table 2). Genes related to phage structure and assembly, DNA modification, transcriptional
23 regulation and host cell lysis were arranged in distinct functional clusters along the genome (Table 3).

1 The similar sequences of AS1 found within the genomes of *M. manganoxydans* isolate
2 UBA5690_contig_21412 (in reverse orientation) and *Marinobacter* sp. N4 KEHDKFFH_1 are
3 depicted in Fig. 4. They include genes encoding the small and large subunits of terminase, capsid
4 proteins, portal protein, tail proteins, integrase, transcriptional regulators, methyl transferase,
5 endonuclease, lysozyme and hypothetical proteins. Phage sequence was broken up and various
6 functional clusters were shuffled within a particular region of the bacterial genomes.

7 The structural and assembly module encodes both small and large subunits of terminase,
8 portal protein, serine peptidase, major capsid protein, Gp6, Gp3, major tail protein, tail fiber proteins
9 and tail tape measure protein. The terminase large subunit protein (Gene 2) contains a P-loop having
10 nucleoside triphosphate hydrolase activity and belongs to Terminase_1 family (PF03354). Gene 3
11 encodes for a protein with a signal peptide (residues 1-21) and a transmembrane helix (residues 29-
12 51) and is homologous to protein Gp3 of *Klebsiella phage PhiKO2*. Besides, a prohead maturation
13 protease (Gene 5) belonging to the MEROPS peptidase family S49 (protease IV family) was also
14 identified in phage AS1. It has a conserved catalytic Serine-Lysine dyad typical of a serine protease
15 domain (cd07022). The proteolytic activity is essential for the phage capsid maturation. The major
16 capsid protein of AS1 (Gene 6) belongs to HK97 family and it forms coiled coil structure at amino
17 acid positions 4-34 and 39-59. The delta domain of HK97 major capsid protein was shown to have
18 coiled coils involved in prohead assembly and maturation and it was also found to be removed by
19 proteases after the assembly (Oh et al. 2014). Gene 8 of the assembled genome codes for a protein
20 containing gp6 domain (cd08054) with 36 conserved oligomerization interface amino acid residues of
21 the domain. Gp6 of bacteriophage HK97 encodes head tail connector protein which forms an
22 oligomeric ring and serves as an interface for head and tail attachment (Cardarelli et al. 2010).

1 Phage AS1 codes for an integrase gene having 97.2 % homology with that of *Marinobacter*
2 *sp.* N4. It possess a C-terminal conserved domain similar to Shufflon-specific DNA recombinase Rci
3 and bacteriophage Hp1_like integrase (cd00796) belonging to the superfamily of DNA breaking-
4 rejoining enzymes. These enzymes contain a catalytic domain with six conserved amino acid
5 residues. The presence of integrase gene indicates the potential for a temperate lifestyle which is
6 further strengthened by the absence of any tRNA genes in the genome. It can be assumed that phage
7 AS1 utilizes its host's tRNAs for translating the mRNA. Presence of tRNA genes provide phage with
8 competitive advantage over other phages through more efficient expression of their own genes
9 independent of their host's tRNAs. Whereas, their absence results in a narrow host range, making the
10 phage more particular about selecting the host with similar codon usage bias (Bailly-Bechet et al.
11 2007).

12 Two sequence specific DNA methyl transferase genes, DNA adenine methyltransferase
13 (DAM) and DNA (Cytosine-5) methyltransferase (DCM), present in AS1 genome control DNA-
14 protein interactions by methylating adenine and cytosine residues of the DNA sequence. Considering
15 this fact, phage encoded DNA adenine methyl transferase may be interpreted as a coevolutionary
16 adaptation of the phage to protect itself from the host restriction enzymes (Murphy et al. 2014).
17 Normally, specific DNA methyl transferases are part of host restriction-modification system which
18 methylate host DNA at specific sites. This protects the host DNA from cleavage by its own restriction
19 endonucleases which on the other hand destroy the incoming foreign phage DNA. These enzymes
20 were also reported to have functions like controlling the expression of host virulence genes and
21 selective silencing of genes that they methylate (Low et al. 2001, Oakey et al. 2002). In bacteriophage
22 lambda, methylation by DNA adenine methyltransferase is associated with packaging of the phage
23 genome (Sternberg and Coulby 1990), whereas, DNA (Cytosine-5) - methyltransferases are found

1 rarely in bacteriophage genomes and their functions in the context of viral lifecycle are unknown.
2 Interestingly, a gene coding for the protein HNH endonuclease (gene 52) having homology to 5-
3 methylcytosine specific restriction endonuclease McrA family (COG1403) with two highly conserved
4 histidine and one asparagine residues was identified in AS1 genome nearby the lysozyme gene within
5 the lysis module. McrA is a member of the superfamily HNHc (cI00083) which includes many
6 homing endonucleases, bacterial colicins, pyocins etc. and are rarely reported in Podoviruses.
7 Modified cytosine restriction (Mcr) systems capable of restricting phage λ modified by sequence
8 specific cytosine methylases has been earlier reported in *Escherichia coli* strain K12 (Raleigh and
9 Wilson 1986, Raleigh et al. 1989). The protein Gp74 from lambda-like phage HK97 was reported to
10 possess HNH endonuclease activity and mediate the cleavage of phage DNA (Moodley et al. 2012).
11 Later, Kala et al. (2014) discovered that these endonucleases were associated with the DNA
12 packaging terminase proteins in HK97 and majority of the large terminase subunits linked with HNH
13 endonucleases belonged to the Terminase_1 family (PF03354). The position of AS1 HNH
14 endonuclease, adjacent to terminase and other morphogenetic genes, suggests its potential
15 involvement in DNA packaging.

16 Genome of phage AS1 encodes pyocin activator protein (gene 36) with a conserved domain
17 belonging to the PrtN family (PF11112) and is involved in the transcriptional activation of the
18 polypeptide endonuclease toxin, pyocin. Other than lysozyme and HNH endonuclease, the lysis
19 module encodes a Cro/CI family transcriptional regulator containing a HTH motif and a phage anti-
20 termination Q type I family protein. The decision of the phage to enter either lytic or lysogenic cycle
21 is determined by Cro/C1 regulatory system which is well characterized and studied in phage λ
22 infecting the bacterium *Escherichia coli* (Ptashne 1967, Eisen et al. 1970, Schubert et al. 2007). The
23 antitermination Q Type I protein positively regulates the phage early and late genes by modifying the

1 host RNA polymerase and making it to proceed transcription past the terminator sequences. A total of
2 nine trans-membrane helices were predicted in seven proteins. 53 promoter sequences and 20 rho-
3 independent terminator sequences were identified in the assembled genome.

4 Phylogenetic analysis revealed a mosaic pattern of inheritance of various proteins encoded by
5 phage AS1 (Fig. 5). Major capsid protein, terminase large subunit and portal protein of AS1 had a
6 Siphoviral lineage, whereas, integrase and major tail proteins were closely related to other Podoviral
7 homologs. Recombination driven exchange of genetic material and the resulting genetic mosaicism is
8 wide-spread among *Caudovirales*. Even though prominent morphological differences exist between
9 the three families (specifically tail morphology), the differentiation is not clear cut due to striking
10 sequence level similarities among several members of these different families. A well-known
11 example for this is the genetic relatedness between phages lambda and P22 which are, however,
12 classified under *Siphoviridae* and *Podoviridae*, respectively, based on their tail morphology. No
13 lineage for the tail proteins of AS1 were deduced due to the limited availability of homologs.

14 **Conclusion**

15 Marinobacter phage AS1, isolated in this study, belongs to the family Podoviridae, infect
16 marine bacteria belonging to genus *Marinobacter*, strain D1S9. The presence of integrase gene,
17 occurrence of DNA methyltransferases and the existence of a significant percentage of phage AS1
18 genes within various *Marinobacter* genomes are suggestive of their intense association with the host.
19 From the genome analysis, it is evident that phage AS1 relies greatly on its host's replication and
20 translation machinery. Its codon usage bias must be similar to that of the host, since the phage
21 genome does not encode any tRNA genes. Another striking feature of the genome is its mosaicism
22 evident from the inheritance pattern of some of the important proteins with high similarity to their
23 siphoviral counterparts. It also contains most of the morphogenetic, assembly, and lysis genes

1 involved in the lytic induction. AS1 genome also encodes a gene which positively regulates the
2 expression of the bacterial toxin, pyocin; an interaction which attributes competitive advantage to the
3 host. All the above arguments indicate the possible coevolutionary interactions between the host and
4 the virus. As the members of genus *Marinobacter* are organisms with both phenotypic and metabolic
5 versatility and reported to have many significant ecological functions like their contribution in
6 biogeochemical cycles and marine snow formation, extensive studies on the involvement of these
7 viruses in impacting host metabolism as well as recruitment of this virus to metagenome need to be
8 carried out. The present study is a detailed analysis of the genomic properties of *Marinobacter* phage
9 AS1, guiding future research on intricate virus-host interactions and the role of phages in the
10 ecological functioning of the genus *Marinobacter*.

11 **References**

- 12 1. Suttle CA (2007) Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol*
13 5:801-812.
- 14 2. Sime-Ngando T (2014) Environmental bacteriophages: viruses of microbes in aquatic
15 ecosystems. *Front Microbiol* 5:355.
- 16 3. Middelboe M, Lyck PG (2002) Regeneration of dissolved organic matter by viral lysis in
17 marine microbial communities. *Aquat Microb Ecol* 27:187-194.
- 18 4. Bouvier T, Del Giorgio PA (2007) Key role of selective viral-induced mortality in
19 determining marine bacterial community composition. *Environ microbiol* 9:287-297.
- 20 5. McDaniel LD, Young E, Delaney J, Ruhnau F, Ritchie KB, Paul JH (2010) High frequency of
21 horizontal gene transfer in the oceans. *Science* 330:50.
- 22 6. Winget DM, Helton RR, Williamson KE, Bench SR, Williamson SJ, Wommack KE (2011)
23 Repeating patterns of virioplankton production within an estuarine ecosystem. *Proc Natl Acad*
24 *Sci USA* 108:11506-11511.
- 25 7. Stern A, Sorek R (2011) The phage-host arms race: shaping the evolution of microbes.
26 *Bioessays* 33:43-51.
- 27 8. Xia G, Wolz C (2014) Phages of *Staphylococcus aureus* and their impact on host evolution.
28 *Infection, Genet Evol* 21:593-601.

- 1 9. Buckling A, Brockhurst M (2012) Bacteria–virus coevolution. In: Soyer O (ed) Evolutionary
2 systems biology, Springer, New York, NY pp. 347-370.
- 3 10. Hall AR, Scanlan PD, Morgan AD, Buckling A (2011) Host–parasite coevolutionary arms
4 races give way to fluctuating selection. *Ecol Lett.* 2011 Jul;14(7):635-42.
- 5 11. Gokhale CS, Papkou A, Traulsen A, Schulenburg H (2013) Lotka–Volterra dynamics kills the
6 Red Queen: population size fluctuations and associated stochasticity dramatically change
7 host-parasite coevolution. *BMC Evol Biol* 13:254.
- 8 12. Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF (1999) Evolutionary relationships
9 among diverse bacteriophages and prophages: all the world’s a phage. *Proc Natl Acad Sci U S*
10 *A* 96:2192-2197.
- 11 13. Hatfull GF, Hendrix RW (2011) Bacteriophages and their genomes. *Curr Opin Virol* 1:298-
12 303.
- 13 14. Gauthier MJ, Lafay B, Christen R, Fernandez L, Acquaviva M, Bonin P, Bertrand JC (1992)
14 *Marinobacter hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely halotolerant,
15 hydrocarbon-degrading marine bacterium. *Int J Syst Evol Microbiol* 42:568-576.
- 16 15. Martin S, Márquez MC, Sánchez-Porro C, Mellado E, Arahal DR, Ventosa A (2003)
17 *Marinobacter lipolyticus* sp. nov., a novel moderate halophile with lipolytic activity. *Int J Syst*
18 *Evol Microbiol* 53:1383-1387.
- 19 16. Yoon JH, Lee MH, Kang SJ, Oh TK (2007) *Marinobacter salicampi* sp. nov., isolated from a
20 marine solar saltern in Korea. *Int J Syst Evol Microbiol* 57:2102-2105.
- 21 17. Zhang DC, Li HR, Xin YH, Chi ZM, Zhou PJ, Yu Y (2008) *Marinobacter psychrophilus* sp.
22 nov., a psychrophilic bacterium isolated from the Arctic. *Int J Syst Evol Microbiol* 58:1463-
23 1466.
- 24 18. Xu XW, Wu YH, Wang CS, Yang JY, Oren A, Wu M (2008) *Marinobacter pelagius* sp. nov.,
25 a moderately halophilic bacterium. *Int J Syst Evol Microbiol* 58:637-640.
- 26 19. Wang CY, Ng CC, Tzeng WS, Shyu YT (2009) *Marinobacter szutsaonensis* sp. nov., isolated
27 from a solar saltern. *Int J Syst Evol Microbiol* 59:2605-2609.
- 28 20. Wang H, Li H, Shao Z, Liao S, Johnstone L, Rensing C, Wang G (2012) Genome sequence of
29 deep-sea manganese-oxidizing bacterium *Marinobacter manganoxydans* MnI7-9. *J Bacteriol*
30 194:899-900.

- 1 21. Chua MJ, Campen RL, Wahl L, Grzymiski JJ, Mikucki JA (2018) Genomic and physiological
2 characterization and description of *Marinobacter gelidimuriae* sp. nov., a psychrophilic,
3 moderate halophile from Blood Falls, an antarctic subglacial brine. *FEMS Microbiol Ecol*
4 94:fiy021.
- 5 22. Green DH, Bowman JP, Smith EA, Gutierrez T, Bolch CJ (2006) *Marinobacter algicola* sp.
6 nov., isolated from laboratory cultures of paralytic shellfish toxin-producing dinoflagellates.
7 *Int J Syst Evol Microbiol* 56:523-527.
- 8 23. Lee OO, Lai PY, Wu HX, Zhou XJ, Miao L, Wang H, Qian PY (2012) *Marinobacter*
9 *xestospongiae* sp. nov., isolated from the marine sponge *Xestospongia testudinaria* collected
10 from the Red Sea. *Int J Syst Evol Microbiol* 62:1980-1985.
- 11 24. Huu NB, Denner EB, Ha DT, Wanner G, Stan-Lotter H (1999) *Marinobacter aquaeolei* sp.
12 nov., a halophilic bacterium isolated from a Vietnamese oil-producing well. *Int J Syst Evol*
13 *Microbiol* 49:367-375.
- 14 25. Shieh WY, Jean WD, Lin YT, Tseng M (2003) *Marinobacter lutaoensis* sp. nov., a
15 thermotolerant marine bacterium isolated from a coastal hot spring in Lutao, Taiwan. *Can J*
16 *Microbiol* 49:244-252.
- 17 26. Kaepfel EC, Gärdes A, Seebah S, Grossart HP, Ullrich MS (2012) *Marinobacter adhaerens*
18 sp. nov., isolated from marine aggregates formed with the diatom *Thalassiosira weissflogii*.
19 *Int J Syst Evol Microbiol* 62:124-128.
- 20 27. Lupette J, Lami R, Krasovec M, Grimsley N, Moreau H, Piganeau G, Sanchez-Ferandin S
21 ((2016) *Marinobacter* dominates the bacterial community of the *Ostreococcus tauri*
22 phycosphere in culture. *Front Microbiol* 7:1414.
- 23 28. Sandhya SV, Preetha K, Vijayan KK (2017) Phylogenetic diversity of culturable bacteria in
24 *Chaetoceros gracilis* mass culture system of a marine finfish hatchery. *J Mar Biol Ass India*
25 59:12-18.
- 26 29. Handley KM, Lloyd JR (2013) Biogeochemical implications of the ubiquitous colonization of
27 marine habitats and redox gradients by *Marinobacter* species. *Front Microbiol* 4:136.
- 28 30. Kaye JZ, Sylvan JB, Edwards KJ, Baross JA (2011) *Halomonas* and *Marinobacter* ecotypes
29 from hydrothermal vent, seafloor and deep-sea environments. *FEMS Microbiol Ecol*
30 75:123-133.

- 1 31. Bonis BM, Gralnick JA (2015) *Marinobacter subterrani*, a genetically tractable neutrophilic
2 Fe (II)-oxidizing strain isolated from the Soudan Iron Mine. *Front Microbiol* 6:719.
- 3 32. Singer E, Webb EA, Nelson WC, Heidelberg JF, Ivanova N, Pati A, Edwards KJ (2011) The
4 genomic potential of *Marinobacter aquaeolei*– A biogeochemical opportunitroph. *Appl*
5 *Environ Microbiol* 77:2763-2771.
- 6 33. Zhu M, Wang M, Jiang Y, You S, Zhao G, Liu Y, Yang Q, Liu Q, Liu Z, Gong Z, Shao H
7 (2018) Isolation and Complete Genome Sequence of a Novel *Marinobacter* Phage B23. *Curr*
8 *microbiol* 75:1619-1625.
- 9 34. Middelboe M, Chan A, Bertelsen SK (2010) Isolation and life cycle characterization of lytic
10 viruses infecting heterotrophic bacteria and cyanobacteria. In: Wilhelm SW, Weinbauer MG,
11 Suttle CA (ed) *Manual of aquatic viral ecology*, American Society of Limnology and
12 Oceanography, pp. 118-133.
- 13 35. Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection
14 of individual microbial cells without cultivation. *Microbiol Mol Biol Rev* 59:143-169.
- 15 36. Pradeep Ram AS, Arnous B, Danger M, Carrias JF, Lacroix G, Sime-Ngando T (2010) High
16 and differential viral infection rates within bacterial ‘morphopopulations’ in a shallow sand pit
17 lake (Lac de Créteil, France). *FEMS Microbiol Ecol* 74:83-92.
- 18 37. Lawrence JE, Steward GF (2010) Purification of viruses by centrifugation. In: Wilhelm SW,
19 Weinbauer MG, Suttle CA (ed) *Manual of aquatic viral ecology*, American Society of
20 Limnology and Oceanography, pp. 166-181.
- 21 38. Yang Y, Cai L, Ma R, Xu Y, Tong Y, Huang Y, Jiao N, Zhang R (2017) A novel
22 roseosiphophage isolated from the oligotrophic South China Sea. *Viruses* 9:109.
- 23 39. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data.
- 24 40. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM,
25 Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV (2012) SPAdes: a new genome assembly
26 algorithm and its applications to single-cell sequencing. *J Comput biol* 19:455-477.
- 27 41. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S,
28 Glass EM, Kubal M, Meyer F (2008) The RAST Server: rapid annotations using subsystems
29 technology. *BMC genomics* 9:75.

- 1 42. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer
2 RC, Gonzales NR, Gwadz M (2016) CDD/SPARCLE: functional classification of proteins via
3 subfamily domain architectures. *Nucleic Acids Res* 45:D200-D203.
- 4 43. Lowe TM, Chan PP (2016) tRNAscan-SE On-line: integrating search and context for analysis
5 of transfer RNA genes. *Nucleic Acids Res* 44:W54-W57.
- 6 44. Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D (2011) ARNold: a web tool for
7 the prediction of Rho-independent transcription terminators. *RNA Biol* 8:11-13.
- 8 45. Reese MG (2001) Application of a time-delay neural network to promoter annotation in the
9 *Drosophila melanogaster* genome. *Comput Chem* 26:51-56.
- 10 46. Kumar S, Stecher G, Tamura K ((2016) MEGA7: molecular evolutionary genetics analysis
11 version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870-1874.
- 12 47. Bradley DE (1967) Ultrastructure of bacteriophage and bacteriocins. *Bacteriol Rev* 31:230.
- 13 48. Oh B, Moyer CL, Hendrix RW, Duda RL (2014) The delta domain of the HK97 major capsid
14 protein is essential for assembly. *Virology* 456:171-178.
- 15 49. Cardarelli L, Lam R, Tuite A, Baker LA, Sadowski PD, Radford DR, Rubinstein JL, Battaile
16 KP, Chirgadze N, Maxwell KL, Davidson AR (2010) The crystal structure of bacteriophage
17 HK97 gp6: defining a large family of head–tail connector proteins. *J Mol Biol* 395:754-768.
- 18 50. Bailly-Bechet M, Vergassola M, Rocha E (2007) Causes for the intriguing presence of tRNAs
19 in phages. *Genome Res* 17:1486-1495.
- 20 51. Murphy J, Klumpp J, Mahony J, O’Connell-Motherway M, Nauta A, van Sinderen D (2014)
21 Methyltransferases acquired by lactococcal 936-type phage provide protection against
22 restriction endonuclease activity. *BMC genomics* 15:831.
- 23 52. Low DA, Weyand NJ, Mahan MJ (2001) Roles of DNA adenine methylation in regulating
24 bacterial gene expression and virulence. *Infect Immun* 69:7197-7204.
- 25 53. Oakey HJ, Cullen BR, Owens L (2002) The complete nucleotide sequence of the *Vibrio*
26 *harveyi* bacteriophage VHML. *J Appl Microbiol* 93:1089-1098.
- 27 54. Sternberg N, Coulby J (1990) Cleavage of the bacteriophage P1 packaging site (pac) is
28 regulated by adenine methylation. *Proc Natl Acad Sci U S A* 87:8070-8074.
- 29 55. Raleigh EA, Wilson G (1986) *Escherichia coli* K-12 restricts DNA containing 5-
30 methylcytosine. *Proc Natl Acad Sci U S A* 83:9070-9074.

- 1 56. Raleigh EA, Trimarchi R, Revel H (1989) Genetic and physical mapping of the *mcrA* (*rglA*)
2 and *mcrB* (*rglB*) loci of *Escherichia coli* K-12. *Genetics* 122:279-296.
- 3 57. Moodley S, Maxwell KL, Kanelis V (2012) The protein gp74 from the bacteriophage HK97
4 functions as a HNH endonuclease. *Protein Sci* 21:809-818.
- 5 58. Kala S, Cumby N, Sadowski PD, Hyder BZ, Kanelis V, Davidson AR, Maxwell KL (2014)
6 HNH proteins are a widespread component of phage DNA packaging machines. *Proc Natl*
7 *Acad Sci U S A* 111:6022-6027.
- 8 59. Ptashne M (1967) Specific binding of the λ phage repressor to λ DNA. *Nature* 214:232.
- 9 60. Eisen H, Brachet P, Da Silva LP, Jacob F (1970) Regulation of repressor expression in λ . *Proc*
10 *Natl Acad Sci U S A* 66:855-862.
- 11 61. Schubert RA, Dodd IB, Egan JB, Shearwin KE (2007) Cro's role in the CI-Cro bistable
12 switch is critical for λ 's transition from lysogeny to lytic development. *Genes Dev* 21:2461-
13 2472.

Table 1. Genome features of *Marinobacter* phage AS1

Feature	<i>Marinobacter</i> phage AS1
Genome size	36,994 bp
GC content	57 %
Total no. of proteins	52
No. of proteins with putative function	21
tRNA genes	None
Proteins with trans-membrane helices	7

Table 2. Genome annotation of Marinobacter phage AS1

Gene	Start (bp)	Stop (bp)	Length (bp)	Strand	Putative function	E-value
1	70	531	462	+	Phage terminase small subunit P27 family	3.2E-111
2	531	2216	1686	+	Terminase large subunit	0.0
3	2210	2404	195	+	Gp3	2.8E-19
4	2404	3669	1266	+	Putative portal protein	0.0
5	3657	4589	933	+	Serine peptidase	0.0
6	4663	5946	1284	+	Major capsid head protein	0.0
7	6000	6224	225	+	Hypothetical protein	1.8E-48
8	6262	6795	534	+	Gp6	6.4E-116
9	6797	7348	552	+	Hypothetical protein	4.5E-128
10	7348	7902	555	+	Hypothetical protein	5.6E-127
11	7899	8348	450	+	Hypothetical protein	6.2E-98
12	8352	9107	756	+	Major tail protein	8.3E-8
13	9172	9321	150	+	Hypothetical protein	-
14	9385	9600	216	+	Hypothetical protein	1.3E-30
15	9665	12883	3219	+	Tail tape measure protein	2.0E-28
16	12880	13278	399	+	Hypothetical protein	8.3E-88
17	13275	14765	1491	+	Tail fiber	3.1E-81

18	14779	16893	2115	+	Tail fiber	1.4E-125
19	16893	18011	1119	+	Hypothetical protein	1.1E-82
20	18016	19926	1911	+	Hypothetical protein	0.0
21	19913	20146	234	+	Hypothetical protein	5.2E-56
22	20955	20143	813	-	Putative lipoprotein	9.2E-76
23	21434	20979	456	-	Hypothetical protein	2.4E-96
24	21810	22895	1086	+	Integrase	0
25	23007	23207	201	+	Hypothetical protein	4.0E-07
26	24260	23265	996	-	DNA methyl transferase	9.7E-131
27	25730	24270	1461	-	Hypothetical protein	-

Table 2. Genome annotation of Marinobacter phage AS1 (continued)

28	26213	25878	336	-	Hypothetical protein	2.8E-46
29	26407	26210	198	-	Hypothetical protein	1.3E-43
30	26661	26404	258	-	Hypothetical protein	6.5E-57
31	28046	26658	1389	-	DNA (Cytosine-5)-methyl transferase 1	0.0
32	28417	28043	375	-	Hypothetical protein	2.7E-27
33	28711	28421	291	-	Hypothetical protein	2.2E-37
34	29094	28720	375	-	Hypothetical protein	3.3E-811
35	29299	29105	195	-	Hypothetical protein	3.5E-48
36	29617	29354	264	-	Pyocin activator protein	3.0E-49
37	29892	29614	279	-	Protein containing phage anti repressor protein domain	3.2E-15
38	30317	29889	429	-	Hypothetical protein	5.6E-81
39	30981	30502	480	-	Hypothetical protein	1.5E-69
40	31316	30978	339	-	Hypothetical protein	6.5E-53
41	32428	31313	1116	-	Hypothetical protein	5.2E-82
42	32880	32473	408	-	Hypothetical protein	5.1E-61
43	32981	33217	237	+	Cro/cl family transcriptional regulator	2.0E-41
44	33282	34235	954	+	Hypothetical protein	0.0
45	34285	34629	345	+	Hypothetical protein	1.3E-70

46	34619	34969	351	+	Phage anti termination Q type I family- like protein	1.5E-15
47	35096	35527	432	+	Lysozyme	3.6E-91
48	35524	35802	279	+	Hypothetical protein	5.5E-53
49	35786	36016	231	+	Hypothetical protein	6.0E-57
50	36013	36462	450	+	Hypothetical protein	1.6E-43
51	36464	36592	129	+	Hypothetical protein	2.8E-13
52	36592	36975	384	+	HNH endonuclease	2.3E-87

Table 3. Functional categorization of phage AS1 proteins

Function	No. of proteins
Phage structure and assembly	13
Phage defense/DNA modification	2
Transcriptional regulation	4
Host lysis	1
Phage integration	1
Hypothetical proteins	31

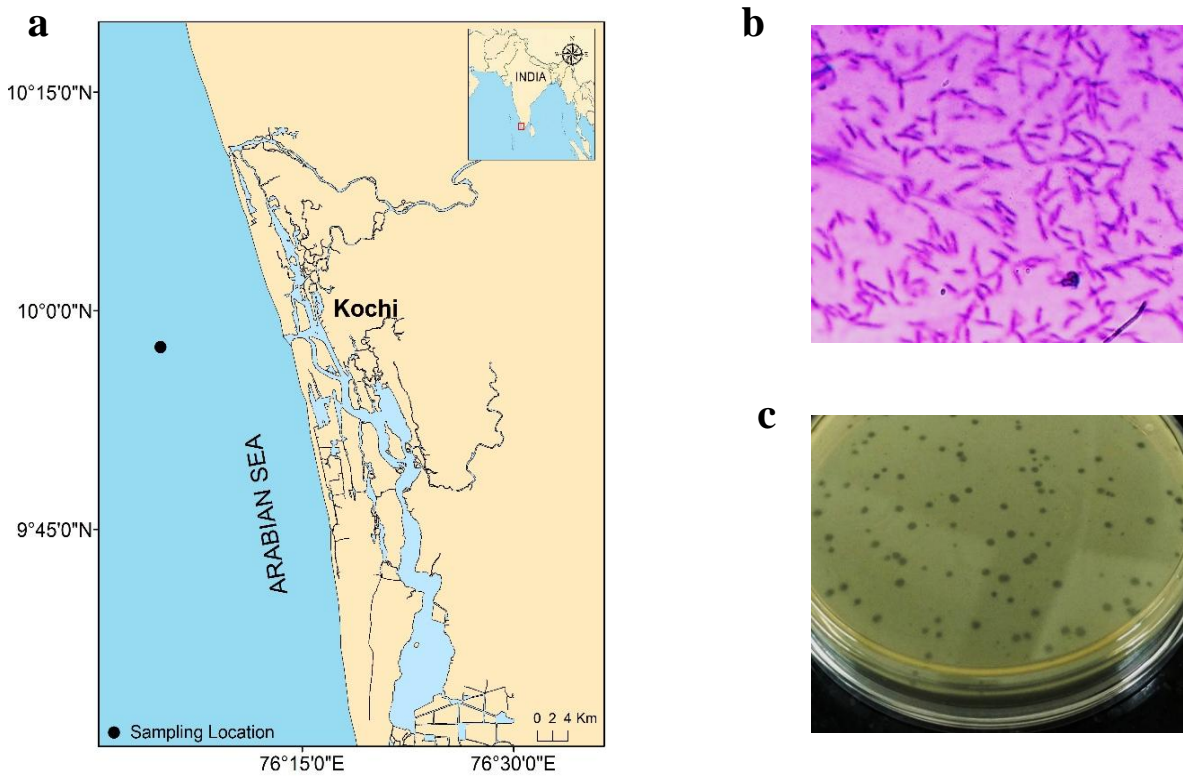


Fig. 1 Isolation of phage AS1. (a) Sampling location in the Arabian Sea, (b) rod shaped *Marinobacter* host cells after staining with crystal violet, (c) clear round plaques of phage AS1 formed on a lawn of host

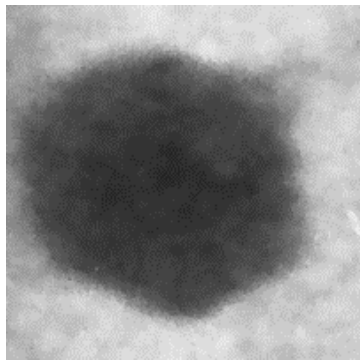


Fig. 2 Morphology of the isolated phage. Electron micrograph of phage AS1. Scale bar equals 50 nm.

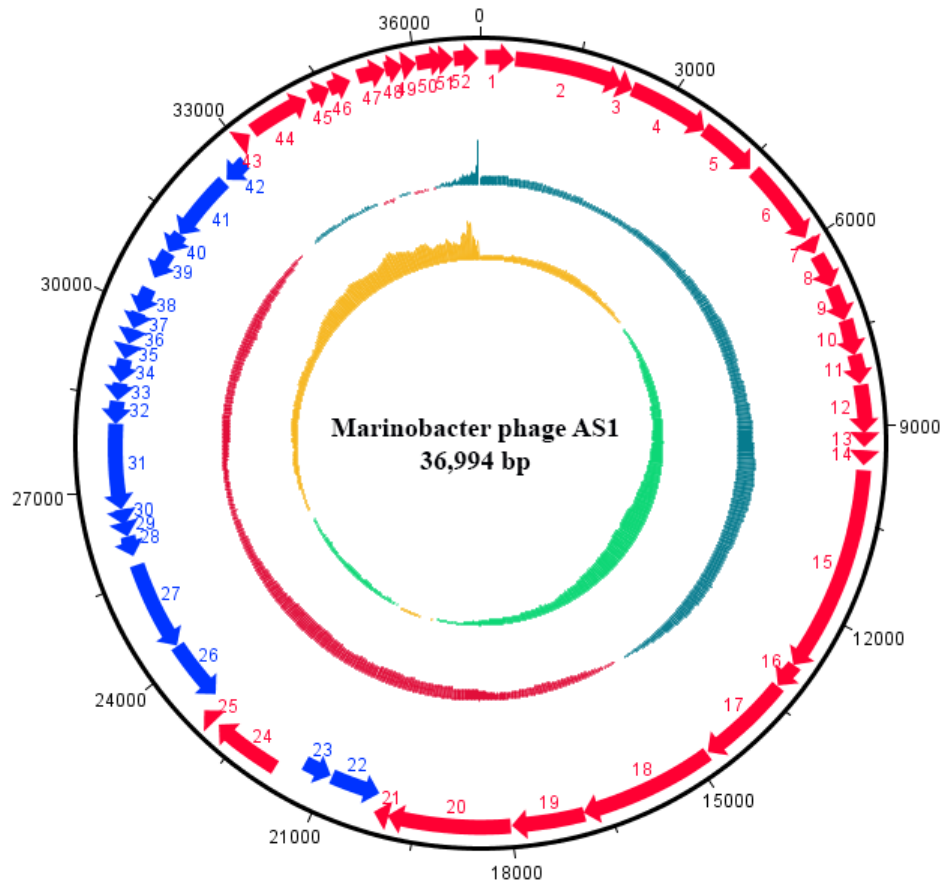


Fig. 3 Circular representation of the double stranded linear genome of Marinobacter phage AS1 featuring (from outside to inside) coding DNA sequences with predicted functions in the forward strand (red), reverse strand (neon blue), GC content (maroon & blue) and GC skew (green & yellow).

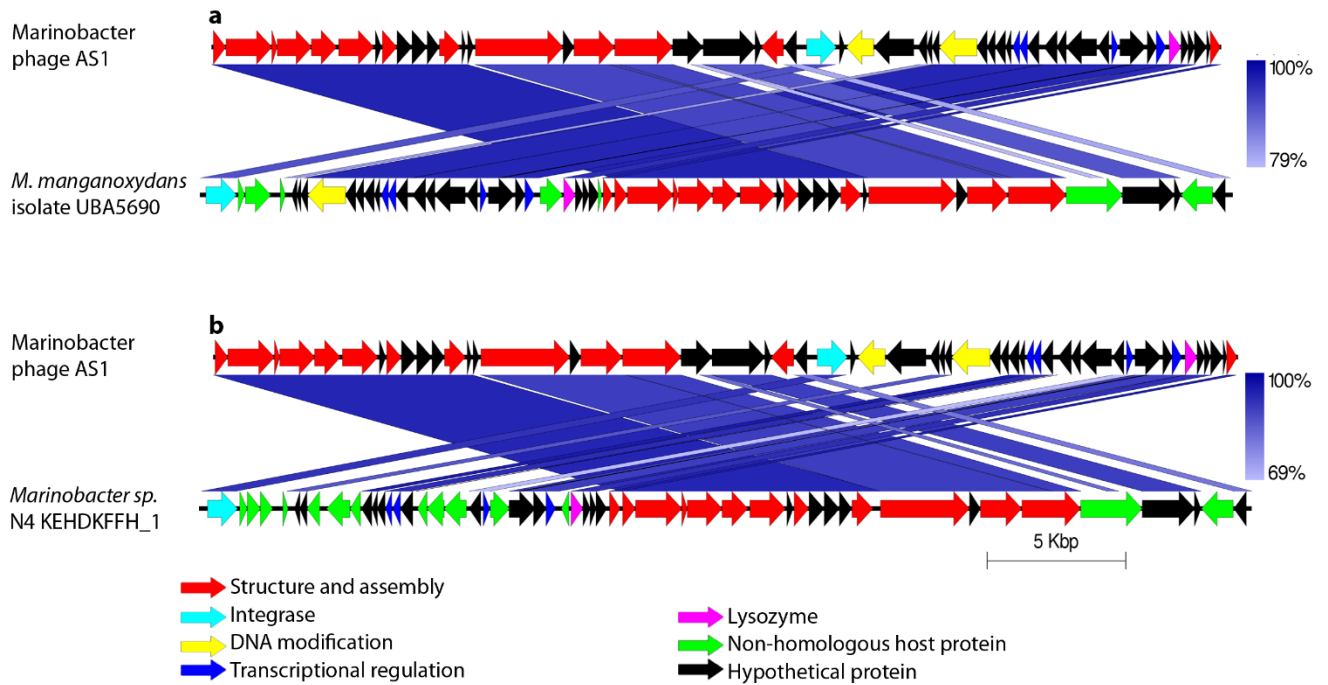


Fig. 4 Comparison of similar sequences found within the sequences of host genus. Comparison with (a) *M. manganoxydans* isolate UBA5690 and (b) *Marinobacter sp.* N4 KEHDKFFH_1. The intensity of blue colour indicates the percentage of similarity between the two sequences.

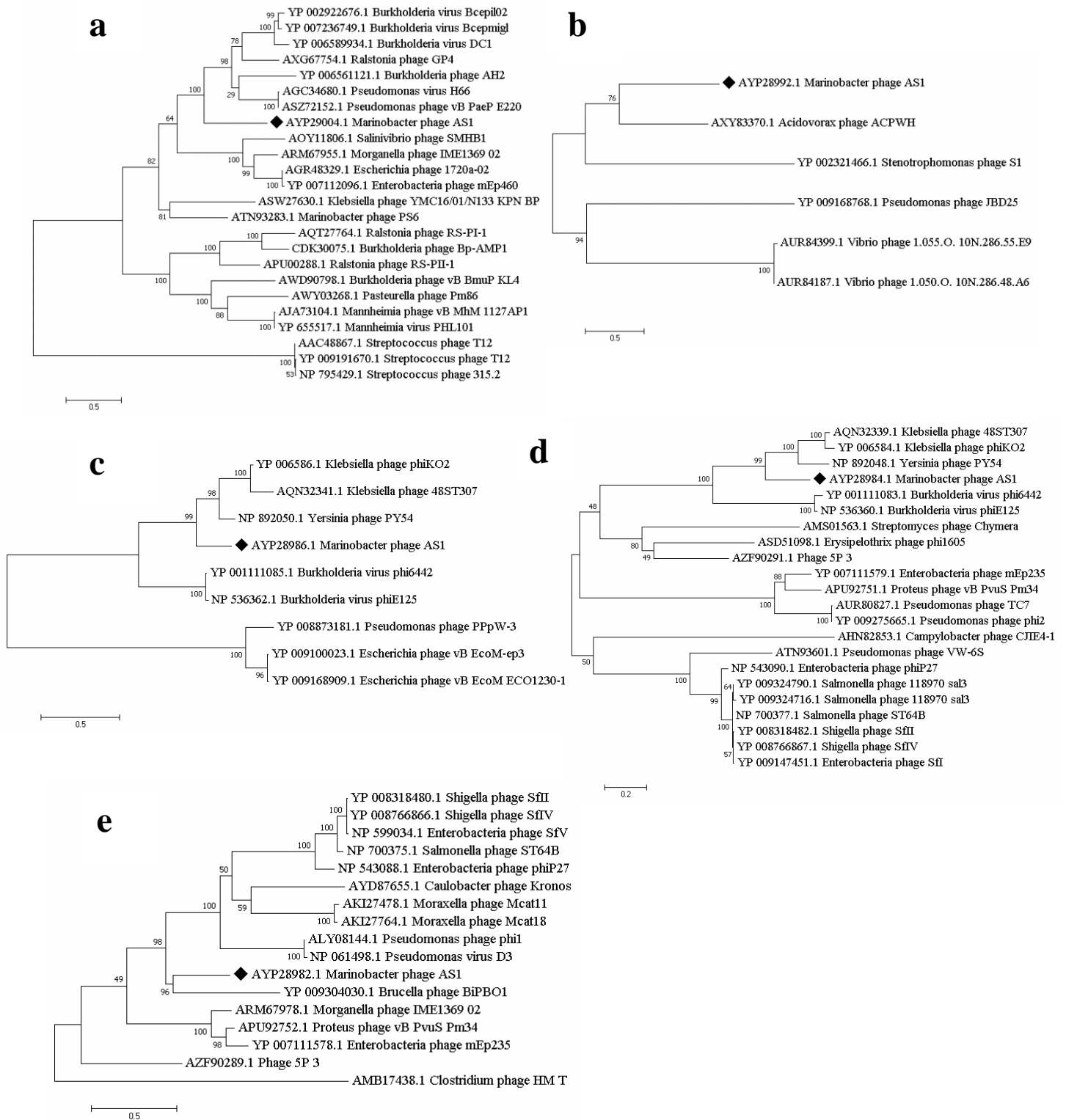


Fig. 5 Phylogenetic trees of AS1 proteins. (a) Integrase, (b) major tail protein, (c) major capsid protein, (d) portal protein and (e) terminase, large subunit.