Vol. 76, No. 2

Searching for a "Hidden" Prophage in a Marine Bacterium[▽]

Yanlin Zhao,[1,2] Kui Wang,[2] Hans-Wolfgang Ackermann,[3] Rolf U. Halden,[4,5]
Nianzhi Jiao,[1]* and Feng Chen[2]*

*State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen 361005, China[1]; Center of Marine Biotechnology,
University of Maryland Biotechnology Institute, Baltimore, Maryland 21202[2]; Department of Microbiology, Infectiology and
Immunology, Faculty of Medicine, Laval University, Quebec, QC G1K 7P4, Canada[3]; Department of
Environmental Health Sciences, Center for Water and Health, Johns Hopkins Bloomberg School of
Public Health, Baltimore, Maryland 21205-2179[4]; and Center for Environmental Biotechnology,
the Biodesign Institute at Arizona State University, Tempe, Arizona 85287-5701[5]*

Prophages are common in many bacterial genomes. Distinguishing putatively viable prophages from non-viable sequences can be a challenge, since some prophages are remnants of once-functional prophages that have been rendered inactive by mutational changes. In some cases, a putative prophage may be missed due to the lack of recognizable prophage loci. The genome of a marine roseobacter, *Roseovarius nubinhibens* ISM (hereinafter referred to as ISM), was recently sequenced and was reported to contain no intact prophage based on customary bioinformatic analysis. However, prophage induction experiments performed with this organism led to a different conclusion. In the laboratory, virus-like particles in the ISM culture increased more than 3 orders of magnitude following induction with mitomycin C. After careful examination of the ISM genome sequence, a putative prophage (ISM-pro1) was identified. Although this prophage contains only minimal phage-like genes, we demonstrated that this "hidden" prophage is inducible. Genomic analysis and reannotation showed that most of the ISM-pro1 open reading frames (ORFs) display the highest sequence similarity with *Rhodobacterales* bacterial genes and some ORFs are only distantly related to genes of other known phages or prophages. Comparative genomic analyses indicated that ISM-pro1-like prophages or prophage remnants are also present in other *Rhodobacterales* genomes. In addition, the lysis of ISM by this previously unrecognized prophage appeared to increase the production of gene transfer agents (GTAs). Our study suggests that a combination of *in silico* genomic analyses and experimental laboratory work is needed to fully understand the lysogenic features of a given bacterium.

Lysogeny is a symbiotic relationship between bacteria and prophages (1). DNA of these prophages is integrated into the host bacterial chromosomes and can be induced to lyse the host cells and become free phage particles under certain circumstances. Prophages can be viewed from a bacterial perspective as "dangerous molecular time bombs" (27), and they also play important roles in biological properties and evolutionary changes of their bacterial hosts (5, 7–10, 12, 24, 27). Ackermann and DuBow (1) reported about 47% of isolated terrestrial bacteria are lysogenic. Similarly, lysogeny is also common in bacteria living in the seas. A wide range (from 28 to 71%) of marine bacterial isolates have been reported to contain prophages inducible by mitomycin C or UV irradiation (17, 19, 23, 32), and lysogeny is also widely present in marine bacterial communities (17, 18, 33–36). After a genomic search of 113 different marine bacteria, 43% were reported to contain prophage-like elements in their genome sequences (27).

Prophages are known to play an important role in genomic diversification of bacteria via horizontal gene transfer. In order to understand the evolution of bacterial genomes, it is essential to search for the presence and location of prophage(s) when an annotated bacterial genome sequence becomes available. Many bioinformatic methods and criteria have been developed to detect the prophage in bacterial genomes, namely, a BLAST-based protein similarity approach (PSA) method (28, 30), the dinucleotide relative abundance (DRA) method (31), and assessment of GC content, codon usage pattern, and amino acid usage (11). When screening a bacterial genome for prophage(s), it is a common practice to manually examine the bacterial genome sequence and search for the clusters of phage-like genes (11, 26, 27). Alternatively, Prophage Finder (4) and Phage_Finder (http://phage-finder.sourceforge.net) (15), based on BLAST sequence comparison, can be used for quick identification of prophages. It is known that a large fraction of the prophages identified solely by genomic analysis appear to be defective and inactive (8, 11). Defective prophages may lack the ability to lyse the cell, but certain functional genes in these defective prophages could be beneficial to the evolution and ecological fitness of their hosts.

However, identifying a functional prophage in a bacterial genome sequence is challenging for the following reasons: (i) bacterial genomes may not be fully annotated, and some phage-like open reading frames (ORFs) may therefore be missed; (ii) only a few phage-like genes may be found within a short sequence region; (iii) phage-like genes are spread over a reasonably compact region, but the whole element represents

* Corresponding author. Mailing address for Feng Chen: Center of Marine Biotechnology, University of Maryland Biotechnology Institute, 701 E. Pratt St., Baltimore, MD 21202. Phone: (410) 234-8866. Fax: (410) 234-8898. E-mail: chenf@umbi.umd.edu. Mailing address for Nianzhi Jiao: State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen 361005, China. Phone and fax: 86-592-2187869. E-mail: jiao@xmu.edu.cn.

only the remnant of a once-functional prophage that has been subject to mutational decay and thus is no longer inducible; (iv) owing to their great diversity, yet-undiscovered prophages may be hidden within bacterial genomes currently considered fully annotated. We reported here the finding of a "hidden" inducible prophage from a marine bacterium, *Roseovarius nubinhibens* ISM, isolated from the surface water of the Caribbean Sea (16). The genome of ISM was sequenced and deposited at the Roseobase (http://www.roseobase.org/roseo /nubinhibens.html) and NCBI (http://www.ncbi.nlm.nih.gov) databases. Our study shows that a putative prophage can be easily missed with *in silico* analysis. Even with manual analysis, no intact prophage was found in ISM genome (27). How did we find the ISM prophage?

MATERIALS AND METHODS

Bacterial strain and medium. *Roseovarius nubinhibens* ISM was kindly provided by Mary Ann Moran at the University of Georgia. ISM was grown in YTSS medium (4 g liter^{-1} yeast extract, 2.5 g liter^{-1} tryptone, 20 g liter^{-1} Crystal Sea salt) at 28°C and was shaken at 200 rpm during induction.

Mitomycin C induction and virus-like particle (VLP) counts with SYBR gold. Exponentially growing culture of ISM (optical density at 600 nm [OD$_{600}$] = 0.2) was split equally into two flasks. One was induced by the addition of mitomycin C (final concentration, 0.5 µg ml^{-1}) (Sigma Chemical Co., St. Louis, MO), and the other served as a control. After 24 h, the quantity of VLPs was enumerated by epifluorescence microscopy, following the method described by Chen et al. (13). Briefly, 1 to 10 µl of each sample was suspended in 1 ml TE buffer (Tris 10 mM, EDTA 1 mM, pH 7.5) and then filtered through a 0.02-µm-pore-size 25-mm Anodisc membrane filter (Whatman, Clifton, NJ) under vacuum pressure at *ca*. 10 mm Hg. The cells were stained with 1× SYBR gold (Molecular Probes, Eugene, OR) for 20 min in the dark. VLPs were counted under blue excitation (485 nm) at magnification ×1,000 using a Zeiss Axioplan epifluorescence microscope. At least a total of 300 VLPs were counted from 10 random view fields. When a highly fluorescent background was encountered, lysates were treated with DNase and RNase at a final concentration of 50 µg ml^{-1} (each) for 30 min prior to SYBR gold staining.

Purification of induced VLPs. One liter of the ISM culture was induced by mitomycin C. VLPs were concentrated using polyethylene glycol 8000 (PEG 8000) as described previously (14), with modifications. The PEG-concentrated VLPs were layered on a 10 to 50% iodixanol (OptiPrep; Sigma-Aldrich, St. Louis, MO) gradient and centrifuged for 2 h at 100,000 × g using a Sorvall Discovery 100S centrifuge. The visible viral band (*ca*. 0.5 ml) was extracted using an 18-gauge needle syringe and then dialyzed twice in TM buffer (20 mM Tris [pH 7.5], 10 mM MgSO$_4$) overnight at 4°C. DNase and RNase were then used to remove nucleic acids outside the viral particles (final concentration, 10 µg ml^{-1} each). Purified VLPs were stored at 4°C in the dark for later analysis.

Transmission electron microscopy (TEM). For electron microscopy, VLPs were sedimented for 60 min at 25,000 × g in a Beckman (Palo Alto, CA) J2-21 high-speed centrifuge using a JA-18.1 fixed-angle rotor and were washed twice in 0.1 M ammonium acetate buffer (pH 7.0). Purified phage particles were deposited on carbon-coated copper grids for 1 min, stained with 2% potassium phosphotungstate (pH 7.0), and examined in a Philips EM 300 electron microscope operated at 60 kV. Magnification was monitored using T4 phage tails.

PFGE. Pulsed-field gel electrophoresis (PFGE) analysis, including sample digestion and gel preparation, was carried out according to the method described previously (37). PFGE was run using a Chef DR-III clamped homogeneous electric field system (Bio-Rad, Richmond, CA) with a 1% agarose gel (pulsed-field certified agarose; Bio-Rad), a 1- to 25-s pulse ramp, a voltage rate of 6.0 V cm^{-1} with an included angle of 120° at a constant temperature of 14°C, and a run time of 20 h. The gel was stained with 1× SYBR gold and visualized with a Typhoon 9410 imager (Amersham Biosciences, Piscataway, NJ).

Primer design and PCR amplification. One PCR primer set, ISM-150F (5′-GAG GCG GAT TTC AAA AAG GC-3′) and ISM-970R (5′-GGA AGT TCT TGG TGC TGG CA-3′), was designed based on the major capsid gene located in the ISM-pro1 region (accession no. ZP_00960168).

The primer sets MCP-109F/368R (39), SO-1 (5′-CCT GTC GCA CAA TGG CTT TG −3′)/SO-2 (5′-GTA ATA GGG GTC GGG GTT GAA G-3′), and 16S 27F (5′-AGAGTITGATCCTGGCTCAG-3′)/1492R (5′-TACGGCTACCTTGT TACGA CTT-3′) were based on conserved regions of the gene transfer agent (GTA) major capsid gene *g5*, *sob* gene, and 16S rRNA gene. These three primers were included to detect the presence of host genomic DNA.

PCR was performed in a 50-µl volume containing 1× reaction buffer (GenScript, Piscataway, NJ) with 1.5 mM MgCl$_2$, a 100 µM concentration of each deoxynucleoside triphosphate, 10 pmol of each primer, and 1 U *Taq* DNA polymerase (GenScript). PFGE bands a and b (see Fig. 2A) were excised, resuspended in 100 µl of PCR water, and stored at 4°C overnight. Two microliters of the supernatant was used as a PCR template. For purified VLPs, 1 µl was used as a PCR template. The PCR program for all reactions included an initial denaturing step at 94°C for 3 min, followed by 30 cycles of 94°C for 1 min, annealing at 50°C for 30 s, and 72°C for 1 min, followed by a final extension step at 72°C for 10 min.

Analysis of ISM induced VLP proteins. Purified VLP suspension was mixed with equal volumes of 2× Laemmli buffer (125 mM Tris-HCl [pH 6.8], 10% 2-mercaptoethanol, 4% sodium dodecyl sulfate [SDS], 20% glycerol, 0.004% bromphenol blue) and boiled for 5 min (20). Proteins were separated by 12% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) gel using a universal mutation detection system (Bio-Rad Laboratories, Hercules, CA). The gel was stained with Coomassie blue. Protein molecular mass standards (Bio-Rad) with molecular masses ranging from 10 to 250 kDa were used to determine the molecular mass of the protein. The major protein band was excised and analyzed via matrix-assisted laser desorption ionization–time-of-flight mass spectrometry (MALDI-TOF MS). The resulting peptide mass fingerprints were search against the GenBank database using a Web-based peptide mass fingerprints analysis tool, Mascot (http://www.matrixscience.com/search_form_select.html).

Sequence analysis. Putative prophage sequences in the ISM genome were identified using Prophage Finder (4). The ISM genome was manually analyzed for the presence of putative prophage. First, the genome was scanned for the phage-related genes. When a phage-related gene was found, the genes in the vicinity were also examined. Reannotation of putative prophage (ISM-pro1) ORFs was done by performing a BLASTP search against GenBank database (2). The beginning and end of a specific prophage genome were estimated based on the annotation of the surrounding genes.

tRNA was searched in the ISM-pro1 genome by using tRNAscan-SE (25). Putative ORFs from *Roseobacter* sp. CCS2, *Sulfitobacter* sp. NAS 14.1, and *Rhodobacter sphaeroides* 2.4.1 were analyzed using BLASTP.

RESULTS AND DISCUSSION

ISM produces inducible virus-like particles (VLPs). After addition of mitomycin C (0.5 µg ml^{-1}) to an exponentially growing culture of ISM, lysis of host cells and a dramatic increase in VLP abundance (from *ca*. 0.3×10^6 to 1.2×10^{10} ml^{-1}) was observed 24 h posttreatment. In contrast, VLPs in the control increased only to *ca*. 0.6×10^7 ml^{-1}, possibly due to the spontaneous induction. The induction efficiency was comparable to that of another roseobacter, *Silicibacter* sp. TM1040, in which VLPs increased from 10^6 to 10^9 ml^{-1} after mitomycin C induction (14). The induced VLPs were siphoviruses with long, flexible tails (of about 150 by 8 nm) and polyhedral heads (of about 55 nm in diameter) (Fig. 1A). The vast majority of VLPs consisted of empty heads (with no stainable materials) and broken tails (Fig. 1B), suggesting that the VLPs produced by strain ISM contain a high proportion of defective phages.

PFGE analysis showed that the induced VLPs were associated with a DNA fragment of *ca*. 30 kb which formed concatemers (Fig. 2A). Interestingly, a few smaller DNA fragments (*ca*. 3, 4, and 12 kb in size) were also detected on the PFGE gel (Fig. 2A). It is likely that these small bands resulted from the DNA fragments contained by gene transfer agent (GTA) particles (see the GTA section).

Identification of a putative prophage in the ISM genome. A search for potential prophages in the ISM genome using Prophage Finder yielded several putative prophages. However, none of these putative prophages was in the 30-kb size range as

FIG. 1. Virus-like particles generated by mitomycin C induction of *R. nubinhibens* ISM. Siphovirus particles were commonly seen in the induced lysate (A and B), and many of them appeared to be "empty" (B). Scale bars, 50 nm.
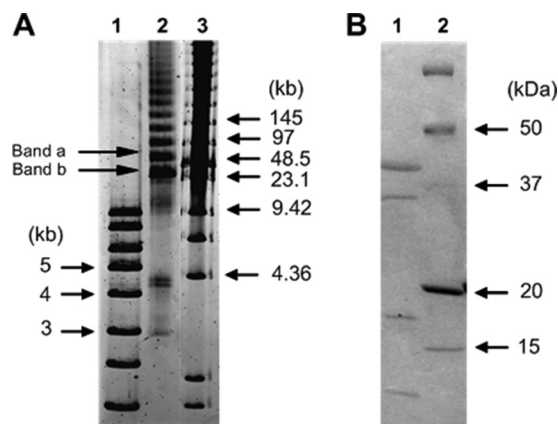

FIG. 2. (A) PFGE analysis of viral genomic DNAs isolated from induced ISM VLPs. Lane 1, 1-kb DNA ladder; lane 2, DNA obtained from the induced VLPs; lane 3, low-range PFGE marker. (B) SDS-PAGE analysis of total proteins from induced ISM VLPs. Lane 1, major proteins from induced ISM VLPs; lane 2, protein marker.

evidenced in the PFGE gel. We then carefully scanned the ISM genome to locate all phage-related genes. One 27-kb region that contains an integrase gene and some other phage-related genes drew our attention. Most of the genes in this region were first annotated as hypothetical proteins, and very few were assigned a function. We then applied BLASTP to reanalyze all the ORFs in this region, and more recognizable phage-like genes were found (Table 1 and Fig. 3). It should be noted that this putative prophage (ISM-pro1) contains many genes related to bacteria (Table 1), which makes this prophage less recognizable for Prophage Finder or other approaches. The ISM-pro1 genome likely extends from ISM_12685 (phage integrase) to ISM_12880 in the ISM genome.

In order to confirm that the induced VLPs matched the prophage detected in the ISM genome, a primer set targeting the major capsid gene (ORF 19) of ISM-pro1 was designed. PCR products of the expected size (*ca.* 820 bp) were amplified from templates prepared from either induced prophage lysates or gel-purified VLP genomic DNA bands (Fig. 2A) (data not shown). Sequences of these PCR amplicons were 100% identical to the major capsid gene of ISM-pro1. Three host-specific genes, *g5*, *sob*, and the 16S rRNA gene, could not be PCR amplified, indicating that both the induced VLPs and the gel-purified prophage genomic DNA were free of host DNA contamination.

In addition, the protein profile of induced VLPs was analyzed using an SDS-PAGE gel (Fig. 2B). Four major polypeptides were visible (Fig. 2B, lane 1), and the protein with a molecular mass of ~45 kDa was characterized by mass spectrometry. The mass fingerprint of this polypeptide matched the ISM-pro1 major capsid protein (45.3 kDa; accession number ZP_00960168). Both DNA and protein analyses confirmed that the inducible VLPs with a DNA content of *ca.* 30 kb are indeed produced from ISM-pro1.

Genome analysis of prophage ISM-pro1. Genomic analysis revealed that the genome size of ISM-pro1 is 26,858 bp with a G+C content of 60%, which is slightly lower than that of ISM (63%). In total, ISM-pro1 contains 40 predicted open reading frames (ORFs). The location of ISM-pro1 in the bacterial genome also considered the surrounding bacterial genes. Twenty-eight ORFs yield significant hits in the GenBank da-

tabase, and 25 ORFs are most closely related to the genes found in *Rhodobacterales* bacteria. Only 11 ISM-pro1 ORFs were assigned putative functions (Table 1), including phage integrase (ORF 1), terminase subunits (ORFs 15 and 16), HNH nuclease (ORF 14), lysozyme gene (ORF 35), and several phage-like structural genes (ORFs 17, 18, 19, 22, and 28) (Table 1 and Fig. 3). These ORFs are putatively involved in the main steps of the prophage's life cycle: integration, DNA packaging, morphogenesis, and bacterial lysis (Fig. 3). Since the majority of the ISM-pro1 ORFs have unknown functions, it is difficult to subdivide the genome into functional modules, and the repressors, transcriptional regulator, and DNA replication genes have not been identified in the ISM-pro1 genome. In addition, no tRNA gene was found in ISM-prol by using tRNAscan-SE.

ORF 1 encodes an integrase gene which is most similar to the integrase of *Paracoccus denitrificans* PD1222. ORF 14 encodes the HNH nuclease, which shares sequence similarity with nucleases encoded by other bacteria and siphoviruses. ORF 15 and ORF 16 encode the phage terminase small and large subunits, respectively, which are involved in phage DNA packaging. ORFs 17, 18, 19, 22, and 28 are identified as phage morphogenesis genes. Sequence analysis revealed that ISM-pro1 morphogenesis genes show distant relatedness to the lambda group of siphoviruses. ISM-pro1 head morphogenesis genes including ORFs 17, 18, and 19 (encoding portal-protease-major capsid protein) share weak sequence identity and the same gene order with those of the lambda-group siphoviruses, which includes *Burkholderia* phage phi644-2, *Burkholderia* phage phiE125, and *Klebsiella* phage phiKO2. For example, the products of ISM-pro1 ORFs 17, 18, and 19 share 22, 28, and 39% amino acid identity with *Burkholderia* phage phiE125 gp4, 5, and 6, respectively (38). In addition, ISM-pro1 ORF 28 also shows partial sequence identity with tail tape measure protein of lambdoid *Pseudomonas* phage D3 (51% amino acid identity over the N-terminal 213-amino-acid [aa] region). ORF 35 encodes the putative phage lysozyme and contains the glycoside hydrolase family 19 chitinase domain (cd00325).

TABLE 1. Bioinformatics analysis of ISM-pro1 ORFs

ORF	GC content (%)	Size (aa)	Predicted function	Best database hit; % aa identity (% similarity)
1	50.2	352	Phage integrase	Phage integrase family protein [*Paracoccus denitrificans* PD1222]; 44 (58)
2	56.3	92	Unknown	Hypothetical protein rsp_6015 [*Rhodobacter sphaeroides* 2.4.1]; 47 (64)
3	63.5	220	Putative phage-related gene	Conserved hypothetical protein [*Rhodobacteraceae* bacterium KLH11]; 79 (87)
4	59.4	59	Unknown	
5	60.1	116	Unknown	Hypothetical protein R2601_23860 [*Roseovarius* sp. HTCC2601]; 60 (77)
6	61.5	95	Unknown	
7	56.4	112	Unknown	Hypothetical protein EE36_14048 [*Sulfitobacter* sp. EE-36]; 54 (72)
8	56.7	83	Unknown	
9	55.7	72	Unknown	Hypothetical protein R2601_23895 [*Roseovarius* sp. HTCC2601]; 50 (64)
10	62.5	88	Unknown	
11	61.3	204	Unknown	Hypothetical protein SSE37_25308 [*Sagittula stellata* E-37]; 44 (61)
12	65.3	193	Unknown	Hypothetical protein Rsph17025_1746 [*Rhodobacter sphaeroides* ATCC 17025]; 39 (59)
13	65.2	246	Unknown	Hypothetical protein RCCS2_17656 [*Roseobacter* sp. CCS2]; 45 (60)
14	61.7	132	HNH nuclease	Hypothetical protein RCCS2_17671 [*Roseobacter* sp. CCS2]; 55 (67)
15	60.5	151	Putative phage terminase, small subunit	Hypothetical protein RCCS2_17676 [*Roseobacter* sp. CCS2]; 59 (73)
16	59.7	555	Putative phage terminase, large subunit	Hypothetical protein RCCS2_17681 [*Roseobacter* sp. CCS2]; 74 (82)
17	61.3	425	Putative portal protein	Hypothetical protein RSP_1648 [*Rhodobacter sphaeroides* 2.4.1]; 58 (73)
18	64.7	309	Periplasmic serine protease (ClpP class)	Peptidase [*Rhodobacter sphaeroides* 2.4.1]; 52 (68)
19	64.6	438	Putative phage major capsid	Putative phage-related protein [*Roseobacter* sp. CCS2]; 50 (67)
20	59.8	72	Unknown	
21	62.3	188	Putative phage-related gene	Hypothetical protein RCCS2_17711 [*Roseobacter* sp. CCS2]; 37 (51)
22	57	113	Putative head-tail adaptor	Hypothetical protein RSP_1641 [*Rhodobacter sphaeroides* 2.4.1]; 46 (65)
23	59.7	172	HK97 family phage protein	Hypothetical protein RSKD131_0362 [*Rhodobacter sphaeroides* KD131]; 48 (63) Hypothetical protein RSP_1640 [*Rhodobacter sphaeroides* 2.4.1]
24	60.2	127	Unknown	Hypothetical protein EE36_14123 [*Sulfitobacter* sp. EE-36]; 54 (70)
25	63.3	147	Unknown	Hypothetical protein NAS141_19229 [*Sulfitobacter* sp. NAS-14.1]; 74 (82)
26	59.4	95	Unknown	Hypothetical protein EE36_14133 [*Sulfitobacter* sp. EE-36]; 47 (72)
27	59.6	74	Unknown	Hypothetical protein NAS141_19239 [*Sulfitobacter* sp. NAS-14.1] 50 (74)
28	62.4	1129	Putative phage tail tape	Lambda family phage tail tape measure protein [*Rhodobacter sphaeroides* 2.4.1]; 39 (53)
29	60.1	312	Carbohydrate-binding family V/XII protein	Hypothetical protein NAS141_19254 [*Sulfitobacter* sp. NAS-14.1]; 28 (48)
30	58.5	327	Unknown	
31	54.6	112	Unknown	
32	55.8	94	Unknown	Hypothetical protein MED193_18929 [*Roseobacter* sp. MED193]; 51 (69)
33	63.1	74	Unknown	
34	60.7	128	Unknown	
35	62.7	200	Putative lytic enzyme	Hypothetical protein Oant_0261 [*Ochrobactrum anthropi* ATCC 49188]; 46 (63)
36	67	105	Unknown	Hypothetical protein dvul_1461 [*Desulfovibrio vulgaris* subsp. *Vulgaris* dp4]; 38 (45)
37	61.4	75	Unknown	
38	47	21	Unknown	
39	49.7	48	Unknown	Hypothetical protein retlb5_26363 [*Rhizobium etli* Brasil 5]; 80 (82)
40	50.5	32	Unknown	

Close relatives of ISM-pro1. A GenBank BLASTP search showed that 25 of 40 ISM-pro1 ORFs have homologues in other members of the *Rhodobacterales*, suggesting that it may be possible to find ISM-pro1 relatives in some *Rhodobacterales* genomes. A close examination of these bacterial genomes identified prophage-like elements in *Roseobacter* sp. CCS2, *Sulfitobacter* sp. NAS-14.1, and *Rhodobacter sphaeroides* 2.4.1. The putative prophages in these bacterial genomes were similar to ISM-pro1 in genome size and gene arrangement (Fig. 4). ISM-pro1 shares 18, 11, and 18 ORFs with prophages from *Roseobacter* sp. CCS2, *Sulfitobacter* sp. NAS-14.1, and *Rhodobacter sphaeroides* 2.4.1, respectively. These homologous ORFs are arranged in a similar gene order with the exception of ORFs 23 and 24 in *Roseobacter* sp. CCS2 prophage, whose order was different from those in other related prophages (Fig. 4A). These prophages share identity in DNA packaging, morphogenesis, and some unknown genes. Compared to *Roseobacter* sp. CCS2 prophage and *Rhodobacter sphaeroides* 2.4.1 prophage, *Sulfitobacter* sp. NAS-14.1 prophage shares fewer ORFs with ISM-pro1. In addition, the *Sulfitobacter* sp. NAS-14.1 prophage also has morphogenesis genes that are very distantly related to those of ISM-pro1, although they all show relatedness to the siphoviruses of the lambda group. This relatedness of these individual prophages suggests that this type of prophage is common in *Rhodobacterales* and may play important roles in horizontal gene transfer during evolution. However, the genomes of these prophages also appear to be mosaics, implying that they evolved with much genetic exchange (Fig. 4).
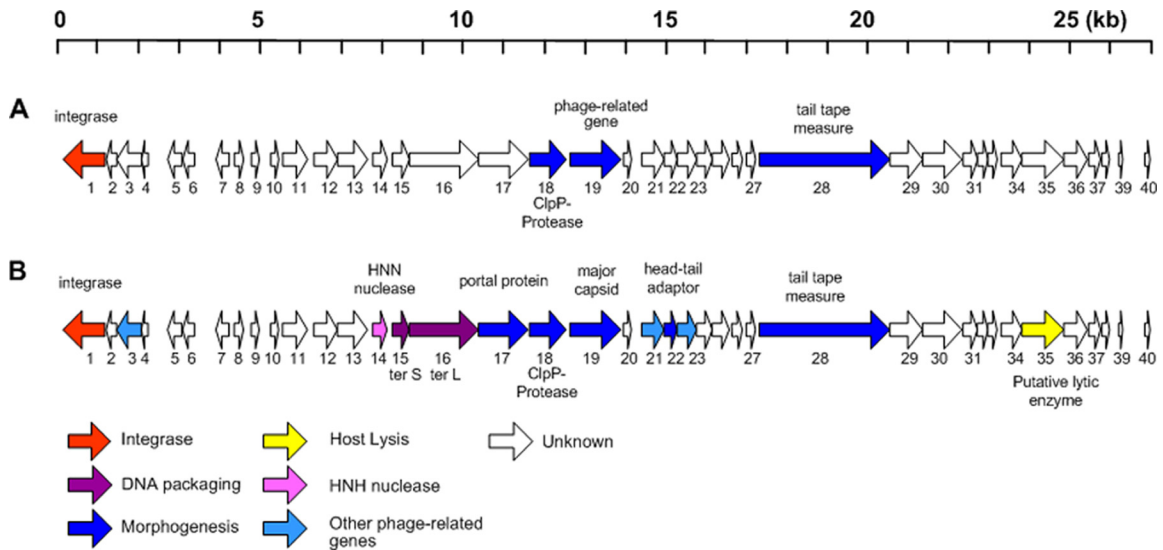
FIG. 3. Comparison of genome maps of ISM-pro1 before and after reannotation. ORFs are oriented according to the direction of transcription (the direction of arrows). (A) Original annotation of ISM-pro1 based on the genome sequence of *Roseovarius nubinhibens* ISM. (B) Reannotation of ISM-pro1.

Our comparative analysis suggests that *Roseobacter* strain CCS2 may also contain a prophage similar to ISM prophage, which may have gone undetected in a previous screening of the genome of *Roseobacter* sp. CCS2 by Paul (27). The induction experiment should

be conducted to confirm the lysogeny of *Roseobacter* sp. CCS2, *Sulfitobacter* sp. NAS-14, and *Rhodobacter spaeroides* 2.4.1.

Small size of encapsidated DNA. One interesting phenomenon observed after prophage induction is the presence of
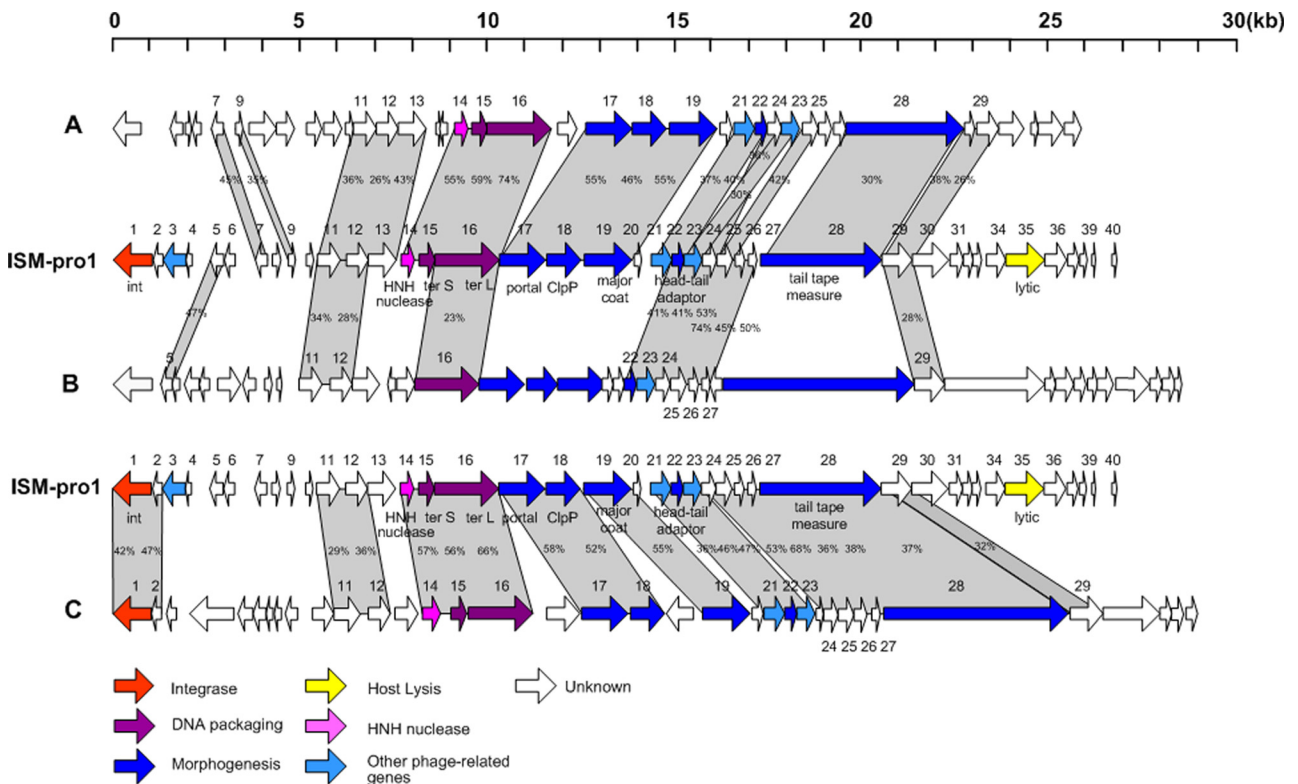

FIG. 4. Genomic comparison between ISM-pro1 and three representative prophage-like elements found in other *Rhodobacterales* genomes. Genes with sequence identity are connected by gray shading. The numbers in the shaded areas represent the sequence similarities at the amino acid level. For comparison, similar ORFs were numbered as in ISM-pro1. (A) *Roseobacter* sp. CCS2: from RCCS2_17591 (phage integrase) to hypothetical protein RCCS2_17776. (B) *Sulfitobacter* sp. NAS-14.1: from NAS141_19114 (site-specific integrase/recombinase-like) to hypothetical protein NAS141_19304. (C) *Rhodobacter sphaeroides* 2.4.1: from hypothetical protein RSP_1630 to hypothetical protein RSP_1662 (integrase/recombinase).

small DNA fragments (*ca*. 3, 4, and 12 kb) within the induced VLPs (Fig. 2A). These bands were invisible when ISM was not induced by mitomycin C. We hypothesize that these small DNA fragments were from gene transfer agent (GTA) particles released by ISM. A GTA is a small phage-like particle containing a random, small fragment of bacterial genomic DNA (29). The genes that encode *R. capsulatus* GTA have been characterized (21). Recently, related, conserved GTA gene clusters were found in, and appear to be limited to, the *Rhodobacterales* and other *Alphaproteobacteria* (3, 22). The genome of ISM contains a conserved GTA gene cluster similar to that of many marine roseobacters (3, 27). Do these small PFGE bands result from the DNA fragments enclosed in the GTA particles? GTA particles in *R. capsulatus* are known to contain an ~4.5-kb DNA fragment (29). A 10-kb DNA fragment associated with GTA has been reported in the alphaproteobacterium *Azospirillum lipoferum* (6). Recently it was found that virus-like particles released from *Silicibacter* sp. TM1040 also contain a *ca*. 3-kb DNA fragment (Y. Zhao et al., unpublished data). Earlier studies showed that GTAs are not inducible by mitomycin C (29). Without mitomycin C induction, we could not detect these small DNA bands in PFGE (data not shown). It is possible that induction of the ISM prophage and viral lysis may help to release more GTA particles from ISM cells, thereby making the small DNA bands from GTA more visible on PFGE gels. We also cannot exclude the possibility that these small DNA fragments come from additional cellular DNA segments packed in prophage capsids. Further studies are warranted to understand the source of these small DNA fragments.

Conclusion. Using a combination of biological and genomic approaches, we confirmed that a marine bacterium, *Roseovarius nubinhibens* ISM, contains an inducible prophage. Our study demonstrated that identification of prophages based solely on known bacterial or phage genomes could be misleading. Prophage identification is still facing unexpected challenges, and we are far from a complete understanding of the structure and function of prophages. Reannotation of prophage-like regions in bacterial genomes is necessary for proper identification of prophages. In addition, prophage induction experiments should be performed to test for functional prophages.

REFERENCES

1. **Ackermann, H. W., and M. S. DuBow.** 1987. Viruses of prokaryotes, vol. I. General properties of bacteriophages. CRC Press, Inc., Boca Raton, FL.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Biers, E. J., K. Wang, C. Pennington, R. Belas, F. Chen, and M. A. Moran.** 2008. Occurrence and expression of gene transfer agent (GTA) genes in marine bacterioplankton. Appl. Environ. Microbiol. **74:**2933–2939.
4. **Bose, M., and R. D. Barber.** 2006. Prophage Finder: a prophage loci prediction tool for prokaryotic genome sequences. In Silico Biol. **6:**223–227.
5. **Boyd, E. F., and H. Brüssow.** 2002. Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. Trends Microbiol. **10:**521–529.
6. **Boyer, M., J. Haurat, S. Samain, B. Segurens, F. Gavory, V. González, P. Mavingui, R. Rohr, R. Bally, and F. Wisniewski-Dyé.** 2008. Bacteriophage prevalence in the genus *Azospirillum* and analysis of the first genome sequence of an *Azospirillum brasilense* integrative phage. Appl. Environ. Microbiol. **74:**861–874.
7. **Brüssow, H., C. Canchaya, and W. D. Hardt.** 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol. Mol. Biol. Rev. **68:**560–602.
8. **Canchaya, C., C. Proux, G. Fournous, A. Bruttin, and H. Brüssow.** 2003. Prophage genomics. Microbiol. Mol. Biol. Rev. **67:**238–276.
9. **Canchaya, C., G. Fournous, and H. Brussow.** 2004. The impact of prophages on bacterial chromosomes. Mol. Microbiol. **53:**9–18.
10. **Canchaya, C., G. Fournous, S. Chibani-Chennoufi, M. L. Dillmann, and H. Brüssow.** 2003. Phage as agents of lateral gene transfer. Curr. Opin. Microbiol. **6:**417–424.
11. **Casjens, S.** 2003. Prophages and bacterial genomics: what have we learned so far? Mol. Microbiol. **49:**277–300.
12. **Chen, Y., I. Golding, S. Sawai, L. Guo, and E. C. Cox.** 2005. Population fitness and the regulation of *Escherichia coli* genes by bacterial viruses. PLoS Biol. **3:**1276–1282.
13. **Chen, F., J. R. Lu, B. Binder, Y. C. Liu, and R. E. Hodson.** 2001. Application of digital image analysis and flow cytometry to enumerate marine viruses stained with SYBR gold. Appl. Environ. Microbiol. **67:**539–545.
14. **Chen, F., K. Wang, J. Stewart, and R. Belas.** 2006. Induction of multiple prophages from a marine bacterium: a genomic approach. Appl. Environ. Microbiol. **72:**4995–5001.
15. **Fouts, D. E.** 2006. Phage-Finder: automated identification and classification of prophage regions in complete bacterial genome sequences. Nucleic Acids Res. **34:**5839–5851.
16. **Fuhrman, J. A., S. H. Lee, Y. Masuchi, A. A. Davis, and R. M. Wilcox.** 1994. Characterization of marine prokaryotic communities via DNA and RNA. Microb. Ecol. **28:**133–145.
17. **Jiang, S. C., and J. H. Paul.** 1994. Seasonal and diel abundance of viruses and occurrence of lysogeny/bacteriocinogeny in the marine environment. Mar. Ecol. Prog. Ser. **104:**163–172.
18. **Jiang, S. C., and J. H. Paul.** 1996. Occurrence of lysogenic bacteria in marine microbial communities as determined by prophage induction. Mar. Ecol. Prog. Ser. **142:**27–38.
19. **Jiang, S. C., and J. H. Paul.** 1998. Significance of lysogeny in the marine environment: studies with isolates and a model of lysogenic phage production. Microb. Ecol. **35:**235–243.
20. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
21. **Lang, A. S., and J. T. Beatty.** 2000. Genetic analysis of a novel bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus.* Proc. Natl. Acad. Sci. U. S. A. **97:**859–864.
22. **Lang, A. S., and J. T. Beatty.** 2007. Importance of widespread gene transfer agent genes in alpha-proteobacteria. Trends Microbiol. **15:**54–62.
23. **Leitet, C., B. Riemann, and Å. Hagström.** 2006. Plasmids and prophages in Baltic Sea bacterioplankton isolates. J. Mar. Biol. Assoc. U.K. **86:**567–575.
24. **Levin, B. R., and R. E. Lenski.** 1983. Coevolution in bacteria and their viruses and plasmids. *In* D. J. Futuyma and M. Slatkin (ed.), Coevolution. Sinauer, Sunderland, MA.
25. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
26. **Mehta, P., S. Casjens, and S. Krishnaswamy.** 2004. Analysis of the lambdoid prophage element e14 in the E. coli K-12 genome. BMC Microbiol. **4:**4.
27. **Paul, J. H.** 2008. Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? ISME J. **2:**579–589.
28. **Rao, G. V., P. Mehta, K. V. Srividhya, and S. Krishnaswamy.** 2005. A protein similarity approach for detecting prophage regions in bacterial genomes. Genome **6:**11.
29. **Solioz, M., and B. Marrs.** 1977. The gene transfer agent of *Rhodopseudomonas capsulata*—purification and characterization of its nucleic acid. Arch. Biochem. Biophys. **181:**300–307.
30. **Srividhya, K. V., G. V. Rao, L. Raghavenderan, P. Mehta, J. Pirulsky, S. N. Manicka, J. L. Sussman, and S. Krishnaswamy.** 2006. Database and comparative identification of prophages. Lect. Notes Control Inf. Sci. **344:**863–868.
31. **Srividhya, K. V., V. Alaguraj, G. Poornima, D. Kumar, G. P. Singh, L. Raghavenderan, A. V. S. K. M. Katta, P. Mehta, and S. Krishnaswamy.** 2007. Identification of prophages in bacterial genomes by dinucleotide relative abundance difference. PLoS One **2:**e1193.
32. **Stopar, D., A. Cerne, M. Zigman, M. Poljisal-Prijatelj, and V. Turk.** 2003. Viral abundance and high proportion of lysogens suggest that viruses are important members of the microbial community in the Gulf of Trieste. Microb. Ecol. **46:**249–256.
33. **Weinbauer, M. G., and C. A. Suttle.** 1996. Potential significance of lysogeny to bacteriophage production and bacterial mortality in coastal waters of the Gulf of Mexico. Appl. Environ. Microbiol. **62:**4374–4380.
34. **Weinbauer, M. G., and C. A. Suttle.** 1999. Lysogeny and prophage induction in coastal and offshore bacterial communities. Aquat. Microb. Ecol. **18:**217–225.
35. **Williamson, K. E., M. Radosevich, D. W. Smith, and K. E. Wommack.** 2007. Incidence of lysogeny within temperate and extreme soil environments. Environ. Microbiol. **9:**2563–2574.

36. **Williamson, S., L. McDaniel, L. Houchin, and J. H. Paul.** 2002. Seasonal variation in lysogeny as depicted by prophage induction in Tampa Bay, Florida. Appl. Environ. Microbiol. **68:**4307–4314.

37. **Wommack, K. E., J. Ravel, R. T. Hill, J. Chun, and R. R. Colwell.** 1999. Population dynamics of Chesapeake Bay virioplankton: total-community analysis by pulsed-field gel. Appl. Environ. Microbiol. **65:**231–240.

38. **Woods, D. E., J. A. Jeddeloh, D. L. Fritz, and D. DeShazer.** 2002. Burkholderia thailandensis E125 harbors a temperate bacteriophage specific for Burkholderia mallei. J. Bacteriol. **184:**4003–4017.

39. **Zhao, Y. L., K. Wang, C. Budinoff, A. Buchan, A. Lang, N. Z. Jiao, and F. Chen.** 2009. Gene transfer agent (GTA) genes reveal diverse and dynamic *Roseobacter* and *Rhodobacter* populations in Chesapeake Bay. ISME J. **3:**364–373.